

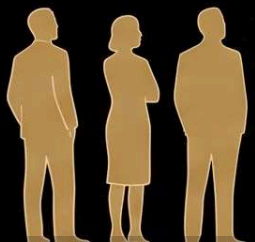


LLMs Unplugged

Does 'Unplugged' still matter in the age of AI?

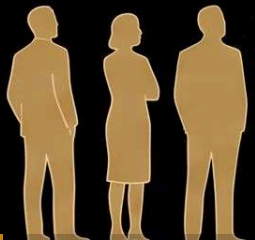
Everyday life is increasingly regulated by complex technologies that most people neither understand nor believe they can do much to influence.

— Albert Bandura, 2001



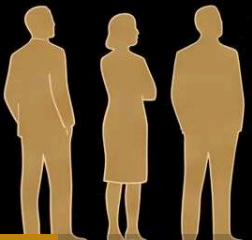
Everyday life is increasingly regulated by complex technologies that most people neither understand nor believe they can do much to influence.

— Albert Bandura, 2001



Everyday life is increasingly regulated by complex technologies that most people neither understand nor believe they can do much to influence.

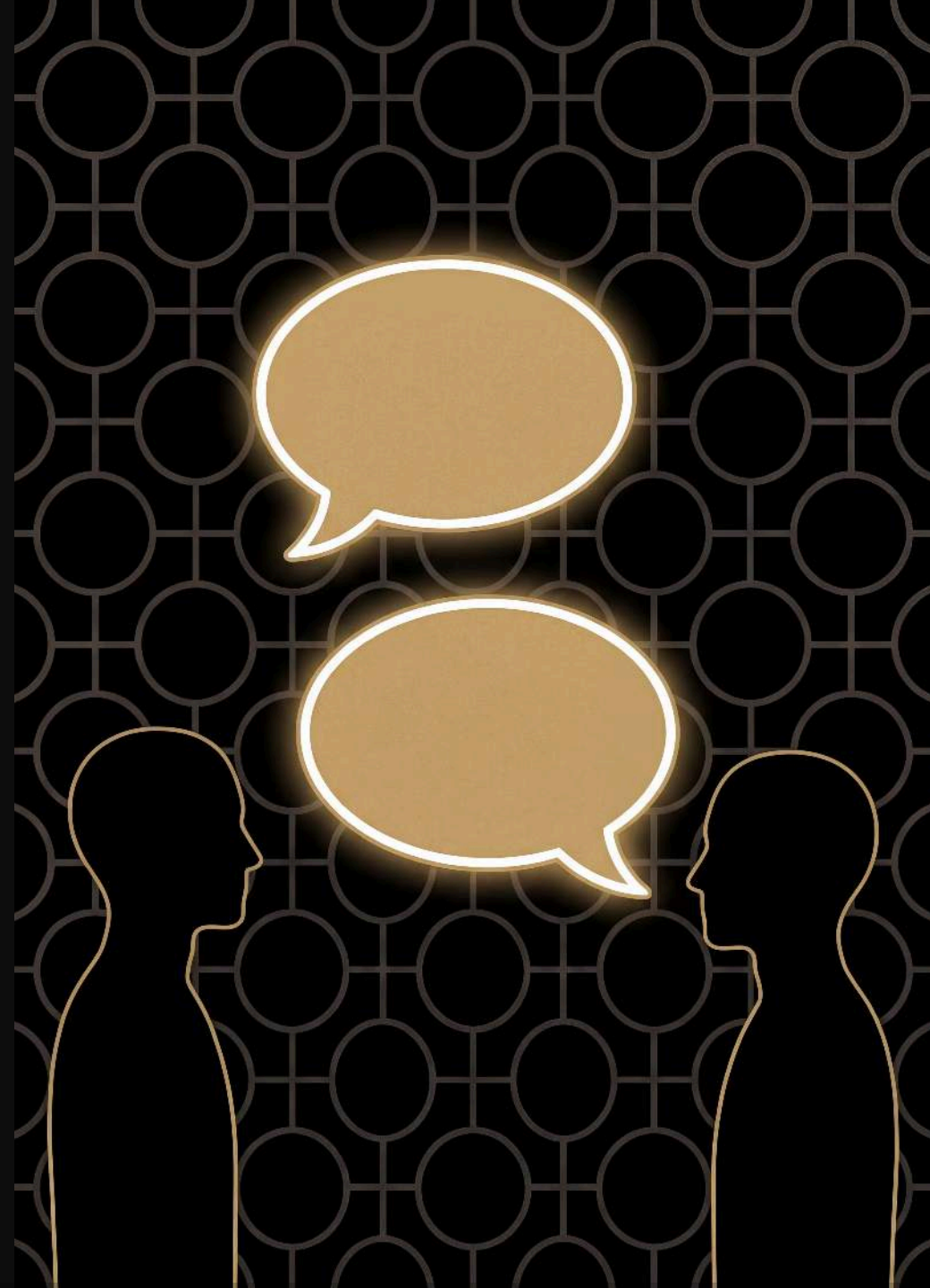
— Albert Bandura, 2001

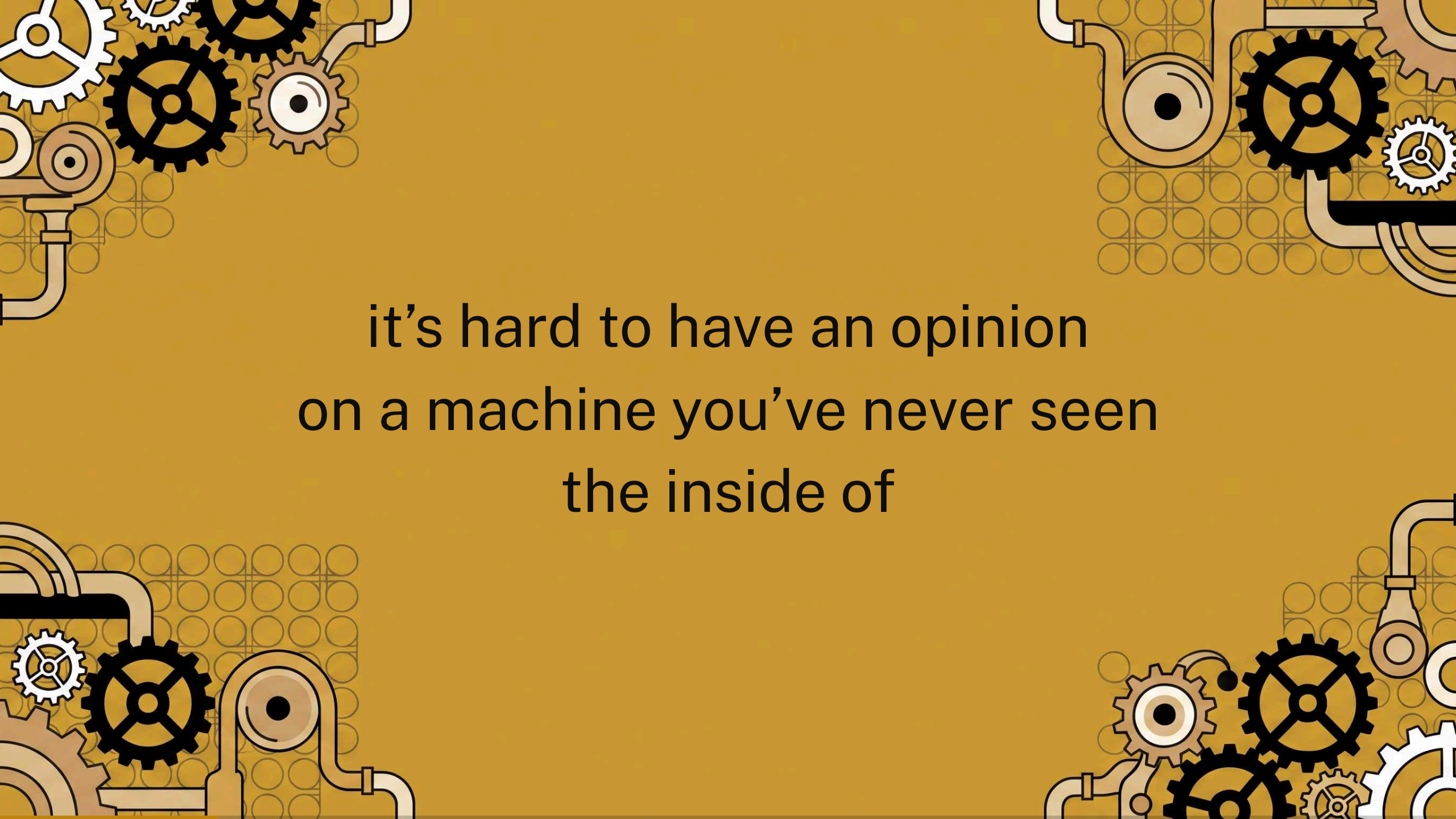


what we get told about AI

it **understands** you — it's basically a mind

it's **just autocomplete** — it's basically a parlour
trick



The background is a solid mustard yellow color. It is decorated with stylized, hand-drawn