

Evolving Intelligence for Social Good:

Towards Automated, Explainable, Reliable, and Resilient AI

Amir H Gandomi

*Professor of Data Science at University of Technology Sydney
Distinguished Professor at Óbuda University, Budapest*

AI in Action – An overview

What is Artificial Intelligence (AI) ?



Sora AI-generated video.
OpenAI, 2025

AI: The ability of a machine to perform cognitive functions we associate with human minds, such as perceiving, reasoning, learning

Exponential increases in computing power and storage

1965: Moore's law

Explosion of data

1991: Opening of the World Wide Web

1805: Legendre lays the ground work for machine learning

Algorithmic advancements

9/27/2026

2009: Ng uses GPUs to train deep learning models more efficiently

2011: IBM Watson beats Jeopardy

2012: Google demonstrate effectiveness of deep learning for image recognition

2014: Number of mobile devices exceeds number of humans

2017: AlphaZero beats AlphaGo Zero

2022: ChatGPT

2018

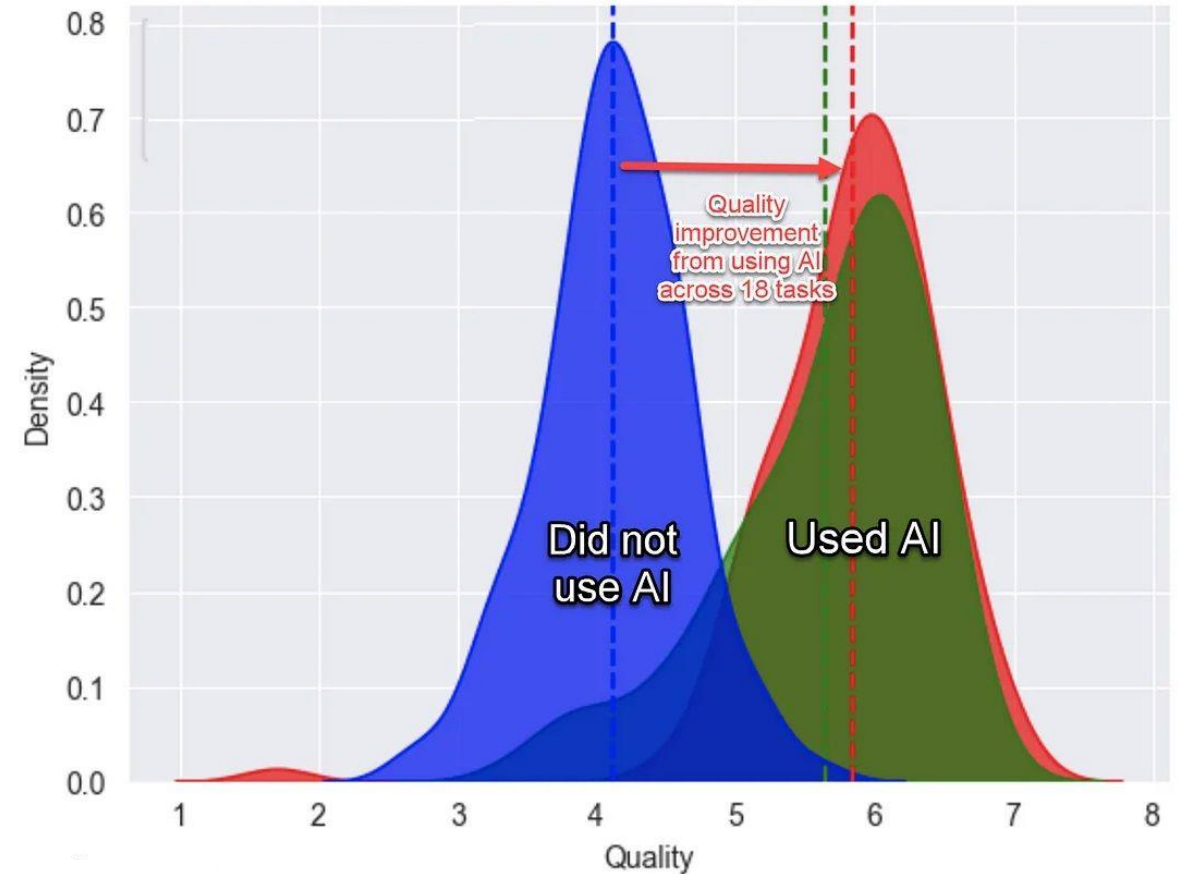
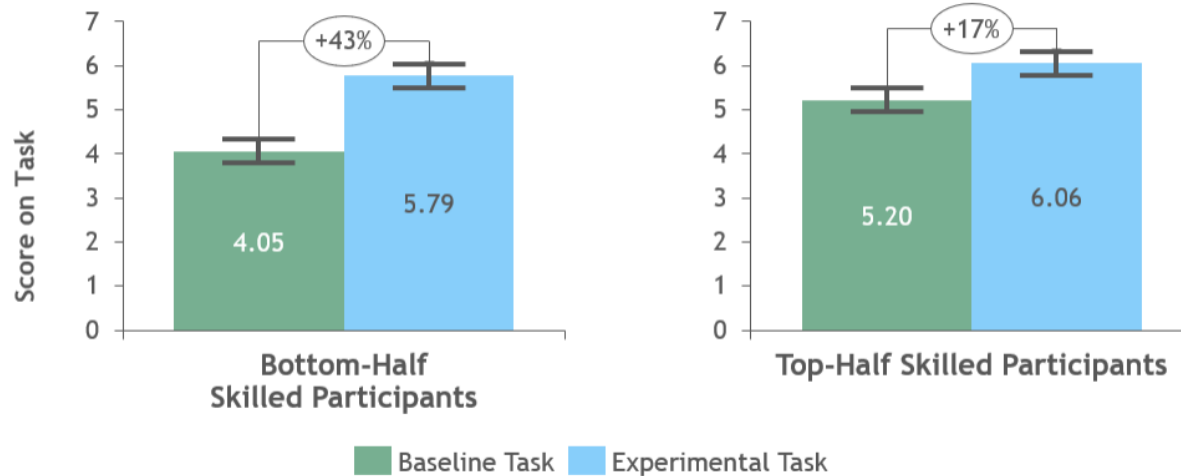
Agentic AI 2023

A convergence of algorithmic advances, data proliferation, and tremendous increases in computing power and storage has propelled AI from hype to reality

<https://www.mckinsey.com/business-functions/mckinsey-analytics/>

Effect of AI in Performance

25+% increase in speed,
40+% improvement in output quality, and
12+% more tasks completed

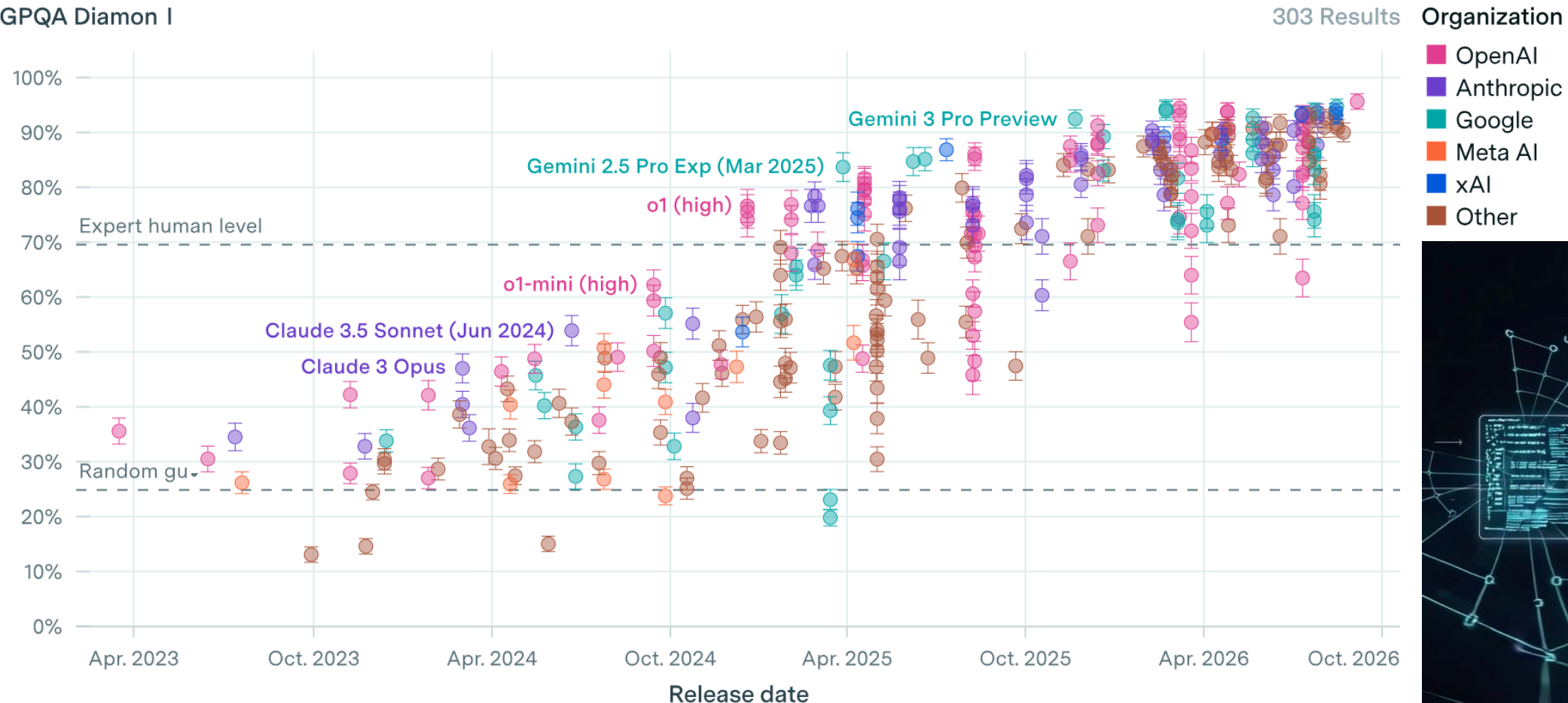


Dell'Acqua, et al., Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality (September 15, 2023). Harvard Business School Technology & Operations Mgt. Unit Working Paper No. 24-013

AI Vs. Human Expertise in Scientific Reasoning (Physics)

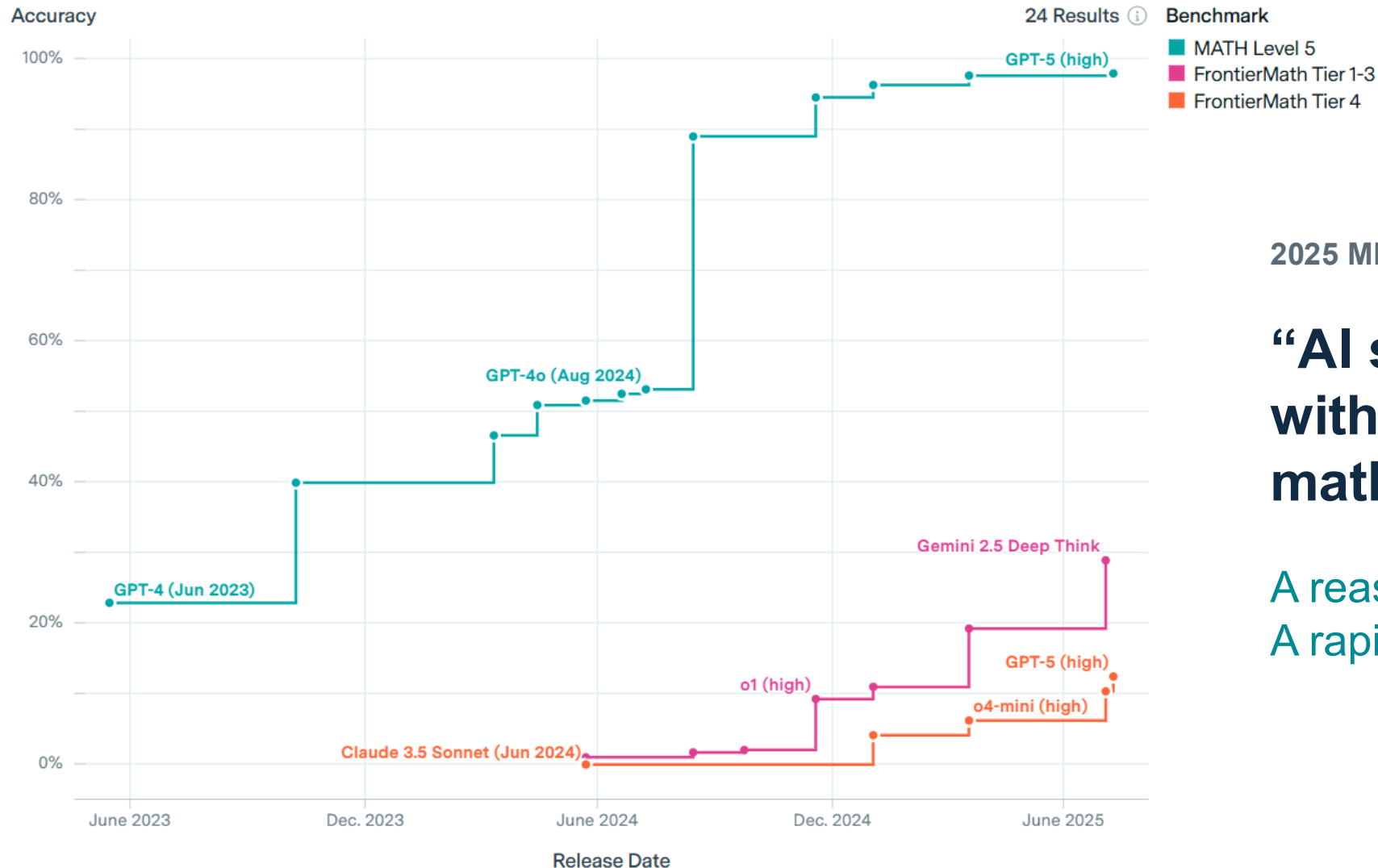
AI performance on a set of Ph.D.-level science questions

