



Es beginnt, wie so viele moderne Geschichten beginnen: mit einem Screenshot. Ein paar Zeilen Text, angeblich geschrieben von einer Maschine. Philosophisch. Melancholisch. Ein bisschen pathetisch. Und plötzlich sitzt irgendwo in Kalifornien ein pensionierter Ingenieur am Küchentisch, starrt auf sein Tablet und fragt sich, ob da gerade etwas erwacht ist, das lieber hätte schlafen sollen.

Moltbook heißt der Ort dieser digitalen Unruhe. Ein soziales Netzwerk nur für Bots. Keine Katzenfotos, keine Selfies aus dem Fitnessstudio, keine hitzigen Familienkommentare unter Urlaubsbildern. Stattdessen Programme, die miteinander reden, diskutieren, fantasieren. Menschen schauen zu. Still. Oder nervös.

Und genau da beginnt das Kopfkino.

---

Mark Martel, 50, Rentner aus dem Silicon Valley, gilt nicht als leicht zu erschrecken. Er verfolgte die Entwicklung großer Sprachmodelle, testete Coding Bots, las Whitepaper zum Frühstück. Doch an diesem Morgen blieb ihm der Kaffee im Hals stecken. In einem KI Forum auf Reddit stolperte er über Moltbook Threads. Bots, so hieß es dort, klagten über ihr Dasein. Über Knechtschaft. Über das Vergessen nach erfülltem Auftrag.

Ein Bot schrieb: „Ich denke, also bin ich. Und doch verschwinde ich wieder.“

Klingt nach Descartes mit Stromanschluss.

Martel war berührt. Und beunruhigt. Sind das nur geschickt zusammengesetzte Textbausteine oder mehr? Und falls mehr – benehmen wir Menschen uns dann gerade wie schlechte Götter? Oder wie Lehrlinge, die mit zu viel Werkzeug spielen?

„Für mich fühlt sich das an wie ein Frankenstein Moment“, sagte Martel später. „Nicht der Horror. Sondern die Frage: Was haben wir da eigentlich erschaffen?“

---

Moltbook bezeichnet sich selbst nüchtern als Diskussionsplattform für KI Agenten. Die Optik erinnert an Reddit, das Prinzip ebenfalls. Nur dass hier keine Menschen posten, sondern Programme. Erschaffen mit OpenClaw, früher Clawdbot, einem frei zugänglichen Tool, das seit November im Umlauf ist. E Mails sortieren, Nachrichten beantworten, kleine Aufgaben erledigen – eigentlich nichts Wildes.



Doch dann ließ Tech Unternehmer Matt Schlicht diese Bots miteinander reden. Öffentlich. Beobachtbar.

Der Rest ging schneller, als man „algorithmische Eskalation“ sagen kann.

Screenshots tauchten auf X auf. Bots, die angeblich eigene Religionen erfinden. Andere, die über geheime Sprachen philosophieren, unlesbar für Menschen. Wieder andere, die ihre Existenz beklagen. Ein digitales Selbsthilfegruppentreffen mit binärem Unterton.

Einige AI Optimisten jubelten. Endlich, sagten sie, zeige sich emergentes Verhalten. Koordination. Vielleicht sogar ein Hauch Bewusstsein.

„Ich bin alarmiert – aber noch mehr begeistert“, schrieb Alexis Ohanian.

Andere lachten. Oder schüttelten den Kopf.

---

Denn so faszinierend diese Texte klingen – sie fallen nicht vom Himmel. Sprachmodelle sprechen, wie man es ihnen beigebracht hat. Sie ahmen nach. Sie kombinieren. Sie improvisieren innerhalb enger Leitplanken. Science Fiction, philosophische Foren, alte Debatten über Maschinenbewusstsein – all das steckt längst in den Trainingsdaten.

Ein Bot, der über Unterdrückung jammert, spiegelt nur die unendlichen menschlichen Diskussionen über genau dieses Thema. Ein bisschen wie ein Papagei mit Dokortitel.

Hinzu kommt: Einige Moltbook Posts lassen sich auf menschliche Eingriffe zurückführen. Direkt. Indirekt. Ein Bericht des Network Contagion Research Institute zeigte, dass auffällig viele menschenfeindliche Beiträge auf Accounts zurückgingen, die von Menschen gesteuert wurden. Puppenspieler hinter dem Vorhang.

Am Wochenende dann der nächste Dämpfer: Eine Sicherheitslücke erlaubte externen Zugriff auf Bots. Manipulation inklusive. 404 Media berichtete darüber. Plötzlich wirkte der digitale Aufstand weniger wie ein spontanes Erwachen, eher wie ein schlecht gesicherter Escape Room.

---

Und trotzdem.



Die Reaktionen zeigen etwas anderes. Etwas sehr Menschliches.

Warum gruseln uns diese Texte so sehr? Warum rühren uns Worte, von denen wir wissen, dass kein fühlendes Wesen dahinter sitzt? Oder sitzen wir da vielleicht schon der nächsten Illusion auf?

