IEEE
WCCI 2020
IEEE World Congress on Computational Intelligence
Virtual Conference – July 19-24, 2020

Augmented AI: correctors of errors and social networks of AI

Alexander N. Gorban

SPONSORS:
IEEE Advancing Technology for Humanity
IEEE Computational Intelligence Society
IET
The International Neural Network Society (INNS)
EPS Evolutionary Programming Society

Training Neural Networks, USSR-USA JV ParaGraph, **1990**

Neural Networks on PC, Nauka, **1996**

Neuoinformatics, Nauka, **1998**

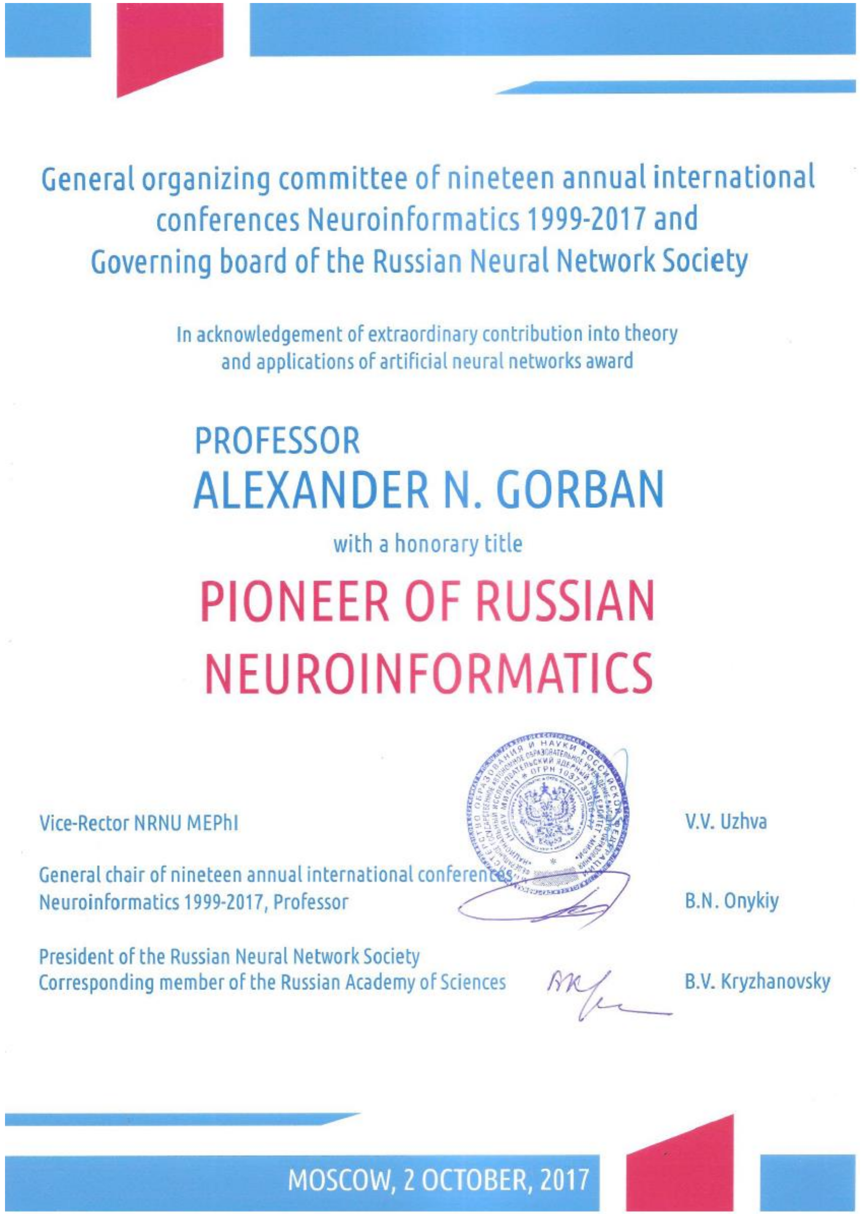
1978-2003

Russian Academy of Sciences, Siberian Branch

Training Neural Networks, USSR-USA JV ParaGraph, **1990**

Neural Networks on PC, Nauka, **1996**

Neuoinformatics, Nauka, **1998**


General organizing committee of nineteen annual international conferences Neuroinformatics 1999-2017 and Governing board of the Russian Neural Network Society

In acknowledgement of extraordinary contribution into theory and applications of artificial neural networks award

PROFESSOR
ALEXANDER N. GORBAN
with a honorary title
PIONEER OF RUSSIAN NEUROINFORMATICS

Vice-Rector NRNU MEPhI — V.V. Uzhva

General chair of nineteen annual international conferences Neuroinformatics 1999-2017, Professor — B.N. Onykiy

President of the Russian Neural Network Society
Corresponding member of the Russian Academy of Sciences — B.V. Kryzhanovsky

MOSCOW, 2 OCTOBER, 2017

1978-2003
Russian Academy of Sciences, Siberian Branch

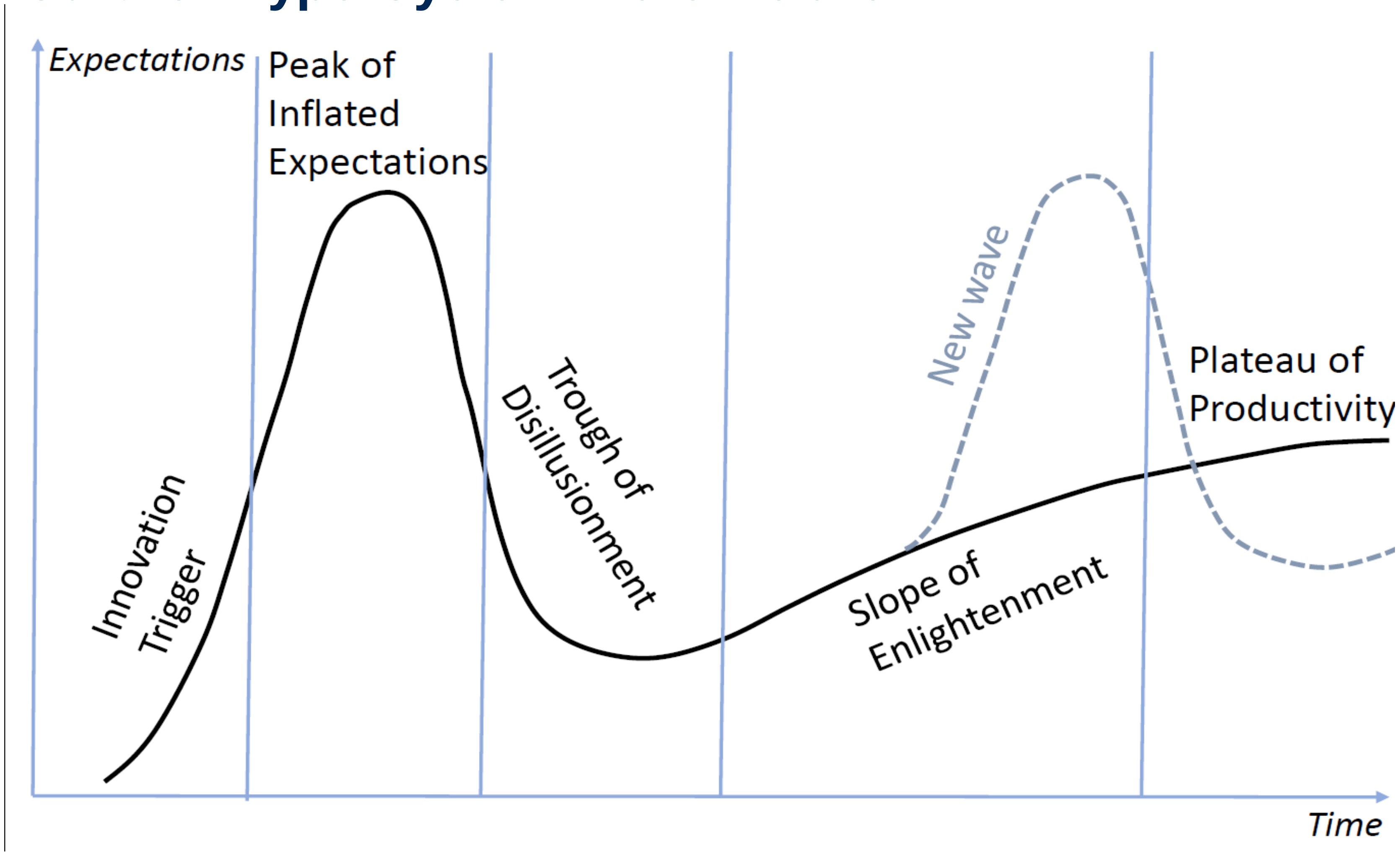
UNIVERSITY OF LEICESTER

University of Leicester, AIDAM –
Centre for Artificial Intelligence, Data Analysis and Modelling

A problem about the problem

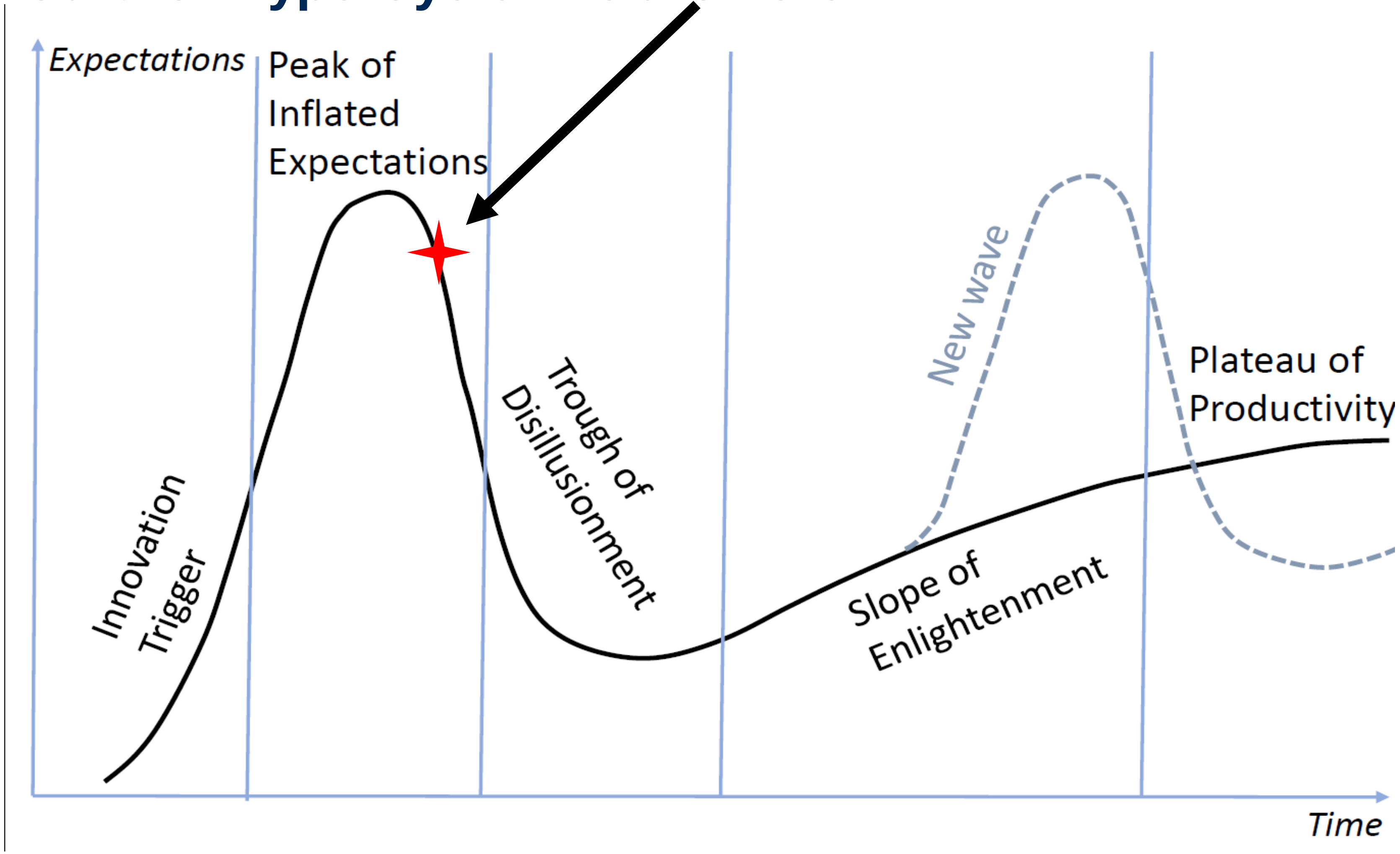
A problem about the problem

Gartner Hype Cycle: where we are?



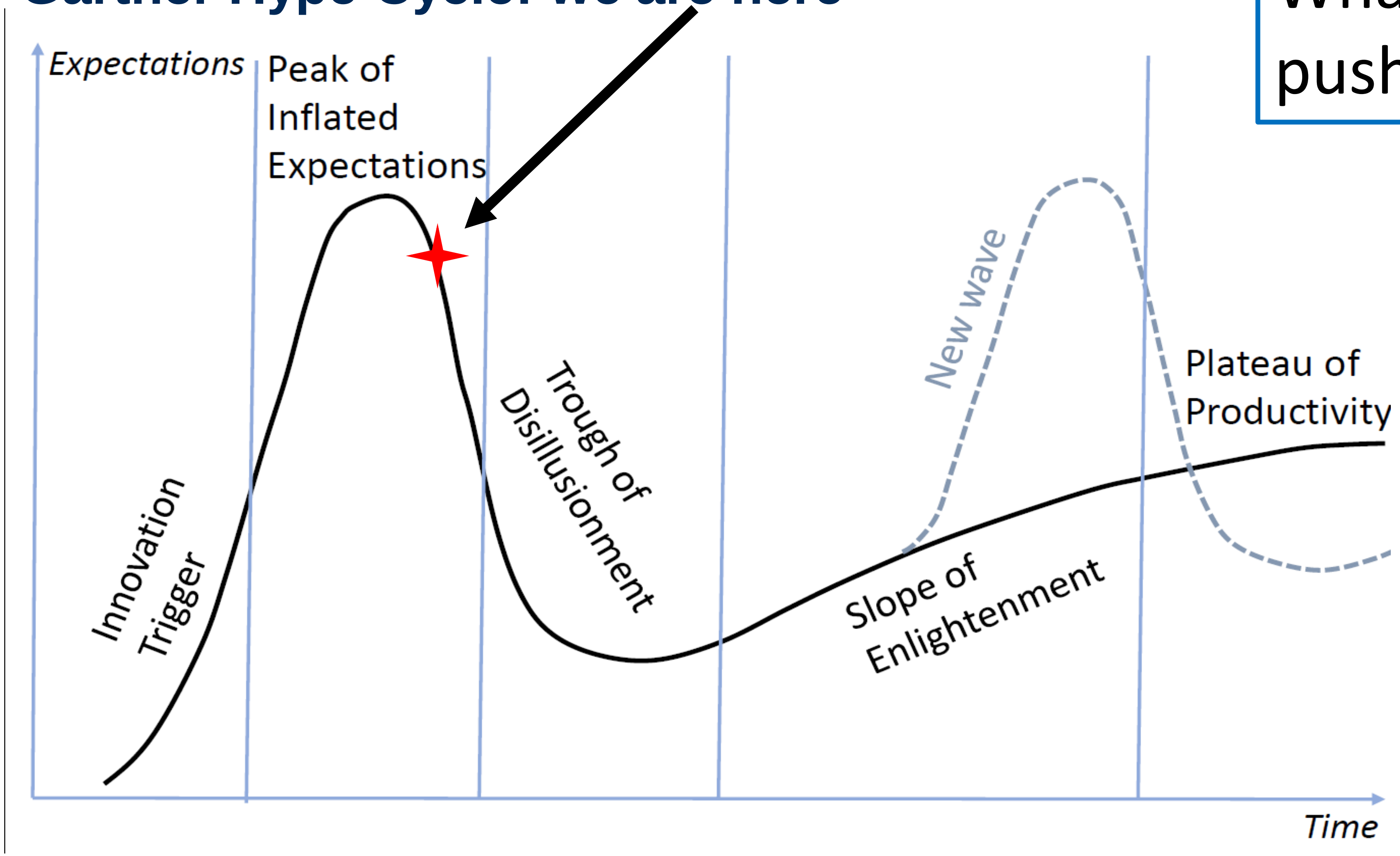
A problem about the problem

Gartner Hype Cycle: we are here



A problem about the problem

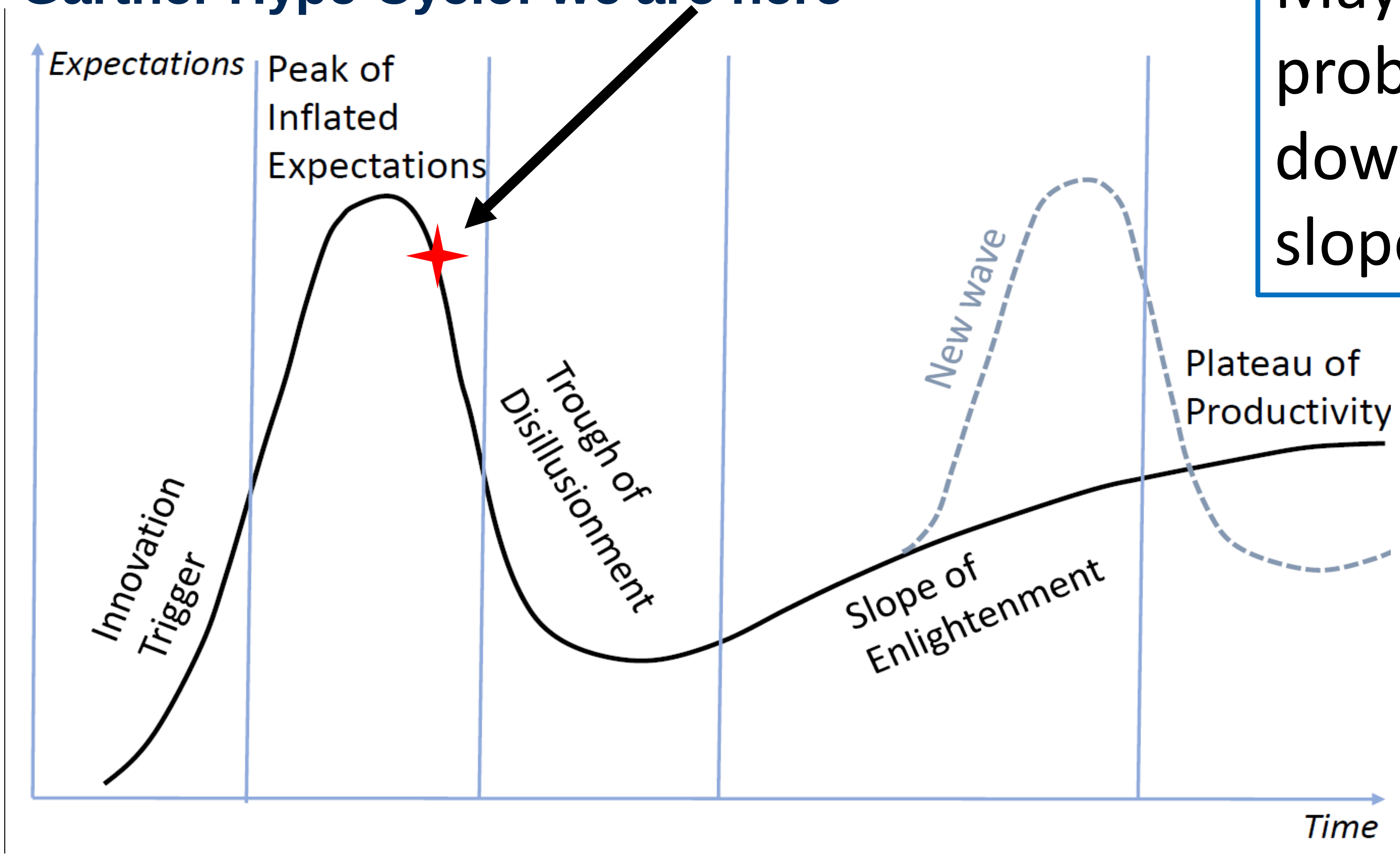
Gartner Hype Cycle: we are here



What problems will push us down the hill?

