

# ELMED219 - 2022

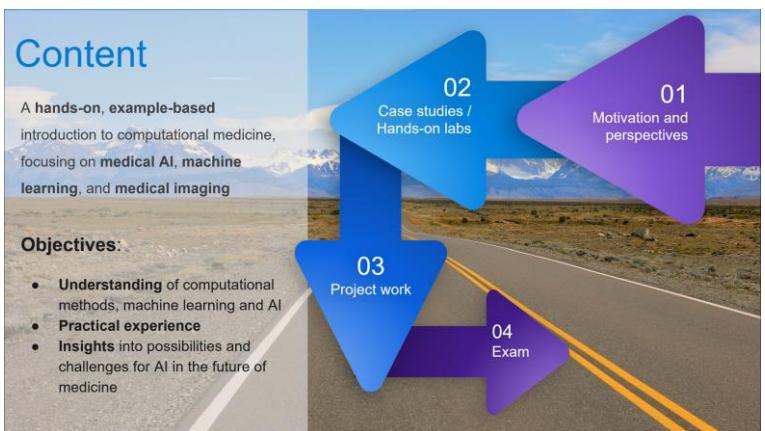
Fri Jan 21st, 10:15-12

Perspectives on medical AI: innovation,  
education, epistemology, ethics and impact

# Objectives

- *A very quick overview of what we've discussed so far in ELMED219*
- *A look at some challenges and limitations of current AI systems*
- *Also*  
*Innovation, education, epistemology, ethics and impact*

$$y \approx f(x; \theta)$$

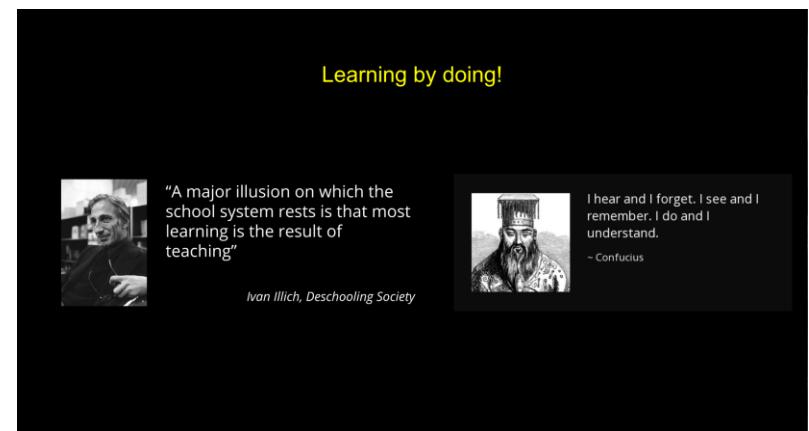


## ELMED219

Will learn about medical AI and computational medicine through labs and corresponding lectures

- Theory and tools from **machine learning** and **deep learning**
- Natural language processing** and analysis of **medical health records**
- Image-based diagnostics** and **MRI**. **Brain tumors** and **MRI**
- Perspectives** for medical AI and computational medicine

introduction, "hands-on", state-of-the-art



MMIV-ML / ELMED219-2022 Public

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

2 stars 1 branch 0 tags Go to file Add file + Code v About

Material for the 2022 version of the course ELMED219, UBI and HVL

169 commits

Lab01-ML: link to vids added 2 days ago

Lab02-MRI (optional): rename and cleaning of Lab02 16 days ago

Lab1-EH: colab 6 days ago

Lab2-NLP: colab message 6 days ago

Lab3-DL: lab3 vid added 6 days ago

Lab4-BRATS: links to slides and recording 2 days ago

assets: pdf slides 2 days ago

project: Update README.md 6 days ago

testdata: Instructions + tests for local installs added 16 days ago

using MONAI Segformer for BRATS 3 days ago

ignore: Initial commit 3 months ago

LICENSE: vid added 6 days ago

README.md: Update environment-img.yml 13 days ago

environment-img.yml: tweaks 16 days ago

environment.yml: tweaks 16 days ago

setup-img.yml: Update setup-img.md 16 days ago

setup.md: tweaks 16 days ago

Languages

### How do you do machine learning these days?

Machine learning: Python, pandas, scikit-learn

Deep learning: Python, PyTorch, TensorFlow

Processing / compute: Cloud computing diagram, ANACONDA

- Why Python ?**

Teaching biologists the tools they need to use computers to do cool science

<http://www.programmingforbiologists.org/about/why-python> <https://www.upgrad.com/blog/reasons-why-python-popular-with-developers>

- Why Jupyter notebooks ?**

Jupyter notebooks provide an environment where you can freely combine human-readable narrative with computer-readable code.

<https://www.nature.com/articles/tb2020002> [\[ link \]](#)

- Why GitHub ?**

Github is like facebook for programmers. Everyone's on there. You can look at what they're working on and easily peruse their code and make suggestions or changes.

It's really open source. "Open source" is not so open if you can't easily study it. With github, all of the code is easily inspected, as is its entire history.

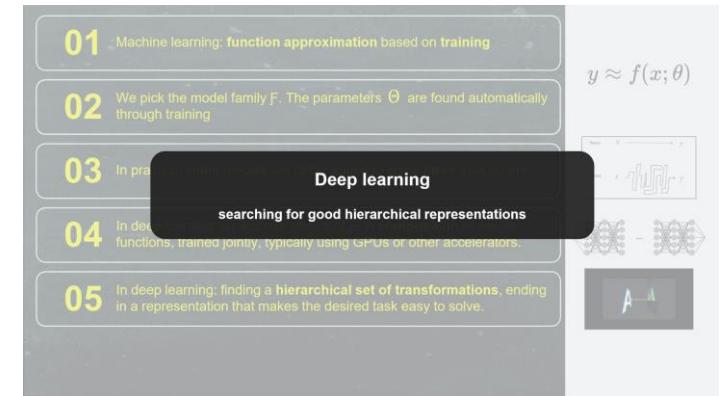
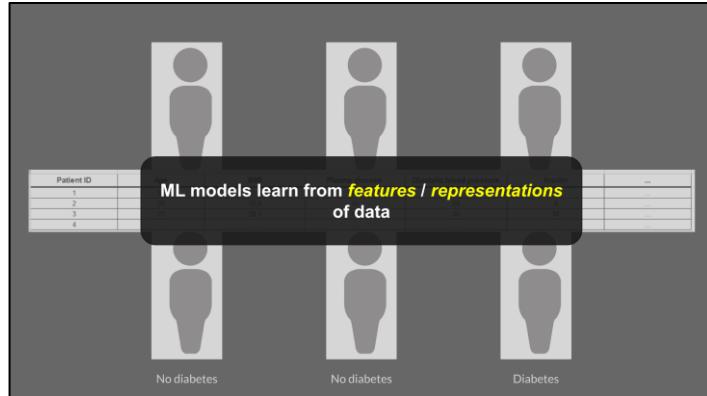
Github lowers the barriers to collaboration. [\[ link \]](#)

Ten simple rules for writing and sharing computational analyses in Jupyter Notebooks [\[ link \]](#)

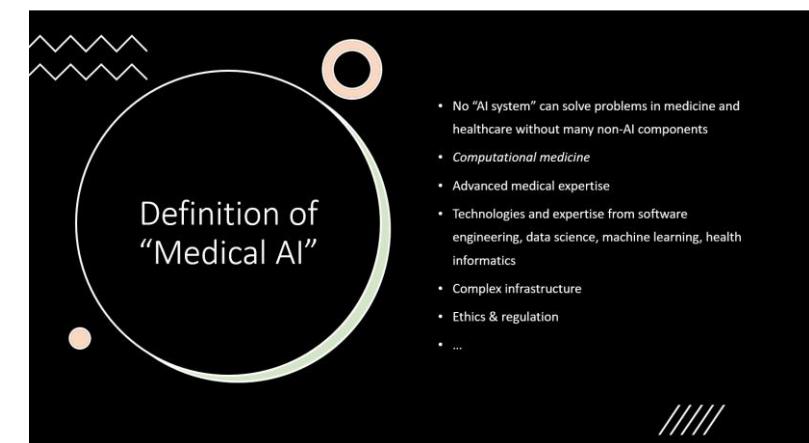
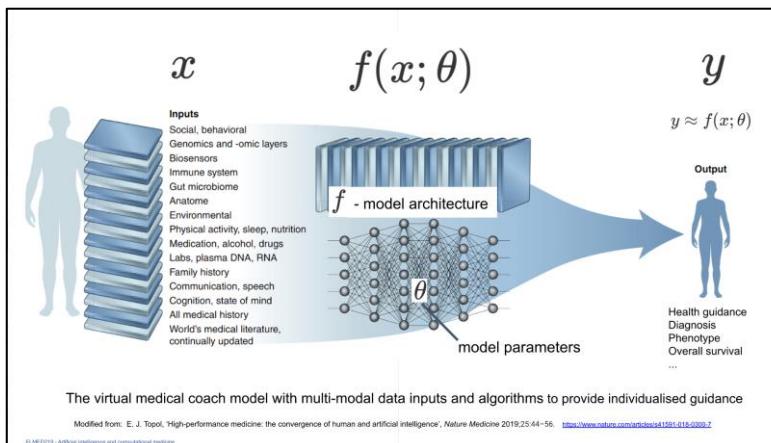


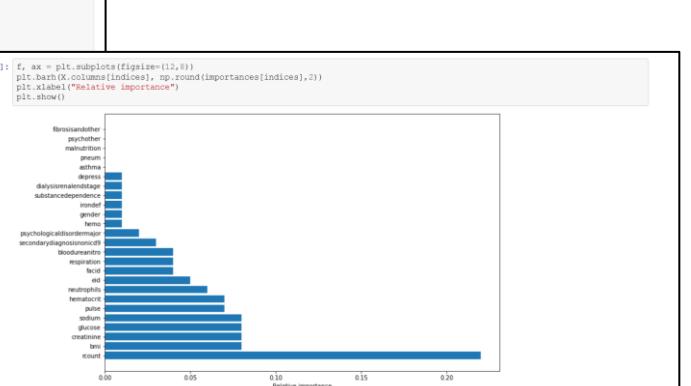
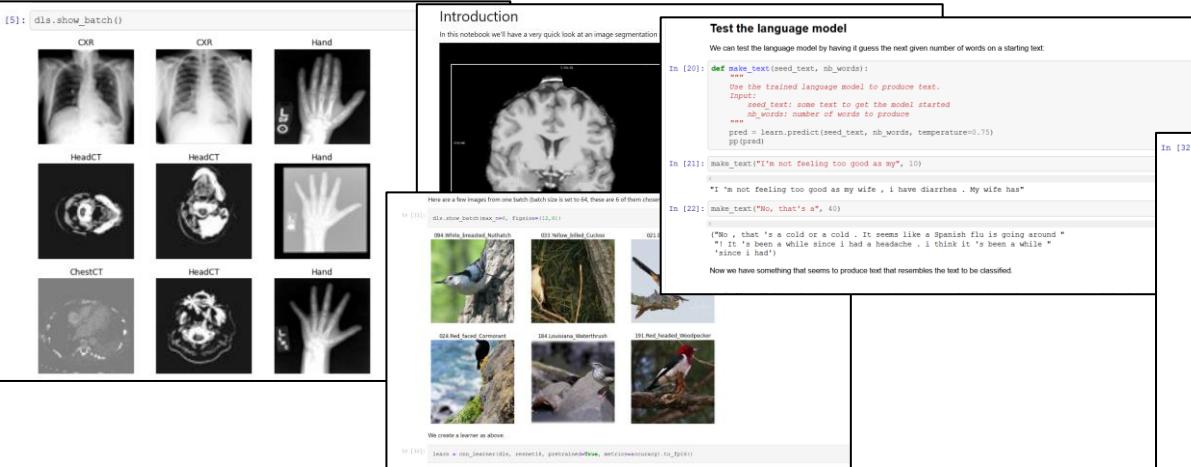
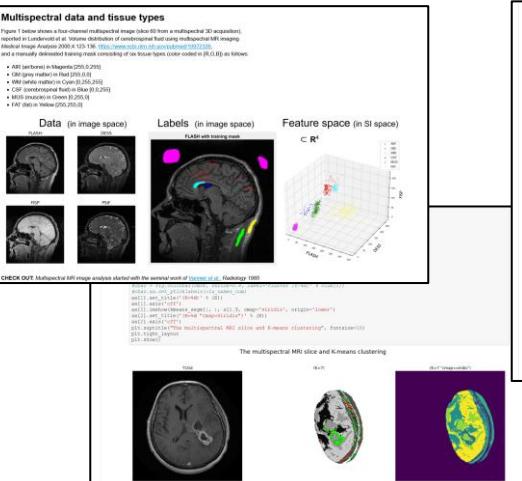
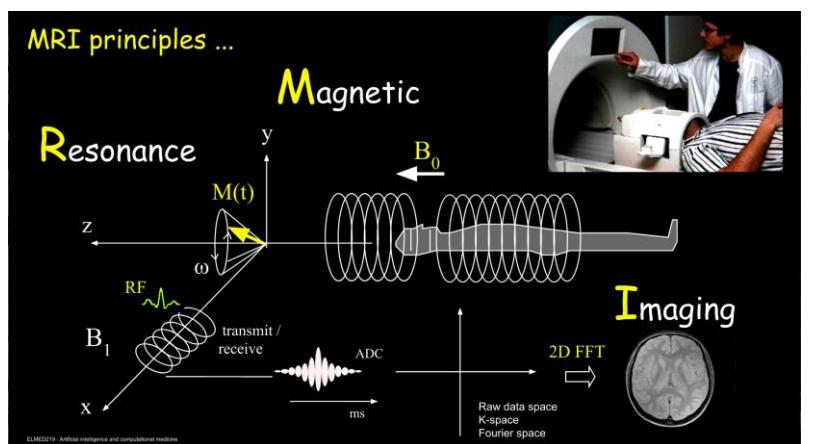
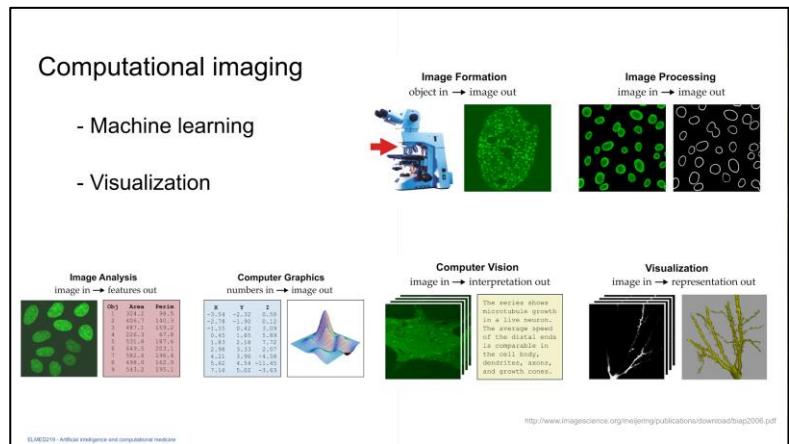
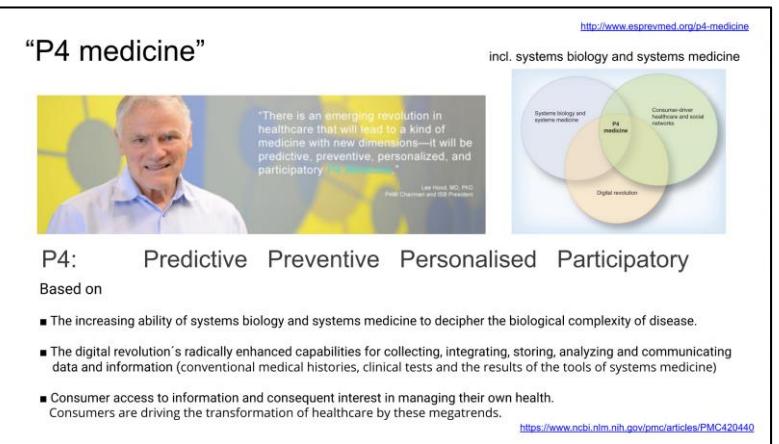
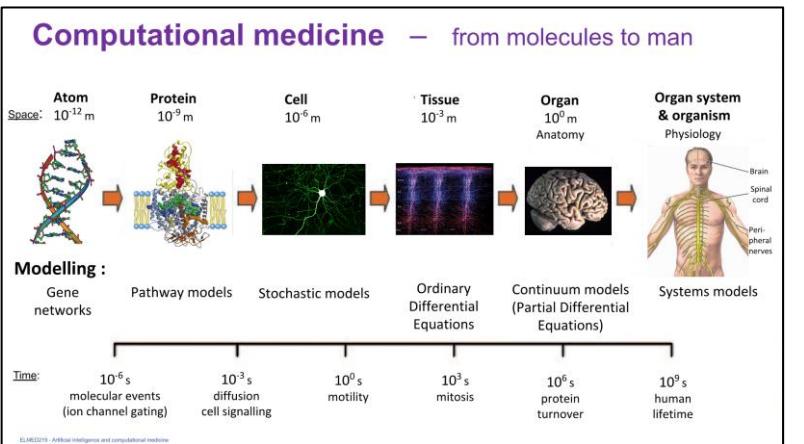
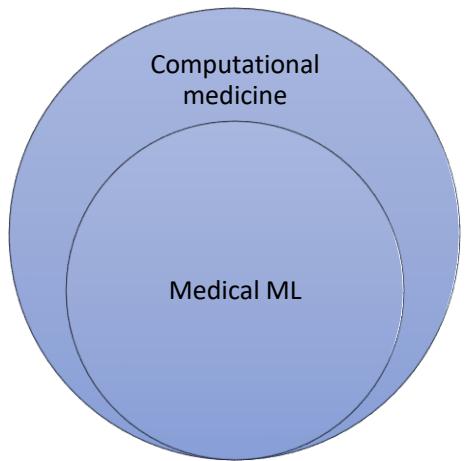
## Machine learning

Two (very) simple examples



## AI in medicine





# Content

A **hands-on, example-based** introduction to computational medicine, focusing on **medical AI, machine learning, and medical imaging**

## Objectives:

- **Understanding** of computational methods, machine learning and AI
- **Practical experience**
- **Insights** into possibilities and challenges for AI in the future of medicine

02

Case studies /  
Hands-on labs

01

Motivation and  
perspectives

03

Project work

04

Exam

# Content

A hands-on, example-based introduction to computational medicine, focusing on **medical AI**, **machine learning**, and **medical imaging**

## Objectives:

- **Understanding** of computational methods, machine learning and AI
- **Practical experience**
- **Insights** into possibilities and challenges for AI in the future of medicine

02

Case studies /  
Hands-on labs

01

Motivation and  
perspectives

03

Project work

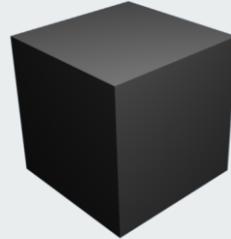
04

Exam

Some challenges for medical AI, especially  
systems based on deep learning



"Explainable AI"



Data hungry!



Privacy Data access  
Data labeling Protecting data  
Fairness

Lack of theory



What makes one model better than another?

Software engineering best practices

Deployment, maintenance, debugging, hidden technical debt, ...

Black box

De-anonymization-angrep

Model inversion  
Membership attack

And more...

Differential privacy

Models lack of understanding

Trust

Clinical validation!

Uncertainty

Education

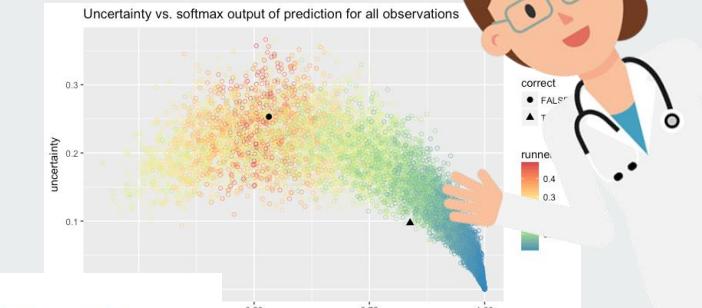
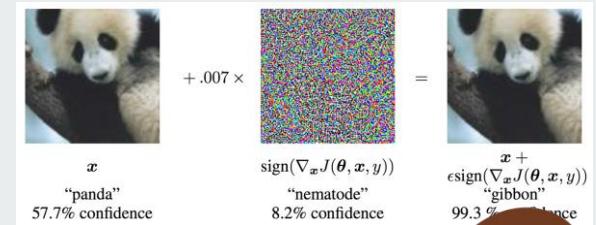
Workflow integration



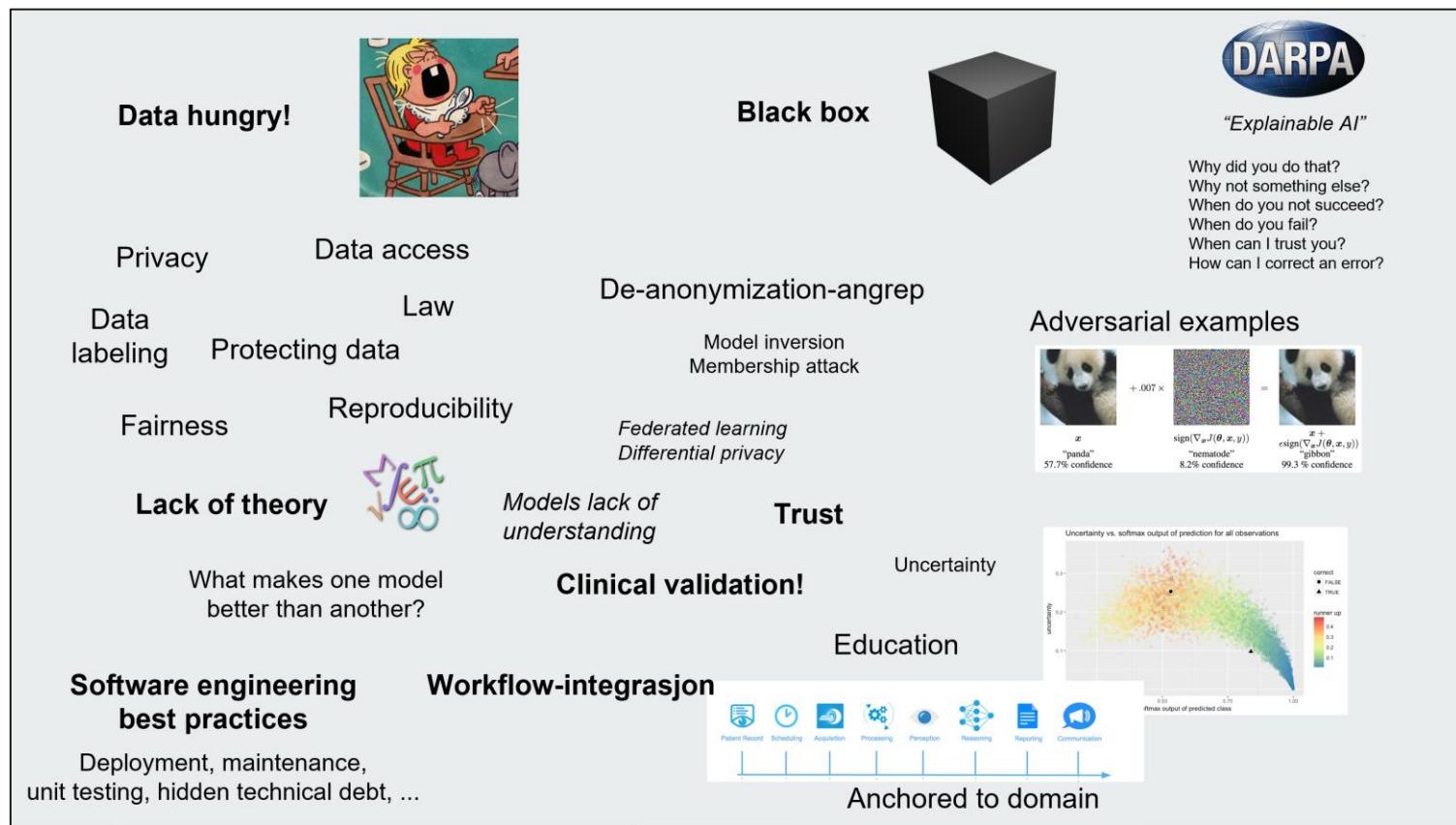
Anchored to domain

Why did you do that?  
Why not something else?  
When do you not succeed?  
When do you fail?  
When can I trust you?  
How can I correct an error?

