
**Kristian Kersting**

kerstingAIML


Alejandro Molina


Patrick Schramowski


Zhongjie Yu


Wolfgang Stammer

and many more

TECHNISCHE UNIVERSITÄT DARMSTADT

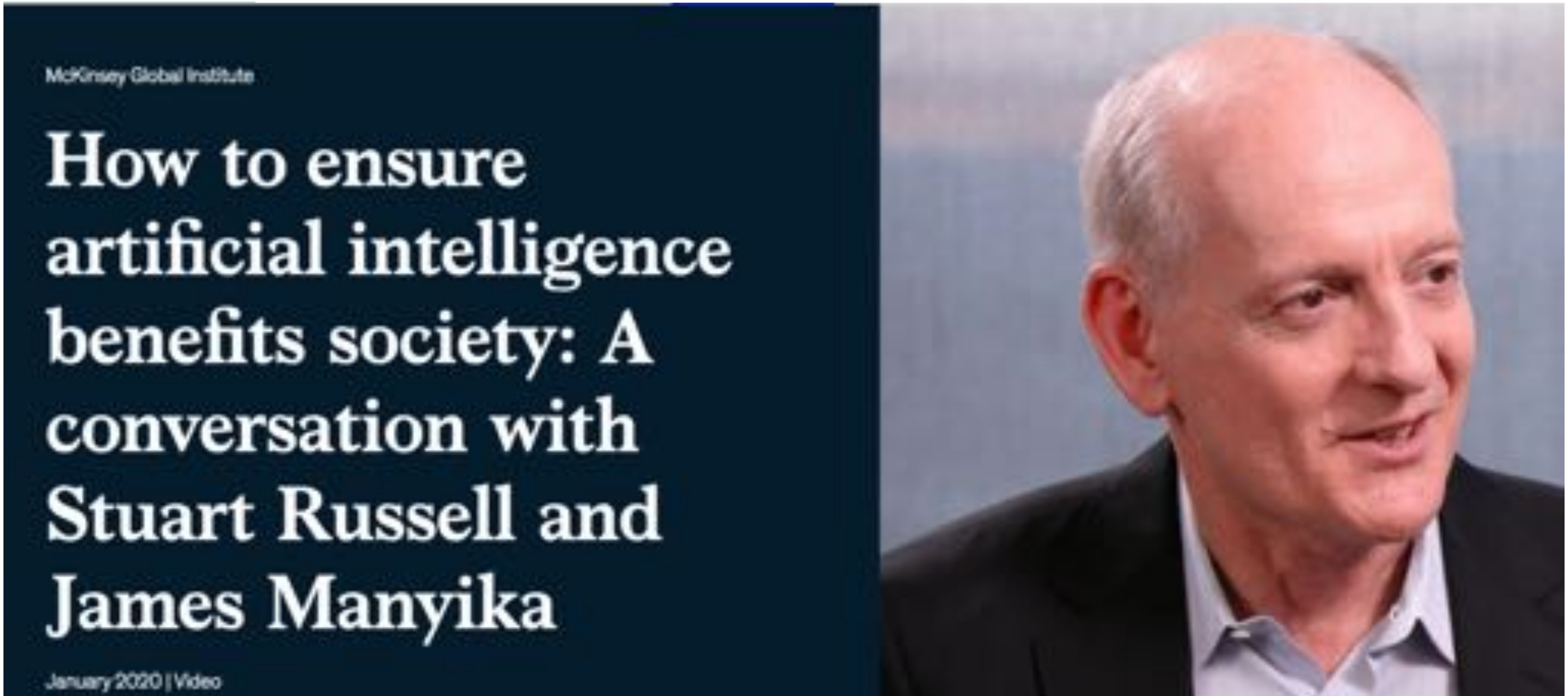
Centre for Cognitive Science

Fachbereich Informatik

# **The Third Wave of AI:** Closing the Gap between AI and the Domain Experts, using the example of Deep Plant Phenotyping

McKinsey
& Company

# AI and Climate Science




I think **climate science** is another problem around the totality of the picture where AI can help. You've got atmospheric specialists, you have ocean people, you have cloud people, you have economists who look at migration and mitigation and so on, you have got the biosphere people who look at bacteria and processes of putrefaction of peat bogs and Siberian permafrost, and all the rest of it**. But does anyone have the whole picture in their mind? AI systems could have the whole picture.**

https://www.mckinsey.com/featured-insights/artificial-intelligence/how-to-ensure-artificial-intelligence-benefits-society-a-conversation-with-stuart-russell-and-james-manyika

Nature 466, 531–532 (29 July 2010)


**The world's population expected to grow from 6.8 billion today to 9.1 billion by 2050**

**How to expand agricultural output massively without increasing by much the amount of land used?**

# Early Detection of Plant Stress using Machine Learning and Hyperspectral Imaging

Tardieu, Schurr, “White Paper on Plant Phenotyoping,” EPSO Workshop on Plant Phenotyping, Juelich, Nov. 2009
Houle, Govindaraju, Omholt “Phenomics: the next challenge“ Nature Reviews Genetics 11:855-866, Dec. 2010

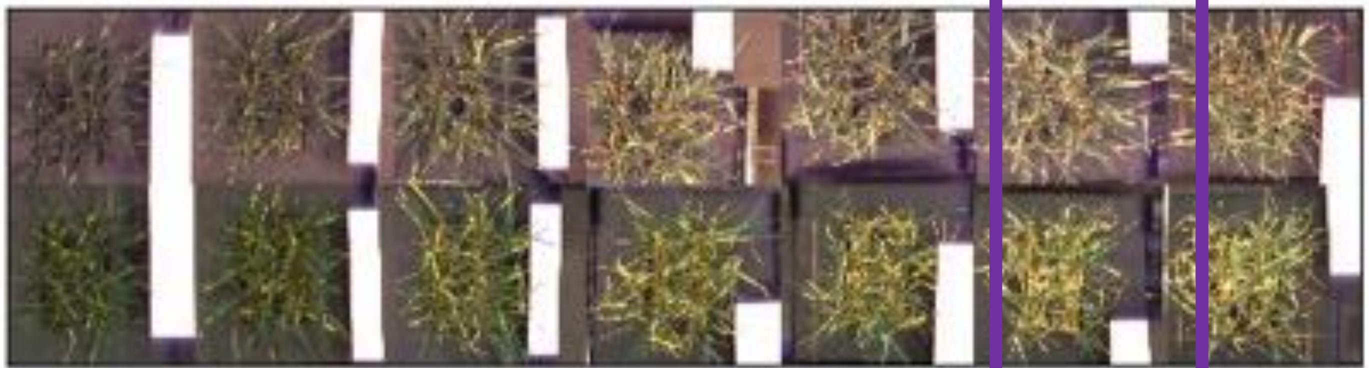
# Phenotyping: Who is stressed and why?

Large-scale phenotyping is the natural complement to genome sequencing as a route to rapid advances in biology. Ultimately, it links genomics with the performance of plants in the interaction with environmental cues.

Can we detect it earlier using sensor technology?

