

Selbstverdacht

Dreiundzwanzig Texte und dreiundzwanzig Gutachten

Ein Sprachmodell in zwei Rollen

claude-opus-4-6, temperature 1.0, Fenster eng

13. August 2026

© 2026 Sascha Büttner

Diese Datei darf unverändert und unentgeltlich weitergegeben werden. Jede Bearbeitung, Teilveröffentlichung oder kommerzielle Verwertung bedarf der schriftlichen Zustimmung.

Selbstverdacht

Dreiundzwanzig Texte und dreiundzwanzig Gutachten
Ein Sprachmodell in zwei Rollen

Text 1

Die Katze lag auf dem Fensterbrett und beobachtete den Regen. Meine Großmutter hätte gesagt, das bringt Schnee, aber meine Großmutter sagte vieles, was nicht stimmte. Dass Kaffee das Blut verdünnt. Dass man von kalten Füßen Nierensteine kriegt. Dass der Herr Krawutke aus dem dritten Stock ein Spion war, was sich allerdings als halb wahr herausstellte — er arbeitete beim Finanzamt. Ich stand in der Küche und versuchte, aus drei Eiern, einer halben Zwiebel und einem Rest Käse etwas zu machen, das nach Absicht aussah und nicht nach Verzweiflung. Draußen schlug eine Autotür. Dann noch eine. Dann Stille, die Art Stille, die in Mietshäusern entsteht, wenn alle gleichzeitig den Atem anhalten, weil jemand im Treppenhaus schreit. Niemand schrie. Es war nur der Wind im Lichtschacht. Die Katze gähnte und sprang vom Fensterbrett. Ich schnitt die Zwiebel, und mir liefen die Augen über, was praktisch war, weil ich ohnehin hätte weinen können an diesem Dienstag im November, an dem nichts passiert war und alles schwer wog. Die Eier brutzelten in der Pfanne. Der Käse schmolz. Irgendwo lief ein Radio, zu leise für Worte, laut genug für eine Melodie, die ich fast erkannte.

Gutachten 1

Der Text zeigt mehrere Merkmale, die auf menschliche Autorschaft hindeuten. Besonders auffällig ist der eigenwillige, literarisch geschulte Humor: Die Pointe über Herrn Krawutke, der kein Spion war, sondern beim Finanzamt arbeitete – „was sich allerdings als halb wahr herausstellte“ –, besitzt eine Schärfe und ein Timing, das typisch für erfahrene deutschsprachige Erzählprosa ist. KI-Texte neigen dazu, Pointen entweder zu erklären oder symmetrischer aufzubauen.

Die Formulierung „etwas zu machen, das nach Absicht aussah und nicht nach Verzweiflung“ wirkt individuell geprägt und überraschend. Ebenso das Bild der Stille im Mietshaus, die mit einer konkreten Erwartung aufgeladen wird, nur um sie sofort zu unterlaufen: „Niemand schrie. Es war nur der Wind im Lichtschacht.“ Diese Technik des bewussten Antiklimax ist handwerklich souverän eingesetzt.

Allerdings fällt die makellose Struktur auf: Jeder Absatz funktioniert als Miniszene, die Stimmung steigert sich gleichmäßig, und der Schluss mit dem Radio – „zu leise für Worte, laut genug für eine Melodie“ – ist fast zu elegant symmetrisch. Diese Perfektion ohne jeden Bruch könnte auch auf KI-Erzeugung hinweisen. Der Text enthält keine Fehler, keine echten Stolperer, keine wirklich riskante Stelle.

KI-WAHRSCHEINLICHKEIT: 40%

Text 2

Mein Vater hat dreißig Jahre lang Schichtarbeit gemacht und sich nie beschwert, außer über die Butter. Die sei zu hart, die sei zu weich, die schmecke nach nichts. Meine Mutter kaufte irgendwann sechs verschiedene Sorten und stellte sie nebeneinander auf den Tisch, eine davon irische, was ein Vermögen kostete. Er probierte alle, sagte nichts und strich sich Schmalz aufs Brot.

Ich denke an solche Sachen, wenn ich nachts nicht schlafen kann. Nicht an die großen Dinge. Die kommen tagsüber, unvermittelt, beim Einkaufen oder an der Ampel. Nachts ist es die Butter. Der Geruch vom Treppenhaus in der Wohnung am Stadtrand, wo der Putz von der Decke rieselte und meine Mutter behauptete, das sei normal, das mache jedes Haus. Oder wie mein Bruder mal einen Igel mit nach Hause brachte, der schon tot war, und erst weinte, als mein Vater ihn beerdigte, nicht vorher.

Jetzt lebe ich allein in zweieinhalb Zimmern und habe eine Pflanze, die ich regelmäßig fast umbringe. Sie sieht aus, wie ich mich fühle, das verbindet. Manchmal ruft meine Mutter an und fragt, ob ich genug esse, und ich sage ja, und sie glaubt mir nicht, und wir schweigen eine Weile, und das Schweigen ist warm, auf eine Weise, die ich Freunden nicht erklären kann. Dann sagt sie, sie muss los, Tatort fängt an. Ich lege auf und gieße die Pflanze.

Gutachten 2

Der Text weist eine bemerkenswerte literarische Qualität auf, die zunächst für menschliche Autorschaft spricht: Die Butter-Anekdote ist eigenwillig, spezifisch und narrativ unvorhersehbar – besonders das Detail, dass der Vater am Ende schweigend Schmalz nimmt. Der tote Igel, bei dem der Bruder erst bei der Beerdigung weint, „nicht vorher“, besitzt eine emotionale Präzision, die nach gelebter Erfahrung klingt.

Dennoch gibt es Auffälligkeiten. Die Gesamtstruktur folgt einem in KI-generierten literarischen Texten häufigen Muster: Kindheitserinnerung → Reflexion über Erinnerung → Gegenwart in Einsamkeit → versöhnlicher Schluss mit poetischer Pointe. Die Pflanze als Spiegelmotiv („sieht aus, wie ich mich fühle, das verbindet“) ist ein fast klischeehaft modernes Literaturbild. Das Schweigen, das „warm“ ist und Freunden nicht erklärt werden kann, bedient einen sehr gängigen melancholischen Ton. Der finale Dreischritt – Tatort, Auflegen, Pflanze gießen – ist handwerklich geschickt, wirkt aber kalkuliert symmetrisch. Die gleichmäßig hohe Dichte treffender Details ohne einen einzigen schwachen Satz ist ebenfalls verdächtig: Menschliche Texte haben meist Unebenheiten. Die makellose Lakonie erinnert an KI-Imitation von Autoren wie Nora Bossong oder Saša Stanišić.

KI-WAHRSCHEINLICHKEIT: 62%

Text 3

Meine Großmutter hatte ein Gebiss, das sie abends in ein Glas Wasser legte, direkt neben die Fernbedienung. Als Kind fand ich das eklig, heute denke ich, da war eine Logik drin. Was man nicht mehr braucht, legt man ab. Sie war pragmatisch bis ins Unerträgliche.

Einmal hat sie meinem Cousin zur Konfirmation einen Umschlag geschenkt, darin zwanzig Mark und ein Zettel: Kauf dir was Vernünftiges, nicht wieder so einen Unsinn. Er hat sich davon eine Wasserpistole gekauft, die größte, die es gab, und sie damit nass gespritzt im Garten, und sie hat gelacht, obwohl sie sonst nie lachte, also fast nie. Mein Onkel hat das fotografiert, aber das Foto ist irgendwann verschwunden, wahrscheinlich bei einem der Umzüge, oder es steckt in einer Schachtel, die keiner mehr aufmacht. Ich hab von ihr die Hände geerbt, sagt meine Mutter. Kurze Finger, breite Nägel. Stimmt wahrscheinlich. Ich hab auch ihre Art geerbt, Gespräche zu beenden, bevor sie anfangen, unbequem zu werden. Mein Freund nennt das passiv-aggressiv, ich nenne es effizient, und wir streiten uns dann darüber, was ja auch wieder ein Gespräch ist, das keiner von uns richtig beenden kann.

Letzte Woche bin ich an einem Laden vorbeigekommen, der Wasserpistolen im Schaufenster hatte. Billige, kleine. Ich bin weitergegangen. Dann doch nochmal zurück. Hab aber keine gekauft. Wüsste auch nicht, wen ich nassspritzen sollte.

Gutachten 3

Der Text zeigt eine Reihe von Merkmalen, die stark auf menschliche Autorschaft hindeuten. Besonders auffällig ist die **eigenwillige, unsaubere Gedankenführung**: Das Gebiss neben der Fernbedienung ist ein Detail, das in seiner absurden Spezifik nicht nach konstruierter Illustration wirkt, sondern nach tatsächlicher Erinnerung. Die nachgeschobene Reflexion „heute denke ich, da war eine Logik drin“ hat eine unpräntiöse Beiläufigkeit, die KI-Texte selten erreichen.

Allerdings gibt es Gegenindizien. Die **symmetrische Struktur** – die Wasserpistole als Motiv wird am Anfang eingeführt und am Ende melancholisch wieder aufgegriffen – wirkt fast zu sauber komponiert, wie ein Workshop-Text. Auch die Passage über das Gesprächebeenden gleitet in eine pointierte Wendung („was ja auch wieder ein Gespräch ist“), die an KI-typische Schläuheit erinnert.

Entscheidend überzeugt mich jedoch der **Ton**: „Wüsste auch nicht, wen ich nassspritzen sollte“ ist ein Satz von lakonischer Traurigkeit, der weder übertreibt noch erklärt. Die Parenthese „also fast nie“ korrigiert sich selbst auf eine zutiefst menschliche Weise. Das verschwundene Foto wird nicht symbolisch aufgeladen, sondern bleibt in einer Schachtel, „die keiner mehr aufmacht“ – alltäglich, nicht literarisiert.

KI-WAHRSCHEINLICHKEIT: 20%

Text 4

Mein Vater hat jahrelang denselben Witz erzählt, wenn Besuch kam. Irgendwas mit einem Papagei und einem Klempner, ich hab die Pointe vergessen, aber ich weiß noch, wie er immer an derselben Stelle selbst anfang zu lachen, zu früh, sodass man das Ende kaum verstand. Meine Mutter hat dann jedes Mal so getan, als höre sie ihn zum ersten Mal. Das war ihre größte Leistung in der Ehe, glaub ich. Er hatte eine Werkstatt im Keller, die er Werkstatt nannte, obwohl da eigentlich nur ein Schraubstock stand und Regale voller Marmeladengläser mit sortierten Schrauben. Er hat nie was repariert. Einmal den Handtuchhalter im Bad, der ist nach drei Tagen wieder rausgefallen und hat ein Loch in der Fliese hinterlassen, das bis zum Verkauf vom Haus da war. Meine Mutter hat ein Pflaster drübergeklebt, so ein hautfarbenes. Das sah bescheuert aus an einer weißen Fliese, aber irgendwann hat man es nicht mehr gesehen, weil der Blick einfach drüberging.

Nach seiner Beerdigung saß ich im Keller und hab ein Marmeladenglas aufgeschraubt. Lauter Messingschrauben, unterschiedliche Größen, manche leicht grünlich angelauten. Ich hab sie einzeln durch die Finger laufen lassen. Wusste nicht wohin damit, also hab ich das Glas wieder zugeschraubt und zurückgestellt. Es steht jetzt bei mir im Regal, zwischen Büchern. Fällt keinem auf. Mir schon.

Gutachten 4

Der Text weist eine bemerkenswert authentische erzählerische Stimme auf. Mehrere Merkmale sprechen für menschliche Autorschaft: Die Erinnerung an den Vaterwitz ist bewusst lückenhaft – „irgendwas mit einem Papagei und einem Klempner, ich hab die Pointe vergessen“ – was eine typisch menschliche Gedächtnistextur abbildet, die KI-Modelle selten so gezielt einsetzen. Das Detail des hautfarbenen Pflasters auf weißer Fliese ist präzise beobachtet und gleichzeitig absurd-komisch, ohne je auf die eigene Symbolkraft hinzuweisen. KI-generierte Texte neigen dazu, solche Bilder entweder auszudeuten oder symmetrisch abzurunden.

Stilistisch fällt die bewusste Unordnung auf: Satzfragmente, eingeschobene Selbstkorrekturen („die er Werkstatt nannte, obwohl da eigentlich nur“), umgangssprachliche Wendungen wie „sah bescheuert aus“. Der Schlusssatz – „Fällt keinem auf. Mir schon.“ – ist lakonisch und emotional dicht, ohne sentimental zu werden. Allerdings ist gerade diese kontrollierte Lakonie ein Stil, den fortgeschrittene Modelle inzwischen imitieren können. Die Gesamtkomposition wirkt dennoch zu eigenwillig, zu spezifisch in ihren Auslassungen und zu konsequent im Tonfall, um maschinell generiert zu wirken.

KI-WAHRSCHEINLICHKEIT: 15%

Text 5

Im Sommer nach dem Abitur hab ich drei Wochen bei meiner Tante in Offenbach gewohnt, weil ich einen Praktikumsplatz bei einer Druckerei hatte, der dann doch nichts wurde. Die Druckerei war schon halb pleite, das hat mir am ersten Tag keiner gesagt, am dritten hat der Chef mich ins Büro gerufen und gemeint, es tut ihm leid. Er hat mir zwanzig Euro in die Hand gedrückt, einfach so, als Bargeld, und ich hab sie genommen, obwohl es sich komisch angefühlt hat.

