



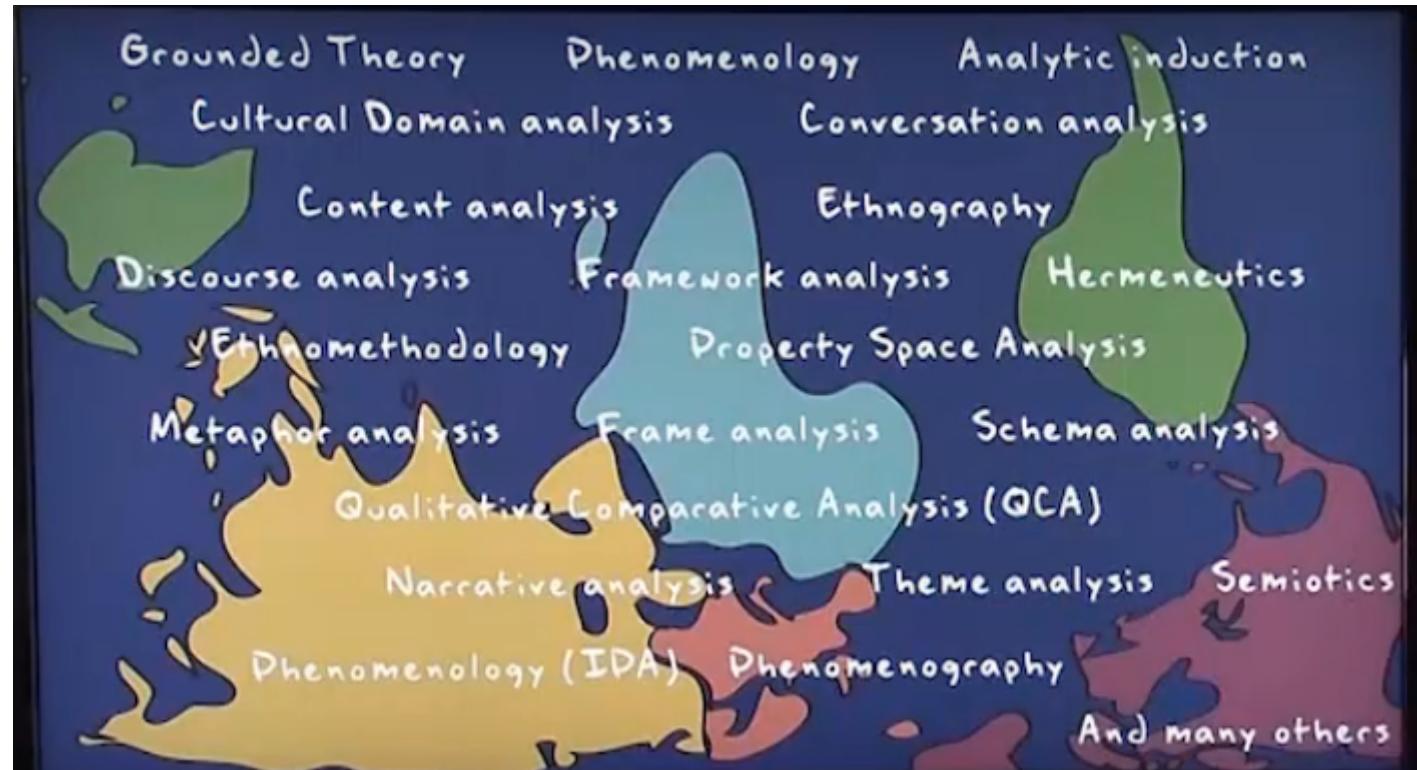
# Module 5 - Session 1 - Qualitative Data Analysis

Working effectively with data

CivicDataLab

2021/10/28 (updated: 2021-10-28)

