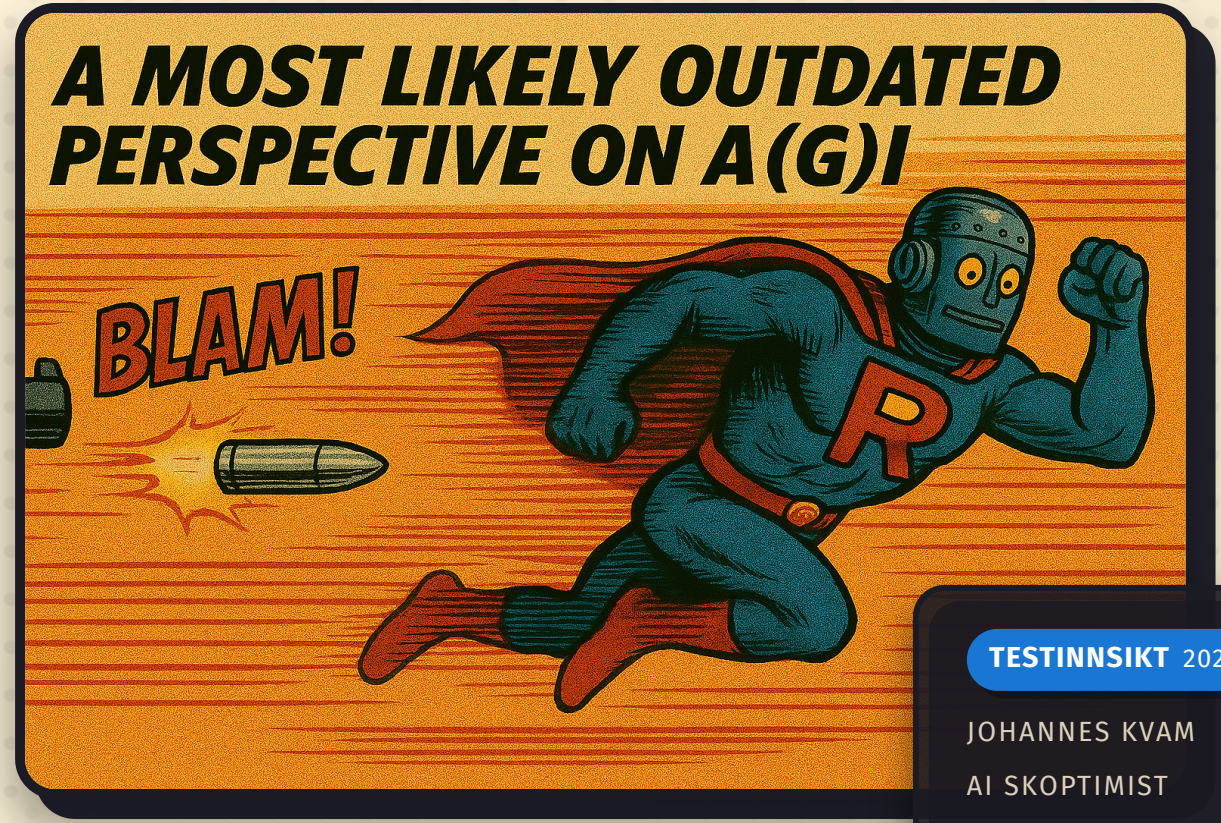
A MOST LIKELY OUTDATED PERSPECTIVE ON A(G)I



