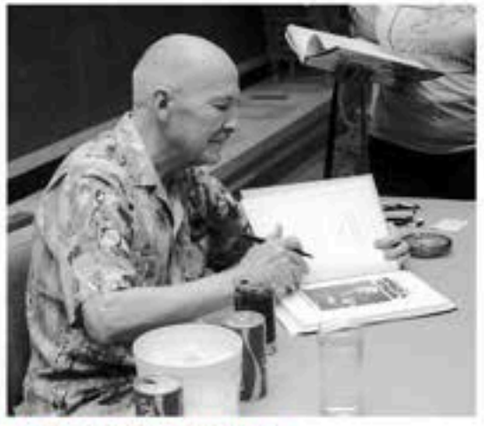
KI im Unternehmen

Andreas Windisch
Joanneum Research, DIGITAL
Intelligent Vision Applications

Deutschlandsberg
15. Juni 2026

Defining (artificial) intelligence

“A human being should be able to change a diaper, plan an invasion, butcher a hog, conn a ship, design a building, write a sonnet, balance accounts, build a wall, set a bone, comfort the dying, take orders, give orders, cooperate, act alone, solve equations, analyze a new problem, pitch manure, program a computer, cook a tasty meal, fight efficiently, die gallantly. Specialization is for insects.”

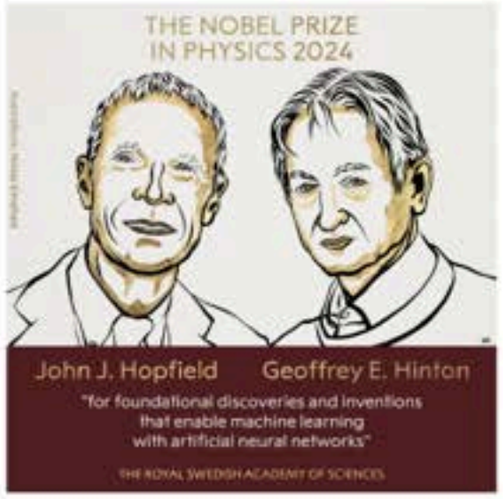


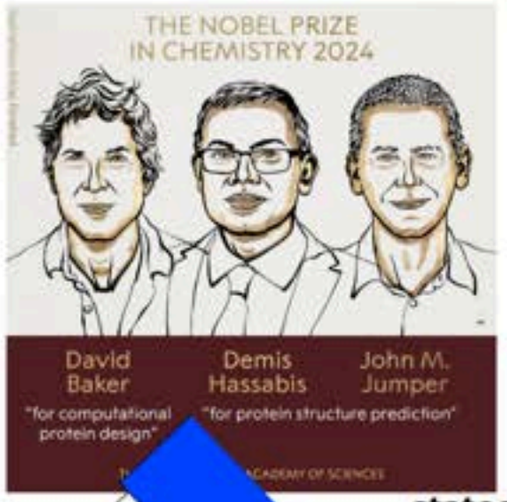
Source: wikipedia

Robert A. Heinlein, SciFi Author, 1907 - 1988

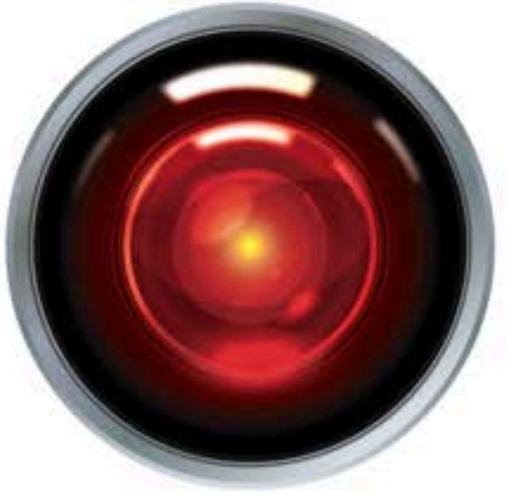
Global AI Revolution

stateof.ai 2023

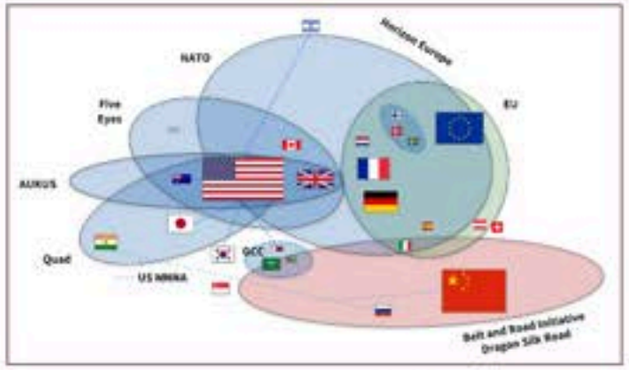




stateof.ai 2024

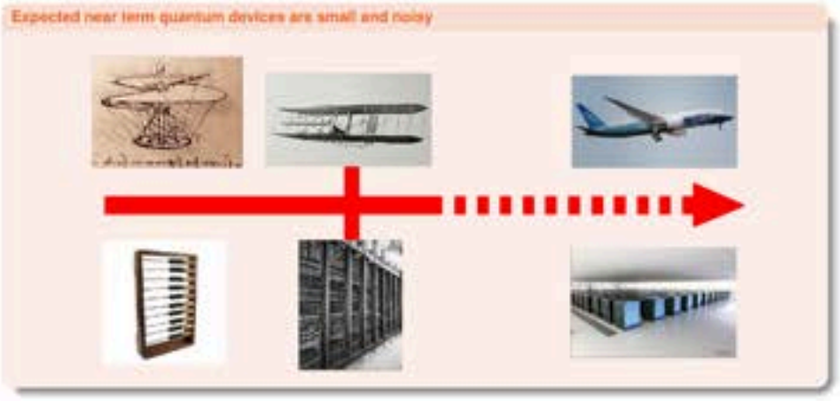


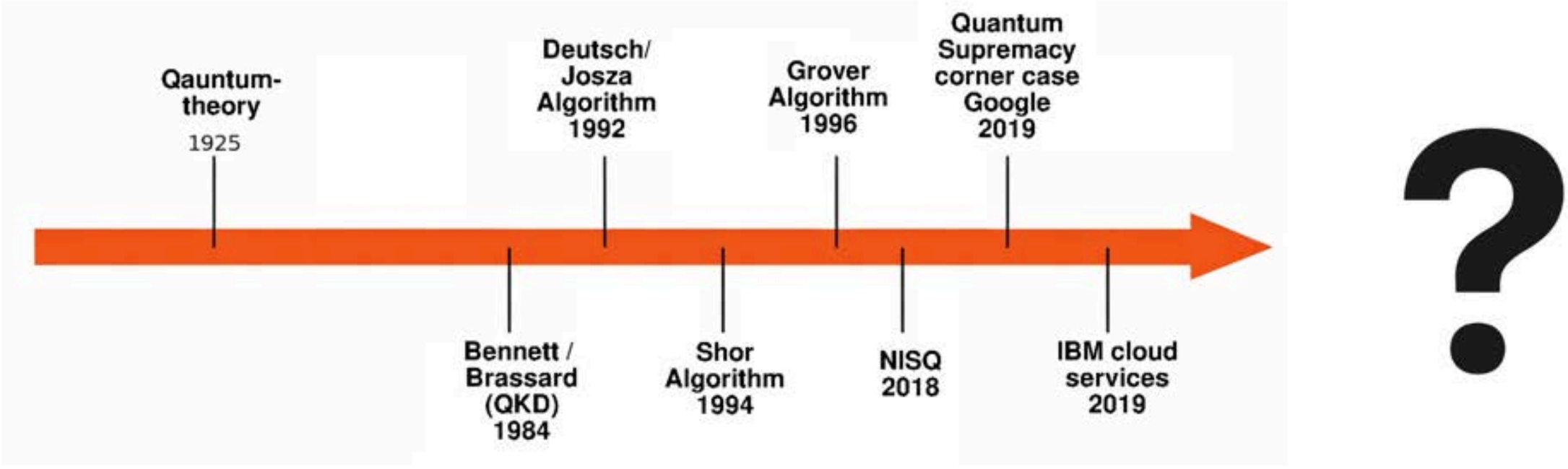
Global Quantum Developments



Quantum Computing Report
by CQI

Alain Aspect, John F. Clauser and Anton Zeilinger. Credit: Ill. Niklas Elmehed © Nobel Prize Outreach





Trends in - Research (State of AI Report Oct 24)

JOANNEUM RESEARCH DIGITAL

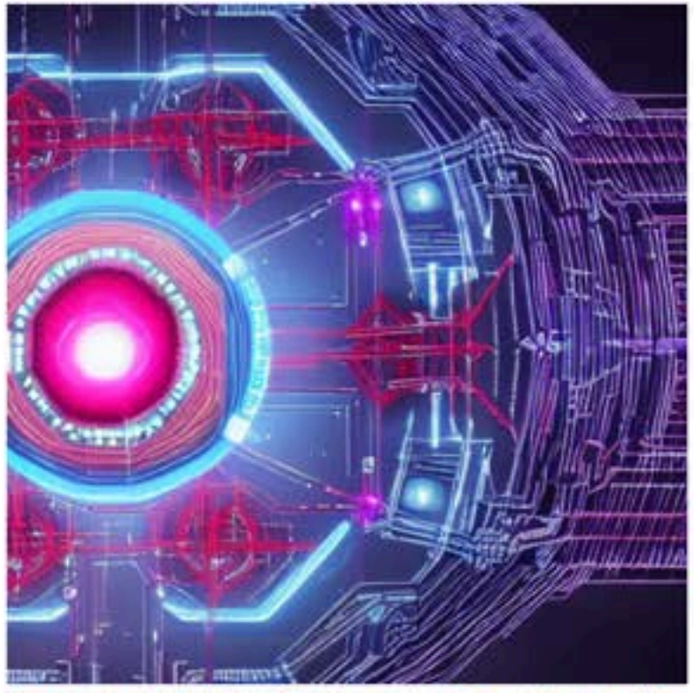


Image generated with Stable Diffusion model v1.5-pruned-emaonly.ckpt using ComfyUI and prompt ‚ai research'

Input/Output types:

- 📝: Text
- 🖼️: Image
- </> : Code
- 🛠️ : Software tool use
- 🎥 : Video
- 🎵: Music
- 🧊: 3D
- 🤖: Robot state

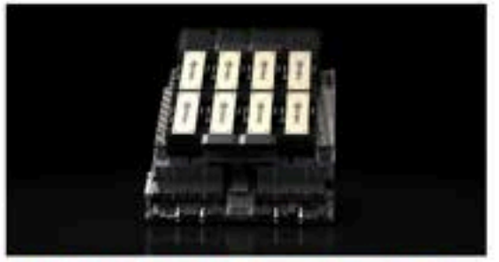
Model types:

- 📝 → 📝 : LLMs
- 📝+ 🖼️ → 📝 : Multimodal LLMs
- 📝+ 🖼️+ 🤖 → 📝 : Multimodal LLMs for Robotics
- 📝 → </> : Text to Code
- 📝 → 🛠️ : Text to Software tool use
- 📝 → 🖼️ : Text to Image
- 📝 → 🎥 : Text to Video
- 📝 → 🎵 : Text to Music
- 🖼️ → 🧊 : Image to 3D
- 📝 → 🧊 : Text to 3D

- Scientific and algorithmic discovery
  - AlphaEvolve
- new frontier: reasoning and planning
  - OpenAI o1: complex, reinforcement learning based, self improved chain-of-thought prompting
  - OpenAI o1 scores 84 on AIME 2024 (math competition) vs score of 13 of GPT-4
- LlaMA3 closes gap between open and closed models
  - LlaMA3 405B largest model to date, keeping up with GPT-4o and other across many tasks
  - LlaMA3 405B trained on over 16000 H100 GPUs
  - LlaMA models 440M downloads on HuggingFace
- Neuro-symbolic systems
  - AlphaGeometry: symbolic deduction engine + LLM, performs like human gold medalist on Math Olymp.
- Efforts to shrink models, also deploying distilled models
- LLMs and multimodal models for smart phones (large scale end user deployment)

Trends in AI 2025 - Industry

Image generated with Stable Diffusion model v1.5-pruned-emaonly.ckpt using ComfyUI and prompt ‚laidessy'ch'



NVIDIA H200 Tensor Core, source: nvidia.com

- NVIDIA is king, hits > $4T market cap in summer/fall 2025
  - H200 currently prime compute HW, Blackwell B200 cut costs and energy consumption (25x)
  - AMD, Intel way behind in revenue/earnings per share
  - Austria: Project MUSICA (VSC Research): Vienna, Linz, Innsbruck, 40 Pflops. GPU+Quantum, €40M
- Saudi Arabia: Data is the new oil: >3000 H100 GPUs (~$30000 / unit)
- Europe: Mistral in the lead, AI Factories, AI Giga Factories
- GitHub CoPilot still in the lead, but new AI coding companies emerging
  - CoPilot leading to productivity increase for developers (~30% increase less experienced users)
- Self driving companies Wayve and Waymo power ahead
  - After years of disappointment, industry seems to be booming
- Text-to-speech booming
  - ElevenLabs unicorn status
- Avatar video generation products growing strongly

Trends in AI - Politics

Image generated with Stable Diffusion model v1.5-pruned-emaonly.ckpt using ComfyUI and prompt ‚politics'

- US: Executive Order 14110
  - Prevent AI enabled threats to civil rights and national security, signed Oct 2023
  - Several federal organizations have to create a Chief AI Office positions
  - States puruse own rules (California SB 1047, safety and liability regime for foundationmodels)
  - US labs struggling with EU regulation (e.g. Apple and EU Digital Market Act)
- Europe: EU AI Act
  - Added regulation: foundation models and general purpose AI systems
  - High risk definition of use of AI, all commercial foundation model providers subject to rules
  - Risk assessment, disclosing AI generated content, preventing illegal content
  - Prevent summaries of copyrighted data used for training
  - EU AI Act passed legislative stages March 13 2024
- UK: signals to deviate from regulating with existing legislation
  - but ruling our EU style regulation
  - signals to codify previously voluntary commitments
- China: Strong regulation framework enters enforcement era
  - specific regulation for recommender systems
  - generative AI guardrails in place
  - protections against misuse, no anonymous access, cencorship requirements
  - developers have to register algos with government for a security assessment

Trends in AI - Safety

Image generated with Stable Diffusion model v1.5-pruned-emaonly.ckpt using ComfyUI and prompt ‚ai saefty'

