Een visie op het AI landschap

Ann Nowé

AI lab
Vrije Universiteit Brussel

ai.vub.ac.be

ARTIFICIAL INTELLIGENCE LAB

The birth of AI

IJCAI 1956

" We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer."

John McCarthy

Marvin Minsky

Nathaniel Rochester

Claude Shannon

Ann Nowé

The birth of AI

IJCAI 1956

" We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer."

John McCarthy

Marvin Minsky

Nathaniel Rochester

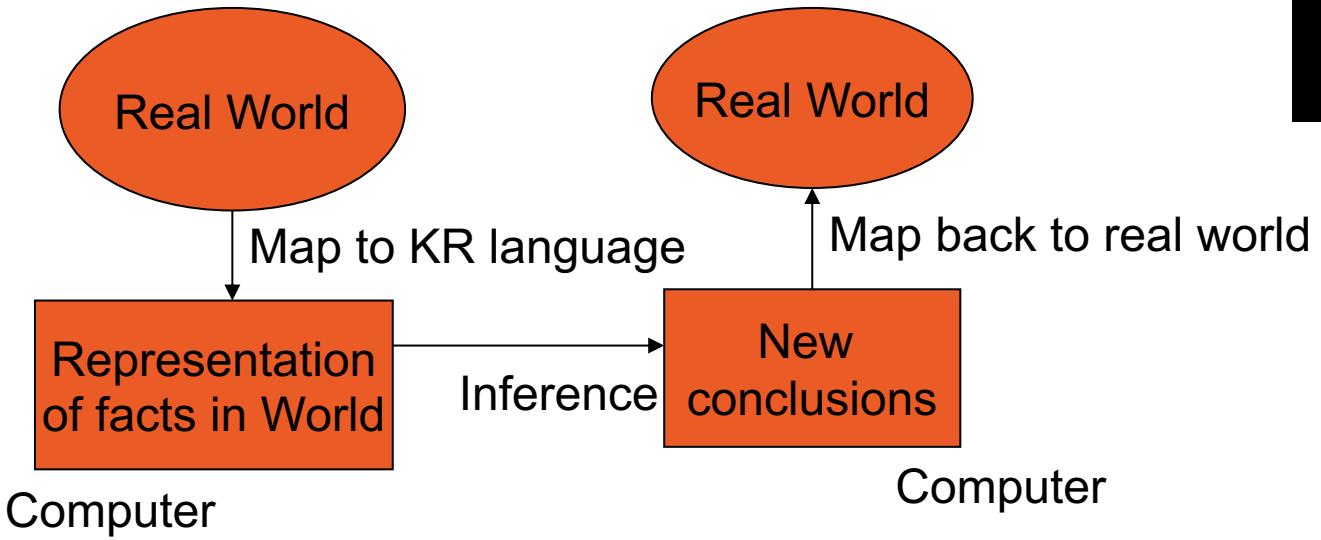
Claude Shannon

Ann Nowé

The symbolic approach

Knowledge representation and state space search

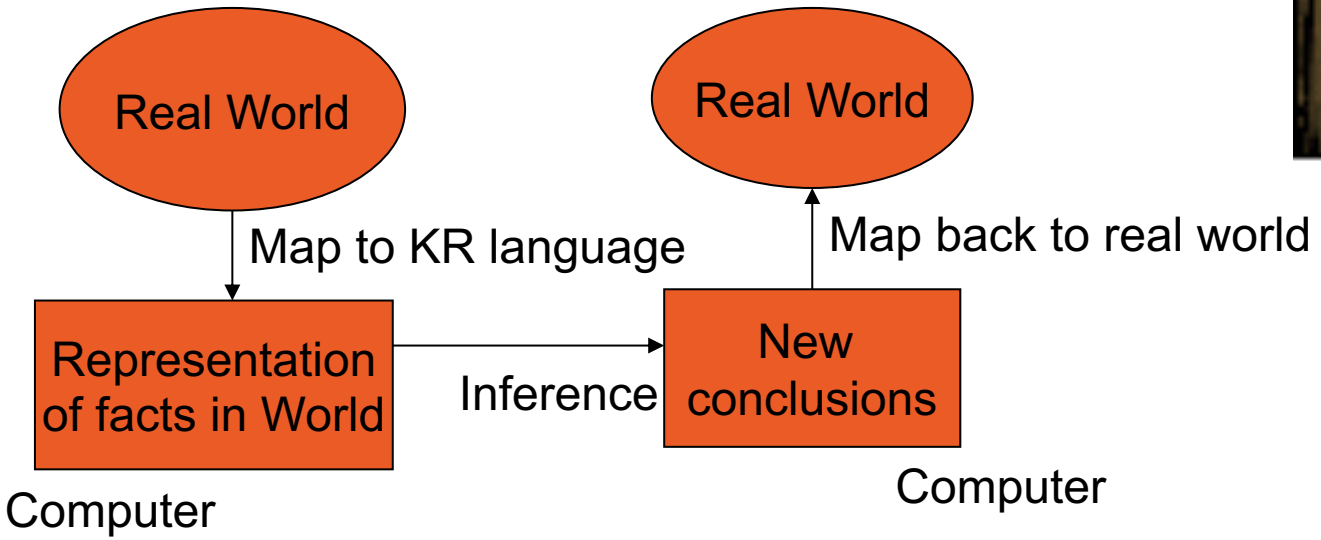
- Knowledge representation languages should have precise syntax and semantics.
- You must know exactly what an expression means in terms of objects in the real world.



