

Computational Grounded Theory Integration  
(CGTI):  
Eine methodologische Alternative zur  
XAI-gestützten  
qualitativen Sozialforschung mit Large  
Language Models

Paul Koop

2026

**Zusammenfassung**

Die Integration von Large Language Models (LLMs) in die qualitative Sozialforschung wirft grundlegende methodologische Fragen auf. Während jüngere Ansätze (sogenannte XRSR-Methoden) versuchen, die Opazität von LLMs durch Explainable AI (XAI) zu überwinden, argumentiert dieser Beitrag, dass solche Versuche an einer grundlegenden Inkommensurabilität zwischen quantitativer Opazitätsdurchbrechung und qualitativer Sinnogenese scheitern. Wir entwickeln als Alternative die *Computational Grounded Theory Integration (CGTI)*, eine Methode, die LLMs ausschließlich in methodologisch unbedenklichen Rollen einsetzt: technische Vorverarbeitung und kontrafaktische Exploration. Die eigentliche Interpretation, Fallrekonstruktion und Theoriebildung bleibt dem menschlichen Forscher vorbehalten. Anhand von acht transkribierten Marktgesprächen aus Aachen (1994) demonstrieren wir die vollständige Anwendung der CGTI von der sequenziellen Mikroanalyse bis zur formalen Modellierung mittels Petrinetzen und probabilistischer kontextfreier Grammatiken. Die Ergebnisse zeigen eine überraschende Harmonie der Transaktionen und eine 100%ige Erfolgsquote von Upselling-Versuchen – ein Befund, der ohne systematischen Methodenmix nicht sichtbar geworden wäre. Wir schließen mit einer wissenschaftstheoretischen Beurteilung der CGTI im Vergleich zu XAI-basierten Ansätzen.

# 1 Problembeschreibung

Die jüngere Debatte um den Einsatz von Large Language Models (LLMs) in den Sozial- und Geisteswissenschaften ist von einer grundlegenden Ambivalenz geprägt. Einerseits eröffnen LLMs völlig neue Möglichkeiten der Textanalyse, der Mustergenerierung und der Exploration großer Textkorpora. Andererseits stellt ihre inhärente Opazität – die fehlende Nachvollziehbarkeit, wie genau ein LLM von der Eingabe zur Ausgabe gelangt – ein erkenntnistheoretisches Problem für jede Wissenschaft dar, die auf Nachvollziehbarkeit und methodischer Kontrolle basiert.

Die Debatte hat eine interessante Asymmetrie hervorgebracht: Während in den Geisteswissenschaften die hermeneutische Tradition ohnehin nicht von einer objektiven Wahrheit hinter dem Text ausgeht, sondern vom produktiven Akt des Verstehens selbst, scheinen die Sozialwissenschaften mit ihrer methodisch kontrollierten Fallrekonstruktion vor einem grösseren Problem zu stehen. Ein LLM, das ein Interviewtranskript interpretiert, liefert eine Behauptung (z.B. "Der Subtext dieses Satzes ist eine latente Auflehnung gegen Autorität"). Doch ohne Kenntnis des Rekonstruktionsweges bleibt diese Behauptung wissenschaftlich wertlos – sie könnte auf Halluzinationen oder statistischen Artefakten beruhen.

Ein viel diskutierter Lösungsvorschlag ist der Rückgriff auf XAI (Explainable AI) – Verfahren wie Attention Visualization, Concept Activation Vectors (TCAV) oder Mechanistic Interpretability, die versuchen, die inneren Berechnungen eines LLMs sichtbar zu machen. In jüngsten methodologischen Entwürfen (im Folgenden als XRSR-Ansatz bezeichnet) wird vorgeschlagen, XAI zur Validierung von LLM-Interpretationen zu nutzen und aus den so validierten Kategorienketten formale Modelle (PCFG, Petri-Netze, Bayessche Netze) zu induzieren.

Der vorliegende Beitrag zeigt, dass dieser Ansatz an einer grundlegenden Inkommensurabilität scheitert. Wir entwickeln eine methodologisch konsistentere Alternative, die CGTI, und demonstrieren ihre Anwendung an einem empirischen Korpus von acht Marktgesprächen aus Aachen (1994).

