

The Analysis of Secondary Case Studies as a Starting Point for Grounded Theory Studies: An Example from the Enterprise Software Industry

Abilio Avila, Orestis Terzidis

Abstract—A fundamental principle of Grounded Theory (GT) is to prevent the formation of preconceived theories. This implies the need to start a research study with an open mind and to avoid being absorbed by the existing literature. However, to start a new study without an understanding of the research domain and its context can be extremely challenging. This paper presents a research approach that simultaneously supports a researcher to identify and to focus on critical areas of a research project and prevent the formation of prejudiced concepts by the current body of literature. This approach comprises of four stages: Selection of secondary case studies, analysis of secondary case studies, development of an initial conceptual framework, development of an initial interview guide. The analysis of secondary case studies as a starting point for a research project allows a researcher to create a first understanding of a research area based on real-world cases without being influenced by the existing body of theory. It enables a researcher to develop through a structured course of actions a firm guide that establishes a solid starting point for further investigations. Thus, the described approach may have significant implications for GT researchers who aim to start a study within a given research area.

Keywords—Grounded theory, qualitative research, secondary case studies, secondary data analysis, interview guide.

I. INTRODUCTION

THE research in some domains of the enterprise software industry is still in a formative phase and a broad variety of different phenomena and aspects are being explored. In order to get an insight into this domain with an open mind and to avoid getting absorbed too quickly by existing theories, the authors decided to use an inductive research approach that enables the emergence of the primary research questions driven by data. GT represents an appropriate research approach to identify research questions, and generate in the further course, a theory that is fully grounded in data. A key principle of this research approach is to prevent the unreflecting adoption of preconceived theories, and in particular, to avoid an intensive literature review before the building blocks of the own GT are described. Overall, GT is a research approach that focuses on the inductive generation of theory, well grounded in empirical data [1]–[4].

GT is primarily based on the following fundamentals:

- 1) The development of theory based on data-driven evidence
- 2) The prevention of a preconceived theory at the initiation of the research.

Abilio Avila Albez is with the Karlsruhe Institute of Technology, Germany (e-mail: abilio.avila@kit.edu).

- 3) The application of iterative and simultaneously used coding procedures.
- 4) The constant comparison of data with already identified concepts.
- 5) The selection of adequate data resources driven by the emerging theory (theoretical sampling).
- 6) Theoretical saturation, where the collection of data continues until additional data no longer contributes to the identification of new aspects [2], [4], [5].

II. RESEARCH QUESTION

In the present study, the researchers began with a broad research area guided by an initial research question as the starting point: What are the critical factors for successful products in the enterprise software industry?

The researcher analyzed 15 secondary case studies addressing this research question with the aim to identify and aggregate an initial cross-case pattern of success among these case studies, which serves the researcher as a starting point for further investigations. It allowed a preliminary structuring of the research domain based on real business cases.

Primarily based on the analysis of these case studies, the researcher constructed an initial conceptual framework consisting of six main categories that formed the framework for the guideline for expert interviews. This inductive approach allowed the researcher to start the interviews very open mindedly without being influenced by theories described in the current literature.

III. PURPOSE OF THIS PAPER

One of the key fundamentals of GT is to reduce the risk to begin a research project with a preconceived theory in mind. The objective is to allow theory to emerge from data, as [3] argued that it is not unusual for a researcher to fall in love with a previous study, before or during his or her own research. Therefore, a guiding element of GT is to avoid an intensive upfront literature review. This guideline is “brought about by the concern that literature might contaminate, stifle or otherwise impede the researcher’s effort to generate categories...” [6]. A specific literature review could lead to a research journey where data are forced into a preconceived structure, instead of the discovery of new concepts [3]. Reference [7] argues similarly:

“Only when the theory has substance, or when no new findings arise from freshly collected data, should the

researcher review the work in the field and relate the theory to it through the integration of ideas”.

However, [8] pointed out, “there is a difference between an open mind and an empty head”. On the contrary, the development of the context, structure and content of a GT research project from scratch without sensitivity for the research domain and further input is extremely challenging. Starting a new project without guiding elements or sensitizing concepts is bristling with serious obstacles [9].

