CAUSAL MAP GARDEN

Articles

Contents

🌻 Intro

Chapter contents.

23 Dec 2025

Some brief one-page bullet-point summaries of some of our key published papers. Enjoy.

It's all Work in Progress!

PAGES IN THIS CHAPTER

🌻 A workflow for collecting and understanding stories at scale – Summary (eval2025)

📅 23 Dec 2025

Summary

(Powell et al. 2025)

Source: *Evaluation* 31(3), 394–411 (2025).

- **What problem this paper solves**
- Evaluations often start with a ToC and then collect evidence for each link (e.g. Contribution Analysis), but in many real settings the ToC is uncertain, contested, or incomplete.
- The paper proposes collecting evidence about **structure/theory** (what influences what) and **contribution** simultaneously, using a scalable workflow that stays open-ended.

- **Core idea: "AI-assisted causal mapping pipeline"**

- Treat causal mapping as **causal QDA**: each coded unit is an ordered pair **(influence → consequence)** with provenance, rather than a theme tag.
- Use AI as a **low-level assistant** for interviewing + exhaustive extraction, leaving high-level judgement (prompt design, clustering choices, interpretation) with the evaluator.

- **Pipeline (end-to-end)**

- **Step 1 — AI interviewer**: a single LLM "AI interviewer" conducts semi-structured, adaptive chat interviews at scale.
- **Step 2 — Autocoding causal claims**: an LLM is instructed (radical zero-shot) to list *each* causal link/chain and to ignore hypotheticals.
- **Step 2c — Clustering labels**: embed factor labels and cluster them; then label clusters and optionally do a second "deductive" assignment step to ensure cluster cohesion.
- **Step 3 — Analysis via maps/queries**: produce overview maps, trace evidence for (direct/indirect) contributions, compare subgroups/timepoints.

- **Demonstration study (proof-of-concept)**

- Respondents: online workers recruited via Amazon MTurk; topic: "problems facing the USA" (chosen to elicit causal narratives without a specific intervention).
- Data collection repeated across **three timepoints**; data pooled.
- This is an analogue demonstration; not intended to generalise substantively about "the USA".

- **Key results (reported metrics)**

- **AI interviewing acceptability (proxy)**: 78.5% of interviewees did not ask for changes to the AI's end-of-interview summary; 4.29% asked for changes; 15.3% had no summary (drop-off).
- **Autocoding effort/cost**: ~5 hours to write/test coding instructions; ~$20 API cost (in the reported experiment set-up).
- **Autocoding recall/precision**:
  - Ground-truth link count (authors' assessment): 1154 links.
  - AI-identified links: 1024 (≈ 89%) before precision screening.
  - Precision scoring (0–2 on four criteria: correct endpoints; true causal claim; not hypothetical; correct direction): 65% perfect; 72% dropped only one point.
- **Overview-map "coding coverage"**

  - An 11-factor overview map (plus filters) covered ~42% of raw coded claims while remaining readable.
  - Coarse clustering can collapse opposites/valence (e.g. "military strengthening" and "military weakening" both under "International conflict").
- **Interpretation claims**

- The approach is good for sketching **"causal landscapes"** and triaging hypotheses; it is not reliable enough for high-stakes single-link adjudication without human checking.
- Many outputs depend on **non-automated clustering decisions** (number of clusters, labelling intervention), analogous to researcher degrees of freedom in variable construction.

- **Caveats / ethics**

- Not suitable for **sensitive data** when using third-party LLM APIs; risks of bias and hegemonic worldviews are highlighted.
- Differential response/selection into AI interviewing may not be random.
- Causal mapping shows **strength of evidence**, not **effect size**; forcing magnitudes/polarity is risky.

Related

- chapter intro
- Auto-coding with AI

References

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.

🌻 AI-assisted causal mapping – Summary (validation)

📅 23 Dec 2025

(n.d.)

- **Goal / research question**
- Test whether an **untrained LLM** can **identify and label causal claims** in qualitative interview "stories" well enough to be useful, compared with **human expert coding** (a criterion study).
- Focus is on **validity/usefulness of causal-claim extraction**, not causal inference.

- **Core framing: causal mapping vs systems modelling**

- In systems mapping, an edge $X \to Y$ is often read as "X causally influences Y".
- In causal mapping (as used here), an edge means **there is evidence that X influences Y / a stakeholder claims X influences Y**.
- Output is therefore a **repository of evidence with provenance**, not a predictive system model.

- **"Naive" (minimalist) causal coding definition**

- Deliberately avoids philosophical detail; codes **undifferentiated causal influence** only.
- Does **not** encode effect size/strength; does **not** do causal inference; does **not** encode polarity as a separate field (left implicit in labels like "employment" vs "unemployment").
- Coding decision reduced to: **where is a causal claim, and what influences what?**

- **Data and criterion reference**

- Corpus from a **QuIP** evaluation (2019) of an "Agriculture and Nutrition Programme".
- Dataset previously hand-coded by expert analysts (used as a **criterion study**).
- Validation subset: **3 sources**, **163 statements**, **~15 A4 pages**.

- **Extraction procedure (AI as low-level assistant)**

- Implemented via the **Causal Map** web app using **GPT-4.0**.
- Temperature set to **0** for reproducibility.
- AI instructed to produce an **exhaustive, transparent** list of claims with **verbatim quotes**; synthesis is done later by causal mapping algorithms.
- Exclusions: **ignore hypotheticals/wishes**.
- Output per claim: statement ID + quote + influence factor + consequence factor.

