Artificial Intelligence, machine learning and clinical data mining

Prof. Dr. Bart De Moor
KU Leuven, Belgium

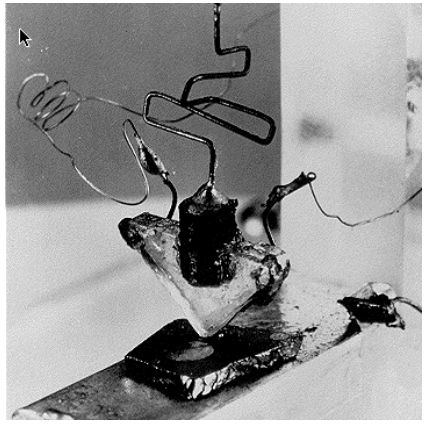
KU LEUVEN

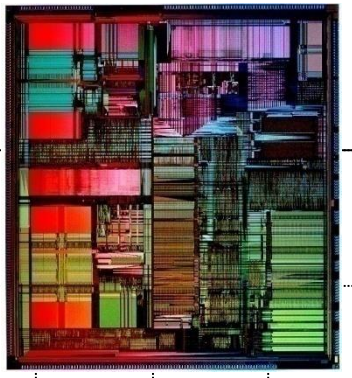
STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

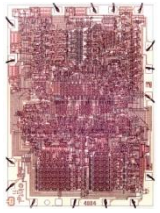
Outline

- AI: Why now ?
- Some clinical examples
- Future opportunities

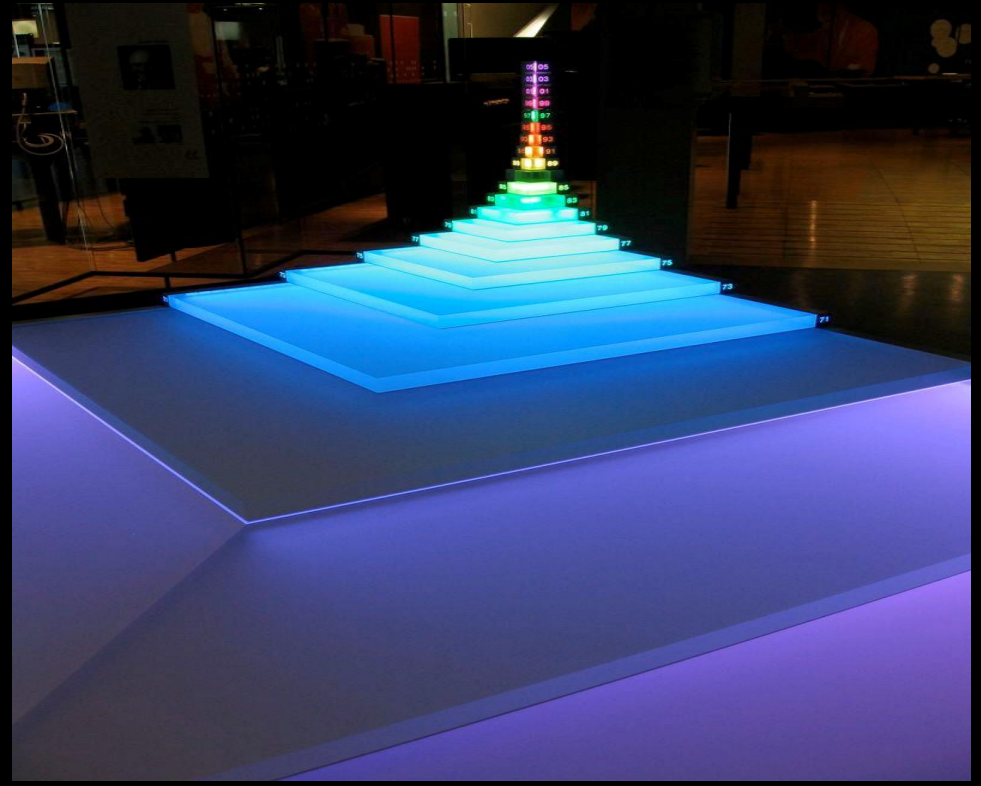
Technology and Engineering Design:
The third industrial revolution (1945…)