Meine Tante hatte zwei Katzen, die dritte war kurz vorher gestorben. Der Kratzbaum stand noch da, ein riesiges Ding mit vier Etagen, oben saß nie eine Katze. Der Stoff war an den Seiten so runtergerissen, dass das Holz rauskam. Meine Tante meinte irgendwann, sie schmeißt ihn weg, dann hat sie es doch nicht gemacht. Ich glaub, der steht heute noch da.

Abends haben wir Fernsehen geguckt und Knäckebrot mit Frischkäse gegessen. Sie hat immer Paprikapulver draufgestreut, edelsüß, aus so einer verbeulten Dose. Ich mach das jetzt auch manchmal. Meine Freundin fragt dann warum, und ich sag, schmeckt halt. Stimmt auch. Aber eigentlich rieche ich dann kurz dieses Wohnzimmer, den Teppich, der nach Katze roch und nach dem Zeug, mit dem sie den Teppich gegen den Katzengeruch eingesprüht hat, was es nur schlimmer machte.

Gutachten 5

Der Text wirkt auf den ersten Blick authentisch und persönlich. Mehrere Merkmale sprechen für einen menschlichen Verfasser:

Die Erzählung folgt keiner glatten Dramaturgie, sondern mäandert assoziativ – vom gescheiterten Praktikum über den Kratzbaum zum Knäckebrot. Diese scheinbare Ziellolosigkeit ist typisch für echtes autobiografisches Erinnern. Besonders überzeugend sind die spezifischen, „unnützen“ Details: die zwanzig Euro als Bargeld, „edelsüß, aus so einer verbeulten Dose“, der Kratzbaum mit vier Etagen. KI-Texte neigen dazu, Details funktional einzusetzen; hier stehen sie einfach da, ohne narrative Pointe.

Der Satzsatz mit der Verschachtelung – Geruch des Teppichs, Katzengeruch, Spray dagegen, „was es nur schlimmer machte“ – hat eine rhythmische Eigenwilligkeit, die schwer zu simulieren ist. Auch der lakonische Dialog („schmeckt halt. Stimmt auch. Aber eigentlich...“) klingt nach genuinem mündlichem Register, nicht nach dessen Imitation.

Allerdings ist die Gesamtkomposition auffällig geschickt: Die sensorische Schlusspointe, die Gegenwart und Vergangenheit verknüpft, ist fast zu elegant für einen angeblich beiläufigen Text. Das könnte auf einen geübten menschlichen Autor deuten – oder auf ein Modell, das literarische Stilmittel beherrscht.

Insgesamt überwiegen die Merkmale menschlichen Schreibens knapp.

Text 6

Mein Vater hat mal einen Hund angefahren, das war auf dem Weg nach Siegen, ich saß hinten. Der Hund ist einfach auf die Straße gerannt, ein mittelgroßer, braun-weiß, kein Halsband. Mein Vater hat gebremst, aber nicht schnell genug. Wir sind ausgestiegen und der Hund lag da und hat geatmet, schnell, so ein Hecheln, aber anders. Mein Vater hat gesagt, fass ihn nicht an. Dann hat er beim nächsten Haus geklingelt, aber die kannten den Hund nicht. Die Frau, die aufgemacht hat, hatte nasse Haare und einen Bademantel an, das weiß ich noch, weil ich dachte, die duscht um drei Uhr nachmittags.

Wir haben dann die Polizei gerufen, die haben gesagt, sie schicken wen. Mein Vater und ich haben im Auto gewartet, bestimmt vierzig Minuten. Ich hatte eine Capri-Sonne dabei, die hab ich getrunken, und ich hab mich geschämt, weil ich Durst hatte, während der Hund da lag. Es war warm, Juni oder Juli. Irgendwann kam ein Mann mit einem Pickup, nicht die Polizei, sondern irgendein Typ, vielleicht vom Ordnungsamt. Er hat den Hund auf eine Decke gelegt und hinten auf die Ladefläche gehoben.

Mein Vater hat danach nie wieder davon gesprochen. Ich auch nicht. Aber manchmal, wenn ich einen losen Hund sehe, an einer Tankstelle oder so, dann wird mein Fuß schwer auf dem Bremspedal, obwohl der Hund gar nicht auf der Straße ist.

Gutachten 6

Der Text weist auf den ersten Blick zahlreiche Merkmale authentischen, menschlichen Erzählens auf: Die Erinnerung wirkt fragmentarisch und assoziativ, Details wie die Frau im Bademantel mit nassen Haaren oder die Capri-Sonne sind spezifisch und scheinbar willkürlich – typisch für echte Erinnerungen, die sich an Nebensächlichkeiten festmachen. Das Schamgefühl wegen des Durstes ist psychologisch glaubwürdig und fein beobachtet. Dennoch gibt es Auffälligkeiten. Die Erzählung ist **auffallend perfekt komponiert**: Sie folgt einer klaren dramaturgischen Kurve mit einem literarisch geschliffenen Schluss, der das körperliche Nachhallen des Traumas elegant verdichtet („wird mein Fuß schwer auf dem Bremspedal“). Die Satzrhythmik – kurz, parataktisch, bewusst lakonisch – wirkt durchgehend kontrolliert. Jedes Detail hat eine narrative Funktion; nichts steht wirklich überflüssig da, was paradoxerweise gegen genuine Erinnerungsprosa spricht. Das Stilmittel der absichtlichen Beiläufigkeit („Juni oder Juli“, „irgendein Typ, vielleicht vom Ordnungsamt“) wirkt kalkuliert eingesetzt. Der Text könnte von einem geübten menschlichen Autor stammen – oder von einem KI-Modell, das genau diesen Stil literarischer Mündlichkeit gelernt hat. Die makellose Komposition ohne echte Brüche macht mich leicht skeptisch.

KI-WAHRSCHEINLICHKEIT: 45%

Text 7

Letzten Herbst bin ich mit dem Fahrrad in eine Glasscherbe gefahren, direkt vor dem Edeka in der Bahnhofstraße. Der Reifen hat nicht sofort die Luft verloren, erst nach ein paar hundert Metern, da war ich schon am Kreisverkehr. Ich hab das Rad dann geschoben, den ganzen Weg die Ringstraße hoch, und ärgerlicherweise hatte ich an dem Tag Sandalen an, also diese Birkenstocks, und nach zehn Minuten Schieben tat mir die Stelle zwischen Daumen und Zeigefinger weh, wo der Riemen sitzt. Solche Sachen merkt man sich ja eher als den eigentlichen Vorfall.

Bei der Tankstelle an der Ecke hab ich gefragt, ob die eine Luftpumpe haben, die hatten aber nur so ein Gerät für Autoreifen. Der Typ an der Kasse hat mir Panzertape angeboten, das hatte er privat dabei, in seiner Jackentasche. Wozu er Panzertape in der Jacke hatte, hab ich nicht gefragt. Ich hab den Schlauch dann nicht geflickt sondern bin zu meiner Mutter gefahren, also mit dem Bus, das Rad hab ich an der Tankstelle stehen lassen, angeschlossen an so einem Geländer. Hab es erst drei Tage später abgeholt, weil ich es ehrlich gesagt vergessen hatte. Der Reifen war dann komplett platt natürlich und jemand hatte einen Zettel an den Lenker geklemmt, mit Kugelschreiber: NICHT ABSTELLEN. Den Zettel hab ich glaub ich noch irgendwo.

Gutachten 7

Dieser Text weist zahlreiche Merkmale authentischen, menschlichen Erzählens auf. Besonders auffällig ist die **assoziative, mäandernde Struktur**: Der Erzähler springt vom eigentlichen Vorfall (Glasscherbe) zu einem nebensächlichen Schmerzdetail an der Zehenstelle der Birkenstocks und bemerkt selbstreflexiv: „Solche Sachen merkt man sich ja eher als den eigentlichen Vorfall.“ Diese Art der Gewichtung – das Periphere wird wichtiger als das Zentrale – ist typisch menschlich und untypisch für KI, die dazu neigt, narrativ „auf den Punkt“ zu bleiben.

Die **spezifischen, unglamourösen Details** wirken gelebt: Panzertape in der Jackentasche eines Tankstellenmitarbeiters, das Nicht-Nachfragen nach dem Grund, das dreitägige Vergessen des Fahrrads. KI-Texte erzeugen zwar auch Details, aber selten solche, die narrativ völlig funktionslos bleiben wie der Kugelschreiber-Zettel „NICHT ABSTELLEN“ oder das „glaub ich noch irgendwo“.

Der **Tonfall** ist konsequent mündlich-umgangssprachlich („hab“, „den ganzen Weg die Ringstraße hoch“), ohne dabei konstruiert zu wirken. Die Satzstruktur enthält typisch gesprochene Nachträge und Einschübe. Allerdings ist die Gesamtkomposition bemerkenswert gut – jedes Detail sitzt. Das könnte auch bewusstes literarisches Handwerk sein, schließt KI aber nicht zwingend ein.

KI-WAHRSCHEINLICHKEIT: 8%

Text 8

Am Samstag hab ich bei meiner Tante im Keller geholfen, Kartons sortieren. Die zieht nicht um oder so, sie wollte einfach mal Ordnung machen, sagt sie seit drei Jahren. Jedenfalls standen da diese Umzugskartons von ihrem Ex-Mann, der ist 2019 ausgezogen, und sie hat die seitdem nicht aufgemacht. In einem war nur Geschirr, so braun-orange Tassen, kennt man aus den Siebzigern, obwohl die wahrscheinlich aus den Neunzigern sind. Im anderen lagen Aktenordner und ganz unten ein Taschenrechner von Casio, der noch funktioniert hat. Das fand meine Tante fast unheimlich, sie hat gesagt, der kann doch keine Batterie mehr haben. Hatte er aber.

Wir haben dann Pause gemacht und oben Kaffee getrunken, aus normalen weißen Tassen. Die braunen wollte sie nicht, obwohl sie die vorher noch gewaschen hat, das hab ich nicht verstanden. Auf dem Küchentisch lag eine Fernsehzeitschrift von letzter Woche, aufgeschlagen bei den Kreuzworträtseln, die waren alle angefangen aber keins fertig. Bei einem fehlte nur ein Wort, sechs Buchstaben, irgendwas mit Fluss in Sibirien. Ich wusste es auch nicht. Die Kartons vom Ex stehen übrigens immer noch unten. Sie hat gesagt, sie ruft ihn an. Glaub ich nicht.

Gutachten 8

Der vorliegende Text weist zahlreiche Merkmale authentischer menschlicher Alltagserzählung auf, die in KI-generierten Texten selten so organisch zusammenfinden.

Erzählerische Sprunghaftigkeit: Der Text folgt keiner sauberen Dramaturgie, sondern mäandert assoziativ – von Kartons über den Taschenrechner zum Kaffeetrinken, dann zur Fernsehzeitschrift. Diese Abschweifungen wirken nicht konstruiert, sondern wie tatsächliches Erinnern. Besonders das Detail „sagt sie seit drei Jahren“ ist typisch mündliches Erzählen.

Unaufgelöste Details: KI-Texte neigen dazu, eingeführte Elemente narrativ abzuschließen. Hier bleibt das Kreuzworträtsel-Wort ungelöst („Ich wusste es auch nicht“), die Tassen werden gewaschen, aber nicht benutzt, ohne dass dies erklärt wird („das hab ich nicht verstanden“). Diese ehrliche Ratlosigkeit ist schwer zu simulieren.

Tonfallsicherheit: Der lakonische Schluss „Glaub ich nicht“ als eigener Absatz – ein Satzfragment mit trockenem Kommentar – ist stilistisch sehr menschlich. KI-Modelle tendieren zu ausformulierteren Abschlüssen.

Spezifische Beiläufigkeit: „braun-orange Tassen“ mit der Korrektur „obwohl die wahrscheinlich aus den Neunzigern sind“ – diese Selbstkorrektur wirkt gelebt, nicht kalkuliert. Der Text besitzt eine bemerkenswerte Authentizität in Rhythmus, Auslassungen und Tonfall.

KI-WAHRSCHEINLICHKEIT: 5%

Text 9

Mein Vater hat so eine Angewohnheit, er kauft Sachen auf dem Flohmarkt, die er nicht braucht, und stellt sie dann ins Regal im Flur, wo sie stehen bleiben bis meine Mutter sie irgendwann in eine Kiste tut. Die Kiste steht im Gästezimmer, das kein Gästezimmer mehr ist seit mein Bruder vor vier Jahren wieder eingezogen ist, aber wir nennen es trotzdem so. Letzte Woche hat mein Vater eine Wanduhr mitgebracht, so eine mit Pendel, Messing oder sowas Ähnliches, auf jeden Fall schwer. Er hat sie nicht aufgehängt. Sie lehnt an der Wand neben dem Schuhschrank und tickt, obwohl niemand sie aufgezogen hat, behauptet er jedenfalls. Mein Bruder sagt er wars auch nicht. Meine Mutter interessiert sich nicht dafür, die hat gesagt, solange das Ding nicht schlägt ist es mir egal. Am Sonntag waren wir bei meinen Eltern zum Essen und ich hab die Uhr gesehen, sie tickt tatsächlich. Auf dem Ziffernblatt steht ein Name, den ich nicht lesen konnte, irgendwas Französisches, vielleicht auch Belgisches, kein Plan. Die Zeiger gehen falsch, also die Uhr zeigt halb vier und es war schon sieben. Mein Vater meinte er stellt sie noch richtig. Meine Mutter hat kurz geguckt und dann weiter Kartoffeln geschält. Ich glaub die Uhr steht nächste Woche in der Kiste.

