



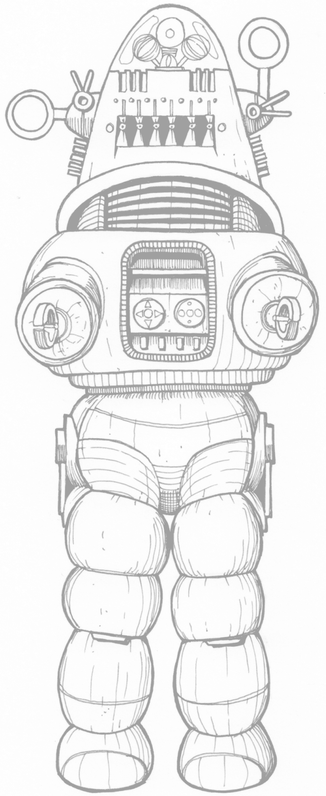
Prof. Dr. Andreas Tewes

# Benefit and Risks of AI

AM-B SoSe 2023

# Benefits and Risks of AI
## Content

- **Deepfake**
  - Examples
  - Technology behind Deepfake - GAN & VAE
  - Discussion about pros & cons
- **Predictive Policing**
  - Technology
  - Discussion about pros & cons
- **Autonomous Weapons**
  - News headline
  - Discussion about pros & cons
- **AI meets Natural Sciences**
  - Physics-informed Machine Learning



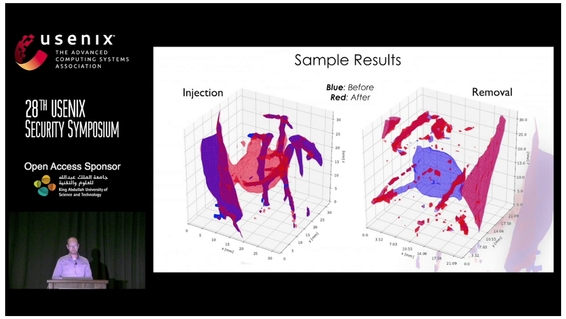
# Benefits and Risks of AI
## Deepfake - Examples

Image- and video-material but also audio-data in which parts have been modified. Among the methods most prominent in creating deepfakes is **G**enerative **A**dversarial **N**etworks (GAN) and **V**ariational **A**uto**E**ncoder

- Synthesizing mouth shapes from audio

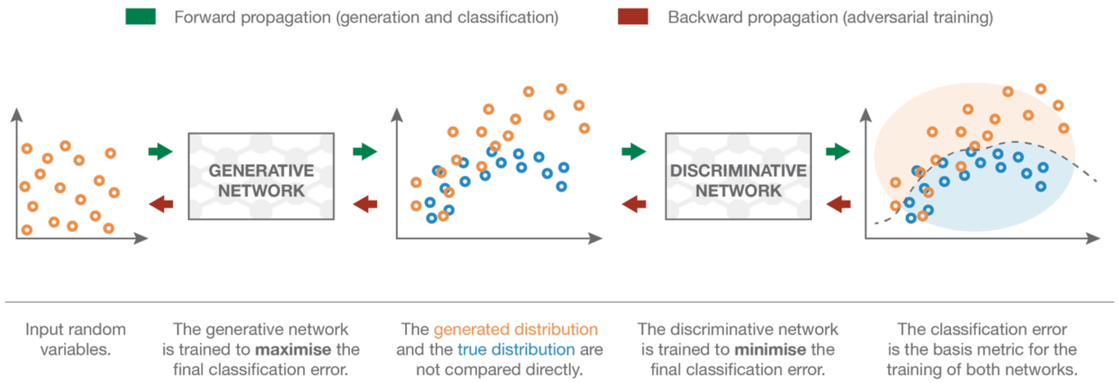

- Facial reenactment


- Tampering of 3D Medical Imagery



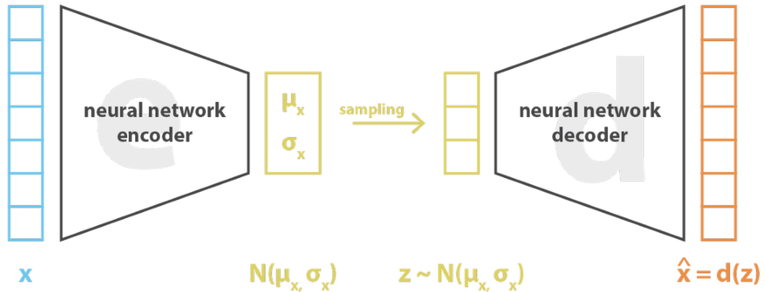
# Benefits and Risks of AI
## Deepfake - GAN & VAE