# Qualitative Data Analysis



# Content Analysis - The Process

**Step 1** - Conducting the Interview

# Content Analysis - The Process

**Step 1** - Conducting the Interview

**Step 2** - Transcribing interviews

# Content Analysis - The Process

**Step 1** - Conducting the Interview

**Step 2** - Transcribing interviews

**Step 3** - Assigning Codes to each interview

# Content Analysis - The Process

**Step 1** - Conducting the Interview

**Step 2** - Transcribing interviews

**Step 3** - Assigning Codes to each interview

**Step 4** - Analysing Codes

# Content Analysis - The Process

**Step 1** - Conducting the Interview

**Step 2** - Transcribing interviews

**Step 3** - Assigning Codes to each interview

**Step 4** - Analysing Codes

**Step 5** - Repeat the process

# Coding in Qualitative Data Analysis

<sup>1</sup> Mrs. Jackson rises from her desk and announces, “OK, you guys, let’s get lined up for lunch. Row One.” Five children seated in the first row of desks rise and walk to the classroom door. Some of the seated children talk to each other.

<sup>2</sup> Mrs. Jackson looks at them and says, “No talking, save it for the cafeteria.

<sup>3</sup> Row Two.” Five children seated in the second row of desks rise and walk to the children already standing in line.

<sup>1</sup> LINING UP FOR LUNCH

<sup>2</sup> MANAGING BEHAVIOR

<sup>3</sup> LINING UP FOR LUNCH

# What is a code

A code in qualitative inquiry is most often a word or short phrase that symbolically assigns a summative, salient, essence -capturing, and/or evocative attribute for a portion of language-based or visual data<sup>1</sup>

[1]An Introduction to Codes and Coding

# Codes and Coding

1. Creating the Codes is an **iterative process** for a qualitative researcher.

# Codes and Coding

1. Creating the Codes is an **iterative process** for a qualitative researcher.
2. They will often **change, lump together, split or re organize** codes as they go through more and more data.

# Codes and Coding

1. Creating the Codes is an **iterative process** for a qualitative researcher.
2. They will often **change, lump together, split or re organize** codes as they go through more and more data.
3. The first round of coding can be **very open-ended**, but the codes can get more specific in additional rounds.  
For E.g. A conversation between two people can be tagged as *A-B* in the first round, *A-B - Access to Courts* in the second and *A-B - Access to Courts - Financial Hurdles* in the third

# Codes and Coding

1. Creating the Codes is an **iterative process** for a qualitative researcher.
2. They will often **change, lump together, split or re organize** codes as they go through more and more data.
3. The first round of coding can be **very open-ended**, but the codes can get more specific in additional rounds.  
For E.g. A conversation between two people can be tagged as *A-B* in the first round, *A-B - Access to Courts* in the second and *A-B - Access to Courts - Financial Hurdles* in the third
4. Later rounds involve **conflating codes** that might mean the same thing, **relating codes** to one another, and **eliminating codes** that no longer make sense.

# Grounded Theory - Case Study

How agile teams make self-assignment work<sup>1</sup>

[1] How agile teams make self-assignment work: a grounded theory study

# Research - Making self assignment work

## Features

1. Data collected through interviews with 42 participants representing 28 agile teams from 23 software companies and supplemented these interviews with observations.
2. Coding procedures used - **Open**, **Axial**, and **Selective coding**
3. Grounded Theory builds on research question which is open ended and drives the direction of research.

# Research - Making self assignment work

## Features

1. Data collected through interviews with 42 participants representing 28 agile teams from 23 software companies and supplemented these interviews with observations.
2. Coding procedures used - **Open**, **Axial**, and **Selective coding**
3. Grounded Theory builds on research question which is open ended and drives the direction of research.