mechanical elements in the corners. These include various sizes of black and white gears, tan-colored pipes, and pulleys. Some of the gears have a simple cross-hatch pattern on their faces. The mechanical elements are arranged in a way that suggests a complex, interconnected system, typical of steampunk or industrial design aesthetics. The text is centered in the middle of the image.

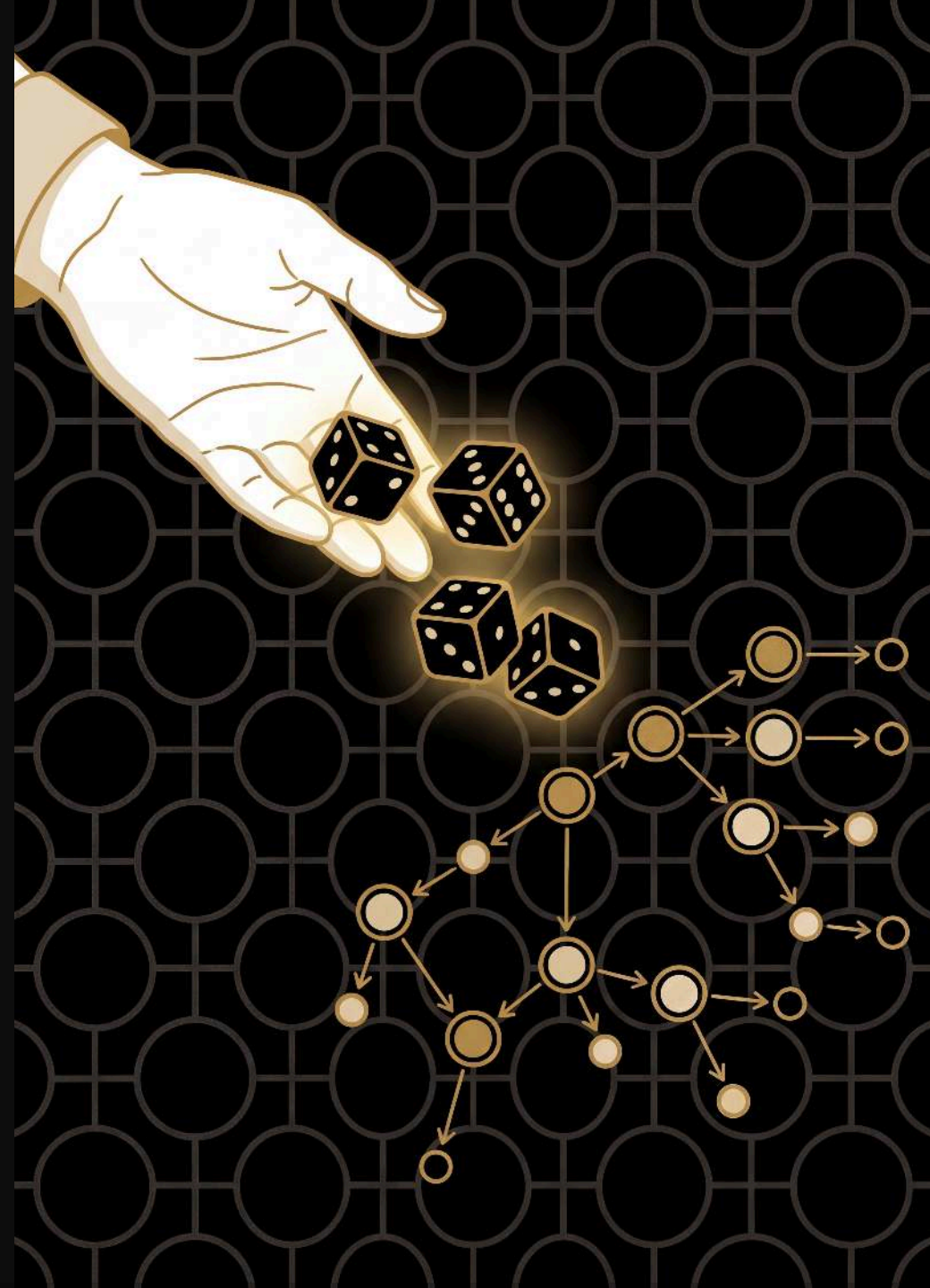
it's hard to have an opinion
on a machine you've never seen
the inside of



