Data Harmonization for Robust and Generalizable Artificial Intelligence Models

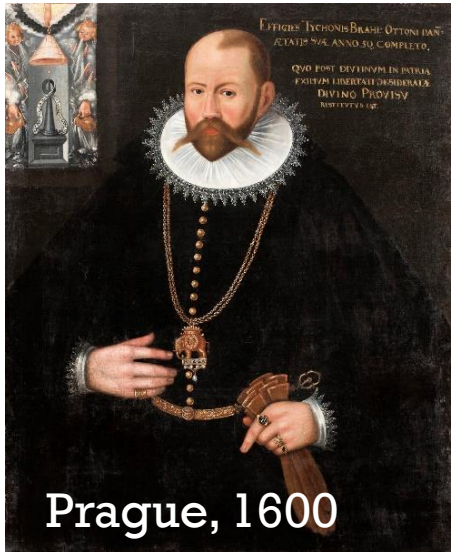
JOAQUIN LOPEZ HERRAIZ

Grupo de Física Nuclear

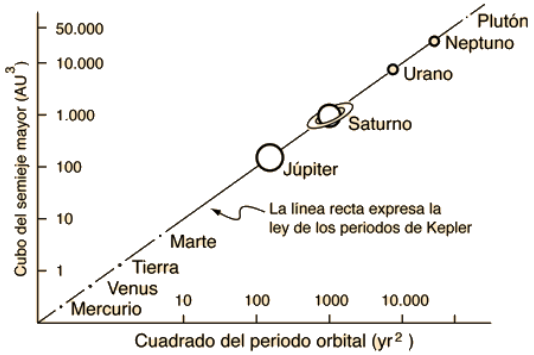
Classic Scientific Method

Scientific Method

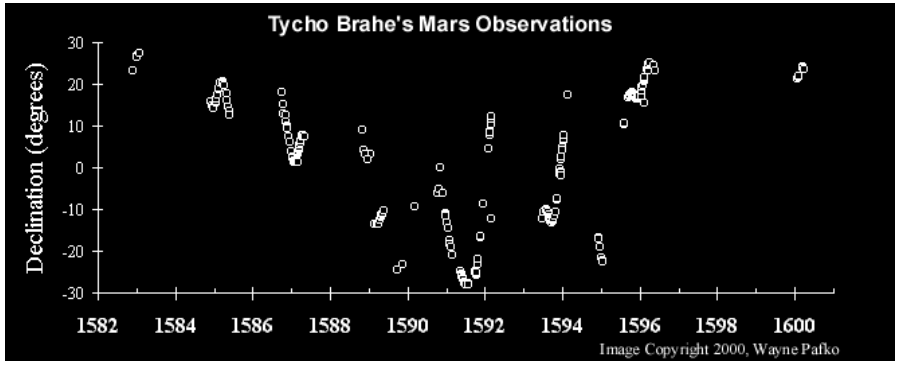
1) Measurements (Data)
2) Propose a Model
   (with some parameters)
3) Fit Parameters with the Data
4) Validation with new Data

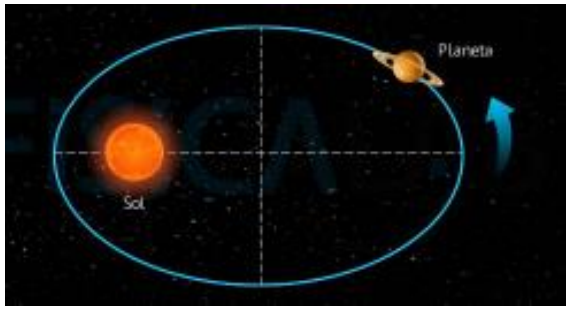


$$T^2 = ka^3$$



Tycho Brahe and Johannes Kepler





LIMITATIONS:

1) MODELS HAVE TO BE SIMPLE

2) MANY PROBLEMS NOT SOLVABLE

Artificial Intelligence vs Clasical Scientific Method

Scientific Method

1) Measurements
2) Propose a Model
   (with some parameters)
3) Fit Model to the Data
4) Model Validation with New Data

Fitted Model = Knowledge

Artificial Intelligence

1) Measurements
2) ----------
3) Fit the data with ANY model
4) Validation with new data

