

Case Study Preparation and Data Collection



**Idaho State
University**

Computer
Science

Isaac Griffith

CS 6620

Department of Informatics and Computer Science
Idaho State University

ROAR



Outline

- Case study preparation
- Data sources and collection techniques
- Interviews
- Observation
- Archival data
- Metrics



Case Study Preparation

- When preparing a case study we need to:
 - Consider the number of data sources
 - Types of data sources
 - Methods of data collection



Multiple Data Sources

- Multiple data sources
 - Reduce the effects of a single interpretation from a single data source
 - sharing the same conclusion result, via triangulation, a stronger conclusion than a single data source
- Additionally, we should take into account
 - multiple roles of subjects
 - differences between subjects/situations



Data Collection Techniques

- First Degree: Direct method placing researchers with the subjects
 - Interviews
 - Focus Groups
 - Surveys
 - Observations
- Second Degree: Indirect collection of raw data without interacting with subjects
 - logging the use of tools
 - Observation via video
- Third Degree: Independent analysis of artifacts or compiled data
 - Archival data
 - Metrics
- Cost/Effort
 - First Degree - Most expensive
 - Third Degree - Least expensive
- Control
 - First and Second Degree
 - High control of data collection
 - High control of form of data collected
 - Third Degree
 - Low control of data collection
 - Low control of data quality



Interviews

- **Idea:** Researcher asks a series of questions to a set of subjects about the area of interest of the study
- Interview questions are based on research questions
 - **Open Questions** - allow for a broad range of answers
 - **Closed Questions** - limit answers to a set of known alternatives
- Interview types:
 - **Unstructured** - General concerns and interests
 - **Semi-structured** - Planned questions, no specific order
 - **Fully structured** - Planned questions, specific order



Interview Planning

- Decide whom to interview
 - focus on subject differences rather similarities
 - involve different roles, personalities, etc.
- Select number of interviewees
 - sufficient such that interviews reach **saturation**
 - that is no new info is learned by end of interviews

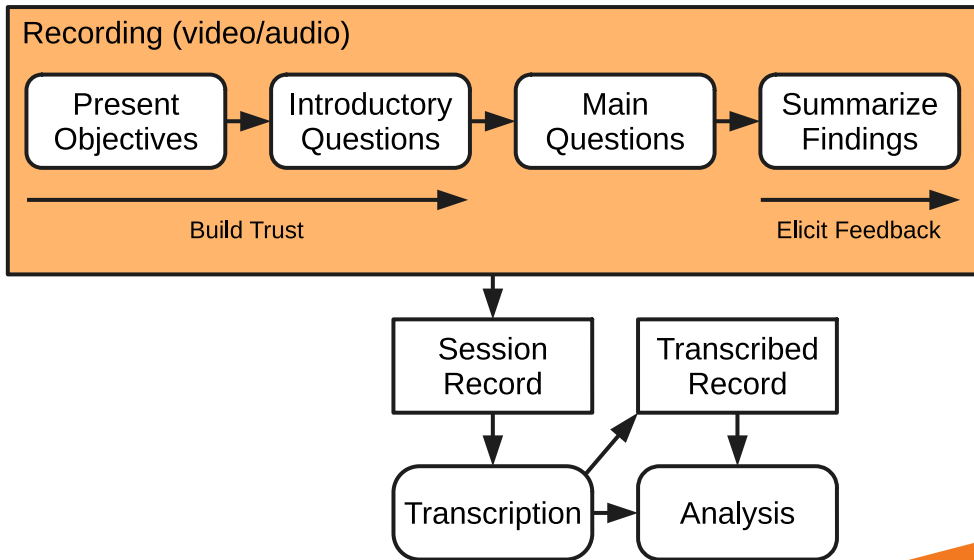


Interview Types

	Unstructured	Semi-structured	Fully Structured
Typical foci	How individuals quantitatively experience the phenomena	How individuals qualitatively and quantitatively experience the phenomena	Research 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



Interview Session Process





Interview Session Process

- Present Objectives - explain how the data will be used
- Introductory Questions - simple to answer questions about the background of the subject
- Main Questions - largest part of interview
 - ensure the confidentiality of the interviewee
- Summarize Findings - summarize key findings and elicit feedback to avoid misunderstandings



Observations

- Observations provide a deeper understanding of a phenomena
- Sources of observational data:
 - “Think aloud protocols”
 - Audio or Video recordings
 - Keylogging
 - Observing meetings



Observation Approaches

	High awareness of being observed	Low awareness of being observed
High degree of interaction by the researcher	Category 1	Category 2
Low degree of interaction by the researcher	Category 3	Category 4

- Category 1 or 2 – Action research or Ethnographic Studies
- Category 1 vs 2 – researches are seen as “observing” rather as “normal” participants
- Category 3 – Researcher seen only as a researcher
- Category 4 – Subjects observed using second degree methods such as video recordings



Archival Data

- A third degree type of data
- Refers to data such as
 - Meeting minutes
 - Development documentation
 - Organizational charts
 - Financial records
 - Previously collected measurements
- Must keep in mind that the data was not originally intended for the study



Metrics

- Quantitative data is also important
- Can be defined for the study or be repurposed data
- Data explicitly for the study is more flexible and more likely suitable for answering research questions
 - Utilize the GQM
- If repurposed data is used need to possibly handle the problem of missing data



Are there any questions?