A problem about the problem

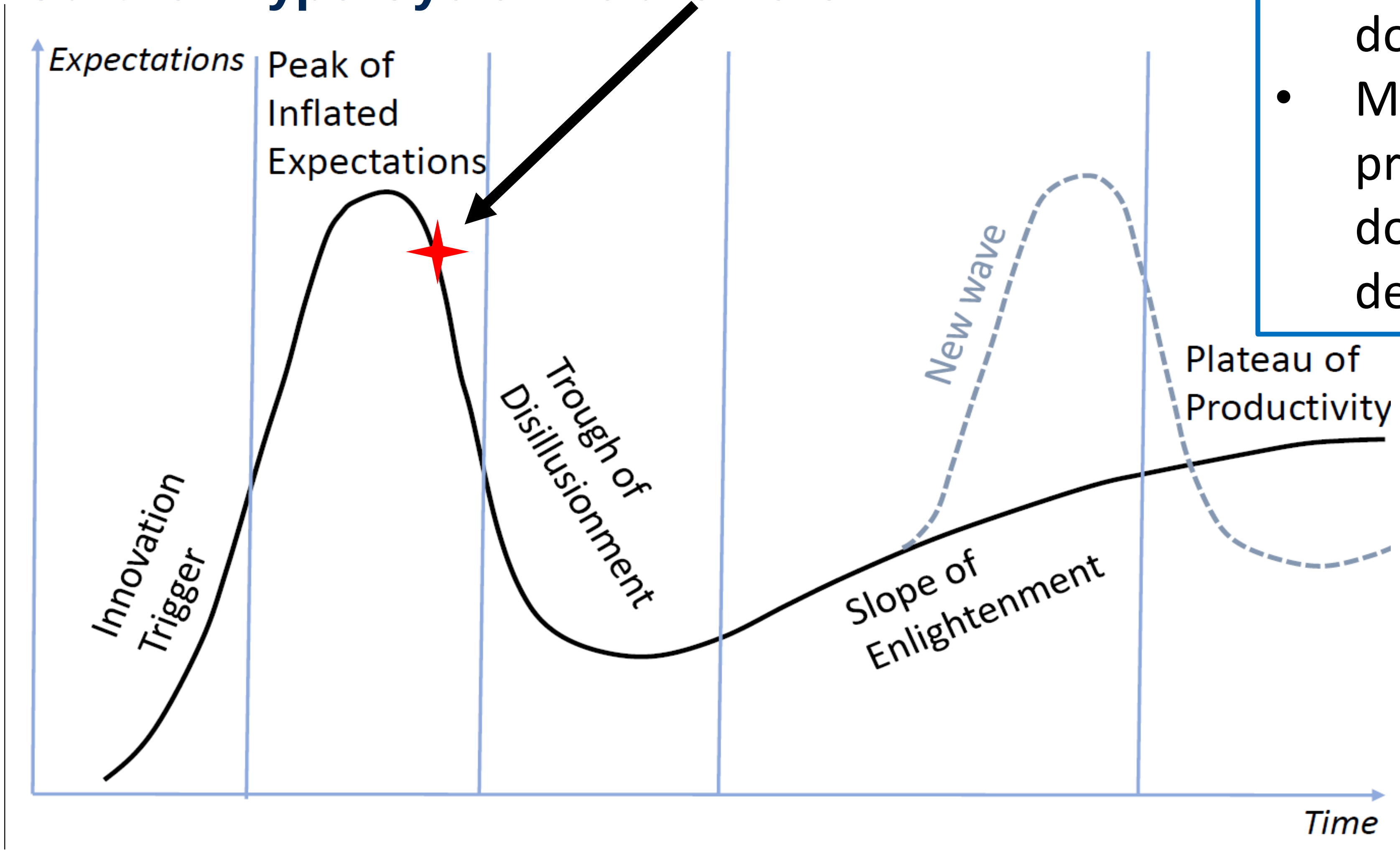
Gartner Hype Cycle: we are here



Maybe this is the same problem that pushed us down the previous slopes, decades ago?

A problem about the problem

Gartner Hype Cycle: we are here

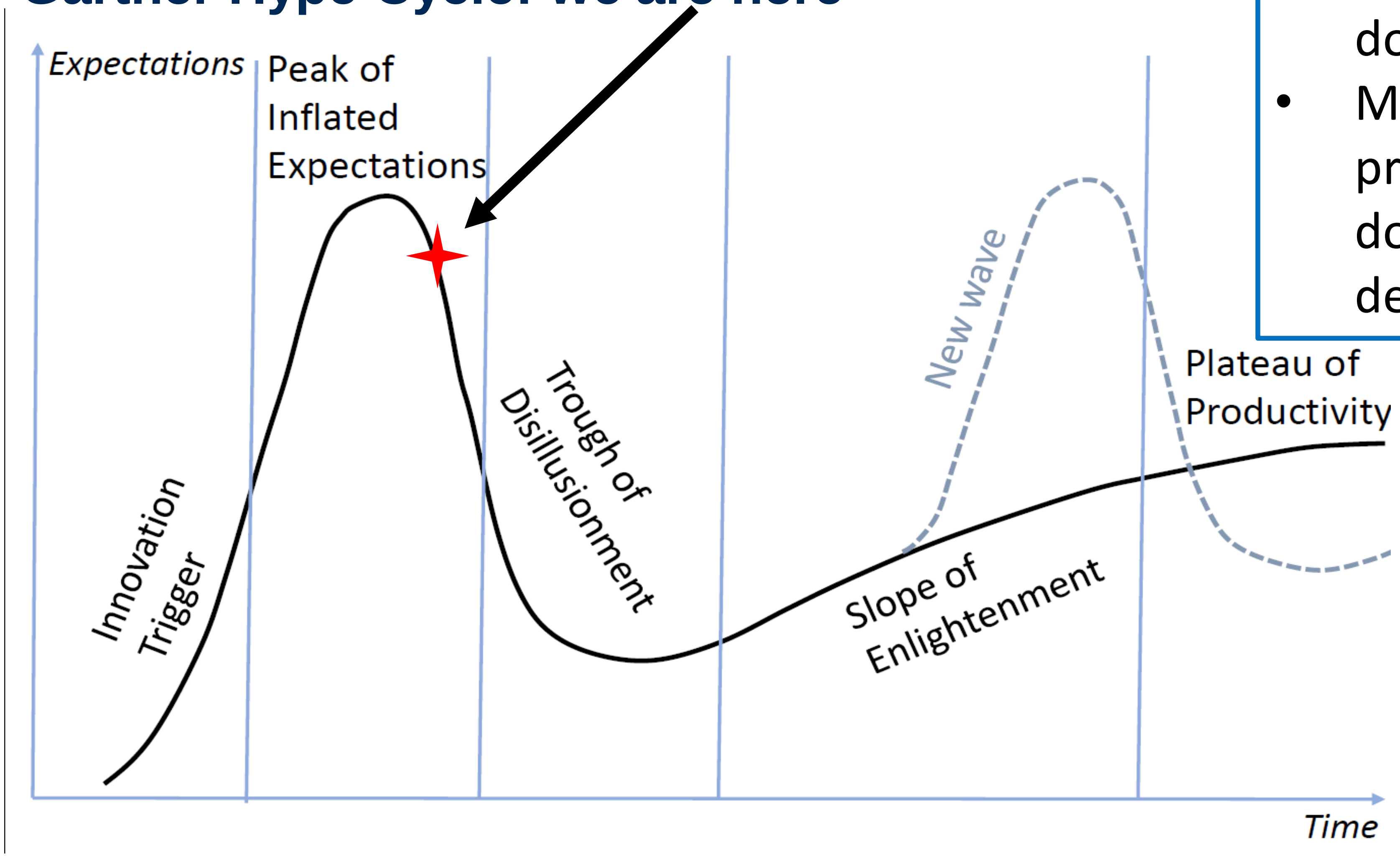


- What problems will push us down the hill?
- Maybe this is the same problem that pushed us down the previous slopes, decades ago?

I try to answer these questions about problems

A problem about the problem

Gartner Hype Cycle: we are here



- What problems will push us down the hill?
- Maybe this is the same problem that pushed us down the previous slopes, decades ago?

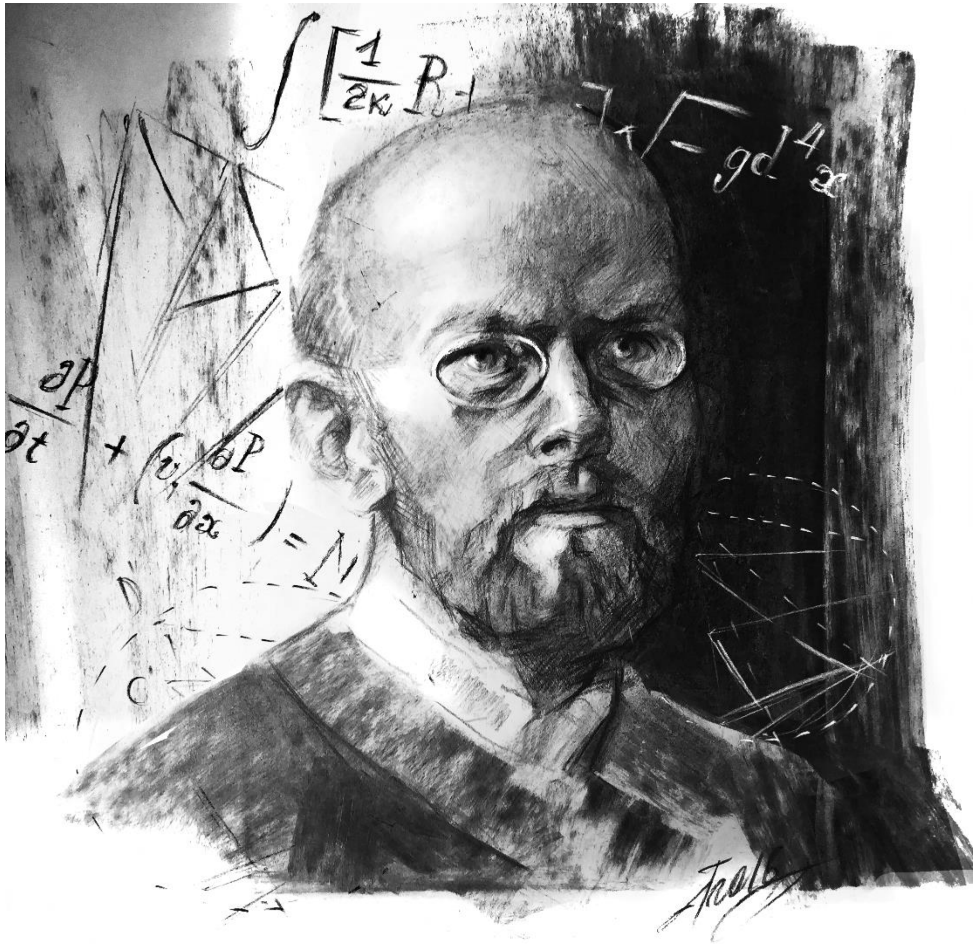
and give some ideas on how to solve these problems.

World congress is a proper place to discuss problems

In Mathematics, this is a long tradition.

David Hilbert, Portrait by Anna Gorban

World congress is a proper place to discuss problems



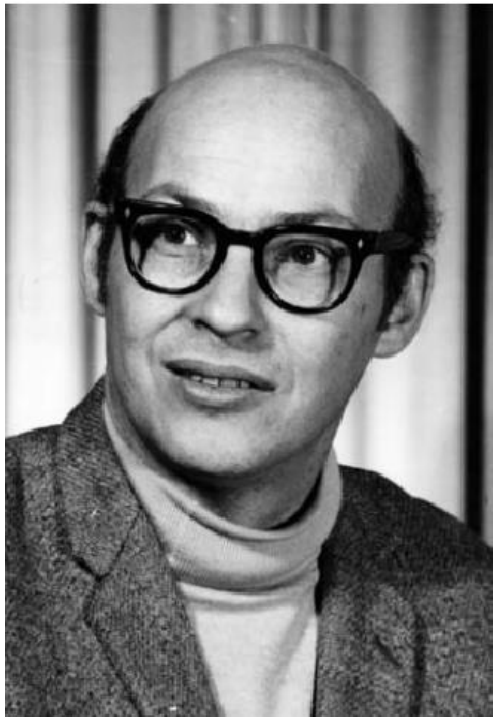
David Hilbert in 1900, Portrait by Anna Gorban

In Mathematics, this is a long tradition.

A famous example:

- In 1900, Hilbert presented his problems at the International Congress of Mathematicians.
- These problems stimulated the development of mathematics and focused the efforts of many mathematicians

Marvin Minsky's Problems for AI

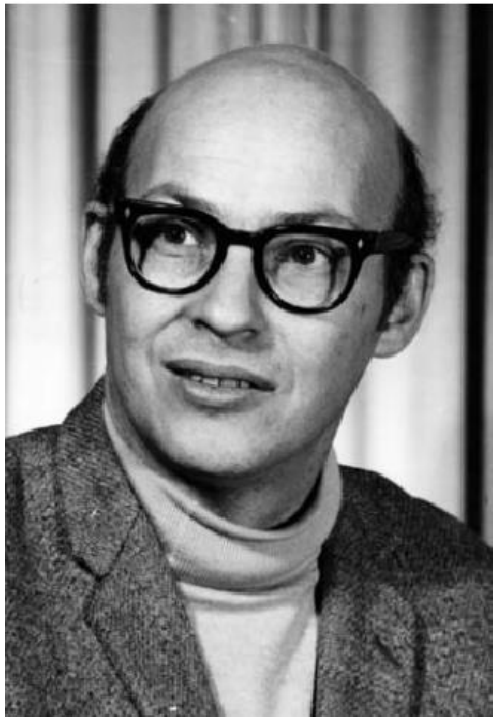


Marvin Minsky,
Semantic Scholar
courtesy MIT Museum,

"It is convenient to divide the problems into five main areas:

- Search,
- Pattern-Recognition,
- Learning,
- Planning,
- and Induction… .

Minsky, M. (1961). Steps toward artificial intelligence. *Proceedings of the IRE*, *49*(1), 8-30.



Marvin Minsky,
Semantic Scholar
courtesy MIT Museum,

Marvin Minsky's Problems for AI

"It is convenient to divide the problems into five main areas:

- Search,
- Pattern-Recognition,
- Learning,
- Planning,
- and Induction… .

- Minsky's problems focused research on the most important technical issues.
- We are currently seeing great success in these areas.

Minsky, M. (1961). Steps toward artificial intelligence. *Proceedings of the IRE*, *49*(1), 8-30.

Two pebbles in a shoe

- A mountain of success of Neural AI inspired us!

Two pebbles in a shoe

- A mountain of success of Neural AI inspired us!
- However, two pebbles in our shoes can slow down this fantastically successful movement:

Two pebbles in a shoe

- A mountain of success of Neural AI inspired us!
- However, two pebbles in our shoes can slow down this fantastically successful movement:

AI makes unexpected mistakes, and will make them in the future;

Two pebbles in a shoe

- A mountain of success of Neural AI inspired us!
- However, two pebbles in our shoes can slow down this fantastically successful movement:

AI makes unexpected mistakes, and will make them in the future;

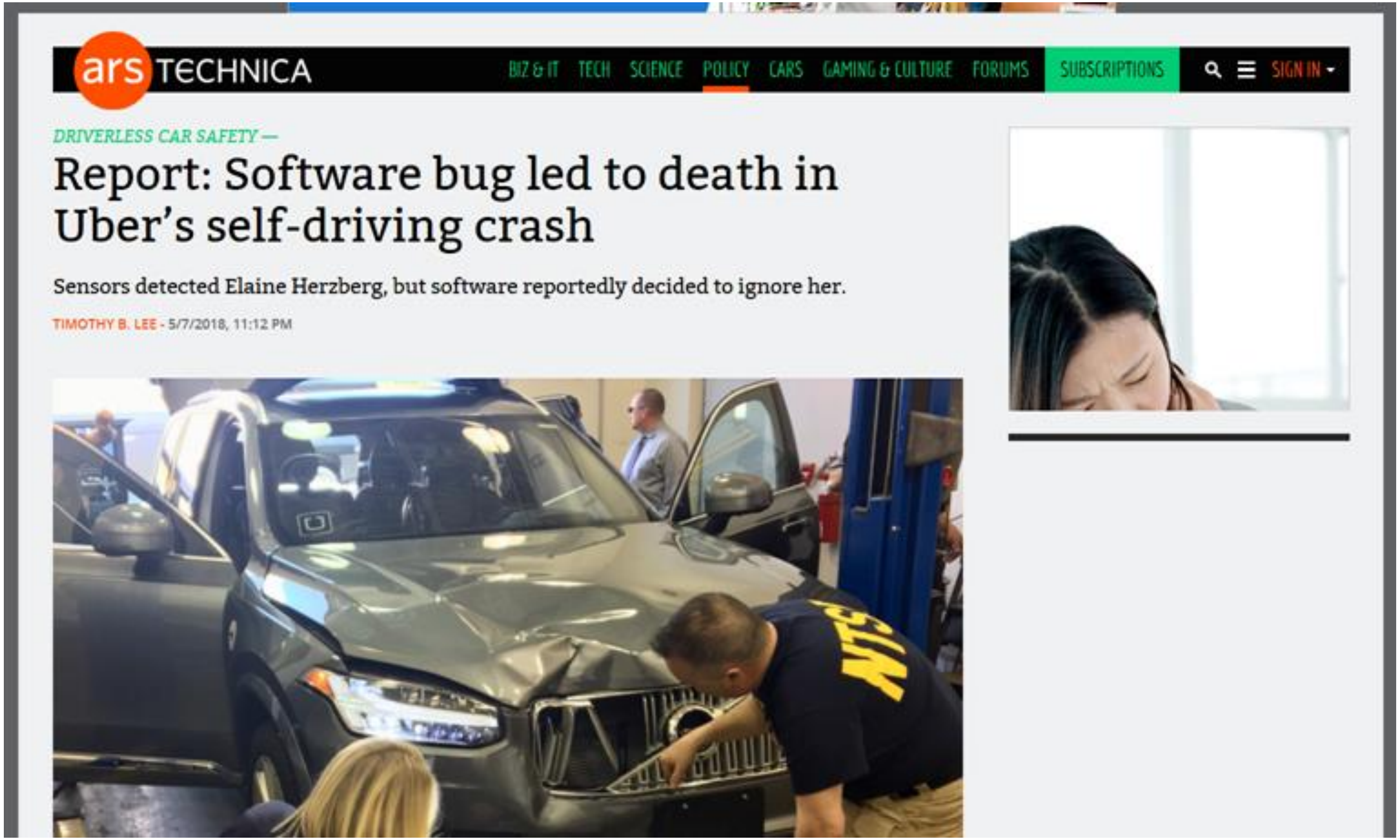
Decisions of Neural AI are not transparent and, therefore, cannot be explained logically.