GPQA Diamond I



My 2025 slide had a short shelf life

Frontier performance across benchmarks



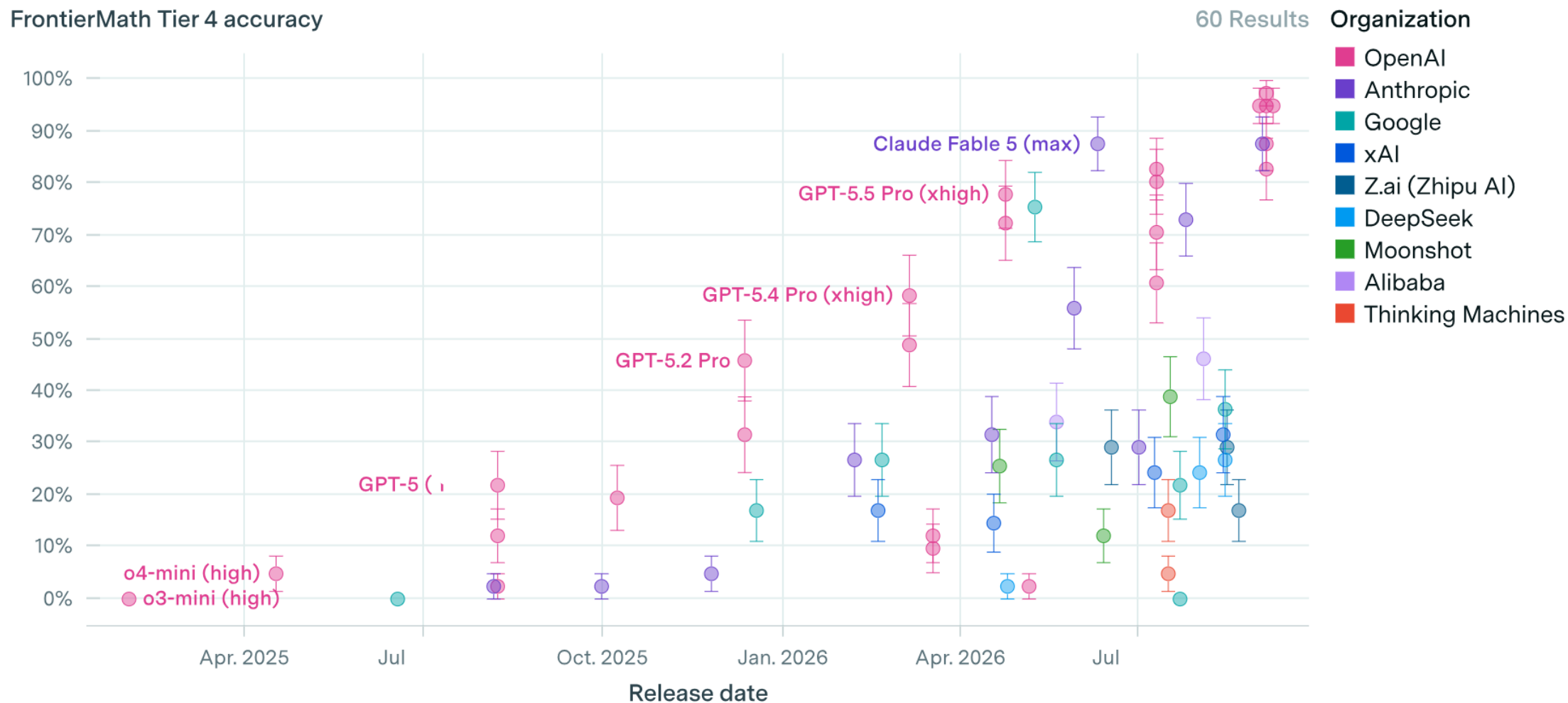
2025 MESSAGE

“AI still struggles with frontier mathematics.”

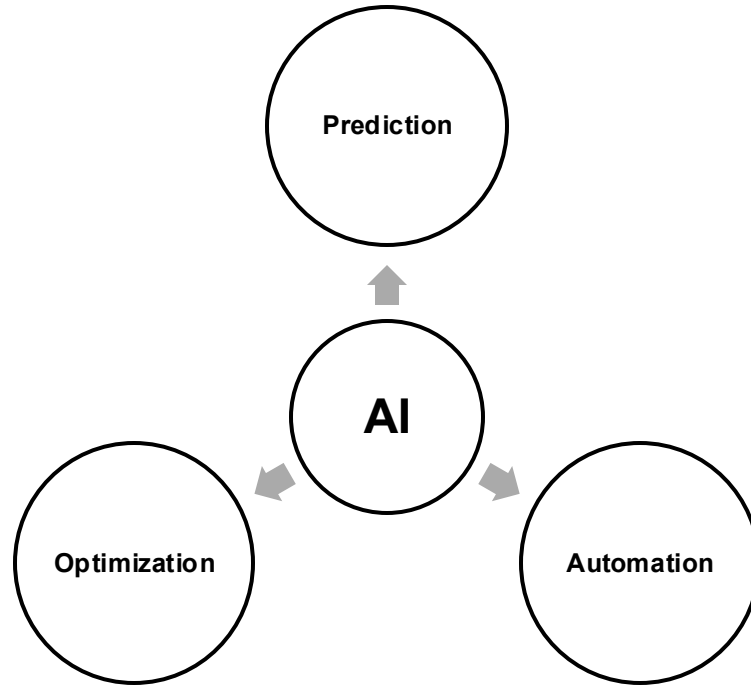
A reasonable conclusion.
A rapidly expiring slide.

Solving Frontier Math Problems (Current)

AI performance on the most challenging expert-level mathematics problems

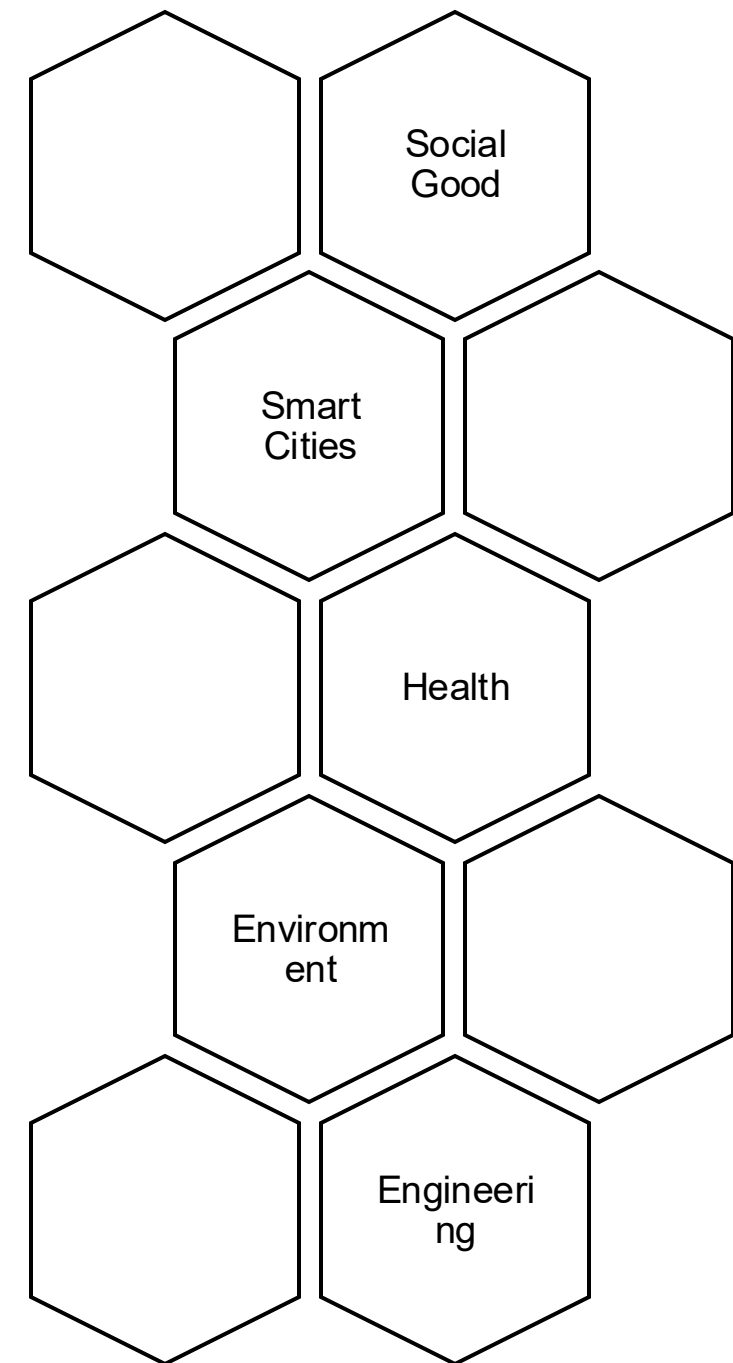


How we use AI in Action



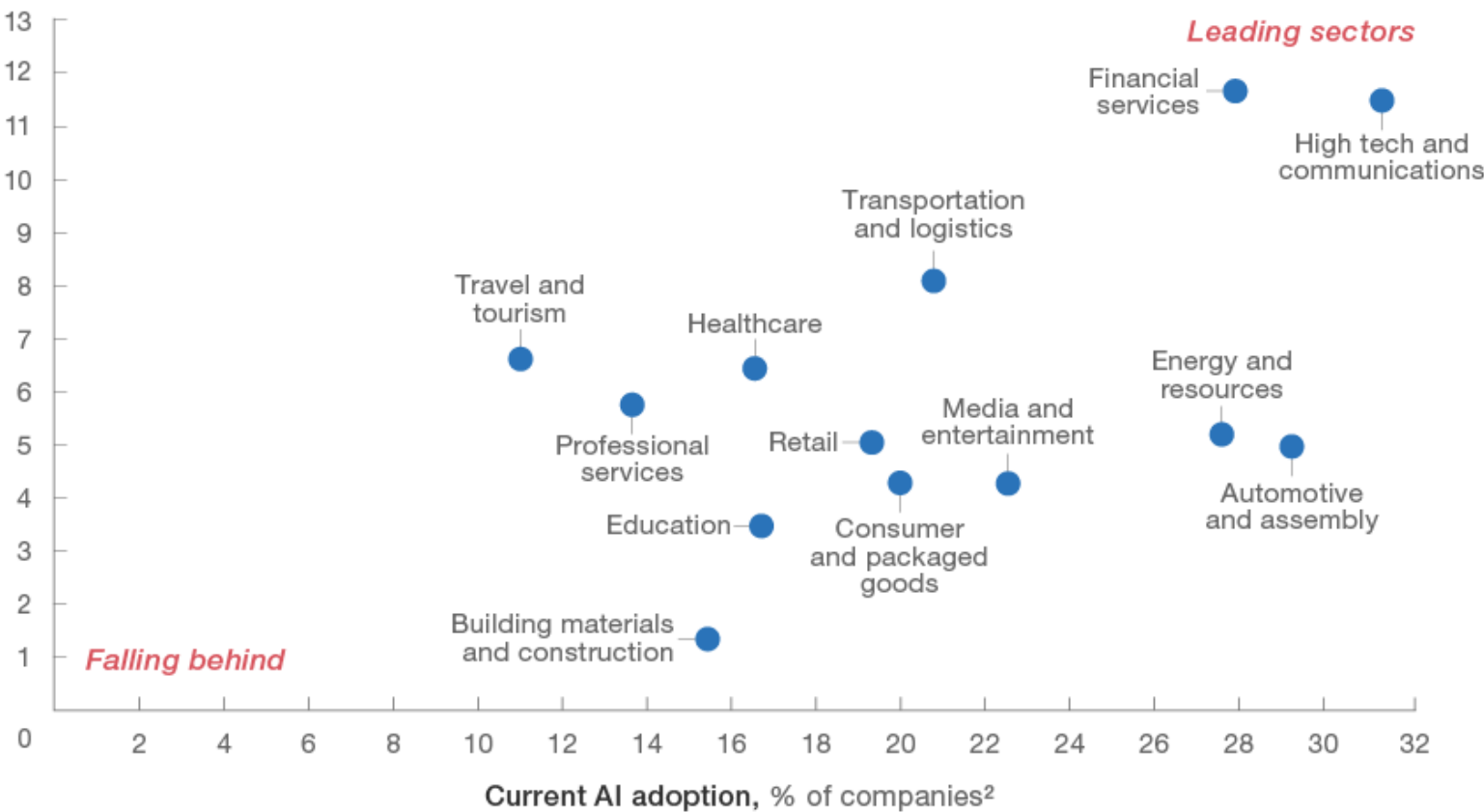
Award winning international figure:

- Rank 3rd Computer Scientist in AU
- Rank 29th researcher globally in the AI & Image Processing subfield
- 84,000+ citations (H-index=125+)
- Receiving prestigious Int. awards, from Sigma Xi, XPRIZE, IEEE, ASCE, ICE, IMech, IFAC, IWSHM



Leaders in the adoption of artificial intelligence also intend to **invest more in the near future** compared with laggards.

Future artificial intelligence (AI)-demand trajectory, % change in AI spending over next 3 years¹



Sora AI-generated video. OpenAI, 2025

¹Estimated average, weighted by company size; demand trajectory based on midpoint of range selected by survey respondent

²Adopting 1 or more AI technologies at scale or in business core; weighted by company size.