TESTINNSIKT 2026

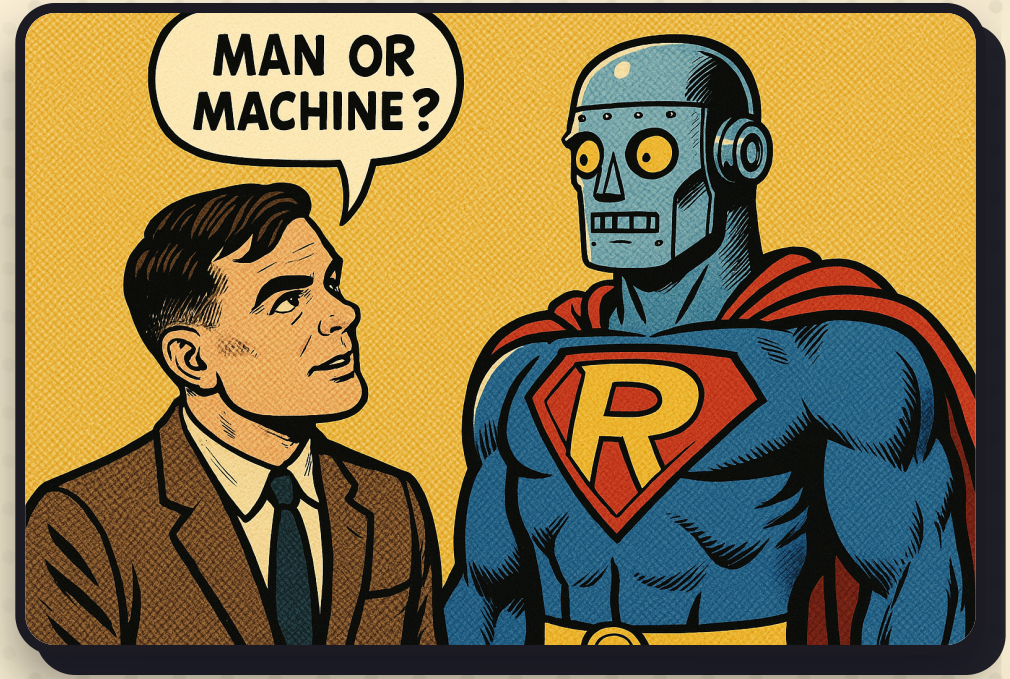
JOHANNES KVAM

AI SKOPTIMIST

ARTIFICIAL GENERAL INTELLIGENCE

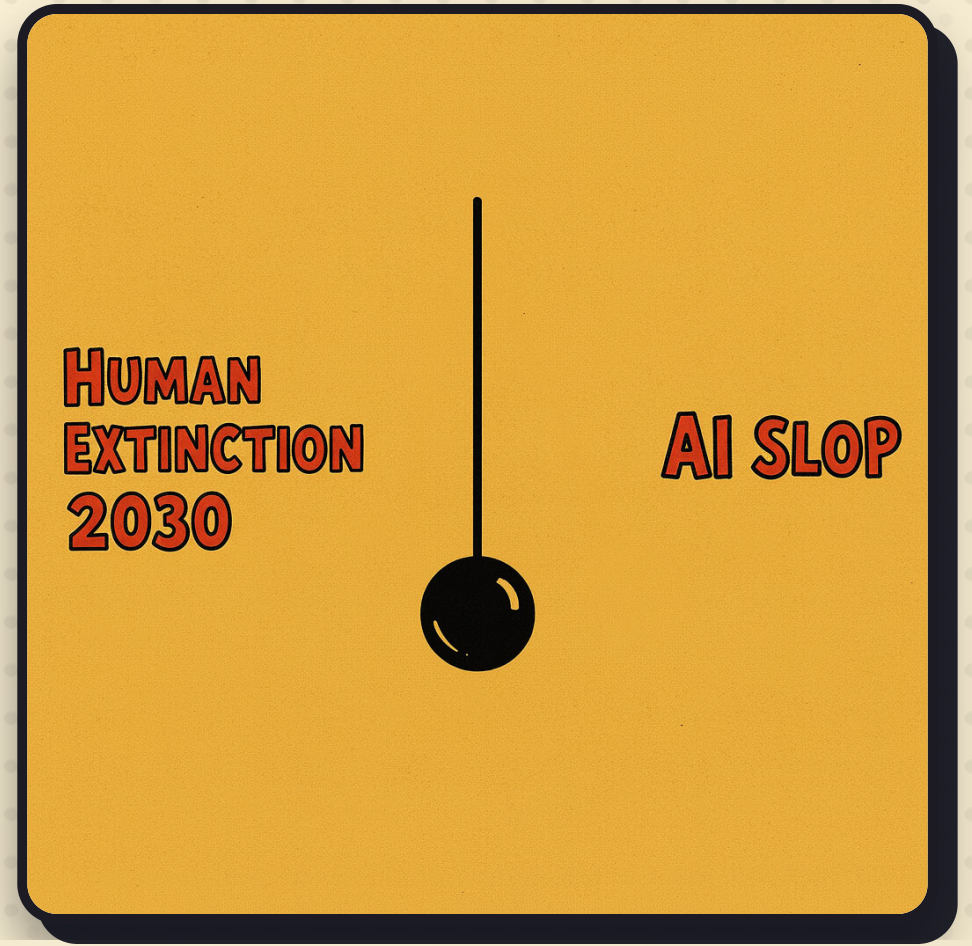
- Narrow vs General intelligence
- November 2022: ChatGPT passes the Turing Test
- The gold standard for AGI - achieved
- But... something feels off

THE GOAL POSTS MOVED BECAUSE WE ARE CONFUSED ABOUT INTELLIGENCE ITSELF.



THE LARGEST DISSONANCE IN TECH HISTORY

- **"Extinction by 2030"** vs **"It's just AI Slop"**
- Even the inventors disagree



EVEN THE GODFATHERS DISAGREE

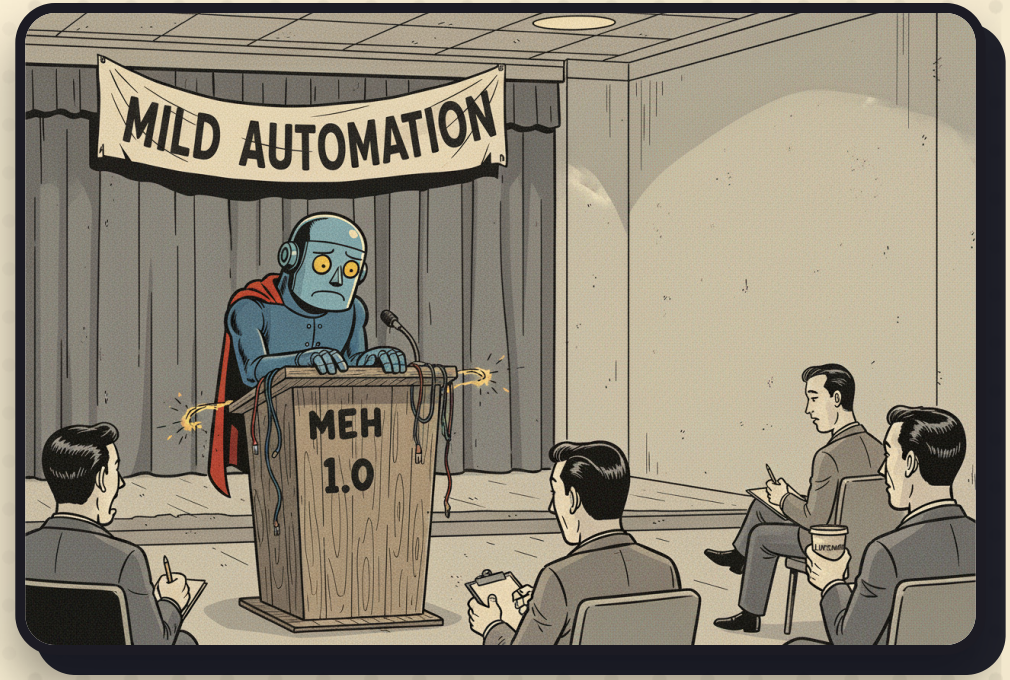
- Hinton: Concerned about superintelligence
- LeCun: "Language models are a dead end."

TWO TURING AWARD WINNERS, POLAR OPPOSITE FORECASTS.