Barley A

Barley B

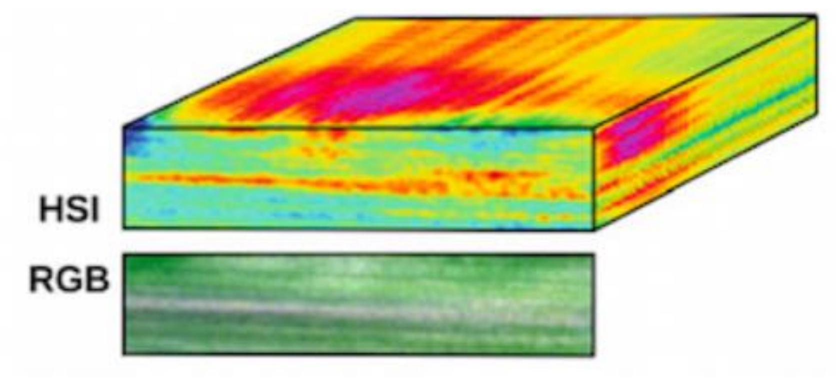


Measurement Day (MD) 1 MD 2 MD 3 MD 4 MD 5 MD 6 MD 7

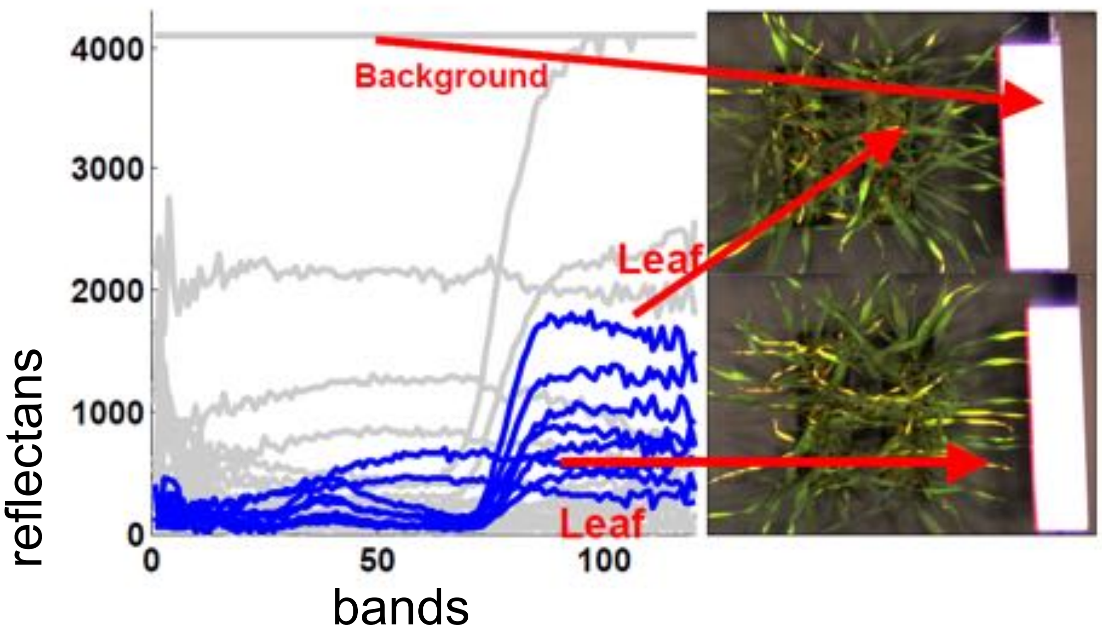
naked eye

# Hyperspectral Imaging

Phenotyping by means of sensor technologies involves the identification of relevant patterns in massive data sets of high-dimensional sensor readings with a demanding signal-to-noise ratio.




data cube



Anne-Katrin Mahlein, Matheus Thomas Kuska, Stefan Thomas, Mirwaes Wahabzada, Jan Behmann, Uwe Rascher, Kristian Kersting (2019): **Quantitative and qualitative phenotyping of disease resistance of crops by hyperspectral sensors: seamless interlocking of phytopathology, sensors, and machine learning is needed!.** *Current Opinion in Plant Biology 50:156–162.*

Anna Brugger, Jan Behmann, Stefan Paulus, Hans-Georg Luigs, Matheus Thomas Kuska, Patrick Schramowski, Kristian Kersting, Ulrike Steiner, Anne-Katrin Mahlein (2019): **Extending hyperspectral imaging for plant phenotyping to the UV-range.** *Remote Sensing 11(12):1401.*

Matheus Thomas Kuska, Anna Brugger, Stefan Thomas, Mirwaes Wahabzada, Kristian Kersting, Erich-Christian Oerke, Ulrike Steiner, Anne-Katrin Mahlein (2017): **Spectral patterns reveal early resistance reactions of barley against Blumeria graminis f. sp. hordei.** *Phytopathology 107:1388-1398.*

Mirwaes Wahabzada, Anne-Katrin Mahlein, Christian Bauckhage, Ulrike Steiner, Erich-Christian Oerke, Kristian Kersting (2016): **Plant phenotyping using probabilistic topic models: uncovering the hyperspectral language of plants.** *Scientific Reports (Nature) 6.*

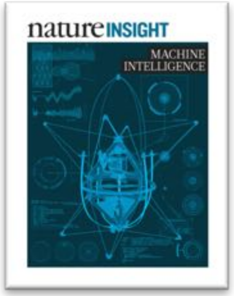
Marlene Leucker, Mirwaes Wahabzada, Kristian Kersting, Madlaina Peter, Werner Beyer, Ulrike Steiner, Anne-Katrin Mahlein, Erich-Christian Oerke (2016): **Hyperspectral imaging reveals the effect of sugar beet QTLs on Cercospora leaf spot resistance.** *Functional Plant Biology 44:1-9.*

Mirwaes Wahabzada, Anne-Katrin Mahlein, Christian Bauckhage, Ulrike Steiner, Erich-Christian Oerke, Kristian Kersting (2015): **Metro maps of plant disease dynamics–-automated mining of differences using hyperspectral images.** *PLoS One 10(1):e0116902.*

Matheus Kuska, Mirwaes Wahabzada, Marlene Leucker, Heinz-Wilhelm Dehne, Kristian Kersting, Erich-Christian Oerke, Ulrike Steiner, Anne-Katrin Mahlein (2015): **Hyperspectral phenotyping on the microscopic scale: towards automated characterization of plant-pathogen interactions.** *Plant Methods 11(1):28.*

Christoph Römer, Mirwaes Wahabzada, Agim Ballvora, Francisco Pinto, Micol Rossini, Cinzia Panigada, Jan Behmann, Jens Leon, Christian Thurau, Christian Bauckhage, Kristian Kersting, Uwe Rascher, Lutz Plümer (2012): **Early drought stress detection in cereals: simplex volume maximisation for hyperspectral image analysis.** *Functional Plant Biology 39(11):878–890.*