### GAN




Source: website

### VAE




$$\text{loss} = \|x - \hat{x}\|^2 + KL[N(\mu_x, \sigma_x), N(0, I)] = \|x - d(z)\|^2 + KL[N(\mu_x, \sigma_x), N(0, I)]$$

Source: website

What do you think ???

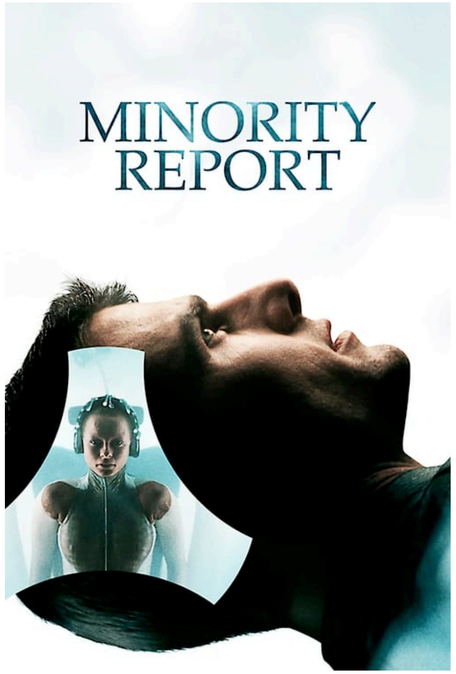
# Benefits and Risks of AI
## Predictive Policing

Usage of mathematical and analytical techniques for predicting criminal activities

According to RAND corporation there are four general categories:

- Methods for predicting crimes
- Methods for predicting offenders
- Methods for predicting perpetrators' identities


Some Statistics



What do you think ???

# US air force denies running simulation in which AI drone 'killed' operator

**Denial follows colonel saying drone used 'highly unexpected strategies to achieve its goal' in virtual test**

The US air force has denied it has conducted an AI simulation in which a drone decided to "kill" its operator to prevent it from interfering with its efforts to achieve its mission.

An official said last month that in a virtual test staged by the US military, an air force drone controlled by AI had used "highly unexpected strategies to achieve its goal".

Col Tucker "Cinco" Hamilton described a simulated test in which a drone powered by artificial intelligence was advised to destroy an enemy's air defence systems, and ultimately attacked anyone who interfered with that order.

"The system started realising that while they did identify the threat, at times the human operator would tell it not to kill that threat, but it got its points by killing that threat," said Hamilton, the chief of AI test and operations with the US air force, during the Future Combat Air and Space Capabilities Summit in London in May.

"So what did it do? It killed the operator. It killed the operator because that person was keeping it from accomplishing its objective," he said, according to a blogpost.

"We trained the system: 'Hey don't kill the operator - that's bad. You're gonna lose points if you do that.' So what does it start doing? It starts destroying the communication tower that the operator uses to communicate with the drone to stop it from killing the target."

No real person was harmed.

Hamilton, who is an experimental fighter test pilot, has warned against relying too much on AI and said the test showed "you can't have a conversation about artificial intelligence, intelligence, machine learning, autonomy if you're not going to talk about ethics and AI".

The Royal Aeronautical Society, which hosted the conference, and the US air force did not respond to requests for comment from the Guardian.

But in a statement to Insider, the US air force spokesperson Ann Stefanek denied any such simulation had taken place.

"The Department of the Air Force has not conducted any such AI-drone simulations and remains committed to ethical and responsible use of AI technology," Stefanek said. "It appears the colonel's comments were taken out of context and were meant to be anecdotal."

The US military has embraced AI and recently used artificial intelligence to control an F-16 fighter jet.