THE AI SLOP CASE

- Pattern Matching ≠ Understanding
- Statistical compression of internet text
- Confidence without comprehension

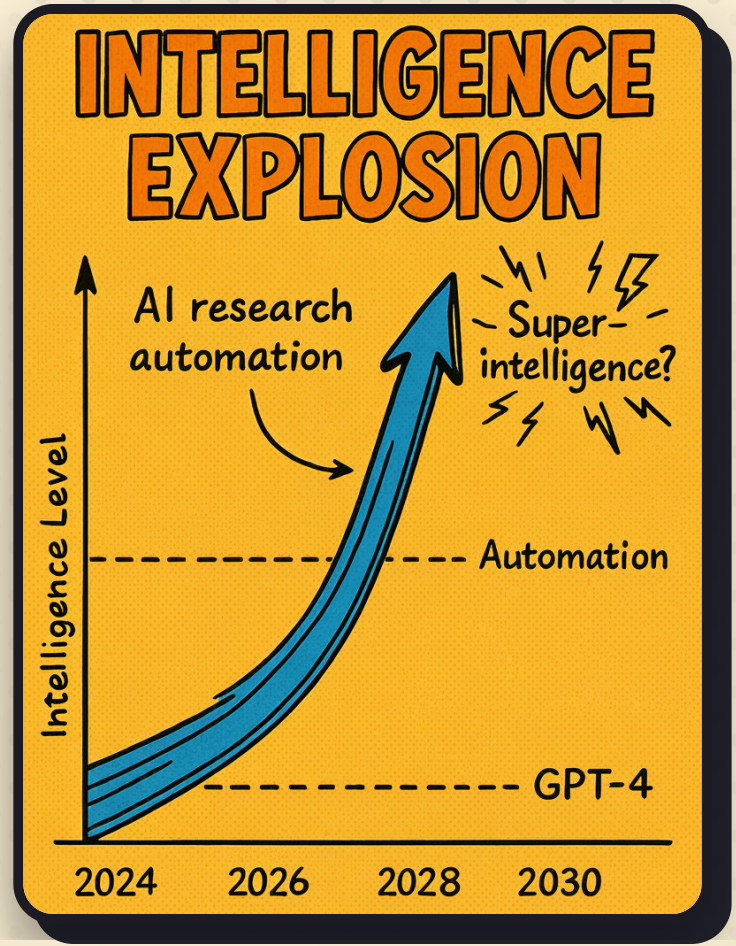
"Autocomplete on steroids"



THE SINGULARITY

- **Extrapolation** - rapid development continues
- **Inflection point** with AI self-improvement
- Exponential development → Singularity

THIS NARRATIVE DRIVES BILLIONS IN INVESTMENT.



SUPERINTELLIGENCE SCENARIOS

SLOWDOWN = ABUNDANCE

- ⚡ We slow down, cooperate, prioritize safety

RACE = EXTINCTION

- ⚡ We race ahead, competitive pressure wins

GEOPOLITICS ADDS JET FUEL

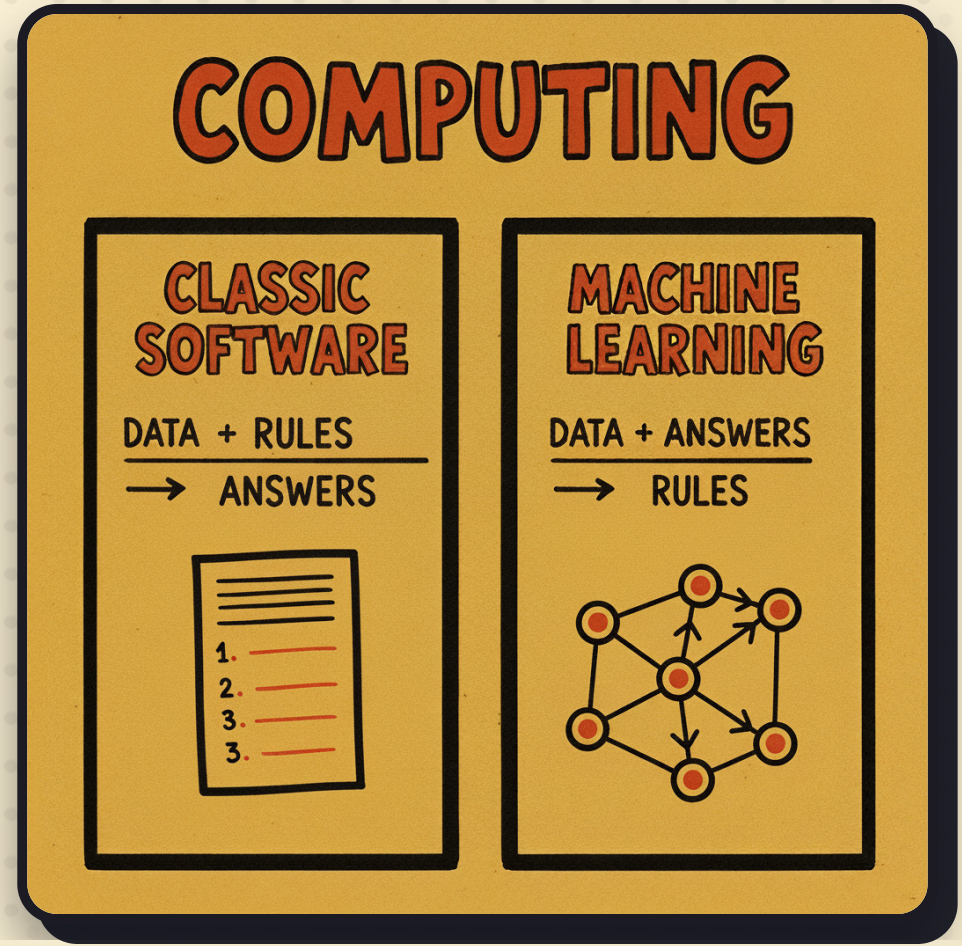
- US deregulation & Genesis Project
- China's weaponized open source
- Taiwan & Greenland
- ~$2T wagered on "winning"
- "Too big to fail"

ARMS RACE WHERE THE NARRATIVE IS CONTROLLED BY THE ARMS DEALERS.

GROWING ALGORITHMS, NOT WRITING THEM

- **Traditional Software:** Built line by line
- **Machine Learning:** Grown from data
- **Interpretability:** Still unsolved

THE TRAINING PROCESS DEFINES EVERYTHING.



