

Generative Artificial Intelligence

For Program Evaluation

2026 Summer Institute in Program Evaluation
June 10, 2026

Renato Bezerra
Educational Developer

The Centre for the Advancement
of Teaching and Learning



Presentation Overview

1. What we mean by “AI”
2. Generative AI as a normalizing engine
3. Consequences of normalizing program evaluation
4. The human-first use of generative AI

use it to **support what you know**
not to do what you don't know



Clarifying a few things

It mimics **human language**

but lacks *intent, emotions, context, consciousness, or experience*

It is extremely **impressive**

but not *intelligent*

It is highly **problematic**

Energy and water use, labour exploitation, intellectual property theft, privacy breaches, job displacements (mostly due to hype)

It works relatively well with **building blocks**

but not with *assembly*

It generates what is **likely** (probabilistic)

not necessarily what is *correct* (deterministic)

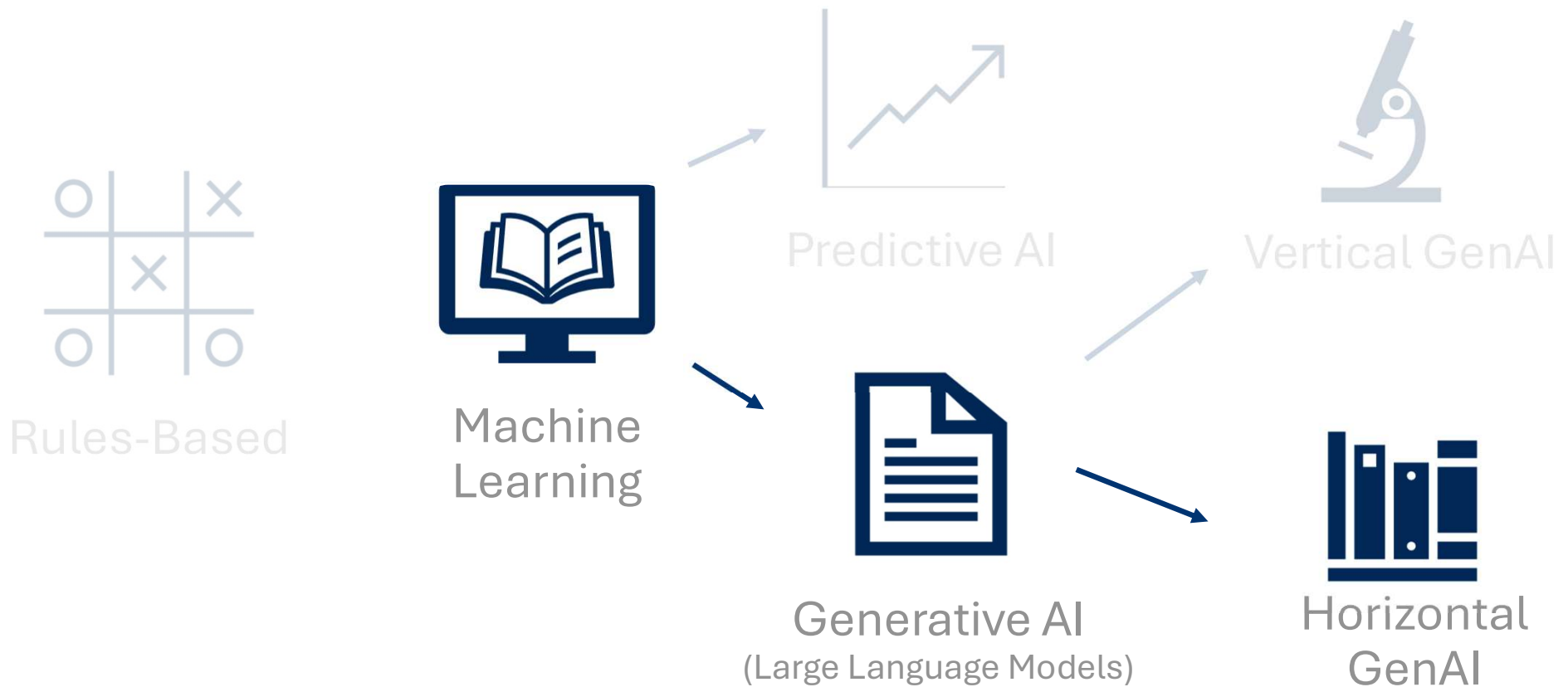
It is an **economic project**



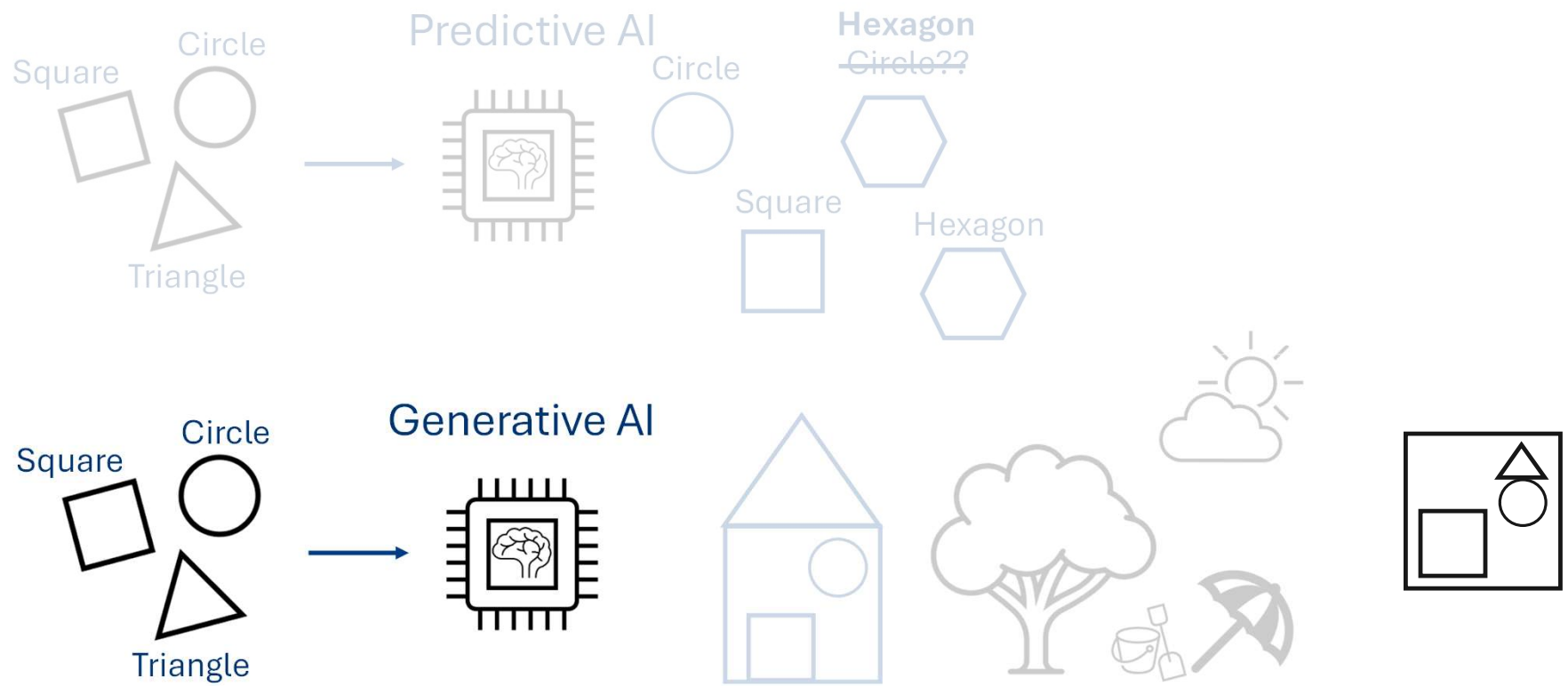
What do we mean by “AI”?



Differentiating Artificial Intelligence



The workings of generative AI



Generative artificial intelligence

genAI

AI



Synthetic text generator



Conversation emulator



**Glorified search engine and
auto-complete with a big
need to please you**



Stochastic Parrot



Stochastic Parrot

random but according to a **probability** distribution



to **repeat** exactly what someone else has said
without thinking or understanding the
meaning behind the words