Adversarial examples

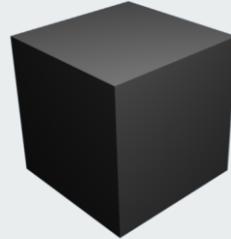


- «Explainability»
  - Fairness, bias, resource allocation
  - Responsibility
  - Skewed incentives
  - "Quality metrics" and commercial interests versus patient care
  - The Hippocratic Oath
  - "Best practices" constantly in flux
  - Robustness
  - Causality
  - Human-in-the-loop, patient-in-the-loop?
  - Data management, governance & provenance
  - Responsible research and innovation
  - Medical, cultural and professional community with a rich history and considerable societal impact.
  - ...





"Explainable AI"



Data hungry!



Privacy Data access  
Data labeling Protecting data  
Fairness Reproducibility  
**Lack of theory**

What makes one model better than another?

**Software engineering best practices**

Deployment, maintenance, unit testing, hidden technical debt, ...

Black box

De-anonymization-angrep

Model inversion  
Membership attack

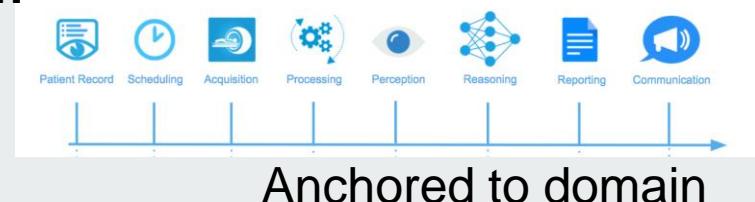
*Federated learning*  
*Differential privacy*

*Models lack of understanding*

**Clinical validation!**

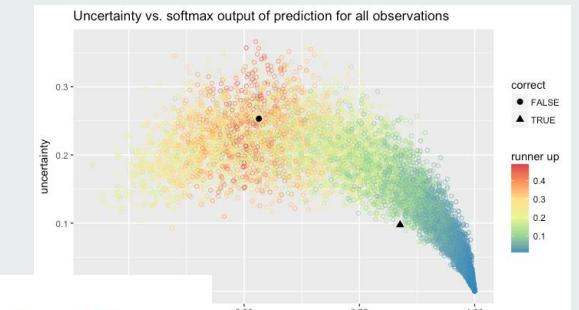
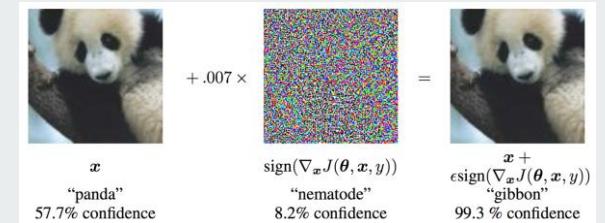
Uncertainty

Education



Why did you do that?  
Why not something else?  
When do you not succeed?  
When do you fail?  
When can I trust you?  
How can I correct an error?

Adversarial examples





"Explainable AI"



Data hungry!



Privacy

Data access

Data labeling      Protecting data

Fairness

Reproducibility

Lack of theory



What makes one model  
better than another?

Software engineering  
best practices

Deployment, maintenance,  
unit testing, hidden technical debt, ...

Black box

Why did you do that?  
Why not something else?  
When do you not succeed?  
When do you fail?  
When can I trust you?  
How can I correct an error?

De-anonymization-angrep

Model inversion  
Membership attack

Federated learning  
Differential privacy

Models lack of  
understanding

Trust

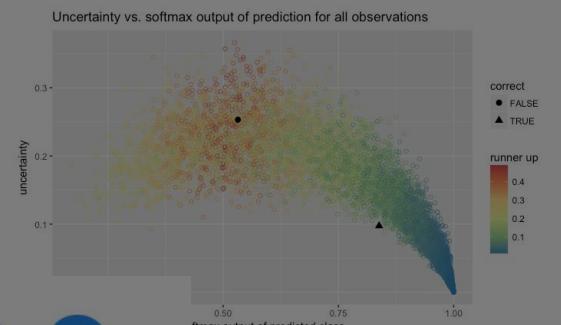
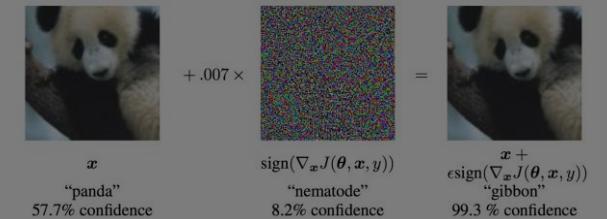
Clinical validation!

Uncertainty

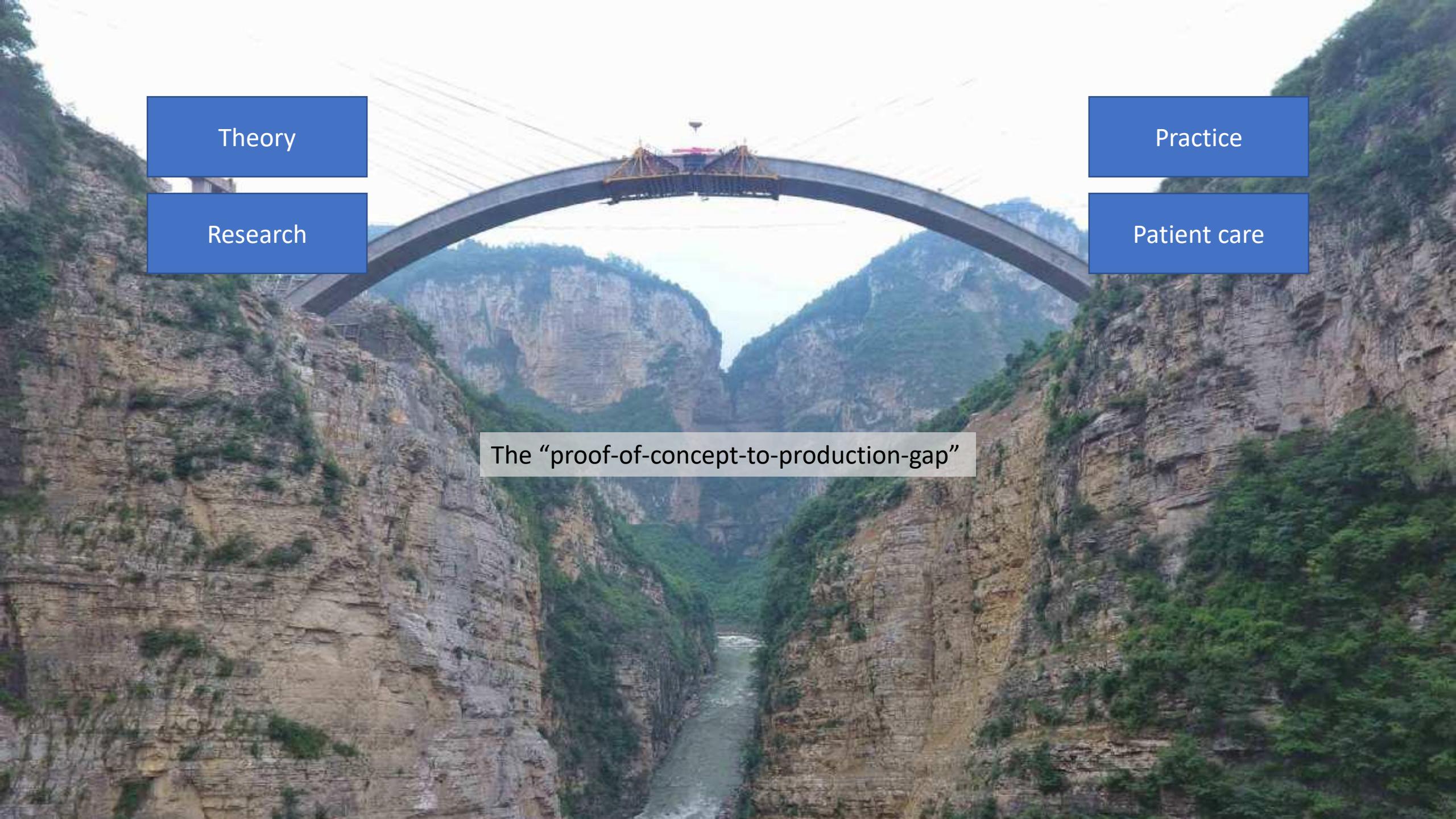
Education



Adversarial examples







Theory

Research

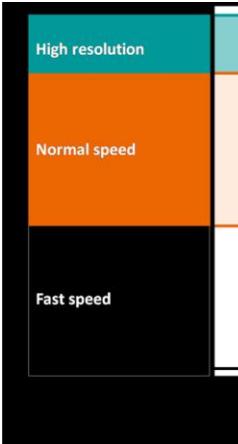
Practice

Patient care

The “proof-of-concept-to-production-gap”



# Siemens Biogra



## Products

Find the artificial intelligence based software for radiology that you are looking for.  
All products listed are available for the European market (CE marked).

Subspecialty: Modality: CE: All CE class: All FDA class: All

Search...

Search

185/185 results



AI4MedImaging

### AI4CMR

Heart ventricles segmentation, quantification of ventricular volumes, myocardial mass, and ejection ...

The AI4CMR software performs a automatic cardiac segmentation and interpretation of Cardiac Magnetic Resonance (CMR), allowing quantification of various parameters without the need for intervention ...

Read more

CE:

Class IIa - MDR



FDA:

Information source:  
Certification verified:

Vendor  
Yes

Subspecialty: Cardiac  
Modality: MR



Siemens Healthineers

### AI-Rad Companion Brain MR

brain volume quantification, segmentation, normative comparison, report generation

AI-Rad Companion Brain MR performs an automatic segmentation of the different brain areas, including an individual volumetric analysis. It compares the different volumes to a normative database and ...

Read more

CE:

Class IIa - MDR



FDA:

Class II

Subspecialty: Neuro  
Modality: MR



Siemens Healthineers

### AI-Rad Companion Chest CT

Segmentation and volume quantification of lungs, lung lobes, lung lesions, heart, thoracic aorta, ...

Read more

CE:

Class IIa - MDR



FDA:

Class II

## Share: Edison

Health with Edison

Applications

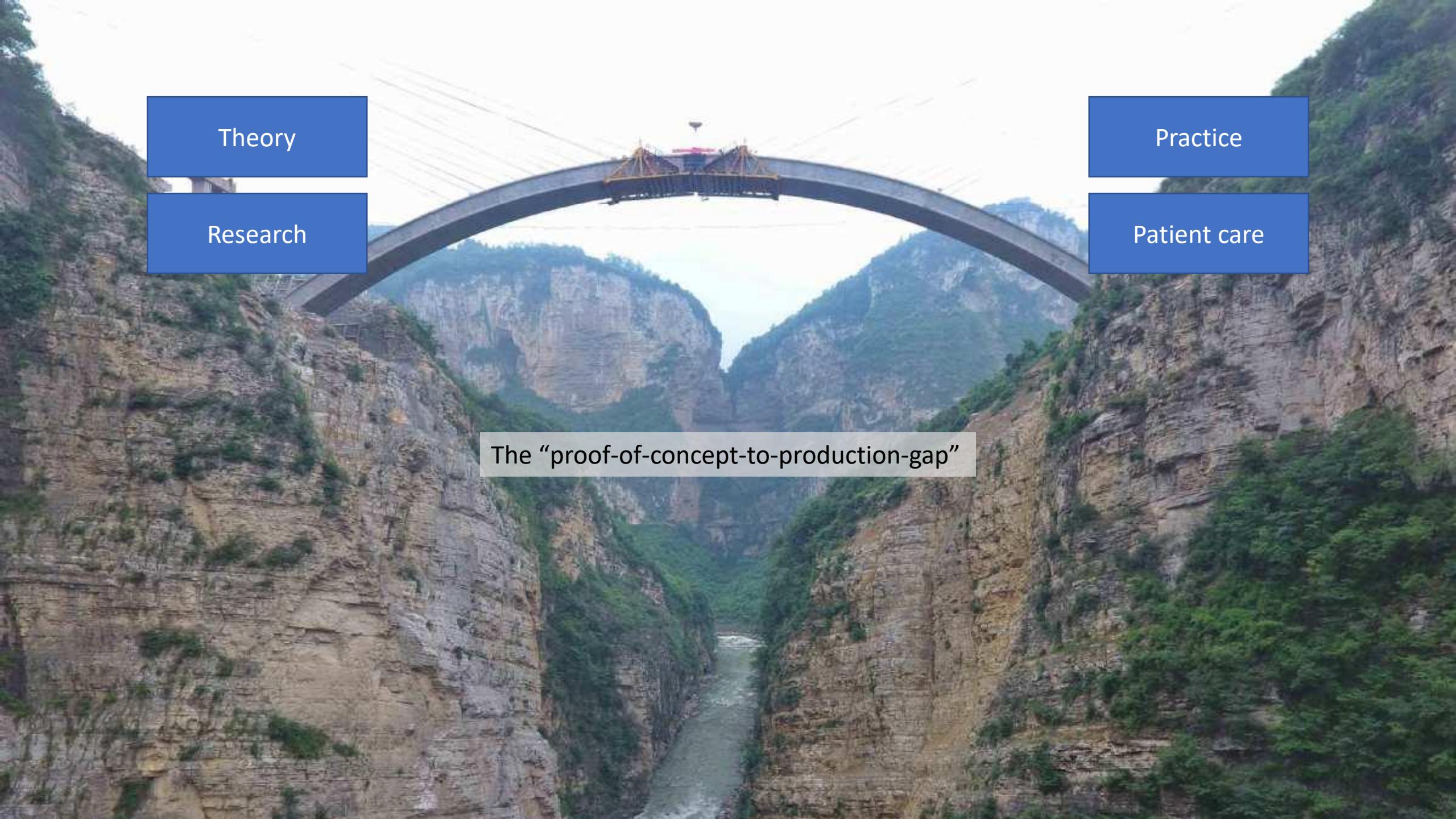
Manager • AIRx • Mural • SonoCNS • Motion • Bone VCAR •

EF 2.0

Smart Devices

Applications are seamlessly embedded in devices, enabling improved workflow, productivity, diagnostics and more





Theory

Research

Practice

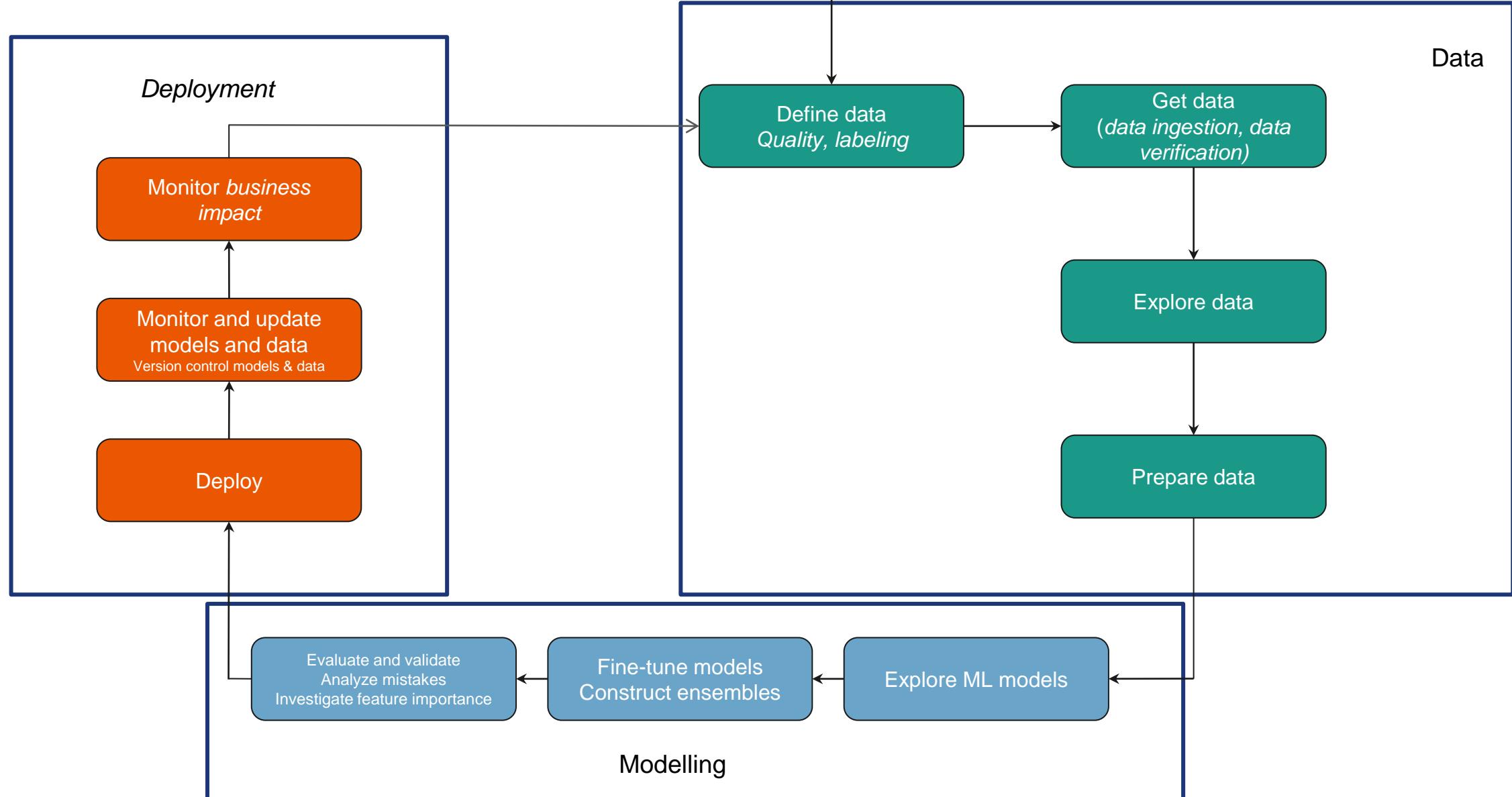
Patient care

The “proof-of-concept-to-production-gap”

*How do you create machine learning solutions?*

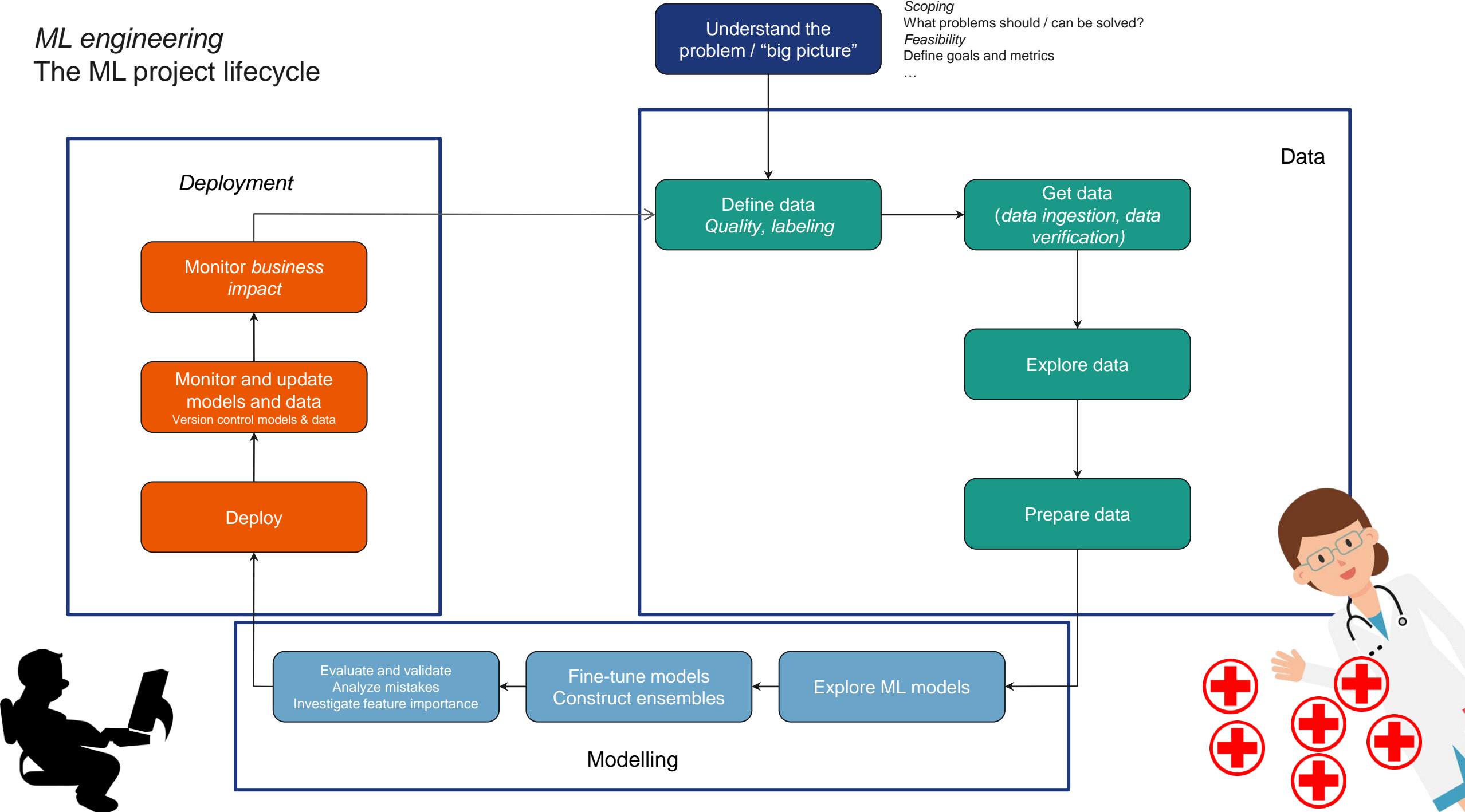
# ML engineering

## The ML project lifecycle



# ML engineering

## The ML project lifecycle



Model

# *Machine learning engineering*

*CompSci + Data Science + Software Engineering*

Data Collection

Data Verification

Feature Extraction

Model

Process Management Tools

Configuration

Machine

ML engineering  
The ML project lifecycle

Understand the problem / "big picture"  
Scoping  
What problems should / can be solved?  
Feasibility  
Define goals and metrics  
...

Define data  
Quality, labeling

Get data  
(data ingestion, data verification)

Explore data

Prepare data

Monitor business impact

Monitor and update models and data  
Version control models & data

Deploy

Evaluate and validate  
Analyze mistakes  
Investigate feature importance

Fine-tune models  
Construct ensembles

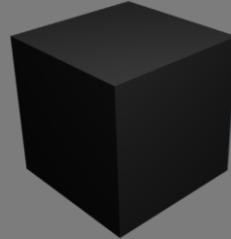
Explore ML models



5 min break



"Explainable AI"



Data hungry!



Privacy

Data access

Data labeling

Protecting data

Fairness

Reproducibility

Lack of theory



What makes one model  
better than another?

Software engineering  
best practices

Deployment, maintenance,  
unit testing, hidden technical debt, ...

Black box

Why did you do that?  
Why not something else?  
When do you not succeed?  
When do you fail?  
When can I trust you?  
How can I correct an error?

De-anonymization-angrep

Model inversion  
Membership attack

Federated learning  
Differential privacy

Models lack of  
understanding

Trust

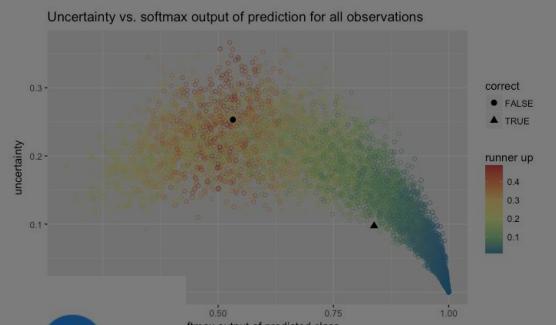
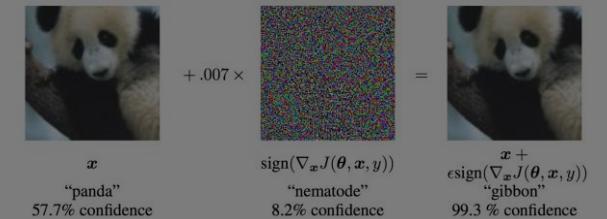
Clinical validation!