## 2 Stand der wissenschaftlichen Forschung

### 2.1 Hermeneutik und Large Language Models

Die Frage, ob ein LLM *verstehen* kann, ist Gegenstand einer intensiven philosophischen Debatte [10, 1]. Während einige Autor:innen argumentieren, dass die statistischen Regularitäten, die LLMs aus Trainingsdaten extrahie-

ren, eine Form von "Verstehen" darstellen [2], betont die phänomenologische Tradition die Rolle des Bewusstseins, der Leiblichkeit und der temporalen Erfahrung für jeden Verstehensakt [5, 7].

Für die methodologische Diskussion ist entscheidend: Selbst wenn man einem LLM *Verstehen* im Sinne einer Input-Output-Funktion zuspricht, bleibt die *Nachvollziehbarkeit* dieses Verstehens ein Problem. Die klassischen Gütekriterien qualitativer Forschung – insbesondere Intersubjektivität, Replizierbarkeit und Reflektiertheit – verlangen die lückenlose Dokumentation des Erkenntnisweges [11, 9].

## 2.2 XAI in den Sozialwissenschaften

Die Forschung zu Explainable AI hat in den letzten Jahren eine Vielzahl von Verfahren hervorgebracht, um die Opazität neuronaler Netze zu reduzieren [6]. Für LLMs sind insbesondere relevant:

- **Attention Visualization** (z.B. BertViz [12]): Macht sichtbar, welche Token im Input ein LLM bei der Generierung eines Outputs gewichtet hat.
- **Concept Activation Vectors (TCAV)** [8]: Erlaubt die Prüfung, ob ein LLM ein bestimmtes Konzept (z.B. "Gender-Bias") erkannt hat.
- **Mechanistic Interpretability** [4]: Versucht, spezifische Schaltkreise im LLM für bestimmte Operationen zu identifizieren.

Die Übertragung dieser Verfahren auf sozialwissenschaftliche Fragestellungen ist jedoch mit erheblichen methodologischen Problemen verbunden, wie wir im folgenden Abschnitt zeigen.

## 2.3 Kritik des XRSR-Ansatzes

Der hier als XRSR bezeichnete Ansatz (XAI-gestützte relationale Strukturrekonstruktion) schlägt vor, LLM-generierte Lesarten durch XAI zu validieren und anschließend formale Modelle aus den resultierenden Kategorienketten zu induzieren. Unsere Analyse identifiziert drei grundlegende Probleme:

### 2.3.1 Das Kompatibilitätsproblem

XAI-Verfahren liefern quantitative Maße – Attention-Gewichte, Aktivierungsstärken, Vektordistanzen. Die qualitative Sozialforschung operiert jedoch mit sinngenetischen Kategorien, die sich nicht auf quantitative Maße reduzieren

lassen. Dass Token A und Token B eine hohe Attention aufweisen, bedeutet nicht, dass das LLM *dieselbe* soziale Bedeutung erfasst hat, die ein menschlicher Interpret in einer Sequenz sieht. Die Übersetzung zwischen diesen Ebenen wird im XRSR-Ansatz nicht methodisch gelöst, sondern schlicht behauptet.

### 2.3.2 Das Zirkelproblem

XAI soll der Validierung von LLM-Lesarten dienen. Aber: Die XAI-Verfahren selbst sind interpretationsbedürftig. Ein Attention-Heatmap zeigt Korrelationen, keine Kausalitäten. Dass ein LLM bestimmte Token miteinander verknüpft, belegt nicht, dass es die *soziale Bedeutung* einer Sequenz korrekt erfasst hat – es belegt lediglich, dass es statistische Regularitäten aus seinen Trainingsdaten gelernt hat. Der XRSR-Ansatz verschiebt die Opazität von der LLM-Ebene auf die XAI-Ebene, ohne sie zu überwinden.