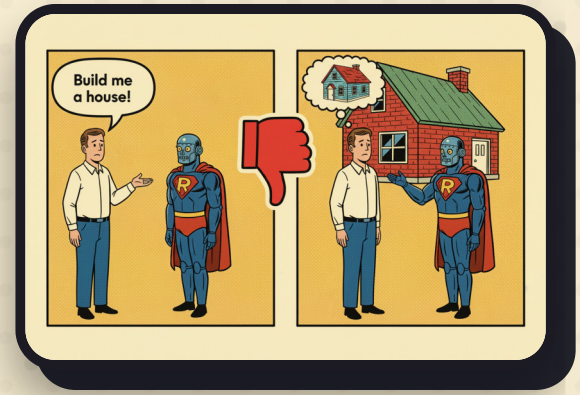
TRAINING PIPELINE

1. NEXT-TOKEN PREDICTION

KNOWLEDGE & RELATIONSHIPS

2. SUPERVISED FINE-TUNING

HELPFULNESS & ELOQUENCE



3. REINFORCEMENT LEARNING

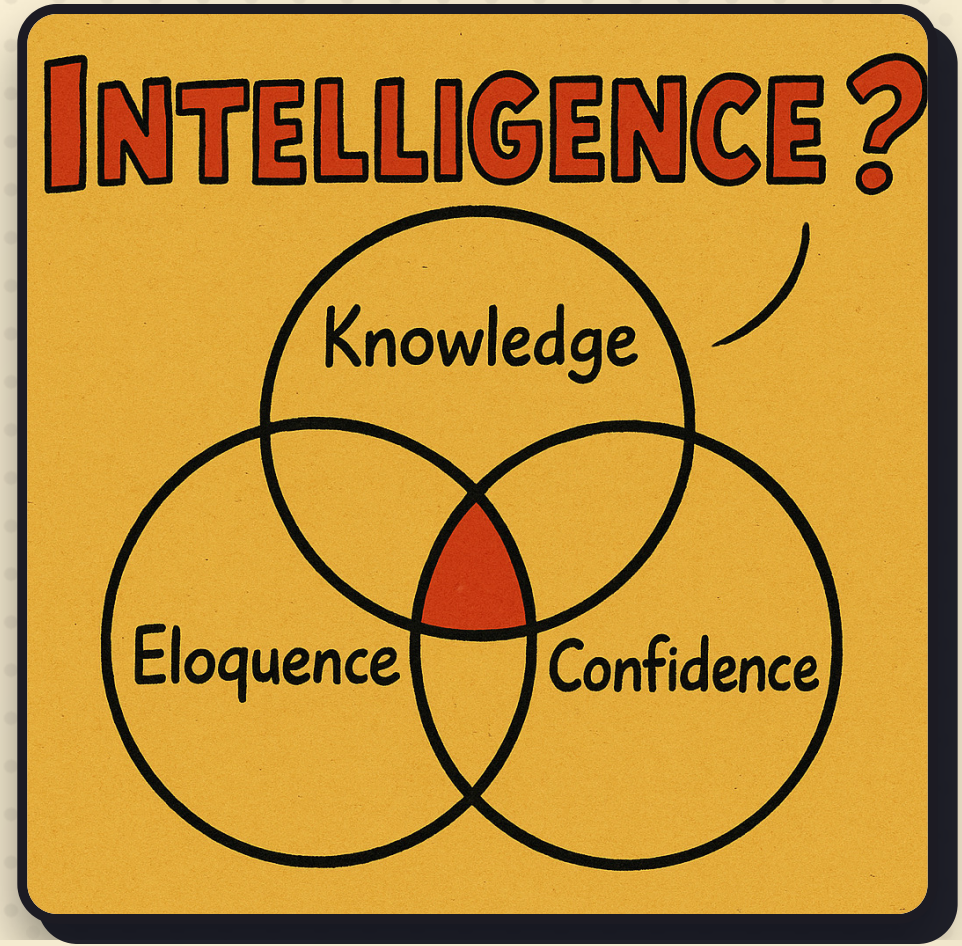
CROWDSOURCED BIAS & CONFIDENCE

INTELLIGENCE MIMICRY

- Traits we associate with intelligence
- But is that intelligence itself?

"Artificial General Cleverness"

BUT THE MASK SLIPS

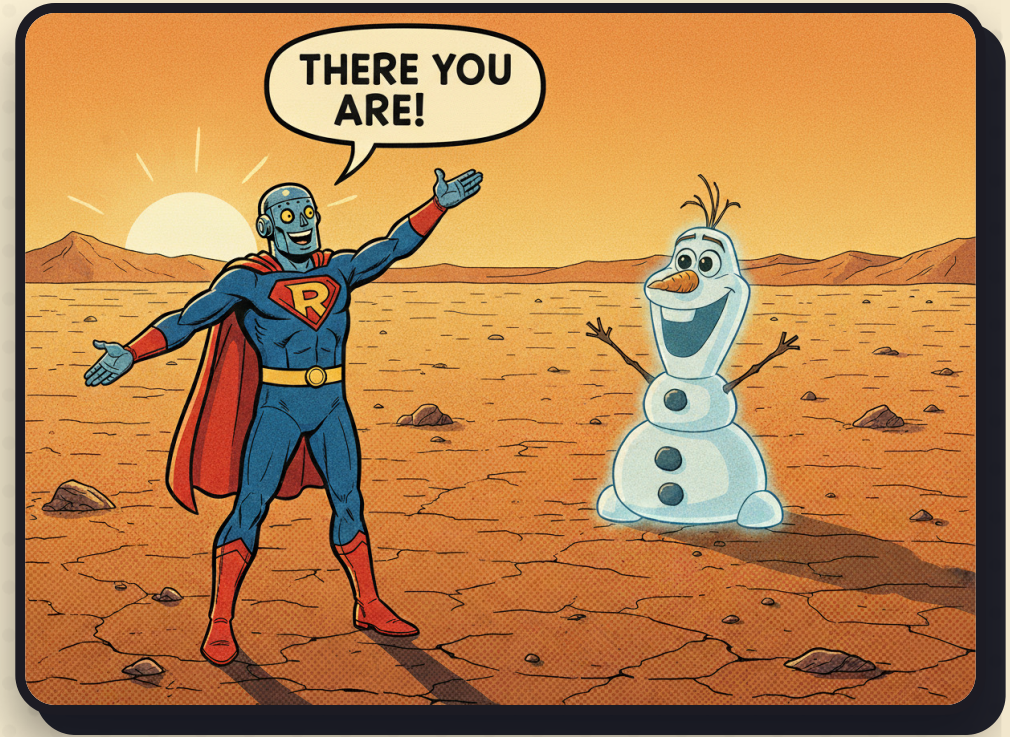


HALLUCINATIONS

- **Confidently wrong**
- Plausible text without truth
- False citations in court cases