The main problem for the widespread use of AI around the world are unexpected errors

- The mistakes can be dangerous;
- Usually, it remains unclear, who is responsible for them;
- To avoid a recurrence of a detected error, a quick, non-iterative system fix is required;
- Correction of errors should not damage existing skills;
- The real world is not a good i.i.d. sample,

  -the types of errors are very numerous and often unpredictable,

  -so we cannot rely on a statistical estimate of the probability of errors.



ars TECHNICA

BIZ & IT TECH SCIENCE POLICY CARS GAMING & CULTURE FORUMS SUBSCRIPTIONS SIGN IN

DRIVERLESS CAR SAFETY —

Report: Software bug led to death in Uber's self-driving crash

Sensors detected Elaine Herzberg, but software reportedly decided to ignore her.

TIMOTHY B. LEE - 5/7/2018, 11:12 PM

The main problem for the widespread use of AI around the world are unexpected errors

- **The mistakes can be dangerous;**
- Usually, it remains unclear, who is responsible for them;
- To avoid a recurrence of a detected error, a quick, non-iterative system fix is required;
- Correction of errors should not damage existing skills;
- The real world is not a good i.i.d. sample,

  -the types of errors are very numerous and often unpredictable,

  -so we cannot rely on a statistical estimate of the probability of errors.

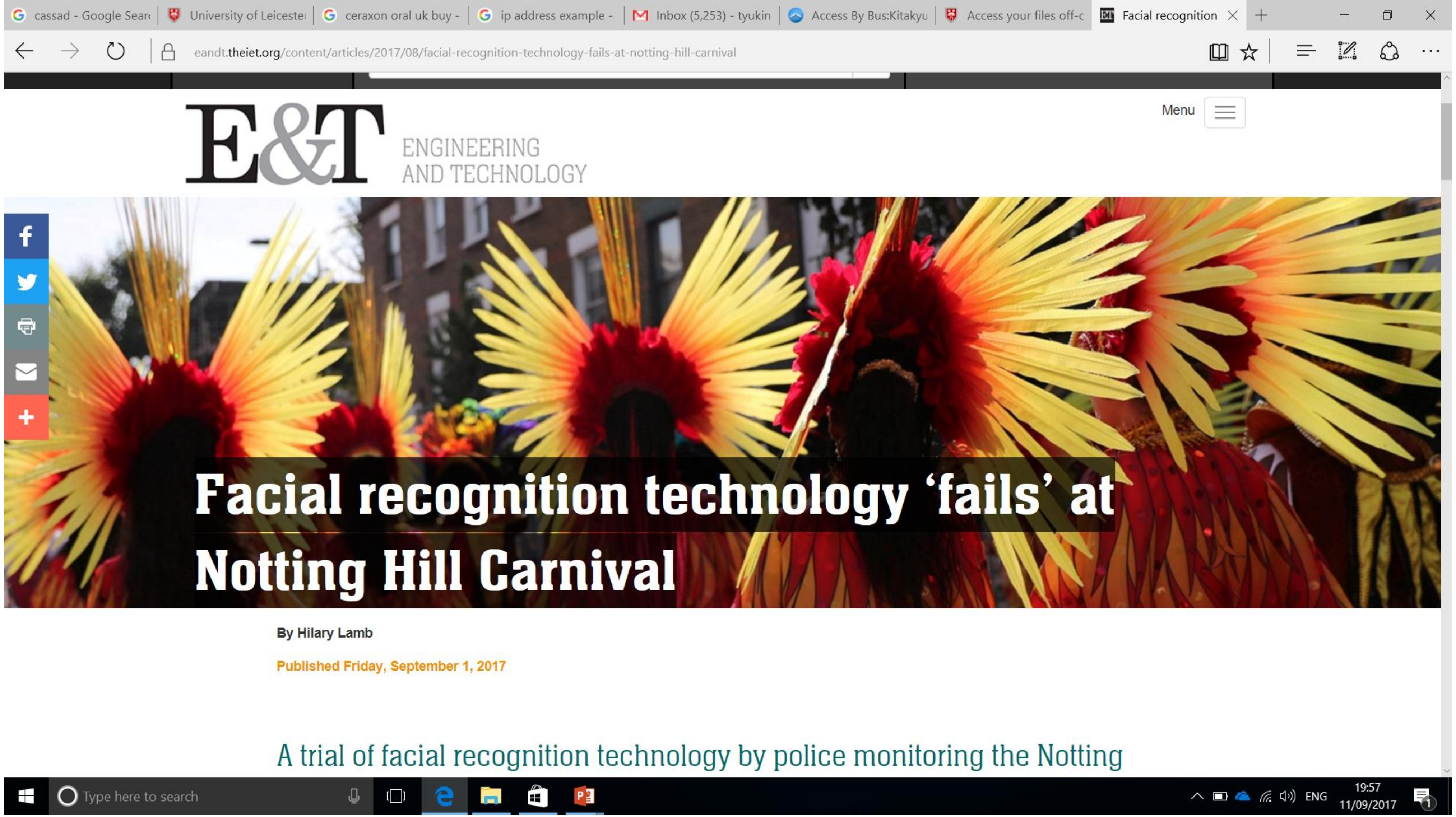
A Tesla electric car crashed into a highway barrier in Mountain View, California, on March 23, 2018. Investigators confirmed that Autopilot was partially to blame.

The main problem for the widespread use of AI around the world are unexpected errors

- The mistakes can be dangerous;
- **Usually, it remains unclear, who is responsible for them;**
- To avoid a recurrence of a detected error, a quick, non-iterative system fix is required;
- Correction of errors should not damage existing skills;
- The real world is not a good i.i.d. sample,

  -the types of errors are very numerous and often unpredictable,

  -so we cannot rely on a statistical estimate of the probability of errors.



eandt.theiet.org/content/articles/2017/08/facial-recognition-technology-fails-at-notting-hill-carnival

E&T ENGINEERING AND TECHNOLOGY

Facial recognition technology 'fails' at Notting Hill Carnival

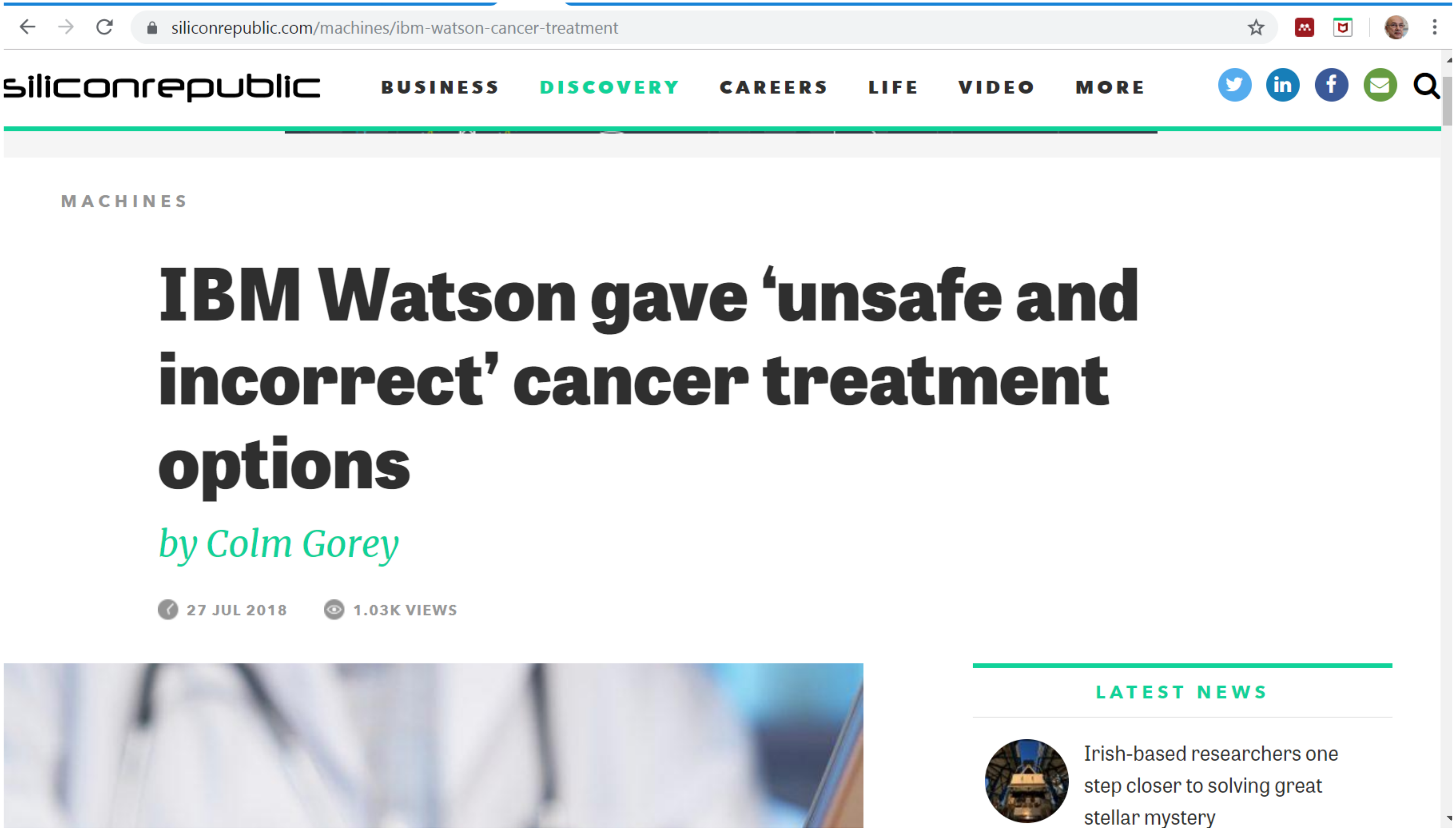
By Hilary Lamb

Published Friday, September 1, 2017

A trial of facial recognition technology by police monitoring the Notting

The main problem for the widespread use of AI around the world are unexpected errors

- The mistakes can be dangerous;
- Usually, it remains unclear, who is responsible for them;
- **To avoid a recurrence of a detected error, a quick, non-iterative system fix is required;**
- Correction of errors should not damage existing skills;
- The real world is not a good i.i.d. sample,
  - -the types of errors are very numerous and often unpredictable,
  - -so we cannot rely on a statistical estimate of the probability of errors.



siliconrepublic.com/machines/ibm-watson-cancer-treatment

siliconrepublic

BUSINESS DISCOVERY CAREERS LIFE VIDEO MORE

MACHINES

IBM Watson gave 'unsafe and incorrect' cancer treatment options

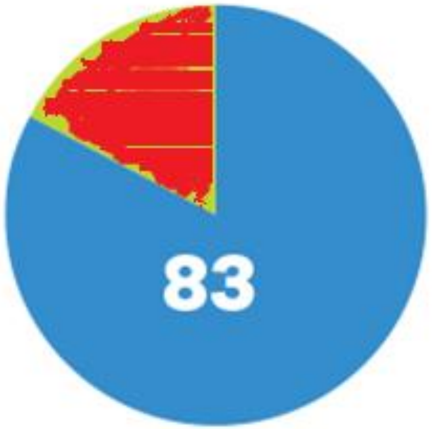
by Colm Gorey

27 JUL 2018 1.03K VIEWS

LATEST NEWS

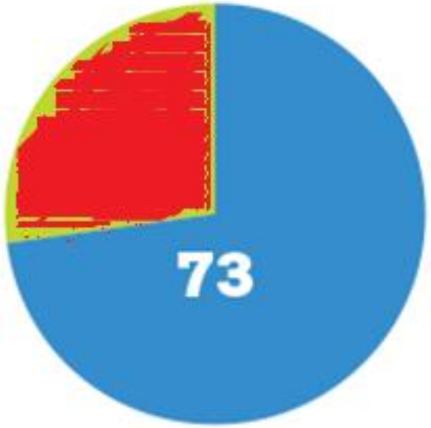
Irish-based researchers one step closer to solving great stellar mystery

Bumrungrad International Hospital, Thailand: 83% concordance



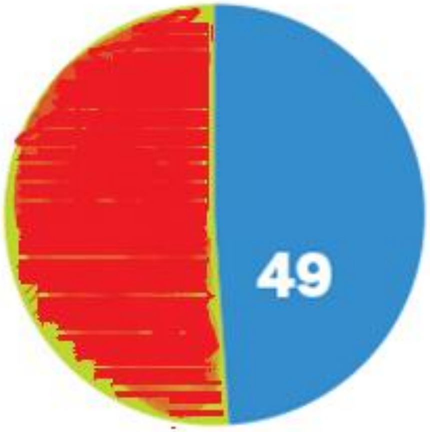
211 patients with breast, colorectal, gastric, and lung cancer

Manipal Comprehensive Cancer Center, India: 73% concordance



638 patients with breast cancer

Gachon University Gil Medical Center, South Korea: 49% concordance



656 patients with colon cancer

A 2016 audit by the University of Texas found that
the cancer centre spent $62 million on the project before canceling it.

The main problem for the widespread use of AI around the world are unexpected errors

- The mistakes can be dangerous;
- Usually, it remains unclear, who is responsible for them;
- To avoid a recurrence of a detected error, a quick, non-iterative system fix is required;
- **Correction of errors should not damage existing skills;**
- The real world is not a good i.i.d. sample,
  - -the types of errors are very numerous and often unpredictable,
  - -so we cannot rely on a statistical estimate of the probability of errors.

AI Errors: and the list continues ...

- **InsiperoBot** Poor Testing/Coding
- **Cortana** Not Working Bias/Limited Training Set Poor Testing/Coding
- **Alexa** Blasting Music Unexpected Human Behaviour
- **Passport Checker** Bias/Limited Training Set Poor Testing/Coding
- **Knightscope Robot** Hits Toddler Poor Testing/Coding Unexpected Human Behaviour
- **Facebook Translate** Arrest, Bias/Limited Training Set Unexpected Human Behaviour
- **Google Tag** Racist Bias/Limited Training Set
- **WeChat** Racist Bias/Limited Training Set Unexpected Human Behaviour
- **Microsoft - Tay** Unexpected Human Behaviour Bias/Limited Training Set

……………………………

(Thanks to Rosie Fenwick and Eliyas Woldegeorgis for collecting the list)

The main problem for the widespread use of AI around the world are unexpected errors

- The mistakes can be dangerous;
- Usually, it remains unclear, who is responsible for them;
- To avoid a recurrence of a detected error, a quick, non-iterative system fix is required;
- Correction of errors should not damage existing skills;
- **The real world is not a good i.i.d. sample**,
  - -the types of errors are very numerous and often unpredictable,
  - -so we cannot rely on a statistical estimate of the probability of errors.

Fundamental origins of AI errors

- Software errors
- Unexpected human behaviour
- Non-intended use
- ….
- **Uncertainty in training data (100 attributes require $> 2^{100} > 10^{30}$ records for completeness)**
- **Uncertainty in the training process**

Errors are unavoidable companions of data driven AI

Fundamental origins of AI errors

- Software errors
- Unexpected human behaviour
- Non-intended use
- ….
- **Uncertainty in training data (100 attributes require $> 2^{100} > 10^{30}$ records for completeness)**
- **Uncertainty in the training process**

<u>Errors are unavoidable companions of data driven AI</u>

Message 1

The main AI 2020 problem:
Errors of AI and their processing

Errors are unavoidable companions of data driven AI

A new AI winter will come if we don't focus on the problem of AI errors

And only those applications will survive where errors are not very dangerous

Can we correct errors by systematic re-training?

- To preserve existing skills we must use the full set of training data.
- This approach requires much recourses for each error.
- However, new errors may appear.
- The preservation of existing skills is not guaranteed.
- The probability of damage is a priori unknown.

Can we correct errors by systematic re-training?

- **To preserve existing skills we must use the full set of training data.**
- This approach requires much recourses for each error.
- However, new errors may appear.
- The preservation of existing skills is not guaranteed.
- The probability of damage is a priori unknown.

Can we correct errors by systematic re-training?

- To preserve existing skills we must use the full set of training data.
- **This approach requires much recourses for each error.**
- However, new errors may appear.
- The preservation of existing skills is not guaranteed.
- The probability of damage is a priori unknown.

Can we correct errors by systematic re-training?

- To preserve existing skills we must use the full set of training data.
- This approach requires much recourses for each error.
- **However, new errors may appear.**
- The preservation of existing skills is not guaranteed.
- The probability of damage is a priori unknown.

Can we correct errors by systematic re-training?

- To preserve existing skills we must use the full set of training data.
- This approach requires much recourses for each error.
- However, new errors may appear.
- **The preservation of existing skills is not guaranteed.**
- The probability of damage is a priori unknown.

Can we correct errors by systematic re-training?

- To preserve existing skills we must use the full set of training data.
- This approach requires much recourses for each error.
- However, new errors may appear.
- The preservation of existing skills is not guaranteed.
- **The probability of skill damage is a priori unknown.**

Can we correct errors by systematic re-training?

- To preserve existing skills we must use the full set of training data.
- This approach requires much recourses for each error.
- However, new errors may appear.
- The preservation of existing skills is not guaranteed.
- **The probability of skill damage is a priori unknown.**

Why?

D. Donoho, from Stanford University webpage

High-dimensional post-classical world

- **The number of attributes *p* >> The number of examples *N***
- This *post-classical* world is different from the '*classical world*'.
- The classical methodology was developed for the 'classical world' based on the assumption of $p < N$, and $N \to \infty$.
- These results all fail if $p > N$.
- The $p > N$ case is not anomalous; it is the generic case.

Donoho, D.L. High-Dimensional Data Analysis: The Curses and Blessings of Dimensionality. Invited Lecture at Mathematical Challenges of the 21st Century, AMS.

D. Donoho, from Stanford University webpage

High-dimensional post-classical world

- **The number of attributes *p* >> The number of examples *N***
- This *post-classical* world is different from the ‘*classical world*’.
- The classical methodology was developed for the ‘classical world’ based on the assumption of $p < N$, and $N \to \infty$.
- These results all fail if $p > N$.
- **The *p* > *N* case is not anomalous; it is the generic case.**

Donoho, D.L. High-Dimensional Data Analysis: The Curses and Blessings of Dimensionality. Invited Lecture at Mathematical Challenges of the 21st Century, AMS.

