KU LEUVEN

Impact of AI on research, education, innovation at research universities

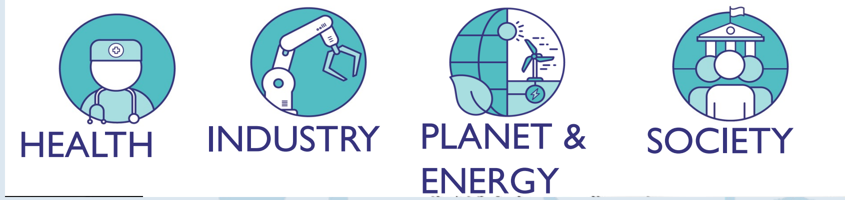
Prof. Dr. Bart De Moor

Co-coordinator Flanders AI Program
Chairman athumi (Flanders Data Utility Company)

Menu

- AI: what, when, how ?

- Inspiring examples of AI in science



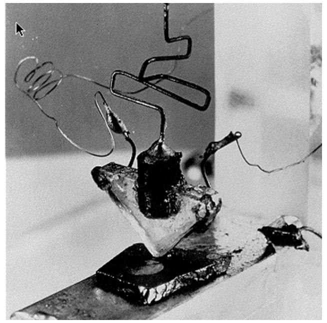
- Issues in ‘AI in science’ policies and science policy for AI

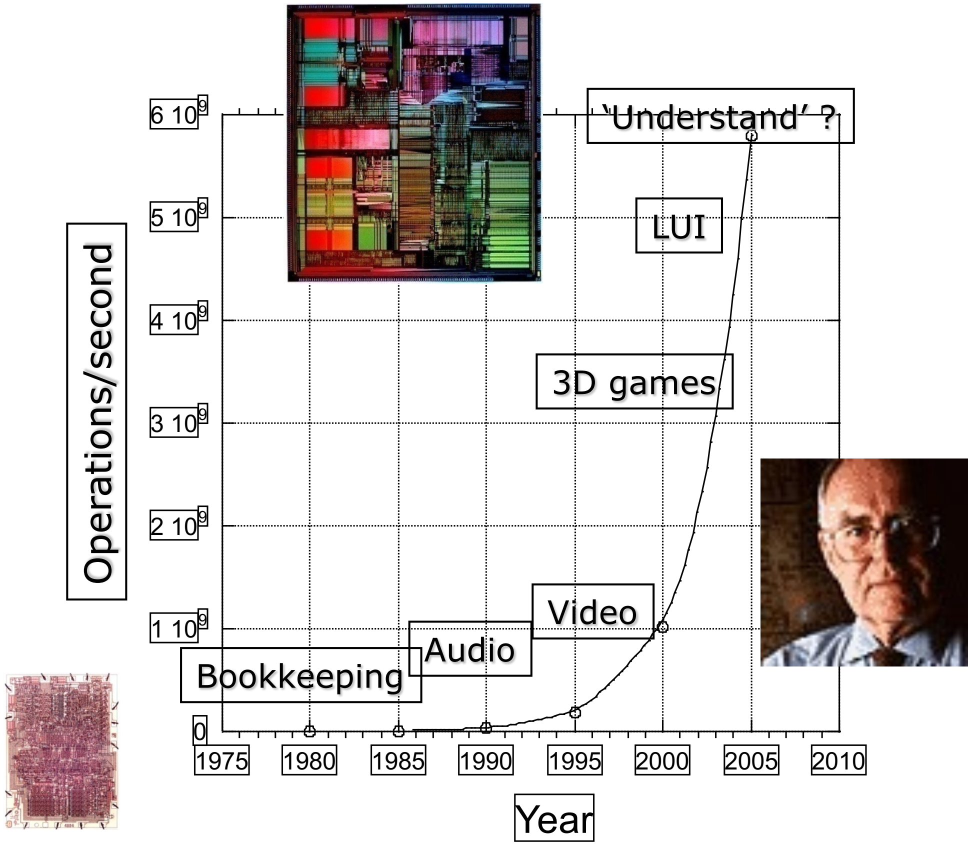
What is Artificial Intelligence and Machine Learning?

- Artificial intelligence
  - “Intelligence” as demonstrated by a machine unlike ‘natural (animal, human) intelligence’
  - Mimic the human mind in ‘cognitive functions’ and ‘problem solving’
  - Mimic = by massive computing power, exploiting tsunami of data
  - Inter/cross-disciplinary: mathematics, computer information science, control engineering, psychology, linguistics,…
  - Emotionality ? (Self-)consciousness ?

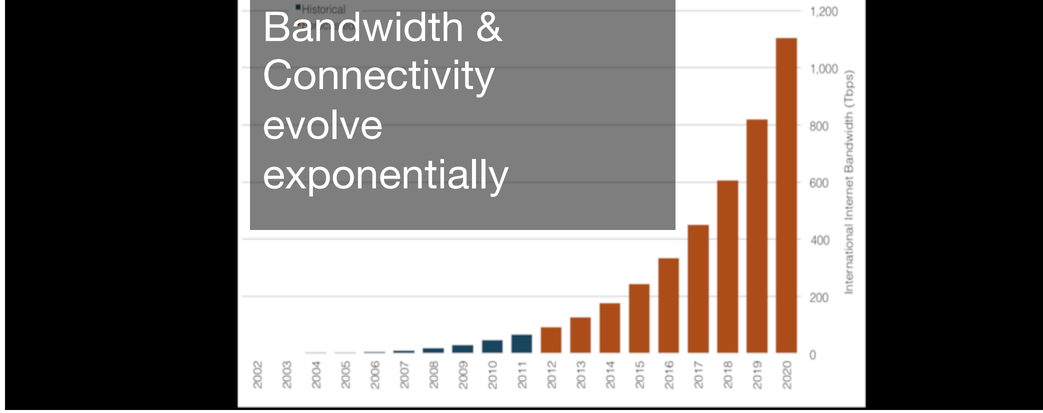
- Machine Learning
  - Computer algorithms that ‘improve’ their performance through experience/data processing
  - Supervised (e.g. by providing classification labels) or unsupervised (e.g. data reduction)
  - Interdisciplinary: mathematics, statistics, numerical optimization, …
  - Training and validation data
  - Generalization, Transfer Learning, Generative AI, …

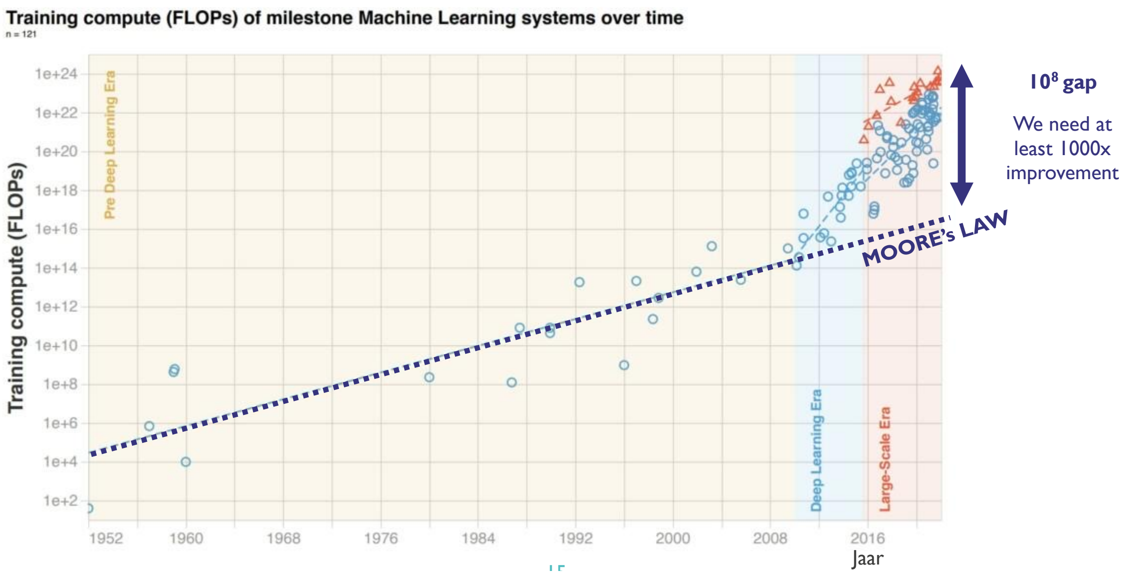
Technology and Engineering Design:
The third industrial revolution (1945…)