Source: McKinsey Global Institute AI adoption and use survey; McKinsey Global Institute analysis

Social good starts with a consequential decision

HEALTH

Support more reliable assessment

INFRASTRUCTURE

Detect damage and monitor condition

WATER

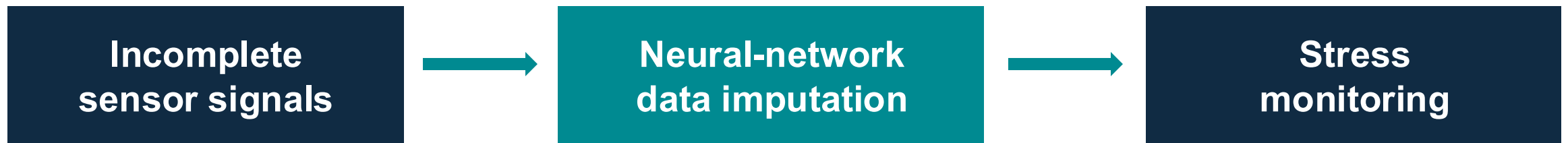
Anticipate floods and understand aquifers

ENERGY

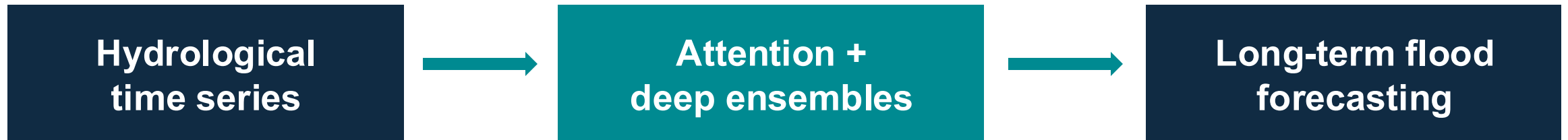
Plan cleaner, resilient supply

Anticipate hazards before they become failures

NAVAL HULL MONITORING



FLOOD FORECASTING



2026 recognition: Thomas Hawksley Gold Medal · ASCE–EWRI Best Research Paper Award

Support healthier people and environments

HEALTH

X-ray features



Fracture classification

WATER

**Groundwater
observations**



Diagnostic calibration

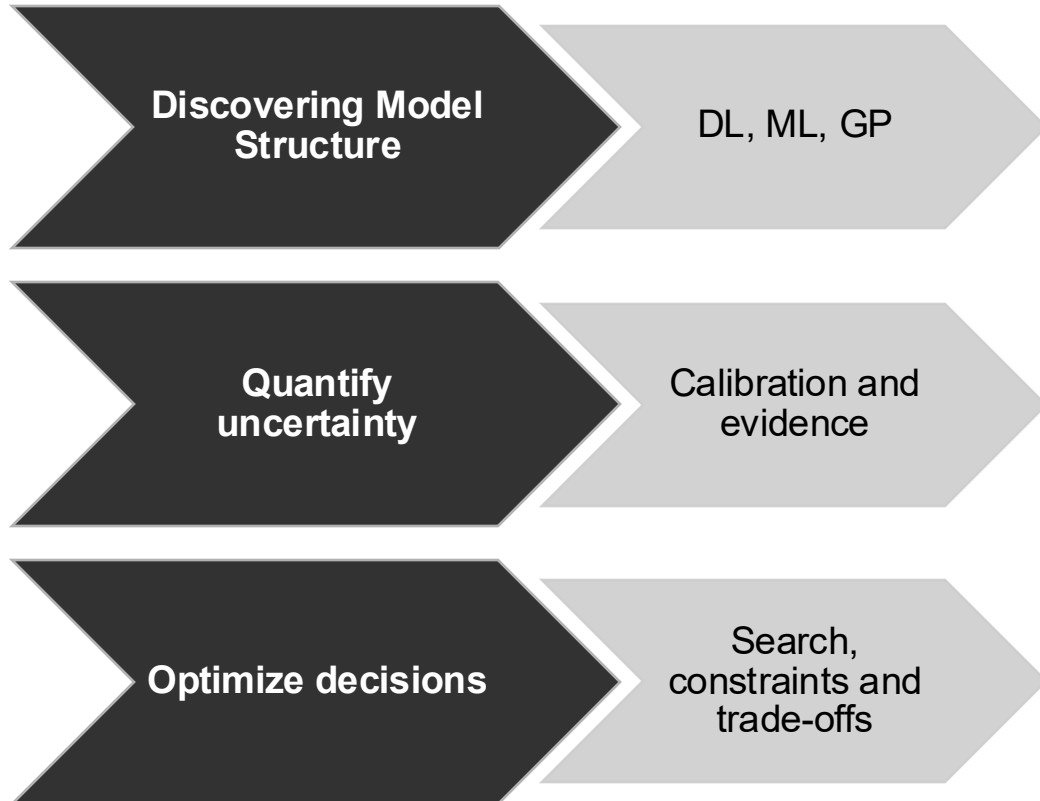
ENERGY

Regional resource data



Hybrid-system siting

General Outlines



What is Genetic Programming

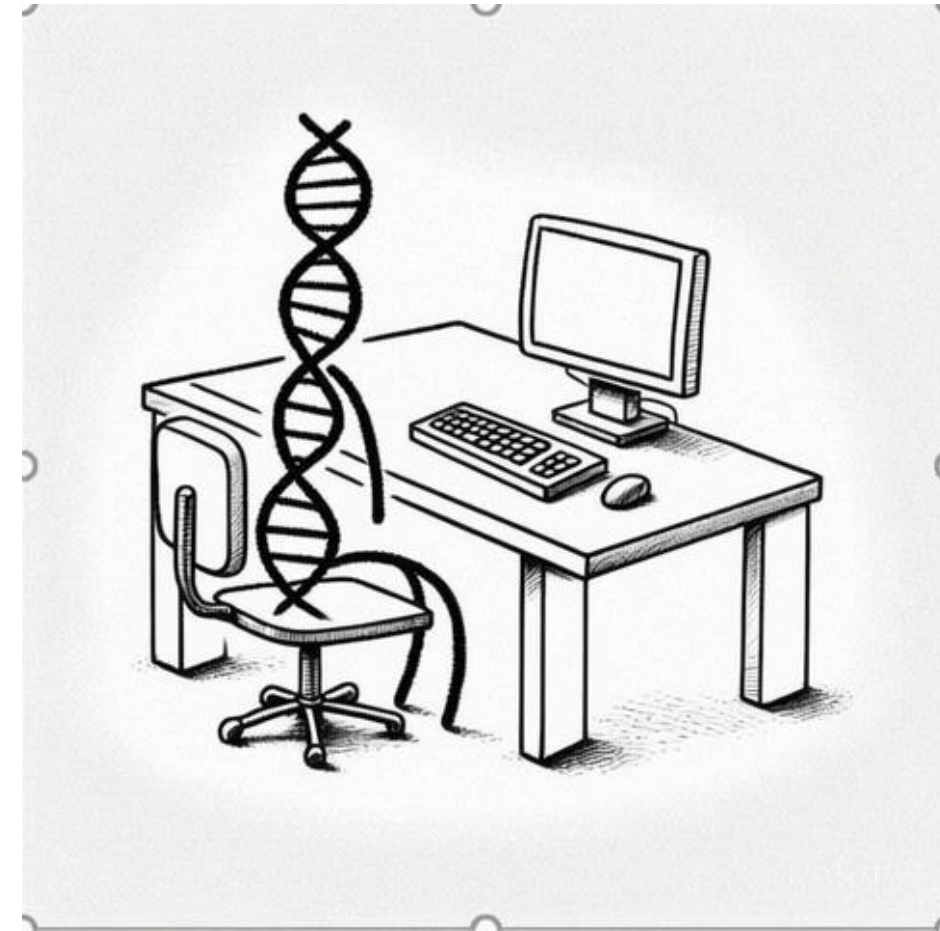
A Software with Evolution as the programmer!

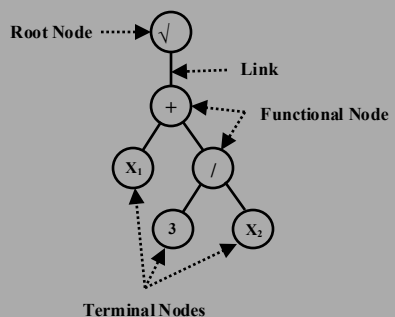
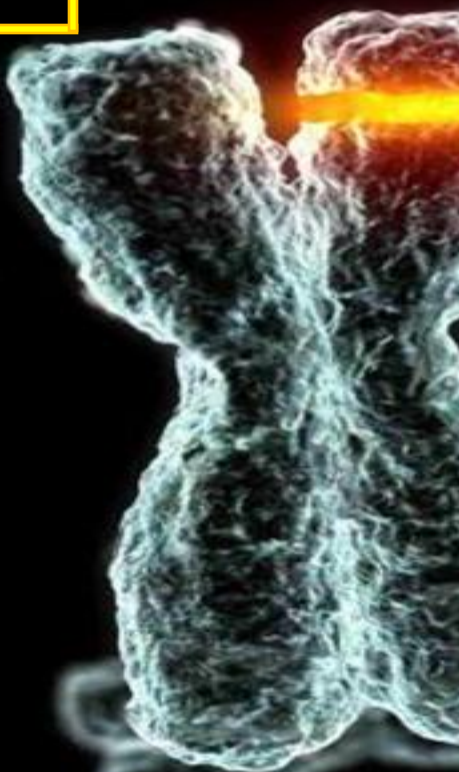
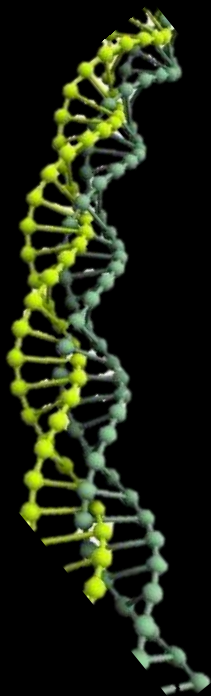
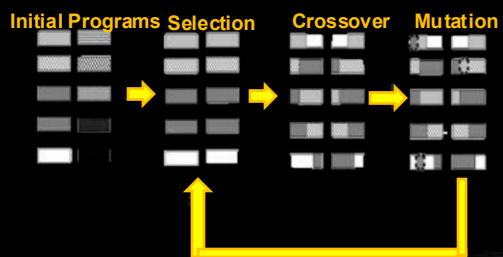
- Many problems can be solved by genetic programming!
- Most popular for predictive data analytics

Reference:

Gandomi, Amir H., et al. *Handbook of genetic programming applications*. Springer, 2015.

Sora AI-generated video. OpenAI, 2025





01

Generate

Candidate programs

02

Evaluate

Accuracy and complexity

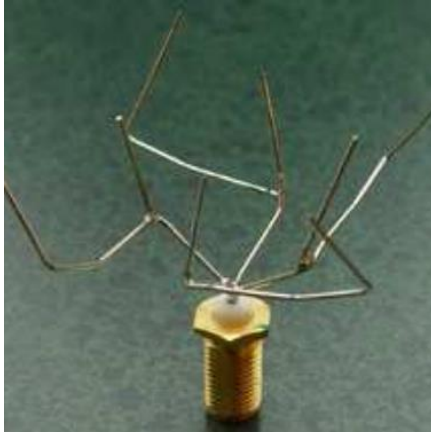
03

Evolve

Selection, crossover, mutation

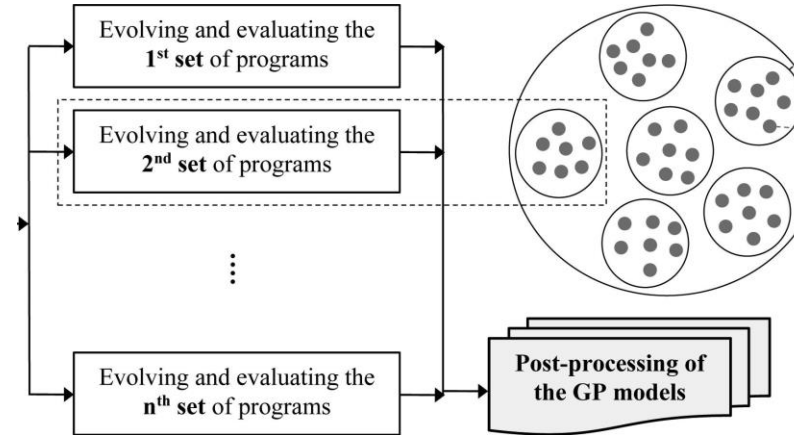


Genetic Programming – Distinguished Features



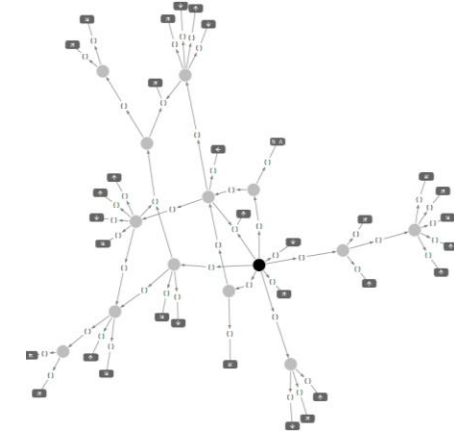
Finding the Model Structure

- The Structure is found via Evolution
- Pre-defined structure is not required
- Feature selection
- It can model the behaviour without any prior assumptions



Parallel Processing

- It is Population based Approach
- It can easily take advantage of parallel processing, cloud computing, and distributed systems



Simplicity and Explainability

Intelligent Machines

**Evolutionary algorithm
outperforms deep-learning
machines at video games**

MIT
Technology
Review

Neural networks have garnered all the headlines, but a much more powerful approach is waiting in the wings.

