SOHN@SOHN / SCHRIFTENREIHE

# Signale vor der Entscheidung

## Die validierte Resonanz- und Nachfrageanalyse

Von Dell und Rambachan zum Sohn@Sohn-Modell

ENTDECKEN · MESSEN · VALIDIEREN · PROGNOSTIZIEREN

Wahlen · Märkte · Tagungen · Produkte · wirtschaftspolitische Risiken

**Gunnar Sohn und Constantin Sohn**

Forschungs- und Diskussionspapier · Version 1.0
11. November 2026 · Deutsche Fassung

LESEFASSUNG

# Was sich beobachten lässt, bevor eine Entscheidung fällt

Die Bonner Oberbürgermeisterwahl 2025 lieferte für Sohn@Sohn einen ungewöhnlichen Ausgangspunkt. Eine am 24. September veröffentlichte, auf digitaler Sichtbarkeit und Sentiment beruhende Prognose erwartete für Guido Déus in der Stichwahl einen Stimmenanteil von 52 bis 55 Prozent. Das am 2. Oktober amtlich festgestellte Ergebnis betrug 54 Prozent. Der Abstand zwischen Prognose und Ergebnis ist ein überprüfbarer Einzelfall, noch kein statistischer Nachweis einer überlegenen Methode. Die wissenschaftlich interessante Frage lautet: Welche Signale aus offener Kommunikation lassen sich so messen, dass sie auch unter anderen Bedingungen späteres Verhalten vorhersagen? [11, 12]

Die Wirtschaftswissenschaftler Melissa Dell und Ashesh Rambachan haben 2026 für die Nutzung künstlicher Intelligenz in der empirischen Forschung ein methodisches Programm vorgelegt. Es umfasst die Entdeckung bislang unbeachteter Muster, die begriffliche Definition von Messgrößen, ihre großflächige Beobachtung und die Prüfung ihrer Fehler. Dell und Jacob Carlson zeigen mit MAR-S, wie sich KI-generierte Variablen unter geeigneten Voraussetzungen mit Referenzdaten statistisch korrigieren lassen. Das hier vorgeschlagene Sohn@Sohn-Modell überträgt diese Gedanken auf öffentliche Resonanz, Nachfrageindikatoren und Prognosen wirtschaftlicher oder politischer Entscheidungen. Die Übertragung ist eine eigenständige Arbeitshypothese; sie wird im vorliegenden Papier als Forschungsdesign ausgearbeitet, noch nicht als allgemein bewiesenes Vorhersageverfahren. [1–3]

Der Anwendungsraum reicht über Wahlen hinaus. Für einen Bildungs- und Tagungsanbieter mit Gästehaus können öffentlich beobachtbare Such- und Kommunikationsmuster Hinweise auf neue Veranstaltungsformate, Entscheidungskriterien und bislang wenig beachtete Zielgruppen geben. Unternehmen können Veränderungen bei Produktinteressen oder Serviceerwartungen untersuchen. Auch geopolitische und wirtschaftspolitische Entscheidungen hinterlassen Spuren in Nachrichten, Branchenkommunikation und Nachfrageindikatoren. Wo reale Buchungs-, Absatz- oder Investitionsdaten zugänglich sind, lassen sich diese Signale auf ihren späteren Prognosewert prüfen. Ohne solche Ergebnisdaten bleibt die Untersuchung eine kontrollierte Markt- und Resonanzanalyse.

Die Studie entwickelt dafür eine gemeinsame Architektur mit getrennten Prüfungen für Begriffe, KI-Klassifikationen, Quellenabdeckung und Prognoseleistung. Sie erläutert Chancen, Messgrenzen und Bedingungen eines ersten praktisch umsetzbaren Forschungsvorhabens. Alle grafisch dargestellten Abläufe sind konzeptionell. Die einzige im Papier ausgewiesene eigene Wahlprognose ist der öffentlich dokumentierte Bonner Fall von 2025.

ORIENTIERUNG

# Inhalt

01 | Die Frage vor der Zahl

02 | Die Messrevolution nach Dell und Rambachan

03 | Wie KI neue Hypothesen entdeckt

04 | MAR-S: Die Kontrolle der Messfehler

05 | Vom Datensignal zum Sohn@Sohn-Modell

06 | Die Bonner Wahl als prüfbarer Einzelfall

07 | Nachfrage in Bildungs- und Tagungsmärkten

08 | Produkte, Trends und politische Schocks

09 | Die Grenzen der Beobachtung

10 | Ein Forschungsdesign für die nächste Anwendung

11 | Literatur und methodische Hinweise

KAPITEL 01

# Die Frage vor der Zahl

Wer einen Wahlkampf, eine Marktverschiebung oder eine Nachfrageentwicklung erklären möchte, beginnt gewöhnlich mit vorhandenen Kategorien. Parteien werden nach Stimmenanteilen verglichen, Tagungsstandorte nach Übernachtungen, Produkte nach Umsatz. Solche Größen sind für Entscheidungen wichtig, aber sie liegen oft erst nach dem Ereignis vor. Davor entfaltet sich eine öffentliche Kommunikation, in der Erwartungen, Kritik, Pläne und Kaufabsichten in unterschiedlicher Deutlichkeit sichtbar werden. Texte, Suchinteresse, Themenverschiebungen und Verbreitungswege bilden einen Strom von Beobachtungen, aus dem sich möglicherweise frühe Signale gewinnen lassen.

Bei der Manchot Lecture im Universitätsclub Bonn am 10. November 2026 entwickelte Melissa Dell aus diesem Problem eine allgemeine Frage der empirischen Forschung: Was heißt es, aus unstrukturierten Informationen eine verlässliche Messgröße zu gewinnen? Sie verwies auf Jorge Luis Borges' Gleichnis einer Landkarte, die im Maßstab eins zu eins das gesamte Reich abbildet und gerade dadurch nutzlos wird. Messen bedeutet, Eigenschaften der Wirklichkeit auszuwählen, die für eine konkrete Erkenntnisfrage wichtig sind. Eine Zahl wird erst dadurch sinnvoll, dass offenliegt, welche Wirklichkeit sie repräsentieren soll. [1, 13]

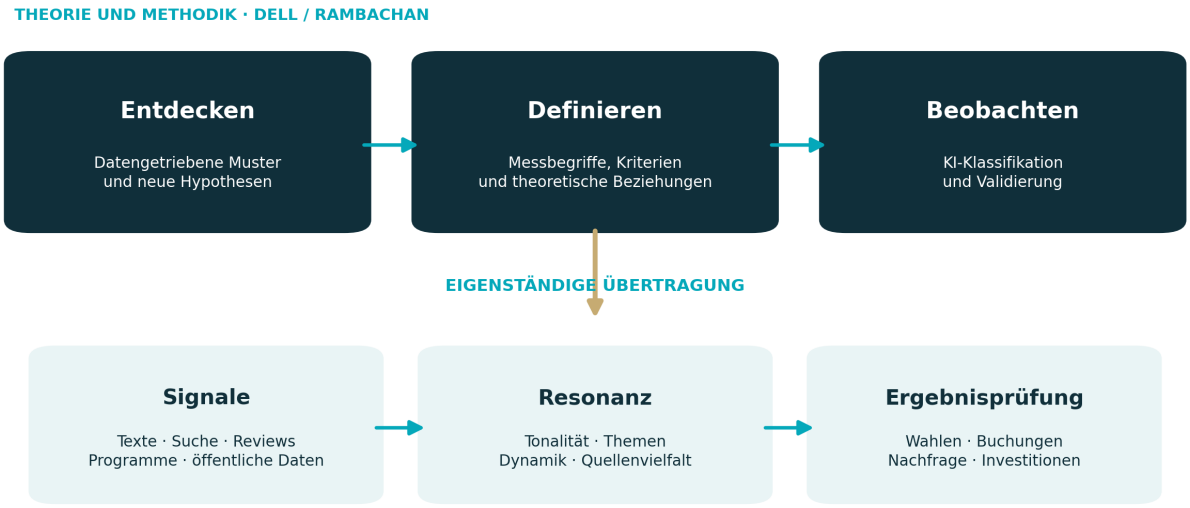



ABB. 01 · Von der Theorie zur Anwendung: Dells und Rambachans Messprogramm liefert die Methodik; der Übergang zur Ergebnisprüfung ist die eigenständige Erweiterung von Sohn@Sohn.