- **Two validation variants**

- **Variant 1 — open coding ("radical zero-shot")**
  - No codebook; includes an "orientation" so the AI understands the research context.
  - Uses a multi-pass prompting process (initial extraction + revision passes).
- **Variant 2 — codebook-assisted ("closed-ish")**

  - Adds a partial codebook (most-used top-level labels from the human coding).
  - Uses hierarchical labels `general concept; specific concept`.
- **Validation metrics and headline results**

- **Precision** (human-rated, four criteria): correct endpoints; correct causal claim; not hypothetical; correct direction.
  - Variant 1: 180 links; perfect composite score (8/8) for **84%** of links.
  - Variant 2: 172 links; perfect composite score (8/8) for **87%** of links.
- **Recall (proxy)**: compared link counts vs the human-coded set (acknowledging no true ground truth because granularity is underdetermined).

- **Utility check (overview-map similarity)**

- Detailed maps differ (expected in qualitative coding).
- When zoomed out to top-level labels and filtered to the most frequent nodes/links, AI and human overview maps are **broadly similar**.

- **Scope limits / risks**

- Small sample; single (relatively "easy") dataset; informal rating process.
- Label choice/consistency remains a major source of variation; batching can introduce inconsistency across prompts.
- Suitable for mapping "how people think" and building auditable evidence sets; not suitable for high-stakes adjudication of specific links without checking.

Related

- chapter intro
- Auto-coding with AI

🌻 Causal mapping for evaluators – Summary (eval2024)

📅 23 Dec 2025

(Powell et al. 2024)

Source: (DOI: `10.1177/13563890231196601`)

- **History / lineage (why this isn't "new", just under-used in evaluation)**: Causal mapping (diagramming "what causes what" using directed links between factors) has been used since the **1970s** across disciplines (e.g. Axelrod-style **document coding** of causal assertions; management/OR traditions emphasising maps for **decision support**; comparative methods like Laukkanen's work on **standardising factor vocabularies** and combining maps). The evaluation literature has relatively sparse, inconsistent "causal mapping" usage; this paper synthesises the wider literature and re-specifies it for evaluators.
- **How we pitch it to evaluators (the niche)**: treat causal mapping as a **discrete evaluation task**: (i) systematically **assemble causal evidence from narrative sources** into an explicit link database with provenance, then (ii) separately use that assembled evidence to make evaluative judgements about "what is really happening". This is positioned as a way to work with large bodies of **messy, heterogeneous** qualitative causal data (different boundaries, contexts, specificity, and ambiguity) without forcing early convergence on a single prior ToC.
- **How causal mapping differs from adjacent approaches**:
- **Primary object is evidence-with-provenance**: causal mapping is explicitly about *who/what source said what link*, not a modeller's best estimate of system structure.
- **Epistemic first, ontic later**: unlike approaches mainly aimed at simulation/prediction (e.g. SD/BBNs/CLDs/FCMs as typically used), causal mapping foregrounds **organising claims/evidence**; inference about reality is a later step.
- **Lightweight causal typing**: it usually does not require consistent weights/functional forms/necessity-sufficiency labels; it can incorporate them when elicited, but warns about spurious precision.

- **How causal-mapping approaches differ among themselves (key axes)**:

- **Mode of construction**: coding **documents** vs coding **interviews** vs **group** map-building (consensus/problem-structuring) vs hybrids.
- **Elicitation openness**: **closed** (pre-specified factor lists) vs **open** (respondent-generated factors), with chaining variants (forward/back).

- **Single-source vs multi-source & context handling**: idiographic maps vs aggregated multi-source maps; whether and how you track **case/context metadata** to avoid invalid transitive inferences.
- **Coding philosophy**: "factors as variables" vs "factors as **changes**" (e.g. QuIP-style); whether polarity/opposites are represented as separate factors/links or handled differently; extent of factor-name **standardisation/merging/nesting**.

- **Problem / motivation**: Evaluators need to represent (a) what causally influences what **in the world**, and (b) what different stakeholders **claim/believe** causally influences what. Causal mapping—defined as the **collection, coding, and visualisation of interconnected causal claims** with explicit **provenance**—is widely used outside evaluation, but under-specified in evaluation practice/literature.

- **Core argument (the "Janus" dilemma + resolution)**:
- **Janus dilemma**: Causal mapping faces two directions—maps can be read as **models of beliefs** or as **models of causal reality**; in practice these get blurred unless source information and analysis steps are explicit.
- **Resolution**: Treat causal maps primarily as **repositories of causal evidence** (epistemic objects), not as direct models of either beliefs or reality. Maps then support structured questions like: *Is there evidence X influences Z? Directly/indirectly? How much evidence? How many sources? How reliable?* The *evaluation* step that judges "what is really happening" is distinct and subsequent.

- **What causal maps encode (and don't)**:

- **Epistemic content**: Map elements are claims/perceptions/evidence, not facts.
- **Causal semantics are usually coarse**: ordinary language claims typically encode **partial influences**, not total/necessary/sufficient causation; coding a link need not assert evidence quality (though you may later weight/filter by quality).
- **Multiple sources + contexts**: maps may be single-source or multi-source; inference across sources requires care about **which case/context** each link refers to.
- **Boundaries are often messy**: system boundaries are frequently loose/implicit; mapping can proceed, but ambiguity must be managed rather than hidden.