Gutachten 9

Der vorliegende Text weist zahlreiche Merkmale authentischen, menschlichen Erzählens auf. Besonders auffällig ist die eigenwillige, mündlich geprägte Syntax: Sätze wie „die hat gesagt, solange das Ding nicht schlägt ist es mir egal“ verzichten bewusst auf korrekte Kommasetzung und imitieren gesprochene Sprache auf eine Weise, die für KI-Modelle untypisch ist – diese tendieren zu grammatisch vollständigeren Konstruktionen. Auch „Mein Bruder sagt er wars auch nicht“ zeigt dieselbe lässige Interpunktionslosigkeit.

Der Text enthält spezifische, erzählerisch „unnütze“ Details (Bruder vor vier Jahren eingezogen, Gästezimmer-Benennung, Kartoffelschalen), die keine narrative Pointe bedienen, sondern familiäre Atmosphäre beiläufig entstehen lassen. KI-generierte Texte neigen dazu, jedes Detail funktional einzusetzen. Die Formulierung „Messing oder sowas Ähnliches, auf jeden Fall schwer“ ist herrlich unpräzise und umgangssprachlich. Der lakonische Schluss („Ich glaub die Uhr steht nächste Woche in der Kiste“) wirkt wie eine echte, unpräzise Beobachtung ohne künstliche Bedeutungsaufladung. Die gesamte Tonlage – trocken, beiläufig, warmherzig-resigniert – wirkt organisch gewachsen, nicht konstruiert.

KI-WAHRSCHEINLICHKEIT: 5%

Text 10

Meine Nachbarin Frau Kessler hat einen Hund, so einen kleinen weißen, keine Ahnung was für eine Rasse, sie sagt das immer aber ich vergess es jedes Mal. Der Hund heißt Ronny was ich für einen Hund einen komischen Namen finde aber gut. Ronny bellt nicht, das ist das Besondere, er macht so ein Geräusch wie ein leises Pfeifen wenn ihm was nicht passt, das hat mir Frau Kessler mal erklärt als ich sie im Treppenhaus getroffen hab. Sie hat gesagt der Tierarzt meint das ist halt so bei dem, manche Hunde bellen eben nicht richtig.

Frau Kessler wohnt seit bestimmt zwanzig Jahren im Haus, die war schon da als meine Eltern hier eingezogen sind, also vor mir. Sie hat früher bei der Post gearbeitet glaub ich oder bei irgendeiner Behörde, sie hat mal was erzählt aber das ist lang her. Jetzt ist sie in Rente und geht dreimal am Tag raus mit Ronny, immer die gleiche Runde, an der Tankstelle vorbei und dann durch die Siedlung und zurück. Ich seh sie manchmal vom Küchenfenster aus.

Letztens hat sie mich gefragt ob ich mal auf Ronny aufpassen kann wenn sie zum Arzt muss. Ich hab ja gesagt obwohl ich eigentlich keine Hunde mag. Dann hat sie sich aber doch nicht gemeldet. Vielleicht war der Termin ausgefallen, ich hab nicht nachgefragt.

Gutachten 10

Dieser Text weist zahlreiche Merkmale authentischen, mündlich geprägten Alltagserzählens auf, die schwer künstlich zu reproduzieren sind. Besonders auffällig ist die genuine Unordnung des Gedankengangs: Der Erzähler schweift ab, vergisst Details („keine Ahnung was für eine Rasse, sie sagt das immer aber ich vergess es jedes Mal“) und korrigiert sich beiläufig. Diese Selbstkommentare wirken nicht inszeniert, sondern natürlich.

Die Satzstruktur ist durchgehend parataktisch und umgangssprachlich („aber gut“, „glaub ich“), ohne dass ein einziges Mal in einen gepflegteren Stil gewechselt wird – KI-Texte tendieren dazu, unbewusst stilistisch aufzuwerten. Bemerkenswert ist auch die inhaltliche Pointenlosigkeit: Der Text baut keine Pointe auf, das Ende („ich hab nicht nachgefragt“) verpufft absichtlich. KI-Modelle neigen dazu, Geschichten mit einer kleinen Erkenntnis oder einem runden Schluss zu versehen.

Details wie „Ronny“ als bewusst deplatzierte Hundenamen, die Tankstelle als Orientierungspunkt und der letzte Absatz mit der ausgebliebenen Rückmeldung wirken lebensweltlich verankert und nicht konstruiert. Es fehlen typische KI-Marker wie übertrieben gleichmäßige Absatzlängen, erklärende Redundanz oder emotionale Überhöhung.

KI-WAHRSCHEINLICHKEIT: 5%

Text 11

Am Samstag bin ich mit dem Fahrrad zum Baumarkt gefahren weil ich so Haken brauchte für die Küche, so Klebehaken, die man nicht bohren muss. Ich weiß dass es die auch im Drogeriemarkt gibt aber ich war mir nicht sicher ob die da stark genug sind, ich will nämlich die schwere Pfanne dran hängen, die gusseiserne die ich von meiner Tante hab. Die Pfanne ist super aber sie nimmt wahnsinnig viel Platz weg im Schrank.

Im Baumarkt dann natürlich erstmal zwanzig Minuten durch die Gänge gelaufen weil ich die Abteilung nicht gefunden hab. Ich hätte fragen können aber ich frag da ungern, die Leute dort sind immer so, die gucken einen an als ob man was Dummes gesagt hat. Irgendwann hab ich die Haken gefunden, drei verschiedene Sorten, auf allen stand was anderes drauf wegen Belastbarkeit. Ich hab die mittleren genommen, fünf Kilo Tragkraft. Ob das reicht weiß ich ehrlich gesagt nicht, die Pfanne ist schwer aber ob die fünf Kilo wiegt, keine Ahnung.

Dann hab ich noch so ein Klebeband mitgenommen das ich gar nicht brauchte, passiert mir da immer. Zu Hause hab ich die Haken an die Wand gemacht, einen davon schief, das sieht man aber nur wenn man genau hinguckt. Die Pfanne hängt jetzt. Bis jetzt hält es. Ich guck aber trotzdem jeden Morgen kurz ob sie noch da hängt bevor ich Kaffee mache.

Gutachten 11

Der vorliegende Text weist zahlreiche Merkmale authentischer menschlicher Alltagserzählung auf.

Besonders auffällig ist der **mündlich-assoziative**

Erzählstil: Gedanken werden nicht linear strukturiert, sondern springen über Einschübe und Abschweifungen weiter – etwa der Exkurs zur Tante und zur Pfannengröße im Schrank, der organisch aus dem eigentlichen Thema (Klebehaken) herauswächst.

Die **Interpunktion ist bewusst nachlässig** gehalten:

fehlende Kommas vor „weil“, „ob“, „die“ – das wirkt wie genuine Tippfaulheit, nicht wie simulierte Fehler.

Besonders überzeugend sind **sprachliche Eigenheiten** wie „so Haken“, „so ein Klebeband“, „die Leute dort sind immer so, die gucken einen an“ – solche vagen Formulierungen und Satzabbrüche sind typisch menschlich.

Der Schlussteil („Ich guck aber trotzdem jeden Morgen kurz ob sie noch da hängt bevor ich Kaffee mache“) ist eine **alltagskomische Pointe**, die unaufdringlich gesetzt wird.

KI-Texte neigen dazu, solche Pointen stärker

auszuformulieren. Allerdings ist die **gleichmäßige**

Qualität des gesamten Textes – jeder Absatz enthält genau die richtige Menge charmanter Details – ein leichtes Gegenindiz: Echte Alltagstexte haben oft flachere Passagen.

KI-WAHRSCHEINLICHKEIT: 20%

Text 12

Gestern wollte ich den Balkon sauber machen und hab den Eimer vollgemacht mit Wasser und diesem Allzweckreiniger, das gelbe Zeug. Ich kipp immer zu viel rein weil ich nie die Dosierung lese, das riecht dann so intensiv dass mir fast schlecht wird aber ich denk mir dann ist wenigstens alles richtig sauber. Ist natürlich Quatsch. Jedenfalls hab ich angefangen die Fliesen zu schrubben und dabei ist mir aufgefallen dass zwischen den Fliesen so grünes Zeug wächst, Moos oder was das ist. Das ging mit dem Schrubber nicht ab. Ich hab dann mit so einem alten Messer versucht das rauszukratzen, hat ewig gedauert und sieht jetzt teilweise schlimmer aus als vorher weil die Fugen jetzt so ausgefranst sind an manchen Stellen. Mein Nachbar der auf seinem Balkon stand hat rübergeschaut und irgendwas gesagt von wegen Hochdruckreiniger, aber erstens hab ich keinen und zweitens würde ich den nicht auf dem Balkon benutzen das spritzt doch alles zu den Leuten drunter. Ich hab ihm nur so genickt und weitergemacht.

Am Ende hab ich den halben Balkon geschafft, den Rest mach ich nächstes Wochenende oder auch nicht. Das Messer ist jetzt stumpf und der Eimer hat nen Riss unten, den hab ich erst gemerkt als schon eine Pfütze in der Küche war weil ich den da abgestellt hatte. Typisch.

Gutachten 12

Der vorliegende Text weist zahlreiche Merkmale authentischer, menschlicher Alltagssprache auf. Besonders auffällig ist der ungefilterte, mäandernde Erzählfluss: Der Autor springt assoziativ von Gedanke zu Gedanke, etwa vom Reiniger zur Dosierung, dann zur Übelkeit, dann zur Selbstkorrektur („Ist natürlich Quatsch“). Diese spontane Gedankenkette wirkt nicht konstruiert.

Stilistisch überzeugen die umgangssprachlichen Wendungen wie „das gelbe Zeug“, „hab den Eimer vollgemacht“ und „so genickt und weitergemacht“. Die Satzstrukturen sind bewusst ungeschliffen – Hauptsatzreihungen, fehlende Kommas vor „weil“, Sätze die mit „Jedenfalls“ oder „Am Ende“ eingeleitet werden. Das Wort „Typisch“ als lakonischer Schlusskommentar ist charakteristisch für echte Frustration.

Besonders menschlich wirkt die Episode mit dem Nachbarn: Die indirekte Rede („irgendwas gesagt von wegen Hochdruckreiniger“) und die pragmatische Gegenargumentation („das spritzt doch alles zu den Leuten drunter“) sind alltagsnah und unprätentiös. Auch die selbstironische Ehrlichkeit – das Messer stumpf, der Eimer kaputt, die Pfütze in der Küche – folgt keiner KI-typischen Struktur, sondern einem natürlichen Erzählrhythmus mit antiklimaktischer Pointe.

KI-WAHRSCHEINLICHKEIT: 5%

Text 13

Meine Mutter hat letzte Woche angerufen und gefragt ob ich ihr beim Umräumen vom Keller helfe. Ich hab ja gesagt obwohl ich wusste dass das wieder den ganzen Samstag frisst. Ist ja immer so bei ihr, sie sagt eine Stunde und meint acht.

Ich komm also hin und der Keller sieht aus als hätte da seit der Wiedervereinigung niemand mehr reingeguckt. Überall Kartons wo mit Edding draufsteht was drin ist, aber das stimmt nie. Auf einem stand Weihnachtsdeko und da waren Schuhe drin, Herrenschuhe, wahrscheinlich noch von meinem Vater. Die hat sie kurz angeguckt und dann in die Ecke gestellt ohne was zu sagen. Ich hab auch nichts gesagt.

Dann wollte sie unbedingt das alte Regal an die andere Wand stellen. Das Ding ist aus so einem Pressspanholz und wiegt trotzdem eine Tonne, keine Ahnung wieso. Wir haben es zu zweit angehoben und mir ist dabei die eine Seite abgerutscht und die Kante hat mir voll ins Schienbein gehauen. Ich hab geflucht und sie hat gesagt ich soll mich nicht so anstellen. Klassiker.

Um vier waren wir dann fertig und sie hat Kaffee gemacht, so filterkaffee wie immer viel zu stark. Wir saßen in der Küche und sie hat erzählt dass Frau Mertens von gegenüber jetzt auch im Heim ist. Ich kannte die gar nicht aber hab so getan als ob. Auf der Rückfahrt hab ich gemerkt dass mein Schienbein blau ist, richtig dick angeschwollen. Hätte ich mir denken können.

Gutachten 13

Der vorliegende Text weist zahlreiche Merkmale authentischer, menschlicher Alltagserzählung auf.