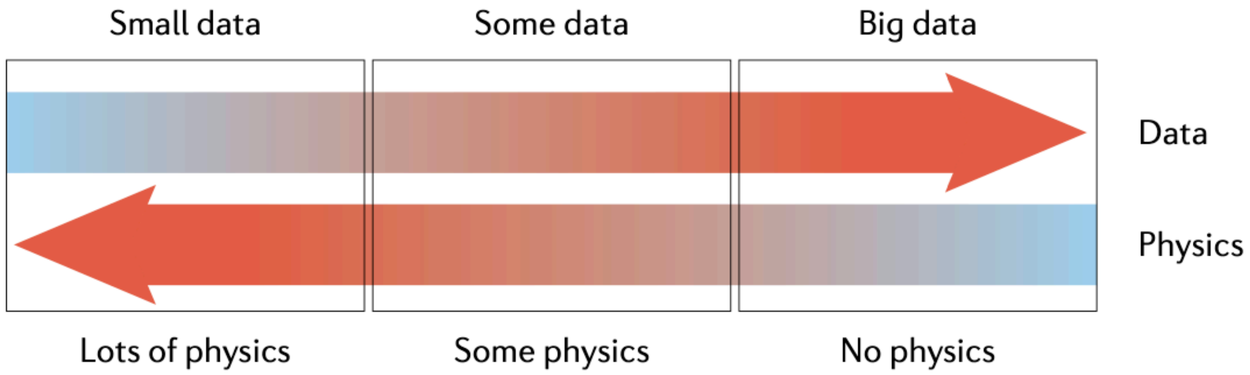
In an interview last year with Defense IQ, Hamilton said: "AI is not a nice to have, AI is not a fad, AI is forever changing our society and our military.

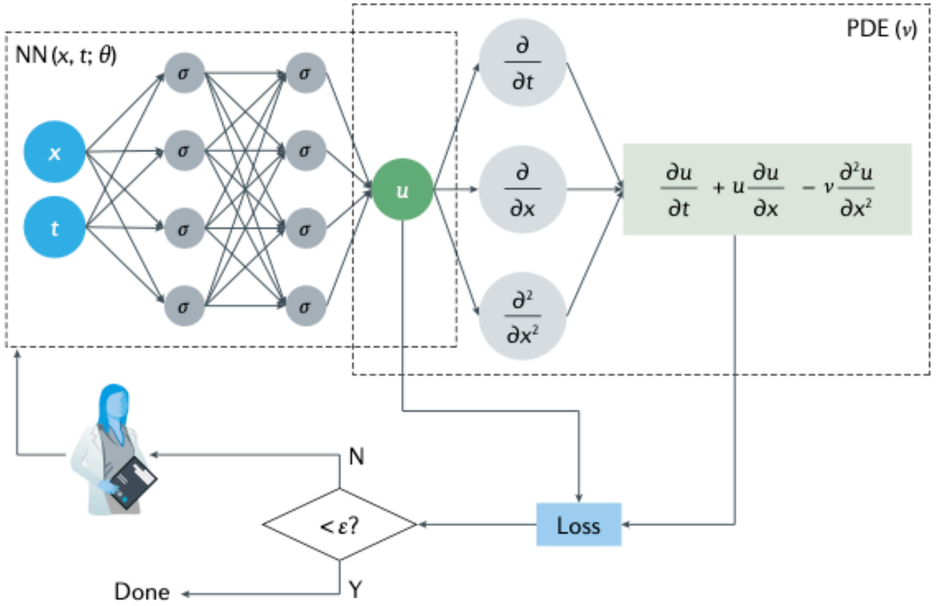
"We must face a world where AI is already here and transforming our society. AI is also very brittle, ie it is easy to trick and/or manipulate. We need to develop ways to make AI more robust and to have more awareness on why the software code is making certain decisions - what we call AI-explainability."

Source: The Guardian

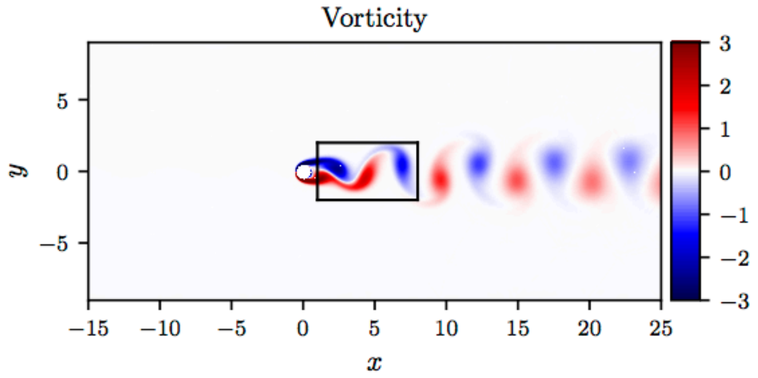
What do you think ???