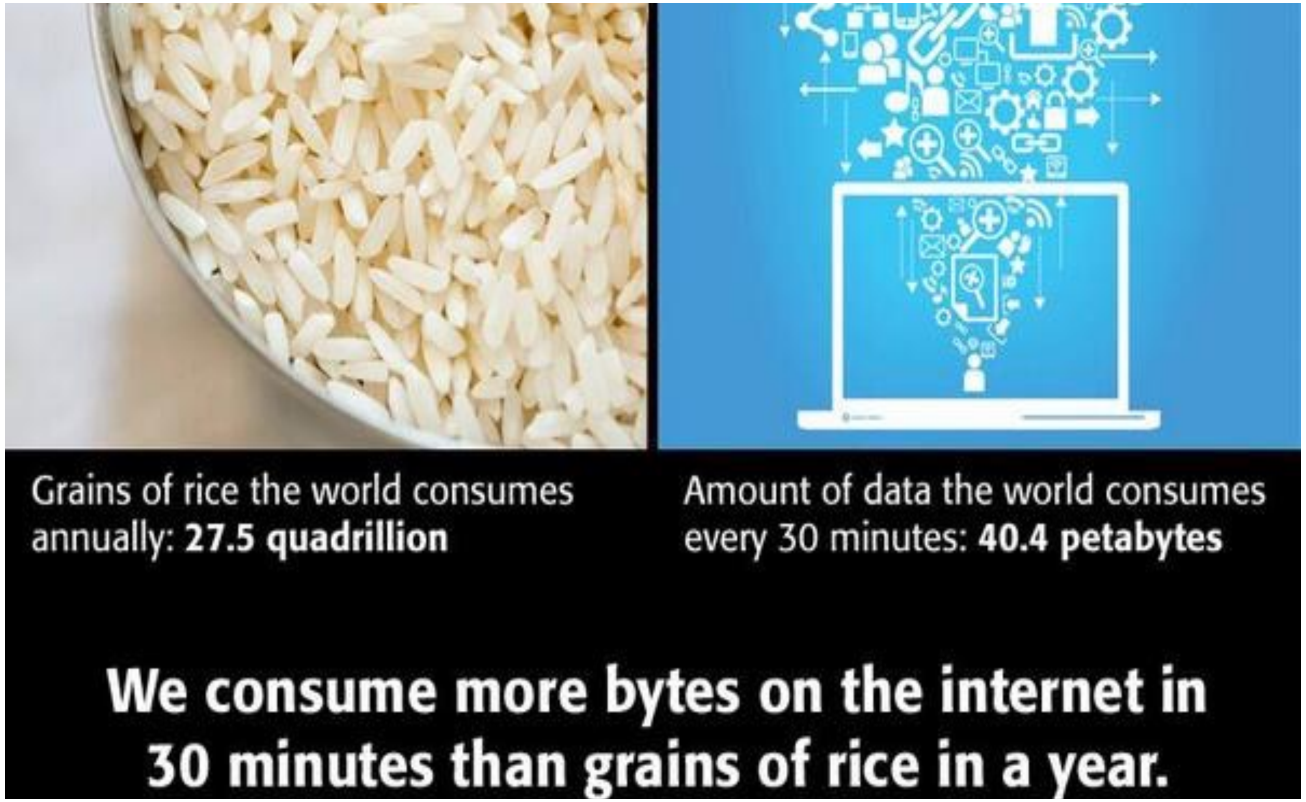


Computational power x 2 every 18 months





KU LEUVEN



1 million = 1 000 000	1 kB = 1 000	1 TB
1 billion = 1 000 000 000	1 MB = 1 000 000	= large university library
1 trillion = 1 000 000 000 000	1 GB = 1 000 000 000	= 212 DVD discs
1 quadrillion =	1 TB = 1 000 000 000 000	= 1430 CDs
1 000 000 000 000 000	1 PB = 1 000 000 000 000 000	= 3 year music CD quality

Data

The Fourth Paradigm

Paradigm	Time Ago	Method
First	A millenium	Empirical
Second	A few centuries	Theoretical
Third	A few decades	Computational
Fourth	Today	Data-driven

Evolution

From

I have a hypothesis

⇩

I need data to check it

⇨

To

I have data

⇩

Which hypotheses can I check?

⇨ Data first!

review articles

DOI:10.1145/3495256

Given the complexity of data science projects and related demand for human expertise, automation has the potential to transform the data science process.

BY TIJL DE BIE, LUC DE RAEDT, JOSÉ HERNÁNDEZ-ORALLO, HOLGER H. HOOS, PADHRAIC SMYTH, AND CHRISTOPHER K.I. WILLIAMS

Automating Data Science

Data Engineering:
data wrangling,
data integration,
data preparation,
data transformation,
…

Model Building:
algorithm selection,
parameter optimization,
performance evaluation,
model selection,
…

Data Exploration:
domain understanding,
goal exploration,
data aggregation,
data visualization,
…

Exploitation:
model interpretation and visualization,
reporting and narratives,
predictions and decisions,
monitoring and maintenance,
…

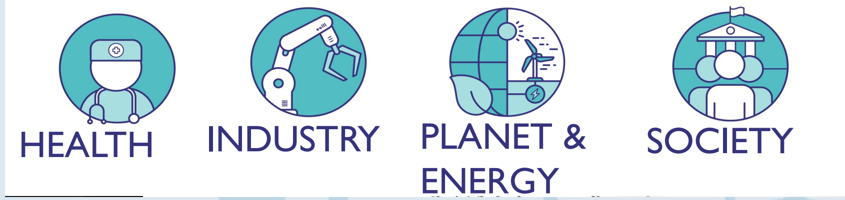
Generic data processing tasks

- Data preprocessing, denoising, normalization
- Clustering and classification; feature detection; profiling;
- Relevance detection, ranking
- Dynamic modelling, time series, longitudinal modelling
- Decorrelation, modelling, (Kalman) filtering
- Predictive analytics
- Vizualisation
- Heterogeneous data fusion
- Prediction, processing and monitoring

Menu

- AI: what, when, how ?

- Inspiring examples of AI in science



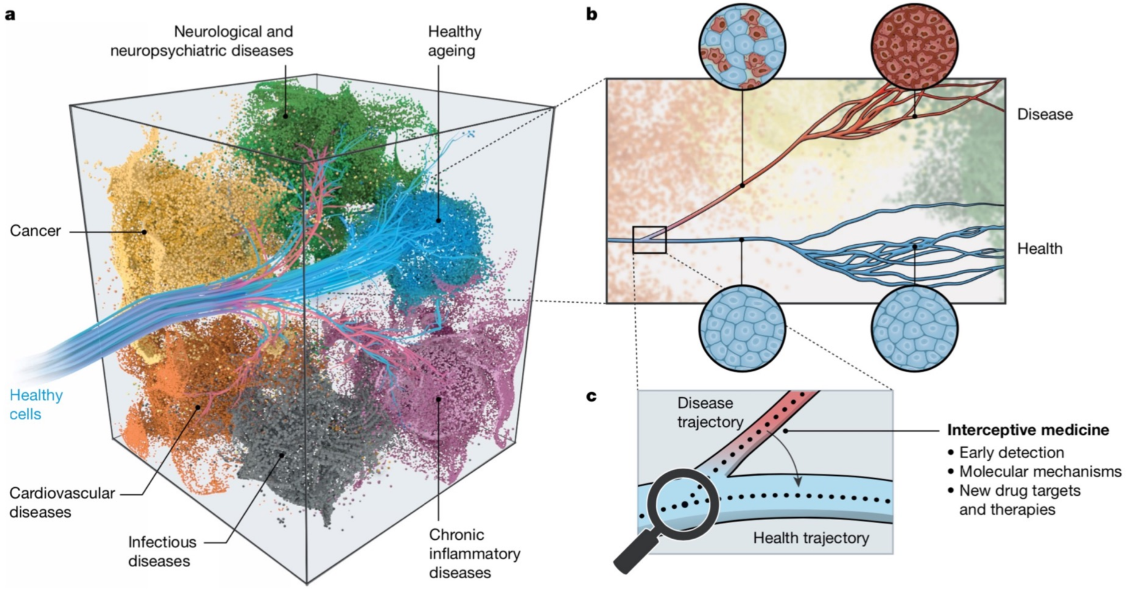
- Issues in ‘AI in science’ policies and science policy for AI

KU LEUVEN

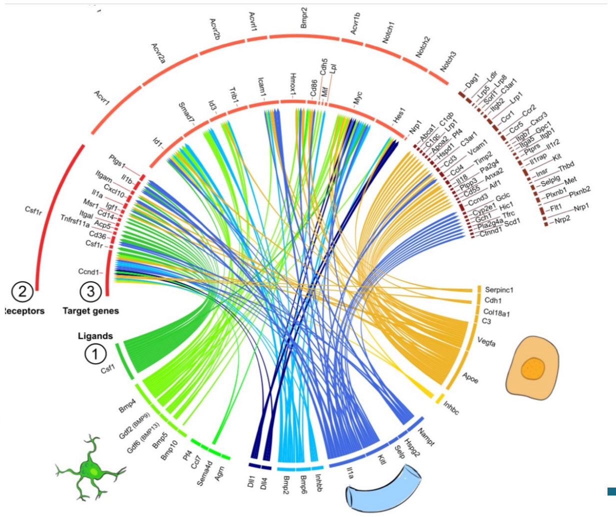
Single Cell Technologies

How normal and disease cells evolve and interact, including spatial resolution

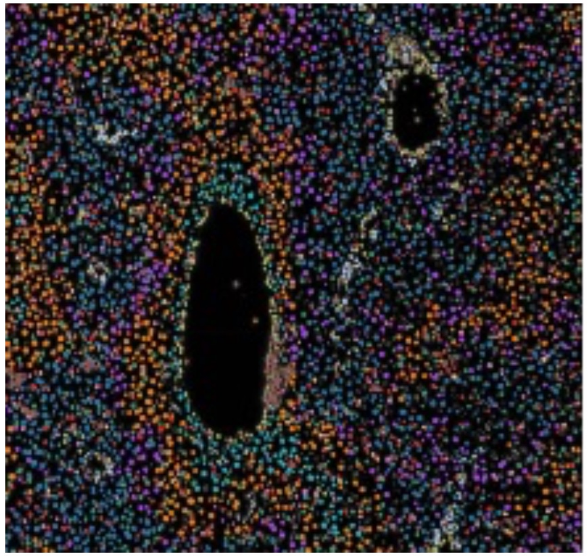
Evolution over time



Cell interactions



New spatial dimension



- New technologies create novel and much richer data types that should be integrated and interpreted to yield novel biological insights
- Multimodal measurements at single cell resolution (genomics, transcriptomics, epigenomics, …)
- Spatial omics (convergence of ‘omics’ and imaging) and imaging cytometry

