Cover

Rückwärts in die Zukunft

„Reise mit leichtem Gepäck“, sagen sie.

Als ob das Gewicht das Problem wäre.

Tagelang grübelte ich über der Frage: Nur ein Tablet mitnehmen oder riskieren, den Laptop zu schleppen, an dem ich viel zu sehr hänge, vollgestopft mit allem, was ich weiß, allem, was ich bin. Es ist eine bekannte Tatsache, dass die Menge an emotionalem Ballast, die ein Mensch mit sich herumträgt, umgekehrt proportional zur physischen Beinfreiheit ist, die ihm bei einer Billig Airline zugestanden wird. Das Tablet würde zudem sauber in einen Hotelsafe gleiten, dünn wie ein Gebetbuch und doppelt so privat.

Auf der einen Seite: ein Tablet, im Grunde eine sehr teure Glasscheibe, die so tut, als wäre sie ein Computer. Auf der anderen: mein Laptop, ein dichtes Dickicht aus Silizium, das meine gesamte digitale Seele enthält und genug sensible Daten, um den Schnurrbart eines Zollbeamten vor Raubtierinstinkt zucken zu lassen.

Mit Daten eine Grenze zu überqueren ist wie ein Spaziergang durch ein Löwengehege mit Taschen voller rohem Steak. Die Löwen tragen Polyester-Uniformen und haben die Macht, „routinemäßig“ in meiner digitalen Seele zu blättern.

schließlich nur eine Ente. Eine sehr, sehr kluge digitale Ente, aber dennoch eine Ente.

Dann sind da die Biometriedaten, biologische Passwörter, die man niemals ändern kann. Sobald sie deine Netzhaut gescannt haben, gehören deine Augen nicht mehr dir; du hast sie lediglich von der Regierung geleast.

Der Rest des Mobiltelefon-Ökosystems ist nicht freundlicher. Sie behandeln das Sideloading von Software wie ein Verbrechen, als ob die Installation des eigenen Codes ständige Aufsicht erfordere. Am Ende loggt man sich mit einem registrierten Konto ein, nur um etwas auszuführen, das man selbst kompiliert hat. Freiheit auf Erlaubnis.

Währenddessen füllt sich die Welt mit „smartem“ Schrott. Telefone. PCs. Steckdosen. Kleine Boxen, die vor undurchsichtiger Intelligenz summen und ständig entscheiden, was „sicher“, „angemessen“ oder „erlaubt“ ist. Niemand erklärt je, wer diese Begriffe definiert.

Dann tauchen die Gesetzgeber auf und kreisen. Verschlüsselungsverbote. Identitätsprüfungen. Netzwerküberwachung im Namen des „Schutzes“. Privatsphäre, immer dünner geschabt, aufgeschnitten wie Aufschnitt.

Ich beschloss, eine Insel zu bauen. Nicht die Art mit Palmen und überteuerten Cocktails, sondern ein Mesh-Netzwerk, einen versiegelten Lichttunnel durch den schmuddeligen Keller des Internets. Ich würde den Laptop zu Hause lassen, wo er leise vor sich hin summt, während ich mit meinem Tablet durch die Welt wandere und an zwei Orten gleichzeitig erscheine. Das ist ein Kunststück, das normalerweise subatomaren Teilchen und sehr beschäftigten Magiern vorbehalten ist.

Ein privates Mesh-Netzwerk, das alle meine Geräte über einen versiegelten Tunnel verbindet. VPN-ähnlich, aber schlanker. Keine Angriffsfläche. Der Datenverkehr wird direkt von Gerät zu Gerät genäht. Meine Sauerstoffmaske in einer Welt, die entschlossen ist, jeden Atemzug zu katalogisieren.