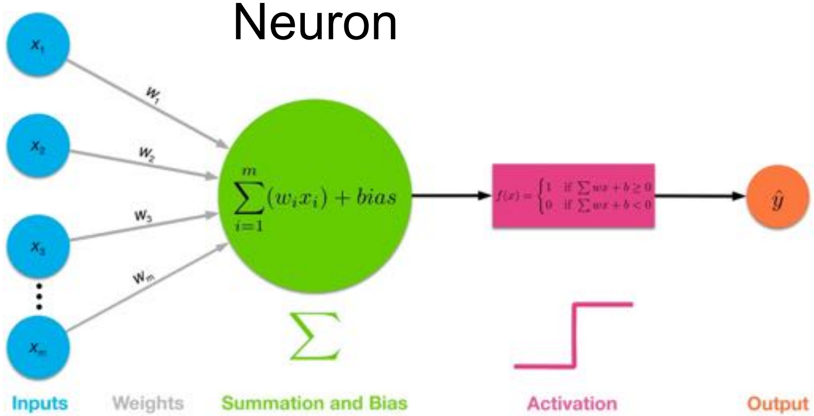
# Deep Neural Networks

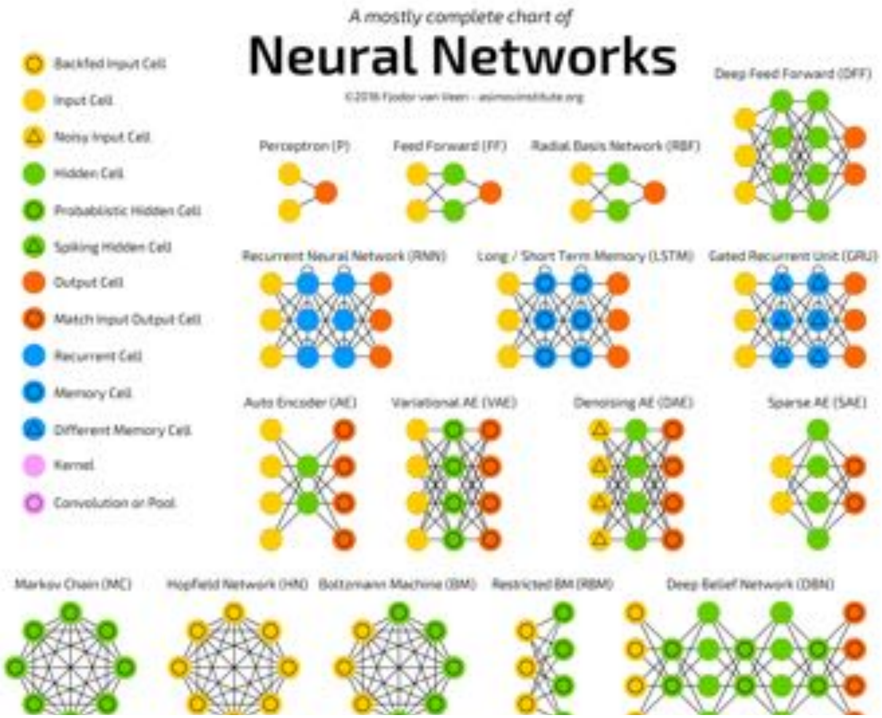



Potentially much more powerful than shallow architectures, represent computations

[LeCun, Bengio, Hinton Nature 521, 436–444, 2015]

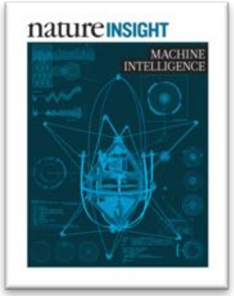
Neuron



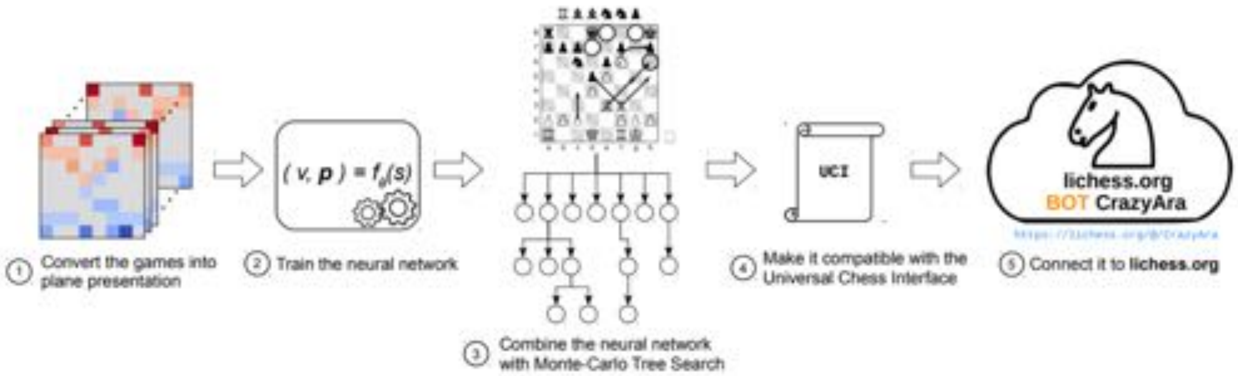

**Differentiable Programming**

# Deep Neural Networks




Potentially much more powerful than shallow architectures, represent computations

[LeCun, Bengio, Hinton Nature 521, 436–444, 2015]

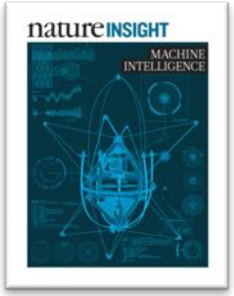




**They can beat the world champion in CrazyHouse**

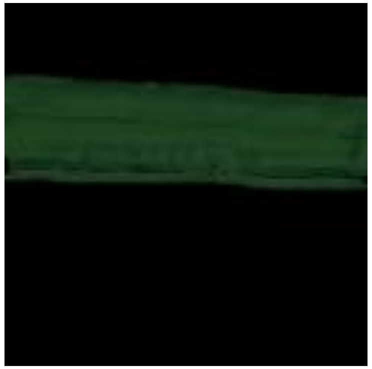
[Czech, Willig, Beyer, Kersting, Fürnkranz *arXiv:1908.06660 2019*]

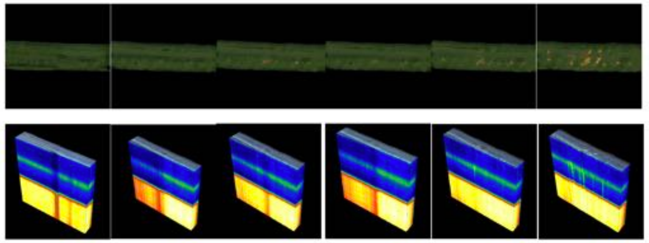
# Deep Neural Networks

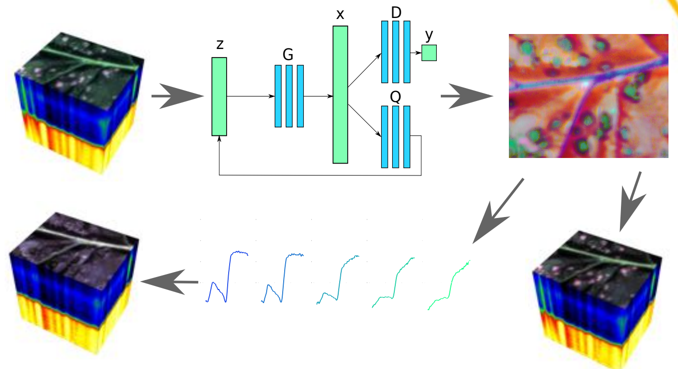


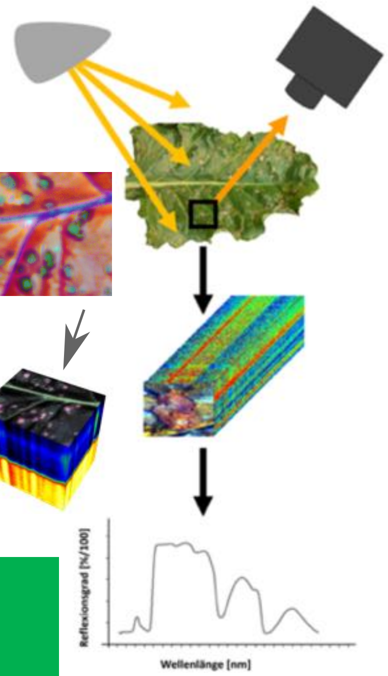

Potentially much more powerful than shallow architectures, represent computations

[LeCun, Bengio, Hinton Nature 521, 436–444, 2015]







