Application of Open Coding using the Grounded Theory Method to Identify the Profile of Information and Communication Technology Companies in the State of Pará from Brazil

Elziane Monteiro Soares and Sandro Ronaldo Bezerra Oliveira
Graduate Program in Computer Science, Institute of Exact and Natural Sciences, Federal University of Pará, Belém, Pará, Brazil

Keywords: Experimental Software Engineering, Grounded Theory, Open Coding, Customer and Market, MOSE.

Abstract: The use of methods inserted in the context of Experimental Software Engineering as the experimental study, case studies, opinion surveys and controlled experiments, has intensified in recent years. Considering this context, this study aims to present the application of the Open Coding proposal in the experimental method of Grounded Theory (GT), in order to contribute to the definition of the profile of Information and Communication Technology (ICT) companies in the State of Pará from Brazil, since the results generated in this work favor the development of the next steps of the GT method. The context explored happened under the perpescence of the MOSE model, more specifically in the Customer and Market (CM) dimension, once the investigation was based on information related to the concepts present in this dimension, aiming to know the form that occurs the development of the CM competence in the daily life of each company participating in this research.

1 INTRODUCTION

According to SOFTEX (2016), the changes that are occurring in business environments have motivated companies to modify their organizational structures and production processes, moving from the traditional view based on functional areas towards customer-centric process networks. Therefore, for companies to improve the quality of software products and related services, such as the processes of production and distribution of software, leads to obtaining competitive advantage in the current market of software development companies.

In view of the above, this work aims to contribute, from the application of the open coding proposed in the experimental method GT, to define the ICT companies’ profile in the State of Pará, from Brazil, since the results generated in this work favours the development of the next steps that compose the GT method.

To structure that perception, this work will provide a consistent view better of the way these companies operate under CM’s perspective, and thus to contribute to making the reality of the companies of the State of Pará in the current market being known by all, in order to foster investments of public organizations and / or private in their projects, aiming potentially to increase the competitiveness of the ICT companies of the State of Pará in the brazilian market.

The MOSE model (Rouiller, 2017) was selected as a reference for the development of the research in order to deal with the lack of visibility of the companies from Pará in relation to other states about the competitive market, by treating in detail the procedures to maintain a good organizational environment, quality management of products and / or services, good relationship with customers, competitiveness of the company in the market and innovation, among other factors. MOSE is composed of five competences dimensions: Society and Sustainability, Human Talent, Quality, Customer and Market and Innovation, however this work focuses on the Customer and Market (CM) competence, which is seen as substantial for a company which is starting or already has a few years of experience in the market, helping to remain competitive and innovative in its services offered and / or goods produced, while other skills can be worked after the company acquires stability in the market.

The definition of the context in this work, based on the information about the MOSE model, favored the development of the Grounded Theory
methodology, since it requires a well-defined context 
of the phenomenon to be studied, since this context provides the basis for definition problem or the issues that guided the research.

According to Bandeira-de-Melo and Cunha (2003), the method of grounded research is a qualitative research method that uses a set of systematic data collection and analysis procedures to generate, validate substantive theories about essentially social phenomena or comprehensive social processes.

In the quest to create a theory that more accurately portrays the profile of companies in the scenario of Pará, the Grounded Theory method was selected precisely because the empirical data base provides a closer approximation to reality by allowing behavior analysis to be performed with patterns and more appropriate procedures for the discovery of the theory.

The core element of Grounded Theory is coding, which is the process of data analysis (Conte, 2009). Coding refers to the procedures used to label and analyze the data collected (Glaser and Strauss, 1967). Therefore, the interpretation of the data is the essence of qualitative research, whose function is to develop the theory, helping in the decisions about which data will be worked (Flick, 2004).

To assist in the collection of empirical data, a survey was developed and applied in ICT companies in the state of Pará in order to collect information related to the Customer and Market competence area, which is part of the MOSE model, in order to know the form that the development of CM competence occurs in the daily life of each one.