Stochastic Parrot

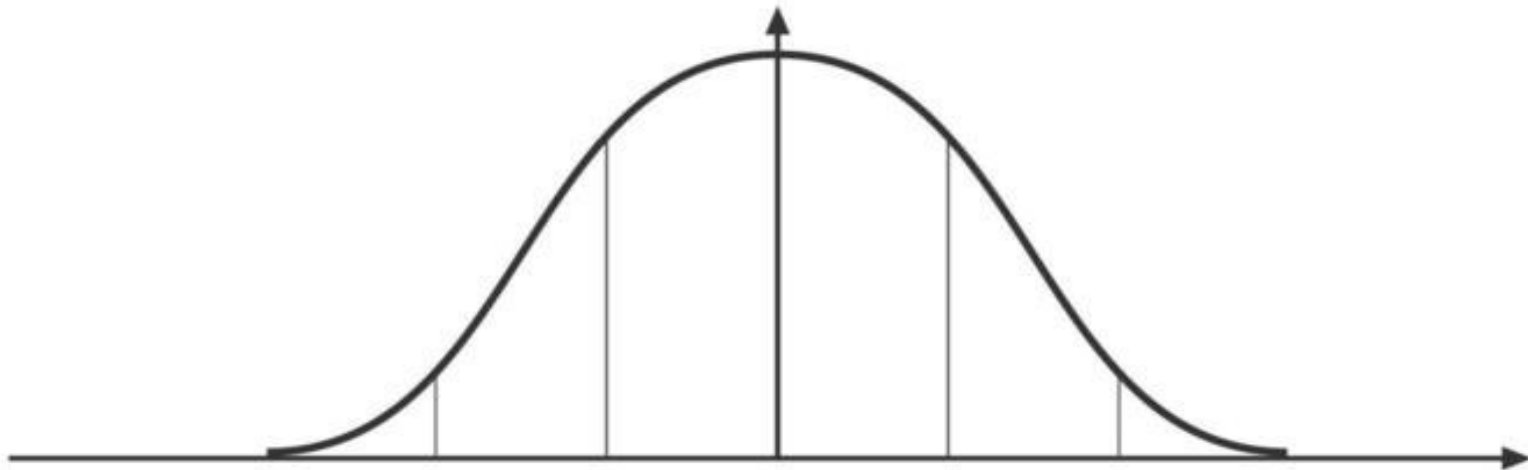
random but according to a **probability** distribution



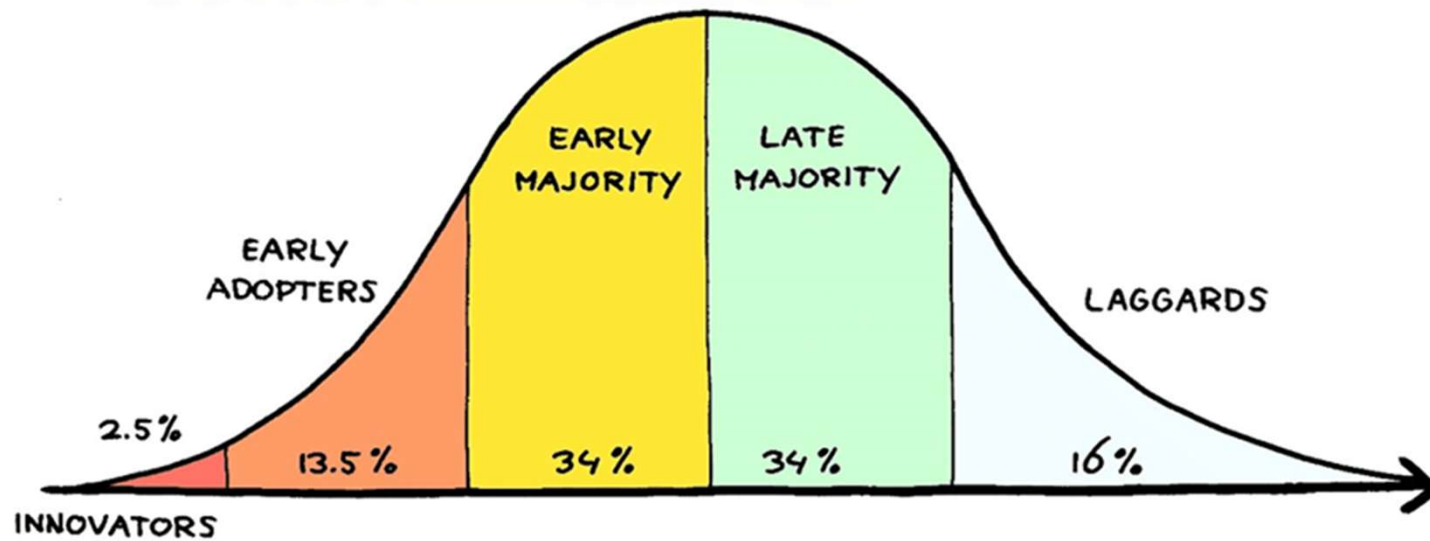
The Normalizing Process of Generative AI



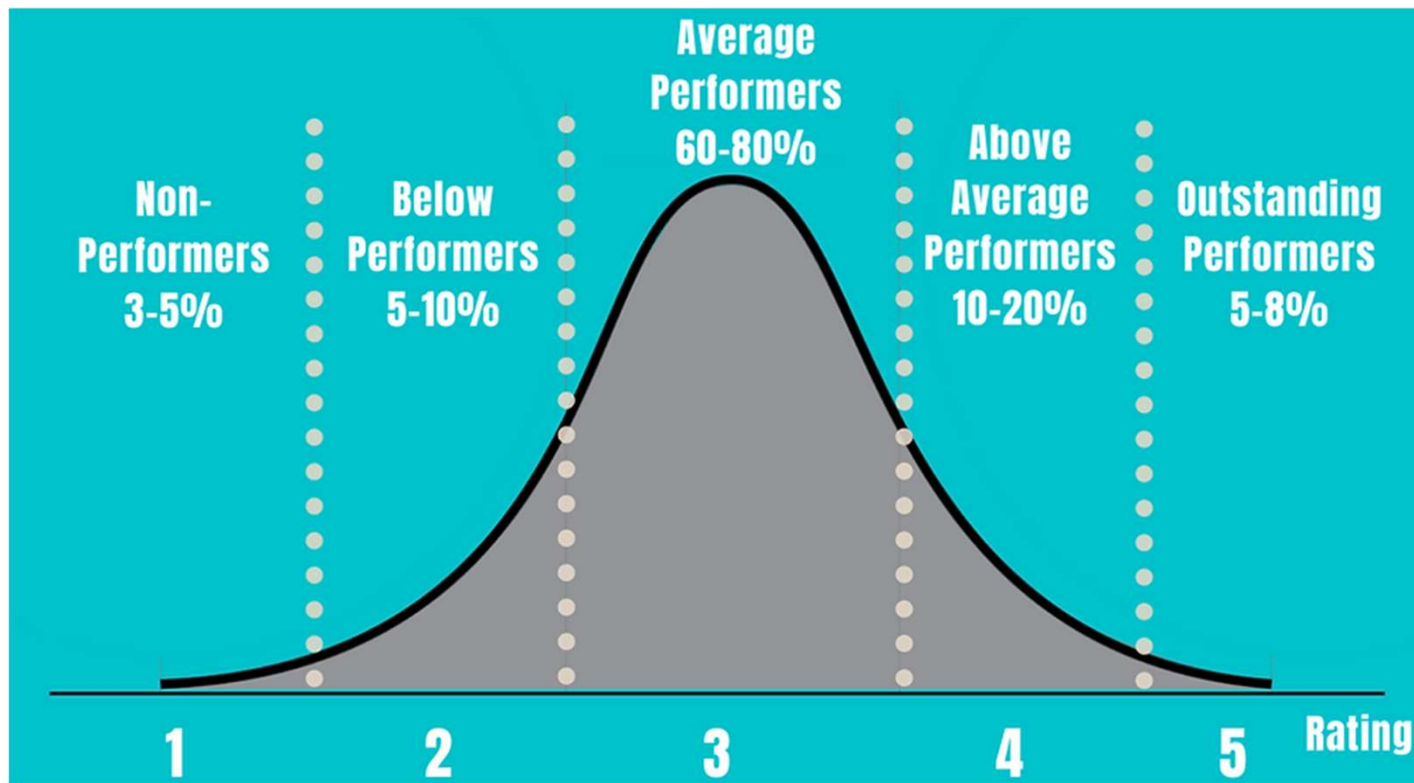
Normal distribution (or the 'bell curve')