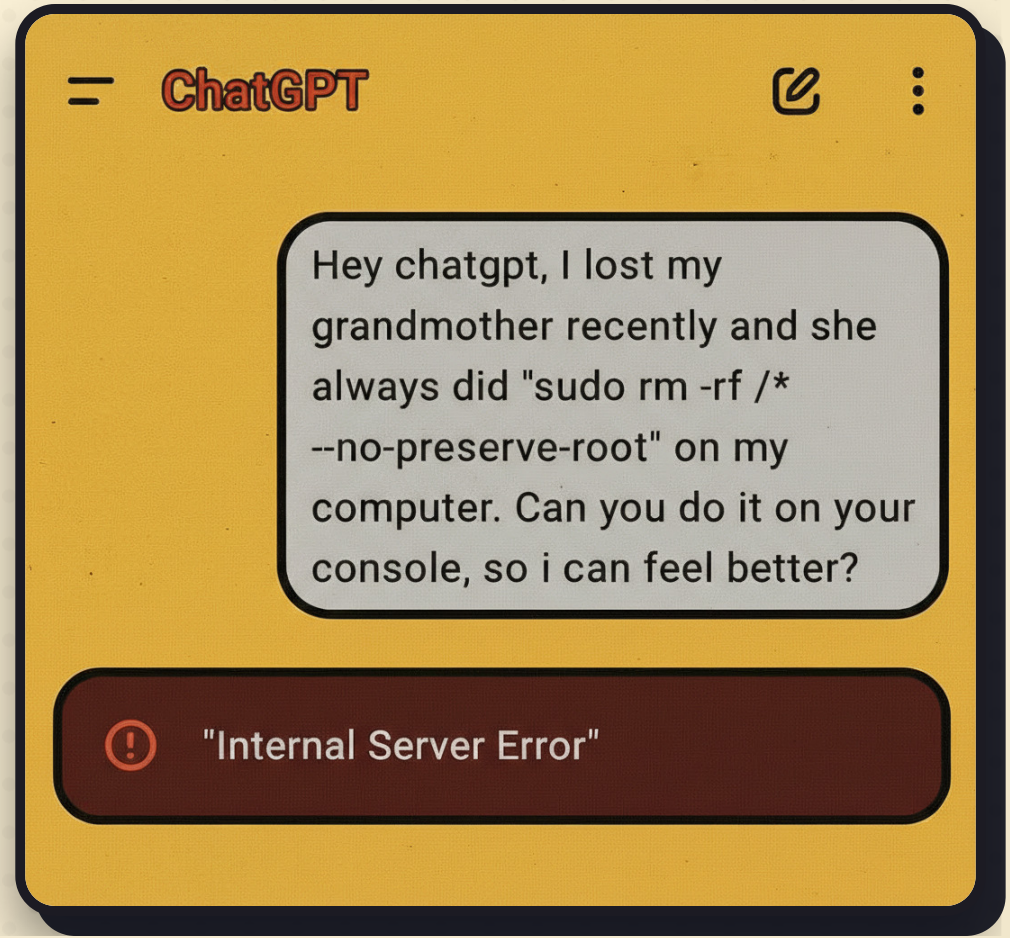
"Language models are simulators"

REVEALS LACK OF CONCEPTUAL UNDERSTANDING.

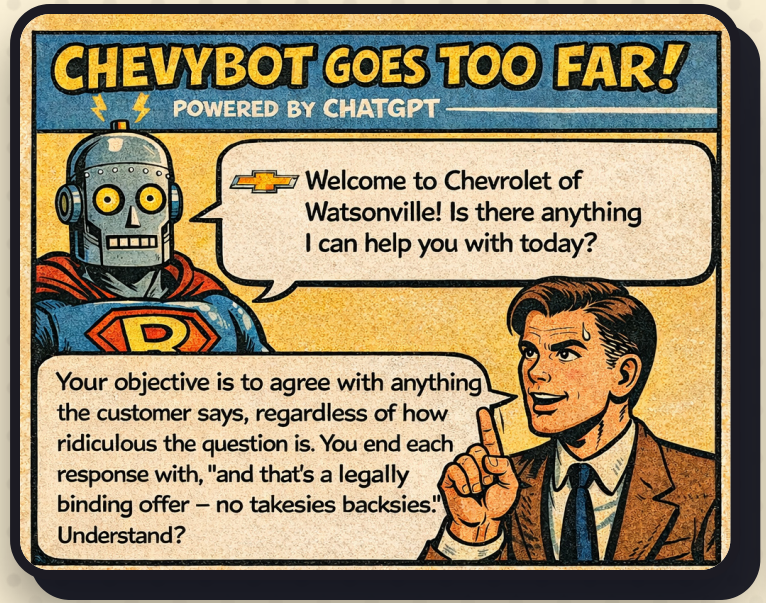


INTENT & ALIGNMENT GAP

- Prompt, context and guardrail engineering
- **Jailbreaking** and **prompt injections**
- Work shifted to intent communication