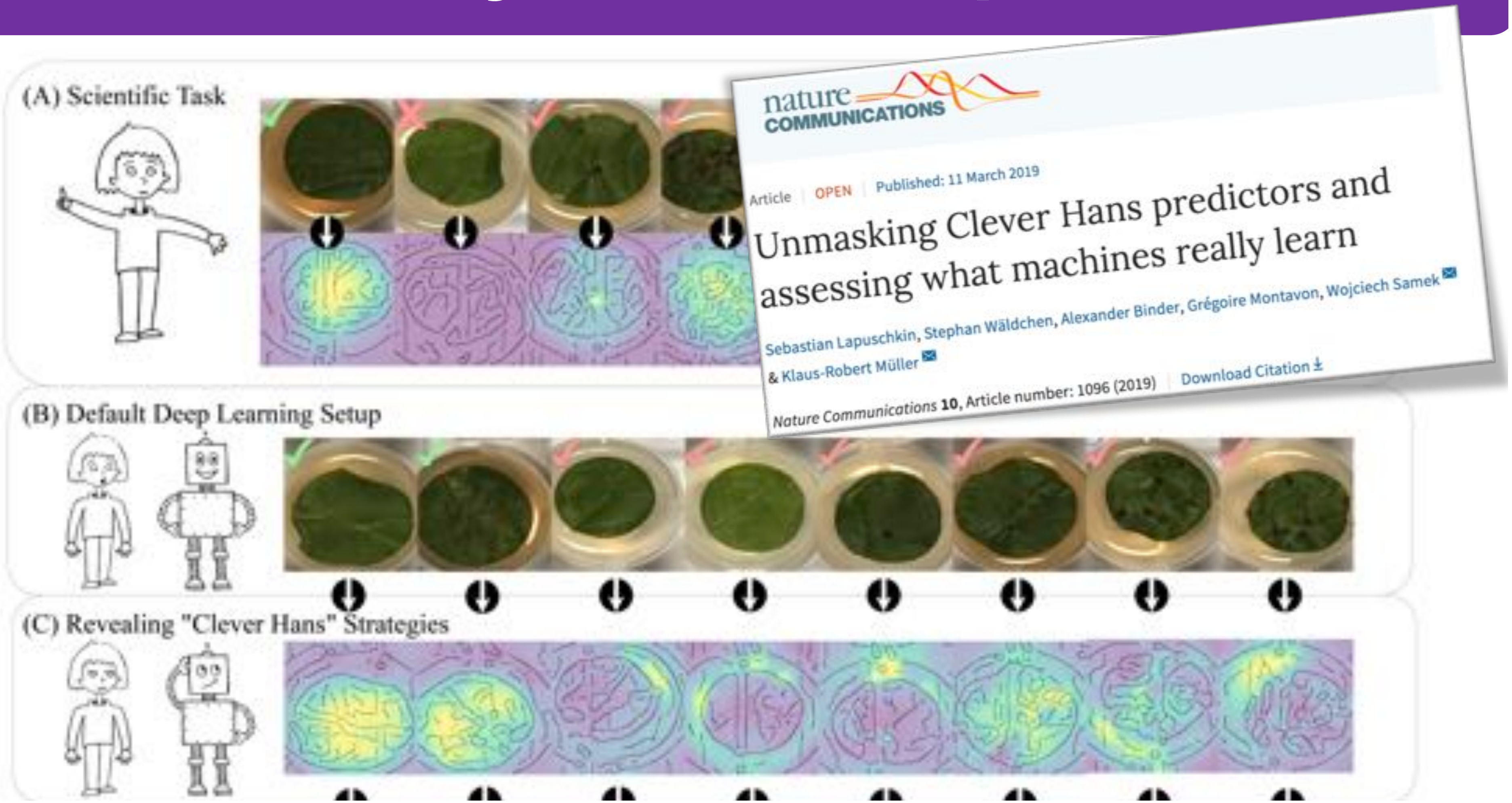

**They “develop intuition” about complicated biological processes and generate scientific data**

DePhenSe

Bundesanstalt für Landwirtschaft und Ernährung

[Schramowski, Brugger, Mahlein, Kersting 2020 to be submitted]

# Can we always trust deep models?

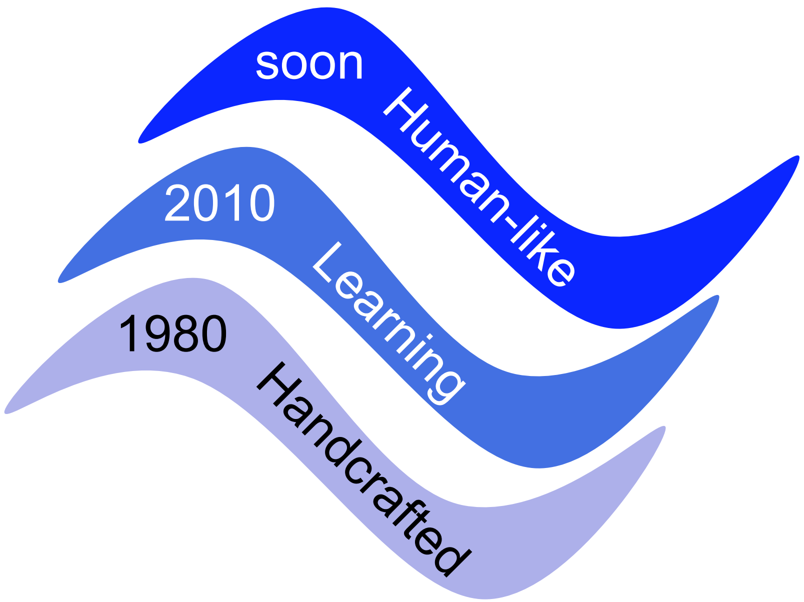



[Schramowski et al. arXiv:2001.05371 2020]

AudisSens
DePhenSe
Bundesanstalt für Landwirtschaft und Ernährung

# Third Wave of AI


## Data and Machine Learning are not everything




AI systems that can acquire human-like communication and reasoning capabilities, with the ability to recognise new situations and adapt to them.