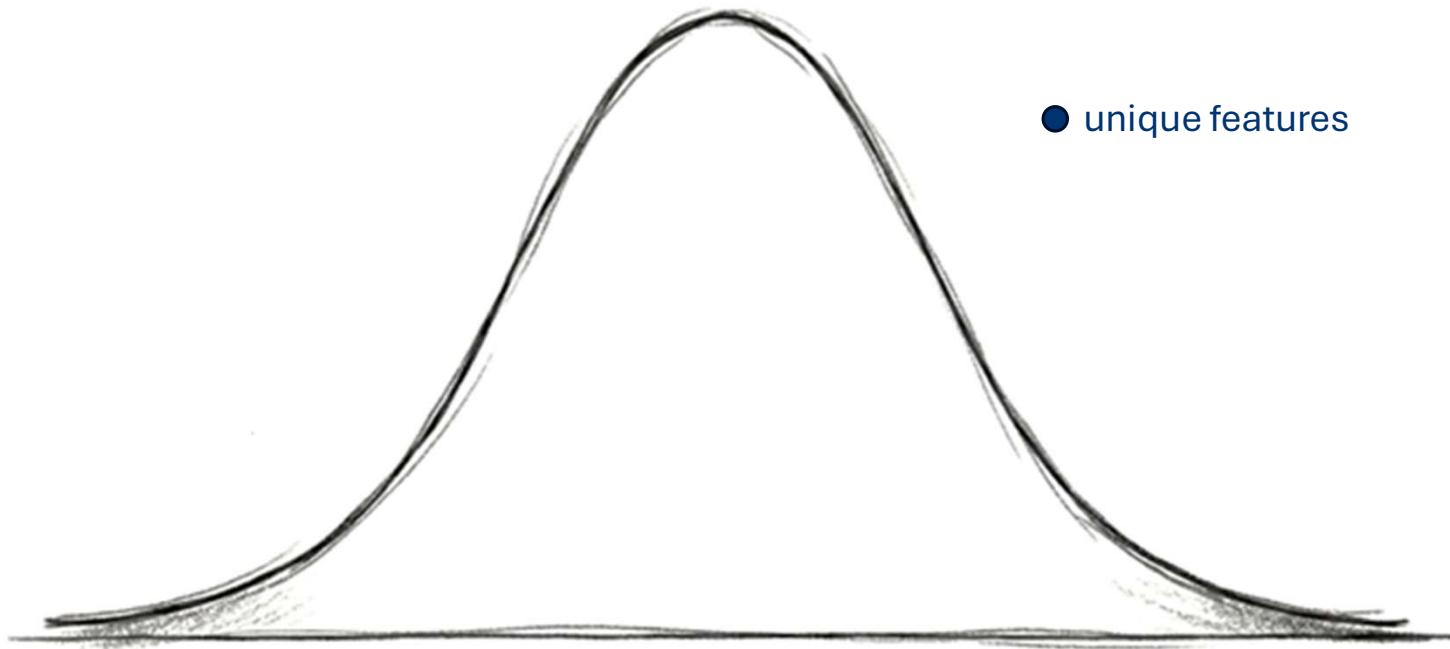
The bell curve: technology adoption



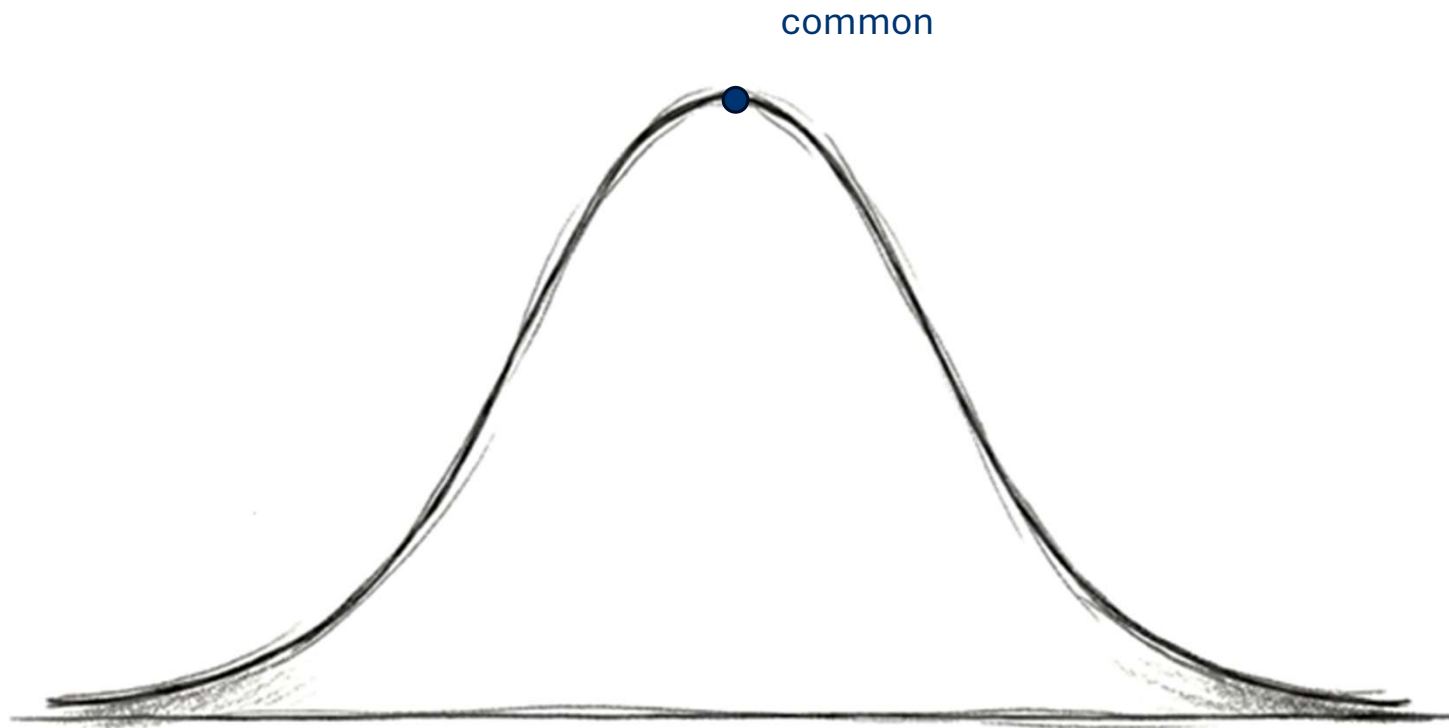
The bell curve: performance



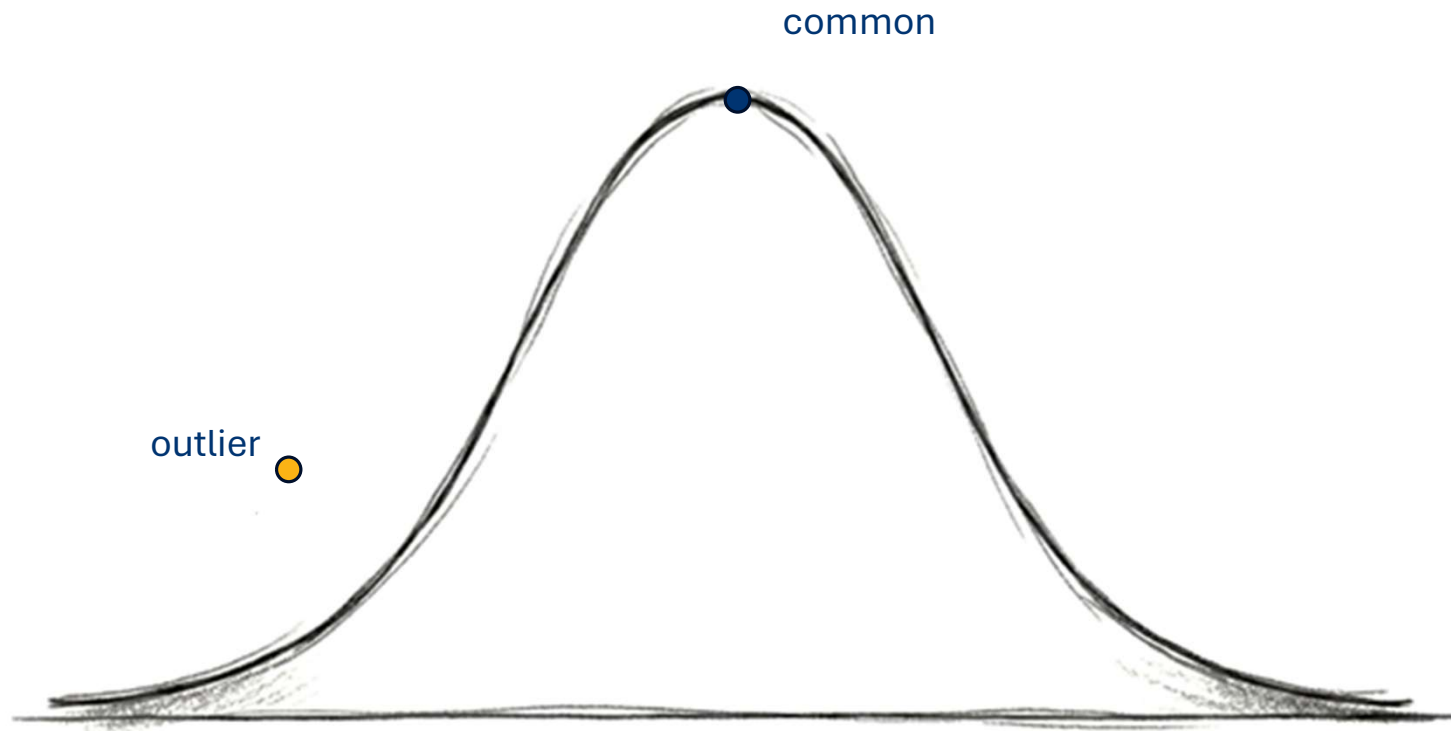
Erasure or flattening of unique contexts