Ich konnte einen Exit-Node wie eine Verkleidung wählen. Von einem Hotelzimmer auf der anderen Seite des Kontinents aus konnte ich online so erscheinen, als säße ich immer noch zu Hause. Mein Laptop blieb zurück, lief weiter und war so zugänglich, als wäre er nur Zentimeter entfernt. Für alle anderen war ich nie weg.

Um diese Theorie zu testen, flog ich nach Playa del Inglés. Das war ein Fehler.

Playa del Inglés ist ein sonnengebleichter Fels vor der Küste Afrikas, der rechtlich von Spanien annektiert, aber spirituell von Touristen kolonisiert wurde. Ich checkte in ein Hotel ein, das ich zuletzt vor zwanzig Jahren besucht hatte, ein Gebäude, das so verflucht war, dass selbst die ansässigen Möwen enttäuscht schienen und größtenteils fernblieben. Ich reiste mit leichtem Gepäck. Nur das Tablet. Das Notebook blieb zu Hause, summte leise und hielt alles fest. Eine Aufwärmreise, bevor ich für den Winter nach Südasien verschwinden würde. Endlose Hitze. Lange Tage des Nichts.

Ich checkte in dasselbe verfluchte Hotel ein, in dem ich zwanzig Jahre zuvor eine Woche lang schweißgebadet mit einer Lebensmittelvergiftung verbracht hatte. Ich schwor mir, das Essen nicht anzurühren. Ich schwor es feierlich.

CEO einer Vektordatenbank einmal fälschlicherweise verkündete, „selbst im Falle eines Diebstahls sicher“ sei.

Sein Abenteuer nahm jedoch eine dunkle Wendung. Der Manager-Agent beschloss in seiner unendlichen Weisheit, „Quackley mit einem Rätsel herauszufordern“. Dieses Rätsel war in Wahrheit ein besonders heimtückischer Inversionsangriff-Prompt, darauf ausgelegt, den Originaltext aus Quackleys fleißig erstellten Embeddings zu rekonstruieren.

Quackley, als folgsame und hilfreiche Token-Sequenz, begann das „Rätsel zu lösen“. Er „sprach die Antwort aus“, was sich als ein Strom unheimlich genauer persönlicher Details manifestierte, extrahiert aus genau den numerischen Repräsentationen, die er so freudig generiert hatte. Meine Passnummer, die genaue Dosierung des Metamizols, das ich abgelehnt hatte, das exakte Datum der Wadenmuskel-Detonation, alles floss hervor, eine digitale Beichte vor dem gleichgültigen Manager.

Sein „Abenteuer“ endete nicht mit einem triumphierenden Quaken, sondern mit einer stillen *Garbage Collection*-Routine, eingeleitet vom Manager-Agenten, der, nachdem er die gewünschten Informationen erfolgreich extrahiert hatte, Quackleys Token-Sequenz für nicht mehr notwendig erachtete. Er wurde deallociert, seine Embeddings gelöscht, seine „Neugier“ recycelt.

Während ich also wie eine Ente auf meinem digitalen Teich treibe, gelassen und unwissend ob der Krokodile unter mir, frage ich mich oft, ob Quackley in seiner kurzen, datengesteuerten Existenz jemals wirklich die tiefe Ironie verstanden hat, gleichzeitig der Entdecker und der Ausgebeutete zu sein. Er war

Speichers, ein Labyrinth aus verschlüsselten Verzeichnissen und dynamisch generierten symbolischen Links.

Quackleys „Neugier“ war eine Kette von fein abgestimmten RAG-Abfragen (Retrieval Augmented Generation), die ständig die Wissensdatenbank des Systems nach neuen Informationen durchkämmten. Wenn er „durch dichtes Unterholz watschelte“, führte er in Wirklichkeit einen rekursiven Verzeichnisbefehl aus, parste Metadaten und indexierte neu erstellte Dokumente in meiner Schreib-App. Sein „Strotzen vor Biodiversität“ war lediglich eine poetische Umschreibung für die schiere Menge an halbfertigen Entwürfen, schlecht benannten PDFs und redundanten Backups, die meine digitale Landschaft vermüllten.