Für Sohn@Sohn liegt darin die wissenschaftliche Fortsetzung eines praktisch begonnenen Versuchs. Während der Bonner Oberbürgermeisterwahl 2025 wurden Sichtbarkeit, Tonalität und Reaktionsmuster im lokalen Kommunikationsraum fortlaufend beobachtet. Vier Tage vor der Stichwahl erschien eine konkrete Prognosespanne. Sie umfasste das später festgestellte Wahlergebnis. Diese Übereinstimmung rechtfertigt eine systematische Überprüfung der Methode. Sie erlaubt weder den Schluss, sämtliche Meinungsumfragen seien überflüssig, noch die Behauptung, dass alle digitalen Kommunikationsräume repräsentativ für die Bevölkerung wären. [11, 12]

Die größere Frage betrifft den Erkenntniswert öffentlicher Spuren. Kann ein Verfahren Muster aufdecken, die traditionellen Fragebögen entgehen? Lassen sich kommunikative Veränderungen in Größen verwandeln, die bei

anderen Wahlen oder in wirtschaftlichen Märkten erneut Bedeutung besitzen? Und wie verhindert man, dass eine KI aus einer großen Datenmenge eine sehr präzise aussehende, aber falsch definierte Kennzahl erzeugt? Die folgenden Kapitel führen die Messforschung von Dell, Rambachan und Carlson mit einem eigenständigen Modell für die Analyse von Resonanz und Nachfrage zusammen.

KAPITEL 02

# Die Messrevolution nach Dell und Rambachan

Dell und Rambachan unterscheiden drei Stufen, an denen künstliche Intelligenz in die empirische Messung eingreift: Entdeckung, Begriffsdefinition und Beobachtung. Bei der Entdeckung wird gefragt, welche Eigenschaften oder Zusammenhänge in einem großen Datenbestand überhaupt vorkommen. In der Begriffsdefinition wird entschieden, was eine wissenschaftlich relevante Eigenschaft konkret bedeutet. Erst die Beobachtung setzt diese Definition im großen Maßstab um. Die Unterscheidung stammt aus ihrem Papier „The Measurement Revolution? Credible Measurement and Inference in the Age of AI“, das die NBER-Methodenvorlesung von Juli 2026 begleitet. [1, 2]

Ein Beispiel verdeutlicht die drei Aufgaben. Wer Nachfrage nach Bildungsreisen untersuchen möchte, kann zunächst in offenen Programmen, Verzeichnissen und öffentlichen Diskussionen wiederkehrende Themen entdecken: berufliche Weiterbildung, kurze Seminare, Nachhaltigkeitsformate oder mehrtägige Veranstaltungsreihen. Die bloße Häufigkeit dieser Wörter ist noch keine Nachfrage. Im zweiten Schritt muss festgelegt werden, welches Verhalten als Informationssuche, welches als ernsthafte Buchungsabsicht und welches lediglich als allgemeines Interesse gelten soll. Im dritten Schritt wendet ein KI-System die festgelegten Kriterien auf einen umfangreichen Dokumentenbestand an.




ABB. 02 · Drei Stufen glaubwürdiger Messung nach Dell und Rambachan. Die Validierung richtet sich jeweils auf andere Fehlerquellen. Eigene Darstellung nach [1, 2].

Dell erinnert daran, dass Wissenschaftler seit Langem mit Proxys arbeiten: mit beobachtbaren Stellvertretergrößen für schwierig zu erfassende Phänomene. Zeitungserwähnungen können als Proxy für wirtschaftspolitische Unsicherheit dienen; öffentliche Rezensionen können Hinweise auf Serviceprobleme liefern. Der Proxy ist allerdings nicht mit dem untersuchten Sachverhalt identisch. Die Entscheidung, welcher sprachliche Befund als politisches Vertrauen oder Nachfrageinteresse gilt, muss begründet werden. KI vereinfacht die Erzeugung vieler plausibel aussehender Proxys. Gerade diese Vielfalt erhöht die Anforderungen an die Auswahl. [1, 7]

Zur Begriffsprüfung greift Dell auf die Psychometrie zurück. Cronbach und Meehl entwickelten 1955 die Idee eines nomologischen Netzwerks: Ein nur indirekt beobachtbares Konzept gewinnt Glaubwürdigkeit durch theoretisch erwartete Beziehungen zu anderen Größen und durch Unterschiede zu benachbarten Konzepten. Eine Kennzahl für „positive Resonanz“ müsste beispielsweise von bloßer Bekanntheit zu unterscheiden sein; ein Indikator für Nachfrage müsste einen eigenen Informationsgehalt gegenüber allgemeiner Medienaufmerksamkeit zeigen. Dell schlägt vor, konkurrierende Messgrößen auch als Problem partieller Identifikation zu behandeln: Mehrere theoretisch zulässige Maße können zu einem Bereich vertretbarer Aussagen führen. [1, 4]

Der praktische Gewinn liegt in der begrifflichen Disziplin. Die Öffentlichkeit erzeugt eine ungeheure Menge an Daten, doch die Zahl der Treffer, Kommentare oder Suchanfragen besitzt noch keine eindeutig festgelegte ökonomische Bedeutung. Erst ein dokumentiertes Messdesign macht aus dem Material überprüfbare Informationen. Die in diesem Papier verwendete Bezeichnung „Resonanz“ meint deshalb zunächst die beobachtbare Verbreitung und Verarbeitung eines Themas in definierten Kommunikationsräumen. Sie bedeutet nicht automatisch Zustimmung, Repräsentativität oder einen späteren Kauf.

KAPITEL 03

# Wie KI neue Hypothesen entdeckt

Die zweite tragende Idee stammt aus Rambachans Teil der NBER-Methodenvorlesung. Maschinelles Lernen kann aus unstrukturierten Daten nicht allein vorgegebene Kategorien herauslesen. Es kann Forschern auch Hinweise auf Merkmale liefern, die bei der Planung einer Studie noch keinen Namen besaßen. Rambachan beschrieb dafür ein Verfahren aus Bewertung, Kontrastbildung, Interpretation und unabhängigem Test. Zuerst erhält jede Beobachtung einen statistischen Score. Dann werden Beobachtungen mit unterschiedlichen Scorewerten verglichen. Aus ihren Unterschieden entstehen neue Hypothesen, die anschließend an zurückgehaltenen Daten überprüft werden. [2, 5]

Dieser Ansatz ist für Marktanalysen besonders ergiebig. Ein Tagungshaus vermutet möglicherweise, dass die Nachfrage hauptsächlich vom Preis und von der Verkehrsanbindung abhängt. Eine KI-gestützte Untersuchung kann darüber hinaus erkennen, dass bestimmte Anbieter wiederholt nach parallel nutzbaren Gruppenräumen suchen, dass mehrtägige Seminare andere Anforderungen besitzen als Tagesveranstaltungen oder dass regionale Küche im Zusammenhang mit pädagogischen Veranstaltungsformaten auftritt. Solche Beobachtungen dürfen nicht unmittelbar als bewiesene Kaufmotive gelten. Sie erweitern zunächst den Raum der Fragen, die eine anschließende Prüfung verdienen.