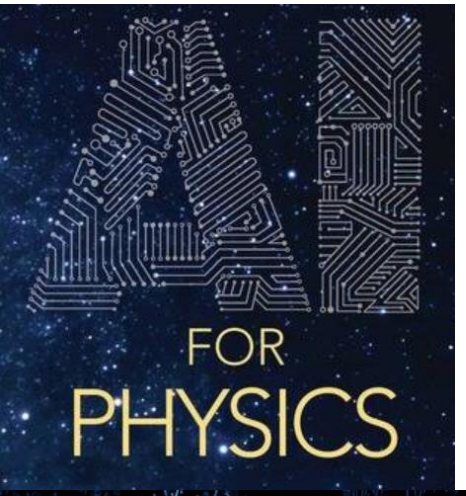
LIMITATIONS:

1) MODELS HAVE TO BE SIMPLE
2) MANY PROBLEMS NOT SOLVABLE

LIMITATIONS:

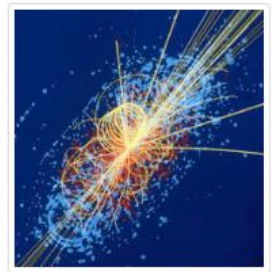
1) HOW TO INTERPRET THE RESULTS?
2) RISK OF BIAS, OVERFITTING…

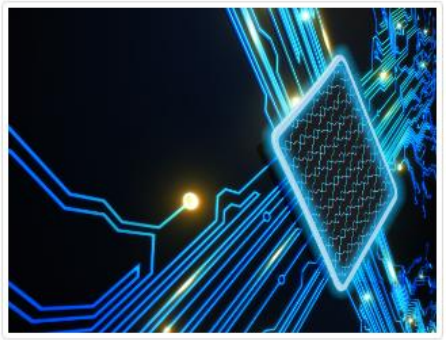
Artificial Intelligence in Physics



AI for Physics.
By Volker Knecht

We're using machinelearning tools to analyze particle physics data from the Large Hadron Collider.