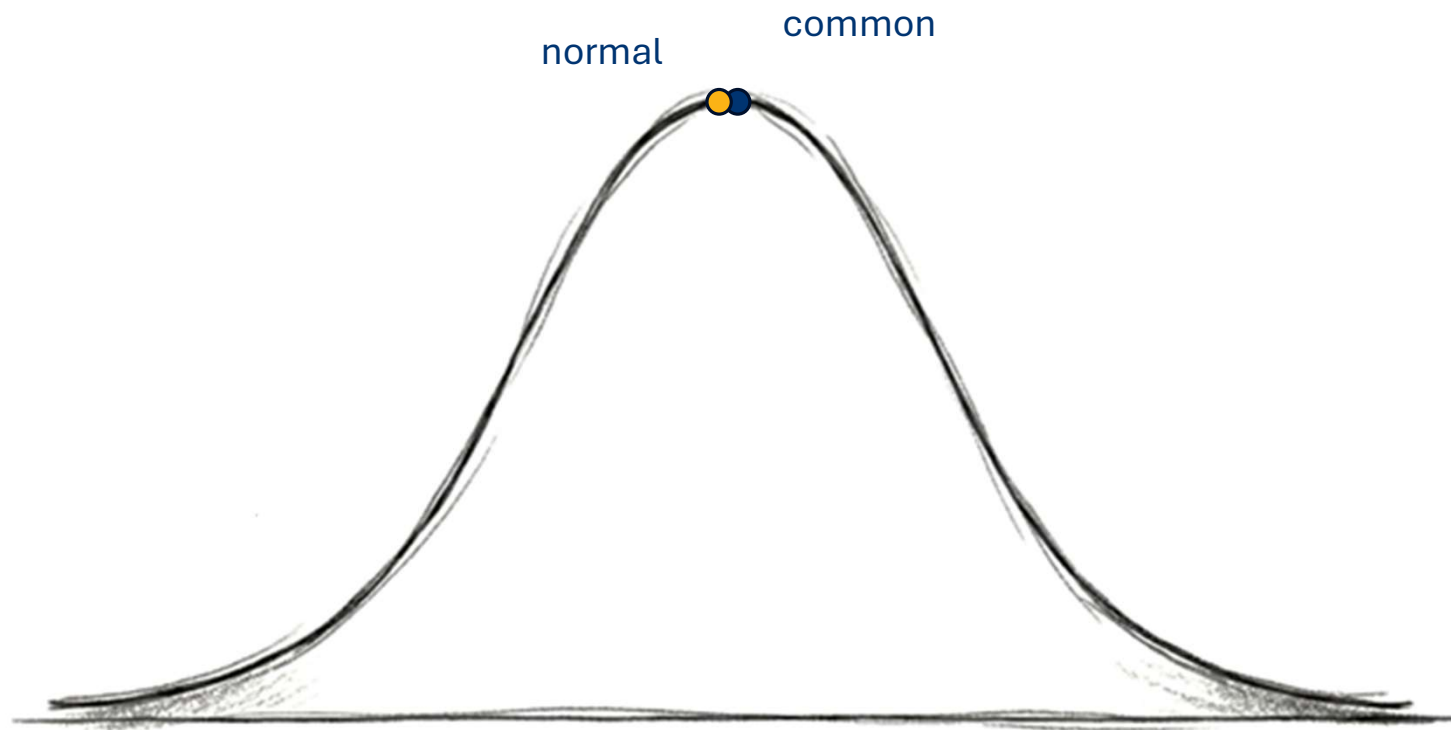
Erasure or flattening of unique contexts



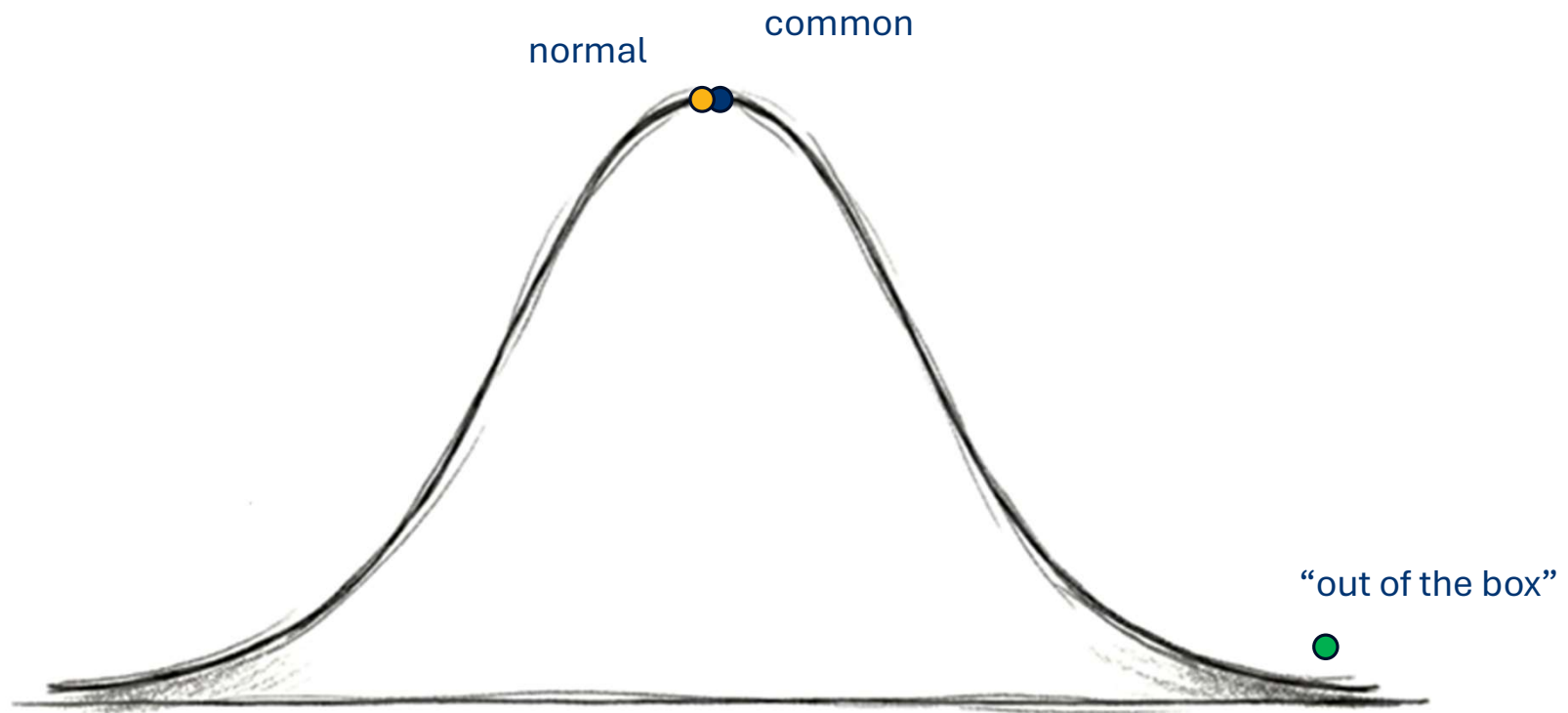
Erasure or flattening of unique contexts



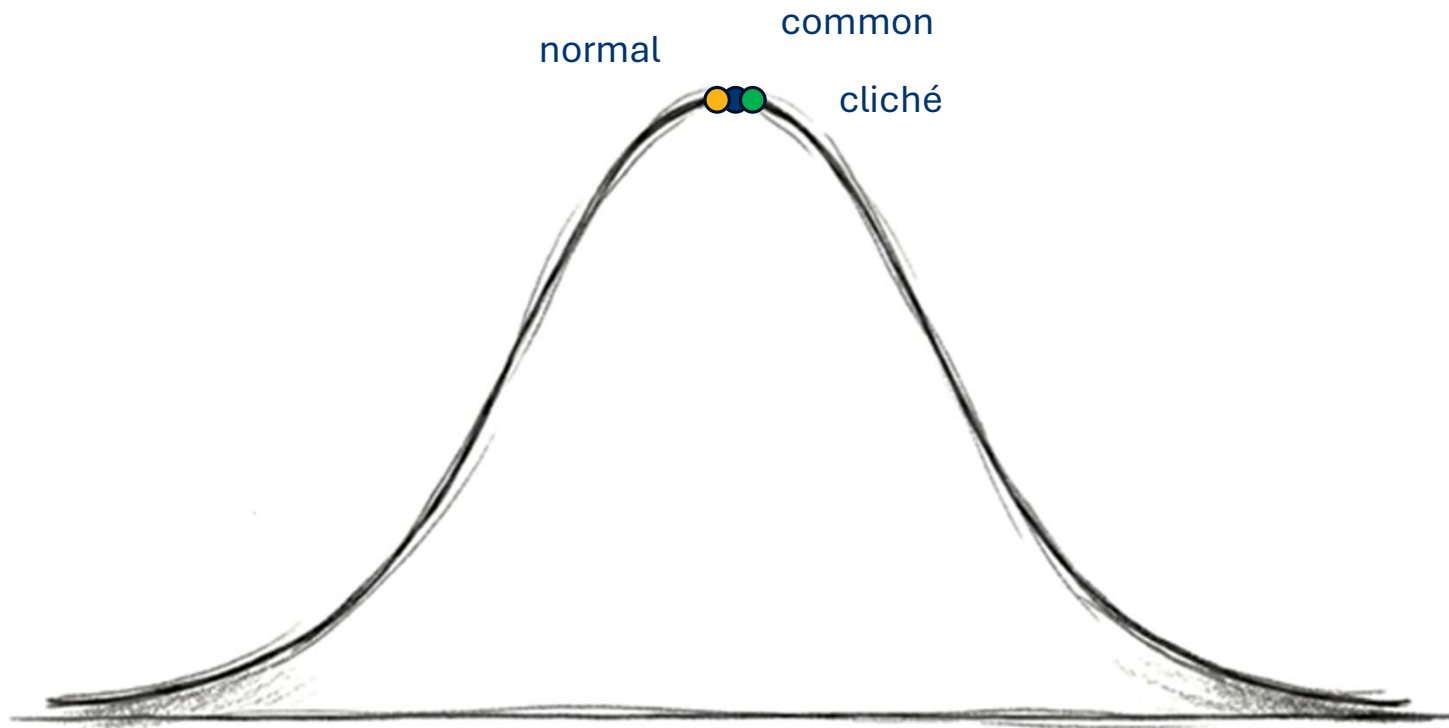
Erasure or flattening of unique contexts