Researchers new in the field of GT may find themselves confused about how to begin a GT research project. On the one hand, a research project has to avoid to get absorbed by existing literature and to prevent the formation of preconceived concepts, while on the other hand it is crucial to develop a point of departure and to mark out the research context. The examination of the corresponding literature e.g. [10]-[13] indicates that this twofold problem represents an unresolved methodological challenge with different, often contradictory, views on whether and how to involve existing theory at the beginning of a GT study.

The purpose of this paper is to present a structured procedure that simultaneously supports a researcher to identify and focus on the critical areas of its research project and prevents the formation of preconceived concepts by the current body of literature. The present paper aims to offer an approach for the initiation of a GT study that enables a researcher to develop an understanding of the research domain, understand the context of the research, mark out the research area and identify key elements of interest for the further research activities, without being influenced by the current body of theories.

The research approach for the initiation of the focal GT study described in this paper has four stages:

- 1) Selection of secondary case studies.
- 2) Analysis of secondary case studies.
- 3) Development of an initial conceptual framework.
- 4) Development of an initial interview guide.

Correspondingly, the present paper falls into associated sections, each dealing with a stage of the research approach. Stage 1 covers the selection of suitable secondary case studies and the use of selection criteria. Stage 2 is devoted to the coding and analysis of the selected cases. Stage 3 illustrates the initial conceptual framework as the result of multiple iterations of data coding. Derived from the previously outlined initial conceptual framework, Stage 4 discusses the construction of an initial interview guide as the fundament for further research activities.

IV. SELECTION OF SECONDARY CASE STUDIES (STAGE 1) - DATA COLLECTION

In order to prevent the formation of preconceived concepts by the current body of literature, we avoided at the beginning of the research project an intensive literature review. Instead, we started the research project with the analysis of secondary case studies by using the coding scheme of GT. The researchers describe the analysis of secondary case studies as the use of existing case studies in order to address a research

question that is in line with, but differs from the purpose of the original case study [14]. The selected cases offered evidence based on real business experiences.

Reference [15] defines case study as “an empirical inquiry that investigates a contemporary phenomenon (the “case”) in-depth and within its real-world context, especially when the boundaries between phenomena and context may not be clearly evident”. Hereby, case studies usually combine multiple data collection methods such as archives, interviews, observations and questionnaires [16]. Since case studies are conducted in real-world settings, they have a high degree of realism and offer insights in real business situations [17]. Thus, case studies are a suitable data source to investigate the complex conditions of the development and commercialization of software products in the enterprise software industry. The case studies allowed to access aggregated data, that offered an insight into the determinants for successful software products.

The researchers selected suitable case studies for the subsequent analysis guided by three simultaneously applicable main criteria:

- 1) Companies in the Enterprise Software Industry (B2B market).
- 2) Companies that offered complex software products to solve complex business problems for their customers.
- 3) Companies that needed to provide complementary business services in order to offer their customers a satisfactory solution for their business need.

In addition, an obligatory selection criterion was the contribution of the cases to the addressed research question: What are the critical factors for successful products in the enterprise software industry? Consequently, the researchers concentrated on cases covering the development and commercialization of software products in the enterprise software industry. Based on these criteria, a diverse selection of cases in the Enterprise Software Industry was collected. By using a heterogeneous sample of case studies, the research aimed to set up the base for the identification of research topics of overall importance for the enterprise software industry, across different business fields and product areas. The selected cases represent real business cases from areas such as virtualization, CRM and middleware. Irrespective of the diversity of the selected cases, all of the cases offered the researchers aggregated data from the enterprise software industry associated with the addressed research question. The cases comprise between 15 pages and 41 pages (besides one exception with 10 pages) and were written in English.

In addition to the selection criteria regarding the appropriateness to address the focal research question, the researcher evaluated the completeness and quality of the secondary cases. In accordance with [14], the researcher used the following general criteria to assess the conditions of the cases:

- 1) Condition of the case study report e.g., related documents and data are intact, no relevant documents are missing.
- 2) Accuracy of the case study report and transcription e.g., no vague and incomplete interview quotes.

- 3) Comprehensibility of the case study report e.g., no misleading or ambiguous descriptions.
- 4) Interpretability of the case study report e.g., clear relationships between objective of the case study and their elements.
- 5) Explorability of the case study report e.g., necessary depth and breadth, sufficient details for an analysis, enough context and data that invite to explore a phenomenon.