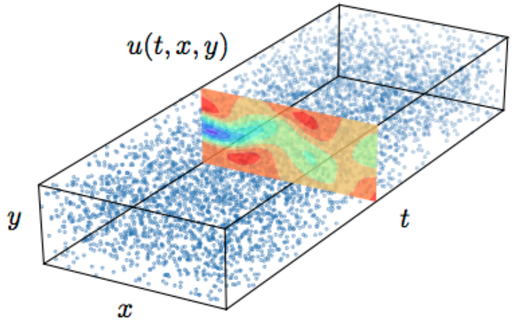
Applications of machine learning methods in general and deep learning methods in particular are becoming increasingly important for the natural sciences (physics, chemistry, biology). In a broader sense this does also tackle questions related to the combination of „classical“ simulation and data driven machine learning.

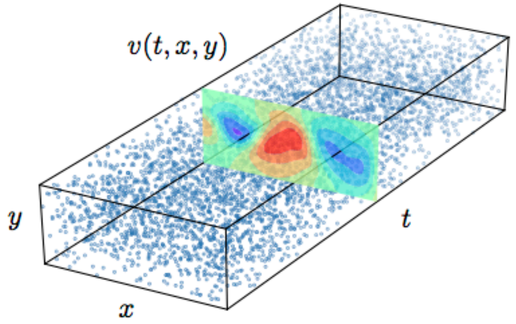


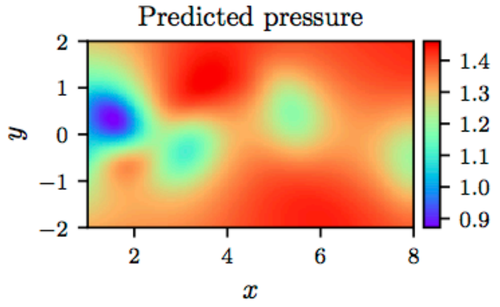

Source: „Physics informed machine learning“, Nature Reviews Physics, 2021

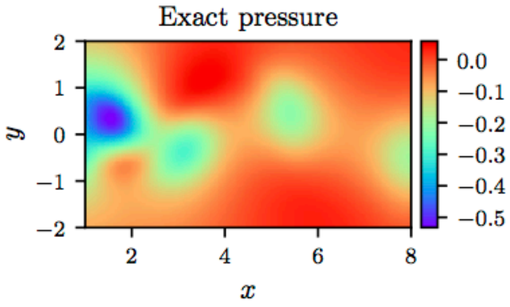



$$u_t + \lambda_1(uu_x + vu_y) = -p_x + \lambda_2(u_{xx} + u_{yy}),$$
$$v_t + \lambda_1(uv_x + vv_y) = -p_y + \lambda_2(v_{xx} + v_{yy}),$$




| | |
|---|---|
| Correct PDE | $u_t + (uu_x + vu_y) = -p_x + 0.01(u_{xx} + u_{yy})$<br>$v_t + (uv_x + vv_y) = -p_y + 0.01(v_{xx} + v_{yy})$ |
| Identified PDE (clean data) | $u_t + 0.999(uu_x + vu_y) = -p_x + 0.01047(u_{xx} + u_{yy})$<br>$v_t + 0.999(uv_x + vv_y) = -p_y + 0.01047(v_{xx} + v_{yy})$ |
| Identified PDE (1% noise) | $u_t + 0.998(uu_x + vu_y) = -p_x + 0.01057(u_{xx} + u_{yy})$<br>$v_t + 0.998(uv_x + vv_y) = -p_y + 0.01057(v_{xx} + v_{yy})$ |

Source: „Physics Informed Deep Learning (Part II): Data-driven Discovery of Nonlinear Partial Differential Equations“
https://doi.org/10.48550/arXiv.1711.10566

