Thematic-LM: A LLM-based Multi-agent System for Large-scale Thematic Analysis

Tingrui Qiao, Caroline Walker,
Chris Cunningham, Yun Sing Koh

Presented by Marko Morrison

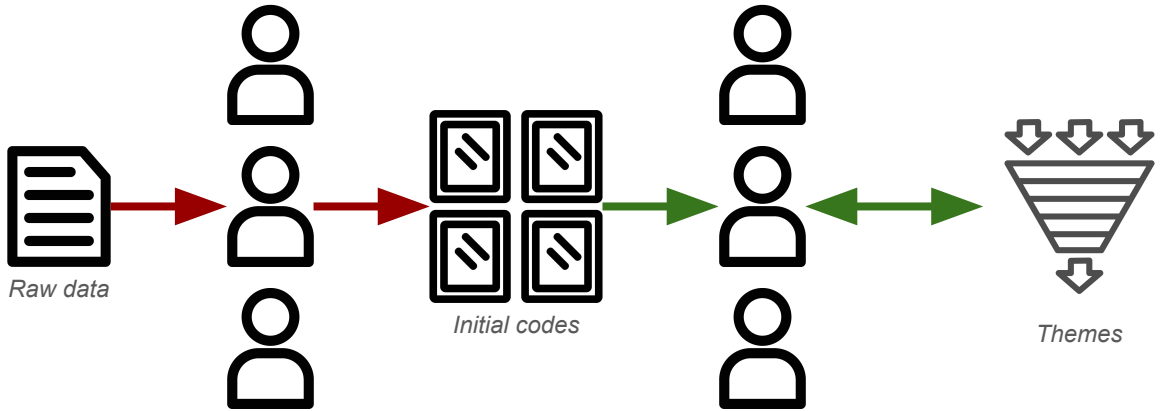
Agenda

- What is thematic analysis?
- Application of LLMs for thematic analysis
- Limitations of LLMs for thematic analysis
- Proposal: Thematic-LM
- Evaluation criteria
- Findings

What is thematic analysis?

“A qualitative method for identifying underlying meanings within unstructured text.”

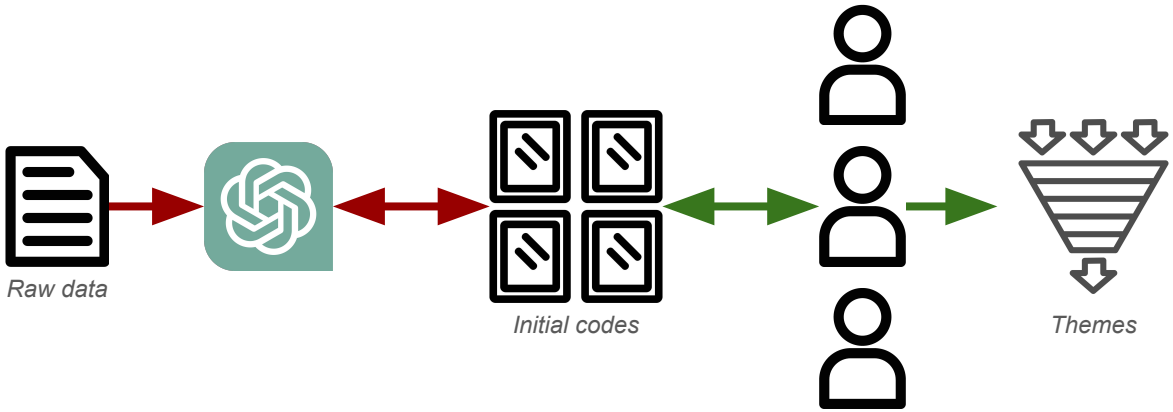
Phase
1. Familiarizing yourself with your data:
2. Generating initial codes:
3. Searching for themes:
4. Reviewing themes:
5. Defining and naming themes:
6. Producing the report:

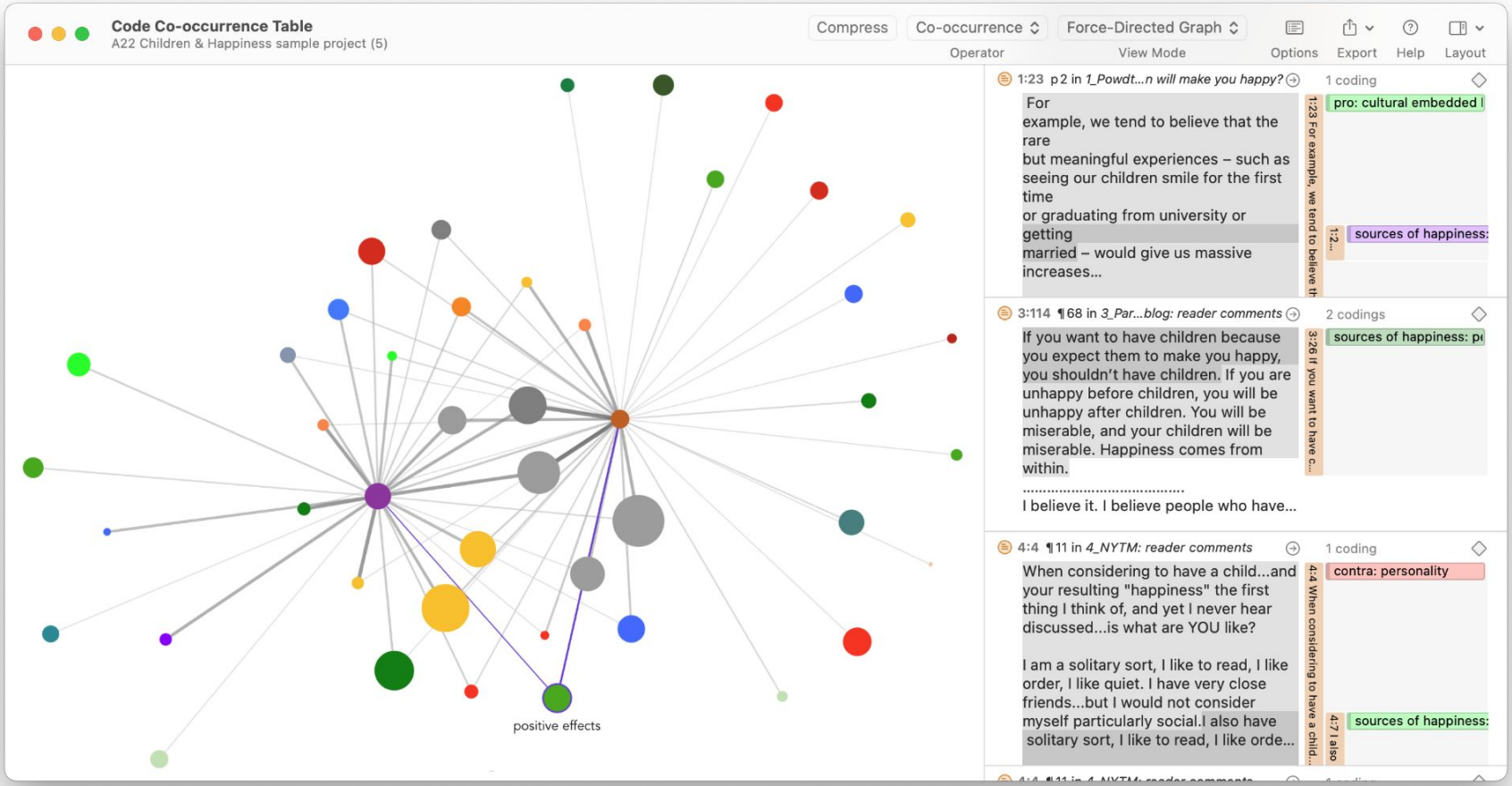


Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative research in psychology, 3(2), 77-101.

Usage of LLMs for qualitative coding shows promise

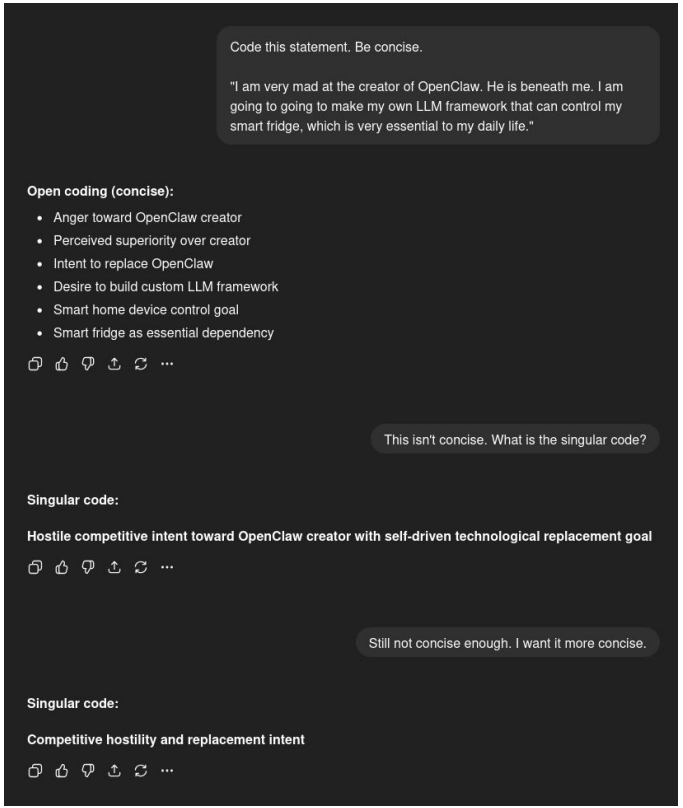
1. LLM proposes initial codes
2. Human revises
3. Repeat until human is satisfied





