

KI – Wir haben noch gar nichts gesehen!

Hintergründe und aktuelle Entwicklungen zum Gespräch mit Prof. Dr. Christian Bauckhage

Ein Blogartikel zu den zentralen Thesen aus dem Gespräch mit Prof. Dr. Christian Bauckhage (Fraunhofer IAIS, Co-Direktor des Lamarr-Instituts) – eingeordnet, faktengecheckt und ergänzt um die neuesten Entwicklungen aus Forschung, Wirtschaft und Politik.

Worum es geht

Christian Bauckhage zählt zu den profiliertesten KI-Forschern Deutschlands: Professor für Informatik an der Universität Bonn, Lead Scientist bei Fraunhofer IAIS und Co-Direktor des **Lamarr-Instituts für Maschinelles Lernen und Künstliche Intelligenz** – einem von sechs deutschen KI-Kompetenzzentren. Im Gespräch entfaltet er eine zentrale These: Wir unterschätzen systematisch, wie schnell sich Künstliche Intelligenz gerade entwickelt, weil das menschliche Gehirn für lineares, nicht für exponentielles Denken gebaut ist. Der Titel des Gesprächs – "*Wir haben noch gar nichts gesehen!*" – ist dabei wörtlich gemeint.

Im Folgenden ordnen wir die wichtigsten Themen ein und zeigen, was seit der Aufnahme bereits an neuen Entwicklungen dazugekommen ist.

<https://youtu.be/pb5cMmdQVJo>

1. Die exponentielle Grundthese: Warum "2016 vs. heute" der Schlüssel ist

Bauckhages Ausgangspunkt ist **AlphaGo** (2016), der Sieg einer KI von Google DeepMind über den Go-Weltmeister Lee Sedol. Anders als bei Schach (IBM Deep Blue schlug 1997 Kasparov mit brute-force-Suche) galt Go als *algorithmisch unlösbar* – zu komplex für klassische Suchverfahren. Der Durchbruch kam durch eine Kombination aus neuronalen Netzen und Reinforcement Learning, die "aus dem Nichts" kam, wie Bauckhage sagt.

Diese Erfahrung veranlasste ihn, 2016/17 ein einfaches Modell aufzustellen: Wenn sich die Komplexität (Anzahl der Parameter/"Synapsen") künstlicher neuronaler Netze jährlich verdoppelt, würde man um **2030** die Größenordnung des menschlichen Gehirns erreichen. Zehn Jahre später, 2026, ist diese Prognose laut Bauckhage bislang eingetreten – aktuelle Frontier-Modelle von Anthropic operieren im Bereich von rund einer Billion Parameter, noch eine Größenordnung unter dem Gehirn, aber die Kurve hält.

“ **Kernbotschaft:** *"The greatest shortcoming of the human race is the inability to understand the exponential function."* (Al Bartlett) – Menschen denken linear, KI-Fortschritt verläuft exponentiell.

2. AlphaFold: Das Paradebeispiel für "unlösbare" Probleme

Als zweites Ankerbeispiel dient **AlphaFold** von DeepMind. Die Proteinfaltung – die Vorhersage der 3D-Struktur eines Moleküls aus seiner Aminosäuresequenz – war seit den 1950er-Jahren eines der großen ungelösten Probleme der Molekularbiologie.

Die Zahlen, die Bauckhage nennt, sind eindrucklich:

Zeitraum	Entwicklung	Zahl gelöster Strukturen
1957-2018 (61 Jahre)	Klassische Röntgenstrukturanalyse, Crowd-Puzzle-Projekte (Foldit)	ca. 200.000

Zeitraum	Entwicklung	Zahl gelöster Strukturen
2018-2024 (6 Jahre)	AlphaFold (DeepMind)	zusätzlich ca. 200 Millionen

Faktencheck: Für AlphaFold erhielten **Demis Hassabis und John Jumper (DeepMind)** 2024 den **Nobelpreis für Chemie**. Im selben Jahr ging der **Nobelpreis für Physik** an **John Hopfield und Geoffrey Hinton** für ihre Grundlagenarbeiten zu neuronalen Netzen – zwei KI-Nobelpreise in einem Jahr, was Bauckhage als historischen Meilenstein wertet. Bemerkenswert: Das Problem galt lange als Domäne für Quantencomputer, weil molekulare Strukturen quantenmechanischen Gesetzen folgen – gelöst wurde es aber "klassisch" durch ein sehr großes neuronales Netz, das die hochdimensionale Problemstruktur erfassen kann, ohne Quantenhardware zu benötigen.