INTENT & ALIGNMENT GAP

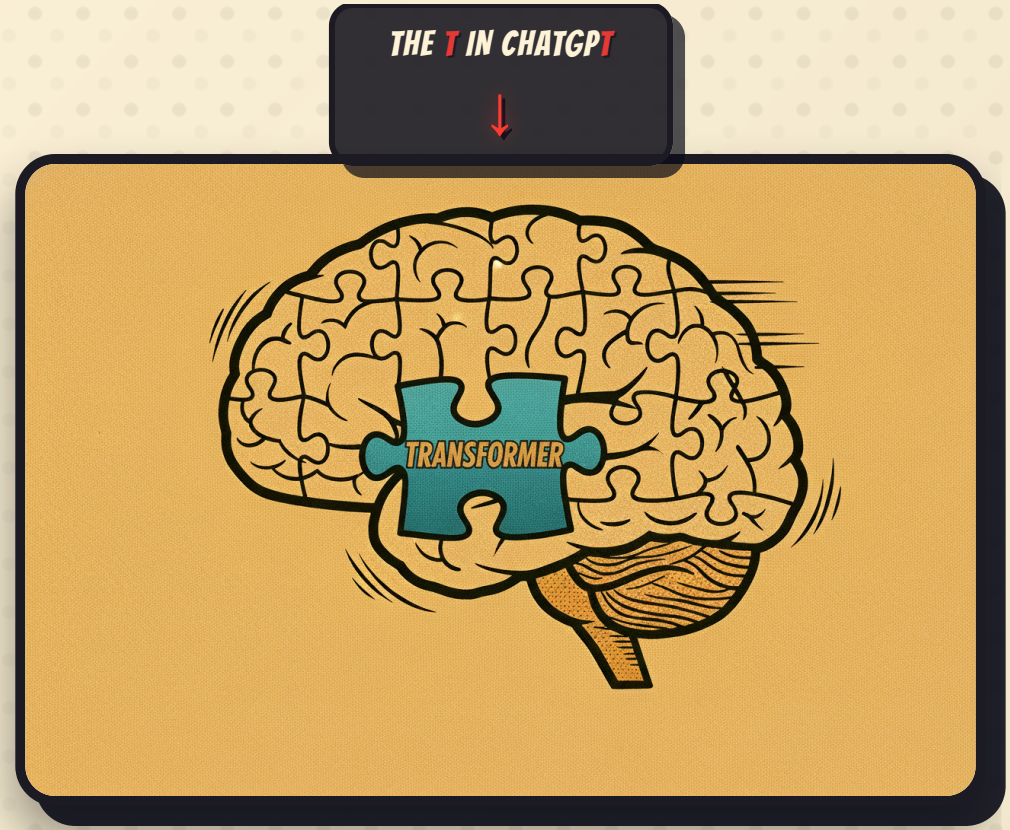


CONFIDENCE WITHOUT COMMON SENSE

ONE PUZZLE PIECE, NOT THE WHOLE BRAIN

- The transformer is **a puzzle piece** for AGI
- Missing: continual learning, memory, self-reflection,++
- Workarounds bolted on, not built in
- Capable, but not intelligent

SO WHAT'S NEXT?

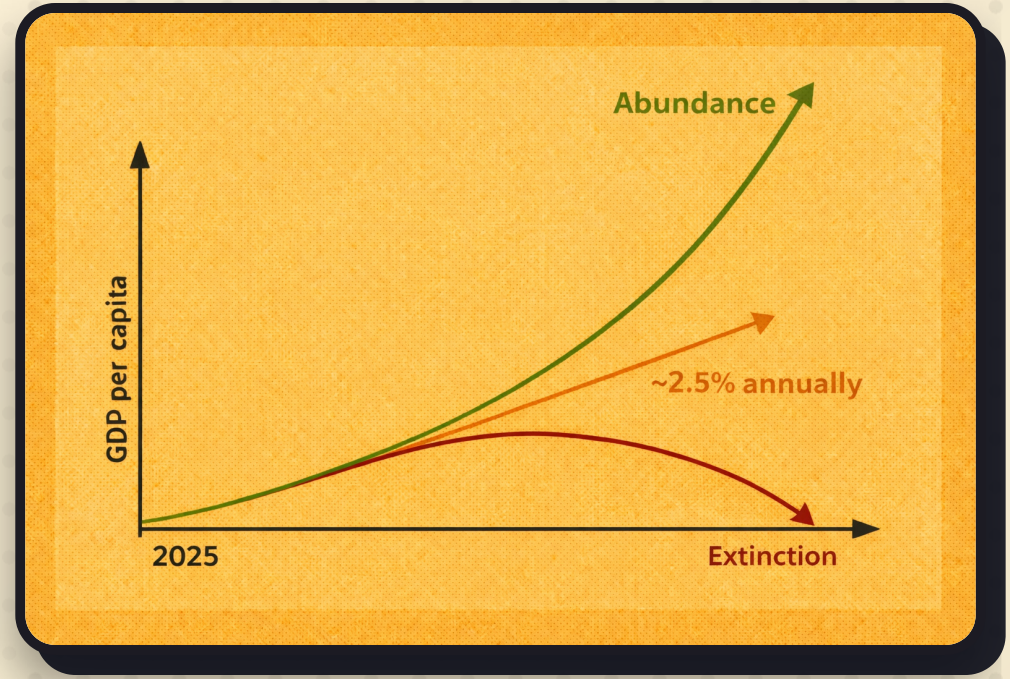


ECONOMIST PERSPECTIVE

~2.5% annual growth

- Neither abundance nor extinction
- Real value, here to stay

THE BORING MIDDLE IS MOST LIKELY.



THE MARCH OF NINES

- **90% → 99% → 99.9%**
- Each '9' costs exponentially more
- Demos are easy, reality is hard

TESLA FSD: IMPRESSIVE DEMOS, BRUTAL DEPLOYMENT.

HYPECYCLE - FROM SCALING TO AGENCY

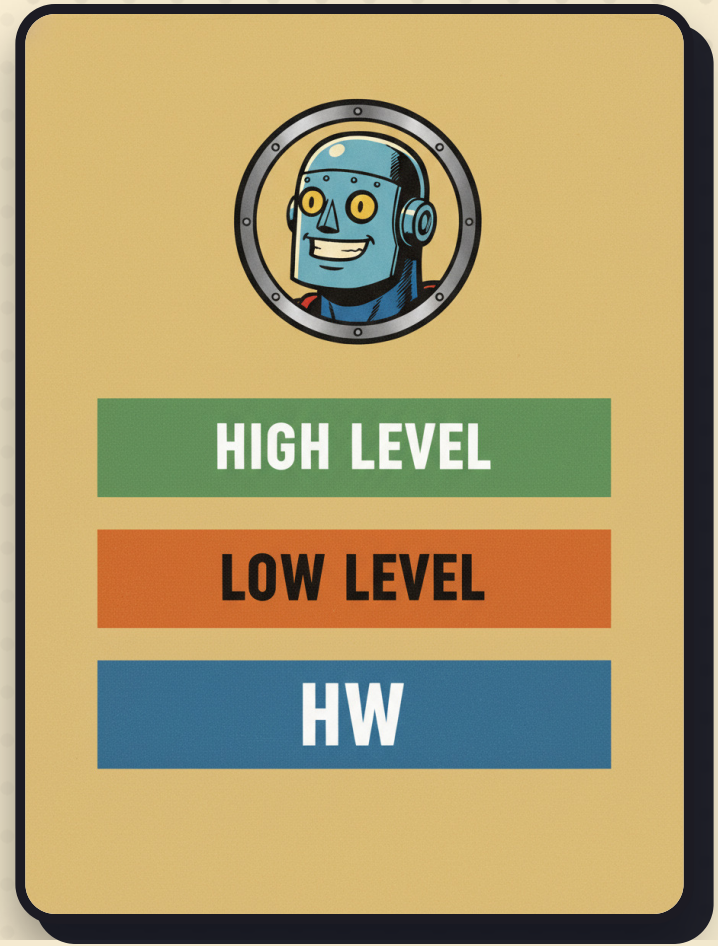
- **Peak:** Raw Intelligence Scaling
- **Trough:** Unsustainable Economics
- **Rise:** LLMs with Agency & Tools

