ReuseECOS: An Approach to Support Global Software Development through Software Ecosystems

Rodrigo Pereira dos Santos, Cláudia Maria Lima Werner
Computer Science and System Engineering Department
COPPE – Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
{rps, werner}@cos.ufrj.br

Abstract – Economic and social issues were pointed out as Software Engineering (SE) challenges for the next years, since the field needs to treat issues beyond the technical side, which requires observing it in another perspective. In this sense, Software Ecosystems (SECOs) have emerged as an approach to improve software reuse in global software development (GSD) industry considering relations among companies and stakeholders around the world. Companies and organizations have opened up their platforms and artifacts to others, including partners and third-part developers. This scenario changes the traditional software industry and requires mature research in software architecture, component-based software engineering and software product line (SPL) in a global market and business environment. So, GSD requires linking an architectural, a business and a social-based environment in an integrated way, as well as a focus on SECO-based management and engineering. This concern motivated a proposal of a “3+1” framework for SECOs management and engineering called ReuseECOS. The approach aims at outlining a set of steps that combines those three dimensions and joins different perspectives in SECOs literature from a survey through a research strategy and a solution proposal for GSD with SECO and SPL. It was observed that SECO concepts can be merged in a broader GSD approach in SE industry.

Keywords – software ecosystems; software reuse; component-based development; global software development; social networks; value-based software engineering

I. INTRODUCTION

According to Bosch and Bosch-Sijtsema [5], for most software systems companies, large-scale software development is complicated, expensive, slow and unpredictable. So, Boehm [3] has pointed out economic and social issues in Software Engineering (SE) as a challenge for the next years, since the field needs to treat issues beyond the technical side, which requires observing it in another perspective. An example is the fact that software teams have been working geographically dispersed since industry decided to explore the distribution of software development processes and resources as a mean to reduce cost and development time, and also to expand to new markets [11][23]. This context modifies the process of decision making in the global software industry because it requires not only the understanding of investments and costs, but benefits, risks and opportunities [2]. It also requires more software engineers who focus on the system as a whole, that is, to combine software, hardware and “peopleware” in a platform [4]. Thus, global software development (GSD) involves better thinking about platforms and their nets of artifacts and stakeholders worldwide (i.e., developers, project managers, business staff, engineers, users, clients etc.).

In order to analyze and understand the resulting challenge, three trends accelerate the complexity of software development [5]: (i) the wide-spread adoption of software product lines (SPLs) and the sub challenge of exposing reusable assets and opening the platform architecture; (ii) the broad globalization of software development in many organizations and the sub challenge of elevated complexity of sociotechnical dependency management; and (iii) the importance of building a software ecosystem (SECO), i.e., a community of external developers around a successful product as a popular and central hub to the strategy of many companies, and the sub challenge of structuring the dependencies between the platform company and the third-part developers since it is an evolution of the software reuse topic [4]. Especially in (ii), literature has shown that geographical distance has a significant impact on the coordination and communication processes among members of dispersed development teams, affecting their performance. Despite the challenges faced by these teams, distributed software development (DSD) has become a reality that developers have to deal with every day [9][11][23].

In this sense, component-based development (CBD) represents a fundamental pillar of global software platforms considering an engineering mindset and the three mentioned sub challenges. Components can be internally developed or acquired as COTS, or open source. Also, mature industrial application development is reuse-driven and uses existent artifacts when possible, contributing to improvements and maturing repositories [Ove04]. Thus, the development of a new system is focused on specific requirements and components that represent a competitive advantage for the platform, considering the system domain peculiarities, clients’ needs, existent technologies etc. In other words, the traditional strategy for developing a unique product has been changed into developing multiple products, based on a common system architecture that consists in exploring reuse in GSD [13].

On the other hand, due to the tendency of opening global platforms to third-part developers in order to get a competitive advantage and maintain a community around the platform and systems [8], CBD critical issues can be listed:
standardization, information visualization, guidelines to support customizations, and intellectual property [24]. Some questions emerge: “open source or proprietary: what is the best scenario?”; “what is the most attractive strategy?”; “what is the most innovative one?”; “how to perform the opening process?” etc. Considering these changes in research and practice of GSD, the SECO point of view has motivated main players in industry to rethink their operational actions in order to open their platforms to external agents and keep their business alive. This fact changes global software industry because it requires linking an architectural, a business and a social-based environment in an integrated way, and it focuses on SECO-based Management and Engineering (M&E). Since SECO has focused on large systems of systems, this concept is not formally applied to Brazilian platforms yet, although the governmental platform can represent a significant case study [20], considering initiatives in relation to DSD and Software Reuse Processes.