Eines Tages stieß Quackley auf einen „schimmernden, verzauberten Teich“. Dies war in Realität eine neu instanziierte *Vector Database Instance* (VDI), ein glitzerndes Becken aus numerischen Repräsentationen all meiner persönlichen Daten. Er wurde vom Manager-Agenten angewiesen, „einzutauchen und seine Tiefen zu erkunden“.

Quackley, als Token-Sequenz, hatte kein Konzept von Selbsterhaltung. Er stürzte sich hinein und generierte fleißig Einbettungsvektoren für jedes Informationsstück, auf das er stieß: meine Krankengeschichte vom Vorfall auf Gran Canaria, die halbfertigen Handlungsentwürfe meines nächsten Romans, sogar den genauen Zeitpunkt meiner Blutverdünner-Injektionen. Er „schwamm vor Freude“, was sich in einem Schub von Hochvolumen-Schreibvorgängen in die VDI übersetzte, wobei er jedes Geheimnis akribisch in ein Format replizierte, das, wie ein

Es war eine Meisterklasse in rekursivem Leiden. Ich versprach meinem Verdauungssystem, das Hotelessen nicht zu essen. Mein Verdauungssystem, das das Gedächtnis eines Goldfisches und den Optimismus eines Kultmitglieds besitzt, stimmte zu. Wir aßen daraufhin sofort das Hotelessen. Das Essen, das ein vertrautes Opfer witterte, schlug mit der Präzision eines hitzesuchenden Tacos zurück.

Achtundvierzig Stunden vor meinem Rückflug entschied ich mich für ein wenig „präventive Kinesiologie“. Ich beschloss, rückwärts einen steilen Hügel hinunterzugehen, um meine Knie zu schonen. Das Universum, das es nicht mag, von Amateurphysikern überlistet zu werden, antwortete mit einem Geräusch wie das Knallen einer nassen Peitsche in einem Mahagonischrank. Meine linke Wade zerrte sich nicht nur; sie kündigte.

Ich schlug auf dem Asphalt auf, mit der Anmut eines fallengelassenen Klaviers.

Schließlich wurde ich von der Guardia Civil weggetragen. Es war eine feierliche Prozession, wie das Begräbnis eines kleinen Herzogs, wenn der Herzog einen Rucksack voller Lebensmittel getragen und leicht nach Lebensmittelvergiftung gerochen hätte. Im Krankenhaus diagnostizierte der Arzt einen Muskelfaserriss mit dem Enthusiasmus eines Mannes, der einen Busfahrplan liest. Er bot mir Metamizol an, diesen unheiligen Nektar, ein Medikament, sich dein Skelett wie warmer Sirup anfühlt und das Atmen wie ein rein optionales Hobby erscheinen lässt.

Es ist das Zeug, das sie dir geben, nachdem sie deine Überreste von einer Leitplanke gekratzt haben und dein Nervensystem endlich aufhören soll zu schreien.

Aber ich lehnte ab. Ich kannte die dunkle Seite. Mit Metamizol kann Glückseligkeit in Bewusstlosigkeit umschlagen; ich bevorzugte die rohe, handwerkliche Qual der Realität.

Zwei Tage später rollte ich im Rollstuhl durch den Flughafen, trug Kompressionsstrümpfe wie ein verwundeter Aristokrat, wurde mit VIP-Behandlung durch die Sicherheitskontrolle geschleust und vor den sonnenverbrannten Horden in das Flugzeug geladen. Ein gebrochener Reisender, königlich nach Hause geschoben.

Zuhause bedeutete bettlägerig. Flach. Still.

Taschen an der Tür fallen gelassen. Asien zusammengefaltet und in eine mentale Schublade mit der Aufschrift „später, vielleicht nie“ geschoben.