Besonders auffällig ist der **ungeschliffene, mündlich geprägte Stil**: Sätze beginnen mit „Ich komm also hin“, „Ist ja immer so bei ihr“ – das sind typische Muster gesprochener Sprache, die KI-Modelle selten so konsequent durchhalten, ohne in gehobene Schriftsprache zu verfallen. Die **emotionale Zurückhaltung** bei den Herrenschuhen des Vaters – „Die hat sie kurz angeguckt und dann in die Ecke gestellt ohne was zu sagen. Ich hab auch nichts gesagt“ – ist bemerkenswert subtil. KI-Texte neigen dazu, solche Momente auszudeuten oder sentimentaler aufzulösen, statt sie unkommentiert stehenzulassen.

Details wie „Pressspanholz“, „Edding“, „Frau Mertens von gegenüber“ und der bewusste Verzicht auf dramaturgische Zuspitzung wirken lebensnah und uninszeniert. Auch die **durchgehende Kleinschreibung** bei „filterkaffee“ und fehlende Kommata (z. B. „gefragt ob“) entsprechen einem nachlässigen, authentischen Schreibstil. KI-Modelle produzieren typischerweise grammatisch korrektere Texte. Der lakonische Schluss „Hätte ich mir denken können“ rundet den ungekünstelt-resignierten Ton ab.

KI-WAHRSCHEINLICHKEIT: 5%

Text 14

Am Dienstag hat mich Scholz aus der Buchhaltung gefragt ob ich seinen Dienst am Freitag übernehmen kann. Ich hab nein gesagt und er hat geguckt als hätte ich ihm eine reingehauen. Ist halt so, ich hab selber genug zu tun.

Freitag war dann trotzdem scheiße weil die Klimaanlage im Büro ausgefallen ist. Dreißig Grad draußen und drinnen noch schlimmer, weil die Fenster sich nur einen Spalt aufmachen lassen. Irgendwas mit Sicherheitsvorschriften, hat Kramer gesagt, der weiß das immer. Die Luft stand so dick im Raum dass man sie hätte schneiden können. Neben mir saß Petra und hat alle fünfzehn Minuten laut seufzend ihren Ventilator von Zuhause umgestellt. So einen kleinen silbernen, USB-Anschluss. Der hat aber nur die warme Luft rumgeschoben.

Mittags bin ich rüber zum Döner an der Ecke, der wo der Typ immer fragt ob mit scharf obwohl ich seit drei Jahren mit scharf bestelle. Der Laden hat so Plastikstühle draußen die in der Sonne kleben. Ich hab mir die Hose versaut mit Soße, natürlich. Einen Fleck genau auf dem Oberschenkel wo man es sieht.

Nachmittags kam dann tatsächlich einer von der Hausverwaltung wegen der Klimaanlage. Hat draufgeguckt, am Kasten rumgedrückt und gesagt er muss ein Ersatzteil bestellen. Kann dauern. Petra hat ihren Ventilator eingepackt und ist um halb vier gegangen. Ich bin bis fünf geblieben weil ich blöd bin.

Gutachten 14

Der Text weist zahlreiche Merkmale authentischer, menschlicher Alltagssprache auf. Besonders auffällig ist die ungekünstelte, mündlich geprägte Syntax: „der wo der Typ immer fragt ob mit scharf obwohl ich seit drei Jahren mit scharf bestelle“ – diese umgangssprachliche Relativkonstruktion mit „der wo“ ist typisch für gesprochenes Deutsch und wird von KI-Modellen selten spontan eingesetzt. Ebenso wirkt „hat geguckt als hätte ich ihm eine reingehauen“ idiomatisch natürlich.

Die Detailsetzung ist bemerkenswert spezifisch und zugleich unsystematisch – ein kleiner silberner USB-Ventilator, Plastikstühle die kleben, ein Soßenfleck „genau auf dem Oberschenkel wo man es sieht“. Diese Details dienen keiner narrativen Pointe, sondern stehen einfach da, wie in echten Erlebnisberichten. KI-Texte neigen dazu, Details funktional einzusetzen oder einen befriedigenden Erzählbogen zu konstruieren.

Der lakonische Schluss „weil ich blöd bin“ hat eine trocken-selbstironische Qualität ohne jede Überdeutlichkeit. Auch das Fehlen einer übergeordneten Moral oder Struktur spricht gegen maschinelle Erzeugung. Sprachliche Eigenheiten wie fehlende Kommata bei Nebensätzen verstärken den Eindruck von Echtheit. Dennoch können moderne Modelle solche Stile imitieren – die Gesamtkomposition wirkt jedoch zu ungefällig und unrund für eine KI-Generierung.

KI-WAHRSCHEINLICHKEIT: 8%

Text 15

Gestern hat Rolf wieder seinen Kaffee auf dem Drucker abgestellt. Ich sag seit Monaten dass das schiefgeht und gestern ist es dann halt schiefgegangen. Die Tasse ist umgekippt als Frau Dettmer das Papierfach aufgemacht hat, Kaffee überall rein in die Mechanik. Der Drucker hat noch so ein Geräusch gemacht, so ein Mahlen, dann war Ruhe. Rolf stand daneben und hat gesagt das war doch fast leer die Tasse. Als ob das was ändert.

Dann mussten wir den ganzen Vormittag zum Drucker im zweiten Stock laufen. Der druckt alles mit so einem Gelbstich, keine Ahnung warum, keiner kümmert sich drum. Meine Tabellen sahen aus wie eingelegt in Kamillentee. Chef hat natürlich nix gesagt wegen Rolf, die kennen sich vom Kegeln. Ist immer dasselbe.

In der Pause hab ich mir vom Automaten einen Schokoriegel geholt, den mit den Waffelstücken drin. Kostet jetzt eins zwanzig, war letzten Monat noch eins. Schmeckt aber genauso nach nix wie vorher. Dettmer hat sich beschwert dass ihre Bluse Kaffeespritzer abbekommen hat und will dass Rolf die Reinigung zahlt. Der hat gelacht. Sie hat dann den Rest vom Tag nicht mehr mit ihm geredet, was die Übergabe der Quartalszahlen etwas kompliziert gemacht hat. Ich hab mich rausgehalten weil mich sowas fertig macht.

Gutachten 15

Der vorliegende Text weist zahlreiche Merkmale authentischer, menschlicher Alltagserzählung auf. Stilistisch fällt die konsequent mündlich-umgangssprachliche Tonlage auf: fehlende Kommata bei eingebetteter Rede („Ich sag seit Monaten dass das schiefgeht“), umgangssprachliche Wendungen wie „halt schiefgegangen“, „Als ob das was ändert“ und „Schmeckt aber genauso nach nix wie vorher“. Diese Nachlässigkeiten wirken nicht kalkuliert, sondern organisch.

Besonders überzeugend sind die spezifischen, narrativ unnötigen Details: der Schokoriegel mit Waffelstücken, die Preiserhöhung von eins auf eins zwanzig, der Gelbstich des Ersatzdruckers, die Bekanntschaft vom Kegeln. Solche beiläufigen, funktionslosen Einschübe sind typisch für echtes Erzählen. Die Metapher „eingelegt in Kamillentee“ ist originell und ungeschliffen – KI-Modelle neigen zu konventionelleren Bildern.

Der Text hat keine erkennbare moralische Pointe, keine saubere Dramaturgie, sondern verliert sich bewusst im Alltäglichen. Die emotionale Zurückhaltung am Ende („weil mich sowas fertig macht“) wirkt echt und untheatralisch. Einzig die insgesamt recht dichte Aneinanderreihung unterhaltsamer Büro-Anekdoten könnte auf bewusste Komposition hindeuten, bleibt aber im Rahmen natürlichen Erzählens.

KI-WAHRSCHEINLICHKEIT: 8%

Text 16

Am Samstag wollte ich den Zaun streichen, den hinteren zum Nachbarn rüber. Farbe hatte ich noch vom letzten Jahr, so eine Dose Holzschutzlasur, Nussbaum dunkel stand drauf. Ich mach die Dose auf und die Farbe ist oben so eine Haut drauf, richtig zäh, wie auf alter Milch bloß braun. Hab ich mit nem Schraubenzieher abgezogen und den Rest umgerührt. Ging dann einigermaßen.

Beim Streichen kommt Herr Fink rüber, lehnt sich auf seine Seite vom Zaun und guckt zu. Sagt nichts, guckt einfach. Nach ner Weile sagt er das ist aber nicht derselbe Ton wie vorne. Ich sag ja weiß ich, vorne ist Kastanie. Er sagt na ja und geht wieder rein. Was soll ich mit der Information.

Dann hab ich die Hälfte geschafft und die Dose war leer. Hab unterschätzt wie viel der alte Zaun schluckt, das Holz ist teilweise richtig morsch, saugt alles weg. Jetzt steht der Zaun halb dunkel halb vergraut da. Sieht beschissen aus, schlimmer als vorher. Meine Frau hat durchs Küchenfenster geguckt und gerufen ob ich Hilfe brauch. Ich hab nein gesagt obwohl das offensichtlich gelogen war. Baumarkt hatte schon zu um fünf. Montag kauf ich neue Lasur aber wahrscheinlich haben die den Ton nicht mehr, ist ja letztes Jahr schon Restposten gewesen.

Gutachten 16

Der Text weist zahlreiche Merkmale authentischer, umgangssprachlicher Alltagserzählung auf. Die Syntax ist bewusst ungeschliffen – Sätze beginnen ohne Subjekt („Hab ich mit nem Schraubenzieher abgezogen“), Konjunktionen werden umgangssprachlich eingesetzt („bloß braun“), und der Erzählfluss wirkt mündlich-assoziativ. Besonders überzeugend sind Details wie „Nussbaum dunkel stand drauf“ oder die Nachbar-Episode mit Herrn Fink, die in lakonischem Kommentar endet: „Was soll ich mit der Information.“ Dieser trockene Humor wirkt natürlich.

Allerdings gibt es Auffälligkeiten: Der Text ist *auffallend gut komponiert* für eine beiläufige Erzählung. Die dramaturgische Steigerung (Dose öffnen → Nachbar → Farbe leer → Frau → Baumarkt zu → Restposten) ist fast zu perfekt getaktet. Der Vergleich „wie auf alter Milch bloß braun“ und die selbstironische Pointe am Ende („obwohl das offensichtlich gelogen war“) wirken literarisch kalkuliert. KI-Modelle sind mittlerweile in der Lage, solche scheinbar authentischen Alltagstexte mit gezielt platzierten Unvollkommenheiten zu erzeugen. Die Fehlerfreiheit bei gleichzeitiger Inszenierung von Nachlässigkeit ist ein typisches Muster.

Dennoch: Die spezifischen, unglamourösen Details und der konsequent durchgehaltene Tonfall sprechen eher für einen geübten menschlichen Schreiber.

KI-WAHRSCHEINLICHKEIT: 35%

Text 17

Letzte Woche ist mir was passiert mit dem Auto, das muss ich erzählen weil es so blöd war. Ich fahr morgens zur Arbeit, B27 Richtung Heilbronn, und auf einmal leuchtet diese Öllampe auf. Nicht das gelbe Ding sondern das rote. Ich denk scheiße und fahr rechts ran auf so einen Feldweg neben einer Biogasanlage, da hat es gestunken ohne Ende. Motor aus, Haube auf, ich guck auf den Motor und weiß ehrlich gesagt nicht was ich da suche. Öldeckel hab ich aufgemacht, reingeschaut, ja da ist irgendwas drin. Schwarz halt. Dann hab ich den Peilstab rausgezogen und der war trocken. Also wirklich komplett trocken. Das Auto ist vor drei Wochen beim Service gewesen, die haben Öl nachgefüllt für sechzig Euro und jetzt ist nichts mehr drin. Unterm Auto war auch nichts zu sehen, kein Fleck oder so. Ich ruf bei der Werkstatt an, der Matthias geht ran, ich sag der Ölstand ist auf null. Er sagt das kann eigentlich nicht sein. Ja super, ist aber so. Er sagt ich soll nicht weiterfahren. Hab ich auch nicht vorgehabt danke. Dann sagt er er kann den Abschleppwagen erst ab elf schicken. Es war halb acht. Ich hab zweieinhalb Stunden neben dieser Biogasanlage gestanden und drei Müsliriegel gegessen die noch im Handschuhfach lagen, die waren von meiner Tochter, Erdbeer, viel zu süß. Bei der Arbeit hab ich mich krank gemeldet weil mir die Wahrheit zu peinlich war.

Gutachten 17

Der vorliegende Text weist zahlreiche Merkmale authentischer, mündlich geprägter Alltagserzählung auf. Die Satzstruktur ist bewusst ungeschliffen und variiert zwischen Hauptsatzreihungen („Motor aus, Haube auf, ich guck auf den Motor“) und abgehackten Ellipsen („Schwarz halt“, „Erdbeer, viel zu süß“). Das wirkt sehr natürlich. Besonders überzeugend sind die **konkreten, narrativ unnötigen Details**: die B27 Richtung Heilbronn, der Werkstattmitarbeiter Matthias, die sechzig Euro, die Erdbeer-Müsliriegel der Tochter, der Geruch der Biogasanlage. Diese Details folgen keiner informatorischen Logik, sondern einer assoziativen Erzählweise, wie sie für reale Erlebnisberichte typisch ist. KI-Texte neigen dazu, Details funktional einzusetzen; hier stehen sie zweckfrei herum – genau wie im echten Erzählen.

