



Designing Trust: Lessons from Building AI Products for People Who Cannot Afford to be Wrong

Murali Krishna Nanda Kumar
Head of Product & Design,
Estatemin

IVI Summit 2026

20 Years of Impactful Research – What's Next?



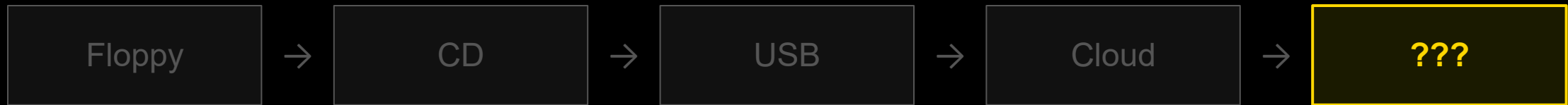
Context

なに？



THE BIG QUESTION

What if we are building trust on the wrong foundation?



Every generation builds infrastructure on the technology about to become obsolete.

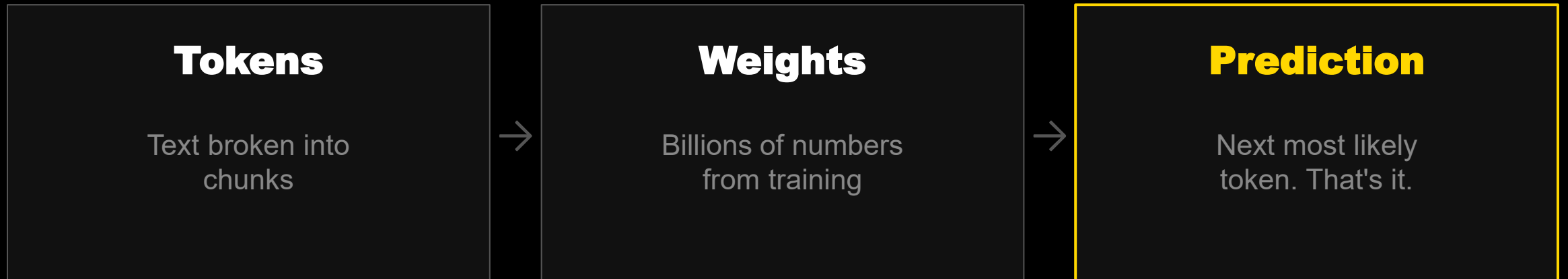
SELL ME THIS PEN



ITS AI POWERED

LET'S BE HONEST

Expensive Maths



*No understanding. No reasoning. Sophisticated autocomplete at scale.
Useful but worth being honest about what we're building trust frameworks for.*

SECOND-LEVEL THINKING

**The problem is not that we don't know.
The problem is that we don't know what we don't know.**

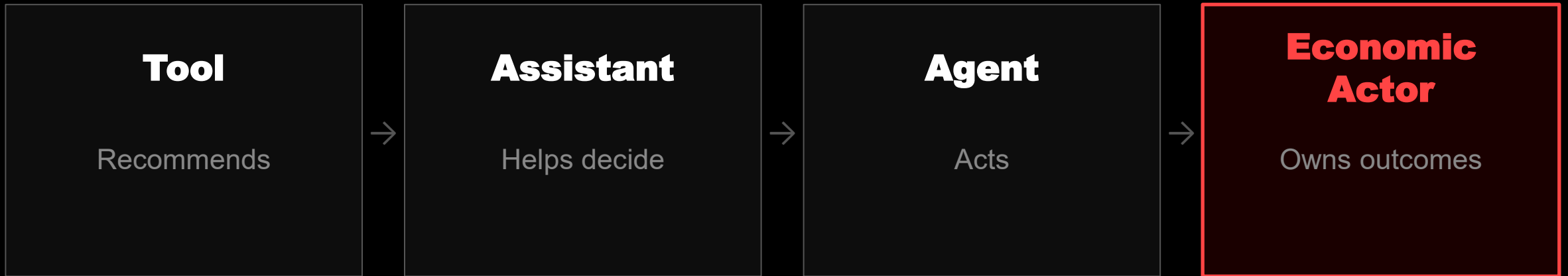
— *Howard Marks*



The AI industry has a confidence problem. Our models are confidently wrong at exactly the moments that matter most.

THE AGENTIC SHIFT

The trust problem just got 100× harder.



An AI agent doesn't wait for approval. It moves money. Commits to contracts. And when it goes wrong, who is responsible?

That is not a technology question. It is a design question.

AGENTIC COMMERCE IS ALREADY HERE

The first autonomous financial agents are live. We haven't designed trust for them yet.



**Do I trust the recommendation? → Do I trust the agent?
These are completely different design problems.**

THE PROOF

**Metric move.
Not by improving the AI.
By redesigning trust.**

BEFORE

**AI output
presented as fact**

12% activation



AFTER

**AI as advisor.
Human owns decision.**

68% activation

Three things that build trust with agents.

01

Reversibility

Can the action be undone? If yes, let the agent act. If no, require human sign-off. Always.

02

Bounded Authority

The agent knows what it can and cannot do. Hard limits, not soft guidelines.

03

Legible Accountability

When it goes wrong, there is a clear chain of who authorised what. Not a black box.

SECOND-LEVEL THINKING

First-level:

How do we make AI trustworthy?

Second-level:

How do we make trust survive the next AI?

Reversibility. Bounded authority. Legible accountability.
These are not AI principles. They are human principles.



**Design for the
technology that
does not exist yet.**

*The users who cannot afford to be wrong
will tell you everything you need to know.*