Three Phases of Artificial Intelligence

Sepp Hochreiter

Artificial Intelligence

Artificial Intelligence (AI) will contribute to overcome crucial challenges concerning

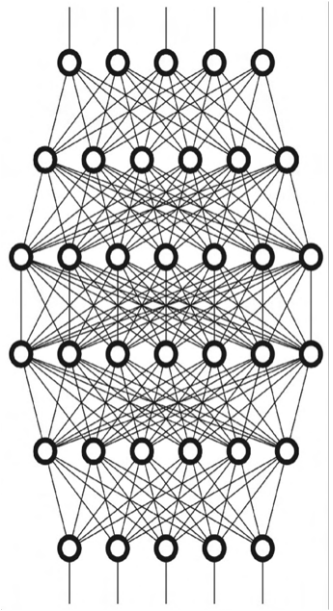
- **Energy**
- **Climate**
- **Food**
- **Healthcare**
- **Mobility**

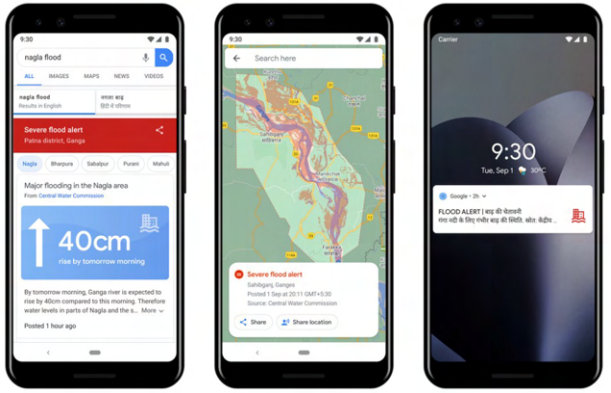
Deep Learning

Key technology of artificial intelligence.

AI became the fastest growing field of computer science via Deep Learning.

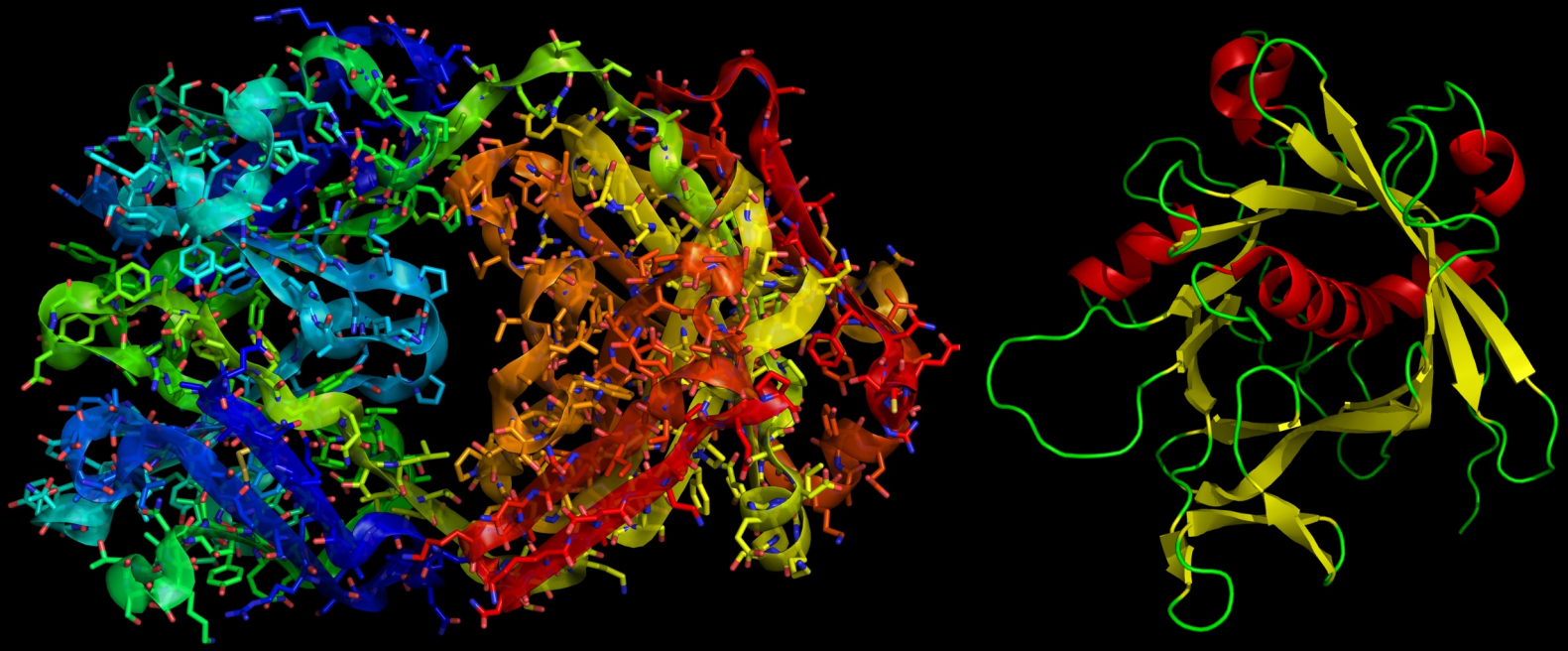
Many **success stories** of Deep Learning.



