An Infrastructure to Support Socialization, Monitoring and Analysis of Software Ecosystems

Thaiana Lima, Gabriel Barbosa, Rodrigo Santos, Cláudia Werner

PESC/COPPE/UFRJ – System Engineering and Computer Science Department
Federal University of Rio de Janeiro – 21945-970 – Rio de Janeiro, RJ, Brazil

{thaiana, gabrielsb, rps, werner}@cos.ufrj.br

Abstract. *The software ecosystems context brings additional complexity to the organizations' activities since IT management decisions can strengthen (or weaken) relationships in the software supply network. Several acquirers lack common, structured documentation to allow visualizing and analyzing impacts of demands and solutions in their asset bases over time. We researched how the ecosystem perspective affects IT management activities, more specifically demand and solution analysis, aiming at helping IT managers and architects to make acquisition decisions. This paper presents an infrastructure to support socialization, monitoring and analysis of software ecosystems named Brechó-SocialSECO. A feasibility study was executed to evaluate our proposal.*

1. Introduction

According to Seichter et al. (2010), the artifacts shared among different stakeholders are the means for making interactions and communication concrete throughout the software development, and they create a common environment based on a technological platform. Such environment composed by actors, artifacts, and their relationships has been known as a Software Ecosystem (SECO). There is a network that represents those interactions, neither totally social (among actors) nor technical (among artifacts), i.e., the interactions involve actors manipulating artifacts (Lima et al., 2015).

Software organizations can benefit from information about the SECO elements, either a supplier or an acquirer. In this context, a critical acquirer's process is software acquisition, because it requires information about dependencies, technologies, suppliers, licensing, and support. Based on such elements, an acquirer should be able to choose which demands to prioritize and which solutions to incorporate into the organization's asset base (Finkelstein, 2014). However, several acquirers lack common documentation that is structured to allow visualizing and analyzing impacts of demands and solutions in their asset bases over time. So, the community participation (SECO *socialization*) and the asset base management (SECO *analysis*) are harmed so that obstacles are created for SECO monitoring. As such, an automated aid for managing and monitoring SECO elements becomes important (Santos, 2016).

We researched how the SECO perspective affects IT management activities, more specifically demand and solution analysis, aiming at helping IT managers and architects to make acquisition decisions. This paper presents an infrastructure to support socialization, monitoring and analysis of SECO that was evaluated through a feasibility study. Section 2 presents the background and related work; Section 3 describes our proposal; Section 4 discusses the feasibility study; and Section 5 concludes the paper.

2. Background

In the globalized software industry, there is a network composed by the SECO's elements and their relationships, including the involved organizations and community's demands. As such, the SECO context brings additional complexity to the organizations' activities since IT management decisions can strengthen (or weaken) relationships in the software supply network. For the traditional IT management, requirement specifications and available budget serve as key criteria to help IT managers and architects to analyze demands, lacking a structured software asset base to explore other indicators, especially the 'hidden effects' of their long-term decisions. In this intertwined network, acquirers operate relationships with/between suppliers and need to manage and monitor SECO elements in order to make decisions and take business advantages (Gartner, 2007).

The SECO context comes up with the importance of combining market information (inter-organizational) with those obtained from the software asset base (intra-organizational) to help IT managers and architects in daily activities (Finkelstein, 2014). To cope with this challenge, we have been investigating SECO management and extended a component library named Brechó (Werner et al., 2009). Brechó is a web information system for storage, publishing, search and retrieval, download, purchase and negotiation of software artifacts. Brechó is a technical repository that can serve as a software asset base, supporting licensing and SECO analysis. The lack of social resources to encourage stakeholders' interactions and communication had been pointed out as the Brechó's weakness until we conducted this research (Lima et al., 2015).