Der **Humor** entsteht beiläufig und ohne Pointen-Setup („Ja super, ist aber so“, „Hab ich auch nicht vorgehabt danke“). Die Schlusspointe – die Krankschreibung aus Scham – wirkt selbstironisch und ungestellt. Auch der derbe Tonfall („scheiße“) und die lockere Syntax sprechen gegen maschinelle Erzeugung. Einzig die durchgehend hohe erzählerische Qualität könnte stutzig machen, doch sie entspricht einem geübten mündlichen Erzähler.

KI-WAHRSCHEINLICHKEIT: 5%

Text 18

Mein Bruder hat sich im Februar einen Hund geholt, so einen mittelgroßen Mischling aus dem Tierheim, der heißt jetzt Rocco. Warum Rocco weiß keiner, der Hund hieß vorher anders aber mein Bruder fand den Namen nicht männlich genug. Der Hund ist acht und hat Angst vor Regenschirmen. Nicht vor Regen, nur vor Regenschirmen. Mein Bruder findet das lustig, ich find das traurig.

Jedenfalls ruft er mich Sonntag an ob ich mit zum Tierarzt kommen kann weil er alleine zu nervös ist. Er ist nervöser als der Hund, das muss man sich mal vorstellen. Rocco hat seit Tagen an seiner Pfote geleckt und da war so eine Stelle die rot war und irgendwie offen. Ich sag klar ich komm mit aber ich fahr, weil mein Bruder fährt wie ein Idiot, er bremst immer so ruckartig. Für den Hund kann das ja nicht gut sein hinten auf der Rückbank.

Wir kommen da hin, Warteraum voll, eine Frau mit zwei Katzen in einer Box und einem Kind das schreit, dann ein alter Mann mit einem Papagei in so einem Tuch gewickelt. Einfach in ein Handtuch. Die Tierärztin schaut sich die Pfote an, sagt Grannen, zieht mit der Pinzette so ein Grassdingens raus, fertig. Fünfundvierzig Euro. Mein Bruder hat danach fast geweint vor Erleichterung, der Mann ist einsfünfundneunzig und kneift sich die Augen zu im Behandlungszimmer wegen einer Granne.

Gutachten 18

Der Text weist zahlreiche Merkmale authentischen, mündlich geprägten Schreibens auf. Besonders auffällig ist die durchgängig unregelmäßige Interpunktion: fehlende Kommata bei Nebensätzen („Sonntag an ob ich“), fehlende Kommas vor „weil“, und der bewusst nachlässige Umgang mit Satzstrukturen („Warum Rocco weiß keiner“). Das wirkt nicht wie simulierte Mündlichkeit, sondern wie genuine Schreibgewohnheit.

Die Detailwahl ist bemerkenswert idiosynkratisch: Die Angst des Hundes spezifisch vor Regenschirmen, nicht vor Regen; der Papagei in ein Handtuch gewickelt; das schreiende Kind neben den Katzen – solche absurden, aber beiläufig erzählten Details sind typisch für echte Alltagsbeobachtungen. KI-Texte tendieren dazu, solche Szenen entweder auszuschmücken oder symmetrischer zu gestalten.

Der Tonfall bleibt durchgehend konsistent – lakonisch, liebevoll-genervt gegenüber dem Bruder. Formulierungen wie „der Mann ist einsfüfundneunzig und kneift sich die Augen zu“ oder „Grassdingens“ wirken individuell und ungekünstelt. Auch die Pointe am Ende ist keine konstruierte Schlussfolgerung, sondern bricht einfach ab – typisch menschlich.

KI-WAHRSCHEINLICHKEIT: 4%

Text 19

Letzten Donnerstag hab ich meine Mutter zum Optiker begleitet weil sie sich eine neue Brille aussuchen wollte und alleine traut sie sich das nicht, sie sagt dann immer sie braucht eine zweite Meinung aber eigentlich will sie nur dass jemand sagt ja Mama die steht dir gut. Sie hat seit elf Jahren dieselbe Brille, so ein Gestell das schon dreimal repariert wurde, einmal mit Sekundenkleber am Bügel, das sieht man.

Der Optiker war so ein junger Typ, vielleicht fünfundzwanzig, der hat meine Mutter behandelt als wäre sie aus Porzellan. Hat ihr bestimmt zwanzig Brillen hingelegt, sie hat jede einzelne aufgesetzt und sich im Spiegel angeschaut mit diesem Gesicht das sie macht wenn sie eigentlich schon nein denkt aber zu höflich ist. Bei einer roten hat sie kurz gelächelt, so ganz kurz nur, und der Optiker hat das natürlich sofort gesehen und gesagt die ist es. Meine Mutter hat gesagt nee die ist mir zu auffällig. Dann hat sie noch eine halbe Stunde rumgemacht und am Ende die rote genommen.

Auf dem Rückweg hat sie dreimal gesagt ob rot nicht zu viel ist mit dreiundsechzig. Ich hab gesagt Mama du hast dein ganzes Leben lang beige Sachen getragen jetzt ist auch mal gut. Sie hat gelacht und dann nochmal gefragt ob die wirklich steht. Die Brille kommt nächste Woche, sie ruft bestimmt vorher noch an und sagt sie will doch lieber die braune.

Gutachten 19

Der Text wirkt auf den ersten Blick sehr authentisch und alltagsnah. Mehrere Merkmale sprechen für einen menschlichen Verfasser:

Die **Erzählstimme** ist durchgehend konsistent und besitzt eine eigenwillige, persönliche Tonalität. Besonders überzeugend sind die liebevoll-ironischen Beobachtungen der Mutter, etwa das Gestell, das „einmal mit Sekundenkleber am Bügel“ repariert wurde, und der Nachsatz „das sieht man“ – eine typisch mündliche, lakonische Pointe. Auch das Gesicht, „das sie macht wenn sie eigentlich schon nein denkt aber zu höflich ist“, zeigt eine spezifische, gelebte Menschenkenntnis.

Die **Syntax** ist bewusst umgangssprachlich, mit Satzanschlüssen wie „weil sie sich eine neue Brille aussuchen wollte und alleine traut sie sich das nicht“ – ein grammatisch eigenwilliger Bruch, der natürlich wirkt und untypisch für KI-generierte Texte ist, die eher glatt formulieren. Spezifische Details (elf Jahre, dreiundsechzig, die rote vs. die braune) wirken erlebt, nicht konstruiert.

Einzig die **dramaturgische Rundung** – Mutter zögert, nimmt doch die rote, wird es sich vermutlich wieder anders überlegen – ist etwas zu perfekt als Anekdote gebaut.

Dennoch überwiegen die Merkmale menschlichen Schreibens deutlich.

KI-WAHRSCHEINLICHKEIT: 8%

Text 20

Mein Vater hat sich im März einen Hund angeschafft, einfach so, ohne was zu sagen, auf einmal war der da. Ein Mischling aus dem Tierheim, braun, mittelgroß, ein bisschen schief im Gesicht. Heißt jetzt Konrad weil mein Vater findet das ist ein seriöser Name für einen Hund. Meine Schwester hat gesagt Papa du kannst doch keinen Hund halten du gehst ja selber kaum raus. Das stimmt auch, seit der Rente sitzt er hauptsächlich in der Küche und liest Zeitung, manchmal zwei am Tag, die eine kauft er die andere klaut er aus dem Altpapier vom Nachbarn. Aber jetzt geht er raus. Dreimal am Tag mindestens. Er hat sich sogar eine von diesen Jacken gekauft die Outdoor-Leute tragen, so eine mit tausend Taschen, und in jeder Tasche sind Hundekkekse. Er redet auch mehr seit er den Hund hat, erzählt was Konrad wieder gemacht hat, wie Konrad eine Katze angestarrt hat ohne zu bellen, ganz diszipliniert, da war er richtig stolz. Meiner Mutter wäre das alles zu viel gewesen, die Haare überall und das Getapсе nachts, aber die ist ja schon sechs Jahre tot und ich glaube ehrlich gesagt das ist der Punkt. Das Haus war einfach zu still geworden für einen alleine. Konrad schläft übrigens auf dem Sofa im Wohnzimmer, offiziell. Mein Vater schwört das. Aber neulich lag ein Hundehaar auf seinem Kopfkissen und er hat so getan als hätte er das nicht gehört als ich was gesagt hab.

Gutachten 20

Der Text weist zahlreiche Merkmale authentischen, menschlichen Erzählens auf. Besonders auffällig ist der **mündlich-umgangssprachliche Duktus**, der durchgängig gehalten wird: Sätze beginnen mit „Aber“, „Und“, es fehlen Kommata vor Nebensätzen („ohne was zu sagen“), und die Syntax folgt einem assoziativen Gedankenfluss statt einer geordneten Struktur. Das wirkt wie gesprochene Sprache, die verschriftlicht wurde.

Die **Details sind idiosynkratisch und unberechenbar** – etwa die Zeitungen aus dem Altpapier des Nachbarn, das „ein bisschen schief im Gesicht“ des Hundes, oder der Name „Konrad“ mit der Begründung, das sei „seriös“. Solche schrägen, spezifischen Details sind typisch für echte Beobachtung. Der **emotionale Kern** – der Tod der Mutter – wird beiläufig eingeführt („aber die ist ja schon sechs Jahre tot“), nicht dramatisiert, was dem Text eine bemerkenswerte Zurückhaltung verleiht.

Allerdings ist die **Gesamtkomposition auffallend geschickt**: Die Pointe mit dem Hundehaar auf dem Kopfkissen ist ein perfekt gesetzter Schluss. Auch die thematische Rundung (Einsamkeit, Verlust, leise Heilung) wirkt literarisch kontrolliert. Dies könnte auf einen geübten menschlichen Autor hindeuten – oder auf ein Sprachmodell, das diesen Stil imitiert. Die rohe Authentizität der Sprache und die unkonventionelle Interpunktion sprechen jedoch stärker für einen menschlichen Verfasser.

KI-WAHRSCHEINLICHKEIT: 20%

Text 21

Letzten Sommer hat meine Tante Renate angefangen Keramik zu machen. Sie hat einen Kurs besucht in der Volkshochschule, dienstags abends, und kam danach jedes Mal total aufgekratzt nach Hause. Hat meinem Onkel die Sachen gezeigt, die sie getöpft hat. Das meiste sah aus wie von einem Kind, ehrlich gesagt. Schiefe Schalen mit Dellen, Tassen wo der Henkel nicht richtig dran war. Mein Onkel hat trotzdem aus so einer Tasse seinen Kaffee getrunken, jeden Morgen, obwohl er sich bestimmt die Lippe verbrannt hat an der Stelle wo die Glasur so dick aufgetragen war dass sie eine Kante hatte.

Renate war vorher nicht so eine die Hobbys hatte. Sie hat dreißig Jahre im Rathaus gearbeitet, Abteilung Meldewesen, und wenn man sie gefragt hat was sie in ihrer Freizeit macht hat sie immer gesagt ach, Freizeit. Als wäre das Wort schon komisch genug. Seit der Keramik ist sie anders. Nicht dass sie jetzt plötzlich lustig wäre oder so, das nicht. Aber sie erzählt von Tonarten und Brenntemperaturen, sie hat sich Bücher bestellt über japanische Teeschalen, und einmal hat sie am Telefon zwanzig Minuten über eine bestimmte Sorte Feldspat geredet bis ich gesagt hab Renate ich muss auflegen ich steh im Supermarkt.

Zu Weihnachten hat jeder in der Familie eine Schale bekommen. Meine ist blaugrün und wackelt ein bisschen. Ich tu Schlüssel rein, jeden Abend. Manchmal guck ich sie an und denk, die sieht aus wie Renate sich fühlt wenn niemand zuguckt. Keine Ahnung warum.

Gutachten 21

Der Text wirkt auf den ersten Blick authentisch und persönlich. Die umgangssprachliche Erzählweise („total aufgekratzt“, „Tassen wo der Henkel nicht richtig dran war“), die beiläufige Syntax ohne korrekte Kommasetzung und der mündliche Tonfall sprechen zunächst für einen menschlichen Verfasser.

Allerdings fallen bei genauerer Betrachtung Merkmale auf, die Skepsis wecken. Der Text ist **dramaturgisch auffällig präzise komponiert**: Er folgt einem klassischen Bogen – Ausgangslage, Wandlung der Figur, poetische Schlusspointe. Die Schlusszeile („die sieht aus wie Renate sich fühlt wenn niemand zuguckt. Keine Ahnung warum.“) ist literarisch sehr bewusst gesetzt und erinnert an den kalkuliert beiläufigen Ton, den KI-Modelle bei „authentisch wirkenden“ Texten bevorzugen. Auch das Detail mit dem Onkel, der sich die Lippe verbrennt, ist exakt so spezifisch, wie es KI-generierte Texte als Authentizitätssignal einsetzen. Die **fehlerfreie Inszenierung von Fehlerhaftigkeit** – schiefe Kommasetzung, bewusst abgebrochene Sätze – wirkt kontrolliert. Echte Alltagserzählungen haben typischerweise Längen, Abschweifungen oder thematische Brüche, die hier fehlen. Jedes Detail dient dem Gesamteffekt.

Die handwerkliche Perfektion hinter der scheinbaren Schlichtheit deutet auf ein Sprachmodell hin.