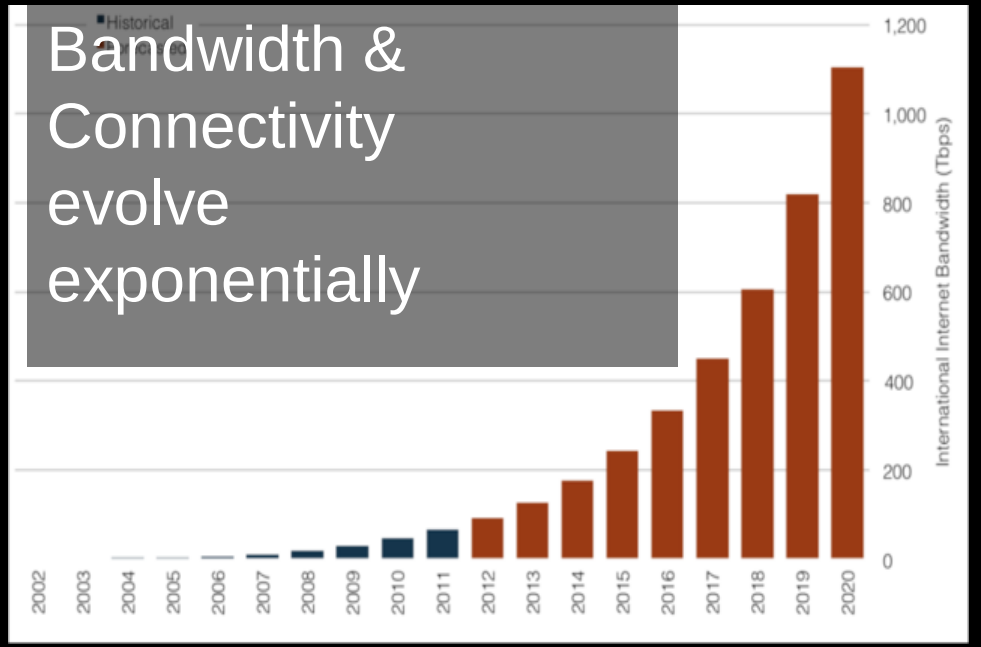


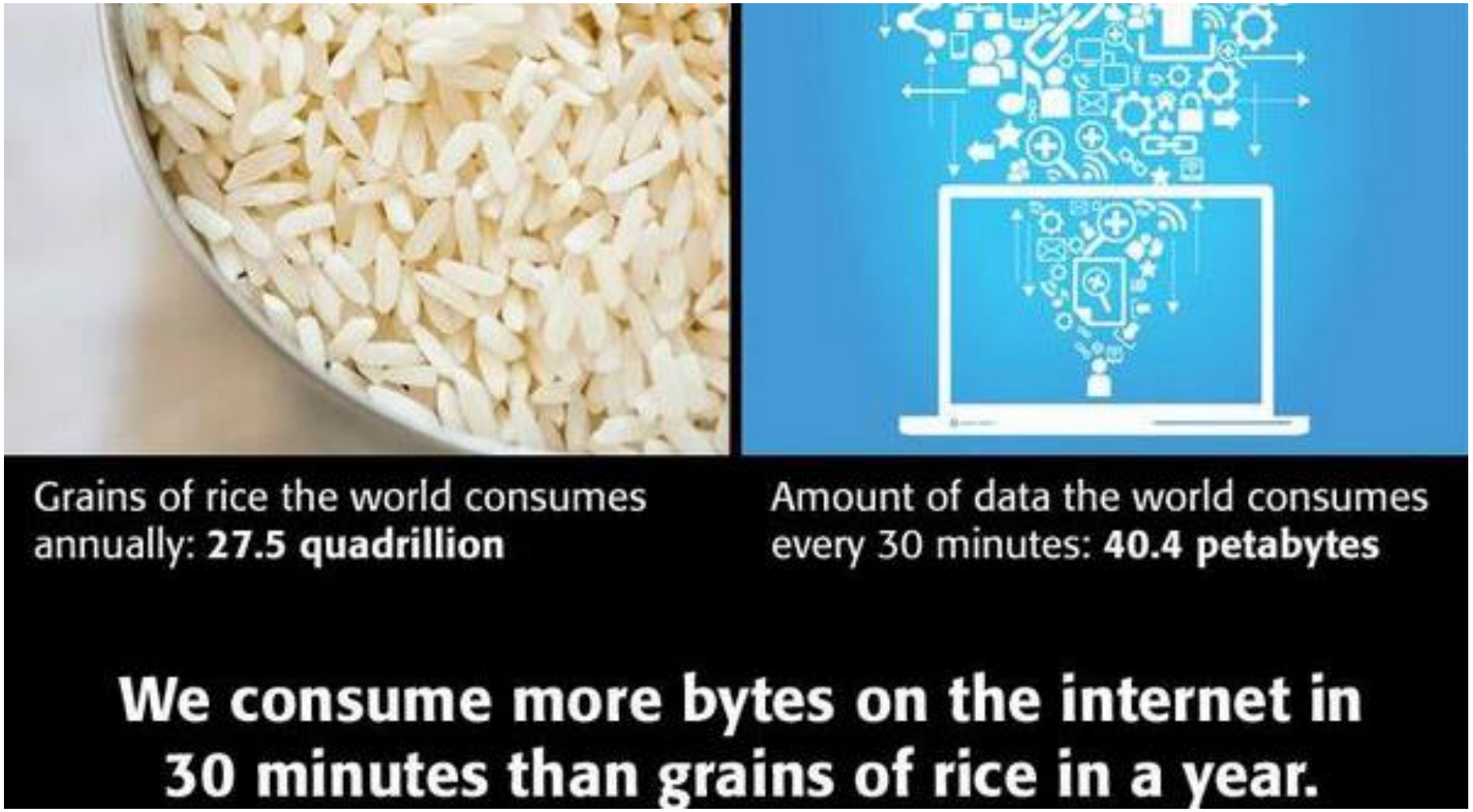
Computational power x 2 every 18 months



Moore's law:
computing power
doubles
every 18 months

Bandwidth &
Connectivity
evolve
exponentially





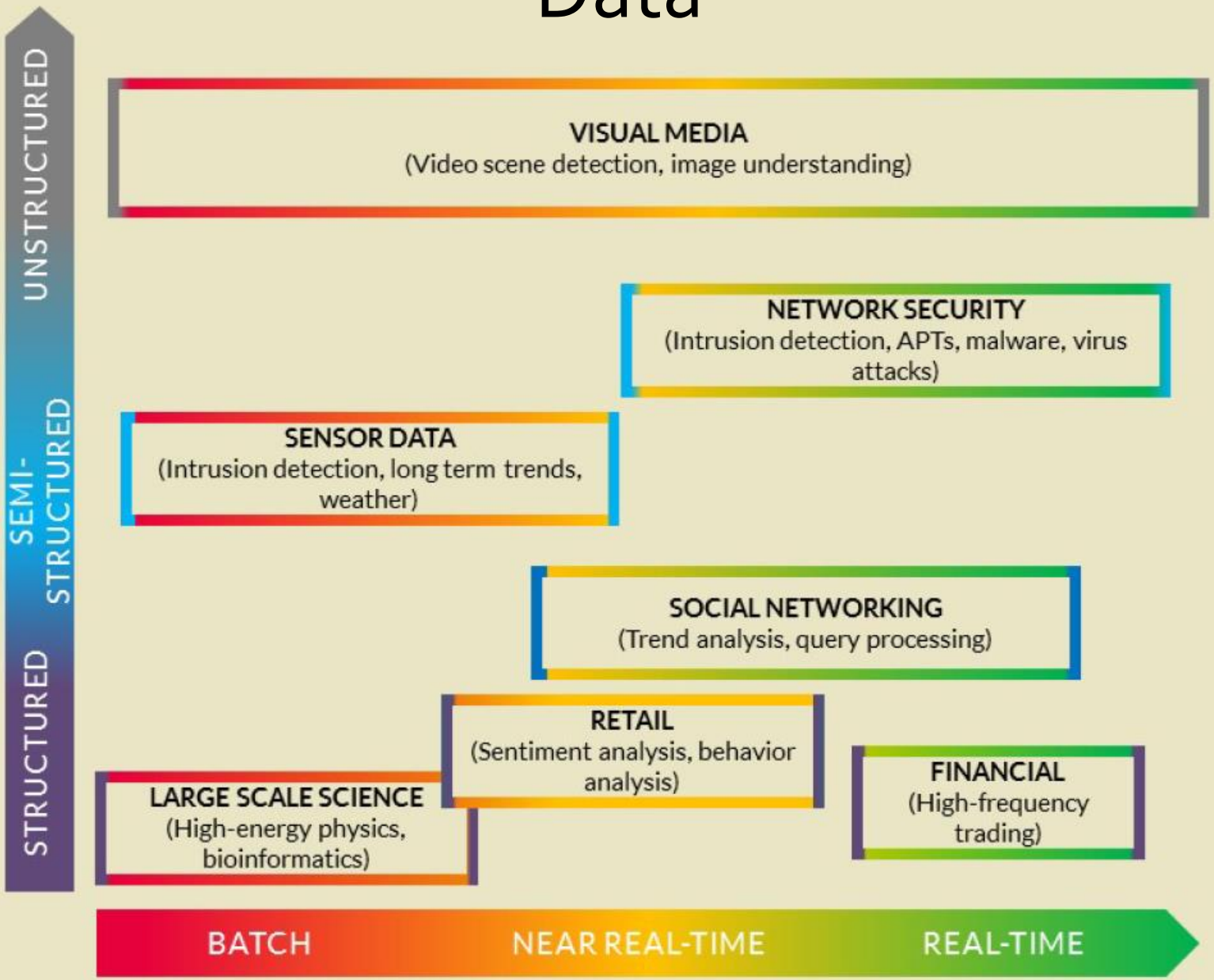
1 million = 1 000 000
1 billion = 1 000 000 000
1 trillion = 1 000 000 000 000
1 quadrillion =
1 000 000 000 000 000

1 kB = 1 000
1 MB = 1 000 000
1 GB = 1 000 000 000
1 TB = 1 000 000 000 000
1 PB = 1 000 000 000 000 000

1 TB
= large university library
= 212 DVD discs
= 1430 CDs
= 3 year music CD quality


1
Data
2
Compute Infrastructure
3
Storage Infrastructure
4
Analytics
5
Visualization
6
Security & Privacy
Big Data

Data


UNSTRUCTURED
SEMI-STRUCTURED
STRUCTURED
VISUAL MEDIA
(Video scene detection, image understanding)
NETWORK SECURITY
(Intrusion detection, APTs, malware, virus attacks)
SENSOR DATA
(Intrusion detection, long term trends, weather)
SOCIAL NETWORKING
(Trend analysis, query processing)
RETAIL
(Sentiment analysis, behavior analysis)
FINANCIAL
(High-frequency trading)
LARGE SCALE SCIENCE
(High-energy physics, bioinformatics)
BATCH
NEAR REAL-TIME
REAL-TIME

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

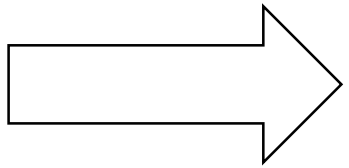
Evolution

From

I have a hypothesis

⇩

I need data to check it



To

I have data

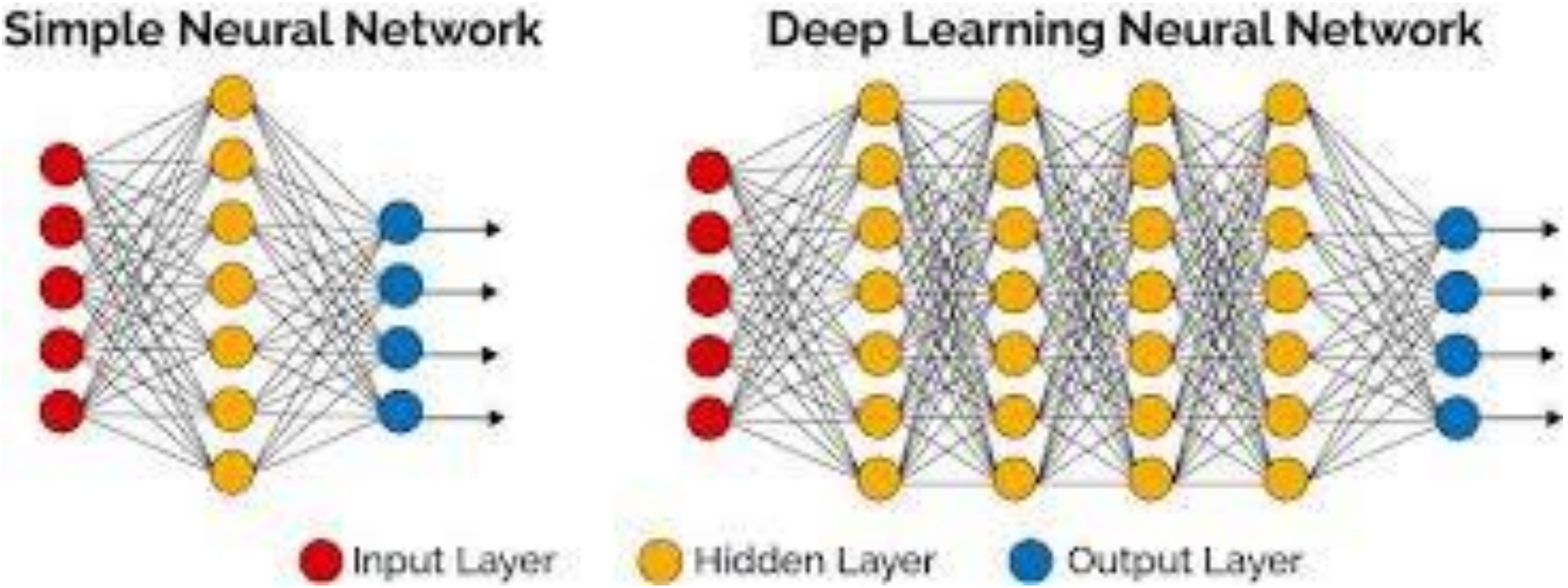
⇩

Which hypotheses can I check?