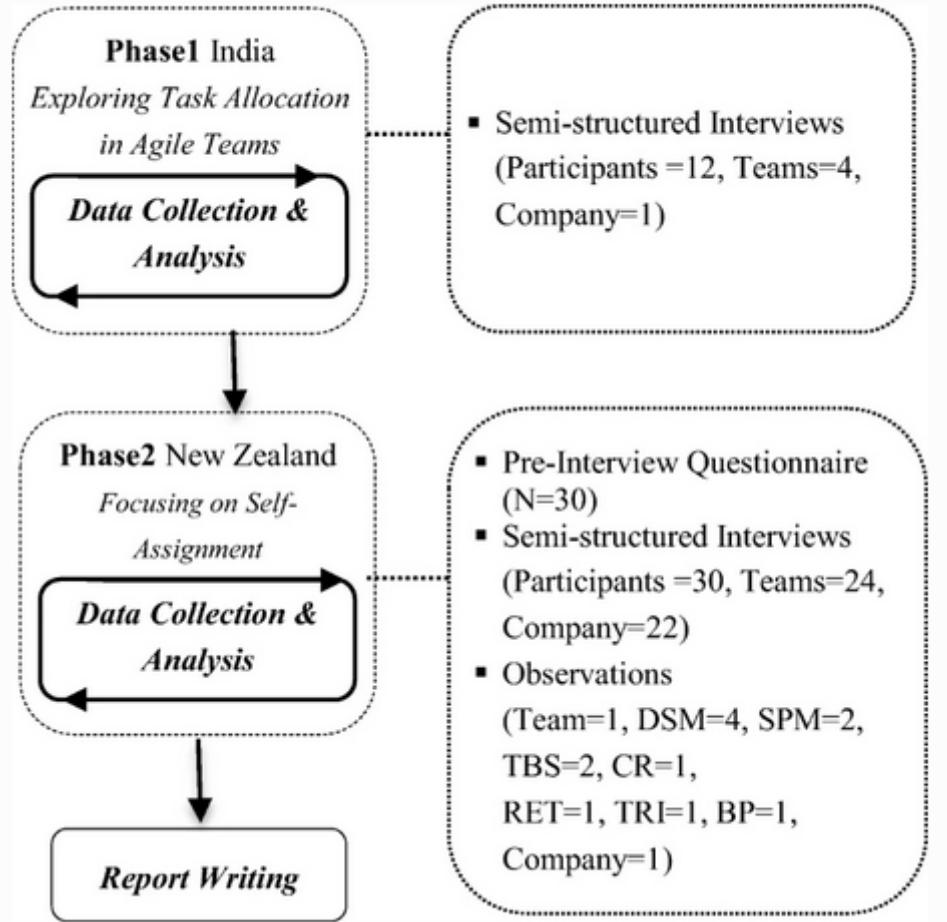
## Results

To understand how self-assignment works, the study shares insights on:

1. Context
2. Causal conditions that give rise to the need for self-assignment
3. A set of facilitating conditions that mediate how self-assignment may be enabled
4. A set of constraining conditions that mediate how self-assignment may be constrained
5. Strategies applied by agile teams to make the central phenomenon, self-assignment, work.