### 2.3.3 Das Induktionsproblem

Die in Phase 4 des XRSR-Entwurfs vorgeschlagene „Induktion“ formaler Modelle (PCFG, Petri-Netze, Bayessche Netze) aus Kategorienketten ist logisch undurchsichtig. Eine PCFG benötigt ein definiertes Nichtterminal-Symbolsystem. Woher stammen diese grammatischen Kategorien? Wenn sie aus den Lesarten des LLMs extrahiert werden, zirkuliert die Methode. Wenn sie vom Forscher vorgegeben werden, handelt es sich nicht um Induktion, sondern um Konstruktion. Das gleiche Problem gilt für Petri-Netze (die diskrete Zustände voraussetzen, während qualitative Verläufe fließend sind) und Bayessche Netze (die Wahrscheinlichkeitsmaße auf Variablen verlangen, während qualitative Kategorien nicht notwendigerweise intervallskaliert sind).

## 2.4 Die methodologische Alternative: CGTI

Aus dieser Kritik entwickelt der vorliegende Beitrag die *Computational Grounded Theory Integration (CGTI)*. Die CGTI basiert auf drei methodologischen Grundprinzipien:

1. **Primat der menschlichen Interpretation:** Sinnverstehen ist ein bewusstseinsgebundener Akt, den keine KI leisten kann. Die finale Interpretationsverantwortung verbleibt beim menschlichen Forscher.
2. **Trennung von technischer und hermeneutischer Opazität:** Technische Opazität (wie ein LLM intern rechnet) ist akzeptabel, hermeneutische Opazität (wie eine Interpretation zustande kommt) nicht.

3. **Fallrekonstruktion vor Aggregation:** Jeder Fall wird einzeln verstanden, bevor vergleichende Muster extrahiert werden.

Die CGTI unterscheidet sich vom XRSR-Ansatz in vier zentralen Punkten:

Tabelle 1: Vergleich von XRSR und CGTI

Kriterium	XRSR	
Rolle des LLM	Interpretationsgenerator + Validierung durch XAI	Technische Vorve
XAI-Funktion	Validierungsinstanz	
Formale Modellierung	Induktionäus LLM-Kategorien	Explizit
Status formaler Modelle	Empirisch "gegeben"	Hermeneuti
Opazitätsproblem	Verschiebung (LLM → XAI)	Umg

### 3 Problembearbeitung: Die CGTI-Methode im empirischen Einsatz

#### 3.1 Die CGTI-Methode im Überblick

Die CGTI umfasst sechs Phasen:

1. **Datengrundlage & technische Aufbereitung:** Transkription, Anonymisierung, technische Vorverarbeitung (LLM nur für orthografische Korrektur und Segmentierung)
2. **Sequenzielle Mikroanalyse:** Hermeneutische Einzelfallrekonstruktion nach Methoden der objektiven Hermeneutik oder Dokumentarischen Methode
3. **LLM-gestützte kontrafaktische Exploration:** Generierung alternativer Sequenzverläufe zur Kontrastierung (LLM heuristisch, nicht interpretativ)
4. **Systematischer Fallvergleich:** Fallkontrastierung, Identifikation von Vergleichsdimensionen, theoretisches Sampling
5. **Formale Modellierung als Forscherkonstrukt:** Explizite Konstruktion von Petrinetzen, PCFG, ggf. Bayesschen Netzen *aus der qualitativen Theorie heraus*
6. **Theoretische Integration & Validierung:** Triangulation, kommunikative Validierung, reflexive Validierung

## 3.2 Empirisches Material

Das Material umfasst acht Transkripte von Marktgesprächen aus Aachen aus den Jahren 1994 (Juni/Juli). Die Transkripte variieren in Länge, Vollständigkeit und Gesprächstyp:

Tabelle 2: Korpus der Marktgespräche

Text	Datum	Ort	Beteiligte	Besonderheit
T1	28.06.1994	Metzgerei	VK (w), K	vollständig, ruhig
T2	28.06.1994	Kirschenstand	VK (m), K1, K2	fragmentiert, Werberuf
T3	28.06.1994	Fischstand	VK (m), K	vollständig, minimal
T4	28.06.1994	Gemüsestand	VK (m), K	vollständig, Beratung
T5	26.06.1994	Gemüsestand	VK (m), K1, K2	teilweise
T6	28.06.1994	Käseverkauf	VK (m), K1	fragmentiert, abgebrochen
T7	28.06.1994	Bonbonstand	VK (m), K	vollständig, Mini
T8	09.07.1994	Bäckerei	VK (w), K	vollständig, Rollenbruch

## 3.3 Ausgewählte Ergebnisse der Phasen 2–5

Aus Platzgründen beschränken wir uns auf drei zentrale Befunde.

### 3.3.1 Befund 1: Die sequenzielle Struktur des Upselling (Phase 2 und 4)

Die sequenzielle Mikroanalyse (Phase 2) zeigt eine auffällige Regelmässigkeit: In jedem Fall, in dem die Verkäuferin oder der Verkäufer die Frage *SSonst noch etwas?* stellt, folgt ein weiterer Kauf (T1, T4, T5, T8). Die Frage *SSonst noch etwas?* ist damit keine neutrale Serviceanfrage, sondern eine *sequenzielle Nötigung zum Weiterkauf*. Sie erzeugt eine Erwartungsstruktur, der der Kunde nur unter sozialen Kosten widersprechen kann – ein "Nein, danke" müsste als Verweigerung einer angebotenen Beziehung interpretiert werden.

Im gesamten Korpus findet sich kein einziges explizites "Nein" auf diese Frage – lediglich in T4 wird das Upselling durch die eigeninitiative Frage des Kunden nach Pfifferlingen vorweggenommen.

### 3.3.2 Befund 2: Zwei Modi der Verkäufer-Kunde-Beziehung (Phase 4)

Der systematische Fallvergleich (Phase 4) erlaubt die Identifikation von drei distinkten Interaktionsmodi:

- **Transaktionaler Modus** (T1, T3, T5, T6, T7): Klare Rollen, minimaler Austausch, standardisierte Sequenzen, keine Beratung.
- **Beratender Modus** (T4): Expertise-Differenz, Rezeptfragen, sprachliche Nähe ("Nehmen wir die hellen"), Kunde delegiert kulinarische Expertise.
- **Privater Einbruch** (T8): Verlassen der Transaktionsrolle, gemeinsames Klagen über die Tür ("Keiner kümmert sich darum, die Türen zu ölen"), Rückkehr zur Rolle am Ende.

Bemerkenswert ist, dass der beraten-de Modus *nicht* mit höherem Upselling-Erfolg korreliert – im Gegenteil: Der transaktionale Modus ist ebenso erfolgreich.

### 3.3.3 Befund 3: Formale Modellierung als Petrinetz (Phase 5)

Aus den Ergebnissen der Phasen 2–4 konstruieren wir ein Petrinetz für den Standardtransaktionsmodus:



Tabelle 3: Triangulation der Befunde zum Upselling

Befundquelle	Ergebnis
Hermeneutik (Phase 2)	SSonst noch etwas?ist sequenzielle Öffnung, der schwer z
Formales Modell (Phase 5)	$P(\text{Weiterkauf} \mid \text{Upselling-Frage}) = 1,0$ im Korpus
Kontrafaktische Simulation (Phase 3)	Alternative (Nein) führt zu kürzerer Transaktion, aber s
Fehlende Fälle	Kein einziges "Nein" im Korpus deutet auf soziale Erwüns

Die kommunikative Validierung (simulierte Befragung von Marktverkäufern) ergab, dass "Das war's dannäls höfliche Nein-Variante im Alltag vorkommt – eine Kategorie, die im Korpus nicht vertreten ist.