⇨ Data first!

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



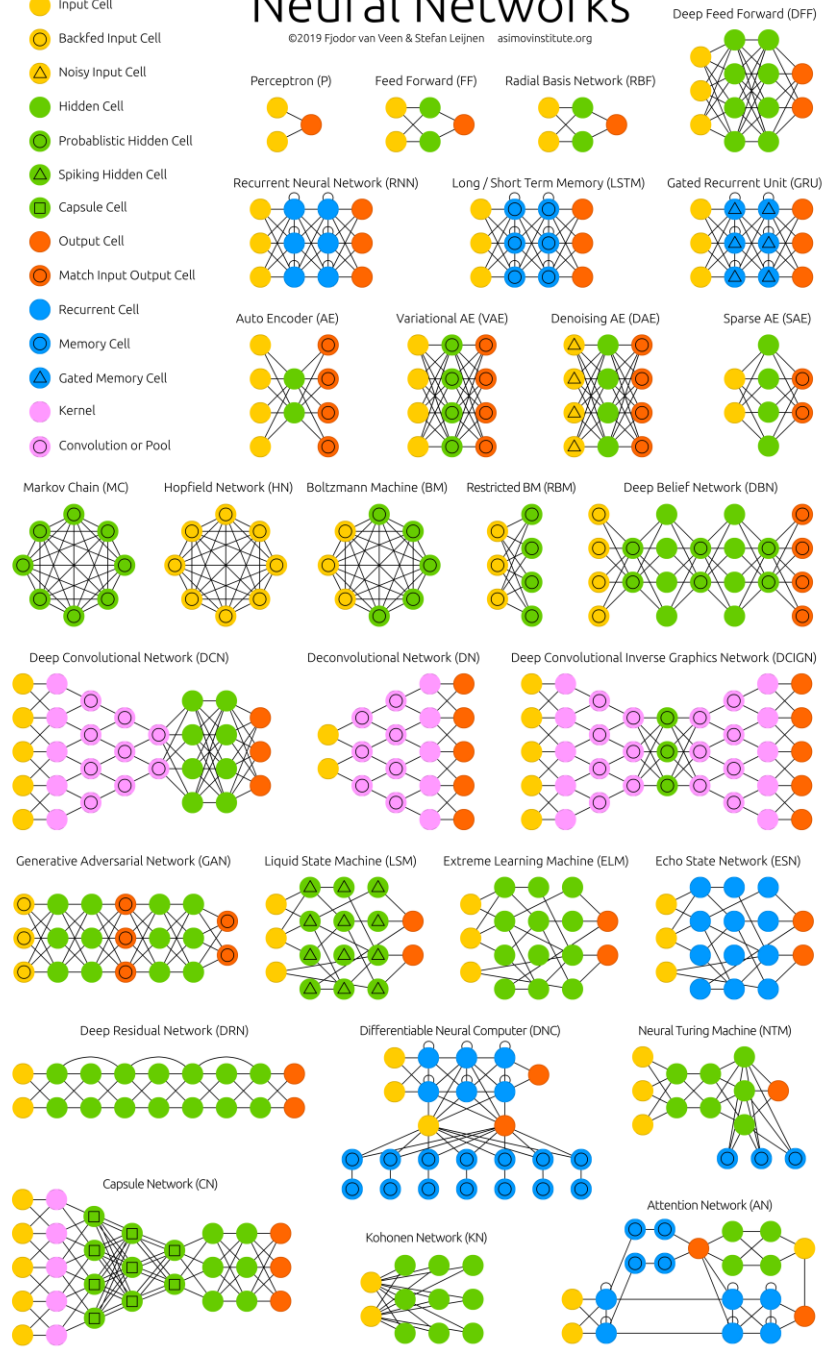
Universal approximation

A mostly complete chart of

Neural Networks

 asimovinstitute.org

- Input Cell
- Backfed Input Cell
- Noisy Input Cell
- Hidden Cell
- Probablistic Hidden Cell
- Spiking Hidden Cell
- Capsule Cell
- Output Cell
- Match Input Output Cell
- Recurrent Cell
- Memory Cell
- Gated Memory Cell
- Kernel
- Convolution or Pool



Outline

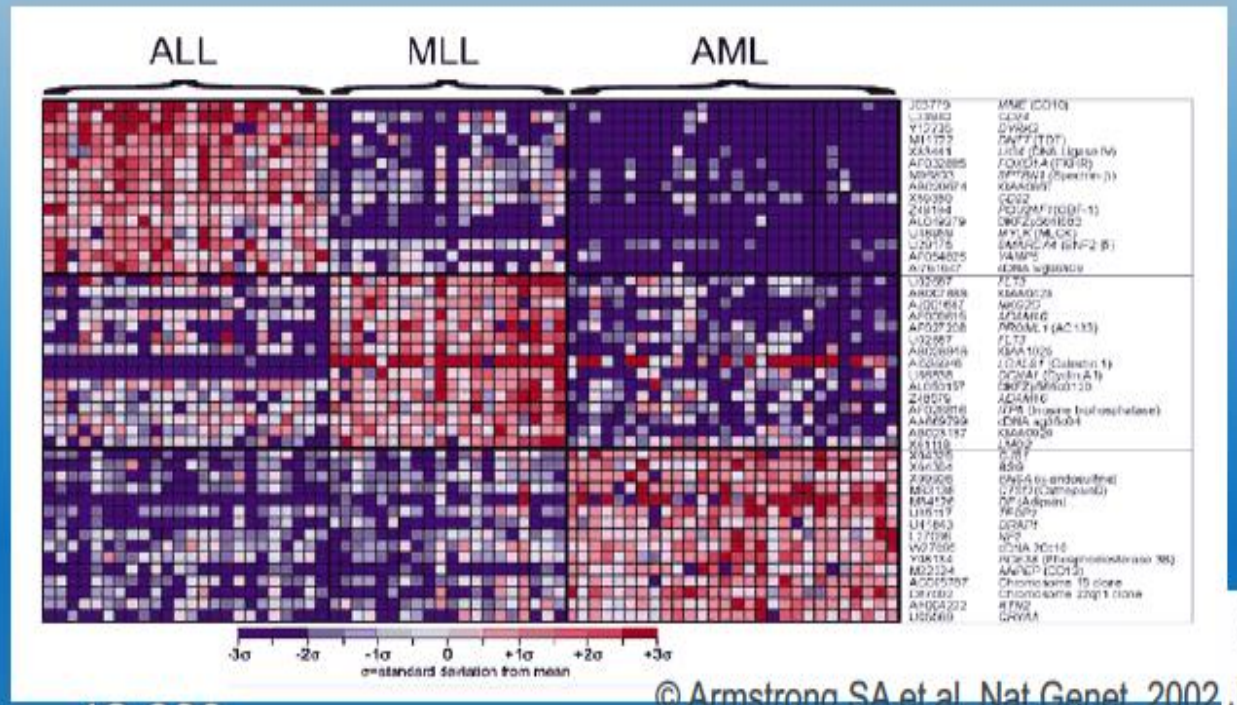
- AI: Why now ?
- Some clinical examples
- Future opportunities