Malicious Use

- ✗ Bioterrorism
- ✗ Surveillance State
- ✓ Access Restrictions
- ✓ Legal Liability

AI Race

- ✗ Automated Warfare
- ✗ Evolutionary Pressures
- ✓ International Coordination
- ✓ Safety Regulation

Organizational Risks

- ✗ Weak Safety Culture
- ✗ Leaked AI Systems
- ✓ Information Security
- ✓ External Audits

Rogue AIs

- ✗ Power-Seeking
- ✗ Deception
- ✓ Use-Case Restrictions
- ✓ Safety Research

Source: arXiv:2306.12001

- **Catastrophic AI risks**
  - No consensus as of what these risks actually are
- **Debate of x-risk*) AI has hit mainstream hard this year**
  - ‘We must slow down the race to God-like AI’
  - ‘How Rogue Ais may Arise’
  - ‘Pausing AI Development Isn’t Enough. We Need to Shut it All Down’
  - AI godfather Geoff Hinton warns of dangers as he quits Google
  - Yoshua Bengio warns of potential goal of self-preservation
- **Critics of x-risk AI state:**
  - x-risk arguments are just a conjecture
  - Yann LeCun: at this early stage it is premature to discuss these risks
- **AI safety gets attention from government**
- **Number of publications in x-risks go up significantly (topped by ethics, though)**

*) x-risk AI: existential risk

What Is AI? A Functional View: Pre-Gen Era

- **Predict** 

They learn patterns in data and forecast what is likely to happen
(e.g., energy demand, machine behavior, price curves, etc)

- **Optimize** 

They find the best settings, actions, or schedules under constraints
(e.g., reducing peaks, balancing loads, improving parameters, etc)

- **Detect** 

They identify anomalies, defects, or inefficiencies in real time
(e.g., abnormal machine states, wear, unstable combustion, etc)

*Together, these capabilities let **AI improve energy efficiency**, stability, and throughput.*

What Is AI? A Functional View

- **Predict**

They learn patterns in data and forecast what is likely to happen
(e.g., energy demand, machine behavior, price curves, etc)

- **Optimize**

They find the best settings, actions, or schedules under constraints
(e.g., reducing peaks, balancing loads, improving parameters, etc)

- **Detect**

They identify anomalies, defects, or inefficiencies in real time
(e.g., abnormal machine states, wear, unstable combustion, etc)

*Together, these capabilities let **AI improve energy efficiency**, stability, and throughput.*

What Is AI? A Functional View

- **Predict**

They learn patterns in data and forecast what is likely to happen
(e.g., energy demand, machine behavior, price curves, etc)

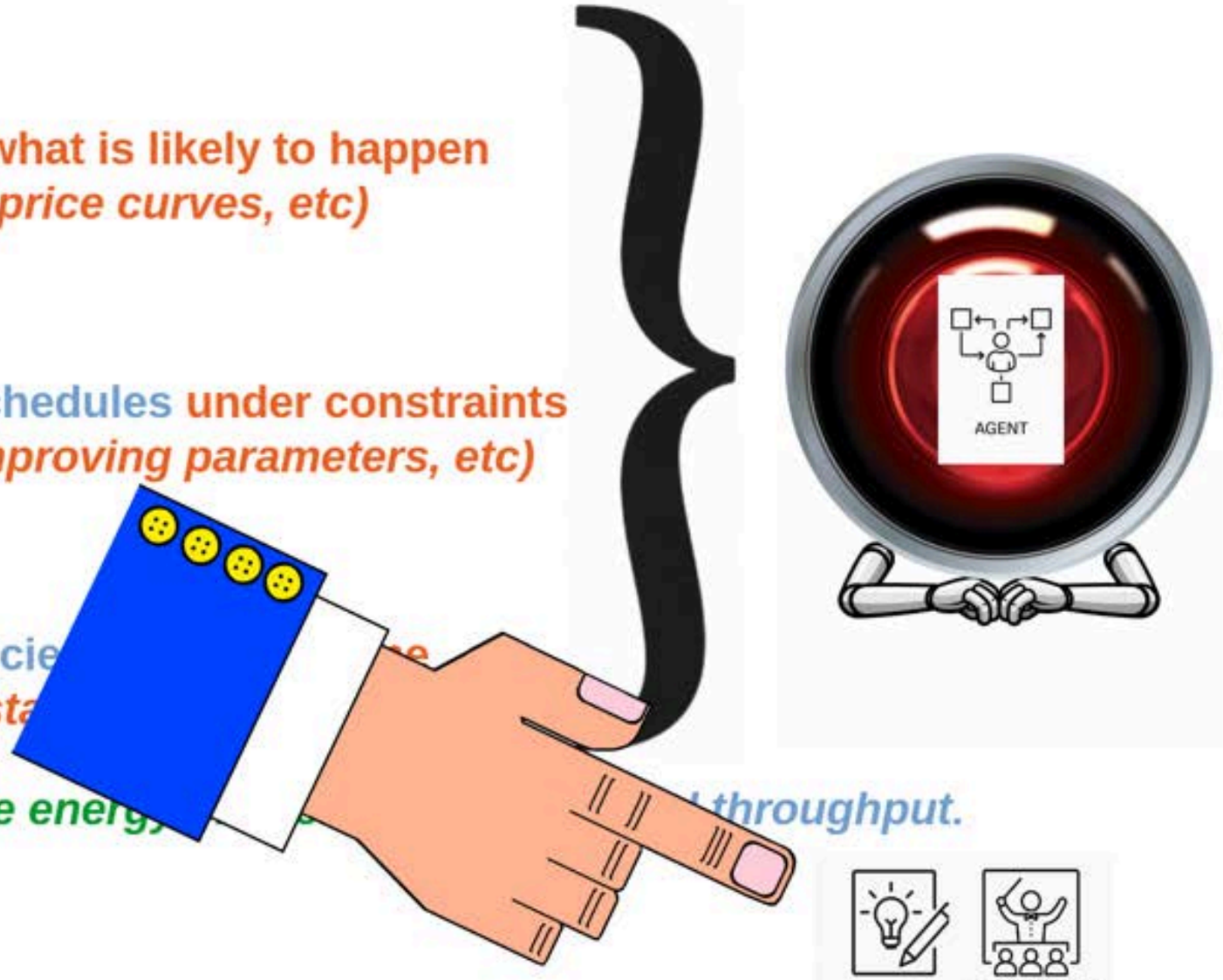
- **Optimize**

They find the best settings, actions, or schedules under constraints
(e.g., reducing peaks, balancing loads, improving parameters, etc)

- **Detect**

They identify anomalies, defects, or inefficie
(e.g., abnormal machine states, wear, unsta

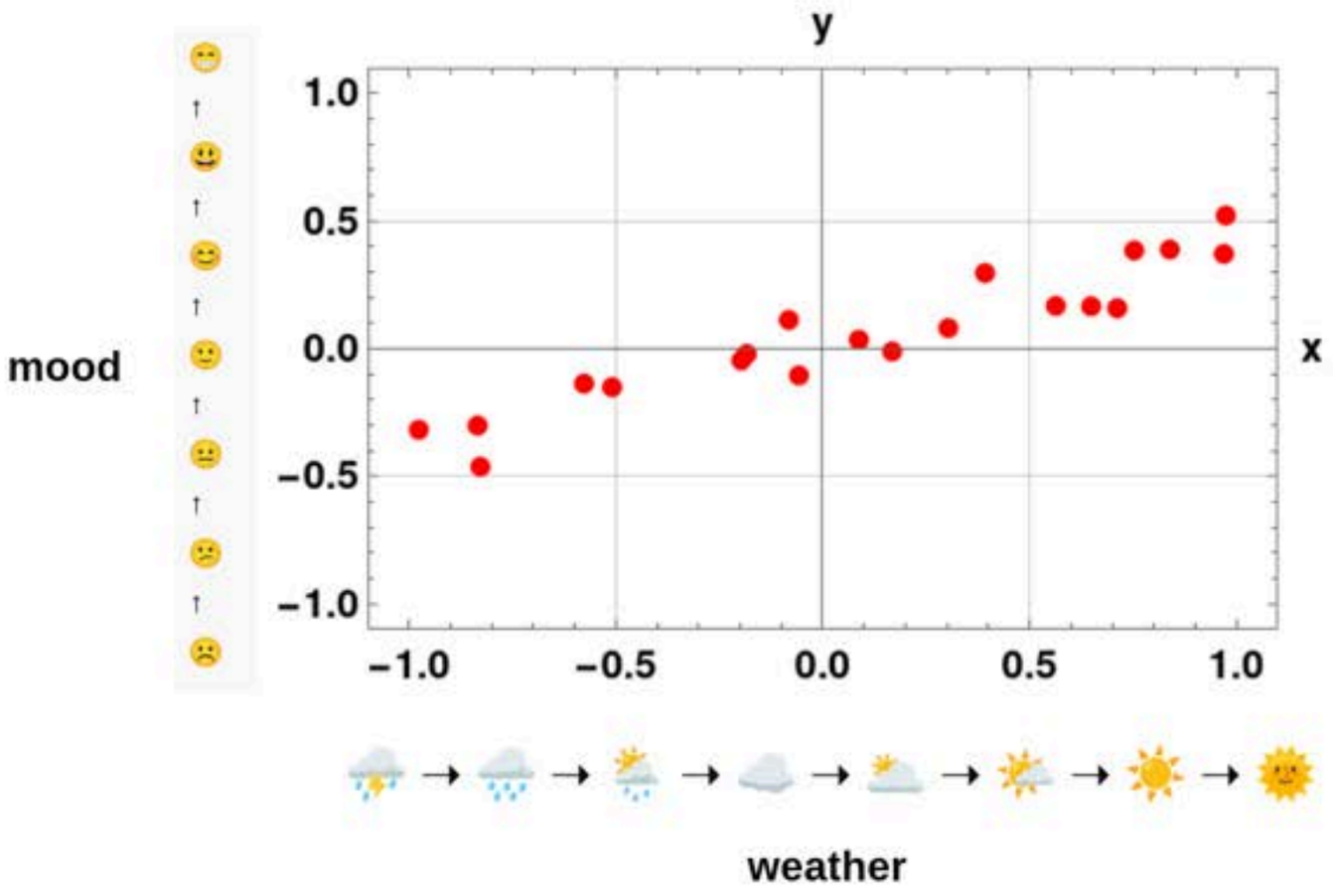
*Together, these capabilities let **AI improve energy** throughput.*