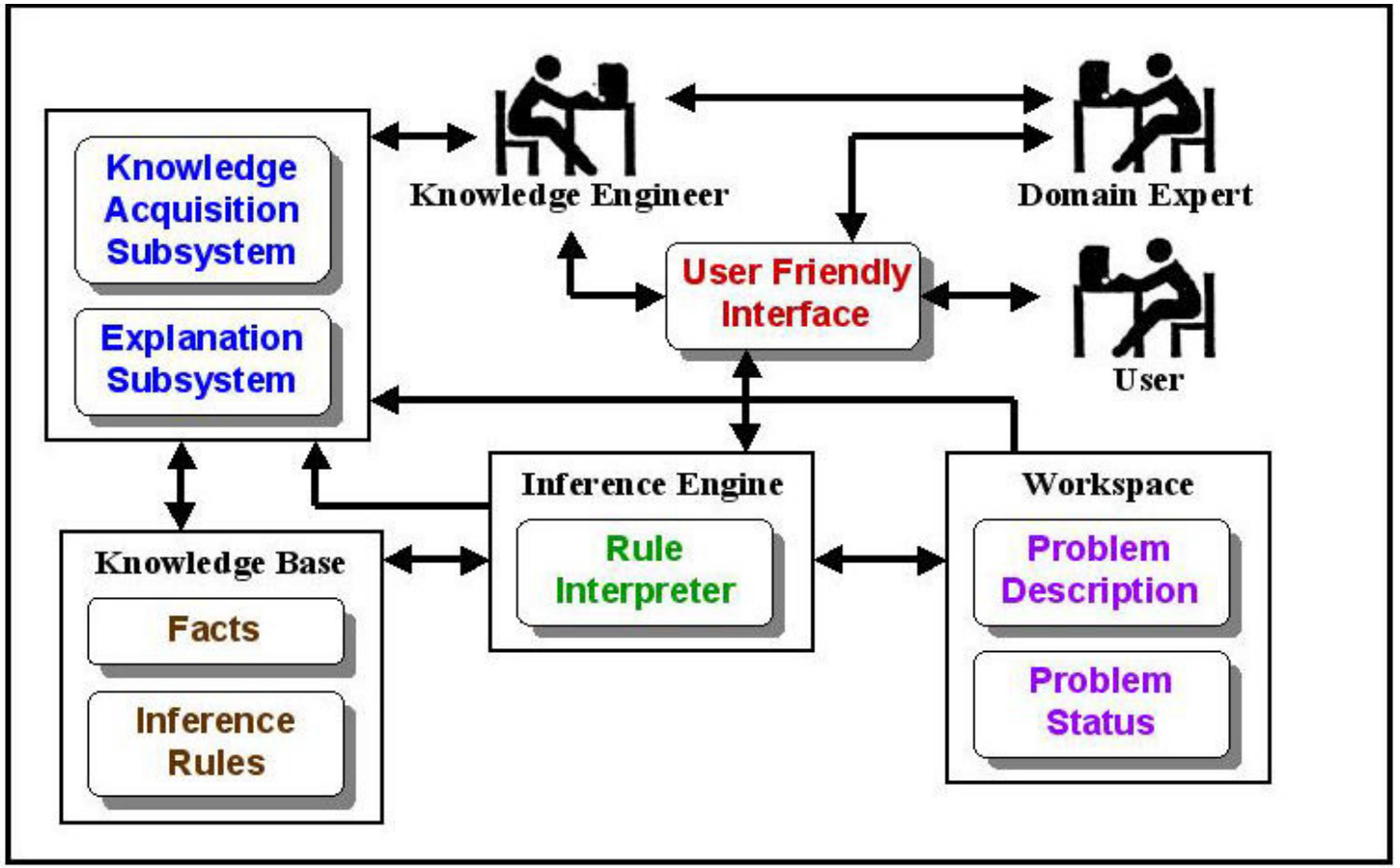
The symbolic approach

Knowledge representation and state space search

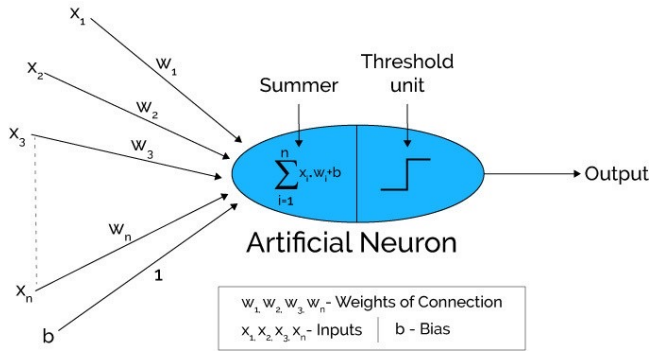
- Knowledge representation languages should have precise syntax and semantics.
- You must know exactly what an expression means in terms of objects in the real world.