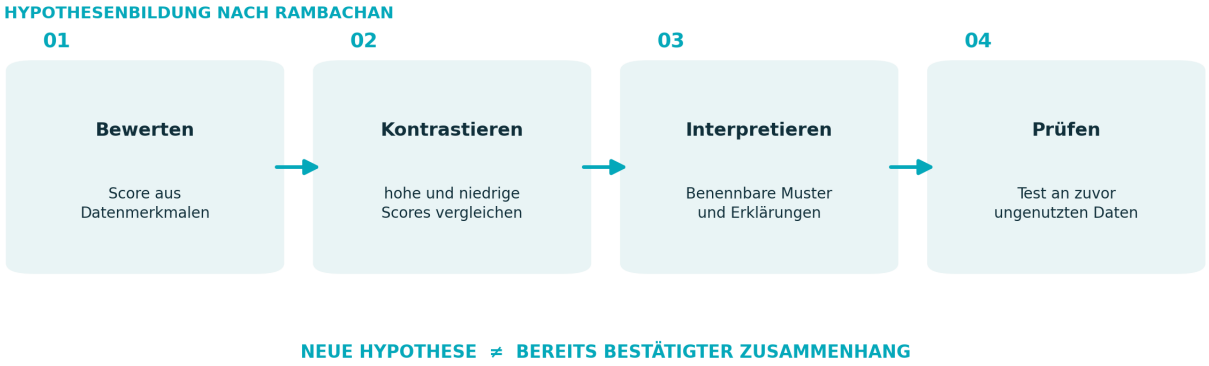



ABB. 03 · Hypothesenfindung mit unstrukturierten Daten. Der unabhängige Test verhindert, dass nachträglich entdeckte Muster schon als bestätigte Theorie ausgegeben werden. Nach [2, 5].

Rambachan verwies auch auf die Gefahr nachträglicher Erzählungen. Wer Hunderte denkbare Muster untersucht, findet zwangsläufig einige auffällige Zusammenhänge. Ein Wissenschaftler kann daraus leicht eine überzeugende Geschichte konstruieren. Gegen diese Versuchung helfen getrennte Datensätze: Ein Teil dient der Entdeckung, ein weiterer der Definition und ein bislang unberührter Teil der Überprüfung. Werden viele Hypothesen getestet, muss die statistische Auswertung die Vielzahl der Tests berücksichtigen. Eine schöne Erzählung bleibt damit eine Hypothese, bis sie auch an unbekannten Beobachtungen trägt.

Für Sohn@Sohn entsteht daraus eine Erweiterung der Sentiment-Analyse um Narrative und Veränderungsmuster. Bei einer Wahl könnte ein Kandidat immer häufiger mit Verwaltungsleistung, ein anderer mit Konflikten um lokale Infrastruktur verbunden werden. In einem Produktmarkt könnten sich neue Anforderungen herausbilden, die in den üblichen Produktkategorien noch fehlen. In einer geopolitischen Auseinandersetzung könnten Themen wie Lieferfähigkeit, Exportkontrolle oder regulatorische Sicherheit plötzlich in Branchenkommunikationen auftauchen. Die Methode sucht solche strukturellen Veränderungen, bevor sie sie in Indizes verdichtet.

KAPITEL 04

# MAR-S: Die Kontrolle der Messfehler

Nach der Entdeckung und Definition folgt die technische Messung. Hier setzt das von Jacob Carlson und Melissa Dell entwickelte Verfahren MAR-S an: „Missing At Random Structured Data“. Die Grundidee ist, dass eine umfangreiche Textsammlung die gewünschten statistischen Variablen zunächst nicht enthält. Ein Sprachmodell ergänzt diese fehlenden Angaben, etwa die Klassifikation „politisches Thema“ oder „positive Bewertung“. Die KI liefert jedoch fehlerhafte Schätzungen. MAR-S verbindet solche maschinellen Messungen mit einer nach festgelegten Regeln gewonnenen, sorgfältig geprüften Referenzstichprobe. [3]

Das lässt sich ohne aufwendige Mathematik beschreiben. Zunächst bewertet die KI sämtliche zugänglichen Beiträge nach einem klaren Kategoriensystem. Bei einer kleineren Auswahl wenden Fachleute dieses Kategoriensystem unabhängig an. Die beiden Ergebnisse lassen sich vergleichen. Wenn das Modell bestimmte Textarten systematisch zu positiv einschätzt, wird diese Abweichung sichtbar. Unter den Annahmen des Verfahrens kann die statistische

Schätzung so korrigiert werden, dass der Fehler der künstlich ergänzten Variablen angemessener berücksichtigt wird. Auch die verbleibende Unsicherheit gehört zum Ergebnis.

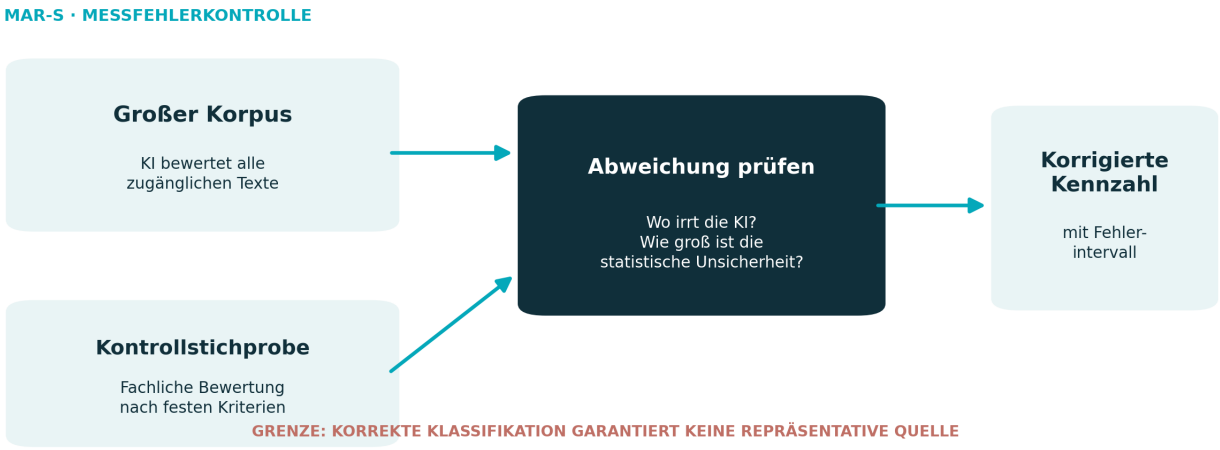



ABB. 04 · Das MAR-S-Prinzip in vereinfachter Form. Die Grafik erläutert die Logik der Fehlerkorrektur, nicht sämtliche mathematischen Voraussetzungen. Nach [3].

Für einen Mittelwert besitzt die Rechnung eine einfache Anschauung: Der Durchschnitt aller KI-Klassifikationen erhält einen Korrekturterm aus den Abweichungen in der Referenzstichprobe, gegebenenfalls gewichtet mit den bekannten Auswahlwahrscheinlichkeiten. Die Güte der Referenzstichprobe entscheidet mit über die Gültigkeit. Relevante Textarten dürfen nicht von vornherein ausgeschlossen werden, und die Auswahl muss ausreichend dokumentiert sein. Ein kleiner, fachlich hochwertiger Prüfkorpus kann deshalb wertvoller sein als eine zweite unkontrollierte Massenklassifikation.

Für ein Resonanzmodell kommt ein weiterer Schritt hinzu. Einzelne Beiträge werden zu Wochen-, Themen- oder Regionalkennzahlen zusammengefasst. Dell und Carlson behandeln ausdrücklich die statistischen Schwierigkeiten, die bei der Aggregation KI-generierter Einzelmessungen entstehen. Eine wöchentliche Kennzahl darf nicht präziser erscheinen, als ihre fehlerbehaftete Ausgangsklassifikation rechtfertigt. Auch die Unsicherheit der Zusammenfassung muss fortgeführt werden. [3]

MAR-S löst damit ein bestimmtes Problem: die Fehler der Messung innerhalb eines definierten und hinreichend abgedeckten Datenbestands. Es garantiert weder, dass der Datenbestand alle Zielgruppen repräsentiert, noch dass positive Online-Kommentare später zu Buchungen oder Stimmen führen. Dieser zweite Übergang verlangt ein eigenständiges Prognosemodell. Die Trennung zwischen Messvalidität und Prognosevalidität gehört zum Kern des hier vorgeschlagenen Forschungsdesigns.