KU LEUVEN

Dr. Algorithm is coming

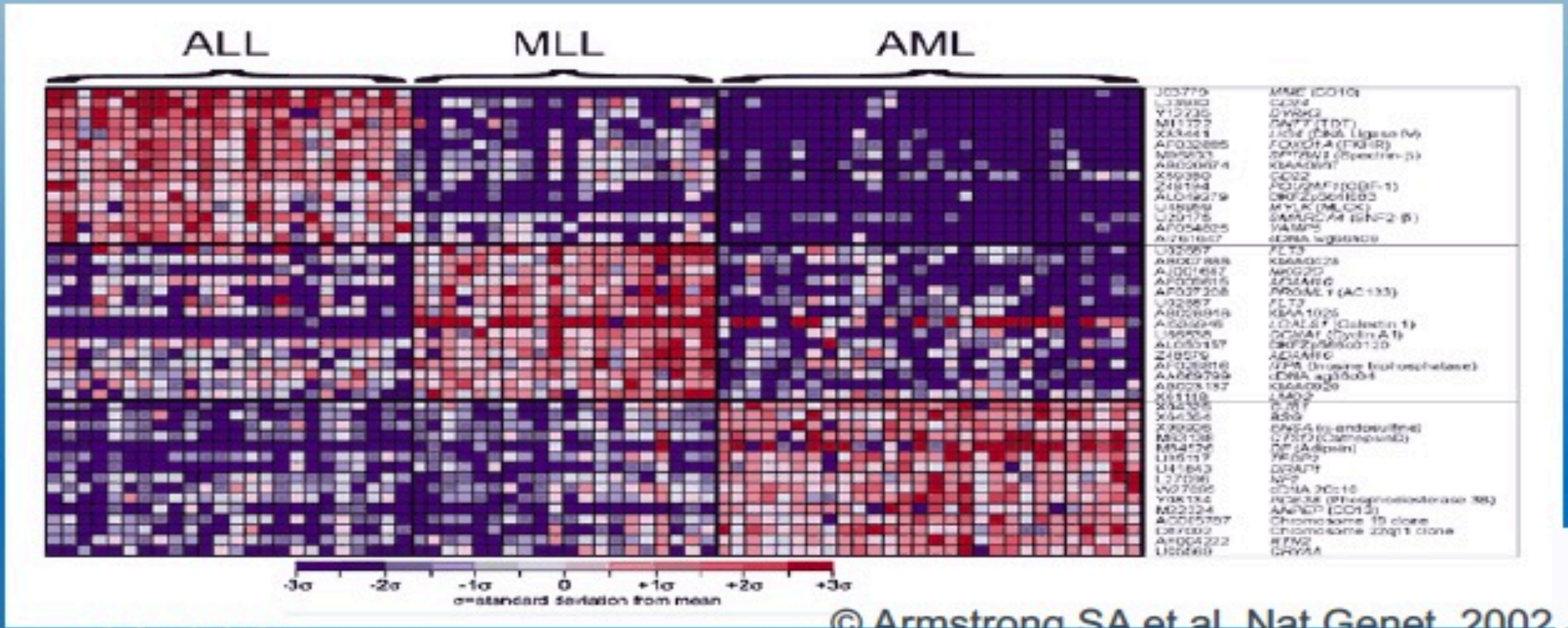
"In the next 10 years, data science and software will do more for medicine than all the biological sciences together."

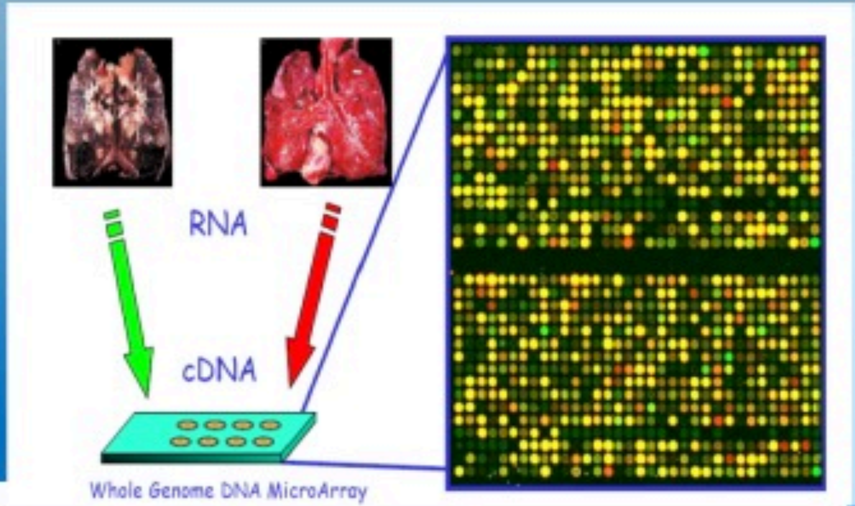
– Vinod Khosla, Khosla Ventures

An AI revolution is brewing in medicine

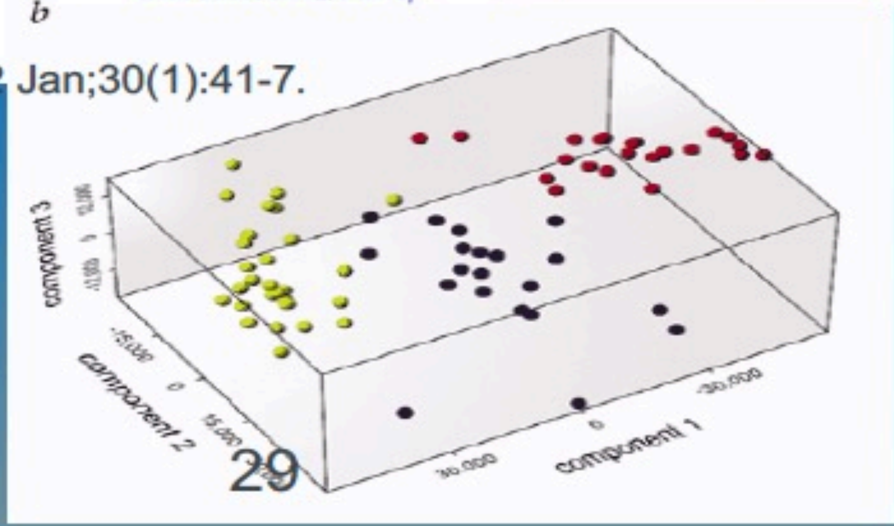
Excitement seems to be growing about generalist medical artificial intelligence (AI). Unlike the tools available now, which serve specific functions such as detecting lung nodules in a chest scan, generalist models would act more like a physician.
They could assess every anomaly in a scan and even create something like a diagnosis. Inspired by the models underlying chatbots such as ChatGPT, they would be trained on massive data sets in the hope that it will give them broad capabilities.

Example: Genomic markers for Leukemia





© Armstrong SA et al. Nat Genet. 2002 Jan;30(1):41-7.

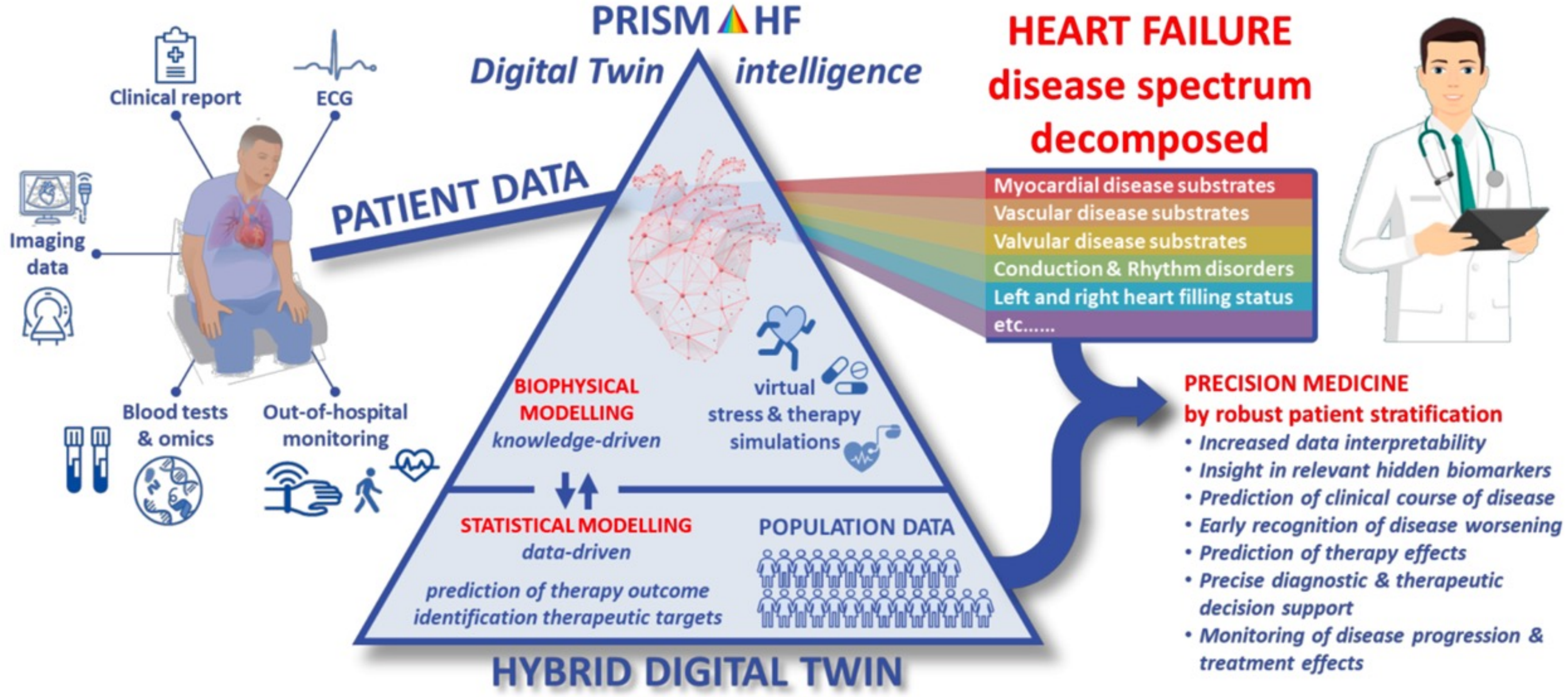


12 600 genes

72 patients

- 28 Acute Lymphoblastic Leukemia (ALL)
- 24 Acute Myeloid Leukemia (AML)
- 20 Mixed Linkage Leukemia (MLL)

