CAUSAL MAP GARDEN

Causal Mapping as QDA

Contents

🌻 Intro

Chapter contents.

21 Sep 2025

> Causal mapping is also a kind of Qualitative Data Analysis (QDQ). How does that even work? This chapter explains.

See also !!Causal mapping is a simple yet powerful form of qualitative coding

Causal Mapping outputs not just codes but a model you can query to answer useful questions

Causal mapping is an interesting QDA approach which is very suitable for scaling with AI

! causal mapping turns QDA on its head

PAGES IN THIS CHAPTER

🌻 Causal Mapping outputs not just codes but a model you can query to answer useful questions

> **SOURCE NOTE (consolidation):** This file is a draft/fragment. The flagship QDA-facing paper is now: Causal mapping as causal QDA.
> Companion methods notes: Magnetisation; A simple measure of the goodness of fit of a causal theory to a text corpus.

In the first part of this series !!Causal mapping is a simple yet powerful form of qualitative coding, we argued that causal mapping is a simple yet powerful form of qualitative coding. We showed how its focus on identifying causal links reduces and simplifies the analytical task and provides a clear, structured approach which is relatively easy to apply. However, the true strength of causal mapping lies not just in the coding process itself, but in its output: a **query-able qualitative model**.

Qualitative Data Analysis (QDA) most typically produces a written report alongside a deeper understanding in the minds of the researcher or research team, and hopefully also in the readers. It may also provide other outputs like tables of frequencies or co-occurrences.

We already argued that causal mapping is an especially useful form of QDA because:

- It produces a **structured graph database of causal claims**. This can be viewed as a table or as a map.
- This output isn't a static summary; it's a **dynamic qualitative model** — a network of interconnected evidence that can be systematically interrogated using standardized, out-of-the-box filters and algorithms.
- The kinds of questions you can answer with a causal map are **particularly useful** because they are about *what causes what — as seen by your sources*.

From a List of Themes to a Network of Evidence

Traditional QDA often produces additional outputs like tables of theme frequencies alongside narrative summaries. It is possible to query these to answer questions like "what themes did the younger respondents most often mention when also talking about the main theme".

Causal mapping provides a particularly rich output: a **causal network** — a qualitative knowledge graph where the factors (which can be understood as themes) are nodes and the causal claims are the links connecting them. This structure is inherently machine-readable and ready for analysis.

Thinking of the output as a **model** is key. Just as a quantitative researcher builds a statistical model to explain relationships in their data, a causal mapper builds a qualitative model of the causal beliefs expressed in texts. This is not unique to our Causal Map app: all applications of causal mapping provide, more or less explicitly, this kind of model, going right back to (Axelrod 1976). This model can then be used to answer new questions, often without needing to go back to the original source texts, though the underlying quotes and context are always available.

Some standard ways to answer useful questions

Because the output is a structured network, we can apply a range of queries to explore the data. This gives us a library of **pre-existing approaches** to ask **practical questions** about the causal landscape described by the participants.

Here are some questions you can answer using causal mapping.

Relevant page: Individual questions — introduction ▸

A corresponding library of filters

The Causal Map app provides about 20 corresponding, ready-to-use filters to answer these kinds of questions, some based on existing causal mapping publications, some new.

In the list of typical queries above:

- Some queries have matching unique pre-defined outputs: a map with a specific filter applied, or a table.
- For some queries, there are different ways to answer them and/or the answer requires more than one filter.
- For a few queries we do not yet have a specific way to answer them in Causal Map.