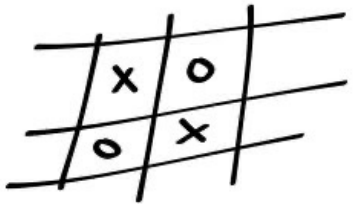


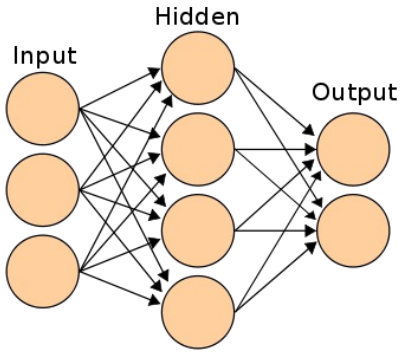
The symbolic approach



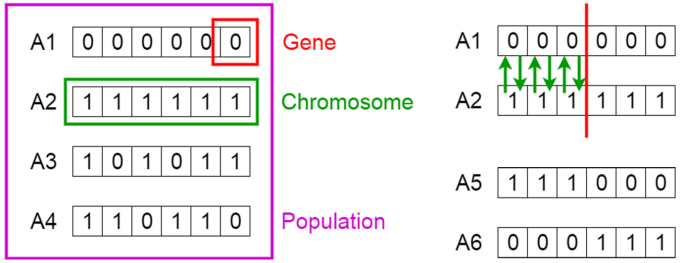
Sub-symbolic approach



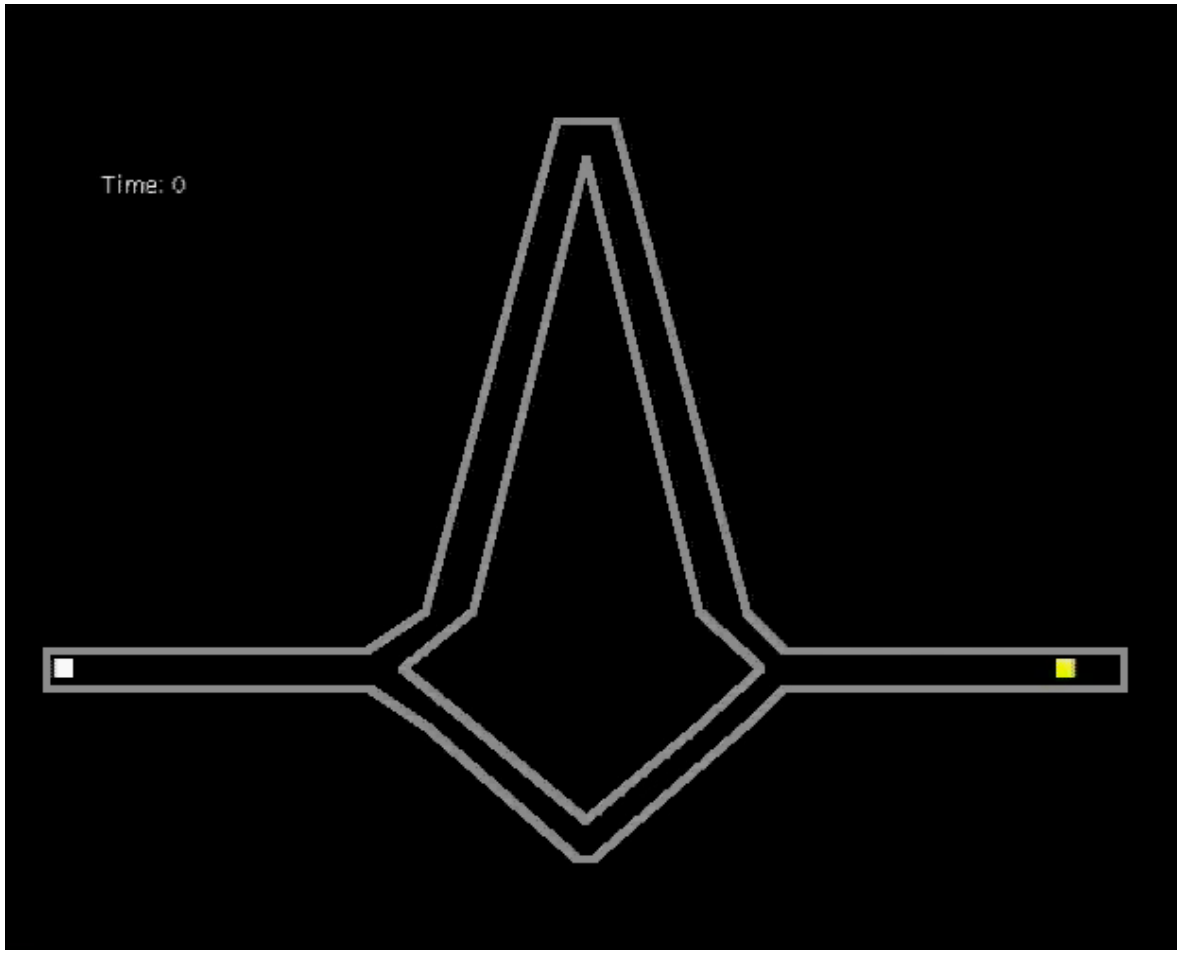


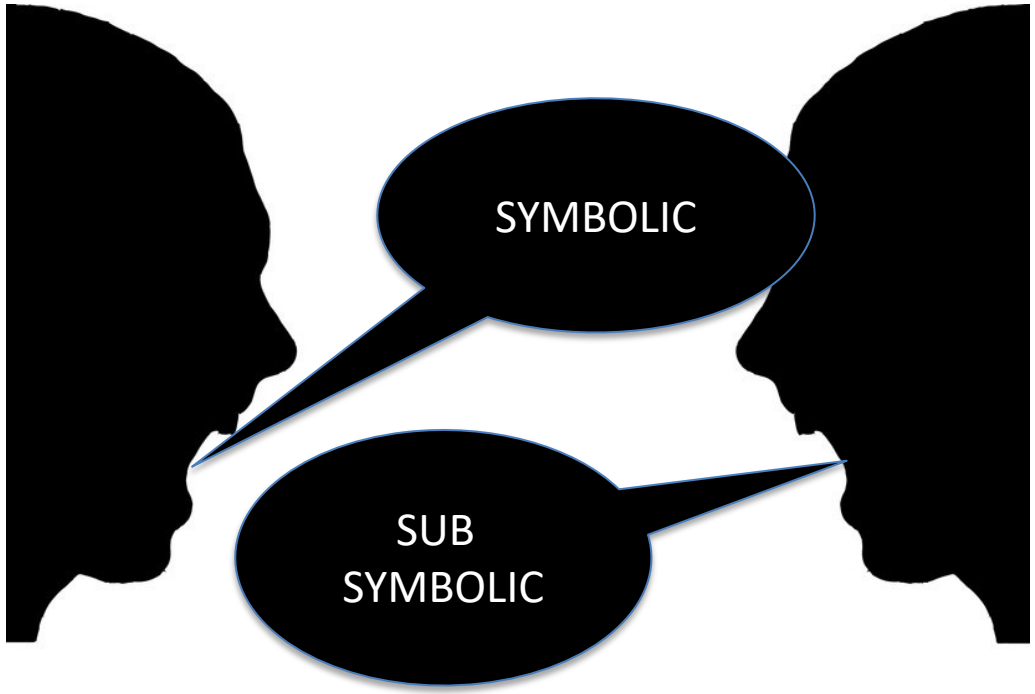


Genetic Algorithms

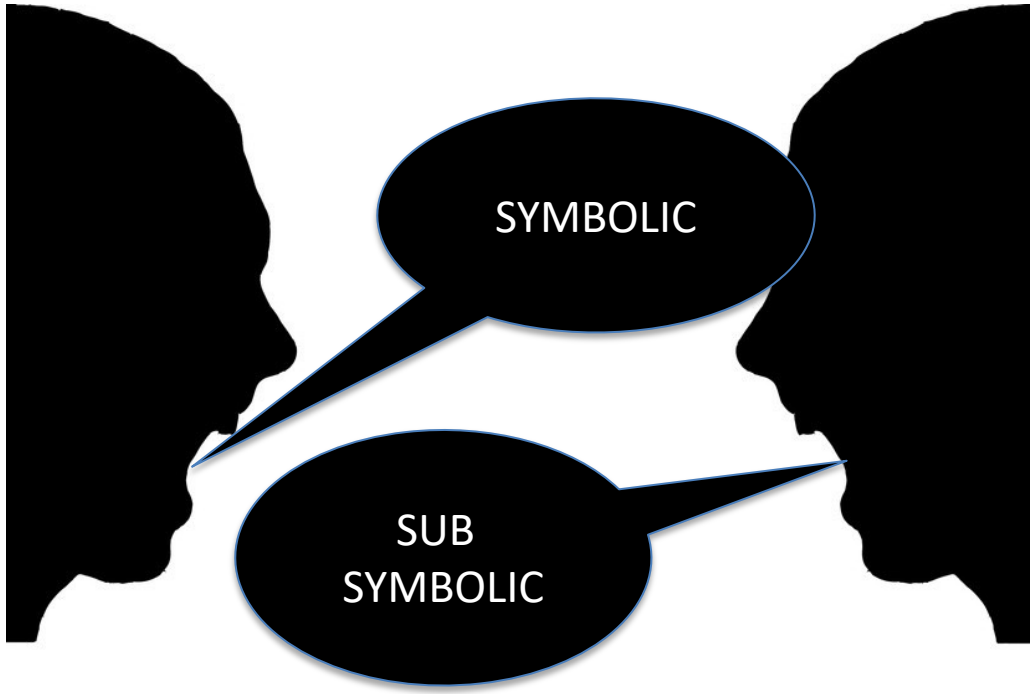


Ant algorithms

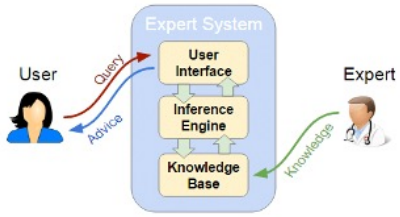


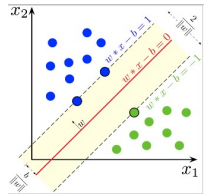

SYMBOLIC
SUB
SYMBOLIC

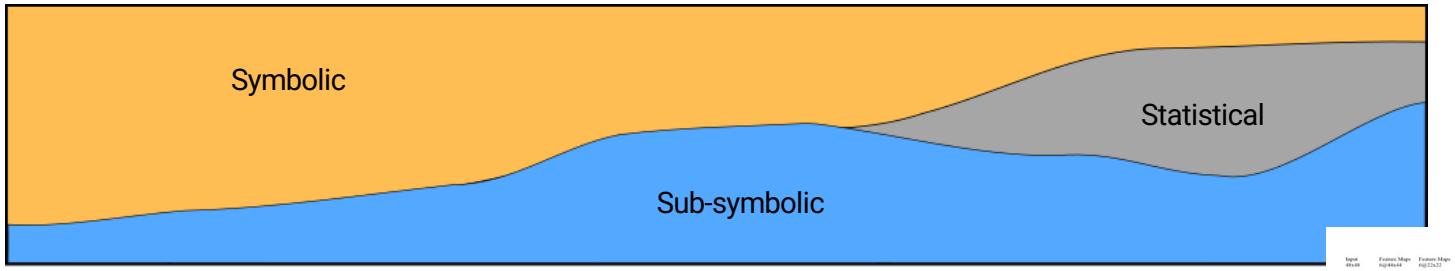
1988 AAAI Symposium on Parallel Models of Intelligence