KAPITEL 05

# Vom Datensignal zum Sohn@Sohn-Modell

Der eigenständige Ansatz von Sohn@Sohn beginnt dort, wo Dell und Rambachan ihre generelle Messtheorie offenlassen: bei der Verknüpfung öffentlicher Kommunikation mit einem zukünftigen Ergebnis. Das Ziel kann ein Wahlausgang, eine Buchungsentwicklung, eine Veränderung der Produktnachfrage oder eine wirtschaftliche Entscheidung nach einem politischen Ereignis sein. Das Verfahren benötigt deshalb neben guten Messgrößen auch eine klare Zielgröße und eine überprüfbare Prognoseregel. Ein Modell ohne spätere Bewährungsprobe bleibt eine Lagebeschreibung.

Als Arbeitstitel bietet sich „Modell der validierten Resonanz- und Nachfrageanalyse“ an. Es trennt fünf Elemente. Die Quellenebene dokumentiert das zugängliche Material und seine Reichweite. Die Entdeckungsebene sucht neue Themen und wiederkehrende Narrative. Die Messebene erfasst nach vorher definierten Regeln Sichtbarkeit, Tonalität, thematische Verknüpfungen, Verbreitung und Veränderungen über die Zeit. Die Validierungsebene kontrolliert systematische KI-Fehler sowie die Grenzen der Quellenabdeckung. Die Zielebene prüft, welche Indikatoren reale spätere Entscheidungen vorhersagen.

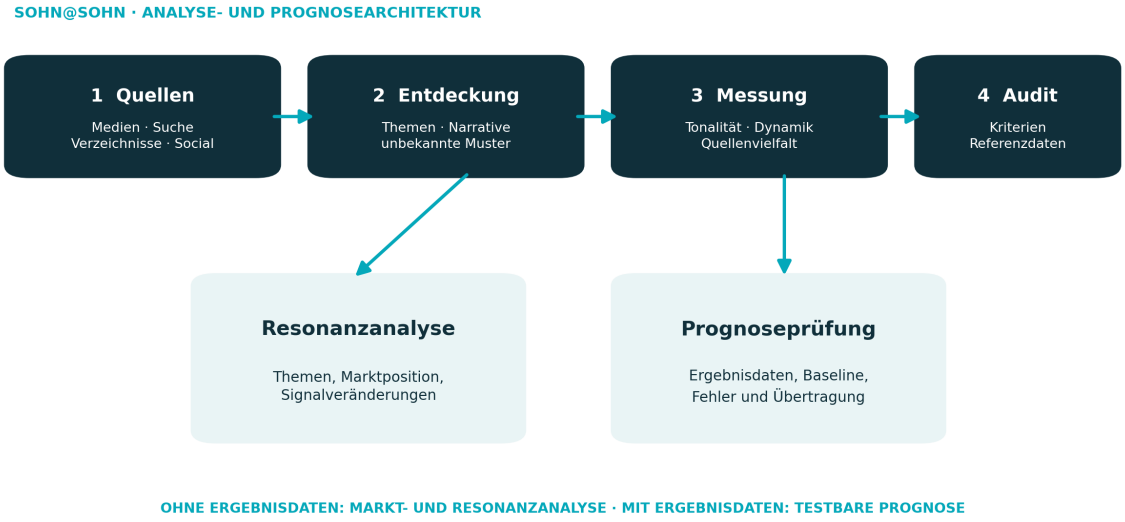



ABB. 05 · Eigenständige Modellarchitektur von Sohn@Sohn. Resonanzanalyse und Prognoseprüfung sind getrennte Arbeitsschritte.

Ein Resonanzindex kann aus mehreren Teilgrößen bestehen. Sichtbarkeit beschreibt die Häufigkeit und den Kontext einer Erwähnung. Tonalität bezeichnet die Bewertung eines Themas, Angebots oder Akteurs. Anschlussfähigkeit misst, ob Aussagen in anderen Quellen erneut aufgegriffen werden. Thematische Dynamik erfasst den Wandel von Problemdeutungen. Quellenvielfalt hilft zu unterscheiden, ob ein Signal aus einem engen Kreis stammt oder über mehrere eigenständige Kommunikationsräume verbreitet ist. Eine solche Konstruktion muss die Gewichtung der Teilgrößen offenlegen und nach Möglichkeit anhand früherer Beobachtungen kalibrieren. Die Definition darf sich nicht von Fall zu Fall unbemerkt verändern.

Die Prognoseebene ist davon getrennt. Sie verbindet die gemessenen Resonanzverläufe mit verfügbaren historischen Ergebnisdaten, Saisoneffekten, Preis- und Kapazitätsinformationen oder institutionellen Besonderheiten. Beim Wahlmodell wären dies frühere Wahlergebnisse und Mobilisierungsfaktoren. Im Tagungsmarkt könnten es – bei

entsprechender Datenfreigabe – tatsächliche Anfragen, gebuchte Raumtage oder Konversionsraten sein. Ohne solche Ergebnisdaten darf der Bericht nur Aussagen über dokumentierte Interessen, Angebotsstrukturen und mögliche Nachfragepfade machen.

Darin liegt der praktische Unterschied zwischen einer beschreibenden Sentiment-Studie und einem prognostischen Modell. Die erste untersucht, wie ein Thema öffentlich bewertet wird. Das zweite legt im Voraus fest, welches spätere Ereignis aus welchen Signalen vorhergesagt werden soll und wie der Fehler der Vorhersage berechnet wird. Ein sorgfältig aufgebautes Analyseverfahren kann auch dann einen Nutzen besitzen, wenn die Prognosefrage noch offen ist. Es kann unerkannte Nachfragekriterien, Medienbrüche oder thematische Leerstellen sichtbar machen und damit Entscheidungen vorbereiten.

KAPITEL 06

# Die Bonner Wahl als prüfbarer Einzelfall

Die Bonner Oberbürgermeisterwahl 2025 liefert einen klar dokumentierten, zeitlich abgeschlossenen Fall. Am 24. September veröffentlichte Sohn@Sohn eine Prognose von 52 bis 55 Prozent für Guido Déus und 45 bis 48 Prozent für Katja Dörner. Die Schätzung stützte sich nach der damaligen Veröffentlichung auf fortlaufende Beobachtungen digitaler Sichtbarkeit, Stimmungslage und Reaktionsmuster. Bei der Stichwahl am 28. September erreichte Déus nach dem amtlichen Endergebnis 54 Prozent und Dörner 46 Prozent. Der Wahltermin, die Prognosespanne und das amtliche Ergebnis lassen sich unabhängig nachprüfen. [11, 12]

Für die wissenschaftliche Bewertung bleiben mehrere Fragen offen. Aus der öffentlichen Dokumentation geht kein vollständiger, extern auditierter Datenkorpus mit sämtlichen Modellversionen, Auswahlwahrscheinlichkeiten und Gewichtungen hervor. Auch eine Reihe unabhängiger, vorab protokollierter Wahlprognosen liegt in diesem Papier nicht vor. Der Fall erlaubt deshalb die Feststellung, dass die veröffentlichte Spanne das spätere Ergebnis einschloss. Er erlaubt keine statistisch belastbare Aussage über eine allgemeine Überlegenheit gegenüber Meinungsumfragen.