- **Causal mapping in evaluation = 3 tasks (workflow)**:

- **Task 1 — Gather narrative causal material**: interviews, open-ended survey questions, document/literature review, archives/secondary text, or consensus processes (e.g., Delphi, participatory systems mapping). Elicitation may use **back-chaining** ("what influenced X?") and **forward-chaining** ("what followed/could follow?"). Question framing affects factor semantics (e.g., QuIP tends to elicit **changes** like "reduced hunger" rather than variables).
- **Task 2 — Code causal claims ("causal QDA")**: unlike standard thematic QDA (codes = concepts), causal QDA codes **links**: each highlighted quote yields an **influence factor → consequence factor** pair; factors mainly exist as endpoints of links. Labelling can be

exploratory/inductive (curate a common vocabulary across sources) or **confirmatory** (codebook from a ToC/prior work), with sequencing cautions to reduce framing/bias. Manual coding is costly; partial automation via NLP/ML is possible but not the focus.

- **Task 3 — Answer evaluation questions using the link database**: global maps become "hairballs", so analysis should generate **selective maps** aligned to questions (e.g., consequences of an intervention; causes of a valued outcome). Techniques include bundling **co-terminal links** (thickness/count), producing frequency-based overview maps (caution: rare-but-important links), rolling-up hierarchical factor taxonomies (with caveats), and limited quantitative summaries (warning: sensitive to coding granularity).

- **Limits / risks**:

- **Inference depends on source credibility**: stronger conclusions require explicit, context-specific **rules of inference** (e.g., independent mentions threshold + theoretical plausibility + bias-mitigation steps).
- **Effect strength/type is hard to capture**: respondents rarely provide consistent magnitudes/necessity/sufficiency/certainty; forcing weights risks **spurious precision**.
- **Transitivity is both payoff and trap**: inferring $C \to E$ from $C \to D$ and $D \to E$ is powerful for indirect effects, but can be invalid when links come from **non-overlapping contexts**; valid inference requires attention to the **intersection of contexts**.

- **Concrete analytic contributions highlighted**:

- Treat diagrams as an **index into the underlying corpus**: tool support should allow tracing from any link/factor back to transcript excerpts + source metadata.
- Quantify robustness of evidence-based "arguments" along paths using **maximum flow / minimum cut** on the causal-claim network (how many claims would need removal to eliminate all paths between C and E), plus **source thread count** (how many distinct sources each provide a complete path).

- **Conclusion / evaluator-facing payoff**:

- Helps evaluators (i) assemble narrative evidence about intervention and contextual influences (direct/indirect, intended/unintended), (ii) search/summarise/select quotations systematically, (iii) increase transparency/peer-reviewability of qualitative causal reasoning, (iv) communicate complexity with readable graphics.
- Key discipline is a **two-step separation**: first assemble and organise causal evidence; then judge what is actually happening—avoiding premature constraint of data collection to fit a prior ToC that stakeholders may not share.

Related

- chapter intro

- Causal mapping in evaluation

References

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

🌻 Causal mapping teams globally – since 2020

📅 20 Feb 2026

Causal mapping teams globally (last 15 years: 2011-2026)

Inclusion rule used here: records from 2011 onward that are clearly causal-mapping work (or close adjacent approaches used to build causal maps), starting from `zotero-bib` and then filtering out obvious noise.

Date of most recent publication	Team members	Location / institution	Papers (title, year, citekey)	What they call the approach	Key methods / fit to categorisation / AI use	Primary evidence source mode	Multi-source handling style	Maturity / status	Tooling stack
2026	Vanessa Hammond; Joseph Lea (plus collaborators)	City of Casey, Melbourne (Australia) context; institution not explicit in record	*Where to Start? Participatory Systems Mapping for Place-Based Service Integration in the City of Casey* (2026) `Hammond et al. (2026)`	Participatory Systems Mapping (PSM)	Group co-production of causal loop maps; strong fit with **group map-building** and **complexity-aware analysis**; uses network metrics + Action Scales; no direct AI	Group workshops	Aggregated group map with facilitated consensus	Preprint	Workshop facilitation + network analysis

Date of most recent publication	Team members	Location / institution	Papers (title, year, citekey)	What they call the approach	Key methods / fit to categorisation / AI use	Primary evidence source mode	Multi-source handling style	Maturity / status	Tooling stack
					in method				
2025	Steve Powell; Gabriele Caldas Cabral; Fiona Remnant; James Copestake; Heather Britt; Rebekah Avard; co-authors	Causal Map / Causal Pathways ecosystem; some records do not specify formal affiliations	*AI-assisted Causal Mapping: A Validation Study* (2025) (n.d.); *Qualitative Causal Mapping in Evaluations* (2025) Remnant et al. (2025); *Strengthening Outcome Harvesting with AI-assisted Causal Mapping* (2025) Britt et al. (2025); *Causal Mapping for Evaluators* (2024) Powell	Causal mapping; AI-assisted causal mapping; causal QDA; QuIP evidence synthesis	Strong fit with **document/interview coding**, **open elicitation**, **provenance-explicit multi-source synthesis**; explicit separation of evidence assembly vs evaluative judgement; substantial AI support for low-level coding/extraction	Narrative interviews, documents, mixed	Explicit provenance and source-thread logic; map-as-evidence-repository approach	Peer-reviewed + chapter + guidance + report + preprint	Causal Map software + LLM-assisted coding/extraction