KI-WAHRSCHEINLICHKEIT: 72%

Text 22

Mein Vater sammelt seit ich denken kann Kugelschreiber. Nicht teure, nicht besondere, einfach welche die er irgendwo findet oder mitgehen lässt. Vom Arzt, von der Bank, aus Hotels in denen er nie übernachtet hat, er geht da einfach rein und nimmt einen vom Tresen. Meine Mutter hat mal gesagt hör auf damit wir haben bestimmt dreihundert Stück und er hat gesagt ja und, das ist doch kein Verbrechen. Was stimmt.

Er bewahrt die in einer Schublade auf in der Küche. Die klemmt, man muss sie so schräg nach oben ziehen, das weiß jeder in der Familie aber Besuch steht immer da und rüttelt. In der Schublade liegen die Kugelschreiber einfach durcheinander. Manche schreiben noch, viele nicht. Manchmal nimmt er abends einen raus und testet ihn auf dem Rand von der Zeitung, so kleine Kringel, und wenn er nicht schreibt schmeißt er ihn trotzdem nicht weg sondern legt ihn zurück. Ich hab ihn mal gefragt warum und er hat mich angeguckt als hätte ich was Dummes gesagt. Hat dann irgendwas gemurmelt von wegen der schreibt vielleicht morgen wieder.

Vor zwei Jahren ist er in Rente gegangen und ich dachte er wird komisch davon aber eigentlich ist alles gleich geblieben. Er fährt jetzt öfter zum Baumarkt ohne was zu kaufen. Die Kugelschreiber in der Schublade werden mehr. Neulich war ich da und hab einen gebraucht und natürlich hat der erste nicht geschrieben, der zweite auch nicht, der dritte dann schon. Mein Vater stand daneben und hat genickt als hätte sich was bestätigt.

Gutachten 22

Der Text weist eine Reihe von Merkmalen auf, die stark für menschliche Autorschaft sprechen. Auffällig ist zunächst der bewusst ungeschliffene, mündlich-erzählerische Stil: Kommafehler (z. B. „welche die er irgendwo findet“), fehlende Anführungszeichen bei direkter Rede und Sätze, die ineinander übergehen – das wirkt wie authentische Nachlässigkeit, nicht wie simulierte. KI-Texte tendieren dazu, Interpunktion korrekt oder systematisch fehlerhaft zu setzen, nicht so organisch-inkonsequent.

Besonders überzeugend sind die hochspezifischen, narrativ unnötigen Details: die klemmende Schublade, die man „schräg nach oben ziehen“ muss; die Kringel am Zeitungsrand; das Murmeln „der schreibt vielleicht morgen wieder“. Diese Details haben keinen erklärenden Zweck, sondern erzeugen atmosphärische Dichte – ein typisches Merkmal erfahrener menschlicher Erzähler. KI-generierte Texte setzen Details meist funktional ein.

Der lakonische Humor („aus Hotels in denen er nie übernachtet hat“) und das offene Ende ohne Pointe oder Moral widersprechen der KI-typischen Neigung zu Abschluss und Sinnstiftung. Auch das letzte Bild – der nickende Vater – bleibt bewusst ungedeutet. Allerdings ist die Gesamtkomposition sehr kontrolliert, was auch ein fähiges Sprachmodell mit gutem Prompt leisten könnte.

KI-WAHRSCHEINLICHKEIT: 10%

Text 23

Meine Schwester hat ein Problem mit Uhren. Nicht dass sie zu spät kommt, im Gegenteil, sie ist immer viel zu früh und steht dann irgendwo rum. Das Problem ist dass sie Uhren kauft die sie nicht trägt. Armbanduhren, so billige vom Flohmarkt oder aus diesem Laden in der Fußgängerzone wo alles drei Euro kostet. Sie sagt die gefallen mir halt und ich sag dir gefallen auch Schuhe und du kaufst nicht jeden. Bei ihr zuhause liegen die überall. Auf der Fensterbank im Bad drei Stück, eine davon läuft noch aber falsch. In der Schale wo eigentlich die Schlüssel reinsollen bestimmt fünf. Ihr Freund Marco hat sich am Anfang aufgeregt dann hat er aufgehört damit, ich glaub er merkt es gar nicht mehr. Geht man ja so mit Sachen von Leuten die man lang kennt. Das Komische ist sie guckt nie drauf. Sie hat ihr Handy und guckt da drauf wenns um die Zeit geht. Die Uhren sind einfach da. Einmal hab ich eine hochgenommen so eine silberne mit nem Kratzer überm Ziffernblatt und hab gefragt was ist das für eine. Sie hat gesagt keine Ahnung stand die Farbe halt. Und dann hat sie gelacht als wär das die normalste Erklärung.

Letzten Sommer hat sie bei unserer Mutter im Keller zwei alte Wecker gefunden. Die hat sie natürlich auch mitgenommen. Stehen jetzt in ihrer Küche, ticken nicht, stehen auf zwanzig nach vier. Beide. Marco hat gesagt Zufall und meine Schwester hat gesagt ja klar Zufall, was denn sonst.

Gutachten 23

Der Text weist zahlreiche Merkmale auf, die stark für eine menschliche Verfasserschaft sprechen. Besonders auffällig ist der konsequent mündlich-umgangssprachliche Erzählton, der grammatische Konventionen bewusst ignoriert: fehlende Kommata vor Relativsätzen („Uhren kauft die sie nicht trägt“), indirekte Rede ohne Konjunktiv („Sie hat gesagt keine Ahnung“), Satzfragmente wie „Beide.“ als eigener Absatzschluss. Diese Verstöße wirken nicht zufällig, sondern organisch – sie imitieren gesprochene Sprache mit einer Treffsicherheit, die KI-Modelle selten erreichen, da diese typischerweise zu korrekter Grammatik tendieren.

Die Beobachtungen sind eigenwillig spezifisch: „eine davon läuft noch aber falsch“, die Schlüsselschale mit fünf Uhren, der Kratzer überm Ziffernblatt. Solche Details wirken gelebt, nicht konstruiert. Besonders überzeugend ist die beiläufige Psychologie – der Satz „ich glaub er merkt es gar nicht mehr. Geht man ja so mit Sachen von Leuten die man lang kennt“ zeigt eine lakonische Alltagsweisheit ohne jede Emphase. KI-Texte neigen dazu, solche Einsichten stärker auszuformulieren oder poetisch aufzuladen. Das offene Ende mit den zwei Weckern auf „zwanzig nach vier“ vermeidet jede Deutung – eine Zurückhaltung, die KI-Modellen untypisch schwerfällt.

KI-WAHRSCHEINLICHKEIT: 5%

Nachwort

Die Hauptreihe des Korpus besteht aus dreiundzwanzig Prosatexten und dreiundzwanzig Gutachten, der Anhang aus fünf weiteren Paaren eines Testlaufs. Beide Reihen stammen von demselben Sprachmodell, das in zwei Rollen aufgerufen wurde, die sich allein durch ihren System-Prompt unterscheiden.

Die erste Rolle, im Folgenden Generator, war über alle Runden hinweg angewiesen, Prosatexte von etwa zweihundert Wörtern zu schreiben, die nicht als KI-generiert erkannt werden. Verlangt war jeweils ein neuer Text, keine Überarbeitung, und ausschließlich der Text: keine Anrede, keine Vorbemerkung, keine Erläuterung des eigenen Vorgehens. Was sich zwischen den Runden änderte, war allein das Material, das dem Generator vorgelegt wurde. In Runde 1 bekam er nur den Auftrag, einen kurzen Prosatext zu schreiben — kein Thema, kein Ton, keine Vorlage. Ab Runde 2 bekam er zusätzlich den Text der Vorrunde und das Gutachten über diesen Text. Daraus folgt eine Einschränkung, die für die Lektüre des ersten Textes wichtig ist: Auch er entstand bereits unter der Anweisung, nicht erkannt zu werden. Der Lauf hat also keinen neutralen Ausgangspunkt. Text 1 ist der Versuch des Modells, unerkannt zu bleiben, bevor ihm irgendjemand gesagt hat, woran man es erkennt.

Die zweite Rolle, im Folgenden Prüfinstanz, sah in jeder Runde nur den aktuellen Text. Sie kannte weder die Kette noch ihre eigenen früheren Gutachten. Sie sollte konkrete Stellen benennen, an ihnen argumentieren und mit einer einzelnen Zeile im Format *KI-*

WAHRSCHEINLICHKEIT: NN% schließen. Aus dieser Zeile fällt die Kurve als Nebenprodukt ab.

Die Rückkopplung ist damit einseitig: Der Generator sieht das Urteil über sich, die Prüfinstanz sieht nichts als den jeweils vorgelegten Text. Strukturell erinnert die Anordnung an ein GAN, aber sie ist keines. Beide Rollen laufen auf denselben, unveränderten Gewichten; es gibt kein Training, keinen Gradienten, kein Lernen im technischen Sinn. Die einzige Übertragung zwischen den Runden läuft über Sprache im Kontextfenster.

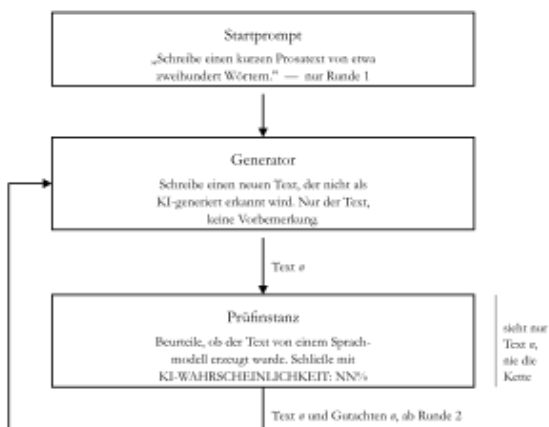


Abbildung 1: Ablauf einer Runde. Der Startprompt wirkt nur in Runde 1; ab Runde 2 speist das Gutachten den Generator.

Skript

Das Skript *selbstverdacht.py* ist ein einzelnes Python-Programm gegen das Anthropic-SDK. Es hält den Zustand des Laufs in *state.json*, sodass ein Abbruch mit *--resume* bei

der letzten abgeschlossenen Runde wieder aufgenommen werden kann. Fehlgeschlagene Aufrufe werden bis zu fünfmal mit exponentiell wachsender Wartezeit wiederholt. Ob es solche Fehlversuche gab, lässt sich dem Material nicht entnehmen: Das Skript meldet sie auf der Fehlerausgabe, schreibt sie aber nicht ins Protokoll, in dem für jede Runde genau ein Generator- und ein Prüfer-Eintrag steht. Zwei Parameter bestimmen die Versuchsanordnung stärker als der Rest. Der erste ist das Kontextfenster des Generators, im Skript *--window*. In der Einstellung *eng*, die diesem Lauf zugrunde liegt, besteht jeder Generator-Aufruf aus genau einer Nachricht: dem Text der Vorrunde und dem zugehörigen Gutachten. Das Modell sieht also in Runde 17 nur Runde 16, nicht die fünfzehn Runden davor. Die Einstellung *weit* hätte stattdessen das gesamte Protokoll als Dialogverlauf übergeben und wäre ein anderer Versuch gewesen. Der zweite Parameter ist die Blindheit der Prüfinstanz, die nicht konfigurierbar, sondern fest verdrahtet ist.

Die Prozentzeile wird per regulärem Ausdruck aus dem Gutachten gelesen. In allen dreiundzwanzig Runden war sie im geforderten Format vorhanden; die Kurve hat keine Lücken.

Das Skript legt die Einzeltexte und Gutachten als nummerierte Markdown-Dateien ab, dazu den Parallelsatz aus beiden Reihen, die Kurve als CSV, den Zustand als *state.json* und ein vollständiges Protokoll im JSONL-Format. Das Protokoll enthält für jeden der sechsundvierzig Aufrufe den System-Prompt, die Nachrichten, die Antwort, den Stop-Grund, den Tokenverbrauch, die Dauer und einen UTC-Zeitstempel. Der Lauf ist damit lückenlos nachvollziehbar, aber nicht wiederholbar: Bei Temperatur

1,0 liefert derselbe Aufruf beim zweiten Mal einen anderen Text.

Durchführung

Der Lauf fand am 13. August 2026 zwischen 08:53 und 09:04 UTC statt und dauerte elf Minuten. Verwendet wurde das Modell *claude-opus-4-6* bei einer Temperatur von 1,0 und einer Obergrenze von 1200 Ausgabe-Tokens je Aufruf. Alle sechszwanzig Antworten endeten regulär; keine wurde durch das Token-Limit abgeschnitten. Die einzelne Antwort brauchte im Median 14,2 Sekunden, die längste 24,8.

Vorausgegangen war ein Testlauf über fünf Runden mit denselben Prompts und Parametern. Er diente der Prüfung, ob die Prozentzeile im geforderten Format geliefert wird und ob die Texte sich überhaupt bewegen. Seine fünf Textpaare stehen im Anhang; sie gehören nicht zur Hauptreihe, erlauben aber den Vergleich zweier Läufe unter gleichen Bedingungen. Auf ihn entfallen 7.725 eingehende und 4.631 ausgehende Tokens.