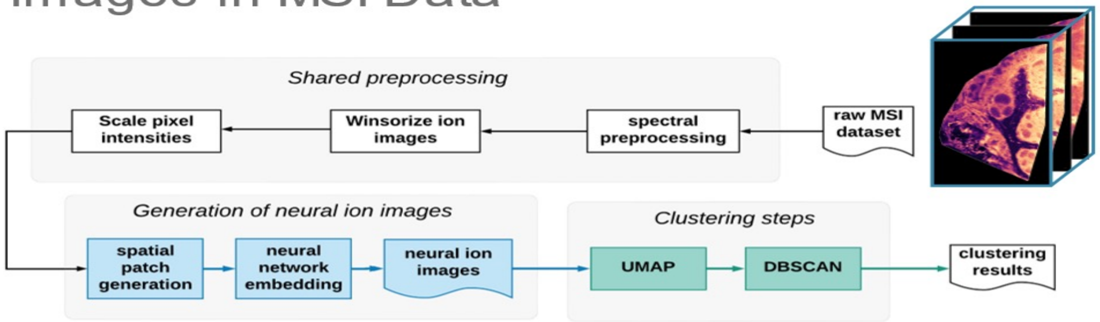
Digital Twin Cardio

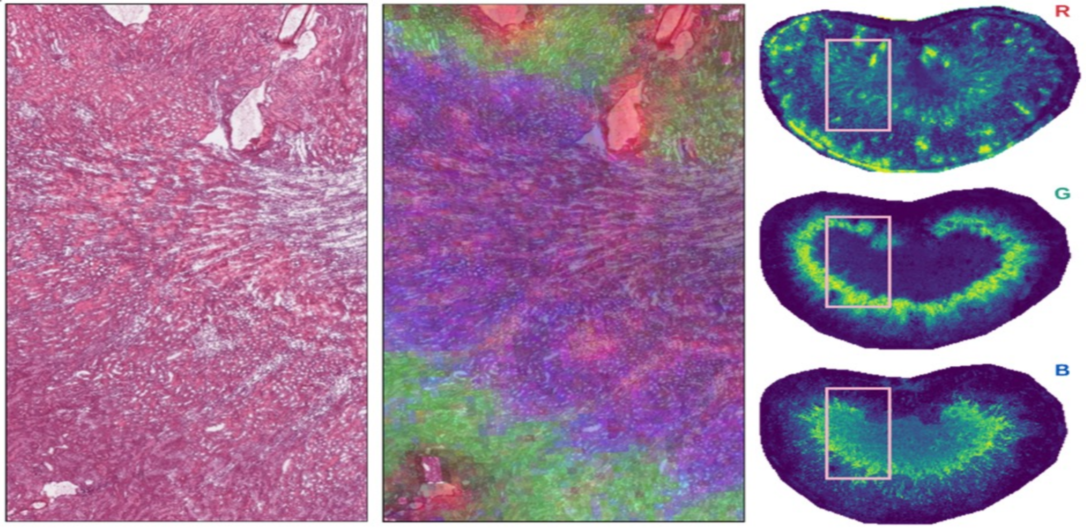


The **objective** of this use case is **to provide disease-tailored treatments for each individual patient at the right time according to their individual pathophysiological disease spectrum** by harvesting multimodal and heterogeneous data and allowing to dynamically update those predictions when additional data becomes available.

cDSS for enhanced interpretation of Mass Spectrometry Imaging data using Deep Learning

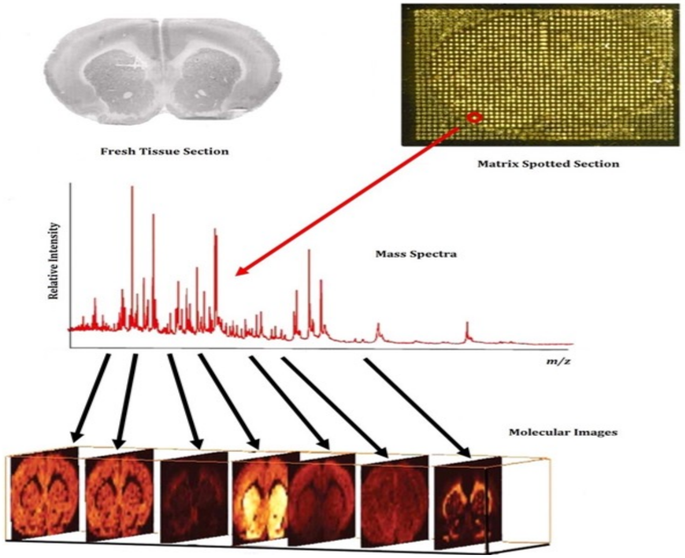
Deep Learning Enables Spatially-aware Clustering of Ion Images in MSI Data

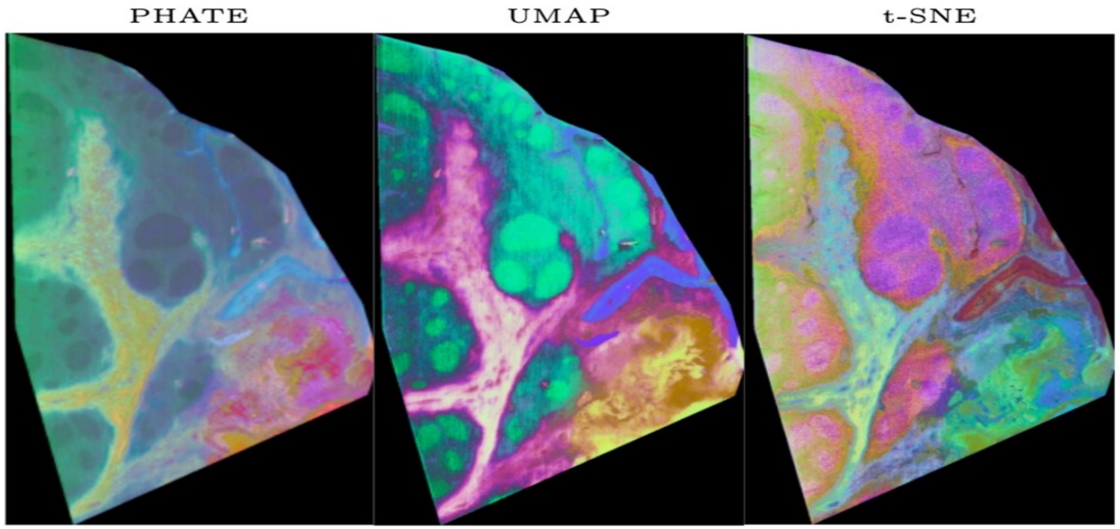




Neuron ion images using pre-trained neural networks enabled direct incorporation of spatial expression patterns, improving clustering of similar ion images, making them readily available as inputs for any downstream machine learning pipeline

Mass Spectrometry Imaging Data Analysis Using Unsupervised Machine Learning





UMAP, with it's very memory efficient GPU implementation, outperforms the other methods in computational time, and gives the most intuitive visualisations on our datasets available.

Data-driven derivation and validation of novel phenotypes for acute kidney transplant rejection using semi-supervised clustering

JASN
JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY

BACKGROUND

The Banff classification for kidney transplant histology is increasingly complex with intermediate and mixed rejection phenotypes.

METHODS

Training cohort | Validation cohorts

1 Belgian kidney transplant center 3510 biopsies

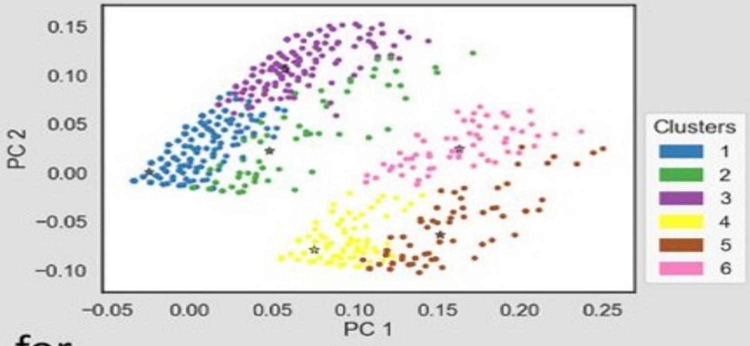
2 French transplant centers with 3835 biopsies

Semi-supervised clustering

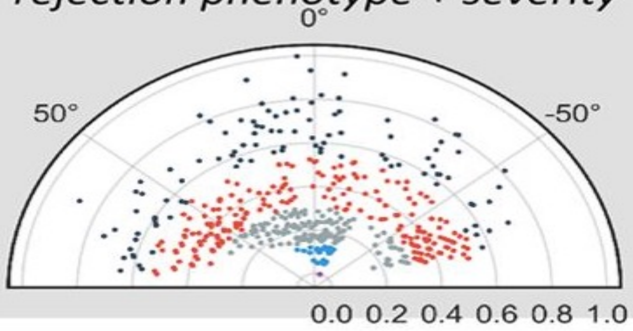
Acute Banff lesion scores
+
Graft survival information

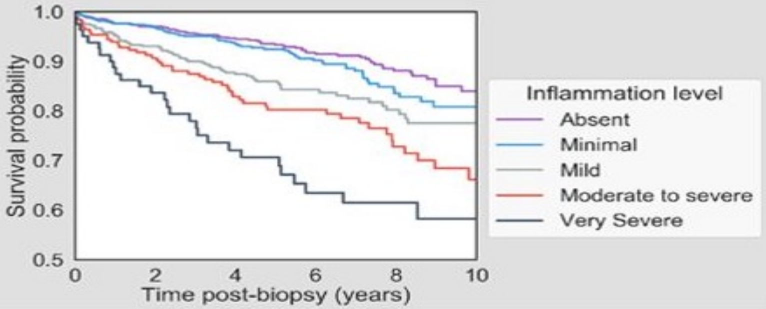
RESULTS

Six data-driven histological (rejection) phenotypes that are clinically sound and that associate with graft failure