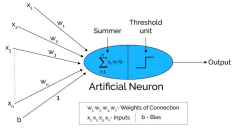
Brief history of AI: the different movements

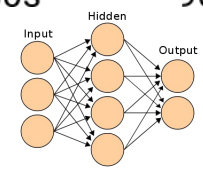


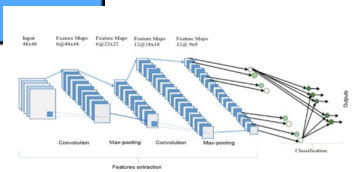




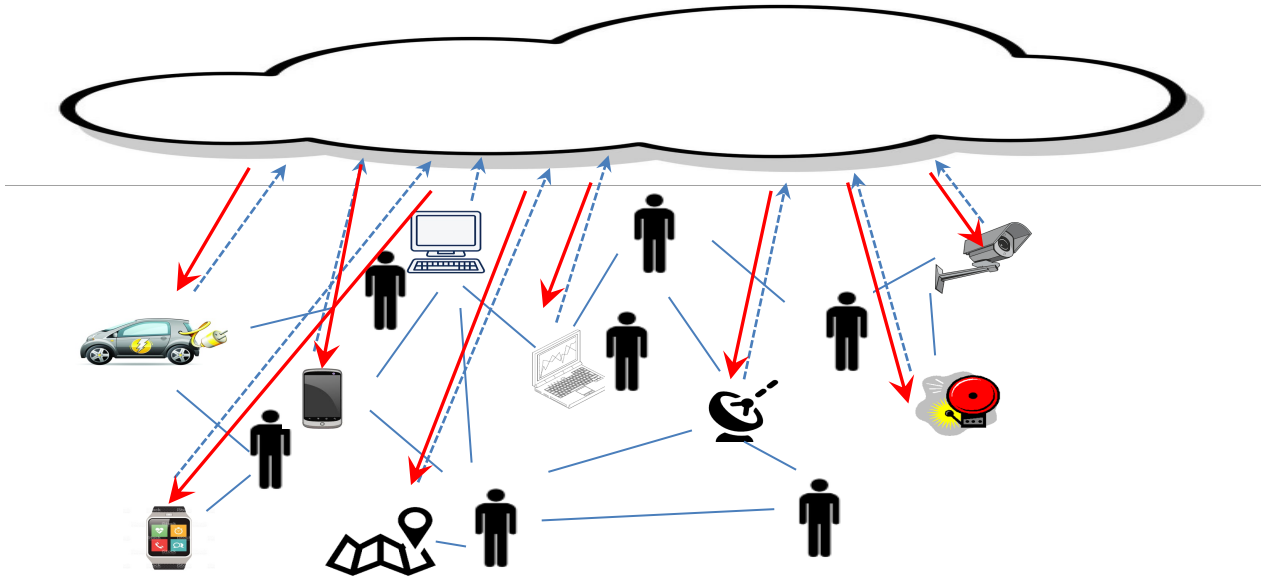
Hybrid AI







From data to big data



and cyber-physical systems

ARTIFICIAL INTELLIGENCE LAB

Ann Nowé

Self-driving cars

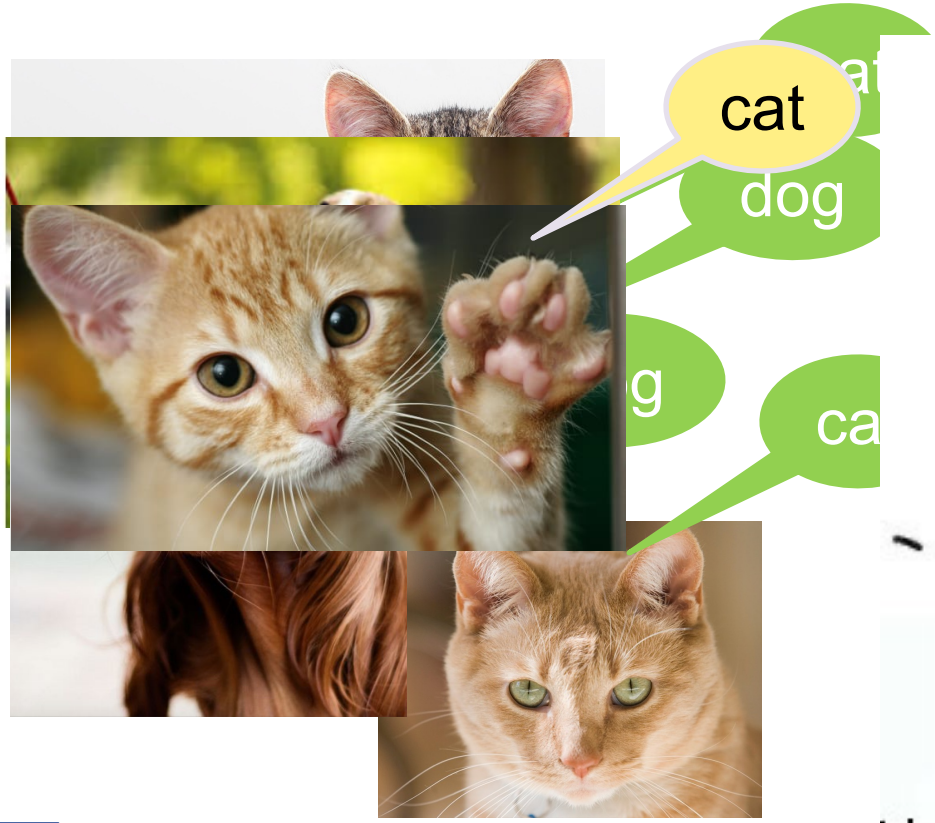
90s

ARTIFICIAL INTELLIGENCE LAB

Ann Nowé

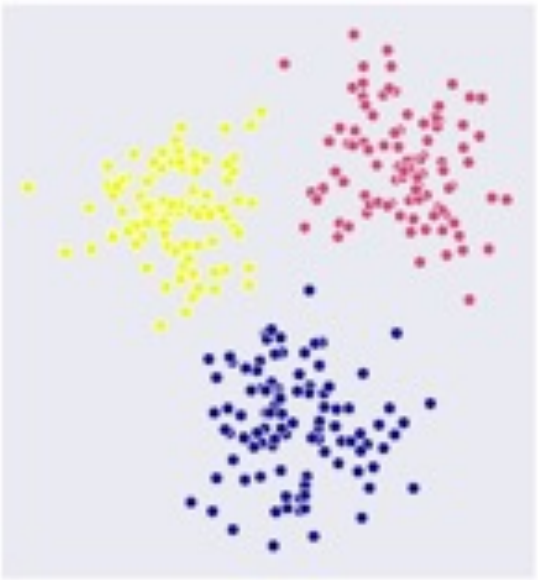
Subfields of Machine Learning

Supervised

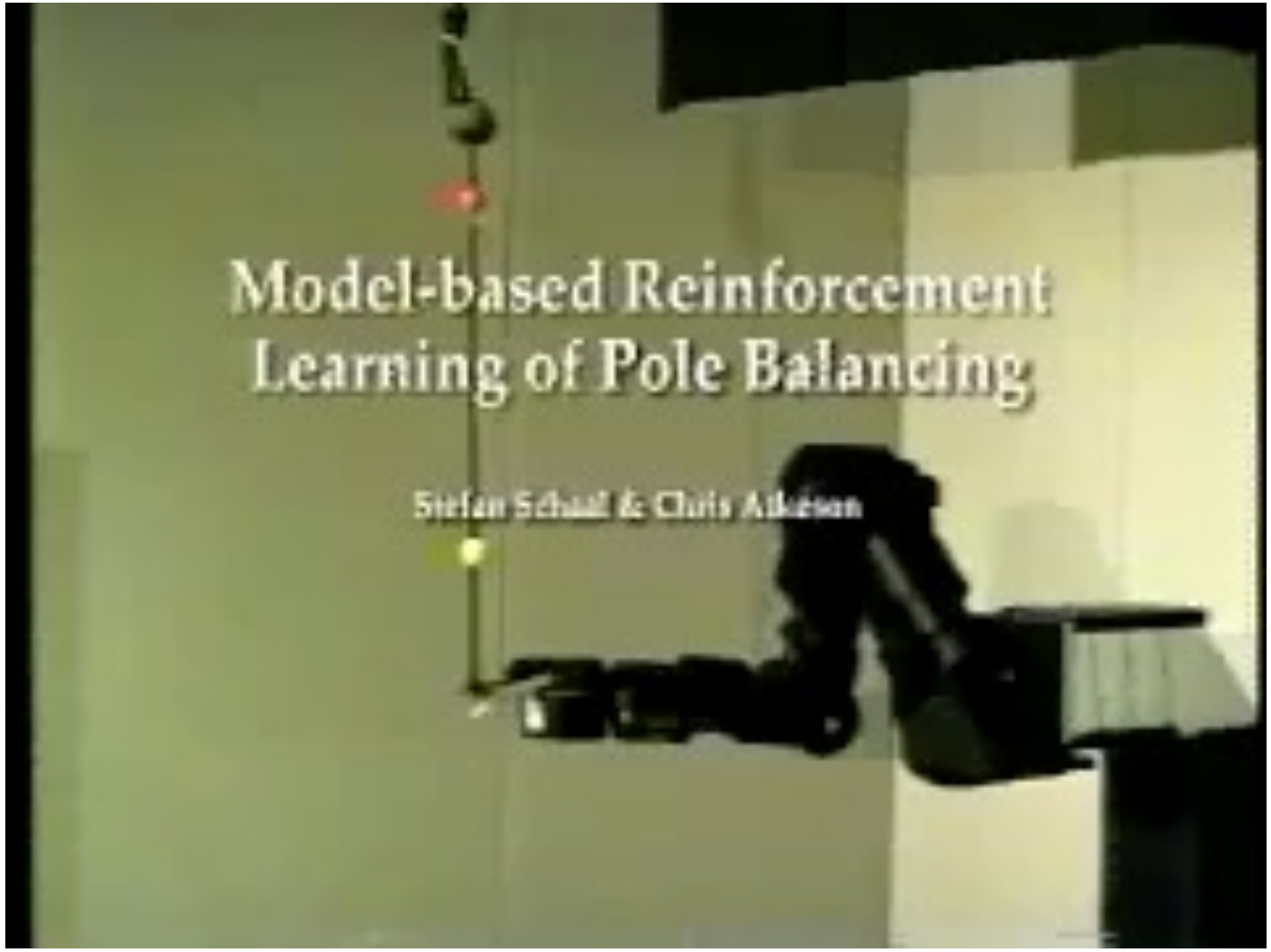


Reinforcement

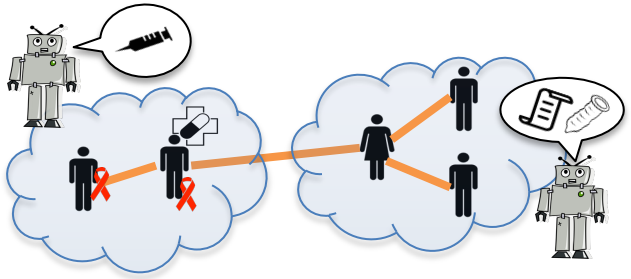
Unsupervised

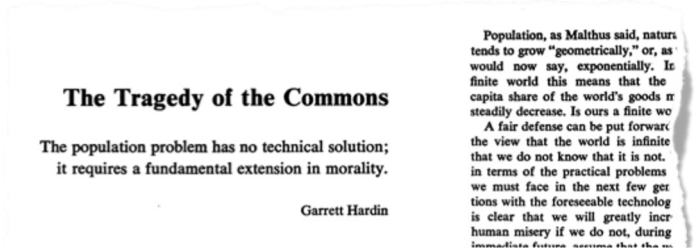


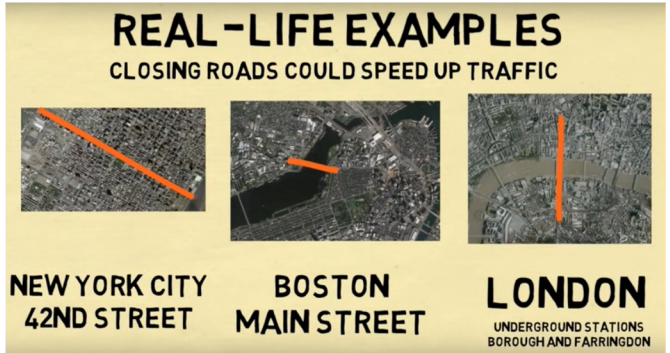
Reinforcement Learning



Multi-agent Systemen en speltheorie



