[simula . research laboratory]

Case Study Research

Bente Anda

Research scientist, Simula Research Laboratory

Associate professor II, IFI

25.09.2007

[simula . research laboratory]

Motivation

- Researchers (including some MSc and PhD students) need knowledge of how to conduct and report case studies

- Practitioners need knowledge of case study research to be able to identify and apply relevant results to their own projects. This includes
    - how to read papers that report case studies, and
    - how to judge their quality.

Overview of Lecture

1. What is a case study

2. Choice of case(s)

3. Case study design
   a. Research question(s)
   b. Unit(s) of analysis.
   c. The case study protocol
   d. Data collection
   e. Analysis
   f. Generalization

4. Summarizing results from case studies

5. Challenges with case studies

A software enginering case study is

a method for learning about a software engineering technology (process, method, technique, tool, language), based on a comprehensive understanding of one instance of use, obtained through extensive description and analysis of the instance taken as a whole and in its context.

When should you do a case study?

- When the phenomenon of interest consists of many, possibly related variables
- When you cannot separate the phenomena under study from context, and hence, the context is important
- When you need to know whether your theory applies to a specific real world setting

  Case studies are well-suited to investigate "how" and "why" questions. Ex.
    - Why and how a technology is used or not used in a specific context
    - Why and how a technology functions or not functions in a specific context

25/9 – 07 Bente Anda   5

Case Study as a Scientific Method

- A case study is an empirical research method.

- It is a defined, scientific method for posing research questions, collecting data, analyzing data and presenting the results.

- Each step of a case study is planned from the outset.

25/9 – 07 Bente Anda   6

3

A Case Study is not

- An exemplar or illustrative example, e.g., a teaching example.

- An experience report, that is a retrospective account of events or interventions accompanied by lessons learned.
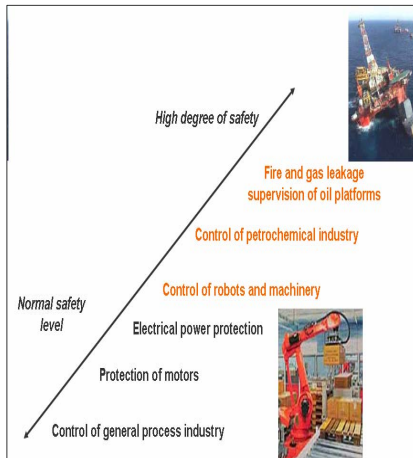

Example Case Study 1

A case study was conducted in cooperation with the company ABB on:

Identifying challenges and benefits when introducing UML-based development in safety-critical software development.

ABB

Example Case Study 1 Cont.:
ABB Safety Systems

High degree of safety

Fire and gas leakage
supervision of oil platforms

Control of petrochemical industry

Control of robots and machinery

Normal safety
level   Electrical power protection

Protection of motors

Control of general process industry

- The Company ABB has a large number of projects involving software development.
  - Safety critical systems with embedded software.
  - The projects and products must satisfy international safety standards, e.g. EN 954 and IEC 61508.
- UML-based development was considered an important means to satisfy these standards.

25/9 – 07 Bente Anda   9

Example Case Study 1 Cont.

Research questions:

- How does the use of UML-based development affect the project and the resulting product?

- Why is the project and product affected in this way?

ABB

25/9 – 07 Bente Anda   10

[simula . research laboratory]

Exercise 1

Which of these research questions are most suitable for, respectively, a case study, experiment and survey?

1. Which of the following design patterns Visitor, Observer and Decorator is the easiest to use?

2. Is pair programming more efficient than individual programming when maintaining Java code?

3. What are the main differences between test driven development and traditional development with ad hoc testing?

4. How efficient is test-driven development?

5. What are major software project and product success factors?

6. What are the costs and benefits of close customer engagement in agile projects?

7. How do software developers view software reuse?

8. Are there more defects in new or in modified source code?

9. What is most effective when evaluating software architecture, face-to-face meetings or distributed meetings?

10. How are software inspections conducted?

[simula . research laboratory]

Choice of Case

- A typical case is chosen to highlight and understand general characteristics of a phenomenon and possibly to construct theory

- A critical case is chosen to test existing theory (seeks falsification of hypotheses)

- In practice, convenience cases are often used, but it is still necessary to describe how the case fits the research question

Choice of Case Cont.

Types of case studies:

- Exploratory – identifies concepts and tentative relationships for further studies.

- Descriptive – describes a sequence of events and underlying mechanisms.

- Explanatory – identifies cause-effect relationships.

