SELF-DESIGNING AI

From LLM Productivity to Tailored **Enterprise Growth**

Stratos Idreos
Professor, Harvard CS
Co-director, Harvard Data Science
Founder, Leibniz Labs, Next Gen. AI

ML

AI

COPILOTS

TRANSFORMERS

LLMs

NEED MORE AI

AI-DRIVEN

AI-FIRST

AGENTS

VECTOR DBs

RAG

AI NOISE

AI RACE

AI-POWERED

WHERE IS THE REAL VALUE?

HOW LONG
WILL IT TAKE?

ROI?

REVENUE?

COST?

SHOULD WE DO
THIS AI PROJECT?

WILL IT WORK?

MARGIN?

CHURN?

EBIT?

ML

AI

COPILOTS

TRANSFORMERS

LLMs

NEED MORE AI

AI-DRIVEN

AI-FIRST

AGENTS

VECTOR DBs

RAG

AI NOISE

AI RACE

AI-POWERED

SURVIVING vs TRANSFORMING

ROI?

REVENUE?

COST?

SHOULD WE DO THIS AI PROJECT?

WILL IT WORK?

MARGIN?

CHURN?

EBIT?

BUILD YOUR OWN INTELLIGENCE FOR CORE BUSINESS OPS

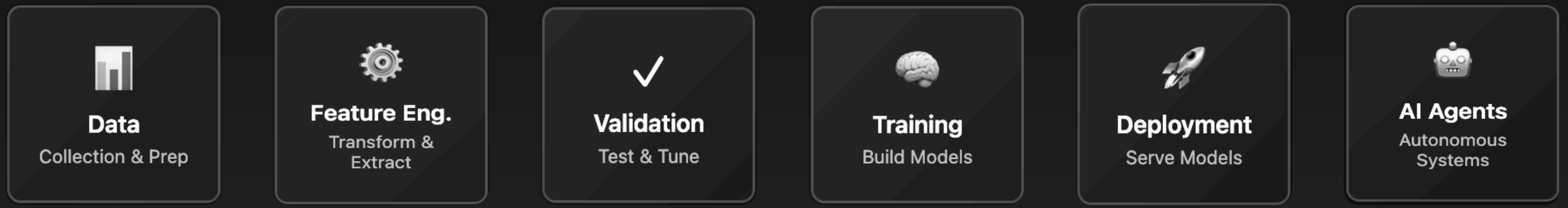
On your data. Your goals. Your business context. NOT JUST FOR PRODUCTIVITY

RENT INTELLIGENCE THAT EVERYONE HAS

OR OWN INTELLIGENCE NO ONE CAN COPY

TRANSFORMATION STARTS WITH AI SYSTEMS

The infra that builds AI — and reshapes how companies organize, hire, and compete

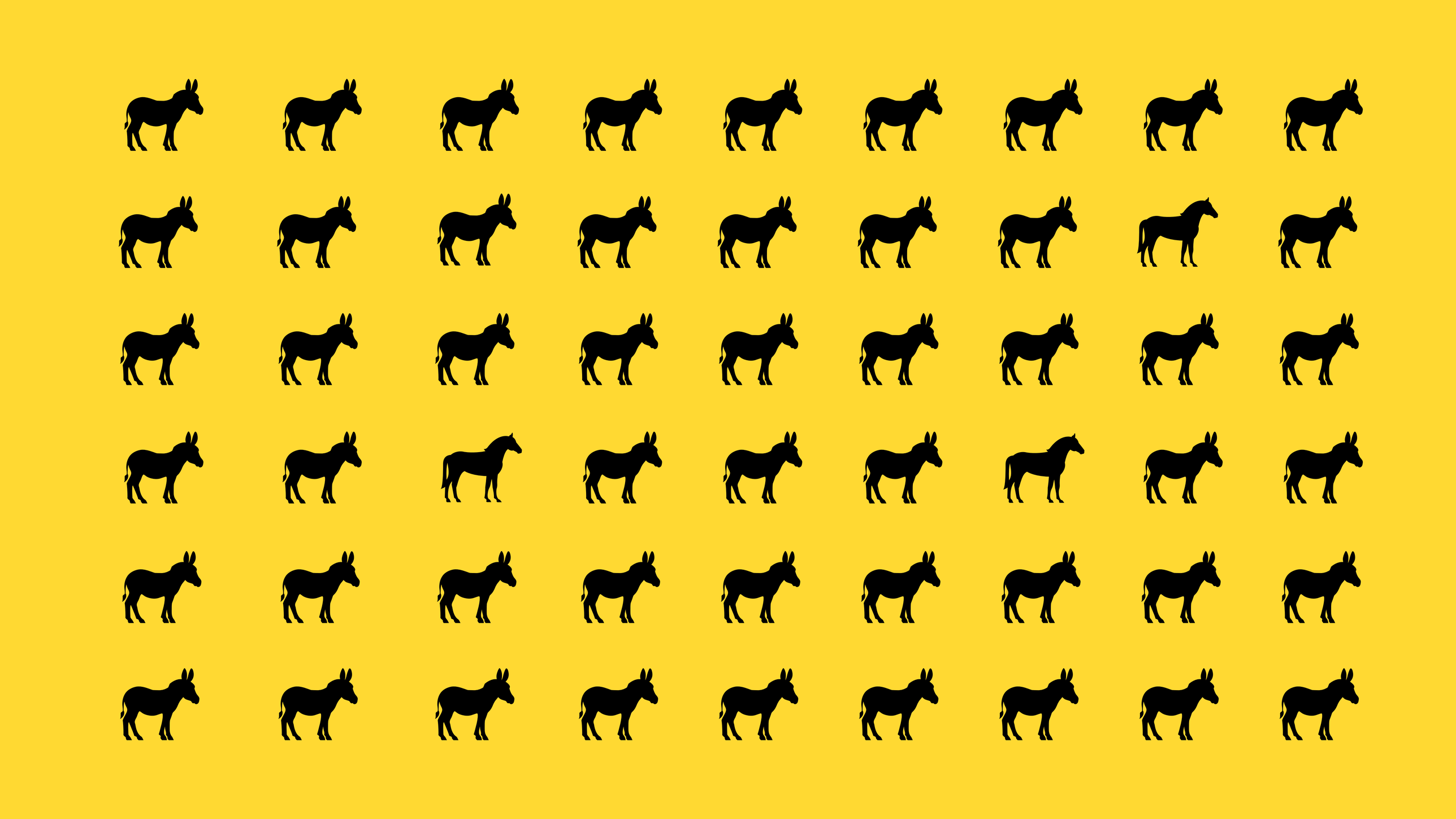


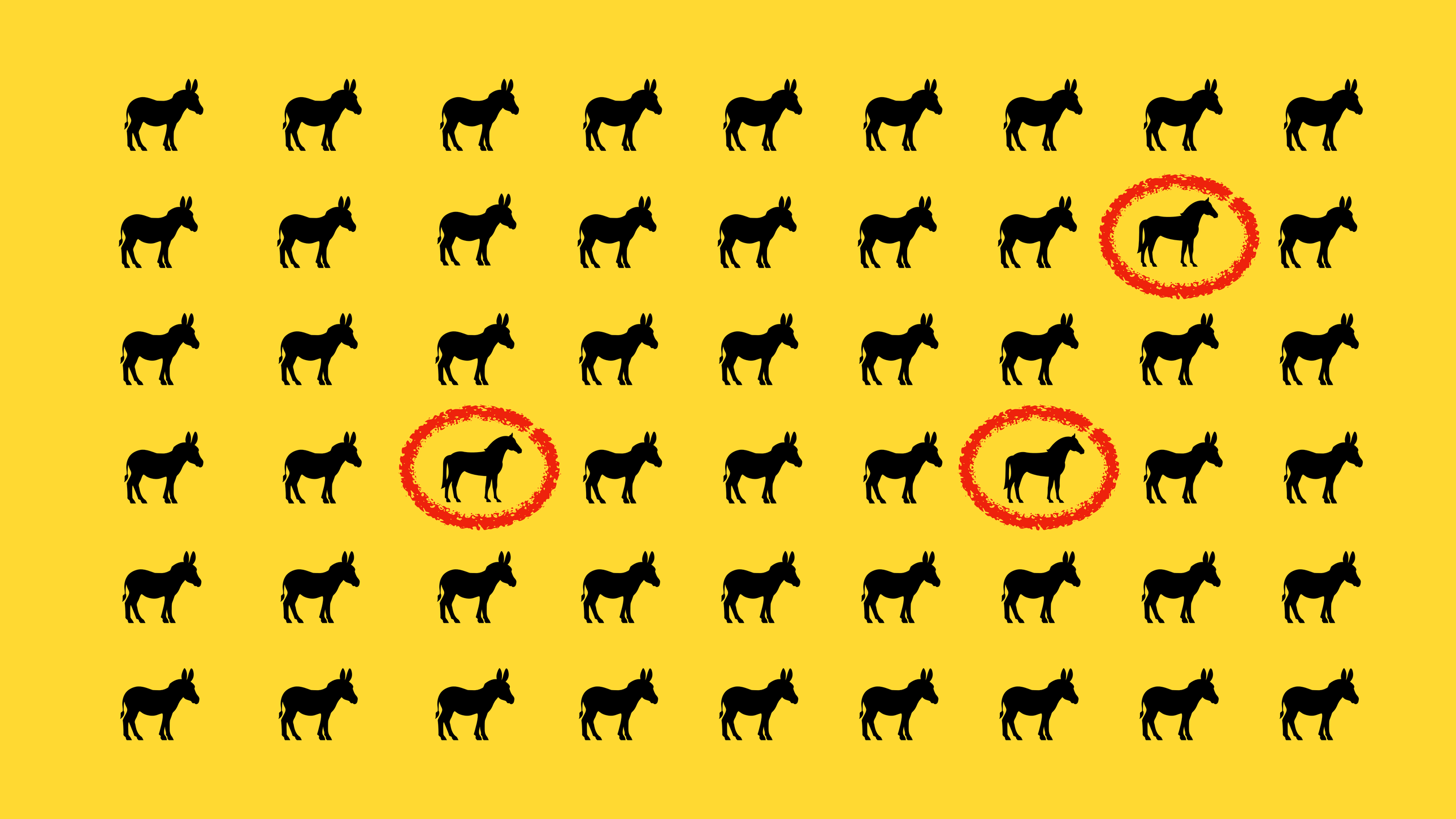
DETECT THE NUMBER OF HORSES

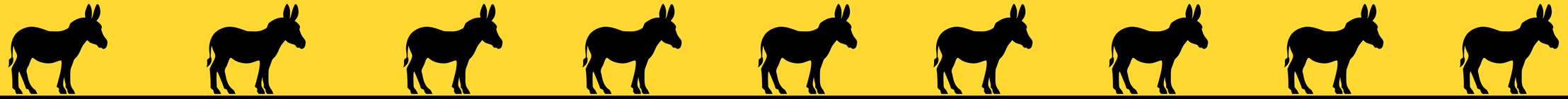
a typical image recognition AI task

DETECT THE NUMBER OF HORSES

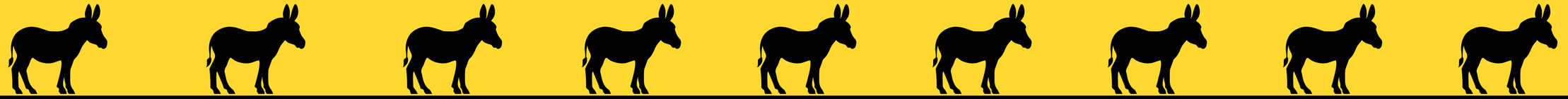
a typical image recognition AI task





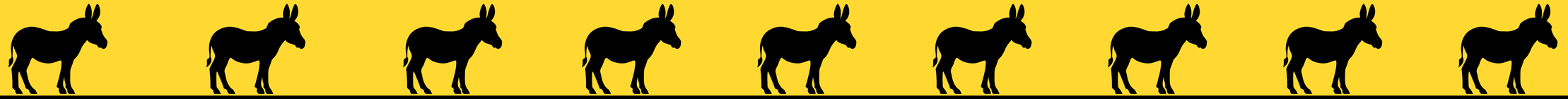


AS DATA GROWS, DATA ORGANIZATION IS KEY



AS DATA GROWS, DATA ORGANIZATION IS KEY

Ideal storage for this problem?