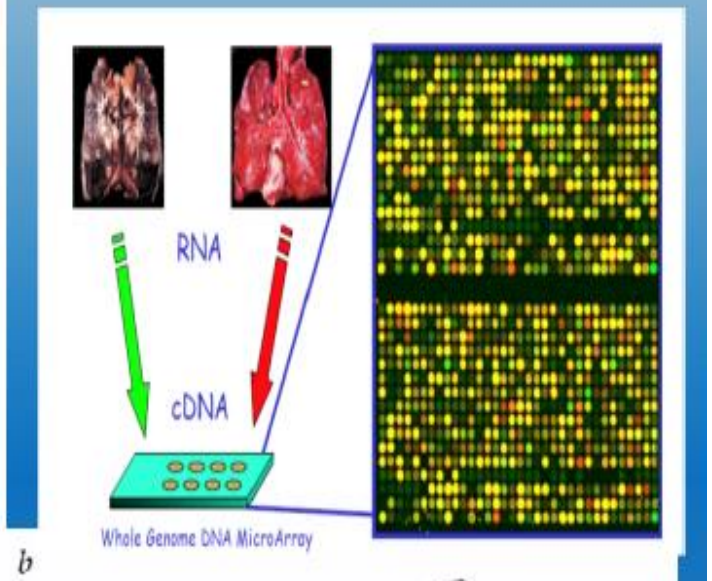
Dr. Algorithm is coming

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

– Vinod Khosla, Khosla Ventures

Example: Genomic markers for Leukemia



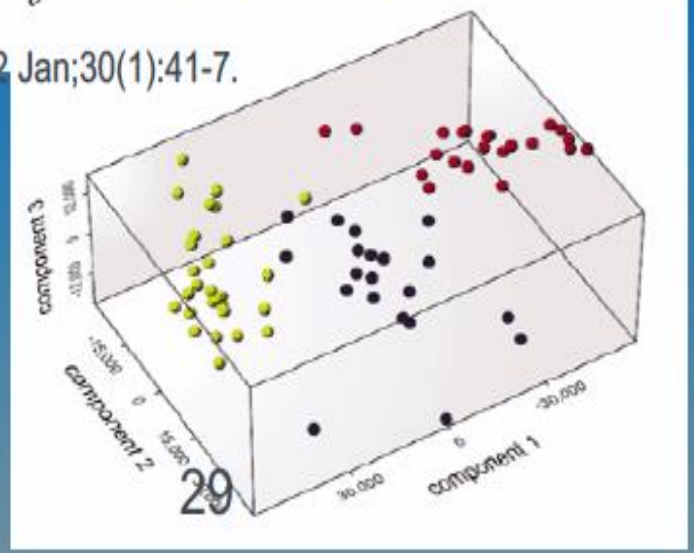


© Armstrong SA et al. Nat Genet. 2002 Jan;30(1):41-7.

12 600 genes

72 patients

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



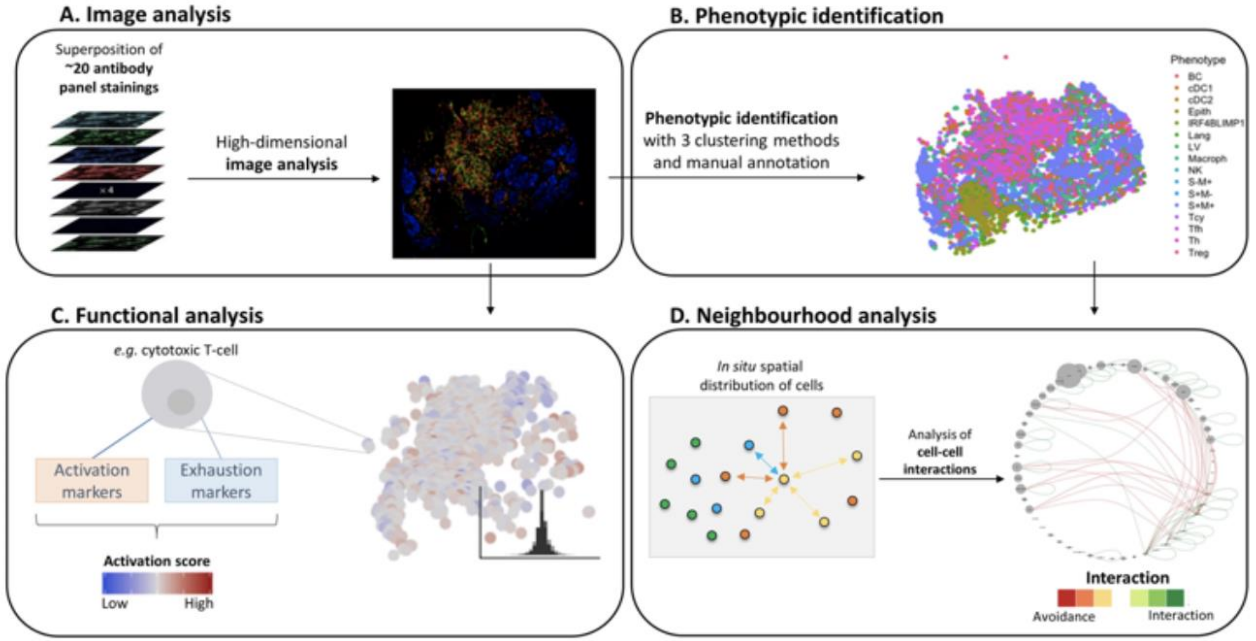
Multiple Iterative Labeling by Antibody Neodeposition **(MILAN)** method for multiplexed IHC analysis – **KOTK**

Problem: morphology-based methods that only allow the simultaneous assessment of 1-2 proteins in a single tissue slide

Aim: the generation of a reference set of 2500 highly characterized tissue samples across 9 cancer types.
the investigation of integrated tissue-based parameters (markers expression, cell types, spatial distribution, neighbourhood analysis) in order to identify two types of promising predictors of response to immunotherapy

Deliverables: A reference database for multiplex analysis of 9 cancer types and their microenvironment allowing for more precise diagnostics and a better, evidence-based design of future clinical trial projects.
Deliverable 2: A set of discriminant parameters for response to immunotherapy prediction.
Deliverable 3: Machine learning algorithms with potential for implementation in the clinics.

Project partners: Frederik De Smet, Francesca Bosisio, Oliver Bechter, Birgit Wynand, Christophe Dooms, Thomas Tousseyn, Giuseppe Floris, Hans Wildiers, Ines Nevelsteen, Raf Sciot, Paul Clement, Steven De Vleeschouwer, Tania Roskams, Chris Verslype, Frederik Nevens, Sabine Tejpar, Xavier Sagaert, Evelyne Lerut, Benoit Beuselinck, Bart De Moor, Daan Dierickx, Ann Smeets, Eric Verbeken, Esther Hauben



Direct detection and identification of neuro-oncology markers in brain tumours (TBM project)

Problem: The discrimination between tumour and normal, healthy tissue is of critical importance to achieve optimal surgical results and an improved prognosis, in any type of oncological surgery but in particular in brain tumour surgery

Aim: Use Mass Spectrometry Imaging (MSI) data and REIMS (surgical iKnife) to assess tumour delineation

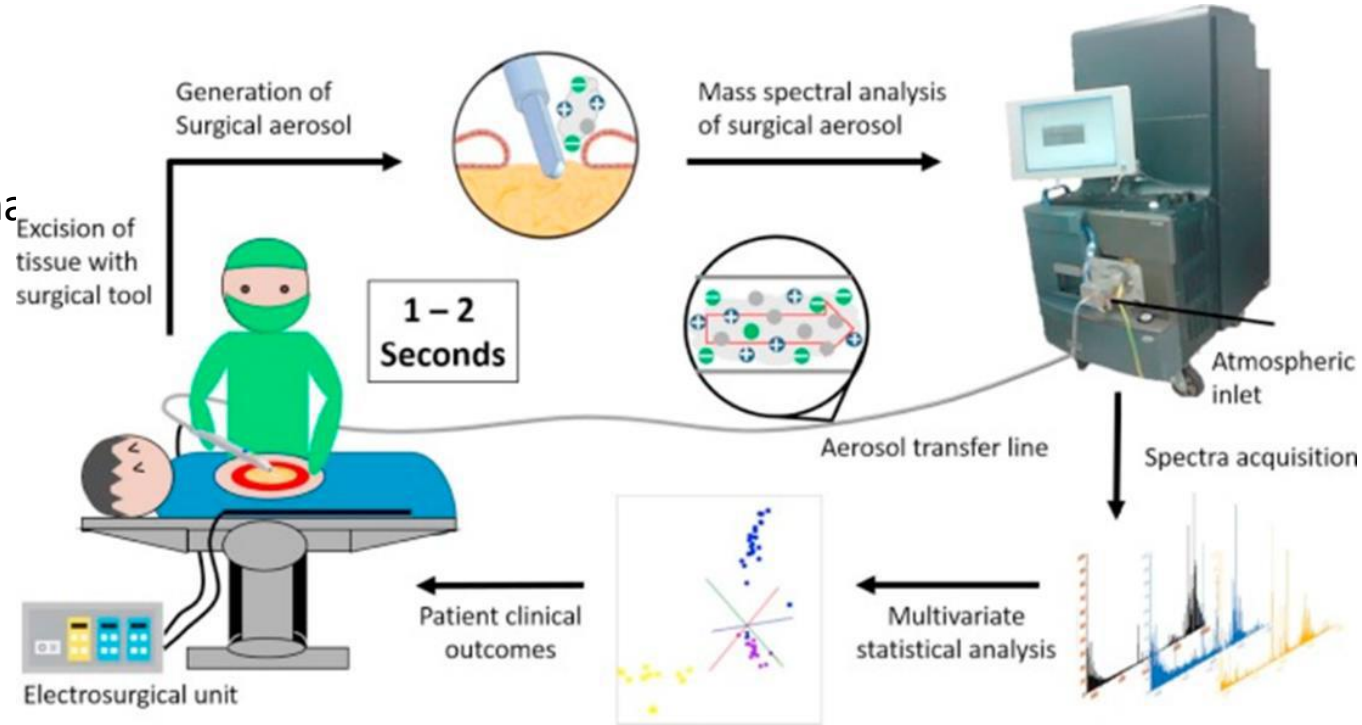
Build a database containing typical profiles of healthy and brain tumour tissue