The researcher concentrated the search on case studies used by business schools. This decision was based on the fact that those case studies are used to transfer experience-based knowledge and are meant to foster the understanding of a business and its domain based on real-world situations. Case study teaching aims to encourage the discussion among the participants and to improve the ability to recognize cross-case patterns. Thus, these cases are usually developed according to the above-mentioned conditions and their quality and completeness are carefully reviewed [18]-[20].

Since the researcher focused on case studies used by business schools, the sample is dominated by US-companies. This is due to the fact that US-Business Schools intensively use case studies as teaching material for their postgraduate degree programs. Consequently, the total amount of available case studies is predominated by US-cases.

The 15 selected secondary case studies represent the initial theoretical sample (theoretical sample zero) and served the researchers as a starting point for further investigations.

An overview of the selected case studies is given in Table I.

V. ANALYSIS OF SECONDARY CASE STUDIES (STAGE 2) - DATA ANALYSIS

In GT, the conceptualization and categorization of the amount of collected data are based on the coding of information chunks. The resulting categories represent the elements of an emerging theory.

“Coding gets the analyst off the empirical level by fracturing the data, then conceptually grouping it into codes that then become the theory that explains what is happening in the data. A code gives the researcher a condensed, abstract view with scope of the data that includes otherwise seemingly disparate phenomena” [21].

The categorization of the data through a structured coding approach intends to identify a main research question driven by data and ultimately to develop an associated theory grounded on empirical data. However, at this research stage the objective of the data analyst was not to develop a full theory, but to identify areas of interest for the further investigation.

Beginning with the foundation of the Grounded Theory Method (GTM) in 1967 by Barney Glaser and Anselm Strauss [22] two main strands of GTM evolved and different coding procedures were suggested. Based on Strauss’s GT version of GT [23] we used a coding approach composed of three main elements: open coding, axial coding, and selective coding.

TABLE I
 SELECTED CASE STUDIES

ID	Case Title	Company	Country
01	Scrum, Sprints, Spikes and Poker	Telerik	Bulgaria
02	Beas Systems, Inc. In 2013: Reaching for the Next Level	Beas Systems	USA
03	Precise Software Solutions	Precise Software Solutions	USA
04	PremiumSoft: Managing Creative People	PremiumSoft	China
05	Oracle Corporation	Oracle Corporation	USA
06	Product Development at OPOWER	OPOWER	USA
07	Nuway Software	Nuway Software	USA
08	SAP AG: Orchestrating the Ecosystem	SAP AG	Germany
09	Salesforce: The Evolution of Marketing Systems	Salesforce.com	USA
10	Siebel Systems	Siebel Systems	USA
11	WebSpective Software, Inc.	WebSpective Software, Inc.	USA
12	Customer Value-Based Pricing	Trilogy Corporation	USA
13	Lean at Wipro Technologies	Wipro Technologies	India
14	VMware, Inc., 2008	VMware	USA
15	SAP and Cloud Computing in 2012 and Beyond	SAP AG	Germany

A. Open Coding

Open coding aims to obtain a first structure of the data by breaking data apart into discrete blocks enabling the further analysis. The purpose is to distinguish categories and their related sub-categories within the data. This is done by a constant comparison of the data and the identification of similarities and differences. The researchers identified phenomena in the data and mark them with codes in order to develop categories. New categories appeared and new phenomena were allocated to existing categories. Categories that represented characteristics or properties of other categories were sub-grouped under corresponding categories. The result was an initial structure of the data composed of categories and associated sub-categories.

B. Axial Coding

Axial coding reassembles the data that were fractured through open coding in order to create a dense composition of relationships. For this purpose, axial coding specifies in more details the associations between the categories and their subcategories and describes the nature of their relationships, to obtain more precise and comprehensive explanations of particular phenomena.

C. Selective Coding

Through this coding stage, the researcher incorporates and sharpens the identified main categories (the categories on the highest abstraction level) in order to create a dense theory. This implies the need to recognize a central category that represents the central topic of the study. This category links all the other categories, with the purpose to obtain an integrated theory. Consequently, through this procedure, the researcher determines which of the initial identified categories contribute to the core theory, and specifies how the selected main categories are related to the all-comprising core category.