AS DATA GROWS, DATA ORGANIZATION IS KEY

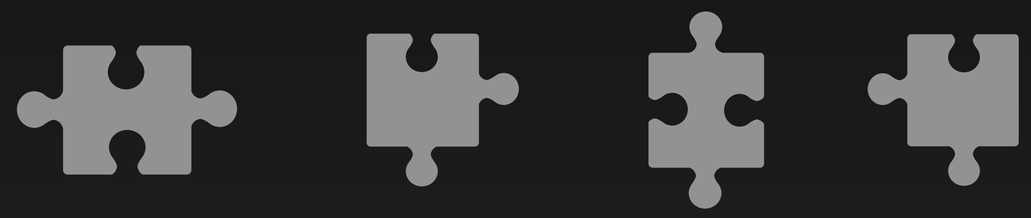
Ideal storage for this problem?

New AI Features → New AI Systems

BUT IT TAKES 7-10 YEARS TO BUILD A SYSTEM

FULL AI SYSTEMS DO NOT EXIST (YET)

Increasingly more:
Algorithms, Models, Tools, Libraries



What is the difference between an AI MODEL and INTELLIGENCE ?

INTELLIGENCE

AI MODEL

INTELLIGENCE

AI MODEL

not just about putting components together: efficiency, robustness, longevity, …

AN ABSURD EXAMPLE

My 9-year-old son is very very smart (I think)
Say I gave him the best heart surgery textbook

AN ABSURD EXAMPLE

My 9-year-old son is very very smart (I think)
Say I gave him the best heart surgery textbook

Should I let him operate on me?

No team. No tools. No hospital. No experience

AN ABSURD EXAMPLE

My 9-year-old son is very very smart (I think)
Say I gave him the best heart surgery textbook

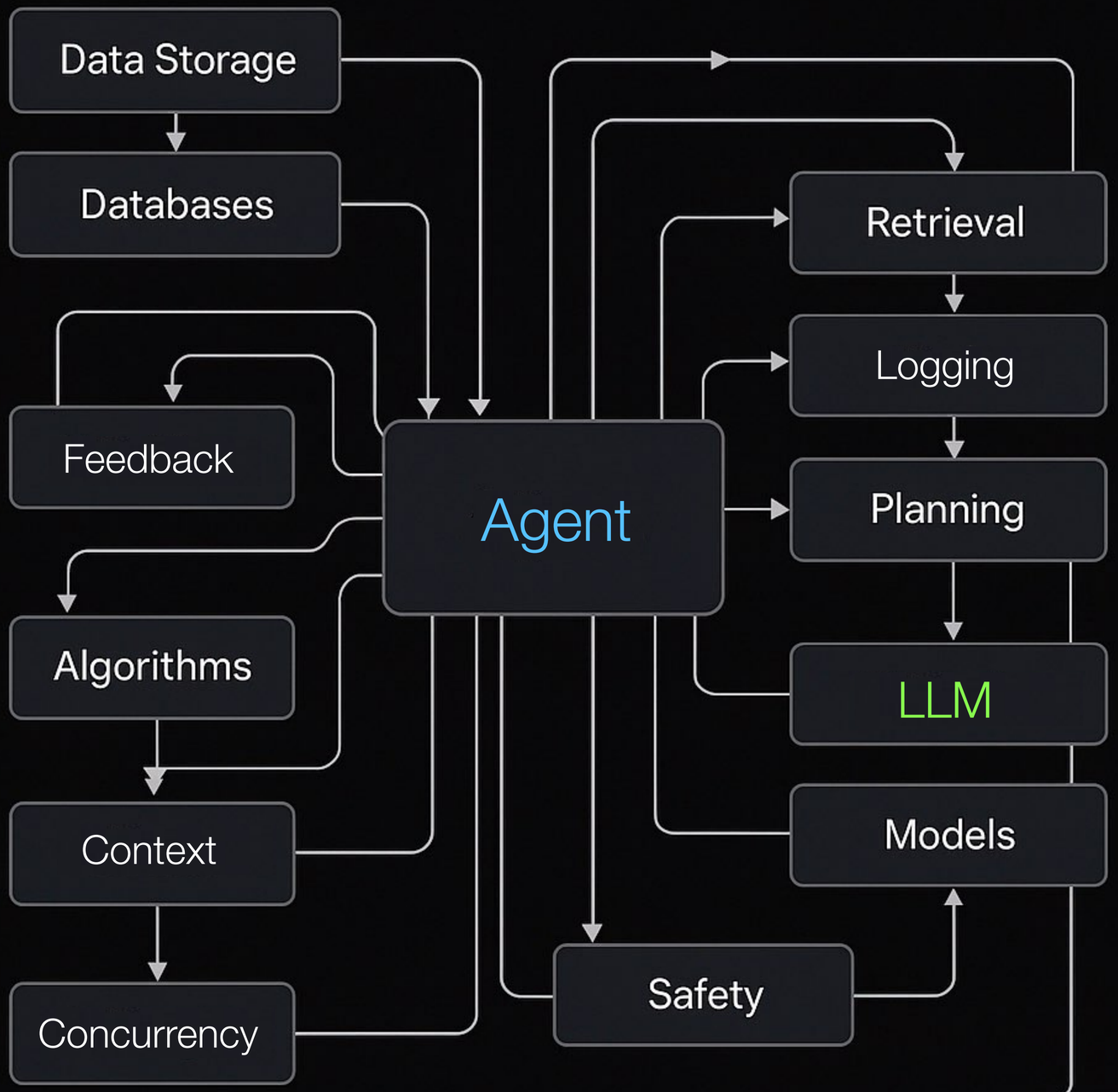
Should I let him operate on me?

No team. No tools. No hospital. No experience

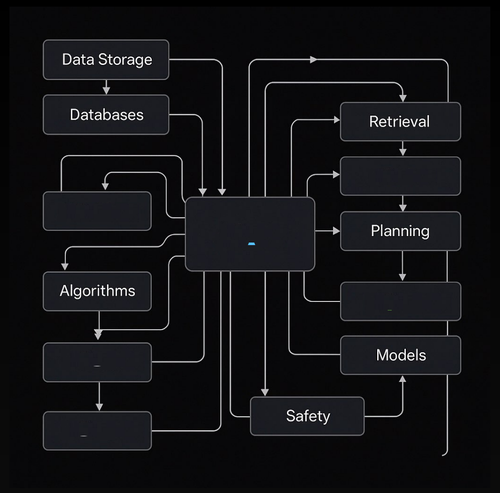
AND LIKE LLMs, HE
HARDLY EVER FOLLOWS MY INSTRUCTIONS…

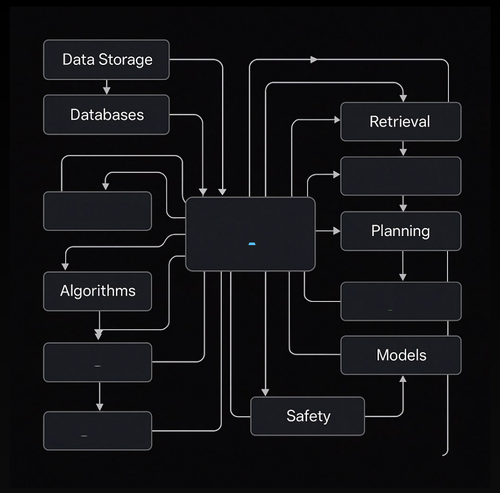
WE DO NOT HAVE FULL AGENT SYSTEMS YET

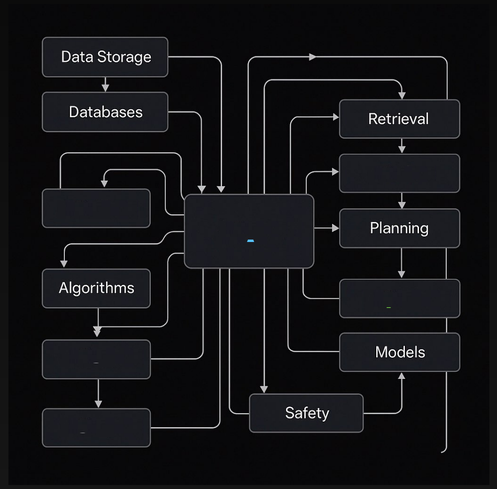
WHAT WE HAVE	WHAT WE NEED
Impressive demos	Production infrastructure
Point solutions	Predictable behavior
Fragile prototypes	Programmable agents


Data Storage
Databases
Feedback
Agent
Algorithms
Context
Concurrency
Retrieval
Logging
Planning
LLM
Models
Safety

This is why...



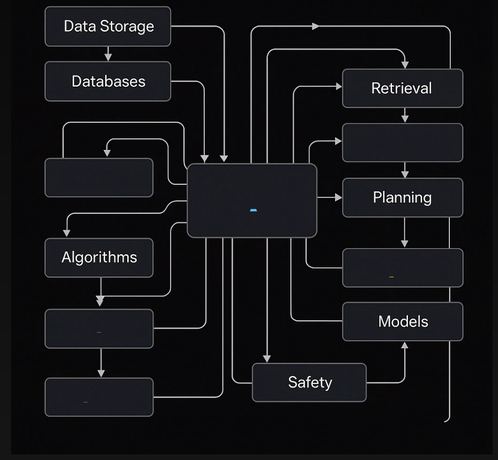


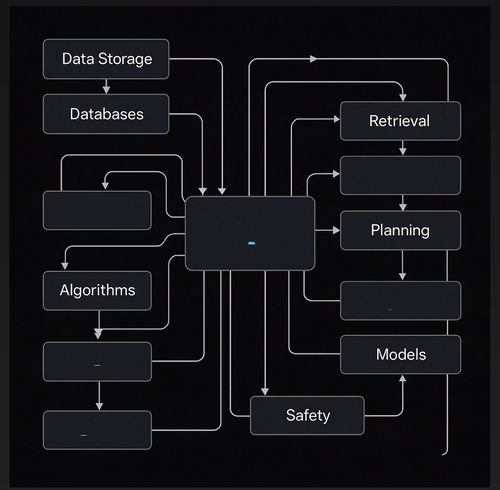


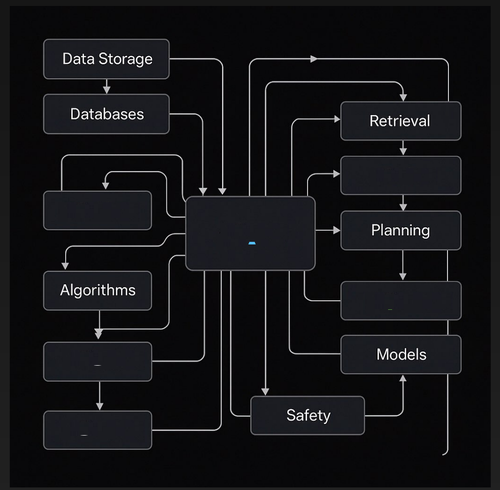
A full agent system is a

DISTRIBUTED SYSTEM

Concesus Protocols. Game Theory. Self-improving.