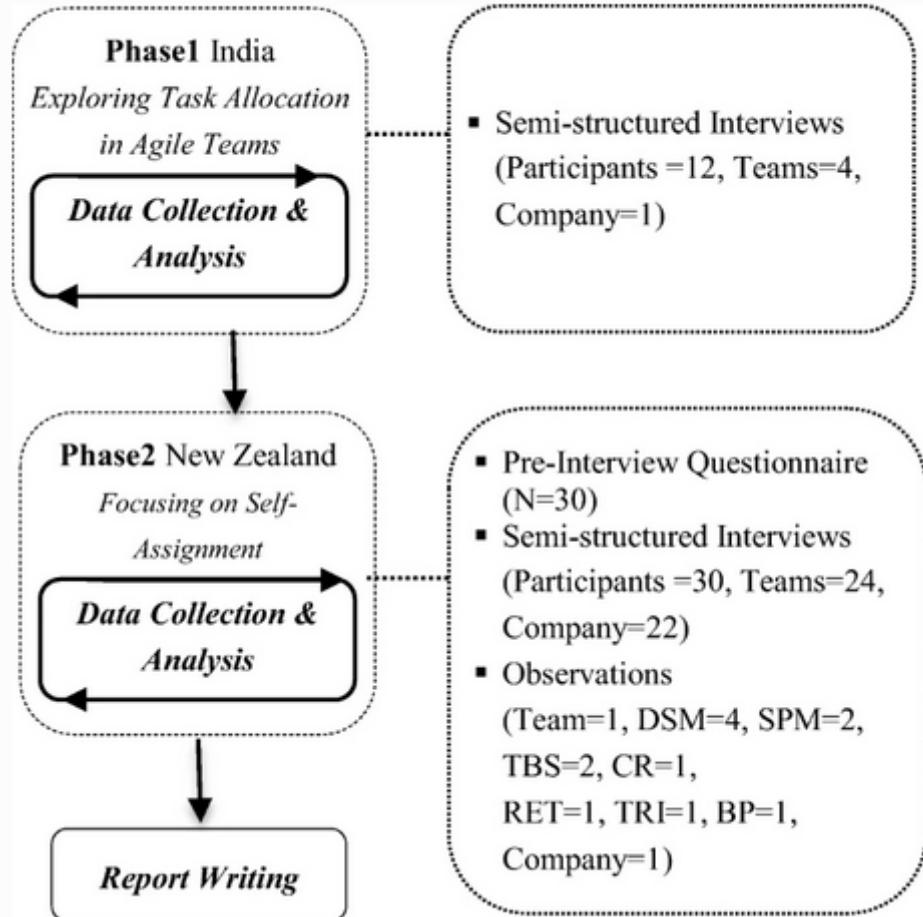
# Process

Fig. 1



# Process

Fig. 1



## Note

1. Iterative data collection. Data collection, followed by interviews, followed by data collection.
2. Each phase includes **multiple iteration of data collection and analysis**
3. Data for each round was analysed before collecting more data until the point of **theoretical saturation**
4. The objective of Phase 1 is **different** from Phase 2
5. The primary data sources for phase1 were *face-to-face interviews* and for phase2 were *pre-interview questionnaires, face-to-face semi-structured interviews, and team observations*

# Phase 1 - Data Analysis

1. The data collected helped in **developing an initial understanding** of task allocation in agile teams.

# Phase 1 - Data Analysis

1. The data collected helped in **developing an initial understanding** of task allocation in agile teams.
2. We applied open coding on participants' transcribed interview responses. During open coding, we **labelled the data with short phrases** that summarize the main key points.

# Phase 1 - Data Analysis

1. The data collected helped in **developing an initial understanding** of task allocation in agile teams.
2. We applied open coding on participants' transcribed interview responses. During open coding, we **labelled the data with short phrases** that summarize the main key points.
3. These were **further condensed into two to three words**, captured as codes.

# Phase 1 - Data Analysis

1. The data collected helped in **developing an initial understanding** of task allocation in agile teams.
2. We applied open coding on participants' transcribed interview responses. During open coding, we **labelled the data with short phrases** that summarize the main key points.
3. These were **further condensed into two to three words**, captured as codes.
4. As a result of data analysis, **different concepts from similar codes emerged**, one the most prominent of which was task allocation through self-assignment. Others included manager-driven, manager-assisted, team-driven, and team-assisted task allocation.

# Phase 1 - Data Analysis

1. The data collected helped in **developing an initial understanding** of task allocation in agile teams.
2. We applied open coding on participants' transcribed interview responses. During open coding, we **labelled the data with short phrases** that summarize the main key points.
3. These were **further condensed into two to three words**, captured as codes.
4. As a result of data analysis, **different concepts from similar codes emerged**, one the most prominent of which was task allocation through self-assignment. Others included manager-driven, manager-assisted, team-driven, and team-assisted task allocation.
5. The **results of phase1 directed us to focus on self-assignment** as the substantive area of the study in the next phase.