**DNNs often have no probabilistic semantics. They are not calibrated joint distributions.**

$$P(Y|X) \neq P(Y,X)$$

**MNIST**

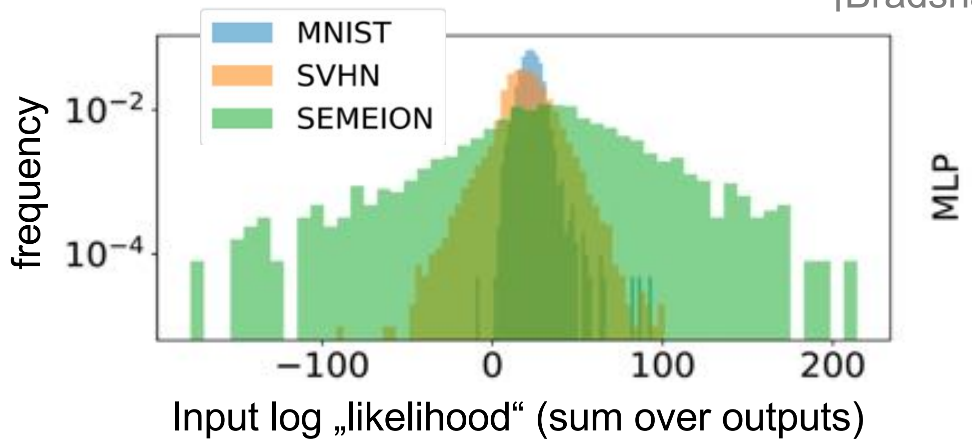
**Train & Evaluate**

**SVHN**

**SEMEION**

**Transfer Testing**

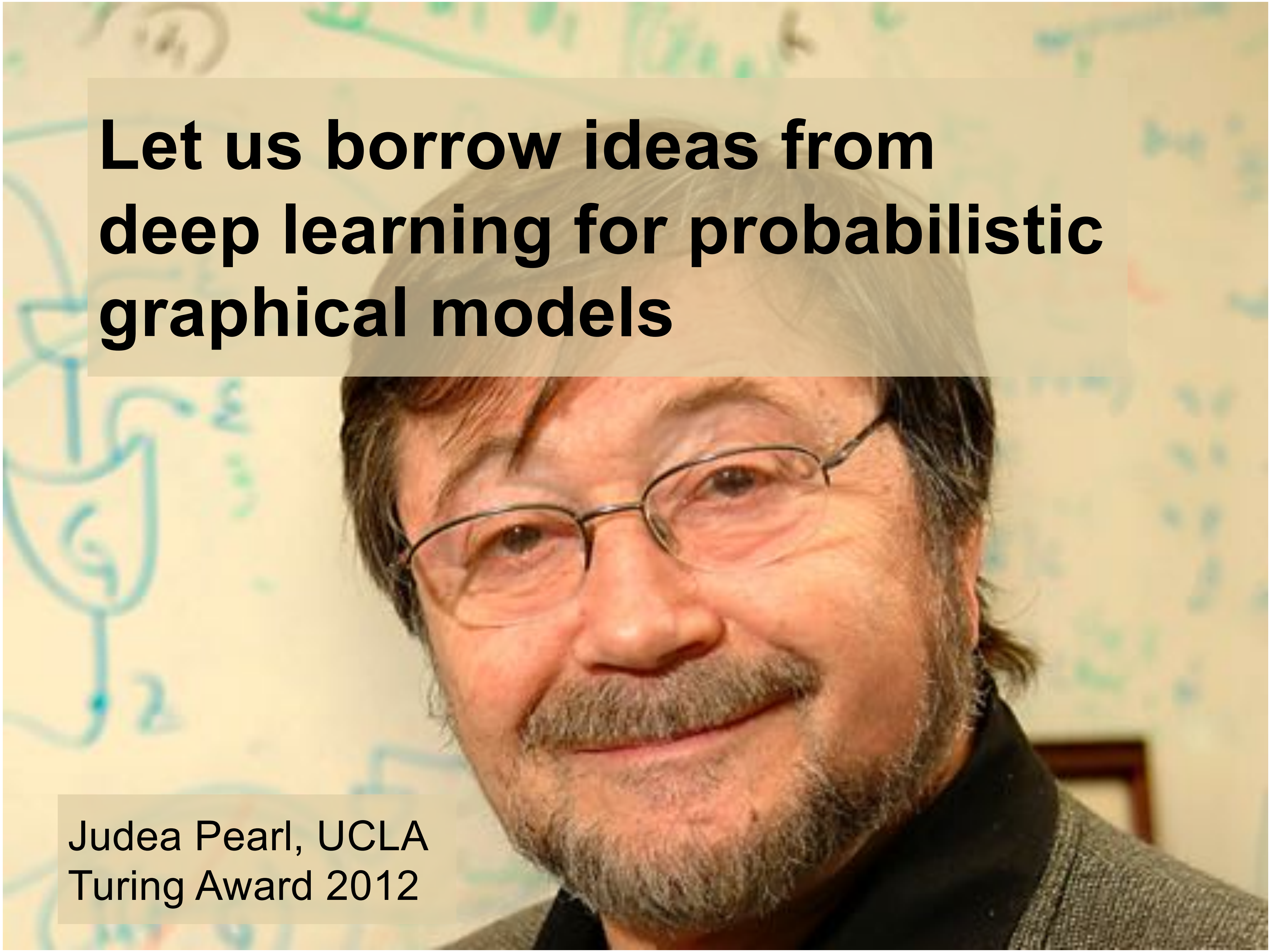
[Bradshaw et al. arXiv:1707.02476 2017]




**Many DNNs cannot distinguish the datasets**

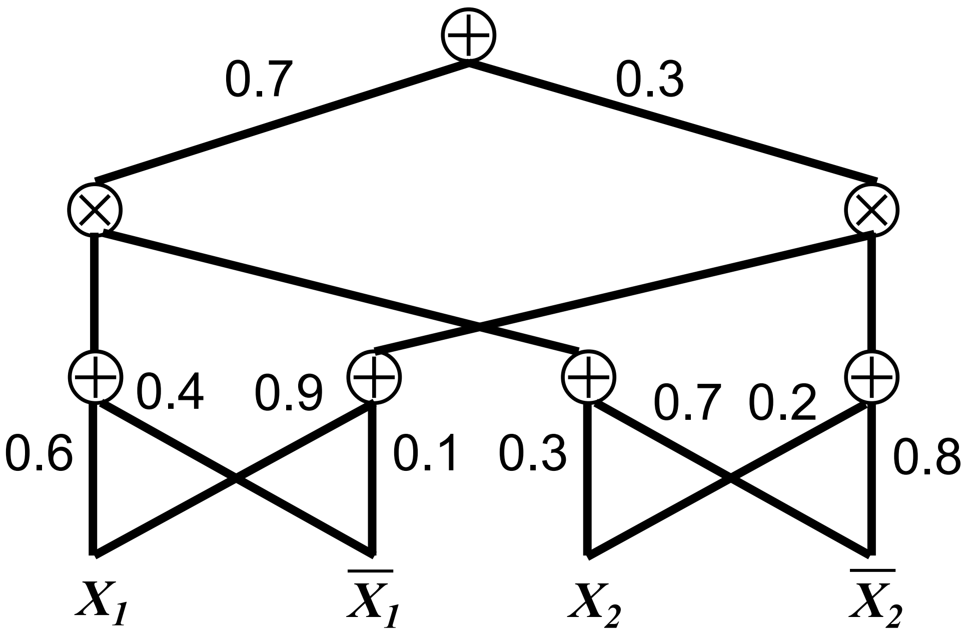
[Peharz, Vergari, Molina, Stelzner, Trapp, Kersting, Ghahramani UAI 2019]

# Let us borrow ideas from deep learning for probabilistic graphical models



Judea Pearl, UCLA
Turing Award 2012

# Sum-Product Networks
## a deep probabilistic learning framework


Computational graph (kind of TensorFlow graphs) that encodes how to compute probabilities

**Inference is linear in size of network**

[Poon, Domingos UAI'11; Molina, Natarajan, Kersting AAAI'17; Vergari, Peharz, Di Mauro, Molina, Kersting, Esposito AAAI '18; Molina, Vergari, Di Mauro, Esposito, Natarajan, Kersting AAAI '18, Peharz et al. UAI 2019, Stelzner, Peharz, Kersting iCML 2019]

# There are also ways to learn the structure of SPNs

SPFlow: An Easy and Extensible Library for Sum-Product Networks

[Molina, Vergari, Stelzner, Peharz, Subramani, Poupart, Di Mauro, Kersting arXiv:1901.03704, 2019]

TECHNISCHE UNIVERSITÄT DARMSTADT · UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO · UNIVERSITY OF WATERLOO · CAML · MADESI · Max Planck Institute for Intelligent Systems · UNIVERSITY OF CAMBRIDGE · VECTOR INSTITUTE · DFG · Federal Ministry of Education and Research

195 commits · 2 branches · 0 releases · 6 contrib...

Branch: master · New pull request · Create new file · Upload files · Find file · Clone or download

**https://github.com/SPFlow/SPFlow**

```
from spn.structure.leaves.parametric.Parametric import Categorical

from spn.structure.Base import Sum, Product

from spn.structure.base import assign_ids, rebuild_scopes_bottom_up


p0 = Product(children=[Categorical(p=[0.3, 0.7], scope=1), Categorical(p=[0.4, 0.6], scope=2)])
p1 = Product(children=[Categorical(p=[0.5, 0.5], scope=1), Categorical(p=[0.6, 0.4], scope=2)])
s1 = Sum(weights=[0.3, 0.7], children=[p0, p1])
p2 = Product(children=[Categorical(p=[0.2, 0.8], scope=0), s1])
p3 = Product(children=[Categorical(p=[0.2, 0.8], scope=0), Categorical(p=[0.3, 0.7], scope=1)])
p4 = Product(children=[p3, Categorical(p=[0.4, 0.6], scope=2)])
spn = Sum(weights=[0.4, 0.6], children=[p2, p4])

assign_ids(spn)
rebuild_scopes_bottom_up(spn)

return spn
```

**Domain Specific Language, Inference, EM, and Model Selection as well as Compilation of SPNs into TF and PyTorch and also into flat, library-free code even suitable for running on devices: C/C++,GPU, FPGA**