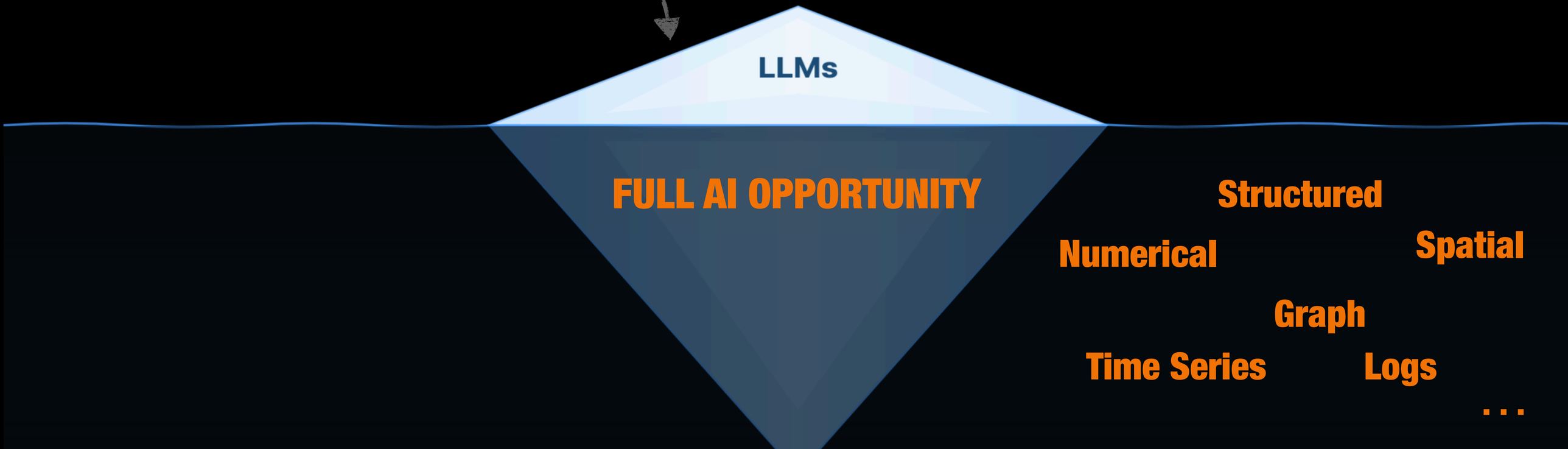


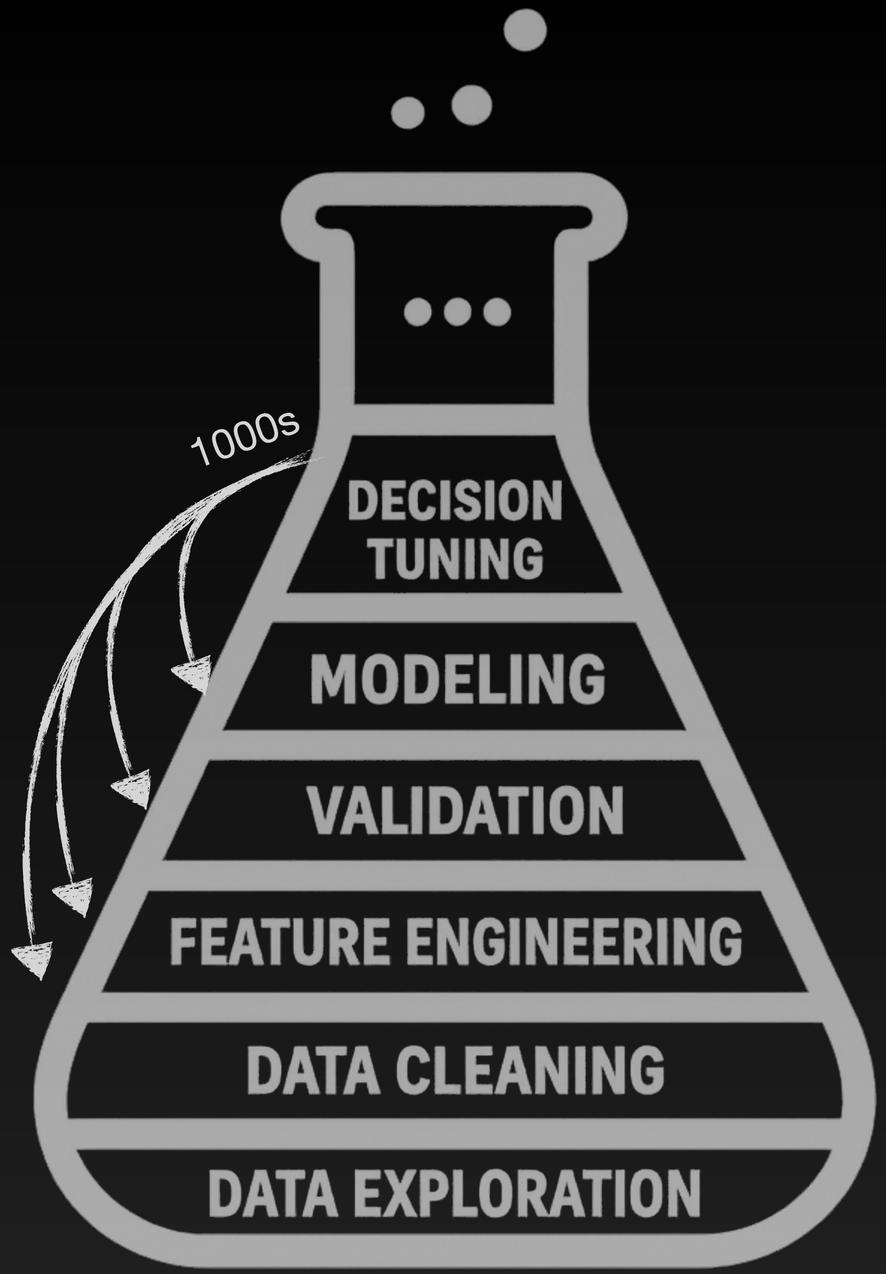




MANY MORE MODELS NEEDED

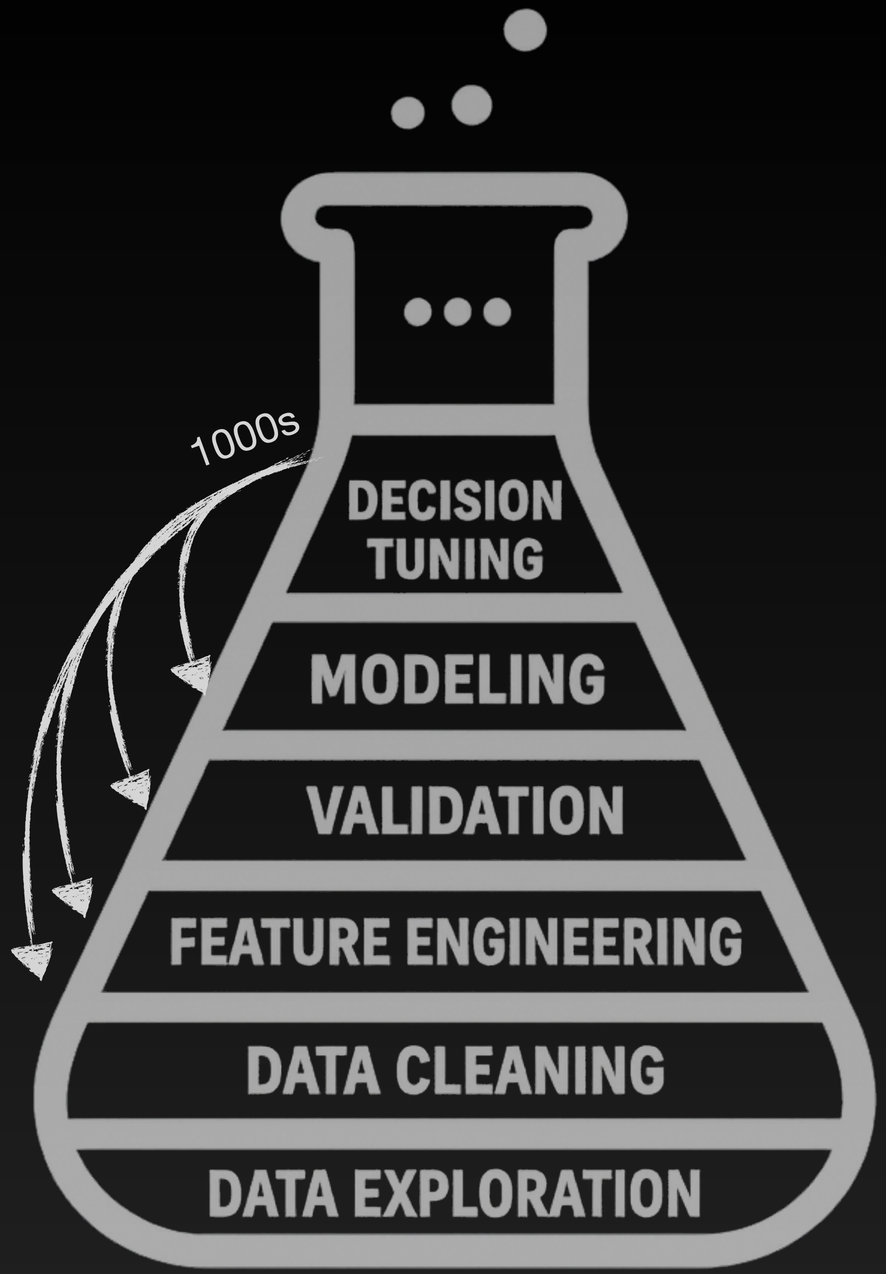
Reasoning, Language, Text
Summaries, Docs, Reports, …





WE DO NOT HAVE FULL MODELING SYSTEMS

SLOW CODING/EXPS $\neq$ ROBUST AI



WE DO NOT HAVE FULL MODELING SYSTEMS

SLOW CODING/EXPS ≠ ROBUST AI

What is needed:

RELENTLESS EXPERIMENTATION

DRIVEN BY BUSINESS CONTEXT

WITHOUT FULL AI SYSTEMS ENTERPRISES ARE LIKE SCIENTISTS WITHOUT LABS

WITHOUT FULL AI SYSTEMS ENTERPRISES ARE LIKE SCIENTISTS WITHOUT LABS

✓ Brilliant people
✓ Massive data
✓ Important questions

✗ Uncertain projects
✗ Manual experiments
✗ Months per iteration

This is why
80% projects fail

BUT HERE IS THE CHALLENGE WITH AI SYSTEMS

Every business
problem is UNIQUE

AI transformation is
about EXPLORATION

BUT HERE IS THE CHALLENGE WITH AI SYSTEMS

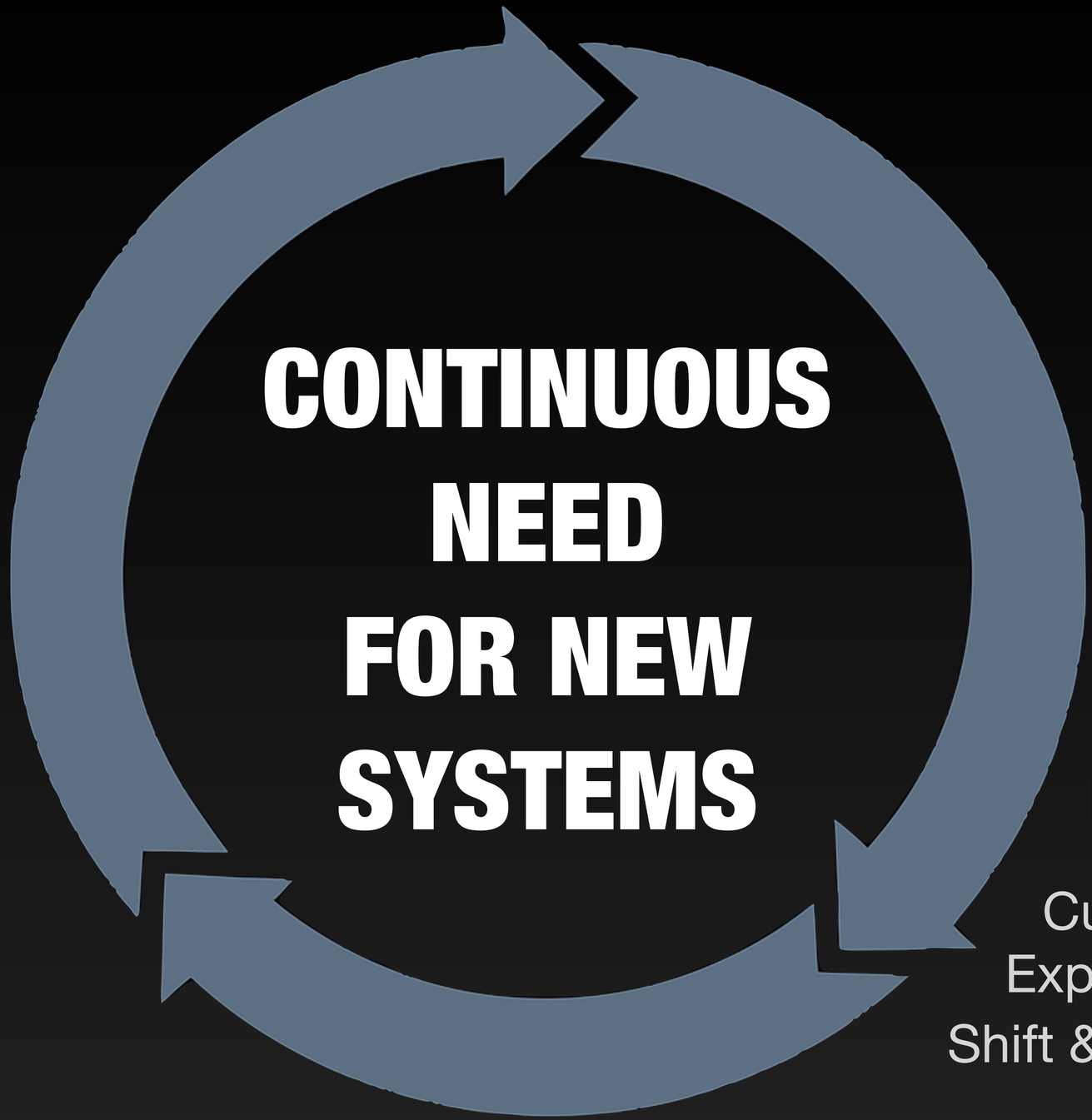
Every business problem is UNIQUE

AI transformation is about EXPLORATION

HOW CAN WE BUILD FULL SYSTEMS IF WE DO NOT KNOW THE GOAL?

Every business problem is UNIQUE

AI transformation is about EXPLORATION

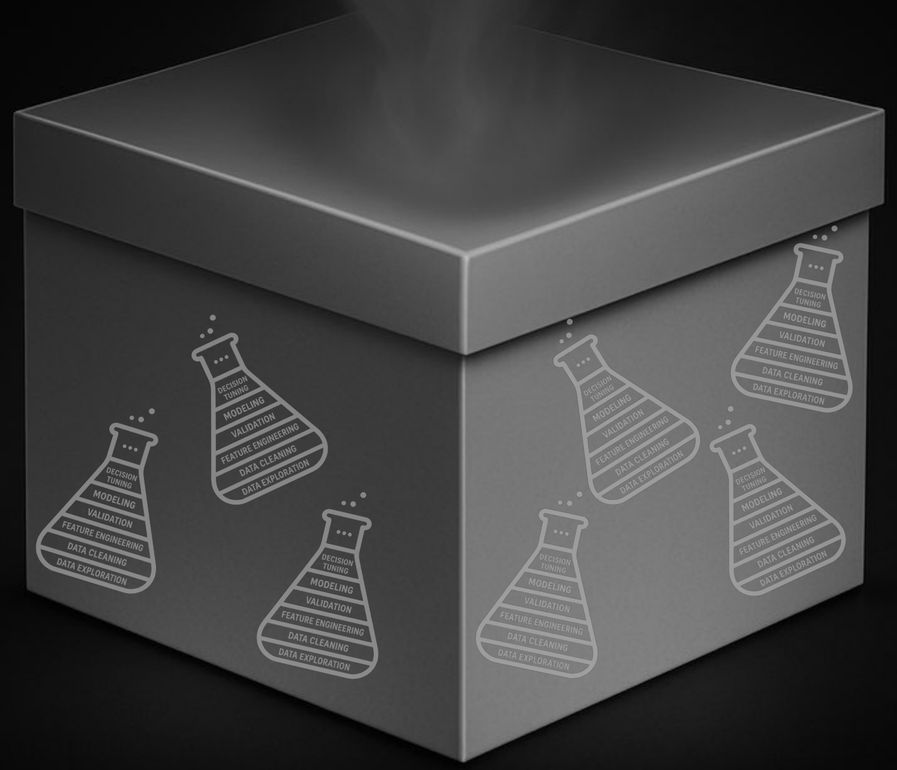


New Goals

New Hardware

Customer Expectations Shift & Grow

What if it was possible



What if it was possible

to simply state the AI goal

What if it was possible

to simply state the AI goal

& automatically/fast
produce the perfect

INTELLIGENCE

INTELLIGENCE

Explainable. Extensible.
Self-improving. Always On.

Agent 1

LLMs

New Tailored Small Models

Agent N

...

Agent 4

Agent 2

Agent 3

DATA. TOOLS. PEOPLE

SELF-DESIGNING AI SYSTEMS

automatically invent & build the perfect system for any (new) goal

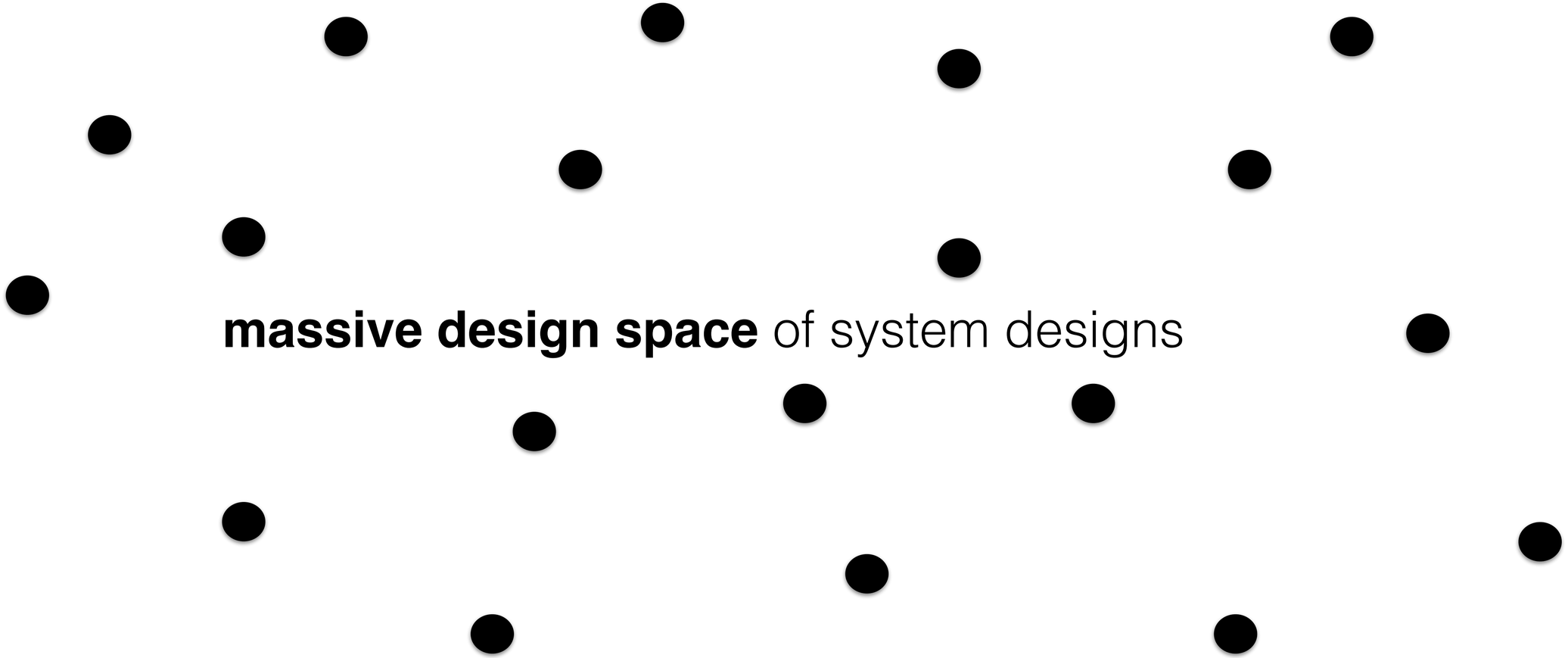
DASlab
@ Harvard SEAS

SELF-DESIGNING AI SYSTEMS

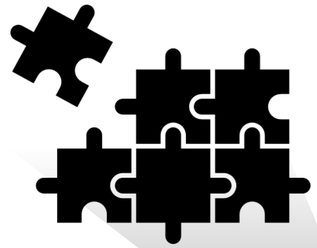
automatically invent & build the perfect system for any (new) goal

Started this journey in 2005 - my PhD thesis:

“Every query should be treated as advice on how the data should be stored”