Thus, the use of an experimental study is applied precisely for the purpose of discovering something unknown or to test a hypothesis (Zendler, 2001 apud Cruzes, 2007), where the population, the research method, the method of collection and the form of analysis of the data collected.

Thus, Experimental Software Engineering (ESE) is the subject of Software Engineering that investigates theories, methods and techniques by experimentation. Typical approaches to experimental evaluation are case studies, opinion surveys and controlled experiments (Wohlin et al., 2000).

In view of the above, the benefits of the application of experimental studies are related to obtaining results that justify the use or not of a technology, based on some indication that it can contribute to the improvement of product quality or software process. The use of a process to conduct the studies contributes to the control and execution of the same, validation, generalization of results and replication, since the theory generated must be able to be applied in another context.

In addition to this introductory section, this paper is organized as follows: Section 2 presents some related works on the study carried out, Section 3 briefly presents the survey conducted, Section 4 presents in detail the open coding of the Grounded Theory method, Section 5 presents the evaluation used in this work and Section 6 discusses the results obtained with the research and possible future work.

2 RELATED WORKS

From a bibliographical review in the literature, in order to find previous works that have themes related to the one proposed by this work, some works were analyzed, which will be described next.

A study conducted by Santos (2011) showed the importance of analyzing the human aspects, as these can influence the success of a Software Process Improvement (SPI) program, in order to obtain a better understanding of these influences. Therefore, in this work a research was presented on the influence of human aspects in SPI programs, in order to deepen the understanding of these influences, considering the human aspects, and with the purpose of helping these programs in future successes. For data collection, we chose to use semi-structured interviews and for data analysis the researchers used the Grounded Theory method, precisely to create the substantive propositions about the influence of human aspects.

Melo (2008) reports that studies show that activities that are inadequate to the requirements engineering process are a critical factor for the success of software projects. Based on this, he presented an empirical research carried out with Requirements Engineers, aiming at a better understanding of the importance and difficulties existing during the requirements elicitation phase. Thus, it provides the main characteristics and necessary profile, which involves a good requirements elicitor, as well as aid to the project managers during the selection of the most appropriate profile of their team to carry out the elicitation task, obtaining, therefore, the better results for software design.

Pereira (2001) developed a research with the entrepreneurs the western of State of Paraná from Brazil, based on a qualitative study, with qualitative approaches to the analysis of the data collected. The general objective of this work was to analyze the profile of the successful entrepreneur from western Paraná and to identify the influences of the profile in
its business in Mercosur. The data collected came from the interviews with entrepreneurs selected from a population universe from industry, wholesale and retail.

In his work, Pereira was able to conclude that the interviewees live in constant professional instability, due to the economic and political factors coming from the international market, mainly from countries with border connections.

The work presented show approaches to data collection and analysis with the aim of creating approaches for improvement during the software development stage, either from analyzes on the influences of human factors and even support mechanism, taking into account the reports of people who participate in the stages in the development process. In other works, the objective was to create a profile for requirements elicitors and entrepreneurs, thus following another research context.

Unlike the related works, this research considers using the Grounded Theory research method to create a replicable approach of a process for the identification of the profiles of companies focused in the area of information technology, from the perspective of Customer and Market competence area belonging to MOSE model.

3 SURVEY

As a way to reach the objective of the work, a survey was developed and applied in information and communication technology companies in the state of Pará in order to collect information related to the Customer and Market competence area, which is part of the MOSE model, in order to know the development of these skills in the daily life of each one.

For the definition of the subjective questions that compose this questionnaire, an exploratory research was carried out based on the bibliographic survey about the assets of the MOSE model, more specifically in the analysis of the goals and expected results of the CM competence dimension that is part of this model, selected to support the research of this research.

The survey was applied in order to collect qualitative data by structured interviews and applied in a face-to-face manner in companies. The application was made with the representatives of each ICT company participating in this research, which had as target audience 14 companies, of which 11 were the sample of participants, all of them operating in the state of Pará, specifically those located in the capital of state of Pará, with a time in the ICT market between 4 and 30 years, but the people selected as respondents were those who have knowledge of business strategies, customer understanding of opportunities, that is, a broad management vision of the products and / or services provided by the companies in question, which favors the achievement of more significant results, within the context investigated, in the competencies chosen in the MOSE model. All the companies are registered in ParaTIC - Association of Companies the State of Pará of Software and ICT (http://www.paratic.com.br).