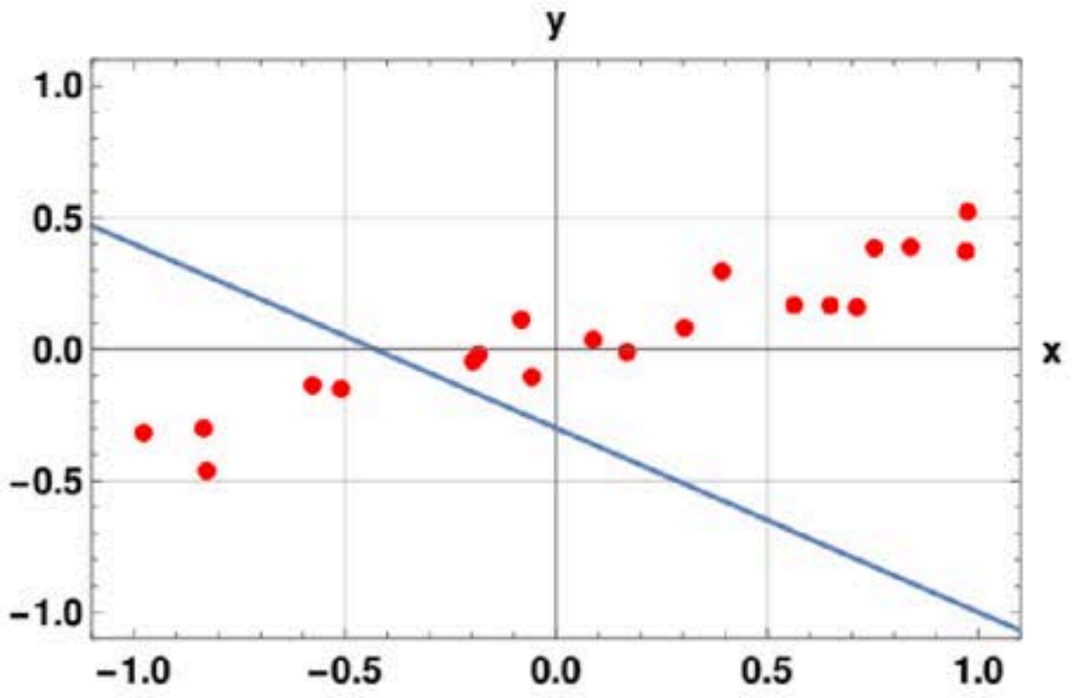


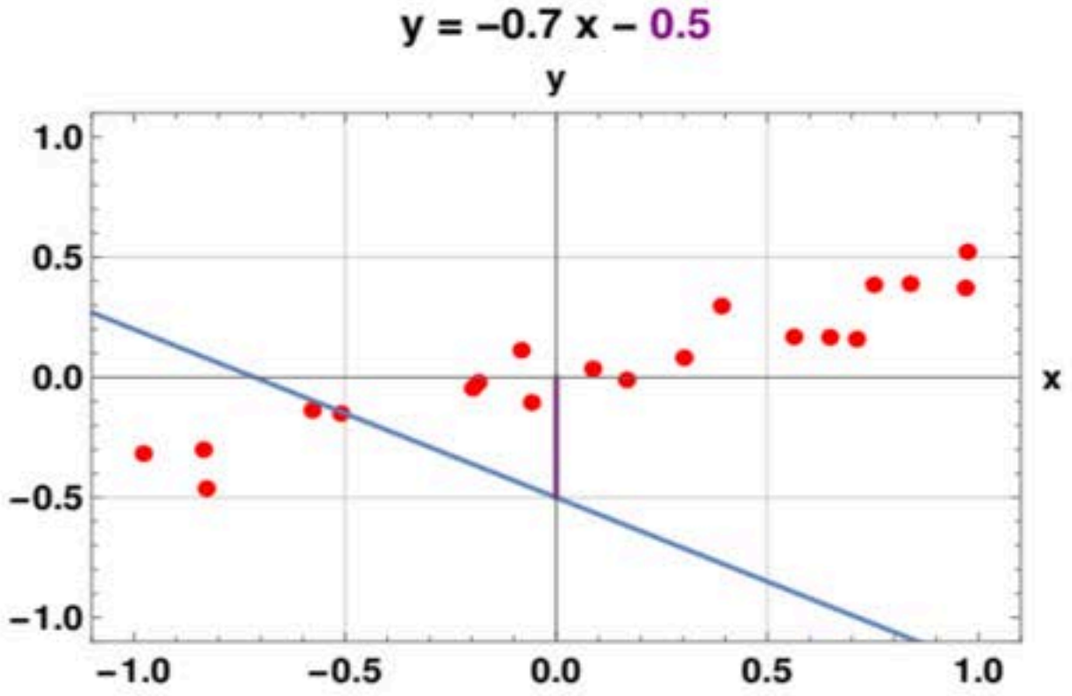
Predictive ML/AI

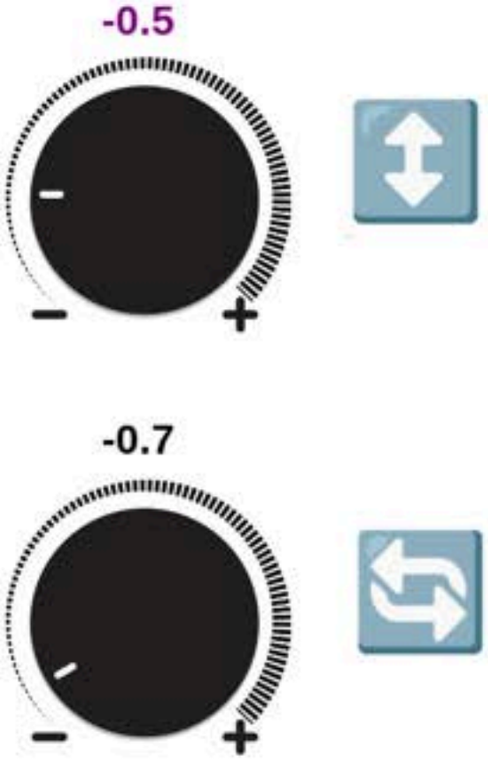
$$y = f(x)$$

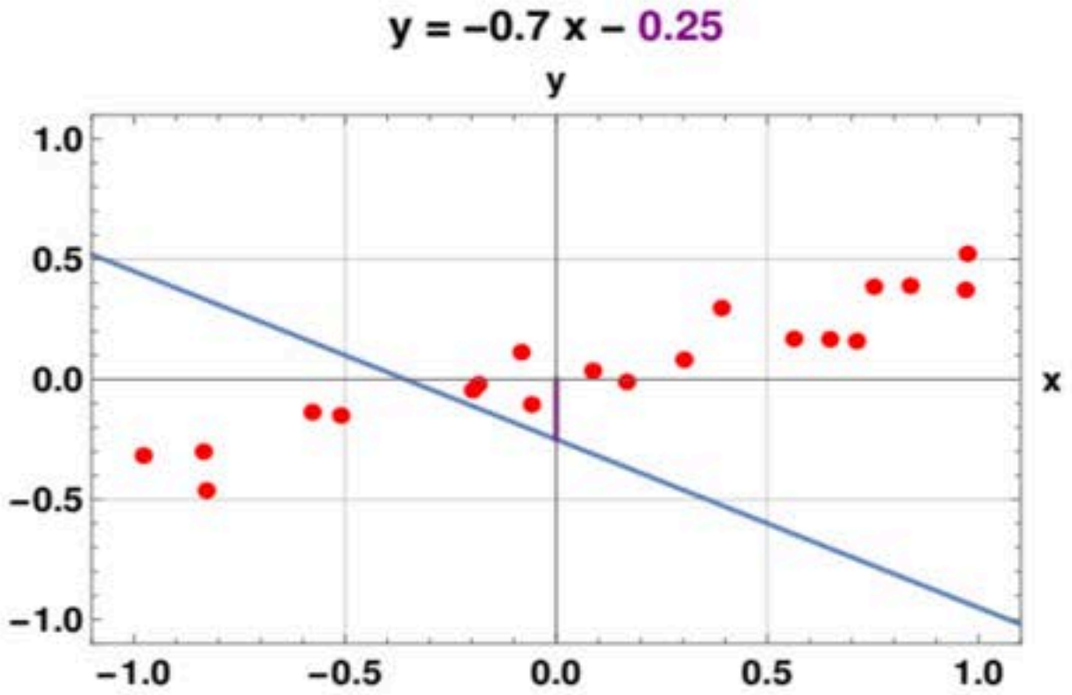
- Predictive maintenance
- Process forecasting
- Demand prediction
- Energy consumption
- Anomaly detection