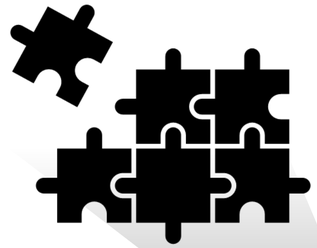
massive design space of system designs



system design=
a set of low-level
design decisions

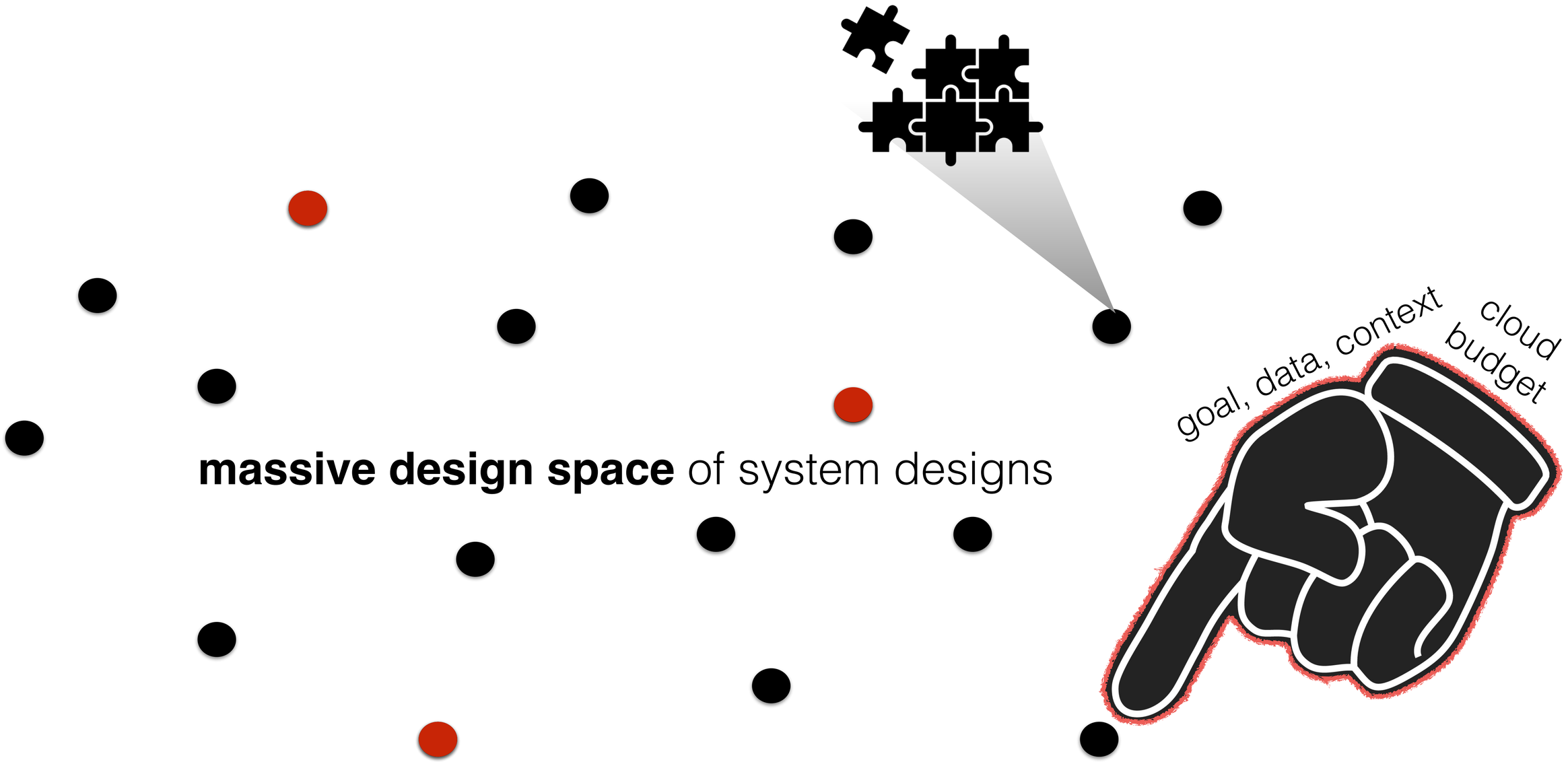
massive design space of system designs

few existing designs



system design=
a set of low-level
design decisions

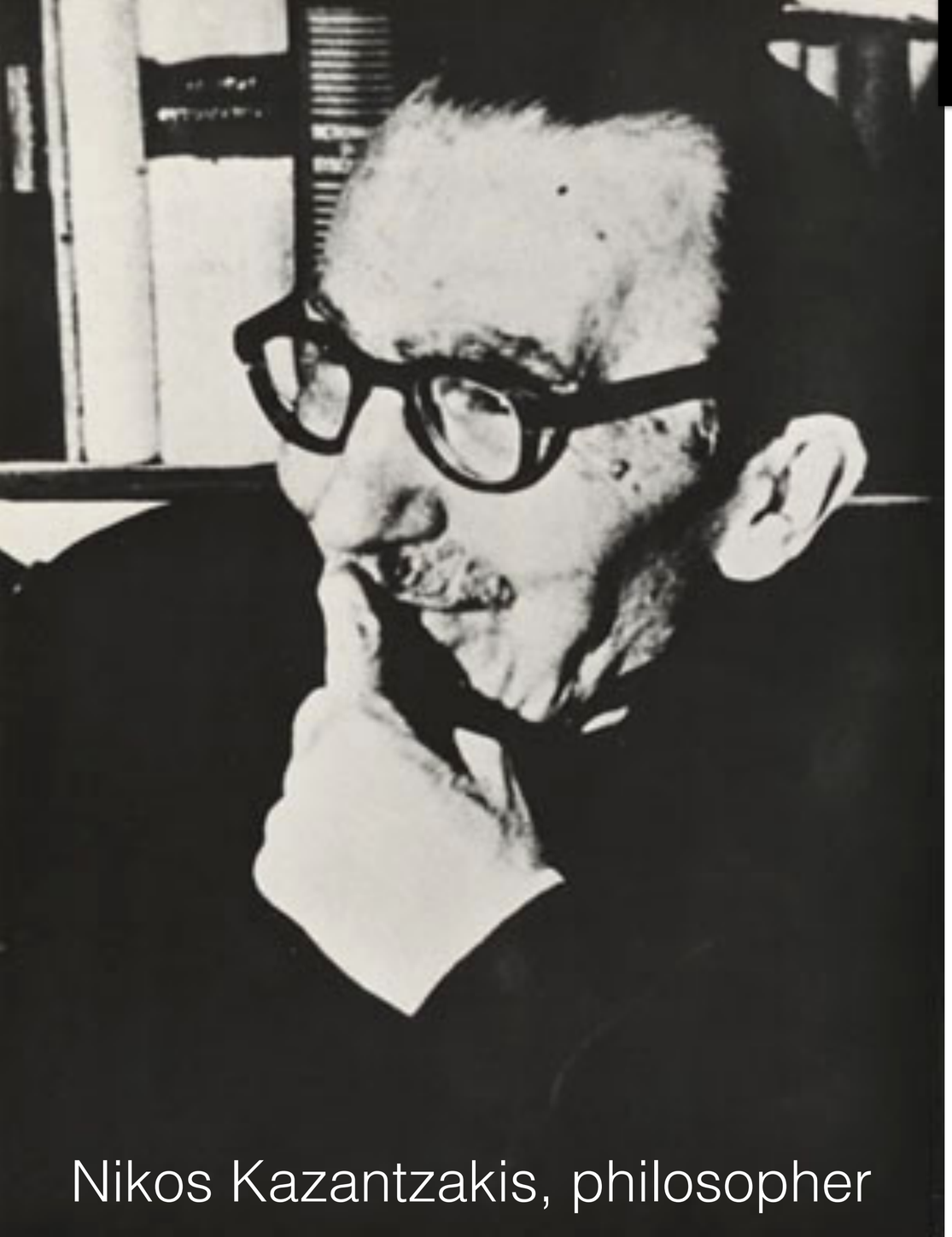
massive design space of system designs



massive design space of system designs

the **grammar** of AI systems design

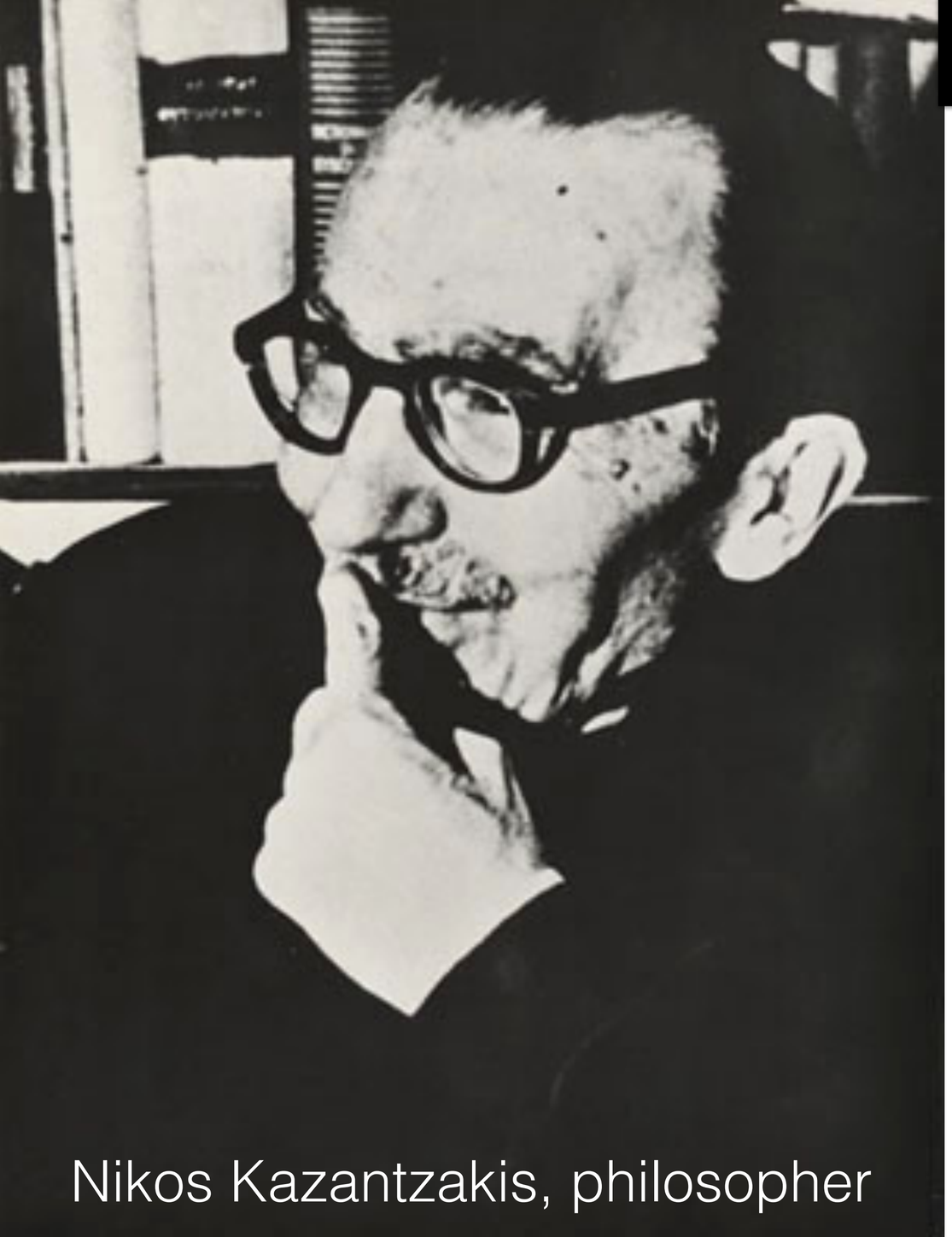
just as DNA encodes all possible proteins,
we are encoding the system design principles that
generate **all possible AI systems**