In GT, the coding approach is not a linear process. Rather, it represents an iterative procedure based on continuous comparison and analysis of the data. Through this coding, the three coding modes are applied iteratively and usually simultaneously.

As described above, at this point the objective was to set up the fundament for the further course of actions through the coding and analysis of the secondary case study. The objective at this research stage was not to develop a theory, but to prepare primary data collection. Thus, the focus through the coding process was on structuring the elements of the secondary case studies and to identify and group the important data units (in order to define relevant research areas, rather than explaining the nature of their relationships).

The coding procedure is not a linear process. Rather, it is an iterative process based on constant comparison of the data, where the coding modes are applied iteratively and often simultaneously.

VI. INITIAL CONCEPTUAL FRAMEWORK AND INTERVIEW GUIDE: RESULTS

A. Initial Conceptual Framework (Stage 3)

As a result of multiple iterations of data coding, the secondary case study research synthesized six main areas of interest for the further primary research activities, as illustrated in Fig. 1. These six areas represent the building blocks and the fundament for the construction of the interview guide.

Reference [3] acknowledges that the development of a “middle-range theory, a previously identified theoretical framework can provide insights, direction, and a useful list of initial concepts”. However, he emphasizes that it is a prerequisite for the use of an initial framework to remain open minded and avoid to be forced in a predefined structure. The use of an initial framework should not lead to the predetermination of concepts. The researcher argued that the analysis of secondary case studies allowed the researcher to start very open minded without being influenced by the current body of theories. The development of this sensitizing concept provided the researcher, in line with [9], with a point of departure “to form interview questions, to look at data, to listen to interviewees, and to think analytically about data” while retaining the openness for new emerging topics.

Instead of a descriptive text, the documentation of the main and sub-categories has been done by the development of corresponding interview questions addressing these categories and through the preparation of a structured interview guide. The objective was to be able to ask the right questions in the further course of research.

B. Initial Interview Guide (Stage 4)

The development of an interview guide represents the foundation for the interviews and prepares the researcher for conducting the interviews with the experts. Reference [9], for example, strongly advocates the use of an interview guide to begin a GT study:

“Starting a new project without a working guide is, however, fraught with pitfalls, particular for novice. This approach invites asking awkward, poorly timed, intrusive questions that may fill with unexamined preconceptions. Even experienced interviewers (...) may construct interview guides”.

The interview guide is a structured list of well-constructed interview questions. These questions address the research topics that are intended to be covered during the course of the interviews. As pointed out by [24], “an interview guide plainly sets out my ‘domains of inquiry’”. Thus, the interview guide represents a data collection tool that directs the researcher on what data to collect, how to obtain the data and how to ask the interview questions. Nevertheless, rather than a fixed structure forcing the data in a pre-developed skeleton, it should be treated as a flexible tool. An elaborate and well-prepared set of questions supports the researcher to conduct fluid and spontaneous interviews, while covering all the relevant topics. The course of the interview may of course lead to different expressing, deviating from the original wording or the order of the questions.

Developing, reviewing and refining an interview guide improves the understanding of how and when to ask specific questions throughout the interviews. It supports the researcher to reflect on the research process and to become aware of the actual point of attention and the use of adequate language [9], [25].

In the focal research, the initial interview guide was crafted according to the previously identified components of the conceptual framework. It allowed the researcher to start the expert interviews with a set of educated questions and to effectively and efficiently start dialogs that led to an in-depth understanding of the research context. Furthermore, the aim was to set up the fundament to discover the primary research question driven by data (expert interviews).

The result of this research step was an initial catalogue of 186 modular questions grouped into six discrete sections that cover the topic of interest plus two introductory sections. Thus, the interview guide comprises eight sections:

- 1) Introduction: The researcher provides the interview participant with an overview of the research context and the objective of the study. The setting of the interview regarding data protection, recording, transcription and further processing are described. Furthermore, the researcher asks the participants explicitly for permission to record the conversation.
- 2) Questions regarding the interview participant: The participant is asked to introduce himself and to offer a short overview of his professional career and his experience in the Enterprise Software Industry. In addition, the researcher asks for a short description of his current job position and employer.
- 3) Company related factors: This group of questions addresses the influence of company related factors (e.g. culture, organization structure and competencies) on the success of a software product in the Enterprise Software Industry. The associated subcategories of the

corresponding main category in the conceptual framework are covered through this set of questions.