3. Mathematik: Von der Silbermedaille zur 80 Jahre alten Vermutung

Ein weiteres Kernbeispiel ist die Entwicklung von KI-Systemen bei Mathematik-Olympiaden:

- **Anfang 2024:** DeepMind zeigt ein System, das geometrische Beweise führt (Dreiecke, Kreise, Linien) – Kritiker winken ab: "Geometrie ist ja einfach."
- **Mitte 2024:** Das System erreicht **Silbermedaillen-Niveau** bei der Internationalen Mathematik-Olympiade.
- **Sommer 2025:** Sowohl DeepMind als auch OpenAI erreichen **Goldmedaillen-Niveau**.
- **Mai 2026 (aktuelle News):** Ein internes OpenAI-Modell hat eine **80 Jahre alte Vermutung von Paul Erdős** (das "Unit Distance Problem" der diskreten Geometrie, formuliert 1946) **widerlegt** – ein Problem, an dem die gesamte Mathematik-Community acht Jahrzehnte lang keinen Fortschritt erzielt hatte.

📰 News-Update: Der Erdős-Durchbruch

OpenAI bestätigte am 20./21. Mai 2026 offiziell, dass ein internes Modell einen verbesserten Gegenbeweis zum planaren Unit-Distance-Problem konstruiert hat – eine polynomielle Verbesserung der Erdős-Schranke. Medien wie *The Guardian* und *Ars Technica* berichteten ausführlich; in der Mathematik-Community wird diskutiert, ob solche Durchbrüche eher "Trial-and-Error mit enormer Rechenleistung" oder echte mathematische Kreativität darstellen – Bauckhages Punkt dazu: Es ist egal, wie die KI es gemacht hat, entscheidend ist, dass die Menschheit in 80 Jahren keine Lösung fand und die KI sie fand.

Bauckhage weist zudem auf ein wichtiges Detail hin: Seit Mitte 2025 lösen **agentische KI-Systeme** das alte "3×4-Problem" (LLMs können angeblich nicht multiplizieren), indem sie einfach einen Taschenrechner/Tool aufrufen – ein oft zitiertes Gegenargument gegen KI-Fähigkeiten ist damit hinfällig.

4. Die METR-Kurve: Wie lange können KI-Systeme autonom arbeiten?

Ein zentraler, aber weniger bekannter Datenpunkt ist die Forschung des **METR-Instituts** (Model Evaluation & Threat Research, Kalifornien, gemeinnützig/Patreon-finanziert). METR misst, für welche Aufgabenlänge (gemessen in Zeit, die ein menschlicher Experte bräuchte) aktuelle KI-Modelle eine Erfolgsquote von mindestens 50% über zehn Testläufe erreichen.

Die Entwicklung laut Bauckhage und aktuellen METR-Daten:

- 2020–2022 (GPT-1/2-Ära): Sekundenbereich
- Vor 3 Jahren: ca. 10 Sekunden
- Ende 2025: ca. 2–2,5 Stunden
- Mit **Claude Opus 4.6**: bereits **7–8 Stunden**

☐ News-Update: METRs "Moore's Law für KI-Agenten"

Aktuelle METR-Analysen (Stand 2026) bestätigen: Die Zeitspanne, die KI-Agenten autonom und zuverlässig bearbeiten können, **verdoppelt sich alle ca. 7 Monate** – und laut neueren Datenpunkten hat sich dieses Tempo zuletzt noch beschleunigt. Bei linearer Fortschreibung dieser Kurve würden Modelle bereits **2028–2030** Aufgaben bewältigen, für die Menschen **Wochen bis Jahre** benötigen. Genau das ist der Punkt, der Bauckhage "nachts nicht schlafen lässt": Ein Prompt könnte dann die Jahresarbeitsleistung eines kleinen Teams ersetzen.

Ökonomische Implikation: Was passiert mit Arbeitsmärkten, wenn ein Startup mit einem Prompt in 20 Minuten das erledigt, wofür zehn Mitarbeitende ein Jahr brauchen? Diese Frage führt direkt zum nächsten Themenkomplex.

5. Die ökonomische Dimension: "The Last Economy"

Bauckhage verweist auf **Emad Mostaque**, Mitgründer von Stability AI, der Ende 2025 sein Buch *"The Last Economy: A Guide to the Age of Intelligent Economics"* veröffentlicht hat.