Date of most recent publication	Team members	Location / institution	Papers (title, year, citekey)	What they call the approach	Key methods / fit to categorisation / AI use	Primary evidence source mode	Multi-source handling style	Maturity / status	Tooling stack
			et al. (2024); *An M&E Time Machine* (2024) Powell et al. (2024); *Does Our Theory Match Your Theory?* (2023) Powell et al. (2023); *Measuring the Women's Economic Empowerment... testing QuIP* (2022) Avard et al. (2022); *Chapter 1 Overview | Guide to Causal Mapping* (2022) Powell et al. (2022); *Guide to*						

Date of most recent publication	Team members	Location / institution	Papers (title, year, citekey)	What they call the approach	Key methods / fit to categorisation / AI use	Primary evidence source mode	Multi-source handling style	Maturity / status	Tooling stack
			Causal Mapping (2021) `Powell & Ltd. (2021)`; *From Narrative Text to Causal Maps* (2021) `Remnant (2021)`						
2025	Jordan White; Pete Barbrook-Johnson	Institute for New Economic Thinking (University of Oxford); CECAN / University of Surrey	*Guidance on Using Large Language Models to Extract Cause-and-Effect Pairs from Texts for Systems Mapping* (2025) `White & Barbrook-Johnson (n.d.)`	Systems mapping with LLM extraction	Fit with **document coding** + **semi-automation**; uses LLM prompts to extract cause-effect pairs then build preliminary maps for later human refinement	Documents/text corpora	Aggregated extraction over multiple texts; human validation step	Report/guidance	GPT-based extraction + external mapping tools (e.g., PRSM)
2024	Fran Ackermann; Colin	Project studies / management	*Overlooked and Underused?*	Causal mapping for projects	Fit with **individual/group**	Interviews, workshops,	Comparative and syntheti	Peer-reviewed articles	Decision Explorer / Group Explorer

Date of most recent publication	Team members	Location / institution	Papers (title, year, citekey)	What they call the approach	Key methods / fit to categorisation / AI use	Primary evidence source mode	Multi-source handling style	Maturity / status	Tooling stack
	Eden; James Alexander; Eunice Maytorena-Sanchez	lineage; institutions not explicit in these records	(2024) Ackermann & Maytorena-Sanchez (2024); *Researching Complex Projects* (2016) Ackermann & Alexander (2016); *Using Causal Mapping to Support Information Systems Development* (2011) Ackermann & Eden (2011)	and IS development	**elicitation**, **document-supported mapping**, and **idiographic systemic analysis**; strong continuity from 2011 to 2024; AI not central	documents	c use across sources		+ manual coding
2023	Rory Hooper; Nihit Goyal; Kornelis Blok; Lisa Scholten	Institution not explicit in record (policy evidence synthesis context)	*A Semi-Automated Approach to Policy-Relevant Evidence Synthesis* (2023)	Semi-automated causal mapping for policy evidence synthesis	Hybrid NLP + causal mapping + graph analytics; fit with **document coding** and **multi-**	Policy/research documents	Aggregated multi-document synthesis	Preprint	NLP pipeline + graph analytics + causal-map post-processing

Date of most recent publication	Team members	Location / institution	Papers (title, year, citekey)	What they call the approach	Key methods / fit to categorisation / AI use	Primary evidence source mode	Multi-source handling style	Maturity / status	Tooling stack
			Hooper (2023)		**source synthesis**; AI used directly for extraction pipeline				
2025	Philippe Giabbanelli; Tyler Gandee; Ameeta Agrawal; Niyousha Hosseinichimeh	Applied ontology / systems mapping	*Benchmarking and Assessing Transformations Between Text and Causal Maps via Large Language Models* (2025) Giabbanelli et al. (2025)	Text-to-map and map-to-text for causal maps	Benchmarking and evaluation datasets for LLM transformation between prose and causal maps; AI core method	Documents and causal-map corpora	Multi-dataset benchmark aggregation	Peer-reviewed article	LLMs + benchmark notebooks/metrics
2025	Melissa Valdivia Cabrera; Michael Johnstone; Joshua Hayward; Kristy Bolton; Douglas	Community health systems modelling	*Integration of Large-Scale Community-Developed Causal Loop Diagra*	NLP-assisted CLD factor merging	Semantic similarity / NLP used to merge community-generated CLD factors	Participatory CLDs + associated text labels	Aggregated map integration across communities	Peer-reviewed articles	NLP semantic matching + network analysis/DEMATEL

Date of most recent publication	Team members	Location / institution	Papers (title, year, citekey)	What they call the approach	Key methods / fit to categorisation / AI use	Primary evidence source mode	Multi-source handling style	Maturity / status	Tooling stack
	Creighton		*ms... NLP Approach* (2025) Valdivia Cabrera et al. (2025); foundational merge workflow in Hayward et al. (2020)		at scale; strong fit with **multi-source map integration**				
2024	Raquel Buzogany; Birgit Kopainsky; Paulo Goncalves	System Dynamics Review / policy narrative analysis	*Developing Theoretically Grounded Causal Maps to Examine and Improve Policy Narratives* (2024) Buzogany et al. (2024)	Theoretically grounded causal maps / CLDs	Grounded-theory coding from qualitative corpora into CLDs; fit with **document coding** and **policy narrative synthesis**	Policy and academic texts	Aggregated cross-domain causal synthesis	Peer-reviewed article	Qualitative coding + CLD modelling
2024	Mohammad S. Jalali; Ali Akhavan	System Dynamics Review	*Integrating AI Language Models in Qualitative*	AI-assisted replication of interview analysis	ChatGPT-assisted replication of interview-to-CLD	Interview transcripts	Replication against prior coded analyses	Peer-reviewed article	ChatGPT-assisted qualitative coding