In the SECO literature, we can find few related work on the socialization, monitoring and/or analysis of SECOs. Peréz et al. (2012) present a tool named SECONDA that aims to aid SECO analysis and evolution according to project metrics and based on SECO behaviors and characteristics. However, this tool has some limitations, such as the lack of support for social relationships and for trend analysis from the collected information. The Brazilian Public Software (BPS) Portal is a web platform that offers a plethora of social interaction mechanisms, such as forums, communities, and a repository. BPS holds information on the actors' relationships, but there is no support for SECO monitoring and visualization. Goeminne et al. (2010) propose a framework for capturing SECO information based on extensions. By using data mining, SECO's graphics and a lifecycle chart are generated, but it is a generic framework that requires effort to program the information one needs to mine from it. All those mentioned gaps are part of the solution described in the proposed infrastructure.

3. Brechó-SocialSECO

Based on a survey conducted to investigate socio-technical resources for SECOs (Lima et al., 2015), we identified some social mechanisms as solutions for the problems of providing software artifact information and helping stakeholders to communicate: *artifact/team forum*, *team creation*, *actor/artifact profile*, *trend topics*, *suggestions and recommendations*, and *demand and solution analysis*. Such mechanisms can help IT managers and architects to be aware of SECO trends, and discussions would be of great value for the identification of new demands over time. In order to treat socialization in a SECO platform and aid an acquirer to satisfy objectives, stimulate collaboration and identify demands, Brechó-SocialSECO infrastructure was developed. Its objective is to

help an acquirer to understand the community surrounding the SECO platform as well as map relationships, manage demands and visualize the evolution of social interaction.

It is essential to record how demands and candidate solutions, also labeled as *software assets*, are added, deleted, and maintained. Considering the lack of research that combines socio-technical network with data obtained from the acquirer's asset base, Brechó-SocialSECO was built upon SECO socialization and analysis mechanisms to support IT management activities, as shown in Figure 1. Social mechanisms are used as inputs to capture information from actors, artifacts and interactions. Such information is processed to generate the SECO network graph. Based on this graph, measures can be extracted so that IT managers and architects can make decisions based on different configurations of the SECO's network (SECO analysis). The infrastructure was implemented as an extension of Brechó (Section 2). All the mechanisms for the SECO socialization, monitoring and analysis are joined in the "My Network" panel (Figure 2).

