

Grounded Theory in Software Engineering Research: A Critical Review and Guidelines

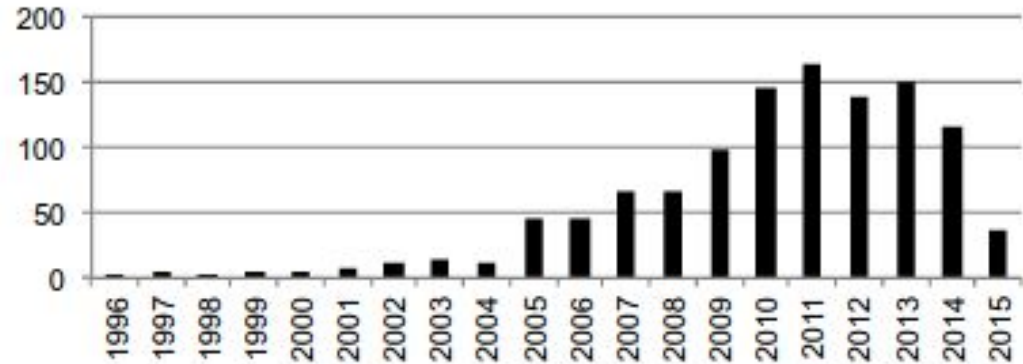
Alexandre Alphonsos



Introduction

Introduction

- SE must consider social, cultural and human aspects of software construction.
- Qualitative research methods.
- GT popularity growth.



Research Question

What is the state of practice of grounded theory research in software engineering?

What is GT?

A method of inductively generating theory from data.

Traditionally you create a hypothesis and put it to test, in GT you generate a hypothesis.

Often focus on unstructured text but can handle all type of data.

- Limit exposure to literature: Avoid confirmation bias
- Treat everything as data.
- Immediate and continuous data analysis: No need to finish collection to begin analysis.
- Theoretical sampling: Find more data to fill the gaps
- Theoretical sensitivity: Establish relationships between concepts, be creative.
- Coding: No preconceived scheme.
- Memoing: Write memos with notes, diagrams, sketches.
- Constant comparison: Iterate over data until the theory 'fit' the data.
- Memo/Theoretical sorting: Oscillate between memos to consolidate.
- Cohesive theory: Move beyond categories and coding.
- Theoretical saturation: New data does not trigger revision

Philosophical Foundations and Versions

- Positivism vs. Interpretivism.
- Versions: Classic/Glaserian, Straussian, Constructivist.

Research Design

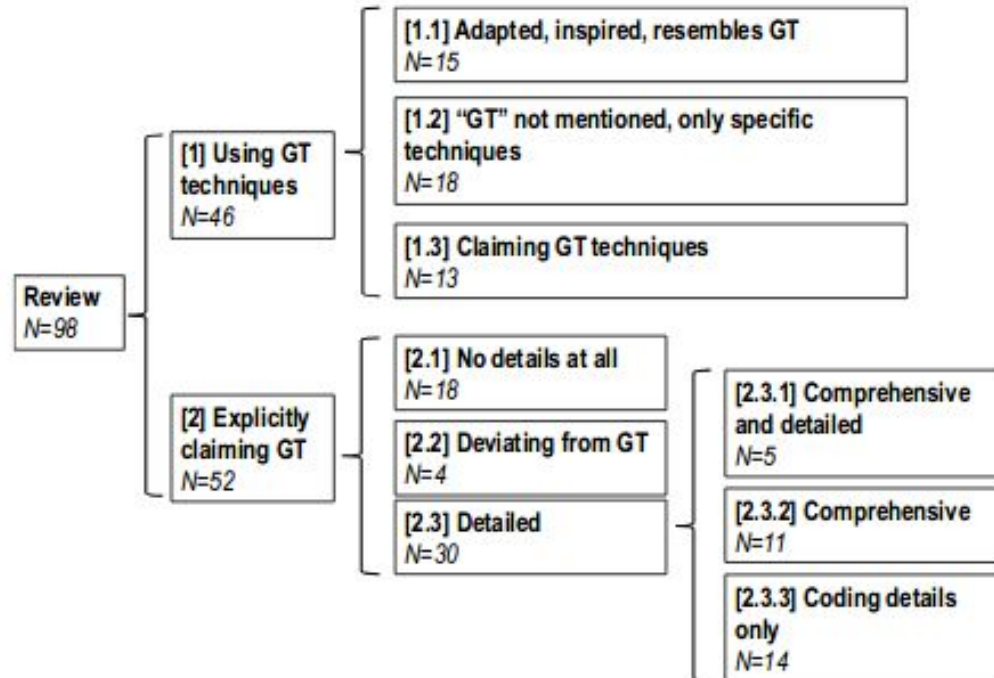
- Search String: “grounded theory” AND “software engineering”
- Filter well-know journals

Journal	Articles
Information and Software Technology	42
Journal of Systems and Software	16
IEEE Transactions on Software Engineering	11
Empirical Software Engineering	10
Software Process: Improvement and Practice ^a	8
Journal of Software: Evolution and Process ^b	4
Software Quality Journal	3
ACM Trans. Software Engineering and Methodology	3
Journal of Software Maintenance and Evolution:	1

Analysis and Results

- GT “Use” is Ambiguous.

Analysis and Results



Analysis and Results

- Many studies present little to no detail.

Analysis and Results

Practice	Papers reporting
GT Practice details reported	30
Simultaneous data collection and analysis	13
Data sources and collection	29
Theoretical sampling	12
Coding	29
Memoing	12
Memo sorting	4
Constant comparison	13
Theoretical saturation	15

Discussion - Method slurring

- Confer legitimacy
- Avoid literature review
- Simplicity
- Sugestion

Discussion - Challenges

- Large amounts of data
- Unconventional types

Conclusion

- GT is complex.
- Highly suitable methodology.
- SE literature is missuse.