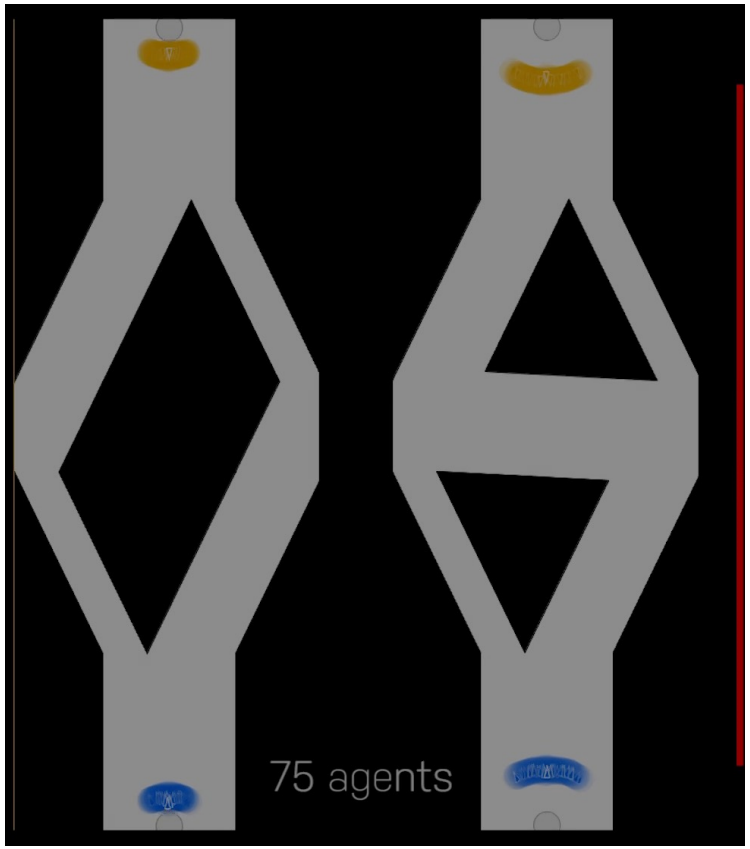




Braess paradox

Ann Nowé

Braess paradox

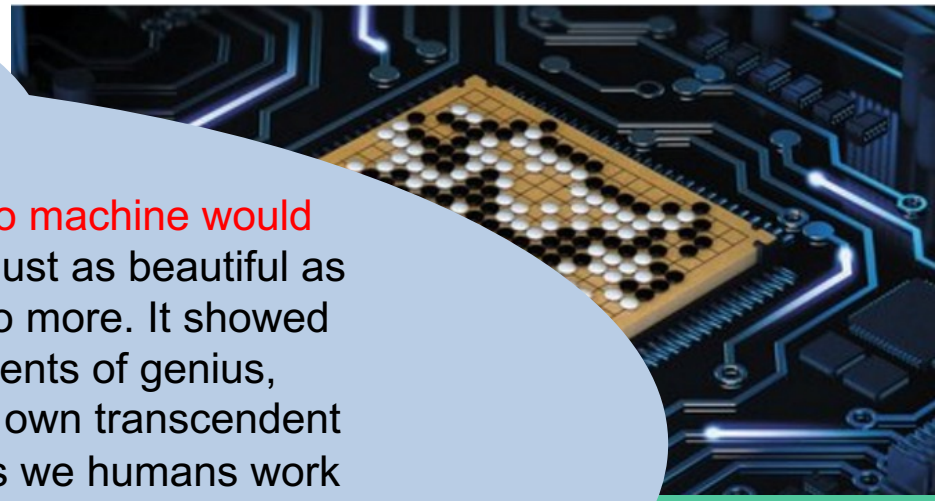


Ann Nowé

AI slimmer dan de mens?

AlphaGo

In Game Two, the Google machine made a move that no human ever would. And it was beautiful. As



But in Game Four, the human made a move that no machine would ever expect. And it was beautiful too. Indeed, it was just as beautiful as the move from the Google machine---no less and no more. It showed that although machines are now capable of moments of genius, humans have hardly lost the ability to generate their own transcendent moments. And it seems that in the years to come, as we humans work with these machines, our genius will only grow in tandem with our creation.