Intuitive visualization algorithm for *rejection phenotype + severity*





Conclusion

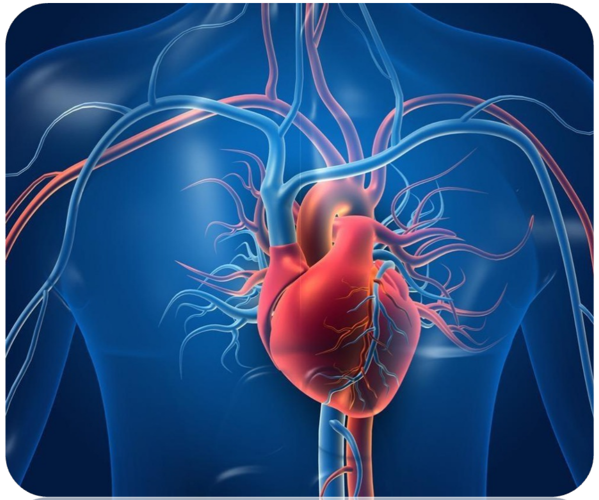
Based on data-driven clustering, six novel clinically and histologically meaningful phenotypes are discovered and validated on external data. Each of them significantly associates with graft failure and overcomes the current limitations of intermediate and mixed Banff phenotypes.

doi: 10.1681/ASN.2020101418

Automate and generalize the development and implementation of AI solutions for Real-World-Evidence generation for several complex clinical finalities

3 clinical finalities


Multiple sclerosis


Cardiovascular disorders


Risk pregnancy

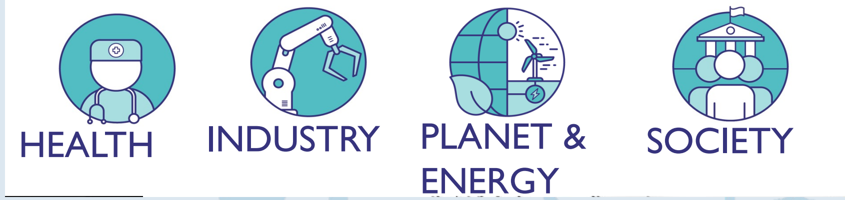
- During the recent years, several (isolated) demonstrator projects have shown the potential of the use of AI for improved RWE generation - Technical challenges that are tackled in specific demonstrators are not always use-case specific. It remains unclear how easily lessons learned from one use case can be leveraged to a variety of clinical research questions
- Urgent need to scale and speed-up the development of so called “AI solutions” and reduce recurring costs and increased time to address several clinical research questions

KU LEUVEN

Menu

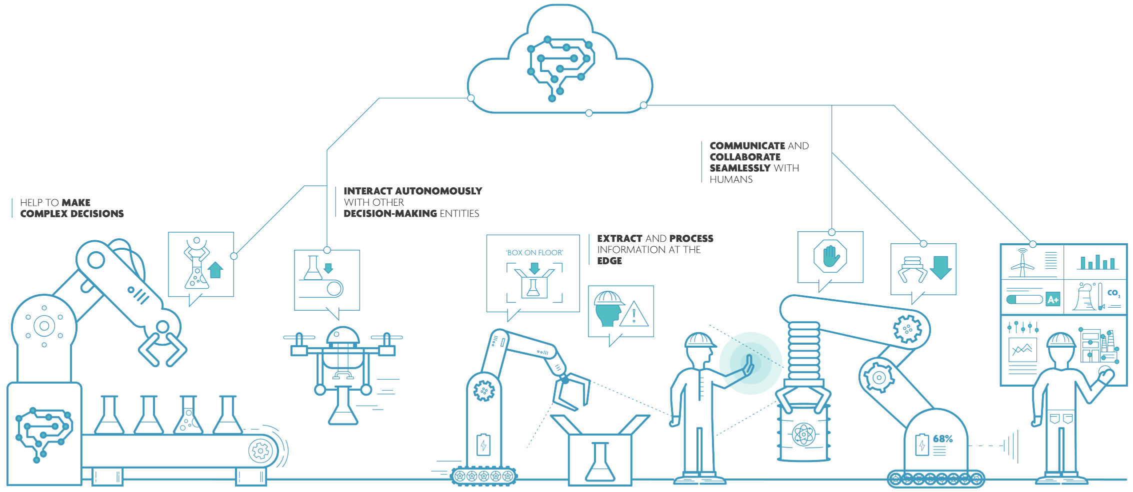
- AI: what, when, how ?

- Inspiring examples of AI in science



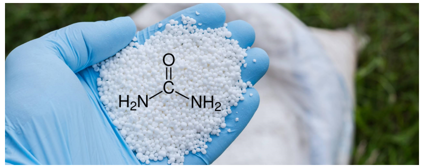
- Issues in ‘AI in science’ policies and science policy for AI

KU LEUVEN

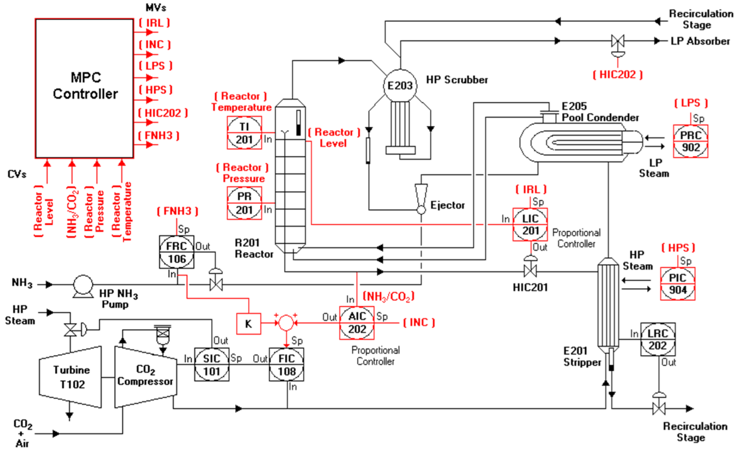


Control of the synthesis section of a Urea plant using MPC control techniques

Urea plant of Yara at Brunsbüttel (Germany)



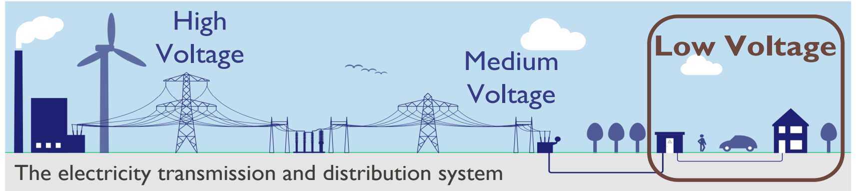
Urea: nitrogenous fertilizer



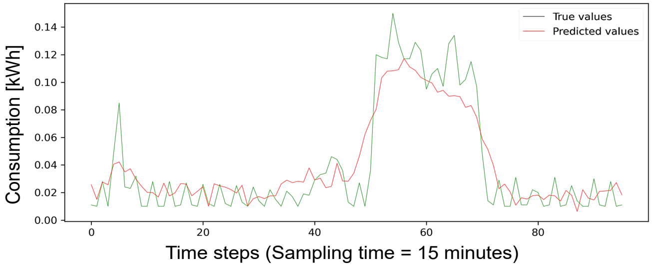
Model Predictive Control (MPC) strategies were used for stabilizing and maximizing the throughput of the synthesis section of a urea plant, while satisfying the process constraints.

Day-ahead consumption forecasting in the Low Voltage (LV) grid using Deep learning

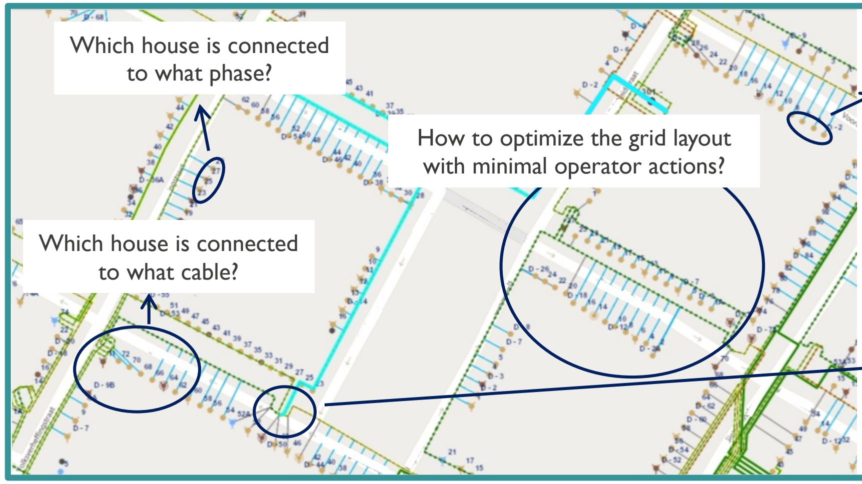
AI FLANDERS
BUILDING OUR DIGITAL FUTURE

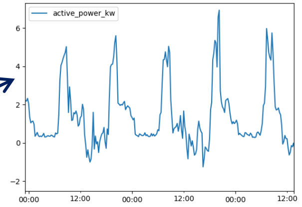


Model: NARX Feedforward Neural Network



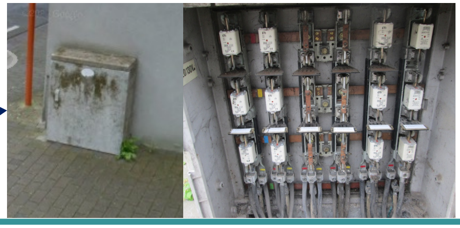
The goal is to build data-driven mathematical models to forecast the electricity consumption of households one day ahead.

