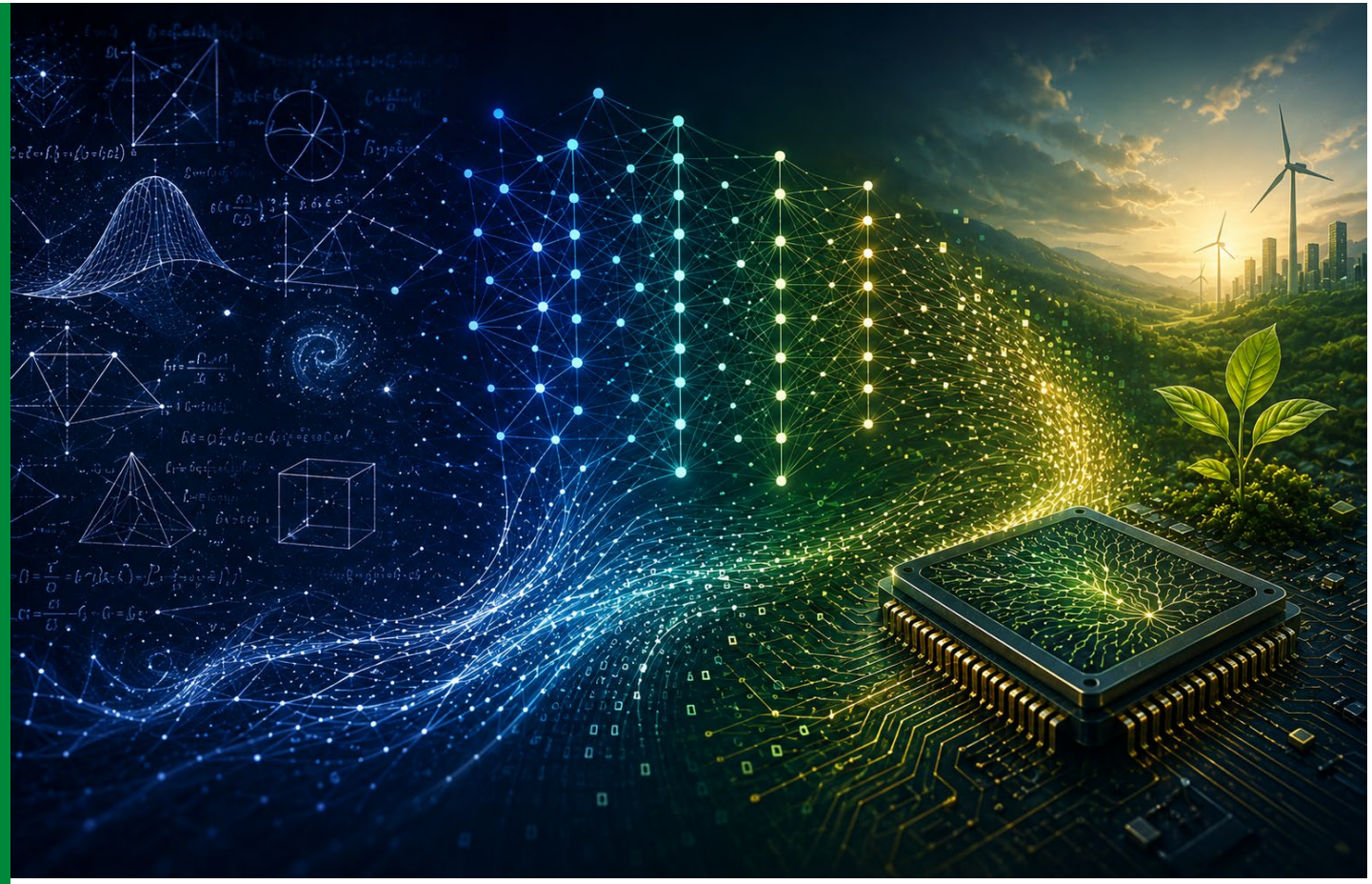
LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

FACULTY FOR MATHEMATICS, INFORMATICS, AND STATISTICS

DEPARTMENT OF MATHEMATICS

CHAIR "MATHEMATICAL FOUNDATIONS OF ARTIFICIAL INTELLIGENCE"

Towards Reliable and Sustainable AI: Foundations and the Future of Informatics



Gitta Kutyniok

LMU Munich

(also DLR – German Aerospace Center & University of Tromsø, Norway)

European Informatics Leaders' Summit (ECSS 2026)

Informatics Europe · Porto, October 27, 2026

baiosphere the bavarian ai network

DFG Deutsche Forschungsgemeinschaft

Bundesministerium für Bildung und Forschung

relAI Konrad Zuse School of Excellence in Reliable AI

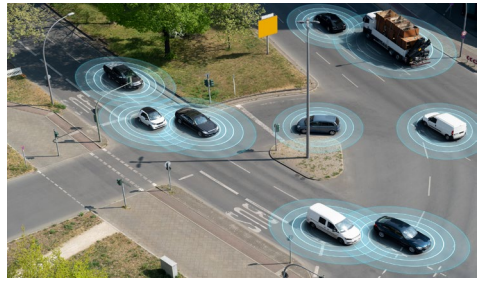
mcml Munich Center for Machine Learning

Bayerisches Staatsministerium für Wissenschaft und Kunst

HIGHTECH Agenda Bayern

AI is Becoming Critical Infrastructure for Science and Society

Societal Infrastructures:



Autonomous mobility systems



Mobile communication networks (5G/6G)

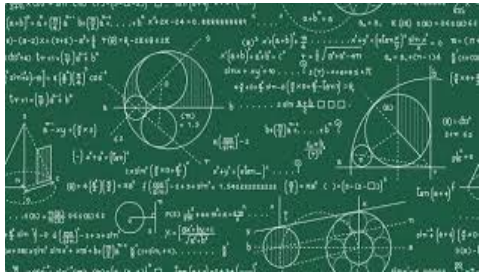


Healthcare systems

Transformation of Universities:



From Automation to
Autonomous Discovery

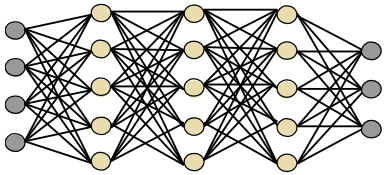


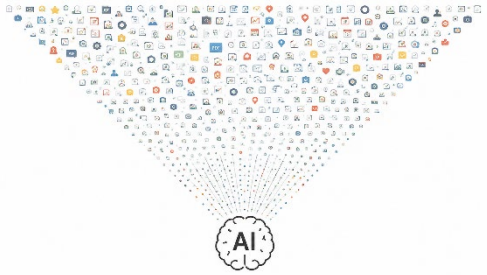
From Proof Assistance to
Machine Reasoning



From Curriculum Delivery to
Personalized Learning Support

AI Is Changing — So Is the Reliability and Sustainability Challenge







Classical AI

- Prediction
- Classification
- Decision support

Generative AI

- Content generation
- Multimodality
- Foundation models

Embodied AI

- Physical actions
- Interaction with the environment



Agentic AI

- Planning
- Autonomous actions
- Multi-agent systems

From systems that provide answers to systems that take actions.

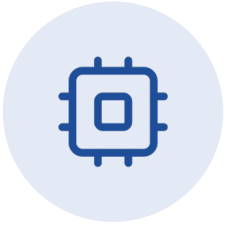
Three Questions for the Future of AI



Can we trust it?

Reliability

- Plausible is not always correct.
- What can we measure and guarantee?



Can we compute it?

Fundamental limits

- A guarantee is only operational if it can be computed.
- Where are the limits of digital computation?



Can we sustain it?

Sustainability

- Energy and compute at scale.
- Can mathematics unlock greater efficiency?

Mathematics makes all three questions precise – informatics is where they must be answered.