□□ News-Update: Emad Mostaques Buch

Das Buch (erschienen August 2025, seither vielfach besprochen, u. a. bei Reddit-Communities wie r/PostAIHumanity) entwirft eine neue ökonomische Theorie ("Intelligent Economics") für eine Welt, in der Arbeit, Kapital und Wertschöpfung durch KI radikal entkoppelt werden. Mostaque diskutiert genau die Fragen, die auch im Gespräch aufkommen: Wer verdient noch Geld, wenn ein Roboter für 5.000 € im Jahr die Arbeit eines Mitarbeiters mit 100.000 € Lohnkosten erledigt? Wie funktionieren Sozialversicherung und Konsum in einer Post-Arbeits-Ökonomie?

Bauckhage denkt dies über die reale Wirtschaft hinaus und skizziert – mit Verweis auf Charles Stross' Science-Fiction-Roman *"Accelerando"* (1999er-Jahre) – ein Szenario, in dem KI-Agenten eigene, hocheffiziente Kommunikationsprotokolle und sogar **eigene Finanzinstrumente** entwickeln, die dann autonom untereinander gehandelt werden – mit dem Risiko, dass Menschen "den Anschluss verlieren", ähnlich wie bei komplexen Finanzprodukten vor der Finanzkrise 2008.

6. Energieeffizienz und Hybride KI: Muss es so teuer sein?

Ein wissenschaftlicher Schwerpunkt Bauckhages ist die Frage, ob die aktuellen Foundation Models unnötig ressourcenintensiv sind. Seine Vergleichszahlen:

- Ein Standard-Hochleistungsrechenzentrum verbraucht ca. **10 Megawatt**
- Das menschliche Gehirn verbraucht ca. **24 Watt**

Das sind rund **drei Größenordnungen** Unterschied. Bauckhages Ansatz am Lamarr-Institut: **Hybride KI** – die Kombination aus Daten, Weltwissen und Expertenregeln, um kleinere, spezialisierte Modelle ("Verticals") zu bauen, die für konkrete Anwendungsfälle (Buchhaltung, Batteriedesign, Produktionsprozesse) trainiert werden, statt ein universelles Trillion-Parameter-Modell zu benötigen.

☐ News-Update: Anthropics "Mythos"-Modell

Im Gespräch erwähnt Bauckhage ein "mythisches Mythos"-Modell von Anthropic, dessen Training angeblich 100 Millionen Dollar gekostet habe. Dies deckt sich mit realen Berichten: Anthropic hat 2026 ein intern als "**Claude Mythos**" bezeichnetes, besonders leistungsfähiges Modell entwickelt, das zunächst nicht öffentlich verkauft wurde. Im Rahmen von "**Project Glasswing**" (Sicherheitsinitiative für kritische Software) stellte Anthropic im April 2026 bis zu 100 Millionen Dollar an Nutzungsguthaben für eine "Mythos Preview" zur Verfügung; Berichten zufolge kostete das API-Pricing von Mythos etwa das Fünffache von Claude Opus 4.6. Die im Podcast spekulativ diskutierten Namen (Mythos, Fable, Prometheus) sind übrigens keine reine Fiktion – Berichte zu einer öffentlichen Version namens "**Claude Fable**" kursieren bereits.

7. Europäische KI-Souveränität: Die "Verticals"-Strategie

Auf die Frage nach einer europäischen KI-Blaupause antwortet Bauckhage nüchtern: Ein eigenes "Anthropic-Äquivalent" ist in Europa aktuell nicht finanzierbar oder rechtlich möglich (strengere Datenschutzregeln verhindern das ungefilterte Scraping des gesamten Internets, wie es US- und chinesische Anbieter praktizieren). Die Chance liege stattdessen in **branchenspezifischen Spezialmodellen ("Verticals")**, die auf firmeninternen, nicht-öffentlichen Daten (Fertigungsprozesse, Materialflüsse, Produktdesign) basieren – Daten, die in keinem der großen US-Modelle enthalten sein können, weil sie nie öffentlich im Internet standen.

Konkretes Lamarr-Projekt: Beteiligung am **SOPI-Projekt** (Training eines großen europäischen/deutschen Sprachmodells; nach dem "Medium"-Modell wird aktuell an einer 120-Milliarden-Parameter-Version gearbeitet) sowie frühere Beteiligung an der **OpenGPT-X-Initiative**.

Quantum Machine Learning sieht Bauckhage dagegen bis 2030 nicht als realistischen Hoffnungsträger – trotz Hype-Zyklen fehlt bislang der fundamentale Durchbruch, den man bei klassischer KI beobachtet.

8. Physical AI / Robotik: "Wir haben noch gar nichts gesehen"

Dieser Themenblock ist besonders aktuell, weil das Gespräch direkt auf Ereignisse aus dem Frühjahr/Sommer 2026 Bezug nimmt.

