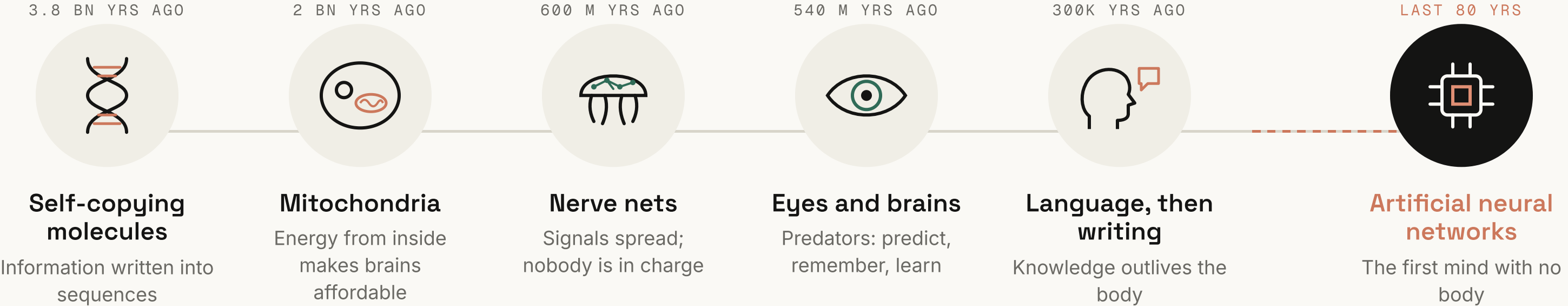


AI · PART I · 8-SLIDE SUMMARY

From biological intelligence to artificial intelligence: a brain for hire

How intelligence went from biology to silicon, and why today's large language models are best understood as a brain you can hire.

For four billion years, intelligence lived in bodies.



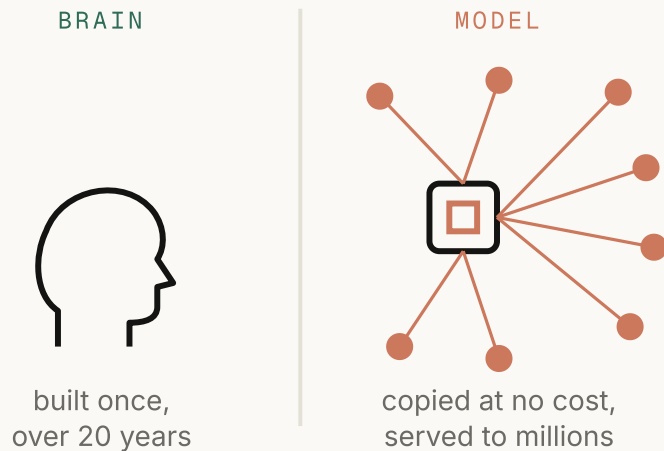
If life's history were one day



Where artificial and biological intelligence part ways.

01

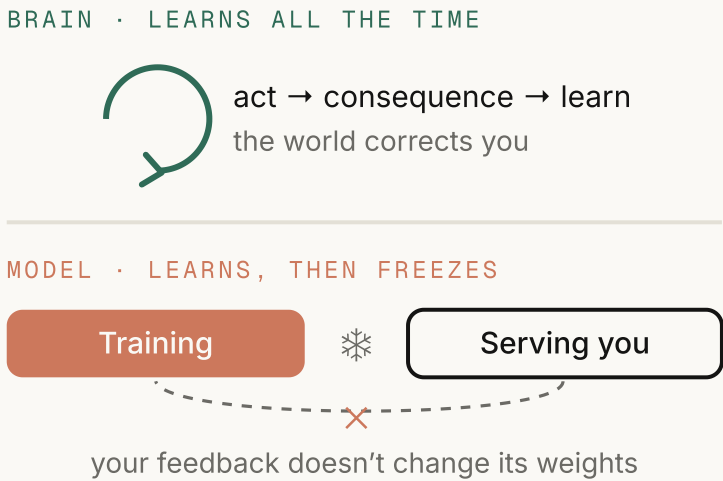
It can be copied



Every human mind is built from scratch. A model is copied at no cost, and APIs make it a brain for hire for millions at once.