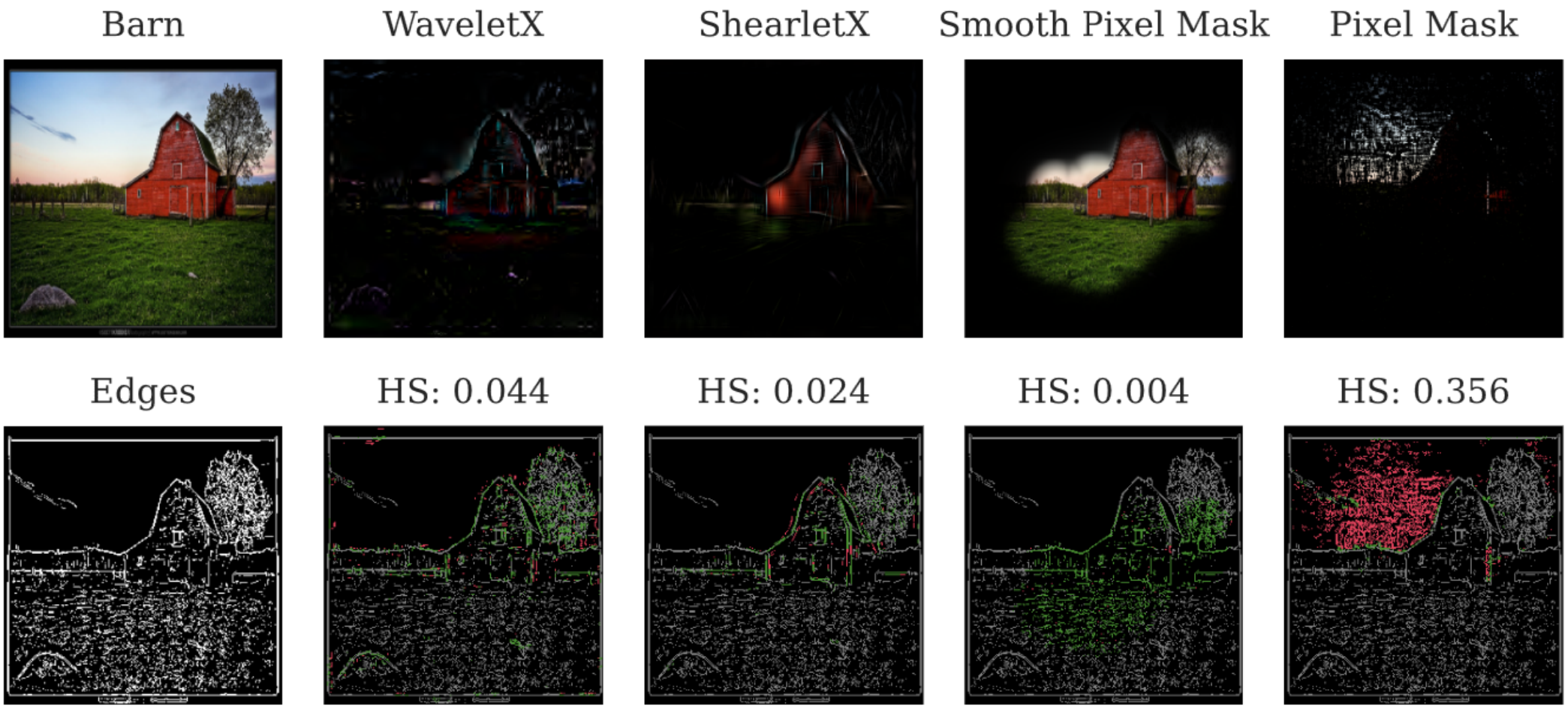
LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

MATH 4AI

Can we trust it?

From plausible to provable

Even Explanations Can Hallucinate

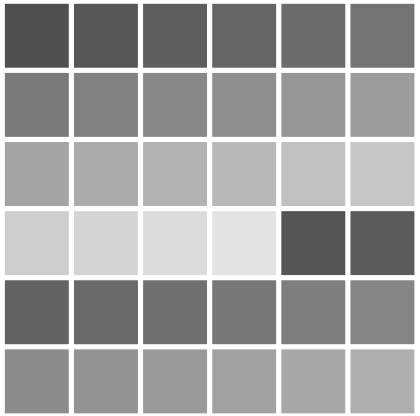


HS: hallucination score of Kolek et al. (CVPR 2023), fraction of newly created edges.

Representation Matters

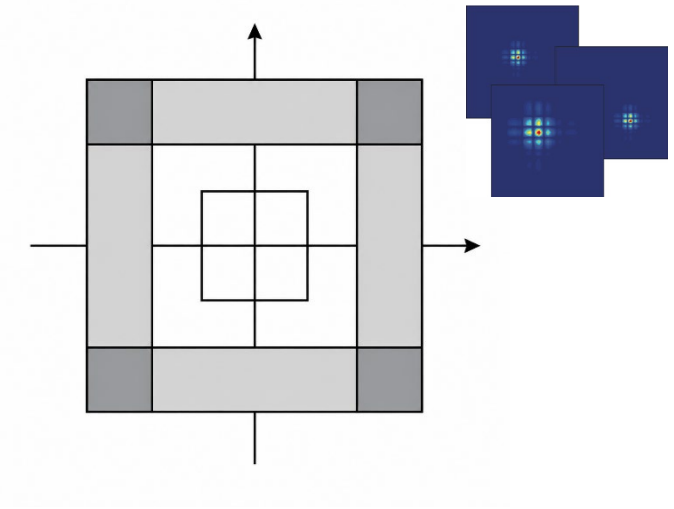
Pixels

local information



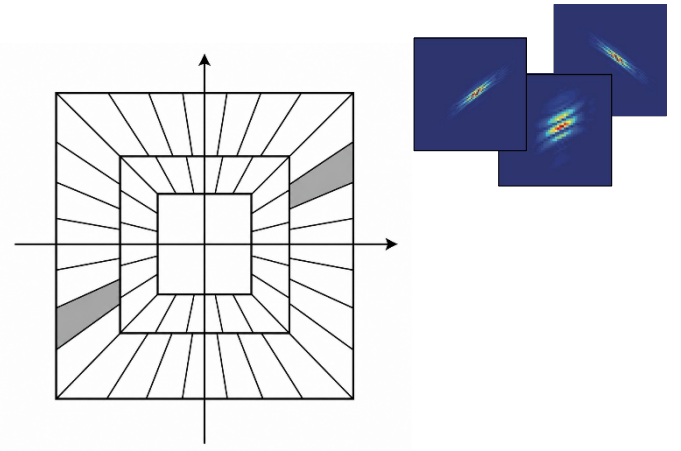
Wavelets

scale + location



Shearlets

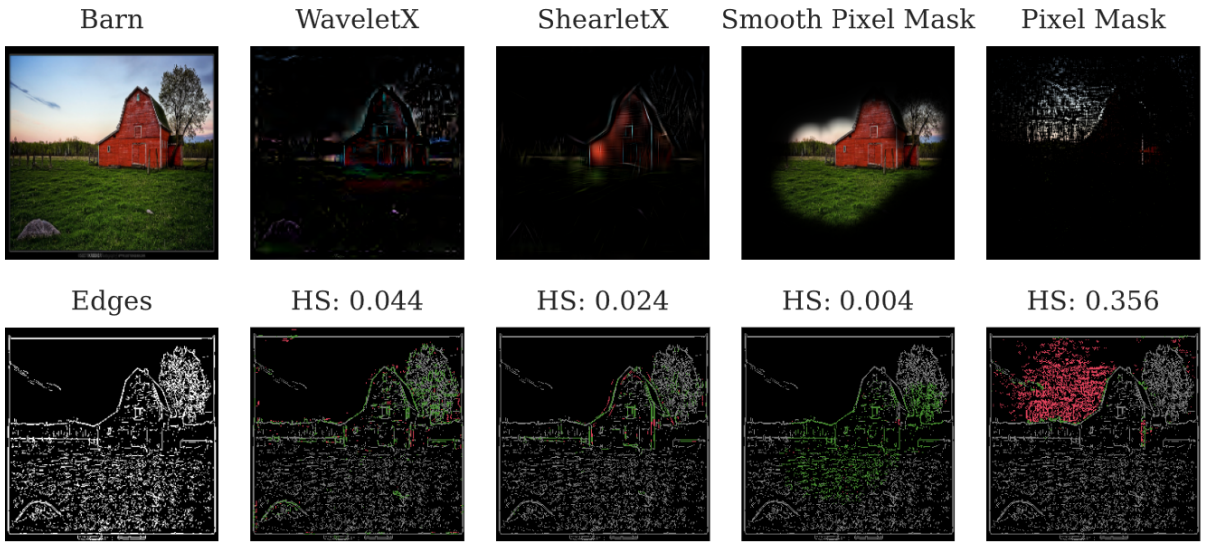
scale + direction + location



$$\mathcal{SH}_\psi f(a,s,t) = \langle f, \psi_{a,s,t} \rangle$$

What counts as detectable structure depends on the representation.