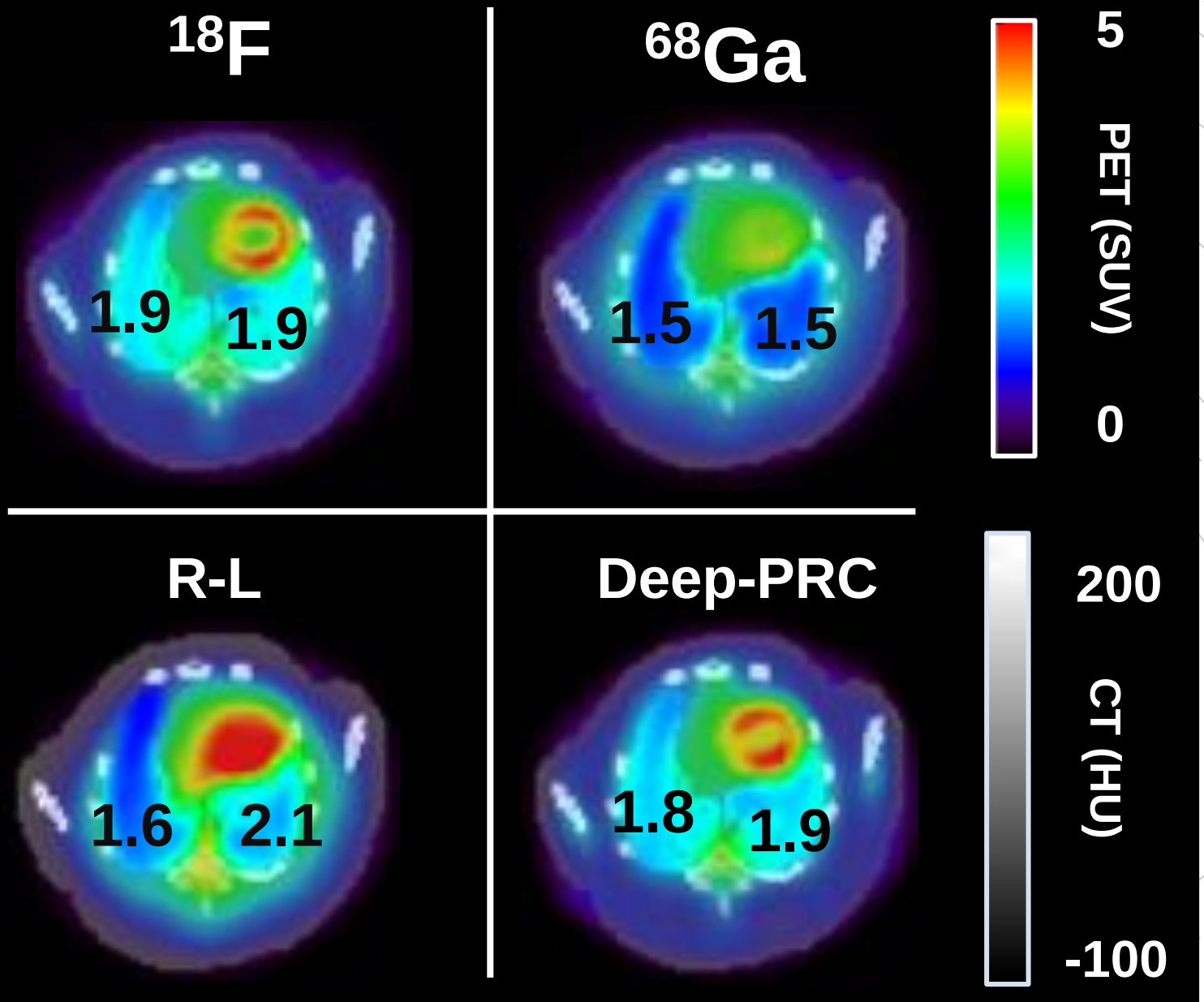


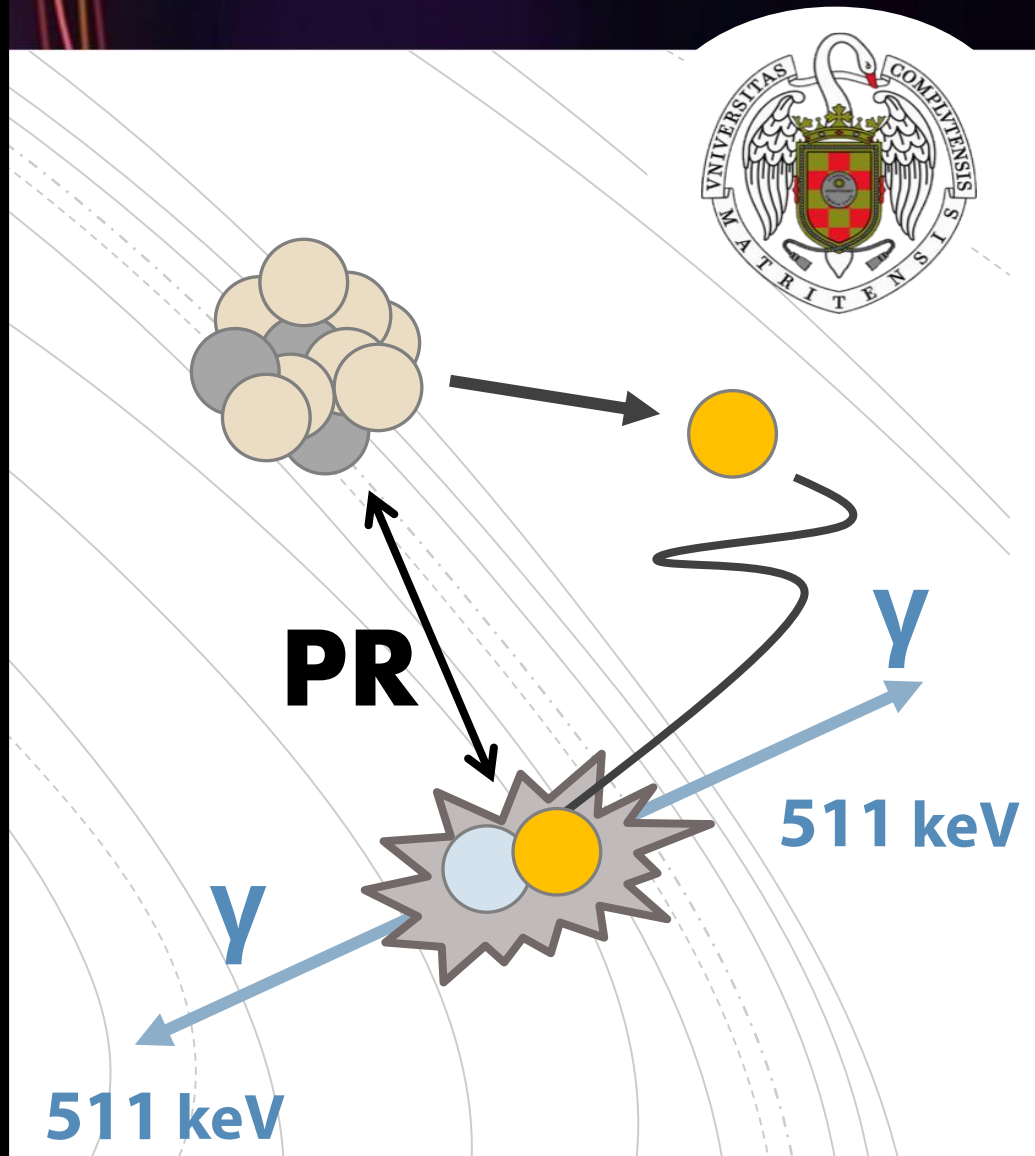
We're developing technology for faster and more energy-efficient deep learning using optical chips, that compute using photons instead of electrons.