Uncertainty

Education



Adversarial examples



**FACEBOOK AI**  

Home Public Leaderboard Challenge Leaderboard The Dataset Submission Guidelines Discussion Login

## fastMRI

Accelerating MR Imaging with AI

Latest News & Updates

09-17-2020 The 2020 fastMRI challenge opens for submissions on October 1 [Read More](#)

What is fastMRI?

fastMRI is a collaborative research project between Facebook AI Research (FAIR) and NYU Langone Health. The aim is to investigate the use of AI to make MRI scans up to 10 times faster.

By producing accurate images from under-sampled data, AI image reconstruction has the potential to improve the patient's experience and to make MRIs accessible for more people.

Download our research paper that describes baselines, evaluation metrics, and the dataset.

Download Dataset Paper

Download fastMRI Papers

PATIENT CARE EDUCATION & RESEARCH ▾

MyChart Patient Login Patient Forms Contact Us Give

 **NewsHub**

Filter News by Category | AUGUST 18, 2020

RESEARCH, PATIENT CARE, PRESS RELEASES | AUGUST 18, 2020

# New Research Finds FastMRI Scans Generated with Artificial Intelligence Are as Accurate as Traditional MRI

POPULAR SCIENCE

Menu Sign Up

Best portable air conditioner: Cool off where you need it most

Maple snow candy is the easiest—and most delicious—winter treat

Quickly de-ice your car's windshield with these sprays

Binoculars for your birdwatching adventure

Great tasting your mornin

TECHNOLOGY

## Artificial intelligence creates better, faster MRI scans

You may someday be able to spend much less time in an imaging machine. Just let AI make the picture.

Rob Verger | September 4, 2020

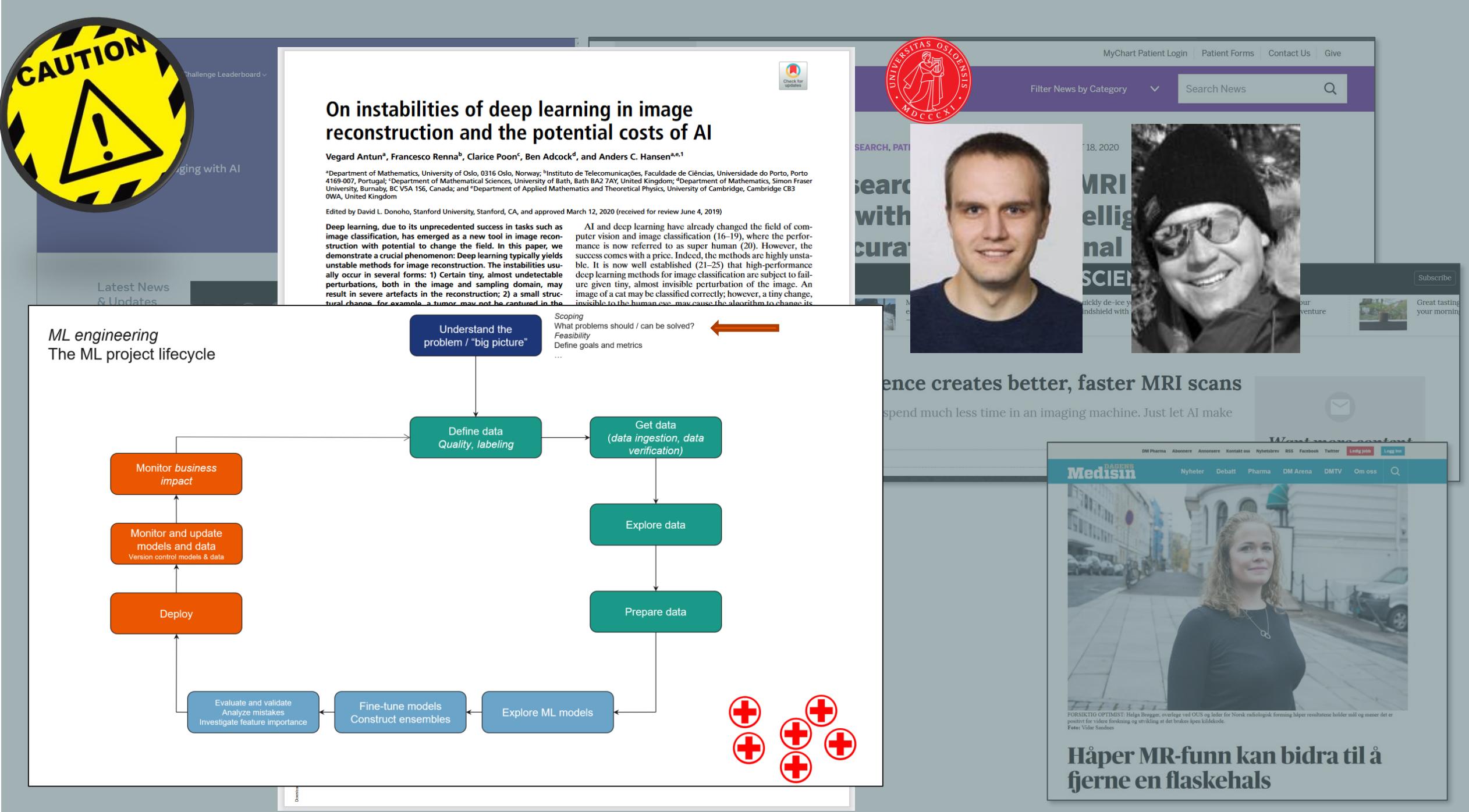
DAGENS Medisin

Nyheter Debatt Pharma DM Arena DMTV Om oss

FORSIKTIG OPTIMIST: Helga Brøgger, overlege ved OUS og leder for Norsk radiologisk forening håper resultatene holder mål og meiner det er positivt for videre forskning og utvikling at det brukes åpen kildekode.

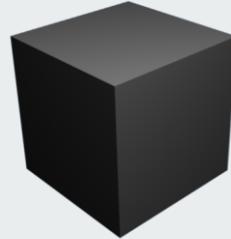
Foto: Vidar Sandnes

## Håper MR-funn kan bidra til å fjerne en flaskehals





"Explainable AI"



Data hungry!



Privacy Data access  
Data labeling Protecting data  
Law

Fairness Reproducibility

Lack of theory

What makes one model better than another?

Software engineering best practices

Deployment, maintenance, unit testing, hidden technical debt, ...

Black box

De-anonymization-angrep

Model inversion  
Membership attack

*Federated learning*  
*Differential privacy*

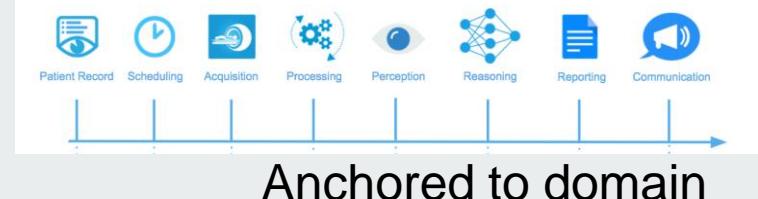
Models lack of understanding

Trust

Clinical validation!

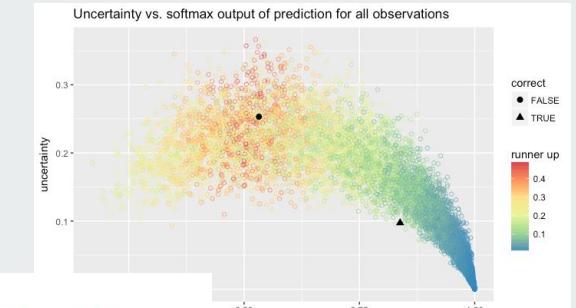
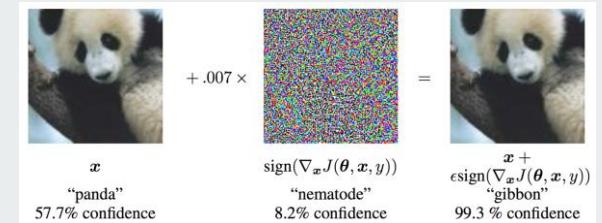
Uncertainty

Education



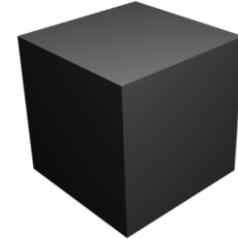
Why did you do that?  
Why not something else?  
When do you not succeed?  
When do you fail?  
When can I trust you?  
How can I correct an error?

Adversarial examples





"Explainable AI"



Black box

Why did you do that?  
Why not something else?  
When do you not succeed?  
When do you fail?  
When can I trust you?  
How can I correct an error?

Data hungry!



Privacy

Data access

Law

Data labeling

Protecting data

Fairness

Reproducibility



Lack of theory

What makes one model  
better than another?

Software engineering  
best practices

Deployment, maintenance,  
unit testing, hidden technical debt, ...

Workflow integration



~~De-anonymization-angrep~~

Model inversion  
Membership attack

*Federated learning*  
*Differential privacy*

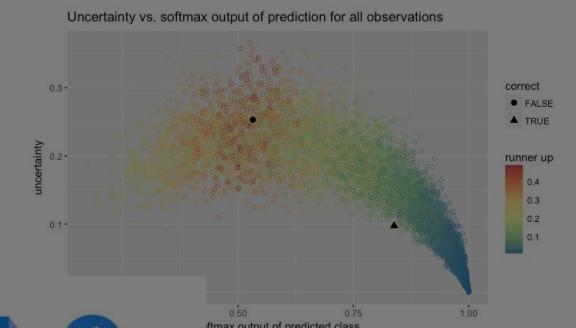
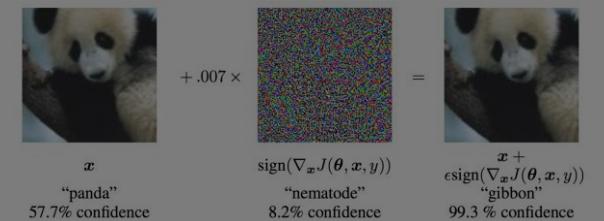
*Models lack of  
understanding*

Trust

Uncertainty

Education

Adversarial examples

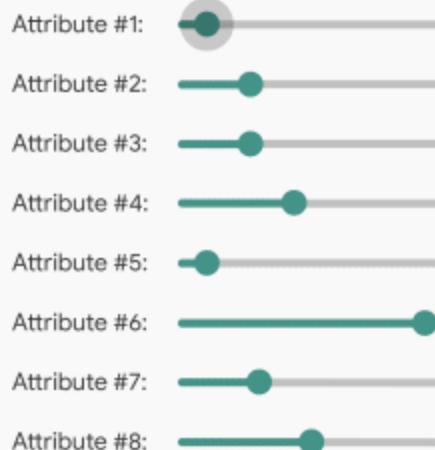


Input Image:

0.827



Detected Classifier Attributes:



$$y \approx f(x; \theta)$$

image → young vs. old



## Explaining in Style: Training a GAN to explain a classifier in StyleSpace

Sept. 2021

Oran Lang<sup>\*1</sup>

Gal Elidan<sup>1</sup>

Yossi Gandelsman<sup>\*1</sup>

Avinatan Hassidim<sup>1</sup>

Michal Yarom<sup>\*1</sup>

William T. Freeman<sup>1,3</sup>

Yoav Wald<sup>\*1,2</sup>

Phillip Isola<sup>1,3</sup>

Amir Globerson<sup>1,4</sup>

Michal Irani<sup>1,5</sup>

Inbar Mosseri<sup>1</sup>

<sup>1</sup> Google Research

<sup>2</sup> Hebrew University

<sup>3</sup> MIT

<sup>4</sup> Tel Aviv University

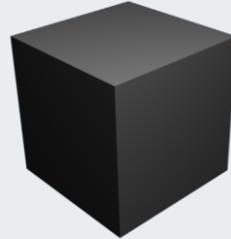
<sup>5</sup> Weizmann Institute of Science

*The current kind of **artificial intelligence**  
requires some kind of **artificial neuroscience***

Break



"Explainable AI"



Data hungry!



Privacy Data access  
Data labeling Protecting data  
Fairness Reproducibility  
**Lack of theory**

What makes one model better than another?

**Software engineering best practices**

Deployment, maintenance, unit testing, hidden technical debt, ...

Black box

De-anonymization-angrep

Model inversion  
Membership attack

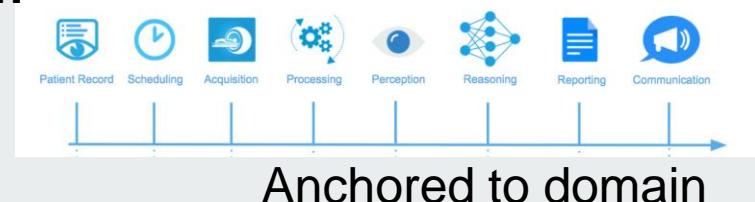
*Federated learning*  
*Differential privacy*

*Models lack of understanding*

**Clinical validation!**

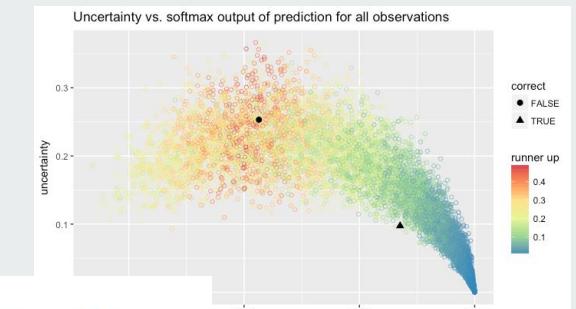
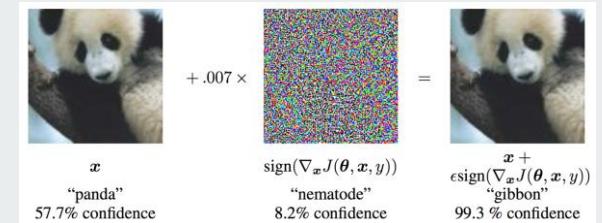
Uncertainty

Education



Why did you do that?  
Why not something else?  
When do you not succeed?  
When do you fail?  
When can I trust you?  
How can I correct an error?

Adversarial examples



# BIAS

AI and machine  
learning is showing  
up in more and more  
areas in society

---

Who gets a mortgage?  
Who's hired?  
Who's convicted?  
Who gets what news?

---

*Who gets what?*

There are some obvious problems with  
this, based on the systems strengths and  
weaknesses..



Bobby Vankavelaar

EXCLUSIVE

TAKING  
ACTION  
FOR YOU

INVESTIGATION FOCUSED ON TESLA AUTOPILOT



ACTION  
NEWS

11:02 83°



# Chinese authorities collecting DNA from all residents of Xinjiang

Officials build database of iris scans and blood types of everyone aged 12 to 65 in region home to 11 million Muslim Uighurs

**China's Huawei tested A.I. software that could identify Uighur Muslims and alert police, report says**

PUBLISHED WED. DEC 9 2020 4:39 AM EST | UPDATED WED. DEC 9 2020 11:07 PM EST

Judges Now Using Artificial Intelligence to Rule on Prisoners



CAMPAIGN TO STOP KILLER ROBOTS

LEARN ACT ABOUT NEWS

THE THREAT OF FULLY AUTONOMOUS WEAPONS

REUTERS

World Business Markets Breakingviews Video More

RETAIL OCTOBER 11, 2018 / 1:04 AM / UPDATED 2 YEARS AGO

Amazon scraps secret AI recruiting tool that showed bias against women

January 24, 2018

**AI Could Trigger WWIII, Alibaba CEO Warns**

Alex Woodie

Facebook looks to AI to monitor harmful behaviours

UK PM seeks 'safe and ethical' artificial intelligence



Chinese police are using facial recognition sunglasses to track citizens

The glasses are being used by officers in police stations to oversee travelers during the Lunar New Year

By James Vincent | @jvincent | Feb 8, 2018, 4:51am EST

f t SHARE



NOW TRENDING



Officials confirm that a cyberattack took place during the Winter Olympics opening ceremonies



New national security unit set up to tackle fake news in UK

**AI Could Set Us Back 100 Years When It Comes to How We Consume News**



Fake videos could become so convincing that we may have to get used to getting our news without them.

▶ Questioning AI: does artificial intelligence need an off switch? - Science Weekly podcast

...but one can encounter problems even  
*with no ill intent!*

*"If we let algorithms decide then society becomes more just. Everyone's treated the same, with the same result."*



Machine learning systems are trained on data collected from “the real world”.

*The real world is unfair!*

With no bad intentions one can easily end  
up designing unfair AI systems

“Dog”



# “Dog”

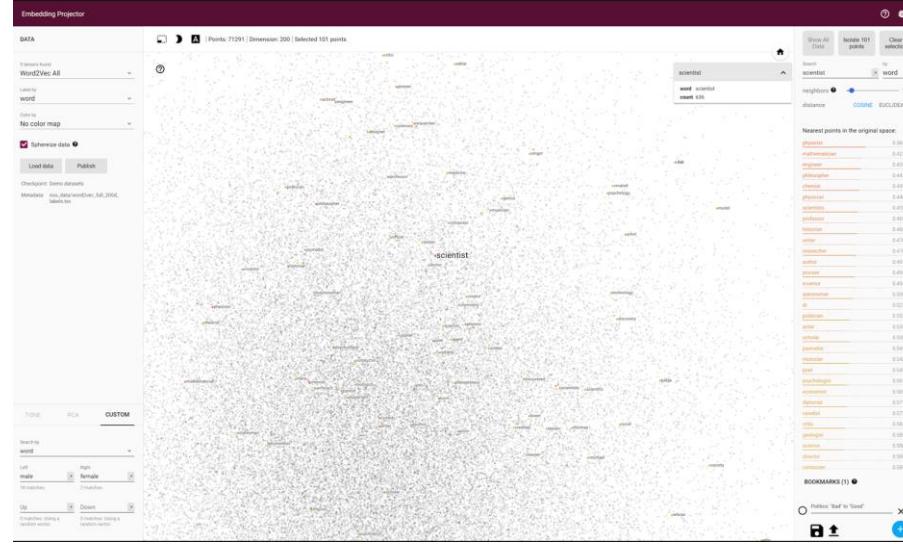


## Word embeddings

word → vector of numbers

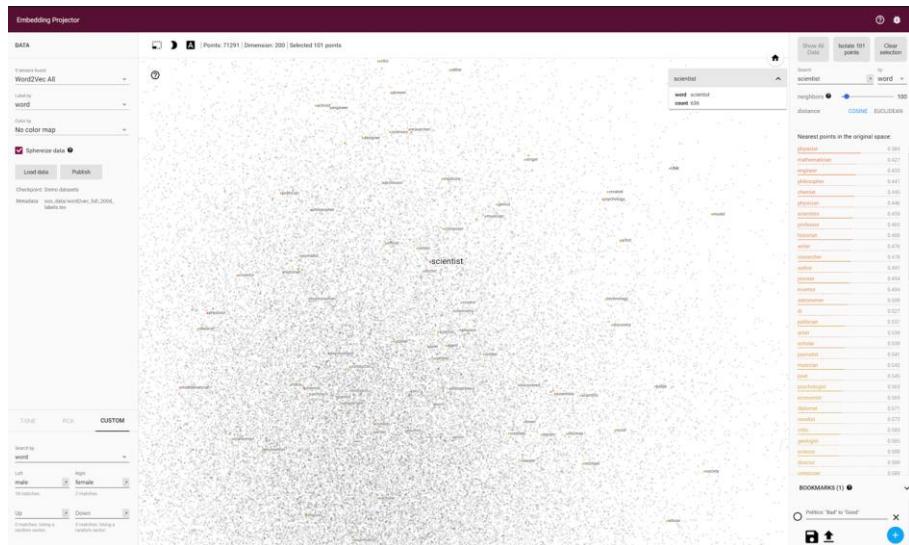
Similar things can be said for other representations of language  
used in AI

l	5.12695312e-02	-2.23388672e-02	-1.72851562e-01	1.61132812e-01
-8.44726562e-02	5.73730469e-02	5.85937500e-02	-8.25195312e-02	
-1.53808594e-02	-6.34765625e-02	1.79687500e-01	-4.23828125e-01	
-2.25830078e-02	-1.66015625e-01	-2.51464844e-02	1.07421875e-01	
-1.99218750e-01	1.59179688e-01	-1.87500000e-01	-1.20117188e-01	
1.55273438e-01	-9.91210938e-02	1.42578125e-01	-1.64062500e-01	
-8.93554688e-02	2.00195312e-01	-1.49414062e-01	3.20312500e-01	
2.28125000e-01	2.44140625e-02	9.71679688e-02	-8.20312500e-02	
-3.63769531e-02	-8.59375000e-02	9.86328125e-02	7.78198242e-03	
-1.34277344e-02	5.27343750e-02	1.48437500e-01	3.39843735e-01	
1.66015625e-02	-2.12890625e-01	-1.50756836e-02	5.24902344e-02	
-1.07421875e-01	-8.88671875e-02	2.49923438e-01	-7.03125000e-02	
-1.59912109e-02	7.56635938e-02	-7.03125000e-02	1.19140625e-01	
2.29492188e-01	1.41601562e-02	1.15234375e-01	7.50732422e-03	
2.75390625e-01	-2.44140625e-01	2.96875000e-01	3.49121094e-02	
2.42187500e-01	1.35742188e-01	1.42578125e-01	1.25781250e-02	
2.92968750e-02	-1.21528031e-01	2.28271484e-02	-4.76074219e-02	
-1.55273438e-01	3.14331055e-03	3.45703125e-01	1.22558594e-01	
-1.95312500e-01	8.10546875e-02	-6.83593750e-02	-1.47094727e-02	
2.14843756e-01	-1.216937350e-01	1.57226562e-01	-2.07031250e-01	
1.36718750e-01	-1.29882812e-01	5.29785156e-02	-2.71484375e-01	
-2.98828125e-01	-1.84570312e-01	-2.29492188e-01	1.19140625e-01	
1.53198242e-02	-2.61718750e-01	-1.23046875e-01	-1.86767578e-02	
-6.49414062e-02	-8.151429688e-02	7.86132812e-02	-3.53515625e-01	
5.24902344e-02	-2.45361328e-02	-5.43212891e-03	-2.08984375e-01	
-2.10937500e-01	-1.79687500e-01	2.42187500e-01	2.57812500e-01	
1.37695312e-01	-2.10937500e-01	-2.17285156e-02	-1.38671875e-01	
1.84326172e-02	-1.239013367e-02	-1.59179688e-01	1.61132812e-01	
2.08807812e-01	1.63027344e-01	9.81445312e-02	-6.83593750e-02	
8.72802734e-03	-2.89062500e-01	-2.14843750e-01	-1.14257812e-01	
-2.21679688e-01	4.12597656e-02	-3.12500000e-01	-5.59082031e-02	
-9.76562500e-02	5.81054688e-02	-4.05273438e-02	-1.45273438e-01	
1.64062500e-01	-2.53906250e-01	1.54296875e-01	-2.31933594e-02	
-2.38281250e-01	2.07519531e-02	-2.73437500e-01	3.90625000e-03	
1.13769531e-01	-1.73828125e-01	2.57812500e-01	2.35351562e-01	
5.22460938e-02	6.83953750e-02	-1.75781250e-01	1.60156250e-01	
-5.98907471e-04	5.98144531e-02	-2.11914062e-01	-5.54199219e-02	
-7.51953125e-02	-3.06640625e-01	4.27734375e-01	5.32265625e-02	
-2.08984375e-01	-5.71289062e-02	-2.09960938e-01	3.29589844e-02	
1.05468750e-01	-1.50390625e-01	-9.37500000e-02	1.16699219e-01	
6.44531250e-02	2.80761719e-02	2.41210938e-01	-1.25795625e-01	
-1.00585938e-01	-1.22680664e-02	-2.36156616e-04	1.58691406e-02	
1.27929688e-01	-3.32031250e-02	4.07714844e-02	-1.31835938e-01	
9.81445312e-02	1.74804688e-01	-2.36328125e-01	5.17578125e-02	
1.83593750e-01	2.42919922e-02	-4.31640625e-01	2.46093750e-01	
-3.03955078e-02	-2.47802734e-02	-1.17187500e-01	1.61132812e-01	
-5.71289062e-02	1.16577148e-02	2.81250000e-01	4.27734375e-01	
4.56542969e-02	1.01674219e-01	-3.95587812e-02	1.77001953e-02	
-8.98437500e-02	1.35742188e-01	2.080007812e-01	1.88476562e-01	
-1.52343750e-01	-2.37304688e-01	-1.90429688e-01	7.12890625e-02	
-2.46093750e-01	2.61718750e-01	-2.34375000e-01	-1.45507812e-01	
-1.17187500e-02	-1.50390625e-01	1.13281250e-01	1.82617188e-01	
2.63671875e-01	-1.37695312e-01	-4.58984375e-01	-4.68750000e-02	
-1.26953125e-01	4.22363281e-02	1.66992188e-01	1.26953125e-01	
2.59765625e-01	-2.44140625e-01	2.19726562e-01	-8.69140625e-02	
1.59179688e-01	-3.78417969e-02	8.97216797e-03	-2.77343750e-01	
-1.84988469e-01	-1.75781250e-01	-2.28515625e-01	-2.70996694e-02	
2.85156250e-01	-2.73437500e-01	1.61132812e-02	5.90820312e-02	
-2.39257812e-01	1.77734375e-01	-1.34765625e-01	1.38671875e-01	
3.53515625e-01	1.22070312e-01	1.43554688e-01	9.22851562e-02	
2.29492188e-01	-3.008781250e-01	-4.88281250e-02	-1.79687500e-01	
2.96875000e-01	1.75781250e-01	4.80957031e-02	-3.38745117e-03	
7.91015625e-02	-2.38281250e-01	-2.31445312e-01	1.66015625e-01	
-2.13867188e-01	-7.03125000e-02	-7.56835938e-02	1.9628962e-01	
-1.29882812e-01	-1.05957031e-01	-3.53515625e-01	-1.16699219e-01	
-5.10253906e-02	3.39355469e-02	-1.43554688e-01	-3.90625000e-03	
1.73828125e-01	-9.96093750e-02	-1.66015625e-01	-8.54492188e-02	
-3.82812500e-01	5.90820312e-02	-6.22558594e-02	8.83789625e-01	
-8.88671875e-02	3.28125000e-01	6.83593750e-02	-1.91406250e-01	
-8.35418701e-04	1.04003906e-01	1.52343750e-01	-1.53350830e-03	
4.16015625e-01	-3.32031250e-02	1.49414062e-01	2.42187500e-01	
-1.76757812e-01	-4.93164062e-02	-1.24511719e-01	1.25976562e-01	
1.74804688e-01	2.81250000e-01	-1.80664062e-01	1.03027344e-01	
-2.75390625e-01	2.61718750e-01	2.46993750e-01	-4.71191466e-02	
6.25000000e-02	4.16015625e-01	-3.55468750e-01	2.22656250e-01	