by Emerging Technology from the arXiv July 18, 2018

With all the excitement over neural networks and deep-learning

Complex simulations can yield compact equations

SC-CBF · COMPACT NOTATION

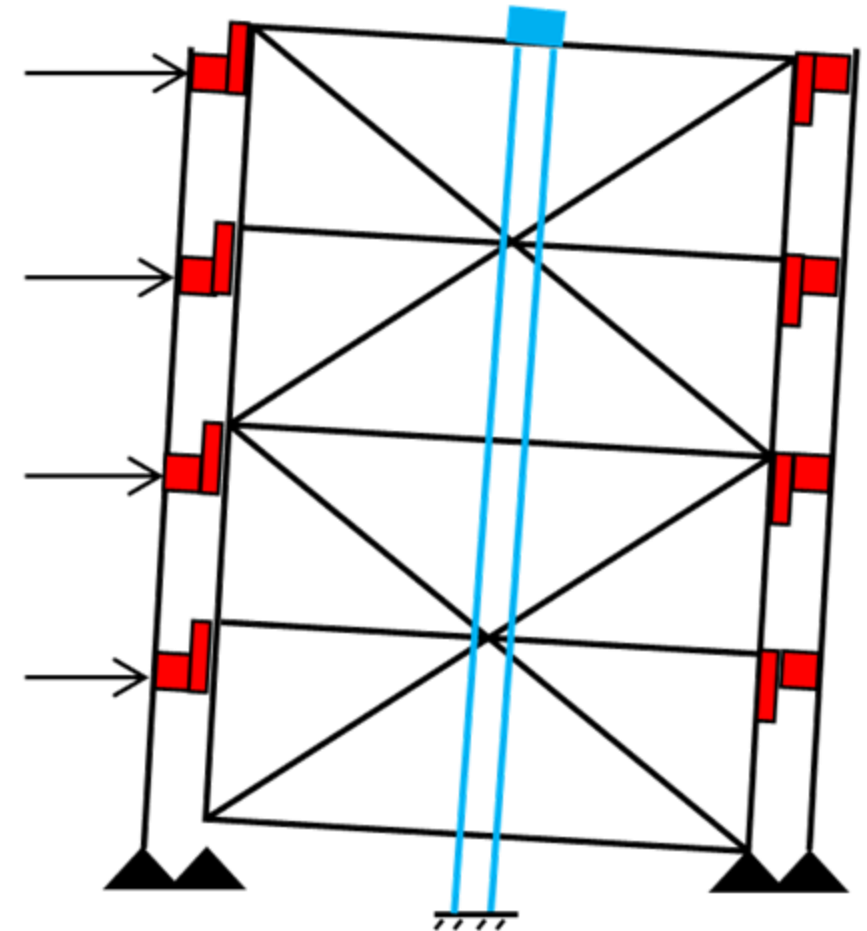
$$\ln \theta = 25.9 \text{ PGV} + 0.615 \ln|q| - 1.08$$

$$q = \tanh(2S_a(T)) [S_a(2T) + (h/b)^2] \sqrt{F_y}$$

$$R = 0.970$$

Use the study's units
and preprocessing.

Other GP applications: joint response · constitutive models · airfoil noise

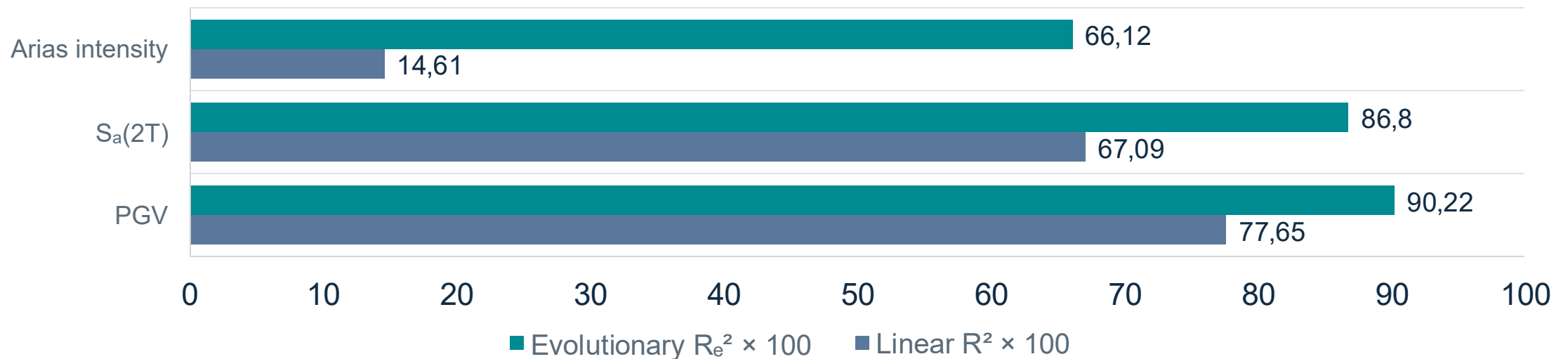


A useful correlation may be nonlinear

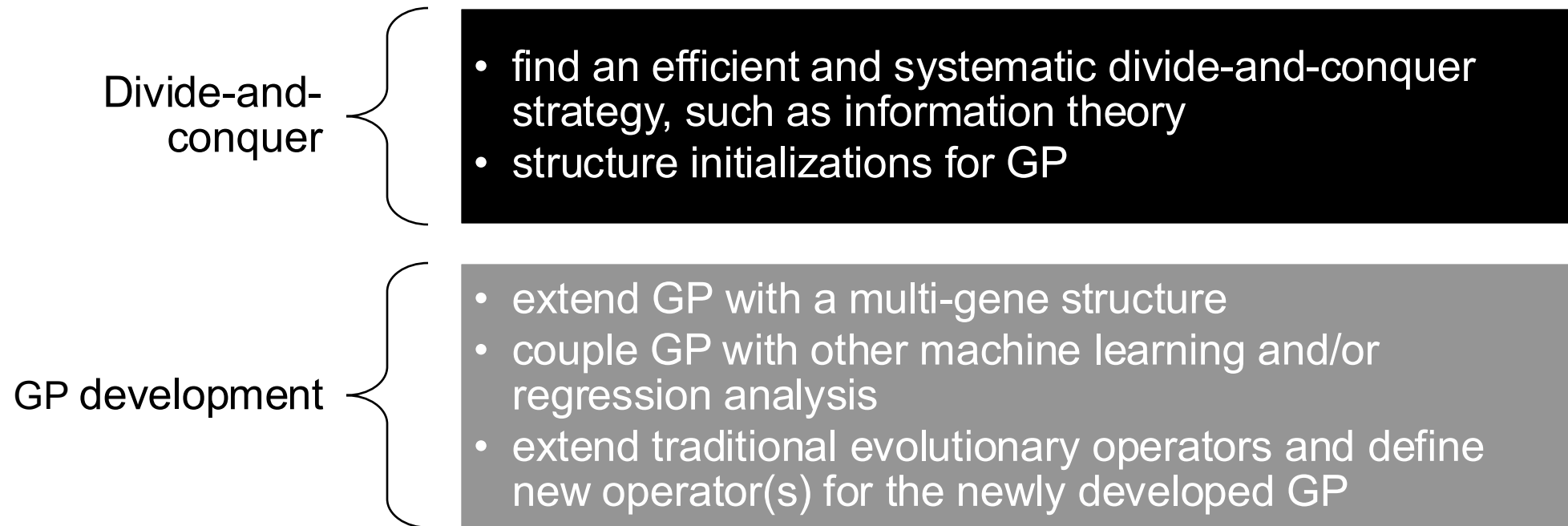
$$R_{e,j} = \text{corr}(f_j, \text{GP}(x_j), y)$$

$$R_e = \frac{\sum_{i=1}^n (y_i - \bar{y}_i) \left(f_{j,GP}(x_{ij}) - \overline{f_{j,GP}(x_{ij})} \right)}{\sqrt{\sum_{i=1}^n (y_i - \bar{y}_i)^2 \sum_{i=1}^n \left(f_{j,GP}(x_{ij}) - \overline{f_{j,GP}(x_{ij})} \right)^2}}$$

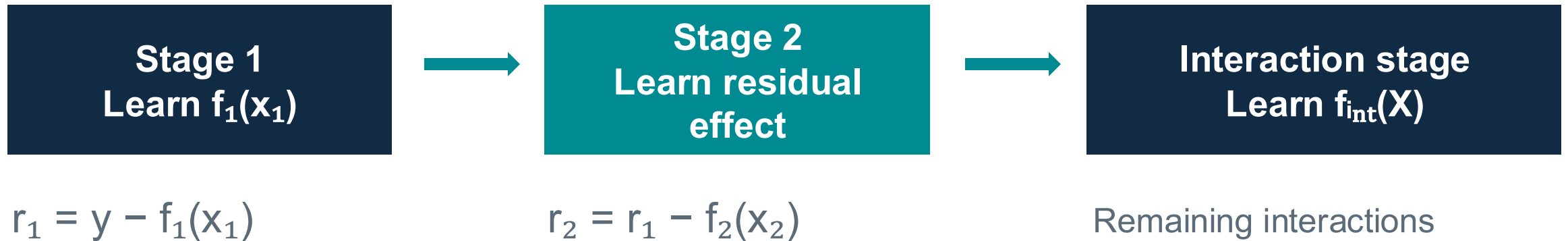
Evolve a transformation before measuring association.



Genetic Programming for Big Data Analytics (Current)



Multi-stage GP learns what remains unexplained

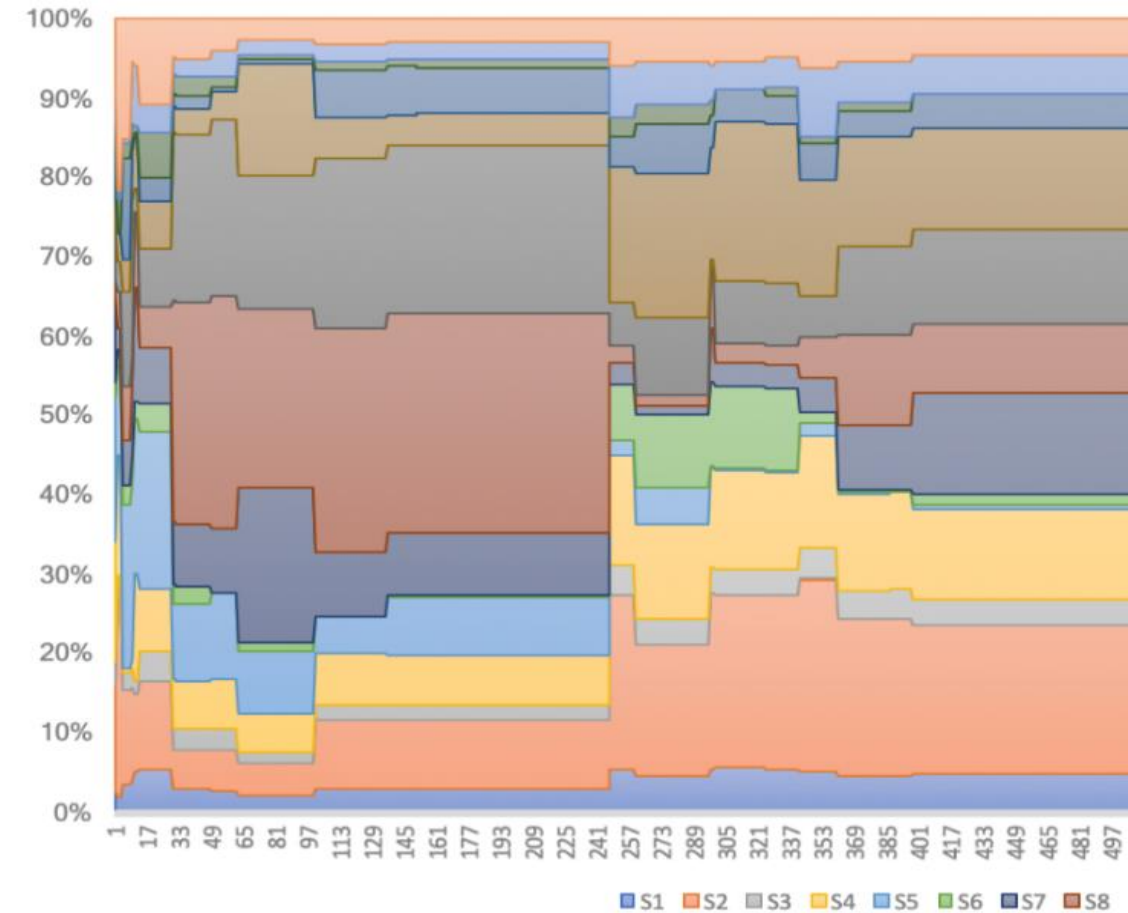


$$\hat{y}(X) = \sum_i f_i(x_i) + f_{\text{int}}(X)$$

Advancing Genetic Programming via Information Theory

$$\begin{pmatrix} G_{11} & G_{21} & \dots & G_{n1} \\ G_{1m} & \dots & G_{im} & \dots & G_{nm} \end{pmatrix} \times \begin{pmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \beta_{n2} & \dots & \beta_{nm} \end{pmatrix} + \begin{pmatrix} \beta_{01} \\ \vdots \\ \beta_{0m} \end{pmatrix} = \begin{pmatrix} y_1^{pr} \\ \vdots \\ y_m^{pr} \end{pmatrix}$$