D. Donoho, from Stanford University webpage

High-dimensional post-classical world

- **The number of attributes *p* >> The number of examples *N***
- This *post-classical* world is different from the '*classical world*'.

Thus, the classical statistical learning theory is mostly useless in the multidimensional post-classical world.

- **The *p* > *N* case is not anomalous; it is the generic case.**

Donoho, D.L. High-Dimensional Data Analysis: The Curses and Blessings of Dimensionality. Invited Lecture at Mathematical Challenges of the 21st Century, AMS.

High-dimensional post-classical world

Comment 1

- All the geometric blessing and curse of dimensionality effects appear already when

$$p \gg \log N,$$

- That is, the “post-classical world” effects begin long before $p > N$.
- The non-classical world is around!

AN Gorban, VA Makarov, IY Tyukin, **High-Dimensional Brain in a High-Dimensional World: Blessing of Dimensionality** *Entropy* 22 (1), 82, https://doi.org/10.3390/e22010082

High-dimensional post-classical world

Comment 2

- The number of attributes differs significantly from the dimensionality of data.
- The definition of the “post-classical world” should be modified to

$$\mathrm{Dim(DataSet)} \gg N,$$

- or, according to Comment 1,

$$\mathbf{Dim(DataSet)} \gg \mathbf{log}\ \boldsymbol{N}.$$

About different definitions of data dimensionality related to blessing of dimensionality:
Bac, J., Zinovyev, A. (2020). **Lizard brain: tackling locally low-dimensional yet globally complex organization of multi-dimensional datasets**. *Frontiers in Neurorobotics*, *13*, 110. https://doi.org/10.3389/fnbot.2019.00110

How to cope with the non-classical world?

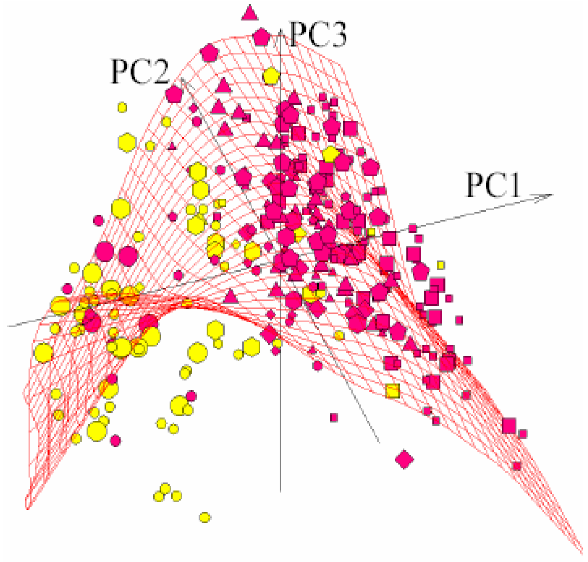
How to cope with the non-classical world?

Return to classics?

How to cope with the non-classical world?

Return to classics:

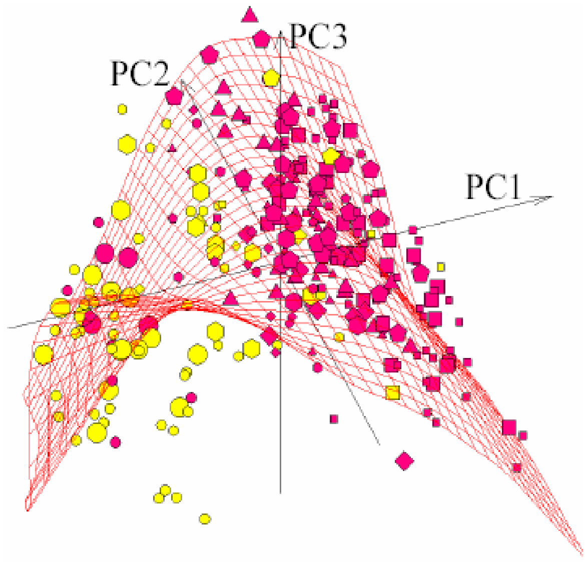
- Reduce dimensionality
  (PCA, ICA, Principal graphs and manifolds, Autoencoders,… - methods of the previous century with some modern improvements);



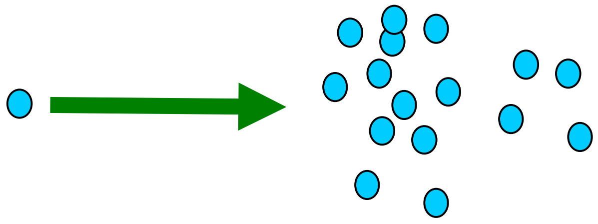
How to cope with the non-classical world?

Return to classics:

- Reduce dimensionality
  (PCA, ICA, Principal graphs and manifolds, Autoencoders,… - methods of the previous century with some modern improvements);



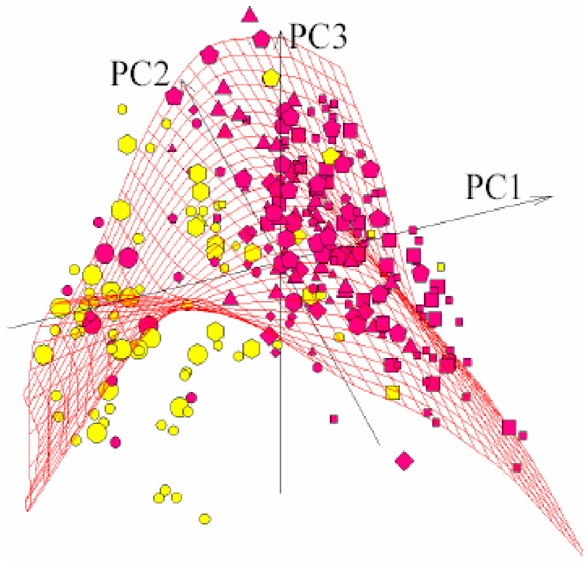
- Data augmentation – smearing data by random perturbations
  (also well-known from 1990s).



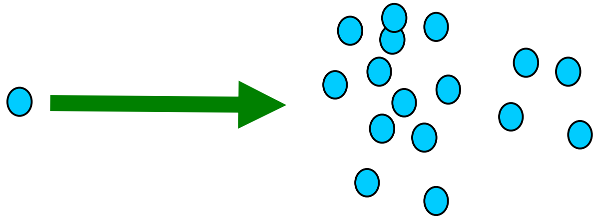
How to cope with the non-classical world?

Return to classics:

- Reduce dimensionality
  (PCA, ICA, Principal graphs and manifolds, Autoencoders,… - methods of the previous century with some modern improvements);



- Data augmentation – smearing data by random perturbations
  (also well-known from 1990s).



Sometimes this works: A. Gorban, B. Kégl, D. Wunsch, A. Zinovyev (Eds.), ***Principal manifolds for data visualization and dimension reduction.*** Springer, (2008). Gorban, A. N., Zinovyev, A. **Principal manifolds and graphs in practice: from molecular biology to dynamical systems**. *Int. J. Neural Syst.*, *20* (2010), 219-232.

How to cope with the non-classical world?

when return to classics is impossible?

How to cope with the non-classical world?

when return to classics is impossible?

An *external device* - Corrector of AIs' mistakes is needed

A.N. Gorban, I.Y. Tyukin, Blessing of dimensionality: mathematical foundations of the statistical physics of data. *Philos. Trans. Royal Soc.* A 376(2118), 20170237, 2018.

How to cope with the non-classical world?

when return to classics is impossible?

An *external device* - Corrector of AIs' mistakes is needed

Technical requirements

Corrector should:

A.N. Gorban, I.Y. Tyukin, [Blessing of dimensionality: mathematical foundations of the statistical physics of data](). *Philos. Trans. Royal Soc.* A 376(2118), 20170237, 2018.

How to cope with the non-classical world?

when return to classics is impossible?

An *external device* - Corrector of AIs' mistakes is needed

Technical requirements
Corrector should:

1. **be simple;**

A.N. Gorban, I.Y. Tyukin, [Blessing of dimensionality: mathematical foundations of the statistical physics of data](). *Philos. Trans. Royal Soc.* A 376(2118), 20170237, 2018.

How to cope with the non-classical world?

when return to classics is impossible?

An *external device* - Corrector of AIs' mistakes is needed

Technical requirements

Corrector should:

1. be simple;
2. **not change the skills of the legacy system;**

A.N. Gorban, I.Y. Tyukin, [Blessing of dimensionality: mathematical foundations of the statistical physics of data](). *Philos. Trans. Royal Soc.* A 376(2118), 20170237, 2018.

How to cope with the non-classical world?

when return to classics is impossible?

An *external device* - Corrector of AIs' mistakes is needed

Technical requirements

Corrector should:

1. be simple;
2. not change the skills of the legacy system;
3. **allow fast non-iterative learning;**

A.N. Gorban, I.Y. Tyukin, [Blessing of dimensionality: mathematical foundations of the statistical physics of data](). *Philos. Trans. Royal Soc.* A 376(2118), 20170237, 2018.

How to cope with the non-classical world?

when return to classics is impossible?

An *external device* - Corrector of AIs' mistakes is needed

Technical requirements

Corrector should:

1. be simple;
2. not change the skills of the legacy system;
3. allow fast non-iterative learning;
4. **allow correction of new mistakes without destroying previous corrections.**

A.N. Gorban, I.Y. Tyukin, Blessing of dimensionality: mathematical foundations of the statistical physics of data. *Philos. Trans. Royal Soc.* A 376(2118), 20170237, 2018.

How to cope with the non-classical world?

when return to classics is impossible?

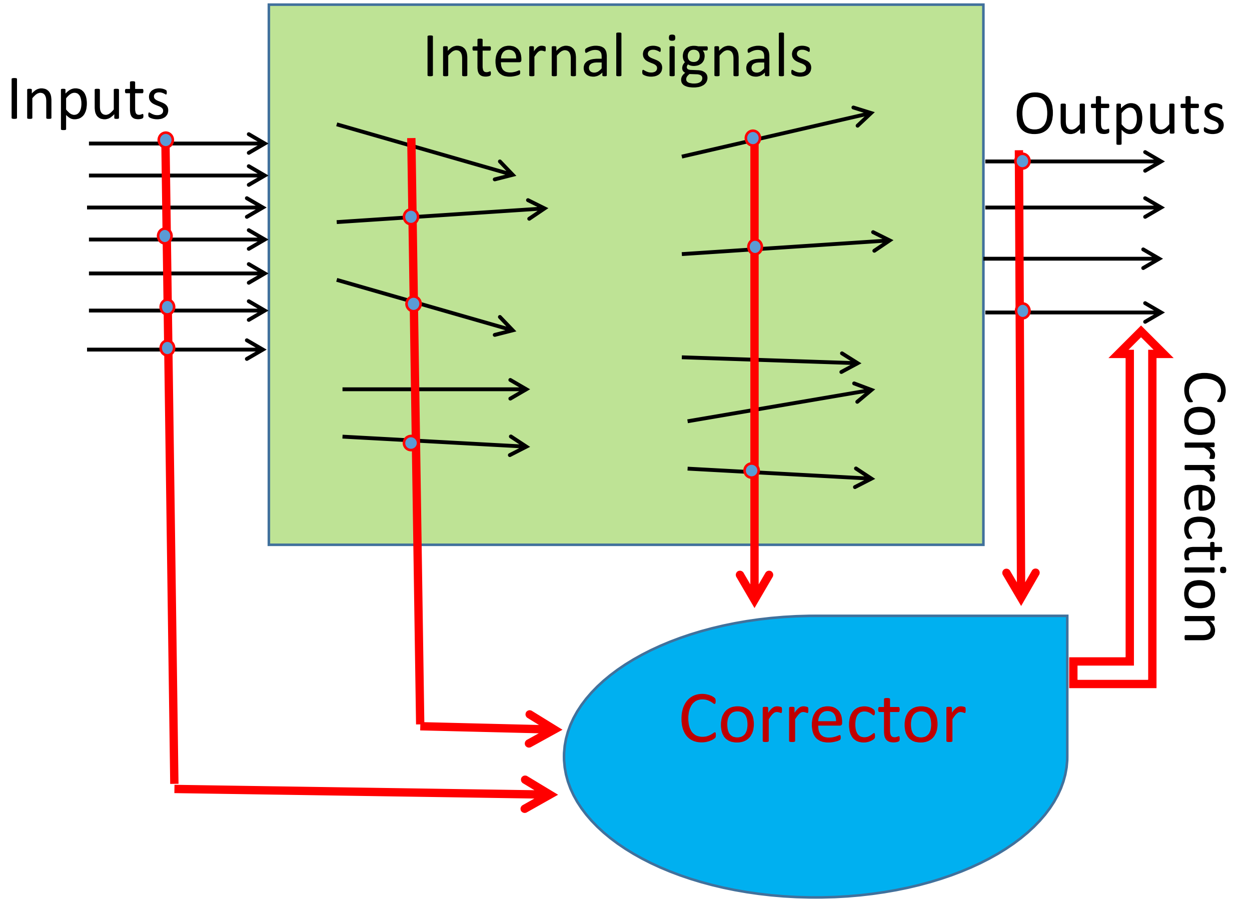
An *external device* - Corrector of AIs' mistakes is needed

Technical requirements
Corrector should:

1. be simple;
2. not change the skills of the legacy system;
3. allow fast non-iterative learning;
4. **allow correction of new mistakes without destroying previous corrections.**

A.N. Gorban, I.Y. Tyukin, Blessing of dimensionality: mathematical foundations of the statistical physics of data. *Philos. Trans. Royal Soc.* A 376(2118), 20170237, 2018.

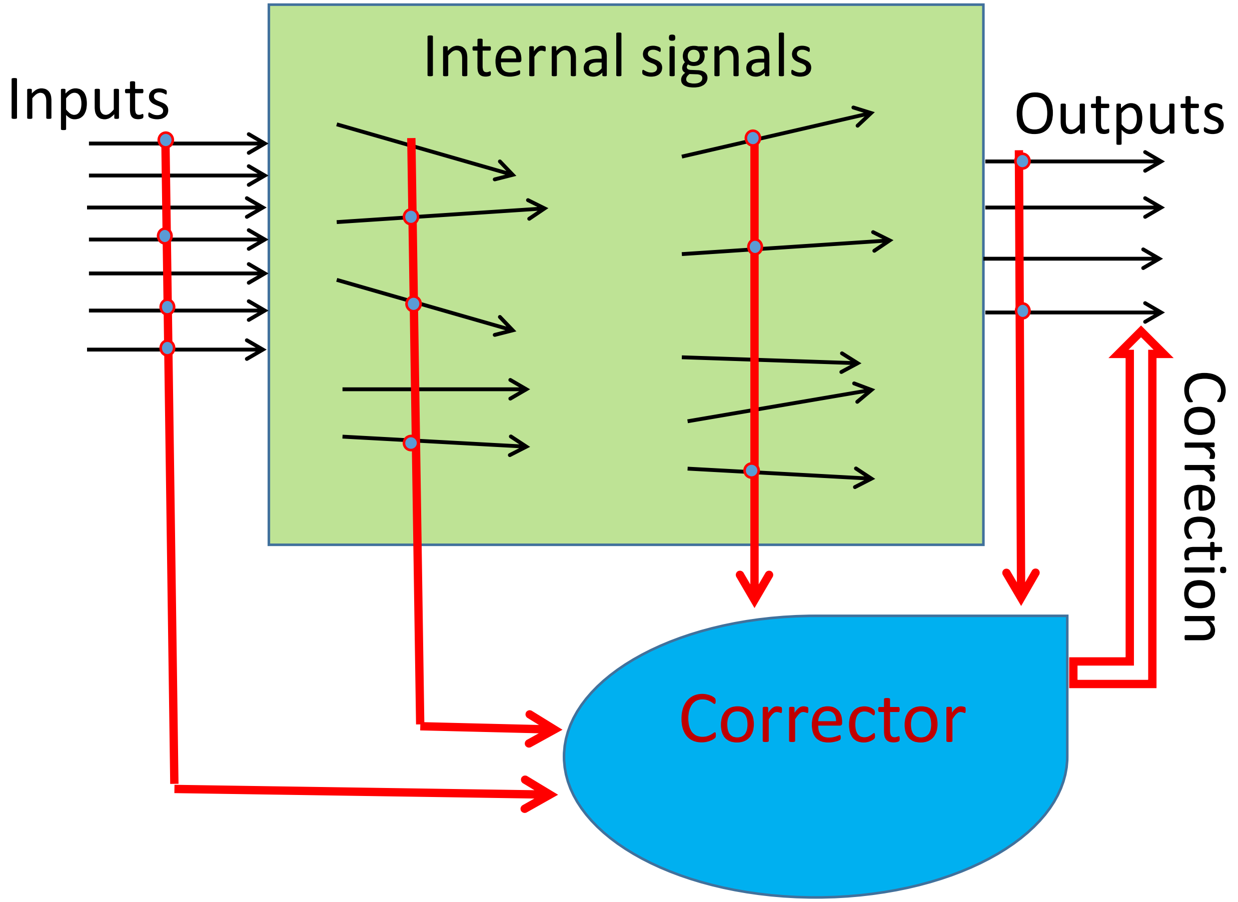
Corrector has to separate mistakes from correctly solved examples and correct them



Corrector is a **binary classifier** *for error diagnosis* equipped by the **modified decision rule** *for high risk situations*.

Gorban, A. N., Golubkov, A., Grechuk, B., Mirkes, E. M., Tyukin, I. Y. [Correction of AI systems by linear discriminants: Probabilistic foundations](). *Information Sciences*, *466* (2018), 303-322.

Corrector has to separate mistakes from correctly solved examples and correct them



Corrector is a **binary classifier** *for error diagnosis* equipped by the **modified decision rule** *for high risk situations*.

Fisher's linear discriminants separate errors from other examples surprisingly well in high dimensions.

Gorban, A. N., Golubkov, A., Grechuk, B., Mirkes, E. M., Tyukin, I. Y. **Correction of AI systems by linear discriminants: Probabilistic foundations**. *Information Sciences*, *466* (2018), 303-322.

Fisher discriminants are explicit and non-iterative classifiers

Let a data set *Y be centralized* (zero mean) and *whitened* (unit covariance matrix)

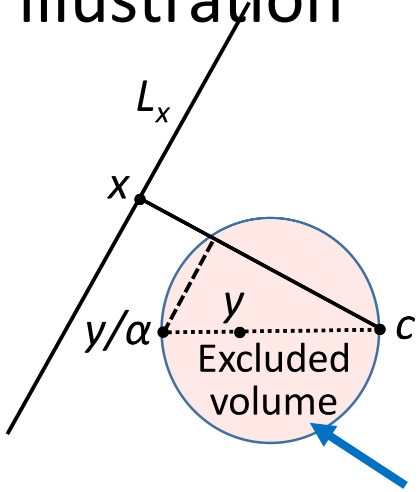
Definition. *A point x is Fisher separable from a finite set Y with a threshold α (0 ≤ α < 1) if*

$$(x,y) \le \alpha (x,x) \qquad (1)$$

for all y from Y.

Where (x,y) is the standard inner product.