<http://projector.tensorflow.org>

Data generated in our societies contain **stereotypes** and **cultural bias**



<http://projector.tensorflow.org>

Data generated in our societies contain **stereotypes** and **cultural bias**

The New York Times

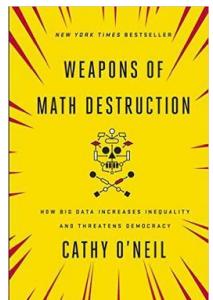
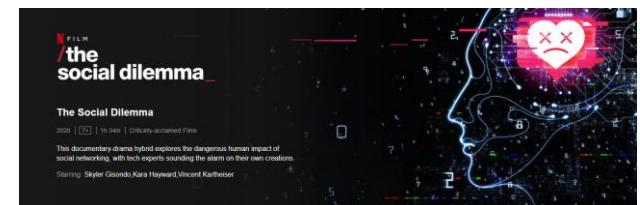
### Facial Recognition Is Accurate, if You're a White Guy

REUTERS

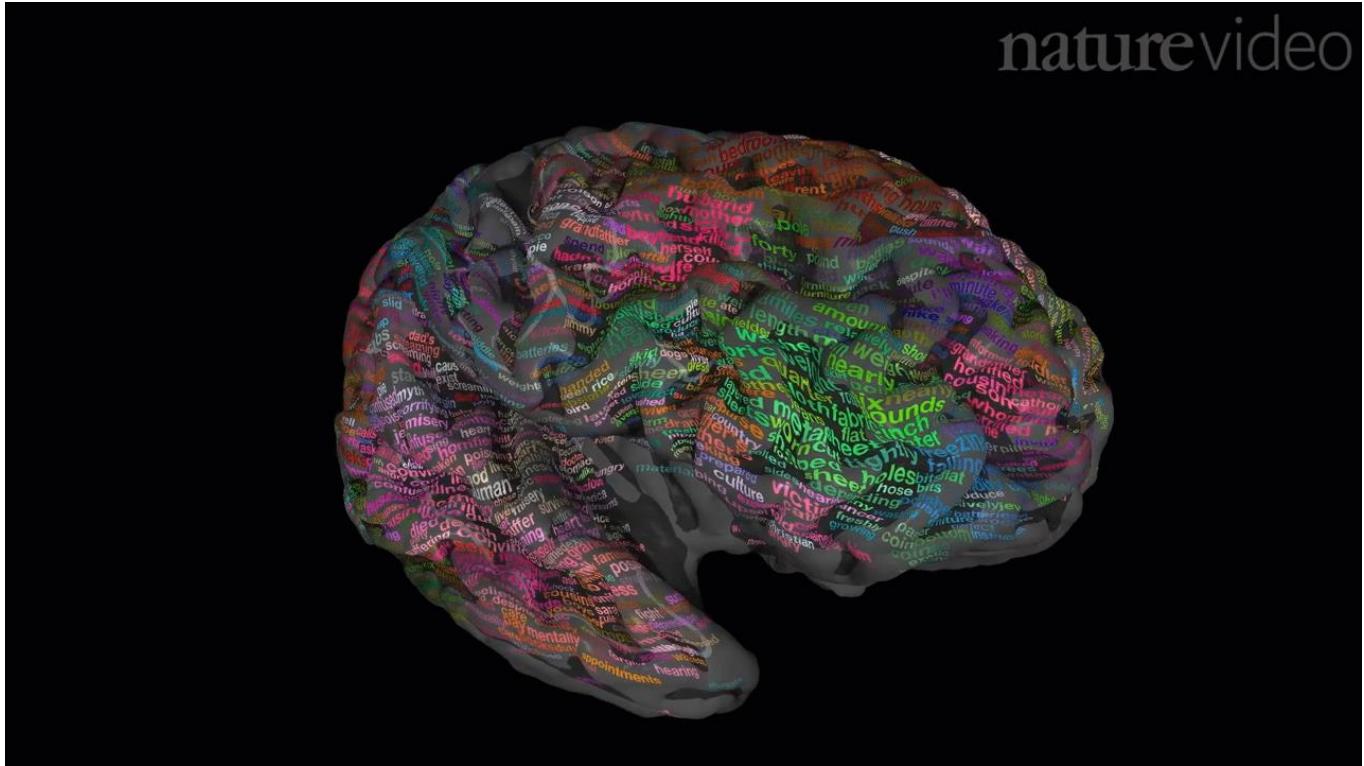
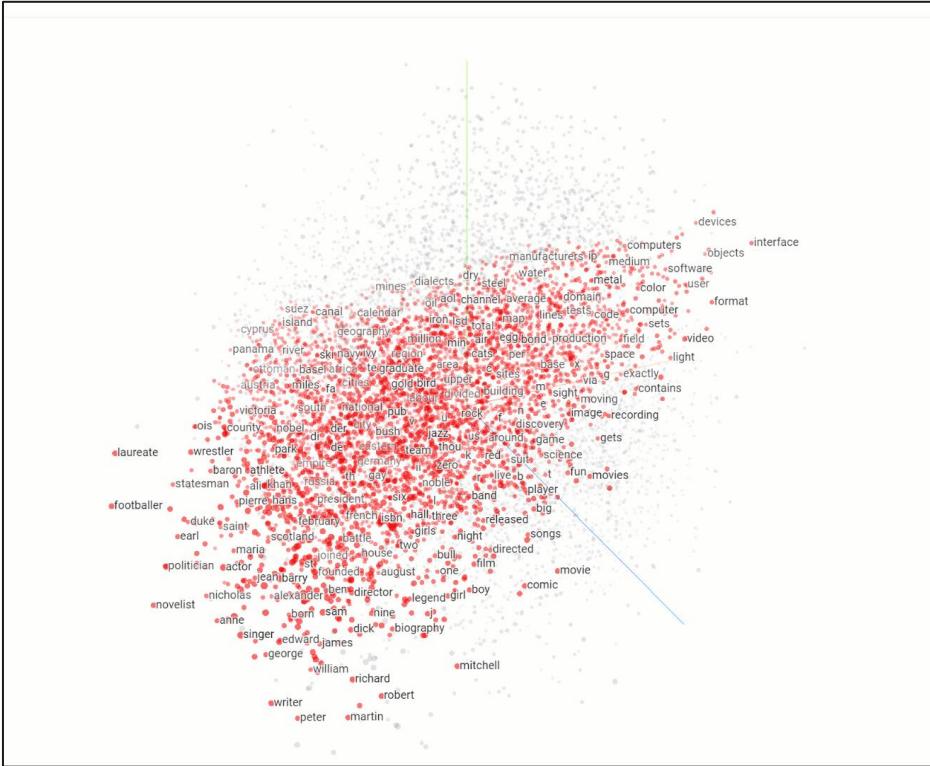
RETAIL OCTOBER 11, 2018 / 1:04 AM / UPDATED 2 YEARS AGO

### Amazon scraps secret AI recruiting tool that showed bias against women

Facebook's response to Netflix documentary



Break



*Natural speech reveals the semantic maps that tile human cerebral cortex, Nature, April 2016*



"Explainable AI"



Data hungry!



Black box

Why did you do that?  
Why not something else?  
When do you not succeed?  
When do you fail?  
When can I trust you?  
How can I correct an error?

Privacy Data access

Data labeling Protecting data

Fairness Reproducibility

Lack of theory



What makes one model better than another?

Software engineering best practices

Deployment, maintenance,  
unit testing, hidden technical debt ...

De-anonymization-angrep

Model inversion  
Membership attack

Federated learning  
Differential privacy

Models lack of understanding

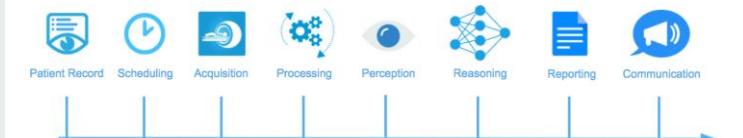
Trust

Clinical validation!

Uncertainty

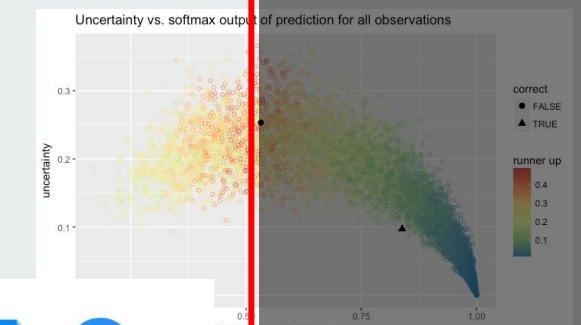
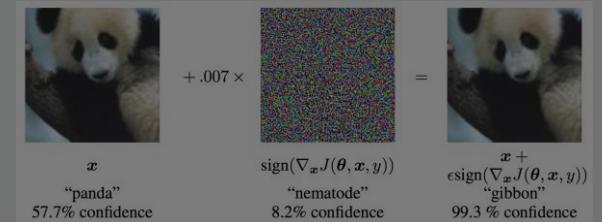
Education

Workflow integration



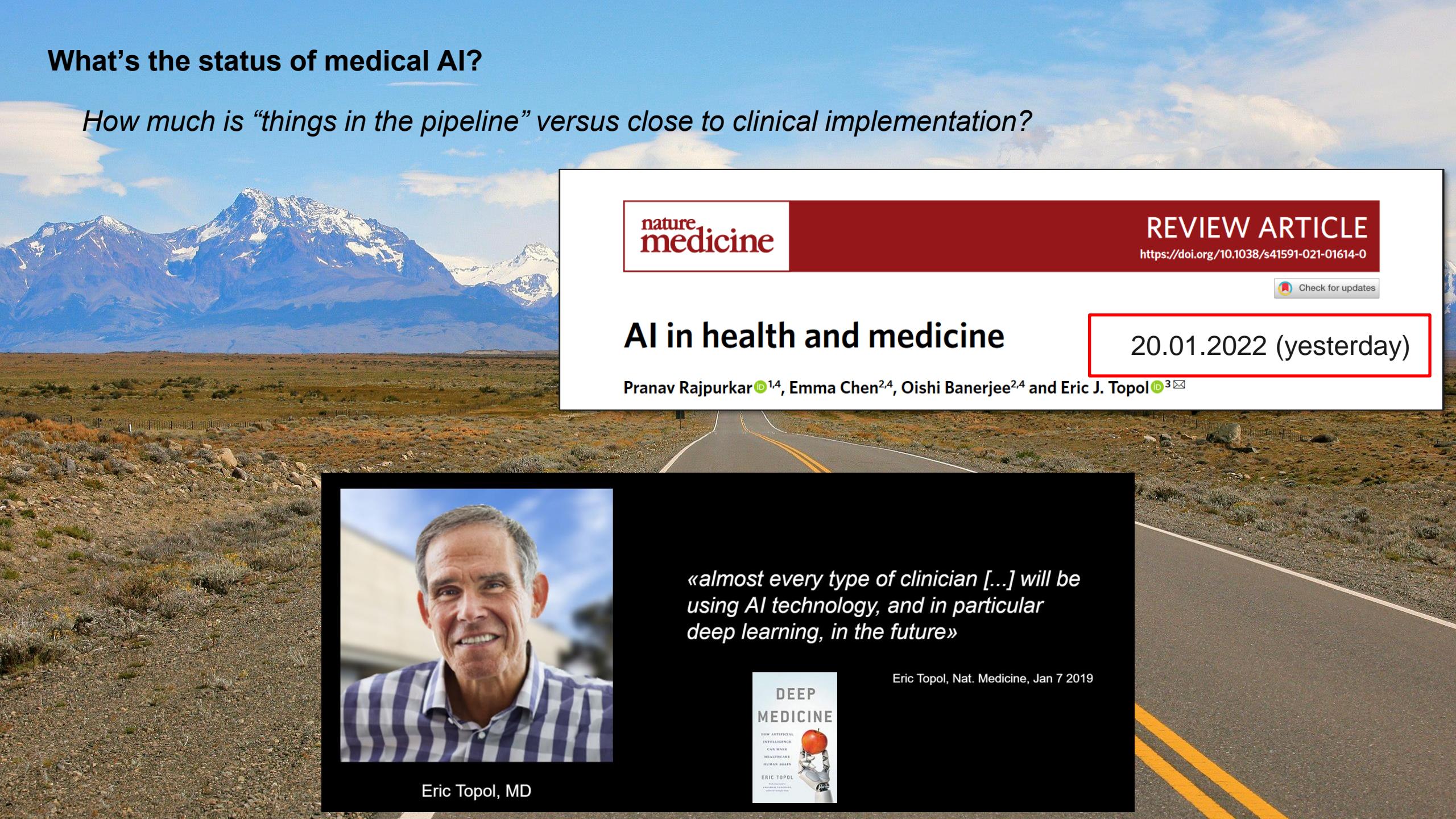
Anchored to domain

Adversarial examples



# What's the status of medical AI?

*How much is “things in the pipeline” versus close to clinical implementation?*



**nature medicine**

**REVIEW ARTICLE**  
<https://doi.org/10.1038/s41591-021-01614-0>

Check for updates

**AI in health and medicine**

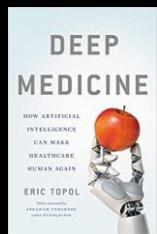
20.01.2022 (yesterday)

Pranav Rajpurkar<sup>1,4</sup>, Emma Chen<sup>2,4</sup>, Oishi Banerjee<sup>2,4</sup> and Eric J. Topol<sup>1,3</sup>✉



Eric Topol, MD

*“almost every type of clinician [...] will be using AI technology, and in particular deep learning, in the future”*

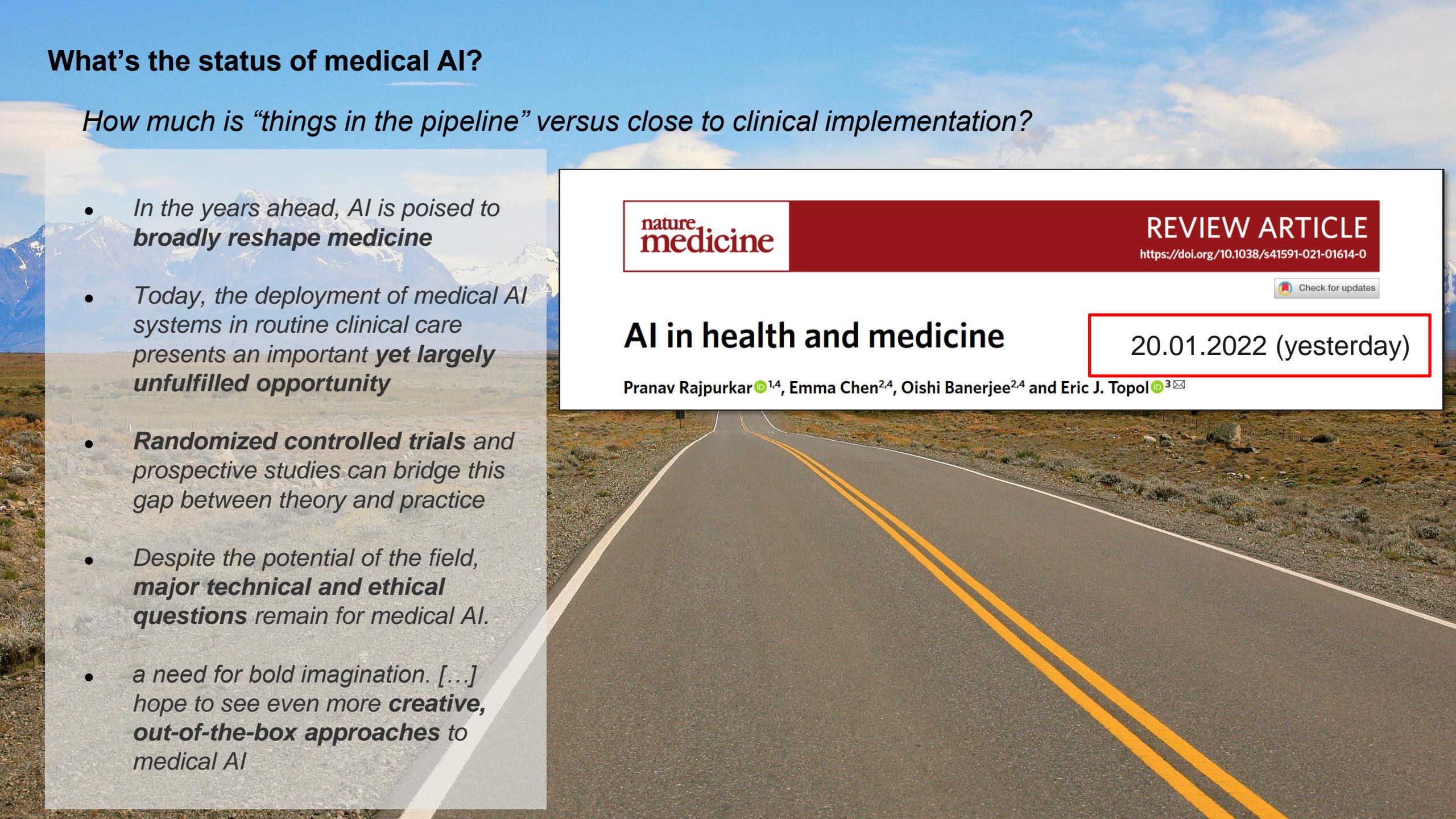


Eric Topol, Nat. Medicine, Jan 7 2019

# What's the status of medical AI?

*How much is “things in the pipeline” versus close to clinical implementation?*

- *In the years ahead, AI is poised to broadly reshape medicine*
- *Today, the deployment of medical AI systems in routine clinical care presents an important yet largely unfulfilled opportunity*
- *Randomized controlled trials and prospective studies can bridge this gap between theory and practice*
- *Despite the potential of the field, major technical and ethical questions remain for medical AI.*
- *a need for bold imagination. [...] hope to see even more creative, out-of-the-box approaches to medical AI*



**nature medicine**

**REVIEW ARTICLE**

<https://doi.org/10.1038/s41591-021-01614-0>



**AI in health and medicine**

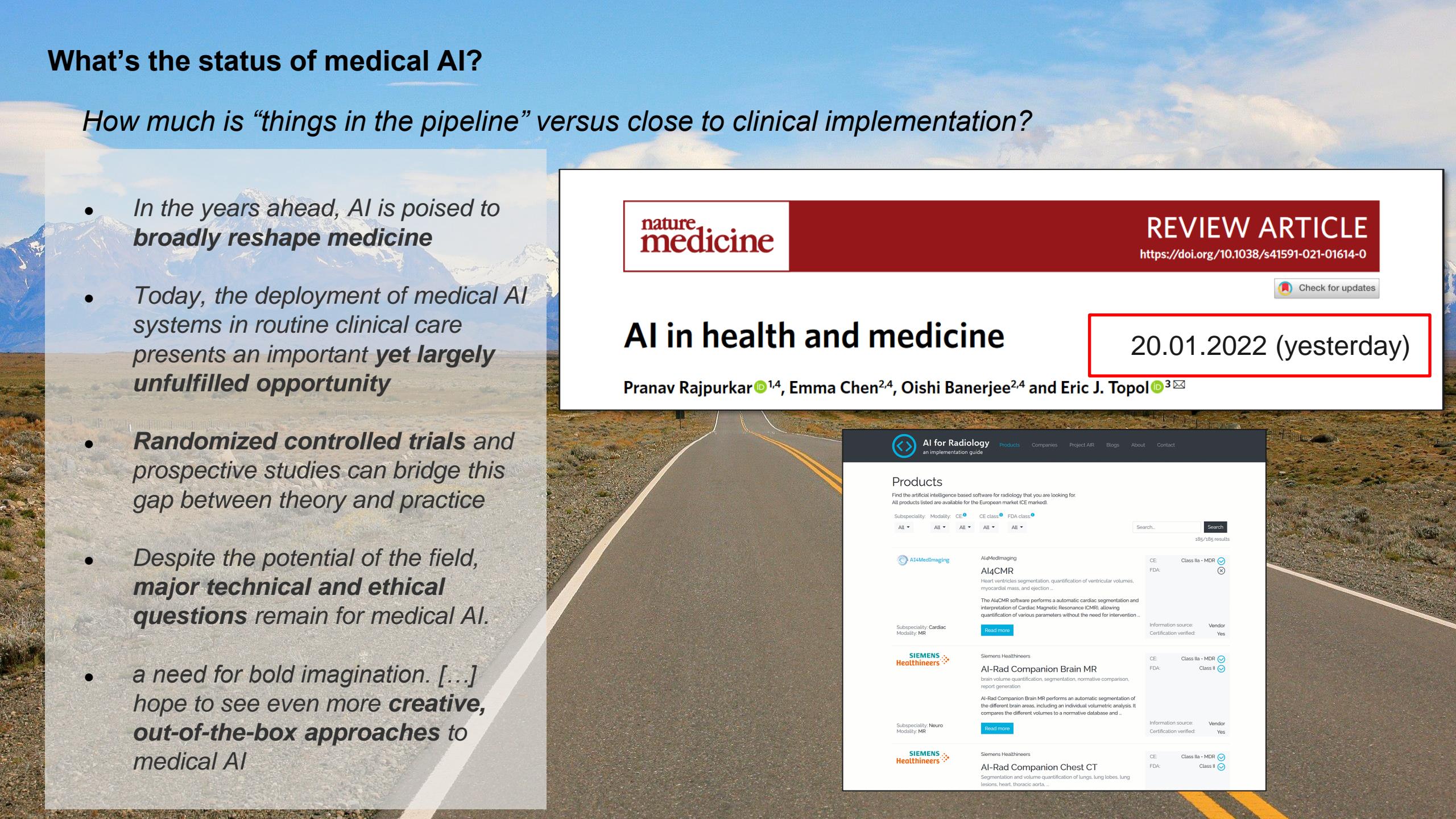
Pranav Rajpurkar<sup>1,4</sup>, Emma Chen<sup>2,4</sup>, Oishi Banerjee<sup>2,4</sup> and Eric J. Topol<sup>1,3</sup>

20.01.2022 (yesterday)

# What's the status of medical AI?

*How much is “things in the pipeline” versus close to clinical implementation?*

- *In the years ahead, AI is poised to broadly reshape medicine*
- *Today, the deployment of medical AI systems in routine clinical care presents an important yet largely unfulfilled opportunity*
- *Randomized controlled trials and prospective studies can bridge this gap between theory and practice*
- *Despite the potential of the field, major technical and ethical questions remain for medical AI.*
- *a need for bold imagination. [...] hope to see even more creative, out-of-the-box approaches to medical AI*



The image shows a road curving away from the viewer, leading towards a range of mountains under a clear blue sky with some white clouds. This visual metaphor represents the "pipeline" of medical AI development, where the road leads from current research and clinical implementation towards distant, unfulfilled opportunities.