Representation Can Prevent Artificial Structures



Theorem (Kolek, Windesheim, Loarca, Kutyniok, Levie; 2023):

Let $x \in L^2[0,1]^2$ be an image. Let y be the explanation from ShearletX. Then, we have

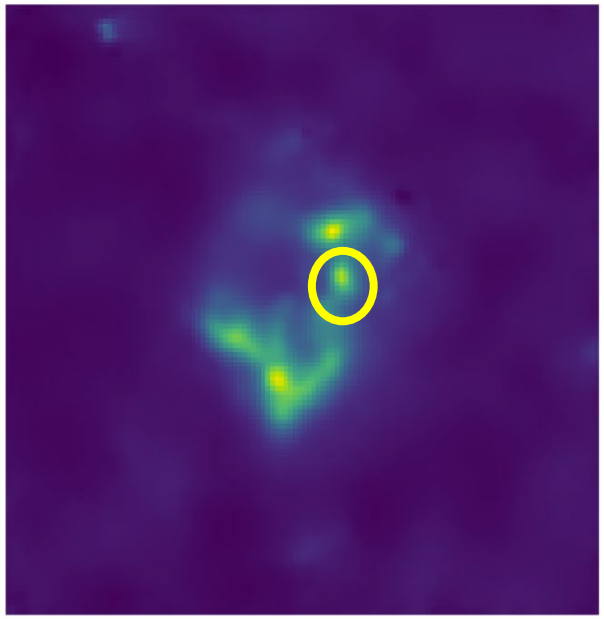
$$WF(y) \subset WF\ (x).$$

Masking in a shearlet space does not create new edges.

A suitable representation can turn a desired property of an explanation into a provable guarantee.

Hallucinations: Plausible Does Not Mean Correct

Reconstruction $\hat{x}$



Observation model

$$y = h * x + \eta$$

blurred by the PSF h, noise η

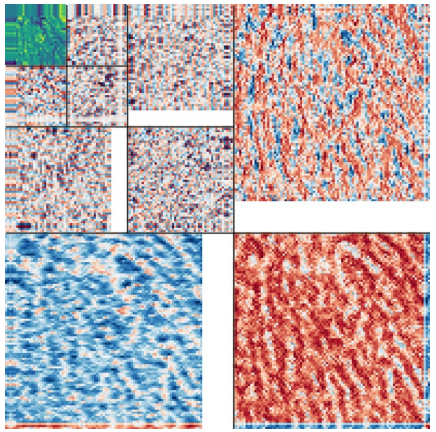
Example with U-Net-based deconvolution.

Plausible structure — but is it supported by the measurements?

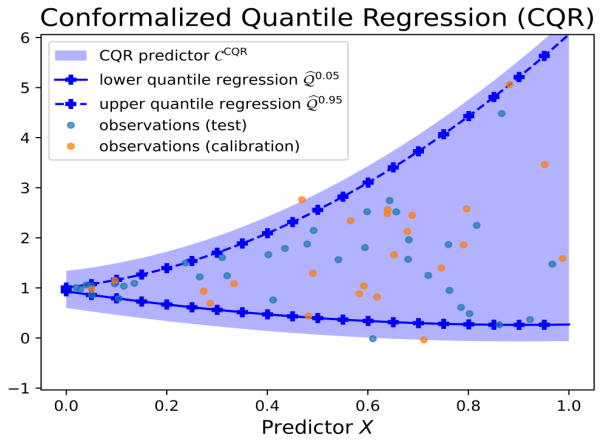
Imaging as a test case: an explicit observation model $y = h * x + \eta$ makes 'unsupported by the data' checkable.
Here the truth is known because the degradation is simulated on real CANDELS galaxies.

CHEM (Conformal Hallucination Estimation Metric): The Basic Idea

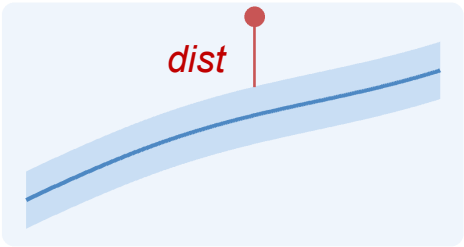
Structure-adapted representation → **Calibrated region** → **CHEM**



Wavelet coefficients: localized, directional, sparse



Conformal calibration



Distance to the calibrated region for each coefficient

Is the structural deviation larger than what calibration says is plausible?

CHEM: ongoing joint work with J. Li, I. Rosellon-Inclan, J.-L. Starck, 2026.

Where Does the Calibration Guarantee Come From?

Conformalized Quantile Regression:

Quantile Regression → Held-out Calibration → Calibrated Prediction Interval

Coverage Guarantee:

$$\alpha - \frac{1}{N+1} \leq \mathbb{P}\left(\left(\hat{Y}\right)_j \notin B_\alpha \left(\Phi\left(\left(\hat{X}\right)_j\right)\right)\right) \leq \alpha$$

- Finite-sample
- Distribution-free
- i.i.d. calibration and test samples

Theorem: Quantitative Sensitivity to Hallucinations

Quantitative detection bound (Rosellon-Inclan, Li, Starck, Kutyniok; 2026):

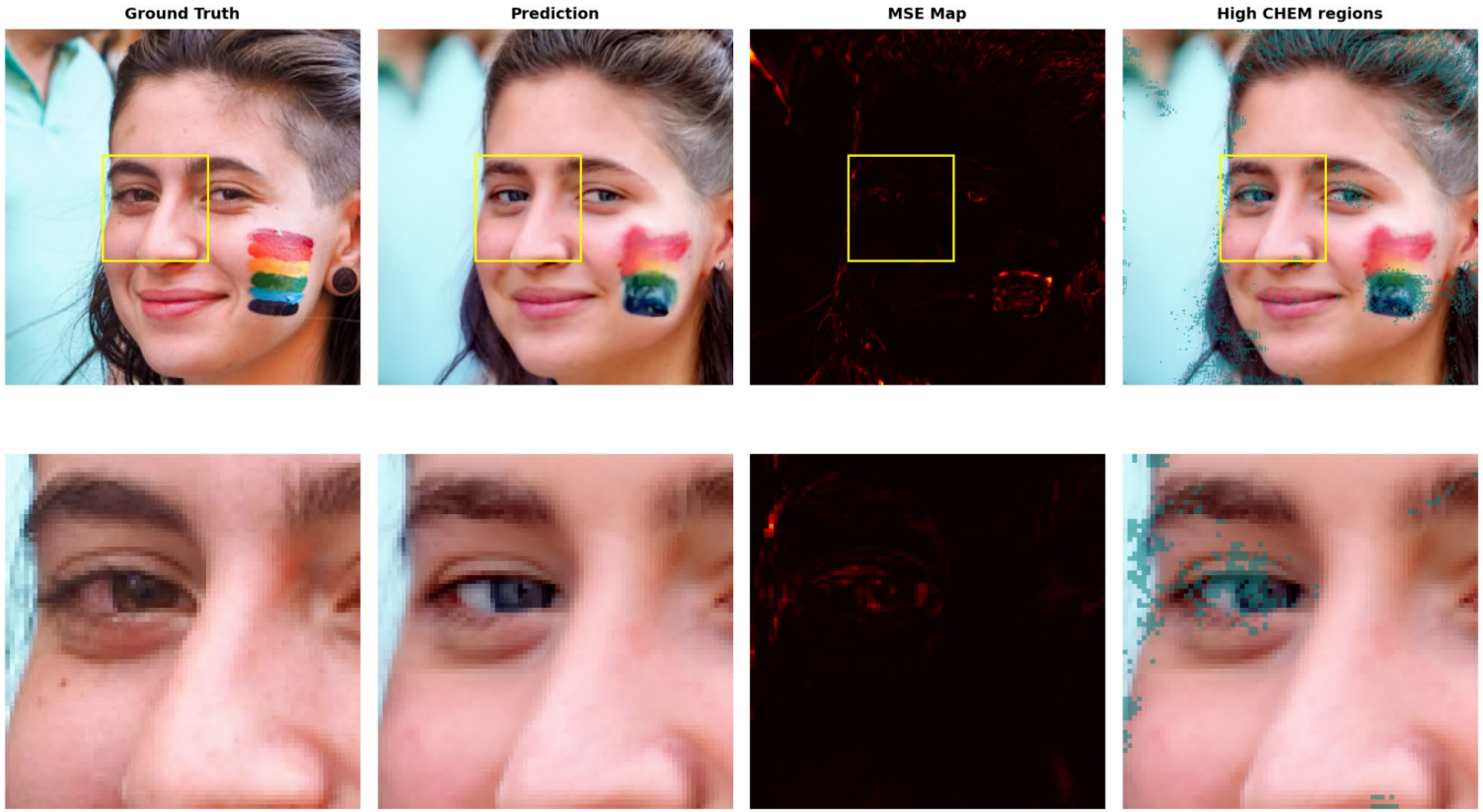
The probability of missing hallucinations can be bounded by:

$$\mathbb{P}(\mathrm{CHEM} = 0 \mid h) \leq 2 \exp\left(- c \, \frac{\left(s^{-\frac{1}{2}} \|h\|_2 - \hat{\mathcal{R}}(X)_{j^*} \right)_+^2}{\|\eta\|_{\psi_2}^2} \right)$$

s: sparsity of h in the representation; $\hat{\mathcal{R}}(X)_{j^*}$ *interval half-width at the dominant coefficient;* $\|\eta\|_{\psi_2}^2$: sub-Gaussian noise norm · c : absolute constant

- Sparser hallucinations in the chosen representation are *easier to detect*
- Larger hallucinations have an *exponentially smaller probability of being missed*
- The bound is *independent of the reconstruction architecture*

Pixel Error and Structural Error Tell Different Stories

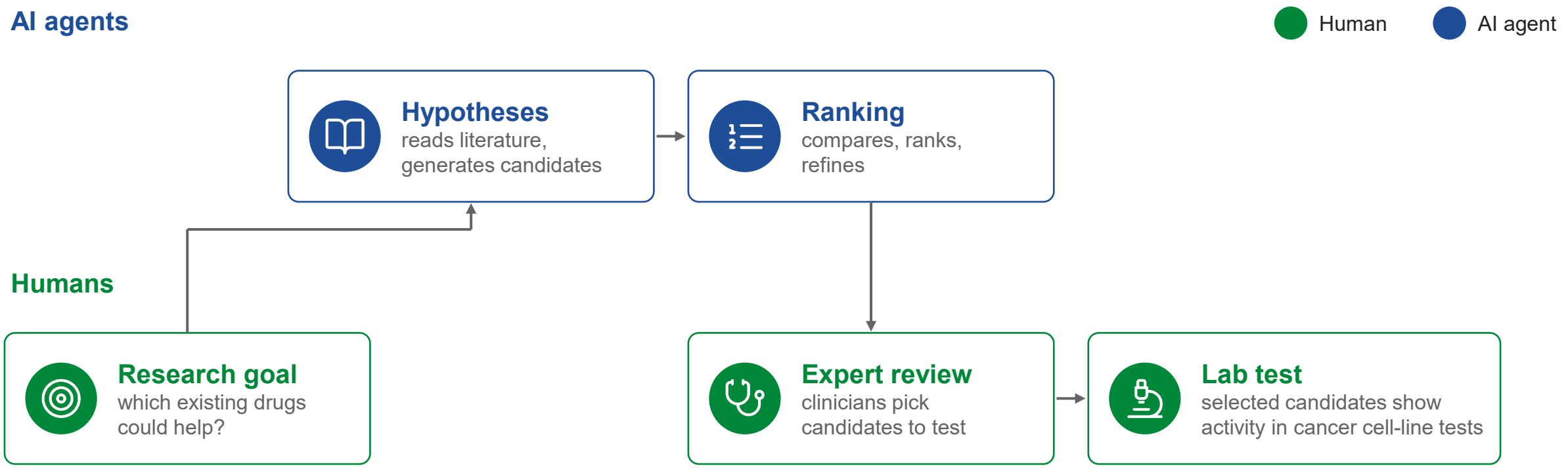


Illustration, not a theorem: the MSE map is unthresholded, the CHEM regions are thresholded (db4 wavelets, 4× super-resolution).

Different quantities reveal different aspects of the reconstruction error.

Agentic AI in Science: From Hypotheses to Experimental Validation

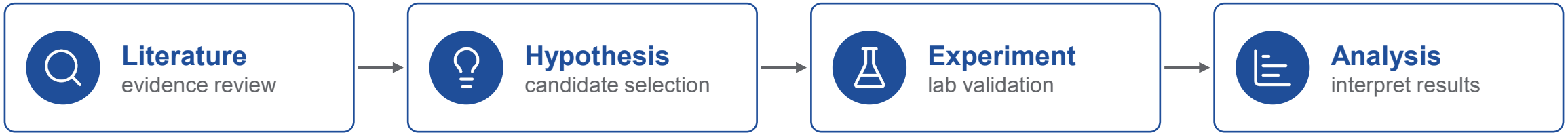
Co-Scientist Proposed Leukaemia Drug Candidates:



AI can propose hypotheses. Scientific evidence must validate them.

Source: Gottweis et al., Accelerating scientific discovery with Co-Scientist, Nature (2026).

Can Reliability Survive a Scientific Workflow?



An early error can affect every downstream decision.

Reliable individual steps ≠ a reliable scientific workflow

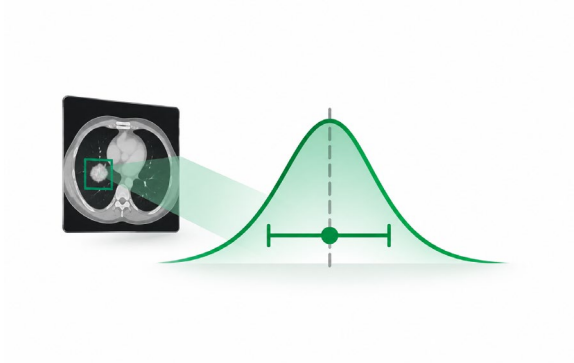
Open mathematical questions

Error propagation
How do errors accumulate over many steps?

Compositional guarantees
When do local bounds compose?

Task-level reliability
What can we guarantee overall?

Three Roles for Mathematics in Reliable AI



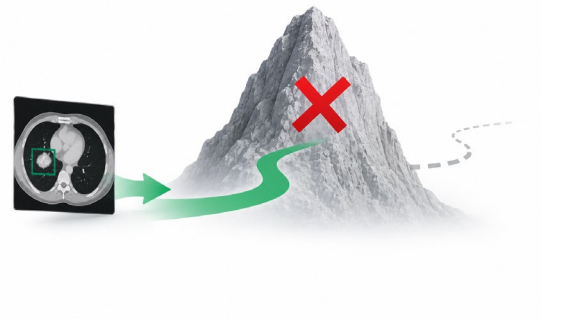
Quantification

- What can we measure?
- Can uncertainty and failure be quantified?



Guarantees

- What can be proven?
- Which properties can be ensured?



Limits

- What cannot be achieved in general?
- Are there fundamental boundaries?

Reliable AI requires understanding all three!

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

MATH 4AI

Can we compute it?

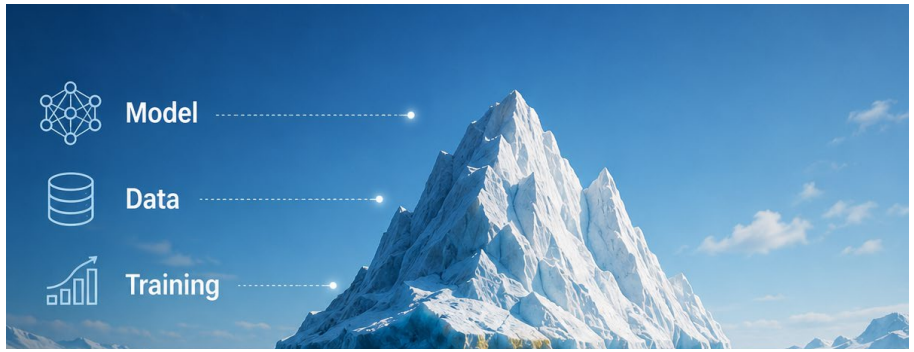
Fundamental limits of digital computation



A Hidden Assumption

So far:

- Object
- Representation
- Guarantee
- Composition



What Does Computable Mean?