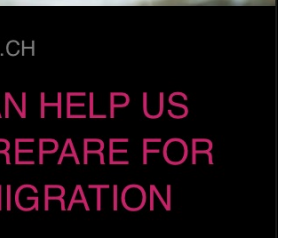
AI & kunst

AI : the good, the bad and the ugly

Ethics Inf Technol. 2021 Feb 9 : 1–7. PMCID: PMC7871022
doi: 10.1007/s10676-020-09567-7 [Epub ahead of print] PMID: 33584129

The CLAIRE COVID-19 initiative: approach, experiences and recommendations

Gianluca Bontempi,[1] Ricardo Chavarriaga,[2] Hans eD Canck.[3] Emanuela Girardi.[4] Holger Hoos.[5]
Iarla Kilbane-Dawe,[6] Tonio Ball,[7] Ann Nowé,[8]
Manlio eD Domenico,[12] Alessandro Saffiotti,[13]



How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

The AI That Predicts Your Sexual Orientation Simply By Looking At Your Face

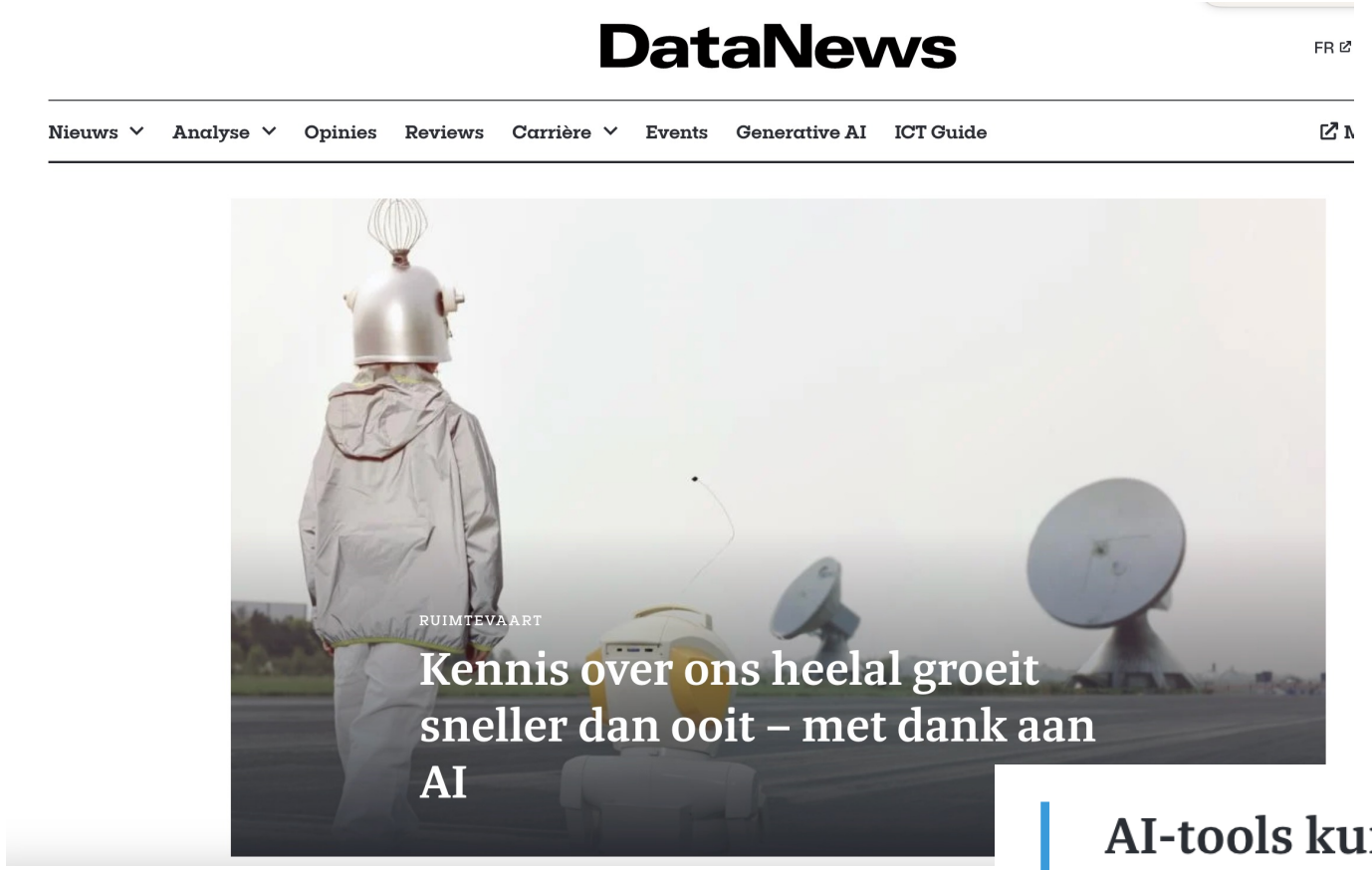
Bernard Marr, CONTRIBUTOR
FULL BIO
Opinions expressed by Forbes Contributors are their own.