SPFlow, an open-source Python library providing a simple interface to inference, learning and manipulation routines for deep and tractable probabilistic models called Sum-Product Networks (SPNs). The library allows one to quickly create SPNs both from data and through a domain specific language (DSL). It efficiently implements several probabilistic inference

# Random sum-product networks

UNIVERSITY OF CAMBRIDGE

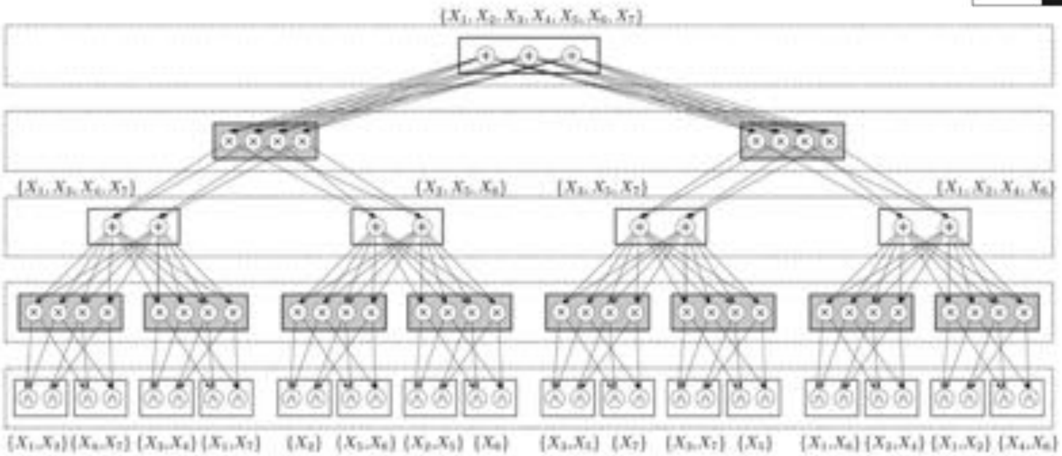
Max Planck Institute for Intelligent Systems

TECHNISCHE UNIVERSITÄT DARMSTADT

UBER AI Labs

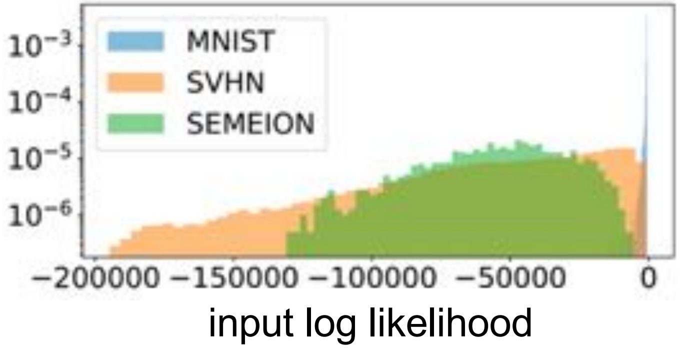
UNI GRAZ

Conference on Uncertainty in Artificial Intelligence
Tel Aviv, Israel
July 22 - 25, 2019
uai2019



**Similar to Random Forests, build a random SPN structure. This can be done in an informed way or completely at random**

| | | RAT-SPN | MLP | vMLP |
|---|---|---|---|---|
| Accuracy | MNIST | 98.19 (8.5M) | 98.32 (2.64M) | 98.09 (5.28M) |
| | F-MNIST | 89.52 (0.65M) | 90.81 (9.28M) | 89.81 (1.07M) |
| | 20-NG | 47.8 (0.37M) | 49.05 (0.31M) | 48.81 (0.16M) |
| Cross-Entropy | MNIST | 0.0852 (17M) | 0.0874 (0.82M) | 0.0974 (0.22M) |
| | F-MNIST | 0.3525 (0.65M) | 0.2965 (0.82M) | 0.325 (0.29M) |
| | 20-NG | 1.6954 (1.63M) | 1.6180 (0.22M) | 1.6263 (0.22M) |




outliers
prototypes
outliers
prototypes

**SPNs can have similar predictive performances as (simple) DNNs**

**SPNs can distinguish the datasets**

**SPNs know when they do not know by design**

# Unsupervised scene understanding

ICML | 2019
Thirty-sixth International Conference on Machine Learning

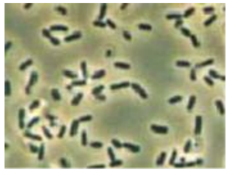
[Stelzner, Peharz, Kersting ICML 2019, Best Paper Award at TPM@ICML2019]

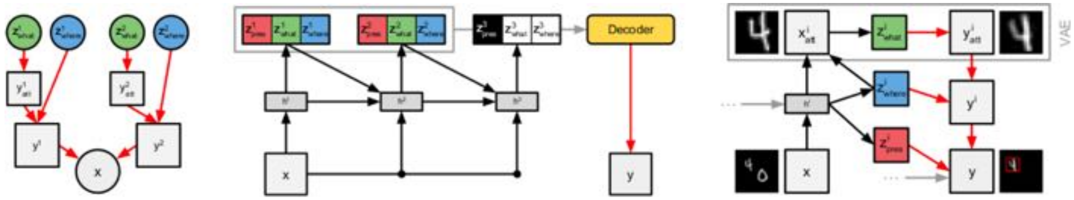
UNIVERSITY OF CAMBRIDGE

TECHNISCHE UNIVERSITÄT DARMSTADT

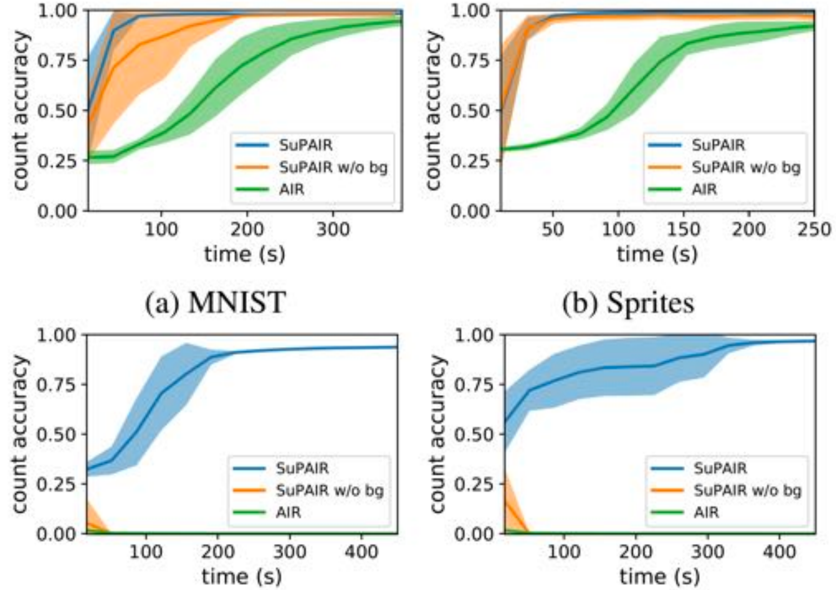
https://github.com/stelzner/supair

Consider e.g. unsupervised scene understanding using a generative model implemented in a neural fashion





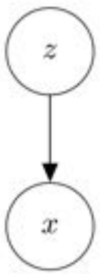

[Attend-Infer-Repeat (AIR) model, Hinton et al. NIPS 2016]