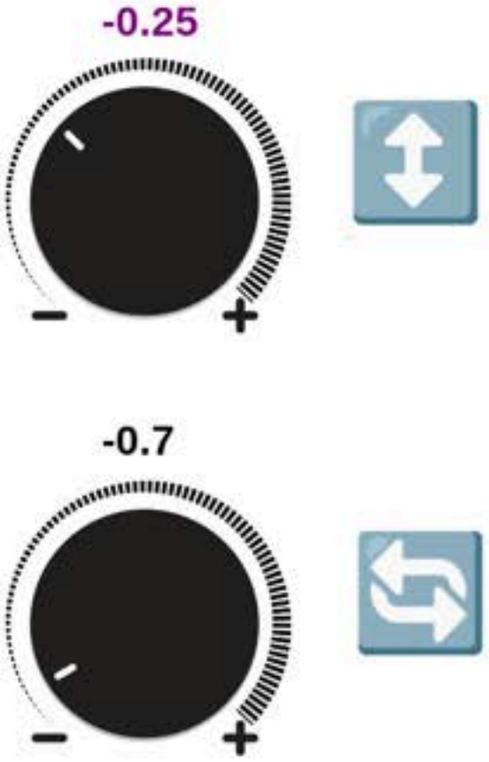

y
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
x
mood
weather

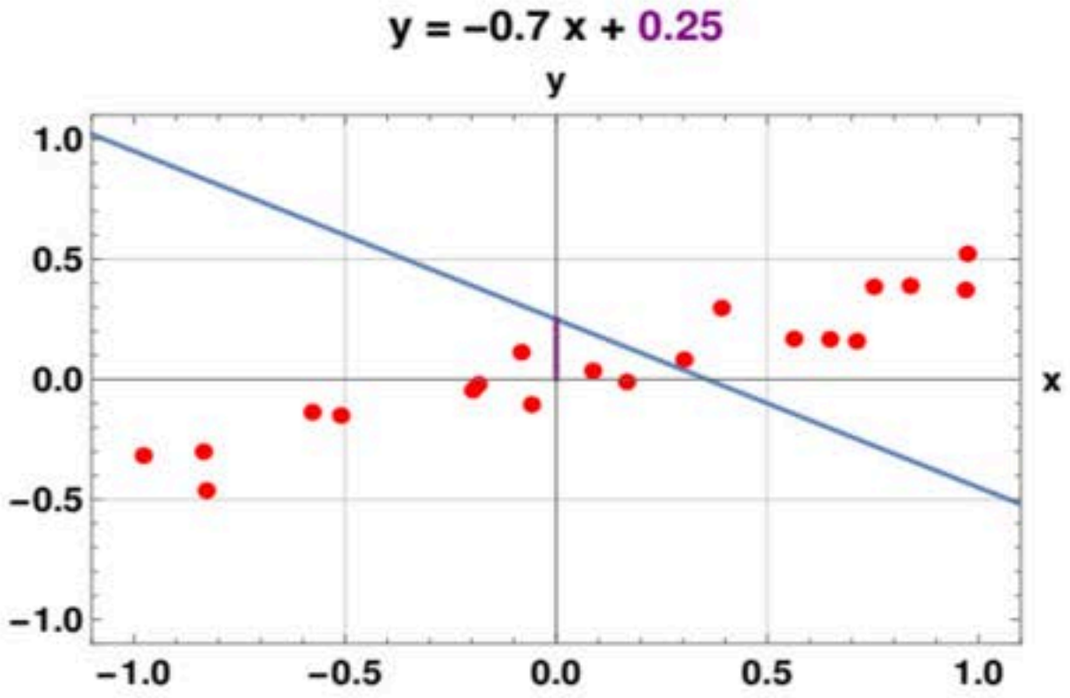

y
1.0
0.5
0.0
−0.5
−1.0
x
−1.0
−0.5
0.0
0.5
1.0

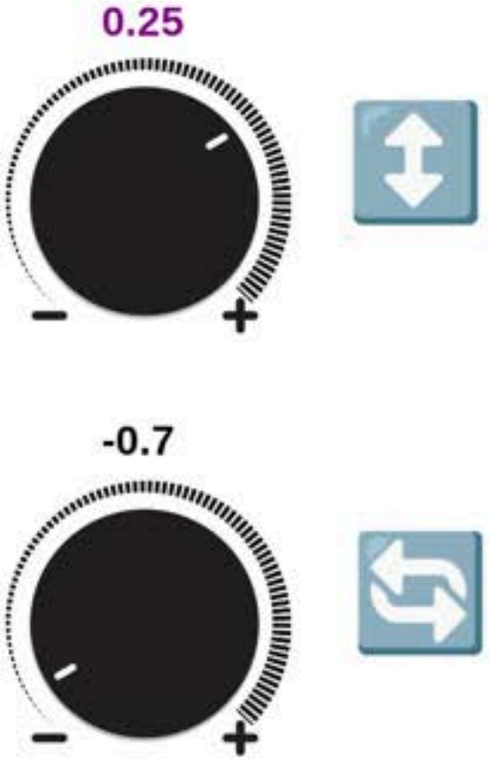

y = –0.7 x – 0.5
y
1.0
0.5
0.0
–0.5
–1.0
–1.0
–0.5
0.0
0.5
1.0
x

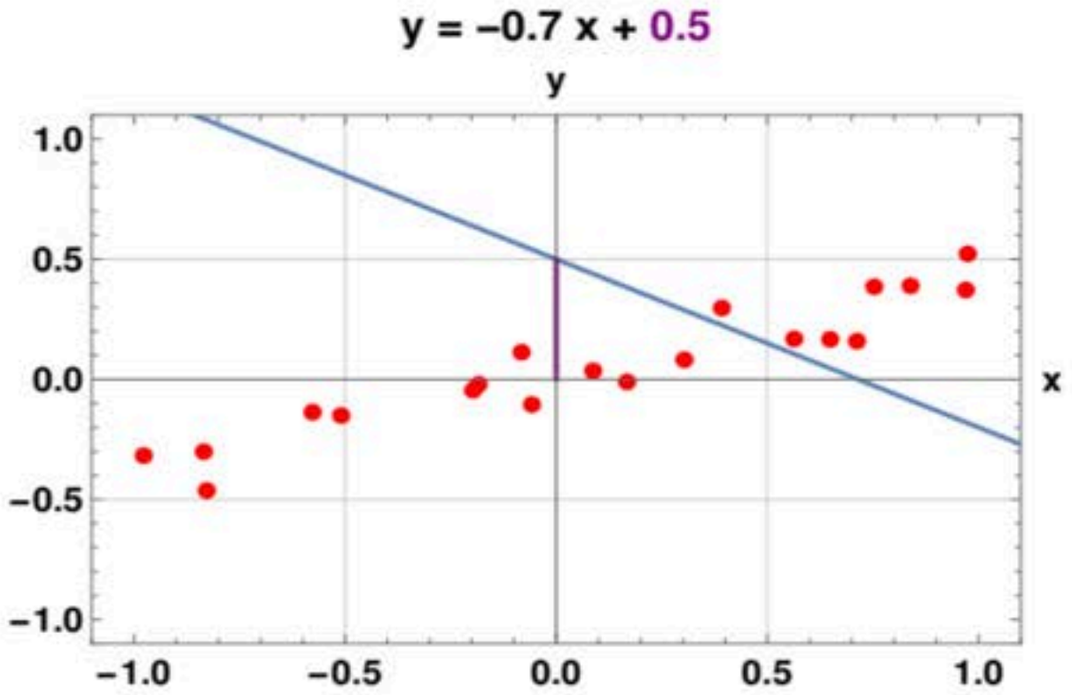

-0.5
-0.7

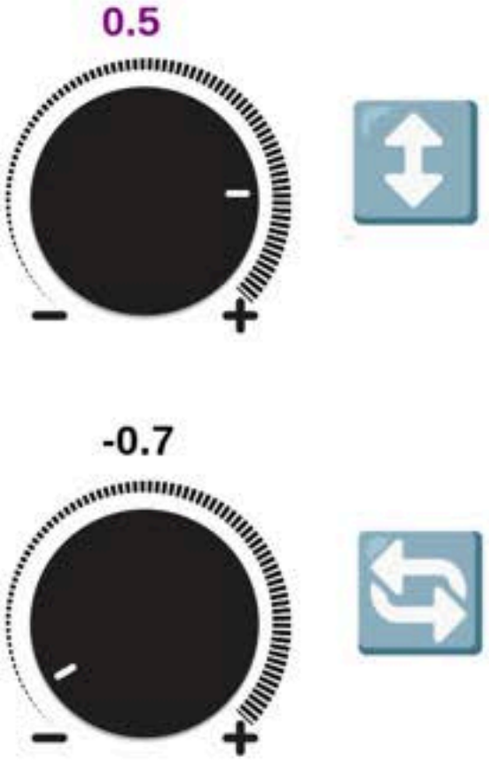

y = −0.7 x − 0.25
y
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
x

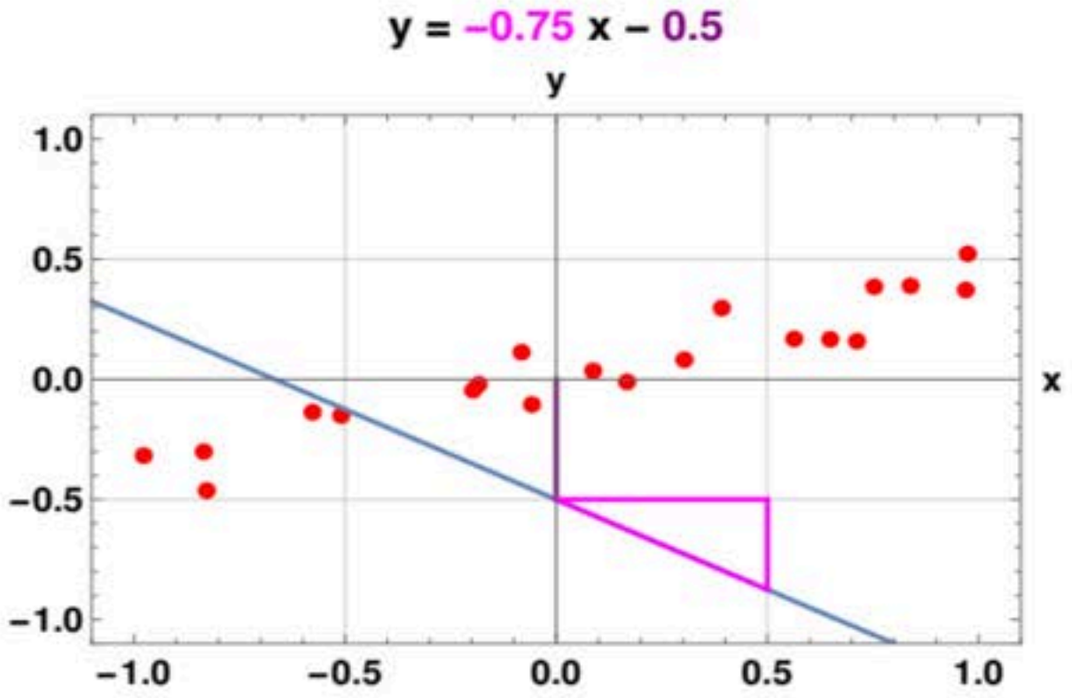

-0.25
-0.7

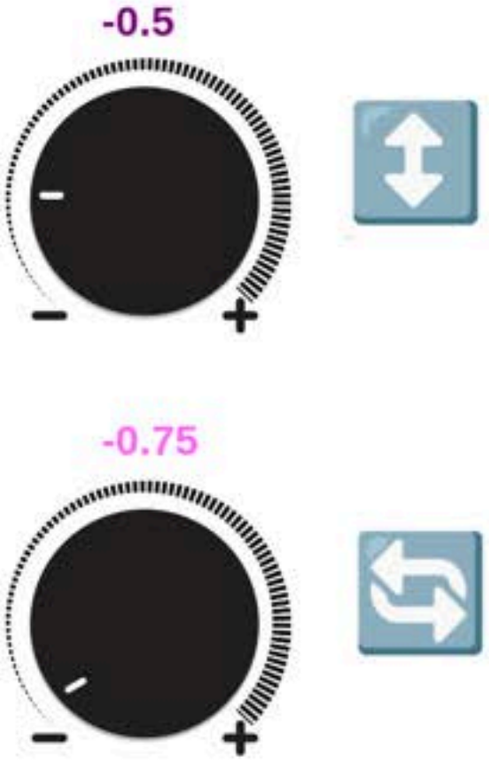

y = −0.7 x + 0.25
y
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
x

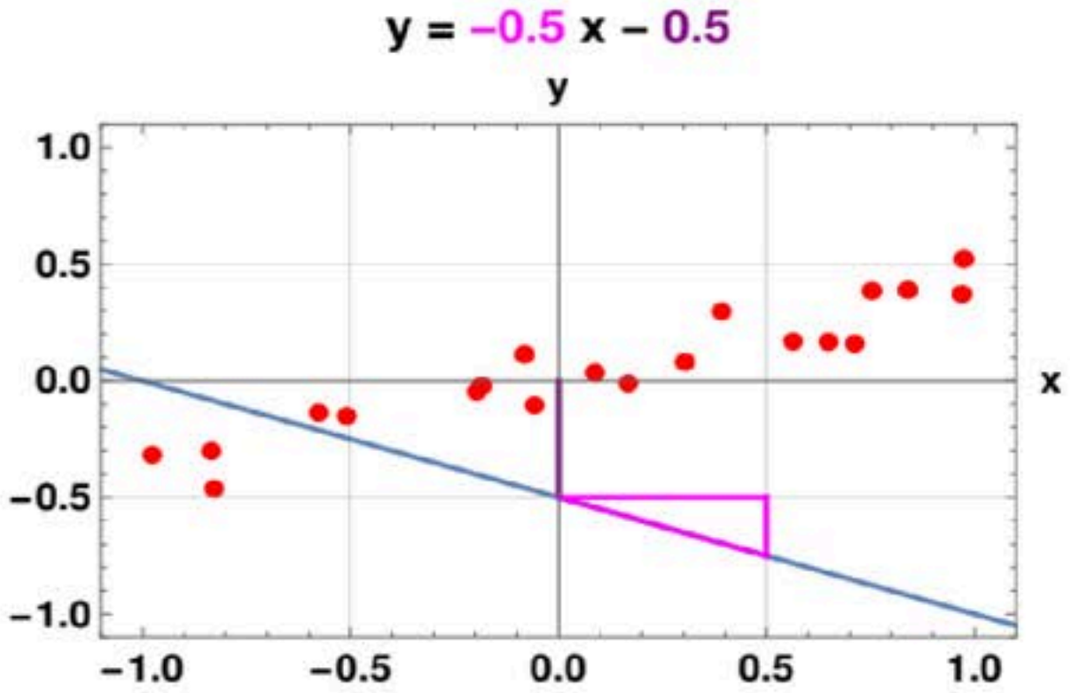

0.25
-0.7

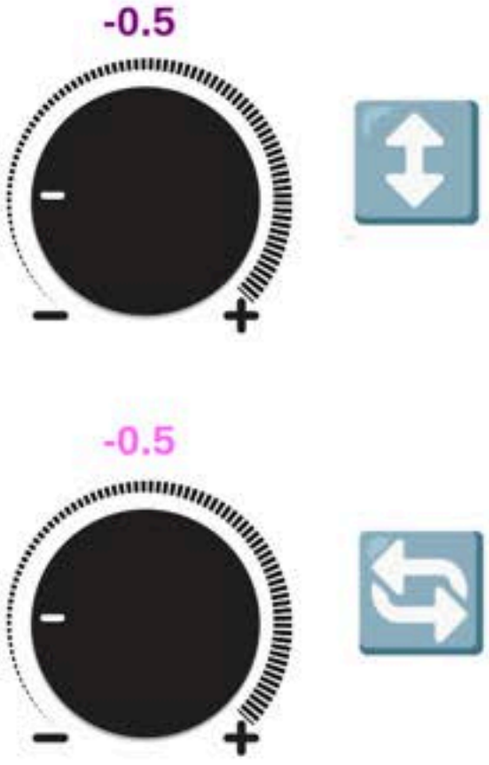

y = −0.7 x + 0.5
y
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
x

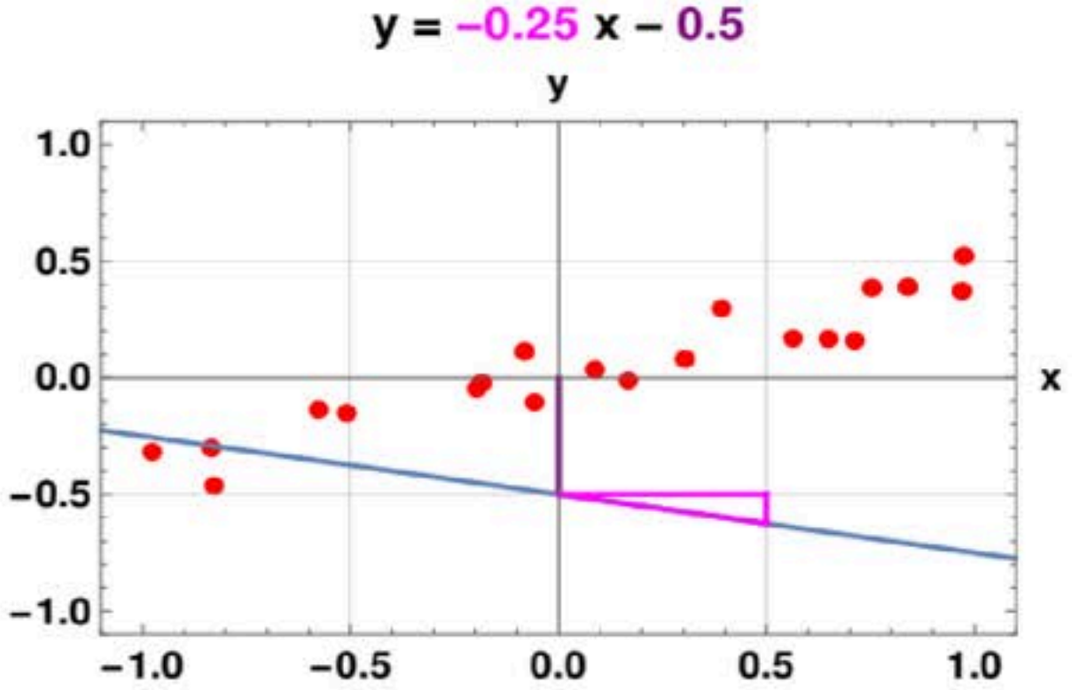

0.5
-0.7

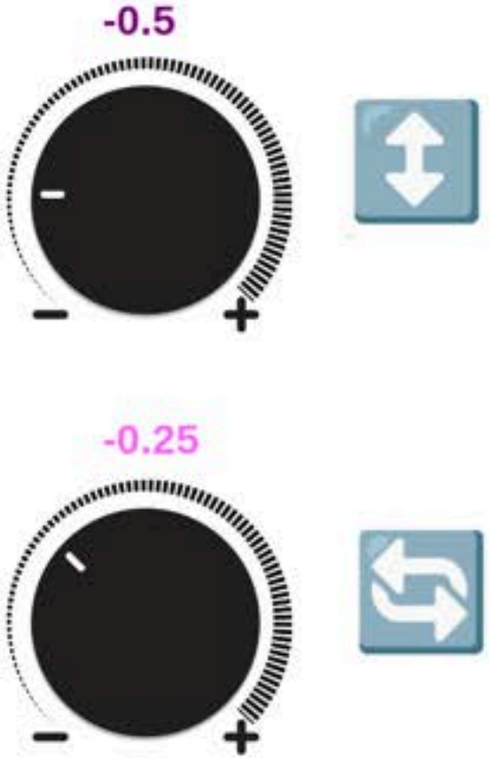

y = −0.75 x − 0.5
y
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
x

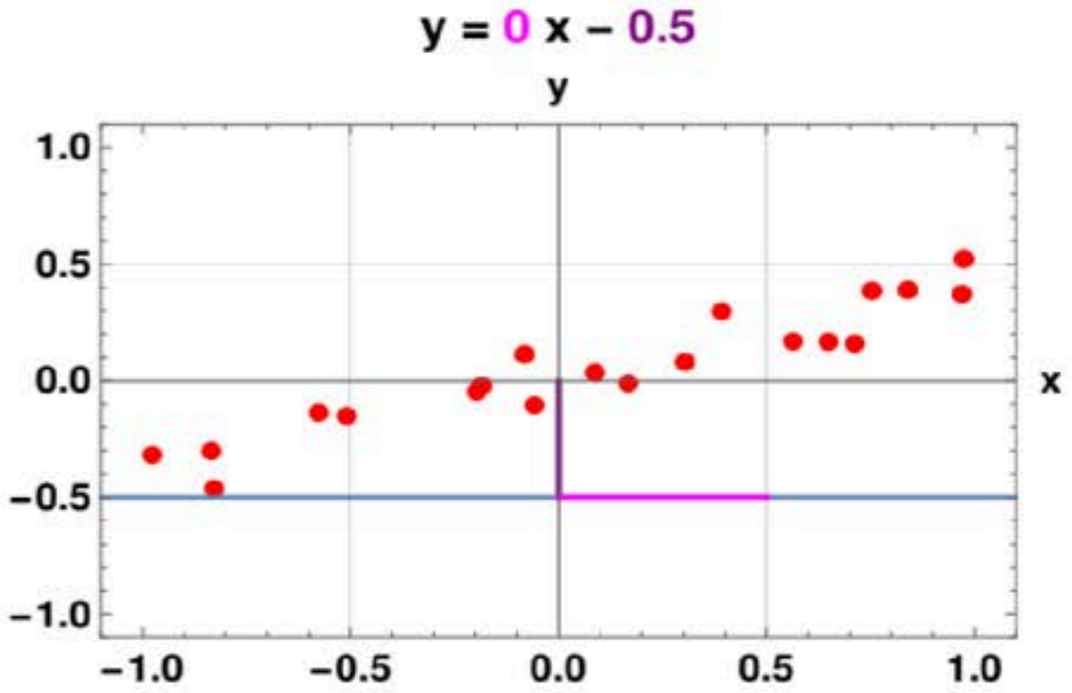

-0.5
-0.75