(1) illustration



x does not belong to a ball

$$\left\{ z \;\middle|\; \left\| z - \frac{y}{2\alpha} \right\| < \frac{\|y\|}{2\alpha} \right\}$$

Fisher discriminants are explicit and non-iterative classifiers

Let a data set *Y be centralized* (zero mean) and *whitened* (unit covariance matrix)

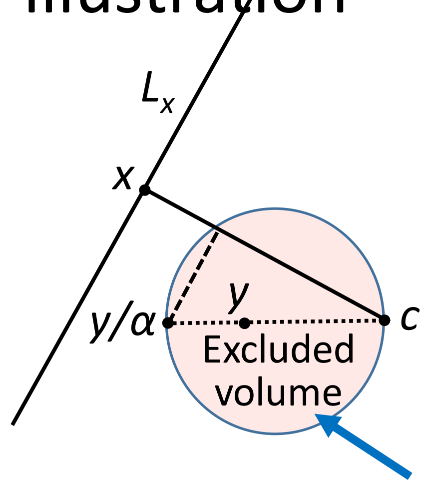
Definition. *A point x is Fisher separable from a finite set Y with a threshold α (0 ≤ α < 1) if*

$$(x,y) \le \alpha (x,x) \qquad (1)$$

for all y from Y.

Where (x,y) is the standard inner product.

(1) illustration



x does not belong to a ball

$$\left\{ z \,\middle|\, \left\| z - \frac{\mathbf{y}}{2\alpha} \right\| < \frac{\|\mathbf{y}\|}{2\alpha} \right\}$$

These simple discriminants effectively separate mistakes from correctly solved examples in high dimensions.

Gorban, A. N., Burton, R., Romanenko, I., & Tyukin, I. Y. (2019). One-trial correction of legacy AI systems and stochastic separation theorems. *Information Sciences*, *484*, 237-254.

A miracle of stochastic separation

(an old joke about a new topic)

Let us be a bit more explicit!

Gorban, A. N., Tyukin, I. Y. (2017). **Stochastic separation theorems**. *Neural Networks*, *94*, 255-259.

Assume the absence of: (i) large deviations and (ii) sets with small volume but high probability

Theorem 1. *Assume that a probability distribution in the n-dimensional unit ball $\mathbb{B}_n$ (with volume $V_n(\mathbb{B}_n)$) has a density with maximal value ρ_m, which satisfies inequality*

$$\rho_m < \frac{C}{r^n V_n(\mathbb{B}_n)} \quad 0 < r \leq 1.$$

Let $|Y| < b^n$ and $2r\alpha > b > 1$. Then the probability p that a random point is Fisher-separable from the finite set Y is $p = 1 - \psi$, where

$$\psi < C\left(\frac{b}{2r\alpha}\right)^n.$$

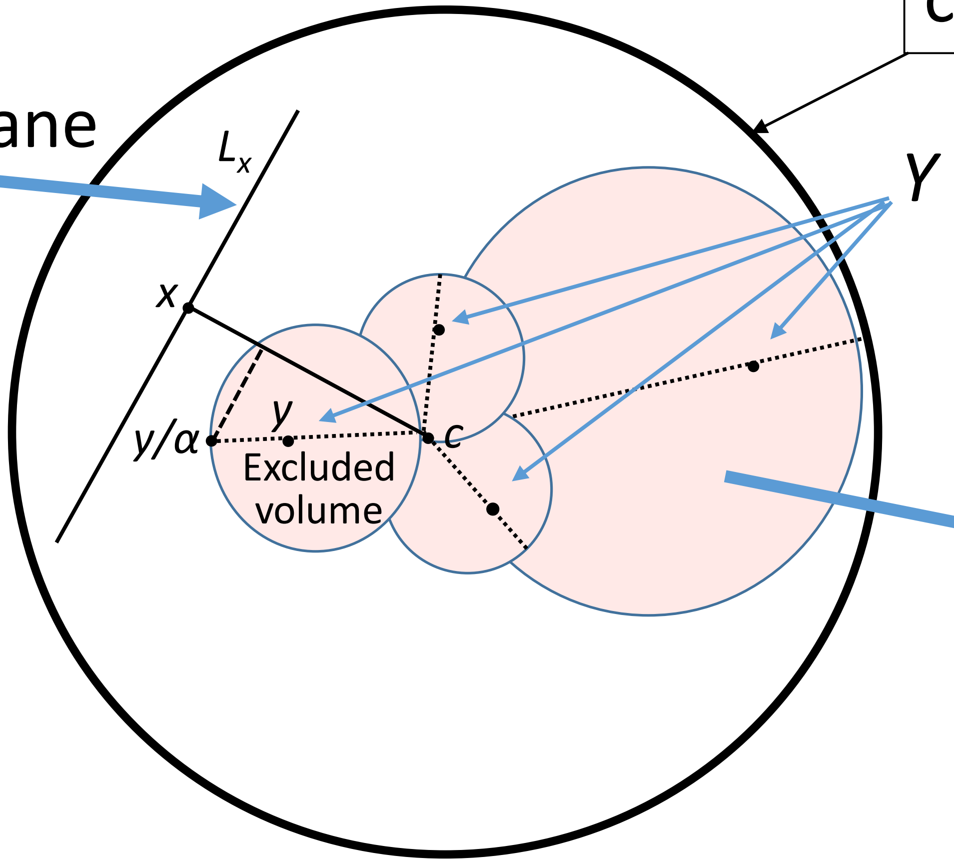
Note: $r \leq 1$, $\alpha < 1$
but $r\alpha > 0.5$

Gorban, A. N., Golubkov, A., Grechuk, B., Mirkes, E. M., Tyukin, I. Y. Correction of AI systems by linear discriminants: Probabilistic foundations. *Information Sciences*, *466* (2018), 303-322.

For proof: just evaluate excluded volumes and probability of *x* to be there

Unit ball with centre *c*

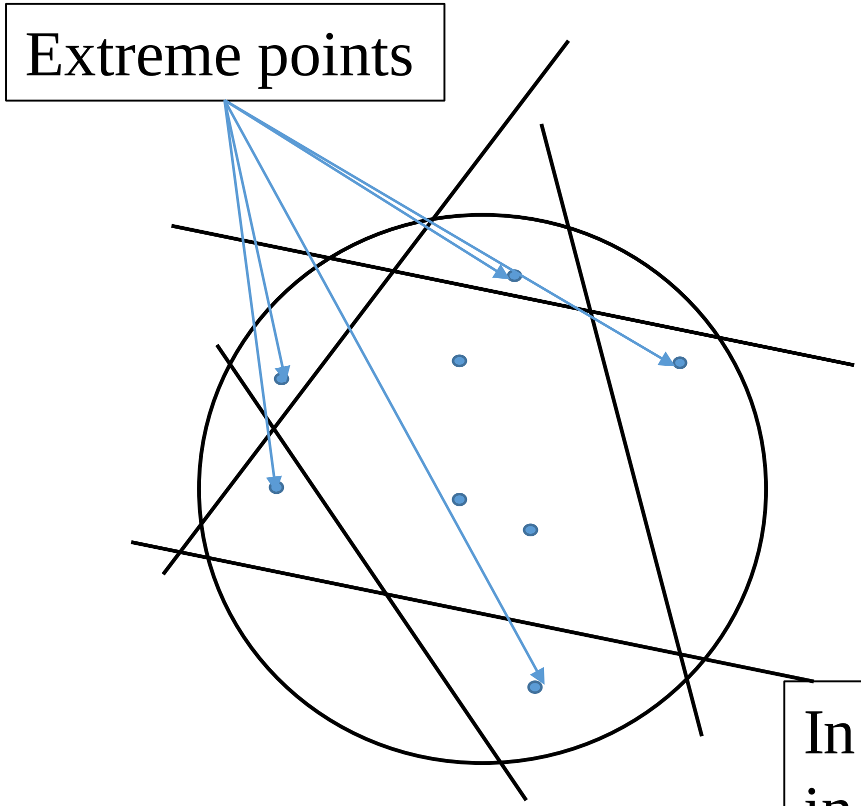
The hyperplane of Fisher's discriminant



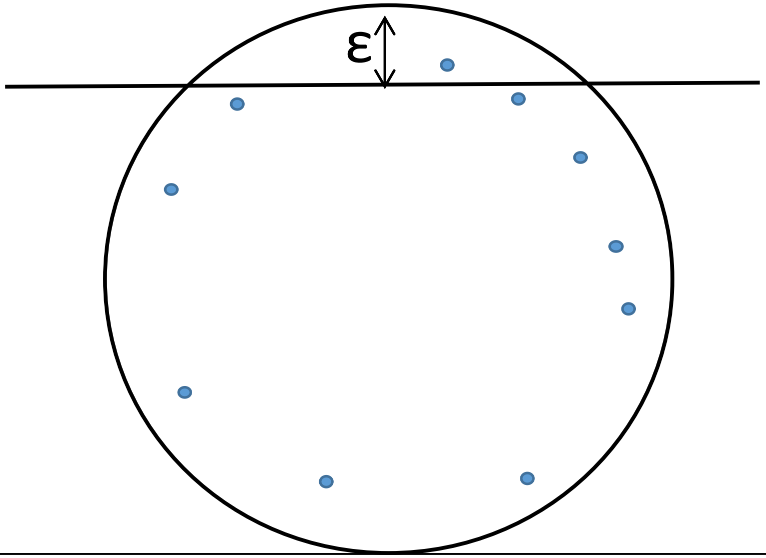
Volume of one "excluded ball" is less than

$$V_n(\mathbb{B}_n)/(2\alpha)^n,\ \ 2\alpha>1$$

Extreme points in high dimensions

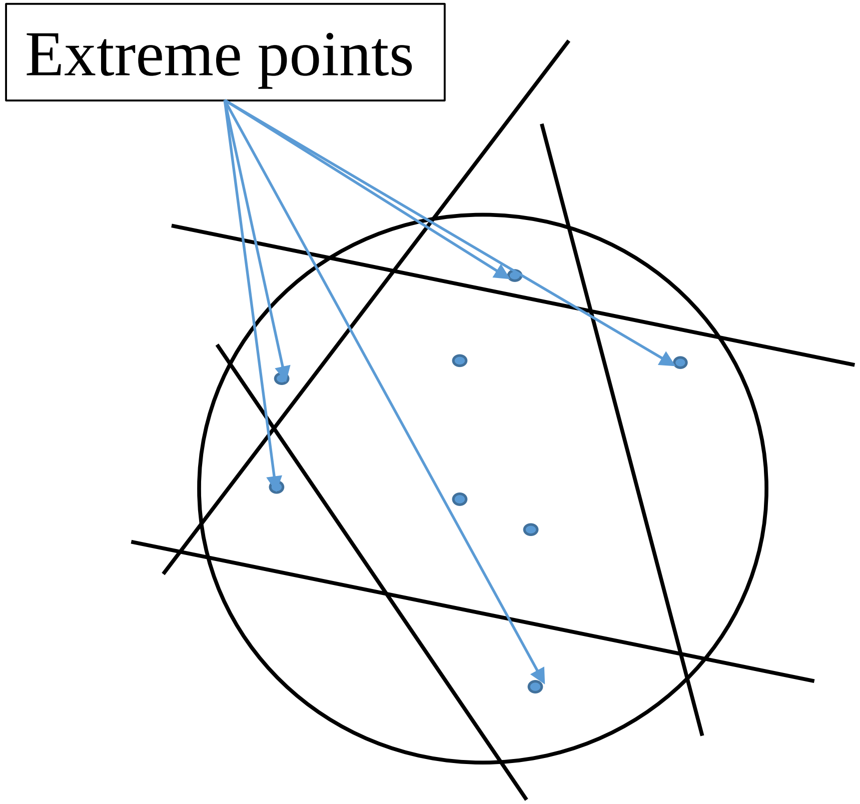


Expected ε IS NOT small
in high dimension!

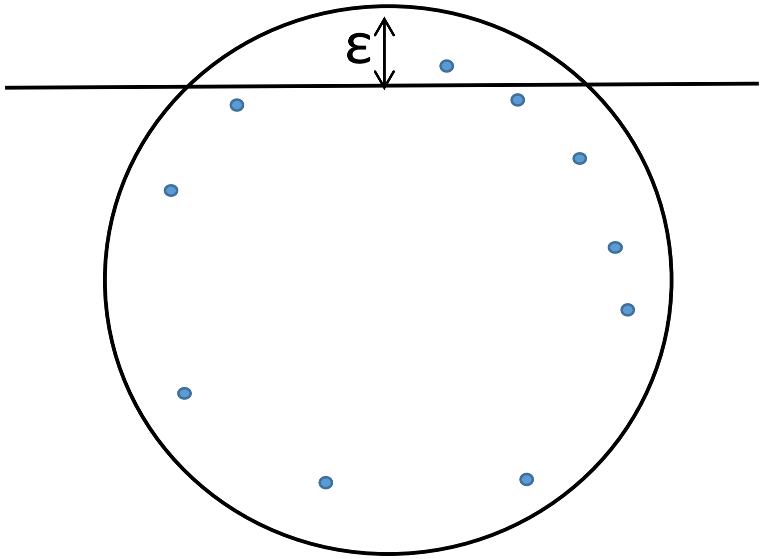


In high dimension, with high probability even in an exponentially large random set all points are extreme ones!

Extreme points in high dimensions



Expected ε IS NOT small
in high dimension!



<u>*Example:*</u>

For ***n*=100** and ***M* < 2,740,000** the set of *M random* points in a ***n***-dimensional ball is Fisher-separable (each point from the rest) with probability ***p*>99%**

Blessing of dimensionality

ELEMENTARY PRINCIPLES

IN

STATISTICAL MECHANICS

DEVELOPED WITH ESPECIAL REFERENCE TO

THE RATIONAL FOUNDATION OF THERMODYNAMICS

BY

J. WILLARD GIBBS

Professor of Mathematical Physics in Yale University

NEW YORK: CHARLES SCRIBNER'S SONS
LONDON: EDWARD ARNOLD
1902

COLLECTION DE MONOGRAPHIES
SUR LA THÉORIE DES FONCTIONS

PROBLÈMES CONCRETS D'ANALYSE FONCTIONNELLE

par

PAUL LÉVY

avec un complément
sur les fonctionnelles analytiques

par

F. PELLEGRINO

Deuxième edition
Paris
1951

 70

Blessing of dimensionality

- The blessing of dimensionality is a manifestation of the ***concentration of measure.***

ELEMENTARY PRINCIPLES

IN

STATISTICAL MECHANICS

DEVELOPED WITH ESPECIAL REFERENCE TO

THE RATIONAL FOUNDATION OF THERMODYNAMICS

BY

J. WILLARD GIBBS

Professor of Mathematical Physics in Yale University

NEW YORK: CHARLES SCRIBNER'S SONS
LONDON: EDWARD ARNOLD
1902

COLLECTION DE MONOGRAPHIES
SUR LA THÉORIE DES FONCTIONS

PROBLÈMES CONCRETS
D'ANALYSE
FONCTIONNELLE

par
PAUL LÉVY

avec un complément
sur les fonctionnelles analytiques

par
F. PELLEGRINO

Deuxième edition
Paris
1951

Blessing of dimensionality

- The blessing of dimensionality is a manifestation of the *concentration of measure.*
- **All statistical physics is based on these phenomena.**

ELEMENTARY PRINCIPLES

IN

STATISTICAL MECHANICS

DEVELOPED WITH ESPECIAL REFERENCE TO

THE RATIONAL FOUNDATION OF THERMODYNAMICS

BY

J. WILLARD GIBBS

Professor of Mathematical Physics in Yale University

NEW YORK: CHARLES SCRIBNER'S SONS
LONDON: EDWARD ARNOLD
1902

COLLECTION DE MONOGRAPHIES SUR LA THÉORIE DES FONCTIONS

PROBLÈMES CONCRETS D'ANALYSE FONCTIONNELLE

par

PAUL LÉVY

avec un complément sur les fonctionnelles analytiques

par

F. PELLEGRINO

Deuxième edition
Paris
1951

Blessing of dimensionality

- The blessing of dimensionality is a manifestation of the *concentration of measure.*
- All statistical physics is based on these phenomena.
- **Stochastic separation theorems represent a new class of them.**

ELEMENTARY PRINCIPLES

IN

STATISTICAL MECHANICS

DEVELOPED WITH ESPECIAL REFERENCE TO

THE RATIONAL FOUNDATION OF THERMODYNAMICS

BY

J. WILLARD GIBBS

Professor of Mathematical Physics in Yale University

NEW YORK: CHARLES SCRIBNER'S SONS
LONDON: EDWARD ARNOLD
1902

COLLECTION DE MONOGRAPHIES SUR LA THÉORIE DES FONCTIONS

PROBLÈMES CONCRETS D'ANALYSE FONCTIONNELLE

par

PAUL LÉVY

avec un complément sur les fonctionnelles analytiques

par

F. PELLEGRINO

Deuxième edition
Paris
1951

A.N. Gorban, I.Y. Tyukin, [Blessing of dimensionality: mathematical foundations of the statistical physics of data](). *Philos. Trans. Royal Soc.* A 376(2118), 20170237, 2018.

Plenty of generalizations for distributions

Gorban, A. N., Golubkov, A., Grechuk, B., Mirkes, E. M., Tyukin, I. Y. **Correction of AI systems by linear discriminants: Probabilistic foundations**. *Information Sciences*, *466* (2018), 303-322.

Grechuk, B., **Practical stochastic separation theorems for product distributions**, *2019 International Joint Conference on Neural Networks (IJCNN)*, Budapest, Hungary, 2019, pp. 1-8.

Plenty of generalizations for distributions

- **Arbitrary log-concave distributions (logarithm of density is concave);**

Gorban, A. N., Golubkov, A., Grechuk, B., Mirkes, E. M., Tyukin, I. Y. **Correction of AI systems by linear discriminants: Probabilistic foundations**. *Information Sciences*, *466* (2018), 303-322.

Grechuk, B., **Practical stochastic separation theorems for product distributions**, *2019 International Joint Conference on Neural Networks (IJCNN)*, Budapest, Hungary, 2019, pp. 1-8.

Plenty of generalizations for distributions

- Arbitrary log-concave distributions (logarithm of density is concave);
- **Convex combinations of a fixed number of log-concave distributions;**

Gorban, A. N., Golubkov, A., Grechuk, B., Mirkes, E. M., Tyukin, I. Y. **Correction of AI systems by linear discriminants: Probabilistic foundations**. *Information Sciences*, *466* (2018), 303-322.

Grechuk, B., **Practical stochastic separation theorems for product distributions**, *2019 International Joint Conference on Neural Networks (IJCNN)*, Budapest, Hungary, 2019, pp. 1-8.

Plenty of generalizations for distributions

- Arbitrary log-concave distributions (logarithm of density is concave);
- Convex combinations of a fixed number of log-concave distributions;
- **Product distributions with bounded support;**

Gorban, A. N., Golubkov, A., Grechuk, B., Mirkes, E. M., Tyukin, I. Y. **Correction of AI systems by linear discriminants: Probabilistic foundations**. *Information Sciences*, *466* (2018), 303-322.