# Phase 2 - Data Analysis



# Process Terminologies

Terminologies	Strauss and Corbin Definition	Our study
Phenomenon	The actions of an individual as well as interactions between different people revolve around the phenomenon	Making self-assignment work represents the phenomenon
Causal Conditions	Events, incidents, happenings that lead to the occurrence or development of a phenomenon	Causes leading to self-assignment, reasons why teams adopt self-assignment
Intervening Conditions	The conditions that intervene the strategies taken within a specific context	Facilitating and constraining conditions influencing self-assignment
Context/Contextual Conditions	Context represents the particular set of conditions within which the action/interactional strategies are taken	The specific set of conditions within which the strategies to make self-assignment work occur
Action/Interactional Strategies	Strategies devised to manage, handle, carry out, and respond to a phenomenon under a specific set of perceived conditions	Adopted strategies to work around challenges of self-assignment
Consequences	Outcomes or results of action and interaction	Impact of adopting these strategies to make self-assignment work

# Applying the coding paradigm to data collected



# Building a theory - Selective Coding

1. In Selective Coding, we **started building a storyline** presenting the essence of our study where each sub-category and category captured a part of the whole story of making self-assignment work.

# Building a theory - Selective Coding

1. In Selective Coding, we **started building a storyline** presenting the essence of our study where each sub-category and category captured a part of the whole story of making self-assignment work.
2. How agile teams make self-assignment work emerged as the **most prominent and central phenomenon** from our data analysis process

# Building a theory - Selective Coding

1. In Selective Coding, we **started building a storyline** presenting the essence of our study where each sub-category and category captured a part of the whole story of making self-assignment work.
2. How agile teams make self-assignment work emerged as the **most prominent and central phenomenon** from our data analysis process
3. It was during the selective coding, we confirmed which **relational phrases** such as 'mediates', 'overcome by', 'give rise to' were fitting well to our entire theory model

# Building a theory - Selective Coding

1. In Selective Coding, we **started building a storyline** presenting the essence of our study where each sub-category and category captured a part of the whole story of making self-assignment work.
2. How agile teams make self-assignment work emerged as the **most prominent and central phenomenon** from our data analysis process
3. It was during the selective coding, we confirmed which **relational phrases** such as 'mediates', 'overcome by', 'give rise to' were fitting well to our entire theory model
4. It was also during the selective coding, when **theoretical saturation was reached** and no new concepts, categories or insights were identified.

# Building a theory - Selective Coding

1. In Selective Coding, we **started building a storyline** presenting the essence of our study where each sub-category and category captured a part of the whole story of making self-assignment work.
2. How agile teams make self-assignment work emerged as the **most prominent and central phenomenon** from our data analysis process
3. It was during the selective coding, we confirmed which **relational phrases** such as 'mediates', 'overcome by', 'give rise to' were fitting well to our entire theory model
4. It was also during the selective coding, when **theoretical saturation was reached** and no new concepts, categories or insights were identified.
5. Then, finally we **revisited and refined the categories to make sense of the entire theory** explaining the phenomenon.

# The state of QDA tools

1. Not many good open source tools available

# The state of QDA tools

1. Not many good open source tools available
2. Some popular tools are NVIVO, Atlas TI, MAX QDA, etc. Links to other similar tools are available [here](#)

# The state of QDA tools

1. Not many good open source tools available
2. Some popular tools are **NVIVO**, **Atlas TI**, **MAX QDA**, etc. Links to other similar tools are available [here](#)
3. All of these offer way more features than what is usually needed in an analysis

# The state of QDA tools

1. Not many good open source tools available
2. Some popular tools are **NVIVO**, **Atlas TI**, **MAX QDA**, etc. Links to other similar tools are available [here](#)
3. All of these offer way more features than what is usually needed in an analysis
4. Might be difficult to combine the output from these tools with other tools when doing quantitative data analysis, data visualisations, etc.

# The state of QDA tools

1. Not many good open source tools available
2. Some popular tools are **NVIVO**, **Atlas TI**, **MAX QDA**, etc. Links to other similar tools are available [here](#)
3. All of these offer way more features than what is usually needed in an analysis
4. Might be difficult to combine the output from these tools with other tools when doing quantitative data analysis, data visualisations, etc.
5. Open Source Tools like **qcoder**, **Taguette**, etc are available but are still under active development and might not be stable yet.