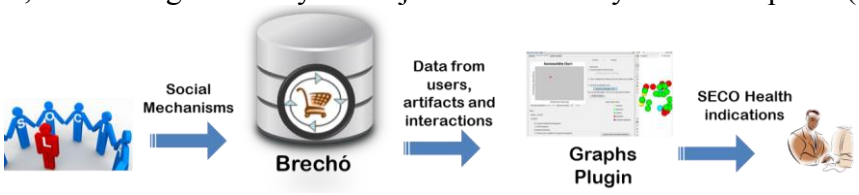

Figure 1. Brechó-SocialSECO's conceptual model

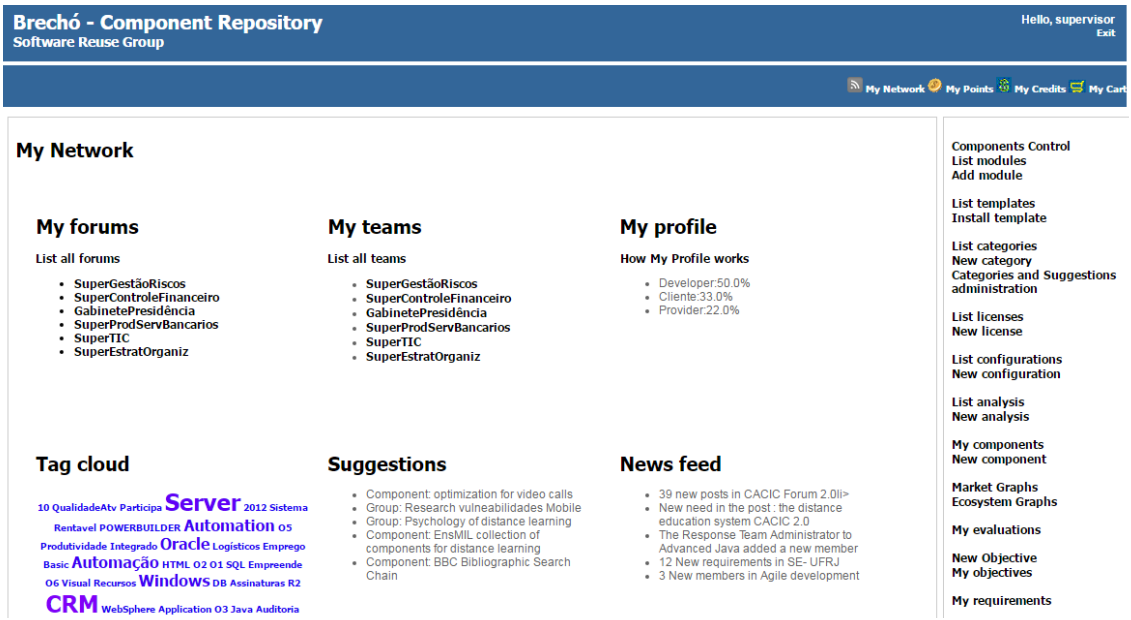

Figure 2. "My Network Panel" at Brechó-SocialSECO infrastructure

3.1. Artifact/Team Forum (demand registering)

The resource *forum* implemented in Brechó-SocialSECO is organized in three sections (Figure 3). The first section is used for general discussion, organized by 'topics', that are theme-free, i.e., users can talk to IT managers and architects about an artifact's function, ask for help, report a bug etc. This section allows a potential user to communicate with others before suggesting new demands. The second section intends to gather 'suggestions' from the organizational units. Anyone who is a member of a forum can contribute with suggestions for new demands, and a forum representative is responsible for managing these suggestions, e.g., department head. Finally, the third section allows a

unit representative to officially register ‘demands’. He/she registers these demands, or he/she can select a suggestion from the second section and make it into a demand.

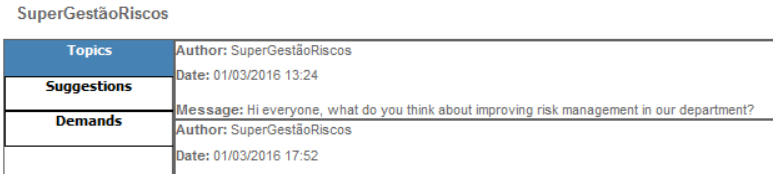

Figure 3. Forum’s discussions and sections

Users can ‘follow’ a registered demand since it involves the SECO community’s discussions and artifacts’ features. In addition, every message/suggestions are subjected to an evaluation system. An actor can vote for positive (+1) or negative (-1), regarding his/her opinion for each message. An actor can collect points for registering a suggestion that has received many positive votes. These points are used to leverage a unit’s collaborative level, as well as to reward users. Then, IT managers and architects can better understand the relevance of suggestions of the units, as well as all the demands emerging from the community. This extra information allows an IT management team to make better decisions on the selection and prioritization of demands, not only relying on the supplier’s word, but on the community’s ‘word of mouth’.

3.2. Team Creation

Actors can create teams and add other actors. Different types of teams can be created, e.g., ‘iOS users’, ‘Financial Sector’, ‘CRM extensions’ etc. Teams are defined as a type of user, which means that they keep a user profile. They can request demands, publish messages in forums, and evaluate artifacts on behalf of the team. Every team has administrators and members, and has a forum associated with it.

3.3. Actor/Artifact Profile (roles)

The SECO’s actors have many possible roles. There are internal and external actors playing as competitors, suppliers, users etc. (Lima et al., 2015). In a software acquirer, users and teams can check the proportion of actions they have performed within the SECO, being a producer, a consumer, or a simple user profile. This proportion is calculated according to their actions, such as publishing, downloading, or evaluating.

3.4. Trend Topics (tag cloud)

Tag clouds improve the way new trends and popular information are visualized in the SECO platform. It is a direct summarization of what is being discussed and evaluated by the community. The tag cloud consists of a data mining function that uses forums discussions, recommendations and demands, as well as teams’ information as inputs.