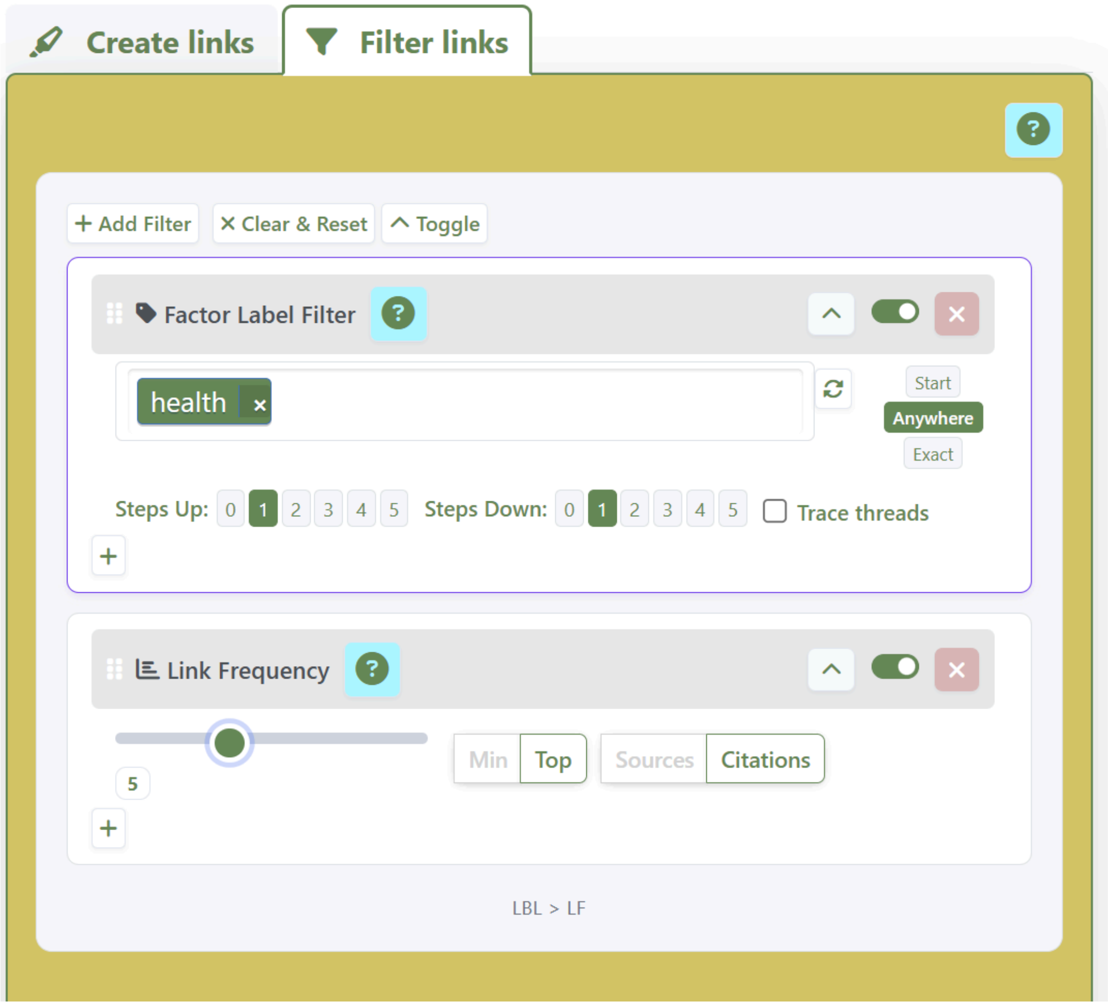
There are two types of filter:

- simple filters which select a subset of the links like links frequency, factor label filter; and path tracing
- "transform filters" like zoom, which temporarily rewrite the cause and/or effect labels.

Here as reference are more details: .

Chaining filters together

Most importantly **we can chain filters together**, to answer corresponding, composite questions like: *what are the most frequently mentioned upstream influences on these key outcomes, according to the* *younger respondents?*



Most of the core filters have nothing to do with AI or large language models. They are straightforward, transparent and deterministic. In the Causal Map app, they are available at a click and can be rearranged with drag and drop. But most of them are simple to reconstruct in a spreadsheet or a graph database, without using Causal Map at all.

Cutting to the chase with causal queries

This ability to systematically query **causal** evidence is what allows causal mapping to **cut to the chase**. For evaluators and researchers, **the core questions are very often about causation**: "Did the program work?" and "How did it work?". Causal mapping structures the evidence provided in the texts to help researchers answer these types of questions.

A Model of Causal Evidence, Not Causal Facts

It's crucial to be clear about one thing: Causal mapping is **not a method of causal inference**. It does not, on its own, tell you what truly causes what in the real world.

Instead, it creates a qualitative model of **causal claims**. The map organizes what people *said* about causation, allowing the researcher to weigh, compare, and synthesize this evidence. The logic we apply is one of evidence management:

- How much evidence is there for a link between X and Z?
- Is that link direct or indirect?
- Do different subgroups of people agree on this causal pathway?

Calling these causal claims "evidence" does not mean that anyone should necessarily believe it or that it has been verified[1]. It is simply raw material which we organise and inspect before drawing any conclusions. Researchers can choose to also code additional properties for each link and/or source such as "doubtful" or "reliable" or "verified" and include or exclude links by filtering on those properties.

If the researcher wants to make a causal judgement, they must interpret the map in context, examine the quotes and consider the source of the claims. The map is a powerful tool for structuring and clarifying that judgment.

In any case, **many colleagues use causal mapping not to make causal inferences but simply to understand** *what people think causes what*, and how, for example as a crucial prerequisite for planning policy, communications or interventions.

See also the other caveats we listed in the previous post: !!Causal mapping is a simple yet powerful form of qualitative coding.

Conclusion

The real power of causal mapping as a QDA method is that it produces a query-able, qualitative model of causal evidence. This structured output allows researchers and evaluators to apply a range of standardized algorithms to answer practically relevant questions about the causal mechanisms at play, as seen by the sources.

Many researchers and evaluators like using causal mapping to explore their data. In the third part of this series, we will explore how these properties — a simplified coding task and a structured, query-able output — make causal mapping especially suited for transparent and verifiable automation with AI.

---

1. Thanks to Stève Duchêne https://www.linkedin.com/in/steve-duchene/ for reminding us to clarify this. ↩

References

Axelrod (1976). *The Cognitive Mapping Approach to Decision Making*. In *Structure of Decision : The Cognitive Maps of Political Elites*.