JASON TASHEA SECURITY 04.17.17 7:00 AM

COURTS ARE USING AI TO SENTENCE CRIMINALS. THAT MUST STOP NOW

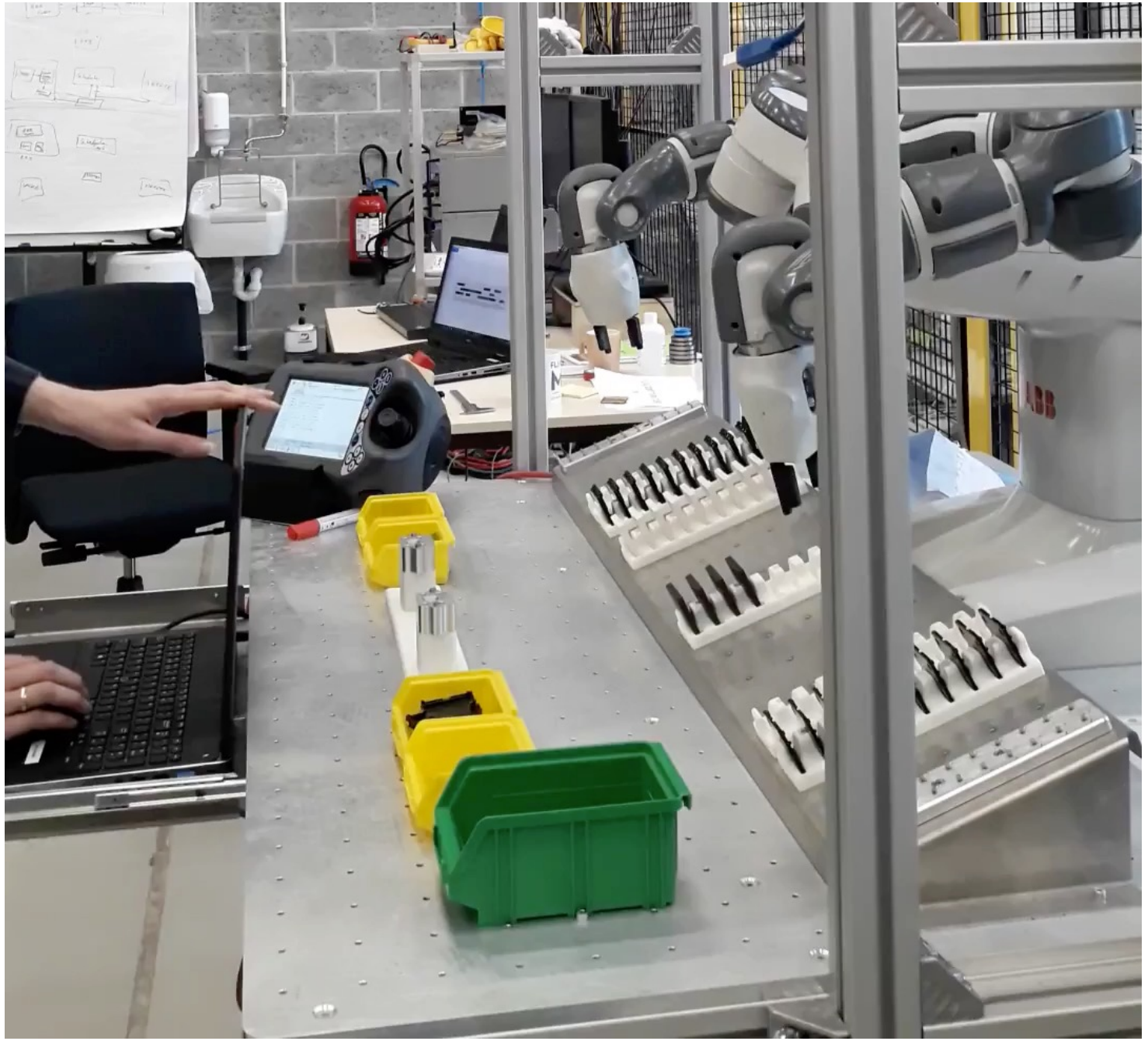
VUB ARTIFICIAL INTELLIGENCE LAB

Ann Nowé



> **AI-tools kunnen vandaag al met 96 procent nauwkeurigheid de signalen van een exoplaneet herkennen.**

https://datanews.knack.be/nieuws/innovatie/ruimtevaart/kennis-over-ons-heelal-groeit-sneller-dan-ooit-met-dank-aan-ai/



VUB ARTIFICIAL INTELLIGENCE LAB

Ann Nowé

In *9th CIRP Conference on Assembly Technology and Systems* (Vol. 106, pp. 227-232)

Cyberlegs

The bad

ARTIFICIAL
INTELLIGENCE
LAB

Ann Nowé

The uggly

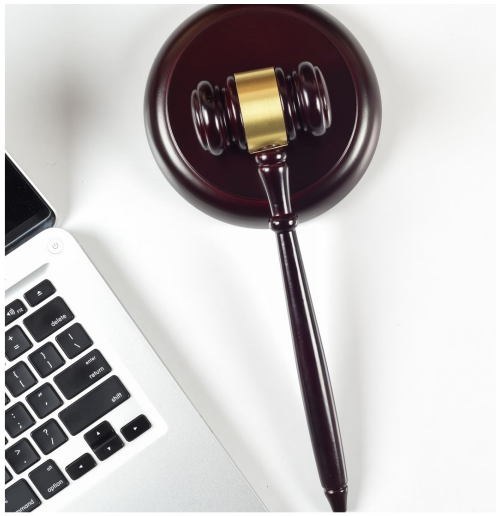
The Washington Post
Democracy Dies in Darkness

MONKEY CAGE

A computer program used for bail and sentencing decisions was labeled biased against blacks. It's actually not that clear.

By Sam Corbett-Davies, Emma Pierson, Avi Feller and Shar

October 17, 2016 at 5:00 a.m. EDT



VUB ARTIFICIAL INTELLIGENCE LAB

Gaat het algoritme in de fout?

Is het algoritme fair?