Identify principal components for the differentiation between healthy brain and brain tumour tissue

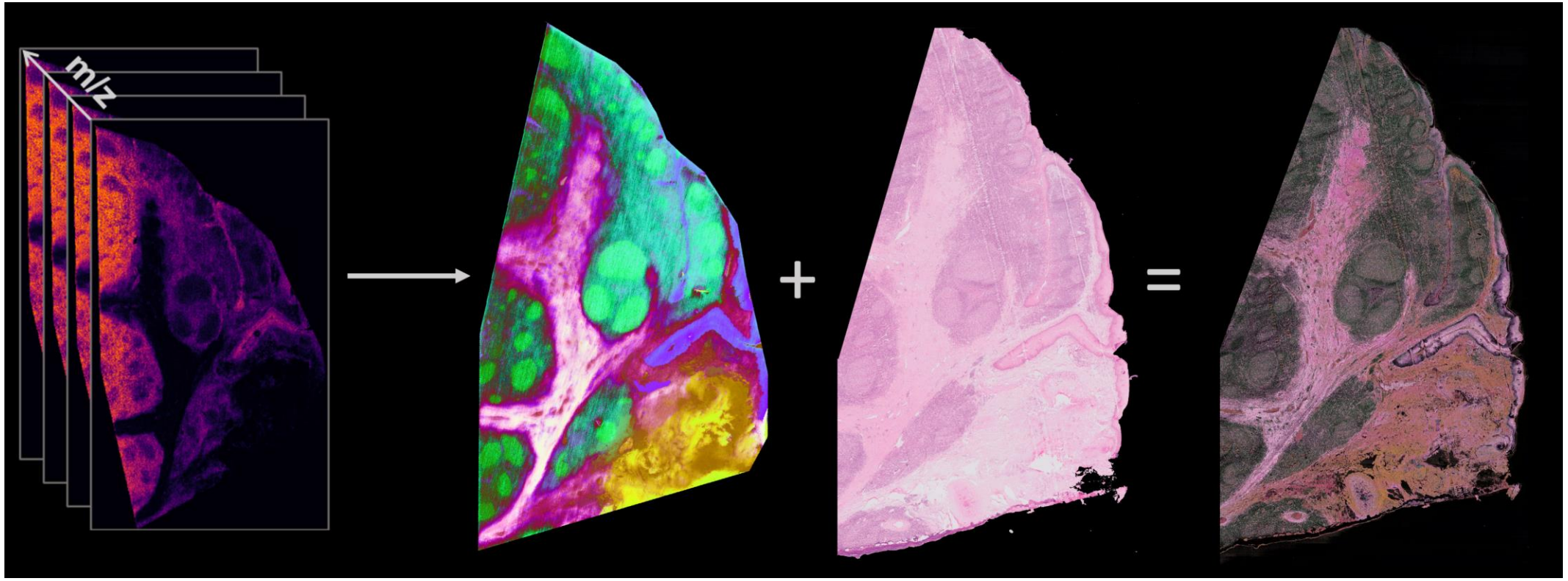
Build pattern recognition program

Deliverables: software system connected to database tha warns the neurosurgeon when approaching normal or benign tissue such that neurosurgeon knows better where tumour boundaries are located

Partners: prof. De Vleeschouwer, prof. De Moor, prof. Cuypers, prof. Sciot, prof. De Smet

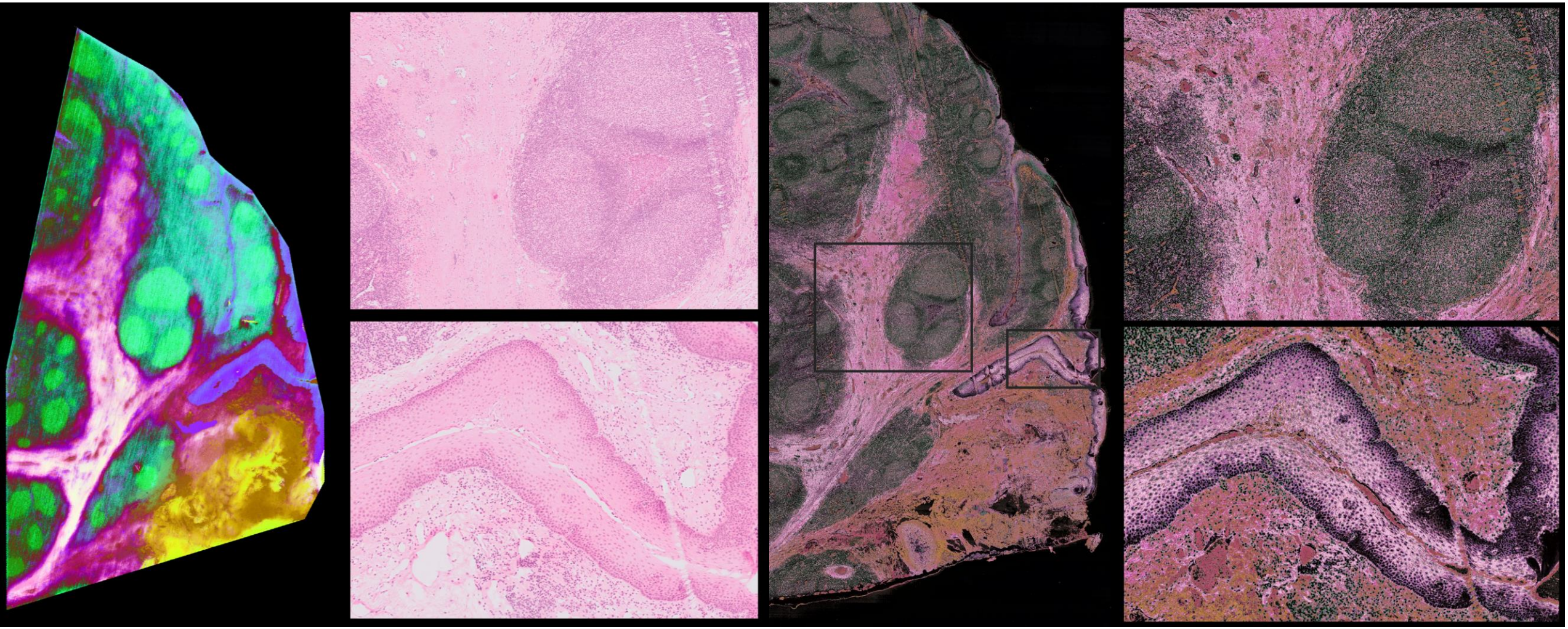


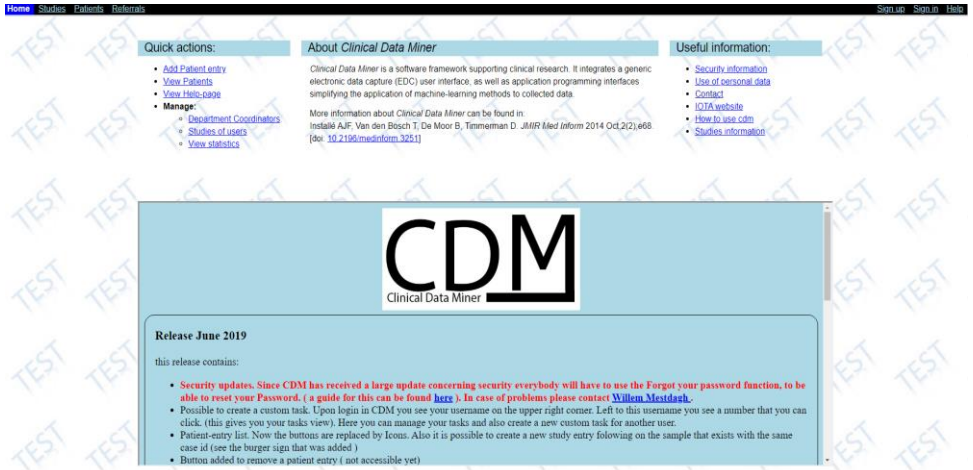
Example application: **Mass Spectrometry Imaging and H&E data fusion**