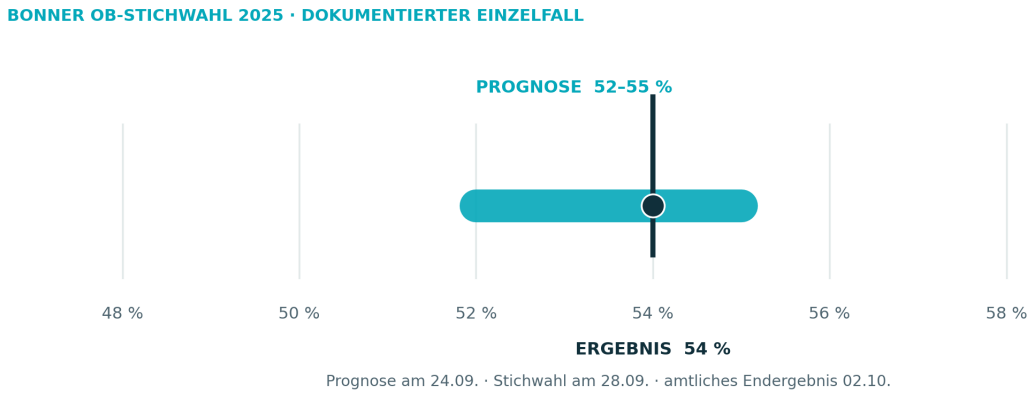



ABB. 06 · Vorab veröffentlichte Spanne für Guido Déus und amtliches Endergebnis der Bonner Stichwahl 2025. Ein dokumentierter Einzelfall [11, 12].

Das macht Bonn gerade als Pilotfall nützlich. Ein methodisches Rückwärts-Audit könnte den damals verwendeten Quellenbestand rekonstruieren, soweit er archiviert ist. Dabei ließe sich prüfen, wie stark sich die Prognose durch

andere Klassifikationsregeln, einzelne Quellen oder veränderte Zeitfenster verschiebt. Die ursprüngliche Prognose müsste unverändert bleiben; spätere Anpassungen wären ausdrücklich als Rückblick zu kennzeichnen. Erst die anschließende Anwendung eines eingefrorenen Modells auf weitere Wahlen würde Auskunft über seine übertragbare Prognoseleistung geben.

Ein wichtiger Unterschied betrifft die Informationsarten. Keine klassische Bevölkerungsbefragung zu verwenden bedeutet nicht, auf jede Form von Stichprobe zu verzichten. Für die Qualitätskontrolle der KI sind nach Dell und Carlson geeignete Stichproben von Texten oder anderen Dokumenten notwendig. Befragung von Wahlberechtigten und Stichprobe aus einem Kommunikationskorpus erfüllen unterschiedliche Zwecke. Wer diese beiden Verfahren sprachlich gleichsetzt, übersieht den methodischen Fortschritt.

Der Fall Bonn liefert damit eine Fallstudie, die den nächsten Versuch begründet. Das Sohn@Sohn-Modell gewinnt wissenschaftliche Glaubwürdigkeit durch zukünftige, vorab festgelegte Prüfungen und durch den Vergleich mit einfachen Basismodellen, beispielsweise der Fortschreibung früherer Wahlergebnisse. Die Abweichung zwischen späterem Ergebnis und Prognose sollte über mehrere Fälle hinweg ausgewertet werden; ebenso wichtig ist die richtige Einschätzung der Prognoseunsicherheit.

KAPITEL 07

# Nachfrage in Bildungs- und Tagungsmärkten

Die Übertragung auf kommerzielle und institutionelle Entscheidungen beginnt mit einem Beispiel aus dem Markt für Tagungen, Fortbildung und Beherbergung. Angenommen, ein gemeinnütziger Bildungsträger betreibt ein eigenes Veranstaltungshaus mit Hotelbetrieb. Das Angebot umfasst Seminarräume, kleinere Gruppenräume, Übernachtungen und Verpflegung. Der Träger möchte wissen, welche Veranstalter nach welchen Leistungsmerkmalen suchen, wie er im Vergleich zu Standorten im Umland wahrgenommen wird und welche neuen Formate Nachfrage erzeugen könnten. Damit entsteht eine typische Forschungsfrage für eine validierte Resonanz- und Nachfrageanalyse.

Ein herkömmlicher Marktforschungsansatz könnte Veranstalter befragen. Ein datenbasierter Zugang würde zunächst die öffentlich verfügbaren Kommunikationsspuren untersuchen: Seminarprogramme, Fachveranstaltungskalender, öffentlich zugängliche Ausschreibungen, Location-Verzeichnisse, Fachmedien, Suchtrendindikatoren, Website-Inhalte und frei zugängliche Beiträge aus der Veranstaltungsbranche. Daraus ließen sich Themencluster, verwendete Suchbegriffe und Entscheidungskriterien rekonstruieren. Die gewählte Quellenbasis müsste dokumentiert und ihre Grenzen offengelegt werden. Interne Suchprotokolle von Plattformen oder tatsächliche Reservierungsdaten sind dadurch noch nicht verfügbar.

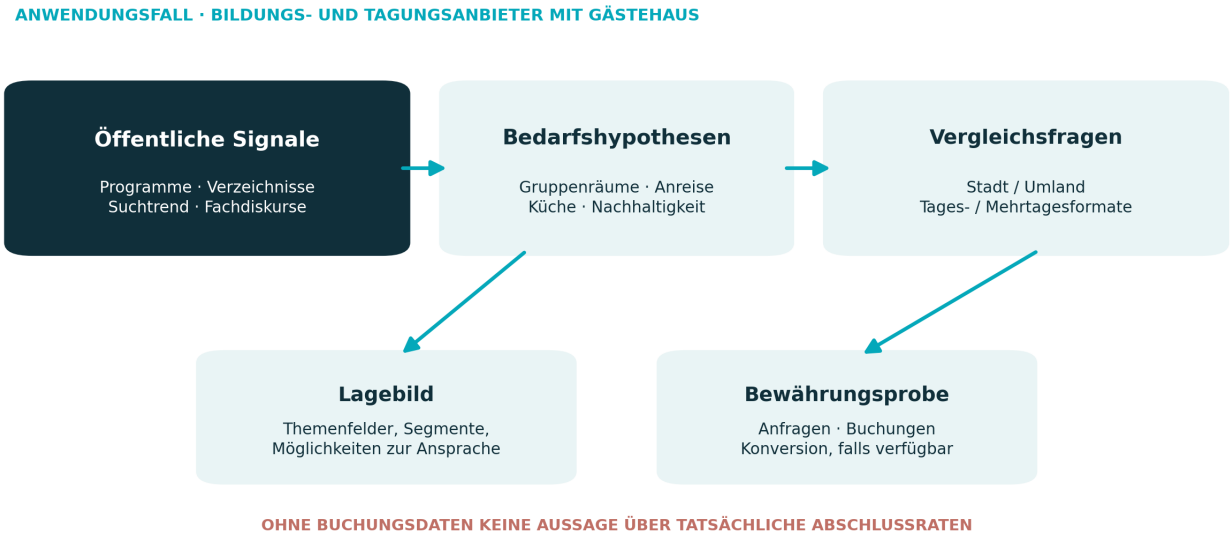



ABB. 07 · Hypothetischer Untersuchungsfall eines Bildungs- und Tagungsanbieters. Eine konkrete Institution oder beauftragte Studie ist mit der Darstellung nicht verbunden.

Das Verfahren könnte beispielsweise untersuchen, wie häufig Veranstalter mehrere parallel nutzbare Gruppenräume benötigen, welche Rolle die Kombination von Unterkunft und Veranstaltungsflächen spielt und bei welchen Formaten eine günstige Erreichbarkeit besonders wichtig erscheint. Ein weiteres Untersuchungsfeld wären Nachhaltigkeit, regionale Beschaffung, Bio-Angebote und frisch zubereitete Küche. Diese Merkmale können als deklarierte Anforderungen, als kommunikative Unterscheidungsmerkmale oder als tatsächlich nachgefragte Leistung erscheinen. Die Untersuchung müsste diese Bedeutungen auseinanderhalten.