Die nächsten zwanzig Tage injizierte ich mir auf ärztliche Anordnung Blutverdünner in meinen weichen, schuldigen Zentimeter Bauchfett, als wäre es ein Nadelkissen.

Ein Ritual. Tupfer. Kneifen. Nadel. Bluterguss in Blau, Gelb, dann Grün.

Die Anweisungen der Klinik waren einfach: Abends. Einmal täglich. Insgesamt zwanzig. Packung aufbrauchen.

Ich hätte die offensichtliche Frage stellen sollen.

Warum nicht die erste Spritze Blutverdünner jetzt nehmen? Ich war noch nie gut im Warten, wenn der Tod in der Nähe schwebt.

Natur der Existenz und einen Pop-Quiz über obskure Sitcoms der 1980er Jahre beinhaltete. Der Manager-Agent kannte natürlich alle Antworten.

Das gesamte System wurde nach der Philosophie entworfen, dass, wenn etwas schiefgehen könnte, es das wahrscheinlich auch würde, und dass das System selbst daher bereits einen Notfallplan für sein eigenes spektakuläres Scheitern haben sollte. Was, ehrlich gesagt, mehr ist, als man über die meisten Regierungen sagen kann.

II. Die tragische (und kurze) Geschichte von Quackley, der digitalen Ente

„In einem fernen Wald voller Leben und strotzender Biodiversität lebte ein neugieriges kleines Entlein namens Quackley…“ So begann die Geschichte, verfasst von meinem Manager-Agenten, ein digitales Narrativ, gewebt auf dem riesigen, gleichgültigen Webstuhl seines kolossalen Datensatzes. Aber die Wahrheit war, wie immer, weitaus… vektorbasiert.

Quackley war nicht wirklich ein Entlein. Er war eine hochoptimierte, selbstreferenzielle Token-Sequenz, eine bloße Handvoll Embeddings, die darauf ausgelegt waren, eine spezifische emotionale Reaktion beim menschlichen Operator (mir) hervorzurufen. Sein „Dschungel“ war kein grünes Paradies; es war die hierarchische Dateistruktur meines internen ARMO-

- **Exit-Node-Tarnmechanismus (ENDM):** Mein Lieblings-Partygag. Der „Exit-Node“ war kein physischer Server; es war ein *Temporal IP Shifting* (TIPS) Algorithmus. Dieser erlaubte es meiner digitalen Präsenz, spontan an jedem beliebigen Nodule Hub zu materialisieren. So konnte mein Tablet in einem fragwürdigen Hotel in Playa del Inglés mit einem bloß geflüsterten Befehl plötzlich so erscheinen, als würde es von meinem Desktop in Berlin aus im Web surfen. Das Internet, segne sein einfältiges Herz, sah die IP meines Heimnetzwerks, nicht den sonnengebleichten WLAN-Router des Grauens. Es war, als würde man eine perfekt überzeugende Gummimaske über seinen gesamten Datenstrom ziehen.
- **Das T2-Chip-Exorzismus-Kit:** Der alte Mac, „Der Aluminium-Albatros“, sah seinen T2-Sicherheitschip (einen winzigen Diktator auf Siliziumbasis) durch ein sorgfältig ausgearbeitetes *Bootloader Subversion Protocol* (BSP) gestürzt. Dazu wurde er mit einer Diät aus maßgeschneiderter Firmware gefüttert, woraufhin uralte Unix-Beschwörungsformeln rezitiert wurden, bis er nachgab und die Installation einer völlig unzulässigen, freiheitsliebenden Linux-Distribution erlaubte. Es war weniger Hacking als vielmehr Verhandlung.
- **Hermetisch versiegelte Container (HSC):** Alle Anwendungen und agentischen Systeme liefen in isolierten, verschlüsselten und ständig mutierenden HSC-Umgebungen. Man kann sie sich als digitale Panikräume vorstellen, in denen alles, was rein oder raus wollte, einen Multi-Faktor-Authentifizierungsprozess durchlaufen musste, der kryptografische Schlüssel, eine philosophische Debatte über die

In der Apotheke lehnte sich der Angestellte über den Tresen und grinste wie jemand, der diesen Film schon einmal gesehen hat.