Erasure or flattening of unique contexts



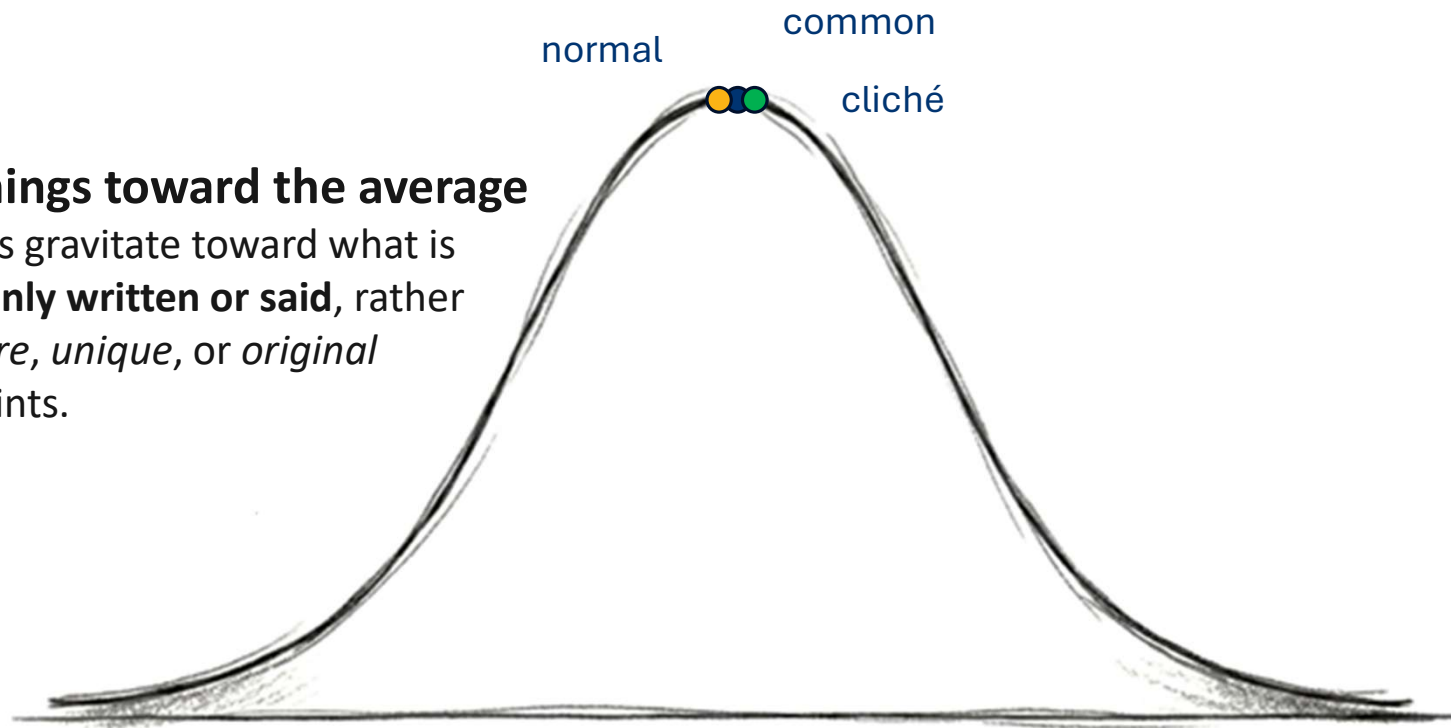
Erasure or flattening of unique contexts



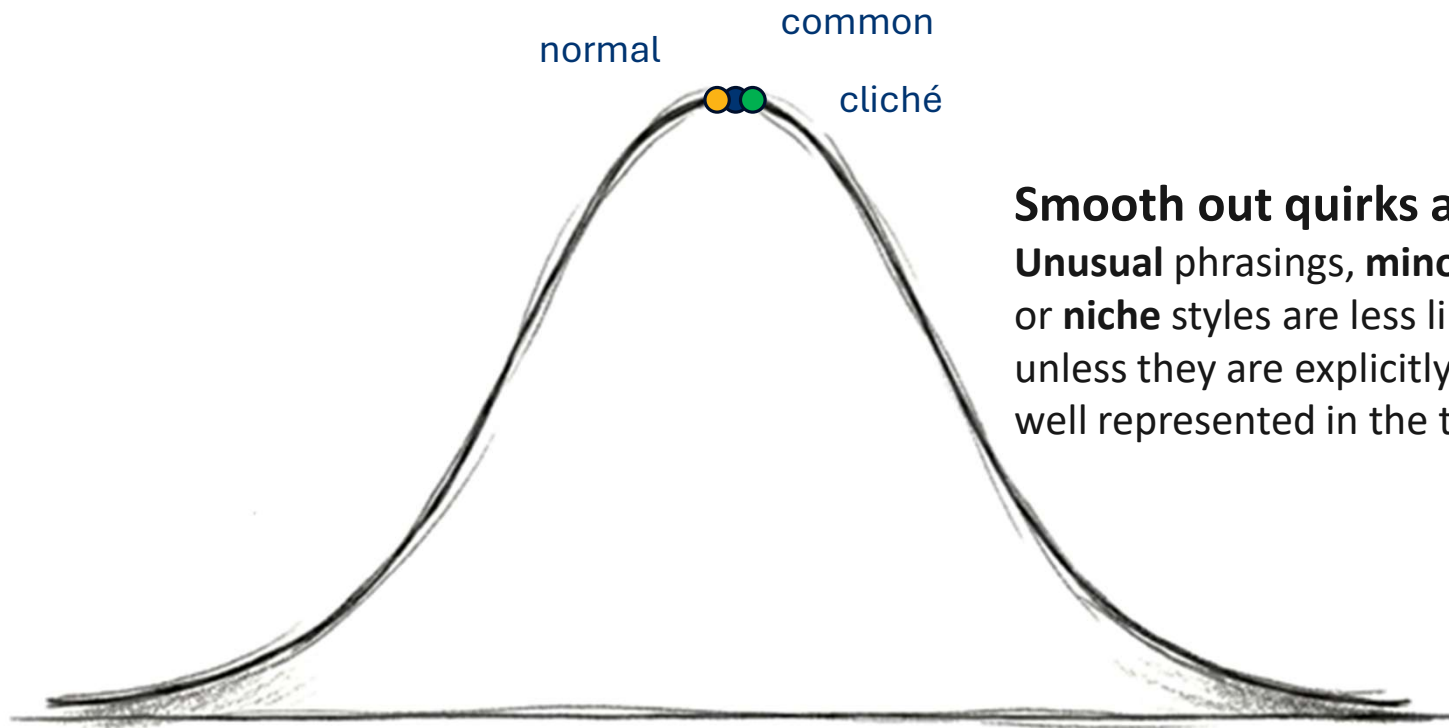
Generative AI as a normalizing engine

Pull things toward the average

Answers gravitate toward what is **commonly written or said**, rather than *rare, unique, or original* viewpoints.



Generative AI as a **normalizing engine**



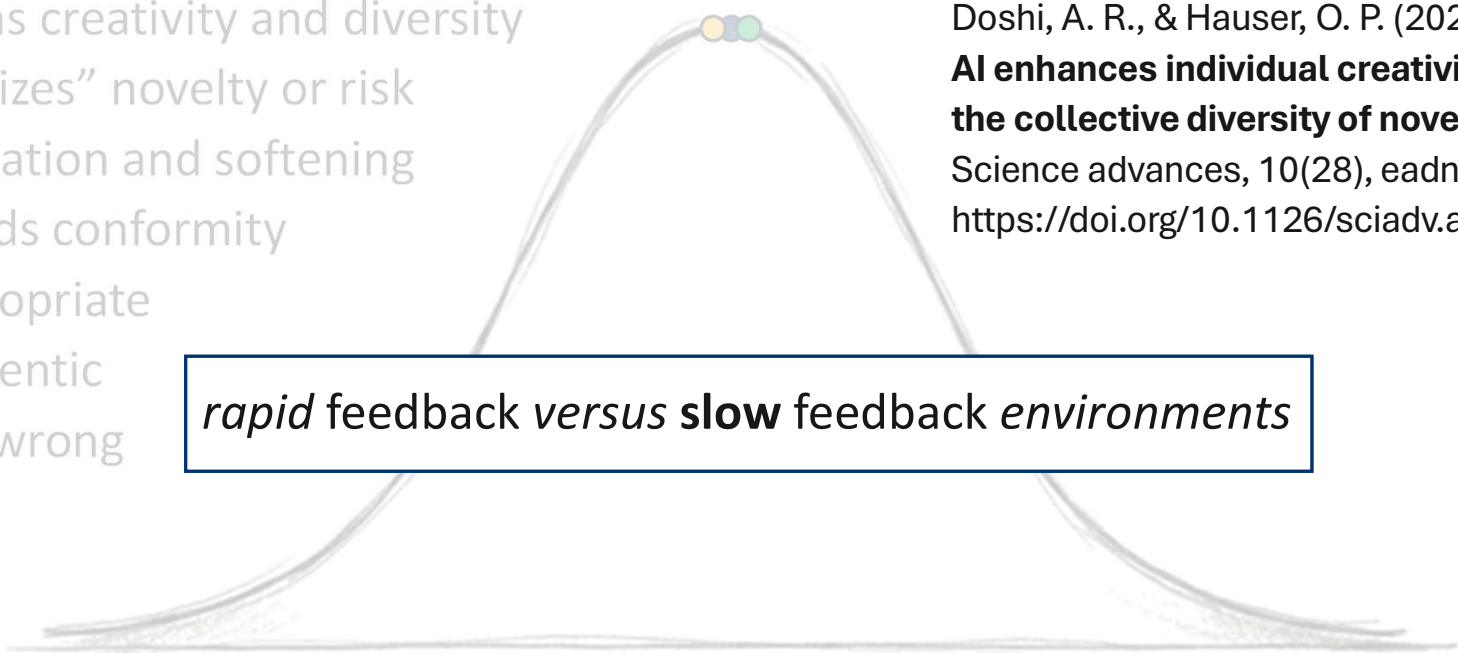
Smooth out quirks and outliers
Unusual phrasings, **minority** opinions,
or **niche** styles are less likely to appear
unless they are explicitly prompted and
well represented in the training data.