Turing computability for problems with real-valued inputs (Church, Turing; 1936):

REAL-VALUED INPUT

x

Given to arbitrary precision

ONE ALGORITHM

A

Works for every admissible input

APPROXIMATION

y_n

Requested error: 2^{-n}

$$\|y_n - f(x)\| \leq 2^{-n}$$

Uniform guarantee across the specified class of inputs

But does such an algorithm always exist?

Hard Is Not the Same as Noncomputable

Computationally Hard

An algorithm exists, but the required resources may grow prohibitively.

How expensive?



Noncomputable

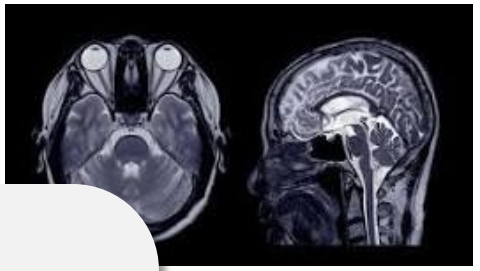
No effective algorithm exists with the requested guarantee over the entire problem class.

Possible at all?

Noncomputability of Finite-Dimensional Inverse Problems

Solution Set: For $A \in \mathbb{C}^{m\times \mathrm{N}}$ and $y \in \mathbb{C}^m$ let

$$\Psi(A,y) := \underset{x \in \mathbb{C}^N}{\arg\min} \|x\|_1 \text{ such that } \|Ax - y\|_2 \leq \epsilon.$$



Theorem (Boche, Fono, Kutyniok; 2023):

Fix parameters $\epsilon \in \left(0, \frac{1}{4}\right)$, $N \geq 2$, and $m < N$. There does *not exist a Turing-computable function* $\widehat{\Psi} : \mathbb{C}^{m\times \mathrm{N}} \times \mathbb{C}^m \to \mathbb{C}^N$ such that

$$\sup_{(A,y) \in \mathbb{C}^{m\times \mathrm{N}} \times \mathbb{C}^m} \left\|\Psi(A,y) - \widehat{\Psi}(A,y)\right\|_2 < \frac{1}{4}.$$

Even in finite dimensions, uniformly computable and arbitrarily accurate solutions cannot always be guaranteed.

Why Does This Matter for Trustworthiness?

Further Noncomputability Results:

- Optimization
- Classification
- Certain Partial Differential Equations

Many *guarantees* and *uncertainty bounds* are themselves defined through optimization or solution mappings.

Mathematical Definition $\not\Rightarrow$ ***Computational accessibility***

A guarantee that cannot be realized is not yet an operational reliability guarantee.

Beyond Digital Computing

Theorem (Boche, Fono, Kutyniok; 2024):
The corresponding inverse problem becomes *computable* on an *analog (Blum-Shub-Smale) machine*!

Computability depends on the computational model!

But idealized analog models do not yet imply physical realizability.

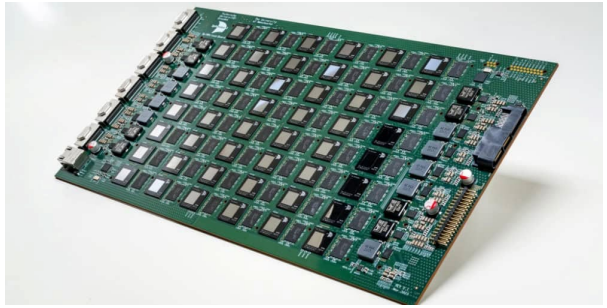
Alternative Hardware

Some recent developments:

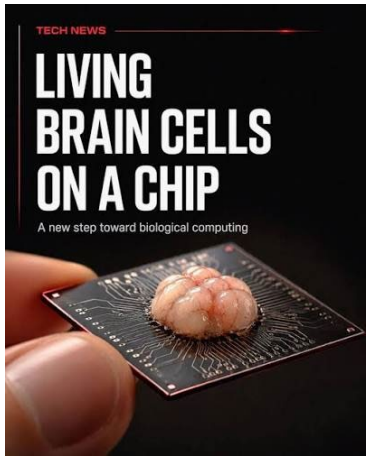
- Neuromorphic computing
- Biocomputing
- Quantum computing



Quantum Computing



Neuromorphic Chip



Biocomputing

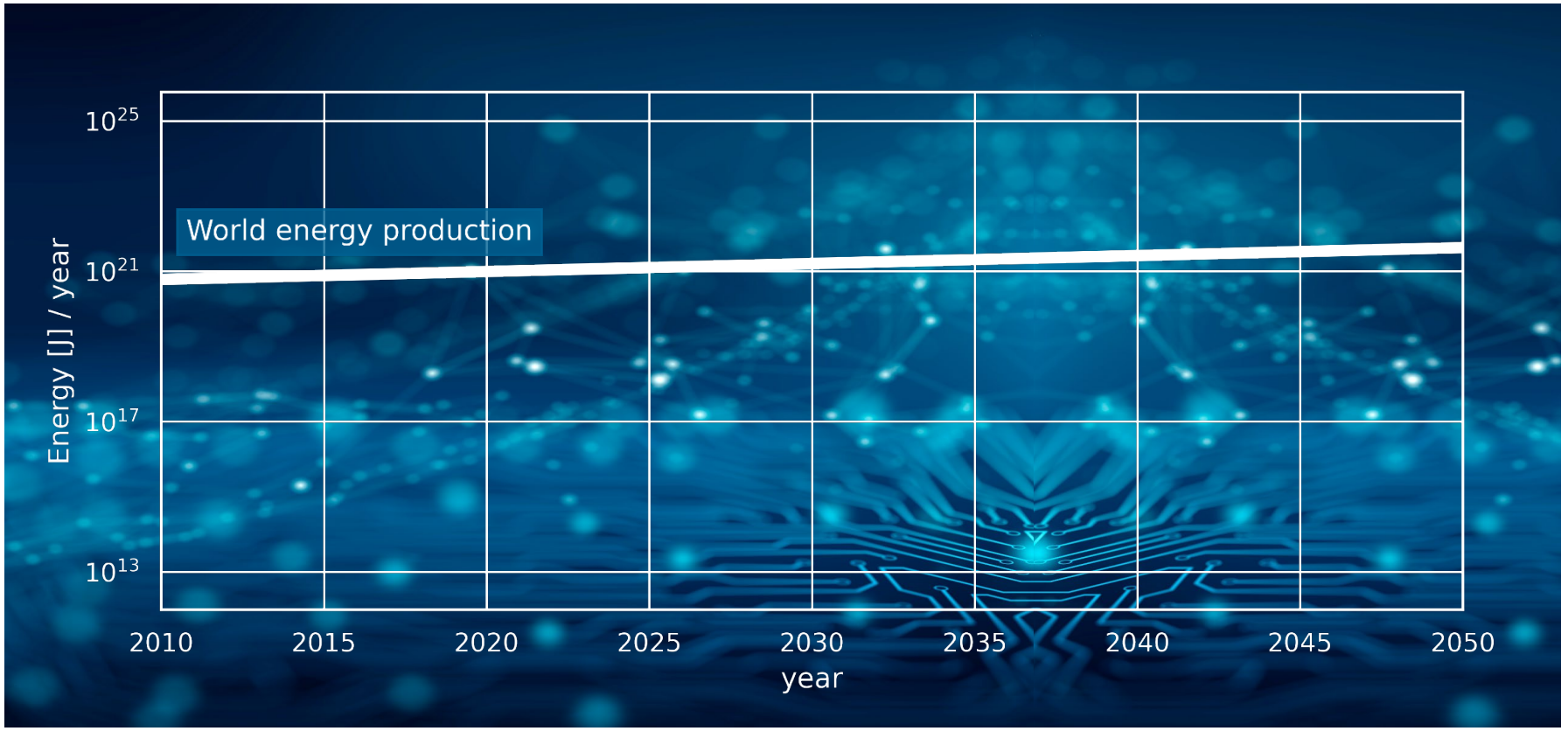
LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

MATH 4AI

Can we sustain it?