MSI dataset:
500 000 pixels x 8000 m/z
30 Gb

Example application: **Mass Spectrometry Imaging and H&E data fusion**





Capturing study data of:

- International Ovarian Tumor Analysis Studies (IOTA)
- International Endometrial Tumor Analysis Studies (IETA)
- International Deep Endometriosis Analysis (IDEA)

Collaboration:

- Prof. Timmerman Dirk (UZ-Leuven)
- Prof. De Moor Bart (ESAT / STADIUS)

Electronic Data Capture system for clinical studies

Cleaned Data

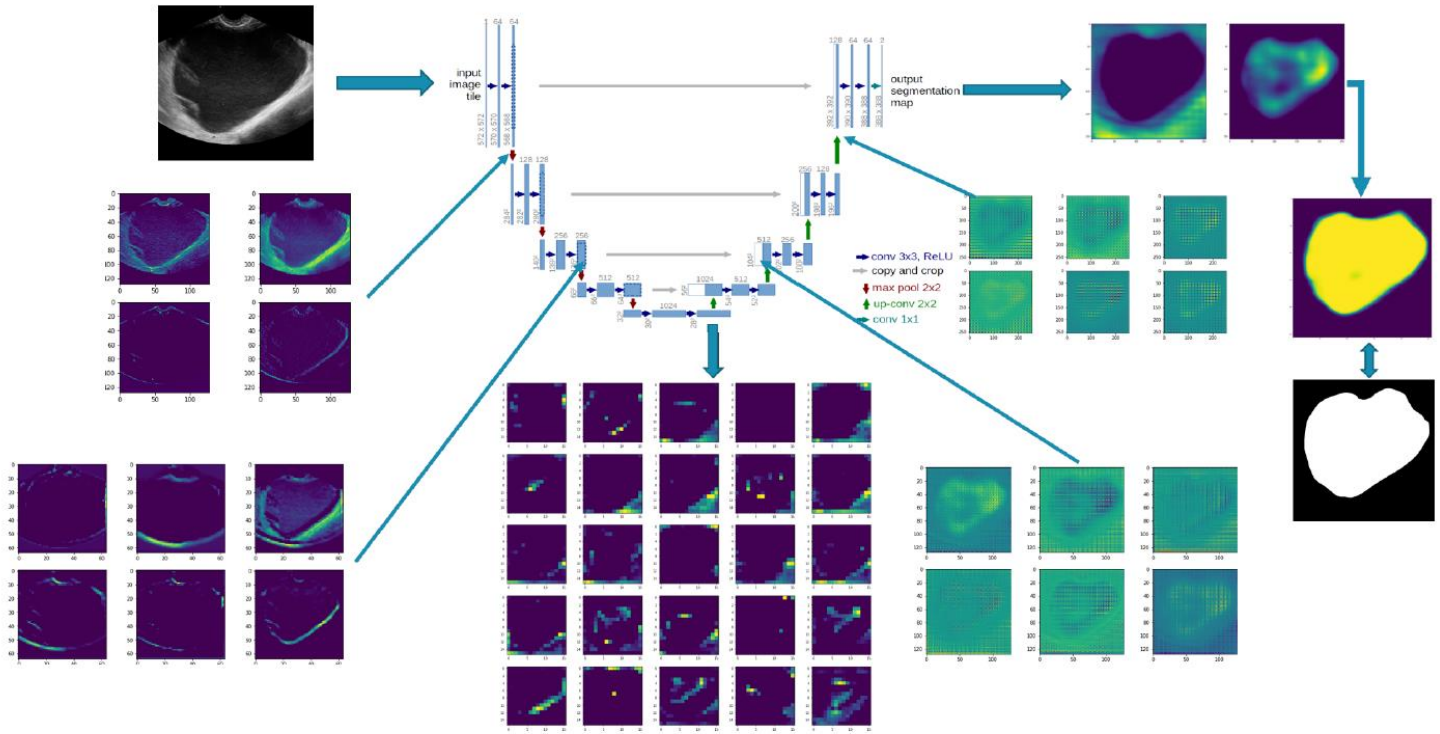
Integrated quality checks on data

Context of study items explanation made possible by the use of pictograms.

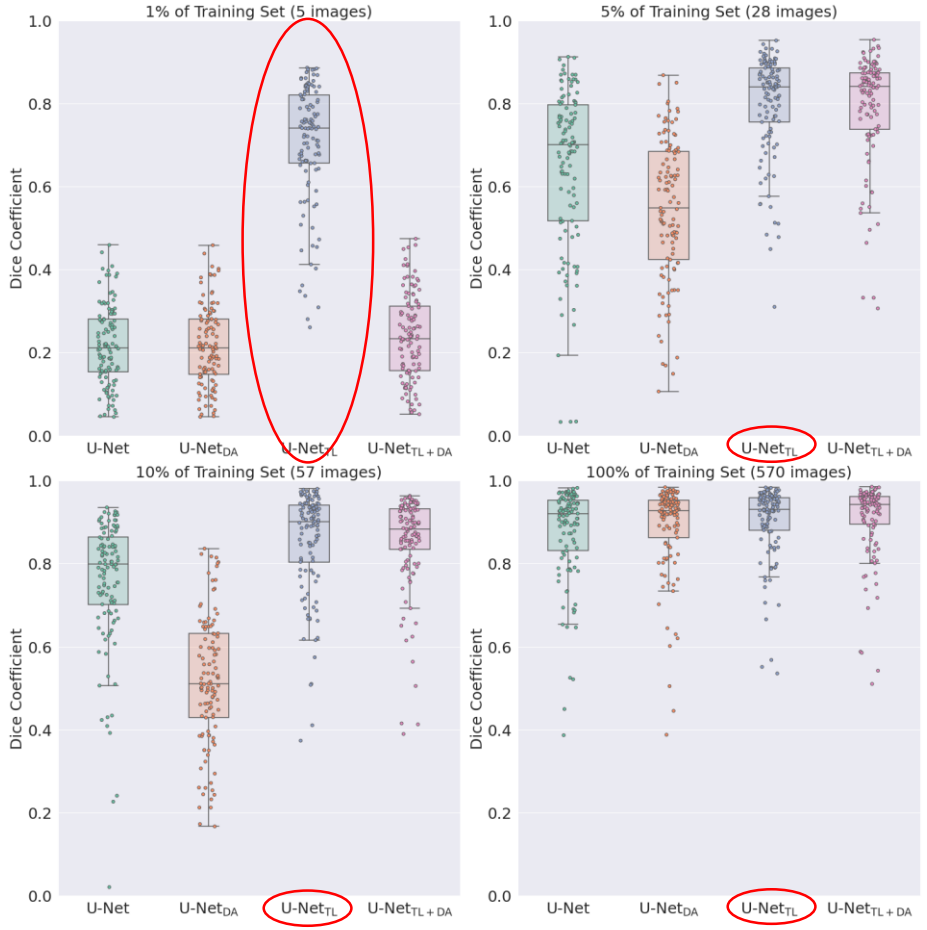
Cloud solution, international accessible

GDPR compliant

Studies	#samples	#centers
IOTA7	>6500	>50
IETA	>6000	28
IDEA	>300	11



Automated segmentation of ovarian ultrasound 2D images via modified U-net[1] convolutional neural network. The top left figure is an ovarian ultrasound image and the most right lower figure is the segmentation prediction from our learned model. The figure shows the intermediate layers' outputs from our model, in order to provide more intuition and interpretation in 'black box' deep learning model, which is very important in medical field.



The potential of exploring transfer learning in medical field. After applying a pre-trained neural network (based on ImageNet[2] dataset) in our original model, the result shows a big potential that the required number of training data for deep learning model could be reduced with the help of utilizing transfer learning.

[1]Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.
[2]Deng, Jia, et al. "Imagenet: A large-scale hierarchical image database." *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009.