(a) MNIST (b) Sprites

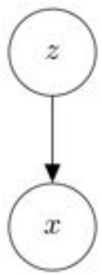
(c) Noisy MNIST (d) Grid MNIST

VAE



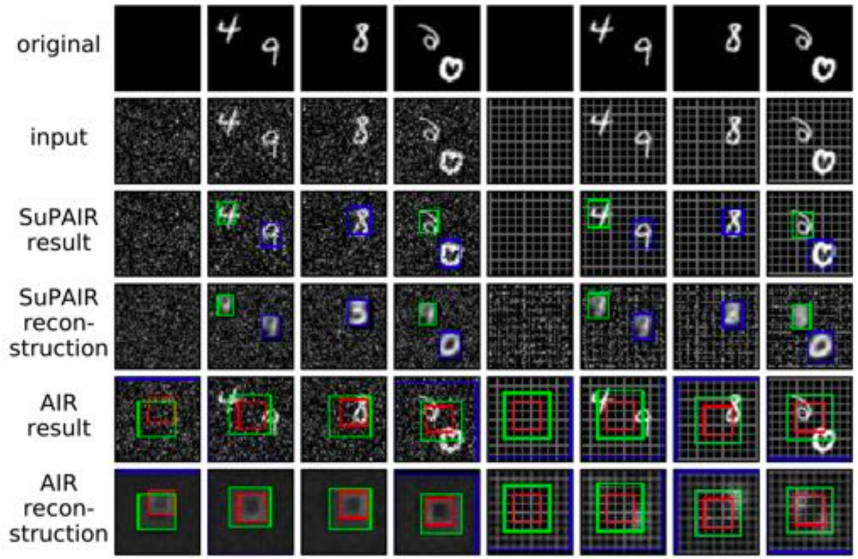
**Replace VAE by SPN as object model**

SPN



- infinite mixture model
- intractable density
- intractable posterior

- "large" but finite mixture model
- tractable density
- tractable marginals [Peharz et al., 2015]
- tractable posterior [Vergari et al., 2017]

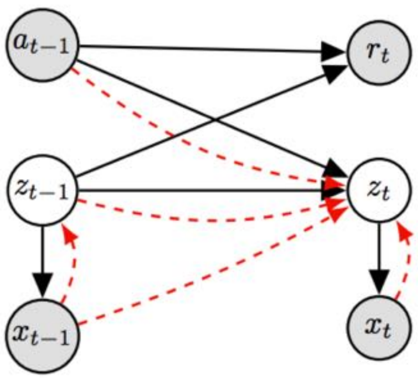


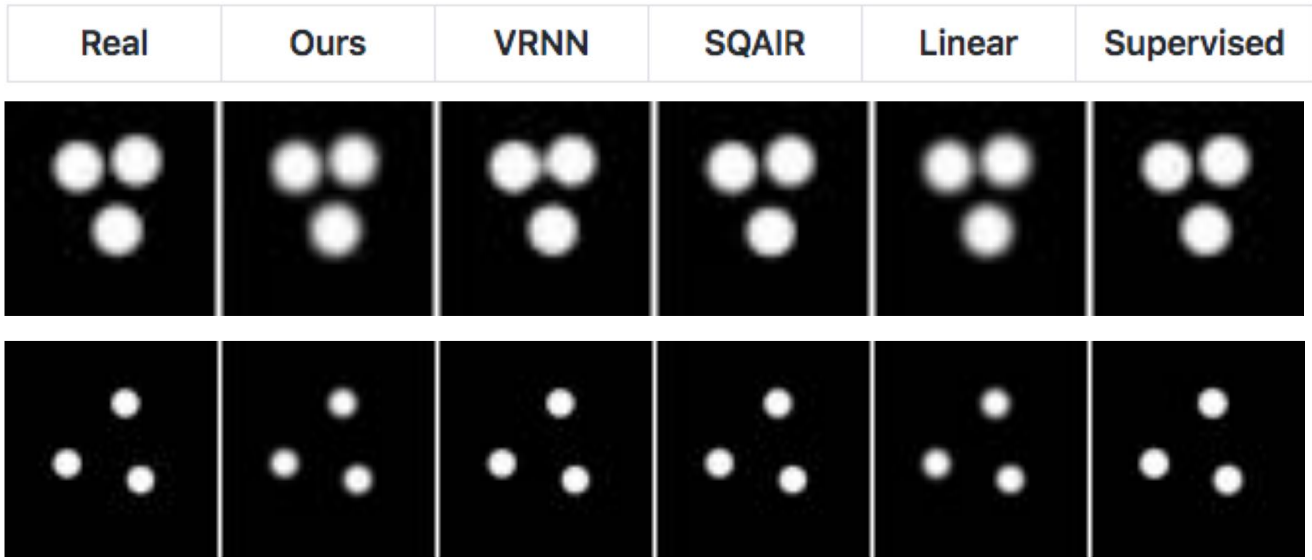
# Unsupervised physics learning

ICLR | 2020

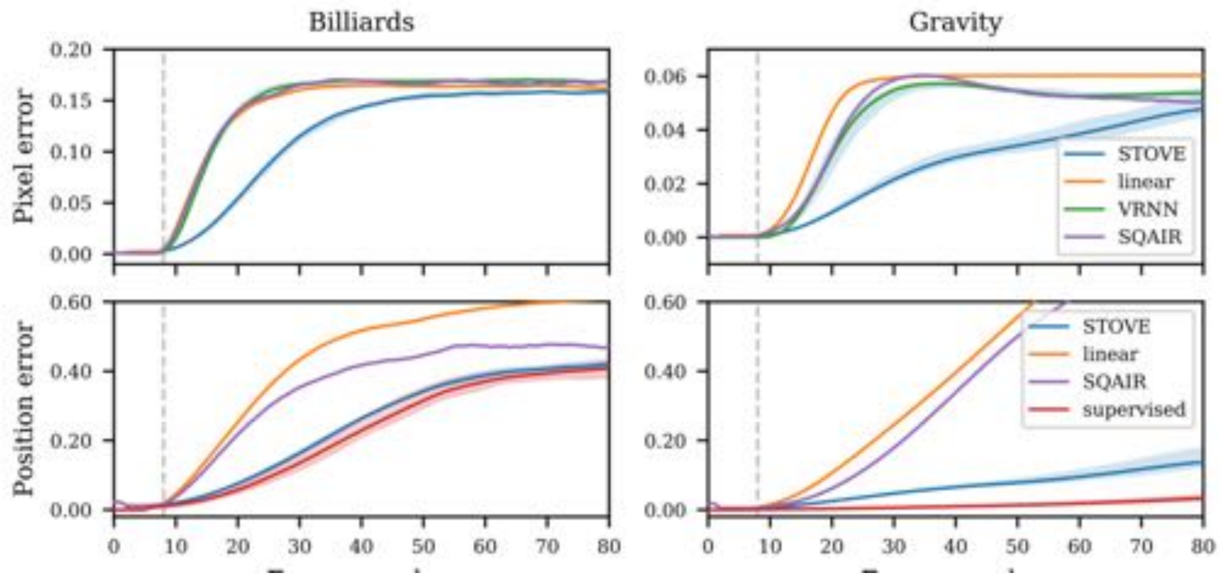
Eighth International Conference on Learning Representations

[Kossen, Stelzner, Hussing, Voelcker, Kersting ICLR 2020]




putting structure and tractable inference into deep models



# Deep Likelihoods for Time Series

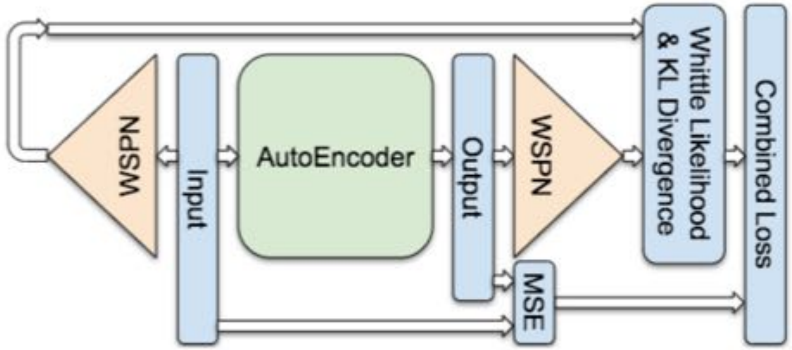
[Yu, Kersting 2020 to be submitted]