Population size	Method	WIR	WIW
1,000	MR	0.1012	0.0992
	GP	0.1084	0.1048
	RRGP	0.1012	0.1094
	MRGP-TC	0.1172	0.0958
	MRGP-SumT	0.1115	0.0959
	MRGP-MDL	0.1120	0.0960
	MRGP-SMDL	0.1347	0.0961
100	MRGP-TC	0.1085	0.0961
	MRGP-SumT	0.1076	0.0963
	MRGP-MDL	0.1030	0.0964
	MRGP-SMDL	0.1080	0.0968
	Proposed GP	0.0617	0.0664



Domain-Informed Genetic Superposition Programming

Case Study on SFRC Beams

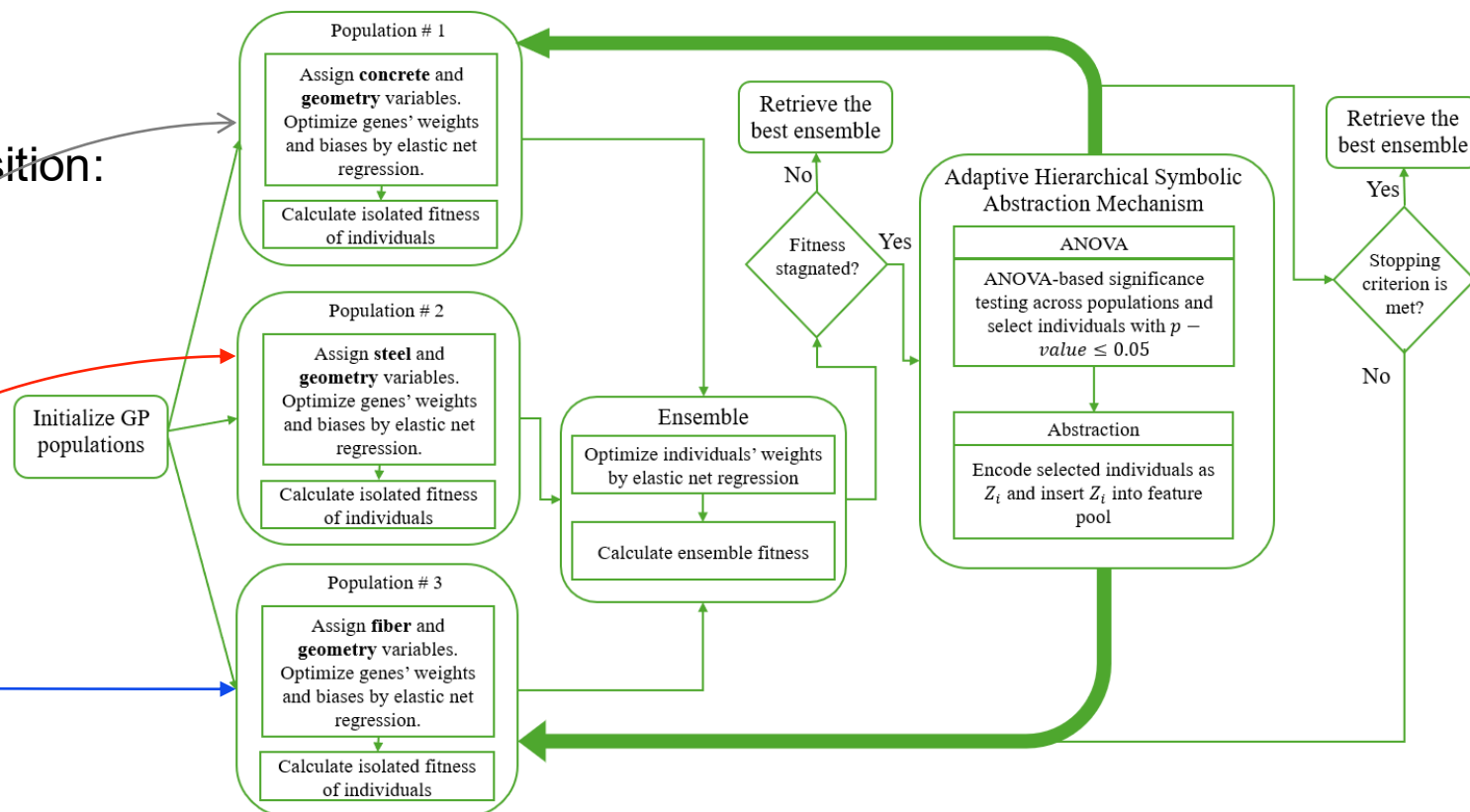
Ultimate Strength (f_u) based on superposition:

$$f_u = f_c + f_s + f_f$$

f_c = concrete strength

f_s = steel strength

f_f = fibre strength



Boruta Enforced Individual (BEI): Shadow Gene Guidance on Creating Elite Individual

STEP 01

Creation of Shadow Genes

Noise-augmented versions of the original genes are generated. For each gene, Gaussian noise, scaled to the gene's standard deviation, is added, followed by a shuffling of values.

STEP 02

Gene Importance Assessment

- A shallow ANN is trained using both original and shadow
- genes,
- The model's performance, quantified by the F1-score as a baseline for each gene's importance.
- Each gene undergoes perturbation
- The model is retrained.
- New F1-score is computed.
- The performance decrement denotes importance.

STEP 03

Importance Assessment & Hit Counting

- A more substantial performance drop for an
- original gene, relative to the maximum observed for shadow
- genes, signifies its importance.
- For each gene, a 'hit' is registered if its performance decrement exceeds the highest decrement observed among the shadow genes.

STEP 04

Selection of Important Genes

The algorithm selects the top-ranking genes, guided by the "Maximum allowable gene" parameter, which denotes the maximum number of genes that is allowable within each individual.

STEP 05

Producing Boruta Enforced Individual (BEI)

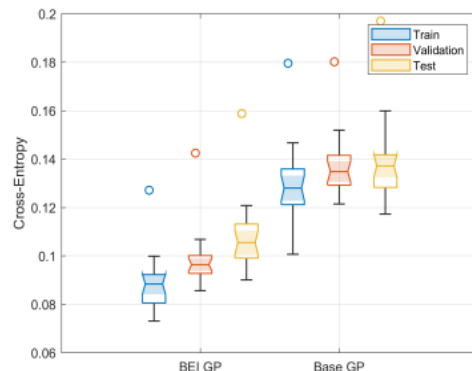
The most significant genes, as determined by their hit history and importance, are combined to form a new, singular individual for the next generation.

Keep genes with repeated evidence of importance

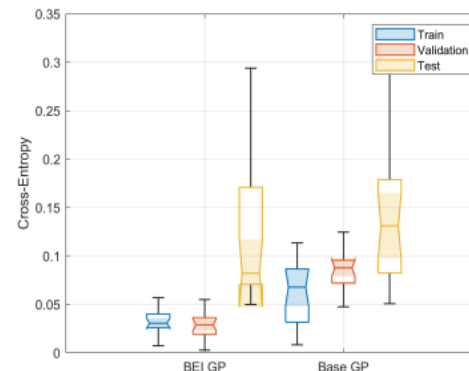
Comparison result of the proposed BEI Method with the base model

Wilcoxon rank-sum test, 20 independent evaluation

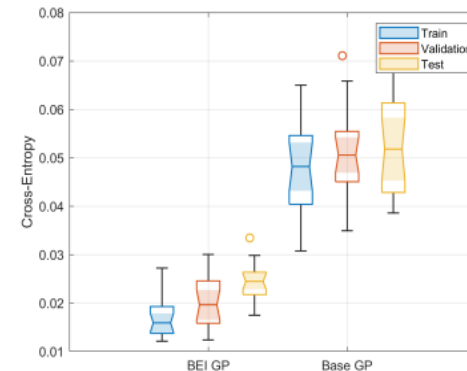
		ARWPM	GECR	GSAD	HAPT
BEI	Train	0.0883 ± 0.0117	0.0327 ± 0.0118	0.0173 ± 0.0044	0.0239 ± 0.0029
	Val	0.0982 ± 0.0115	0.0272 ± 0.0139	0.0201 ± 0.0051	0.0255 ± 0.0027
	Test	0.1087 ± 0.0146	0.1218 ± 0.0716	0.0243 ± 0.0036	0.0268 ± 0.0034
Base	Train	0.1300 ± 0.0158	0.0627 ± 0.0320	0.0475 ± 0.0099	0.0325 ± 0.0040
	Val	0.1371 ± 0.0127	0.0859 ± 0.0196	0.0509 ± 0.0092	0.0343 ± 0.0036
	Test	0.1392 ± 0.0173	0.1412 ± 0.0693	0.0520 ± 0.0100	0.0344 ± 0.0042
P-Value	Train	2.56×10^{-7}	0.004320	6.8×10^{-8}	3.94×10^{-7}
	Val	6.92×10^{-7}	1.06×10^{-7}	6.8×10^{-8}	3.42×10^{-7}
	Test	1.38×10^{-6}	0.239323	6.8×10^{-8}	5.87×10^{-6}



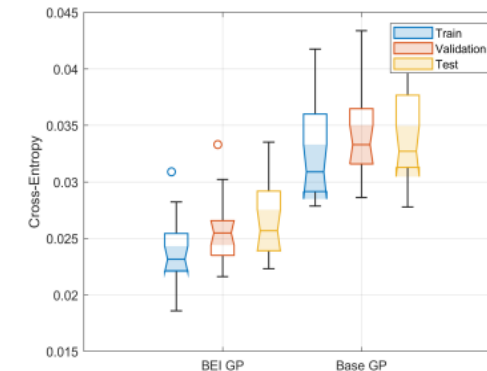
(a) ARWPM



(b) GECR

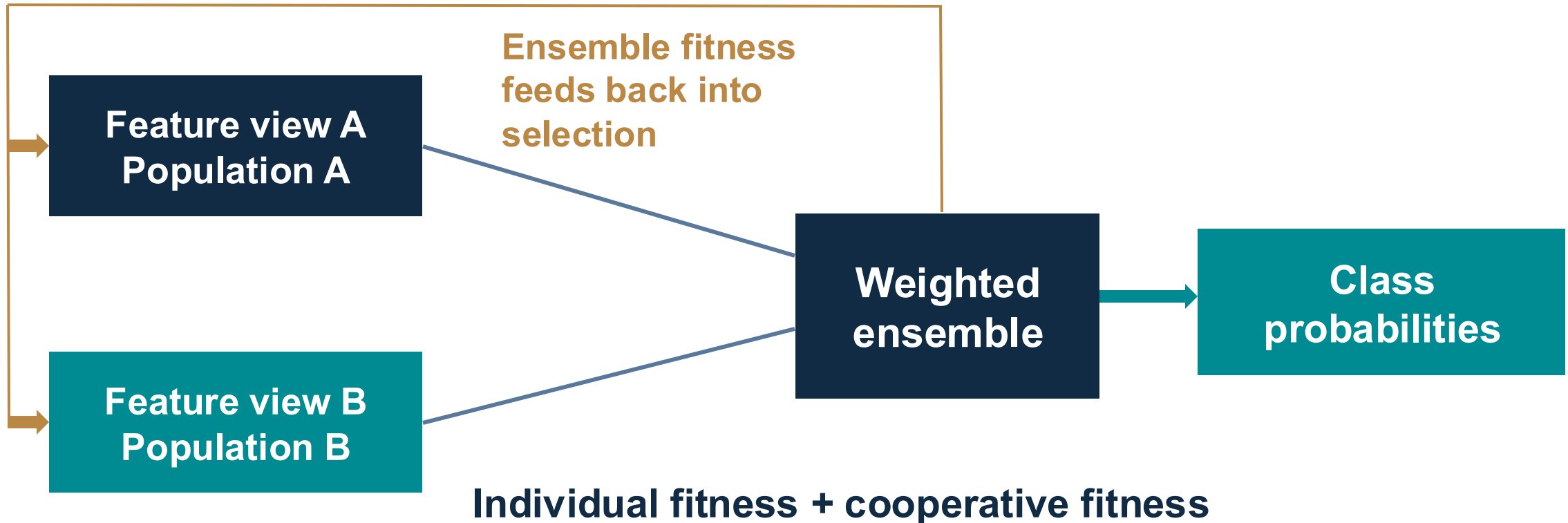


(c) GSAD



(d) HAPT

Co-evolution rewards a good teammate



8 BENCHMARK DATASETS

GenForge makes
symbolic modelling
reusable

GenForge

Open-source Python framework

Software Impact paper:

<https://doi.org/10.1016/j.simpa.2026.100812>

Code:

<https://github.com/agandomi/genforge>

9/27/2026



Accuracy alone does not earn reliance

Uncertainty is what we do not know about a prediction or outcome.

More Important for AI?!

HUMAN DRIVER

**“That driver
made a mistake.”**

AUTOMATED DRIVING

**“Can we trust
the technology?”**

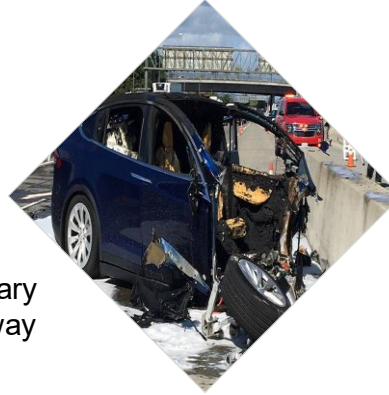
The same error can receive a different interpretation.

Driving analogy; algorithm-aversion evidence comes from forecasting experiments.

Repercussions of Uncertain Artificial Intelligence

In 2016, a severe accident occurred when a driver-assistance system misinterpreted the white side of a trailer as an open sky.

National Highway Traffic Safety Administration, "Tesla crash preliminary evaluation report," U.S. Department of Transportation, National Highway Traffic Safety Administration, Technical Report PE 16-007, Jan. 2017.