Forecasting kidney allograft function

Project cooperator: UZ Leuven

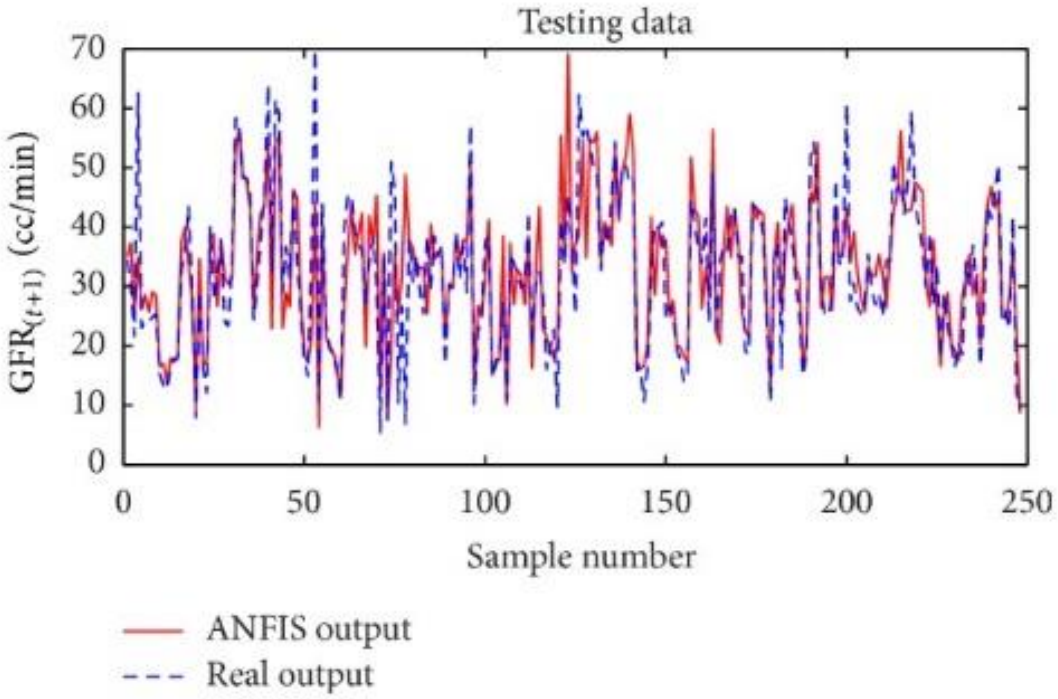
Time: 10/2018 – current (Ongoing)

Aim of Project: Forecast the kidney allograft function after kidney transplantation based on patients' sequential kidney function values

Impact: provide a renal allograft function forecasting model for clinicians to help guide indication for biopsies in transplant recipients in the early stage after kidney transplantation

Related research:

'Predicting Renal Failure Progression in Chronic Kidney Disease Using Integrated Intelligent Fuzzy Expert System' [1]

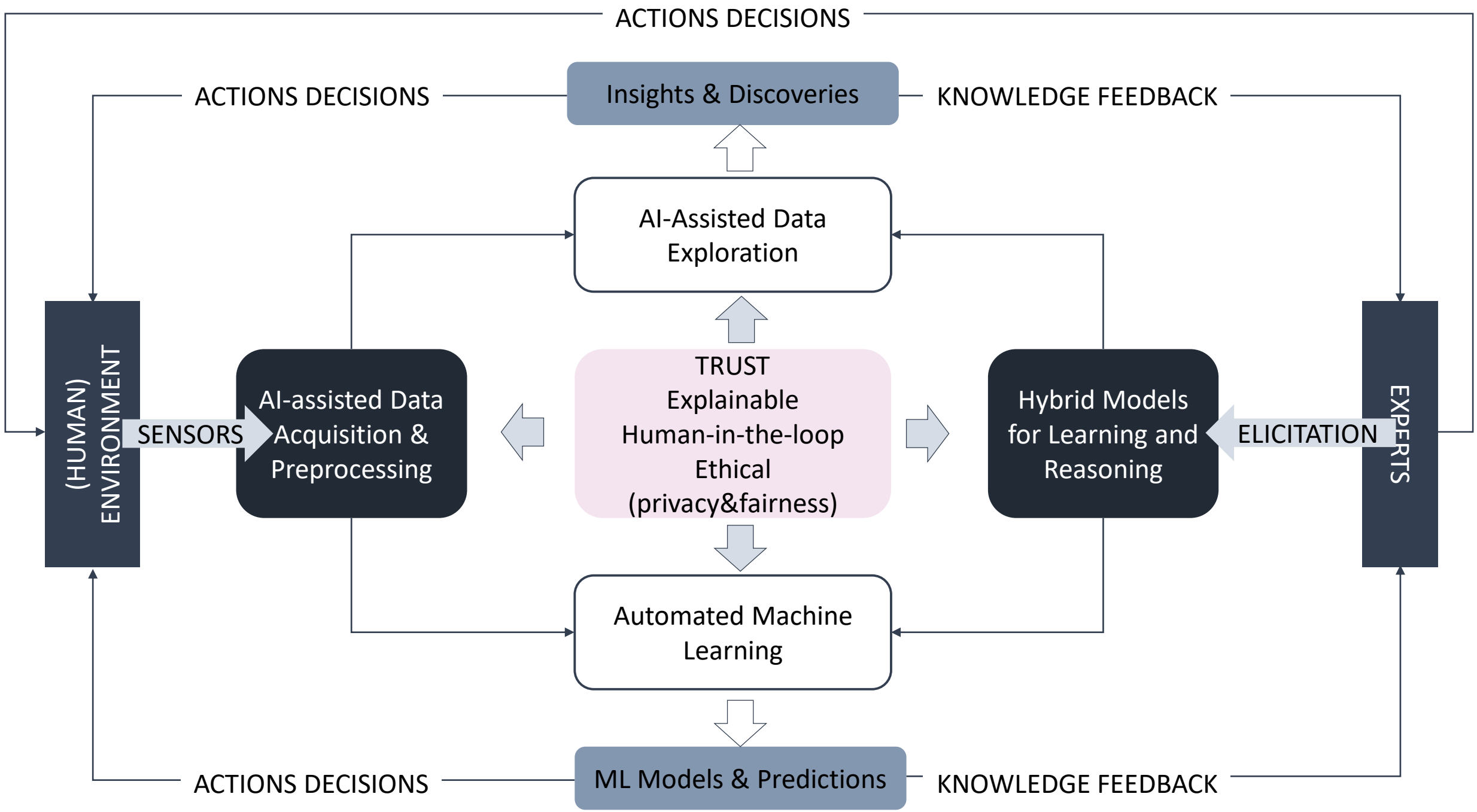


Comparison of the ANFIS prediction and real $GFR_{(t+1)}$ values for the test dataset at 6-month interval [1].