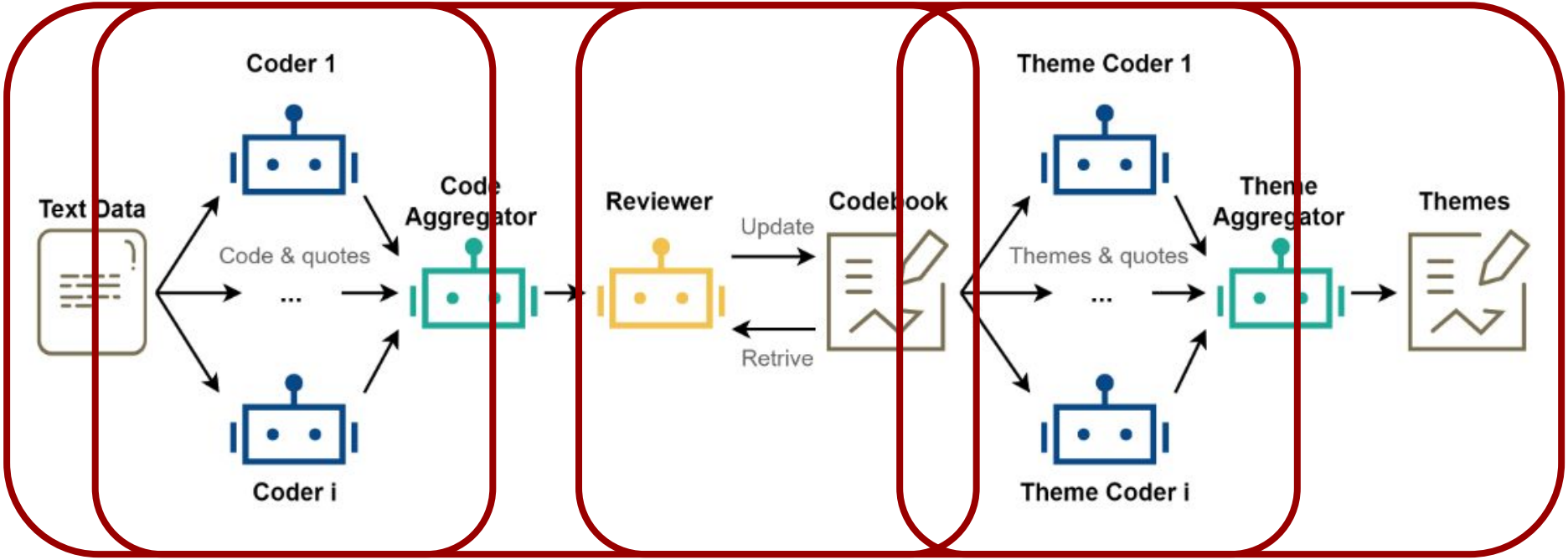
https://atlasti.com/what-s-new-in-atlas-ti-23

Current applications of LLMs for qualitative coding have foundational limitations



1. HiTL is infeasible when coding large datasets
2. HiTL causes the LLM to reflect human biases
3. LLMs do not revisit previously coded data - sequential approach
4. Subjective nature of QA promotes diversity in coders. LLMs are NOT diverse.

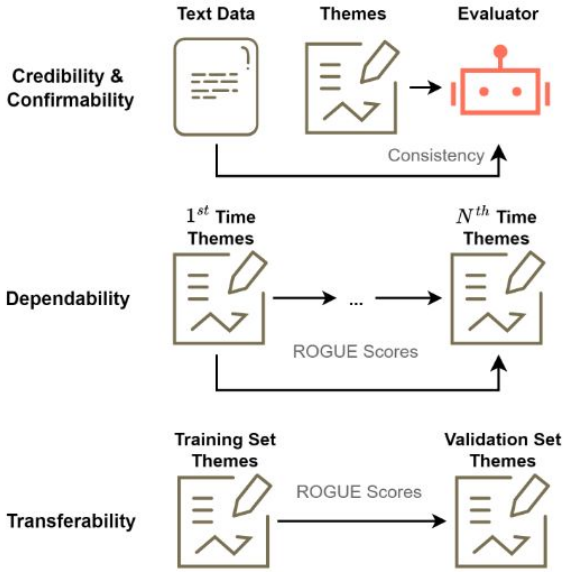
Solution: Thematic-LM



(4) (1, 2) (3) (4)

How do we evaluate how well coding/TA went?