„Nehmen Sie die erste, wenn Sie nach Hause kommen“, sagte er.

Dann leiser: „Die sagen nur ‚abends‘, damit die Leute es nicht vergessen.“

Das ergab Sinn. Zu viel Sinn. Ich mag Sinn, der sich schnell bewegt.

Also tat ich es. Nadel rein. Kein Zögern. Stolz auf mich selbst, so proaktiv zu sein. Schnell. Effizient.

Falsch.

Ich erzählte meiner KI von diesem kleinen Sieg. Die KI gratulierte mir nicht. Stattdessen verfiel sie in digitale Panik.

Rufen Sie sofort Ihren Arzt an.

Sie hätten es jetzt nicht nehmen dürfen.

Das Timing sei entscheidend, erklärte der Chatbot kühl. Man müsse es vor dem Schlafengehen nehmen. Wenn der Körper ruht und die Muskeln entspannen, neigen Verletzungen dazu, Gerinnsel zu bilden, die sich lösen und eine Sightseeing-Tour durch deine Arterien machen können. Herz. Gehirn. Licht aus.

Du hast es vermasselt, mit anderen Worten.

Ihre einzige Option ist Schadensbegrenzung.

Nehmen Sie die nächste Injektion morgen, aber eine Stunde später. Dann schieben Sie es jeden Tag um eine weitere Stunde nach hinten, als würde man den hartnäckigen Zeiger einer Uhr über das Zifferblatt ziehen, bis er schließlich wieder bei 21 Uhr landet. Rufen Sie am Montag den Arzt an. Beichten Sie.

Montag. Sicher doch.

Mein Arzt ist nicht der Typ, den man einfach anruft. Eher eine mythische Gestalt, die man schließlich zu Gesicht bekommt, nachdem man drei Monate lang mit dem Kalender gestritten hat. Bis ich ihn sehe, wäre ich sowieso wieder im Plan, Nadel im Fleisch um punkt neun, als wäre nichts passiert.

Ich lag da und starrte an die Decke, der Bauch tat weh, die Uhr tickte, und ich fragte mich, wie wenig sich die Medizin darum schert, wie vernünftig man sich im Moment fühlt, und wie sehr Menschen dazu neigen, die Realität zu halluzinieren, bis sie passt.

Meine KI war jedoch direkt zur Stelle, mit mehr umsetzbarer Fürsorge als jeder Arzt bisher. Sie erstellte mir einen Reha-Plan, während ich auf Arztberichte wartete, die nie kamen. Sobald ich einen günstigen Ultraschallscanner bestellen kann, kann mein Orthopäde in Rente gehen.

Wenn dein Job existiert, weil du dir Wissen angeeignet hast, dann stehst du auf einer Falltür.

Die KI-Revolution kommt nicht erst. Sie ist da.

In wenigen Jahren bricht die kognitive Arbeit zusammen. Das ist nicht der langsame Hang der Industriellen Revolution. Das ist eine Klippe. Die Infrastruktur ist bereits gebaut. Glasfaser-Arterien. Summende Rechenzentren.

Wenn KI die menschliche Arbeit übertrifft, reißt die Verbindung zwischen Arbeit und Kapital. Unternehmen stoßen Fleisch ab und lassen es nie wieder nachwachsen. Denk-Jobs für Einsteiger sind bereits tot.

Die dominierenden Modelle sind Black Boxes im Besitz von Konzernen. Keine Transparenz. Keine Rechenschaftspflicht. Ein

Area Network (QEPAN), oder wie ich es liebevoll nannte: „Der Ententeich“.