Also, agile methodologies and specific software quality models such as MPS Model can create opportunities to generate a large SECO from a set of micro, small and medium companies [14].

This research explores this issue in GSD and Brazilian DSD in order to develop a proposal of a “3+1” framework for SECOs M&E, called ReuseECOS. The approach outlines a set of steps that combines three different dimensions of SECOs and joins existing perspectives in SECOs research from a survey [14] through a research strategy and a solution proposal for GSD with SECO and SPL. It is also based on some results of a preliminary analysis of the Software Reuse Lab’s Brazilian SECOs at COPPE/UFRJ [14][16][17]. So, this paper integrates an evolution of a Master thesis related to a value-based component market environment [12][13] to a PhD proposal focused on exploring business and social issues in SECOs’ architectures [14][15][16][17], considering SPLs in GSD, as presented in Section IV. Besides this introduction, the paper is organized in the following: Section II summarizes the background related to SECOs and related work; Section III presents an overview of ReuseECOS, based on steps to contemplate a SECO “3+1” view during its lifecycle; Section IV discusses a solution proposal for GSD with SECO and SPL; and Section V concludes the paper.

II. BACKGROUND

A. Software Ecosystems

SECOs consists a new perspective in SE field due to their rapid evolution in the last decade, though the first researches in this topic were done by Business Schools in the 90’s [10]. Jansen et al. [8] defines SECO as a set of actors functioning as a unit and interacting with a shared market for software and services, together with the relationships among them, frequently underpinned by a common technological platform or market, and operating through the exchange of information, resources and artifacts. Some examples are MySQL/PHP SECO, Microsoft SECO, and iPhone SECO. SECOs studies in global software industry were motivated by the SPL approach aiming at allowing external developers to contribute to hitherto closed platforms through GSD [4][5]. Research directions at literature and industrial cases reinforce important SECO perspectives, such as architecture, mobile platforms, GSD, social networks, modeling, business considerations, organizational-based management [8], and a multidisciplinary treatment [13]. In GSD, software vendors resort to virtual integration through alliances to create and keep networks of influence and interoperability, generating SECOs worldwide. The SECOs lifecycles are also analyzed through four phases [8]: (1) the establishment of a market relationship with a dominant and focused organization; (2) the emergence of a preliminary network; (3) the reduction of the dominant (and focused) organization’s power, and the stimulus of new communities of practice; and (4) the existence of a community of creation, where no dominant organizations exist and the power is distributed.

Jansen et al. [8] model SECOs in a three-level view aiming at understanding the mentioned challenges. Software vendor level is the first one where the objects of study are the actors and their relationships in the context of an organization in the SECO. In turn, software supply network level is the second one where the objects of study are the software supply networks (SSNs) as well as their relationships including all stakeholders and the internal characteristics related to the SECO health and stability. Finally, SECO level is the third one where the object of study is the SECO itself and its relationships. In this sense, strategic choices should be made on how a software vendor behaves in a GSD environment to maximize profitability. For example, SECO orchestrators have control over the SECO and can develop strategies to keep a SECO vibrant and profitable for involved organizations.

B. Related Work

The comprehension and realization of time (i.e., evolution) and space (i.e., GSD) dimensions of SECOs were pointed out by Jansen et al. [8], and explored by Hunink et al. [7] and van der Berk et al. [22]. Hunink et al. [7] propose a method to create a SECOs domain specific taxonomy in a wide and complete way, allowing software vendors, scientists and government to have insights and identify the gaps between needed and shared information, called ITAXEM (Industry Taxonomy Engineering Method). This method is divided in steps, based on the behaviorist science, i.e., make the organizational capacity in creating knowledge and artifacts clear. A case study was executed with a European SECO and generated EUSIT (European Software Industry Taxonomy). This result pointed some challenges and weaknesses, such as source quality (a limited number of taxonomies was collected and selected, e.g., Amadeus, ITEuropa, Truffle 100 2008, and databases such as Capterra and Software Network), and stakeholders identification (the method does not explicitly define how stakeholders can be identified among independent software vendors, public companies, researchers etc.).