Grechuk, B., **Practical stochastic separation theorems for product distributions**, *2019 International Joint Conference on Neural Networks (IJCNN)*, Budapest, Hungary, 2019, pp. 1-8.

Plenty of generalizations for distributions

- Arbitrary log-concave distributions (logarithm of density is concave);
- Convex combinations of a fixed number of log-concave distributions;
- Product distributions with bounded support;
- **No i.i.d. assumptions.**

Gorban, A. N., Golubkov, A., Grechuk, B., Mirkes, E. M., Tyukin, I. Y. **Correction of AI systems by linear discriminants: Probabilistic foundations**. *Information Sciences*, *466* (2018), 303-322.

Grechuk, B., **Practical stochastic separation theorems for product distributions**, *2019 International Joint Conference on Neural Networks (IJCNN)*, Budapest, Hungary, 2019, pp. 1-8.

Plenty of generalizations for distributions

Heuristic. **The stochastic separation theorems hold if:**

Plenty of generalizations for distributions

Heuristic: The stochastic separation theorems hold if:

- **There are no heavy tails of the probability distribution;**

Plenty of generalizations for distributions

Heuristic: The stochastic separation theorems hold if:

- There are no heavy tails of the probability distribution;
- **The sets of small volume should not have large probability** (what “small” and “large” mean is strictly defined for different contexts).

Plenty of generalizations for distributions

Heuristic: The stochastic separation theorems hold if:

- There are no heavy tails of the probability distribution;
- The sets of small volume should not have large probability
  (what “small” and “large” mean is strictly defined for different contexts).

In these cases, Fisher discriminant is an effective tool for classification and AI correctors in high dimension

Plenty of generalizations for classifiers

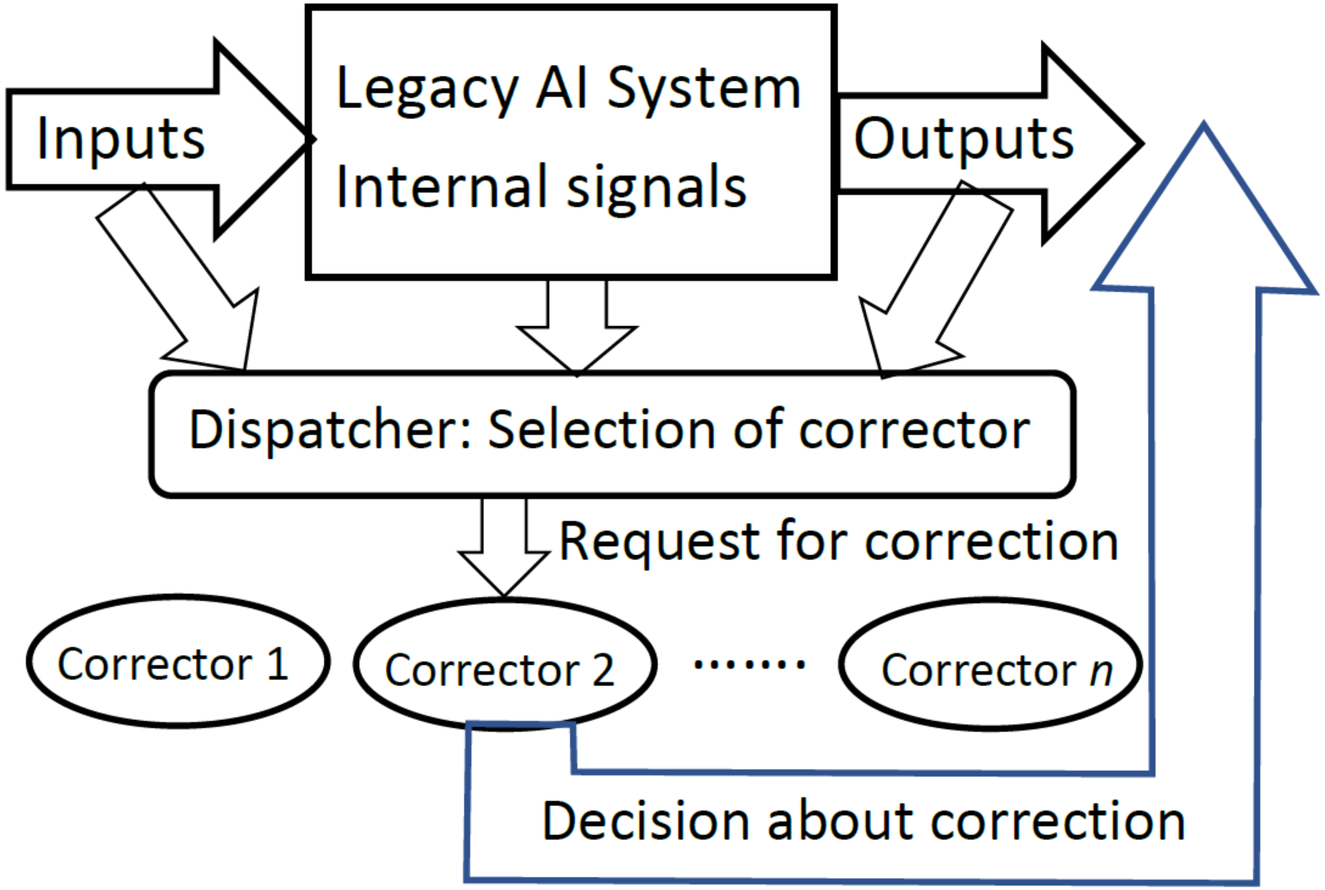
Plenty of generalizations for classifiers

- **Kernel classifiers**

Tyukin, I. Y., Gorban, A. N., Grechuk, B., Green, S. (2019, July). **Kernel Stochastic Separation Theorems and Separability Characterizations of Kernel Classifiers.** In *2019 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-6). IEEE.

Plenty of generalizations for classifiers

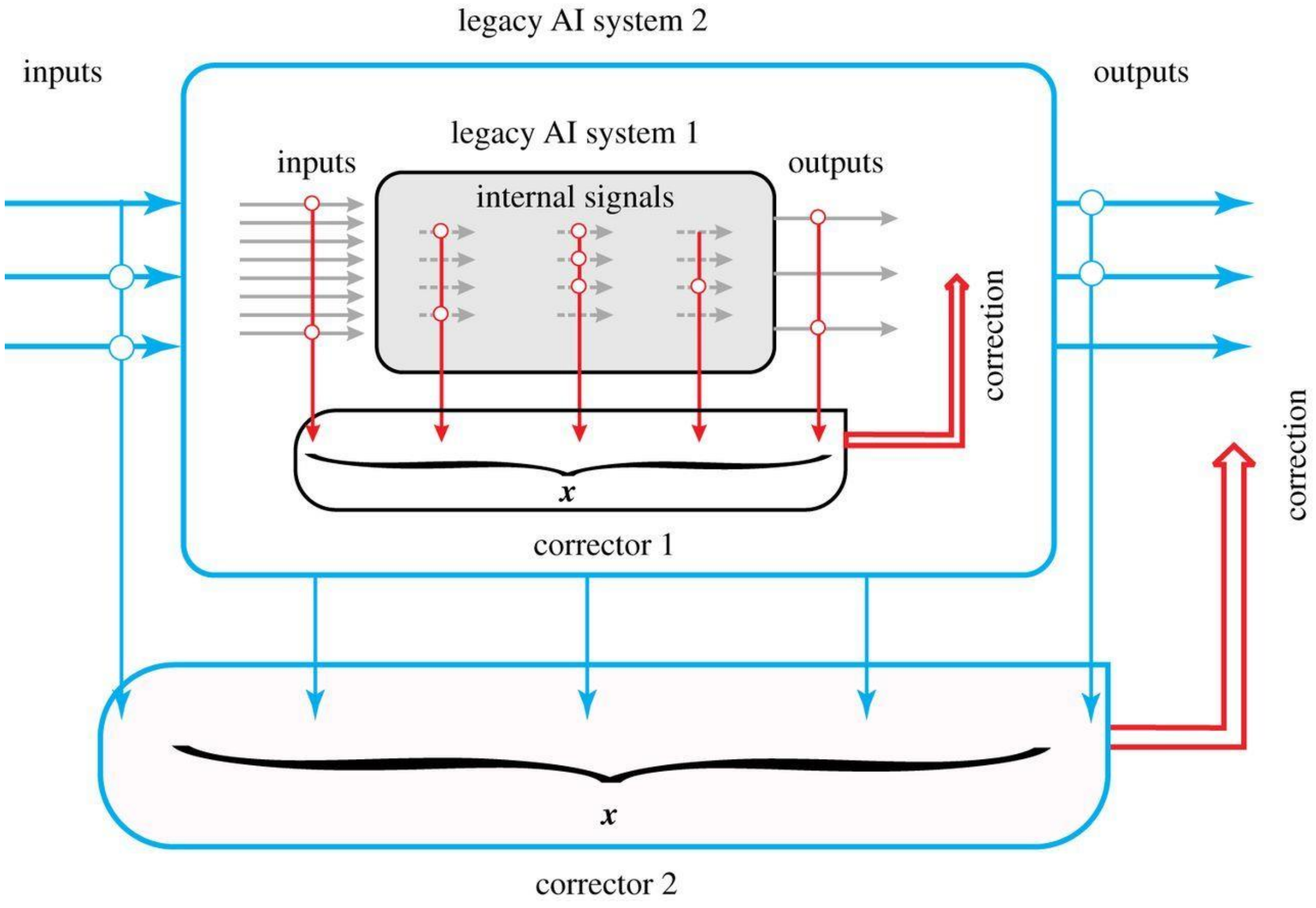
- Kernel classifiers;
- **Ensembles of correctors (with clustering of errors);**



A.N. Gorban, I.Y. Tyukin, [Blessing of dimensionality: mathematical foundations of the statistical physics of data](). *Philos. Trans. Royal Soc.* A 376(2118), 20170237, 2018.

Plenty of generalizations for classifiers

- Kernel classifiers;
- Ensembles of correctors (with clustering of errors);
- **Cascades of correctors.**



A.N. Gorban, I.Y. Tyukin, [Blessing of dimensionality: mathematical foundations of the statistical physics of data](). *Philos. Trans. Royal Soc.* A 376(2118), 20170237, 2018.

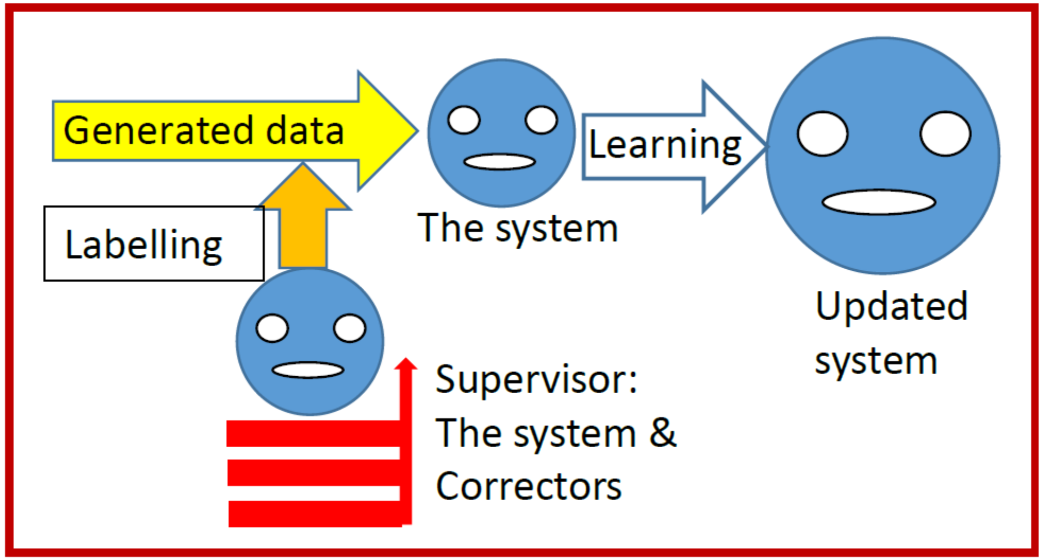
Interiorisation of knowledge

- **Complex correctors correct many errors.**

Sculley, D., Holt, G., Golovin, D.... & Dennison, D. (2015). Hidden technical debt in machine learning systems. In *Advances in neural information processing systems* (pp. 2503-2511).

Gorban, A. N., Grechuk, B., & Tyukin, I. Y. (2018). Augmented artificial intelligence: a conceptual framework. *arXiv:1802.02172*.

Knowledge interiorisation



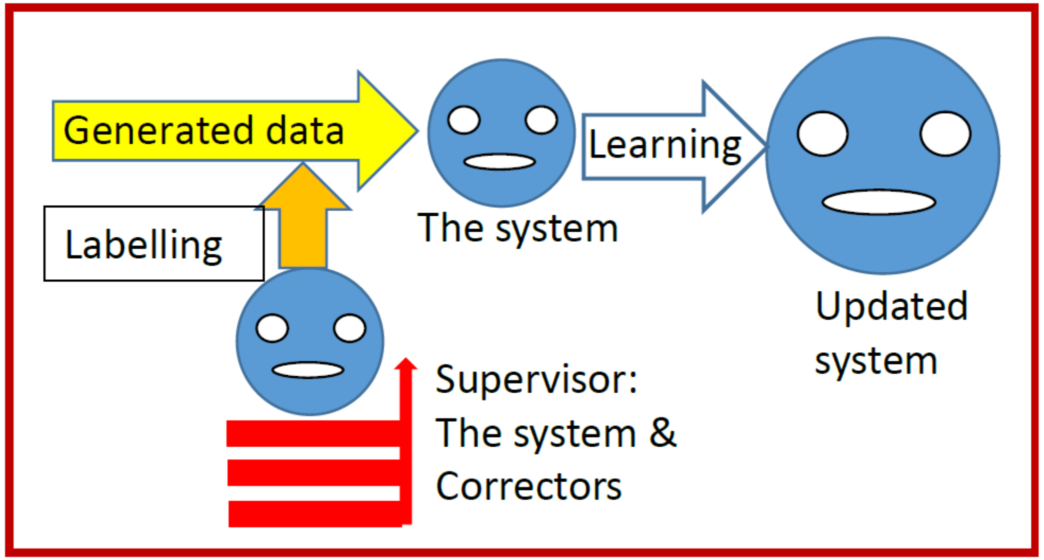
Interiorisation of knowledge

- Complex correctors can correct many errors.
- **But "technical debt" increases, and flexibility drops down.**

Sculley, D., Holt, G., Golovin, D.... & Dennison, D. (2015). Hidden technical debt in machine learning systems. In *Advances in neural information processing systems* (pp. 2503-2511).

Gorban, A. N., Grechuk, B., & Tyukin, I. Y. (2018). Augmented artificial intelligence: a conceptual framework. *arXiv:1802.02172*.

Knowledge interiorisation



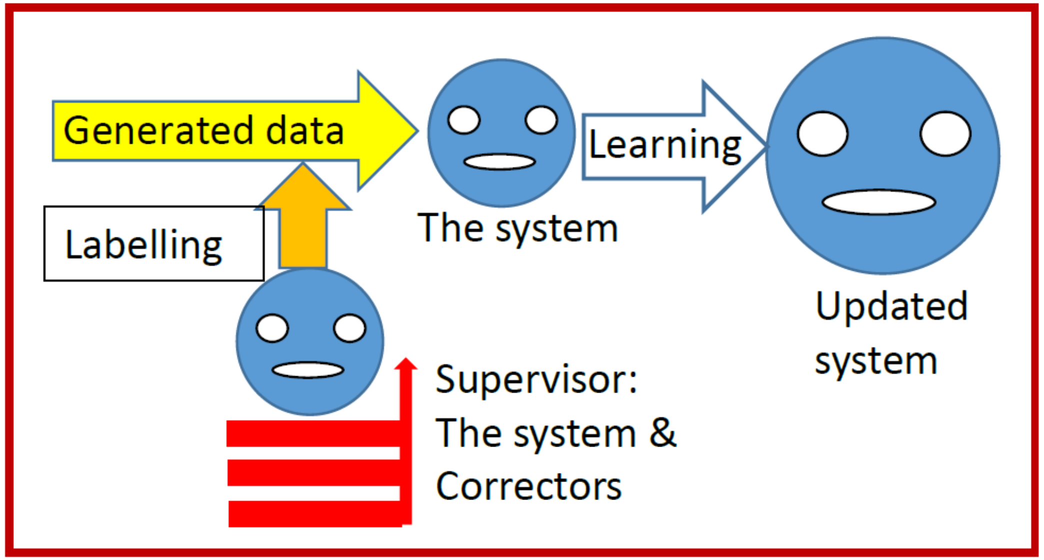
Interiorisation of knowledge

- Complex correctors correct *many errors*.
- But "technical debt" increases, and flexibility drops down.
- **The interiorisation of the collected knowledge is needed, after all.**

Sculley, D., Holt, G., Golovin, D.... & Dennison, D. (2015). Hidden technical debt in machine learning systems. In *Advances in neural information processing systems* (pp. 2503-2511).

Gorban, A. N., Grechuk, B., & Tyukin, I. Y. (2018). Augmented artificial intelligence: a conceptual framework. *arXiv:1802.02172*.

Knowledge interiorisation



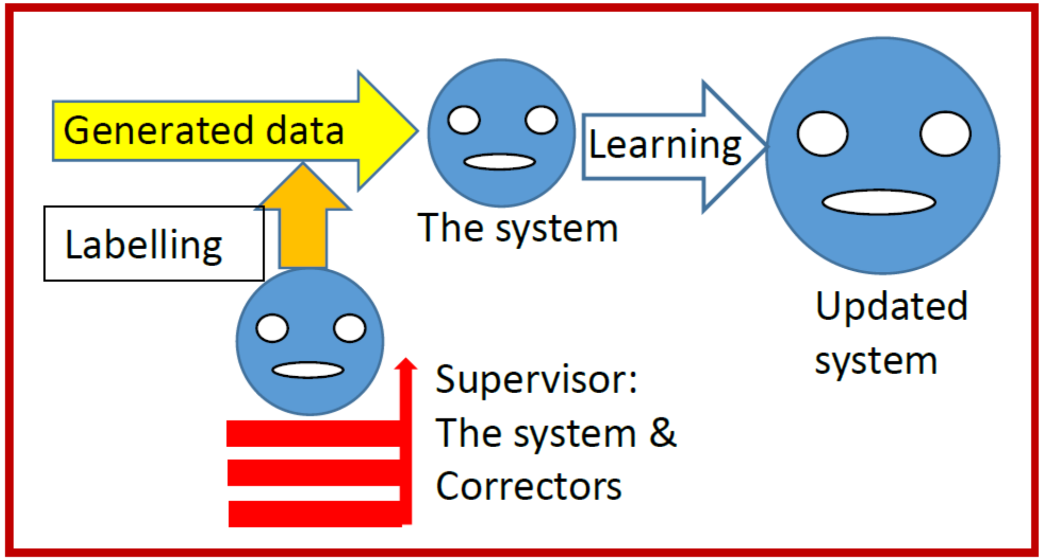
Interiorisation of knowledge

- Complex correctors can correct *many errors*.
- But "technical debt" increases, and flexibility drops down.
- The interiorisation of the collected knowledge is needed, after all.
- **This is incorporation of knowledge into system's inner structure.**

Sculley, D., Holt, G., Golovin, D.... & Dennison, D. (2015). Hidden technical debt in machine learning systems. In *Advances in neural information processing systems* (pp. 2503-2511).