We're using techniques from condensed matter physics to help understand how our brains process information,

http://super-ms.mit.edu/physics-ai.html

AI IN MEDICAL PHYSICS





J.L. Herraiz, et al. (2020). **Deep-Learning Based Positron Range Correction of PET Images**. Applied Sciences, 11(1), 266.

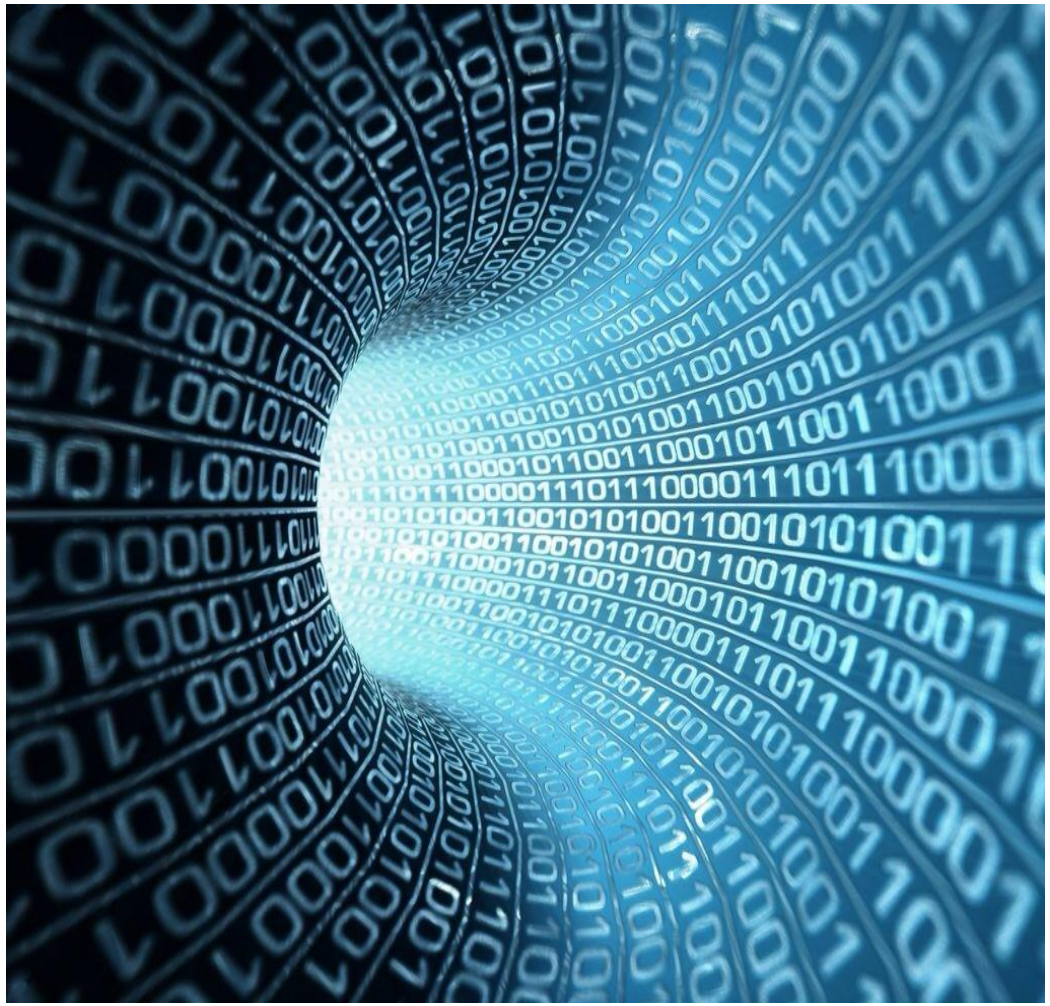
SWOT Analysis of AI in Physics

PRESENT	**Strengths** It sounds cool! It is becoming easy to use It has achieved many successes
	Weaknesses It is too new Not so well understood (heuristics) It changes too fast It can be applied without expertise
FUTURE	**Opportunities** It may solve many opened problems It may open new fields
	Threats It may create biased models Ethical issues Black-box (Do we get knowledge by training an AI tool?)