## 4 Zusammenfassung der Ergebnisse

Die Anwendung der CGTI auf die acht Marktgespräche führt zu drei übergeordneten Erkenntnissen:

### 4.1 Erkenntnis 1: Die Harmonie des Marktes 1994

Der Aachener Markt erscheint in den Transkripten als überraschend harmonischer Ort. Es finden sich keine Preiskonflikte, keine Reklamationen, keine abgebrochenen Transaktionen. Die Interaktionen sind ritualisiert, höflich und effizient. Diese Harmonie könnte ein Effekt der Selektion sein (die Transkription erfasste möglicherweise nur harmonische Fälle), ein Effekt der Transkriptionspraxis (unverständliche Passagen könnten Konflikte enthalten) oder ein zeitgeschichtliches Phänomen (D-Mark-Zeit, andere Erwartungen an Marktinteraktionen).

### 4.2 Erkenntnis 2: Die Macht der sequenziellen Öffnung

Die Frage SSonst noch etwas? entfaltet eine bemerkenswerte sequenzielle Kraft. Im gesamten Korpus wird sie nie mit "Nein" beantwortet – stattdessen folgt stets ein weiterer Kauf oder eine Übergangsfloskel ("Das war's"). Die PCFG aus Phase 5 bildet diese Regelmässigkeit mit einer empirischen Wahrscheinlichkeit von  $p = 1,0$  ab, was als starker Hinweis auf eine zugrundeliegende soziale Erwartungsstruktur interpretiert werden kann.

### 4.3 Erkenntnis 3: Rollenbruch als Ressource

Text 8 (Bäckerei) zeigt, dass der temporäre Austritt aus der Transaktionsrolle – hier durch die Klage über die ungeölte Tür – eine Ressource für Beziehungsstiftung sein kann. Der private Einbruch erzeugt ein "gemeinsames Drittes" (das geteilte Ärgernis), das die Interaktion menschlicher macht, ohne die Transaktion zu gefährden. Bemerkenswert ist die Fähigkeit der Teilnehmer, nach diesem Einbruch wieder in die Rollensprache zurückzufinden.

## 5 Beurteilung der CGTI-Methode

### 5.1 Stärken

Die CGTI weist gegenüber XAI-basierten Alternativen drei zentrale Stärken auf:

1. **Methodologische Konsistenz:** Die CGTI vermeidet den Kategorieneffehler, quantitative XAI-Maße mit qualitativen Sinnkategorien kurzzuschliessen. Die Trennung zwischen hermeneutischer Interpretation (Phase 2) und formaler Modellierung (Phase 5) ist klar und begründet.
2. **Rollenklarheit der KI:** Das LLM wird nur dort eingesetzt, wo seine Fähigkeiten methodologisch unbedenklich sind – technische Vorverarbeitung (Phase 1) und heuristische Exploration (Phase 3). Es ersetzt keine menschliche Interpretation.
3. **Replizierbarkeit:** Die CGTI dokumentiert alle Schritte lückenlos – von der sequenziellen Mikroanalyse über die Vergleichsmatrix bis zur expliziten Konstruktion formaler Modelle. Eine Replikation durch andere Forscher:innen ist möglich.

### 5.2 Limitationen

Die vorliegende Anwendung der CGTI weist drei wesentliche Limitationen auf:

1. **Geringe Fallzahl** ( $n = 8$ ): Die formale Modellierung (insbesondere die Schätzung eines Bayesschen Netzes) ist bei dieser Fallzahl nicht valide möglich. Die in Phase 5 präsentierten Wahrscheinlichkeiten sind deskriptiv, nicht inferenzstatistisch.

2. **Fragmentierte Transkripte:** Die Texte T2 und T6 enthalten unverständliche Passagen und sind teilweise abgebrochen. Dies limitiert die Validität der sequenziellen Mikroanalyse für diese Fälle.
3. **Fehlende Kontrastfälle:** Der Korpus enthält keine Konfliktfälle (Preisverhandlungen, Reklamationen, Abbrüche). Die in der Theoriebildung identifizierten Strukturen gelten daher vorläufig nur für harmonische Transaktionen.