Date of most recent publication	Team members	Location / institution	Papers (title, year, citekey)	What they call the approach	Key methods / fit to categorisation / AI use	Primary evidence source mode	Multi-source handling style	Maturity / status	Tooling stack
			Research (2024) Jalali & Akhavan (2024)		analysis; AI used as augmentative analyst				
2023	Charlotte Matthews; Will Airey; Fiona Remnant; Aurelie Charles	University of Bath / local SDG evaluation context	*The Dynamics of the UN Voluntary Local Review Using Causal Mapping...* (2023) Matthews et al. (2023)	Causal mapping for SDG VLR analysis	Cross-SDG causal mapping to identify leverage points and stakeholders in local policy; fit with **document + stakeholder synthesis**	VLR documents + stakeholder evidence	Within-goal and across-goal map synthesis	Applied report	Causal Map workflow + policy analysis
2022	Pete Barbrook-Johnson; Alexandra Penn; Helen Wilkinson; Dione Hills (overlapping collaborators)	UK policy/evaluation systems-mapping community (institutions not always explicit in these records)	*Participatory Systems Mapping for Complex Energy Policy Evaluation* (2021) Barbrook-Johnson & Penn (2021); *Building a*	Participatory systems mapping; system-based ToC	Strong fit with **group map-building, open elicitation**, and **translation from cyclic systems maps to evaluable ToC**	Group workshops (plus some supporting docs)	Consensus-built maps with subgroup/submap analysis	Peer-reviewed + handbook chapters + practice case	Workshop facilitation + network/submap analysis

Date of most recent publication	Team members	Location / institution	Papers (title, year, citekey)	What they call the approach	Key methods / fit to categorisation / AI use	Primary evidence source mode	Multi-source handling style	Maturity / status	Tooling stack
			System-Based Theory of Change Using Participatory Systems Mapping (2021) Wilkinson et al. (2021); *Participatory Systems Mapping* (2022) Barbrook-Johnson & Penn (2022); *Running Systems Mapping Workshops* (2022) Barbrook-Johnson & Penn (2022); *Participatory Systems Mapping in Action* (2020) Mapping & Incentive (2020)		**submaps**; AI generally not central in core 2021-2022 method papers				

Date of most recent publication	Team members	Location / institution	Papers (title, year, citekey)	What they call the approach	Key methods / fit to categorisation / AI use	Primary evidence source mode	Multi-source handling style	Maturity / status	Tooling stack
2022	Jie Yang; Soyeon Caren Han; Josiah Poon	AI-NLP causality extraction community	*A Survey on Extraction of Causal Relations from Natural Language Text* (2022) Yang et al. (2022)	Causal relation extraction from text	Survey of knowledge-based, ML, and deep-learning pipelines for extracting cause-effect links; strong **document coding / text-to-causal-link** relevance	Text corpora	Cross-dataset method synthesis	Peer-reviewed article	NLP extraction pipelines
2020	Luke Craven (method lineage)	Systems /evaluation research	*System Effects: A Hybrid Methodology for Exploring the Determinants of Food In/Security* (2017) Craven (2017); *Improvi*	System Effects (hybrid SSM + FCM + graph analysis)	Fit with **participant-derived maps + aggregate structural analysis** across cases; no direct AI core component	Interviews/workshops + synthesis	Aggregated comparative map structures	Peer-reviewed + applied report	FCM + graph-theoretic analysis

Date of most recent publication	Team members	Location / institution	Papers (title, year, citekey)	What they call the approach	Key methods / fit to categorisation / AI use	Primary evidence source mode	Multi-source handling style	Maturity / status	Tooling stack
			ng the Health, Wellbeing... (2020) Craven (2020)						
2020	Steven E. Wallis (plus prior collaborators in IPA lineage)	Institution not explicit in record	*Integrative Propositional Analysis for Developing Capacity...* (2020) Wallis (2020)	Integrative Propositional Analysis (IPA)	Causal propositions mapped and scored structurally (breadth/systemicity); fits **analysis-heavy map evaluation** rather than pure elicitation; AI not central	Strategic documents/propositions	Integrated conceptual synthesis	Peer-reviewed article	Manual proposition extraction + structural metrics
2020	Sasha Strelnikoff; Aruna Jammalamadaka; Dana Warmsley	Institution not explicit in record	*Causal Maps for Multi-Document Summarization* (2020) Strelnikoff et	Causal maps for multi-document summarization	Fully unsupervised NLP pipeline for cause-effect extraction and clustering; fits	Multi-document corpora	Aggregated large-scale document synthesis	Peer-reviewed conference paper	DeepCx + embeddings + mixture model + graph pruning