In turn, van der Berk et al. [22] present a model to describe SECO key characteristics, aiming at evaluating the
SECO status and observing how decisions can impact its performance, or generate strategic advantages based on experience (i.e., past), called SECO-SAM (SECO Strategy Assessment Model). In this sense, the work explores the theoretical basis from Business Schools, frequently mapping or instantiating concepts and indicators to the SE context, in four dimensions: Biology, Lifestyle, Environment and Health Care Organization. This model was evaluated through a case study which was executed in Open Design Alliance (ODA). ODA is an organization that promotes open standards for CAD (dwg), focusing on its platform composed by many software libs to support graphical applications. Some issues were identified, such as how ODA affects insights about SECOs; what the main quantitative characteristics observed on SECOs were; and how the evaluation process of SECO-SAM was in the stakeholders’ point of view.

Despite these efforts motivated by the lack of methods, techniques and tools to maximize the awareness in SECOs, no link between a process and a model has been established, aiming at understanding the SECOs lifecycles from their birth to maturity (or disappearance, eventually). This is an orthogonal problem to SECOs and requires exploring information extracted from a set of existent parameters and behaviors in software industry and real cases reported by academy. That is the motivation to SECO-based M&E through the ReuseECOS approach proposed in this research. Moreover, a community consensus exists: architectural, business and social dimensions are the pillars of a SECO analysis, and they involve thinking about software reuse, architecture, computer supported cooperative work, social networks, GSD, sociotechnical aspects in SE, software and system quality, and applied economy [6][14].

III. REUSEECOS APPROACH

The focus on the SECO scope is required to understand SECOs from a three-level perspective, since each level has distinct research challenges, e.g., health, interaction, inputs, outputs, performance, competition, value sharing and coordination methods. When GSD, SECO and SPL are put together in an architecture, business and social-based view, many organizations play with a resulting model from older and newer SE process models in the market. Considering the SECO challenges [14], such as “why do SECOs appear and disappear?” and “how to define and monitor SECO scope, types, roles and characteristics”, this research defines ReuseECOS, a proposal of a “3+1” framework for SECOs M&E, as explained below. The goal is to understand SECOs generated by different SSNs throughout their lifecycle phases in their three levels of scope and allowing the identification of new SECOs. That is, from the calibration of the GSD experience reports and historical data repository based on a research methodology, a SECO M&E mechanism provides a step-by-step process to extract a SECO diagnosis (management) and to interfere in the SECO (engineering).

ReuseECOS methodology aims at guiding and allowing deeper researches related to support SECOs M&E based on empirical studies in a research strategy extended from [19], as detailed in TABLE I and summarized in Figure 1: after the broader ad hoc literature review, the next steps are to plan and execute a systematic review to calibrate the framework, and then a survey to verify it with practitioners. This can maintain a body of knowledge that will support case studies with well-known SECOs such as Android, Blackberry, Force.com, Eclipse, Microsoft, Linux etc., as well as Brazilian candidate SECOs, such as Brazilian Public Software (BPS). Finally, this research strategy can provide a database to make SECOs diagnosis, design, and validation and decision-making processes available based on the fact that ecosystems appear, are developed, mature and/or disappear as well as markets, technologies, platforms and organizations, processes, models, techniques etc.

<table>
<thead>
<tr>
<th>Phases</th>
<th>Objectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ad hoc Literature Review</td>
<td>Aims at identifying basic concepts/terms about SECO and allowing the definition of a broad and more precise systematic review protocol for supporting each SECO dimension.</td>
</tr>
<tr>
<td>Systematic Review</td>
<td>Aims at effectively elaborating and executing the systematic review protocol in each SECO dimension. Based on the results extracted from source analysis, the researchers from each dimension decide if the study needs to be refined, or if the knowledge should be evaluated (survey).</td>
</tr>
<tr>
<td>Survey</td>
<td>Aims at planning and executing studies to evaluate the knowledge acquired in the previous phase, considering SECO community in industry.</td>
</tr>
<tr>
<td>Body of Knowledge</td>
<td>Aims at joining all knowledge to organize a “3+1” SECO body of knowledge (e.g., reports, papers; data from open source software; historical business data collected from Internet; sociotechnical networks inferred from business and social networks; architectural data collected from IDEs and reuse repositories). This phase requires a repository to share information related to M&amp;E.</td>
</tr>
<tr>
<td>Historical Data and Experience Reports Base</td>
<td>Aims to support the SECO platform development and lifecycle through a mechanism integrated to the repository which should allow search, retrieval, manipulation, storage documentation, and publishing of the SECOs data and reports.</td>
</tr>
</tbody>
</table>

Figure 1. Research Strategy [16].