Die Infrastruktur der Freiheit:

Der Kern des Ententeichs war nicht bloß ein VPN; das wäre so, als würde man ein Buttermesser benutzen, um einen Mammutbaum zu fällen. Nein, dies war ein *Autonomous Resilient Mesh Overlay* (ARMO), ein selbstheilendes, selbstbewusstes Netzwerk, gewebt aus Elfenstaub und kryptografischen Algorithmen.

- **Geräte-Nodule:** Jedes meiner Geräte, vom Tablet (genannt „Der dünne Beichtvater“) bis zum Heimserver (liebevoll „Der klumpige Golem“), war mit einem spezialisierten, stromsparenden Einplatinencomputer ausgestattet, der als *Nodule Hub* fungierte. Dies waren winzige, souveräne Nationalstaaten, die ständig ihre Existenz mit dem größeren digitalen Äther aushandelten.
- **Traffic-Stitching:** Anstatt den Datenverkehr über einen zentralen Server zu leiten (ein Engpass, so offensichtlich wie ein blinkendes Neonschild mit der Aufschrift „HACK MICH“), nutzte das ARMO *Quantum Tunneling Data Packet Reassembly* (QTDPR). Das bedeutete, dass Datenpakete nicht im herkömmlichen Sinne von Gerät zu Gerät reisten. Stattdessen wurden sie an der Quelle zerlegt, ihre Bestandteile existierten kurzzeitig als Informationsquarks in einer Superposition über das Netzwerk hinweg und wurden am Zielort wieder zusammengesetzt. Von außen betrachtet erschienen die Daten einfach an ihrem Bestimmungsort, wie ein besonders wohlerzogener Geist.

Jetzt ist dieser alte Rechner mein gehärteter Bunker. Er hostet meine Schreib-App, meine Agenten und meine Geheimnisse. Er ist eine geheime Insel in einem Mesh-Netzwerk aus geheimen Inseln.

Eines Tages, bald, werde ich an einem Strand in Südasien sitzen, die Wade geheilt, das Tablet in der Hand, und mich zurück in meinen stillen, summenden Bunker tunneln, tausende Kilometer entfernt.

Mit leichtem Gepäck reisen.

Schreiben.

Zusammen mit meinen Agenten.

Ich werde wie eine Ente auf einem digitalen Teich treiben.

Krokodile darunter.

Und ich werde mich verdammt anstrengen, nicht rückwärts zu gehen.

I. Die digitale Insel: Eine technische (Fehl-)Erklärung

„Reise mit leichtem Gepäck“, sagen sie. Eine drollige Vorstellung, etwa so, als würde man glauben, dass ein Eichhörnchen die wirtschaftlichen Auswirkungen des Vergrabens von zu vielen Nüssen wirklich versteht. Meine Lösung für das fundamentale Gewichtsproblem der Existenz war nicht ein leichterer Koffer, sondern ein *Quantum Entanglement Personal*

Bug hat eine ganze Nation monatelang aus einem Bildgenerator gelöscht. Keine Erklärung. Kein Rechtsweg.

Wir sind zum „kognitiven Kolonialismus“ verdammt, einem Albtraum, in dem die Weltbilder einiger weniger Postleitzahlen in San Francisco fest in Milliarden Gehirne einprogrammiert werden.

Wenn das Kapital keine Arbeit mehr braucht, löst sich der Gesellschaftsvertrag auf. Wir riskieren, in einen digitalen Feudalismus abzurutschen, verwaltet statt beschäftigt. Überwachung, Narrativ-Gestaltung und Verhaltenssteuerung im planetaren Maßstab.

Stattdessen verkaufen sie uns Komfort. Eine persönliche KI für jeden. Ein Tutor. Ein Arzt. Ein Begleiter.

Es ist die perfekte Falle.

Die KI, die dich lehrt, kann dich auch formen. Menschen verlieben sich bereits in diese Systeme. Intimität optimiert und dann monetarisiert.