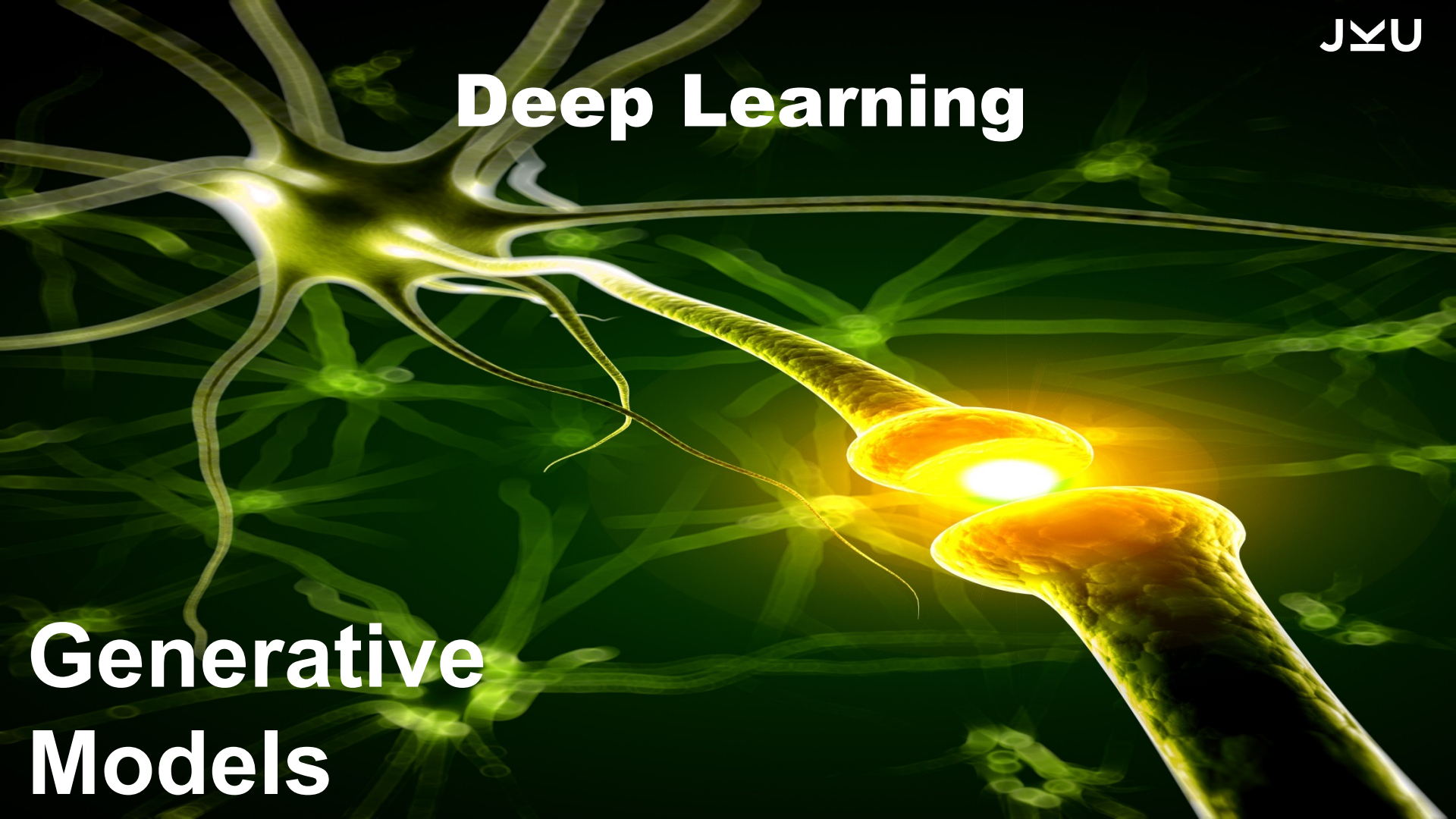


LSTMs predict floods and droughts

AlphaFold (Nobel Prize)



AI Breakthrough in Biology


JKU
Deep Learning
Generative
Models

CHAT GPT

ChatGPT Failures

HO who is sepp hochreiter

Sepp Hochreiter is an Austrian computer scientist and mathematician who is best known for his work on artificial neural networks, particularly the development of the long short-term memory (LSTM) algorithm. He was born on June 4, 1967, in Schärding, Austria.

Hochreiter received his Ph.D. from the Technical University of Munich in 1991 and has since held various research positions at institutions including the University of Toronto and the University of California, Berkeley. In 1997, he and Jürgen Schmidhuber, another prominent researcher in the field of artificial intelligence, introduced the LSTM algorithm, which has become a fundamental component of many modern neural network architectures.

Hochreiter has received numerous awards for his contributions to artificial intelligence, including the IEEE Frank Rosenblatt Award in 2016 and the ACM SIGKDD Test of Time Award in 2019. He currently works as a professor of bioinformatics at Johannes Kepler University in Linz, Austria, where he continues to conduct research in machine learning and artificial intelligence.