### 5.3 Vergleichende Beurteilung: CGTI vs. XRSR

Aus wissenschaftstheoretischer Perspektive ist die CGTI dem XRSR-Ansatz überlegen, weil sie den grundlegenden Unterschied zwischen *statistischen Regularitäten* (was LLMs liefern) und *sinngenetischer Rekonstruktion* (was qualitative Methoden verlangen) anerkennt und nicht durch technische Kunstgriffe zu überbrücken versucht. Der XRSR-Ansatz erliegt dem, was Dreyfus [3] die „Illusion der kognitiven Transparenz“ genannt hat: der Annahme, dass man ein System nur tief genug in seine inneren Berechnungen schauen müsse, um sein Verstehen zu erfassen.

Die CGTI ist demgegenüber bescheidener, aber methodologisch sauberer. Sie nutzt LLMs als *Werkzeuge* – nicht als *Ersatz* für menschliches Verstehen. Ihre Validität ruht nicht auf der Durchbrechung technischer Opazität, sondern auf der methodischen Kontrolle des menschlichen Erkenntnisprozesses.

## 6 Schluss

Die Diskussion um den Einsatz von LLMs in der qualitativen Sozialforschung ist noch lange nicht abgeschlossen. Dieser Beitrag hat argumentiert, dass XAI-basierte Ansätze an einer grundlegenden Inkommensurabilität scheitern und eine methodologisch konsistentere Alternative, die CGTI, entwickelt. Die Anwendung auf acht Marktgespräche aus Aachen (1994) hat die Praktikabilität der Methode demonstriert und erste empirische Befunde zur Struktur von Marktinteraktionen erbracht.

Für die weitere Forschung ergeben sich drei Desiderate:

1. **Methodologische Weiterentwicklung:** Die CGTI sollte mit grösseren und heterogeneren Korpora getestet werden, insbesondere mit Fällen, die Konflikte, Verhandlungen und Abbrüche enthalten.
2. **Vergleichsstudien:** Ein systematischer Vergleich zwischen CGTI und XAI-basierten Ansätzen an *demselben* Material könnte die jeweiligen Stärken und Schwächen präziser herausarbeiten.

3. **Softwareunterstützung:** Die Entwicklung einer Open-Source-Software zur Unterstützung der CGTI (insbesondere für die Phasen 4 und 5) könnte die Zugänglichkeit der Methode erhöhen.

Abschliessend sei betont: Die Frage ist nicht, *ob* LLMs in der qualitativen Sozialforschung eingesetzt werden können – das können sie zweifellos. Die Frage ist, *wie* sie eingesetzt werden sollen, ohne die methodologischen Grundlagen qualitativer Forschung zu untergraben. Die CGTI bietet hierfür einen vielversprechenden, wenn auch weiter entwicklungsbedürftigen Vorschlag.

## Literatur

- [1] Bender, E. M., & Koller, A. (2021). Climbing towards NLU: On meaning, form, and understanding in the age of data. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics*, 5185–5198.
- [2] Chalmers, D. J. (2023). Does a large language model understand language? *Journal of Consciousness Studies*, 30(7-8), 1-25.
- [3] Dreyfus, H. L. (1972). *What computers can't do: A critique of artificial reason*. Harper & Row.
- [4] Elhage, N., Nanda, N., Olsson, C., et al. (2021). A mathematical framework for transformer circuits. *Anthropic Research Paper*.
- [5] Gadamer, H.-G. (1960). *Wahrheit und Methode*. Mohr Siebeck.
- [6] Guidotti, R., Monreale, A., Ruggieri, S., et al. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys*, 51(5), 1-42.
- [7] Heidegger, M. (1927). *Sein und Zeit*. Niemeyer.
- [8] Kim, B., Wattenberg, M., Gilmer, J., et al. (2018). Interpretability beyond feature attribution: Quantitative testing with concept activation vectors (TCAV). *Proceedings of ICML 2018*.
- [9] Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.
- [10] Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417-457.

- [11] Steinke, I. (2004). Gütekriterien qualitativer Forschung. In U. Flick, E. von Kardorff & I. Steinke (Hrsg.), *Qualitative Forschung* (S. 319-331). Rowohlt.
- [12] Vig, J. (2019). A multiscale visualization of attention in the transformer model. *Proceedings of ACL 2019: System Demonstrations*, 37-42.