Causal mapping is an interesting QDA approach which is very suitable for scaling with AI

25 May 2026

Abstract

This paper argues that AI can be used inside a qualitative workflow without becoming the analyst, provided it is based on a method whose coding task is narrow enough to be locally checkable. Causal mapping is such a method: the basic coded unit is a single causal claim made by a source, grounded in a verbatim quote. We describe a workflow in which the AI does one narrow extraction task, pairing each candidate causal link with the exact quote that supports it, leaving almost no room for a black box, while every other analytic decision remains with the human researcher. The vocabulary of factors, the structure of the codebook, the analytic pipeline and the narrative interpretation do not need to be delegated to the AI. The workflow is exploratory in spirit. The AI identifies candidate causal claims, each tied to a quote, and the interpretive work of reading, comparing, reorganising and challenging those candidates still belongs to the analyst. We argue that this division of labour is a clean answer to the call for AI-assisted qualitative work that remains rigorous, transparent and accountable, and that the small but established tradition of causal mapping is well placed to support it.

See also:

- Causal mapping as causal QDA (companion paper, more detail on the QDA framing)
- Minimalist coding for causal mapping (the coding stance)
- Combining opposites, sentiment (label-level extensions)

1. Introduction

The arrival of large language models in qualitative research is producing two reasonable responses. The first is a worry that AI will erode the interpretive practices that give qualitative work its value: close reading, reflexive judgement, a rejection of positivism. The second is enthusiasm for the prospect of working at scales that were previously impractical. Both responses are warranted. The interesting question is whether a workflow can be designed that takes the productivity gains seriously while preserving the methodological commitments.

We argue that one such workflow already exists in causal mapping, a well-established family of approaches in evaluation and applied social research for representing what influences what according to sources (Axelrod 1976; Eden 1988; Powell et al. 2024; Evaluation 2024). The argument has two steps. The first introduces causal mapping as a form of QDA in its own right, sometimes called causal QDA, with a single distinctive coding act and a structured intermediate product. The second shows why this coding act, narrow and locally checkable as it is, is well suited to bounded AI assistance without surrendering judgement to a black box.

The argument is addressed to a methodologically mixed audience: people interested in language, discourse and interaction, but also in the practical question of how to use AI in qualitative work without losing the qualitative.

We should be clear about scope. This is closer to a small-Q than a big-Q argument. We do not engage the wider question of what counts as the AI's positionality, or whether a model can be a participant in interpretation in the same sense that an analyst is. We are not proposing causal QDA as a heroic method for identifying new paradigms in qualitative inquiry. The contribution is more modest: a way of getting useful work done for researchers and evaluators whose questions are about what other people think causes what, where the alternative is either much slower manual coding or a black-box synthesis that is hard to audit.

The workflow does not replace the reflexive, interpretive work that gives big-Q qualitative research its distinctive contribution. It runs alongside that work, and it is most useful when the analyst still wants to read closely but also wants help finding the passages worth reading closely. The aim is to make a corpus explorable rather than to package it as a finding.

2. Causal mapping as practically focused QDA

In ordinary qualitative coding a code typically denotes a theme or concept. In causal mapping the basic coded unit denotes a causal claim made by a source: a cause label, an effect label, a verbatim quote that supports the claim, and a source identifier. A coding act yields an ordered pair, `Cause -> Effect`, attached to evidence and provenance. The dataset of such acts is a links table, and that links table is the core qualitative product rather than a stepping stone to a narrative.

This is a small unit, but by design. We do not code strength, polarity, necessity, sufficiency, role as moderator or any of the other features that systems-dynamics and grounded-theory traditions sometimes attach to links (Kim & Andersen 2012). Most respondents do not state those features explicitly, and most analysts cannot reliably extract them from text. By holding the coded unit to bare causation, we preserve the chain of evidence and keep the act of coding within reach of both human coders and AI assistance. We have called this stance minimalist or barefoot coding (see Minimalist coding for causal mapping).

The links table can be queried directly. Every node is a factor that appears as a cause or effect in at least one claim. Every edge is a claim with a quote behind it. Natural questions include: which factors are the most frequently mentioned upstream influences on a given outcome; how pathways into a target factor differ across subgroups; which links are contested, with both `X -> Y` and `X -> ~Y` appearing in different sources. Each question is answered by an explicit, reversible operation on the table, and each answer remains traceable to the underlying quotes. The full taxonomy of operations and worked analysis pipelines is developed in a companion methods paper (see Causal mapping as causal QDA); here we keep just enough of the methods to make the AI argument.

This makes causal QDA practically focused in a specific sense. The research question is causal, the coded unit is a causal claim, and the analysis is an explicit pipeline of operations on a structured intermediate product. Compared to a thematic analysis (Braun & Clarke 2021), it gives up some breadth of interpretive scope in exchange for a much tighter audit trail from any map back to the original quotes. For research questions about drivers, barriers, mechanisms and pathways, the trade is usually a good one; for worked examples in evaluation practice see (Remnant et al. 2025; Powell et al. 2025).

2.1 A worked example

If an interviewee says,

> After the clinic started opening on Saturdays I did not have to miss work, so I could actually attend.

a thematic pass might code `Access`, `Clinic opening hours`, `Employment constraints` and `Attendance`. A causal-coding pass records two links:

- `Saturday opening -> Not missing work`
- `Not missing work -> Attendance`

Both representations have value. Only the causal one is queryable as a mechanism. A reader can ask which other factors point into `Attendance`, which contexts mention `Saturday opening`, which paths run from clinic operations to attendance, and so on, and every answer remains line-by-line accountable to the original quotes.