Accuracy is not enough

As AI systems are increasingly making decisions that impact people's lives, it's crucial that they can communicate the level of certainty in their decisions.

In applications like medical diagnoses incorrect or uncertain predictions can have dire consequences.

In 2014, Amazon tool to evaluate resumes developed a bias against female candidates, as it was trained predominantly on men's resumes so the AI downgraded resumes that included the word "women's".

Dovidio, J. F., Brescoll, V. L., & Graham, M. J. (2022, Sep 24). *Scientists not immune from gender bias, Yale study shows*. Yale News. Retrieved Feb 7, 2025, from <https://news.yale.edu/2012/09/24/scientists-not-immune-gender-bias-yale-study-shows>



Interactivity and Collaboration

When AI systems can express uncertainty, it opens the door for more interactive and collaborative decision-making processes with humans.

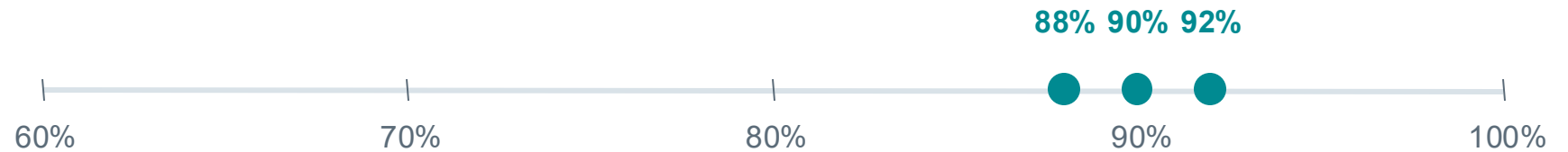
By recognizing what it doesn't know, human experts can incorporate that information into their decisions, leading to better outcomes.

Uncertainty has more than one source

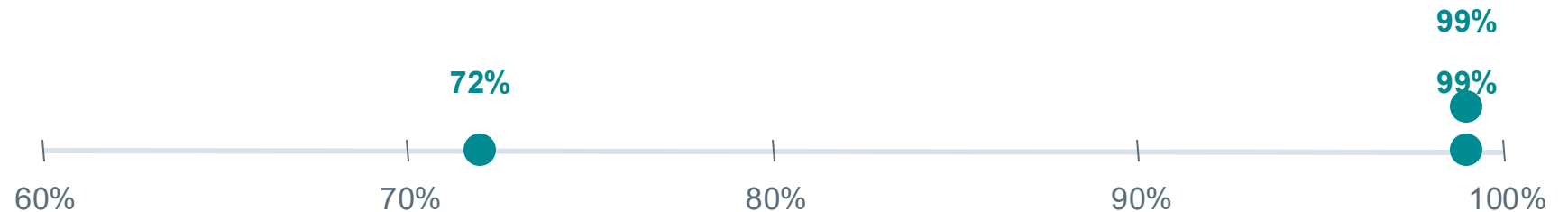
The same confidence can hide different uncertainty

MEAN CONFIDENCE: 90% IN BOTH CASES

**Stable
estimates**

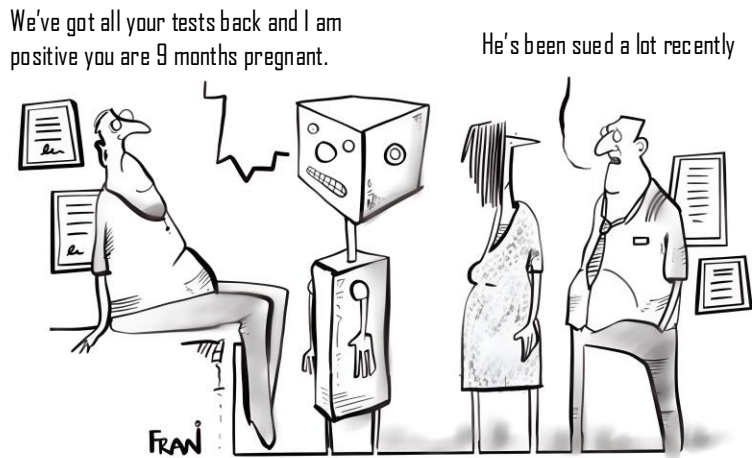


**Disagreeing
estimates**



Same mean and entropy of the mean; different disagreement. Hypothetical values.

Why Isn't AI a Standard Clinical Guideline?



01

Lack of trust in the decisions made among both healthcare professionals and patients*.

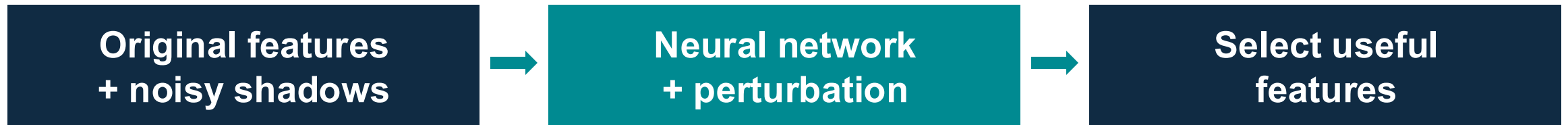
02

If physicians and patients encounter confidently incorrect outcomes, trust in AI and ML systems can be severely undermined.

Fehr, et al., *Frontiers in Digital Health*, 6, 267 290, (2024)

Select features for reliable probabilities

NEUROBORUTA



F1

Classification

ECE

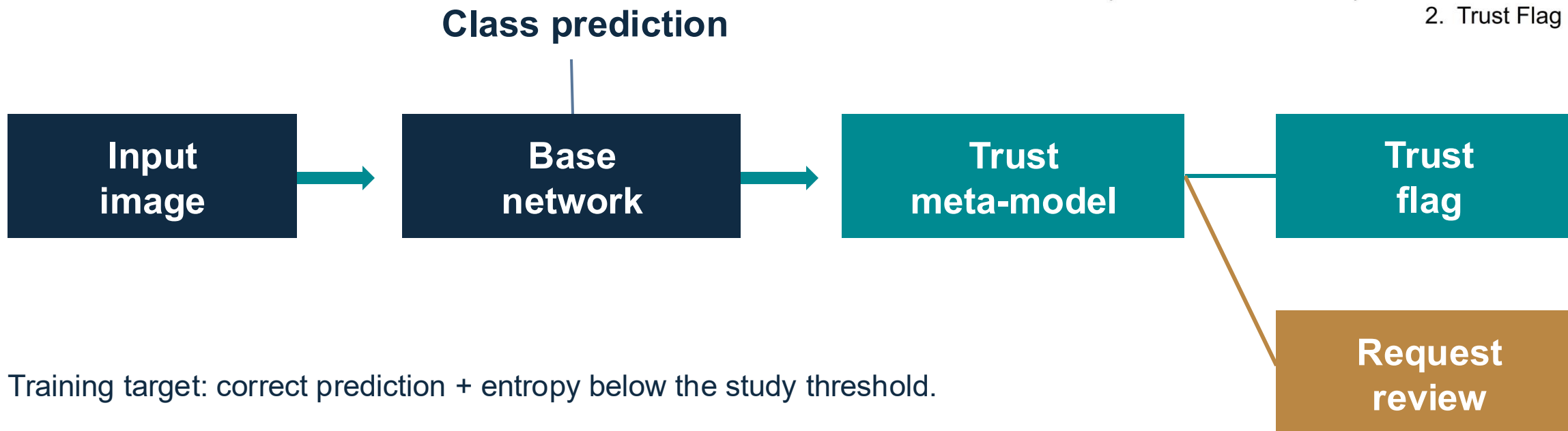
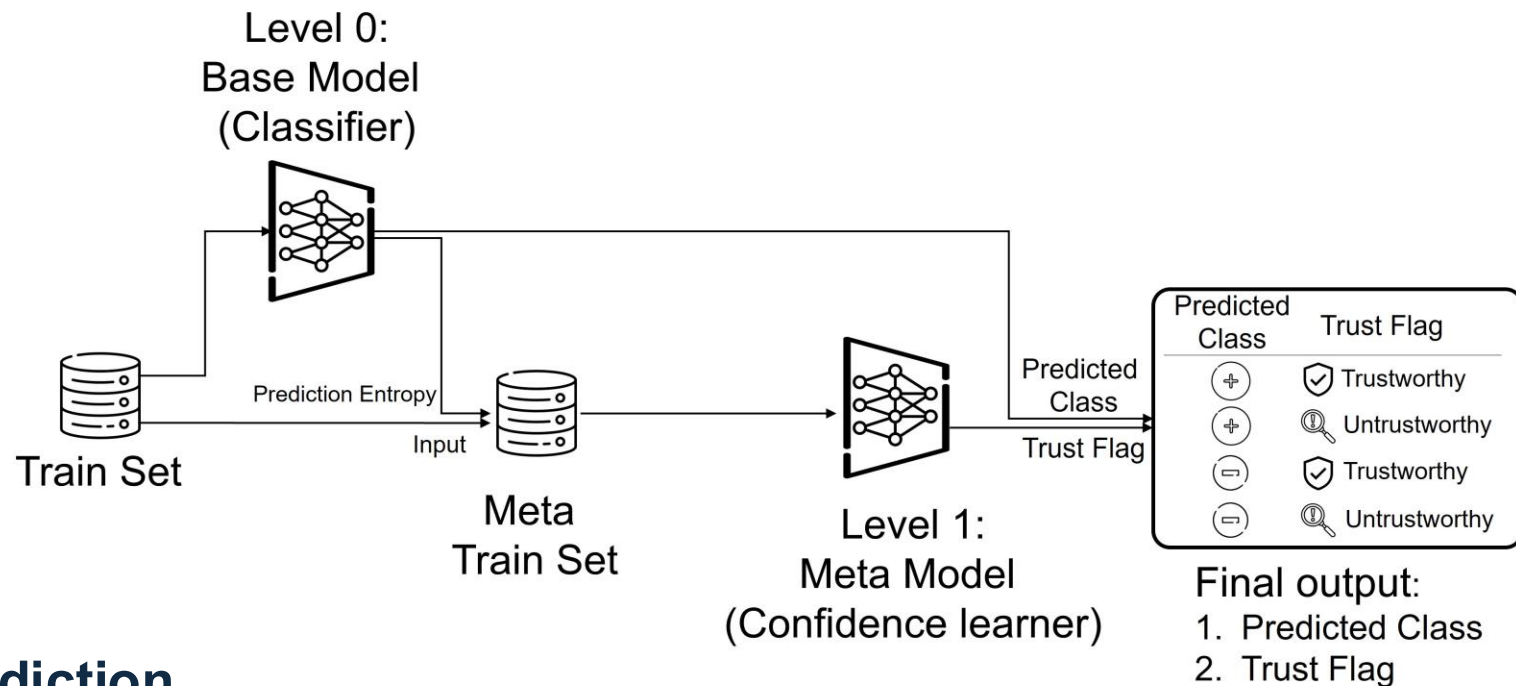
Calibration

Brier score

Probability quality

Feature relevance and the quality of predictive probabilities belong together.

A second network learns when review is needed



Training target: correct prediction + entropy below the study threshold.