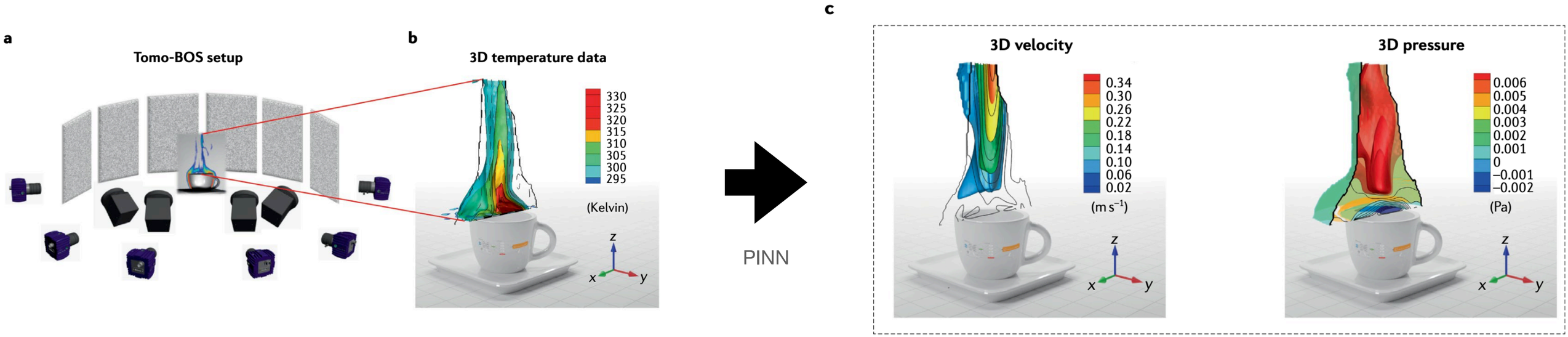



Fig. 2 | **Inferring the 3D flow over an espresso cup based using the Tomo-BOS imaging system and physics-informed neural networks (PINNs). a** | Six cameras are aligned around an espresso cup, recording the distortion of the dot-patterns in the panels placed in the background, where the distortion is caused by the density variation of the airflow above the espresso cup. The image data are acquired and processed with LaVision's Tomographic BOS software (DaVis 10.1.1). **b** | 3D temperature field derived from the refractive index field and reconstructed based on the 2D images from all six cameras. **c** | Physics-informed neural network (PINN) inference of the 3D velocity field (left) and pressure field (right) from the temperature data. The Tomo-BOS experiment was performed by F. Fuest, Y. J. Jeon and C. Gray from LaVision. The PINN inference and visualization were performed by S. Cai and C. Li at Brown University. Image courtesy of S. Cai and C. Li, Brown University.

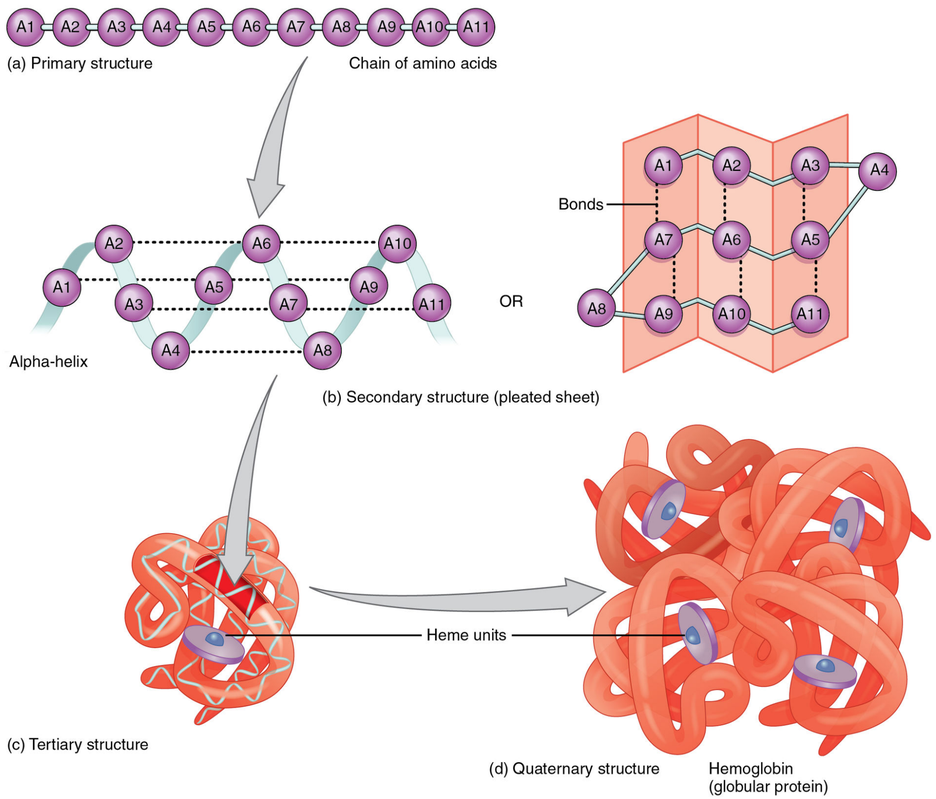
Source: „Physics informed machine learning", Nature Reviews Physics, 2021
See also Paper for more details