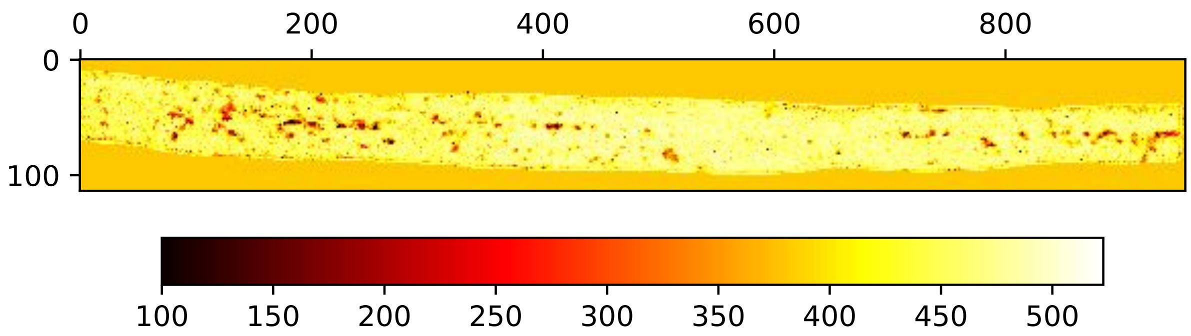
Whittle Likelihood

$$p(\mathbf{X}_{1:N} \mid S_{0:T-1}) \approx \prod_{n=1}^{N} \prod_{k=0}^{T-1} \frac{1}{\pi^p |S_k|} e^{-d_{nk}^* S_k^{-1} d_{nk}}$$

Train a complex-valued SPN in the spectral domain of time series

End2end integration with the deep learning stack




Natural anomaly detection, here for plant data

Federal Ministry of Education and Research

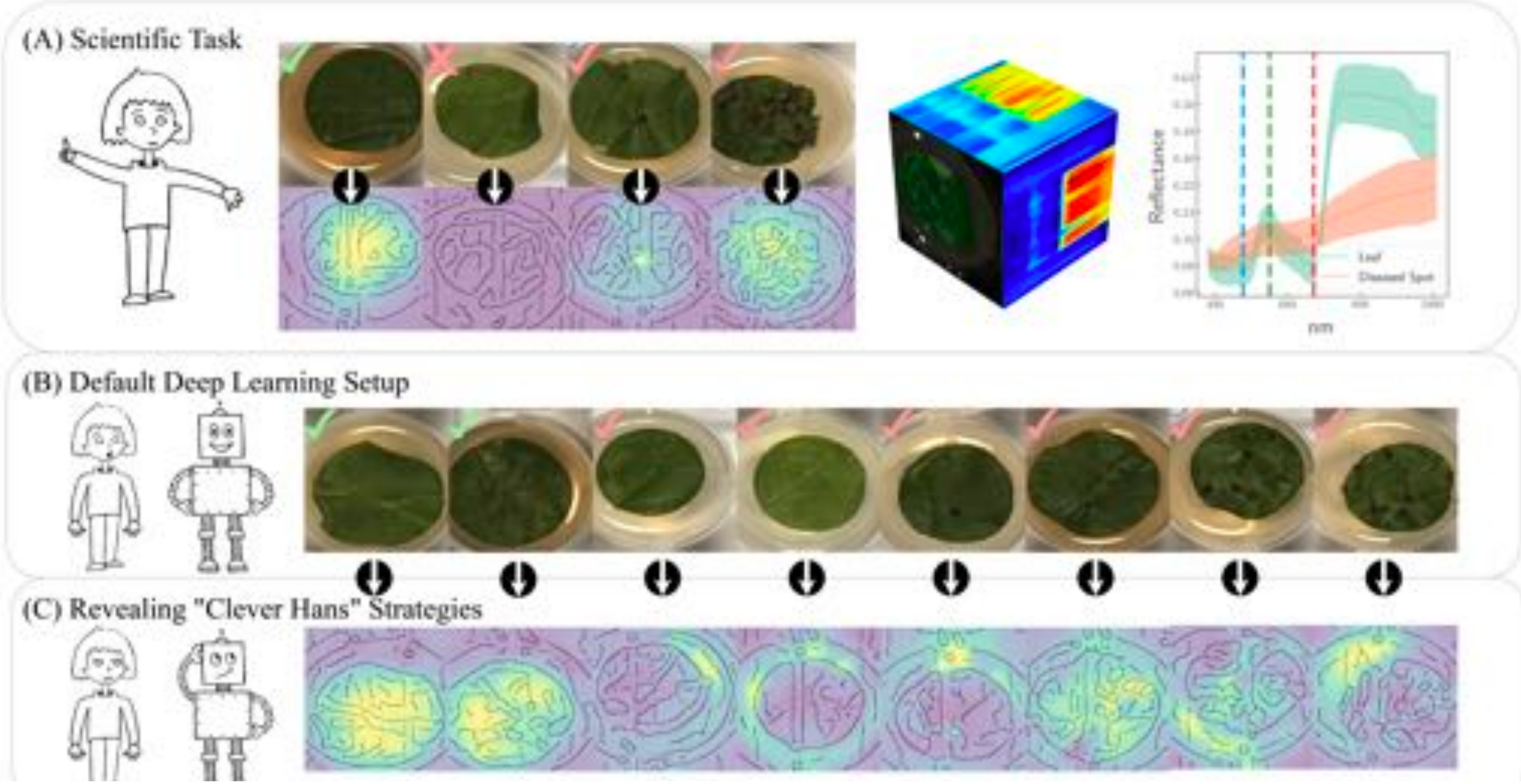
MADESI

# Un-“Hans“-ing Deep Learning

AAAI / ACM conference on
ARTIFICIAL INTELLIGENCE,
ETHICS, AND SOCIETY

**Experts-in-the-loop ML:** expert is changing computer behavior but also adapts in response to what is learned

[Teso, Kersting AIES 2019, Schramowski et al. arXiv:2001.05371 2020]




AudisSens DePhenSe

Bundesanstalt für
Landwirtschaft und Ernährung

UNIVERSITY OF CAMBRIDGE

UBER AI Labs

Max Planck Institute for Intelligent Systems

UNI GRAZ

FLOW

TECHNISCHE UNIVERSITÄT DARMSTADT

Microsoft Research

UBER AI Labs

KATHOLIEKE UNIVERSITEIT LEUVEN

UTD THE UNIVERSITY OF TEXAS AT DALLAS

TECHNISCHE UNIVERSITÄT DARMSTADT

UBC THE UNIVERSITY OF BRITISH COLUMBIA

MORGAN & CLAYPOOL PUBLISHERS

Statistical Relational Artificial Intelligence

*Logic, Probability, and Computation*

Luc De Raedt
Kristian Kersting
Sriraam Natarajan
David Poole

SYNTHESIS LECTURES ON ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

relationalAI

AI for the enterprise

**Getting deep systems that reason and know when they don't know**

**Responsible AI systems that explain their decisions and co-evolve with the humans**

**Open AI systems that are easy to realize and understandable for the domain experts**

„Tell the AI when it is right for the wrong reasons and it adapts its behavior"

(a) Original Image (b) Explaining *Electric guitar* (c) Explaining *Acoustic guitar* (d) Explaining *Labrador*

**Figure 4: Explaining an image classification prediction made by Google's Inception network, highlighting positive pixels. The top 3 classes predicted are "Electric Guitar" ($p = 0.32$), "Acoustic guitar" ($p = 0.24$) and "Labrador" ($p = 0.21$)**

Teso, Kersting AIES 2019

AAAI / ACM conference on ARTIFICIAL INTELLIGENCE, ETHICS, AND SOCIETY

**All this is really a team sport! Let’s join efforts! Thanks for you attention**