3.5. Suggestions and Recommendations (news feeds)

SECO platform can give an actor some suggestions and recommendations (news feeds). It is based on the forums he/she participates, as well as the profiles and demands he/she follows. Any forum member can post suggestions. Recommendations refer to teams, forums and/or artifacts of interest, also including information about the teams that an actor participates and new releases of artifacts. They aim to bring new information to an

actor (motivating to keep him/her updated), and search for further information. From the 'team' resource, actors can send requests to take part in an existing demand.

3.6. Demand and Solution Analysis

In order to organize a given SECO, a module for exploring and manipulating networks was developed and integrated with Brechó-SocialSECO. Gephi (https://gephi.org/) was chosen due to its support to statistics algorithms, integration with Brechó's technologies (Java/MySQL) and IDEs, community support, and ease of use. This module is named SECO-DSA (Santos, 2016) and is responsible for importing SECO elements from Brechó-SocialSECO's database (SECO network graph's nodes), as well as their different types of relationships (SECO network graph's edges). The SECO elements are artifacts (APP), technologies (TEC), demands (DEM), candidate solutions (CAN APP), and business objectives (OBJ). Since the most important TECs and OBJs are those that have more artifacts that depend direct/indirectly on them, and whose artifacts have a high number of licenses, we define our graphs as shown in Figure 4. In this work, a license is the authorization to use an asset, describing rights and limitations. Also, the asset's acquisitions are tracked using licenses.

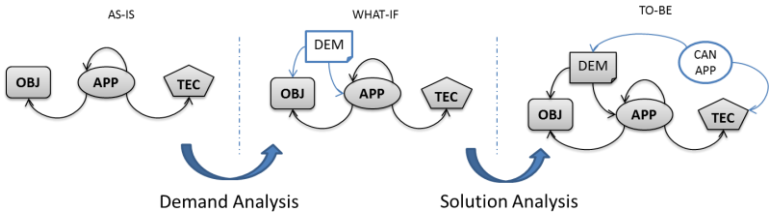

Figure 4. SECO elements at SECO-DSA Module

SECO-DSA considers three SECO platform configurations. *AS-IS* configuration presents the current artifacts, technologies and objectives in the acquirer's asset base. The *demand analysis* activity generates the *WHAT-IF* configuration based on simulation of new configurations after selecting potential demands. Finally, *TO-BE* configuration is produced by the *solution analysis* activity, in which candidate solutions are selected for each demand. All the relationships illustrated in Figure 4 are *dependencies* among SECO elements following the arrow's direction. For example, if selected, a demand X will affect artifacts Y and Z, and it will satisfy objectives A and B. Dependencies are weighted by the number of licenses carried from the origin node in this proposal.

The main goals of SECO-DSA are: (1) *to monitor objective synergy* (OBJ-SYN): allows IT managers and architects monitoring to which extent the existing artifacts in the asset base meet the acquirer's business objectives, before and after they perform *demand* and *solution analyses*; (2) *to monitor technology dependency* (TEC-DEP): allows IT managers and architects monitoring to which extent the existing artifacts depend on the technologies currently adopted by an acquirer, also before and after they perform *demand* and *solution analyses*; and (3) *to analyze the SECO platform*: use a Gephi plug-in to analyze how balanced the asset base is, i.e., how the level of sustainability of a SECO is. The metrics OBJ- SYN and TEC- DEP are measured in a 0-100% scale and allow creating a sustainability chart, as shown in Figure 5.1.

The chart indicates the sustainability status in four proposed quadrants: (a) *subsistence*: the acquirer is highly dependent on few technologies, and few objectives are satisfied by several artifacts at the same time; (b) *diversity*: there is a balanced

dependency on the adopted technologies, but most of the artifacts do not meet many objectives; (c) *fidelity*: the acquirer has high dependency on a small set of technologies, but most of the artifacts are contributing to the objectives; and (d) *sustainability*: it is the ideal situation where an acquirer has low technology dependency and high objective synergy. As exemplified in Figure 5, SECO-DSA workflow begins creating a graph that represents the current asset base registered at Brechó-SocialSECO and then calculates OBJ-SYN and TEC-DEP in order to get the values for the chart's axis. Next, the user can simulate the WHAT-IF configurations when demands are chosen and the SECO network graph is updated, followed by the chart's update. Once the user is satisfied with the selected demands, he/she can select candidate solutions through the same process, but using the filtering tab to choose the candidate solutions for each selected demand.