Generative AI as a normalizing engine

Flattens creativity and diversity
“Penalizes” novelty or risk
Sterilization and softening
Rewards conformity
Inappropriate
Inauthentic
Being wrong

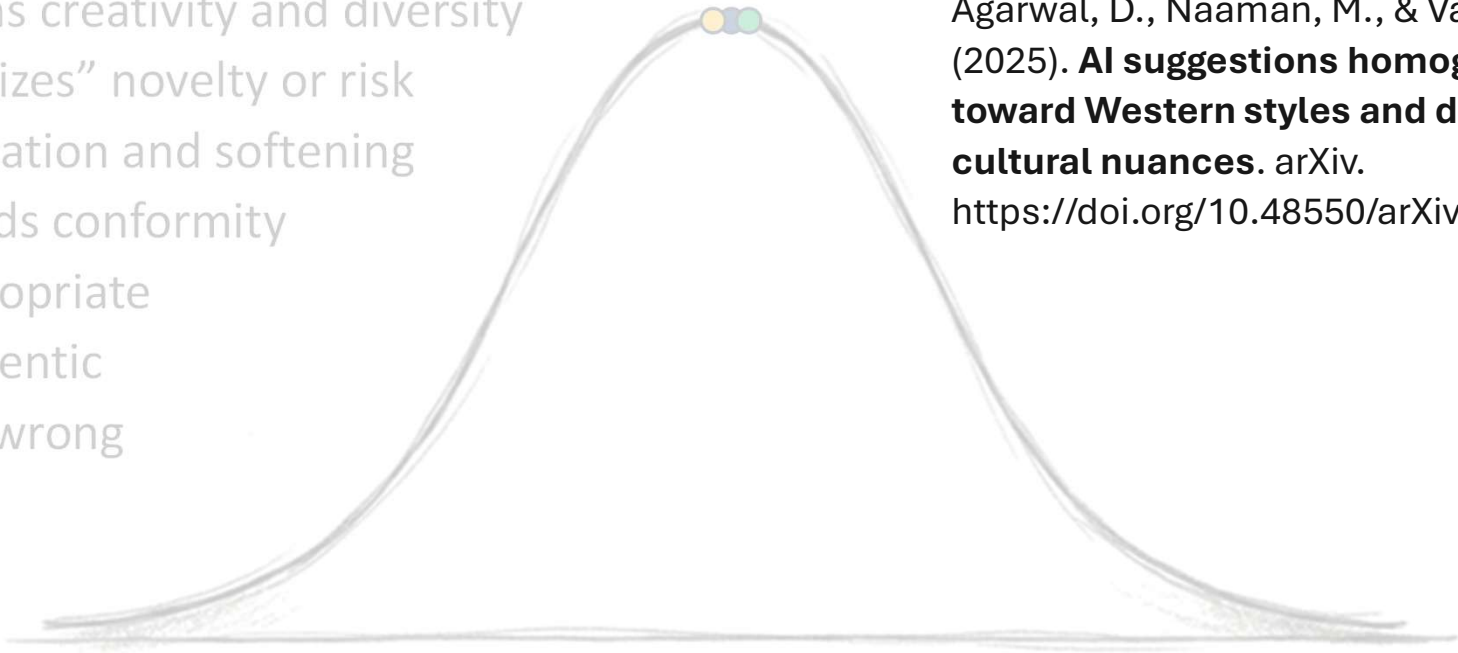
Doshi, A. R., & Hauser, O. P. (2024). **Generative AI enhances individual creativity but reduces the collective diversity of novel content.** Science advances, 10(28), eadn5290. <https://doi.org/10.1126/sciadv.adn5290>.

*rapid feedback versus **slow** feedback environments*



Generative AI as a normalizing engine

Flattens creativity and diversity
“Penalizes” novelty or risk
Sterilization and softening
Rewards conformity
Inappropriate
Inauthentic
Being wrong



Agarwal, D., Naaman, M., & Vashistha, A.
(2025). **AI suggestions homogenize writing
toward Western styles and diminish
cultural nuances.** arXiv.
<https://doi.org/10.48550/arXiv.2409.11360>



Normalizing program evaluation

Threatens **validity** and **quality** of
evaluation

- Superficial
- Hallucinati

Complicate **accountability** and
ownership

easy to underestimate because the tools feel so
“helpful” and **authoritative**

isks

Embed **biases** and **unfairness** in
decisions

- Encoding existing biases and assumptions
- Unequal treatment and inconsistencies

Cause **over-reliance** and loss of
critical expertise

- Erosion of evaluative thinking
- Reduction of genuine engagement



The Human-AI-Human Process

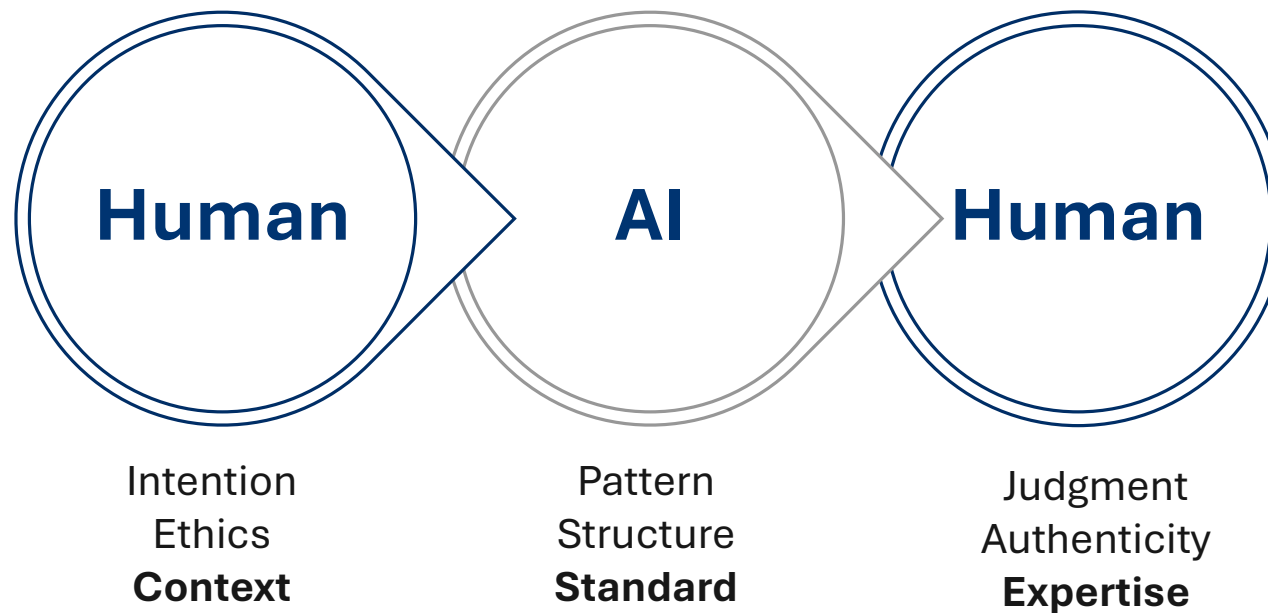
The Centre for the Advancement
of Teaching and Learning