Challenges of AI in Physics

1) Understand and Explain the Results
2) Data and Labeling
3) Robustness and Reliability
4) Harmonization
5) Multidisciplinarity
6) Correlation≠Causation

Artificial Intelligence Requires Massive Amounts of Data



Getting all that Data Requires Merging Multiple Datasets

- In order to get the large number of cases required to train the AI tools, data from multiple sites obtained with a variety of devices and protocols are needed.



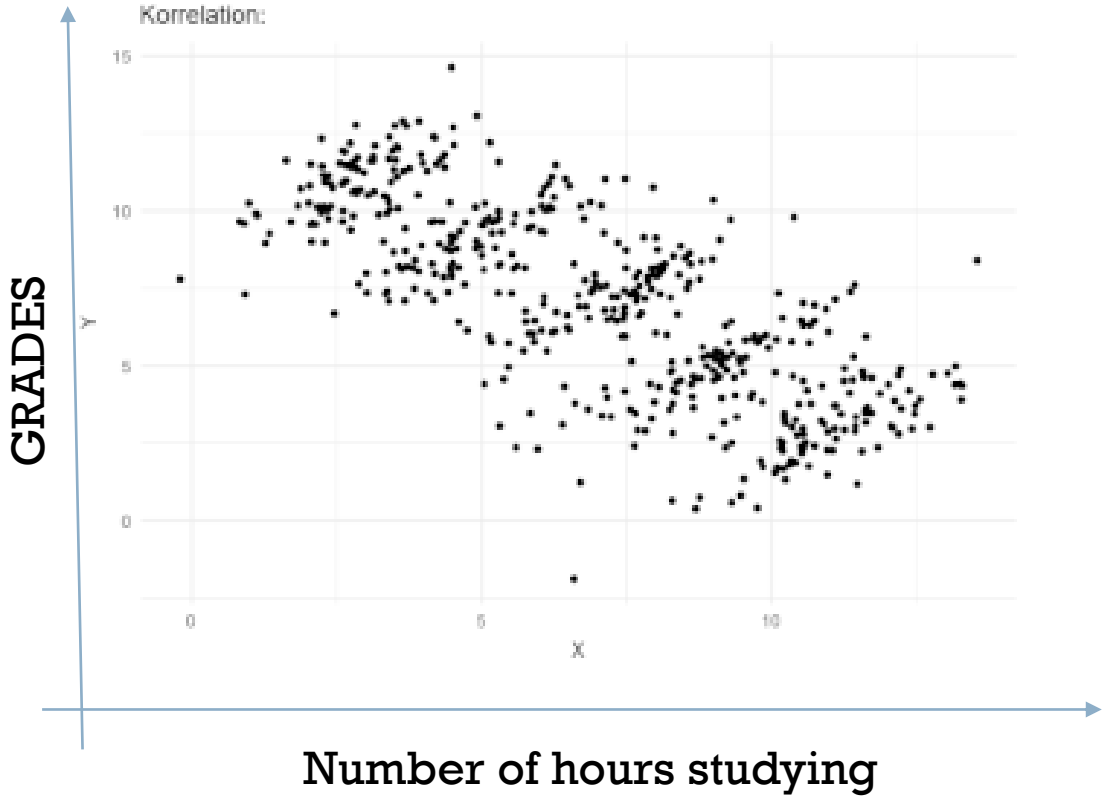
CHEST X-RAYS HAVE A LARGE VARIETY OF IMAGE QUALITY (VOLTAGE USED, INTENSITY, DISTANCE, DETECTOR TYPE)

The Risk of Merging Multiple Datasets: Simpson's Paradox and Covariates

https://en.wikipedia.org/wiki/Simpson%27s_paradox

Lurking or confounding variables

Treatment / Stone size	Treatment A	Treatment B
Small stones	*Group 1* **93% (81/87)**	*Group 2* 87% (234/270)
Large stones	*Group 3* **73% (192/263)**	*Group 4* 69% (55/80)
Both	78% (273/350)	**83% (289/350)**



*C. R. Charig; D. R. Webb; S. R. Payne; J. E. Wickham (29 March 1986). "Comparison of treatment of renal calculi by open surgery, percutaneous nephrolithotomy, and extracorporeal shockwave lithotripsy". Br Med J (Clin Res Ed). **292** (6524): 879–882.*

The Risk of Merging Multiple Datasets: Example

nature machine intelligence

Explore content ∨ About the journal ∨ Publish with us ∨

nature › nature machine intelligence › analyses › article

Analysis | Open Access | Published: 15 March 2021

Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans

Michael Roberts ✉, Derek Driggs, Matthew Thorpe, Julian Gilbey, Michael Yeung, Stephan Ursprung, Angelica I. Aviles-Rivero, Christian Etmann, Cathal McCague, Lucian Beer, Jonathan R. Weir-McCall, Zhongzhao Teng, Effrossyni Gkrania-Klotsas, AIX-COVNET, James H. F. Rudd, Evis Sala & Carola-Bibiane Schönlieb