The execution and the results obtained initially with the survey application were published in (Soares and Oliveira, 2018).

From the production of the data set, obtained with the application of the survey, the process of analysis and codification was started, belonging to the Grounded Theory research method. Coding can be interpreted as an analytical process from which data are broken, conceptualized and integrated to form a theory (Strauss and Corbin, 1998).

Therefore, in order to guarantee the application of the three types of analysis procedures proposed in the Grounded Theory method, a set of steps were defined and executed, and in this work only the Open Coding context is discussed.

4 THE GROUND THEORY FROM OPEN CODING

In the search to construct the substantive theory, in relation to the profile of the ICT companies of the State of Pará, the Grounded Theory (GT) research method was applied considering only the CM competence.

It is emphasized that the Grounded Theory is a methodology of qualitative research, endowed with mechanisms, whose purpose is to aid in the development of theories from data collected in a certain empirical reality.

In order to support and streamline the procedures for coding and interpreting the data in the application of the GT method, Atlas.ti (http://www.atlasti.com) was selected. This tool has extensive mechanisms for managing, extracting, comparing, exploring and reassembling significant pieces of large amounts of data in a creative, flexible and systematic manner that effectively supports the interpreter (researcher) in qualitative research.

Although Atlas.ti facilitates many of the activities involved in the analysis and interpretation of
qualitative data (particularly selection, indexing / coding and annotation), its purpose is not to generate automatic interpretation of the text, since it is the researcher to carry out the creative and systematic work of seeking to conceptualize and associate the data.

Therefore, to facilitate the understanding and development of the application of the GT method, the following steps were defined for the present work:
1) Transcript of the interviews, organized by competence goals and companies,
2) Identification of the codes based on the competence goals and questions applied in the Survey,
3) Definition of the main network between the codes,
4) In each question identify snippets from the codes already defined,
5) Identification of new codes for each specific competence goal,
6) Construction of a new network between the new codes and the specific codes of the competence goal (primary codes),
7) Mark the excerpts from the texts, referring to the transcriptions of the interviews with the new codes,
8) Identification of common factors among codes,
9) Definition of the substantive theory of the profile of companies in the capital of Pará.

In order to ensure that the data coding steps proposed in the Grounded Theory were achieved, a relationship was established between the previously listed steps and the three coding levels of the method, being it open, axial and selective, as may be shown in Table 1.

Table 1: Relationship between the coding type and the steps defined in this work.

<table>
<thead>
<tr>
<th>Coding Types</th>
<th>Steps defined for this Work</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open</td>
<td>Step 1</td>
</tr>
<tr>
<td></td>
<td>Step 2</td>
</tr>
<tr>
<td></td>
<td>Step 3</td>
</tr>
<tr>
<td></td>
<td>Step 4</td>
</tr>
<tr>
<td>Axial</td>
<td>Step 5</td>
</tr>
<tr>
<td></td>
<td>Step 6</td>
</tr>
<tr>
<td></td>
<td>Step 7</td>
</tr>
<tr>
<td>Selective</td>
<td>Step 8</td>
</tr>
<tr>
<td></td>
<td>Step 9</td>
</tr>
</tbody>
</table>

In Step 1, the interviews were recorded in full, organized by competence goals and companies, which provided the initial analysis in the CM context in the companies, with the results presented in the paper about the application of the survey (Soares and Oliveira, 2018), but in this work these data will be further refined by the coding of the GT method.

With the conclusion of Step 1, which dealt with the transcription of the interviews, organized by competence goals and companies, Step 2 was started, in order to continue the other activities listed in the steps of development of the GT method.