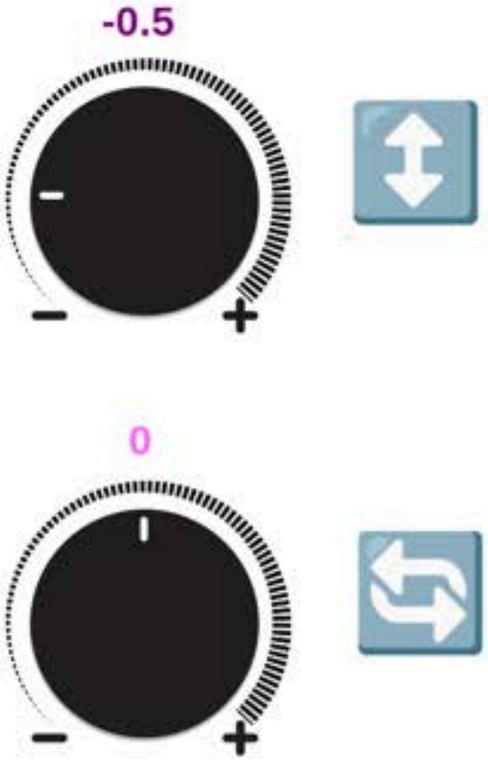

y = −0.5 x − 0.5
y
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
x

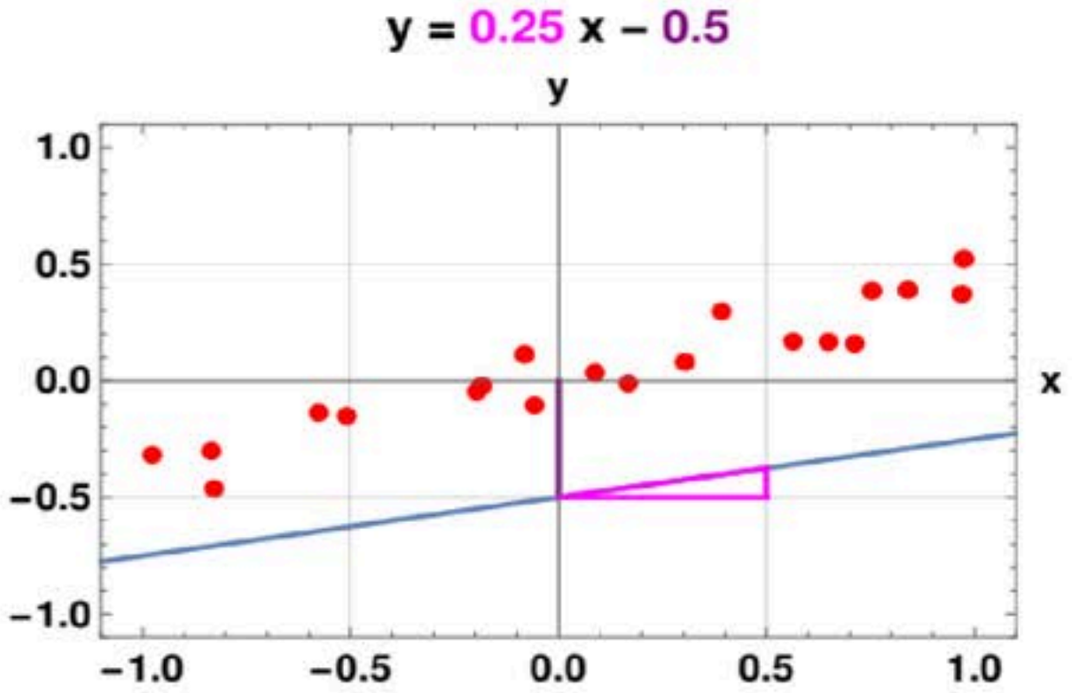

-0.5
-0.5


y = −0.25 x − 0.5
y
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
x


-0.5
-0.25

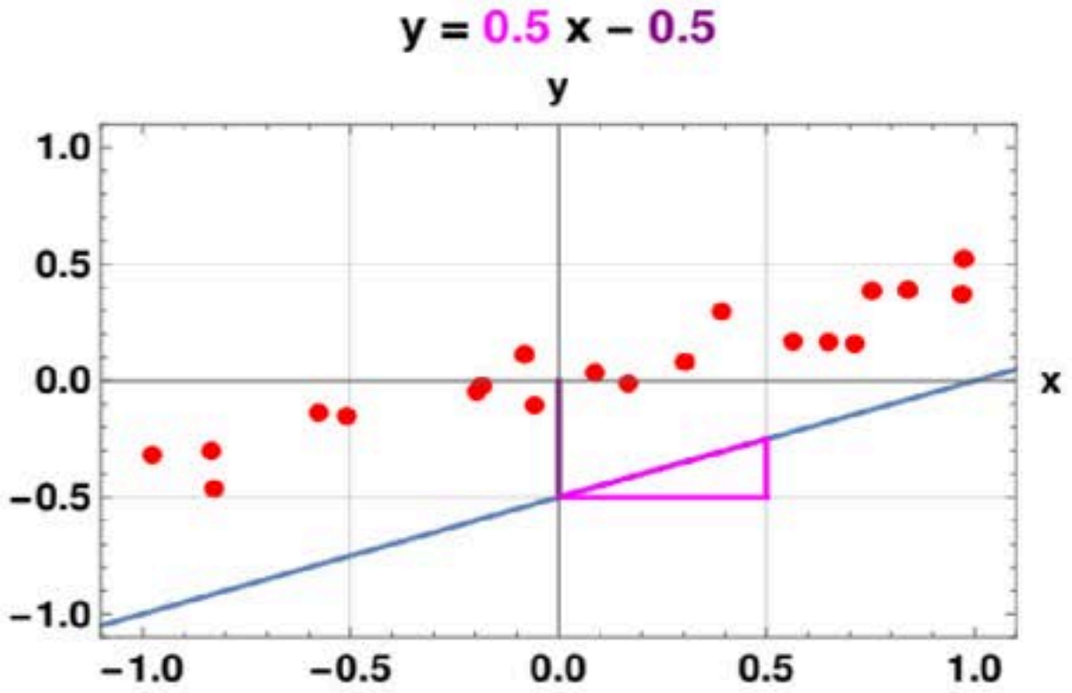

y = 0 x – 0.5
y
1.0
0.5
0.0
–0.5
–1.0
–1.0
–0.5
0.0
0.5
1.0
x

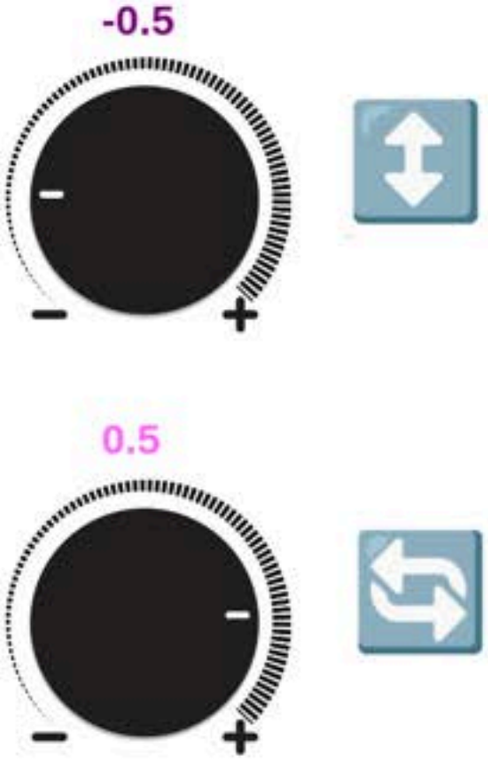

-0.5
0

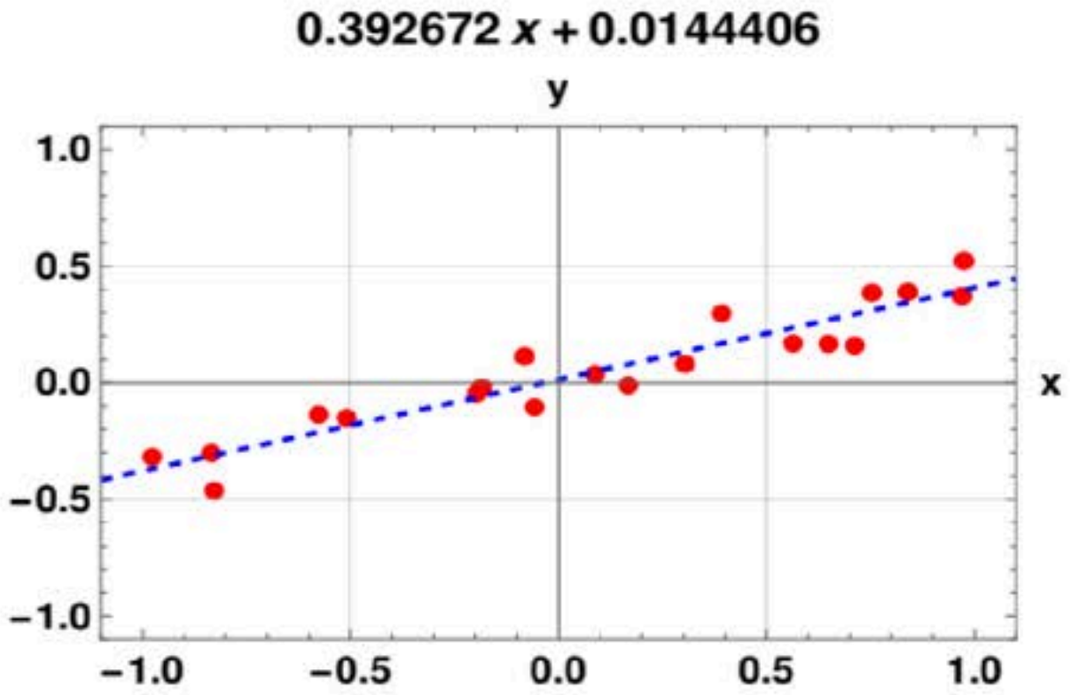

y = 0.25 x – 0.5
y
1.0
0.5
0.0
–0.5
–1.0
–1.0
–0.5
0.0
0.5
1.0
x

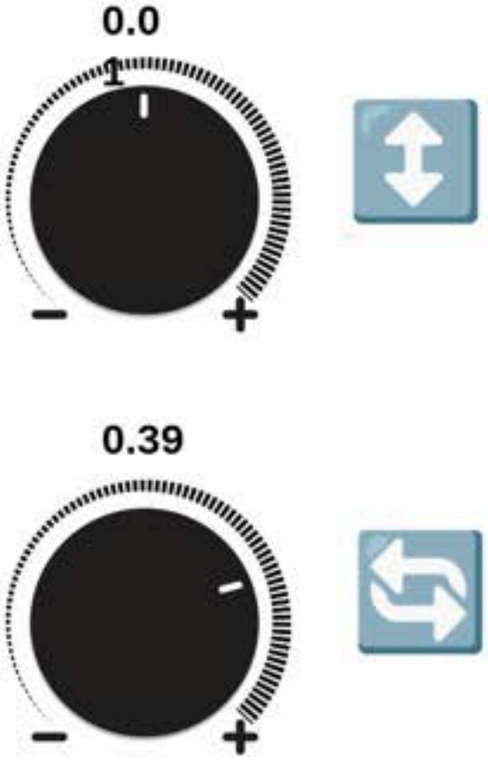

-0.5


0.25

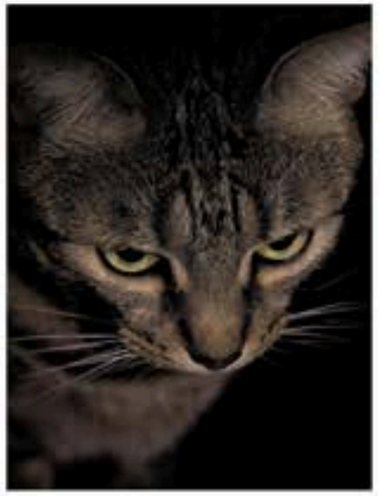

y = 0.5 x – 0.5
y
1.0
0.5
0.0
–0.5
–1.0
–1.0
–0.5
0.0
0.5
1.0
x

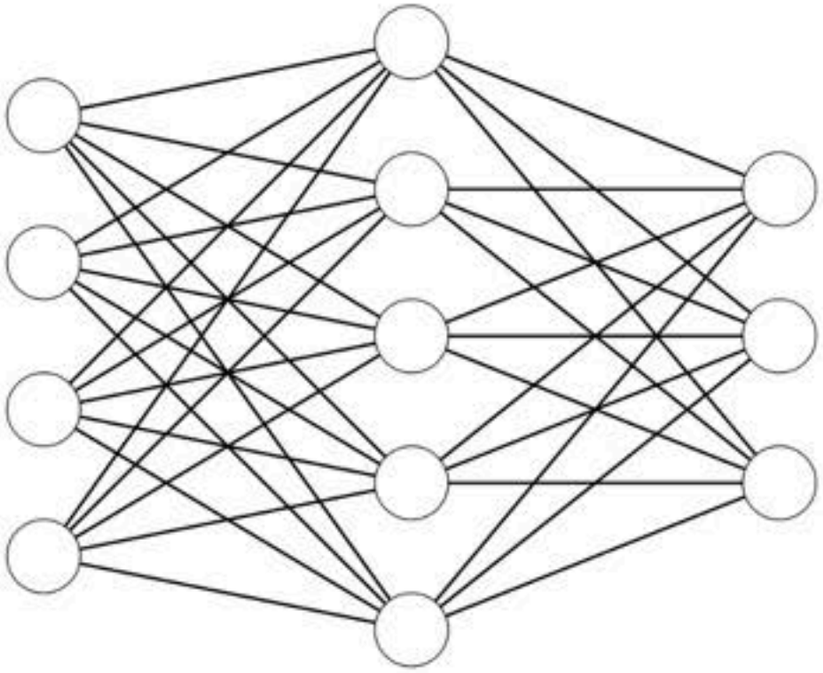

-0.5
0.5


0.392672 x + 0.0144406
y
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
x

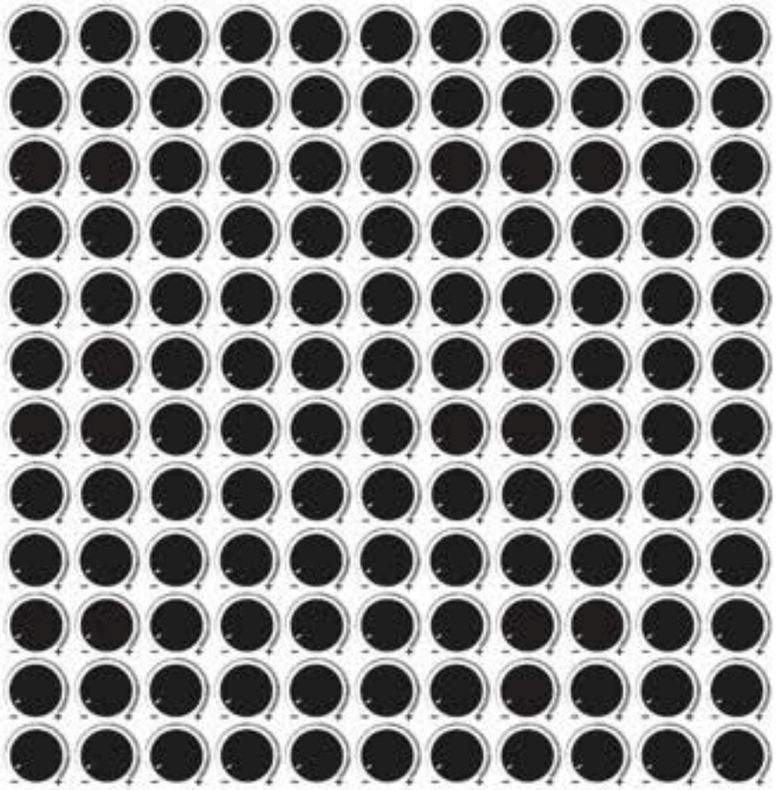

0.0
1
0.39

Concept: Training an artificial neural network (a universal function approximator)





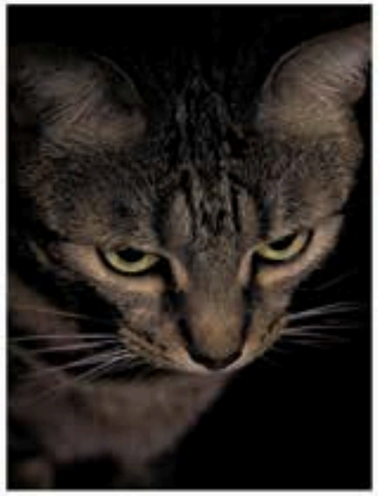
$$\begin{bmatrix} 0.9 \text{ cat} \\ 0.05 \text{ dog} \\ 0.05 \text{ fox} \end{bmatrix}$$