Nature Machine Intelligence **3**, 199–217 (2021) | Cite this article

66k Accesses | **138** Citations | **1121** Altmetric | Metrics

[…] Our search identified 2,212 studies, of which 415 were included after initial screening and, after quality screening, 62 studies were included in this systematic review. ***Our review finds that none of the models identified are of potential clinical use due to methodological flaws and/or underlying biases.***

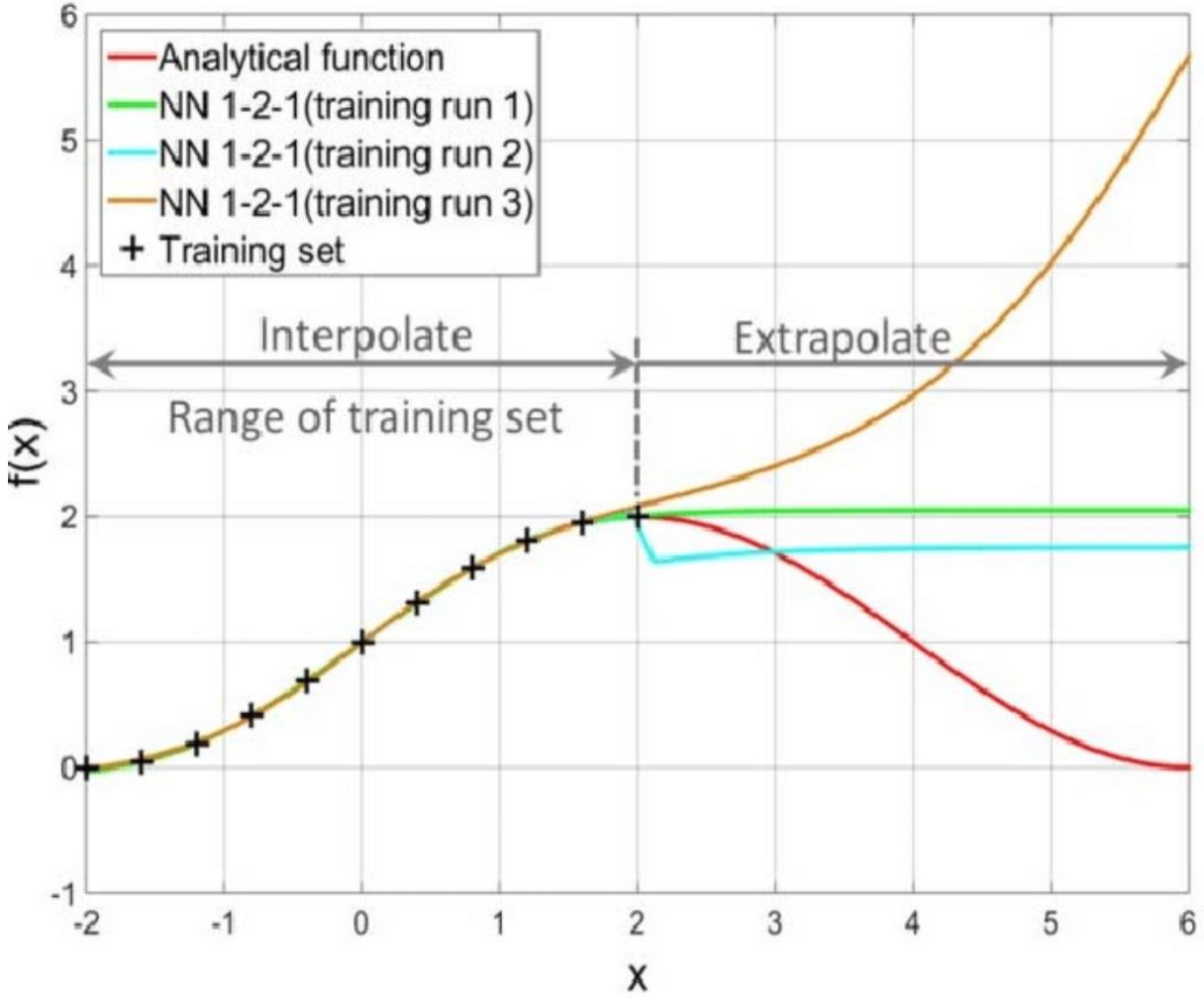
The Risk of Using Pretrained Models

Usually, the range of applicability of models in physics in known.

With AI models is usually not the case..