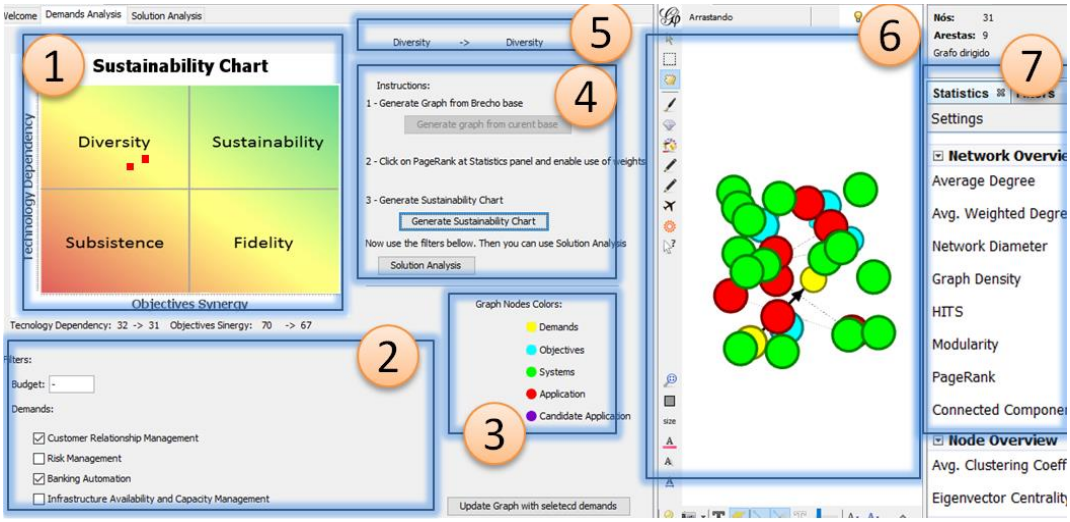

Figure 5. SECO-DSA Module as a Brechó-SocialSECO's Gephi plug-in

The main functions of the SECO-DSA panel are illustrated in Figure 5: (1) *sustainability chart*: is produced as a result of the values of OBJ- SYN and TEC- DEP for current SECO configuration and for the selected demands/solutions; (2) *filtering options*: allow users to select demands/solutions and update the network graph (inserting/removing nodes); (3) *node colors*: reserved area to explain the nodes' types; (4) *instructions and actions buttons*: present the instructions and buttons to guide demand/solution analysis, also (7); (5) *chart information:* shows information about the previous and current charts and notifies the user about the sustainability change caused by demand/solution selections; and (6) *graph visualization*: is a native Gephi's panel and shows the graph built from the asset base registered at Brechó-SocialSECO and also after user's selections. The outputs of SECO-DSA module support IT managers and architects to get insights from the asset base and to better choose demands and solutions.

4. Evaluation

In order to capture potential users' perceptions and suggestions about our infrastructure, experts in IT management activities conducted an initial feasibility study. It was based on the execution of tasks, as well as the fulfillment of feedback questionnaires. The participants should perform activities as IT managers evaluating possible demands and as architects assessing available solutions.

4.1. Planning

For this study, ten participants were invited. All of them worked in different private and public companies and had experience in the software industry and academia. An online questionnaire was applied with four sections: (1) *Informed Consent Form*: participant's rights and responsibilities; (2) *Characterization Form*: participants' background and experience with the study's concepts and similar tools; (3) *Execution Form*: eight tasks to be executed with our infrastructure, based on real data obtained from a Brazilian large banking organization; and (4) *Evaluation Form*: a questionnaire for the participant to provide feedback about ease of use and utility of existing mechanisms. Participants who live out of Rio de Janeiro used a PC remote access software. In any case, the observer took some notes. The two study goals (G) G1 and G2 were defined accordingly to the Goal-Question-Metric (GQM) paradigm (Basili et al., 1999), as described in Table 1.