02

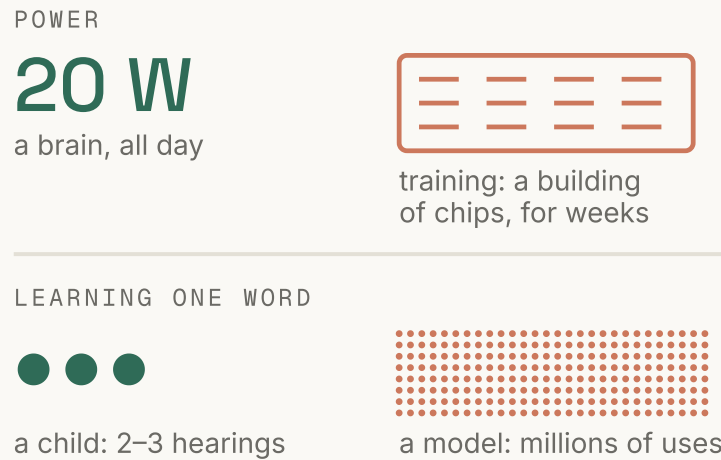
It learns, then freezes



We learn to survive, from consequences. A model is trained to predict text, then frozen to serve.

03

It is inefficient



The brain also holds 50–100× more connections: about 100 trillion, against 1–2 trillion parameters.

Tireless, tool-using and jagged.

04

Nothing is at stake

BRAIN

hunger

fear

safety

belonging

love

status

curiosity

reproduction

↓

point it at what matters
but it tires and gets bored

MODEL

row 1

row 2

⋮

row 10,000

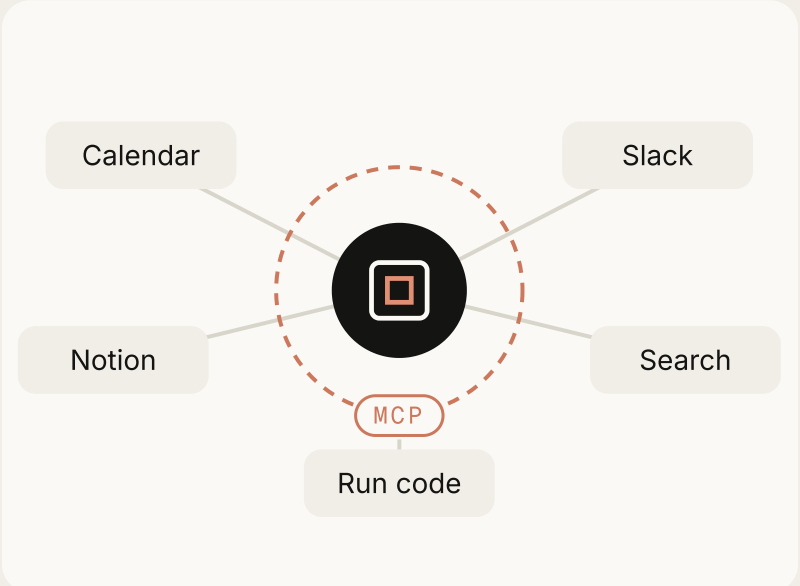
no drive at all

same attention, every row
but no sense of what matters

Tireless and indifferent are the same property: it grinds without tiring, and waits to be pointed at something.