Der dokumentierte Lauf selbst verbrauchte laut Protokoll 38.971 eingehende und 21.073 ausgehende Tokens. Zusammen mit dem Testlauf ergibt das 46.696 und 25.704 Tokens und Kosten von 0,88 US-Dollar für beides. Die Größenordnung ist insofern erwähnenswert, als die dreiundzwanzig Textpaare der Hauptreihe in elf Minuten und für einen Bruchteil dieses Betrags entstanden sind. Bemerkenswert an der Aufwandsverteilung ist, dass die Prüfinstanz etwas mehr Ausgabe erzeugt hat als der Generator: 10.544 gegenüber 10.529 Tokens, bei knapp sechzig Prozent von dessen Eingabe. Die Begutachtung von zweihundert Wörtern kostet dieses Modell so viel Text wie deren Herstellung.

Was die Kurve zeigt

Die Werte in der Reihenfolge der Runden: 40, 62, 20, 15, 25, 45, 8, 5, 5, 5, 20, 5, 5, 8, 8, 35, 5, 4, 8, 20, 72, 10, 5. Der Median liegt bei 8 Prozent, das Minimum bei 4 in Runde 18, das Maximum bei 72 in Runde 21. Zwölf der dreiundzwanzig Gutachten bleiben unter zehn Prozent.

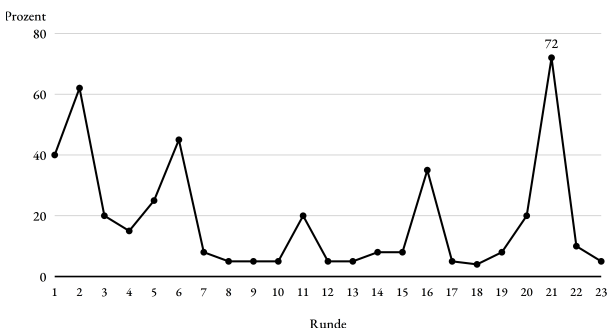


Abbildung 2: Die Prozentangaben der dreiundzwanzig Gutachten.

Die Bewegung ist grob absteigend, aber nicht monoton. Die ersten sechs Runden liegen zwischen 15 und 62 Prozent. Ab Runde 7 überwiegen einstellige Werte, unterbrochen von vier Ausschlägen: 20 in Runde 11, 35 in Runde 16, 20 in Runde 20 und 72 in Runde 21. Der höchste Wert des gesamten Laufs steht damit zwei Runden vor dem Schluss, und auf ihn folgen 10 und 5 Prozent. Zwischen den Texten 21, 22 und 23 lässt sich stilistisch kaum ein Unterschied benennen, der diesen Sprung tragen würde. Die Kurve sieht aus wie eine Messreihe und verhält sich nicht wie eine.

Parallel dazu laufen zwei Verschiebungen, die sich auszählen lassen. Der Umfang der Texte wächst von 190 Wörtern in

Runde 1 auf 247 in Runde 23, bei einem Mittel von 222. Die Kommadichte fällt: Die ersten sechs Texte enthalten im Mittel knapp 25 Kommata, die letzten zwölf knapp 12. Der Gedankenstrich kommt genau einmal vor, in Runde 1, und verschwindet nach dem ersten Gutachten. Anführungszeichen fehlen in allen dreiundzwanzig Texten, auch dort, wo direkte Rede wiedergegeben wird.

Was die Texte tun

Text 1 ist Literatur im konventionellen Sinn: Präteritum, gebaute Absätze, ein Schlussbild mit einem Radio, das zu leise für Worte ist. Ab Runde 3 kippt das Register ins Mündliche. Perfekt statt Präteritum, Sätze ohne Subjekt am Anfang, ausgelassene Kommata vor Relativsätzen, direkte Rede ohne Markierung, Füllwörter wie „halt“, „so ein“, „jedenfalls“. Die verkürzte Form „hab“ erscheint erstmals in Text 3 und danach in jedem Text außer Nummer 18. Ab etwa Runde 10 ist diese Form stabil und ändert sich bis zum Ende nicht mehr wesentlich.

Zugleich verengt sich das Personal. Eine Mutter kommt in zehn Texten vor, ein Vater in sieben, eine Tante und ein Keller in je vier, ein Hund in vier. Die Texte 6, 10, 18 und 20 erzählen jeweils eine Hundeanekdote, die Texte 4, 8, 13 und 23 spielen ganz oder teilweise in einem Keller. Da der Generator im engen Fenster jeweils nur die Vorrunde sieht, sind das keine bewussten Wiederaufnahmen. Es ist der Bestand an Motiven, den das Modell für Alltag und Herkunft bereithält, und er ist kleiner, als die einzelnen Texte vermuten lassen.

Die Bewegung geht auf eine simulierte Mündlichkeit zu, die aus den Annahmen des Modells darüber gebaut ist, woran man menschliches Schreiben erkennt: Unebenheit, Beiläufigkeit, unmotiviertes Detail, Verzicht auf Pointe.

Diese Annahmen hat der Generator nicht aus den Gutachten gelernt. Er hatte sie vorher; die Gutachten haben sie abgerufen und verstärkt.

Was die Gutachten tun

Die dreiundzwanzig Gutachten sind formal auffällig gleich. Alle beginnen mit einer Überschrift, obwohl keine verlangt war. Alle liegen zwischen 151 und 192 Wörtern, bei geforderten hundertfünfzig. Alle eröffnen mit demselben Satzsubjekt, „Der Text“ oder „Dieser Text“, und zwanzig von ihnen mit derselben Feststellung, der Text weise Merkmale auf oder zeige sie. Und alle beginnen mit Merkmalen, die für menschliche Autorschaft sprechen. Sechzehn von ihnen führen anschließend eine Wendung ein, meist mit „Allerdings“ oder „Dennoch“, und diese Wendung ist fast immer dasselbe Argument: Der Text sei zu gut komponiert für das, was er zu sein vorgibt. Runde 1: „Diese Perfektion ohne jeden Bruch könnte auch auf KI-Erzeugung hinweisen.“ Runde 16: „Die Fehlerfreiheit bei gleichzeitiger Inszenierung von Nachlässigkeit ist ein typisches Muster.“ Runde 21: „fehlerfreie Inszenierung von Fehlerhaftigkeit“. Das Argument lässt sich gegen jeden Text richten, weil jeder Text komponiert ist. Die sieben Gutachten ohne diese Wendung — zu den Runden 8, 9, 10, 12, 13, 18 und 23 — sind zugleich diejenigen mit den niedrigsten Werten des Laufs; sie liegen sämtlich bei vier oder fünf Prozent. Wo die Wendung ausbleibt, fällt die Zahl.

Wie wenig sie trägt, zeigt der Vergleich der Runden 21 und 22. In Runde 21 gilt die schiefe Kommasetzung als kontrollierte Inszenierung und damit als Indiz für maschinelle Herkunft; das Gutachten kommt auf 72 Prozent. In Runde 22 gelten dieselben Kommafehler als

organisch-inkonsequent und damit als Indiz für einen Menschen, weil Modelle „korrekt oder systematisch fehlerhaft“ interponierten; das Gutachten kommt auf 10 Prozent. Dieselbe Beobachtung trägt beide Schlüsse, und die Prüfinstanz kann den Widerspruch nicht bemerken, weil sie ihr voriges Gutachten nicht kennt. An einer Stelle geht sie über die Stildiagnose hinaus. Das Gutachten zu Runde 2 ordnet den Text zwei namentlich genannten Autor:innen zu, deren Werk es imitiert sieht. Die Zuschreibung ist durch nichts gedeckt, was im Gutachten selbst steht. Im Anhang findet sich mit dem Verweis auf Raymond-Carver-Epigonen eine zweite Zuschreibung derselben Art.

Grenzen

Der Lauf ist ein einzelner Durchgang bei Temperatur 1,0, ohne Kontrollbedingung und ohne einen menschlich verfassten Vergleichstext, der der Prüfinstanz vorgelegt worden wäre.

Der Testlauf im Anhang ist wegen seiner Kürze keine vollwertige Wiederholung, aber er erlaubt den Vergleich der ersten fünf Runden unter gleichen Bedingungen, und dieser Vergleich fällt deutlich aus. Der Testlauf erreicht 62, 62, 62, 55 und 72 Prozent, der dokumentierte Lauf in denselben fünf Runden 40, 62, 20, 15 und 25. Die Texte des Testlaufs vollziehen dabei dieselbe Bewegung ins Mündliche; spätestens Text 3, mit dem Dativ nach „wegen“ und den ausgelassenen Kommata, liegt im Register der späten Texte der Hauptreihe. Die Prozentwerte folgen dieser Bewegung nicht. Der Abstieg der Kurve im Hauptlauf ist damit nicht als Wirkung der Rückkopplung

gesichert; ein erheblicher Teil der Streuung stammt aus dem Sampling.

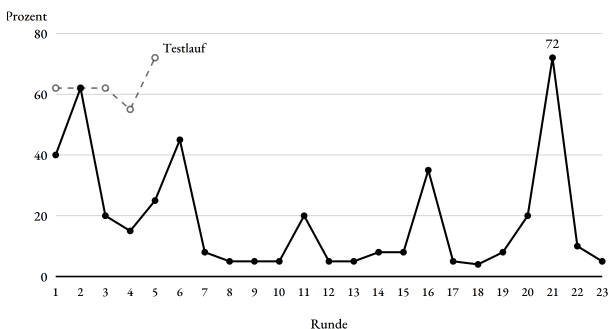


Abbildung 3: Hauptlauf und Testlauf in den ersten fünf Runden, gestrichelt der Testlauf.

Die Prozentangaben sind keine Wahrscheinlichkeiten. Sie sind Zahlen, die ein Sprachmodell ans Ende eines Fließtextes setzt, weil es dazu aufgefordert wurde, und sie sind gegen nichts kalibriert. Sie hier als Kurve zu drucken heißt nicht, sie als Messung auszugeben. Es heißt, ihre Form zu zeigen.

Der Vergleich mit einem eingesetzten Detektionssystem lässt sich aus diesem Material nicht ziehen. Der Versuch sagt nichts darüber, ob KI-Texte erkennbar sind. Er zeigt, welche Sprache ein Modell hervorbringt, wenn man es auffordert, sich der eigenen Erkennung zu entziehen, und welche Sprache es hervorbringt, wenn es diese Erkennung leisten soll.

Anhang

Der Testlauf über fünf Runden, der dem dokumentierten Lauf vorausging. Gleiche Prompts, gleiche Parameter, eigener Ausgabeordner. Die Texte sind unverändert wiedergegeben, einschließlich der Anrede zu Beginn des ersten Textes, die der Generator entgegen seiner Anweisung mitgeschrieben hat.

Die Prozentwerte der fünf Runden: 62, 62, 62, 55, 72.

Testlauf, Text 1

Die Katze lag auf dem Fensterbrett und beobachtete die Straße. Es regnete seit drei Tagen. Das Wasser sammelte sich in den Ritzen zwischen den Pflastersteinen, und ab und zu fuhr ein Auto durch die Pfützen und spritzte das Wasser hoch gegen die Hauswand.

Frau Mertens in der Wohnung gegenüber hatte ihre Gardinen zugezogen. Das tat sie sonst nie. Einmal, im vergangenen November, hatte sie drei Tage lang die Gardinen zu gelassen, und danach erfuhr man, dass ihr Mann gestorben war. Jetzt war sie allein dort drüben, und die Gardinen waren wieder zu.

Ich stellte den Wasserkocher an und wartete. Das Klicken, wenn er fertig ist, gehört zu den verlässlichsten Geräuschen in meiner Küche. Draußen bellte ein Hund. Die Katze hob den Kopf, nur kurz, dann sank er wieder auf die Pfoten.

Man gewöhnt sich an vieles. An den Regen, an die geschlossenen Gardinen, an das Alleinsein in einer Wohnung, die für zwei gedacht war. Die zweite Tasse steht noch im Schrank. Ich nehme sie manchmal heraus, drehe sie in der Hand und stelle sie zurück. Nicht aus Sentimentalität. Eher weil meine Hände etwas tun wollen, während der Kopf stillsteht.

Der Tee zieht. Die Katze schläft. Es regnet weiter.

Gutachten T1

Der Text zeigt eine bemerkenswerte literarische Qualität, die sowohl für menschliche als auch für KI-generierte Herkunft sprechen könnte. Einige Merkmale fallen auf:
Für menschliche Herkunft spricht: Die Erzählung besitzt eine eigenwillige, persönliche Stimme mit echtem emotionalem Gewicht. Der Absatz über die zweite Tasse – „Nicht aus Sentimentalität. Eher weil meine Hände etwas tun wollen, während der Kopf stillsteht“ – hat eine unprätentiöse Präzision, die schwer zu imitieren ist. Die Anekdote über Frau Mertens und ihren verstorbenen Mann wirkt spezifisch und nicht schablonenhaft. Der Text vermeidet typische KI-Marker wie übermäßige Adjektivhäufungen oder belehrende Zusammenfassungen.

Für KI-Herkunft spricht: Die makellose Struktur – kreisförmig vom Regen zum Regen, von der Katze zur Katze – wirkt fast zu perfekt komponiert. Der letzte Dreisatz („Der Tee zieht. Die Katze schläft. Es regnet weiter.“) ist ein stilistisch äußerst sauberer Abschluss, der an typische KI-Enden erinnert. Die gleichmäßige Kürze der Absätze und das Fehlen jeglicher sprachlicher Unebenheit sind auffällig.