Table 1. GQM for the feasibility study

Analyze	*Brechó-SocialSECO (and SECO-DSA module)*
With the purpose of	*characterizing*
With respect to	*the impact of SECO monitoring in the software acquirer's IT management activities (G1) as well as the tool usability (G2)*
The point of view from	*IT managers and architects*
In the context of	*demand and solution analysis*

The main question (Q) was: "*Are the participants able to realize the impact of SECO monitoring in the software acquirer's IT management activities for demand and solution analysis, regarding effectiveness and efficiency?*" (Q0 for G1). In other words, how the socialization mechanisms and the asset base analyses help IT managers to monitor the SECO. This perception was measured by the answers given to the study's tasks based on the following metrics (M):

M1: *Effectiveness* measures the relation between the results and the objectives:

$$Effectiveness = \frac{number\ of\ correct\ answers}{total\ number\ of\ questions}$$

M2: *Efficiency* measures the relation between the results and the resources:

$$Efficiency = \frac{number\ of\ correct\ answers}{time\ taken\ to\ participate}$$

Questions Q1-Q10 for G2 were based on Nielsen's heuristics (Table 2). To answer the questions derived from such heuristics, a 5-point scale were adopted: Totally Disagree, Disagree, Not Agree Nor Disagree, Agree, and Totally Agree. For each question, the percentage of answers was analyzed by means of a frequency chart.

4.2 Execution

The execution was individually performed with five participants in June 2016. Two of them participated in person and the others remotely. Participants signed the informed consent form and answered the characterization form. Next, each of them received an explanation about SECO concepts, as well as a Brechó-SocialSECO tutorial. Finally, each participant received the execution form, and the evaluation form at the end.

4.3. Results

Regarding the academic education, two participants reported to hold a PhD degree, two had a Master degree and one had a Bachelor degree. Figure 6 shows the experience of

each participant regarding the concepts related to the study (time in years). Each concept had at least one participant with more than 4 years of experience (up to 20 years). It also refers to the experience of each participant with related tools. Collaboration tools were the most popular (at least 10 years of experience).

Table 2. Nielsen's heuristics and questions (Nielsen, 1993)

ID	Principle	Question
Q1	*Visibility of system status*	Does the system inform what is happening?
Q2	*Match between system and the real world*	Does the system explore the user language?
Q3	*User control and freedom*	Is the system easy to interact and present clear outputs?
Q4	*Consistency and standards*	Do different situations or actions represent the same thing?
Q5	*Error prevention*	Does the project predict errors instead of using messages?
Q6	*Recognition rather than recall*	Do the screens use metaphors (instructions memorization)?
Q7	*Flexibility and efficiency of use*	Does the system meet experienced users (shortcuts)?
Q8	*Aesthetic and minimalist design*	Is the information summarized and complete?
Q9	*Help users recognize/diagnose/recover from errors*	Are problems and solutions indicated?
Q10	*Help and documentation*	Are there simple and objective manuals?

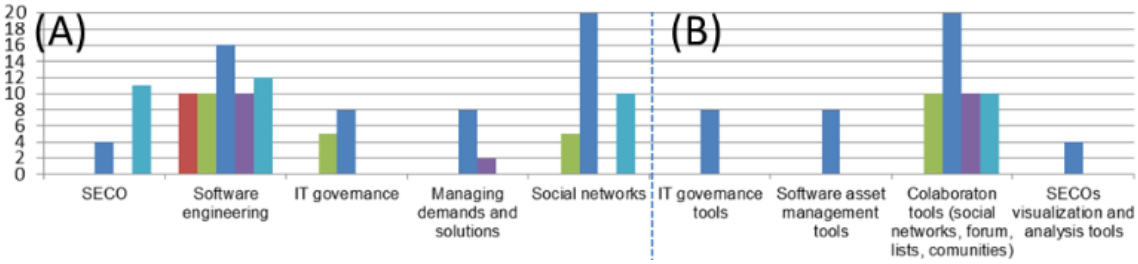

Figure 6. Participants' experience with concepts (A) and related tools (B) (time in years)

