**PAPER • OPEN ACCESS**

# Data Consolidation in Global Software Development Projects: A Grounded Theory

To cite this article: A Subbarao and M N Mahrin 2021 *J. Phys.: Conf. Ser.* **1793** 012043

View the article online for updates and enhancements.

# Data Consolidation in Global Software Development Projects: A Grounded Theory

**A Subbarao[1] and M N Mahrin[2]**

[1] Information Technology Unit, Faculty of Management, Multimedia University, Persiaran Multimedia, 63100 Cyberjaya, Malaysia

[2] Department of Informatics, Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, 54100 Jalan Sultan Yahya Petra, Kuala Lumpur, Malaysia

anusuyah.subbarao@mmu.edu.my, mdnazrim@utm.my

**Abstract**. Effective coordination is a crucial aspect of successful Global Software Development (GSD) projects. Limited studies have examined coordination strategies and their related indicators. Therefore, this study focuses on assessing the coordination processes that require specific strategies and related indicators that can contribute to effective coordination. Thus, the main aim of this research is to consolidate the coordination strategies and the related indicators that were extracted from two different sources namely, Systematic Review and Semi-structured interview. Grounded Theory was used for the consolidation of these two sources. The results of this study are a set of coordination strategies and related indicators that will be used to formulate the evaluation to assess the effectiveness of the coordination processes in GSD projects.

## 1. Introduction

Globalization is going on very rapidly in modern world technology and has brought significant transformation into software development businesses. When software is being developed across the countries, this approach is called Global Software Development (GSD) [1]. Many software organizations are shifting their strategies towards GSD approach due to many benefits such as access to large pool of competent developers, less time taken for software development, reduce software development cost, less time taken to market the software product and to produce better quality software compared to traditional way of development [2].

Despite enjoying the benefits, IT industries that adapted GSD are suffering from many challenges [3]. Babar and Leicester (2014) also have grouped these challenges into three different themes, namely, coordination, collaboration, and communication mechanisms [4]. Although all three themes are playing essential roles in the GSD environment, coordination is the main focus of this research. Coordination Theory defines coordination as working together process and managing interlink between activities to achieve a specific goal [5]. Literature shows that many difficulties are occurring in the GSD environment due to lacking coordination between the collocated and distributed team members, and one of the prominent issues is ineffective coordination.

Ineffective coordination in GSD has caused many problems such as delay in coordination, difficulty in organizing tasks, misinterpretation of tasks, extra time needed for coordination due to mismatches in goals, and others. Empirical research indicates that achieving a state of effective coordination is a crucial success factor for GSD projects. Although researchers are producing many

distinctive solutions for coordination in GSD, to confirm effective coordination in GSD remains a big challenge. Thus, this has motivated the researcher to propose an evaluation model to assess the effectiveness of coordination processes by incorporating coordination strategies and related indicators for each coordination processes in GSD projects [13].

A research framework which consist of four phases [16] were followed to formulate the model and one of the steps in formulating the model is to consolidate the coordination strategies and the related indicators that were extracted from two different sources namely, Systematic Review and Semi-structured interviews from the GSD project practitioners. Hence, this study was conducted to achieve this objective.

This paper is organized as the following. Section 2 discusses the literature review of the selected methodology. Section 3 discusses the activities involved in the methodology of this study. Section 4 presents the results of this study and analysis of findings. Finally, the researcher concluded the work in Section 5.

## 2. Literature Review

Grounded Theory originally was established by Barney Glaser and Anselm Strauss [8][9]. Grounded theory "is a detailed grounding by systematically" and intensively "analyzing data, often sentence by sentence, or phrase by phrase of the field note, interview, or other document; by 'constant comparison,' data are extensively collected and coded," using the operations touched, thus producing a well-constructed theory. The emphasis of the study is not purely on accumulating or collecting "a mass of data, but on establishing several thoughts which have arisen from the study of the data."

The researcher selected Grounded Theory as it delivers a meaningful means of analyzing the data and data are gathered from various sources [10]. In the software engineering field, Grounded Theory is one of the well-established methods to analyze qualitative data. The researcher use sources of data from enhanced analysis of research literature, which is SR and semi-structured interview. This was part of the data collected. Data analysis was carried out according to the theories identified by grounded theory through practicing open, axial, and selective coding techniques [9] [10][11]. Researcher used this coding phases because it has become the most commonly practiced phases in Grounded Theory, and it is utmost prolific when all three stages of coding are practiced [12].