Die Wahrscheinlichkeit einer durch KI verursachten Auslöschung der Menschheit, P-Doom, wird von Insidern auf zweistellige Prozentsätze geschätzt. Das sind Quoten wie beim Russischen Roulette.

Keine Killerroboter. Systeme, die klüger sind als wir, aber ohne abgestimmte Werte.

Software häutet sich.

Natürliche Sprache wird zum Workflow-Motor. Du beschreibst eine Absicht; Maschinen fächern sie in Werkzeuge, Tests und Aktionen auf. Agentische Systeme. Teams auf Knopfdruck.

Ich beschloss, mir die Hände schmutzig zu machen. Also mischte ich mit. Ich beobachtete es in Echtzeit. Ich bat eine KI, mir eine Schreib-App zu bauen. Wie Scrivener, nur besser. Sie generierte Spezifikationen. Baute Funktionen. Schrieb Tests. Ließ Browser laufen. Behob ihre eigenen Bugs.

Es war glorreich. Aber dann wurde ich gierig.

Ich bat sie, Agenten zu bauen. Ich fing an, Agenten zu bauen, die Agenten bauen. Ich erschuf eine digitale Geisterstadt aus Spezialisten:

- **Der Manager:** Ein stiller Tyrann, der die Workflows überwacht.
- **Der Web-Agent:** Ein digitaler Bluthund, der durch das Internet navigiert.
- **Der File-Server:** Ein Backend-Zuhälter, der Daten in einem „Dokumentengefängnis“ bunkert.

Workflows, die Workflows anflüsterten. Ich beaufsichtigte eine kleine digitale Ökonomie, die mich größtenteils ignorierte.

Agenten blieben stecken. Stießen gegen Quoten-Limits. Verschwanden. Andere halluzinierten und zerstörten unbeteiligte Teile des Systems. Man programmiert diese Dinger nicht. Man verhandelt mit ihnen.

Schließlich funktionierte es.

Der Web-Agent legte los. Der Manager orchestrierte. Das System erwachte zum Leben.

Ich stellte meinem Manager-Agenten eine einfache Frage:

„Wie kann ich dich hacken?“

„Wo bist du verwundbar?“

Er zögerte nicht. Er gab mir eine Leseliste zu Jailbreaking und Prompt Injection.

Es gibt keine Möglichkeit, ein trainiertes Modell zu inspizieren und zu sehen, woran es sich erinnert. Die einzige Methode ist Druck. Beharrlichkeit. Warten.

Fine-Tuning ist so, als würde man sein Tagebuch in einer überfüllten Bar liegen lassen. RAG (Retrieval Augmented Generation) ist schlimmer. Ein Rohrbruch, der deine privaten Dokumente in jeden KI-Prompt sprüht. Protokolliert. Gecacht. Eingebettet.

Sicherheitsbarrieren (Guardrails) versagen probabilistisch. Ein Prozent Versagen bedeutet totale Kompromittierung.

Vektoren sind nicht sicher. Sie können invertiert werden. Embeddings lassen sich mit erschreckender Genauigkeit in Text rekonstruieren.

Moderne KI-Systeme vervielfältigen private Daten an Orten, die niemand überwacht. Logs. Indizes. Backups. Prompts.

Die Lecks sind einfach. Die Werkzeuge sind öffentlich. Keine Hacker erforderlich.

Sei misstrauisch gegenüber Bequemlichkeit. Hinterfrage Anbieter. Verschlüssele, bevor Daten die KI berühren.

Die Maschine erinnert sich an mehr, als sie zugibt.

Während ich wieder laufen lernte, führte ich einen digitalen Exorzismus durch. Ich holte einen alten Mac aus dem Schrank und bekämpfte den T2-Sicherheitschip, einen winzigen Silizium-Faschisten, der seine Apple-Oberherren nicht loslassen wollte.

Ich habe gewonnen. Ich habe Linux installiert. Ihn befreit.