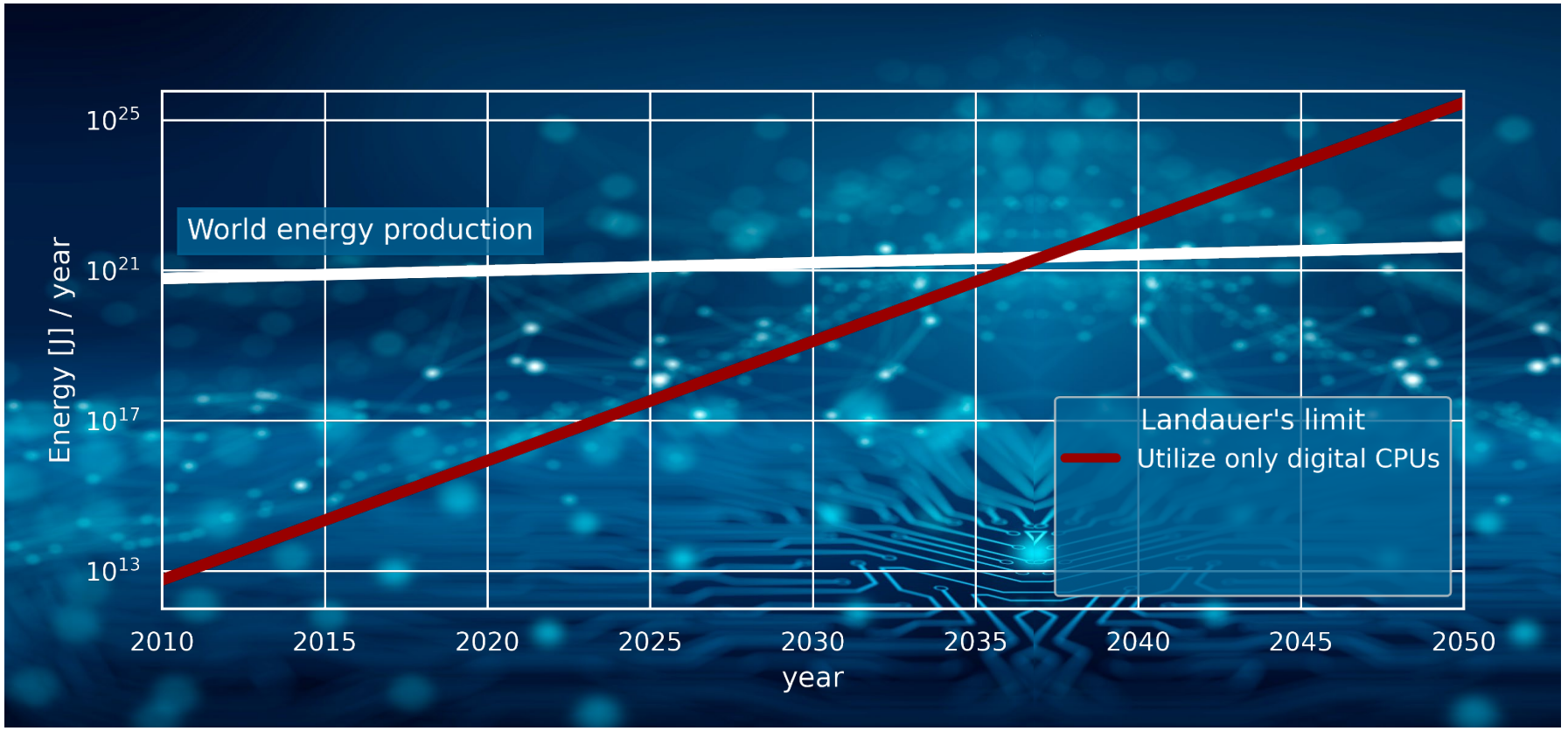
Rethinking computation

Scaling Alone Is Not a Sustainable Strategy



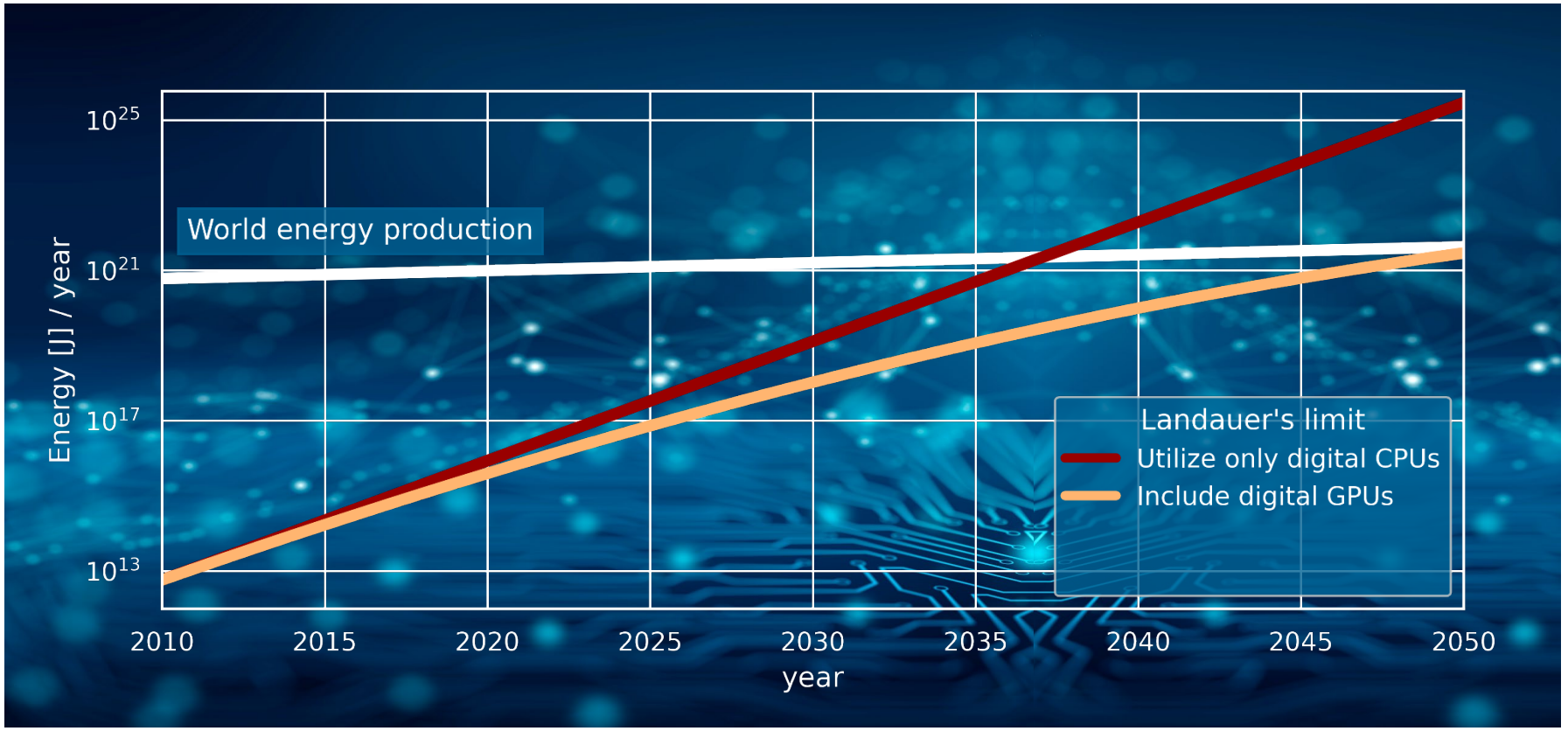
Source: Decadal Plan of the Semiconductor Research Corporation for the Biden (US) Administration, 2021