3. Where AI normally goes wrong in qualitative work

The standard worry about AI in qualitative analysis is about synthesis tasks. "Find the main themes in this corpus" or "Summarise what these interviewees say about X" are open-ended prompts whose output is sensitive to the model's implicit theory of what counts as a theme. The output may read well and may even be roughly right, but it is hard to audit. The model may have downweighted minority views, smoothed over disagreements, drifted off the text, or invented categories that fit its training distribution better than the data. Worse, when the analyst inspects the output, the apparent fluency tends to disarm scrutiny. Or — it might not have wandered off. But how would we know?

There is a second failure mode that is less often noticed. Even when the AI is asked to do something more constrained, such as "code each excerpt against this codebook", the inputs are usually long enough that the model is making implicit selection judgements about what to attend to. Two passes over the same text can produce different codings, not because the model is faulty, but because the task as posed leaves enough latitude for it to vary.

These observations are not arguments against using AI in qualitative work. They are arguments against giving the AI the analytic burden. Once we accept that AI does best on narrow tasks with locally checkable outputs, the question becomes: which qualitative methods have a coding act with that shape?

4. The causal coding task is the right shape for AI

The minimalist coding act in causal mapping is exactly that kind of narrow, locally checkable task. The instruction to the model is:

> Identify each passage where the text says that one thing influenced another. For each, record the cause, the effect and the exact quote that supports the claim.

This instruction has several useful properties. It refers to features of the text that are already present, namely explicit causal claims. It produces outputs whose unit, a link with a quote, is easy to verify by reading the quote. It does not ask the model to weigh, summarise or theorise. Because each link is a separate unit, errors are local: a wrong link can be removed or corrected without unravelling the rest of the analysis.

In practice we run this extraction on short chunks of text, often a single passage at a time, and we recover the corpus-level structure by aggregating the resulting links. The model is not asked to hold the whole corpus in attention or to decide what matters across documents. It does the same small extraction job, repeatedly, on chunk after chunk.

When the output is wrong we usually do not correct individual links. We go back to the prompt, adjust it, recode and iterate. Because the unit cost of a coding pass is low and the output is locally checkable, this experimentation is cheap, and the analyst stays in control of what the AI is being asked to do.

The extracted links are candidates. They are raw material for analysis, ready to be read closely, compared, followed up or set aside. The analyst's job is to do that work: to follow surprising claims, drop dull ones, reorganise the vocabulary where useful, and use the resulting map as an instrument for exploring the corpus rather than as a finding to report.

This way of working is a qualitative variant of the split-apply-combine strategy ({wickham 2011). The split is the minimalist coding act; the apply step is a deterministic pipeline of operations on the links table; the combine step is human-authored synthesis. The clean mapping between strategy and method is what makes the AI question tractable.

5. AI as clerk, human as architect

A useful division of labour follows. The AI is a clerk: fast, consistent, willing to apply a stable rule across thousands of passages, and locally accurate enough for the rule to be worth applying (n.d.). The human is the architect, responsible for everything that requires judgement.

The architect's work includes, at minimum, the following.

- **Framing the research question.** What kind of causal claims matter for this study, and at what level of detail.
- **Choosing the factor vocabulary.** We curate the labels rather than letting the model invent freely. Where the model proposes labels, we cluster and rewrite using a separate operation called magnetisation (see Magnetisation), which is itself a deterministic transform on the links table.
- **Designing the hierarchy.** Where some factors are best read as instances of more general ones, we encode this in the labels themselves using a `;` separator. This is a coding-time decision, not an AI inference.
- **Marking opposites.** Where the data contains paired factors such as employment and unemployment, we mark them as opposites using a tilde convention and let an explicit transform combine evidence across poles when useful. This too is a human-controlled label-level decision (see Combining opposites, sentiment).

- **Building the analytic pipeline.** Which filters to apply, which evidence thresholds to set, which paths to trace, how to render the resulting view. Every step is explicit and reproducible.
- **Writing the interpretation.** The narrative account of what the evidence shows, including its limits, the contested claims and the cases that do not fit.

None of this is delegated to the model in our workflow. The consequence is that the AI never produces an analytic claim. It produces candidate evidence in a fixed format, which the human then assembles into claims.

5.1 What is and is not in the AI's hands

It is useful to be very explicit about the boundary, because this is where critiques of AI-assisted qualitative work tend to land. In our workflow the AI sees one chunk of text at a time and is asked to extract causal claims from that chunk. It does not:

- choose which research questions matter,
- choose which sources to include,
- decide what counts as a factor in this study,
- decide where opposites or hierarchies are warranted,
- decide which paths through the map are interesting,
- summarise across documents,
- write any part of the final account.

It does:

- propose candidate cause and effect labels for each apparent causal claim in the chunk,
- pair each candidate link with the exact quote that supports it,
- optionally flag uncertain cases for human review.

Every analytic move in the rest of the workflow is either the analyst's direct authorship or a deterministic transform whose code can be inspected. The AI is one bounded component of a much larger, mostly non-AI process.

6. What auditability looks like in practice

A causal map produced this way carries a complete audit trail. Pick any edge in the final map, click through, and you see the bundle of underlying claims, each with its source, its verbatim quote and its raw extracted labels. You can also see which transforms have been applied between the raw coding and the rendered view: which sources were filtered, which labels were rewritten, which evidence thresholds were imposed.