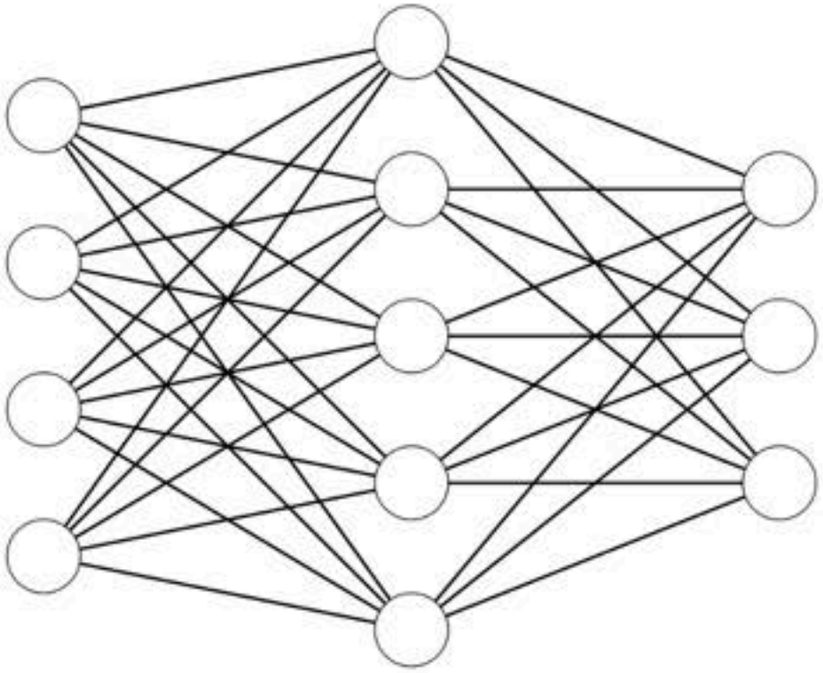


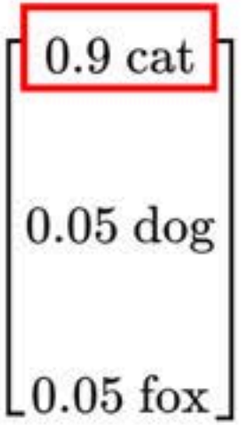
Traning a neural network:

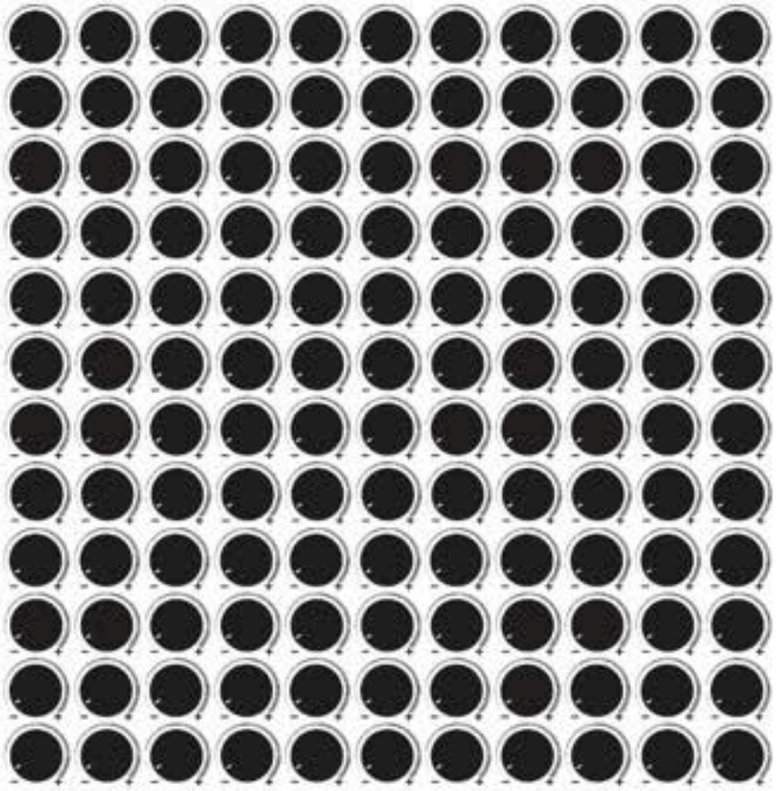
- provide training input
- forward pass
- get prediction
- compute loss function (compare)
- compute gradient, update parameters

Concept: Training an artificial neural network (a universal function approximator)









Number of parameters from 10 thousands to millions

Traning a neural network:

- provide training input
- forward pass
- get prediction
- compute loss function (compare)
- compute gradient, update parameters

Detect

$$\|x - \mu\| > \tau$$

- Anomaly detection
- Monitoring
- Diagnosis
- Event detection

Computer Vision

$$f : \mathbb{R}^{H \times W \times C} \rightarrow \mathcal{Y}$$

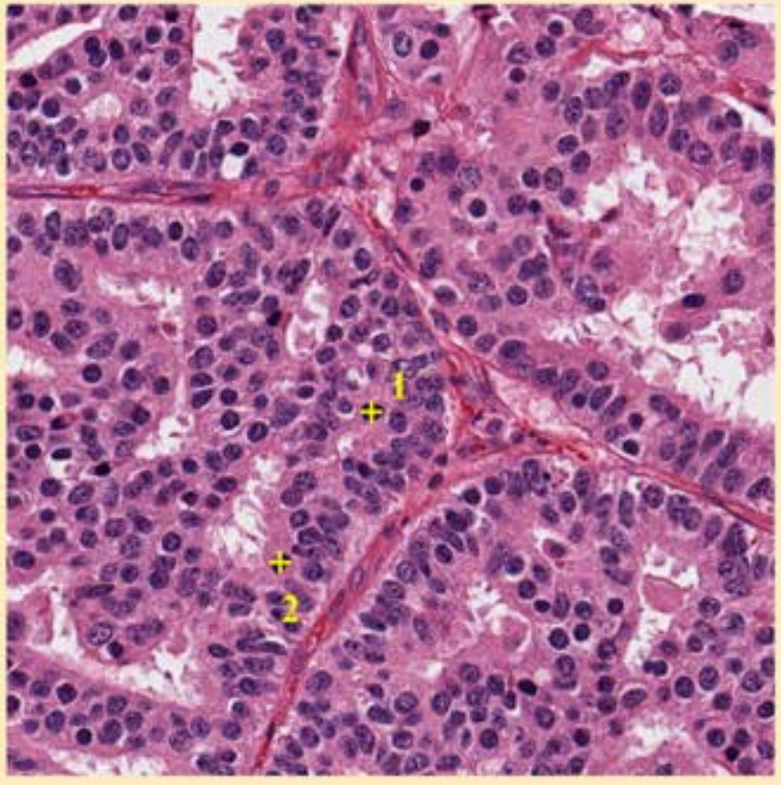
- Quality inspection
- Defect detection
- Robotic perception
- Surveillance
- Diagnosis

- CNNs
- Transformers
- Foundation vision models

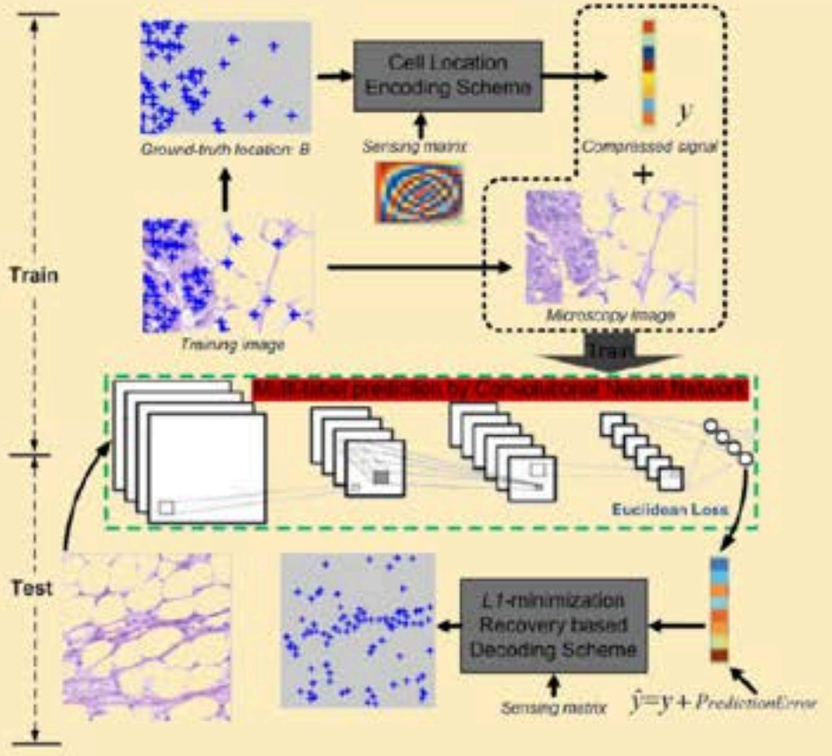
Training a neural network for cell detection

JOANNEUM RESEARCH DIGITAL

[1708.03307cs] Cell Microscopy



[1708.03307cs] Encoding scheme

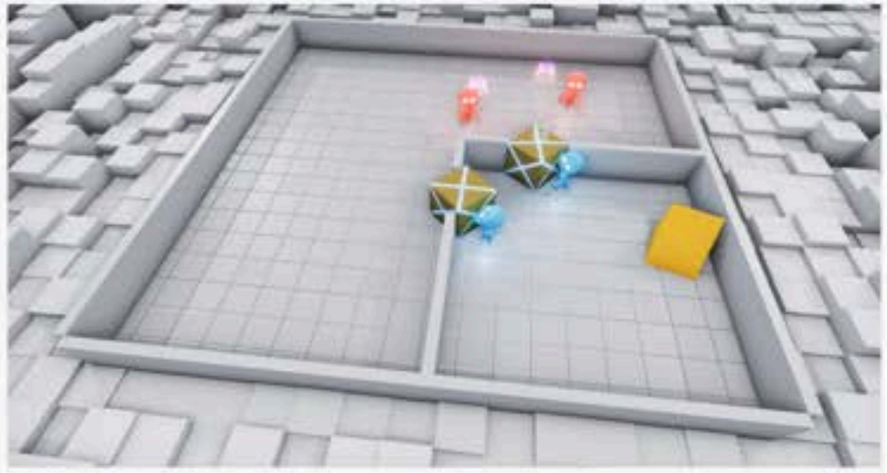


Optimization

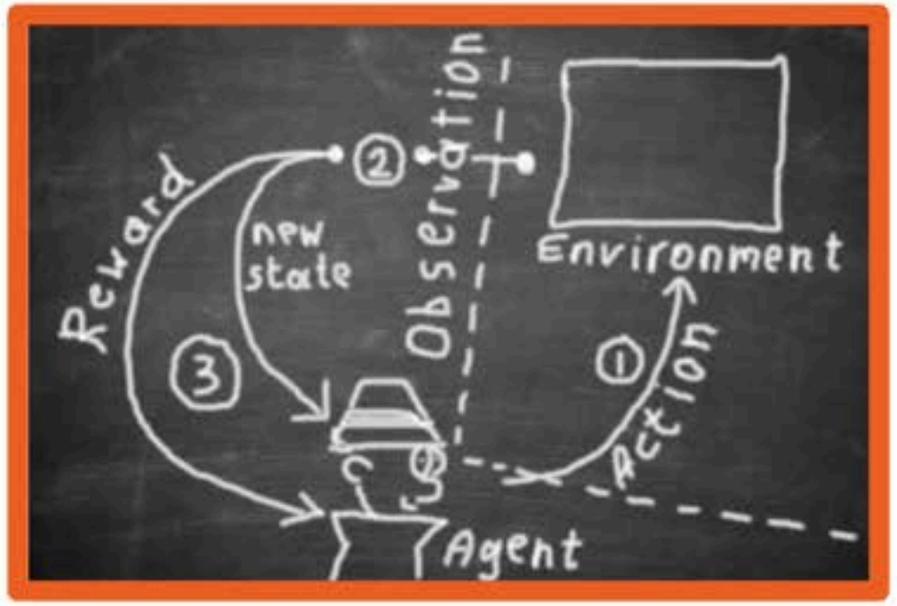
$$x^* = \arg\min_x f(x)$$

- Process optimization
- Resources
- Routes
- Scheduling
- Supply chain
- Control
- Throughput

The Reinforcement Learning Framework



Baker et al., 'Emergent Tool Use From Multi-Agent Autocurricula'
arXiv:1909.07528, 2019



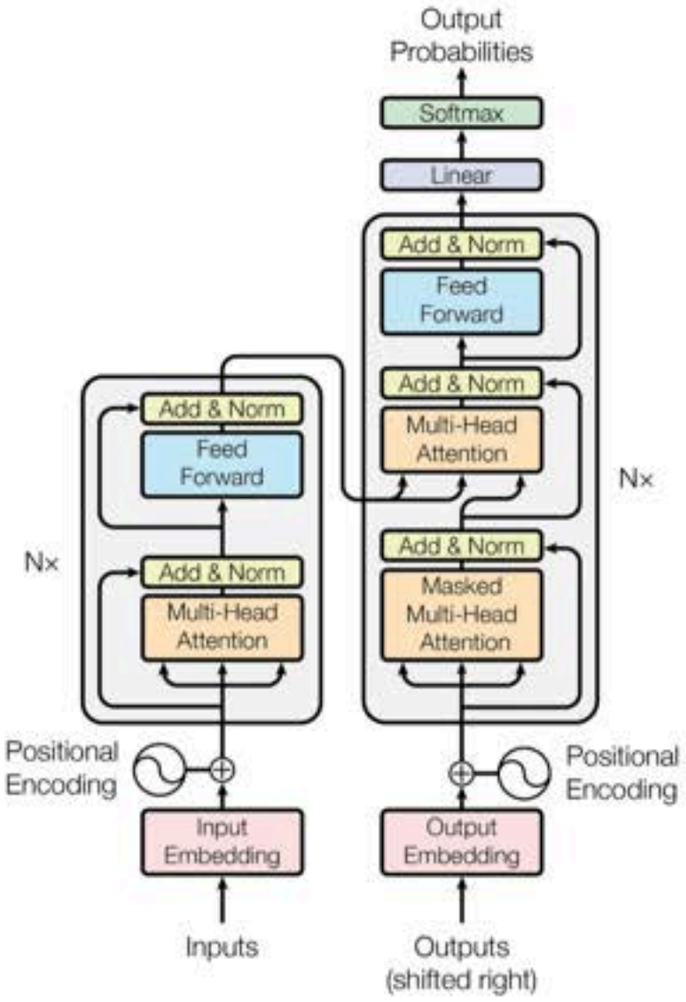
1. Take an action
2. Observe changes
3. Receive reward
4. Maximize reward