Scaling Alone Is Not a Sustainable Strategy



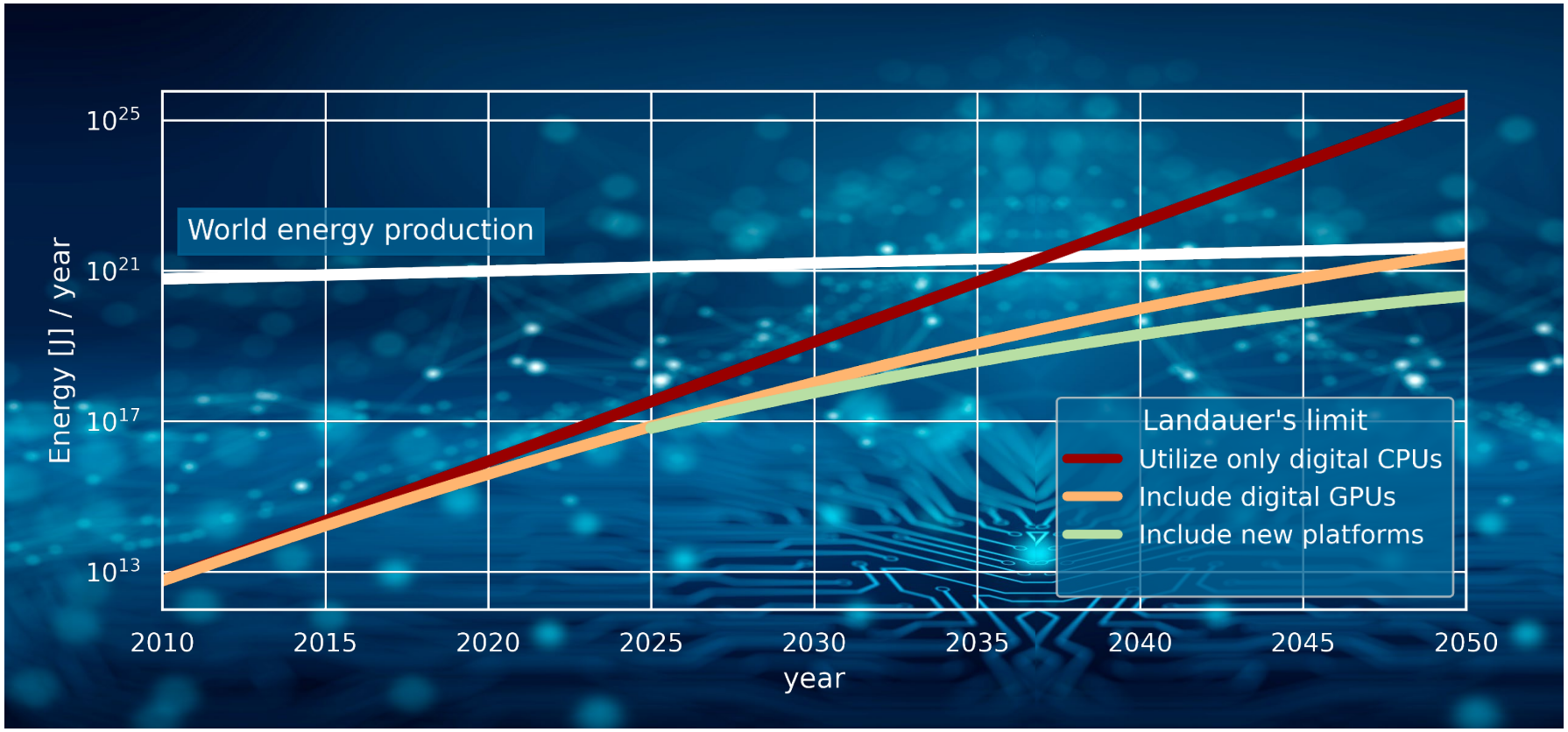
Source: Decadal Plan of the Semiconductor Research Corporation for the Biden (US) Administration, 2021

Scaling Alone Is Not a Sustainable Strategy



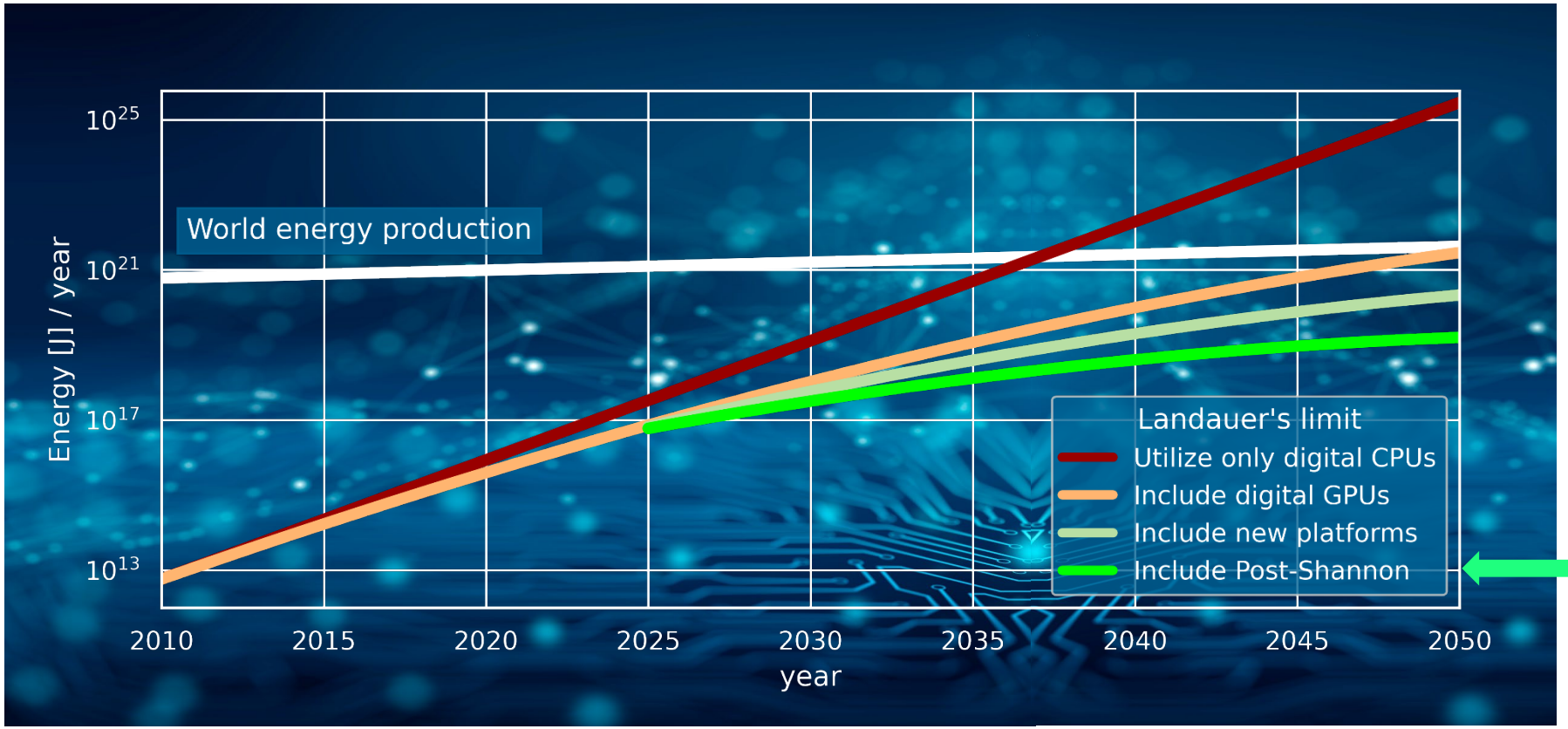
Source: Decadal Plan of the Semiconductor Research Corporation for the Biden (US) Administration, 2021

Scaling Alone Is Not a Sustainable Strategy



Source: Decadal Plan of the Semiconductor Research Corporation for the Biden (US) Administration, 2021

Scaling Alone Is Not a Sustainable Strategy



Novel Algorithmic Approach!

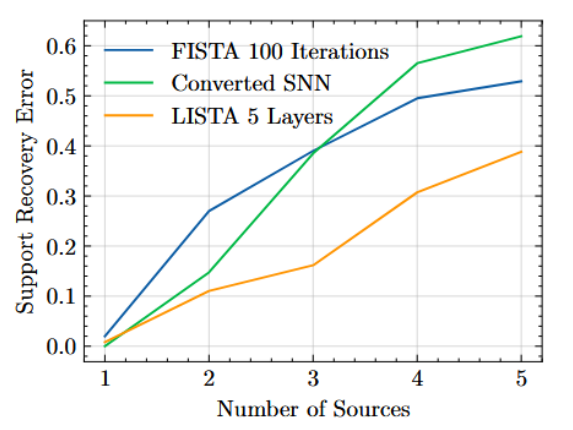
Source: Decadal Plan of the Semiconductor Research Corporation for the Biden (US) Administration, 2021

Neuromorphic Hardware: Measured Energy Efficiency

Comparing neuromorphic and conventional hardware for the same reconstruction task:

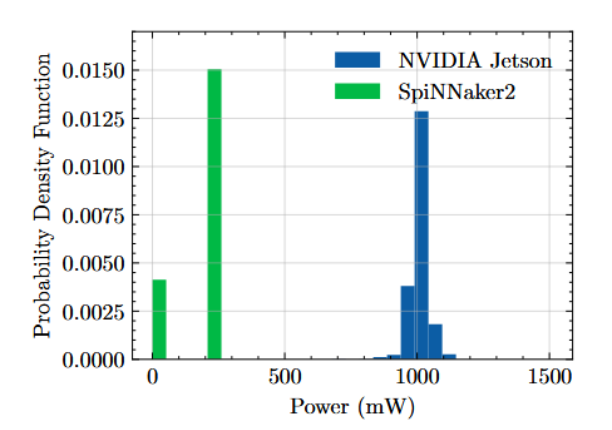
An unrolled iterative reconstruction (LISTA/FISTA) converted to a spiking network and run on neuromorphic hardware.