- 4) Partner network related factors: The questions within this sample focus on the relevance of partners and the company's ability to manage a network of partners. The associated subcategories of the corresponding main category in the conceptual framework are covered through this set of questions.
- 5) Market related factors: The importance of market related factors such as competition, target customers and market barriers are addresses through this compilation of questions. The associated subcategories of the corresponding main category in the conceptual framework are covered through this set of questions.
- 6) Product related factors: The questions of this group aim to address the significance of the characteristics affiliated to

the product, its technology and pricing model. The associated subcategories of the corresponding main category in the conceptual framework are covered through this set of questions.

- 7) Product development process: The attributes of a successful new product development process and its development phases are focused through this part of the conversation. The associated subcategories of the corresponding main category in the conceptual framework are covered through this set of questions.
- 8) Environmental influences: This section of the interview was structured to focus on external influences like technology trends, market trends and regulations. The associated subcategories of the corresponding main category in the conceptual framework are covered through this set of questions.

<p>Company Related Factors</p> <ul style="list-style-type: none"> ▪ Understand and manage the product value chain ▪ Cultivate distinctive competencies along the product value chain ▪ Determine the scope of the software company ▪ Understand dependencies and co-innovation risk ▪ Complementary value enhance services ▪ Complementary Platform Strategy ▪ Intellectual Property (IP) ▪ Organizational structure, processes and internal communication ▪ Company Culture ▪ Human Resources Management ▪ Finance and Investment ▪ Common company vision and business strategy ▪ Software Product Management ▪ Sales organization and sales process 	<p>Product Related Factors</p> <ul style="list-style-type: none"> ▪ Product Platform Strategy ▪ Product Platform Architecture ▪ Product Technology ▪ Pricing / Revenue Model ▪ Unique Superior Product ▪ Complementarity and Interoperability 	<p>Partner Network Related Factors</p> <ul style="list-style-type: none"> ▪ Leverage the ecosystem to accelerate innovations, create business values for the customer and scale ▪ Establish and leverage a partner network around the product ▪ Establish a procedure for partner select ▪ Low barriers for partnership ▪ Offer incentives and a clear partner business proposition for partners ▪ Ensure high quality of partner solutions ▪ Education and service offering for partners ▪ Stimulate complementary innovation for the product ▪ Establish strategic alliances ▪ Establish routines for developing and maintaining relationships with partner ▪ Common product innovation roadmap with partners
<p>Market Related Factors</p> <ul style="list-style-type: none"> ▪ Competition ▪ Target Customers ▪ Market Barriers ▪ Market Opportunity ▪ Community Management ▪ Company reputation and credibility 	<p>New Product Development Process</p> <ul style="list-style-type: none"> ▪ Opportunity Identification and Evaluation ▪ Product Development ▪ Product Commercialization 	<p>Environmental Influences</p> <ul style="list-style-type: none"> ▪ Technology Trend ▪ Market Trends ▪ Government and Regulation

Fig. 1 Six identified main categories and corresponding sub-categories

This interview guide has to be treated as a starting point for the investigation. Those elements that prove to be irrelevant during the subsequent research activities are removed, while new elements extend and adapt the interview guide.

VII. CONCLUSION

The analysis of secondary case studies introduces the possibility to have significant implications for GT researchers who aim to investigate relevant subjects within a broad research area. Using the analysis of existing case studies as a starting point could be a way to mark out the research area, based on real-world examples and thus without being influenced by the existing body of theory. The use of existing

case studies allows a researcher to enter a research domain based on real-world data and to identify areas of interests. It permits a researcher to develop through a course of structured actions a solid interview guide and thus an initial structure for further investigations. However, the results need to be viewed primarily as a preparation for the conducting of interviews, and thus, for the primary data collection. Caution is advised when applying these results beyond this purpose. Both, the derived conceptual framework as well as the interview guide, consistent with [9], has to be treated as a starting point for the identification and investigation of the research area, while retaining an open mind for new topics. The results represent the beginning rather than the end of the journey. Furthermore,