Generate

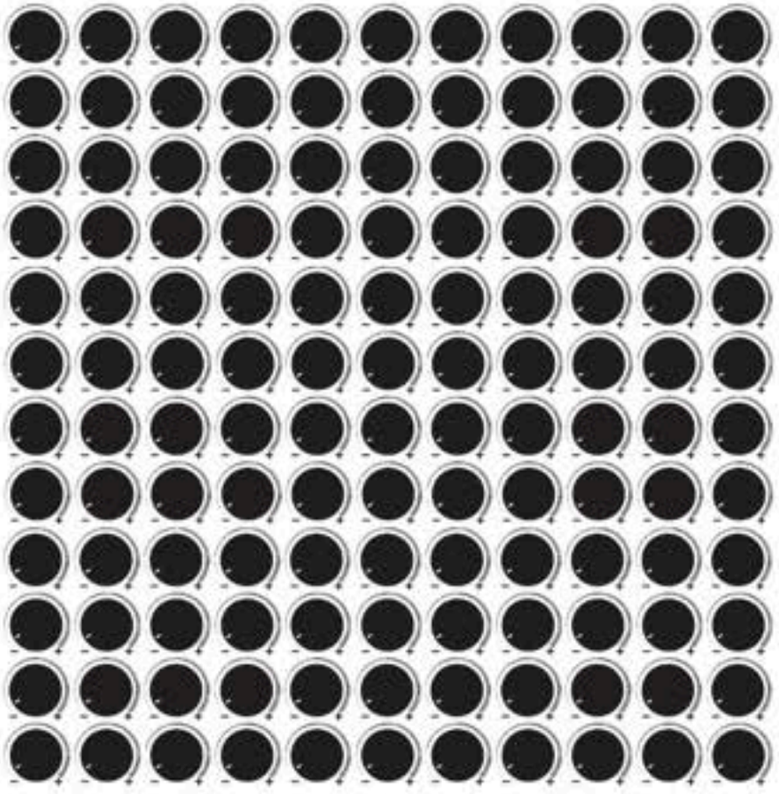
$$x \sim p_\theta(x \mid c)$$

- CAD / model generation
- Code generation
- Text / image / video / audio
- Design exploration
- Workflow genration
- Generative design

The modern era: Transformer

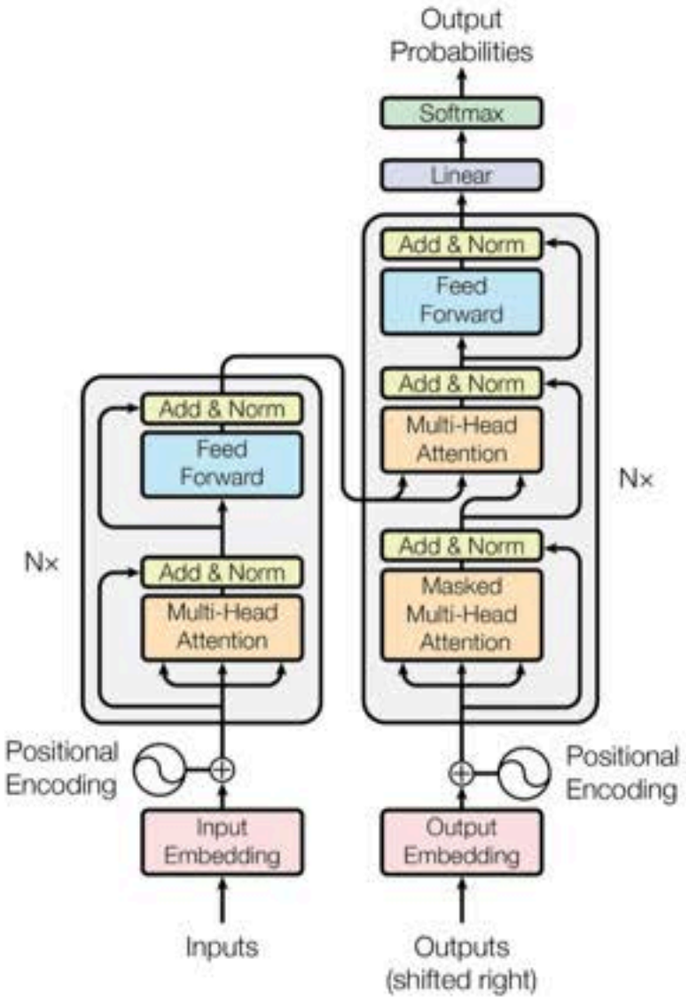


Source: „Attention is all you need“ arXiv:1706.03762



Number of parameters in hundreds of billions to trillions

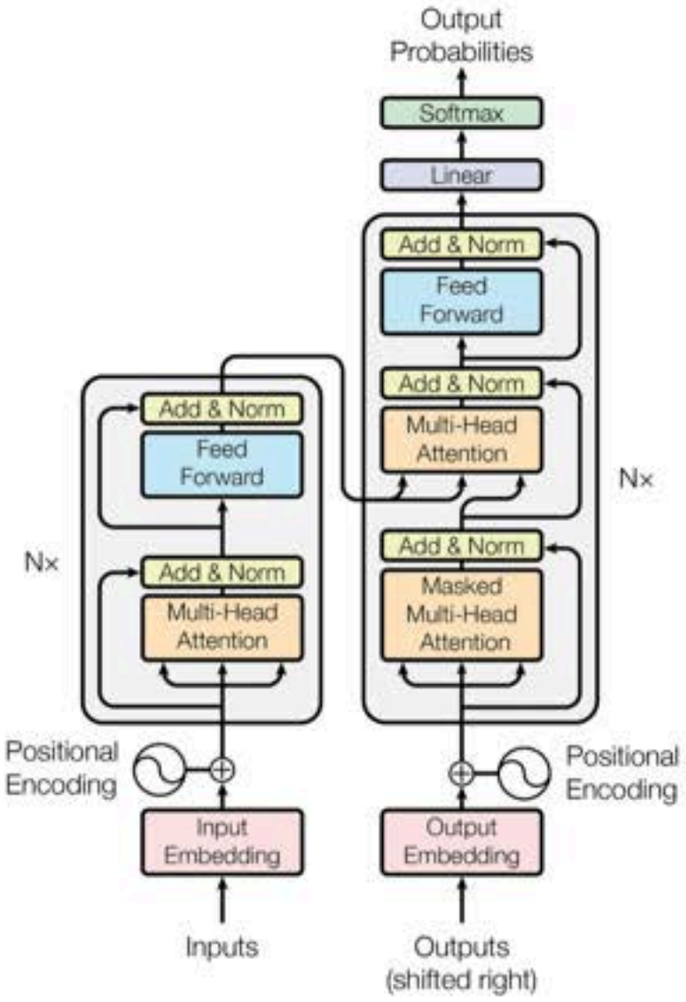
The modern era: Transformer



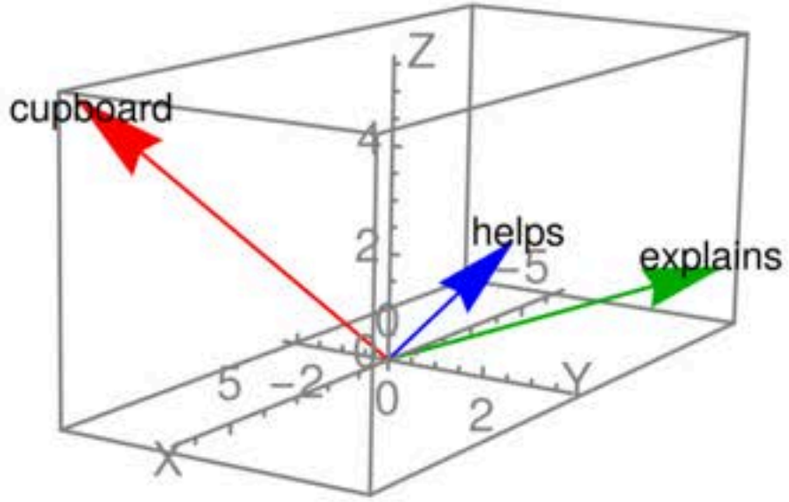
Source: „Attention is all you need“ arXiv:1706.03762

Create an image with huge piles of books. There should be a lectern, on which an open book rests. The floor is strewn with papers that have text printed on them, a candle illuminates the room. The scene has an atmosphere of warmth.

The modern era: Transformer

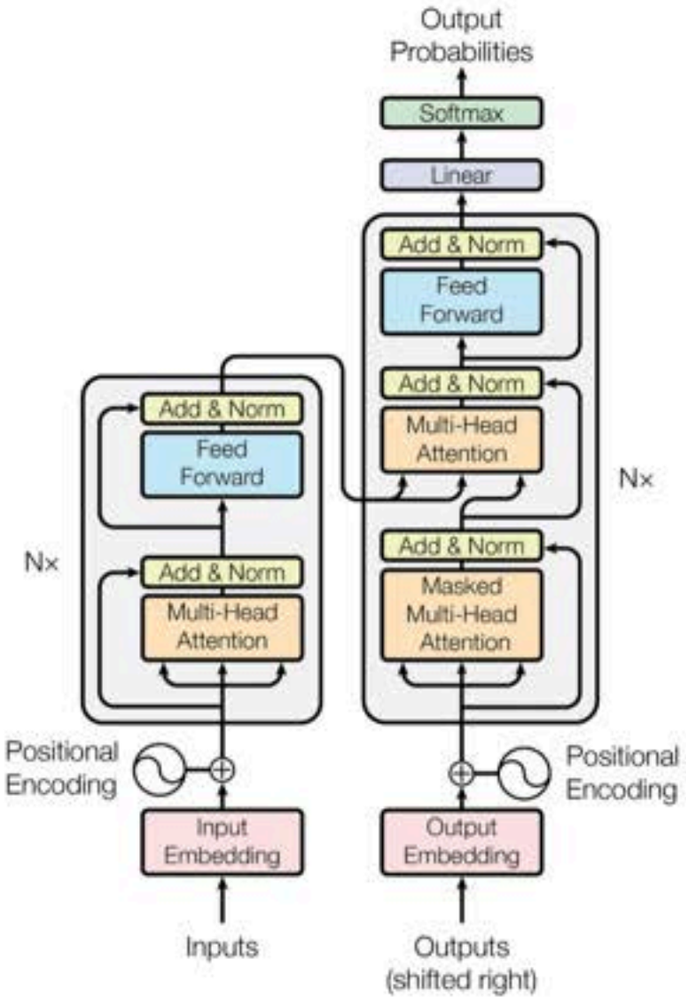


Imagine reading
a long story.
A Transformer is like
a clever helper
who ...



Source: „Attention is all you need“ arXiv:1706.03762

The modern era: Transformer



Source: „Attention is all you need“ arXiv:1706.03762

Imagine reading

a big story.

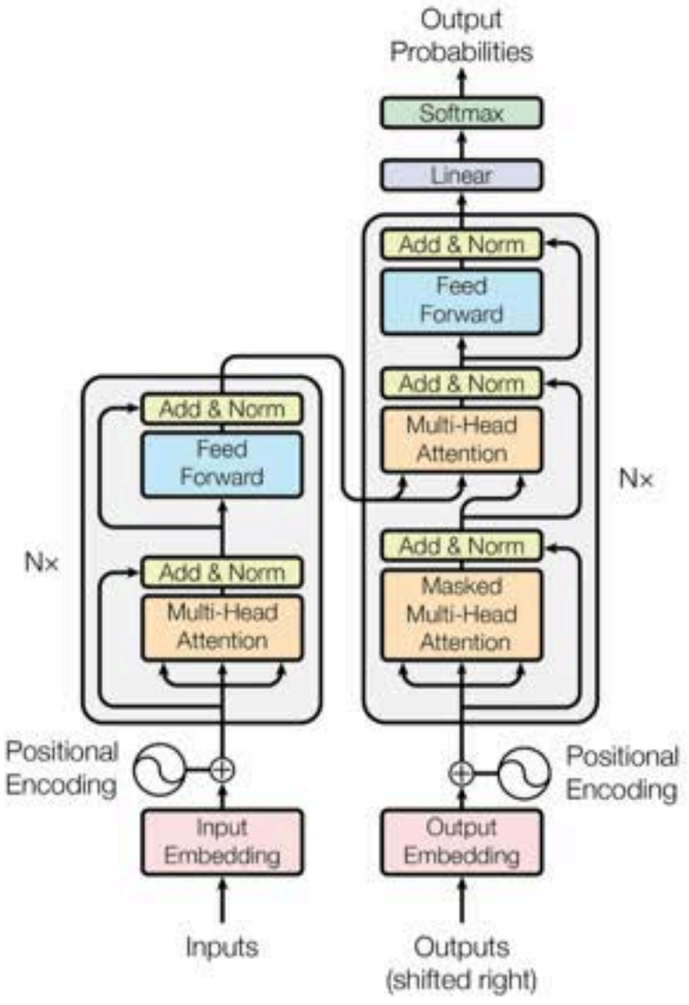
A Transformer is like

a clever helper

who ...