To help analyze the data and identify the codes based on the competence goals and questions applied in the survey (procedure related to Step 2), the documents with the transcripts were imported into the Atlas.ti tool, precisely to assist in the manipulation of the qualitative data.

The documents imported into the tool, with the transcripts, were separated by competence goal, since the CM competence dimension has nine competence goals (Table 2), so nine documents were created, with their respective research question, in these documents, letters of the alphabet were attributed to the companies that participated in the research, as a form of representation in all nine documents, thus ensuring anonymity (Soares and Oliveira, 2018).

Table 2: Goals and expected results of the Customer and Market (CM) competence dimension.

<table>
<thead>
<tr>
<th>CM Competence Goal</th>
<th>Expected results</th>
</tr>
</thead>
<tbody>
<tr>
<td>CM.1 Agreements with customers are established and managed</td>
<td>Agreements with customers are expected to be established and managed. It is also expected to be observed whether goods and services are being delivered and whether customers are making their commitments to signed agreements.</td>
</tr>
<tr>
<td>CM.2 Actions to promote the consumption of goods and / or services are carried out</td>
<td>It is expected that actions are taken to promote the consumption of goods and / or services of the business unit (in some environments are called sales) and that these are preferably registered.</td>
</tr>
<tr>
<td>CM.3 Customer service is performed</td>
<td>In this goal, Customer Service means receiving your request, processing it, asking questions and arranging for the delivery of goods and services purchased. This goal addresses aspects related to ensuring that agreements and contracts with the customer are met.</td>
</tr>
<tr>
<td>CM.4 Relationship with customers is carried out</td>
<td>The business unit is expected to observe the satisfaction of its customers in relation to the experiences of consumption of the goods and services, and that actions are developed to solve problems that may tarnish its image.</td>
</tr>
</tbody>
</table>
Table 2: Goals and expected results of the Customer and Market (CM) competence dimension. (cont.)

<table>
<thead>
<tr>
<th>CM.5</th>
<th>The portfolios of goods and / or services are defined and maintained</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>It is expected that the main goods and / or services offered by the business unit will be organized in the portfolio. It should be emphasized that each good and service must be specified in the portfolio, considering its target audience, its categorization and prioritization, competitors and among others.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CM.6</th>
<th>Prospection of new demands are detailed considering the assets and / or services of the portfolio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Customers (internal or external to the company) are expected to be categorized into profiles and that prospecting, retention and conversion actions of new customers are carried out. It is also necessary to make a detailed analysis of the generation of new demands, now considering specific time goals for the goods and services defined in the portfolios.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CM.7</th>
<th>Approaches to customer service are established and maintained</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>It is expected that channels will be established to meet customer requests and deliver the goods and / or services produced by the business unit.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CM.8</th>
<th>Relationship approaches with customers are established and maintained</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>It is expected that channels of relationship and communication with customers will be established and maintained (and market-focus, where relevant). There should also be studies to see if the form of customer relationship (and market-focus) is being effective and adjustments should be made when this communication has no positive effects on the business.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CM.9</th>
<th>Incidents are recorded, analyzed and preventive actions are carried out</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>It is expected that incidents of delivery of the goods and / or services will be handled in time to generate wear and tear with the customer. In addition to the need to solve these incidents, these should be recorded and periodic reviews should be made with the intent to prevent them.</td>
</tr>
</tbody>
</table>

Also in Step 2, in order to identify the codes, a detailed analysis line by line was made in the sentences transcribed from the interviews, precisely to search for the key words or expressions, which allowed generating the preliminary data and their possible relations. Basically, during the open coding the researcher separates the data into distinct parts, examines and compares them rigorously for similarities and differences.

Therefore, this step must involve a careful and contextualized reading of the transcribed lines, in this way it allows to generate codes, from a part of the text or in some moments of a word for assignment (conceptualization) of a code name that reflects the text segmented (Maciel, 2011).

Thus, from this detailed analysis, the first codes of this work, generated by the researcher’s vision and interpretation, appeared.