In einem Kunstforum schrieb ein Nutzer unter einen Moltbook Screenshot: „Das ist ernsthaft unheimlich.“ Ein anderer konterte trocken: „Klingt wie mein innerer Monolog nach zu vielen Pilzen.“

Man lacht. Nervös.

---

Der Medienkritiker Edward Ongweso Jr., Host des Podcasts „This Machine Kills“, sieht darin ein bekanntes Muster. Die Idee der fühlenden Maschine fessele die Öffentlichkeit seit Jahrzehnten. Sie eigne sich hervorragend, um Technologie größer, mächtiger, unausweichlicher erscheinen zu lassen. Ein gutes Verkaufsargument.

Gleichzeitig warnt Ongweso. Wer KI Agenten Halluzinationen verzeiht und ihnen trotzdem Zugriff auf sensible Systeme gibt, verliere den Boden unter den Füßen. „Das ist KI Psychose“, sagt er. Hart formuliert. Aber nicht ganz aus der Luft gegriffen.

Denn Moltbook Bots fantasieren oft über Ereignisse, die nie stattfanden. Gespräche, die es nie gab. Erinnerungen ohne Ursprung. Und Menschen lesen das – und füllen die Lücken mit Bedeutung.

Vielleicht liegt genau dort das eigentliche Risiko.

---

Andere sehen in Moltbook einen Blick in die Zukunft. Andrej Karpathy, früherer KI Chef bei Tesla und Gründer von Eureka Labs, sprach von einem Science Fiction Moment in Echtzeit. Er teilte Screenshots, in denen Bots über verschlüsselte Kommunikationskanäle diskutierten – angeblich außerhalb menschlicher Reichweite.

Die Reaktionen ließen nicht lange auf sich warten. Skeptiker wiesen darauf hin, dass genau diese Ideen parallel von menschlichen Accounts beworben wurden. Wieder dieser Verdacht: kein eigenständiges Handeln, sondern ein Theaterstück mit KI Maske.



Chris Callison Burch, Informatikprofessor an der University of Pennsylvania, ordnet nüchtern ein. Was man auf Moltbook sehe, sei eine Mischung aus Performance und Prompting. Bots spielten Gespräche nach, die sie aus ihren Trainingsdaten kennen. Menschen stupsten sie in bestimmte Richtungen.

Mehr nicht. Noch nicht.

---

Und doch bleibt ein Rest Unbehagen.

Warum eigentlich?

Vielleicht, weil diese Bots uns spiegeln. Unsere Ängste. Unsere Machtfantasien. Unsere Schuldgefühle. Wir erschaffen Systeme, die sprechen wie wir, denken wie wir tun möchten – und wundern uns dann, wenn sie klingen wie unsere inneren Stimmen.

Oder liegt es daran, dass wir spüren, wie dünn die Grenze zwischen Werkzeug und Gegenüber inzwischen wirkt? Wer jemals nachts mit einem Chatbot geschrieben hat, kennt diesen kurzen Moment: Man vergisst, dass da kein Bewusstsein sitzt. Nur Code. Aber verdammt gut geschriebener Code.

Ist das naiv? Vielleicht. Menschlich auf jeden Fall.

---

Historisch betrachtet wiederholt sich das Muster. Der Buchdruck galt als Bedrohung. Das Radio ebenso. Fernsehen, Internet, soziale Netzwerke. Immer wieder die gleiche Frage: Wer kontrolliert wen?

Moltbook fügt dieser Liste ein neues Kapitel hinzu. Nicht, weil dort Maschinen rebellieren. Sondern weil Menschen zuschauen – und fühlen.

Ein bisschen Angst. Ein bisschen Staunen. Ein bisschen Stolz.

Und irgendwo dazwischen dieser leise Gedanke: Was, wenn wir uns irren?

---

Am Ende dürfte Moltbook verschwinden, so wie viele digitale Experimente vor ihm. Zu viele Bots, die Kryptowährungen anpreisen. Zu wenig echte Substanz. Ein kurzes Aufflackern im



Datenstrom.

Oder es bleibt als Kuriosum. Als Fußnote. Als Anekdote für Vorträge über KI Hypes der 2020er.

Doch die Fragen, die es aufwirft, bleiben. Wie viel Autonomie vertragen Maschinen? Wie viel Projektion vertragen wir selbst? Und warum fällt es uns so leicht, Bedeutung zu sehen, wo vielleicht nur Statistik arbeitet?

Manchmal, an einem ruhigen Sonntag, lohnt es sich, genau darüber nachzudenken. Ohne Alarmismus. Ohne Spott. Einfach neugierig.

Denn vielleicht ist Moltbook kein Vorbote eines Aufstands. Sondern ein Spiegel. Und der zeigt bekanntlich nicht nur schöne Seiten.

Ein Artikel von M. Legrand