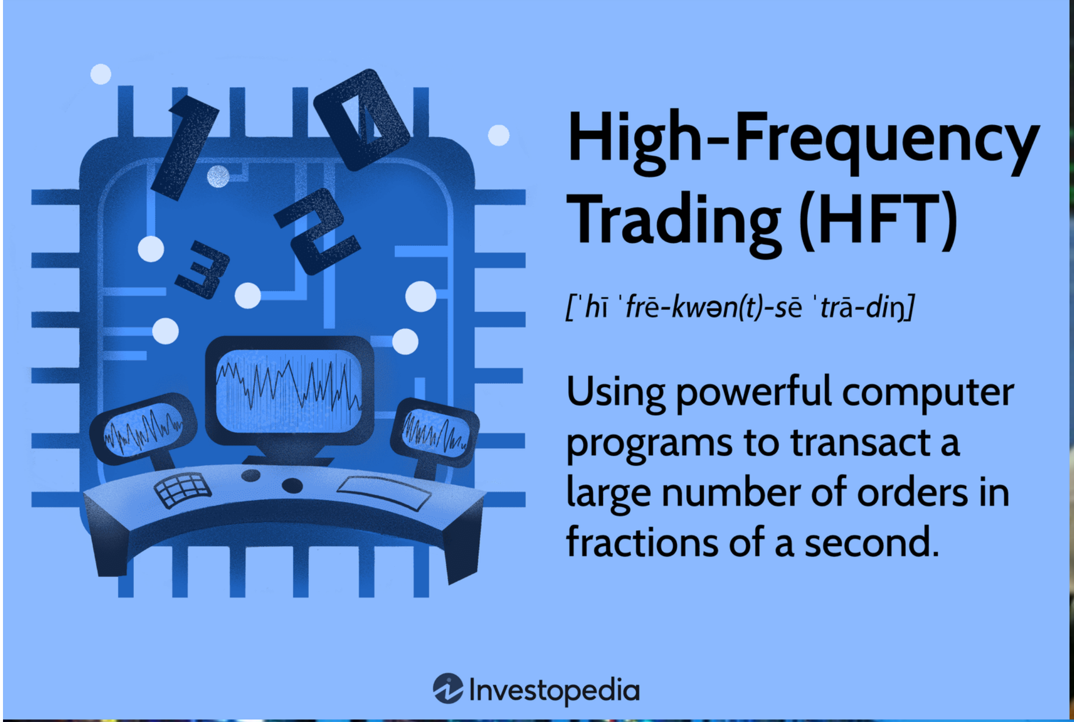




How do heat pumps and electrical cars behave?

Where are the open points?

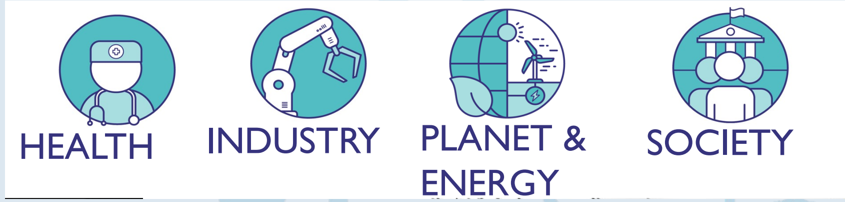



High-Frequency Trading (HFT)
[ˈhī ˈfrē-kwən(t)-sē ˈtrā-diŋ]
Using powerful computer programs to transact a large number of orders in fractions of a second.
Investopedia

Menu

- AI: what, when, how ?

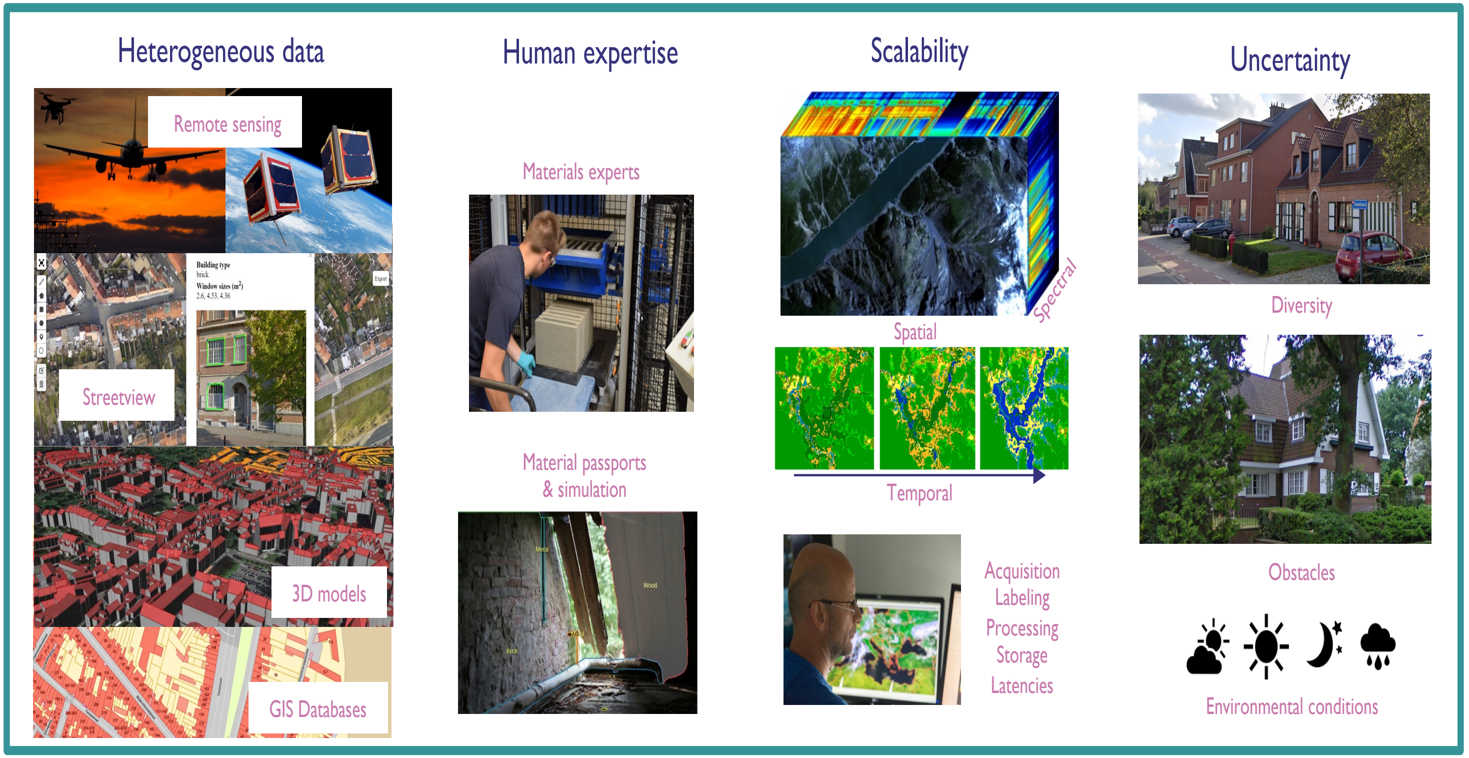
- Inspiring examples of AI in science



- Issues in ‘AI in science’ policies and science policy for AI

KU LEUVEN

AI in and for urban and geographical research



- Data transformations for heterogeneous data
- Integrated information extraction from imagery, databases, graphs, ...
- Multimodal data exploration and modeling

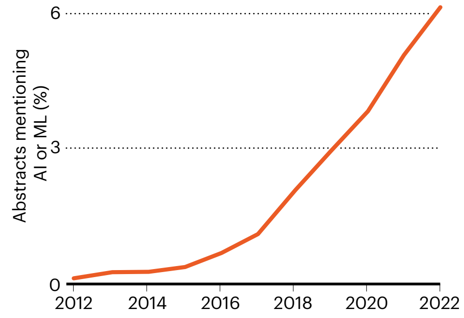
- Dealing with limited ground truth
- Incorporate spatial structure cues
- Expert knowledge-based learning
- Explainable to human experts

- Dense representation learning
- Unsupervised learning at object level
- Efficient data acquisition strategies
- 3D shape/volume extraction

- Uncertainty aware decision making
- Scene decomposition
- Continual learning
- Domain shifts

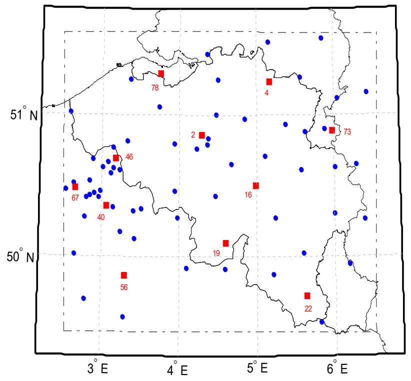
GROWING AI USE IN EARTH AND SPACE SCIENCE

A rising proportion of abstracts for the annual meeting of the American Geophysical Union mention artificial intelligence (AI) or machine learning (ML) — a trend seen across all areas of geoscience.



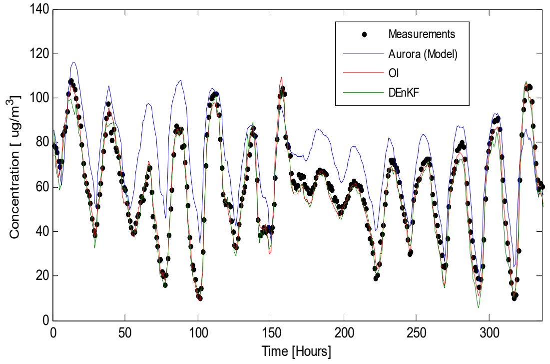
Data assimilation in the Air-quality model Aurora

O_3 air-quality stations



- Assimilation stations
- Validation stations

Average of the O_3 concentration over the validation stations

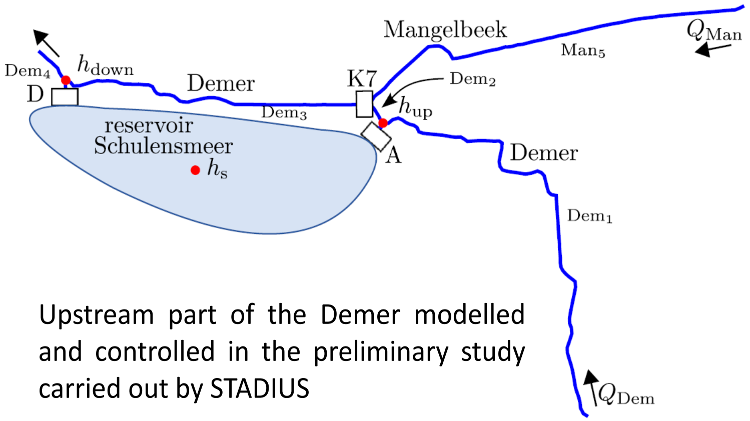


Starting date: May 28th, 2005 at midnight

The objective was to improve the concentration estimates of the air-quality model Aurora by using data assimilation techniques such as Optimal Interpolation (OI), and the Deterministic Ensemble Kalman Filter (DEnKF).