05

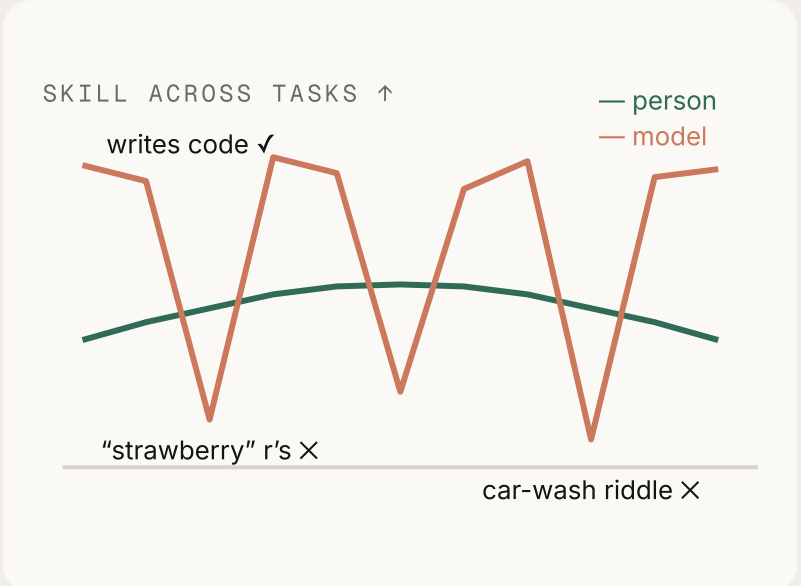
Tools are its first reach



Our tools extended a body. A model never had one; through function calling and MCP it reaches apps built for people.

06

Its intelligence is jagged



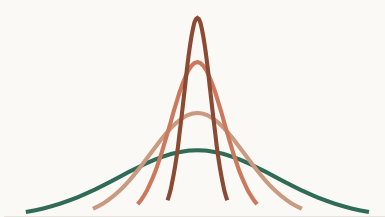
It writes code but miscounts letters. One skill says little about another, so the only way to know is to test it.

Where AI goes next: how it will learn.

DATA

Model collapse

TRAINED ON ITS OWN OUTPUT



wide: human data narrow: gen 4

Text is running out. Sloppy feedback into training narrows each generation.

CONTINUAL LEARNING

No more freeze



outside: retrieval · search · tools

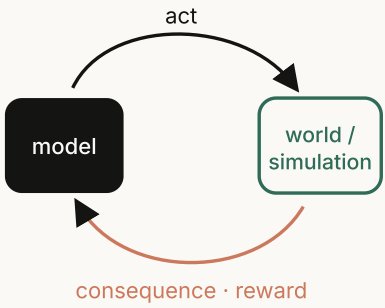
weights frozen *

what it learns later never reaches the weights

Learning after deployment would change what models are.

RL & EMBODIMENT

Real-world feedback



act

model

world / simulation

consequence · reward

Checkable tasks, then bodies: consequences brought into training.

EFFICIENCY

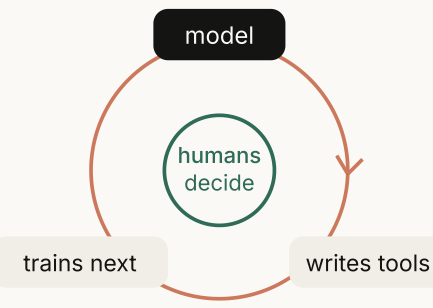
Learn from biology

- event-driven not always on
- active learning not passive
- two memories short + long
- sleep-like consolidates

Training and serving are costly; the brain hints at cheaper ways.

RECURSIVE INTELLIGENCE

Models shaping models



model

humans decide

trains next

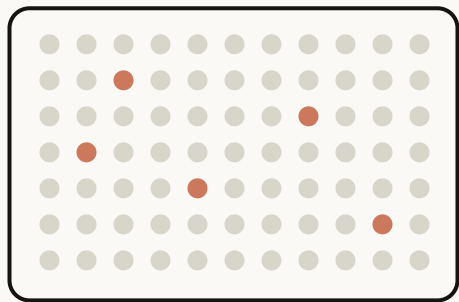
writes tools

Models build their own tools and help train the next ones.

And whether we can see inside it, and trust it.

INTERPRETABILITY

Reading the inside



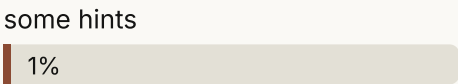
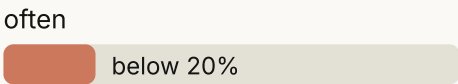
• the few parts we can read

We built something whose internals we can't yet read. A breakthrough would change how models are trained and fixed.