This is the property the special-issue call describes as "accountable links between analytic claims and empirical materials". It is not unique to causal mapping, but the structure of the workflow makes it easy to deliver. Every analytic claim in the final paper reduces to a sequence of explicit operations on a links table whose rows are quote-grounded extractions. A reader who doubts the analysis can rerun the pipeline, change a single step and observe the consequences. A coder who disagrees with a particular extracted link can reject it without disturbing the rest of the analysis.

The contrast with black-box synthesis is sharp. An AI-generated thematic summary is, in effect, a single opaque step from corpus to conclusion. The links-table workflow replaces that opaque step with a sequence of explicit, locally checkable ones, with the AI confined to the first and narrowest of them.

The same audit trail is what makes the map useful for exploration. Picking an edge and reading the underlying quotes is often where analysis begins rather than where it ends. The workflow does not aim to produce a finished diagram and stop. It aims to make a corpus tractable enough that the analyst can decide where to read closely, where to push back, and where to look for the cases that do not fit.

7. Ethical and practical transparency

The Text & Talk call asks authors to declare the type of AI application used, its role in the workflow, and how analytic accountability was maintained. The workflow described here gives clean answers to those three questions.

- **Type.** Commercial large language models are used at the extraction step. Specific model versions are reported, because output behaviour changes with model versions and a replication needs to know which version produced which links.
- **Role.** The model is used only to propose candidate links from short text chunks, each paired with a verbatim quote. It is not used to summarise across documents, to choose the codebook, to write the report, or to make any claim about the world.
- **Accountability.** Every link in the final analysis is traceable to a specific quote and source. Every analytic step beyond extraction is deterministic and human-authored. The links table is an artefact that can be shared with reviewers and co-researchers and inspected line by line.

There are practical ethical considerations beyond declaration. Where source material is sensitive (interview transcripts, clinical narratives, evaluation data from vulnerable groups), the choice of model matters, because some commercial services may train on inputs or store data in ways that conflict with consent and confidentiality commitments. We use settings that disable training on inputs where available, prefer providers with explicit data-handling commitments, and remove personally identifying information at the chunking stage where possible. Consent processes should make AI use explicit, including the kind of model, the role it plays and the safeguards applied.

There is also a softer accountability point. Because the prompt is the codebook in another form, sharing the prompt is sharing the method. We make our extraction prompts available as part of the analytic record. A reviewer can read the prompt and form a view about whether it framed the extraction task fairly.

8. Limits

The workflow does not solve every qualitative research problem, and we do not claim that it should.

- **Non-causal meaning.** Some valuable content in texts is not causal: identity work, norms, emotions, metaphors, turn-by-turn interactional structure. Causal coding ignores these, in the same way that conversation analysis ignores other features. For research questions about how participants account for what makes a difference, causal coding is on point. For research questions about how participants do identity work or manage stance in interaction, it is not.
- **Claims are not facts.** A coded link records what a source claimed. It does not establish what is true in the world. The map is a structured record of evidence, and the move from evidence to truth requires the same epistemic care it would in any qualitative study.
- **Frequency is not effect size.** Counts of sources and citations measure how widely a claim is made in the corpus.

They do not measure the magnitude of any underlying causal effect.
- **The transitivity trap.** If Source A says `Training -> Knowledge` and Source B says `Knowledge -> Adoption`, it is tempting to read a path from `Training` to `Adoption`. Unless within-source thread tracing is imposed and reported, that stitches together a mechanism that no one in the corpus actually claimed. Reasoning over paths is one of the things the links table makes easy, which means it is also one of the things to be careful with.
- **The AI does not eliminate coder bias.** It relocates the bias into the prompt. The advantage of the workflow is that the prompt is an explicit, shareable artefact whose effects can be inspected and revised, but it is still doing the work that a human codebook would otherwise do.
- **Attention is local.** Because the model sees one chunk at a time, it cannot pick up cross-references or implicit qualifications that span chunks. Expanding the context window helps but is slow and expensive. For research questions where cross-document context is the point, this workflow is not on its own enough.

9. Relation to neighbouring traditions

Causal QDA is compatible with thematic analysis and qualitative content analysis (Mayring 2000). It can be used as a front-end that surfaces a causal layer of meaning, with thematic work running alongside for non-causal features. It differs in that it stores ordered pairs and not isolated categories, which makes downstream pathway analysis possible.

In relation to conversation analysis and discourse analysis, causal QDA is doing something different. It is interested in the content of what participants say causes what, rather than in how that talk is constructed turn by turn. The two are complementary rather than competing. A single research project can perfectly well combine a fine-grained interactional analysis of a subset of episodes with a corpus-wide causal-mapping pass.

Narrative analysis is closer. Both approaches take participants' causal accounts seriously as objects of study in their own right, whether or not those accounts are independently verifiable. Causal QDA is one structured way of building a corpus-wide picture of the causal claims that narrative analysts read closely.

A different recent proposal for AI-assisted QDA replaces coding with structured conversation between the analyst and the model (Friese 2025). That is an interesting move for some kinds of interpretive work. The workflow we describe goes in the opposite direction: it preserves a strong intermediate representation, a quote-grounded links table, precisely so that the AI never holds the analytic state in its own working memory. We do not see the two approaches as exclusive, and a hybrid is plausible, but the trade-offs are different.