KU LEUVEN

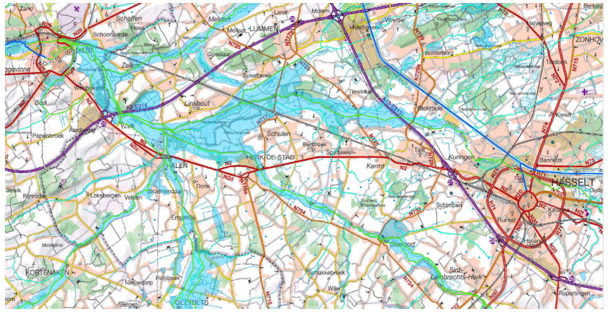
Implementation of a Nonlinear Model Predictive controller (NMPC) for the Demer



Upstream part of the Demer modelled and controlled in the preliminary study carried out by STADIUS

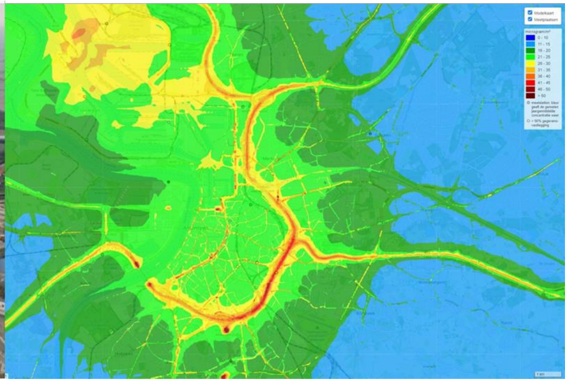
The control goal was to avoid future floodings of the Demer river in Belgium by using NMPC

The Demer in Hasselt



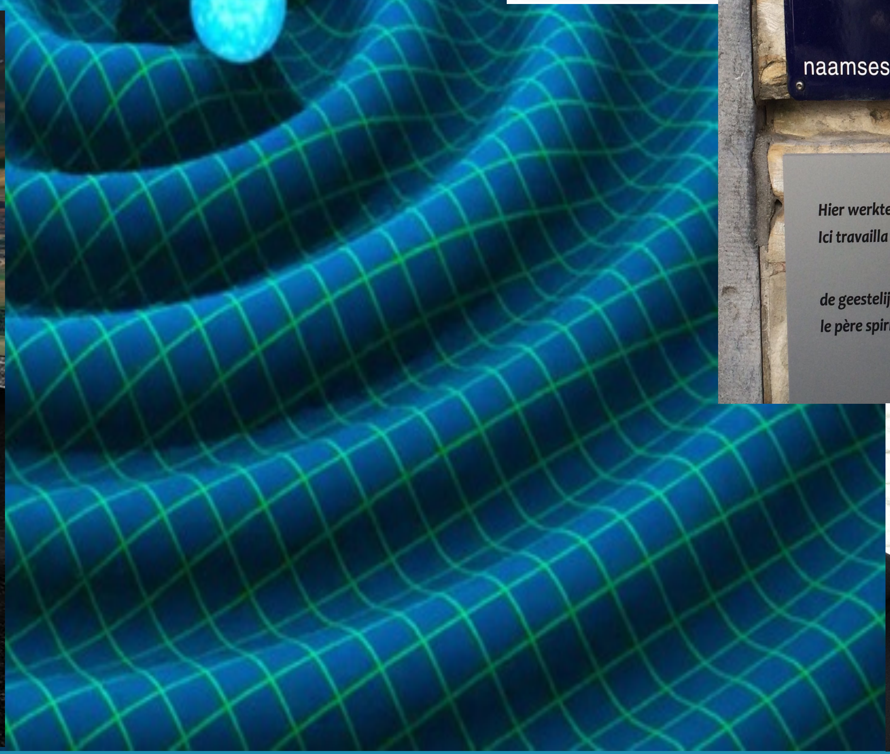
Flooded area during the flood event of 1998

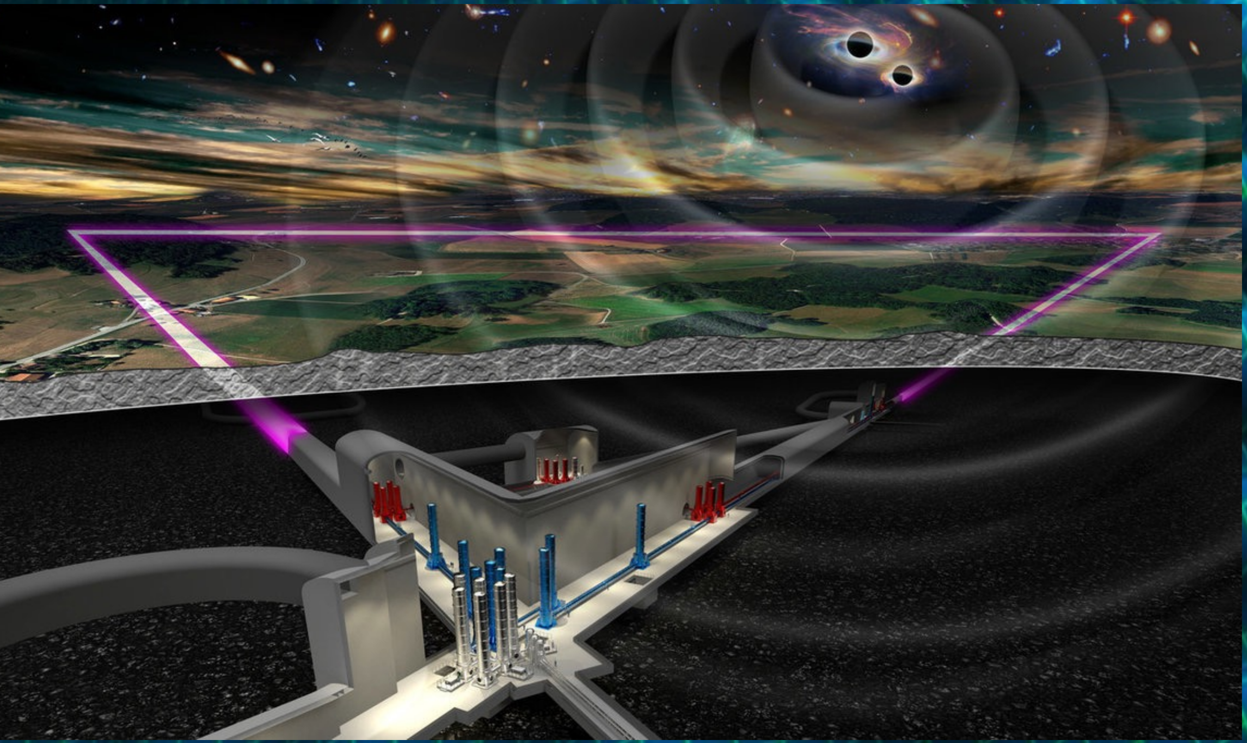
Financial Damage: **> 16 million euro!**



AI helps to save endangered species (Nature)

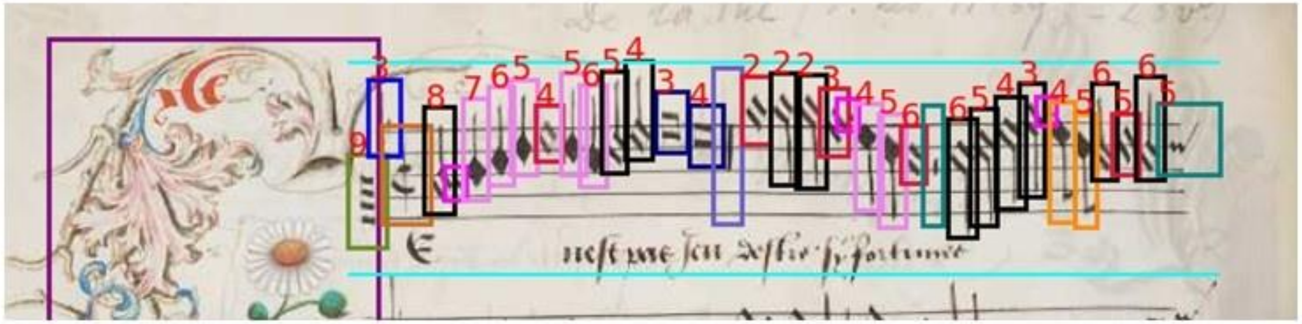
- Artificial intelligence (AI) can help to fight biodiversity loss by analysing rainforest soundscapes or sifting through tens of thousands of camera-trap images. A neural network that can pick out bird species from audio recordings could be crucial for forest-restoration projects that must demonstrate success to secure continued funding. And an AI that analyses footage in real time has already caught a pangolin poacher in the act. Although scientists should be aware that AI is imperfect and has its own environmental impacts, it’s “clearly the way to go”, says Nicolas Miailhe, founder of an AI-governance non-profit.