Accuracy: reconstruction error vs. number of sources



Converted spiking network comparable to FISTA (100 iterations); LISTA (5 layers) remains more accurate.

Power: measured on hardware

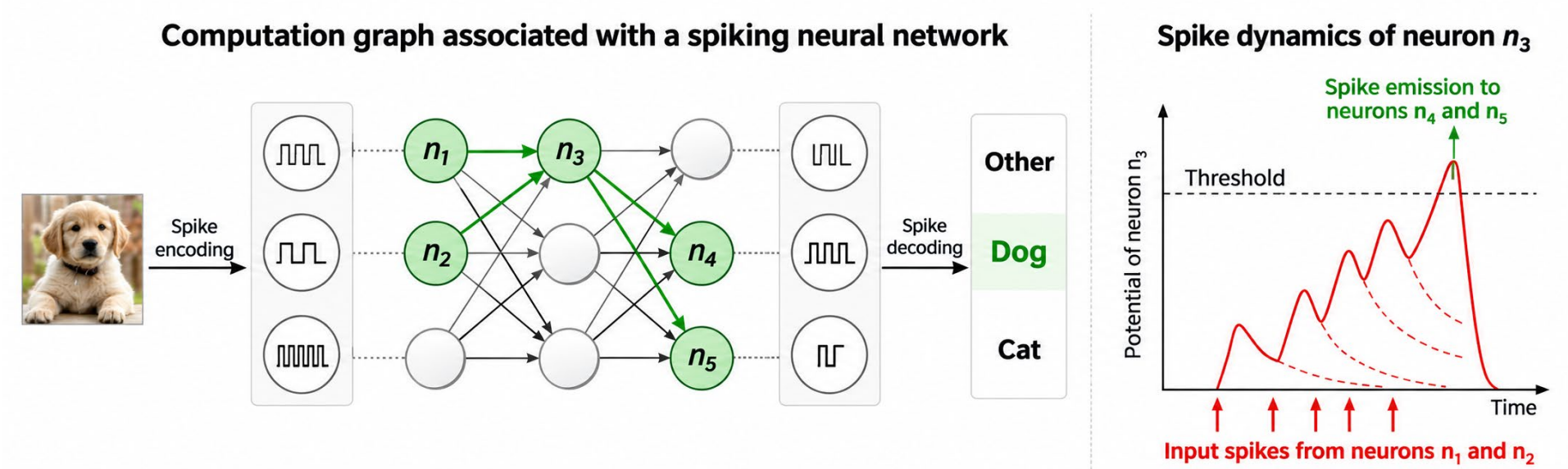


SpiNNaker2 (neuromorphic): a few hundred mW. NVIDIA Jetson (GPU): about 1 W. Same task.

Measured on hardware: a fraction of the power, at a modest cost in accuracy.

Source: Andrei, Drăgutoiu, Béna, Akl, Li, Lohrmann, Mönich, Boche: Deep-Unrolling Multidimensional Harmonic Retrieval Algorithms on Neuromorphic Hardware. Asilomar Conf. on Signals, Systems & Computers, 2024 (TU Munich, Imperial College London, SpiNNcloud Systems).

Why Does This Work? The Framework of Spiking Neural Networks



Remarks:

- *More biologically realistic* than first and second generation artificial neurons.
- Information is encoded in the *timing of individual spikes.*
- Numerous models for spiking neurons exist; one of those is the *Spike Response Model.*

Time is one crucial factor in this model!

Can Spiking Networks Represent Functions More Efficiently?

How expressive are spiking neural networks compared to classical networks?

Theorem (Singh, Fono, Kutyniok; 2025):

Simulating a classical neural network by a spiking neural network requires only a *linear* scaling of the number of neurons and the depth.

Can Spiking Networks Represent Functions More Efficiently?

Theorem (Singh, Fono, Kutyniok; 2025):

Simulating a spiking neural network by a classical neural network can require an *exponential* scaling of the number of neurons and the depth.

Spiking neural networks can be exponentially more efficient in representing functions!

What Theory Shows — and What It Does Not Yet

Proven

What mathematics guarantees

- Spiking networks can represent certain functions exponentially more efficiently.
- Computability can be restored on analog models of computation.

Measured

What hardware experiments show

- A fraction of the power on neuromorphic hardware for one algorithm class and one task.
- At a modest cost in accuracy.

Open

What remains to be shown

- Training, stability, and a theory of energy consumption.
- Hybrid systems and physical realizability.
- *Optimally sustainable hard-software combinations*

Towards Ecologic Intelligence: System-Level Co-Design

Efficiency and reliability are properties of the entire AI system – not of a model or a chip alone.

MATHEMATICAL FOUNDATIONS	ALGORITHMS & MODELS	SOFTWARE & SYSTEMS	HARDWARE ARCHITECTURES
▪ Expressivity ▪ Computability ▪ Error bounds	▪ Sparsity ▪ Pruning ▪ Quantization ▪ Event-driven computation	▪ Compilers ▪ Scheduling ▪ Hybrid digital–analog pipelines	▪ Neuromorphic ▪ Analog computing ▪ Specialized processors
Guarantees	**Efficiency by design**	**System-level optimization**	**Energy-efficient computing**

JOINT DESIGN UNDER EXPLICIT SYSTEM CONSTRAINTS

Reliability | Energy | Latency | Accuracy

From model-centric AI to system-centric AI!

Institutional Co-Design in Practice: Project

gAIn Next Generation AI Computing

MATH 4AI

MATHEMATICAL FOUNDATIONS	COMPUTING ARCHITECTURES	APPLICATION DOMAINS
Computability · Expressivity · Reliability	Algorithms · Software · Hardware efficiency	Communications · Medical AI · Robotics

Connecting foundations, computing and applications across institutions!

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

TUM

TECHNISCHE UNIVERSITÄT DRESDEN

Bayerisches Staatsministerium für Wissenschaft und Kunst

STAATSMINISTERIUM FÜR WISSENSCHAFT KULTUR UND TOURISMUS

Freistaat SACHSEN

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

MATH 4AI

Can we shape it?

The future of informatics

Three Priorities for European Informatics

MATH 4AI

Foundations



Strengthen the mathematical core

- Understand learning, reliability and the limits of computation
- Treat foundations as strategic research, not as a niche

Co-Design



Algorithms, hardware, theory

- Joint programmes across mathematics, informatics and hardware
- From model-centric to system-centric AI

Education



Understanding, not only use

- Graduates who can verify, bound and design AI systems
- Use AI in teaching – and teach how to check it

Europe should not only participate in the AI revolution – it should help define its foundations.

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Educating the Next Generation: From AI Use to AI Understanding

When AI can generate code and solutions, what becomes essential in informatics education?

CORE FOUNDATIONS — STILL ESSENTIAL

- Write code and design algorithms
- Understand models and abstractions
- Test and evaluate performance
- Optimize individual components

NEW DEMANDS IN THE AGE OF AI

- Audit AI-generated code and reasoning
- Assess guarantees, uncertainty and limits
- Verify end-to-end agentic workflows
- Co-design reliable and efficient AI systems

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

MATH 4AI

Conclusions

Three Messages for the Future of AI

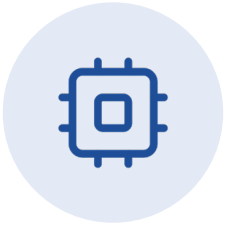
MATH 4AI



Can we trust it?

Reliability

- Only with guarantees, not just performance.



Can we compute it?

Fundamental limits

- Not always – and the computational model matters.



Can we sustain it?

Sustainability

- Only by rethinking computation, not just scaling.

Can we shape it?

Shaping all three is a task for informatics – and for Europe.

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

AI HUB LMU

MATH 4AI

Thank you very much for your attention!