ReuseECOS is structured in a set of related steps classified in three basic dimensions extended from [6] and integrated through a M&E dimension according to a SECO “3+1” view (Figure 2). The analysis of each dimension is divided in large steps, and in turn these are composed by a set of activities. This paper focuses on presenting a ReuseECOS as a roadmap for a SECO research to support a
GSD environment. More specific details should be obtained in papers referred in each dimension:

**Dimension #1: architectural dimension**, focused on the SECO platform (market, technology, infrastructure or organization) through platform domain engineering process (establishing its lifecycle), commonalities and variabilities management (defining platforms features), and developed SPL architecture (treating the platform as SPL). As illustrated by Figure 3, this dimension aims at selecting the GSD’s target platform in order to contextualize its project and development, plan its opening process considering its architecture, and balance modularity (componentization) and transparence (visualization) during its evolution and maintenance, which represent SECO engineering. The activities were extracted from previous works and provide means to instantiate the framework and analyze all three dimensions together, as well as the other dimensions. Details on the architectural dimension are in [14].

![Figure 3. Architectural dimension of SECO M&E](image)

**Dimension #2: business dimension**, focused on SECO knowledge flow, that is, artifacts, resources and information, through a business (establishing GSD goals and action plans by programs and projects), innovation (linking a SECO to a GSD market), and strategic planning (understanding how, when, where and who will perform the goals) views. As presented in Figure 4, this dimension aims at capturing the platform context in order to apply the GQM (Goal-Question-Metric) approach [1]. This approach allows to collect, manipulate and show sustainability and diversity information as health indicators of SECOs. From the set of indicators, a framework of SECOs management can be modeled and instantiated. More details on the business dimension are in [16].

![Figure 4. Business dimension of ReuseECOS](image)

**Dimension #3: social dimension**, focused on SECO stakeholders through balancing proposition and realization of utility (why stakeholders integrate, extend and modify knowledge in a SECO, and interact to each other in GSD), promotion (how stakeholders’ capabilities and engagement are implicit and explicitly recognized in GSD), and knowledge (what collaboration, open source development and other social network opportunities contribute to stakeholders in GSD). As illustrated by Figure 5, this dimension aims at understanding how the social networks creation, organization and maintenance can affect the communities that belong to a SECO, considering acquired, exchanged and transmitted knowledge in these networks (i.e., resources, artifacts and information). More details on the social dimension are in [17].

![Figure 5. Social dimension of ReuseECOS](image)
• **Dimension #4: M&E dimension**, focused on merging the three basic dimensions through three relationships towards establishing of a technological infrastructure:

**Relation #1: platform development and evolution**, motivated by links and traces between the business and architectural dimensions, focusing on understanding relationships and system models in a broader way, not just technically, and also in a GSD point of view;

**Relation #2: platform establishment**, done through links and actions between the social and architectural dimensions, which highlight the involvement of (and attention to) the community around the GSD platform;

**Relation #3: platform value**, maintained through links and communication between the social and business dimensions, which aim to map value propositions (what the concept of value is for all stakeholders from their point of view) and realizations (what the feeling of value is for all stakeholders from the GSD environment and context perspective).

**IV. A Solution Proposal**

In order to make ReuseECOS approach more concrete, this section discusses a solution proposal that integrates the “3+1” framework to a GSD environment for SECOs’ platforms based on the pointed dimensions and relations. As shown in Figure 6, three GSD tools are combined and are respectively mapped to dimensions #1, #2 and #3:

- **Tool #1: GSD environment**, an extension of an IDE to support requirements management, and components and services modeling and development through evolving the SECO platform architecture based on reuse management process in GSD. It is based on collecting, analyzing and visualizing: (i) architectural metrics, e.g., modularity/ transparency of software artifacts; (ii) business indicators, e.g., sustainability and diversity within the GSD platform; and (iii) community participation, e.g., aspects of DSD and crowd computing in a GDS industry;