With the emergence of the first codes, it was possible to define the main network between the codes, corresponding to Step 3. The network visualization allows expressing relations between codes, and was created observing the interrelation or interdependence between the competences goals, that is, the relationship established between the codes in this work considered the structure provided in the MOSE. Figure 1 shows the structure of interdependence between competence goals.

In MOSE, the relationship established among the nine competence goals considers the recommendations required by profile, based on the size and number of employees that the company obtains, in order to meet the diversities of companies in relation to the size and the number of employees, the competence goals are allocated considering three profiles: Small (PEQ), Medium (MED) and Large (GRD), according to relationship shown by Figure 1.

Step 4 proposed in each question to identify excerpts from the codes already defined and was executed in parallel to Step 2, since the generated codes reflect the texts (excerpts) considered important from the perspective of the researcher, in the explored context, with the association between a code and a citation, in order to capture meaning in the data and also serve as identifiers for specific occurrences in the data.

Thus, with the completion of the steps that compose the open coding, the first results were obtained, being these results primordial to the follow up of the next steps referring to the axial and selective
coding. It is remarked that the results of the axial and selective steps are under analysis for publication in the annals of another conferences.

4.1 Results of the Grounded Theory Method for Open Coding

This section presents the results obtained through analyzes performed on the data, that is, resulting from the open codification process, with its due descriptions and foundations, which formed the initial basis for the construction of the substantive theory of definition of the ICT companies profile, considering CM competence.

The main network, seen in Figure 2, emerged as a result of the microanalyses performed in the transcribed texts, with the extraction of the initial codes, occurring in the steps related to the open coding. The initial set of codes refers to the concepts that characterize the participants' responses to the CM competence goals addressed in each question formulated in the survey.

It is also worth noting that the organization of the codes in the diagram does not represent a pre-established order of importance, since their hierarchical order observed in the figure does not express their degree of relevance, for example, the code incident handling, of the diagram does not have a high level of importance compared to others, this code was only positioned to improve the visual presentation field of the diagram, as well as all others present in the primary code diagram, since the main point of this result is the presentation of codes and how the relationship between them occurs.

For Bandeira-de-Mello and Cunha (2003), the codes generated in the open coding should be classified as: a) first order codes, directly associated to the citations (called "in vivo" codes), b) and abstract or theoretical codes associated with other codes. The connectors used to associate the code generated in the Figure 2 were of type “is associated with”, where the source code and destination code have specific concepts related; while in the type “is part of”, the source code is a part that composes together with other parts of destination code.

For a better understanding of the codes presented in the diagram (seen in Figure 2), Table 3 was developed with the purpose of presenting a detailed description of the first codes extracted in the research, with the following specifications: (a) MOSE competence goal, which CM goal in which the code originated, (b) code, presents the name that was assigned to the code in the analysis, (c) grounded (degree of reasoning), indicating the number of times the code was applied, (d) density, indicating the number of links between the code and other codes, (e) related codes, where the names of the codes with which there is the relationship in the diagram are presented.

Posteriorly, a structure graphic, with occurrences of citations, was developed to each code, in order to point out text snippets that provide proper evidence to main concept related to code, which consequently aided in the accounting of the quantity of time that code was applied (grounded).

Thus completing the open coding phase, having already the initial codes estructured in network, it can proceed to next steps that compose the GT method, which is axial and selective coding, developed in future works.

Figure 2: Primary code Network.
Table 3: Primary code specifications.