0.30	helps
0.20	listens
0.12	understands
0.10	learns
0.07	remembers
0.06	thinks
0.05	explains
0.04	guides
0.03	highlights
0.02	reads
0.01	talks

The modern era: Transformer

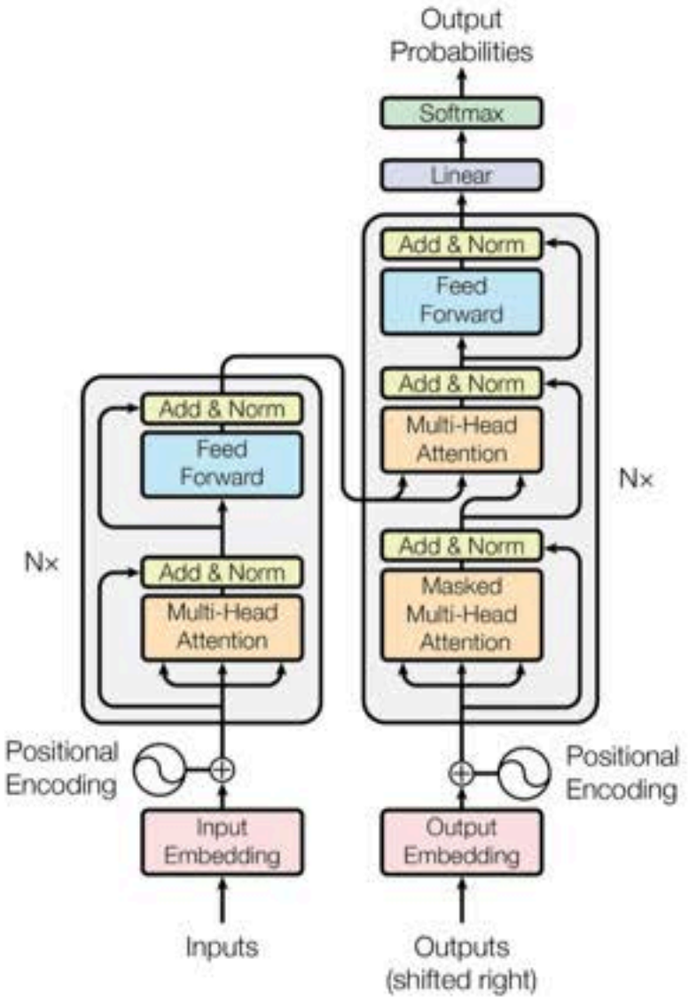


Source: „Attention is all you need“ arXiv:1706.03762

Imagine reading
a long story.
A Transformer is like
a clever helper
who helps ...

0.40	you
0.15	to
0.12	understand
0.10	organize
0.07	find
0.05	connect
0.04	grasp
0.03	identify
0.02	distinguish
0.01	focus
0.01	highlight

The modern era: Transformer



Source: „Attention is all you need“ arXiv:1706.03762

Imagine reading
a long story.
A Transformer is like
a clever helper
who helps you see
which words matter most
and how they fit together,
making the whole story
easy to understand!

Orchestrate

$$a_t = \pi(s_t, m_t, g_t)$$

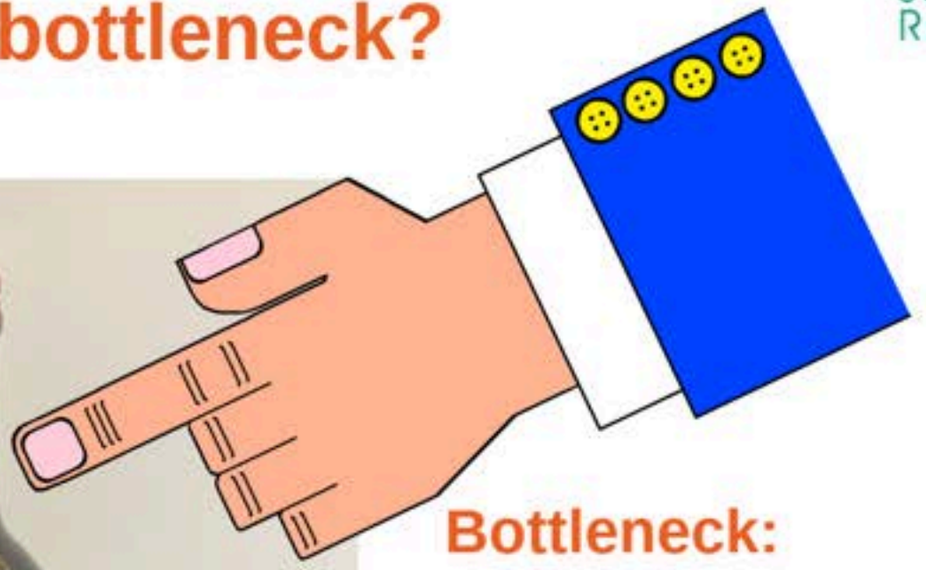
- Tool integration
- Multi agent coordination
- Workflow automation
- Task planning
- Human AI collaboration
- Autonomous operations

Why now? The Status Quo

- Compute: GPUs and HPC
- Data: IoT, sensor networks
- Algorithms: Transformer,...
- Interfaces: Natural language
- Systems convergence:
- AI + {HPC, simulation, robotics…}
- Business / Industry pressure
- Strategic Importance

Where is the bottleneck?

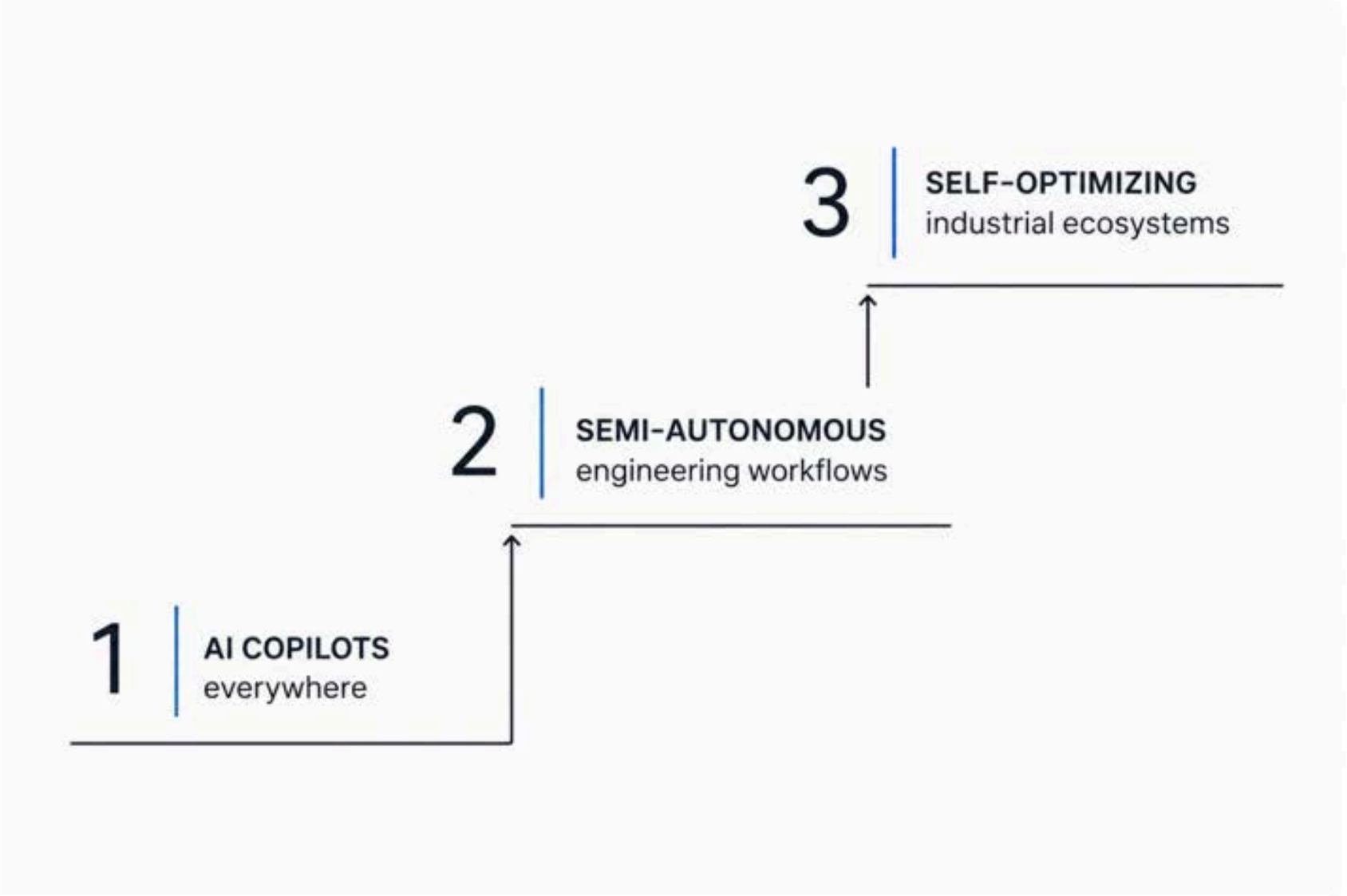
Where is the bottleneck?



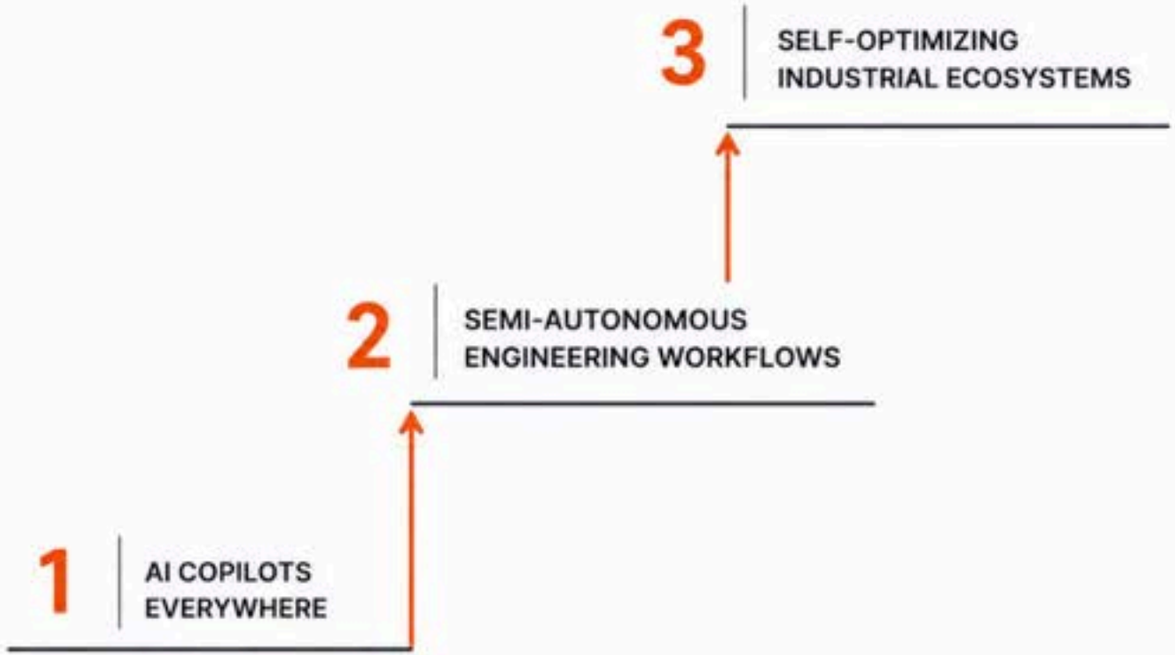
Bottleneck:

- poor / wrong data
- fragmented infrastructure
- unclear ROI
- governance
- lack of integration
- cybersecurity
- safety
- organizational challenges

Potential development



JOANNEUM RESEARCH DIGITAL



Engineering

- AI-native engineering environments
- Simulation-first design
- Autonomous troubleshooting
- Digital twins linked to reality

Robotics

- Embodied AI
- Adaptive robotics
- Reinforcement learning in physical systems

Science & R&D

- Autonomous experimentation
- AI-assisted discovery
- Generative simulation systems

Organizations

- AI operating layers
- AI-mediated decision systems
- Agentic enterprises

The industrial revolutions mechanized muscle.

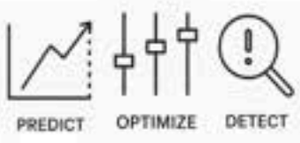
The digital revolution mechanized information.

AI is beginning to mechanize parts of cognition itself.

What all AI use cases have in common:

*Across very different applications, the **pattern** is the same:*

- AI creates value by predicting, optimizing, or detecting
- Energy efficiency improves as a side effect of better decisions
- Benefits often come from reducing waste (time, material, scrap, idle states)
- Improvements do not require new hardware: only data and process access
- AI works best when applied to a specific, well-defined use case



Energy efficiency emerges naturally when processes are optimized through AI

How SMEs can start

- Identify **one clear source** of resource waste
(e.g., energy peaks, idle phases, rework, poor quality, etc)

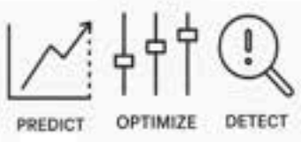
- Start with a **small, contained use case**
(e.g., one machine, one process, one production step, etc)

- Leverage **existing data**
(e.g., machine states, sensor signals, cycle times, energy meters, etc)

- Set a simple, **measurable target**
(e.g., reduce waste, shorten cycles, lower energy per part, etc)

- Run a **short pilot**
(e.g., 8-12 weeks, minimal risk, fast insight)

- **Scale** step-by-step once **value is demonstrated**



AI delivers the best impact when it reduces waste in a targeted and measurable way.

PREDICT OPTIMIZE DETECT

JOANNEUM RESEARCH DIGITAL

Key Takeaways

- AI improves predictability, stability and process quality
- Energy efficiency emerges naturally by reducing waste
- Many benefits can be achieved without new hardware
- The fastest value comes from small, focused use cases
- A short pilot (8-12 weeks) is usually enough to show measurable impact

Q1 — Choose a Company Use Case

⏱ **10 minutes**

Elect: ONE company in your group to act as the use case provider.

Introduce: What does the company do? What are the main products or services?

Discuss: What currently feels inefficient, repetitive, expensive, or frustrating?

Q2 — Identify One AI-Relevant Pain Point

⏱️ **15 minutes**

Choose: ONE concrete pain point where AI could realistically create value in the next 6–12 months.

Look for: repetitive work, communication bottlenecks, knowledge-heavy work, or manual administration.

Be: specific, realistic, and concrete.

Q3 — Design a Realistic AI Pilot

⏱️ **20 minutes**

Design: ONE realistic AI-supported solution.

Clarify: What exactly should AI do?

Discuss: Which tools could help, where humans stay involved, and what success would look like.

Q4 — Implementation Reality Check

⏱ **15 minutes**

Create: a realistic implementation plan.

Define: timeline, resources, ownership, and required skills.

Identify: missing competencies, barriers, and the first concrete next step.

Thank you for your attention!

Andreas Windisch
Head of Research Group Intelligent Vision Applications

JOANNEUM RESEARCH Forschungsgesellschaft mbH
DIGITAL - Institut for Digital Technologies

Steyrergasse 17
8010 Graz

Tel. +43 316 876-5208
andreas.windisch@joanneum.at

www.joanneum.at/digital

JOANNEUM RESEARCH DIGITAL

aiaustria

www.joanneum.at/digital