After charactering the participants' profile, results were analyzed based on their answers, duration of activity, and feedback provided in the evaluation form. The values of M1 and M2 are found in Table 3. For example, for task (2) "*For which combination of demands does the objective synergy get better?*", the participants had to simulate selecting CRM demand, Banking Automation demand, or both, and only one of them answered wrong. As shown in Table 3, the effectiveness to perform IT management activities for demand and solution analysis was positive with our infrastructure in the selected and applied context. However, the efficiency was not so high with its use. In turn, the usability measures shown in Figure 7 suggest that some mechanisms need improvements, especially because of weak evaluations obtained for the heuristics *Error prevention*, *Flexibility and efficiency of use*, and *User control and freedom*.

Table 3. Effectiveness and efficiency measures

Participant	Number of Correct Answers	Time (min)	Effectiveness	Efficiency
P1	6	65	0.750	0.092
P2	6	42	0.750	0.143
P3	5	35	0.625	0.143
P4	6	50	0.750	0.120
P5	5	51	0.625	0.098
Average	5.6	48.6	0.700	0.119

Finally, the qualitative feedback on the user's impressions was analyzed. For the question "*What were the biggest difficulties in performing the proposed tasks?*", some participants pointed out: to perform the steps required for the sustainability chart update (and to ensure that the values were also updated); to understand the sustainability moves since the points were too close and had no marks to facilitate the reading; and to find the list of members of a team. For the question "*If you are missing resources, please write*

here:", some participants mentioned: storage of the historical series of SECO network graphs to evaluate different combinations; and automatic comparison of different SECO configurations (scenarios). Considering the question "*In your opinion, what are the most useful resources?*", some participants highlighted: sustainability chart visualization and simulations of selection's combinations; forum's structure (by specific themes and by discussions/suggestions); and trend topics mechanisms (tag cloud). Finally, the question "*Do you have any suggestion?*" allowed us to collect the following points: usability improvement for the Gephi plug-in (P2: *Richer graphical resources for analysis of the selected items, also tracking the history*) and presentation of relationships in forums (P2: *I missed graphical resources from social networks to support discussions of topics*).

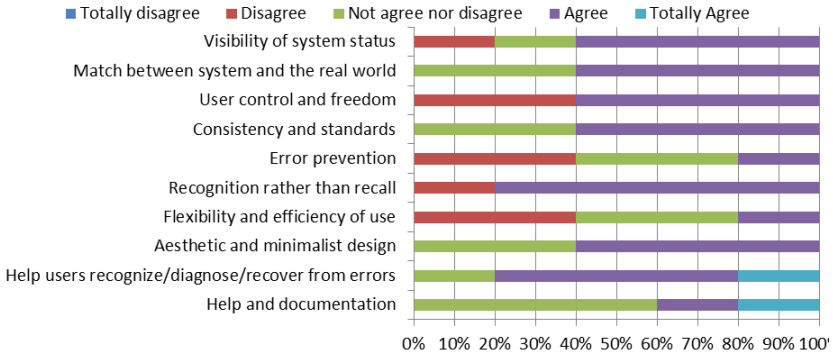

Figure 7. Usability measures

In general, the participants understood the role of our infrastructure and found the proposal useful, considering the mechanisms provided by Brechó (e.g., forum and tag cloud) and by Gephi (e.g., sustainability chart). The most popular suggestions were to improve the steps to generate the chart and to show the simulation history in order to compare and select different options. This study revealed not only corrections (e.g., bugs in button, and resetting options) but also new resources (e.g., to visualize the historic data of simulations) to be implement aiming to better monitor SECOs.