- **Tool #2: value-based components and services repository**, an extension of a CBD infrastructure to support business and quality models for third-part development in a GSD environment with SPL, considering mechanisms such as [12]: (i) visualization of CBD market movements (e.g., offerings, demands, new niches etc.); (ii) marketing analysis; (iii) evaluation of components and services; (iv) negotiation of requirements and artifacts among stakeholders (components or services to producers and consumers, as well as third-part developers and orchestrators); and (v) pricing of software artifacts in a value-based perspective, focused on different value facets;

- **Tool #3: sociotechnical network site**, an extension of a social network site to include software artifacts and support the establishment, combination, dependencies control and evolution of artifacts and stakeholders’ resulting networks in a GSD environment, based on common features and users’ profiles for managing “friends” (with artifacts as citizens in a sociotechnical perspective [18][21]) and relationships in this site.

Besides, another tool is required in order to help SECO orchestrators (i.e., platform administrators) in decision making processes during the SECO lifecycle. This tool, called decision maker for orchestrators, should implement the steps and activities of ReuseECOS framework as wizard-like software in order to collect historical data and experience reports (Phase 5 in ReuseECOS Research Strategy), extract information and generate knowledge. Also, its interface should be contained into the GSD environment since this tool was modeled to communicate to the others.

On the other hand, the three relations identified in dimension #4 are supported by this tool. For example: (i) relation #1 happens from managing new SECO requirements provided (and required) by third-part developers in a GSD environment and stored in the value-based components and services repository, and are linked to architectural elements and functionalities through a traceability matrix and a reuse-based platform (development with and for reuse [24]); (ii) relation #2 happens from monitoring the sociotechnical network through the artifacts and stakeholders’ profiles (site) and the GSD’s community in practice, aiming at using this status to support architectural evolution; and (iii) relation #3 happens from eliciting and conciliating artifacts and stakeholders’ interests from network site and value-based repository, leveraging the SECO value in a GSD environment.

The solution proposal should be built as a prototype integrating existent tools. For example: the GSD environment can be supported by Eclipse IDE extended with plugins for DSD, and a specific plugin to provide the
interface with the other tools in ReuseECOS; the value-based components and services repository can be supported by Brechó-VCM, a web tool focused on exploring CBD in software markets [12]; and the sociotechnical network site can be supported by Facebook social network system and an extension to provide the interface with the other tools.

V. CONCLUSION

Since developing systems involves better thinking about reuse-based platforms, and their networks of artifacts and stakeholders, GSD community has discussed economic and social issues, which are becoming relevant topics in the field, considering the reality of SECOs [10]. GSD as a complex, effort consuming and unpredictable activity can be pointed out as a reason, and after decades of SE research and practice, problems in managing the constantly increasing complexity remain, mainly in GSD [5]. The lack of theoretical and applied research in SECOs M&E through a GSD point of view is making this field attractive in academic research and motivated this research. This paper presented an overview of a “3+1” framework for SECOs M&E, called ReuseECOS. The approach outlines a step of steps that combines three different dimensions of SECOs and joins existing perspectives in SECOs research from a survey through a research strategy and a solution proposal for GSD with SECO and SPL, also based on results of a preliminary analysis of the Software Reuse Lab’s Brazilian SECOs.

The contribution of this paper was to understand how GSD community can understand SECO and SPL as SE elements to GSD. As concluded, it is impossible to treat SECOs with a pure engineering approach, so, distinct dimensions were merged from a management perspective. SECOs requires joining a lot of (in)stable IT elements in an entity (platform), adding GSD elements which alter those elements during the ecosystems creation, development and maintenance. Future work consists of external evaluation of ReuseECOS with empirical studies (i.e., systematic review and survey) to calibrate the “3+1” framework steps and activities and effectively integrate them in a GSD unified approach. Additionally, an extension of the component and services library Brechó-VCM [15] to support a negotiation mechanism in SECOs is ongoing. The goal is to connect a value-based components and services repository (business dimension) to plugin-extended IDEs such as Eclipse and also plugin-extended social network sites such as Facebook, in order to support a GSD solution proposal.

ACKNOWLEDGMENT

The authors thank CNPq and FAPERJ for financial support.

REFERENCES