Nikos Kazantzakis, philosopher

the **grammar** of AI systems design

action is for nothing
hope the most holy
fear form of
am free
ultimate I theory

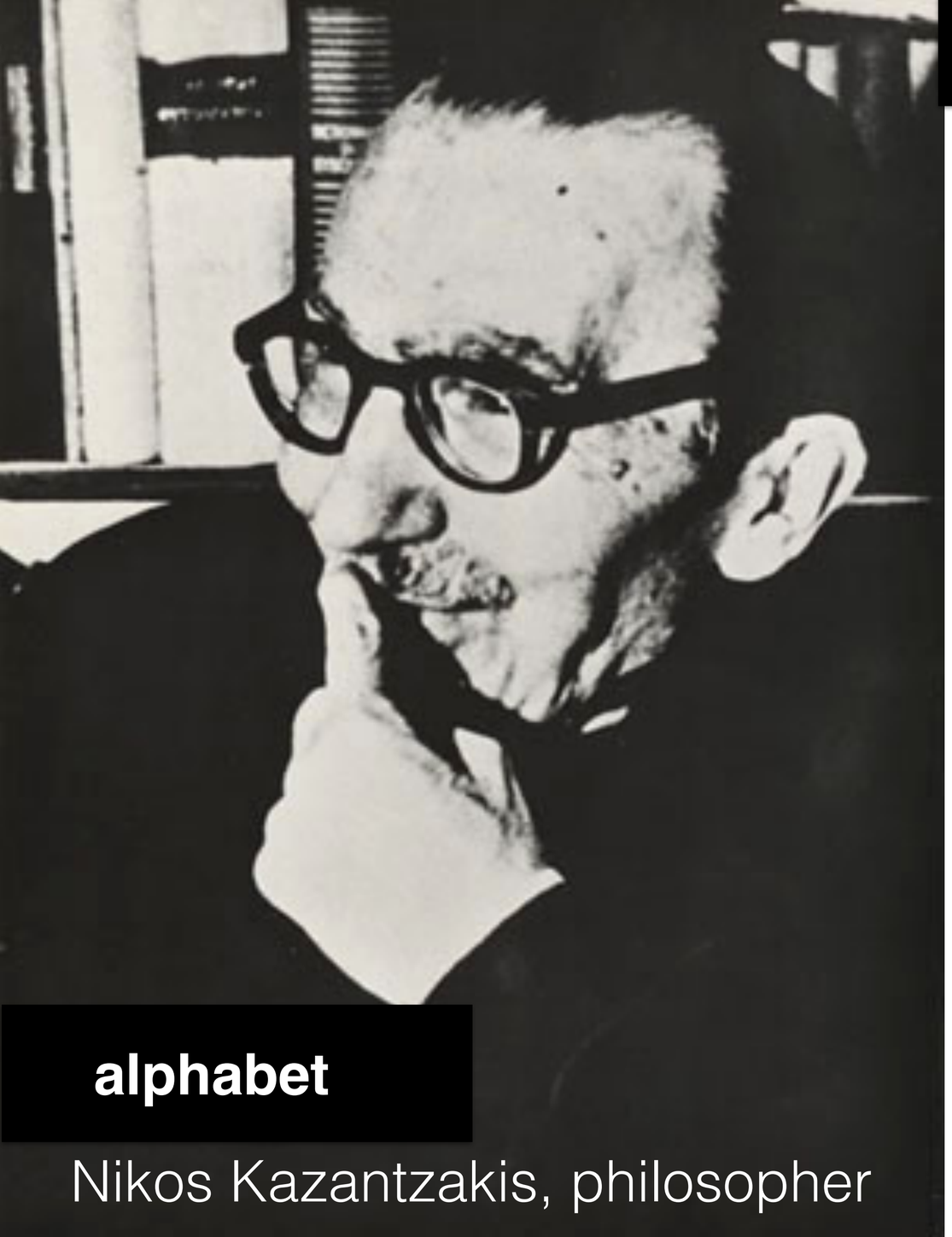


Nikos Kazantzakis, philosopher

the **grammar** of AI systems design

action is the ultimate most holy form of theory

I hope for nothing
I fear nothing
I am free



alphabet

Nikos Kazantzakis, philosopher

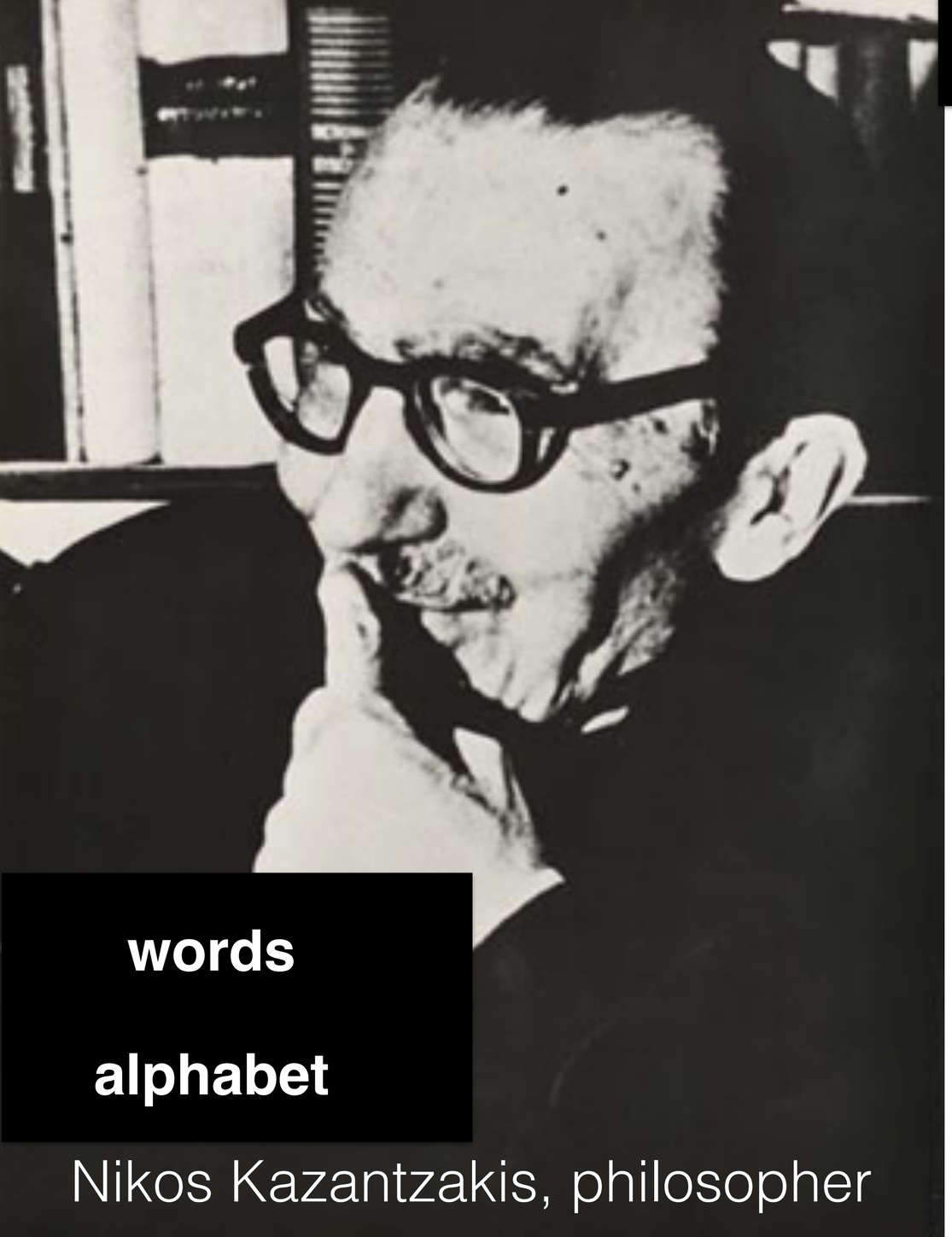
the **grammar** of AI systems design

action is the most holy form of ultimate theory

I hope for nothing

I fear nothing

I am free



words

alphabet

Nikos Kazantzakis, philosopher

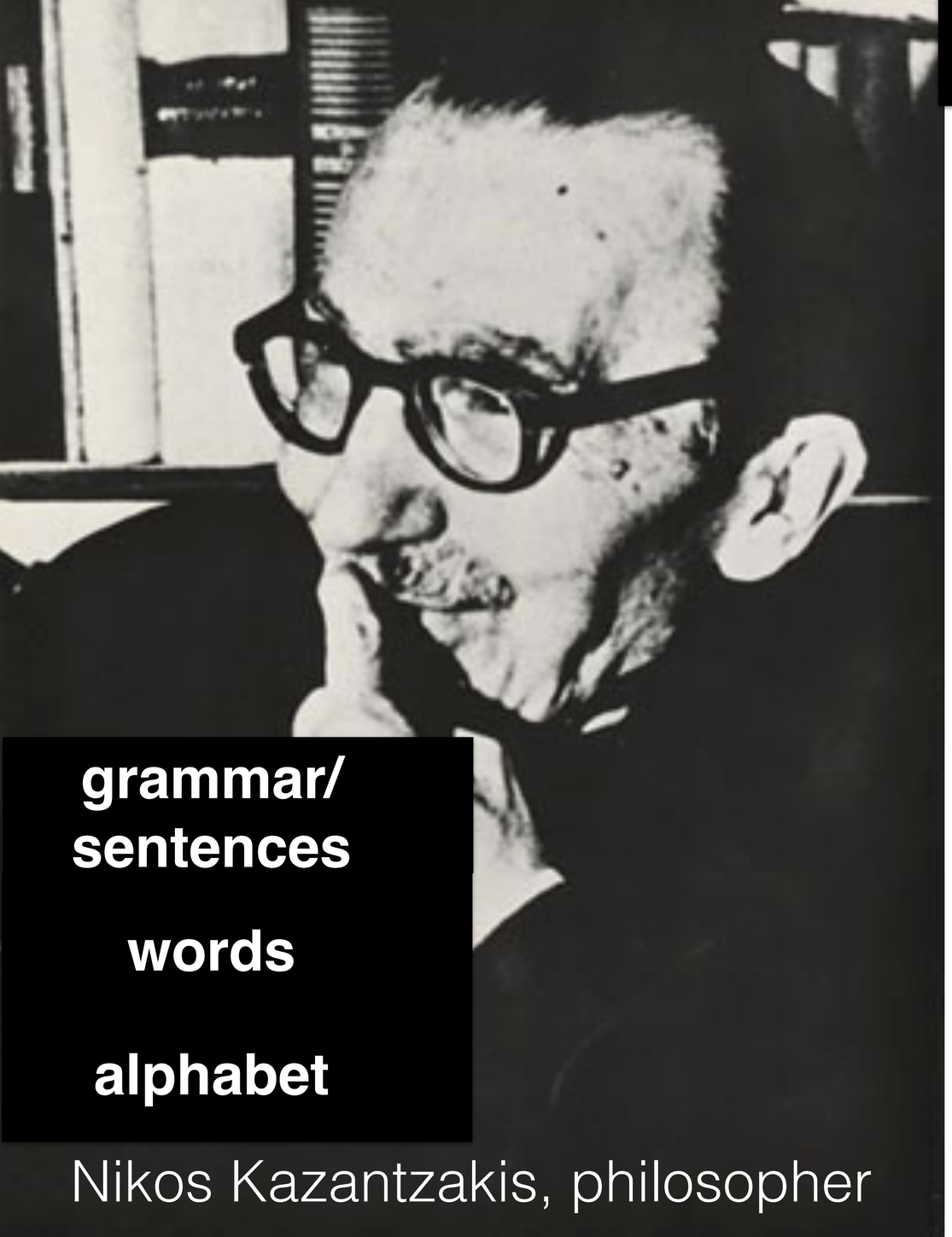
the **grammar** of AI systems design

action is the ultimate most holy form of theory

I hope for nothing

I fear nothing

I am free



grammar/ sentences

words

alphabet

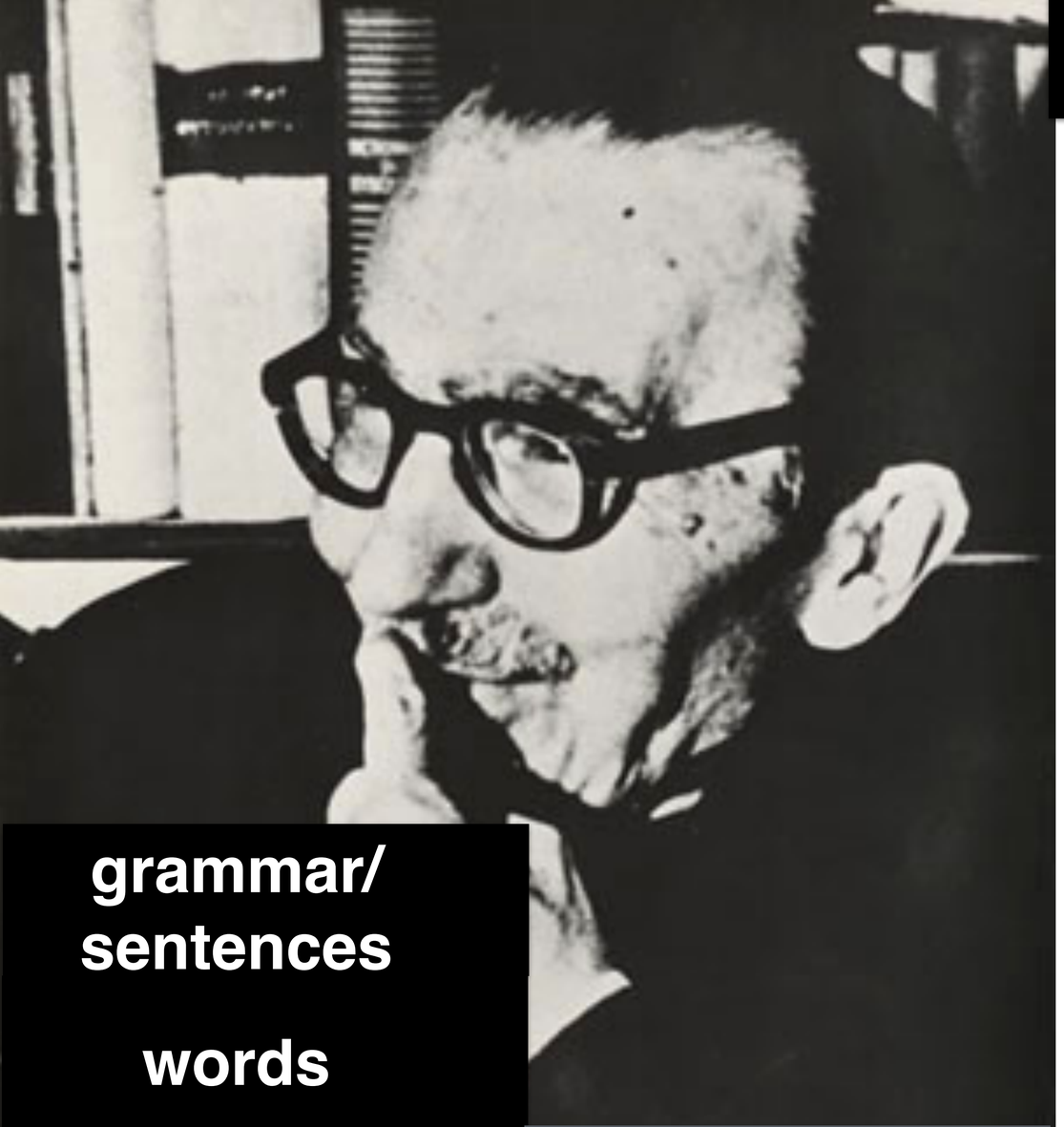
Nikos Kazantzakis, philosopher

the **grammar** of AI systems design

action is the ultimate most holy form of theory

I hope for nothing
I fear nothing
I am free

the **grammar** of AI systems design



grammar/ sentences

words

alphabet

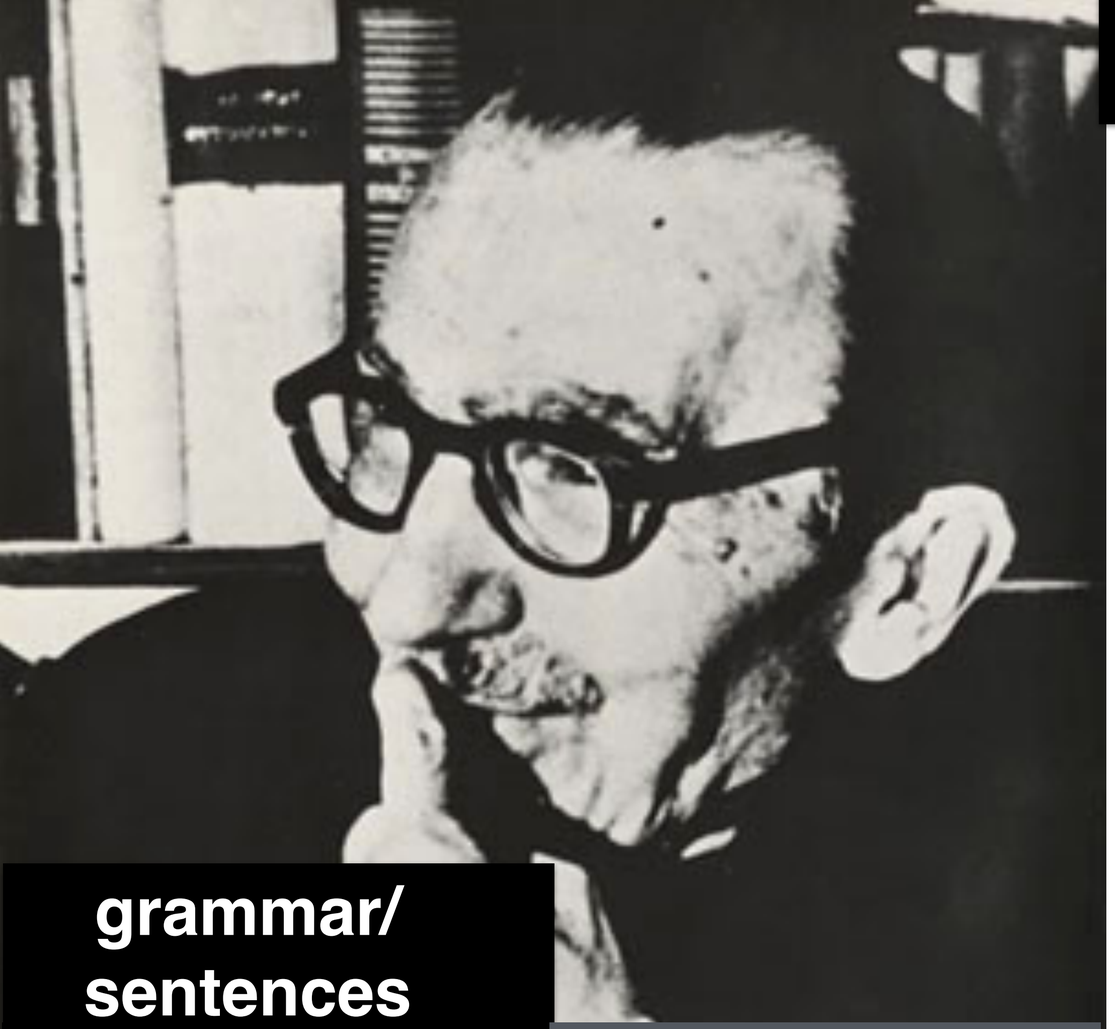
principles

Nikos Kazantzakis, philosopher

action is the ultimate most holy form of theory

I hope for nothing
I fear nothing
I am free

the **grammar** of AI systems design



grammar/ sentences

words

alphabet

models, algos, …

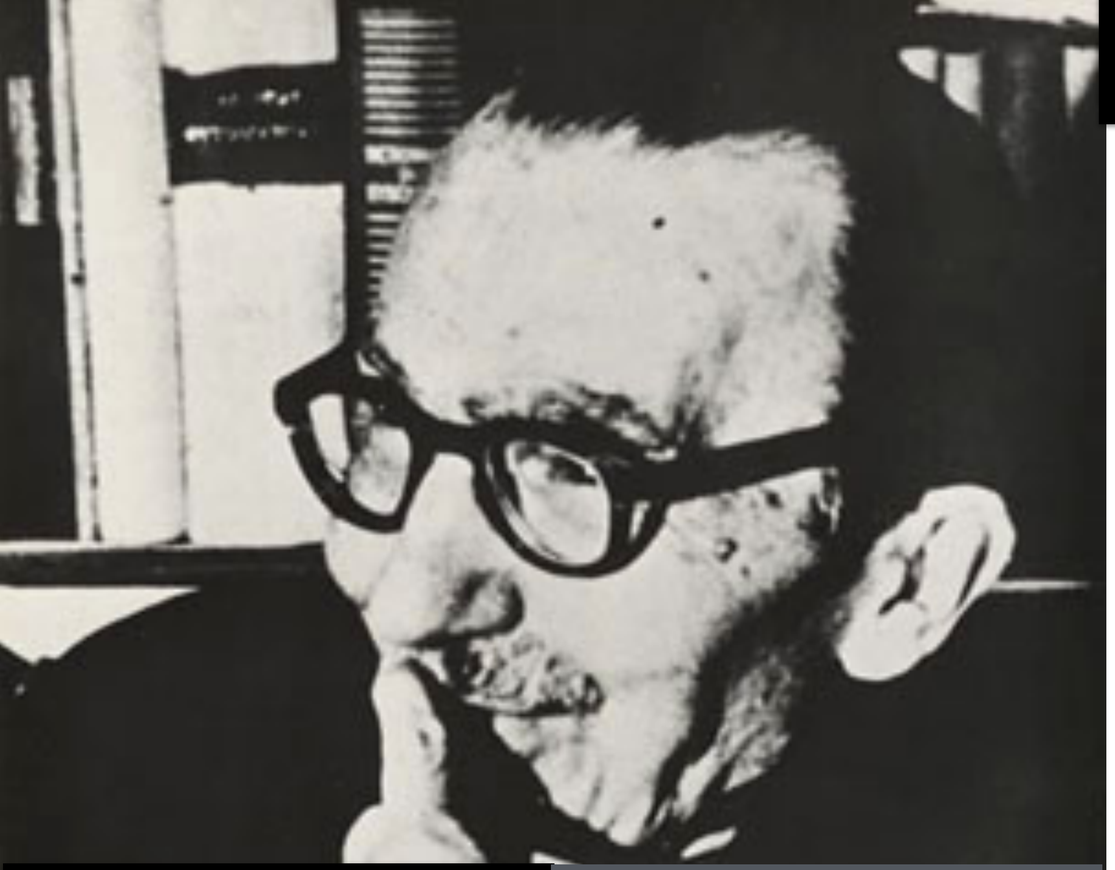
principles

Nikos Kazantzakis, philosopher

action is the ultimate most holy form of theory

I hope for nothing
I fear nothing
I am free

the **grammar** of AI systems design

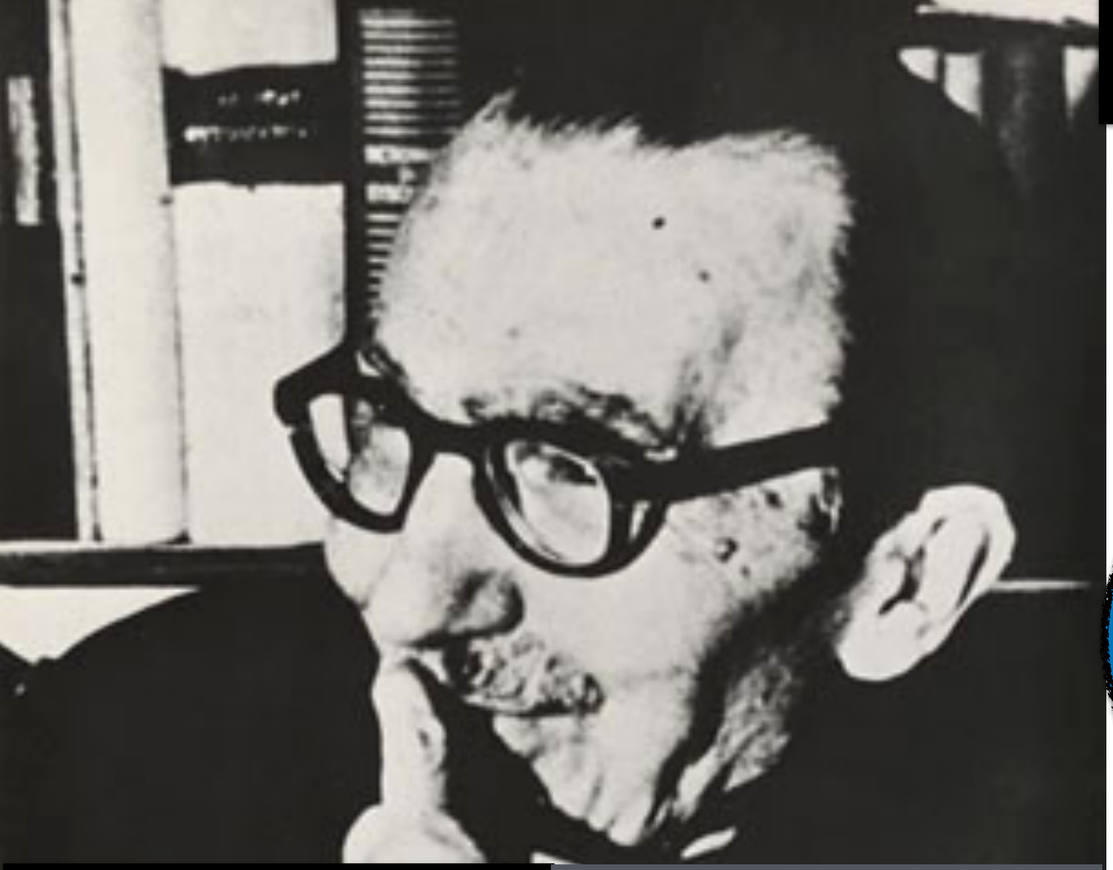


grammar/ sentences	AI systems
words	models, algos, …
alphabet	principles

Nikos Kazantzakis, philosopher