Date of most recent publication	Team members	Location / institution	Papers (title, year, citekey)	What they call the approach	Key methods / fit to categorisation / AI use	Primary evidence source mode	Multi-source handling style	Maturity / status	Tooling stack
			al. (2020)		**document coding at scale** with explicit AI/NLP core				
2018	Igor Pyrko; Viktor Dorfler	Management/organization research context	*Using Causal Mapping in the Analysis of Semi-structured Interviews* (2018) Pyrko & Dorfler (2018)	Eden/Ackermann-style causal mapping for interview analysis	Strong fit with **interview coding**, **map merging across respondents**, and **feedback-chain analysis**; AI not central	Semi-structured interviews	Merged intersubjective maps from individual interview maps	Conference proceedings paper	Manual coding + map-merging analysis
2018	Ricardo Wilson-Grau; Heather Britt	Outcome Harvesting practice community	*Outcome Harvesting Principles in Practice* (2018) Wilson-Grau (2018); *Outcome Harvesting* (2012) Wilson-	Outcome Harvesting (adjacent causal-claim approach)	Adjacent method for harvesting causal contribution claims from narrative evidence; fits **multi-source**	Narrative outcomes/interviews/documents	Multi-source claim harvesting	Guidance/practice documents	Manual harvesting/coding

Date of most recent publication	Team members	Location / institution	Papers (title, year, citekey)	What they call the approach	Key methods / fit to categorisation / AI use	Primary evidence source mode	Multi-source handling style	Maturity / status	Tooling stack
			Grau & Britt (2012)		**evidence assembly**; typically no AI core				
2017	Gerard Hodgkinson; Kristian Sund; Robert Galavan	Managerial and organizational cognition community	*Chapter 1: Exploring Methods in Managerial and Organizational Cognition* (2017) Hodgkinson et al. (2017)	Causal/cognitive mapping methods in MOC	Methodological synthesis of causal mapping in management cognition; mainly **interview/document elicitation + map analysis** tradition; AI not central	Interviews/documents (methodological review)	Comparative methodological synthesis	Book chapter	Method framework synthesis
2016	Mauri Laukkanen; Mingde Wang; Päivi Eriksson	Comparative causal mapping lineage (management/organization)	*Comparative Causal Mapping: The CMAP3 Method* (2016) Laukkanen & Wang (2016); *New*	Comparative Causal Mapping (CCM), CMAP3	Core **comparative/aggregative map methodology**, including standardized	Interviews/documents (structured to low-structured variants)	Comparative aggregation across multiple respondents/cases	Book + peer-reviewed articles	CMAP3 + manual standardization

Date of most recent publication	Team members	Location / institution	Papers (title, year, citekey)	What they call the approach	Key methods / fit to categorisation / AI use	Primary evidence source mode	Multi-source handling style	Maturity / status	Tooling stack
			Designs and Software for Cognitive Causal Mapping (2013) Laukkanen & Eriksson (2013); *Comparative Causal Mapping and CMAP3 Software in Qualitative Studies* (2012) Laukkanen (2012)		concept pools and software-supported cross-case comparison; AI not central				
2014	Michal Sedlacko; Andre Martinuzzi; Inge Ropke; Nuno Videira; Paula Antunes	Sustainability / ecological economics context	*Participatory Systems Mapping for Sustainable Consumption* (2014) Sedlacko et al. (2014)	Participatory systems mapping	Early strong participatory systems-mapping method paper; fits **group map-building** and systemic insight generation; AI	Participatory workshops	Group-built causal map synthesis	Peer-reviewed article	Workshop methods + map analysis

Date of most recent publication	Team members	Location / institution	Papers (title, year, citekey)	What they call the approach	Key methods / fit to categorisation / AI use	Primary evidence source mode	Multi-source handling style	Maturity / status	Tooling stack
					not central				

Notes on team boundaries

- Teams are grouped by overlapping authors and clear method lineage; single-paper rows are kept where overlap is weak or absent.
- Duplicates were collapsed where records represent the same output (for example `(n.d.)` vs `(n.d.)`).
- Obvious noisy keyword matches (items tagged `causal mapping` but not materially about causal-mapping methods) were excluded.

From your supplied list: explicitly not in this table (and why)

- **Pre-2011 foundational classics** (kept out only due to 15-year scope): e.g., Axelrod 1976, Eden 1988/1992, Eden/Ackermann/Cropper 1992, Narayanan/Armstrong 2004, Clarkson/Hodgkinson 2005.
- **Outside causal-mapping core despite causal relevance**: e.g., Pearl 2000, Forrester 1971, Tolman 1948, Wright 1921 (important background, but not causal-mapping method teams).
- **General AI/qualitative-method papers with weak direct mapping focus**: retained only when explicitly tied to causal-map/CLD production or transformation.

Related

- chapter intro

References

Ackermann, & Eden (2011). *Using Causal Mapping to Support Information Systems Development*.

Ackermann, & Alexander (2016). *Researching Complex Projects: Using Causal Mapping to Take a Systems Perspective*. https://linkinghub.elsevier.com/retrieve/pii/S0263786316300072.

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.

Avard, Mukuru, & Liesner (2022). *Measuring the Women's Economic Empowerment Generated by Impact Investing; Testing the QuIP Method on an Investment in Uganda's Cotton Sector*. Taylor \& Francis.

Barbrook-Johnson, & Penn (2021). *Participatory Systems Mapping for Complex Energy Policy Evaluation*. http://dx.doi.org/10.1177/1356389020976153.

Barbrook-Johnson, & Penn (2022). *Running Systems Mapping Workshops*. In *Systems Mapping: How to Build and Use Causal Models of Systems*. https://doi.org/10.1007/978-3-031-01919-7_10.

Barbrook-Johnson, & Penn (2022). *Participatory Systems Mapping*. In *Systems Mapping: How to Build and Use Causal Models of Systems*. https://doi.org/10.1007/978-3-031-01919-7_5.