10. Conclusion

The AI question in qualitative work tends to be framed at the wrong level. Asked whether to trust a model with thematic synthesis or with the interpretation of a corpus, most qualitative researchers reasonably refuse. Asked whether to use a model for one narrow extraction job at which it is locally accurate and locally checkable, while every other analytic decision stays in human hands, the answer can be yes. The workflow described here is what that yes looks like in practice: an AI that stays within the narrow limits given it, a human who never delegates judgement, and a links table that anyone can read in the middle.

Causal mapping earns its modest place in that workflow because the unit of coding is small, the intermediate product is structured, and the chain from any analytic claim back to the supporting quote is short and inspectable (Ackermann & Maytorena-Sanchez 2024; Narayanan 2005). Those are also the properties that make AI assistance honest. The fuller methods treatment of causal mapping as a member of the QDA family is given separately (see Causal mapping as causal QDA); here we have used just enough of it to support the AI argument.

References

Ackermann, & Maytorena-Sanchez (2024). *Overlooked and Underused? The Benefits and Challenges of Using Causal Mapping for Project Studies*. https://doi.org/10.1016/j.plas.2024.100161.

Axelrod (1976). *The Analysis of Cognitive Maps*. In *Structure of Decision : The Cognitive Maps of Political Elites*.

Braun, & Clarke (2021). *Thematic Analysis : A Practical Guide*. SAGE Publications Ltd. https://www.torrossa.com/it/resources/an/5282292.

Eden (1988). *Cognitive Mapping*. https://doi.org/10.1016/0377-2217(88)90002-1.

Evaluation (2024). *Causal Mapping*. https://www.betterevaluation.org/methods-approaches/methods/causal-mapping.

Friese (2025). *Conversational Analysis with AI - CA to the Power of AI: Rethinking Coding in Qualitative Analysis*. https://doi.org/10.2139/ssrn.5232579.

Kim, & Andersen (2012). *Building Confidence in Causal Maps Generated from Purposive Text Data: Mapping Transcripts of the Federal Reserve*. https://doi.org/10.1002/sdr.1480.

Mayring (2000). *Qualitative Content Analysis*. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research. https://doi.org/10.17169/FQS-1.2.1089.

Narayanan (2005). *Causal Mapping: An Historical Overview*. In *Causal Mapping for Research in Information Technology*. https://www.google.co.uk/books/edition/_/61z36j6QgmAC?hl=en&gbpv=1.

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

Powell, Cabral, & Mishan (2025). *A Workflow for Collecting and Understanding Stories at Scale, Supported by Artificial Intelligence*. SAGE PublicationsSage UK: London, England. https://doi.org/10.1177/13563890251328640.

Remnant, Copestake, Powell, & Channon (2025). *Qualitative Causal Mapping in Evaluations*. In *Handbook of Health Services Evaluation: Theories, Methods and Innovative Practices*. https://doi.org/10.1007/978-3-031-87869-5_12.

{wickham (2011). *The Split-Apply-Combine Strategy for Data Analysis*. https://doi.org/10.18637/jss.v040.i01.

🌻 Saturation, scale and attention

📅 9 Nov 2025

"Reached saturation" mostly means "I've found out about as much on this topic as will fit in any human's head, and my supervisor is happy" doesn't it?
And yes, AIs can in many senses fit more in their heads than humans can. But how many answers do you want?

But just because we can interview 80 thousand people, does that mean we should?

What would we do with all that information? There is a competition for attention.

And actually not just attention in the sense of some kind of narrow spotlight we shine on one thing and then another (out of many), but something like "engagement" — understanding, making/sharing sense, even caring.

I think many of us have come across the phenomenon where an AI produces a perfectly adequate or even compelling narrative analysis of something but our eyes kind of glaze over and it's hard to care or follow.

This has nothing to do with the ability of the language model, it's got to do with how its products or outputs fit into our world. It wouldn't be any different if the analyses are produced by humans, angels, aliens or LLMs.

Of course it's perfectly possible to give an LLM more information about who we are and what we care about, and ask it to produce results which fit us better, it can do that, up to a point. But we sometimes still struggle to care. Perhaps because we didn't get our hands dirty enough writing the report, or don't have enough skin in the game. That is not a limitation of the LLMs, or of the angels or aliens. It's a by-product of the fact that we can now get almost free, mostly adequate, sometimes even astounding, results to a completely overwhelming range of questions.

What is the Point of Us. A Sci-Fi Story for Researchers

Related

- chapter intro

🌻 Brief review of S Friese – Conversational Analysis to the Power of AI

📅 7 Oct 2025

> **SOURCE NOTE (consolidation):** This is supporting material. The flagship QDA-facing paper is now: Causal mapping as causal QDA.
> Companion methods notes: Magnetisation; A simple measure of the goodness of fit of a causal theory to a text corpus.

This interesting article (Friese 2025) proposes a methodological shift for qualitative data analysis (QDA) that moves beyond traditional coding by introducing **Conversational Analysis with AI (CAAI)**. This approach can be realised by using Dr. Friese's own software, QInsights, replacing the process of coding — segmenting and labelling data — with a structured, dialogic interaction between the researcher and a large language model (LLM).

The Analytical Process

The method uses a five-step process that focuses on synthesis instead of coding.