Gorban, A. N., Grechuk, B., & Tyukin, I. Y. (2018). Augmented artificial intelligence: a conceptual framework. *arXiv:1802.02172*.

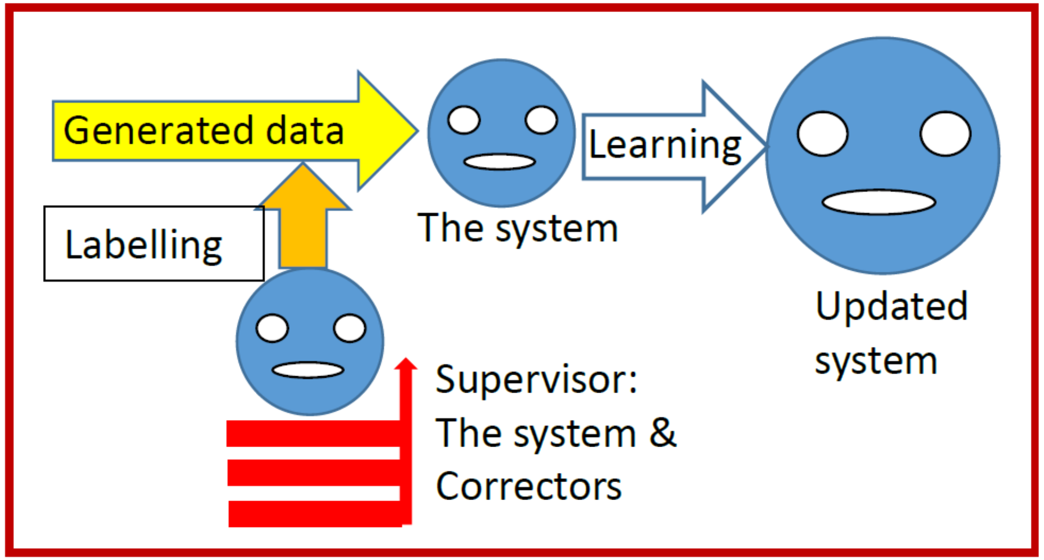
Knowledge interiorisation



Interiorisation of knowledge

- Complex correctors can correct *many errors*.
- But "technical debt" increases, and flexibility drops down.
- The interiorisation of the collected knowledge is needed *after all.*
- This is incorporation of knowledge into system's inner structure.
- **Interiorisation can be organised as supervised learning that uses the system with correctors as the supervisor.**

Knowledge interiorisation



Sculley, D., Holt, G., Golovin, D.... & Dennison, D. (2015). Hidden technical debt in machine learning systems. In *Advances in neural information processing systems* (pp. 2503-2511).

Gorban, A. N., Grechuk, B., & Tyukin, I. Y. (2018). Augmented artificial intelligence: a conceptual framework. *arXiv:1802.02172*.

Adversarial Examples and Stealth Attacks in Artificial Intelligence Systems

Tyukin, I. Y., Higham, D. J., & Gorban, A. N. (2020). **On Adversarial Examples and Stealth Attacks in Artificial Intelligence Systems**. *arXiv:2004.04479*.

Adversarial Examples and Stealth Attacks in Artificial Intelligence Systems

- **The dimensionality of the AI's decision-making space is a major contributor to the AI's susceptibility.**

Tyukin, I. Y., Higham, D. J., & Gorban, A. N. (2020). **On Adversarial Examples and Stealth Attacks in Artificial Intelligence Systems**. *arXiv:2004.04479*.

Adversarial Examples and Stealth Attacks in Artificial Intelligence Systems

- The dimensionality of the AI's decision-making space is a major contributor to the AI's susceptibility.
- **A second crucial parameter is the absence of local concentrations in the data probability distribution.**

Tyukin, I. Y., Higham, D. J., & Gorban, A. N. (2020). **On Adversarial Examples and Stealth Attacks in Artificial Intelligence Systems**. *arXiv:2004.04479*.

Adversarial Examples and Stealth Attacks in Artificial Intelligence Systems

- The dimensionality of the AI’s decision-making space is a major contributor to the AI’s susceptibility.
- A second crucial parameter is the absence of local concentrations in the data probability distribution.
- **Robustness to adversarial examples requires either**
  (a) the data distributions in the AI’s feature space to have concentrated probability density functions or
  (b) the dimensionality of the AI’s decision variables to be sufficiently small.

Tyukin, I. Y., Higham, D. J., & Gorban, A. N. (2020). [On Adversarial Examples and Stealth Attacks in Artificial Intelligence Systems](). *arXiv:2004.04479*.

Adversarial Examples and Stealth Attacks in Artificial Intelligence Systems

- The dimensionality of the AI's decision-making space is a major contributor to the AI's susceptibility.
- A second crucial parameter is the absence of local concentrations in the data probability distribution.
- Robustness to adversarial examples requires either
  (a) the data distributions in the AI's feature space to have concentrated probability density functions or
  (b) the dimensionality of the AI's decision variables to be sufficiently small.
- **The stealth attacks on high-dimensional AI systems are discovered that are hard to spot unless the validation set is made *exponentially large*.**

Tyukin, I. Y., Higham, D. J., & Gorban, A. N. (2020). **On Adversarial Examples and Stealth Attacks in Artificial Intelligence Systems**. *arXiv:2004.04479*.

Message 2: Non-classical alternative

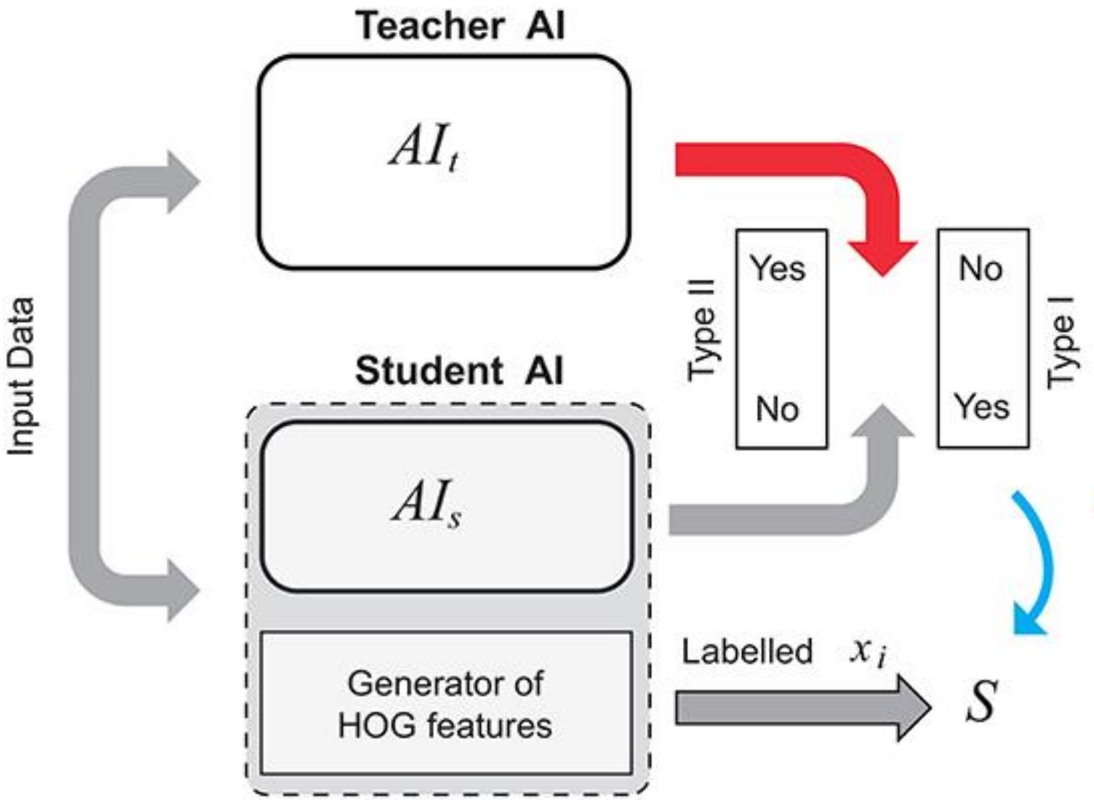
In the non-classical world we have a choice:
either return to classics,
or exploit the blessing of dimensionality
and create one-shot correctors of errors.

When AI can play with AI then the progress is much faster

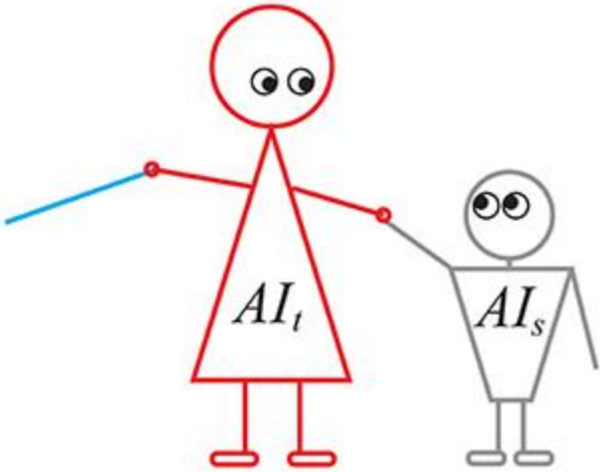
- Correctors are tools for one-shot transferring skills between AIs.

When AI can play with AI then the progress is much faster

- Correctors are tools for one-shot transferring skills between AIs.
- Experiment: training of pedestrian detection



Tyukin, I. Y., Gorban, A. N., Sofeykov, K. I., Romanenko, I. (2018). [Knowledge transfer between artificial intelligence systems](). *Frontiers in neurorobotics*, *12*, 49.



Multiagent neuro-AI systems

Two papers published simultaneously in 1991:

- SE Gilev, AN Gorban, EM Mirkes, **Small experts and internal conflicts in learning neural networks**. *Akademiia Nauk SSSR, Doklady* 320 (1), 220-223, 1991.
- RA Jacobs, MI Jordan, SJ Nowlan, GE Hinton, **Adaptive mixtures of local experts**, *Neural computation* 3 (1), 79-87, 1991.

Multiagent neuro-AI systems

- **A *multiagent system* consists of a number of agents, which *interact*.**

Two papers published simultaneously in 1991:

- SE Gilev, AN Gorban, EM Mirkes, **Small experts and internal conflicts in learning neural networks**. *Akademiia Nauk SSSR, Doklady* 320 (1), 220-223, 1991.
- RA Jacobs, MI Jordan, SJ Nowlan, GE Hinton, **Adaptive mixtures of local experts**, *Neural computation* 3 (1), 79-87, 1991.

Multiagent neuro-AI systems

- A *multiagent system* consists of a number of agents, which *interact*.
- **Agents in a multiagent system can have *different goals* and motivations.**

Two papers published simultaneously in 1991:

- SE Gilev, AN Gorban, EM Mirkes, **Small experts and internal conflicts in learning neural networks**. *Akademiia Nauk SSSR, Doklady* 320 (1), 220-223, 1991.
- RA Jacobs, MI Jordan, SJ Nowlan, GE Hinton, **Adaptive mixtures of local experts**, *Neural computation* 3 (1), 79-87, 1991.

Multiagent neuro-AI systems

- A *multiagent system* consists of a number of agents, which *interact*.
- Agents in a multiagent system can have *different goals* and motivations.
- **For each problem, the agent can evaluate his level of *confidence.***

Two papers published simultaneously in 1991:

- SE Gilev, AN Gorban, EM Mirkes, **Small experts and internal conflicts in learning neural networks**. *Akademiia Nauk SSSR, Doklady* 320 (1), 220-223, 1991.
- RA Jacobs, MI Jordan, SJ Nowlan, GE Hinton, **Adaptive mixtures of local experts**, *Neural computation* 3 (1), 79-87, 1991.

Multiagent neuro-AI systems

- A *multiagent system* consists of a number of agents, which *interact*.
- Agents in a multiagent system can have *different goals* and motivations.
- For each problem, the agent can evaluate his level of *confidence.*
- **They rather learn to be *not too confident in their mistakes* than not to make mistakes at all.**

Two papers published simultaneously in 1991:

- SE Gilev, AN Gorban, EM Mirkes, **Small experts and internal conflicts in learning neural networks**. *Akademiia Nauk SSSR, Doklady* 320 (1), 220-223, 1991.
- RA Jacobs, MI Jordan, SJ Nowlan, GE Hinton, **Adaptive mixtures of local experts**, *Neural computation* 3 (1), 79-87, 1991.

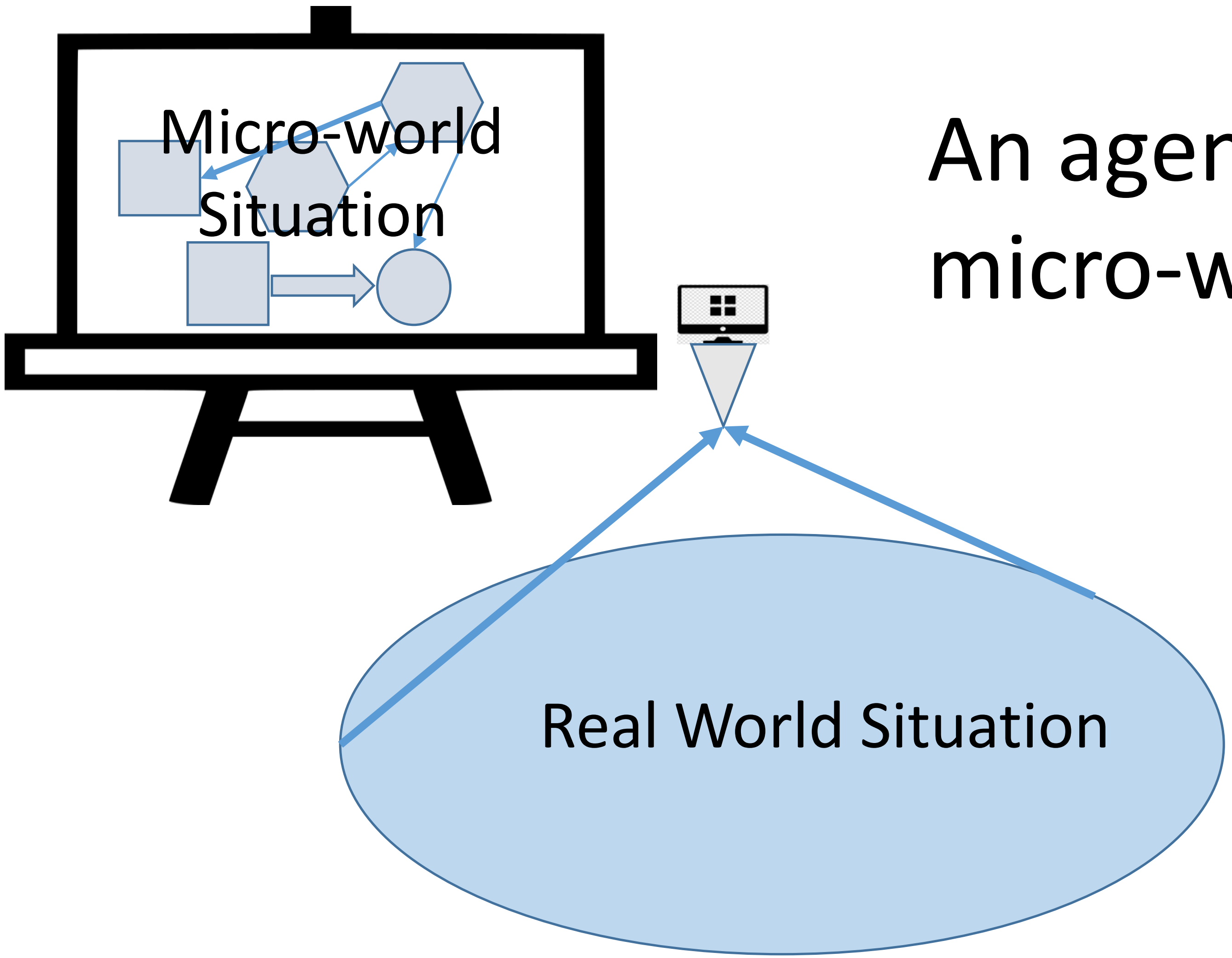
Multiagent neuro-AI systems

- A *multiagent system* consists of a number of agents, which *interact*.
- Agents in a multiagent system can have *different goals* and motivations.
- For each problem, the agent can evaluate his level of *confidence.*
- They rather learn to be *not too confident in their mistakes* than not to make mistakes at all.
- **To successfully interact, agents should *cooperate, coordinate, conflict, negotiate, and correct* each other.**

Two papers published simultaneously in 1991:

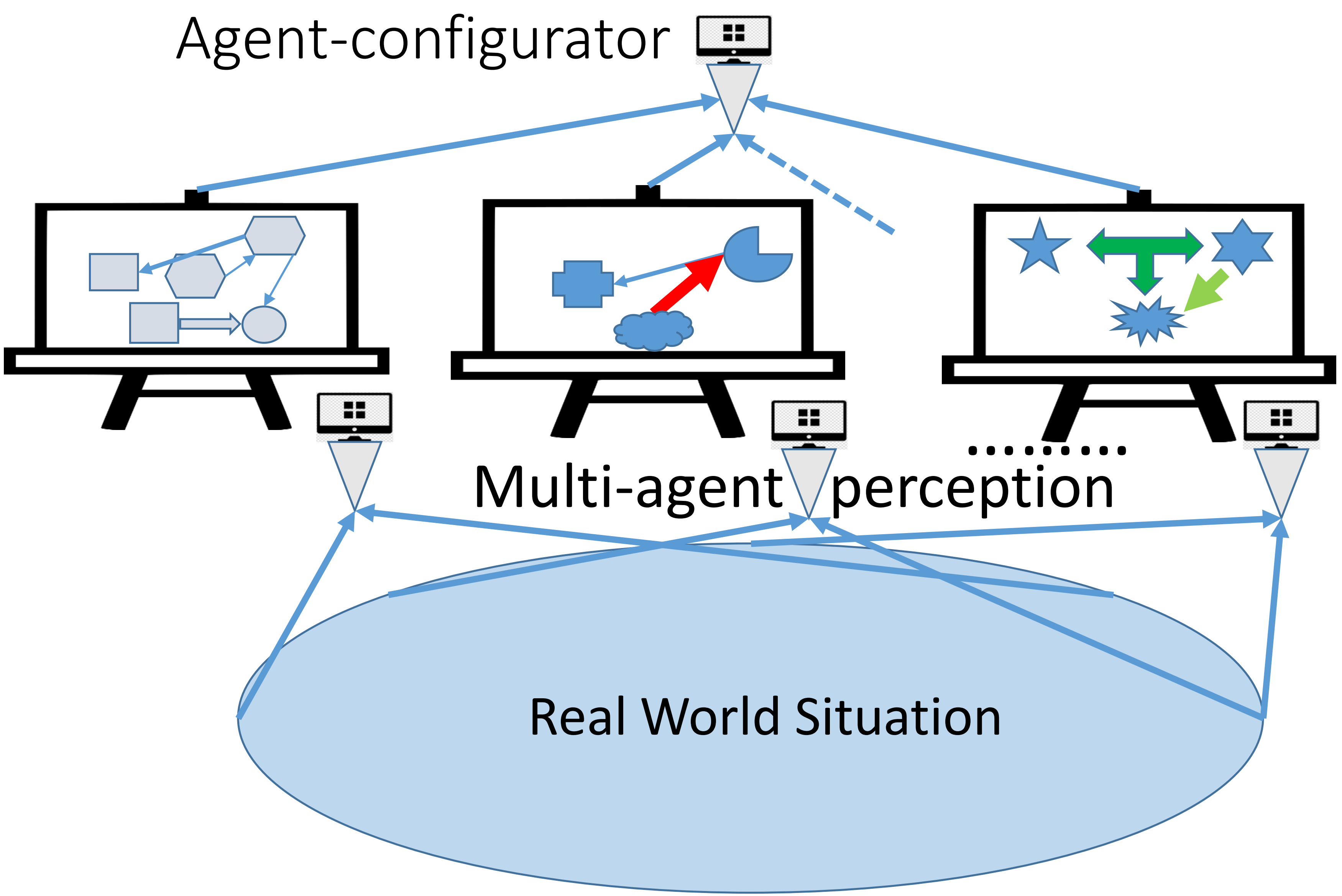
- SE Gilev, AN Gorban, EM Mirkes, **Small experts and internal conflicts in learning neural networks**. *Akademiia Nauk SSSR, Doklady* 320 (1), 220-223, 1991.
- RA Jacobs, MI Jordan, SJ Nowlan, GE Hinton, **Adaptive mixtures of local experts**, *Neural computation* 3 (1), 79-87, 1991.