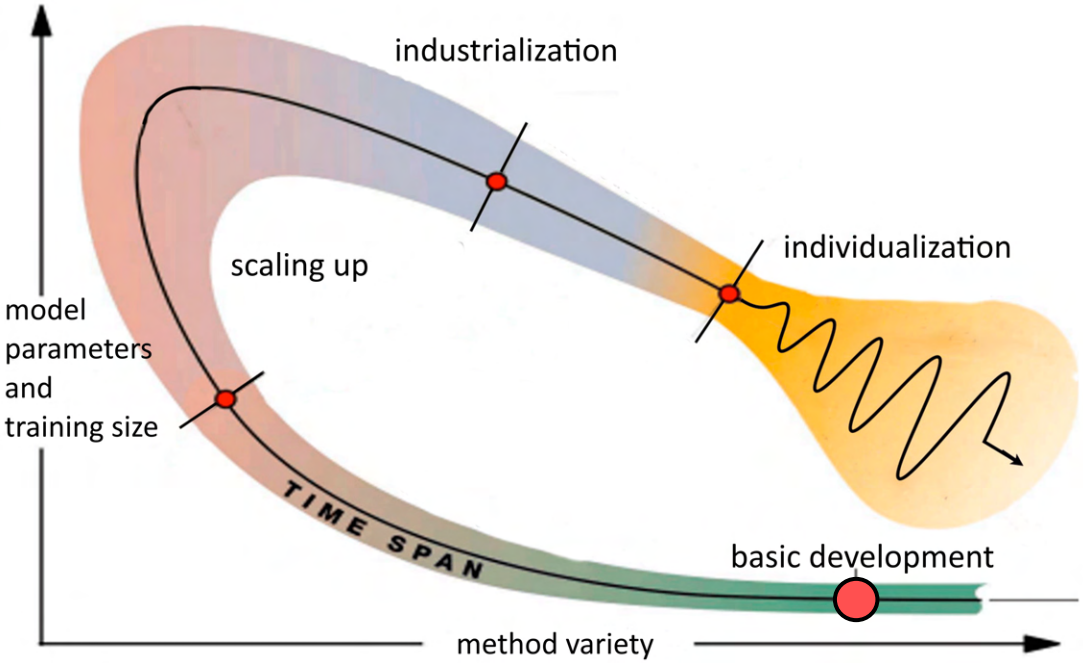
Three Phases of AI

1. Basic Development

2. Scaling Up

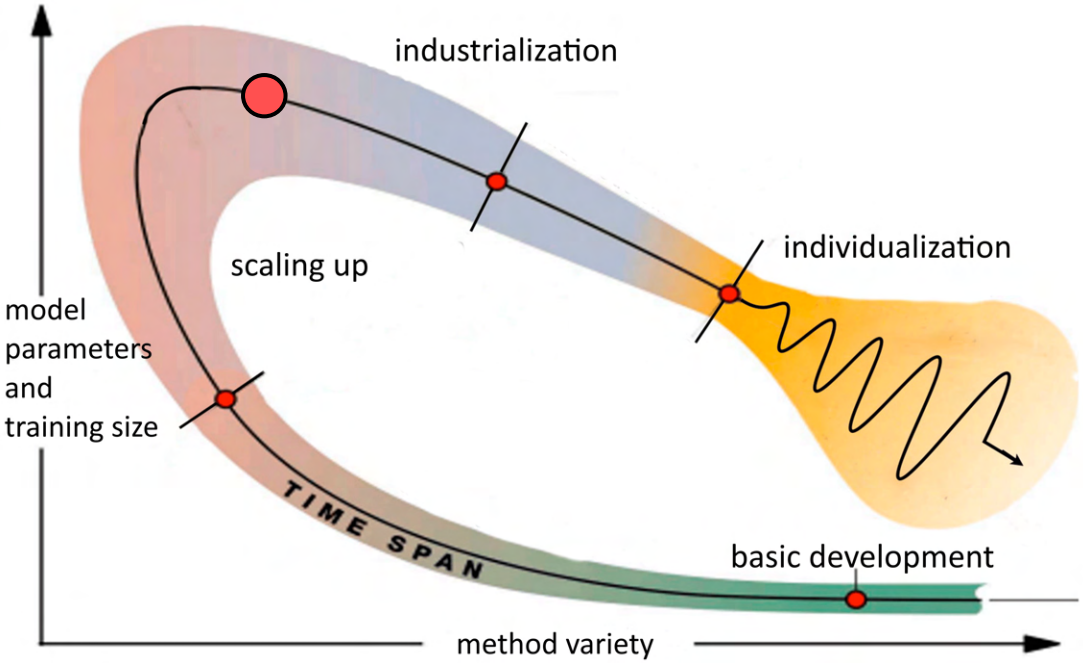
3. Industrialization

Three Phases of AI



According to Gu X, Koren Y. Smart Factories for Mass Individualization. *Encyclopedia*. 2024; 4(1):415-429.

Three Phases of AI

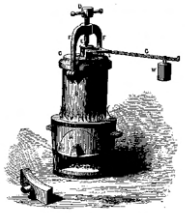


According to Gu X, Koren Y. Smart Factories for Mass Individualization. *Encyclopedia*. 2024; 4(1):415-429.

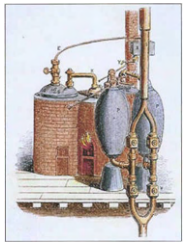
Steam Engine

Basic development

1st century: Hero's engine.

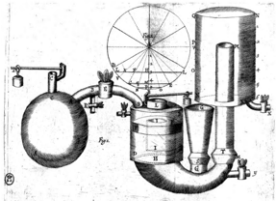


1679: Papin's steam digester.

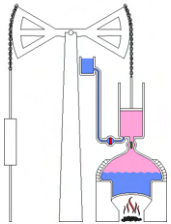


1698: Savery's Engine.

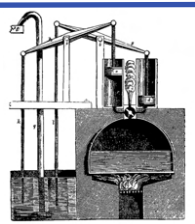
Scaling up



1707: Papin's steam engine.

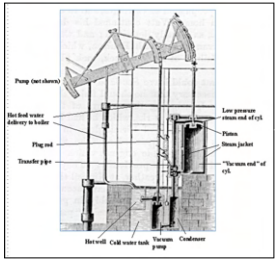


1712: T. Newcomen atmospheric engine.



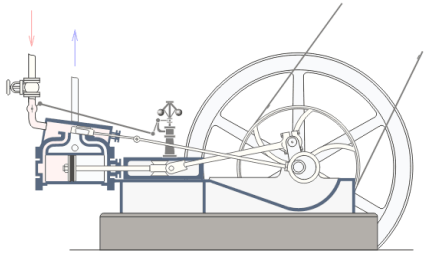
1720: J. Leopold's steam engine.

Industrialization



1878: Watt's steam engine

1788: Boulton & Watt's Lap Engine.