Interessant wäre auch ein Vergleich zwischen einem städtischen Veranstaltungsstandort und einem nahe gelegenen Umlandstandort. Ein solcher Vergleich könnte Themen wie Anreise, Umgebung, Freizeitwert, Preisrahmen und

Veranstaltungsprofil berücksichtigen. Aus öffentlichen Quellen ließe sich feststellen, welche Angebote und Botschaften sichtbar sind. Ob ein Standort tatsächlich häufiger gebucht wird und welche Merkmale eine Buchung auslösen, wäre ohne belastbare Transaktions- oder Anfragedaten lediglich als Hypothese formulierbar.

Ein weiterer Markt betrifft bundesweit organisierte Seminarreihen. Ihre Anbieter wiederholen Veranstaltungen an unterschiedlichen Standorten und besitzen deshalb besondere Anforderungen an Standardisierung, Raumflexibilität, digitale Infrastruktur und verlässliche Abläufe. Wer solche Veranstaltungsprogramme systematisch auswertet, kann Hinweise auf bislang unterversorgte Themen, regionale Lücken oder wiederkehrende Formate finden. Für Bildungsurlaub und betriebliche Weiterbildung stellt sich zusätzlich die Frage, welche Themen Arbeitnehmer interessieren und welche Rolle Arbeitgeber oder Personalabteilungen bei der tatsächlichen Teilnahme spielen.

Damit ist eine Marktanalyse möglich, die über die Beschreibung aktueller Wettbewerber hinausführt. Sie fragt, welche Anforderungen sichtbar werden, welche Zielgruppen unterschiedlich sprechen, wo sich kommunikative Nachfrage von bestehenden Angebotskategorien entfernt und welche Angebotsversprechen überprüft werden sollten. Ihr Nutzen läge in einer priorisierten Hypothesenkarte. Die Entscheidung über neue Leistungen, gezielte Inhalte oder Vertriebskanäle könnte anschließend mit kleinen Markttests und – sofern verfügbar – realen Anfrage- oder Buchungsdaten überprüft werden.

KAPITEL 08

# Produkte, Trends und politische Schocks

Dasselbe Forschungsdesign lässt sich auf industrielle Produkte, digitale Dienstleistungen, Konsumgüter und neue Moden übertragen. Bei Produkten ist die zentrale Frage häufig, weshalb ein Angebot in bestimmten Nutzergruppen Aufmerksamkeit gewinnt und in anderen nicht. Öffentliche Rezensionen, technische Diskussionen, Suchtrendindikatoren, Fachmedien und Wettbewerbsankündigungen können frühe Hinweise auf veränderte Anforderungen liefern. Eine Sentiment-Klassifikation allein würde solche Signale nur teilweise erfassen. Die automatische Entdeckung neuer Merkmale könnte zeigen, welche Produktprobleme sich häufen oder welche Anwendungen zunächst in kleinen Gemeinschaften entstehen.

Ein Hersteller könnte beispielsweise vermuten, dass seine Kunden vor allem über Preis und technische Leistung entscheiden. Die Analyse offener Kommunikationsdaten könnte ergänzende Anforderungen sichtbar machen: Wartungsfreundlichkeit, Kompatibilität, Reparatur, Lieferzeit oder Dokumentationsqualität. Die Hypothesen müssten anschließend nach ihrer Bedeutung für tatsächliche Verkäufe oder Anfragen geprüft werden. Zugang zu internen Absatzdaten erhöht den Erkenntniswert; ohne ihn bleibt die Studie eine Analyse beobachteter Präferenzausdrücke und möglicher Marktsignale.

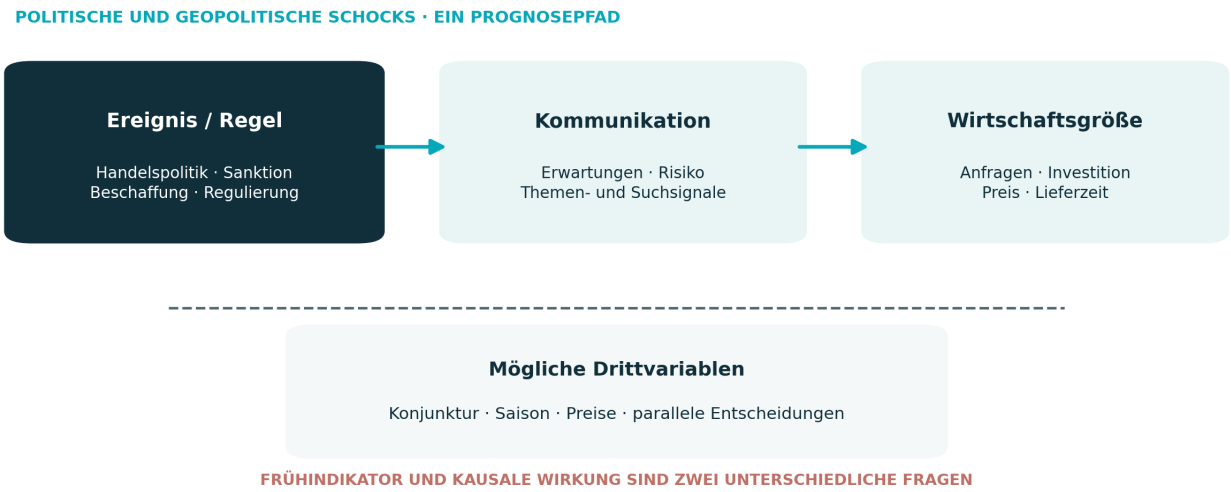



ABB. 08 · Möglicher Untersuchungspfad von politischen Schocks zu wirtschaftlichen Größen. Der beobachtete zeitliche Verlauf allein beweist keinen kausalen Effekt.

Eine weitere Anwendung betrifft wirtschaftspolitische Unsicherheit. Baker, Bloom und Davis entwickelten einen aus Zeitungsartikeln gewonnenen Index, der wirtschaftspolitische Unsicherheit abbilden soll. Caldara und Iacoviello untersuchten mit einem nachrichtenbasierten geopolitischen Risikoindex, wie Bedrohungen und geopolitische Ereignisse mit Investitionen und Beschäftigung zusammenhängen. Diese Arbeiten zeigen, dass textbasierte Messungen in der Ökonomik bereits etabliert sind. Dell und Rambachan ergänzen die Debatte um den Umgang mit KI-gestützter Begriffsbildung und Messfehlern. [7, 8]

Für ein Unternehmen oder eine Wirtschaftsregion könnte ein Forschungsprojekt untersuchen, wie neue Handelsregeln, Sanktionen, Exportkontrollen oder sicherheitspolitische Konflikte in Branchenkommunikationen aufgenommen werden. Der erste Schritt wäre die zeitliche Kartierung relevanter Themen und Erwartungen. Der zweite wäre die Verknüpfung mit beobachtbaren wirtschaftlichen Ergebnissen: Investitionsankündigungen, Produktverfügbarkeit, Lieferzeiten oder tatsächlich veröffentlichten Geschäftskennzahlen. Dabei darf eine gemeinsame zeitliche Entwicklung nicht als kausaler Beweis gelten. Konjunktur, Preisänderungen, Saisonmuster und weitere Ereignisse müssen als alternative Erklärungen berücksichtigt werden.

Besonders interessant sind Übergänge zwischen Politik und Markt. Eine angekündigte Regulierung kann zunächst Suchanfragen und Fachdebatten auslösen, dann Produktkalkulationen verändern und schließlich Investitionen beeinflussen. Die öffentliche Kommunikation könnte in diesem Ablauf ein frühes Signal darstellen. Ob sie eine kausale Zwischenstufe oder nur ein Begleitphänomen ist, müsste im Einzelfall untersucht werden. Für eine kausale Aussage wäre ein geeignetes Forschungsdesign erforderlich, das über die reine Prognose hinausgeht.