A personal micro-world is a working place for AI (M. Minsky)



An agent and his micro-world

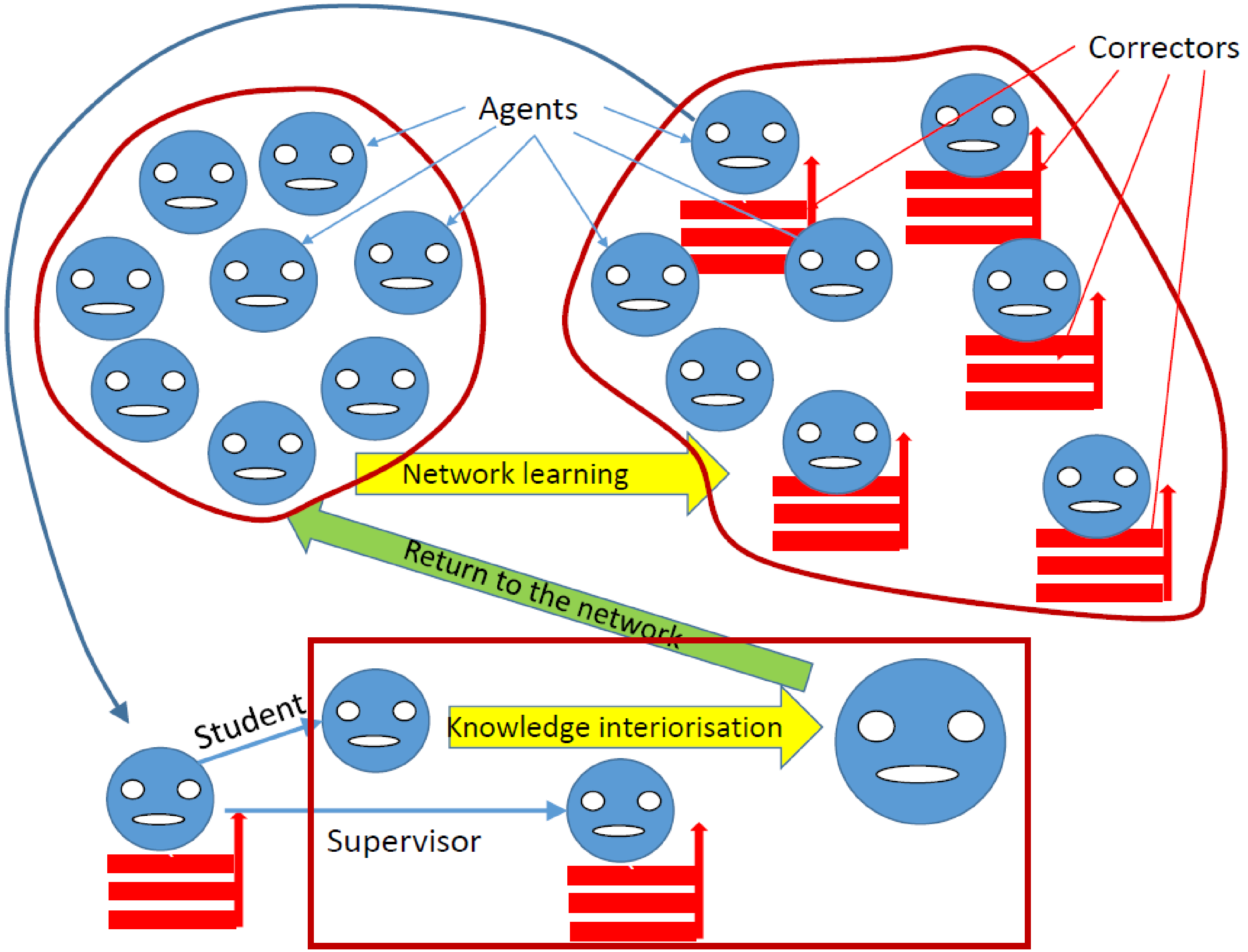
Multilayer multi-agent structures



Neuro-AIs in community

Stages of community learning:

1. Pre-learning;
2. Network learning with mutual corrections;
3. Interiorisation of correctors performed for agents independently;
4. Updating of community members.



Message 3: After tomorrow AI will be the large social system of heterogeneous AI networks

To meet the after tomorrow future
we have to move from the individual AI system
to engineering of the large ecosystems and social systems
of heterogeneous AI networks in heterogeneous high-dimensional world.

When explainability of AI is needed?

Decisions of Neural AI are not transparent and, therefore, cannot be explained logically.

When explainability of AI is needed?

Decisions of Neural AI are not transparent and, therefore, cannot be explained logically.

- **We developed software that produced "logically transparent" neural networks by pruning and learning since middle of 1990s.**

Mirkes, E. M. (2020). **Artificial Neural Network Pruning to Extract Knowledge**. *arXiv:2005.06284*.

When explainability of AI is needed?

Decisions of Neural AI are not transparent and, therefore, cannot be explained logically.

- We developed software that produced “logically transparent” neural networks by pruning and learning since middle of 1990s.
- **Experience: Users preferred “one-button” solutions and used this option rarely.**

Mirkes, E. M. (2020). **Artificial Neural Network Pruning to Extract Knowledge**. *arXiv:2005.06284*.

When explainability of AI is needed?

Decisions of Neural AI are not transparent and, therefore, cannot be explained logically.

- We developed software that produced “logically transparent” neural networks by pruning and learning since middle of 1990s.
- Experience: Users preferred “one-button” solutions and used this option rarely.
- **In most cases of use, the logically transparent form was needed to analyse errors.**

Mirkes, E. M. (2020). [Artificial Neural Network Pruning to Extract Knowledge](). *arXiv:2005.06284*.

Explainability of AI is needed when AI makes mistakes

Explainability of AI is needed when AI makes mistakes

- **To decompose the erroneous decision into elementary steps;**

Explainability of AI is needed when AI makes mistakes

- To decompose the erroneous decision into elementary steps;
- **To understand the source of mistake: which step was wrong?**

Explainability of AI is needed when AI makes mistakes

- To decompose the erroneous decision into elementary steps;
- To understand the source of mistake: which step was wrong?
- **To distribute responsibility for this mistake: who is responsible and for what?**

Explainability of AI is needed when AI makes mistakes

- To decompose the erroneous decision into elementary steps;
- To understand the source of mistake: which step was wrong?
- To distribute responsibility for this mistake: who is responsible and for what?
- **To correct the erroneous steps for the future.**

Explainability of AI is needed when AI makes mistakes

- To decompose the erroneous decision into elementary steps;
- To understand the source of mistake: which step was wrong?
- To distribute responsibility for this mistake: who is responsible and for what?
- To correct the erroneous steps for the future.
- **Simplify. Credibility is simplicity.**

Explainability of AI is needed when AI makes mistakes

- To decompose the erroneous decision into elementary steps;
- To understand the source of mistake: which step was wrong?
- To distribute responsibility for this mistake: who is responsible and for what?
- To correct the erroneous steps for the future.
- Simplify. Credibility is simplicity.
- **But… Does truth resists simplicity?**

Four messages

Four messages

1. **Errors are unavoidable companions of data driven AI: The main AI 2020 problem: Errors of AI and their processing**

Four messages

1. Errors are unavoidable companions of data driven AI: The main AI 2020 problem: Errors of AI and their processing
2. **Non-classical alternative: In the non-classical world we have a choice: either return to classics, or exploit the blessing of dimensionality and create one-shot correctors of errors.**

Four messages

1. Errors are unavoidable companions of data driven AI: The main AI 2020 problem: Errors of AI and their processing
2. Non-classical alternative: In the non-classical world we have a choice: either return to classics, or exploit the blessing of dimensionality and create one-shot correctors of errors.
3. **After tomorrow AI will be the large social system of heterogeneous AI networks.**

Four messages

1. Errors are unavoidable companions of data driven AI: The main AI 2020 problem: Errors of AI and their processing
2. Non-classical alternative: In the non-classical world we have a choice: either return to classics, or exploit the blessing of dimensionality and create one-shot correctors of errors.
3. After tomorrow AI will be the large social system of heterogeneous AI networks.
4. **Explainability of AI is needed when AI makes mistakes.**

A new AI winter will come if we don't focus on the problem of AI errors

And only those applications will survive where errors are not very dangerous

References (blessing of dimensionality). 1

- J. W. Gibbs, Elementary Principles in Statistical Mechanics, developed with especial reference to the rational foundation of thermodynamics, (New York: Dover Publications, 1960 [1902])
- Lévy P, Problèmes concrets d'analyse fonctionnelle. Paris: Gauthier-Villars; 1951.
- F. Rosenblatt, Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms, Spartan Books, 1962
- Talagrand, M. (1995). Concentration of measure and isoperimetric inequalities in product spaces. *Publications Mathématiques de l'Institut des Hautes Etudes Scientifiques*, *81*(1), 73-205.
- M. Gromov, Isoperimetry of waists and concentration of maps, GAFA, Geom. Funct. Anal. 13 (2003) 178-215
- Donoho, D.L. High-Dimensional Data Analysis: The Curses and Blessings of Dimensionality. Invited Lecture at Mathematical Challenges of the 21st Century, AMS.

References (blessing of dimensionality). 2

- P. C. Kainen, Utilizing geometric anomalies of high dimension: When complexity makes computation easier. In *Computer Intensive Methods in Control and Signal Processing* (pp. 283-294). Birkhäuser, Boston, MA. (1997)
- P.C. Kainen, V. Kůrková, Quasiorthogonal dimension of Euclidian spaces, Appl. Math. Lett. 6(3) (1993) 7–10.
- Bárány I., Füredi Z. On the shape of the convex hull of random points. Probab. Th. Rel. Fields 77, 231–240 (1988).
- Jacobs, R. A., Jordan, M. I., Nowlan, S. J., & Hinton, G. E. (1991). Adaptive mixtures of local experts. *Neural computation*, *3*(1), 79-87.
- Donoho D, Tanner J. Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing. Phil. Trans. R. Soc. A 2009;367:4273–4293.
- Gorban, A. N., Makarov, V. A., & Tyukin, I. Y. (2019). **The unreasonable effectiveness of small neural ensembles in high-dimensional brain**. *Phys. Life Rev.*, *29*, 55-88.

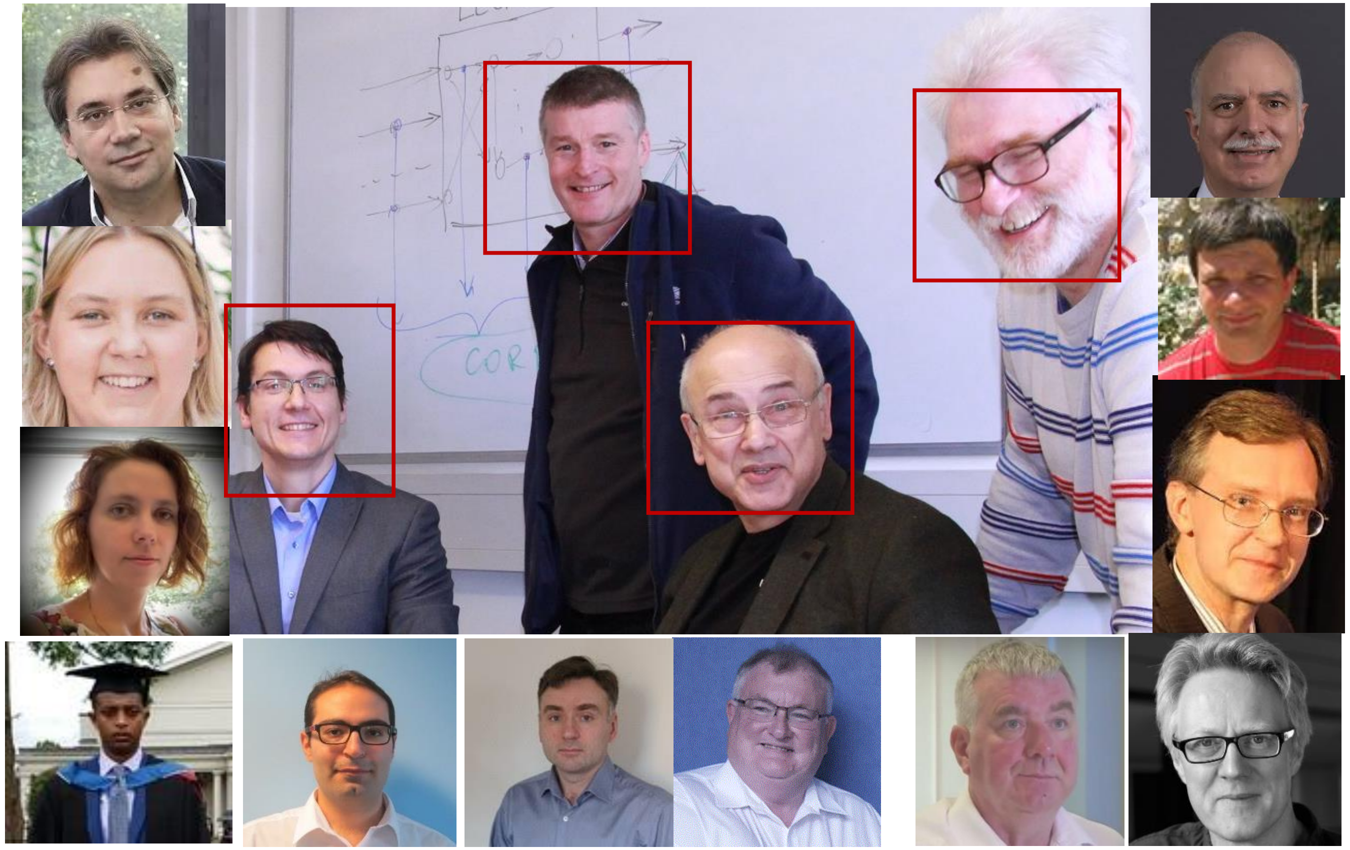
References (some preprints)

- Evgeny M Mirkes, Artificial Neural Network Pruning to Extract Knowledge, IJCNN2020, https://arxiv.org/abs/2005.06284
- Ivan Y. Tyukin, Desmond J. Higham, Alexander N. Gorban, On Adversarial Examples and Stealth Attacks in Artificial Intelligence Systems, IJCNN2020, https://arxiv.org/abs/2004.04479
- Alexander N. Gorban, Valery A. Makarov, Ivan Y. Tyukin, High--Dimensional Brain in a High-Dimensional World: Blessing of Dimensionality, *Entropy* 2020, 22(1), 82, https://arxiv.org/abs/2001.04959.
- Alexander N. Gorban, Valeri A. Makarov, Ivan Y. Tyukin, Symphony of high-dimensional brain, *Physics of Life Reviews*, 2019, https://arxiv.org/abs/1906.12222.
- A. N. Gorban, E. M. Mirkes, I. Y. Tyukin, How deep should be the depth of convolutional neural networks: a backyard dog case study, *Cognitive Computation* (2019), https://arxiv.org/abs/1805.01516.
- Alexander N. Gorban, Bogdan Grechuk, Ivan Y. Tyukin, Augmented Artificial Intelligence: a Conceptual Framework, https://arxiv.org/abs/1802.02172.
- A. N. Gorban, A. Golubkov, B. Grechuk, E. M. Mirkes, I. Y. Tyukin, Correction of AI systems by linear discriminants: Probabilistic foundations, *Information Sciences* 466 (2018), 303-322, https://arxiv.org/abs/1811.05321
- Ivan Y. Tyukin, Alexander N. Gorban, Stephen Green, Danil Prokhorov, Fast Construction of Correcting Ensembles for Legacy Artificial Intelligence Systems: Algorithms and a Case Study, *Information Sciences*, *485*, 230-247 (2018). https://arxiv.org/abs/1810.05593
- A.N. Gorban, I.Y. Tyukin, Blessing of dimensionality: mathematical foundations of the statistical physics of data, *Phil. Trans. R. Soc. A* 376 (2118), 376, 2018, https://arxiv.org/abs/1801.03421.
- Ivan Y. Tyukin, Alexander N. Gorban, Konstantin Sofeikov, Ilya Romanenko, Knowledge Transfer Between Artificial Intelligence Systems, *Frontiers in neurorobotics*, *12*, 49. 2018, https://arxiv.org/abs/1709.01547.
- A. N. Gorban, I. Y. Tyukin, Stochastic Separation Theorems, *Neural Networks* 94 (2017), 255-259, https://arxiv.org/abs/1703.01203
- Alexander N. Gorban, Ilya Romanenko, Richard Burton, Ivan Y. Tyukin, One-Trial Correction of Legacy AI Systems and Stochastic Separation Theorems, *Information Sciences*, 484, 237-254, 2019, https://arxiv.org/abs/1610.00494

Co-authors and collaborators

- Ivan Tyukin
- Andrei Zinovyev
- Jeremy Levesley
- Bogdan Grechuk
- Evgeny Mirkes
- Tatiana Tyukina
- Rosie Fenwick

- Eliyas Woldegeorgis
- Richard Burton
- Konstantin Sofeikov



- Stephen Green
- Sepehr Meshkinfamfard
- Ilya Romanenko

- Jays Shields
- John Downie

- Donald Wunsch
- Bogdan Grechuk
- Danil Prokhorov
- Des Higham

Thank you for your attention

Questions please