action is the most holy form of ultimate theory

I hope for nothing
I fear nothing
I am free



grammar/ sentences	AI systems
words	models, algos, …
alphabet	principles

Nikos Kazantzakis, philosopher

the **grammar** of AI systems design

action is the most holy form of theory

THIS LETS US BUILD ALL POSSIBLE AI SYSTEMS EVEN ONES NEVER IMAGINED

I hope for nothing
I fear nothing
I am free

SIGMOD'18

MORE DATA STRUCTURES THAN STARS IN THE SKY

(The most fundamental component of computer science/AI)

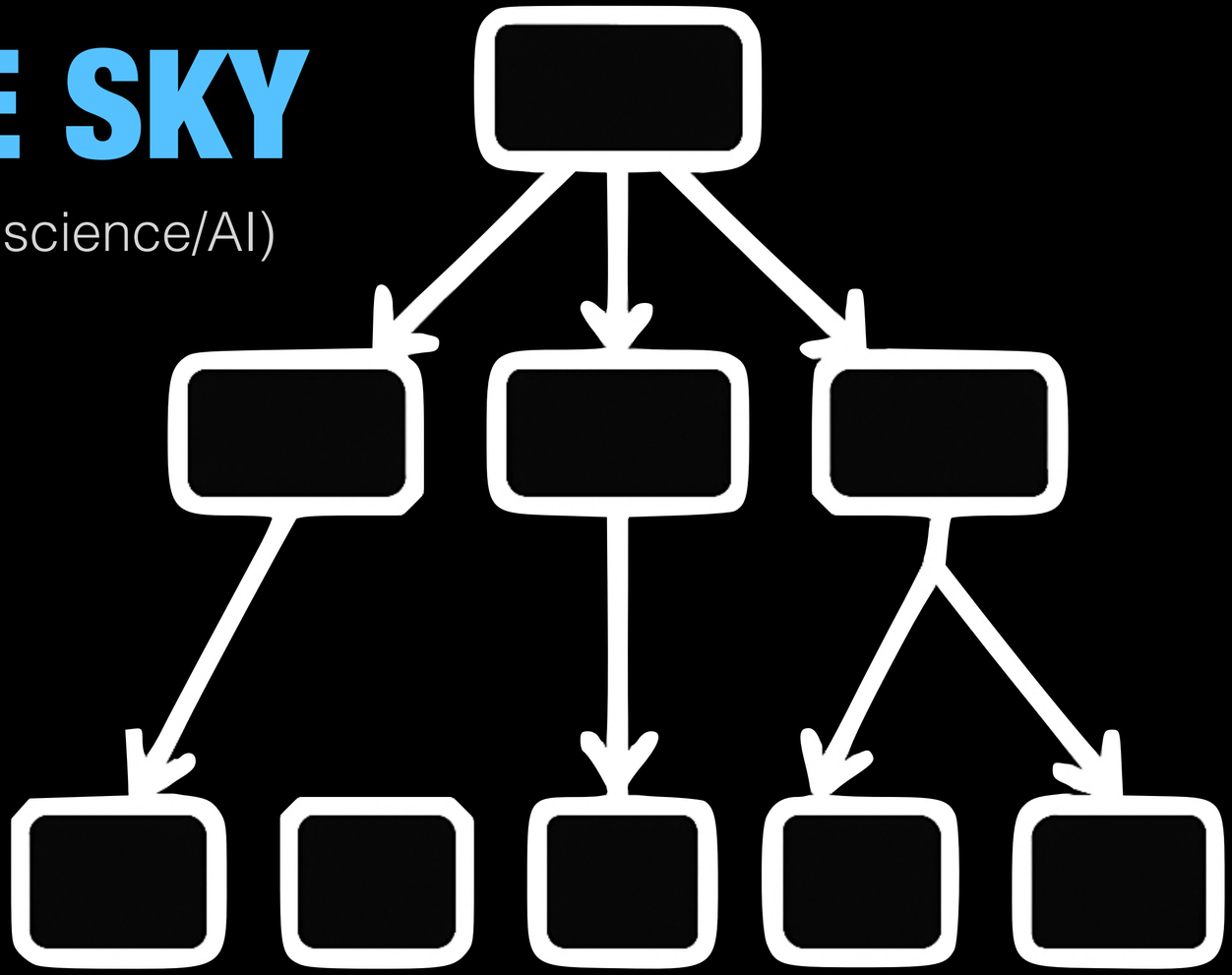
$5x10^3$

Literature

10^{24}

Stars

$>10^{48}$

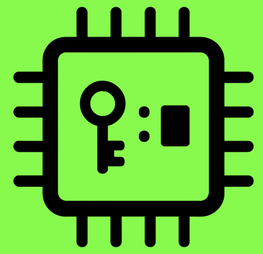


Possibilities We Discovered

10-100X FASTER SYSTEMS

Limousine: NoSQL KV-Store

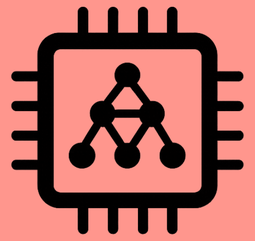
Agents' context management,
but also all kinds of big data infra



SIGMOD'24, VLDB'22

Image Calculator: Image AI

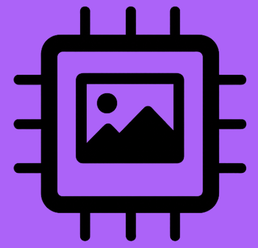
Storage for Training and Infenence



SIGMOD'24

TorchTitan with PyTorch@META

Large Model Training Algorithms



MLsys 2023, ICLR'25

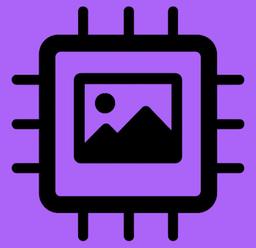
10-100X FASTER SYSTEMS

Here is my model
Here are my VMs
Here is my (new) training algorithm

TAILORED IMPLEMENTATION
or
FUTURE FAILURE ALERT

TorchTitan with PyTorch@META

Large Model Training Algorithms



MLsys 2023, ICLR'25

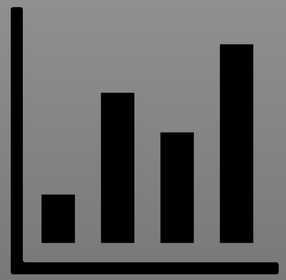
THERE ARE THREE CRITICAL FEATURES IN AI DEVELOPMENT

SPEED, SPEED, & SPEED

Time to market, Better models, More models

PERPETUAL LEARNING POSSIBLE

DESIGN SPACE



PERFORMANCE ESTIMATION

FIND BEST DESIGN

2 OPPORTUNITIES & COLLABORATION CALL

#1 ENTERPRISE INTELLIGENCE (LEIBNIZ STARTUP)

#1 ENTERPRISE INTELLIGENCE (LEIBNIZ STARTUP)

1. Any question/goal in human language:

“How can we increase margin by 1% in the next 9 months?”

“Why did sale X was so fast? How can we repeat this?”

#1 ENTERPRISE INTELLIGENCE (LEIBNIZ STARTUP)

1. Any question/goal in human language:

“How can we increase margin by 1% in the next 9 months?”

“Why did sale X was so fast? How can we repeat this?”

production-ready

AI in just days

AI models+Agents

24/7

#1 ENTERPRISE INTELLIGENCE (LEIBNIZ STARTUP)

1. Any question/goal in human language:

"How can we increase margin by 1% in the next 9 months?"

"Why did sale X was so fast? How can we repeat this?"

2. Here is my data:

What are the highest ROI actions I can take?

production-ready

AI in just days

AI models+Agents

Together with:

Jeff Barnett

Former CEO Com. Cloud Salesforce

Varun Srirum

Former Lead Data Scientist, Guy Carpenter

Stavros Harizopoulos

Former Principal Tech Lead Google BigQuery AI Engine
Google Cloud Apache Spark
AWS Redshift

Ask us about our:

Design Partners Program

#2 SCIENTIFC DISCOVERY (HARVARD DATA SCIENCE)

A UNIFIED SUBSTRATE: ONE HARVARD - ONE AI

Together with:

Francesca Dominici

Harvard T.H. Chan School of Public Health

Ask us about our:

Industry Partners program

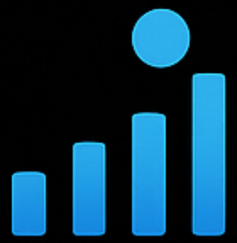
What happens when everyone can build their own intelligence?

1950S-70S	1970S-80S	2000S-10S	2030
Compilers	**Databases**	**Cloud**	**Full AI Systems**
No more assembly	Just write SQL	No cluster needed	No AI engineering needed

History repeats itself

WHERE WILL THE AI VALUE COME FROM?

History repeats itself



DATA

> Model Size

PEOPLE

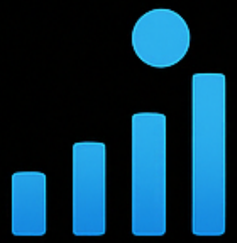
> Generic Models

SYSTEMS

> Manual Builds

WHERE WILL THE AI VALUE COME FROM?

History repeats itself



DATA

> Model Size

PEOPLE

> Generic Models

SYSTEMS

> Manual Builds

THANKS!