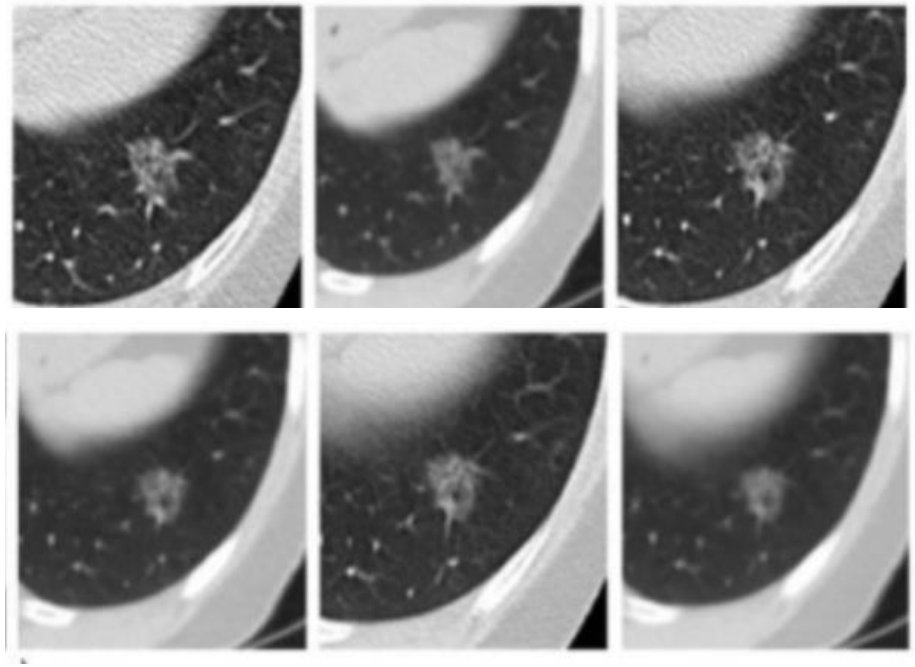
Are you interpolating

or extrapolating?



SOLUTION: Data Harmonization and Domain Transfer

- **OPTION 1 – Data Harmonization**:
  Data preprocessing to make multiple datasets compatible
- **OPTION 2 – Domain Transfer:** Adapt the pretrained model to the specific data



B. Zhao, Scientific Reports 6,23428 (2016)

ELSEVIER

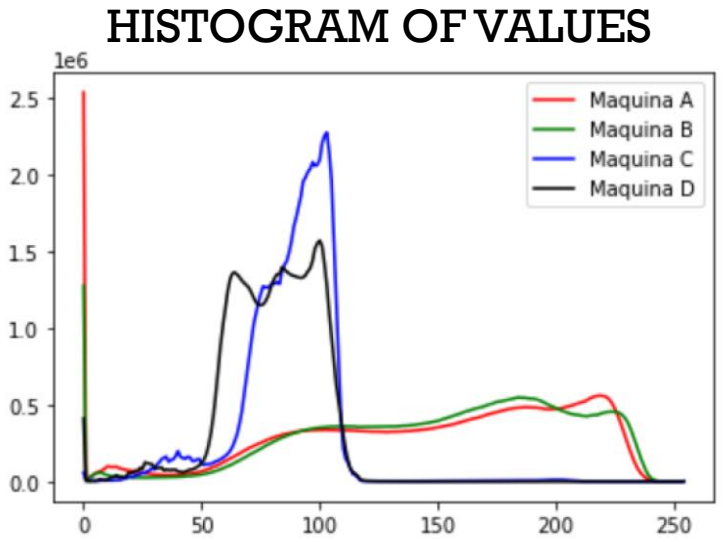
Information Fusion
Volume 82, June 2022, Pages 99-122

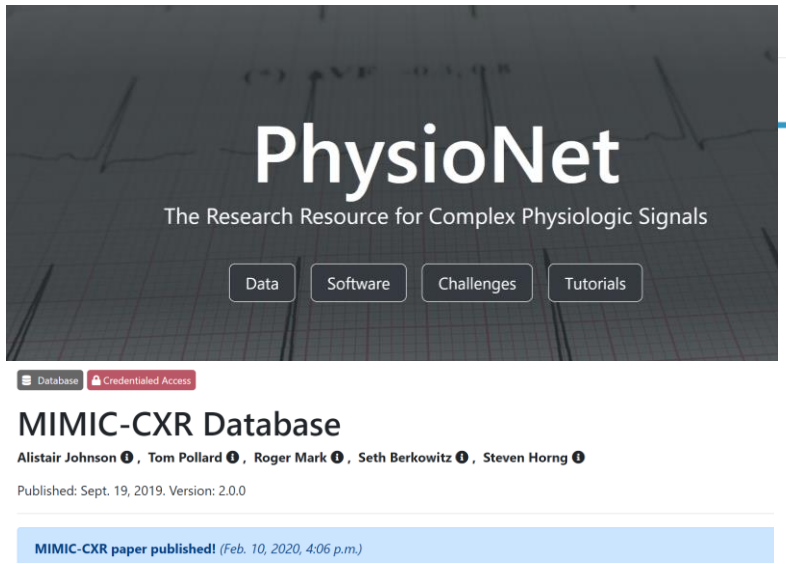
Data harmonisation for information fusion in digital healthcare: A state-of-the-art systematic review, meta-analysis and future research directions