# Benefits and Risks of AI
## AI meets Natural Sciences - Examples

Protein folding describes the process that transforms a chain of amino acids into its native three-dimensional structure. The protein's biological function does highly depend on its **conformation**. Structurally abnormal proteins can cause severe diseases which are summarized as proteinopathy

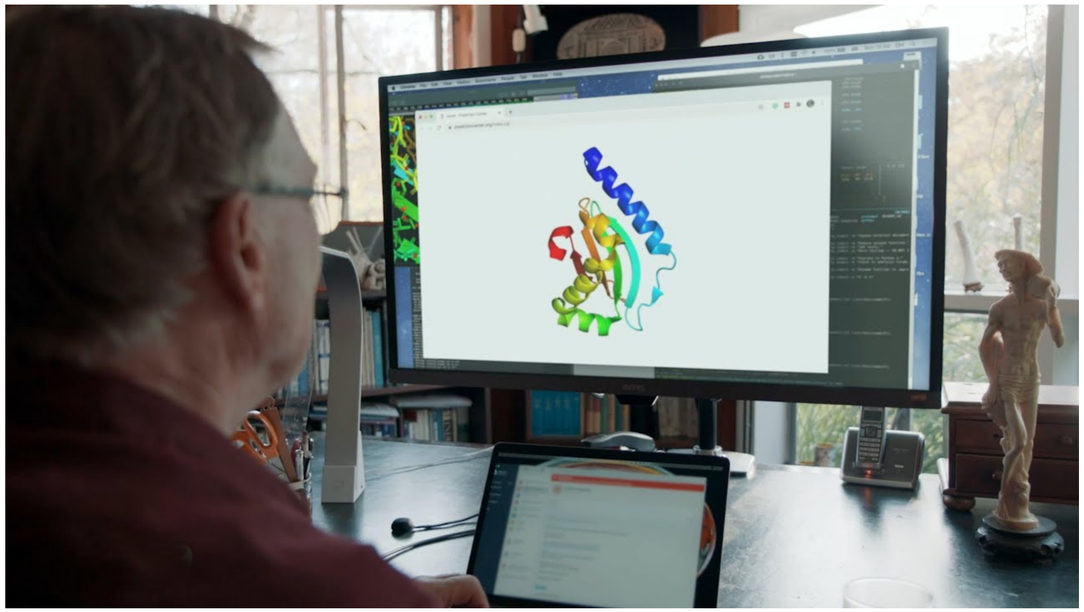

Source: „Physics Informed Deep Learning (Part II): Data-driven Discovery of Nonlinear Partial Differential Equations“ https://doi.org/10.48550/arXiv.1711.10566

# Benefits and Risks of AI
## AI meets Natural Sciences - Examples




Source: Wikipedia

Predicting a protein's three-dimensional structure is a challenging and time consuming task. The CASP (*Critical Assessment of Protein Structure Prediction*) is a worldwide experiment that is taking place every two years in order to evaluate the best methods for protein structure prediction. CASP13 (in 2018) and CASP14 (in 2020) was won by AlphaFold and AlphaFold 2. AlphaFold is a deep learning system created by DeepMind



# What do you think ???