Example Case Study 2

The Project:

- A process control system based on several existing systems. A generic system installed at several locations.
- ABB's largest software project so far
    - 230 people,
    - UML-based development was used in 3 of the subprojects with a total staff of approximately 100 people.

Characteristics:

- Descriptive, single case
- A typical case with respect to large scale safety-critical software development, but from a general software engineering point of view it can be seen as a critical case.
- Can also be seen as a convenience case. ABB had introduced model-driven development with UML and wanted a case study on the effects.
- → Which challenges may this have entailed for the study?

ABB

Case Study Design

Each case study must have a plan for how to collect empirical data and connect them to the research questions and conclusions. The plan should contain:

1. The research questions (how, why, and possibly what, who, where)

2. Unit(s) of analysis (what a 'case' is)

3. How will data collection be performed

4. Analysis - The logic linking of the collected data to results.

5. The criteria for interpreting the findings.

Unit of Analysis

- The unit of analysis is the primary interest of the study

    – Example: an individual, an organization, a program

- The unit of analysis usually depends on the primary research question

- In the ABB case study the main unit of analysis was the project, for some research questions it was the subprojects

ABB

Case Study Protocol

- An overview of the case study

- Description of field procedures (sources of information…)

- Case study questions

- A guide for the case study report

  If possible, conduct a pilot case study to refine the case study protocol.

25/9 – 07   Bente Anda   17

Data Collection

- Data collection in case studies is complex (probably more than for other types of empirical studies).

- Remember to:
  - create a case study database,
  - maintain a chain of evidence, and
  - use multiple sources of evidence

- Data Sources:
  1. Project documentation (possibly archival records) including source code
  2. Interviews
  3. Observations (direct and as participant)
  4. Physical artifacts

25/9 – 07   Bente Anda   18

[simula . research laboratory]

Data Collection Challenges

– Project documentation is often incomplete

– Interviews have the following challenges

- Choice of interviewees
- Development of interview guide
- Conduct of interviews (how to make interviewees talk freely)
- People, and therefore interviewees, have selective memories, and some are more talkative, and hence, may be more influential than others.

→ This challenge can be addressed using triangulation.

Triangulation is a technique for avoiding/reducing bias in the results due to bias in the collected data. Forms of triangulation are data triangulation, investigator triangulation, method triangulation and theory triangulation.


[simula . research laboratory]

Example Case Study 4

1. Interviews with 16 project members.

2. Questionnaire answered by 55 other project members.

3. Project documents: Analysis and design documents and review reports from 6 modules of the system.

ABB

Analyzing Case Study Data

- Strategies and techniques for analysizing case studies have not been well defined.

- Remember to define priorities for what to analyze in-depth and which data that should constitutes the core basis for the evidence

- No collected data should be ignored

- Analysis includes:
    1. Examining the data
    2. Categorizing the data
    3. Tabulating/recombining the data
    4. Testing explanantions

25/9 – 07 Bente Anda   21
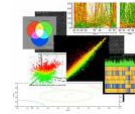
Analysis Techniques

- In **pattern-matching,** an empirically based pattern is compared with a predicted one (the comparison between the predicted and actual pattern may not include any quantitative criteria).

- In **explanation-building**, the analysis of the case study is carried out by building an explanation of the case. It is an iterative process that begins with a theoretical statement, refines it, revises the proposition, and repeats this process from the beginning.

- In **time-series analysis** changes over time are tracked in detail. The essential underlying a time-series design is the match between a trend of data points compared to a trend specified beforehand vs. a rival trend.

25/9 – 07 Bente Anda   22

Example Case Study 6:
Analysis of Interviews

1. Identified categories for coding. The categories were based on the interview guide and on experience from the actual interviews.

2. Coded all statements in the interviews according to the categories.

3. Sorted the statements according to propositions and according to whether they supported or contradicted the proposition.

4. Checked the background of the interviewee behind each statement (not all had relevant experience with respect to each proposition).

5. Identified explanations for each proposition.

ABB

25/9 – 07 Bente Anda   23

Validity

- Construct validity
  - Do our measurements actually represent the concept that we are interested in?

  Ensured through triangulation and participant validation

- Internal validity (for explanatory or causal studies only)
  - Does the effect actually have the cause that we observe?
  - Does the internal logic of the study hold up?

  Ensured by carefully addressing rival theories

- External validity
  - In which contexts and for which domains can this study's findings be generalized?

  Ensured by thorough description of the case study context

- Reliability
  - Can the study be repeated with the same results?