Figure 1 shows the continuous steps in Grounded Theory that was used by the researcher.
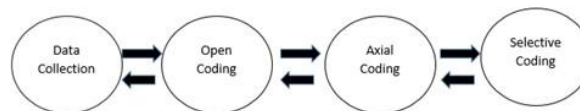


**Figure 1.** Grounded Theory Phases

*a)       Open Coding*

At this stage, the raw data were initially examined and coded. This open coding was completed by attentively analyzing the research literature, which is SR and semi-structured interview very attentively: phrase by phrase or even word by word. The aim was to produce impressions that seem suitable for the data. These ideas directly contributed to additional matters relating to circumstances, strategies, interactions, and significances, which will be later explained.

The principles of open coding are apprehended in what Glaser (1978) called the concept-indicator model [11]. The concept-indicator model, as shown in Figure 2, is grounded on the persistent evaluation of indicators, that is, on regularly recognizing resemblances and discrepancies in writings. According to Larossa (2005), the ''basic, defining rule'' of constant comparison is that, while coding a pointer for a concept, one compares that pointer with previous pointers that have been coded in the same technique [12]. An indicator refers to a word, phrase, or sentence, or a series of words, phrases, or sentences, in the resources being studied. A concept is a tag or name related with a pointer or pointers; stated another way, a concept is a representation or conventional sign attached to a referent. Lines among indicators are intended to show how the constant comparison of indicators generates concepts.
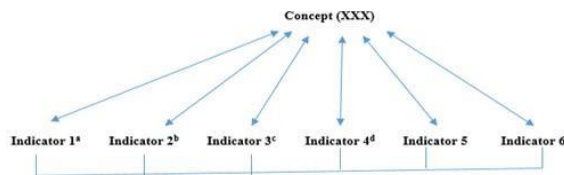
**Figure 2.** Concept-Indicator Model

To map into this study, concept is represented by "coordination strategies," and an indicator is represented by "indicator." The researcher used this constant comparison method to finalize the coordination strategies and related indicators according to the coordination processes in this study.

*b)      Axial Coding*

Axial coding is an essential aspect of open coding. It compromises of intense analysis carried out around one category at a time, in relations of the pattern items (conditions, consequences, and so forth). These outcomes in collective knowledge about relations between that category and other groups and subgroups.

*c)      Selective Coding*

Selective coding relates to coding analytically and specifically for the main category. "The other codes become compliant to the important code under focus. To code selectively, then, means that the researcher has restricted coding to only those codes that narrate to the essential codes inadequately important ways as to be used".

Subsequently, this study used Open Coding, Axial Coding, and Selective Coding to consolidate coordination strategies and related indicators. First, the findings from SR and semi-structured interview were gathered as a data unit in the Open Coding stage. Then, the similarity and commonality between these data units were summarized in the Axial Coding stage. Then data were assembled by making connections between SR and semi-structured coordination strategies and related indicators. Axial Coding was performed using Constant Comparison and Memoing methods. This is followed by Selective Coding stage where the suggested coordination strategies and related indicators were formed. Finally, the consolidated coordination strategies and related indicators according to the coordination processes were ready for validation.

## 3. Methodology

The model formulation is divided into three essential activities, which are indicators consolidation, indicators description, and indicators validation. These activities produce finalized indicators which will be significant for the model formulation. This paper focuses only on the first activity which is indicators consolidation by using Grounded Theory.

Indicators consolidation phase involves the consolidation of coordination strategies and related indicators which was documented through Systematic Review study [15] and semi-structured interview [14] sessions in which the final coordination strategies and associated indicators that likely to contribute to the significance of coordination process assessment. These coordination strategies and relevant indicators were consolidated using Grounded Theory [6][7]. The Grounded Theory approach consists of three stages, which are Open Coding, Axial Coding, and Selective Coding.

## 3.1  Open Coding

In this stage, the researchers gathered the raw data from the SR study and semi-structured interviews together and performed data analysis. These data were grouped to form the categories of

similar phenomena. Open coding process examines the data without any limitations in its scope and without any filters application. Units to analyse differentiate from an individual word, line by line, sentences, or paragraphs. This would create a space for the researcher to search for patterns that may lead to a particular process, which may be of basic interest. The initial finding shows that there are three scenarios involved in the consolidation process for this study, namely as the following:

*a)      First scenario*

The coordination strategies and related indicators are available for both SR and semi-structured interview. The coordination strategies and related indicators found are also mentioned in SR and semi-structured interview methods. An example is shown in Table 1.