<table>
<thead>
<tr>
<th>Customer and Market Competence Goal (MOSE)</th>
<th>Code</th>
<th>Grounded Density</th>
<th>Related Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>CM1 Contracts and agreements established</td>
<td>21 1</td>
<td>Portfolio of Goods and Services</td>
<td></td>
</tr>
<tr>
<td>CM2 Promote Goods and Services</td>
<td>19 1</td>
<td>Prospecting New Customers</td>
<td></td>
</tr>
<tr>
<td>CM3 Make an Attendance</td>
<td>12 1</td>
<td>Maintain an Attendance</td>
<td></td>
</tr>
<tr>
<td>CM4 Conduct Relationship</td>
<td>15 1</td>
<td>Maintain Relationship</td>
<td></td>
</tr>
<tr>
<td>CM5 Portfolio of Goods and Services</td>
<td>11 3</td>
<td>Contracts and agreements established, Prospecting New Customers, Maintain an Attendance</td>
<td></td>
</tr>
<tr>
<td>CM6 Prospecting New Customers</td>
<td>13 3</td>
<td>Promote Goods and Services, Portfolio of Goods and Services, Maintain Relationship</td>
<td></td>
</tr>
<tr>
<td>CM7 Maintain an Attendance</td>
<td>13 4</td>
<td>Make an Attendance, Portfolio of Goods and Services, Maintain Relationship, Incident Treatment</td>
<td></td>
</tr>
<tr>
<td>CM8 Maintain Relationship</td>
<td>11 4</td>
<td>Conduct Relationship, Prospecting New Customers, Maintain an Attendance, Incident Treatment</td>
<td></td>
</tr>
<tr>
<td>CM9 Incident Treatment</td>
<td>15 2</td>
<td>Maintain an Attendance, Maintain Relationship</td>
<td></td>
</tr>
</tbody>
</table>

5 EVALUATION

In order to guarantee the veracity of the results, evaluations were carried out based on the previous knowledge of an expert in the area of experimental software engineering, being certified as a consultant-implementer and leading evaluator of the MPS.BR, CMMI, CERTICS, MEDE-PROS, MOSE, among others, working for more than 5 years in deployments of models to improve the process or software product in organizations from different regions of Brazil.

The evaluations performed on the results occurred at the end of each coding step, guided by specific criteria proposed by Strauss and Corbin (1998) to evaluate substantive theories based on data, which are:

- Fit - the theory must be appropriate to the substantive research area and must correspond to the data,
- Understanding - the theory makes sense for professionals in the area studied,
- Generality - the theory must be abstract enough to serve as a general guide without losing its relevance,
- Control - the theory acts as a general guide and enables a person to fully understand the situation.

In order to standardize and organize the evaluation, a scale was defined and applied to the criteria presented previously, in order to assign a classification for each doubt or inconsistency found in the coding steps, specified below:

- NA (No Attendance): indicates that a problem has been found in an item that, if not changed, will compromise the theory,
- PA (Partial Attendance): indicates that a coding problem has been found which should be changed,
- FA (Full Attendance): indicates that the analysis performed on the data is complete and contributes to the theory.

In general way the expert evaluated the research results of open codification as a full attendance with fit, understanding, generality and control because the research report had consistency about the MOSE assets. This way, the researcher started the axil and selective codding using the results obtained from open coding.

6 CONCLUSION

The objective of this study was to present an application of the open codification proposed in the Grounded Theory experimental method carried out in the data that were collected by the application of the survey with the participation of 11 respondents, with time in the ICT market between 4 and 30 years. The open codification was carried out in order to contribute to the definition of the profile of Information and Communication Technology companies in the capital of Pará from Brazil, as it generated results that favor the development of the next stages of the GT method. The explored context occurred under the perpescence of the MOSE model, more specifically in CM dimension.

Thus, it is intended as future work to further refines this research from the application of the other
stages, axial and selective, that makes up the GT research method. This fact, therefore, will contribute to the creation of the central and substantive theory regarding the profile of the companies that operate in the ICT technology market of the State of Pará.

ACKNOWLEDGEMENTS

The authors would like to thank the Dean of Research and Postgraduate Studies at the Federal University of Pará (PROPESP/UFPA) for providing financial support through the Qualified Publication Support Program (PAPQ).

REFERENCES


Melo, R. A. de, 2008. Um estudo qualitativo do perfil de elicitador de requisitos orientado a obtenção do sucesso de projetos de software. Dissertação de Mestrado – PPGCC, UFPE.


Santos, D. V. dos, 2011. Uma pesquisa sobre a influência dos aspectos humanos em programas de melhoria de processo de software. Dissertação (Mestrado em Informática) - Universidade Federal do Amazonas, Manaus.