Britt, Powell, & Cabral (2025). *Strengthening Outcome Harvesting with AI-assisted Causal Mapping*. https://5a867cea-2d96-4383-acf1-7bc3d406cdeb.usrfiles.com/ugd/5a867c_ad000813c80747baa85c7bd5ffaf0442.pdf.

Buzogany, Kopainsky, & Gonçalves (2024). *Developing Theoretically Grounded Causal Maps to Examine and Improve Policy Narratives about Global Challenges*. https://doi.org/10.1002/sdr.1788.

Craven (2017). *System Effects: A Hybrid Methodology for Exploring the Determinants of Food In/Security*. https://www.tandfonline.com/doi/full/10.1080/24694452.2017.1309965.

Craven (2020). *Improving the Health, Wellbeing , and Chronic Disease Management of the Arabic Speaking Community Data - Through the Culture Well Project , Asthma Australia*.

Giabbanelli, Gandee, Agrawal, & Hosseinichimeh (2025). *Benchmarking and Assessing Transformations Between Text and Causal Maps via Large Language Models*. SAGE Publications. https://doi.org/10.1177/15705838241304102.

Hammond, Lea, & Hammond (2026). *Where to Start? Participatory Systems Mapping for Place-Based Service Integration in the City of Casey*. https://doi.org/10.20944/preprints202601.2371.v1.

Hayward, Morton, Johnstone, Creighton, & Allender (2020). *Tools and Analytic Techniques to Synthesise Community Knowledge in CBPR Using Computer-Mediated Participatory System Modelling*. https://www.nature.com/articles/s41746-020-0230-x.

Hodgkinson, Sund, & Galavan (2017). *Chapter 1: Exploring Methods in Managerial and Organizational Cognition: Advances, Controversies, and Contributions*. In *New Horizons in Managerial and Organizational Cognition*. https://doi.org/10.1108/S2397-52102017002.

Hooper (2023). *A Semi-Automated Approach to Policy-Relevant Evidence Synthesis: Combining Natural Language Processing, Causal Mapping, and Graph Analytics for Public Policy*. https://doi.org/10.21203/rs.3.rs-3285731/v1.

Jalali, & Akhavan (2024). *Integrating AI Language Models in Qualitative Research: Replicating Interview Data Analysis with ChatGPT*. https://doi.org/10.1002/sdr.1772.

Laukkanen (2012). *Comparative Causal Mapping and CMAP3 Software in Qualitative Studies*. https://doi.org/10.17169/fqs-13.2.1846.

Laukkanen, & Eriksson (2013). *New Designs and Software for Cognitive Causal Mapping*. https://www.emerald.com/insight/content/doi/10.1108/QROM-08-2011-1003/full/html.

Laukkanen, & Wang (2016). *Comparative Causal Mapping: The CMAP3 Method*. Routledge.

Mapping, & Incentive (2020). *Participatory Systems Mapping in Action —Supporting the Evaluation of the Renewable Heat Incentive*.

Matthews, Airey, Remnant, & Charles (2023). *The Dynamics of the UN Voluntary Local Review Using Causal Mapping within and across the Sustainable Development Goals: A Case Study of Bath and North East Somerset*.

Powell, & Ltd. (2021). *Guide to Causal Mapping*. https://info.causalmap.app/.

Powell, Remnant, Copestake, Mishan, Arvard, & Goddard (2022). *Chapter 1 Overview | Guide to Causal Mapping*. https://causalmap.app.

Powell, Larquemin, Copestake, Remnant, & Avard (2023). *Does Our Theory Match Your Theory? Theories of Change and Causal Maps in Ghana*. In *Strategic Thinking, Design and the Theory of Change. A Framework for Designing Impactful and Transformational Social Interventions*.

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

Powell, Cabral, & Mishan (2024). *An M\&E Time Machine: Using AI to Measure Changes in a System across a Time Period on Features Which Only Emerge during It.*. https://doi.org/10.31124/advance.171535113.38566027/v1.

Pyrko, & Dorfler (2018). *Using Causal Mapping in the Analysis of Semi-structured Interviews*. https://doi.org/10.5465/AMBPP.2018.14348abstract.

Remnant (2021). *From Narrative Text to Causal Maps- QuIP Analysis and Visualisation 24 Oct 2021*.

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.

Sedlacko, Martinuzzi, Røpke, Videira, & Antunes (2014). *Participatory Systems Mapping for Sustainable Consumption: Discussion of a Method Promoting Systemic Insights*. Elsevier.

Strelnikoff, Jammalamadaka, & Warmsley (2020). *Causal Maps for Multi-Document Summarization*. https://doi.org/10.1109/BigData50022.2020.9377731.

Valdivia Cabrera, Johnstone, Hayward, Bolton, & Creighton (2025). *Integration of Large-Scale Community-Developed Causal Loop Diagrams: A Natural Language Processing Approach to Merging Factors Based on Semantic Similarity*. https://doi.org/10.1186/s12889-025-22142-3.

Wallis (2020). *Integrative Propositional Analysis for Developing Capacity in an Academic Research Institution by Improving Strategic Planning*. https://doi.org/10.1002/sres.2599.

White, & Barbrook-Johnson (n.d.). *Guidance on Using Large Language Models to Extract Cause-and-Effect Pairs from Texts for Systems Mapping Jordan White1 and Pete Barbrook-Johnson1,2 1*

Institute for New Economic Thinking, University of Oxford 2 Centre for the Evaluation of Complexity Across the Nexus, University of Surrey.