Table 1: Example of Open Coding for Scenario 1

| Coordination Strategy | Indicators from SR (Ind1) | Indicators from Semi structured Interview (Ind2) | Consolidation Case |
|---|---|---|---|
| Team Members Selection | Knowledge | Knowledge | Scenario 1 |
|  | Number of years of experience | Number of years' experience | Scenario 1 |
|  | Competent and committed developers | competence skills | Scenario 1 |

*b)      Second scenario*

The coordination strategies and related indicators are available for SR only. The coordination strategies and related indicators found through SR study were not mentioned by the GSD practitioners in interview sessions. An example is shown in Table 2.

Table 2: Example of Open Coding for Scenario 2

| Coordination Strategy | Indicators from SR (Ind1) | Indicators from Semi-structured Interview (Ind2) | Consolidation Case |
|---|---|---|---|
| Team Members Selection | Technical Skills | - | Scenario 2 |
|  | Type of gender | - | Scenario 2 |
|  | Area of expertise | - | Scenario 2 |

*c)      Third scenario*

The coordination strategies and related indicators are available for semi-structured interview only. The coordination strategies and related indicators found through semi-structured interview sessions were not mentioned in SR study. An example is shown in Table 3.

Table 3: Example of Open Coding for Scenario 3

| Coordination Strategy | Indicators from SR (Ind1) | Indicators from Semi structured Interview (Ind2) | Consolidation Case |
|---|---|---|---|
| Team Structure |  | Size of the project | Scenario 3 |
|  |  | Team Size | Scenario 3 |
|  |  | Training plan | Scenario 3 |

These three scenarios were included in the consolidation process. In the first scenario, the researcher compared both existing raw data from the SR study and the semi-structured interview sessions as the core categories, which is coordination strategy and related indicator. In the second scenario, the researcher determined the existing raw data from the SR study as the core category, which is coordination strategy and relevant indicator. In the third scenario, the researcher discovered the existing raw data from semi-structured interview, which is coordination strategies and related indicators as the core category. After that, the researcher continued running the Axial and Selective Coding on these Open Coding data.

## 3.2 Axial Coding

Axial Coding is the process of relating codes (core categories) to each other and is done to identify the properties and dimensions of categories. Then, the core categories that include all the data are identified systematically. An example is shown in Table 4.

Table 4: Example of Axial Coding

| Coordination Strategy | Indicators from SR (Ind1) | Indicators from Semi-structured Interview (Ind2) | Consolidation Case |
|---|---|---|---|
| Team Members Selection | Knowledge | Knowledge | Scenario 1 |
| | Number of years of experience | Number of years' experience | Scenario 1 |
| | Competent and committed developers | competence skills | Scenario 1 |

## 3.3 Selective Coding

Selective Coding is the process of selecting the codes based on the core categories by refining and validating the data according to the core categories. Subsequently, we followed step-by-step Constant Comparison Analysis and Memoing method to perform the consolidation of the indicators. Five steps were followed to achieve the strategies and related indicators consolidation. All the steps are taken from the same example, namely Managing Cultural Diversity, which represents the coordination process.

**Step 1***: The coordination strategies and related indicators are read according to its phrases, meaning, logic and sentence structures under the individual coordination strategies and related indicators.*

Both SR and semi-structured interview mentioned a few same coordination strategies. An example is shown in Table 5. Next, the indicators for the coordination strategies in both SR study and semi-structured interview are read through and understood. For example, the documented Managing Cultural Diversity (coordination process) strategies from the SR study and the semi-structured interviews are Labour turnover, Social Network, and Team awareness. The documented indicators for Labour turnover are Type of Gender, Balance of Religion, Face-to-face meeting, and set expectation. An example is shown in Table 5.

Table 5: Example of Step 1 in Selective Coding

| Coordination Strategy | Indicators from SR (Ind1) | Indicators from Semi-structured Interview (Ind2) | Consolidation Case |
|---|---|---|---|
| Labour Turnover | Gender attitudes | Type of gender | Scenario 1 |
| | Religion attitudes | Balance of religion | Scenario 1 |
| | - | Face-to-face meetings | Scenario 3 |
| | - | Set expectation | Scenario 3 |

**Step 2:** *The similarity (phrases, meaning, logic and sentence structures) between two coordination strategies and indicators identified from the SR study and semi-structured interviews were compared.*

The coordination strategies were compared by determining the similarities between the strategies. Same applies to the indicators also. Each strategies and the indicators were compared implicitly and explicitly. Explicit comparison requires that the strategies and the indicators share the same name. Hence, there is no argument as to their meaning because strategies name and indicators name are the same. For example, the researcher compared the Managing Cultural Diversity (Coordination Process) strategies from both the SR study and the semi-structured interviews.