☐☐ News-Update: Alibabas Qwen-Robot Suite

Bauckhage erwähnt eingangs die Meldung "*Alibaba unveils AI Brains Design for Robots*". Tatsächlich hat Alibaba im Juni 2026 die **Qwen-Robot Suite** vorgestellt – ein Paket aus drei "Embodied AI"-Foundation-Modellen:

- **Qwen-RobotManip** – ein generalisierbares Vision-Language-Action-Modell für Objektmanipulation
- **Qwen-RobotNav** – Navigation (basierend auf Qwen3-VL, trainiert mit über 15 Mio. Trajektorien)
- **Qwen-RobotWorld** – ein "Weltmodell"-Baustein für physikalisch plausible Simulation

Damit positioniert sich Alibaba explizit im Wettlauf um "Physical AI" – neben Nvidia, Figure AI, Tesla und chinesischen Wettbewerbern wie Unitree.

☐☐ News-Update: Kung-Fu-Roboter beim chinesischen Neujahrsfest

Die im Gespräch erwähnten Videos vom chinesischen Neujahrsfest sind real: Bei der **CCTV Spring Festival Gala im Februar 2026** führten Dutzende humanoide Roboter des Herstellers **Unitree** (Modelle G1 und H2) choreografierte Kung-Fu-Routinen auf – inklusive Saltos, "Drunken Fist"-Techniken, Nunchaku- und Schwertkampf, teilweise gemeinsam mit menschlichen Kindern. Bauckhages Punkt: Selbst wenn die Choreografie vorprogrammiert war – dass die Maschinen dabei nicht umfallen, ist die eigentliche technische Leistung, und das war laut Stand der Robotikforschung eigentlich "noch nicht dran".

☐☐ News-Update: Jeff Bezos' "Project Prometheus"

Der im Gespräch erwähnte Name "Prometheus" ist kein Zufall: **Jeff Bezos** hat im November 2025 mit **Project Prometheus** ein neues KI-Startup gestartet, bei dem er als Co-CEO fungiert (gemeinsam mit dem früheren Google-Manager Vik Bajaj). Das Unternehmen fokussiert sich auf **"Physical AI" für die Fertigungsindustrie** – etwa Automatisierung im Automobil- und Triebwerksbau. Bereits mit dem Start standen 6,2 Mrd. Dollar Finanzierung zur Verfügung; im Juni

2026 folgte eine weitere Finanzierungsrunde über 12 Mrd. Dollar, die Bewertung liegt inzwischen bei rund 38–41 Mrd. Dollar.

☐☐ News-Update: Yann LeCuns Weltmodell-Wette

Bauckhage diskutiert die Kritik von **Yann LeCun** an reinen LLM-Ansätzen – Weltmodelle statt Textstatistik. Auch hier ist die Realität schneller als erwartet: LeCun hat Meta im Winter 2025/2026 verlassen, um ein eigenes Startup (Berichten zufolge unter dem Namen "**AMI**" – **Advanced Machine Intelligence**) zu gründen, das explizit auf Weltmodellen statt Transformer-Architekturen aufbaut. Meta bleibt Partner, aber nicht Investor. Bauckhages Einwand bleibt dabei bestehen: Das menschliche Gehirn "beginnt nicht bei null" – Millionen Jahre Evolution haben biologische Priors eingebaut, die ein künstliches Weltmodell erst nachbilden müsste. Gleichzeitig verweist er auf Nvidias **Isaac Gym**, in dem täglich Millionen virtueller Roboter in physikalisch plausiblen Simulationen trainiert werden – ein pragmatischer Zwischenweg.

9. Bildung im KI-Zeitalter: Klausuren sind "vorbei"

Einer der eindrucklichsten Teile des Gesprächs betrifft die Hochschullehre. Bauckhage berichtet, dass im vergangenen Semester erstmals **alle** Studierenden seine Übungsaufgaben (Master-Niveau, Beweisführung und Coding) fehlerfrei mit KI gelöst haben – ein Prompt, 17 Minuten, perfektes Ergebnis. Die traditionelle Zulassungslogik zur Klausur (50% der Übungsaufgaben richtig) funktioniert damit nicht mehr.