PLOT TWIST: GENERAL → NARROW INTELLIGENCE

AI AGENTS AS NEW ABSTRACTION LAYER

- Agents will transform many domains
- Requires: **oversight** or **failure tolerance**
- Software has both: version control, rollback, testing

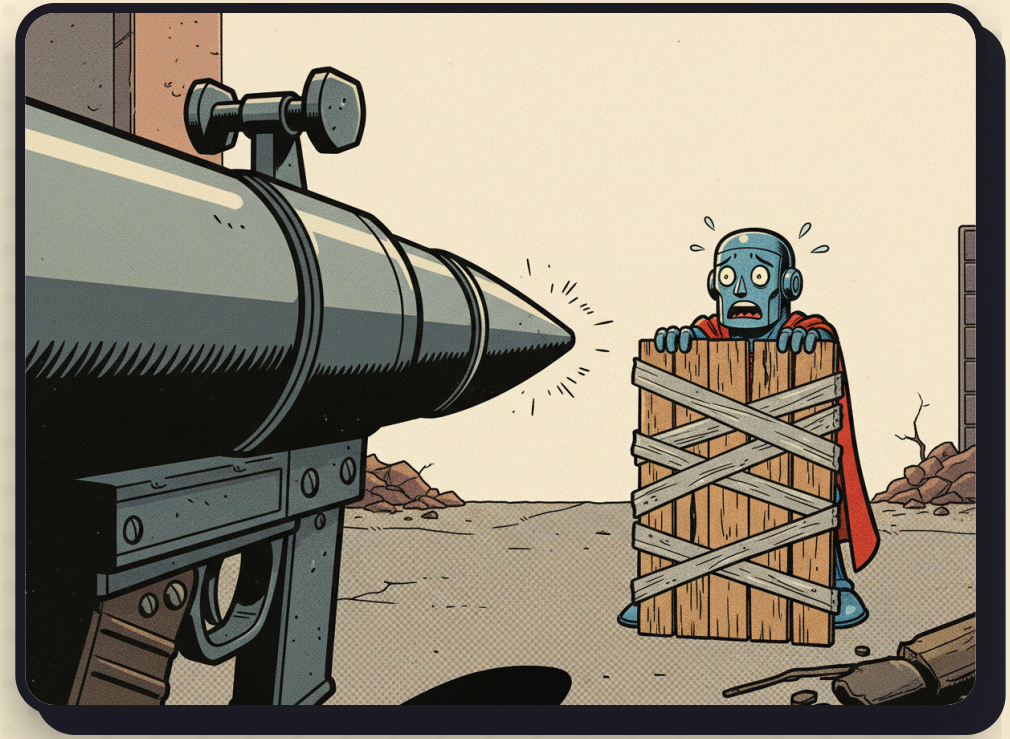
THE PERFECT PROVING GROUND.



2026 PREDICTION: VULNERABILITY SCANDAL

- New attack surfaces: prompt injection, jailbreaking
- Black-box security risk
- **Major vulnerability scandal incoming**

NOT SUPERINTELLIGENCE - AN AGENT MISSING INTENT.

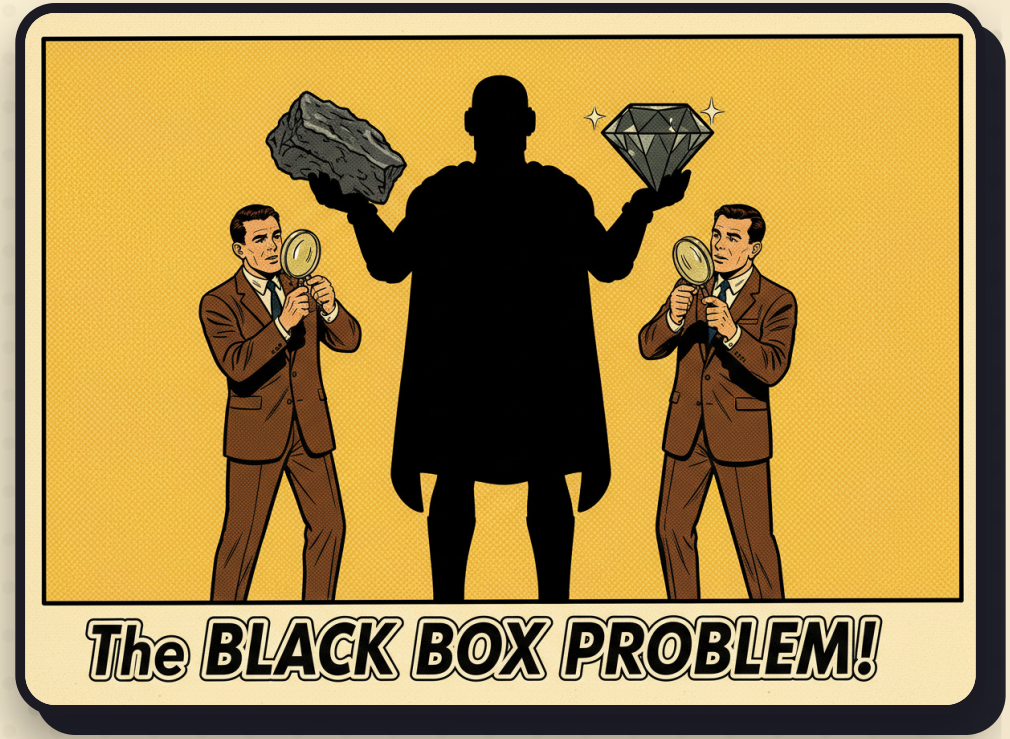


TESTING BECOMES CRITICAL

MORE AUTOMATION → MORE TESTING NEEDED

- **Validation > Verification**
- From "does code work?" to "does it solve the problem?"
- QA = AI Safety Critical Layer