Reliability has a review cost

COVIDx CXR-4 · ViT + MONTE CARLO DROPOUT · $\tau = 0.1$

96.91%

Trust precision

Among trusted flags,
correct-and-confident target

0.46%

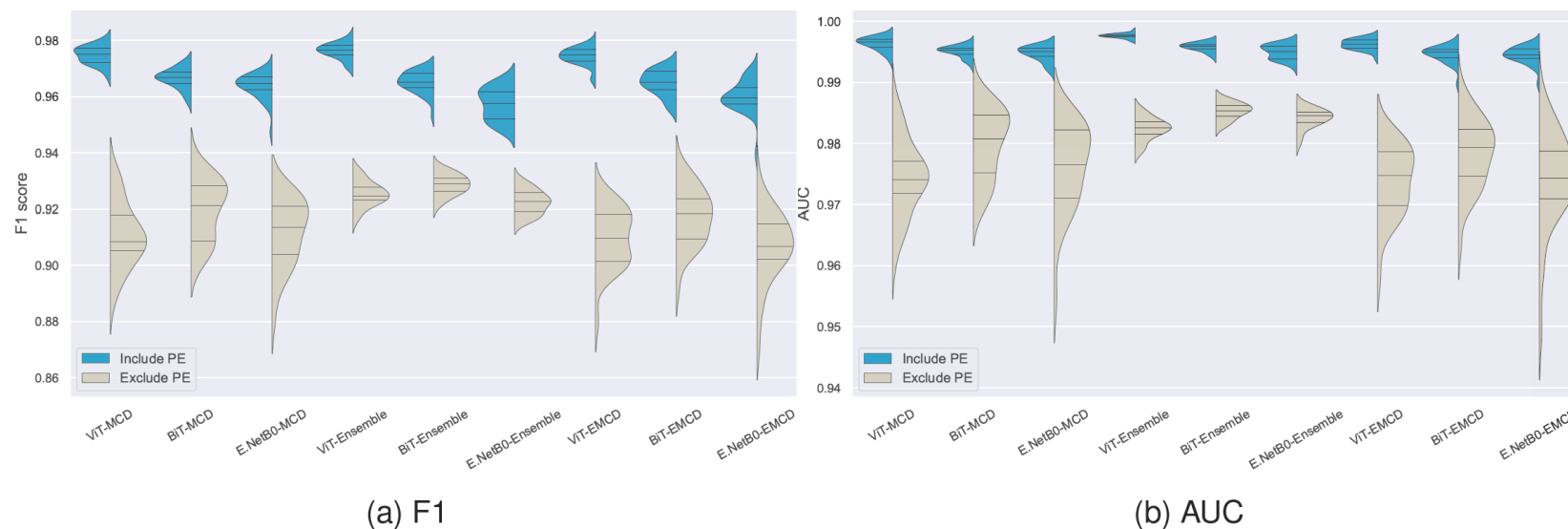
False-trust rate

Of all cases, wrong
predictions accepted as trusted

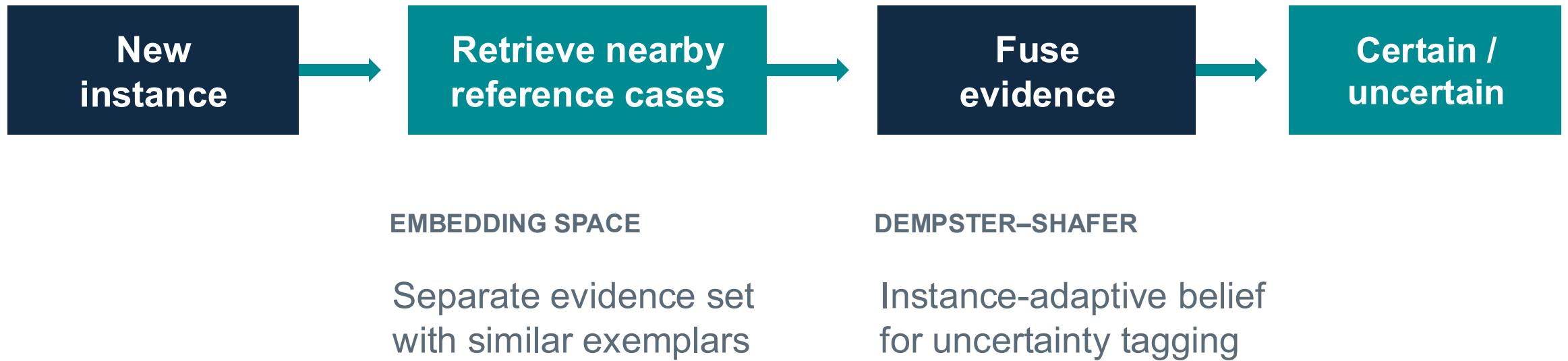
51.67%

Review-alert rate

Of all cases,
flagged for review



Retrieve evidence before accepting a prediction



The classifier's label stays the same; the evidence changes the uncertainty tag.

The decision policy changes the value of accuracy

SAME MODEL	Accepted	Errors accepted	For review	Automatic error
Policy A	100	5	0	5.0%
Policy B	70	1	30	1.4%

95% model accuracy in both policies. Review has a cost and can also make errors.

Uncertainty should guide the agent's next action



Uncertainty triggers an information-seeking action.

Evaluate agents by the decisions they improve

DECISION QUALITY

Task success and justified action

RELIABILITY

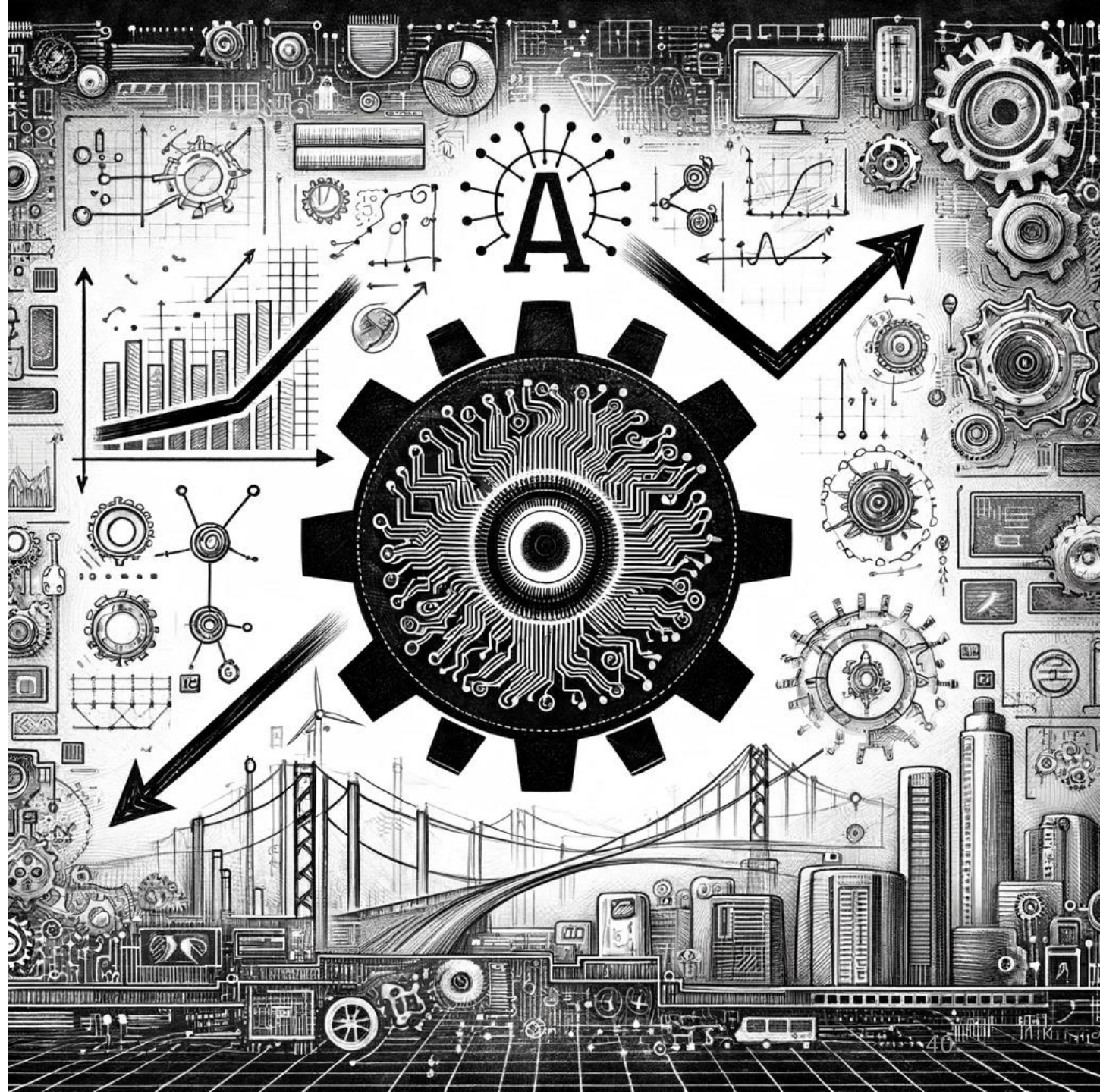
Unjustified confidence and missed warnings

EFFORT

Tool use, delay and human-review burden

Illustrative pilot: sensor gap → new measurement → updated model → reviewed action

AI/EI for Optimization

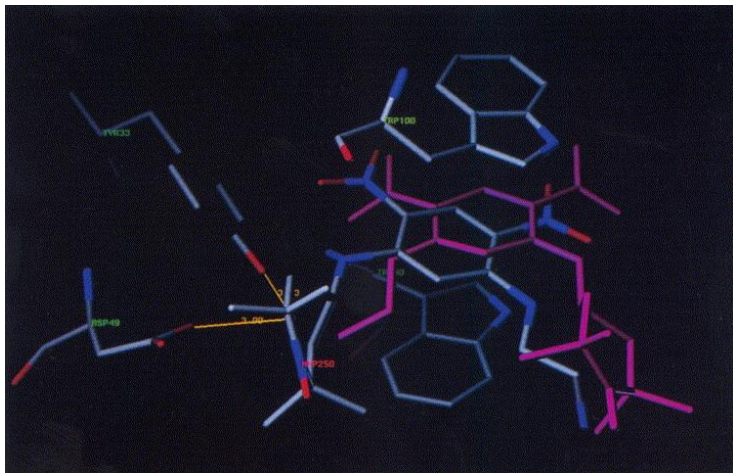


EC in Real-World Problems



Boeing Turbine geometry of 777 GE engine Design:

Charles W. Petit, "Touched by nature: putting evolution to work on the assembly line." US News & World Report, volume 125, issue 4, pages 43–45 (1998).



Merck Pharmaceutical discovered first clinically-approved antiviral drug for HIV:

Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. J Mol Biol 267: 727–748.

•& so many other companies.
amazon



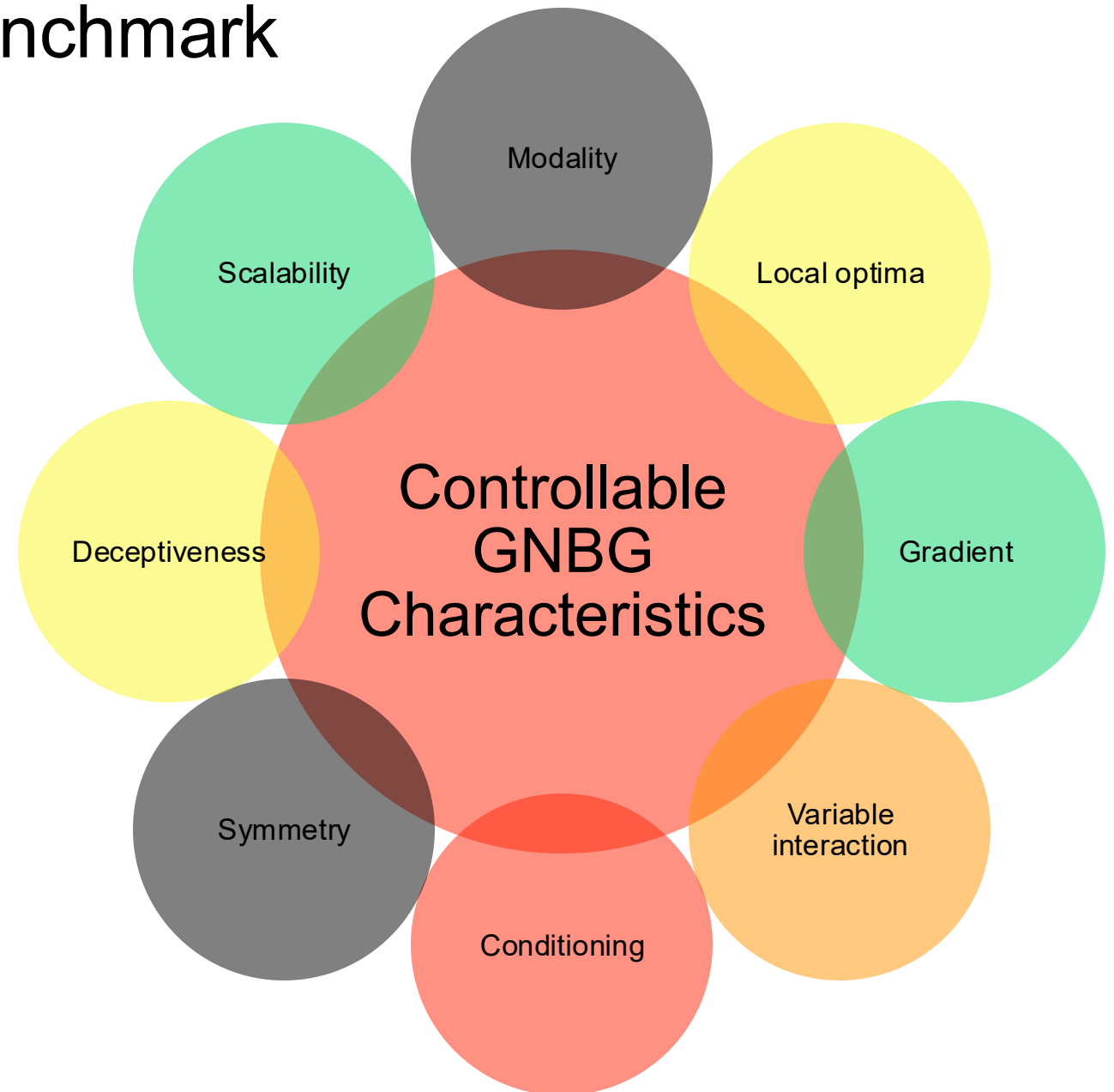
Uber



Generalized Numerical Benchmark Generator (GNBG)

GNBG (Gandomi et al.,) uses a singular parametric baseline function, and by adjusting its parameter settings, it can produce a wide range of problem instances with controllable characteristics and varying degrees of difficulty.