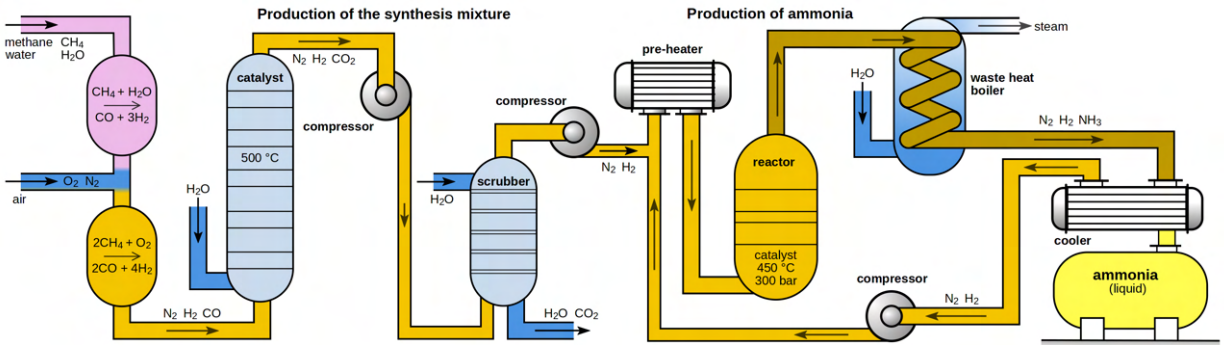


19th century: double acting stationary engine.

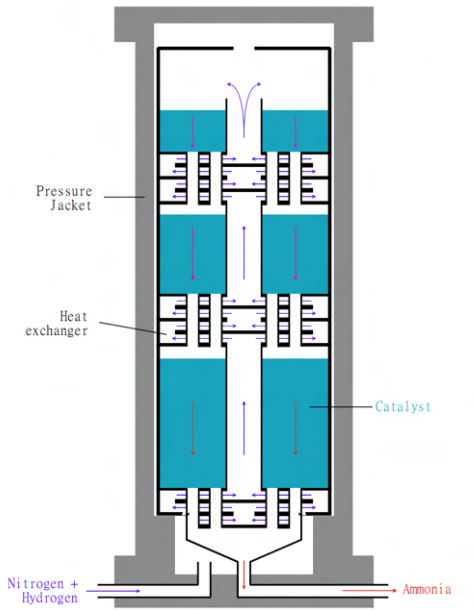
Haber-Bosch Process

The Haber–Bosch process to produce ammonium nitrate for fertilizer led to

- population boom
- concentration in cities
- “nearly 50% of the nitrogen found in human tissues originated from the Haber-Bosch Process” (Solomon, P. M. – 2004).



https://commons.wikimedia.org/wiki/File:Haber-Bosch-En.svg



https://commons.wikimedia.org/wiki/File:Ammoniakreaktor_MS.svg

Basic Research: 1784 Berthollet to 1901 Le Chatelier

Scaling up: 1909 Haber at BASF (Nobel Prize 1918)

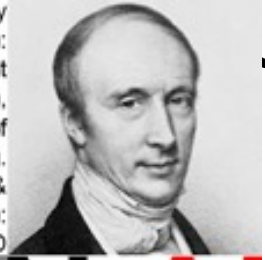
Industrialization: 1910 Carl Bosch (Nobel Prize 1931)

Leibniz (1676): chain rule for backward credit assignment, central ingredient of deep learning

Legendre (1805) and Gauss (1795, unpublished): first linear neural networks (NNs) / linear regression / method of least squares / shallow learning

Famous example of pattern recognition through shallow learning from astronomical data: re-discovery of dwarf planet Ceres (Gauss, 1801)

Cauchy (1847): gradient descent (GD), basic tool of deep learning. Robbins & Monro (1952): Stochastic GD



THE ROAD TO MODERN AI

Ising (1925): 1st recurrent network architecture: Lenz-Ising model (see also McCulloch & Pitts, 1943, Kleene, 1956)

Turing (1948): unpublished ideas related to evolving recurrent NNs (RNNs)

Rosenblatt (1958): multilayer perceptron (MLP) (only last layer learned: no deep learning yet) See also Steinbuch (1961) Joseph (1961)

Kelley (1960): precursor of backprop in control theory (compare Bryson,'61; Dreyfus,'62)

ARTIFICIAL NEURAL NETWORKS UP TO 1979

FROM SHALLOW LEARNING CIRCA 1800 TO DEEP LEARNING

Ivakhnenko & Lapa (1965): first deep learning in deep MLPs that learn internal representations of input data

Amari (1967-68): deep learning by stochastic gradient descent for deep MLPs 1972: 1st published learning RNN based on Ising model (1925)

Linnainmaa (1970): backpropagation or reverse mode of automatic differentiation

First applied to NNs by Werbos (1982)

Fukushima (1979): deep convolutional neural net architecture 1969: rectified linear units. Both now widely used

JKU

AI/Deep Learning

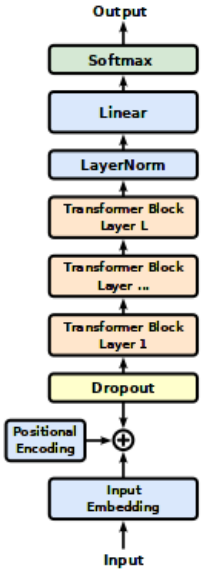
Basic Develop-ment

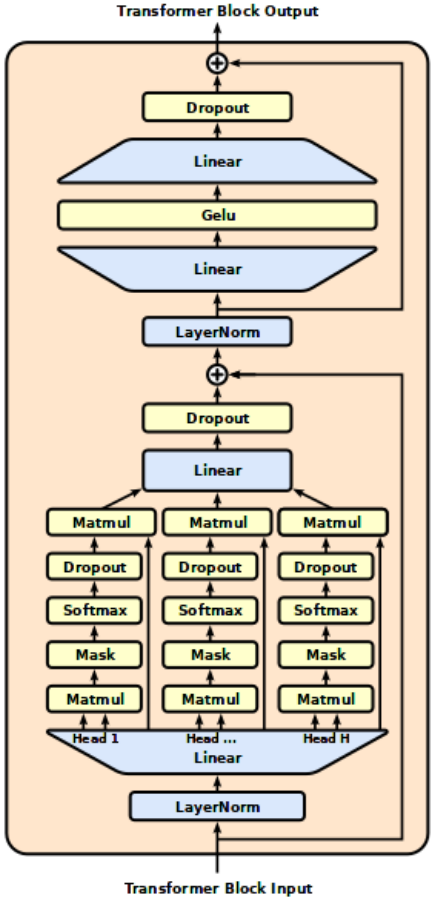
AI Scaling Up: Transformer

Transformer:

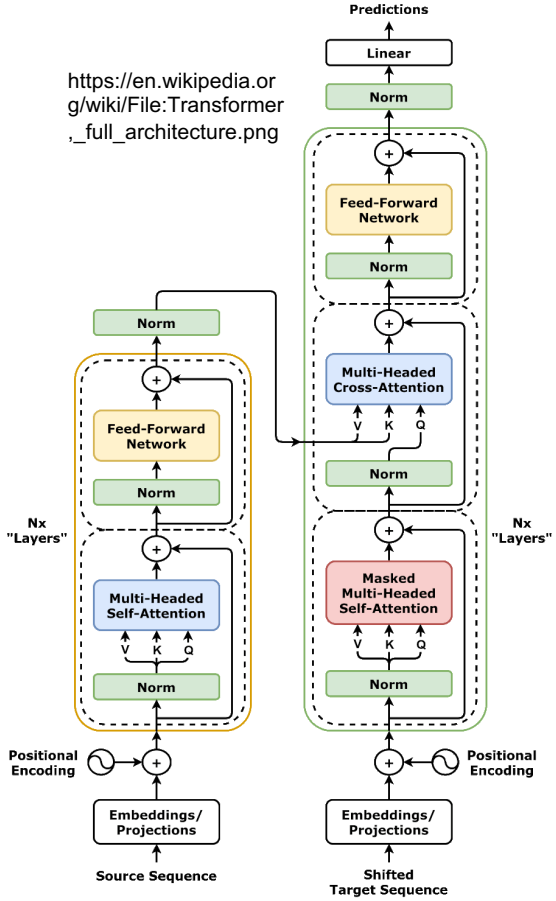
ResNet and

Attention

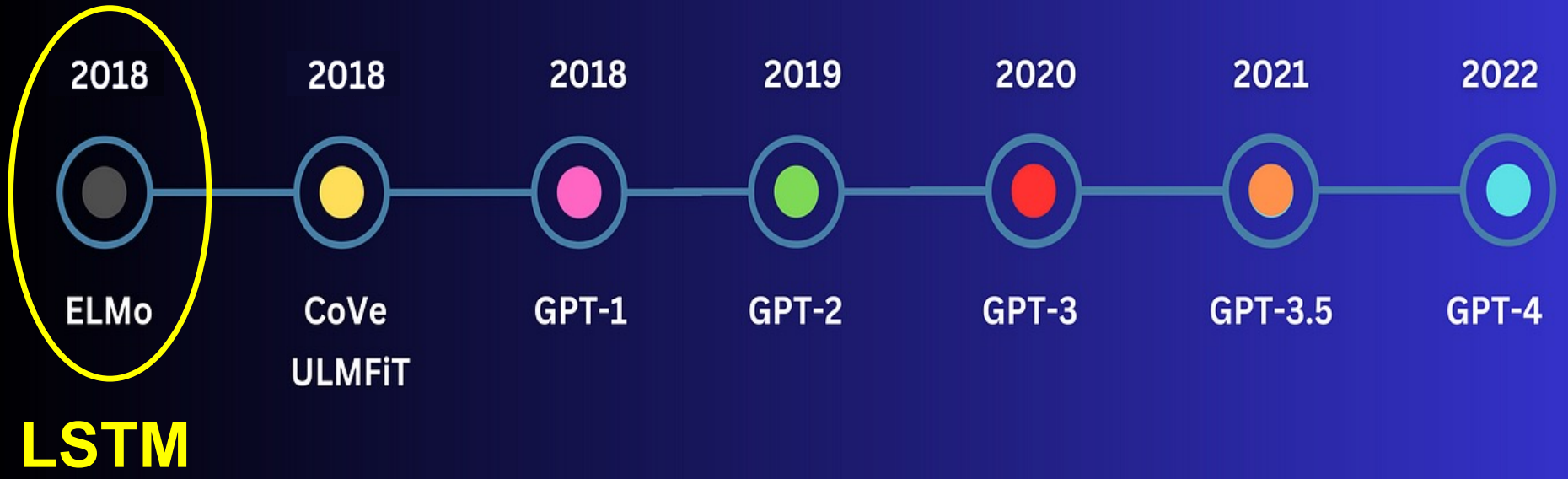




https://en.wikipedia.org/wiki/File:Transformer,_full_architecture.png



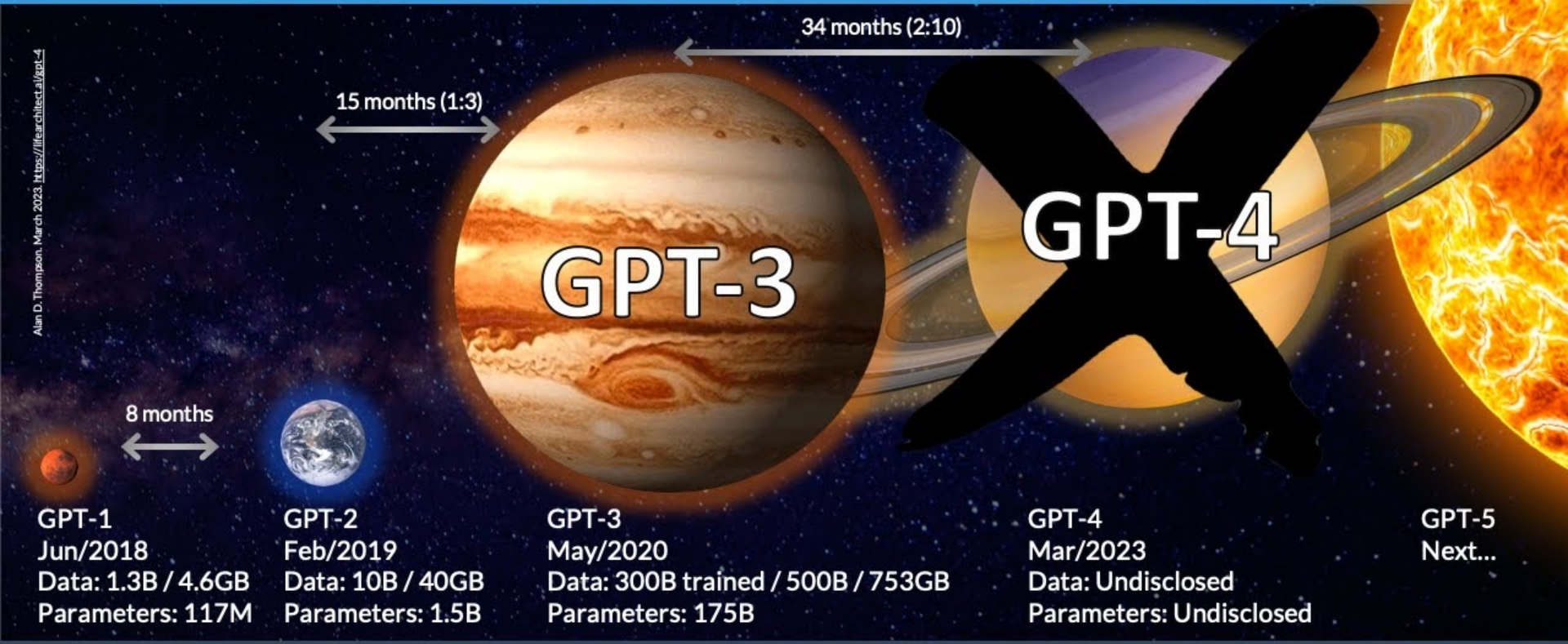
The Evolution of Language Models



RAHUL SINGH

JOURNEY TO GPT-4

ARROWS (RELEASE TIME DELTA) & SPHERES (PARAMS) TO SCALE



GPT-1
Jun/2018
Data: 1.3B / 4.6GB
Parameters: 117M

GPT-2
Feb/2019
Data: 10B / 40GB
Parameters: 1.5B

GPT-3
May/2020
Data: 300B trained / 500B / 753GB
Parameters: 175B

GPT-4
Mar/2023
Data: Undisclosed
Parameters: Undisclosed

GPT-5
Next...

Industrial AI

NXAI GmbH is dedicated to industrial AI:

- AI for industrial applications in engineering, robotics, construction, design, automation, process control, optimization
- AI revolution in industry at scale and with domain expertise
- AI4Simulation: large-scale industrial simulations
- xLSTM: new scaling technology and European LLM (large language model)

AI for Simulations

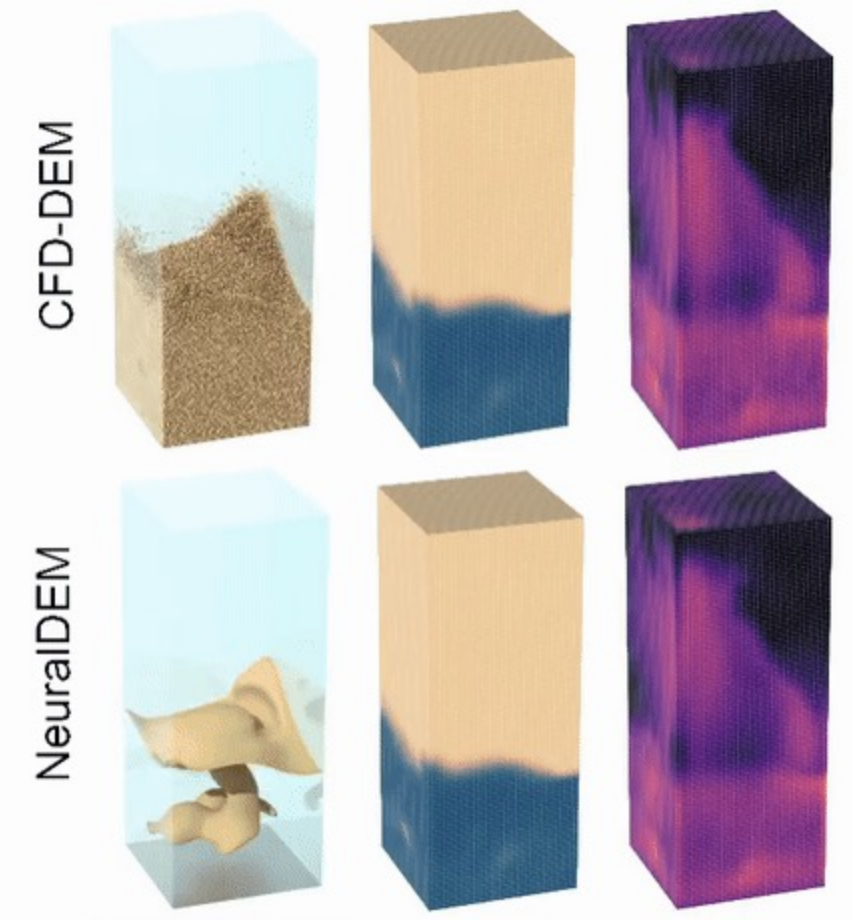
Group of Johannes Brandstetter:

- AI is on the cusp of disrupting industry-scale simulations
- foundation models for language, compute vision, weather modeling, and protein design
- scaling-up deep learning models for every-day engineering and design processes
- disrupts industries