KU LEUVEN
college van premonstreit
naamsestraat 61
116.01
1425
Hier werkte
Ici travailla
Here worked
Hier wirkte
Georges Lemaître
de geestelijke vader van de
le père spirituel du
father of the
der geistige Vater des
Big Bang

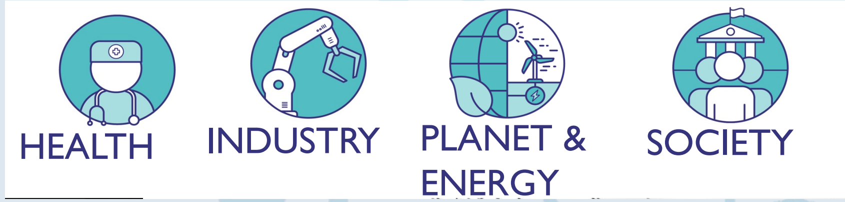

ornate element 0.28
staff 0.77
staff 0.73
staff 0.79
staff 0.78



Menu

- AI: what, when, how ?

- Inspiring examples of AI in science



- Issues in ‘AI in science’ policies and science policy for AI

KU LEUVEN

2024-2028

Foto:
KU Leuven

AI in support of living a healthier life for longer

AI in support of a human-centered prosperous digital future for society

Responsible AI

Foto: Flanders Make

AI in support of Industry 5.0: sustainable, resilient, and human-centered

AI for sustainability, AI taking care of our planet and supporting the energy transition

Foto: VUB

Spin-Off Companies (personal track record)

IPCOS
Creators in Control

Specialized in modelling and control of multivariable industrial processes (chemical and power plants, oil exploration) automation & optimization (1995)
http://www.ipcos.com/en

DATA4s data4s
NORKOM TECHNOLOGIES

Payment fraud detection (2000)
(in 2004 acquired by Norkom Technologies ; in 2011 Norkom Techn. was acquired by Detica NetReveal , Bus. Division of BAE Systems Detica.
http://www.deticanetreveal.com/en/about-us.html

TRANSPORT & MOBILITY LEUVEN

Transport & Mobility research & mgt (2002)
http://www.tmleuven.be/

d square

Data mining software & services for process industry (2008)http://www.dsquare.be/ (won in 2009 the award of "best spin-off of the year")

TrendMiner

TrendMiner: Formerly Dsquare, is the Belgian big-data in process industry company that has been acquired by Software AG
https://www.trendminer.com/

timeseer.AI

Timeseer: Industrial process monitoring
https://www.timeseer.ai/ 1

silicos

In silico drug discovery (2005)

LINDACARE
Enjoy life more.

patient tele-monitoring

Agilent CARTAGENIA

Acquired by Agilent
Data handling & mining for clinical genetics (2008)
https://www.agilent.com/

UgenTec

UgenTec delivers an independent diagnostic software platform to help molecular labs with their DNA (PCR) analyses. It offers The FastFinder 2.0 analysis software for helping lab technicians in their analysis of PCR data. It serves biotech companies, labs, and IT companies.
https://www.ugentec.com/

Aspect Analytics

Software for mass spectrometry imaging
https://www.aspect-analytics.com

Succesful high-tech spinoff = TTT = Technology, Team, Traction

Garbage in, garbage out: mitigating risks and maximizing benefits of AI in research

Six principles to help build trust

Following these best practices will help to avert harm when using AI in research.

Researchers

1. Transparency. Clearly document and report participants, data sets, models, bias and uncertainties.
2. Intentionality. Ensure that the AI model and its implementations are explained, replicable and reusable.
3. Risk. Consider and manage the possible risks and biases that data sets and algorithms are susceptible to, and how they might affect the outcomes or have unintended consequences.
4. Participatory methods. Ensure inclusive research design, engage with communities at risk and include domain expertise.

Scholarly organizations (including research institutions, publishers, societies and funders)

5. Outreach, training, and leading practices. Provide for all roles and career stages.
6. Sustained effort. Implement, review and advance these guidelines.

Can AI systems learn/ be taught to make ethical choices ?

Can AI systems be thaught to deal with differing opinons ?

Einstein Foundation Award 2023 — The Einstein Foundation Berlin awards €500,000 prize to enhance quality in research
The Einstein Foundation Berlin is to honor Belgian bioinformatician Yves Moreau, the Berkeley Initiative for Transparency in the Social Sciences, and the Responsible Research Assessment Initiative with this year's Einstein Foundation Award for Promoting Quality in Research 2023.

The recipient of the **Individual Award** is Yves Moreau from the Katholieke Universiteit Leuven. Moreau ranks among the most ardent advocates for ethical standards in the utilization of human DNA data in the age of artificial intelligence and big data. He designs algorithms that protect personal privacy during the analysis of genetic data.

KU LEUVEN

AI & Science Policy

In June OECD published a report on applying AI to accelerate productivity in research, in which it says, “While AI is penetrating all domains and stages of science, its full potential is far from realised. Policy makers and actors across research systems can do much to accelerate and deepen the uptake of AI in science, magnifying its positive contributions to research.”

National Institutes of Health (NIH) in the US has placed limits on generative AI in its peer review processes

US national Science Foundation has put together an internal working group to find out if there are ways to integrate AI tools in the production of science

A recent workshop at the National Academies explored the present and future of artificial intelligence In advancing discovery across a range of scientific fields, from physics to neurology to meteorology.

While the EU is in the process of finalising its first regulation on artificial intelligence, the scientific community is yet to come up with a unified response on how generative AI could be used higher education and research

European Commision (new unit E4: Industry 5.0 and AI in Science): To develop new policy on how to use AI improve scientific productivity and avoid misuse copyright and academic dishonesty (“CHATGPT detector catches AI-generated papers with unprecedented accuracy”).

KU LEUVEN

EU Commission (policy brief preview)

1. **Introduction**

2. **Context**

(History of AI, European approach to AI)

3. **Why the EU needs a policy for AI in Science**
   3.1 Global competition is under way to harness AI's potential in science
   3.2 The EU needs a competitive advantate in the global AI race
   3.3 AI-boosted science can deliver growth and solutions to societal challenges
   3.4 AI in science comes with new challenges, risks and ethical considerations

4. What AI can do for Science

4.1. AI in Science: a tool for accelerated discovery

4. 2. AI helping science solve global challenges

4. 3. AI-driven collaboration: unlocking innovation

5. A Glimpse into the Future: AI from toolbox to "brainbox

6. **What new challenges from AI in Science**

6.1 AI amplifies ethical issues in science

6.2 The problem of public trust in AI-driven research

6.3. Talent and education system challenges

6.4 Challenges related to data, resources and infrastructures

6.5. Regional disparities and the integration of AI in Science

7. Solutions to enable uptake of AI in Science

7.1 Vision and governance for AI in Science

7.2. Anticipate impact and preserve trust in the scientific practice

7.3 Ramping up access to data, assets and compute infrastructures

7.4 EU R&I funding for AI-driven science

7.5 Enhancing capacities for AI-driven interdisciplinary research

7.6 Education and skills, access to talent

8. **Conclusions**

Annex I: EU initiatives supporting AI in Science

President Biden Issues Executive Order on Safe, Secure, and Trustworthy AI

•Require that developers of the most powerful AI systems share their safety test results and other critical information with the U.S. government.
•Develop standards, tools, and tests to help ensure that AI systems are safe, secure, and trustworthy.
•Protect against the risks of using AI to engineer dangerous biological materials
•Protect Americans from AI-enabled fraud and deception by establishing standards and best practices for detecting AI-generated content and authenticating official content.
•Establish an advanced cybersecurity program to develop AI tools to find and fix vulnerabilities in critical software,
•Order the development of a National Security Memorandum that directs further actions on AI and security.
•Protect Americans' privacy by prioritizing support for accelerating the development and use of privacy-preserving techniques
•Strengthen privacy-preserving research and technologies,
•Evaluate how agencies collect and use commercially available information
•Develop guidelines for federal agencies to evaluate the effectiveness of privacy-preserving techniques,
•Provide clear guidance to landlords, Federal benefits programs, and federal contractors
•Address algorithmic discrimination
•Ensure fairness throughout the criminal justice system
•Advance the responsible use of AI in healthcare and the development of affordable and life-saving drugs.
•Shape AI's potential to transform education
•Develop principles and best practices to mitigate the harms and maximize the benefits of AI for workers
•Produce a report on AI's potential labor-market impacts,

President Biden Issues Executive Order on Safe, Secure, and Trustworthy AI

- Catalyze AI research across the United States
- Promote a fair, open, and competitive AI ecosystem
- Use existing authorities to expand the ability of highly skilled immigrants and nonimmigrants with expertise in critical areas to study, stay, and work in the United States
- Expand bilateral, multilateral, and multistakeholder engagements to collaborate on AI.
- Accelerate development and implementation of vital AI standards
- Promote the safe, responsible, and rights-affirming development and deployment of AI abroad to solve global challenges
- Issue guidance for agencies' use of AI
- Help agencies acquire specified AI products and services
- Accelerate the rapid hiring of AI professionals

The US innovates
China imitates
Europe regulates

True ???

To open or not to open the sources of AI ?

AI risk assessment related to technology security and technology leakage

Guidelines and questionnaire for reporting on findings

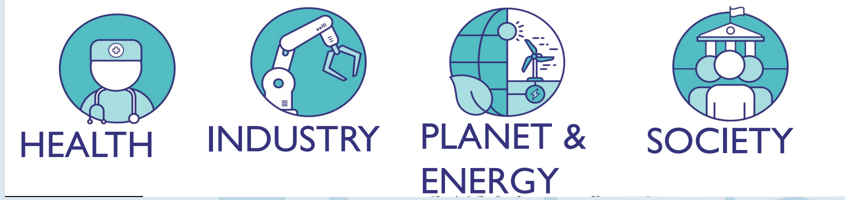
The Commission **Recommendation C(2023) 6689 of 3 October 2023** on critical technology areas for the EU's economic security for further risk assessment with Member States (hereafter 'the Recommendation') proposes that Member States, together with the Commission, perform **a risk assessment**, by the end of 2023, related to technology security and technology leakage risks on the artificial intelligence, high-performance computing, cloud and edge computing technology areas.

Dual use ??

Menu

- AI: what, when, how ?

- Inspiring examples of AI in science



- Issues in ‘AI in science’ policies and science policy for AI

KU LEUVEN