**nature medicine**

**REVIEW ARTICLE**  
<https://doi.org/10.1038/s41591-021-01614-0>



**AI in health and medicine**

Pranav Rajpurkar<sup>1,4</sup>, Emma Chen<sup>2,4</sup>, Oishi Banerjee<sup>2,4</sup> and Eric J. Topol<sup>1,3</sup>

20.01.2022 (yesterday)

**AI for Radiology**  
an implementation guide

Products

Find the artificial intelligence based software for radiology that you are looking for.  
All products listed are available for the European market (CE marked).

Subspecialty: Modality: All CE class: All FDA class: All

All All All All All

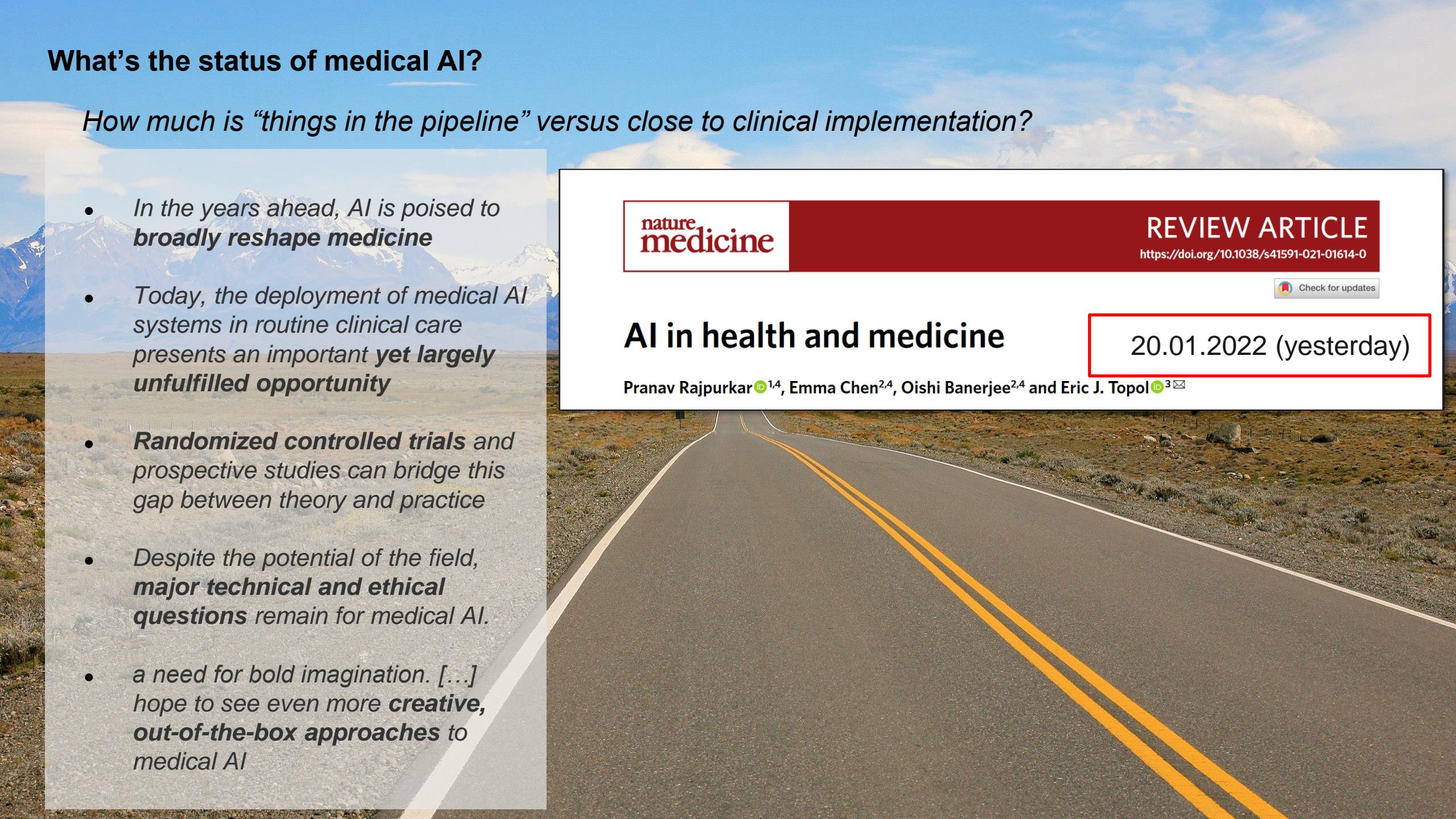
Search... Search 185/185 results

AI4MedImaging	AI4CMR	Siemens Healthineers	Siemens Healthineers	Siemens Healthineers
AI4MedImaging	AI4CMR	AI4MedImaging	AI-Rad Companion Brain MR	AI-Rad Companion Chest CT
Heart ventricles segmentation, quantification of ventricular volumes, myocardial mass, and ejection ...	Automatic cardiac segmentation and interpretation of Cardiac Magnetic Resonance (CMR), allowing quantification of various parameters without the need for intervention ...	Cardiac	brain volume quantification, segmentation, normative comparison, report generation	Segmentation and volume quantification of lungs, lung lobes, lung lesions, heart, thoracic aorta ...
Modality: MR	Modality: MR	Modality: MR	Modality: MR	Modality: CT
Information source: Vendor Yes	Information source: Vendor Yes	Information source: Vendor Yes	Information source: Vendor Yes	Information source: Vendor Yes
CE: Class IIa - MDR	CE: Class IIa - MDR	CE: Class IIa - MDR	CE: Class IIa - MDR	CE: Class IIa - MDR
FDA: Class II	FDA: Class II	FDA: Class II	FDA: Class II	FDA: Class II

# What's the status of medical AI?

*How much is “things in the pipeline” versus close to clinical implementation?*

- *In the years ahead, AI is poised to broadly reshape medicine*
- *Today, the deployment of medical AI systems in routine clinical care presents an important yet largely unfulfilled opportunity*
- *Randomized controlled trials and prospective studies can bridge this gap between theory and practice*
- *Despite the potential of the field, major technical and ethical questions remain for medical AI.*
- *a need for bold imagination. [...] hope to see even more creative, out-of-the-box approaches to medical AI*



**nature medicine**

**REVIEW ARTICLE**

<https://doi.org/10.1038/s41591-021-01614-0>



**AI in health and medicine**

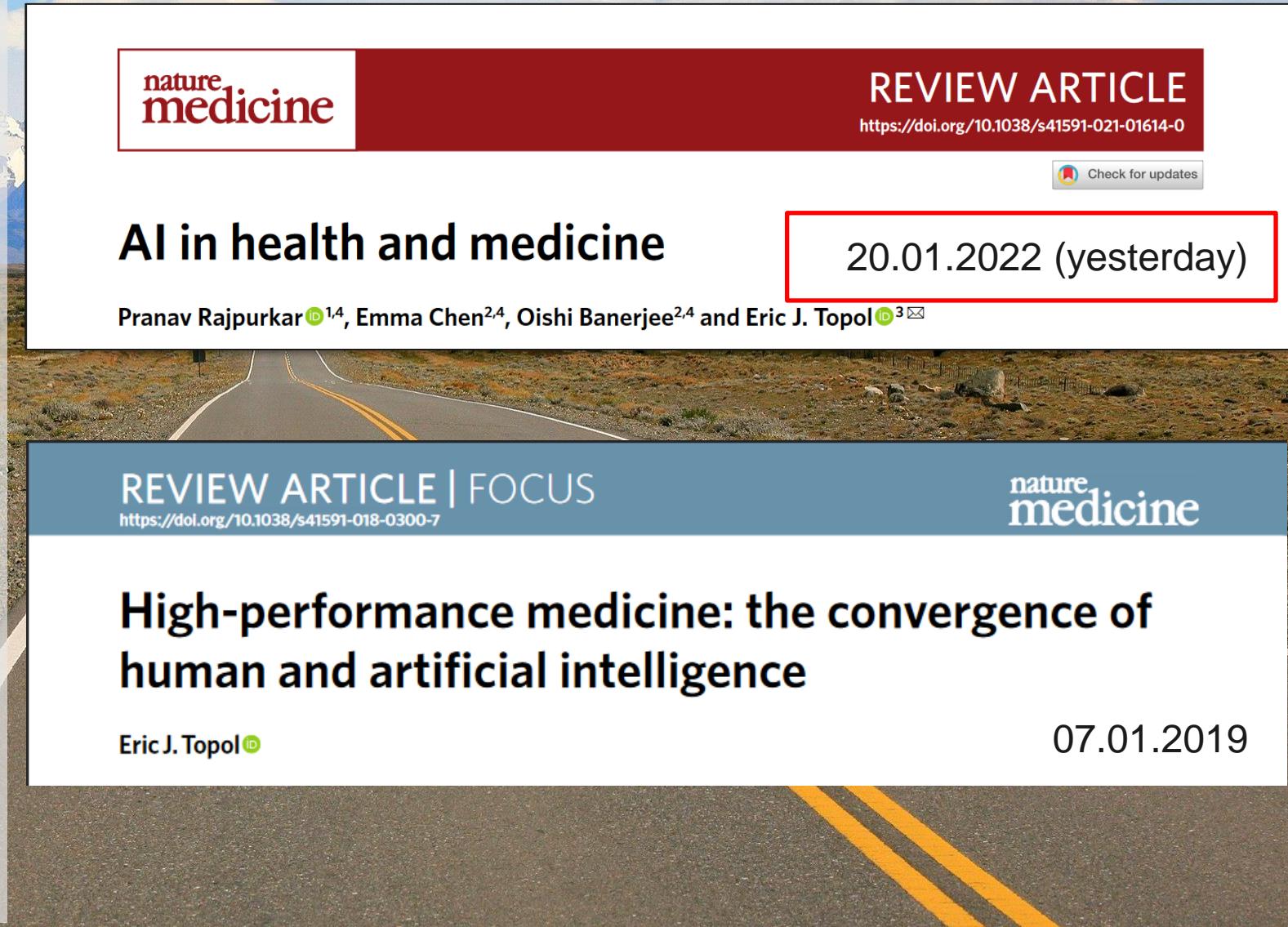
Pranav Rajpurkar<sup>1,4</sup>, Emma Chen<sup>2,4</sup>, Oishi Banerjee<sup>2,4</sup> and Eric J. Topol<sup>1,3</sup>

20.01.2022 (yesterday)

# What's the status of medical AI?

*How much is “things in the pipeline” versus close to clinical implementation?*

- *In the years ahead, AI is poised to broadly reshape medicine*
- *Today, the deployment of medical AI systems in routine clinical care presents an important yet largely unfulfilled opportunity*
- *Randomized controlled trials and prospective studies can bridge this gap between theory and practice*
- *Despite the potential of the field, major technical and ethical questions remain for medical AI.*
- *a need for bold imagination. [...] hope to see even more creative, out-of-the-box approaches to medical AI*



**nature medicine** REVIEW ARTICLE  
<https://doi.org/10.1038/s41591-021-01614-0>  
Check for updates

**AI in health and medicine** 20.01.2022 (yesterday)

Pranav Rajpurkar<sup>1,4</sup>, Emma Chen<sup>2,4</sup>, Oishi Banerjee<sup>2,4</sup> and Eric J. Topol<sup>1,3</sup>

**REVIEW ARTICLE | FOCUS** <https://doi.org/10.1038/s41591-018-0300-7> nature medicine

**High-performance medicine: the convergence of human and artificial intelligence** 07.01.2019

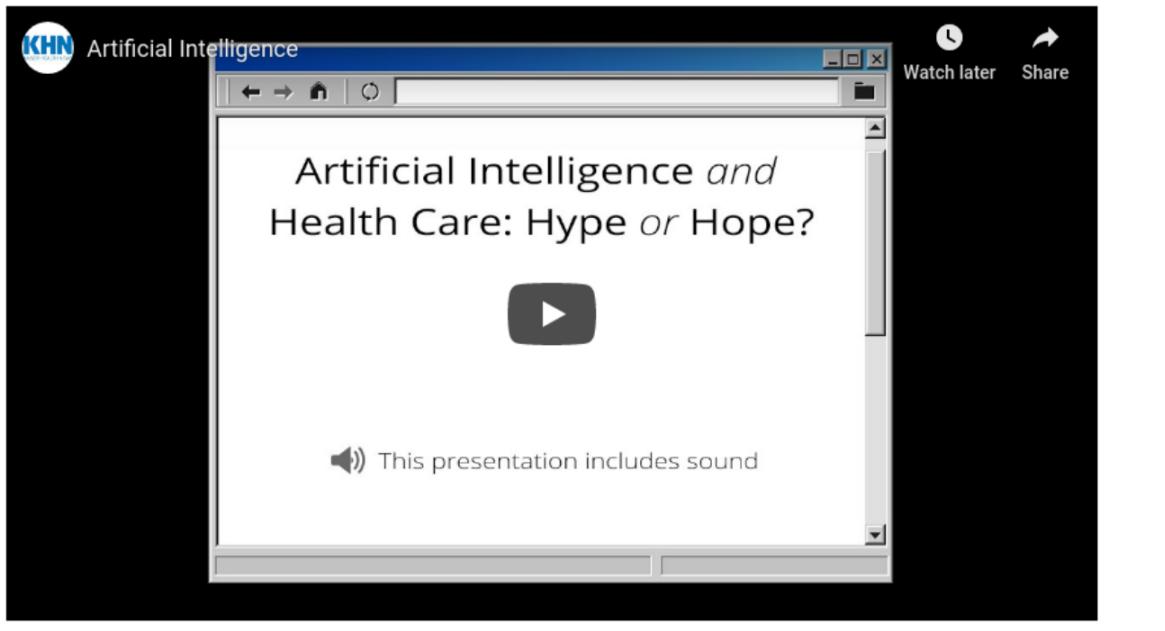
Eric J. Topol

## A Reality Check On Artificial Intelligence: Are Health Care Claims Overblown?

By [Liz Szabo](#)

Dec. 2019

1 week ago



Health products powered by [artificial intelligence](#), or AI, are streaming into our lives, from virtual [doctor apps](#) to [wearable](#)

## Time to reality check the promises of machine learning-powered precision medicine

Jack Wilkinson, Kellyn F Arnold, Eleanor J Murray, Maarten van Smeden, Kareem Carr, Rachel Sippy, Marc de Kamps, Andrew Beam, Stefan Konigorski, Christoph Lippert, Mark S Gilthorpe, Peter W G Tennant

Machine learning methods, combined with large electronic health databases, could enable a personalised approach to medicine through improved diagnosis and prediction of individual responses to therapies. If successful, this strategy would represent a revolution in clinical research and practice. However, although the vision of individually tailored



Sept. 2020



Lancet Digital Health 2020  
Published Online  
September 16, 2020

- “There’s nothing that I’ve seen in my 30-plus years studying medicine that could be as impactful and transformative” as AI, said Dr. Eric Topol
- many health industry experts fear AI-based products won’t be able to match the hype.
- AI systems sometimes learn to make predictions based on factors that have less to do with disease than the **brand of MRI machine used**, the **time a blood test is taken** or whether a patient was **visited by a chaplain**.  
[Or learn] to tell the difference between that hospital’s portable chest X-rays with those taken in the radiology department
- “Most AI products have little evidence to support them”
- AI models are mostly tested on computers, not in hospitals or other medical facilities
- Systems developed in one hospital often flop when deployed in a different facility
- Relaxing standards at the FDA (510(k))
- “While it is the job of entrepreneurs to think big and take risks,” Saini said, “it is the job of doctors to protect their patients.”

## A Reality Check On Artificial Intelligence: Are Health Care Claims Overblown?

By [Liz Szabo](#)

Dec. 2019

1 week ago

Artificial Int Health Care:

This present...

Alexander Selvikvåg Lundervold  
@AlexanderSelvi7

"Thirty four (94%) of 36 AI systems evaluated in these studies were less accurate than a single radiologist, and all were less accurate than consensus of two or more radiologists."

bmj.com

Use of artificial intelligence for image analysis in breast can...  
Objective To examine the accuracy of artificial intelligence (AI) for the detection of breast cancer in mammography ...

Health products powered by [artificial intelligence](#), or AI, ar...

## Time to reality check the promise of AI-powered precision medicine

9:30 PM · Sep 2, 2021 · Twitter Web App

Sept. 2020



Jack Wilkinson, Kellyn F Arnold, Eleanor J Murray, Maarten van Smeden, Kareem Carr, Rachel Sippy, Marc de Kamps, Andrew Beam, Stefan Konigorski, Christoph Lippert, Mark S Gilthorpe, Peter W G Tennant

Machine learning methods, combined with large electronic health databases, could enable a personalised approach to medicine through improved diagnosis and prediction of individual responses to therapies. If successful, this strategy would represent a revolution in clinical research and practice. However, although the vision of individually tailored

*Lancet Digital Health* 2020  
Published Online  
September 16, 2020

- "There's nothing that I've seen in my 30-plus years studying medicine that could be as impactful and transformative" as AI, said Dr. Eric Topol
- many health industry experts fear AI-based products won't be able to match the hype.

es learn to make predictions based on s to do with disease than the **brand** of the **time a blood test is taken** or **as visited by a chaplain**.

difference between that hospital's 's with those taken in the radiology

ave little evidence to support them"

y tested on computers, not in hospitals ilities

in one hospital often flop when ent facility

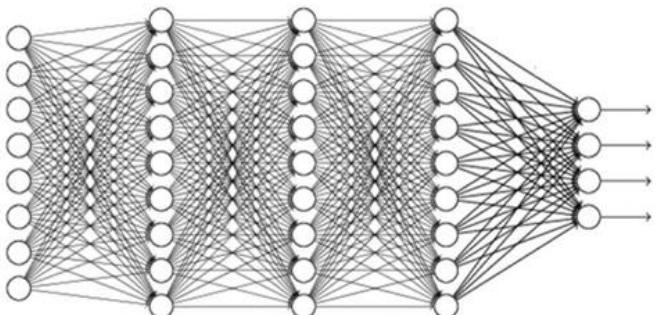
at the FDA (510(k))

f entrepreneurs to think big and take risks, Saini said, "it is the job of doctors to protect their patients."

## What's the status of medical AI?

### From AI Algorithm to Changing Medical Practice

Validate a DNN In Silico



Publish

Clinical Validation  
in Real World Medicine



Publish

FDA, CMS Approval

Implementation in  
Healthcare



but on our way here...

Modified from Eric Topol

# What's the status of medical AI?



Communications  
From the Editor

Algorithm to Changing Me

Clinical Validation  
in Real World Medic

Free Access

## Assessing Radiology Research on Artificial Intelligence: A Brief Guide for Authors, Reviewers, and Readers—From the *Radiology* Editorial Board

David A. Bluemke , Linda Moy, Miriam A. Bredella, Birgit B. Ertl-Wagner, Kathryn J. Fowler, Vicky J. Goh, Elkan F. Halpern, Christopher P. Hess, Mark L. Schiebler, Clifford R. Weiss

but on our way here...

### Key Considerations for Authors, Reviewers, and Readers of AI/ML Manuscripts in Radiology

#### Key Considerations

- Are all three image sets (training, validation, and test sets) defined?
- Is an *external* test set used for final statistical reporting?
- Have multivendor images been used to evaluate the AI algorithm?
- Are the sizes of the training, validation, and test sets justified?
- Was the AI algorithm trained using a standard of reference that is widely accepted in our field?
- Was preparation of images for the AI algorithm adequately described?
- Were the results of the AI algorithm compared with radiology experts and/or pathology?
- Was the manner in which the AI algorithm makes decisions demonstrated?
- Is the AI algorithm publicly available?

Note.—AI = artificial intelligence, ML = machine learning.

28.06.2021

Health informatics  
Protocol

## Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol

Viknes Sounderajah<sup>1, 2</sup>, Hutan Ashrafi<sup>1, 2</sup>, Robert M Golub<sup>3</sup>, Shravya Shetty<sup>4</sup>, Jeffrey De Fauw<sup>5</sup>, Lotty Hooft<sup>6, 7</sup>, Karel Moons<sup>6, 7</sup>, Gary Collins<sup>8</sup>, David Moher<sup>9</sup>, Patrick M Bossuyt<sup>10</sup>, Ara Darzi<sup>1, 2</sup>, Alan Karthikesalingam<sup>11</sup>, Alastair K Denniston<sup>12, 13, 14, 15</sup>, Bilal Akhter Mateen<sup>16</sup>, Daniel Ting<sup>17</sup>, Darren Treanor<sup>18</sup>, Dominic King<sup>19</sup>, Felix Greaves<sup>20</sup>, Jonathan Godwin<sup>5</sup>, Jonathan Pearson-Stuttard<sup>21</sup>, Leanne Harling<sup>1</sup>, Matthew McInnes<sup>22</sup>, Nader Rifai<sup>23</sup>, Nenad Tomasev<sup>5</sup>, Pasha Normahani<sup>1</sup>, Penny Whiting<sup>24</sup>, Ravi Aggarwal<sup>1, 2</sup>, Sebastian Vollmer<sup>16</sup>, Sheraz R Markar<sup>1</sup>, Trishan Panch<sup>25</sup>, Xiaoxuan Liu<sup>12, 13, 14, 15</sup> On behalf of the STARD-AI Steering Committee

## From the Editor



Health Topics ▾

Countries ▾

Newsroom ▾

Emergencies ▾

Data ▾

About WHO ▾

Home / Publications / Overview / Generating Evidence for Artificial Intelligence Based Medical Devices: A Framework for Training Validation and Evaluation

## Generating Evidence for Artificial Intelligence Based Medical Devices: A Framework for Training Validation and Evaluation

Including a use-case study for Cervical Cancer Screening

17 November 2021 | Publication

### Key Considerations for Authors, Reviewers, and Readers of AI/ML Manuscripts in Radiology

#### Key Considerations

Are all three image sets (training, validation, and test sets) defined?

Is an *external* test set used for final statistical reporting?

Have multivendor images been used to evaluate the AI algorithm?

Are the sizes of the training, validation, and test sets justified?

Was the AI algorithm trained using a standard of reference that is widely accepted in our field?

Was preparation of images for the AI algorithm adequately described?

Were the results of the AI algorithm compared with radiology experts and/or pathology?

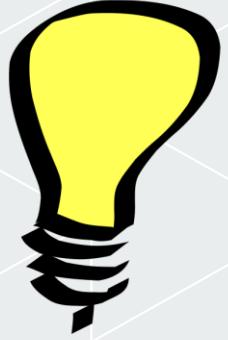
Was the manner in which the AI algorithm makes decisions demonstrated?

Is the AI algorithm publicly available?

Note.—AI = artificial intelligence, ML = machine learning.

$$y \approx f(x; \theta)$$

Who defines the ground truth?  
Is there a ground truth?