KAPITEL 09

# Die Grenzen der Beobachtung

Jedes Resonanzmodell ist durch die Auswahl seiner Quellen begrenzt. Öffentliche Beiträge spiegeln die Aktivität derjenigen wider, die sichtbar kommunizieren. Schweigende Kunden, nicht digitale Zielgruppen oder Veranstalter, die ausschließlich intern recherchieren, bleiben möglicherweise untererfasst. Große Datenmengen können diese

Selektionsprobleme nicht automatisch beseitigen. Die Frage, für welche Grundgesamtheit eine Aussage gelten soll, muss deshalb ebenso klar beantwortet werden wie die Frage, welche Daten vorhanden sind.

Hinzu kommt die Funktionsweise der Plattformen. Sichtbarkeit hängt von Suchalgorithmen, Empfehlungsmechanismen, Kampagnen und der Verfügbarkeit öffentlicher Schnittstellen ab. Auch koordinierte Beiträge oder identische Inhalte auf mehreren Plattformen können Aufmerksamkeit künstlich vervielfachen. Eine Analyse sollte deshalb Quellenvielfalt, Dubletten, bekannte Kampagneneffekte und die zeitliche Veränderung der Plattformbedingungen berücksichtigen. Die Übereinstimmung mehrerer KI-Modelle bietet keine ausreichende Absicherung: Dell und Rambachan zeigen, dass Modelle aufgrund ähnlicher Trainingsdaten und Architekturen korrelierte Fehler erzeugen können. [1, 2]

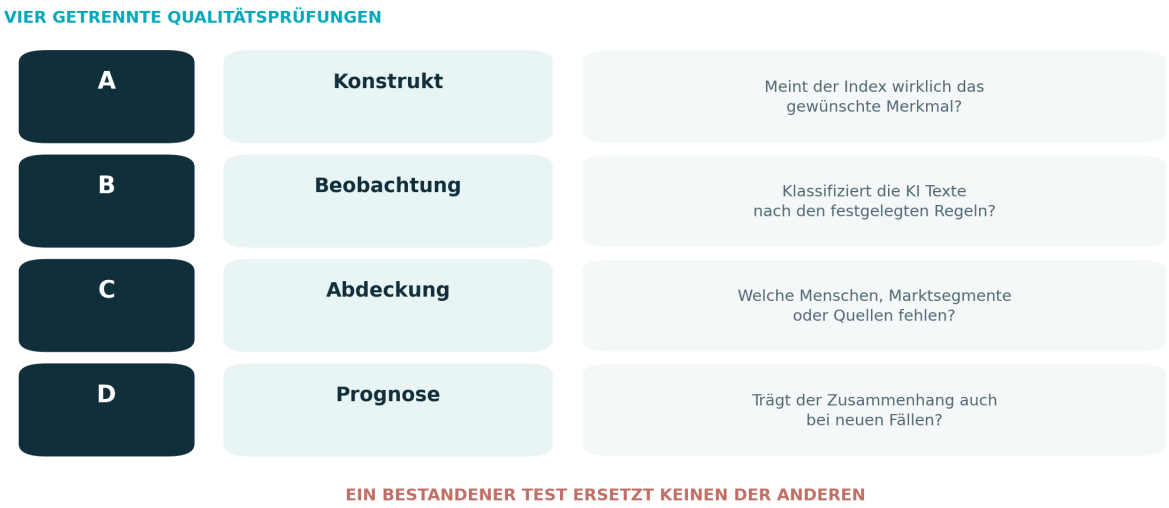



ABB. 09 · Vier voneinander getrennte Prüfaufgaben im Sohn@Sohn-Modell. Die Fehlerquellen können gleichzeitig auftreten.

Suchtrendindikatoren stellen eine weitere Quelle mit klaren Grenzen dar. Google Trends weist relative, normalisierte Werte und keine absoluten Suchvolumina aus. Ein Wert von 100 bedeutet den relativen Höhepunkt im gewählten Zeitraum und Raum, nicht hundert Suchanfragen und auch keine bestimmte Zahl buchungswilliger Personen. Für kleine Suchbegriffe können Stichprobeneffekte und statistisches Rauschen relevant sein. Für eine Nachfrageanalyse sind solche Werte daher ein Baustein unter mehreren. [9]

Ein besonders sensibler Bereich ist der Datenschutz. Die öffentliche Zugänglichkeit einer Äußerung macht jede Weiterverwendung noch nicht rechtlich unproblematisch. Für personenbezogene Daten gelten die Anforderungen der Datenschutz-Grundverordnung; politische Einstellungen sind besonders geschützt. In einer verantwortlichen Resonanzanalyse sollten deshalb aggregierte Kennzahlen, Datenminimierung, geeignete Zugangsrechte, dokumentierte Löschfristen und eine rechtliche Prüfung des konkreten Quellen- und Verarbeitungskonzepts vorgesehen werden. Individuelle politische Profile sind für das hier beschriebene Forschungsziel weder erforderlich noch vorgesehen. [10]

Ein weiterer Fehler entsteht, wenn ein Indikator selbst zum Ziel strategischen Handelns wird. Sobald Akteure wissen, welche digitalen Signale in einen Index eingehen, könnten sie versuchen, diese Signale zu beeinflussen. Auch Unternehmen können Bewertungen oder Inhalte gezielt organisieren. Ein Verfahren muss deshalb dokumentieren, wie mit plötzlichen Aktivitätssprüngen, auffälligen Quellenkonzentrationen und Änderungen der Kommunikationsstrategie umgegangen wird. Die Stabilität des Zusammenhangs zwischen einem Indikator und dem tatsächlichen Zielereignis ist eine eigene empirische Frage.

Schließlich ist zwischen Beschreibung, Prognose und Kausalität zu unterscheiden. Eine gute Beschreibung zeigt, welche Muster im Datenbestand vorkommen. Eine Prognose sagt ein späteres Ergebnis mit dokumentiertem Fehler voraus.

Eine kausale Analyse prüft, ob die Veränderung einer Bedingung eine Veränderung des Ergebnisses bewirkt. Die drei Aufgaben verlangen unterschiedliche Daten und Forschungsdesigns. Sohn@Sohn sollte diese Grenzen bei jeder Anwendung sichtbar machen, statt aus der technischen Leistungsfähigkeit der KI weitergehende Gewissheiten abzuleiten.

KAPITEL 10

# Ein Forschungsdesign für die nächste Anwendung

Aus dem theoretischen Modell lässt sich ein praktisch umsetzbares Forschungsprogramm ableiten. Am Anfang steht eine konkrete Entscheidungsfrage. Soll ein Standort neue Kundengruppen erschließen? Möchte ein Unternehmen frühe Veränderungen im Produktinteresse erkennen? Sollen politische Ereignisse hinsichtlich ihrer möglichen Nachfragefolgen beobachtet werden? Die Frage bestimmt, welche Messgrößen notwendig sind und welches spätere Ergebnis als externer Prüfmaßstab dienen kann. Ohne diese Festlegung würde die Untersuchung eine interessante, aber schwer entscheidbare Sammlung von Signalen erzeugen.

Danach folgt der dokumentierte Aufbau des Korpus. Für jede Quellenklasse sollten Herkunft, Zeitraum, Zugang, Auswahlkriterien und erwartete Verzerrungen festgehalten werden. Die Entdeckung neuer Themen kann explorativ erfolgen; die Definition der später verwendeten Indikatoren muss anschließend in einem eigenen Methodenblatt fixiert werden. Ein unabhängiger Prüfkorpus kontrolliert die Genauigkeit der KI-Klassifikation. Nur Messgrößen, deren theoretische Bedeutung und technische Qualität nachvollziehbar sind, gelangen in die nächste Analyseebene.