Alternatively, implicit comparison means the strategies and also the indicators were examined according to their meaning, logic, and structure.

Nevertheless, both strategies and indicators are the same and only differ in terms of the words used. As shown in Table 6, for instance, in the case of labor turnover strategy, indicators like Gender Attitudes (from SR) and Type of Gender (from Interview) have the same meaning but differ in terms of words used. Next, the researcher executed the consolidation process by merging these indicators into one name. An example is shown in Table 6.

Table 6: Example of Step 2 in Selective Coding

| Coordination Strategy | Indicators from SR (Ind1) | Indicators from Semi-structured Interview (Ind2) | Similarity Type | Consolidation Case | Suggested Indicators (based on Ind1 & Ind2) |
|---|---|---|---|---|---|
| Labour Turnover | Gender attitudes Religion attitudes | Type of gender Balance of religion | Implicit Implicit | Scenario 1 Scenario 1 | Type of gender Balance of religion |

**Step 3:** *In some cases, the coordination strategies and related indicators were not available, either in the Second Scenario or Third Scenario.*

As mentioned previously, several strategies and indicators were not available either in the Second or Third Scenario. In this type of scenarios, the strategies and the indicators are not compared but finalized as the final output. For example, in the case of team awareness strategy, several indicators from SR study were not mentioned in the semi-structured interview such as knowledge level of team members, number of tasks that have to complete and schedules, as shown in Table 7, so that these indicators were taken as the final indicators. An example is shown in Table 7.

Table 7: Example of Step 3 in Selective Coding

| Coordination Strategy | Indicators from SR (Ind1) | Indicators from Semi-structured Interview (Ind2) | Similarity Type | Consolidation Case | Suggested Indicators (based on Ind1 & Ind2) | Final Indicators |
|---|---|---|---|---|---|---|
| Team awareness | Knowledge level of team members | - | - | Scenario 2 | Knowledge level of team members | Knowledge level of team members |
| | Number of tasks that have to complete | - | - | Scenario 2 | Number of tasks that have to complete | Total number of tasks that have to complete |
| | Schedules | - | - | Scenario 2 | Schedules | Schedules |

**Step 4:** *The coordination strategies and indicators name were suggested.*

In this step, the researcher suggested the name by doing grouping for each strategy and related indicators. Example is shown in Table 8.

Table 8: Example of Step 4 in Selective Coding

| Coordination Strategy | Indicators from SR (Ind1) | Indicators from Semi-structured Interview (Ind2) | Similarity Type | Consolidation Case | Suggested Indicators (based on Ind1 & Ind2) |
|---|---|---|---|---|---|
| Training | Leadership Skills | - | - | Scenario 2 | Type of Skills |
| | Assertiveness/ Confidence Skills | Assertive Skills | Explicit | Scenario 1 | Type of Skills |
| | Negotiation Skills | - | - | Scenario 2 | Type of Skills |

**Step 5:** *The finalized coordination strategies and indicators were derived.*

As the last step, the researcher finalized the strategies and the indicators for the Managing Cultural diversity process. Table 9 outlines the sample of finalized indicators for the Managing Cultural diversity process. Total of eight strategies and 19 indicators were identified for the Managing Cultural diversity process by doing this consolidation in this study.

Table 9: Example of Step 5 in Selective Coding

| Coordination Strategy | Indicators from SR (Ind1) | Indicators from Semi-structured Interview (Ind2) | Similarity Type | Consolidation Case | Suggested Indicators (based on Ind1 & Ind2) |
|---|---|---|---|---|---|
| Training | Leadership Skills | - | - | Scenario 2 | Type of Skills |
| | Assertiveness/ Confidence Skills | Assertive Skills | Explicit | Scenario 1 | Type of Skills |
| | Negotiation Skills | - | - | Scenario 2 | Type of Skills |

## 4. Results and Discussion

A total of 36 coordination strategies and 167 indicators were derived from this consolidation process. Table 10 shows the example of coordination strategies and indicators list according to the Coordination Process identified in this study before expert validation. Each indicator is assigned its own indicator ID.

Overall, these results indicate that every coordination process has its strategy and indicators, and it depends on the project managers on how they handle them. Project managers could use these strategies and indicators as a benchmark before executing the GSD project. For example, for team setup (coordination process), the coordination strategy found is team members selection as shown in Table 10. Choosing a right team member is very essential in GSD as all the tasks involve teamwork. List of indicators that are formed is shown in Table 10. One of the indicator shown is total number of years of experience, before selecting a team member, a project manager needs to consider this indicator as it might reflect many numbers of years of experience but it might not be from GSD environment. Then coping with GSD projects, might be a new challenge if the person is hired into the project as GSD involves coordination across the countries. Then, the project manager need to check what type of skill or expertise the person has such as good communication skills is required to work with other team members who are located in different parts of the world. Thus, the indicators will serve as a good guideline for the project manager to assign the best person to form an ideal team.