the results of the described approach depend highly on the selected case studies. Before considering the application of this approach, it must be evaluated if the available case studies in a given research domain are suitable to answer the initial research question of the researcher, and hence, are applicable to the specific research project. The approach may not be applicable if the original objective of the available secondary case study differs significantly from the researchers' purpose and interest. Also the evaluation of completeness and quality of the data from the case study reports may be a challenge. In addition, a limitation of the presented approach is based on its very nature: The retrospective view of secondary cases excludes the possibility for researchers to influence the data collection. Despite these limitations, our results may support researchers with a valuable starting point for the conduction of a GT study. The analysis of secondary case studies as a foundation for further research activities has important implications for qualitative researchers who aim to develop sensitivity and prepare adequately for a research domain and identify areas of interest, without being influenced by current body of theory.

- [22] Glaser, B., Strauss, A.: The discovery of grounded theory. London: Weidenfeld and Nicholson (1967).
- [23] Strauss, A., Corbin, J.: Basics of qualitative research, vol. 15. Newbury Park, CA: Sage (1998).
- [24] Karp, D.A.: Learning How to Speak of Sadness. *Ethnographies Revisited: Constructing Theory in the Field*, 37 (2009).
- [25] Gläser, J., Laudel, G.: *Experteninterviews und qualitative Inhaltsanalyse*. Springer-Verlag (2010).

REFERENCES

- [1] Glaser, B., Strauss, A.: Applying grounded theory. The discovery of grounded theory: strategies of qualitative research. Hawthorne, NY: Aldine Publishing Company, 237–51 (1967).
- [2] Glaser, B.G., Strauss, A.L.: The discovery of grounded theory: Strategies for qualitative research. Transaction Publishers (2009).
- [3] Strauss, A.L., Corbin, J.M.: Basics of qualitative research. Sage (2008).
- [4] Urquhart, C.: Grounded theory for qualitative research: A practical guide. Sage (2013).
- [5] Willett, P., Idrees, I., Vasconcelos, A.C., Cox, A.M. (eds.): The use of grounded theory in PhD research in knowledge management: A model four-stage research design. Emerald Group Publishing Limited (2011).
- [6] Glaser, B.G.: Emergence vs forcing: Basics of grounded theory analysis. Sociology Press (1992).
- [7] Goulding, C.: Grounded theory: the missing methodology on the interpretivist agenda. *Qualitative Market Research: An International Journal* 1(1), 50–57 (1998).
- [8] Dey, I.: *Qualitative data analysis: a user-friendly guide for social scientists* (1993).
- [9] Charmaz, K.: *Constructing grounded theory*. Sage (2014).
- [10] Cutcliffe, J.R.: Methodological issues in grounded theory. *Journal of advanced nursing* 31(6), 1476–1484 (2000).
- [11] Dunne, C.: The place of the literature review in grounded theory research. *International Journal of Social Research Methodology* 14(2), 111–124 (2011).
- [12] McGhee, G., Marland, G.R., Atkinson, J.: Grounded theory research: literature reviewing and reflexivity. *Journal of advanced nursing* 60(3), 334–342 (2007).
- [13] Suddaby, R.: From the editors: What grounded theory is not. *Academy of management journal* 49(4), 633–642 (2006).
- [14] Hinds, P.S., Vogel, R.J., Clarke-Steffen, L.: The possibilities and pitfalls of doing a secondary analysis of a qualitative data set. *Qualitative Health Research* 7(3), 408–424 (1997).
- [15] Yin, R.K.: *Case study research: Design and methods*, vol. 5. Sage (2012).
- [16] Eisenhardt, K.M.: Building theories from case study research. *Academy of Management Review* 14(4), 532–550 (1989).
- [17] Runeson, P., Host, M., Rainer, A., Regnell, B.: *Case study research in software engineering: Guidelines and examples*. Wiley.com (2012).
- [18] Hammond, J.S.: *Learning by the Case Method* (1976).
- [19] Ellet, W.: *The case study handbook: How to read, discuss, and write persuasively about cases*. Harvard Business Press (2007).
- [20] Erskine, J.A., Leenders, M.R., Mauffette-Leenders, L.A.: *Teaching with cases*. Ivey Publishing, Richard Ivey School of Business (1998).
- [21] Glaser, B.G., Holton, J. (eds.): *Remodeling grounded theory* (2004).