a page of **token pairs** —
a word, and the word before it

the algorithm

1. I **call out** the last token we wrote (and its colour)
2. **hands up** if it's the *boxed* token on your page
3. I **pick a hand at random** — you read out the token beside it
4. I **write it down**, and we go again





4:00

start

reset

+1

-1



we just ran a language model

what role did you play?

each word on your sheet → a **token**

the pairs on your page → the **parameters** of the model

the hands in the air → the **probability distribution** over next tokens

my random pick → **sampling**

the sentence I wrote → **generation**



the sentence we wrote
is not in the book



between us, we have the whole book

1,962 words · 240 distinct · 1,961 token pairs
dealt round the room

no single sheet has the whole picture



the same model in grid form

look at me ! look at me ! look at me now !

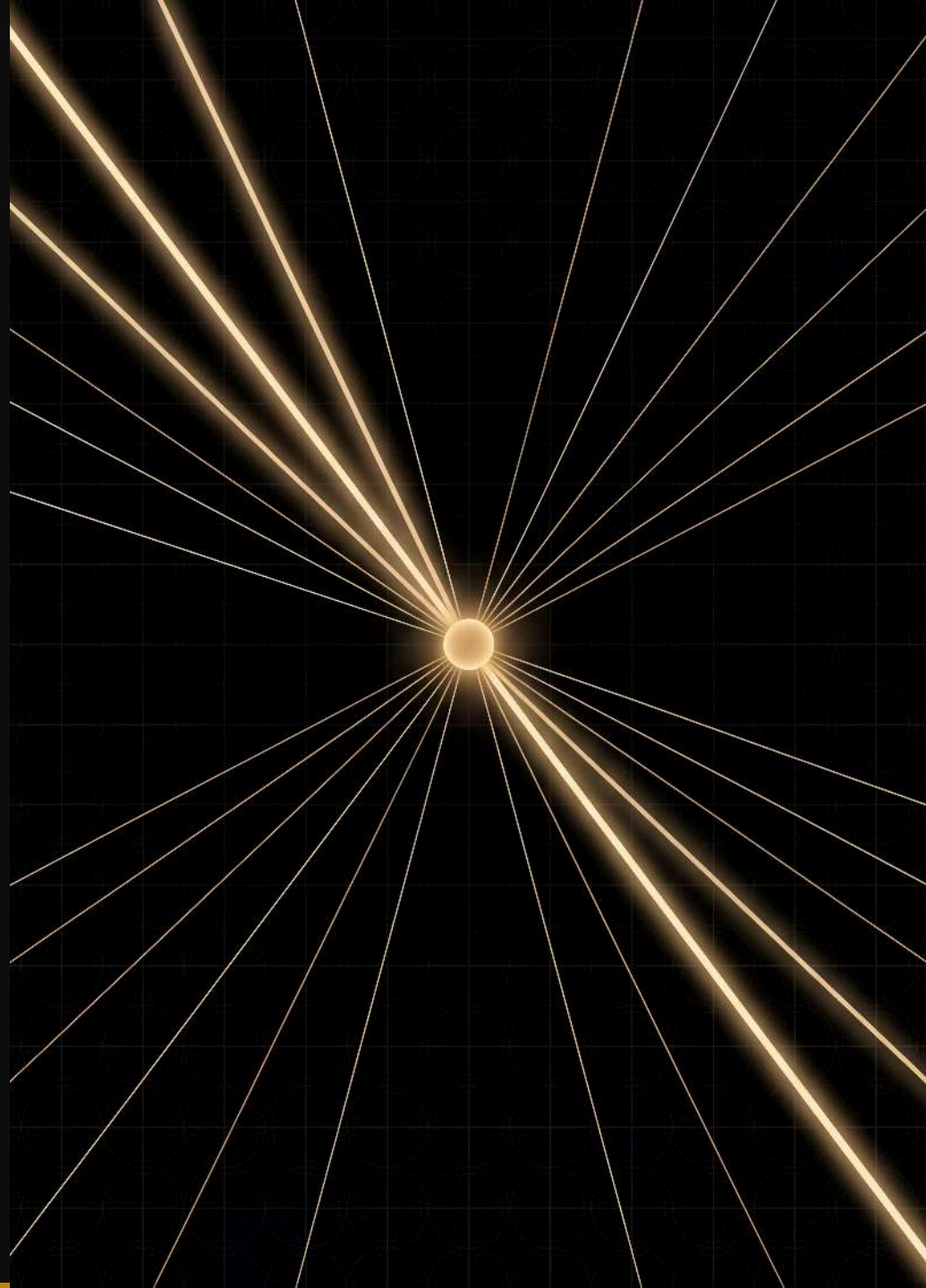
	look	at	me	!	now
look					
at					
me					
!					
now					

three things Claude* does that we didn't

it looks at **more than one word** of context

it can say what it **never saw** — **cat** teaches it
about **dog**

it's been taught to **answer**, not just continue



the line on the last slide



the whole room, as a grid

240 distinct tokens · 57,600 cells

5.8 m² of paper — roughly this stage

today

frontier LLM

context

how much it looks at before each guess



1 token

the whole conversation

training text

how much the model reads

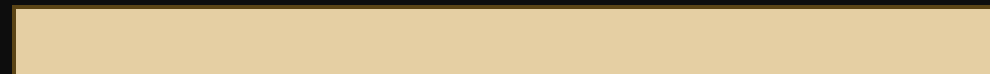


1,962 tokens

15 trillion tokens

model size

how many learned numbers it contains



57,600 grid cells

~1 trillion parameters

three independent scales · different kinds of model

much more context

much more training text

many more learned numbers

a different kind of model



still just **words in** → **words out**



back to Bandura

most people neither understand nor **believe**
they can do much to influence

a room full of people just did, slowly, the thing
you're told needs a data centre



it doesn't make you an expert
it makes you **harder to bullsh*t**





llmsunplugged.org/talks/unplugged-in-the-age-of-ai/



Australian
National
University

School of
Cybernetics

www.llmsunplugged.org