- **Step 1: Get to know the data.** The researcher uses the AI to create summaries and initial themes to identify key topics for exploration.
- **Step 2: Prepare for analysis.** The researcher picks a topic from the previous step and writes a set of questions to guide the AI conversation. This question list replaces a traditional coding frame and makes the analysis transparent. CAAI analysis proceeds topic by topic. This contrasts with conventional coding frame development.
- **Step 3: Ask questions.** Using the prepared questions, the researcher has a dialogue with the AI about a small subset of the data (e.g., 4-6 interviews). This helps the researcher find patterns and explore surprising findings.
- **Step 4: Synthesize insights.** The researcher slows down, reads the text of the conversation and writes a synthesis of the findings. This can be done alone or collaboratively with the AI. Steps 3 and 4 are repeated for each topic.
- **Step 5: Elevate the analysis (Optional).** The researcher can use the AI to help connect findings to broader theories.

How LLMs are Characterized

The author views Large Language Models (LLMs) not as intelligent beings but as useful analytical partners. Their value comes from being trained on vast amounts of human-written, "socially-situated corpora". The article states that models lack true human understanding or lived experience, a concept referred to as "Seinsverbundenheit". However, their outputs are still insightful because they reflect the patterns and "lifeworlds" present in their training data. The AI’s knowledge is described as a "collective and distributed" echo of human meaning-making. In this role, the LLM acts as not only in a deductive and inductive fashion but also as an "abductive catalyst"—a tool that surfaces unexpected connections and provokes new ideas for the researcher, without needing to be intelligent itself.

This approach reframes the researcher's role from a coder to a conductor of an analytic conversation, prioritizing interpretation and synthesis over mechanical categorization. Following (Krähnke et al. 2025), Friese characterises this process as hermeneutical — an approach in which meaning emerges through the researcher engaging in an open interaction with a text, in this case with the addition of a third party, the AI, which provokes and questions the process. "Meaning does not reside in the data itself but is generated by a recursive relationship between the analyst, the context, and the text" (p. 6).

Rigor and trustworthiness are established not through coding frames or inter-coder agreement, but through the transparency of the documented dialogue, traceability to source data via retrieval-augmented generation (RAG) systems (provision of quotes), and the somewhat replicable nature of the question sets that guide the inquiry. Ultimately, CAAI presents a post-coding paradigm where analysis emerges directly from a dynamic and reflexive engagement with the data, mediated by AI.

In the course of the article, Friese revisits the text analysis she conducted for her own PhD and candidly concludes that she could have done better, and much faster, using the method she proposes.

Strengths

Reflection on use of LLMs

Firstly, I'd congratulate Dr. Friese for presenting a **deeply argued explanation of and justification** for the QInsights workflow. As researchers, however we use LLMs, it is crucial that we continue to reflect on what we are doing, rather than letting the muscles of our critical thought go weak through lack of exercise.

LLM-supported QDA as hermeneutics

Secondly, the advent of generative AI is a very good time to ask again: do we really need coding to make QDA rigorous? As she says: "qualitative research is entering a **moment of methodological experimentation**". After all, rigorous coding alone was never enough to make text analysis rigorous, not least because coding can ignore context and the positionality of the researcher and the research task.

Re-positioning LLM-supported QDA in the broader context of hermeneutics is useful and refreshing.

Dangers of consumer-facing LLMs

Thirdly, she reminds researchers that **methodological rigour is very hard to achieve with unreflective use of consumer-facing LLMs** like ChatGPT, not least because they will skim-read the text to find a quick answer and will too often grab a possibly incorrect answer from their training data rather than ground it in the actual text. What's more, a platform like QInsights can provide a completely documented workflow, where every conversation and decision made is recorded — although we have no insight into the AI's 'thought processes' behind each of its contributions to this workflow beyond what it tells us.

A new kind of epistemic actor between human and machine

I usually find arguments about whether generative AIs are "really" conscious or can "really" understand something as simply spurious, and no more interesting than arguing about whether a computer can *really* "copy" or "save" or "read" or even "predict" something. However, Friese has a really interesting angle on this, an angle which takes us further rather than trying to police useful language. She '**treats AI as a new kind of epistemic actor** — neither a mere stochastic parrot nor a conscious knower, but as something else: a dialogic partner capable of generating insight through probabilistic modelling trained on socially situated corpora. “Construction” in CAAI is not solely human; it is distributed, emerging through interaction between human interpretation, data context, and machine-generated associations.' So an AI is more than "an assistant": it is an assistant which possesses, or *is*, a whole world of interconnected meanings.

Some caveats

Against those very notable strengths, here are a couple of caveats.

How new is this? If it is new, is that really because of the inclusion of AI?

So, we can drop coding and use an AI as a sort of meta-assistant to co-create meaning from text. My biggest question is: to the extent that the AI is not that different from a tireless, lightning-fast and knowledgeable human assistant, couldn't we drop the AI from the methodological equation and just say, here's one newish way to analyse texts hermeneutically? (Though you'd have to have plenty of time and either one very good human assistant or a team of normal ones for this to actually work.)

Is this new method really just a speeded-up version of what we could in theory have done with very fast human assistants? Or is does the sheer magnitude of that speed-up mean a kind of Hegelian transformation of quantity into quality: it's just so much faster that the result is qualitatively different? Or is there something else about the method which is fundamentally new?