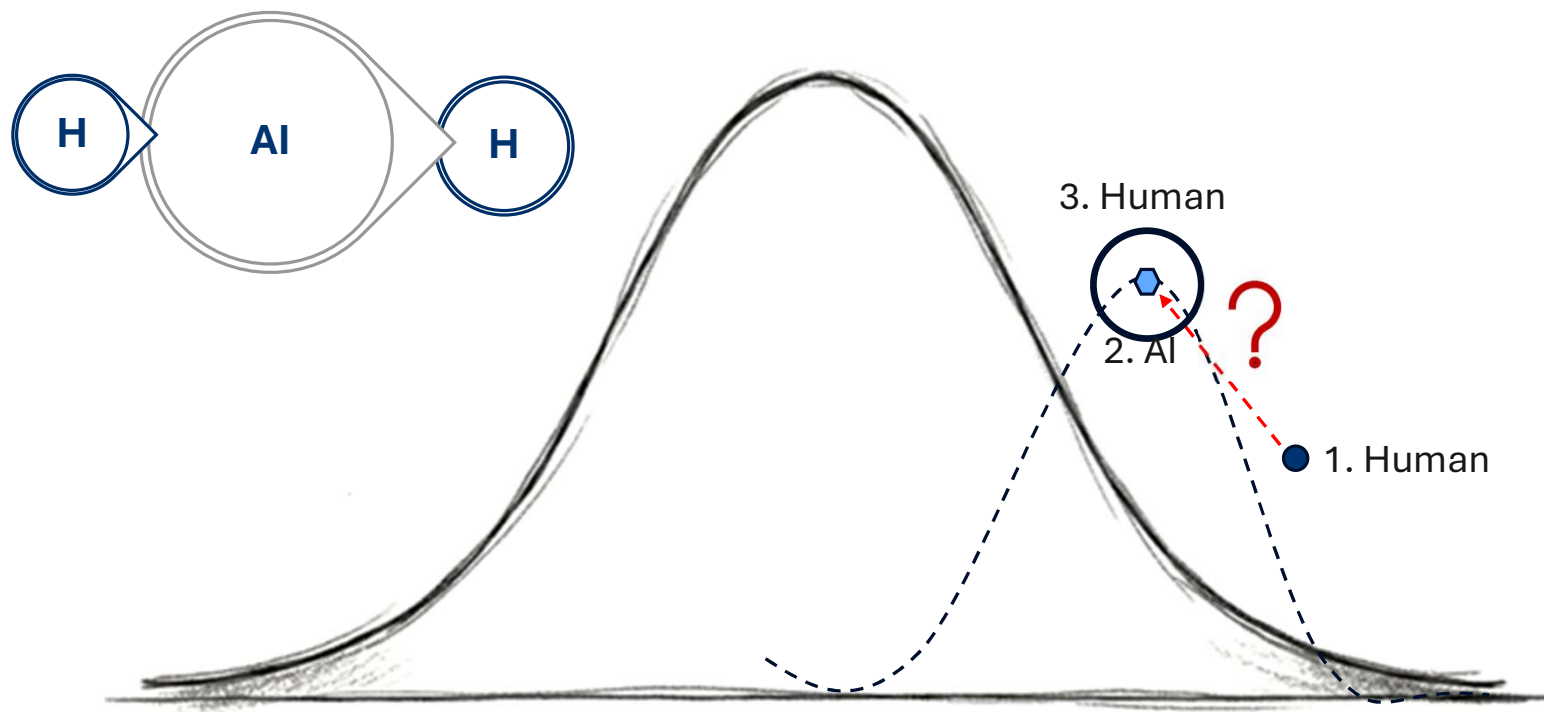
UM

Human-AI-Human Process

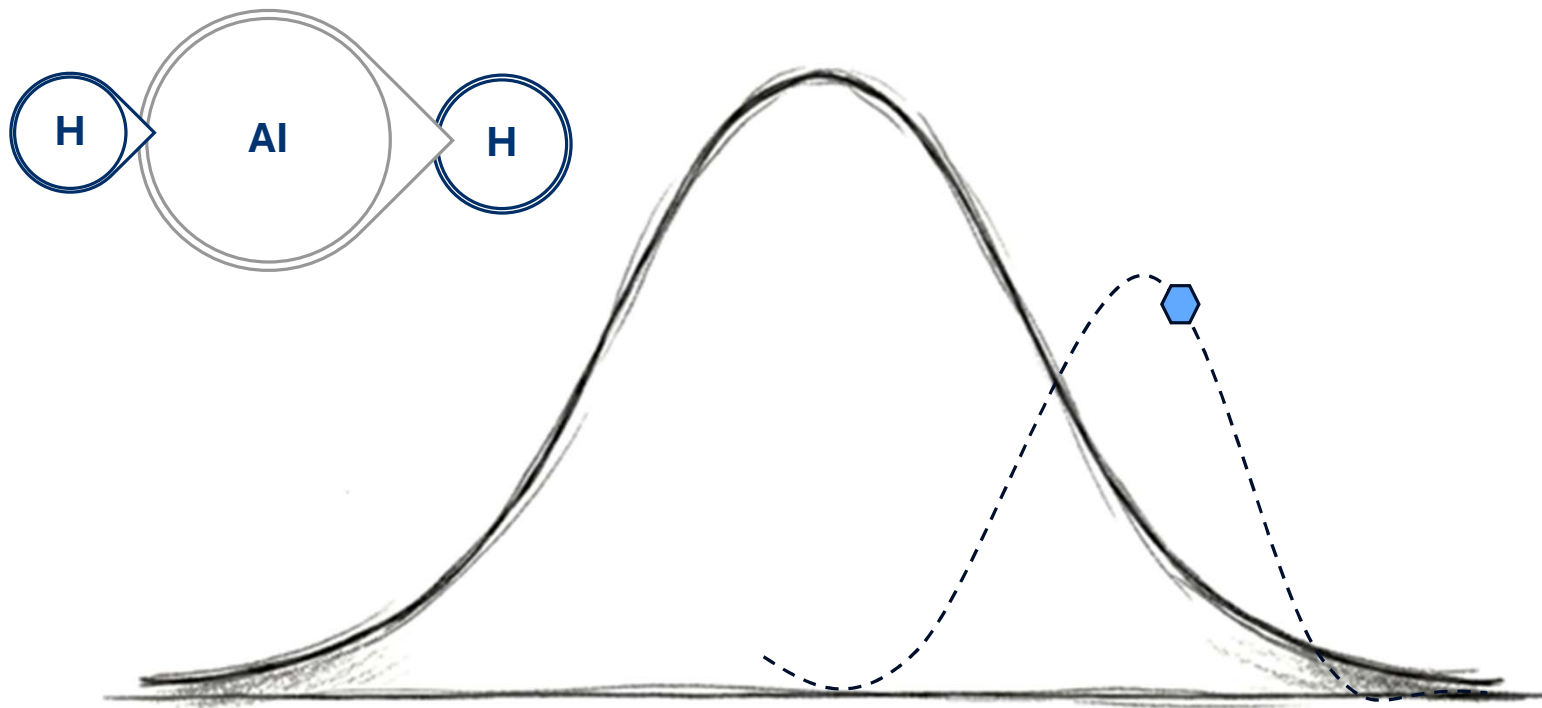
AI may sit between **human acts** of judgment, reflection, and authorship.



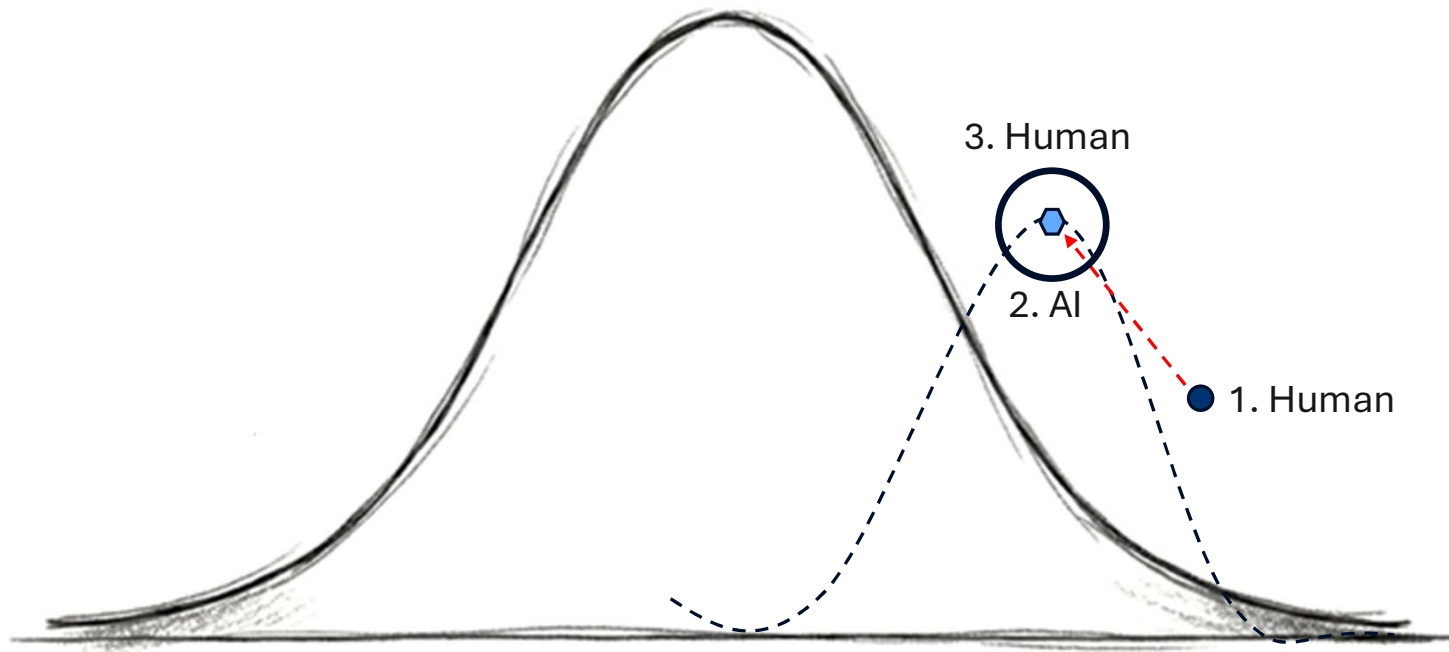
The Bell Curve of GenAI: **Uncritical** Acceptance