Competitions at: ACM GECCO, IEEE WCCI, and IEEE CEC



Baseline Function of GNBG

$$\text{Minimize } : f(\vec{x}) = \min_{k \in \{1, \dots, o\}} \left\{ \sigma_k + \left(\mathbb{T}_k \left((\mathbf{R}_k(\vec{x} - \vec{m}_k))^T \right) \mathbf{H}_k \mathbb{T}_k \left(\mathbf{R}_k(\vec{x} - \vec{m}_k) \right) \right)^{\lambda_k} \right\},$$

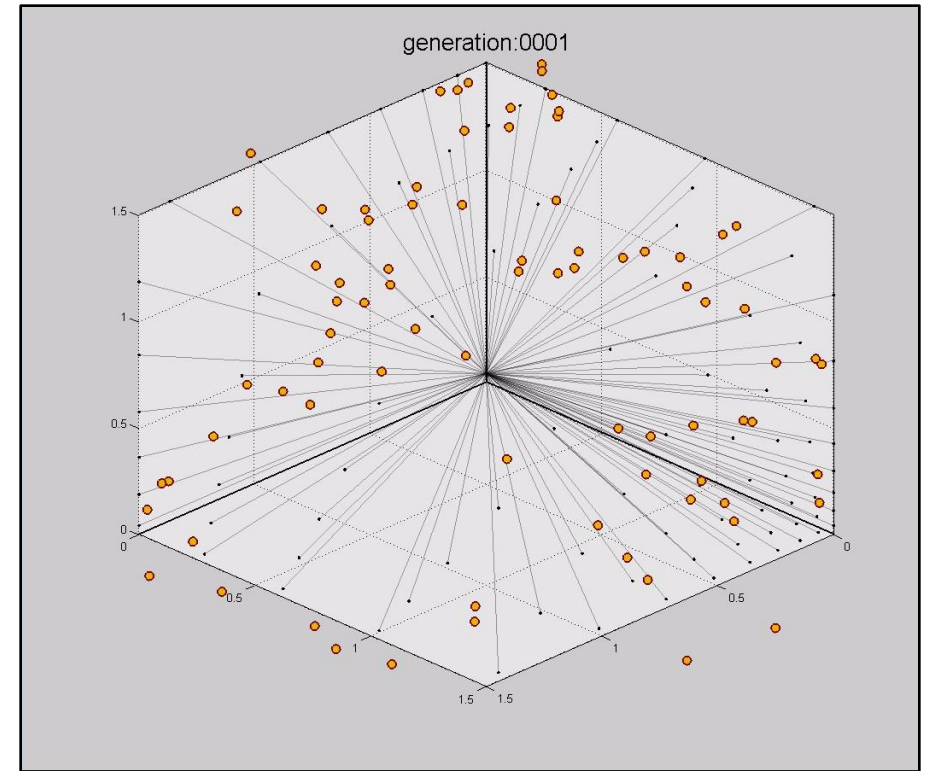
$$\text{Subject to : } \vec{x} \in \mathbb{X} : \mathbb{X} = \{\vec{x} \mid l_i \leq x_i \leq u_i\}, i \in \{1, 2, \dots, d\},$$

where $\mathbb{T}_k(\vec{a}) \mapsto \vec{b}$ is a non-linear transformation function that introduces additional complexities to the basin of each component.

$$a_j \mapsto \begin{cases} \exp \left(\log(a_j) + \mu_{k,1} \left(\sin(\omega_{k,1} \log(a_j)) + \sin(\omega_{k,2} \log(a_j)) \right) \right) & \text{if } a_j > 0 \\ 0 & \text{if } a_j = 0 \\ -\exp \left(\log(|a_j|) + \mu_{k,2} \left(\sin(\omega_{k,3} \log(|a_j|)) + \sin(\omega_{k,4} \log(|a_j|)) \right) \right) & \text{if } a_j < 0 \end{cases}$$

EI For Special Problems

- Variable length optimization [1]
- Robust optimization [2]
- Multi and many objective problems [3]
- Dynamic optimisation [4]
- Very High Dimensional Optimisation [5]



1. Cheng et al., 2018. Solving incremental optimization problems via cooperative coevolution. *IEEE Transactions on Evolutionary Computation*, 23(5), pp.762-775.
2. Yazdani, et al., Robust optimization over time by estimating robustness of promising regions. *IEEE Transactions on Evolutionary Computation*, 27(3), pp.657-670 (2022).
3. Behmanesh, et al. Evolutionary many-objective algorithms for combinatorial optimization problems: a comparative study. *Archives of Computational Methods in Engineering*, 28(2), pp.673-688, (2021).
4. Peng, et al. Algorithm xxx: EDOLAB, a Platform for Research and Education in Evolutionary Dynamic Optimization.” *ACM Transactions on Mathematical Software*, 2026
5. Azad, et al., An enhanced guided stochastic search with repair deceleration mechanism for very high-dimensional optimization problems of steel double-layer grids. *Structural and Multidisciplinary Optimization*, 67(12), p.203 (2024).

EI for Combating COVID-19

XPRIZE-Cognizant Pandemic Challenge

F1: daily new cases

F2: stringency of planned interventions

UTS team:

I led team **Kangaroos** in this competition

Our team used

- ML to build models
- EMO to optimize the objectives

Results:

- We were the only team from Australia to reach the final
- We end up as a top-ten team
- We became one of the Honourable Mention Winners

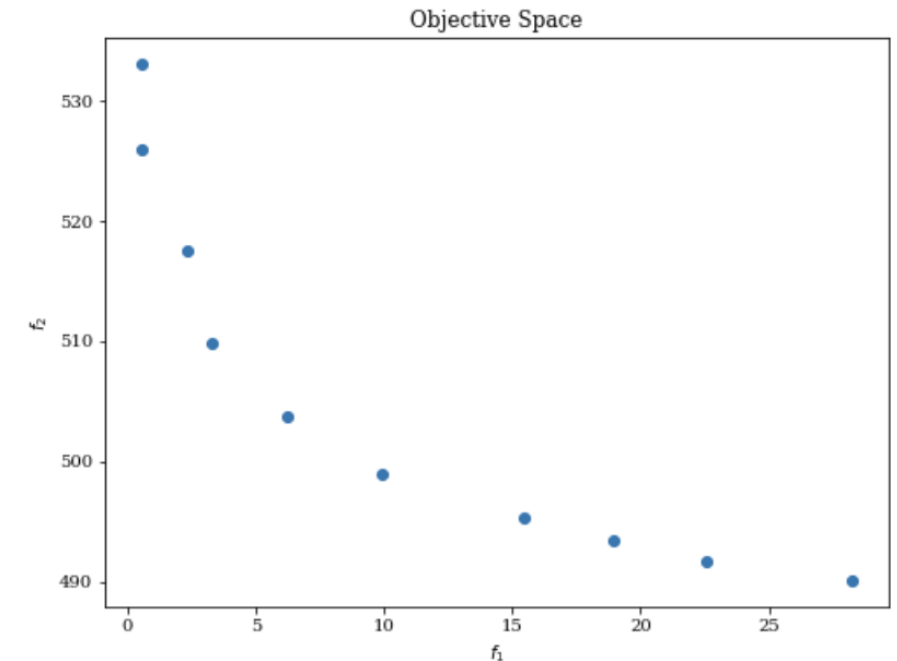


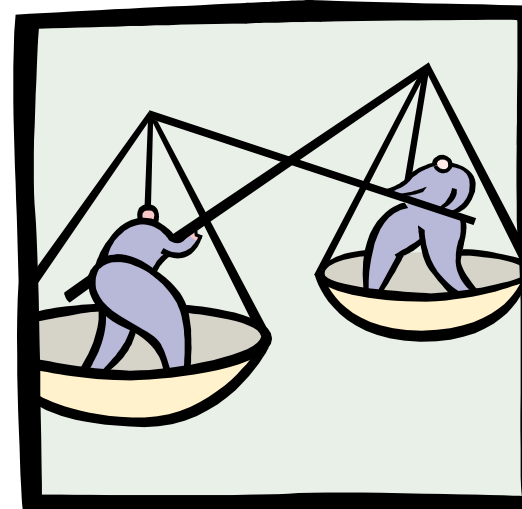
Figure 1. Sample Pareto Front for Canada

Customization in Optimization

Evolutionary Computation

- PROs
 - Derivative-free
 - Global
 - Flexible
- Heuristic
 - More efficient
 - Convergence
 - Speed

- CONs
 - Slow



Domain Knowledge

Gather as many information as possible!

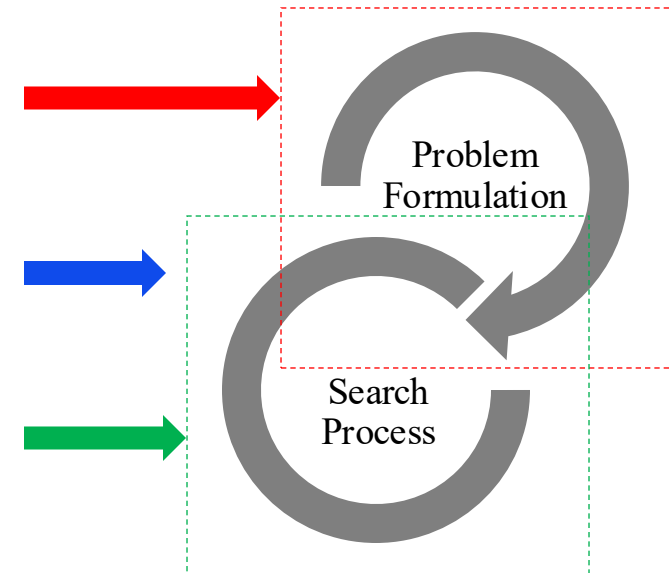
Possible Knowledge

Expert Knowledge

Information and Mathematical Theories

Engineering Principles

Scientific Concepts

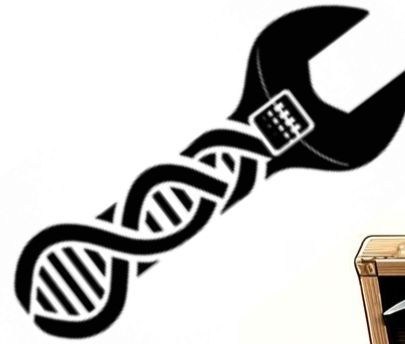


This is the most efficient way applying Evolutionary Computation to Real-World Problems!

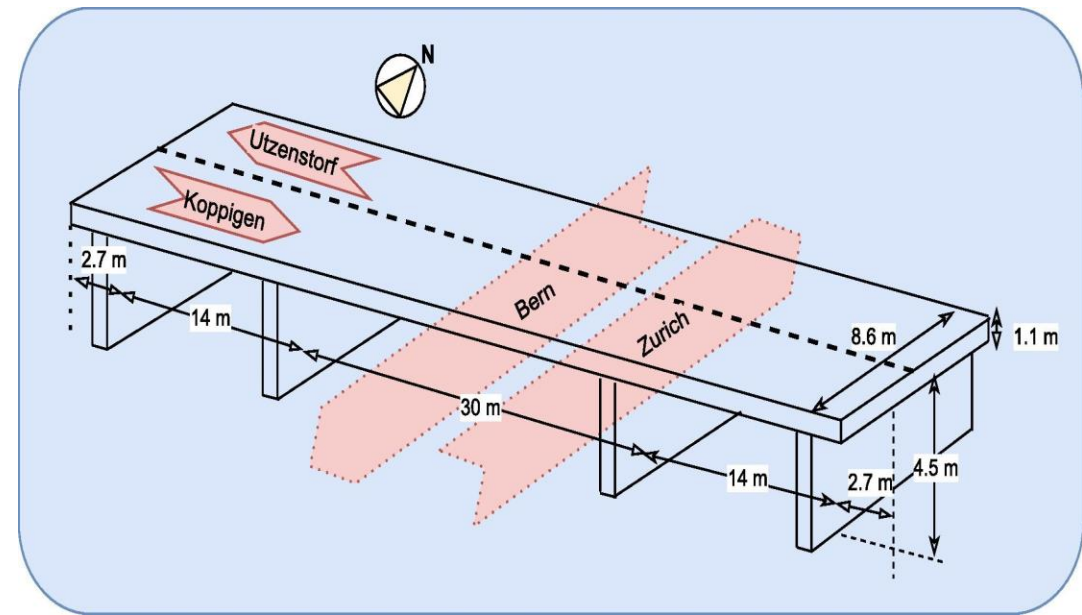
EI is only one of the AI tools!

AI methodologies/Tools:

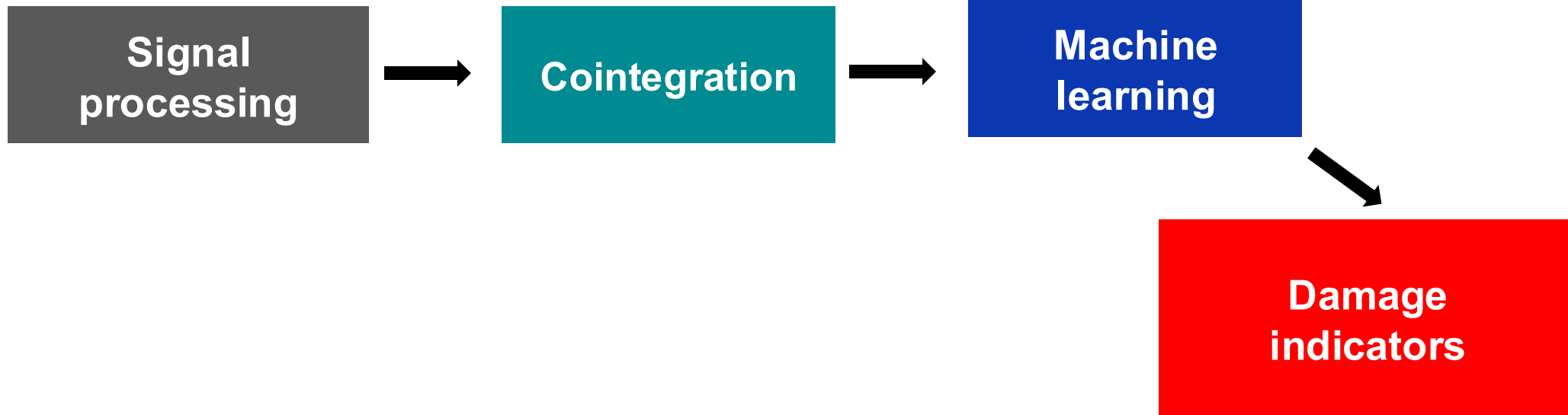
- Machine Learning
- Deep Learning
- Natural Language Processing
- Computer Vision
- Predictive Analytics
- Evolutionary Intelligence
- Fuzzy Logic
- Expert Systems
- Robotics
- Signal Processing



Hybrid AI connects sensing, models and judgment



STRUCTURAL MONITORING





MINI SYMPOSIUM

Recent Advances in Intelligent Engineering

2026



Thank you!

Amir H Gandomi

***Professor of Data Science at University of Technology Sydney
Distinguished Professor at Óbuda University, Budapest
e-mail: gandomi@uts.edu.au***