# Open Data Publication

# What can be published

1. Raw datasets sourced by mining public websites

# What can be published

1. Raw datasets sourced by mining public websites
2. RTI responses

# What can be published

1. Raw datasets sourced by mining public websites
2. RTI responses
3. Aggregated statistics

# What can be published

1. Raw datasets sourced by mining public websites
2. RTI responses
3. Aggregated statistics
4. Data collected using primary methods like surveys, etc

# What can be published

1. Raw datasets sourced by mining public websites
2. RTI responses
3. Aggregated statistics
4. Data collected using primary methods like surveys, etc
5. The process of data collection

# What can be published

1. Raw datasets sourced by mining public websites
2. RTI responses
3. Aggregated statistics
4. Data collected using primary methods like surveys, etc
5. The process of data collection
6. Metadata - Information about the data

# What can be published

1. Raw datasets sourced by mining public websites
2. RTI responses
3. Aggregated statistics
4. Data collected using primary methods like surveys, etc
5. The process of data collection
6. Metadata - Information about the data
7. Observations and Experiences about working with data

# How to publish

1. Choosing a License under which data will be published. [List](#) of open data licences.

# How to publish

1. Choosing a License under which data will be published. [List](#) of open data licences.
2. Prepare a data dictionary, or a [data biography](#)

# How to publish

1. Choosing a License under which data will be published. [List](#) of open data licences.
2. Prepare a data dictionary, or a [data biography](#)
3. Explicitly mention how other users can cite the authors/publishers when using the dataset. [Example](#)

# How to publish

1. Choosing a License under which data will be published. [List](#) of open data licences.
2. Prepare a data dictionary, or a [data biography](#)
3. Explicitly mention how other users can cite the authors/publishers when using the dataset. [Example](#)
4. Publish raw data and processed data as separate files and detail out the entire process one needs to follow to prepare the processed files from the raw datasets

# How to publish

1. Choosing a License under which data will be published. [List](#) of open data licences.
2. Prepare a data dictionary, or a [data biography](#)
3. Explicitly mention how other users can cite the authors/publishers when using the dataset. [Example](#)
4. Publish raw data and processed data as separate files and detail out the entire process one needs to follow to prepare the processed files from the raw datasets
5. Include a document that mentions about the current state of data, update frequency, data-gaps, project-scope, etc.

# How to publish

1. Choosing a License under which data will be published. [List](#) of open data licences.
2. Prepare a data dictionary, or a [data biography](#)
3. Explicitly mention how other users can cite the authors/publishers when using the dataset. [Example](#)
4. Publish raw data and processed data as separate files and detail out the entire process one needs to follow to prepare the processed files from the raw datasets
5. Include a document that mentions about the current state of data, update frequency, data-gaps, project-scope, etc.
6. Mention about all stakeholders involved in the entire process - collaborators, funders, etc.

# How to publish

1. Choosing a License under which data will be published. [List](#) of open data licences.
2. Prepare a data dictionary, or a [data biography](#)
3. Explicitly mention how other users can cite the authors/publishers when using the dataset. [Example](#)
4. Publish raw data and processed data as separate files and detail out the entire process one needs to follow to prepare the processed files from the raw datasets
5. Include a document that mentions about the current state of data, update frequency, data-gaps, project-scope, etc.
6. Mention about all stakeholders involved in the entire process - collaborators, funders, etc.
7. Publish all relevant code that is needed to mine or process the datasets.

# How to publish

1. Choosing a License under which data will be published. [List](#) of open data licences.
2. Prepare a data dictionary, or a [data biography](#)
3. Explicitly mention how other users can cite the authors/publishers when using the dataset. [Example](#)
4. Publish raw data and processed data as separate files and detail out the entire process one needs to follow to prepare the processed files from the raw datasets
5. Include a document that mentions about the current state of data, update frequency, data-gaps, project-scope, etc.
6. Mention about all stakeholders involved in the entire process - collaborators, funders, etc.
7. Publish all relevant code that is needed to mine or process the datasets.
8. A central repository to access all publications/datasets is useful.

# The JALDI Portal

The portal was built with the aim of making available data on the judiciary to researchers, policy makers and the general public who are interested in understanding the working of the judiciary.

# The JALDI Portal - Challenges and Workarounds

# Queries and Feedback