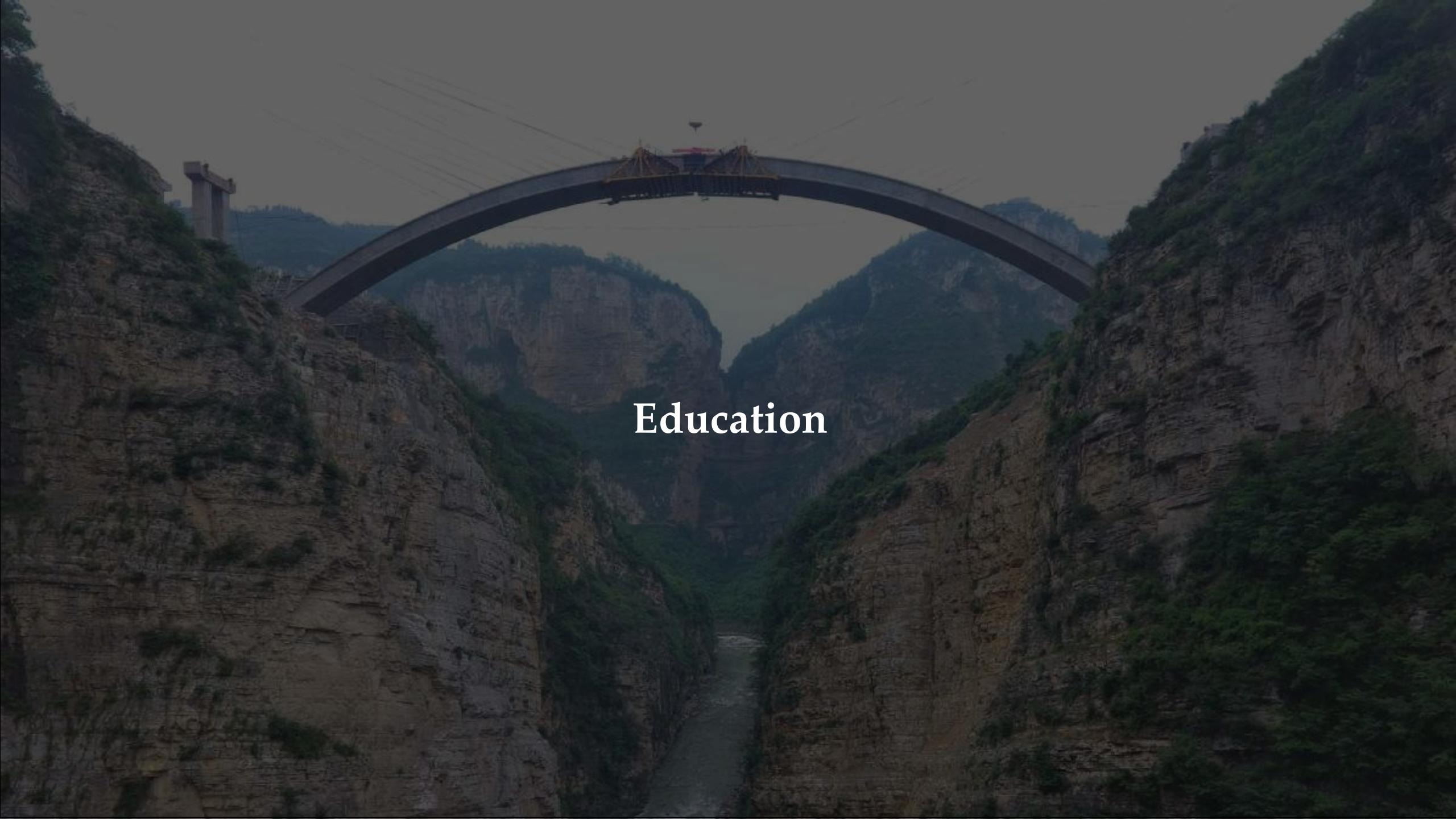
# The development of medical AI should happen with the health-care sector in the driver's seat \*

applications, methods, and infrastructure

Well-placed to *research, evaluate, implement* and *innovate*

\* at least in the front seat

Break



Education



## Nasjonal strategi for kunstig intelligens



Institutions of higher education ought to evaluate how topics with relevance to artificial intelligence can be integrated into their programmes in areas that will be affected by artificial intelligence in the coming years.

### *Need for interdisciplinarity*

Knowledge of artificial intelligence, and related fields such as ethics and data protection associated with AI applications, will be important in study programmes oriented towards the educational sector, health, crime prevention, law and several other fields. For example, an introductory course in artificial intelligence has been established at the University of Bergen, in which students of medicine and bioengineering learn about how AI can be used in clinical practice. One of the aims is to promote interdisciplinary cooperation between physicians and engineers. The universities of Bergen and Oslo also offer courses in artificial intelligence and machine learning oriented towards social scientists.

# AMA Journal of Ethics

*Illuminating the Art of Medicine*

Home Issues Articles Cases Art Multimedia CME

Search



<https://journalofethics.ama-assn.org/article/reimagining-medical-education-age-ai/2019-02>

MEDICAL EDUCATION  
FEB 2019

## Reimagining Medical Education in the Age of AI

Steven A. Wartman, MD, PhD and C. Donald Combs, PhD

### Information Overload

The system for educating medical students is approaching a crisis driven by 2 compelling forces: growing externalization of available medical knowledge outside the minds of physicians and stress-induced mental illness among learners.<sup>2-5</sup>

## AMA Journal of Ethics

*Illuminating the Art of Medicine*

<https://www.journalofethics.org>

MEDICAL EDUCATION  
FEB 2019

## Reimagining Medical Education in the AI

Steven A. Wartman, MD, PhD and C. Donald Combs, PhD

### Information Overload

The system for educating medical students is approaching a crisis driven by the exponential growth of available medical knowledge outside the minds of physicians and stress-

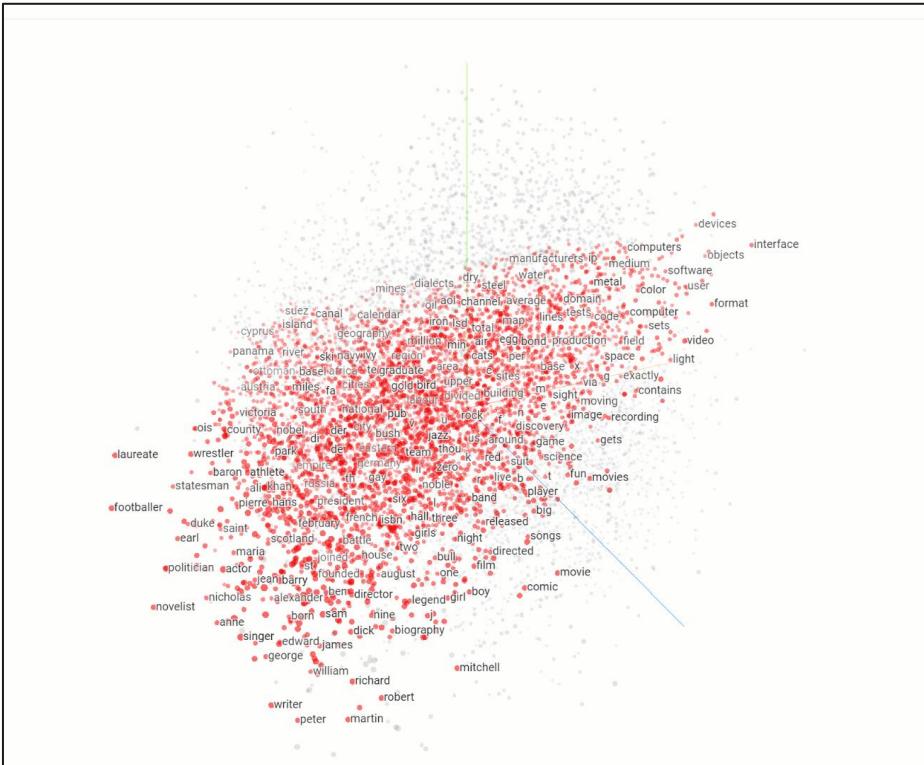
Stanford Medicine 2020 Health Trends Report

## The Rise of the Data-Driven Physician

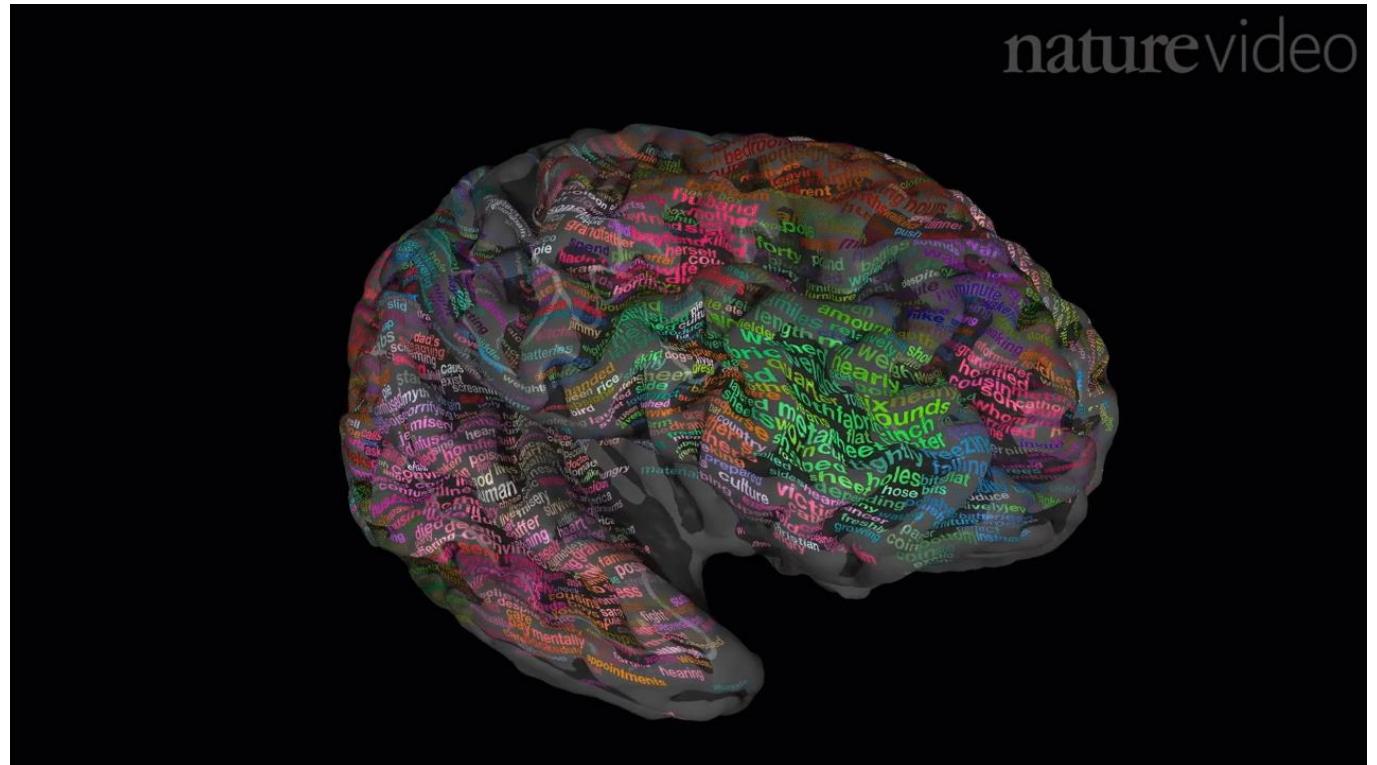
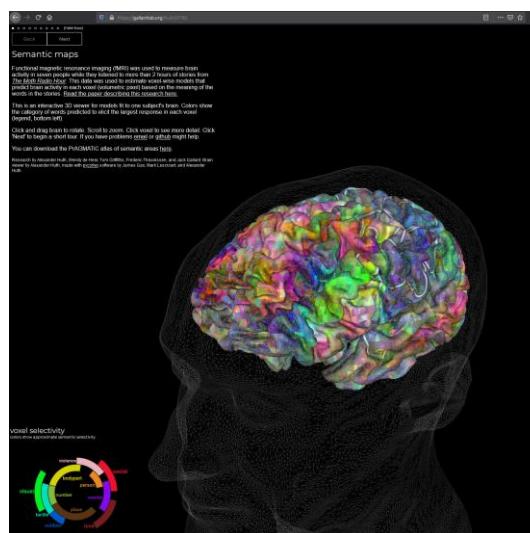


A wide-angle photograph of a massive cable-stayed bridge spanning a deep, narrow canyon. The bridge features two long, sweeping arches supported by numerous cables. It appears to be under construction or recently completed, as evidenced by the yellow construction equipment visible on top of the spans. The canyon walls are steep and composed of layered rock, with patches of green vegetation clinging to them. The sky is overcast and hazy.

# Epistemology



## Representations in brain and machine



*Natural speech reveals the semantic maps that tile human cerebral cortex, Nature, April 2016*

<https://gallantlab.org/huth2016>

[nature](#) > [articles](#) > [article](#)

Article | Open Access | Published: 12 January 2022

# Toroidal topology of population activity in grid cells

Richard J. Gardner , Erik Hermansen, Marius Pachitariu, Yoram Burak, Nils A. Baas , Benjamin A. Dunn  
, May-Britt Moser & Edvard I. Moser 

[Nature](#) (2022)

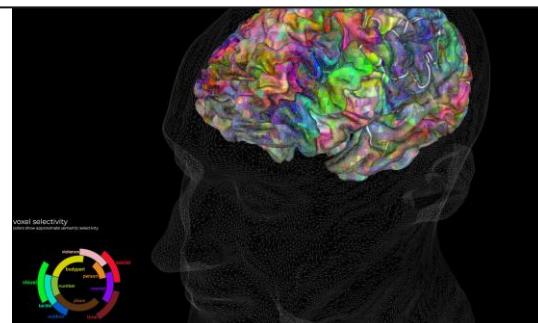
[Cite this article](#)

14k Accesses

241 Altmetric

[Metrics](#)

<https://github.com/erikher/GridCellTorus>

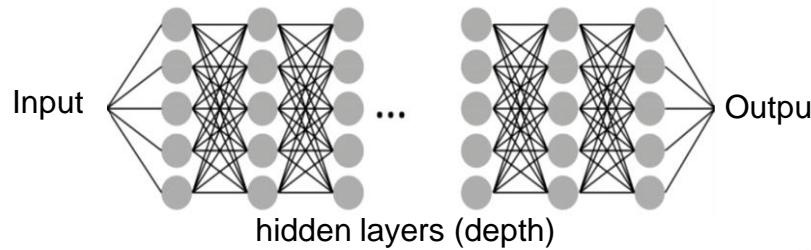


<https://gallantlab.org/huth2016>

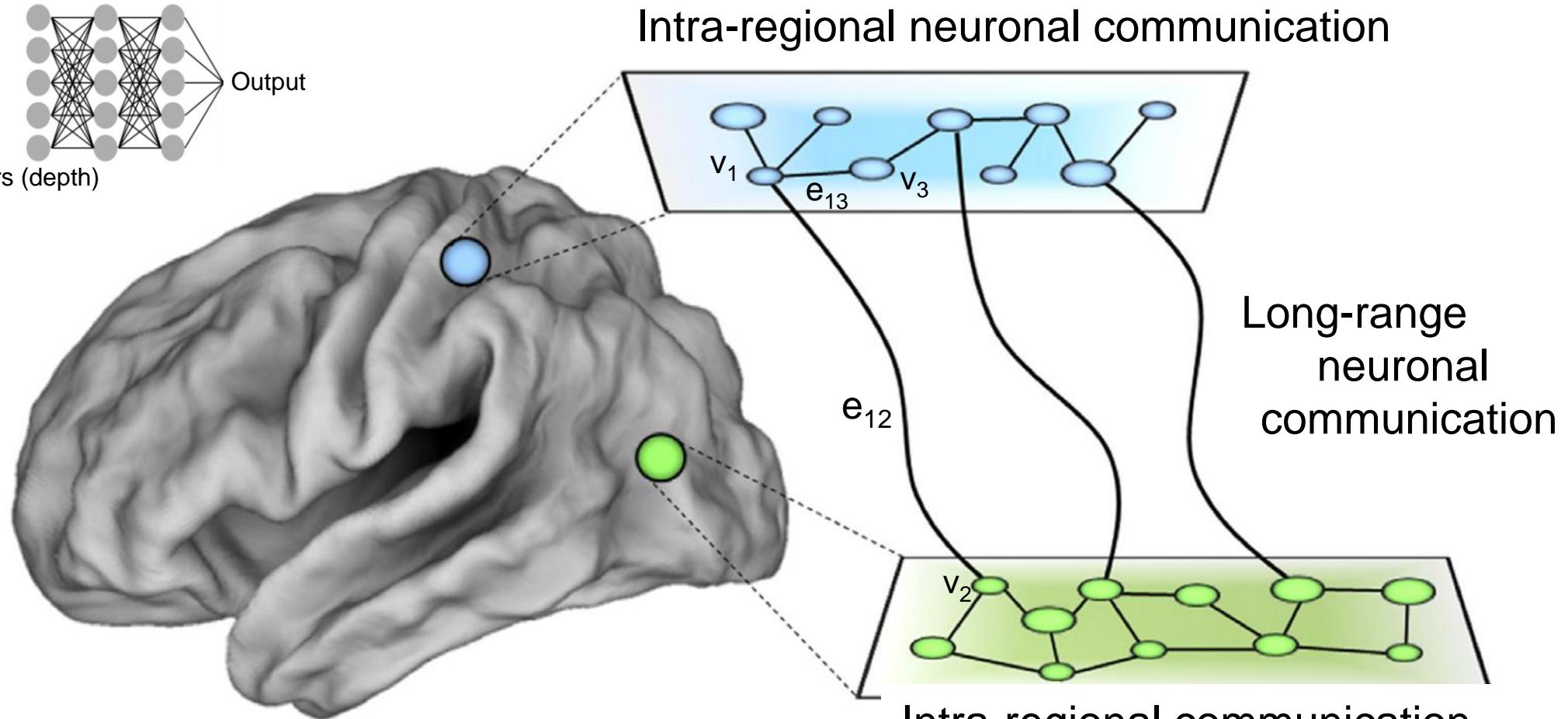
Represen  
brain and

# Neural networks in brain and machine

in machine:



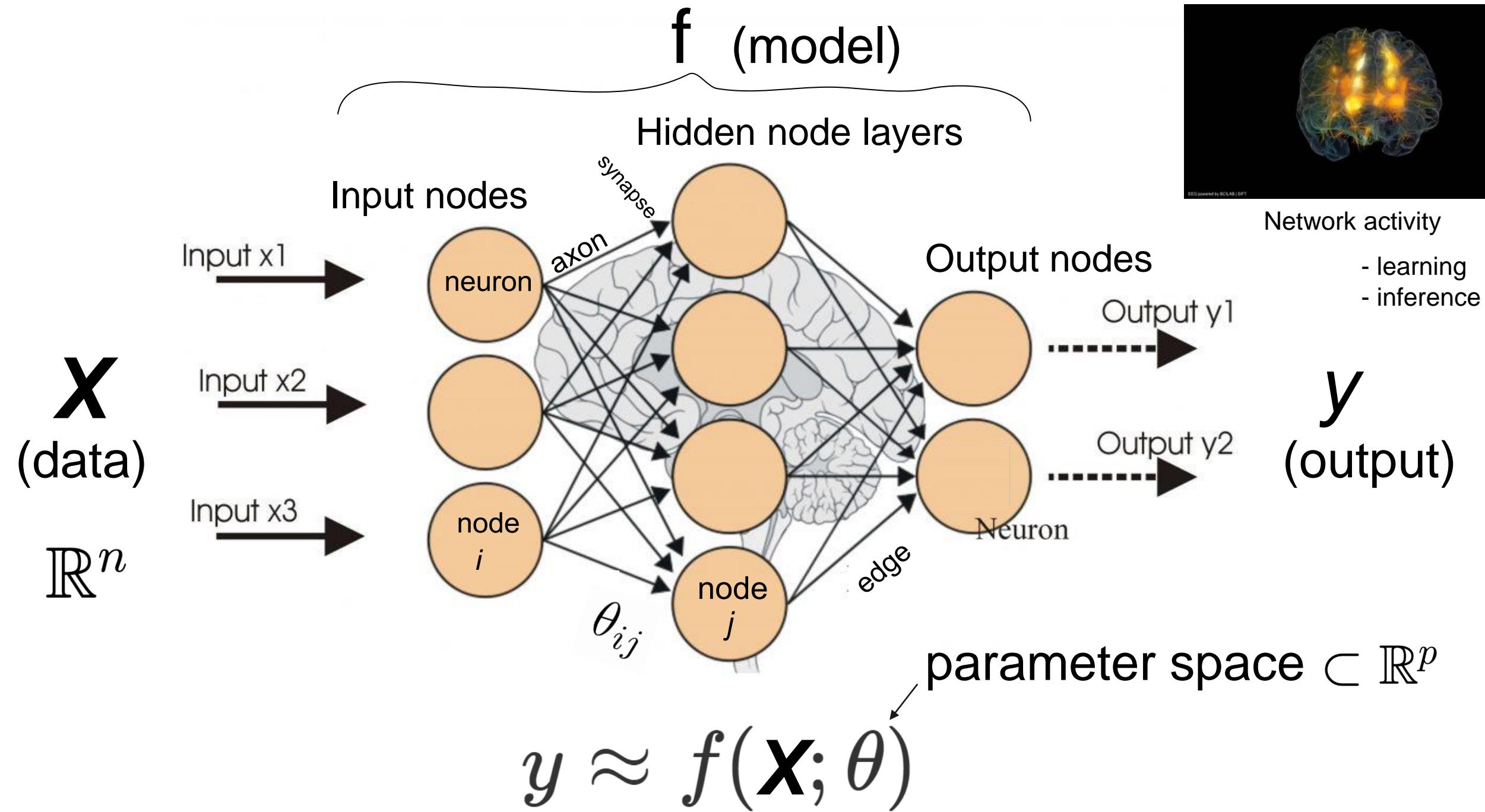
in brain:



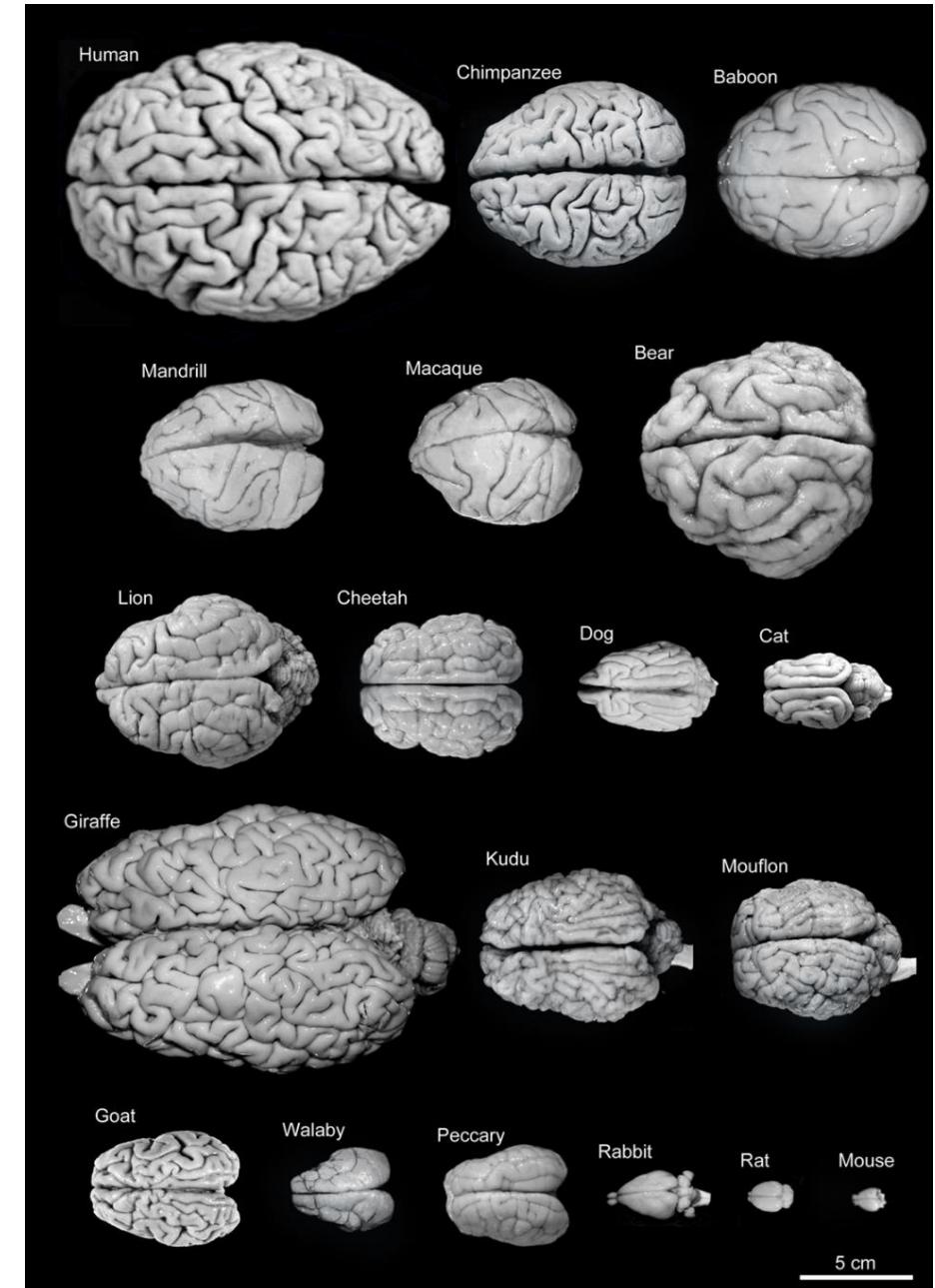
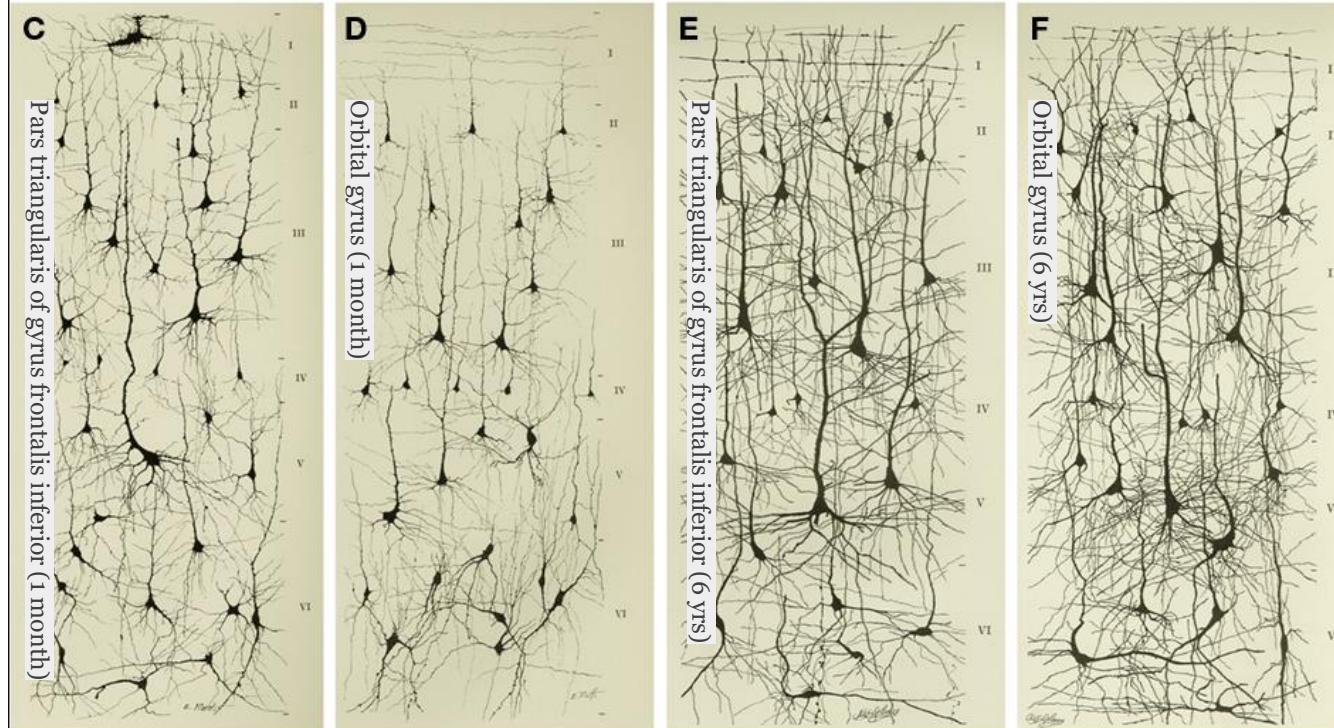
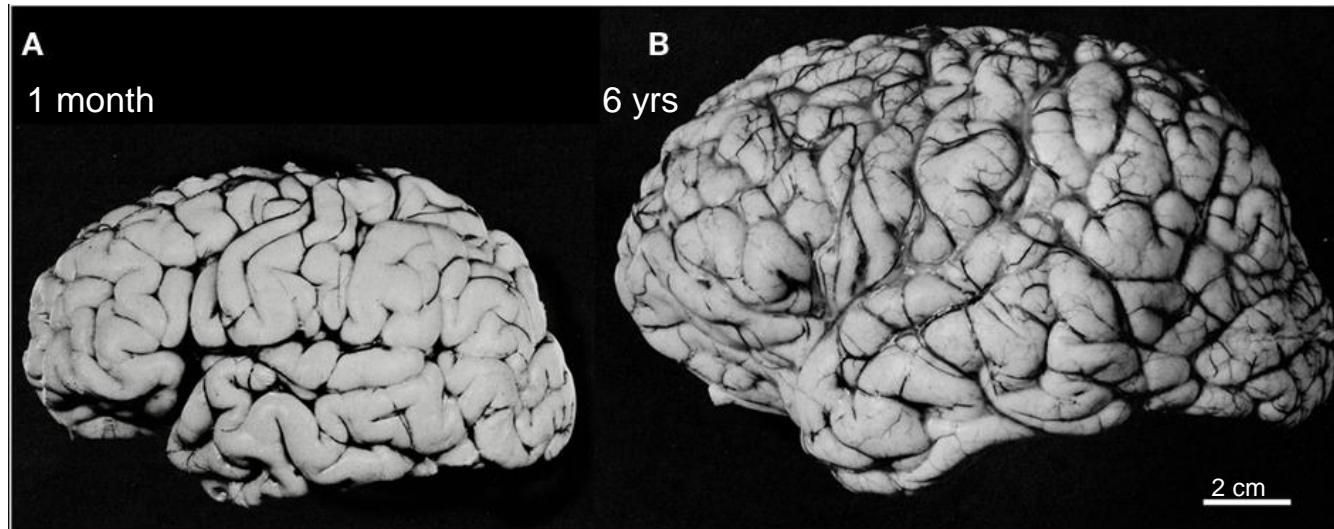
Graph theory:

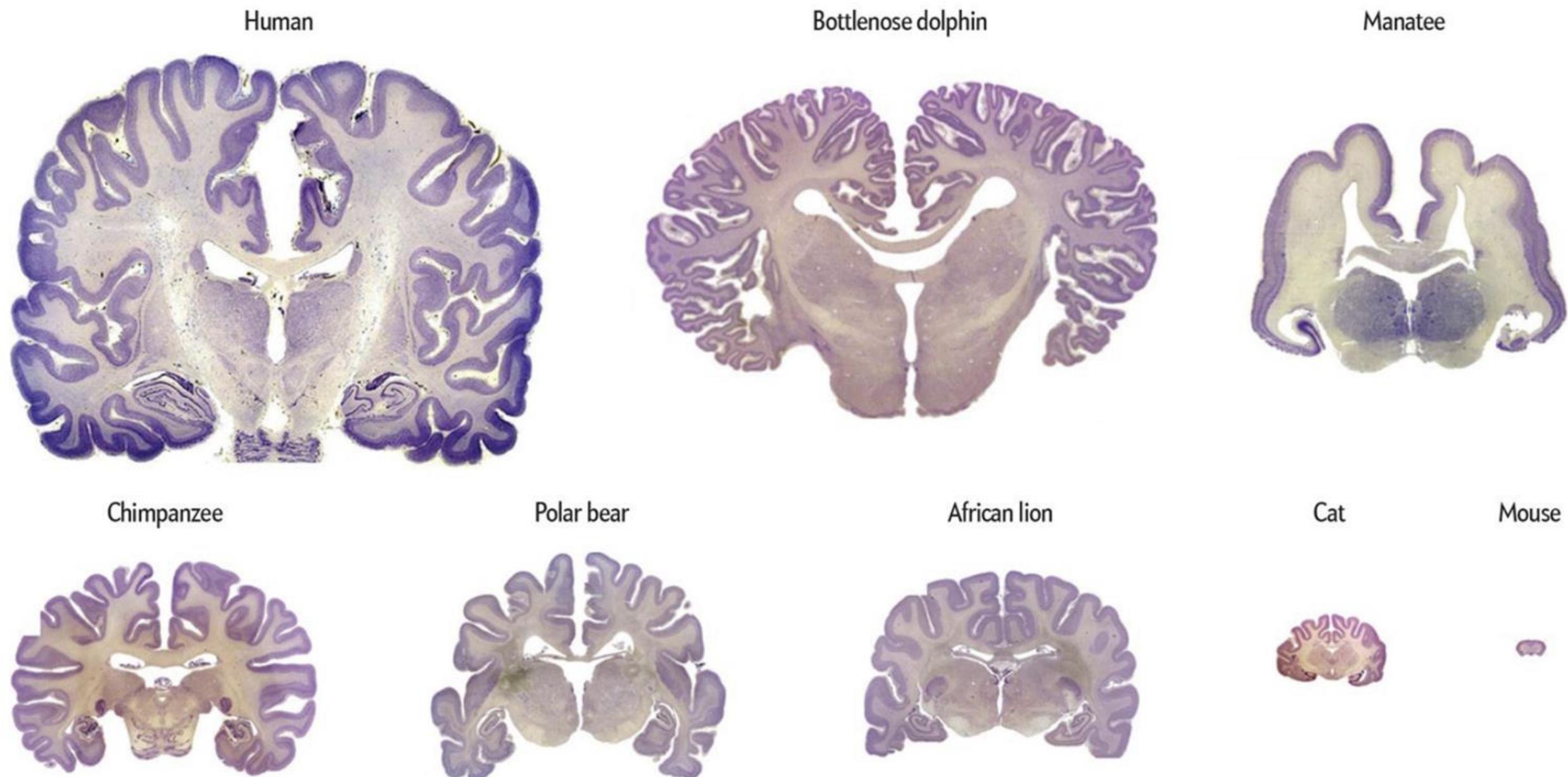
A **graph**  $G$  is an ordered pair  $G=(V, E)$  where  $V$  denote a set of **vertices** (nodes) and  $E \subseteq V \times V$  are the **edges** (arcs, connections) of the graph

e.g.  $v_1, v_2, v_3 \in V$  and  $(v_1, v_2), (v_1, v_3) \in E$

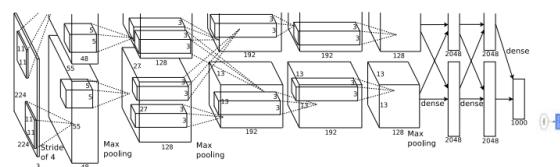


# J. DeFelipe. The evolution of the brain, the human nature of cortical circuits, and intellectual creativity

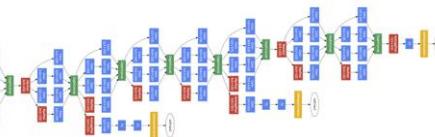




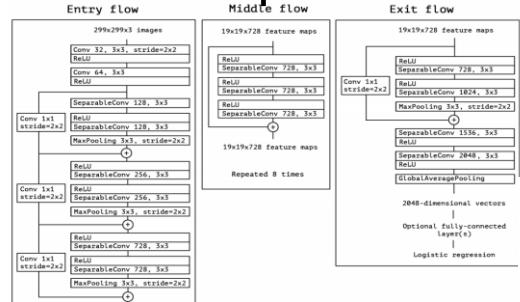
## AlexNet



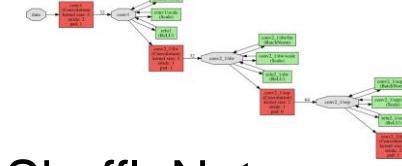
## Inception / GoogLeNet



## Xception

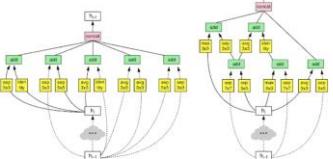


## MobileNet



## ANN architectures

### NASNet/AmoebaNet



## Transformer

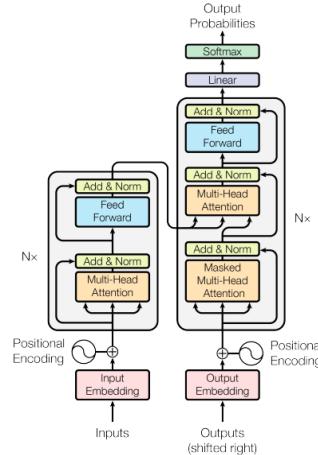
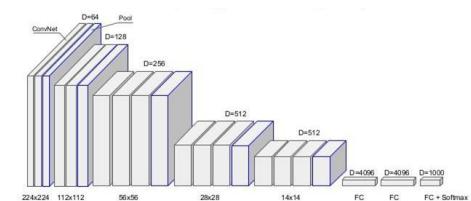


Figure 1: The Transformer - model architecture.

## VGG



## Squeeze and excitation

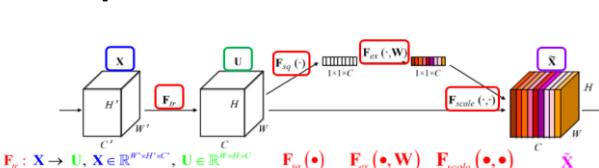
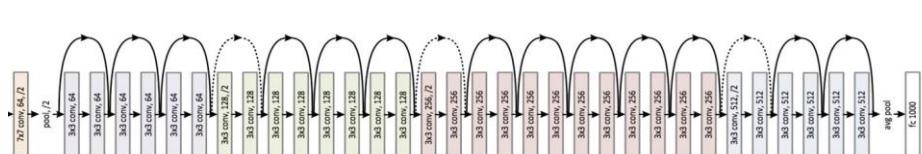
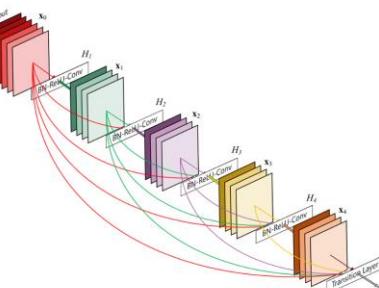


Figure 1. A Squeeze-and-Excitation block.

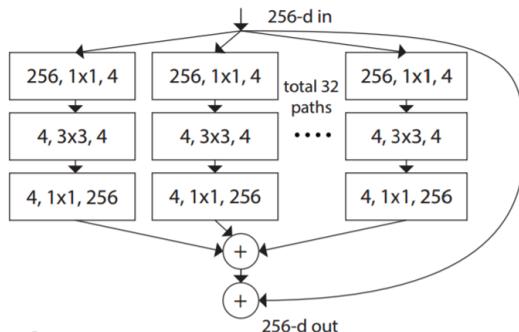
## ResNet



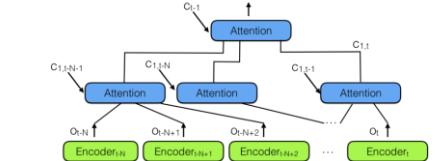
## DenseNet



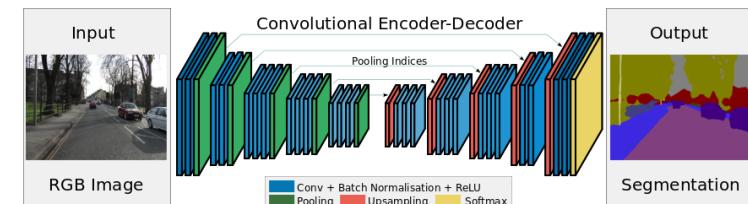
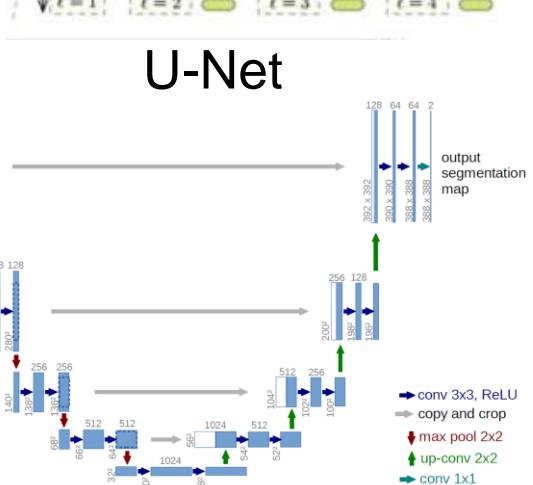
## ResNext



## Hierarchical neural attention encoder



## U-Net

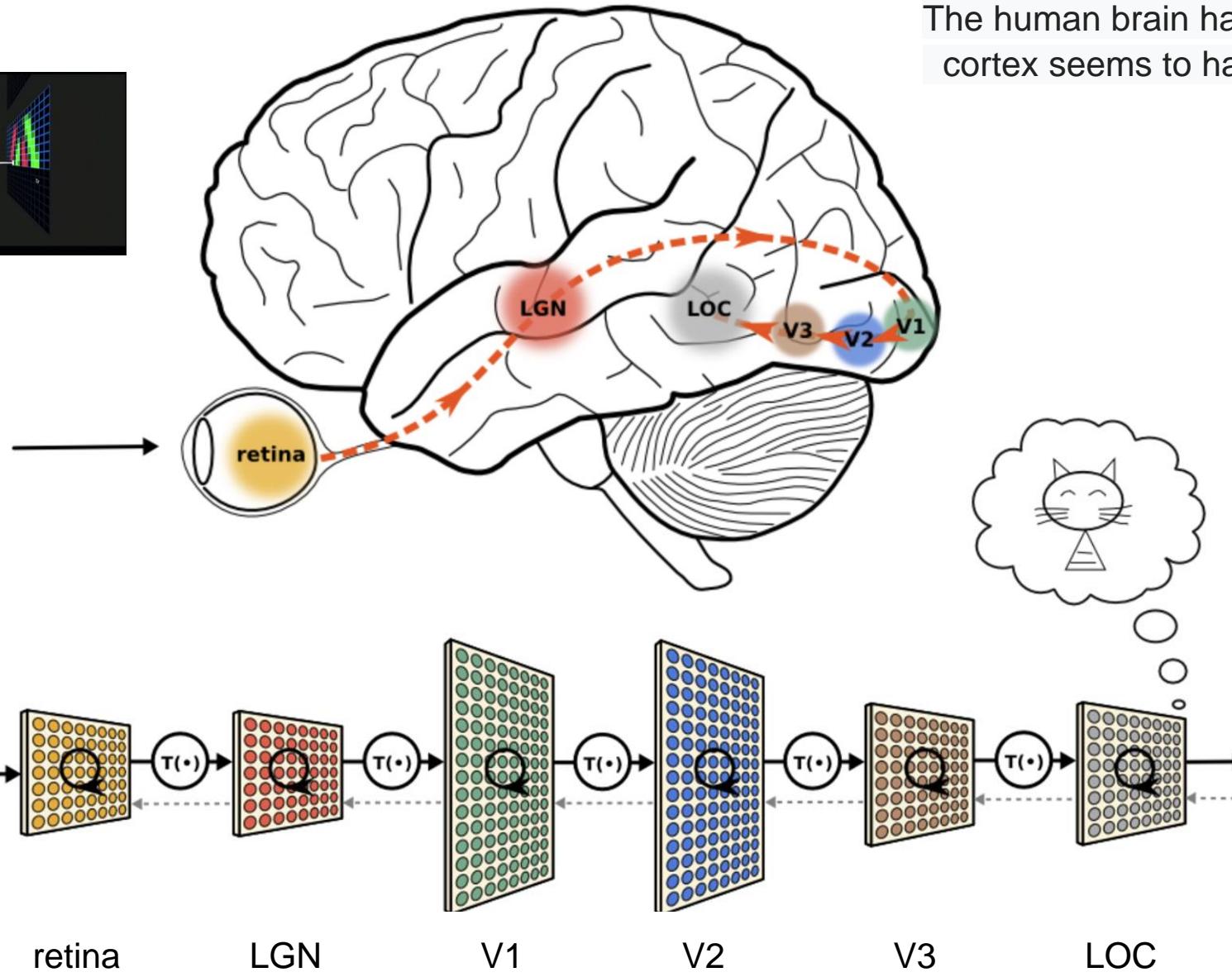
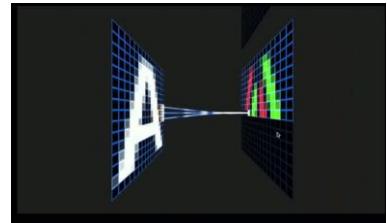


## SegNet

Legend:

- conv 3x3, ReLU
- copy and crop
- max pool 2x2
- up-conv 2x2
- conv 1x1

# Signal pathways from the retina to object detection in the lateral occipital cortex



LGN = [lateral geniculate nucleus](#)

V1 = [primary visual cortex](#)

LOC = [lateral occipital cortex](#)

The human brain has **deep architecture**, where the cortex seems to have a **generic approach to learning**.

A given input is perceived at several **levels of abstraction**. Each level corresponds to a new area of the cortex. We process information in a hierarchical way, with **transformation and representation on several levels**.

Therefore, we learn simple concepts first and then put them together.

This structure of understanding can perhaps be seen most clearly in the human visual system. As shown in the figure, the signaling pathway goes from the retina to the lateral occipital cortex (LOC), which eventually recognizes the object.

The ventral visual cortex (V1, V2, V3) encompasses a set of areas that process images in increasingly abstract ways, from edges, corners and contours, shapes, object-parts to object, so that we can learn, recognize and categorize three-dimensional objects from arbitrary two-dimensional images that hit the retina.

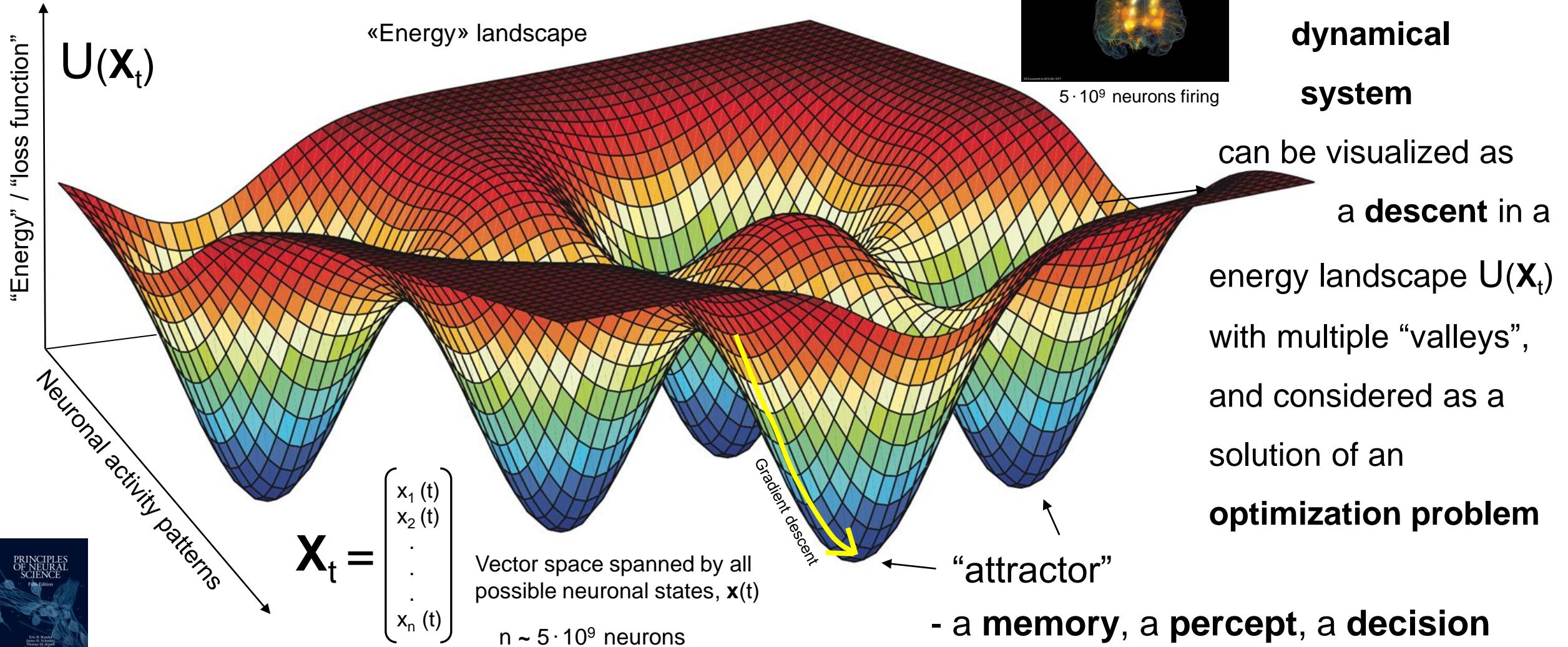
# Neural networks & neurodynamics in brain and machine

... the shape of the landscape (manifold) embedded in  $\mathbb{R}^n$  is dependent of the fitness or loss function and the synaptic weights (learning)

The dynamics of a

**multistable  
dynamical  
system**

can be visualized as  
**a descent** in a  
energy landscape  $U(\mathbf{X}_t)$   
with multiple “valleys”,  
and considered as a  
solution of an  
**optimization problem**



## Neuroscience-Inspired Artificial Intelligence

Dennis Hassabis,<sup>1,2,\*</sup> Dharshan Kumaran,<sup>1,3</sup> Christopher Summerfield,<sup>1,4</sup> and Matthew Botvinick<sup>1,2</sup>

<sup>1</sup>DeepMind, 5 New Street Square, London, UK

<sup>2</sup>Gatsby Computational Neuroscience Unit, 25 Howland Street, London, UK

<sup>3</sup>Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London, UK

<sup>4</sup>Department of Experimental Psychology, University of Oxford, Oxford, UK

\*Correspondence: dhcontact@google.com

<http://dx.doi.org/10.1016/j.neuron.2017.06.011>

The fields of neuroscience and artificial intelligence (AI) have a long and intertwined history. In more recent times, however, communication and collaboration between the two fields has become less commonplace. In this article, we argue that better understanding biological brains could play a vital role in building intelligent machines. We survey historical interactions between the AI and neuroscience fields and emphasize current advances in AI that have been inspired by the study of neural computation in humans and other animals. We conclude by highlighting shared themes that may be key for advancing future research in both fields.