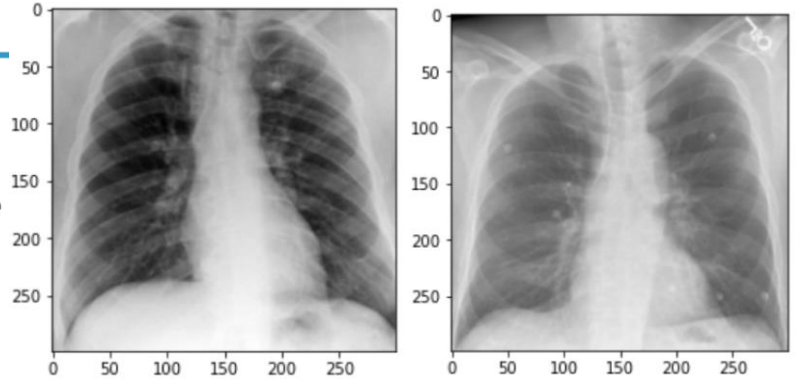
Yang Nan [a], Javier Del Ser [d, e], Simon Walsh [a], Carola Schönlieb [f], Michael Roberts [f, g], Ian Selby [h], Kit Howard [i], John Owen [i], Jon Neville [i], Julien Guiot [j, k], Benoit Ernst [j, k], Ana Pastor [l], Angel Alberich-Bayarri [l], Marion I. Menzel [m, n], Sean Walsh [o], Wim Vos [o], Nina Flerin [o], Jean-Paul Charbonnier [p] ... Guang Yang [a, b, c, #]

Harmonization and Domain Transfer in Action

	Size	Healthy	Pneumonia
SCANNER A	**3056 x 2544**	27695	2028
SCANNER B	**2021 x 2021**	1055	415
SCANNER C	**2022 x 1736**	2588	445







RESULTS OBTAINED WITH ONLY ONE SCANNER

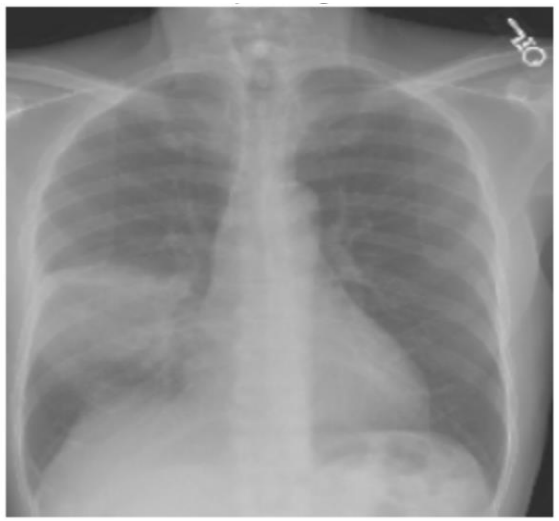
Model A Accuracy= 80 %		**ESTIMATION**	
		HEALTHY	**PNEUMONIA**
ACTUAL	**HEALTHY**	**360**	**76**
	PNEUMONIA	**89**	**287**

The trained model applied directly to cases of another scanner (Scanner C) does not work (**50% accuracy**).

RESULTS OBTAINED WITH ONE SCANNER APPLIED TO ANOTHER ONE (WITH HARMONIZATION)

CXR Scanner A

CXR Scanner C



CXR Scanner A made "similar" to Scanner C with a CycleGAN

With harmonization tools a model trained with one scanner can be used to other scanners with ACCURACY of 73%

(lower than 80%, but much better than 50%)

CONCLUSIONS:

The fact that you can train an AI model without thinking, it does not mean that you should not think.

It is very important to take into account the data that were used to train AI models. Are you interpolating or extrapolating?

Be aware of possible covariates in your data.

Data harmonization techniques can help you reducing the bias of your results.