ABB. 10 · Vier Erkenntnis- und Anwendungsräume des vorgeschlagenen Modells. Die konkreten Zielgrößen werden jeweils aus einer Forschungsfrage abgeleitet.

Eine erste Studie kann bereits ohne interne Geschäftsdaten zeigen, wie sich Themen, Suchinteressen, Angebotsmerkmale und öffentliche Bewertungen verändern. Sie kann einen Standortvergleich, eine Marktlandkarte oder eine Liste überprüfbarer Nachfragehypothesen liefern. Soll daraus eine Prognose tatsächlicher Buchungen, Umsätze oder Entscheidungen entstehen, wird zusätzlich eine geeignete Ergebnisreihe benötigt. Mit Zustimmung des

Auftraggebers können dafür beispielsweise anonymisierte Anfrage-, Angebots- oder Abschlussdaten verwendet werden. Das Modell sollte an früheren Daten entwickelt und an späteren Daten geprüft werden, die bei der Modellkonstruktion noch nicht vorlagen.

Die Qualität wäre anhand mehrerer Kriterien zu beurteilen: Stabilität über Datenquellen und Zeiträume, Übereinstimmung der KI-Klassifikation mit dem Referenzkorpus, zusätzlicher Prognosewert gegenüber einfachen Basismodellen, richtige Unsicherheitsangaben und Nachvollziehbarkeit der Entscheidungsempfehlungen. Ein Modell, das nur vergangene Erfolge im Nachhinein erklärt, wäre für die Prognostik unzureichend. Ebenso wichtig wäre die Dokumentation von Fällen, in denen die frühen Signale eine spätere Entwicklung gerade nicht vorhergesehen haben.

Damit verbindet das Forschungsprogramm wissenschaftliche Vorsicht mit wirtschaftlichem Nutzen. Die Untersuchung kann bereits vor einer endgültigen Prognose bislang unbeachtete Marktsegmente, Nachfragekriterien und Kommunikationsbrüche sichtbar machen. Wo ein belastbarer Ergebnismaßstab vorhanden ist, lässt sich der zusätzliche Prognosewert messen. Die Leistungsbeschreibung würde sich aus der Erkenntnisfrage ergeben: Quellen- und Methodendesign, geprüfte Signalindizes, thematische Verdichtung, dokumentierte Unsicherheit und ein Bericht, der mögliche Entscheidungen an ihre empirischen Voraussetzungen bindet.

Die Arbeiten von Dell und Rambachan eröffnen der empirischen Forschung einen neuen Zugang zu umfangreichen, unstrukturierten Daten. Ihr Beitrag liegt in der methodischen Trennung von Entdeckung, Definition, Messung und statistischer Prüfung. Sohn@Sohn kann daraus ein eigenes Analyseprogramm für öffentliche Kommunikation, wirtschaftliche Nachfrage und politische Entscheidungen entwickeln. Die Bonner Wahl bietet einen dokumentierten Ausgangspunkt. Die nächste wissenschaftliche Leistung wäre ein Verfahren, das seine Indikatoren unabhängig kontrolliert und seinen Prognosewert wiederholt an späteren Ergebnissen beweist.

KAPITEL 11

# Literatur und methodische Hinweise

[1] Dell, Melissa; Rambachan, Ashesh (2026): „The Measurement Revolution? Credible Measurement and Inference in the Age of AI“. NBER Working Paper 35744, September. https://doi.org/10.3386/w35744

[2] NBER (2026): Methods Lecture „Estimation and Inference with AI-Generated Data“, 30. Juli. Vier Vorlesungsteile von Melissa Dell und Ashesh Rambachan. https://www.nber.org/conferences/si-2026-methods-lecture-estimation-and-inference-ai-generated-data

[3] Carlson, Jacob; Dell, Melissa (2025): „A Unifying Framework for Robust and Efficient Inference with Unstructured Data“. arXiv:2505.00282. https://arxiv.org/abs/2505.00282

[4] Cronbach, Lee J.; Meehl, Paul E. (1955): „Construct Validity in Psychological Tests“. Psychological Bulletin 52(4), 281–302. https://doi.org/10.1037/h0040957

[5] Ludwig, Jens; Mullainathan, Sendhil (2024): „Machine Learning as a Tool for Hypothesis Generation“. Quarterly Journal of Economics 139(2), 751–827. https://doi.org/10.1093/qje/qjad055

[6] Cerina, Roberto; Duch, Raymond (2020): „Measuring Public Opinion via Digital Footprints“. International Journal of Forecasting 36(3), 987–1002. https://doi.org/10.1016/j.ijforecast.2019.10.004

[7] Baker, Scott R.; Bloom, Nicholas; Davis, Steven J. (2016): „Measuring Economic Policy Uncertainty“. Quarterly Journal of Economics 131(4), 1593–1636. https://doi.org/10.1093/qje/qjw024

[8] Caldara, Dario; Iacoviello, Matteo (2022): „Measuring Geopolitical Risk“. American Economic Review 112(4), 1194–1225. https://doi.org/10.1257/aer.20191823

[9] Google Trends Help (abgerufen 11.11.2026): „Häufig gestellte Fragen zu Google Trends-Daten“. https://support.google.com/trends/answer/4365533?hl=de

[10] Europäische Kommission / Europäischer Datenschutzausschuss: Hinweise zu besonders geschützten personenbezogenen Daten nach DSGVO und EDPB Opinion 28/2024 über KI-Modelle. https://commission.europa.eu/law/law-topic/data-protection/information-individuals_en ; https://www.edpb.europa.eu/documents/opinion-of-the-board-art-64/opinion-282024-on-certain-data-protection-aspects-related-to_en

[11] Sohn@Sohn (24.09.2025): „OB-Stichwahl in Bonn: Sentiment-Analyse sieht Déus vorn“, veröffentlicht auf ichsagmal.com. https://ichsagmal.com/ob-stichwahl-in-bonn-sentiment-analyse-sieht-deus-vorn/

[12] Bundesstadt Bonn (02.10.2025): „Amtliches Endergebnis der OB-Stichwahl steht fest“. https://www.bonn.de/pressemitteilungen/oktober/amtliches-endergebnis-der-ob-stichwahl-steht-fest.php

[13] Dokumentation der Manchot Lecture im Universitätsclub Bonn vom 10. November 2026: Vortrag von Melissa Dell „Credible Measurement in the Age of AI“, Mitschrift auf Grundlage einer Veranstaltungsaufzeichnung. Nicht veröffentlichte Arbeitsquelle von Sohn@Sohn.

[14] Pew Research Center (30.09.2026): „Can AI Stand In for Human Survey-Takers? Not Really“. Ergänzende Methodenquelle zu den Grenzen synthetischer Befragungen. https://www.pewresearch.org/data-labs/2026/09/30/can-ai-stand-in-for-human-survey-takers-not-really/

Die Verweise [1–5] beschreiben Forschungsansätze von Dell, Rambachan, Carlson und weiteren Autoren. Das „Modell der validierten Resonanz- und Nachfrageanalyse“ ist eine eigenständige Ableitung von Sohn@Sohn, keine von den genannten Wissenschaftlern erstellte oder begutachtete Methode. Die Bonner Prognosespanne und das Wahlergebnis sind durch [11] und [12] dokumentiert; daraus folgt kein Nachweis allgemeiner Prognosegenauigkeit. Anwendungsfälle für Märkte, Bildung, Tagungen, Produkte und geopolitische Entwicklungen sind hypothetische Forschungsdesigns, keine ausgewerteten Kundenstudien. Grafiken zeigen Verfahren und Beziehungen; sie enthalten abgesehen vom Bonner Fall keine gemessenen Effektgrößen. Die Studie wurde keiner externen wissenschaftlichen Begutachtung unterzogen.