  Ensured by a complete case study protocol

25/9 – 07 Bente Anda   24

Example Case Study 6

- Construct validity
  - Triangulation
  - Three ABB employees, not interviewed, but with good knowledge of the project, read through a draft of the report
- Internal validity – not relevant because of exploratory study
- External validity
  - Description of context believed to be relevant for the results
- Reliability
  - Interview guide and analysis procedures
  - Questionnaire and analysis procedures
  - Description of project documents and their analysis

ABB

Generalizing from Case Studies

is obtained through

1. Theory
   - The results of one study, possibly combined with existing literature, is used to build or refine a theory that should apply to other cases.
   - A previously developed theory is used to make predictions in a new situation.
2. Analytical generalization
   - (Good) argumentation for why an observed cause-effect should occur in other cases within the same scope of interest.

Summarizing Several Case Studies

- Requires a good description of the context of a case study to identify comparable cases.

- Unit of analysis and measures should (ideally) be based on a common ontology (common concepts).

- Meta-analysis is applicable when the studies have used comparable measures and report a quantitative measure of effect size.

- One technique for summarizing studies that do not depend on the above (but requires a common research question) is **vote-counting.** Using vote-counting, results are categorized into significant positive, significant negative or non-significant effect. Each study casts a vote.

Why Case Studies?

- Studying cases is important in itself; we learn from individual cases.

- Theories that predict an outcome can seldom be identified for topics involving human behaviour. Concrete, context-dependent knowledge is, therefore, more valuable than the search for predictive theories.

- More discoveries have arisen from intense observation of single cases than from statistics applied to large groups.

[**simula** . research laboratory]

Challenges

- There are few "good" case studies in software engineering

- Case studies often have a bad reputation because
    - case studies often **lack rigour**, and
    - the term **case study** is used for **many different types of studies,** and

Hence, it is difficult to agree on appropriate guidelines for case study research; it is difficult to judge the quality of case studies and it is difficult to replicate such studies.

Furthermore, case studies are suspected to be prone to **biased** results, but experience shows that refutation is more frequent in case studies than, for example, in experiments.

- case studies are **hard** to conduct

Yin: *"... we must all work hard to overcome the problem of doing case study research, including the recognition that some of us were not meant by skill or disposition, to do such research in the first place. Case study research is remarkably hard even though case studies have traditionally been considered 'soft' research, possibly because investigators have not followed systematic procedures."*

25/9 – 07   Bente Anda   29

[**simula** . research laboratory]

Challenges cont.

- There is much risk involved in depending on collaboration with industry, and data from industrial projects may be unreliable.

- When collaborating with industry, researchers have to be opportunistic and it is difficult to focus the research. It is also very difficult to be both consultant (expert helping industry) and researcher (evaluating technologies).

25/9 – 07   Bente Anda   30

[simula . research laboratory]

**Example Case Study 7:
Practical Issues of the ABB case study**

- Difficult to get the project members to spend sufficient time on providing data for the case study as the individual project members may not benefit much from the study.

- Difficult to agree on the expectations and time schedule,.

- Breadth vs. depth:
  – The company expected the researchers to provide help on many aspects of their software development.
  – Good research on the other hand is characterized by focusing in-depth on one specific topic

ABB

[simula . research laboratory]

Exercise 2

Identify differences between

- case studies and experiments, and

- case studies and surveys.

(minimum 5 for each bullet above, differences may be related to:
  – research question
  – unit of analysis
  – differences in data collection
  – differences in analysis)

Exercise 3

Look at

Carolyn B. Seaman and Victor R. Basili: An Empirical Study of Communication in Code Inspection.

And identify:

1. What are the research question(s)?

2. What kind of case study is it?

3. What is the unit of analysis?

4. From which sources was data collected?

5. How was data analyzed, using which techniques?

6. What are the results and what do you think of the validity of the results (and the quality of the study)?

Literature

- **Robert K. Yin. Case Study Research: Design and Methods. 3rd ed. Sage Publications, 2003.**

- Peta Darke, Graeme Shanks and Marianne Broadbent. Successfully completing case study research: combining rigour, relevance and pragmatism. Information Systems Journal, Vol. 8, pp. 273-289, 1998.

- Bent Flyvebjerg. Five Misunderstandings About Case-Study Research. Qualitative Inquiry, 12(2): 219-245, 2006.

- Bente Anda, Kai Hansen, Ingolf Gullesen and Hanne Kristin Thorsen. Experiences from Introducing UML-based Development in a Large Safety-Critical Project. Empirical Software Engineering, 11(4):555-581, 2006.

- **Bente Anda and Kai Hansen. A Case Study on the Application of UML in Legacy Development, In proceedings of ISESE'2006 (Fifth ACM-IEEE International Symposium on Empirical Software Engineering), ACM Press, Rio de Janeiro, Brasil, September 21-22, pp. 124-133, 2006.**