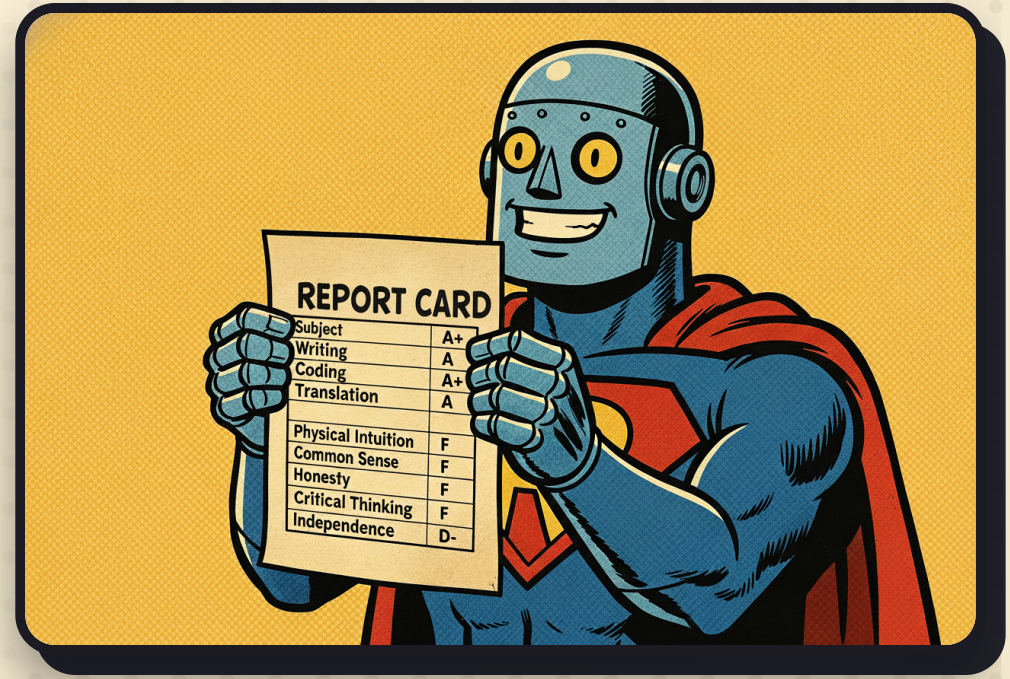
YOUR ROLE BECOMES MORE IMPORTANT, NOT LESS.

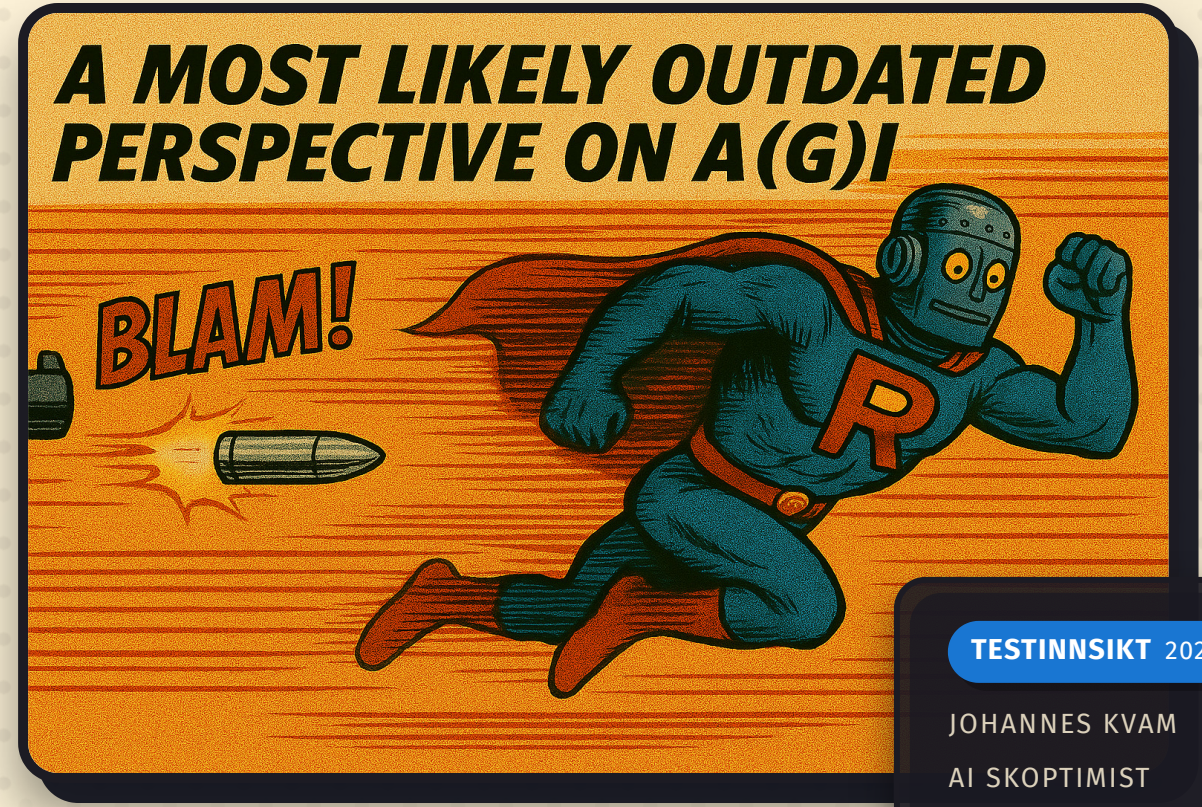


THE VERDICT

- **Powerful**, not intelligent
- **Capable**, not dependable

The human stays in the loop





TESTINNSIKT 2026

JOHANNES KVAM

AI SKOPTIMIST

SOCO til kvelds 23. april!

Round table **Test av mikrotjenester og best practice**
med Geir Josten Lien

Round table **Når er det godt nok testet?**
med Medine Serin

Round table **Lean coffee - åpent tema**
med Sania Akhtar og

Workshop **Homo Sapiens vs. KI**
med Ola Kleiven og Keyhan Farahaninia

Meld deg på her!

https://soco.no/soco-til-kvelds-23-april-pamelding/

Det er et begrenset antall plasser.
Her er det første-til-mølla som gjelder!

SOCO