AI for Simulations

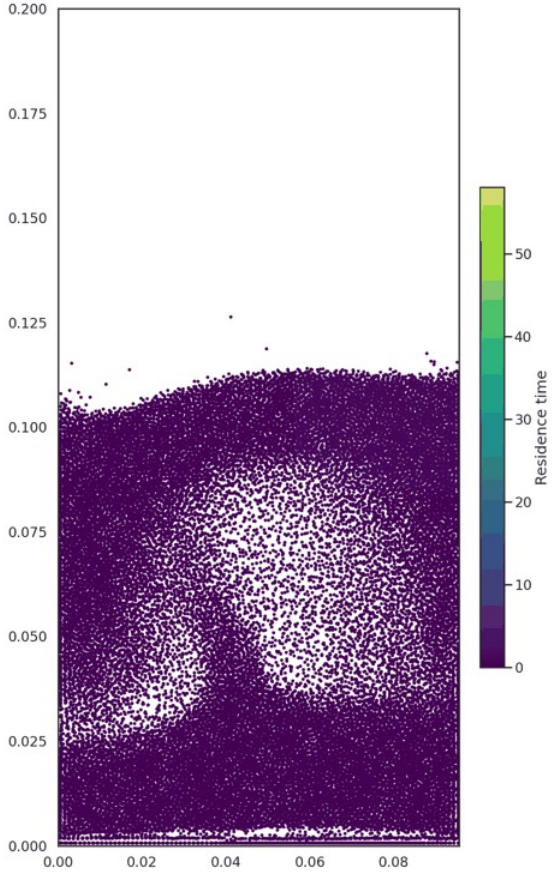
Simulations

- discrete element method (DEM) to simulate particles
- computational fluid dynamics (CFD) for simulating fluid and gas (air)
- neural methods are 1,000 to 10,000 times faster
- neural methods scale to 100 Mio. mesh points or particles while numerical methods are limited to 1 Mio.

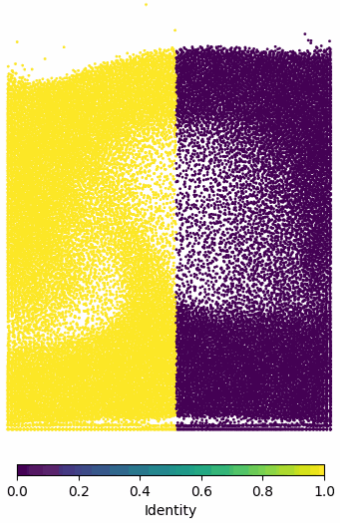


AI for Simulations

Simulations

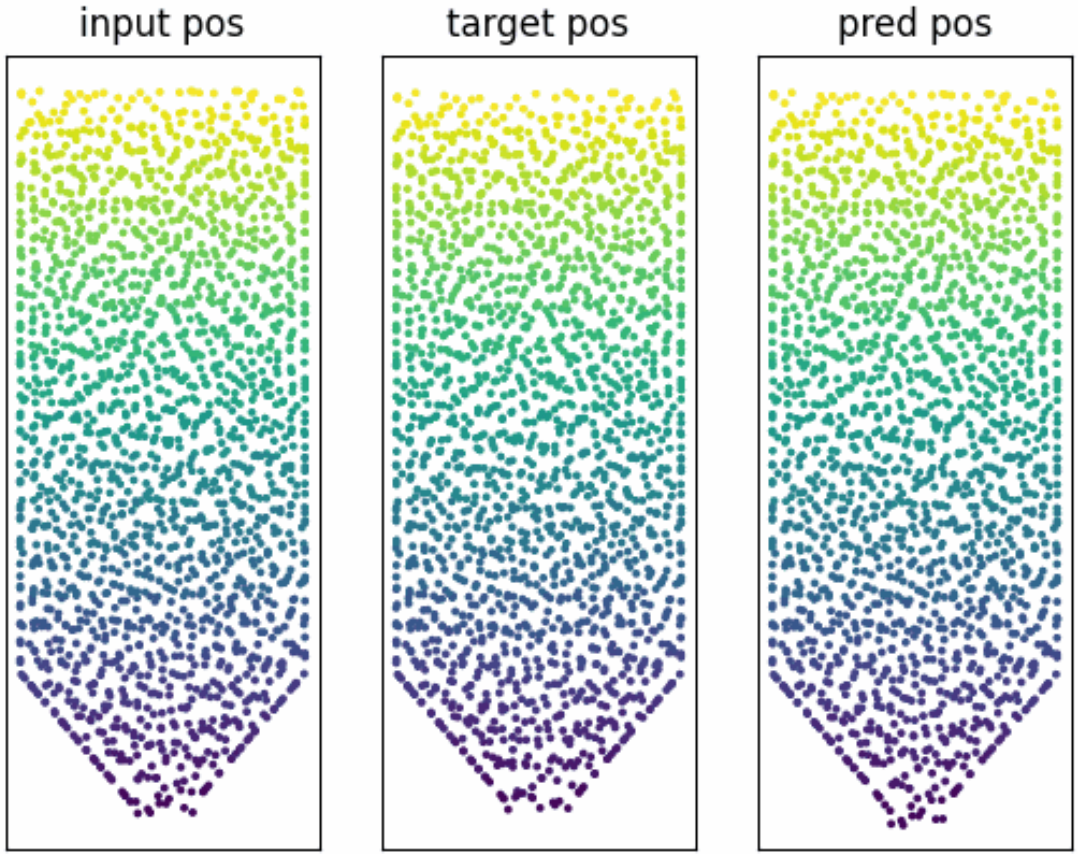


Particle Mixing



AI for Simulations

Simulations



xLSTM

With xLSTM we asked a simple question:

How far do we get in language modeling when

- **scaling** LSTMs to billions of parameters
- leveraging the latest techniques from LLMs
- mitigating known limitations of LSTMs?

xLSTM

xLSTM outperformed all competitors on all tested (small) language datasets.

