



Research Methodology (Software Engineering)

Dr Mohd Nazri Kama
UTM AIS
mdnazri@utm.my
www.ais.utm.my/nazrikama



Research Methodology vs. Research Method

- Research Methodology
 - way to systematically solve the research problem. It may be understood as a science of studying how research is done scientifically
 - what the activity of research is, how to proceed, how to measure progress, and what constitutes success
- Research Method
 - understood as all those methods/techniques that are used for conduction of research
 - Analysis of historical data
 - Analysis of documents
 - Interviews



Types of Purpose for Research

- Exploratory—finding out what is happening, seeking new insights and generating ideas and hypotheses for new research.
- Descriptive—portraying a situation or phenomenon.
- Explanatory—seeking an explanation of a situation or a problem, mostly but not necessary in the form of a causal relationship
- Improving—trying to improve a certain aspect of the studied phenomenon

 **Research Methodology Characteristics**

Methodology	Primary objective
Survey	Descriptive
Case study	Exploratory
Experiment	Explanatory
Action research	Improving

 **Example: Case Study**

- Calling names:
 - Field study (Lethbridge)
 - Observational study (Easterbrook)
 - case study and field study as observational methods (Zelkowitz and Wallace)
- Case study methodology is one of the many kinds of software engineering research

 **Types of CS**

- Positivist
 - A positivist case study searches evidence for formal propositions, measures variables, tests hypotheses and draws inferences from a sample to a stated population, i.e. is close to the natural science research model
- Interpretive
 - An interpretive case study attempts to understand phenomena through the participants' interpretation of their context, which is similar to Robson's exploratory and descriptive types
 - Software engineering case studies tend to lean towards a positivist perspective, especially for explanatory type studies

 **Case Study: Quali vs. Quanti**

- Case studies tend mostly to be based on qualitative data, as these provide a richer and deeper description. However, a combination of qualitative and quantitative data often provides better understanding of the studied phenomenon (Seaman 1999), i.e. what is sometimes called “mixed methods”
- Case studies - flexible design studies,
 - experiments and surveys are fixed design studies
 - fixed design process- all parameters are defined at the launch of the study,
 - flexible design process key parameters of the study may be changed during the course of the study

 **Criteria of Case Study**

- Has research questions set out from the beginning of the study
- Data is collected in a planned and consistent manner
- Explores a phenomenon, or produces an explanation, description, or causal analysis of it
- Threats to validity are addressed in a systematic way.

 **Case Study Process**

- Case study design & planning:
 - objectives are defined and the case study is planned
- Preparation for data collection:
 - procedures and protocols for data collection are defined
- Collecting evidence:
 - execution with data collection on the studied case
- Analysis of collected data
- Reporting

 **Case Study Process**

- ★ Case study design & planning:
 - Preparation for data collection:
 - Collecting evidence:
 - Analysis of collected data
 - Reporting

 **Case Study Design and Planning**

- Objective— what to achieve?
- The case—what is studied?
- Research questions— what to know?
- Methods— how to collect data?
- Selection strategy— where to seek data?
- Ethical consideration –regulatory to comply?

 **Objective - What to achieve?**

- The objective is initially more like a focus point which evolves during the study
- Derived from research problem
- SMART concept

 **The case—what is studied?**

- In software engineering, the case may be a software development project, which is the most straightforward choice. It may alternatively be an individual, a group of people, a process, a product, a policy, a role in the organization, an event, a technology, etc
- The project, individual, group etc. may also constitute a unit of analysis within a case.

 **Research Question - What to know?**

- The research questions state what is needed to know in order to fulfill the objective of the study.
- Similar to the objective, the research questions evolve during the study and are narrowed to specific research questions during the study iterations

 **Methods - How to collect data?**

- The principal decisions on methods for data collection are defined at design time for the case study, although detailed decisions on data collection procedures are taken later.
- Three categories of methods: direct (e.g. interviews), indirect (e.g. tool instrumentation) and independent (e.g. documentation analysis).

 Selection strategy - Where to seek data?

- Case selection is particularly important when replicating case studies.
- A case study may be literally replicated, i.e. the case is selected to predict similar results, or it is theoretically

 Ethical Consideration - Regulatory to comply?

- Informed consent
- Review board approval
- Confidentiality
- Handling of sensitive results
- Inducements
- Feedback

 Case Study Process

- Case study design & planning:
- ★ Preparation for data collection:
- Collecting evidence:
- Analysis of collected data
- Reporting

 **Collecting Data- Different Data Source**

- Different Data Sources
 - It is important to use several data sources in a case study in order to limit the effects of one interpretation of one single data source
- Data collection techniques can be divided into three levels
 - First degree: Direct methods means that the researcher is in direct contact with the subjects and collect data in real time
 - Second degree: Indirect methods where the researcher directly collects raw data without actually interacting with the subjects during the data collection

 **Collecting Data- Different Data Source**

- Third degree: Independent analysis of work artifacts where already available and sometimes compiled data is used. This is for example the case when documents such as requirements specifications and failure reports from an organization are analyzed or when data from organizational databases such as time accounting is analyzed

 **First vs. Second vs. Third**

- First degree methods are mostly more expensive to apply than second or third degree methods, since they require significant effort both from the researcher and the subjects.
- Advantage of the first and second degree methods is
 - large extent exactly control what data is collected, how it is collected, in what form the data is collected, which the context is etc.
- Third degree methods are mostly less expensive, but they do not offer the same control to the researcher;
 - hence the quality of the data is not under control either, neither regarding the original data quality nor its use for the case study purpose.