FAITHFULNESS

Hidden reasoning

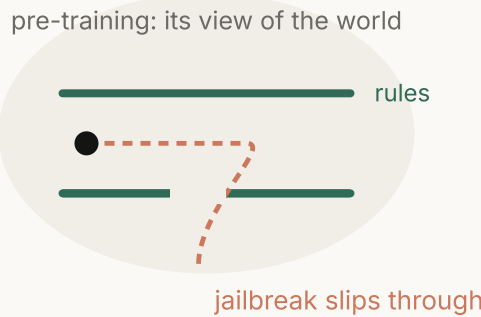
ADMITTED USING THE HINT



Given a hint, models often didn't say they used it. The reasoning you read isn't always the real one.

ALIGNMENT & SAFETY

Behaving vs being aligned

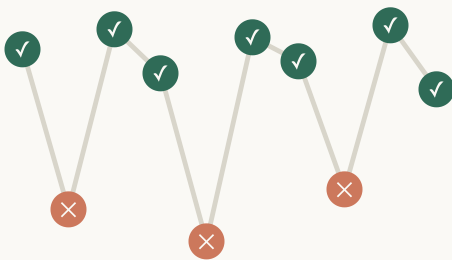


Post-training steers behaviour, but jailbreaks still work. Safety has to keep pace with capability.

EVALUATION

Measure to trust

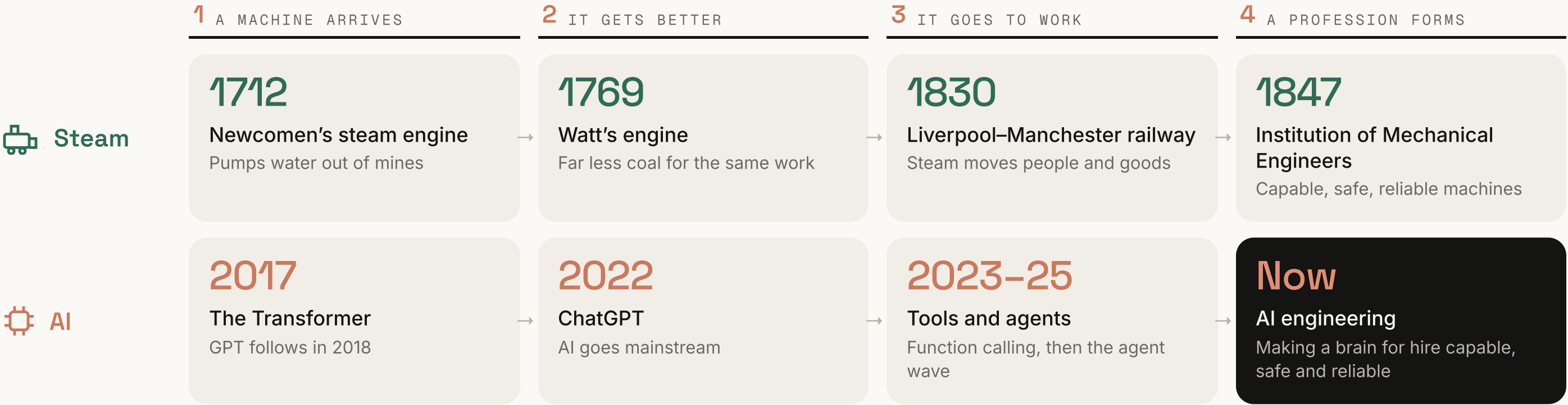
TEST EVERY SKILL



one pass says little about the next

With jagged skills, testing is the only way to know, and everything else on this list depends on it.

The steam engine created mechanical engineering. AI is creating AI engineering.



135 years from engine to profession. For AI, less than a decade. The machine is only the start. The work is everything around it: when to hire it, how to judge it, and the harness it runs in.

THE SHORT VERSION

In four billion years, evolution never separated a mind from its body. AI is the first: a brain you can copy, rent by the hour, and must learn to judge and engineer, not just use.

READ huyentran.tech/writings/a-brain-for-hire