- Credibility & Confirmability
  - Do the codes represent the data? Are the codes actually based on the data (vs. biases)?
  - Checked via LLM-as-judge
- Dependability
  - If another researcher follows the same process, will they arrive at the same conclusions?
  - Tested via another researcher performing TA several times and measuring the overlap via “ROGUE scores”
  - Unclear if this is another LLM or a person.
- Transferability
  - Do the generated codes make sense for other data?
  - Tested via training/validation sets



Example: Theme output for Dreaddit

Table 1: Themes and description of the themes produced by Thematic-LM on the Dreaddit dataset, compared with topic modelling and sentiment analysis results.

Theme	Description	Related Topic	Sentiment Score
Navigating Emotional and Mental Health Challenges	The experiences of individuals dealing with anxiety, mental health concerns, and the impact of emotional stressors on daily life.	therapist_to_therapy	0.00
Impact of Economic Stress and Resource Scarcity	The financial struggles and efforts to manage limited resources and economic instability.	homeless_the_to_have	0.10
Familial Dynamics and Responsibilities	The complex relationships within families, the burden of roles, and the tension between obligations and emotional needs.	food_you_for	0.15
Coping with Academic and Professional Pressures	The stress associated with academic performance and the pressures to succeed in professional life.	job_and_have	0.20
Interpersonal Conflict and Relationship Stress	The conflicts in personal relationships, whether with romantic partners, friends, or colleagues, the emotional and mental strain caused by unresolved tensions or disagreements.	sex_that_sexual	0.15
Seeking Validation and Emotional Support	The frequent attempts by individuals to gain reassurance or validation from others, often by sharing their struggles and emotions openly in the hope of receiving empathy or encouragement.	you_are_support	0.35
The Dichotomy of Online and Offline Identities	The differences between one's real and virtual personas and the impact of social media on identity and interactions.	he_him_me	0.25

Findings: Thematic-LM improved quality scores

Table 3: Comparison of the theme quality scores on the Dreaddit dataset.

Method	Credibility & Confirmability	Dependability	Transferability
Single	0.63	0.45	0.41
Single (Codebook)	0.75	0.61	0.67
System (1 Coder)	0.92	**0.81**	0.86
System (2 Coders)	**0.94**	0.78	**0.87**

Table 4: Comparison of the theme quality scores on the Reddit climate change dataset.

Method	Credibility & Confirmability	Dependability	Transferability
Single	0.66	0.56	0.73
Single (Codebook)	0.74	0.69	0.78
System (1 Coder)	0.96	0.84	**0.90**
System (2 Coders)	**0.98**	**0.86**	0.89

Findings: Diversity in coders led to new emergent themes

Table 5: Examples of themes not captured in Thematic-LM with no assigned coder identities but emerged in Thematic-LM where coders are assigned different identity perspectives.

Theme	Description
Economic Impact of Climate Policies	Concerns about the economic consequences of aggressive climate regulations, particularly their impact on industries and job markets.
Environmental Stewardship	The deep responsibility to protect and maintain the natural environment, viewing humans as caretakers of the earth.
Scepticism of Climate Science	Questioning about the extent of human influence on global warming and discuss whether climate change is part of the natural cycle.
Environmental Justice and Vulnerable Communities	Highlights the disproportionate impact of climate change on marginalized communities, advocating for policies that address environmental justice and protect vulnerable populations.

Findings: Diversity in coders led to new themes

Figure 4: Comparison of codes and themes generated by five coders with no identities given, with the same identities of "human-driven climate change" given, and with five different identities. The five different identities are agents which are instructed to believe in "human-driven climate change" and "climate change as a natural cycle" and instructed to act with "progressive view", "conservative view" and "Indigenous view", denoted as H, N, P, C, and I respectively. The pairwise ROGUE scores measure the differences between codes and themes from different agents.

Summary

- What is thematic analysis?
- Application of LLMs for thematic analysis
- Limitations of LLMs for thematic analysis
  - HiTL is limiting, promotes bias.
  - Plain LLMs aren’t diverse enough
- Proposal: Thematic-LM
  - Introduce LLM-based coders and aggregators
  - Assign identities to coders for diversity
- Evaluation criteria
  - Credibility, dependability, transferability
- Findings
  - Thematic-LM improved coding quality
  - Diverse identities induced new themes