Wilkinson, Hills, Penn, & Barbrook-Johnson (2021). *Building a System-Based Theory of Change Using Participatory Systems Mapping*. SAGE Publications Ltd. https://doi.org/10.1177/1356389020980493.

Wilson-Grau, & Britt (2012). *Outcome Harvesting*.

Wilson-Grau (2018). *Outcome Harvesting: Principles, Steps, and Evaluation Applications*. IAP.

Yang, Han, & Poon (2022). *A Survey on Extraction of Causal Relations from Natural Language Text*. https://doi.org/10.1007/s10115-022-01665-w.

🌻 KLAR Outcome Harvesting AI pilot (DEZIM) – Summary (book chapter draft)

📅 23 Dec 2025

Source: draft chapter in `content/000 Articles/020 !! dezim klar book chapter (DRAFT).md`.

- **Purpose**
- Pilot an AI interviewer ("Harvest Assistant") for Outcome Harvesting (OH) in the KLAR! programme, focusing on scalability, inclusion, and democratic evaluation value.

- **Key method move**

- AI-led interviewing + AI post-hoc transcript analysis to draft a structured OH outcome table.
- Strong emphasis on **traceability**: verbatim short citations + page references; explicit missing-information prompts.
- Human validation checks accuracy and de-duplicates overlapping outcomes across sources (triangulation).

- **Results highlights (as reported)**

- 39 invited; 19 responded; 38 outcome statements; 6 met SMART criteria; others retained as leads.
- Real-time outcome summaries enable respondent validation and transparency.

- **Operational insights**

- Prompt simplicity improves adherence; model choice matters; version prompts/models for comparability.
- Scaling shifts bottlenecks to analysis unless the end-to-end workflow is designed.

- **Risks / responsible scaling**

- GDPR/legal basis, consent, third-party naming pathways; document data flow and model/prompt versions; data sovereignty (EU-hosted inference where possible); attention to equity/digital divides.

Related

- chapter intro
- Causal mapping and Outcome Harvesting

🌻 Qualitative causal mapping in evaluations (health) – Summary (book chapter)

📅 23 Dec 2025

(Remnant et al. 2025)

Source: book chapter draft in `content/000 Articles/020 !! health book chapter.md`.

- **Purpose**
- Position QuIP + causal mapping as a credible, cost-effective way to elicit and analyse perceived drivers/barriers in complex interventions (including health services evaluations).

- **Data collection stance**

- QuIP focuses on *changes* that matter to respondents, and the perceived causes of those changes.
- Goal-free / blindfolded questioning is used to reduce pro-project bias; unprompted mention is treated as important evidence.
- Not designed to estimate effect sizes; complements (rather than replaces) quantitative inference and other theory-based approaches.

- **Coding stance ("natively causal")**

- Coding is not thematic tags that are linked later; coding is **pairs/chains of cause→effect factors** ("causal nuggets").
- Coding is parsimonious: only causal claims are coded; non-causal descriptive text is not.
- Inductive label harmonisation across sources is expected; analyst should manage positionality and avoid over-fitting to prior ToC.

- **Use**

- Compare empirical causal maps against ToCs; compare groups (e.g. men/women; staff cadres) and pathways.
- Keep a traceable link from visual summaries back to underlying quotes for verification/peer review.

- **Relationship to realist ideas**

- Affinity to mechanism/context thinking (multiple pathways), but with broader open capture rather than only a few “hotspots”.

Related

- chapter intro
- Causal mapping in evaluation

References

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.

🌻 ToC and causal maps in Ghana – Summary (book chapter)

📅 23 Dec 2025

(Powell et al. 2023)

Source: book chapter draft in `content/000 Articles/020 !! toc book chapter.md`.

- **Purpose**
- Show how QuIP-style causal mapping can compare an official programme ToC (“their theory”) with empirically coded beneficiary narratives (“our theory”), as a disciplined way to revise ToCs and “middle-level theory”.

- **Minimal definition (what a causal map is)**

- A causal map is nodes + directed links, where a link means (at minimum) *someone believes C influenced E*.
- Links need not encode necessity/sufficiency, nor quantified strength/polarity (though those are sometimes added in other approaches).

- **QuIP as causal mapping**

- Goal-free / (partially) blindfolded elicitation of stories of change reduces confirmation bias.
- “Causal back-chaining” elicits causes, causes-of-causes, etc.
- Coding is inductive and multi-source; maps are then filtered/queried to answer evaluation questions.

- **How analysis is actually done**

- Global maps are too large; use filters (e.g. theme/keyword searches, distance steps, frequency thresholds).
- **Hierarchical coding / zooming out**: encode subfactors in factor labels so detailed factors can be rolled up into higher-level factors for readable summary maps.
- Evidence strength is often shown with counts on links (mentions / sources).

- **Interpretation pitfalls (explicitly listed)**

- **Beliefs about causation are not facts about causation**: evaluator judgement remains separate.

- **Absence of a mentioned link is not evidence of absence** (random-walk conversations; negative cases).
- **Transitivity trap / context overlap**: stitching A→B (source 1) and B→C (source 2) does not justify A→C unless contexts overlap.
- Aggregation/generalisation is non-trivial; counts support confidence but don't convert to "truth percentages".

Related

- chapter intro
- Causal mapping in evaluation

References

Powell, Larquemin, Copestake, Remnant, & Avard (2023). *Does Our Theory Match Your Theory? Theories of Change and Causal Maps in Ghana*. In *Strategic Thinking, Design and the Theory of Change. A Framework for Designing Impactful and Transformational Social Interventions*.