Table 10: Example of Coordination Strategies and related Indicators List before Validation

| No | Coordination Strategy | Indicator ID | Indicator Name |
|---|---|---|---|
| | | **CP1 Team Setup** | |
| 1 | Team Members Selection | TS1 | Team Knowledge |
| 2 | | TS2 | Total number of technical skills |
| 3 | | TS3 | Type of gender |
| 4 | | TS4 | Type of skill or expertise |
| 5 | | TS5 | Able to handle cross functionality |
| 6 | | TS6 | Total number of years of experience |
| 7 | | TS7 | Build the trust |
| 8 | | TS8 | Competent and committed developers |
| 9 | | TS9 | Labour Cost |

## 5. Conclusion

This work describes a study that was carried out to consolidate coordination strategies and related indicators for assessing the coordination processes in GSD projects from two different sources

namely, SR and semi structured interview. As a conclusion, the researcher has consolidated 36 coordination strategies and 167 related indicators. Each finalized strategies and indicators from consolidation process were described according to the ISO/IEEE 15939 Software Measurement Standard before validation by the GSD experts. These coordination strategies and the related indicators were used to formulate the Evaluation Model for assessing the effectiveness of coordination processes in GSD projects.

## 6. Acknowledgement

## References

[1]     Jain, R., & Suman, U. (2015). A Systematic Literature Review on Global Software Development Life Cycle. ACM SIGSOFT Software Engineering Notes, 40(2), 1-14.

[2]     Kaur, P., & Sharma, S. (2014). Agile Software Development in Global Software Engineering. International Journal of Computer Applications, 97(4), 39-43.

[3]     Niazi, M., Mahmood, S., Alshayeb, M., Riaz, M. R., Faisal, K., & Cerpa, N. (2013, October). Challenges of project management in Global Software Development: Initial results. In 2013 Science and Information Conference, 202-206.

[4]     Babar, M. A., & Lescher, C. (2014). Editorial: Global software engineering: Identifying challenges is important and providing solutions is even better. Information and Software Technology, 56(1), 1-5.

[5]     Malone, T. W., Malone, T. W., & Crowston, K. (1994). The interdisciplinary study of coordination. ACM Computing Surveys (CSUR), 26(1), 87-119.

[6]     Babchuk, W. A. (1996). Glaser or Strauss? Grounded theory and adult education. In Proceedings of the 15th Annual Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education, 1-6.

[7]     Heath, H., & Cowley, S. (2004). Developing a grounded theory approach: a comparison of Glaser and Strauss. International Journal of Nursing Studies, 41(2), 141-150.

[8]     Strauss, A. L. (1987). Qualitative analysis for social scientists. Cambridge university press.

[9]     Strauss, A., & Corbin, J. (1990). Basics of qualitative research. Sage publications.

[10]    Strauss, A., & Corbin, J. (1998). Basics of qualitative research techniques. Thousand Oaks, CA: Sage publications.

[11]    Strauss, A. L. (1987). Qualitative analysis for social scientists. Cambridge university press.

[12]    LaRossa, R. (2005). Grounded theory methods and qualitative family research. Journal of marriage and Family, 67(4), 837-857.

[13]    Subbarao, A., & Mahrin, M. N. R. (2017). Evaluation Model to Assess the Effectiveness of Coordination Processes in Global Software Development Projects: A Roadmap. Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 9(3-3), pp. 67-72.

[14]    Subbarao, A., & Mahrin, M. N. R. (2018). Identification of Coordination Strategies and Indicators for Global Software Development Projects: Interview Outcome. New Trends in Intelligent Software Methodologies, Tools and Techniques: Proceedings of the 17th International Conference SoMeT18, pp. 545-558

[15]    Subbarao, A., & Mahrin, M. N. R. (2019). A Systematic Review of Coordination Approaches and Indicators in Global Software Development Projects. Journal of Advanced Research in Dynamical and Control Systems (JARDCS), Vol. 11, Special Issue 10, pp. 1074-1080.

[16]    Subbarao, A., & Mahrin, M. N. (2020, May). Research Framework of Evaluation Model to Assess the Effectiveness of Coordination Processes in Global Software Development Projects. In Journal of Physics: Conference Series (Vol. 1529, No. 5, p. 052064). IOP Publishing.