In recent years, rapid progress has been made in the related fields of neuroscience and artificial intelligence (AI). At the dawn of the computer age, work on AI was inextricably intertwined with neuroscience and psychology, and many of the early pioneers straddled both fields, with collaborations between these disciplines proving highly productive (Churchland and Sejnowski, 1988; Hebb, 1949; Hinton et al., 1986; Hopfield, 1982; McCulloch and Pitts, 1943; Turing, 1950). However, more recently, the interaction has become much less commonplace, as both subjects have grown enormously in complexity and disciplinary boundaries have solidified. In this review, we argue for the critical and ongoing importance of neuroscience in generating ideas that will accelerate and guide AI research (see Hassabis commentary in Brooks et al., 2012).

We begin with the premise that building human-level general AI (or “Turing-powerful” intelligent systems; Turing, 1936) is a daunting task, because the search space of possible solutions is vast and likely only very sparsely populated. We argue that this therefore underscores the utility of scrutinizing the inner workings of the human brain—the only existing proof that such an intelligence is even possible. Studying animal cognition and its neural implementation also has a vital role to play, as it can provide a window into various important aspects of higher-level general intelligence.

The benefits to developing AI of closely examining biological intelligence are two-fold. First, neuroscience provides a rich source of *inspiration* for new types of algorithms and architectures, independent of and complementary to the mathematical and logic-based methods and ideas that have largely dominated traditional approaches to AI. For example, were a new facet of biological computation found to be critical to supporting a cognitive function, then we would consider it an excellent candidate for incorporation into artificial systems. Second, neuroscience can provide *validation* of AI techniques that already exist. If a known algorithm is subsequently found to be implemented in the brain, then that is strong support for its plausibility as an integral component of an overall general intelligence system. Such clues can be critical to a long-term research program when determining where to allocate resources most productively.

For example, if an algorithm is not quite attaining the level of performance required or expected, but we observe it is core to the functioning of the brain, then we can surmise that redoubled engineering efforts geared to making it work in artificial systems are likely to pay off.

Of course from a practical standpoint of building an AI system, we need not slavishly enforce adherence to biological plausibility. From an engineering perspective, what works is ultimately all that matters. For our purposes then, biological plausibility is a guide, not a strict requirement. What we are interested in is a systems neuroscience-level understanding of the brain, namely the algorithms, architectures, functions, and representations it utilizes. This roughly corresponds to the top two levels of the three levels of analysis that Marr famously stated are required to understand any complex biological system (Marr and Poggio, 1976): the goals of the system (the computational level) and the process and computations that realize this goal (the algorithmic level). The precise mechanisms by which this is physically realized in a biological substrate are less relevant here (the implementation level). Note this is where our approach to neuroscience-inspired AI differs from other initiatives, such as the Blue Brain Project (Markram, 2006) or the field of neuromorphic computing systems (Esser et al., 2016), which attempt to closely mimic or directly reverse engineer the specifics of neural circuits (albeit with different goals in mind). By focusing on the computational and algorithmic levels, we gain transferable insights into general mechanisms of brain function, while leaving room to accommodate the distinctive opportunities and challenges that arise when building intelligent machines *in silico*.

The following sections unpack these points by examining the past, present, and future of the AI-neuroscience interface. Before beginning, we offer a clarification. Throughout this review, we employ the terms “neuroscience” and “AI.” We use these terms in the widest possible sense. When we say neuroscience, we mean to include all fields that are involved with the study of the brain, the behaviors that it generates, and the mechanisms by which it does so, including cognitive neuroscience, systems neuroscience and psychology. When we say AI, we mean work

## A deep learning framework for neuroscience

Blake A. Richards<sup>1,2,3,4,42\*</sup>, Timothy P. Lillicrap<sup>5,6,42</sup>, Philippe Beaudoin<sup>7</sup>, Yoshua Bengio<sup>1,4,8</sup>, Rafal Bogacz<sup>9</sup>, Amelia Christensen<sup>10</sup>, Claudia Clopath<sup>11</sup>, Rui Ponte Costa<sup>12,13</sup>, Archy de Berker<sup>7</sup>, Surya Ganguli<sup>14,15</sup>, Colleen J. Gillon<sup>16,17</sup>, Danijar Hafner<sup>18,19</sup>, Adam Kepes<sup>20</sup>, Nikolaus Kriegeskorte<sup>21,22</sup>, Peter Latham<sup>23</sup>, Grace W. Lindsay<sup>22,24</sup>, Kenneth D. Miller<sup>22,24,25</sup>, Richard Naud<sup>26,27</sup>, Christopher C. Pack<sup>3</sup>, Panayiota Poirazi<sup>28</sup>, Pieter Roelfsema<sup>29</sup>, João Sacramento<sup>30</sup>, Andrew Saxe<sup>31</sup>, Benjamin Scellier<sup>1,8</sup>, Anna C. Schapiro<sup>32</sup>, Walter Senn<sup>13</sup>, Greg Wayne<sup>5</sup>, Daniel Yamins<sup>33,34,35</sup>, Friedemann Zenke<sup>36,37</sup>, Joel Zylberman<sup>4,38,39</sup>, Denis Therien<sup>40</sup>, and Konrad P. Kording<sup>4,40,41,42</sup>

Systems neuroscience seeks explanations for how the brain implements a wide variety of perceptual, cognitive and motor tasks. Conversely, artificial intelligence attempts to design computational systems based on the tasks they will have to solve. In artificial neural networks, the three components specified by design are the objective functions, the learning rules and the architectures. With the growing success of deep learning, which utilizes brain-inspired architectures, these three designed components have increasingly become central to how we model, engineer and optimize complex artificial learning systems. Here we argue that a greater focus on these components would also benefit systems neuroscience. We give examples of how this optimization-based framework can drive theoretical and experimental progress in neuroscience. We contend that this principled perspective on systems neuroscience will help to generate more rapid progress.

**M**ajor technical advances are revolutionizing our ability to observe and manipulate brains at a large scale and to quantify complex behaviors<sup>1</sup>. How should we use this data to develop models of the brain? When the classical framework for systems neuroscience was developed, we could only record from small sets of neurons. In this framework, a researcher observes neural activity, develops a theory of what individual neurons compute, then assembles a circuit-level theory of how the neurons combine their operations. This approach has worked well for simple computations. For example, we know how central pattern generators control rhythmic movements<sup>2</sup>, how the vestibulo-ocular reflex promotes gaze stabilization<sup>3</sup> and how the retina computes motion<sup>4</sup>. But can this classical framework scale up to recordings of thousands of neurons and all of the behaviors that we may wish to account

for? Arguably, we have not had as much success with the classical approach in large neural circuits that perform a multitude of functions, like the neocortex or hippocampus. In such circuits, researchers often find neurons with response properties that are difficult to summarize in a succinct manner<sup>5</sup>.

The limitations of the classical framework suggest that new approaches are needed to take advantage of experimental advances. A promising framework is emerging from the interactions between neuroscience and artificial intelligence (AI)<sup>6–10</sup>. The rise of deep learning as a leading machine-learning method invites us to revisit artificial neural networks (ANNs). At their core, ANNs model neural computation using simplified units that loosely mimic the integration and activation properties of real neurons<sup>11</sup>. Units are implemented with varying degrees of abstraction, ranging from

<sup>1</sup>Mila, Montréal, Québec, Canada. <sup>2</sup>School of Computer Science, McGill University, Montréal, Québec, Canada. <sup>3</sup>Department of Neurology & Neurosurgery, McGill University, Montréal, Québec, Canada. <sup>4</sup>Canadian Institute for Advanced Research, Toronto, Ontario, Canada. <sup>5</sup>DeepMind, Inc., London, UK. <sup>6</sup>Centre for Computation, Mathematics and Physics in the Life Sciences and Experimental Biology, University College London, London, UK. <sup>7</sup>Element AI, Montréal, QC, Canada. <sup>8</sup>Université de Montréal, Montréal, Québec, Canada. <sup>9</sup>MRC Brain Network Dynamics Unit, University of Oxford, Oxford, UK. <sup>10</sup>Department of Electrical Engineering, Stanford University, Stanford, CA, USA. <sup>11</sup>Department of Bioengineering, Imperial College London, London, UK. <sup>12</sup>Computational Neuroscience Unit, School of Computer Science, Electrical and Electronic Engineering, and Engineering Maths, University of Bristol, Bristol, UK. <sup>13</sup>Department of Physiology, Universität Bern, Bern, Switzerland. <sup>14</sup>Department of Applied Physics, Stanford University, Stanford, CA, USA. <sup>15</sup>Google Brain, Mountain View, CA, USA. <sup>16</sup>Department of Biological Sciences, University of Toronto Scarborough, Toronto, Ontario, Canada. <sup>17</sup>Department of Cell & Systems Biology, University of Toronto, Toronto, Ontario, Canada. <sup>18</sup>Department of Computer Science, University of Toronto, Toronto, Ontario, Canada. <sup>19</sup>Vector Institute, Toronto, Ontario, Canada. <sup>20</sup>Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA. <sup>21</sup>Department of Psychology and Neuroscience, Columbia University, New York, NY, USA. <sup>22</sup>Zuckerman Mind Brain Behavior Institute, Columbia University, New York, New York, USA. <sup>23</sup>Computational Neuroscience Unit, University College London, London, UK. <sup>24</sup>Center for Theoretical Neuroscience, Columbia University, New York, NY, USA. <sup>25</sup>Department of Neuroscience, College of Physicians and Surgeons, Columbia University, New York, NY, USA. <sup>26</sup>University of Ottawa Brain Institute, Ottawa, Ontario, Canada. <sup>27</sup>Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario, Canada. <sup>28</sup>Institute of Biotechnology (IMBB), Foundation for Research and Technology-Hellas (FORTH), Heraklion, Crete, Greece. <sup>29</sup>Department of Psychology and Neuroscience, University of Amsterdam, Amsterdam, Netherlands. <sup>30</sup>Institute of Neuroinformatics, ETH Zürich and University of Zurich, Zürich, Switzerland. <sup>31</sup>Department of Experimental Psychology, University of Oxford, Oxford, UK. <sup>32</sup>Department of Psychology, Stanford University, Stanford, CA, USA. <sup>33</sup>Department of Computer Science, Stanford University, Stanford, CA, USA. <sup>34</sup>Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland. <sup>35</sup>Centre for Neural Circuits and Behaviour, University of Oxford, Oxford, UK. <sup>36</sup>Department of Physics and Astronomy York University, North York, Ontario, Canada. <sup>37</sup>Center for Vision Research, York University, Toronto, Ontario, Canada. <sup>38</sup>Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, USA. <sup>39</sup>Department of Neuroscience, University of Pennsylvania, Philadelphia, PA, USA. <sup>40</sup>These authors contributed equally. <sup>41</sup>Blake A. Richards, Timothy P. Lillicrap, Denis Therien, Konrad P. Kording. <sup>42</sup>e-mail: blake.richards@mcgill.ca

A wide-angle photograph of a massive cable-stayed bridge spanning a deep, narrow canyon. The bridge features two long, sweeping arches supported by tall pylons, with numerous cables connecting the arches to the deck. The canyon walls are steep and made of layered rock, with patches of green vegetation. The sky is overcast and hazy.

Ethics



The AI HLEG is an independent expert group that was set up by the European Commission in June 2018. Contact Nathalie Smuha - AI HLEG Coordinator E-mail CNECT-HLG-AI@ec.europa.eu European Commission B-1049 Brussels Document made public on 8 April 2019.

Trustworthy AI has **three components**, which should be met throughout the system's entire life cycle:

1. it should be **lawful**, complying with all applicable laws and regulations
2. it should be **ethical**, ensuring adherence to ethical principles and values
3. it should be **robust**, both from a technical and social perspective, since, even with good intentions, AI systems can cause unintentional harm.

<https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines>



# Ethical Principles in the Context of AI Systems

## i. Respect for human autonomy



Humans interacting with AI systems must be able to keep **full and effective self-determination over themselves**, and be able to partake in the democratic process. The allocation of functions between humans and AI systems should follow **human-centric design principles** and leave meaningful opportunity for human choice.

## ii. Prevention of harm



AI systems should neither cause nor exacerbate harm or otherwise adversely affect human beings. This entails the protection of **human dignity** as well as **mental and physical integrity**. Vulnerable persons should receive greater attention and be **included in the development, deployment** and use of AI systems. Preventing harm also entails consideration of the **natural environment** and all living beings.

## iii. Fairness



Ensuring equal and just distribution of both benefits and costs, and ensuring that **individuals and groups are free from unfair bias, discrimination** and **stigmatisation**. Equal opportunity in terms of **access to education**, goods, services and technology should also be fostered.

## iv. Explicability



This means that processes need to be **transparent** (incl. **traceability**), the capabilities and purpose of AI systems openly communicated, and decisions – to the extent possible – **explainable** to those directly and indirectly affected. The degree to which explicability is needed is highly dependent on the context and the severity of the consequences if that output is erroneous or otherwise inaccurate.

# How can this be implemented in a concrete medical domain?



# Ethics of AI in Radiology

## European and North American Multisociety Statement

- Ethics of AI in Radiology. European and North America Multisociety Statement 2019 [<https://www.acr.org/-/media/ACR/Files/Informatics/Ethics-of-AI-in-Radiology-European-and-North-American-Multisociety-Statement--6-13-2019.pdf>]

Key Considerations for Authors, Reviewers, and Readers of AI/ML Manuscripts in Radiology	
Key Considerations	
Are all three image sets (training, validation, and test sets) defined?	
Is an <i>external</i> test set used for final statistical reporting?	
Have multivendor images been used to evaluate the AI algorithm?	
Are the sizes of the training, validation, and test sets justified?	
Was the AI algorithm trained using a standard of reference that is widely accepted in our field?	
Was preparation of images for the AI algorithm adequately described?	
Were the results of the AI algorithm compared with radiology experts and/or pathology?	
Was the manner in which the AI algorithm makes decisions demonstrated?	
Is the AI algorithm publicly available?	
Note.—AI = artificial intelligence, ML = machine learning.	

Summary	6
Introduction	8
About this Statement	10
Ethics of Data	11
Clinical radiology data	12
Business operational and analytic data	13
Pre-training, synthetic, and augmented data	13
Raw image data	14
Data ownership	14
Data sharing and data use	16
Data privacy	20
Bias and data	21
Data labeling and ground truth	23
Ethics of Algorithms and Trained Models	25
Algorithm selection	26
Algorithm training	27
Model evaluation and testing	27
Transparency, interpretability, and explainability	29
Open source software	31
Replicability	31
Algorithm bias	31
Security	32
Ethics of Practice	32
Computer - human interaction: Keeping humans in the loop	33
Education	35
Automation bias	35
Patient preferences	36
Traceability	37
AI and workforce disruption	37
Resource inequality	38
Liability	39
Conflicts of interest	40



Search or jump to...

/

Pulls Issues Marketplace Explore



# The Institute for Ethical Machine Learning

The Institute for Ethical Machine Learning is a think-tank that brings together technology leaders, policymakers & academics to develop standards for ML.

📍 United Kingdom

🔗 <http://ethical.institute>

🐦 @EthicalML

✉️ a@ethical.institute

⌂ Overview

/repos 16

📦 Packages

👤 People

/projects



Search or jump to...

Pulls Issues Marketplace



# The Institute for Ethical Machine Learning

The Institute for Ethical Machine Learning is a think-tank that brings together policymakers & academics to develop standards for ML.

📍 United Kingdom

🔗 <http://ethical.institute>

🐦 @EthicalML



## 1. Human augmentation

I commit to assess the impact of incorrect predictions and, when reasonable, design systems with human-in-the-loop review processes

## 2. Bias evaluation

I commit to continuously develop processes that allow me to understand, document and monitor bias in development and production.

## 3. Explainability by justification

I commit to develop tools and processes to continuously improve transparency and explainability of machine learning systems where reasonable.

## 4. Reproducible operations

I commit to develop the infrastructure required to enable for a reasonable level of reproducibility across the operations of ML systems.



## 5. Displacement strategy

I commit to identify and document relevant information so that business change processes can be developed to mitigate the impact towards workers being automated.



## 6. Practical accuracy

I commit to develop processes to ensure my accuracy and cost metric functions are aligned to the domain-specific applications.



## 7. Trust by privacy

I commit to build and communicate processes that protect and handle data with stakeholders that may interact with the system directly and/or indirectly.



## 8. Data risk awareness

I commit to develop and improve reasonable processes and infrastructure to ensure data and model security are being taken into consideration during the development of machine learning systems.

README.md

## Awesome production machine learning

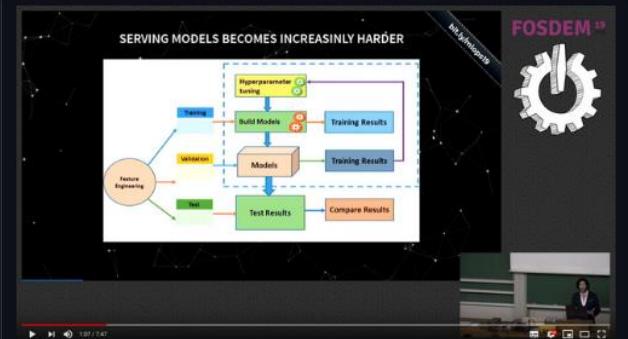
This repository contains a curated list of awesome open source libraries that will help you deploy, monitor, version, scale, and secure your production machine learning 🚀

### Quick links to sections in this page

🔍 Explaining predictions & models	🔒 Privacy preserving ML	📜 Model & data versioning
🚩 Model Training Orchestration	💪 Model Serving and Monitoring	🤖 Neural Architecture Search
💻 Reproducible Notebooks	📊 Visualisation frameworks	💡 Industry-strength NLP
💻 Data pipelines & ETL	🏷️ Data Labelling	📅 Metadata Management
🔧 Functions as a service	🌐 Computation distribution	💻 Model serialisation
📊 Optimized computation frameworks	🌟 Data Stream Processing	🔴 Outlier and Anomaly Detection
🌐 Feature engineering	🎁 Feature Stores	✖️ Adversarial Robustness
📝 Commercial Platforms	💻 Data Storage Optimization	

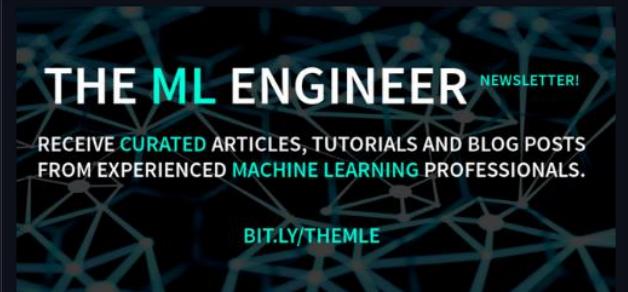
### 10 Min Video Overview

This 10 minute video provides an overview of the motivations for machine learning operations as well as a high level overview on some of the tools in this repo.



Want to receive recurrent updates on this repo and other advancements?

You can join the [Machine Learning Engineer](#) newsletter. You will receive updates on open source frameworks, tutorials and articles curated by machine learning professionals.







## The objectives of ELMED219

- **Understanding** of computational methods, machine learning and AI
- **Practical experience**
- **Insight** into possibilities and challenges for AI in the future of medicine

## ELMED219-2022 (team-based project)

*Precision medicine and quantitative imaging in glioblastoma*

### Description

Imagine that you are a group of established successful scientists that will team up to tackle an important biomedical and medical challenge. There is an open call for research proposals under a new umbrella program entitled "Artificial intelligence and computational (bio)medicine", where your multidisciplinary group are aiming for a project on "Precision medicine and quantitative imaging in glioblastoma - a multiscale approach".

The focus of the assignment is (i) description of relevant imaging technologies and modalities - possibly at different scales, (ii) proposal of imaging-derived biomarkers for glioblastoma, (iii) machine learning techniques for segmentation, classification, treatment stratification and prediction, (iv) the novelty and expected impact of your approach, and (v) the evaluation of the ethics of your project together with a data management plan (and not so much the basic science of brain tumors per se).

### Organization of your report

Fri, Jan 21

10:15-12:00 Perspectives on medical AI: innovation, education, epistemology, ethics and impact

*Arvid Lundervold / Alexander Selvikvåg Lundervold*

Mon, Jan 24

--- 17:00 --- Team project report due

Wed, Jan 26

10:15-12:00 Team project presentations (15 min)

*"Why should this project be funded?"*

Fri, Jan 28

09:00-11:00 Digital home exam (using Inspera)



# Eksamens

Digital hjemme-eksamen på Inspera.

Fredag 28. januar klokken 9 til 11.

## Mulige eksamsoppgaver

8 Ubalanserte klasser

Anta du skal løse et binært klassifikasjonsproblem med svært ubalanerte klasser. Den hyppigst forekommende klasse opptrer i 99 % av treningsdata. Modellen din viser 99 % korrekthet (accuracy) etter prediksjon på testdata. Hvilket av følgende utsagn er riktig i dette tilfellet?

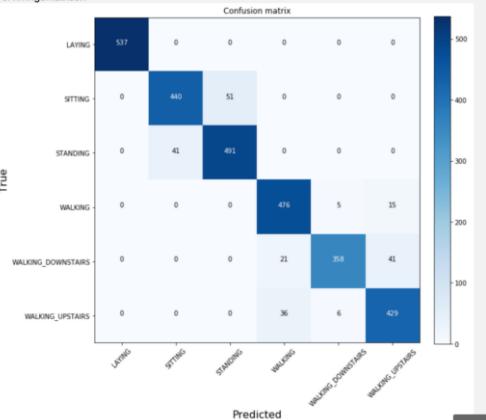
1. Accuracy-metrikken er ikke godt egnet i situasjoner med ubalanserte klasser
  2. Accuracy er godt egnet i situasjoner med ubalanserte klasser
  3. Presisjon (precision) og gjennkalling (recall) er godt egnet ved ubalanserte klasser
  4. Precision og recall er ikke velegnede metrikker i situasjoner med ubalanserte klasser

## **Velg ett alternativ**

- Alternativ 2 og 4
  - Alternativ 2 og 3
  - Alternativ 1 og 3
  - Alternativ 1 og 4

## 27 Forvirringsmatrise 2

### Förvärningsmatrisen



20 Dry lab

Angi svaralternativet som best dekker begrepet "dry lab" (tørr-lab) i kontekst av ELMED219.

## **Velg ett alternativ**

- Laboratorium med tørt inneklima, forenlig med drift av servere og tungt datautstyr
  - Laboratorium for fotografisk fremstilling (photo printing) uten bruk av "våte" kjemikalier
  - Laboratorium som utfører analyse, modellering og computer-basert simulering av biologiske eller fysiske målinger og systemer
  - Laboratorium som ikke tar i bruk kjemikalier eller "våte" reagenser i eksperimentelle studier, men bruker elektroniske sensorer (chips) til sekvensering, imaging o.l.

23 Generalisering

En maskinlæringsmodell kan ha **generaliseringsevne** ("generalization"). Hva innebærer dette? (velg det svar du synes er mest korrekt)

## **Velg ett alternativ**

- Oppnå perfekt accuracy på treningsdata
  - Ha god ytelse på nye usette data
  - Overtilpasser treningsdata
  - Kunne brukes i både styrt og ikke-styrt læring, samt i reinforcement learning

32 Explainable AI

Utdyp begrepet "explainable AI" (XAI) og gi eksempler relevant i medisinske anvendelser.

**Skriv ditt svar her**

# ELMED219 - 2022

Talk to you on Discord and see you for the  
team project presentations on Wednesday  
Jan 26 at 10.15

Perspectives on medical AI: innovation,  
education, epistemology, ethics and impact