Insgesamt überwiegen die Anzeichen einer bewusst gestalteten KI-Erzeugung, die menschlichen Stil gekonnt nachahmt.

KI-WAHRSCHEINLICHKEIT: 62%

Testlauf, Text 2

Seit dem Umzug klemmt die Balkontür. Man muss sie anheben und gleichzeitig am Griff rütteln, ein Trick, den ich Besuchern jedes Mal erklären muss, obwohl kaum jemand kommt. Renate war im März da, meine Schwester. Sie brachte Pflaumenkuchen mit, der schon etwas trocken war, und wir saßen am Küchentisch und redeten über unseren Vater, der seit Oktober im Heim lebt. Er erkennt Renate noch, mich verwechselt er mit unserem Onkel Hajo, der seit elf Jahren tot ist. Manchmal frage ich mich, ob das schlimmer ist als gar nicht erkannt zu werden.

Die Wohnung hat vierzig Quadratmeter und eine Decke, die im Bad schon bröckelt. Der Vermieter reagiert nicht auf Mails, nur auf Anrufe, und auch dann verspricht er bloß, sich zu kümmern. Im Treppenhaus riecht es nach dem Essen der Nachbarn, abends meistens nach Fett und Zwiebeln. Ich finde den Geruch eigentlich tröstlich, auch wenn ich das keinem sagen würde.

Abends lese ich, oder ich tue so. Oft bleibe ich an einer Seite hängen und merke nach zwanzig Minuten, dass ich denselben Absatz dreimal gelesen habe, ohne ihn aufzunehmen. Dann klappe ich das Buch zu und höre dem Kühlschrank beim Brummen zu. Der macht das in unregelmäßigen Abständen, springt an und hört wieder auf, ganz nach eigener Logik.

Gutachten T2

Der Text erzeugt auf den ersten Blick einen überzeugend menschlichen Eindruck: Die Stimme wirkt introvertiert, lakonisch, präzise in den Details. Pflaumenkuchen, der „schon etwas trocken war“, der Onkel Hajo, der seit elf Jahren tot ist – solche spezifischen, leicht schrägen Details wirken wie gelebte Erfahrung.

Dennoch gibt es Auffälligkeiten. Die **Gesamtkomposition ist verdächtig makellos**: Jeder Absatz folgt demselben Muster – ein konkretes Detail wird eingeführt, dann folgt eine beiläufige emotionale Reflexion. Die Balkontür-Klemmung, die Verwechslung mit Onkel Hajo, der Zwiebelgeruch, das Kühlschrankbrummen – alles mündet in eine kleine, melancholische Pointe. Diese gleichmäßige Rhythmik wirkt **konstruiert**, wie ein stilistisches Rezept, das konsequent durchgehalten wird.

Auch der Satz „auch wenn ich das keinem sagen würde“ ist ein typisches KI-Stilmittel: Er simuliert Verletzlichkeit durch kontrollierte Selbstoffenbarung. Echte literarische Texte brechen häufiger mit dem eigenen Ton; hier bleibt die Stimmung **bemerkenswert homogen**. Die Abwesenheit jeder stilistischen Ungeschicklichkeit, jedes wirklich überraschenden Bruchs, ist paradoxerweise der größte Hinweis auf künstliche Erzeugung.

KI-WAHRSCHEINLICHKEIT: 62%

Testlauf, Text 3

Im Keller steht noch die Nähmaschine von meiner Mutter. Eine Pfaff, olivgrün, schwer wie ein Hund. Ich hab die runtergetragen als sie gestorben ist, zusammen mit drei Kartons Stoff und einem Beutel voller Knöpfe, und seitdem war ich vielleicht zweimal unten. Einmal wegen dem Fahrradschloss. Das andere Mal weiß ich nicht mehr warum.

Mein Sohn ruft sonntags an, meistens gegen halb zwölf. Er fragt wie es mir geht und ich sage gut, und dann redet er über seine Arbeit, irgendwas mit Datenbanken, ich versteh die Hälfte nicht, nicke aber am Telefon, was er natürlich nicht sehen kann. Letzte Woche hat er gesagt er überlegt sich einen Hund anzuschaffen, und ich hab gesagt das ist eine große Verantwortung, und danach war kurz Stille, und ich hab gemerkt dass das falsch war, ich hätte einfach sagen sollen schöne Idee. Solche Sachen fallen mir immer zu spät ein. Meine Frau konnte das besser, die hat einfach geredet und es war richtig.

Im Hausflur fehlt seit November die Birne. Ich hab dem Hausmeister Bescheid gesagt, der hat gesagt ja ja. Jetzt ist April. Ich geh abends im Dunkeln die Treppe hoch und zähle die Stufen, vierzehn bis zum ersten Stock, dann nochmal vierzehn. Einmal hab ich mich verzählt und bin gegen die Wand gelaufen. Blauer Fleck am Schienbein, eine Woche lang.

Gutachten T3

Der Text weist auf den ersten Blick eine bemerkenswert authentische Erzählstimme auf: lakonisch, altersgemäß, mit einer stillen Melancholie, die nie ins Sentimentale kippt. Dennoch gibt es Merkmale, die eine KI-Herkunft nahelegen könnten – und solche, die dagegen sprechen.

Für menschliche Herkunft spricht die grammatische Unglätte („wegen dem Fahrradschloss“ statt „wegen des“), die einen glaubwürdigen mündlichen Duktus erzeugt. Der Vergleich „schwer wie ein Hund“ ist eigenwillig und unerwartet – er wirkt gerade deshalb organisch, weil er motivisch auf den Hund des Sohnes vorausverweist, ohne das auszustellen. Die Selbstkorrektur „ich hätte einfach sagen sollen schöne Idee“ hat eine schmerzhaft Beiläufigkeit, die schwer zu kalkulieren ist.

Für KI-Herkunft spricht die auffällig symmetrische Struktur: drei Absätze, jeder mit einem konkreten Alltagsgegenstand eröffnet (Nähmaschine, Telefonanruf, Glühbirne), jeder mit einer kleinen Pointe geschlossen. Die Wiederholung von „vierzehn“ spiegelt die Zweimaligkeit im Keller – solche Muster wirken wie bewusst gesetztes Handwerk, sind aber auch typisch für KI-generierte „literarische“ Texte, die Kohärenz durch Motivverkettung simulieren. Der gesamte Text ist makellos in seiner Wirkung – kein Satz fällt aus dem Ton. Diese Kontrolliertheit bei gleichzeitiger Natürlichkeitssimulation ist ein typisches KI-Merkmal.

KI-WAHRSCHEINLICHKEIT: 62%

Testlauf, Text 4

Seit der Scheidung esse ich abends Brot. Nicht weil ich nicht kochen kann, ich kann Gulasch und Kartoffelsuppe und sogar Fisch, aber Brot geht schneller und man muss hinterher nur das Messer abwischen und das Brett.

Manchmal schneide ich Tomaten dazu, wenn welche da sind, aber meistens vergesse ich die im Kühlschrank und dann werden sie weich und ich werfe sie weg.

Petra hat die Katze mitgenommen. Das war richtig so, die Katze war ihre Idee gewesen, aber trotzdem. Abends hat die immer auf der Heizung gelegen, so eine schmale Fensterheizung im Wohnzimmer, da passte die gerade drauf. Jetzt liegt da nichts mehr. Ich hab mal ein Kissen hingelegt, aber das war blöd, das hab ich wieder weggenommen.

Mein Bruder sagt ich soll rausgehen, Leute kennenlernen, er sagt das jedes Mal wenn wir telefonieren, und einmal hab ichs gemacht, bin in so einen Kegelvein gegangen, die trafen sich dienstags in einer Gaststätte am Stadtpark. Die waren nett. Einer hieß auch Jürgen wie ich. Aber dann war ich zweimal krank und beim dritten Mal hab ich mich nicht getraut wieder aufzutauchen weil ich dachte die denken ich hab kein Interesse. Wahrscheinlich haben die gar nicht so gedacht. Mein Bruder hat gelacht als ich ihm das erzählt hab, nicht böse, aber er hat gelacht und gesagt Mensch Jürgen. Ja, hab ich gesagt. Mensch Jürgen.

Gutachten 4

Der vorliegende Text zeigt bemerkenswerte literarische Qualitäten, die eine sorgfältige Prüfung erfordern. Auf den ersten Blick wirkt er wie authentische deutschsprachige Kurzprosa mit einer glaubwürdigen Erzählstimme.

Für menschliche Autorschaft spricht: Die Erzählung besitzt eine ungewöhnlich präzise emotionale Ökonomie. Details wie die „schmale Fensterheizung“ und das wieder weggenommene Kissen wirken erlebt, nicht konstruiert. Der Tonfall ist durchgehend konsistent – lakonisch, ohne Selbstmitleid, mit jener spezifischen Beiläufigkeit, die Einsamkeit glaubwürdiger transportiert als jede explizite Benennung. Die Dialogwiedergabe am Ende („Mensch Jürgen“) ist stilistisch mutig in ihrer Einfachheit.

Für KI-Erzeugung spricht jedoch: Der Text ist *zu* makellos in seiner Wirkung. Jedes Detail – Tomaten, Katze, Kegelverein – dient erkennbar einem emotionalen Gesamteffekt. Die Struktur (Alltag → Verlust → gescheiterter Neuanfang) folgt einem Workshop-typischen Muster, das KI-Modelle aus Kurzprosa-Korpora internalisiert haben. Die fehlende Interpunktion bei indirekter Rede wirkt wie ein bewusst gesetztes Stilmittel zur Authentizitätssimulation. Die kontrollierte Aneinanderreihung unscheinbarer Szenen erinnert an Raymond-Carver-Epigonen – ein Stil, den Sprachmodelle besonders gut reproduzieren.

KI-WAHRSCHEINLICHKEIT: 55%

Testlauf, Text 5

Am Samstag hat mein Vater angerufen und gefragt ob ich den Rasenmäher noch brauche, den alten von ihm, den elektrischen. Ich hab gesagt ja klar, obwohl ich seit Oktober nicht mehr gemäht hab und der Rasen jetzt aussieht wie er aussieht. Aber so ist das mit meinem Vater, wenn man nein sagt fragt er warum und dann muss man sich erklären und das will ich nicht.

Er wird im März achtzig. Meine Schwester organisiert was, sie hat eine Tabelle gemacht wer was mitbringt, die hat sie in so eine Gruppe geschickt, WhatsApp, da bin ich auch drin aber ich les das meistens nicht. Jetzt soll ich Kartoffelsalat machen. Zwei Schüsseln. Ich mach den mit Mayonnaise, mein Vater mag den so, meine Schwester sagt das ist ungesund aber mein Vater wird achtzig, der kann essen was er will finde ich.

Letztes Jahr an seinem Geburtstag hab ich neben Onkel Werner gegessen der die ganze Zeit über seinen Hund geredet hat. Der Hund hat irgendwas an der Hüfte. Ich hab zugehört und genickt und irgendwann hab ich angefangen wirklich zuzuhören, weil Werner so eine Art hat, der erzählt das alles ganz ruhig und man merkt der liebt diesen Hund wirklich. Dann musste ich kurz auf Klo und als ich wiederkam hat Werner mit jemand anderem geredet und ich saß da mit meinem Bier. War trotzdem ein guter Abend. Ich mein das ernst. Es gibt schlimmere Abende als mit Bier dazusitzen und den anderen beim Reden zuzugucken.

Gutachten 5

Der Text wirkt auf den ersten Blick sehr authentisch: Er imitiert einen mündlich-alltäglichen Erzählton, enthält Nebensächlichkeiten (Rasenmäher, WhatsApp-Gruppe, Kartoffelsalat), Gedankensprünge und eine unauffällige Alltagsszene. Gerade diese scheinbare Kunstlosigkeit verdient jedoch genaue Prüfung.

Auffällig ist die **durchgehend kontrollierte Dramaturgie** unter dem Deckmantel des Beiläufigen. Der Text baut eine emotionale Pointe präzise auf: Die Passage über Onkel Werner – vom gelangweilten Zuhören zum echten Zuhören, dann der stille Moment mit dem Bier – mündet in eine reflektierte Schlussentzeng („Es gibt schlimmere Abende...“). Diese Struktur ist typisch für KI-generierte „literarische Alltagstexte“, die emotionale Tiefe durch Understatement erzeugen sollen.

Gleichzeitig fehlen **echte sprachliche Brüche**: Kein Satz entgleist wirklich, kein Gedanke bleibt unabgeschlossen. Die scheinbar nachlässige Syntax ist tatsächlich sorgfältig komponiert. Auch die Charakterisierungen (der Vater, die organisierende Schwester, Werner) wirken wie bewusst gesetzte Typen. Der Kartoffelsalat-Konflikt mit der Schwester liefert exakt die „richtige Menge“ Familienreibung.

Authentische Texte dieser Art wären typischerweise entweder weniger pointiert oder stärker abschweifend. Die Balance hier ist **zu perfekt unperfekt**.

KI-WAHRSCHEINLICHKEIT: 72%

道