Skipping the coding step does lose some reproducibility — does the AI really change this?

If the addition of AI support does not make this way of working fundamentally new, I don't really see how it gives us a free pass on all the original reasons for using coding in the first place. This method is presumably, if everything else is held equal, less reproducible than a method which does employ coding.

Dr. Friese does mention "Reapplication of refined questions across subgroups; independent synthesis by multiple researchers; replication over time" as possible ways to *assess* reproducibility.

This is where our approach to causal coding at Causal Map differs most strongly from CAAI: we only use the AI as a low-level assistant with narrowly defined tasks, so the workflow is less dependent on the stochastic nature of the AI's behaviour and even confines human input to the most crucial high-level decisions, and so is easier to reproduce.

Why this particular set of steps?

It isn't hard at least in theory to think of endless variations of the 4 or 5 steps set out in the article — for example the team-based approach suggested by (Krähnke et al. 2025), so what are the specific arguments for this particular set of steps? Why for example do we work topic by topic? I'm sure it's not the case that "anything goes", but why not?

Whose world?

I am sure Dr. Friese would acknowledge that the world of "machine-generated associations" underpinning AI responses is not just given but is constructed in a very specific way by very specific organisations on a very specific set of training data. This fact has been repeated almost to exhaustion in recent writings on AI, but she does elevate this world to a kind of new meta-assistant, so it would be good to reflect once more on its make-up and provenance. You could see the meta-assistant as a combination of all human lifeworlds, except that of course it isn't — because many people, and certain continents, and whole swathes of actual human non-digital life are not equally included.

In conclusion

As Dr. Friese says, generative AI potentially opens up a whole world of possibilities for qualitative text analysis and indeed for social science in general. We need to be acutely aware of the multiple social, political, methodological and environmental risks of this technology, and *at the same time* not miss out on its benefits.

And, try out QInsights!

References

Friese (2025). *Conversational Analysis with AI - CA to the Power of AI: Rethinking Coding in Qualitative Analysis*. https://doi.org/10.2139/ssrn.5232579.

Krähnke, Pehl, & Dresing (2025). *Hybride Interpretation textbasierter Daten mit dialogisch integrierten LLMs: Zur Nutzung generativer KI in der qualitativen Forschung*. DEU. https://www.ssoar.info/ssoar/handle/document/99389.

🌻 GenAI is distracted by shiny things

7 Oct 2025

I just read another great post from Susanne Friese. I love the way she is forging different ways to combine reflexive qualitative research and GenAI. GenAI surely opens up many new possibilities for qualitative research, including using it for coding and without coding.

But I don't really buy the main argument. Here's why.

My worry with conversational approaches can be centred on this sentence:

> "This shift becomes tangible when researchers conduct a theme analysis and see that major categories—previously developed through weeks of coding—reappear as coherent, evidence-linked themes within minutes."

Braun & Clarke rightfully criticise human researchers for saying "the themes emerged from the text". But this is just the same, isn't it?

"What are the major themes in this document" is often presented as easy, low-level tasks suitable for an AI. But they are not trivial. They are fundamentally creative and meaning-making and involve a whole mass of evaluative judgements. What on earth is a theme? For whom?

Susanne continues:

> "Researchers are repositioned. Their central task becomes judgment: asking meaningful questions, evaluating plausibility, integrating context, articulating theory, and remaining reflexively aware of technological mediation."

But you already handed a mass of judgement-making to your AI when you asked it to identify themes. And in particular Susanne does this right at the start of the workflow, in step 1 (Friese 2025).

Coding is presented as a necessary evil, some kind of crutch on the way to sensemaking which we can now throw away. But where does coding come from? Long before GenAI, a researcher could immerse themselves in a set of texts and eventually emerge with any number of declarative, findings illustrated by quotes. Were they too distracted by the first shiny thing they saw in the text? Who knows? Did they pay too much attention to some parts of it and ignore other uncomfortable parts? In order for qualitative analysis to count as any kind of science, different frameworks and guidelines were constructed to ensure that at least some parts of the analysis were more systematic, retraceable (nachvollziehbar) and transparent (and perhaps even reproducible). There are endless ways to do that, to show that the research process was not distracted by the first shiny thing it saw and ignored the uncomfortable parts. A human, who is essentially distractible by shiny things, can delegate parts of the process to "mere coding" in order to partially mitigate that risk. But **we cannot mitigate humans' propensity to be distracted by shiny things by delegating tasks to a GenAI which is at least as distractible by shiny things**. I know that conversational approaches **do** bring in other methods and procedures to mitigate these risks, having read Friese (2025), but I think this could be better spelled out in Susanne's post.

I agree with Susanne that GenAI may be bringing about a paradigm shift in qualitative research, also in its relationship to quantitative research. But that paradigm shift can take many forms. The way we use GenAI to scale causal mapping is another really different way. Causal mapping happens to be based much more on coding. We would argue that it is therefore more systematic than one which starts by asking "what are the themes here?"

But there are also surely **hundreds** of other ways to use GenAI in qualitative research, most yet to be discovered.

Related

- chapter intro

References

Friese (2025). *Conversational Analysis with AI - CA to the Power of AI: Rethinking Coding in Qualitative Analysis*. https://doi.org/10.2139/ssrn.5232579.