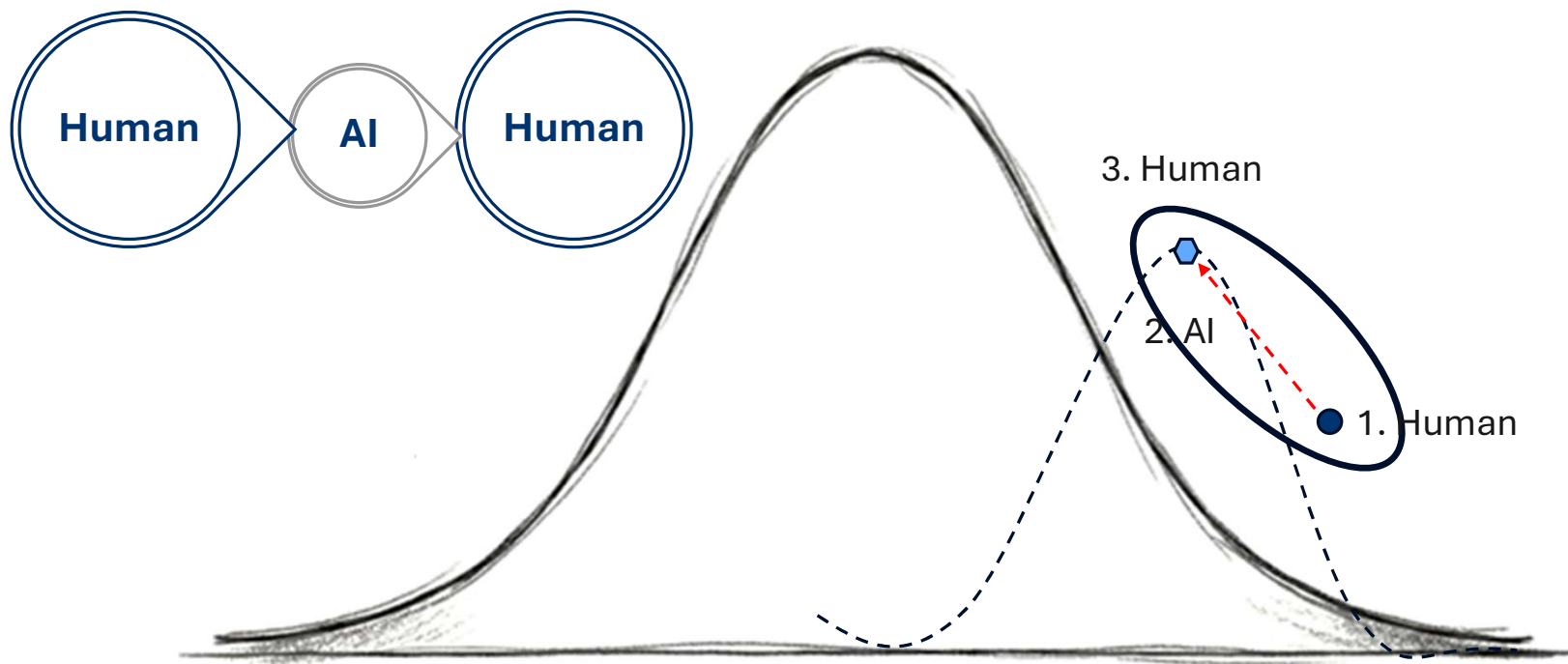
The Bell Curve of GenAI: **Uncritical** Acceptance



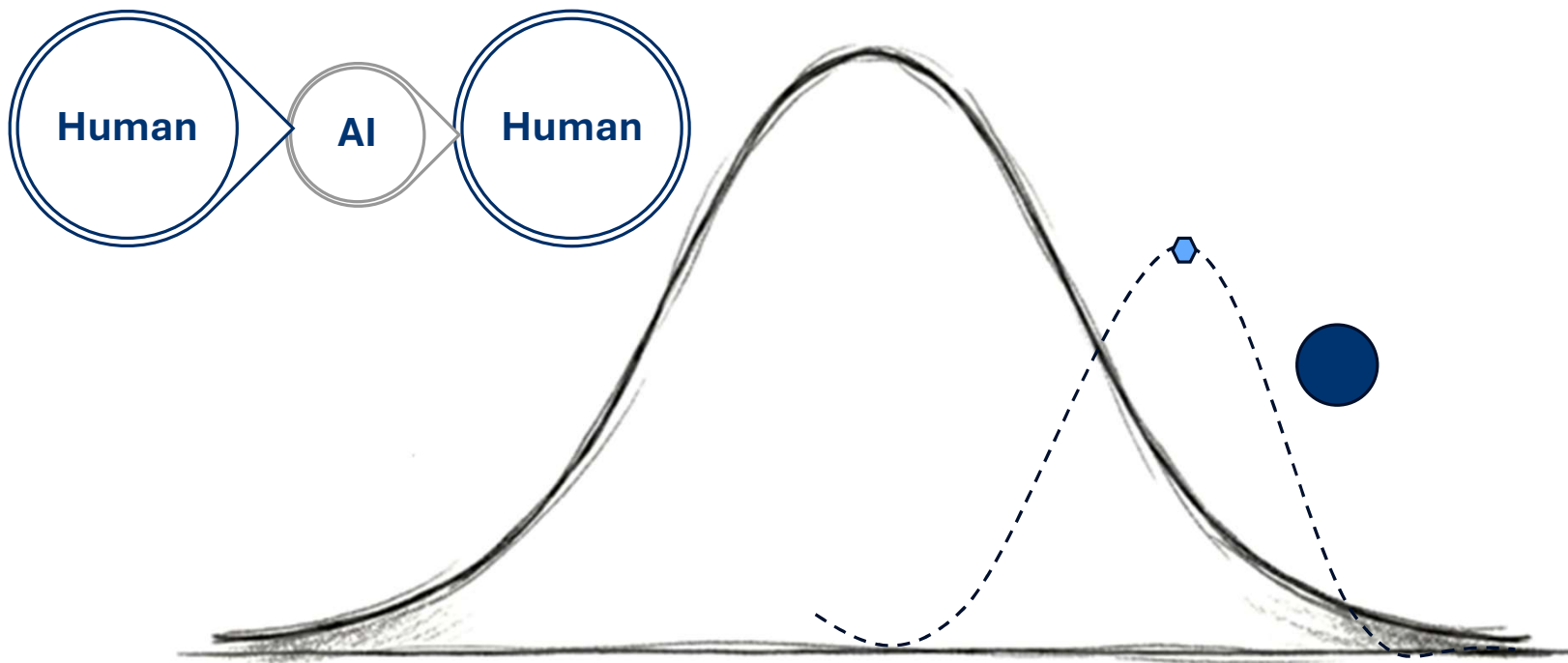
The Bell Curve of GenAI: **Uncritical** Acceptance



The Bell Curve of GenAI: **Critical** Evaluation

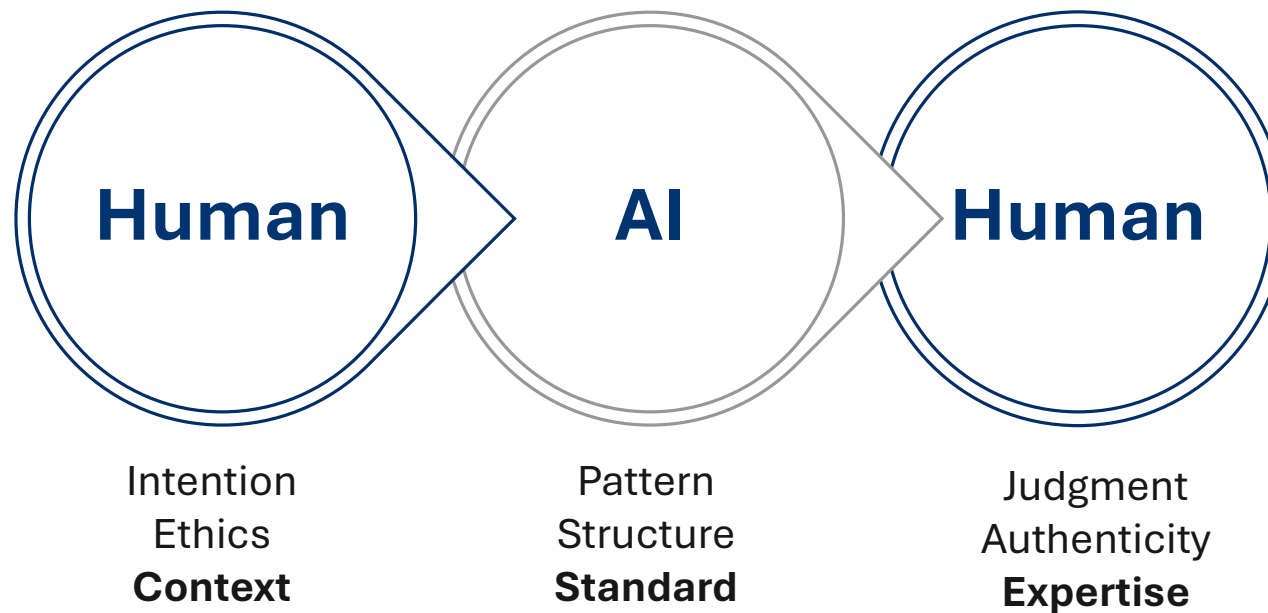


The Bell Curve of GenAI: **Critical** Evaluation



Human-AI-Human Process

AI may sit between **human acts** of judgment, reflection, and authorship.



Some advice

- step back from the hype
- learn about the technology, try it out, stress test it
- be realistic and skeptical
- know your context, know the process
- use it to **support what you know**, *not to do what you don't know*
- slow down





Generative Artificial Intelligence

For Program Evaluation