 **Case Study Process**

- Case study design & planning:
- Preparation for data collection:
- ★ Collecting evidence:
 - Analysis of collected data
 - Reporting

 **Collecting Evidence- Interview**

- Data collection through interviews is important in case studies.
- Interviews can, for example, be divided into
 - Unstructured
 - the interview questions are formulated as general concerns and interests from the researcher. In this case the interview conversation will develop based on the interest of the subject and the researcher.
 - semi-structured
 - fully structured interviews

 **Collecting Evidence- Interview**

- semi-structured
 - questions are planned, but they are not necessarily asked in the same order as they are listed. The development of the conversation in the interview can decide which order the different questions are handled, and the researcher can use the list of questions to be certain that all questions are handled
- fully structured interviews
 - In a fully structured interview all questions are planned in advance and all questions are asked in the same order as in the plan. In many ways, a fully structured interview is similar to a questionnaire-based survey.

UTM

Collecting Evidence- Interview

	Unstructured	Semi-structured	Fully structured
Typical focus	How individuals qualitatively experience the phenomenon	How individuals qualitatively and quantitatively experience the phenomenon	Researcher seeks to find relations between constructs
Interview questions	Interview guide with areas to focus on	Mix of open and closed questions	Closed questions
Objective	Exploratory	Descriptive and explanatory	Descriptive and explanatory

UTM

Collecting Evidence- Observation

- Investigate how a certain task is conducted by software engineers.
- Many different approaches for observation.
 - One approach is to monitor a group of software engineers with a video recorder and later on analyze the recording,
 - apply a "think aloud" protocol, where the researcher are repeatedly asking questions like "What is your strategy?" and "What are you thinking?" to remind the subjects to think aloud.
 - Observations in meetings is another type, where meeting attendants interact with each other, and thus generate information about the studied object.

UTM

Collecting Evidence- Archival Data

- Review meeting minutes, documents from different development phases, organizational charts, financial records, and previously collected measurements in an organization.
- Archival data is a third degree type of data that can be collected in a case study. For this
 - type of data a configuration management tool is an important source, since it enables the
 - collection of a number of different documents and different versions of documents

 **Case Study Process**

- Case study design & planning:
- Preparation for data collection:
- Collecting evidence:
- ★ Analysis of collected data
- Reporting

 **Data Analysis**

- Data analysis is conducted differently for quantitative and qualitative data.
 - quantitative data: includes analysis of descriptive statistics, correlation analysis, development of predictive models, and hypothesis testing. All of these activities are relevant in case study research.
 - Descriptive statistics, such as mean values, standard deviations, histograms and scatter plots, are used to get an understanding of the data that has been collected.
 - Correlation analysis and development of predictive models are conducted in order to describe how a measurement from a later process activity is related to an earlier process measurement

 **Qualitative**

- The basic objective of the analysis is to derive conclusions from the data, keeping a clear chain of evidence
- The chain of evidence means that a reader should be able to follow the derivation of results and conclusions from the collected data
- Analysis must be carried out in parallel with the data collection since the approach is flexible and that new insights are found during the analysis

 **Qualitative**

- There are two different parts of data analysis of qualitative data,
 - hypothesis generating techniques
 - Hypothesis generation is intended to find hypotheses from the data
 - Examples of hypotheses generating techniques are “constant comparisons” and “cross-case analysis”
 - hypothesis confirmation techniques
 - to confirm that a hypothesis is really true
 - Negative case analysis tries to find alternative explanations that reject the hypotheses

 **Tools**

- There are specialized software tools available to support qualitative data analysis, e.g.
- NVivo and Atlas. However, in some cases standard tools such as word processors and spreadsheet tools are useful when managing the textual data

 **Validity**

- The validity of a study denotes the trustworthiness of the results, to what extent the results are true and not biased by the researchers’ subjective point of view.
- It is, of course, too late to consider the validity during the analysis. The validity must be addressed during all previous phases of the case study.

 **Validity Aspect**

- **Construct validity:**
 - to what extent the operational measures that are studied really represent what the researcher have in mind and what is investigated according to the research questions. If, for example, the constructs discussed in the interview questions are not interpreted in the same way by the researcher and the interviewed persons, there is a threat to the construct validity
- **Internal validity:**
 - when causal relations are examined.
 - When the researcher is investigating whether one factor affects an investigated factor there is a risk that the investigated factor is also affected by a third factor. If the researcher is not aware of the third factor and/or does not know to what extent it affects the investigated factor, there is a threat to the internal validity

 **Validity Aspect**

- **External validity:**
 - This aspect of validity is concerned with to what extent it is possible to generalize the findings, and to what extent the findings are of interest to other people outside the investigated case. During analysis of external validity, the researcher tries to analyze to what extent the findings are of relevance for other cases.
- **Reliability:**
 - This aspect is concerned with to what extent the data and the analysis are dependent on the specific researchers. Hypothetically, if another researcher later on conducted the same study, the result should be the same.

 **Case Study Process**

- Case study design & planning:
- Preparation for data collection:
- Collecting evidence:
- Analysis of collected data
- ★ Reporting

 **Reporting**

- A set of characteristics which a case study report should have tell what the study was about communicate a clear sense of the studied case provide a "history of the inquiry" so the reader can see what was done, by whom and how.
- Provide basic data in focused form, so the reader can make sure that the conclusions are reasonable
- Articulate the researcher's conclusions and set them into a context they affect

 **Quantitative??**

- Please propose based on your research..

 **END**