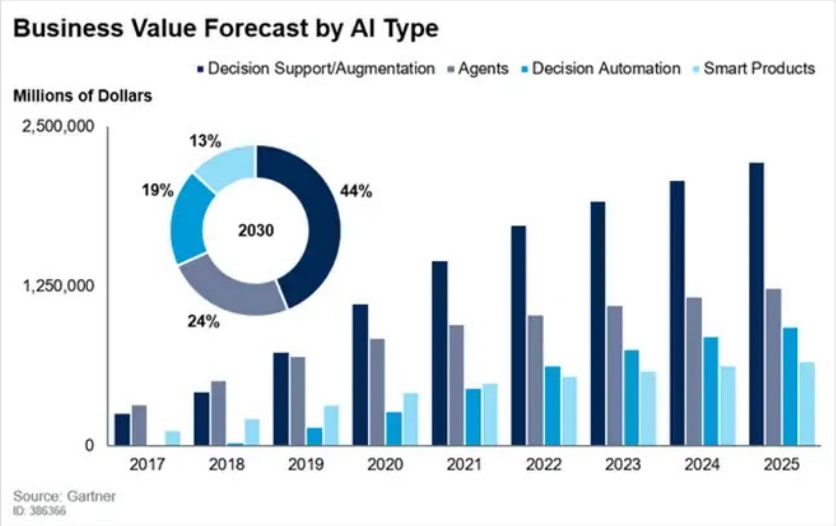
xLSTM is faster than GPTs and has less compute, therefore **is more energy efficient.**

Productivity Increase by AI

>apagen<

odoo

Artificial Intelligence is Beneficial to productivity?



The **European AI ecosystem** is strong, but remains fragmented

The leaders,

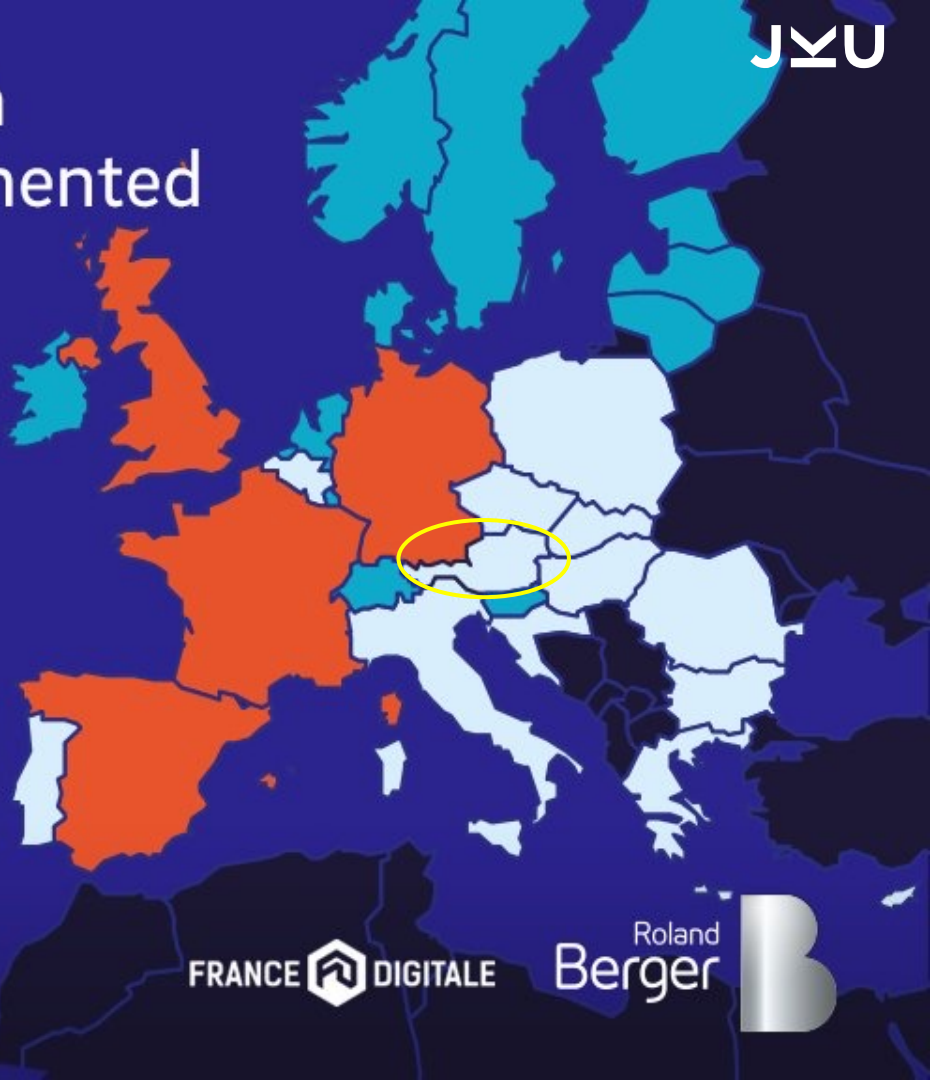
4 countries contributing 60% of startups, labs and communities across the 30 countries.

The rising stars,

comprising 12 countries, mainly Nordic and Baltic states. Very dense ecosystems, +high level of private sector AI research.

The followers,

14 remaining countries. No critical mass, no specific density.



Source France Digitale, Roland Berger

BILAI

Bilateral Artificial Intelligence

Martina Seidl
Symbolic AI
SAT Solving
Formal methods

Sepp Hochreiter
Machine Learning
LSTM
Vanishing gradient

- Institute for Machine Learning
- ELLIS Unit Linz
- LIT AI Lab
- Institute for Symbolic Artificial Intelligence

Christoph Lampert
Machine Learning
Trustworthy Learning

- Machine Learning and Computer Vision group
- ELLIS Unit ISTA

Axel Polleres
Knowledge Graphs

- Institute for Data Process and Knowledge Management

BILAI Consortium:
35 Key researchers
13 ELLIS members
11 ERC grants
2 FWF Wittgenstein awards
3 FWF START prizes

Agata Ciabattoni
Logic
Reasoning

- Institute for Logic and Computation

Gerhard Friedrich
Symbolic AI
Model-based reasoning

- Institute for Artificial Intelligence and Cybersecurity

Robert Legenstein
Machine Learning
Computational Neuroscience

- Institute of Theoretical Computer Science

Thomas Eiter
Symbolic AI
Knowledge representation

END