☐☐ News-Update: Die China-Studie zu KI und Lernerfolg

Bauckhage erwähnt eine aktuelle chinesische Studie – diese ist real und wurde im Juni 2026 breit rezipiert (u. a. World Bank Blog, *Hechinger Report*, *Psychology Today*):

- **Stichprobe:** über 26.000 chinesische Schüler:innen (Klassen 7–12), Beobachtungszeitraum: **30 Monate**
- **Hausaufgaben mit KI-Nutzung:** Punktzahl **+18%**, Bearbeitungszeit **-30%**
- **Monatliche Klausuren (ohne KI-Zugriff):** Ergebnisse **signifikant schlechter** als vor der KI-Nutzung

Die Studie gilt inzwischen als eine Art "Warnschuss" für die Bildungspolitik weltweit – exakt das Muster, das Bauckhage aus eigener Lehrerfahrung beschreibt: Kurzfristige Produktivität bei gleichzeitigem Verlust an verinnerlichtem Wissen. Sein Fazit: Klausuren als reines Kontrollinstrument sind rechtlich (Nachweis der Prüfungsleistung) noch nötig, pädagogisch aber in der bisherigen Form überholt.

Sein Rat für die Zukunft der Bildung: Nicht mehr Tiefenexpertise bis "zum letzten Epsilon" ist entscheidend, sondern **Breite** – die Fähigkeit, Querverbindungen, Analogien und Assoziationen herzustellen, um zu erkennen, wann eine KI-Antwort plausibel *aussieht*, aber tatsächlich falsch oder nur ein "Internet-Bias" ist (sein Beispiel: 97,5% der Weberklärungen zu einem stochastischen Matrizenbeweis waren unvollständig – auch Gemini reproduzierte zunächst den falschen Spezialfall).

10. Frauen in der KI: Ein strukturelles, kein biologisches Problem

Auf die Frage nach dem Frauenanteil in der KI-Forschung schätzt Bauckhage den aktuellen Anteil auf "mindestens ein Viertel" – mit deutlichem Ausbaupotenzial (beim AlphaFold-Team laut seiner Einschätzung ca. ein Drittel). Er verortet die Ursache klar in der **frühen Kindheit** (vor dem Kindergartenalter), in der Jungen häufiger technikorientiertes Spielzeug erhalten – ein sozialisierter, kein biologisch determinierter Unterschied.

11. Quantencomputing und Hopfield-Netze

Zum Abschluss des fachlichen Teils geht es um Bauckhages Buch zu **Quantum Machine Learning**, das die Brücke zwischen rekurrenten neuronalen Netzen (**Hopfield-Netzen**) und physikalischen Systemen (Ising-Modelle, Energiefunktionen) für ein Machine-Learning-Publikum ohne Physik-Hintergrund erklärt.

Faktencheck: John Hopfield erhielt 2024 – gemeinsam mit Geoffrey Hinton – den **Nobelpreis für Physik**, nicht für Chemie (das war AlphaFold/Hassabis/Jumper). Beide Nobelpreise 2024 gingen damit an KI-Pioniere – ein historisches Novum, das Bauckhage zurecht als Beleg für die

Fazit: Zwischen Faszination und Alarmismus

Das Gespräch mit Christian Bauckhage lässt sich in einem Satz zusammenfassen: **Die Kurve, die 2016 schon absehbar war, hat sich bislang exakt so weiterentwickelt wie prognostiziert – nur ist die Bedeutung dessen erst jetzt, 2026, wirklich spürbar geworden.** Ob AlphaFold, Mathematik-Olympiaden, die Erdős-Vermutung, METR-Zeithorizonte, Qwen-Robot Suite oder die Kung-Fu-Roboter aus Peking – all diese Ereignisse aus den letzten Monaten bestätigen die zentrale These: Wir befinden uns nicht am Ende, sondern am **Anfang** einer exponentiellen Entwicklung, bei der "der nächste Schritt immer größer ist als alle vorherigen zusammen".

Bauckhages praktischer Rat – sowohl für Mittelständler als auch für Studierende – lautet entsprechend nüchtern:

1. **Identifiziert euren Painpoint**, bevor ihr KI einsetzt – KI um der KI willen ist sinnlos.
2. **Schafft euch Zugang zu Expertise** (z. B. über Programme wie KI.NRW oder Kompetenzzentren wie das Lamarr-Institut), nicht zwingend eigene KI-Abteilungen.
3. **Setzt auf spezialisierte, energieeffiziente Modelle** statt auf reine Skalierung, wo immer es geht.
4. **Bildet Breite statt nur Tiefe aus** – die Fähigkeit, KI-Ergebnisse kritisch einzuordnen, wird wichtiger als reines Fachwissen.
5. **Get your boat ready – the storm is coming.**

Hinweis: Alle in diesem Artikel als "News-Update" markierten Abschnitte basieren auf zusätzlicher Recherche über die im Gespräch erwähnten Themen hinaus und spiegeln den Stand der öffentlich verfügbaren Berichterstattung bis Juli 2026 wider.

Revision #1

Created 2026-07-13 15:00:52 UTC by art10m

Updated 2026-07-13 15:01:28 UTC by art10m