There are some limitations of our study, e.g., the low number of participants, the lack of a control group for comparisons, the lack of more and diverse scenarios, and the subjective collection of the research impressions. As threats to internal validity, we highlight: the infrastructure can influence the user's experience, and the execution form can interfere in the participant understanding. In turn, as threats to external validity, we point out: data are specific for an acquirer's reality, participants not necessarily work in the same company, and it is not possible to represent all the SECO context in a scenario.

5. Final Considerations

Current business environments and communities are continuously evolving. This reality contributes to changes in the organizational processes, affecting business objectives and demands/solutions selection. This fact motivated us to research how SECO perspective affects IT management activities, more specifically demand and solution analysis, aiming at helping IT managers and architects to make acquisition decisions. This paper presented an infrastructure to support socialization, monitoring and analysis of SECO named Brechó-SocialSECO. The asset base's socialization and analysis mechanisms were built upon a component library to aid the SECO monitoring. The main motivation is the lack of initiatives that combine market data with those obtained from the software

asset database, as reported in a literature review on SECO (Manikas & Hansen, 2013). To evaluate our proposal, we conducted an initial feasibility study to observe the impact of SECO monitoring in the software acquirer's IT management activities as well as the tool usability. As a result, the effectiveness to perform IT management activities for demand and solution analysis was positive with our infrastructure, in the selected and applied context. The same happened for the most usability concerns. However, the efficiency was not high with the use of the tool.

In summary, our study brings initial indications that our proposal helps IT management teams in their daily activities. Our contributions are: (a) implementation of an infrastructure for socialization, monitoring and analysis of SECO, as well as the sustainability indicators that are critical for IT management activities; and (b) evaluation of some mechanisms of our proposal with practitioners. Future work involves a complete evaluation of Brechó-SocialSECO. The level of collaboration resulted from the implemented social mechanisms, as well as the relevance of suggestions turned into demands will be investigated. Despite several limitations of this work, we believe that a long road is right ahead and the topic can contribute to the Software Engineering area regarding the treatment of business and social challenges discussed by the research and industrial communities.

Acknowledgement

We thank CAPES and CNPq (Proc. No. PDJ 150539/20016-9) for the financial support.

References

Basili, V.R., Shull, F., Lanubile, F. (1999) "Building Knowledge through Families of Experiments". IEEE Transactions on Software Engineering 25(4):456-473.

Finkelstein, A. (2014) "Rethinking Software: Business Change and the Consequences for Software Engineering", In: 22nd IEEE RE, Keynote, Karlskrona, Sweeden.

Gartner (2007) "Evaluating Global-Class Software Ecosystems". Available at: <https://www.gartner.com/doc/506608/evaluating-globalclass-software-ecosystems>.

Goeminne, M., Mens, T. (2010) "A Framework for Analysing and Visualising Open Source Software Ecosystems", In: Joint ERCIM EVOL & IWPSE, Antwerp, Belgium, 42-47.

Lima, T., Santos, R., Werner, C. (2015) "A Survey on Socio-Technical Resources for Software Ecosystems", In: 7th ACM MEDES, Caraguatatuba, Brazil, 72-79.

Manikas, K., Hansen, K.M. (2013) "Software Ecosystems – A Systematic Literature Review". The Journal of Systems and Software 86(5):1294-1306.

Nielsen, J. (1993) "Usability Engineering". Cambridge: Academic Press.

Pérez, J., Deshayes, R., Goeminne, M., Mens, T. (2012) "Seconda: Software Ecosystem Analysis Dashboard", In: 16th CSMR, Szeged, Hungary, 527-530.

Santos, R. (2016) "Managing and Monitoring Software Ecosystem to Support Demand and Solution Analysis". PhD Thesis, COPPE/UFRJ, Rio de Janeiro, Brazil, 246f.

Seichter, D. *et al.* (2010) "Knowledge Management in Software Ecosystems: Software Artefacts as First-class Citizens", In: 4th ECSAW, Copenhagen, Denmark, 119-126.

Werner, C. *et al.* (2009) "Towards a Component and Service Marketplace with Brechó Library", In: IADIS International Conference on WWW/Internet, Rome, Italy, 567-574.