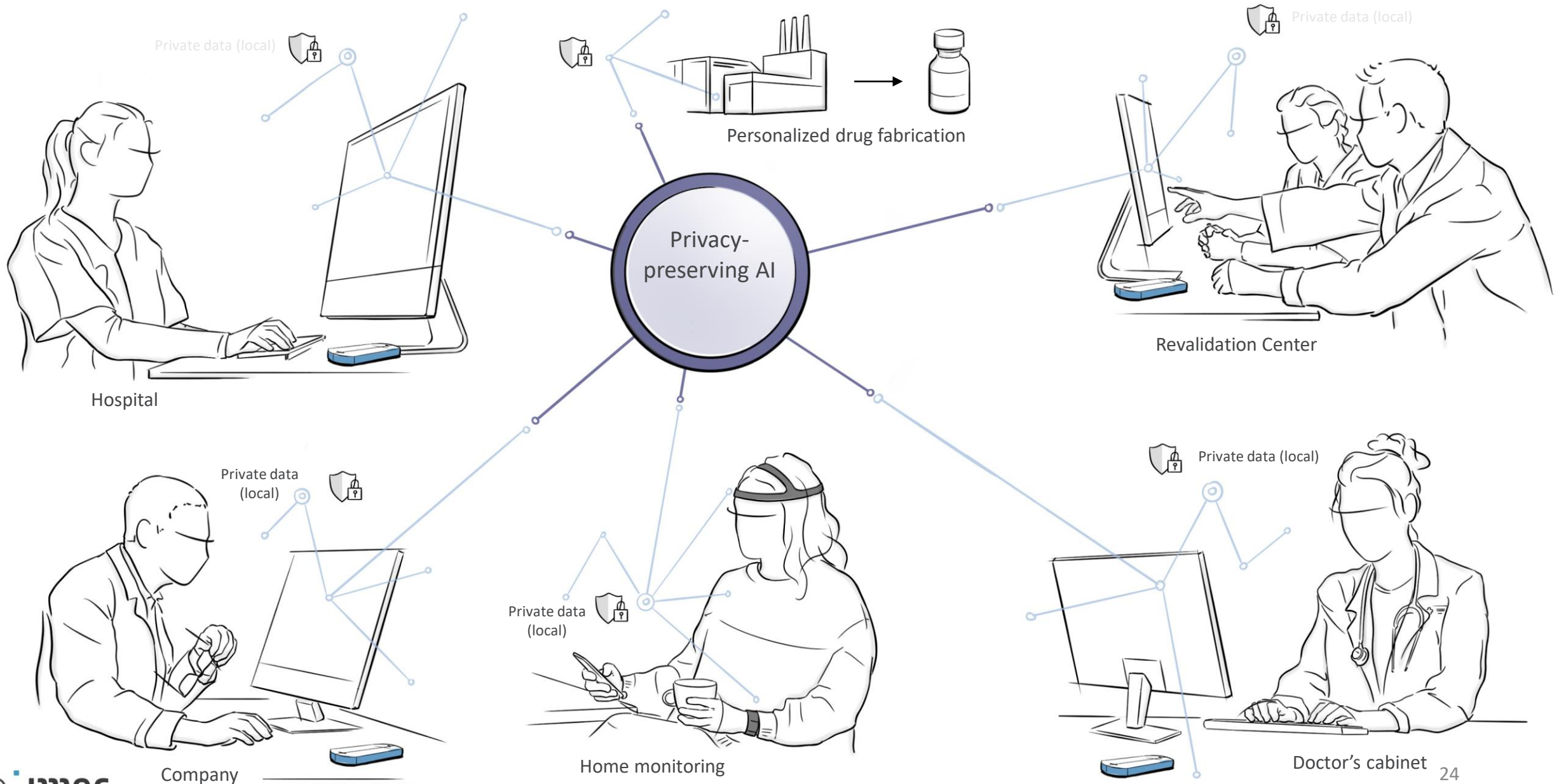
[1] Norouzi, Jamshid, et al. "Predicting renal failure progression in chronic kidney disease using integrated intelligent fuzzy expert system." *Computational and mathematical methods in medicine* 2016 (2016).

Outline

- AI: Why now ?
- Some clinical examples
- Future opportunities


ACTIONS DECISIONS
ACTIONS DECISIONS
Insights & Discoveries
KNOWLEDGE FEEDBACK
AI-Assisted Data Exploration
(HUMAN) ENVIRONMENT
SENSORS
AI-assisted Data Acquisition & Preprocessing
TRUST
Explainable
Human-in-the-loop
Ethical
(privacy&fairness)
Hybrid Models for Learning and Reasoning
ELICITATION
EXPERTS
Automated Machine Learning
ACTIONS DECISIONS
ML Models & Predictions
KNOWLEDGE FEEDBACK

Data interaction on a need to know basis - privacy preserving machine learning



Flanders AI Impulse Program

Program Structure with 3 pillars, funded by the Flemish Government

		Yearly budget
1	FLANDERS AI RESEARCH PROGRAM	12 Mio €
2	FLANDERS AI IMPLEMENTATION PROGRAM	15 Mio €
3	FLANDERS AI SUPPORTING ACTIVITIES: ETHICS, EDUCATION AND TRAINING	5 Mio €

Funding for R&D Projects
Digital transformation trajectories in companies,...

Knowledge Centre Data & Society

Flanders AI Academy

Our consortium in Numbers

5	Universities
5	(Strategic) Research Centers
40	Research Teams
89	Professors
500	PhD students

100+	R&D projects with funding of Flemish government
200+	Companies in Collaborative funded R&D projects
18	“PhD interns” in companies in 2020
400+	Publications in peer-reviewed journals
40+	European Funded Projects

Challenge-Based Research with Demand-Driven Impact

CHALLENGE BASED RESEARCH

WITH DEMAND-DRIVEN IMPACT

AI-DRIVEN DATA SCIENCE

AI IN THE EDGE

MULTI-AGENT COLLABORATIVE AI

HUMAN-LIKE AI

PROOFS-OF-CONCEPTS (Demonstrators)

FOCUS DOMAINS and USE CASES

Management Flanders AI Research Program

Sabine Demey

Director Flanders AI Research Program

imec

AI-driven Data Science

KU LEUVEN

Prof. Bart De Moor

ESAT, KULeuven

UNIVERSITEIT GENT

Prof. Piet Demeester

IDLab, Ghent University-imec

Multi-agent Collaborative AI

Ann Nowé

Professor AI Lab, VUB

VUB VRIJE UNIVERSITEIT BRUSSEL

AI in the Edge

Mieke De Ketelaere

Program Director AI, imec

imec

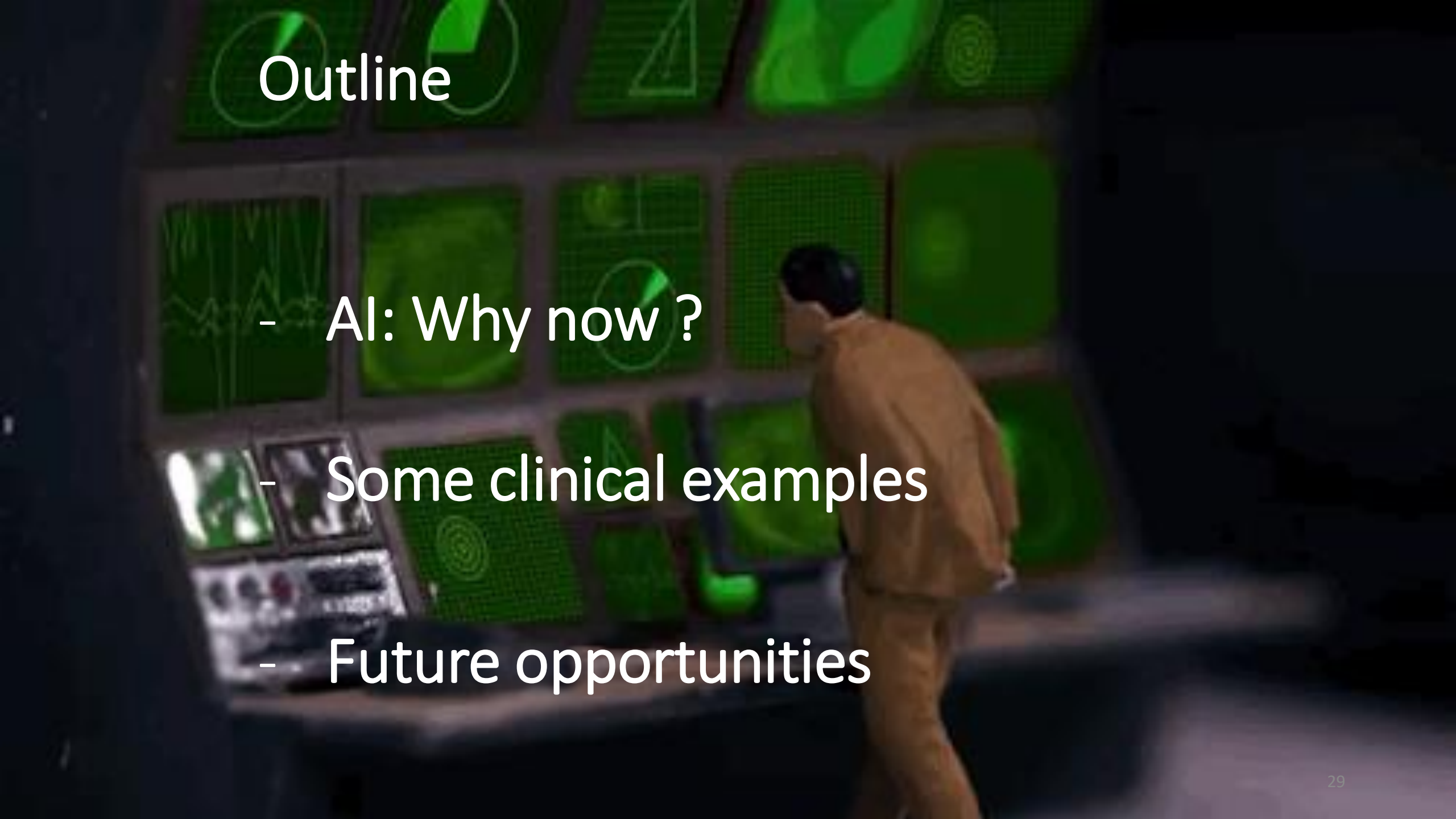
CONFIDENTIAL

Human-Like AI

Prof. Steven Latré

IDLab, University of Antwerp - imec

Universiteit Antwerpen



Outline

- AI: Why now ?
- Some clinical examples
- Future opportunities