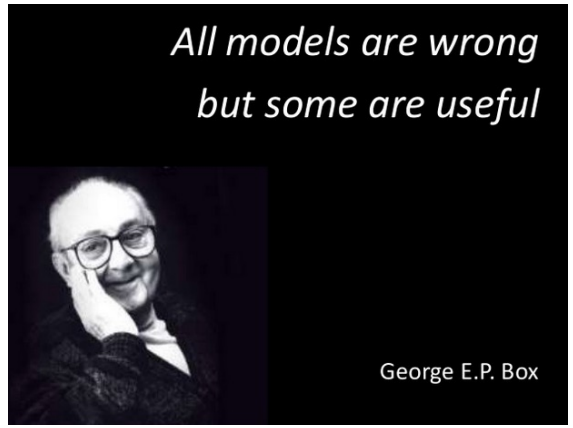


Table 15: Classification of fairness notions. (∗ notion newly defined in this paper)

Fairness Notion	Ref.	Formulation	Classification	Type
Statistical parity	[29]	$P(\hat{Y} \mid A=0) = P(\hat{Y} \mid A=1)$	Independence (equivalent or relaxed★)	Group
Conditional statistical parity	[21]	$P(\hat{Y}=1 \mid E=e, A=0) = P(\hat{Y}=1 \mid E=e, A=1)$★		
Equalized odds	[44]	$P(\hat{Y}=1 \mid Y=y,\ A=0) = P(\hat{Y}=1 \mid Y=y,\ A=1) \quad \forall y \in \{0,1\}$	Separation (equivalent or relaxed★)	
Equal opportunity		$P(\hat{Y}=1 \mid Y=1, A=0) = P(\hat{Y}=1 \mid Y=1, A=1)$★		
Predictive equality	[21]	$P(\hat{Y}=1 \mid Y=0, A=0) = P(\hat{Y}=1 \mid Y=0, A=1)$★		
Balance for positive class	[58]	$E[S \mid Y=1, A=0)] = E[S \mid Y=1, A=1]$★		
Balance for negative class		$E[S \mid Y=0, A=0] = E[S \mid Y=0, A=1]$★		
Overall balance	*	$E[S \mid Y=y, A=0] = E[S \mid Y=y, A=1] \quad \forall y \in \{0,1\}$		
Conditional use acc. equality	[10]	$P(Y=y \mid \hat{Y}=y, A=0) = P(Y=y \mid \hat{Y}=y, A=1) \quad \forall y \in \{0,1\}$	Sufficiency (equivalent or relaxed★)	
Predictive parity	[18]	$P(Y=1 \mid \hat{Y}=1, A=0) = P(Y=1 \mid \hat{Y}=1, A=1)$★		
Negative predictive parity	*	$P(Y=1 \mid \hat{Y}=0, A=0) = P(Y=1 \mid \hat{Y}=0, A=1)$★		
Calibration	[18]	$P(Y=1 \mid S=s, A=0) = P(Y=1 \mid S=s, A=1) \quad \forall s \in [0,1]$		
Well-calibration	[58]	$P(Y=1 \mid S=s, A=0) = P(Y=1 \mid S=s, A=1) = s \quad \forall\, s \in [0,1]$		
Overall accuracy equality	[10]	$P(\hat{Y}=Y \mid A=0) = P(\hat{Y}=Y \mid A=1)$	Other metrics from confusion matrix	
Treatment equality		$\frac{FN}{FP}_{(A=0)} = \frac{FN}{FP}_{(A=1)}$		
Total fairness		–	Independence, Separation and Sufficiency	
Total effect	[76]	$TE_{a_1,a_0}(\hat{y}) = P(\hat{y}_{A \leftarrow a_1}) - P(\hat{y}_{A \leftarrow a_0})$	Causality	
Effect of treatment on treated		$ETT_{a_1,a_0}(\hat{y}) = P(\hat{y}_{A \leftarrow a_1} \mid a_0) - P(\hat{y} \mid a_0)$		
No unresolved discrimination	[53]	–		
No proxy discrimination		$P(\hat{Y} \mid do(P_x = p)) = P(\hat{Y} \mid do(P_x = p')) \quad \forall P_x \quad and \quad \forall\, p, p'$		
Counterfactual fairness	[59]	$P(\hat{Y}_{A \leftarrow a}(U) = y \mid X=x, A=a) = P(\hat{Y}_{A \leftarrow a'}(U) = y \mid X=x, A=a)$		Individual
Causal discrimination	[38]	$X_{(A=0)} = X_{(A=1)} \ \wedge\ A_{(A=0)} \neq A_{(A=1)} \ \Rightarrow \hat{y}_{(A=0)} = \hat{y}_{(A=1)}$	Similarity Metric	
Fairness through awareness	[29]	$D(M(v_i), M(v_j)) \leq d(v_i, v_j)$		

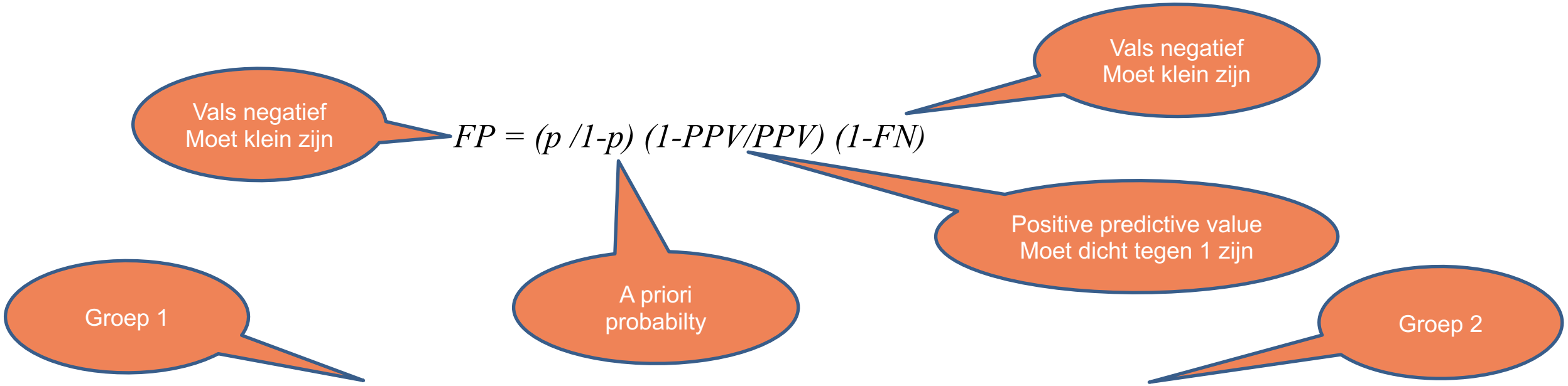
27

Karima Makhlouf, Sami Zhioua, and Catuscia Palamidessi. 2021. On the Applicability of Machine Learning Fairness Notions. SIGKDD Explor. Newsl. 23, 1 (June 2021), 14–23. https://doi.org/10.1145/3468507.3468511

VUB ARTIFICIAL INTELLIGENCE LAB

Ann Nowé

Kan een algoritme (model) fair zijn?



Model is fair als

$FP_1 = FP_2$

$FN_1 = FN_2$

$PPV_1 = PPV_2$

Impliceert

$p_1 = p_2$

Explainable & Accountable AI

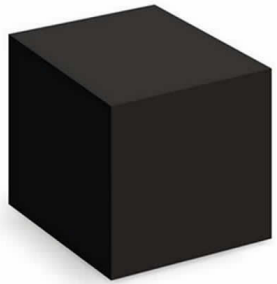
General Data Protection Regulation
GDPR

Article 22 :

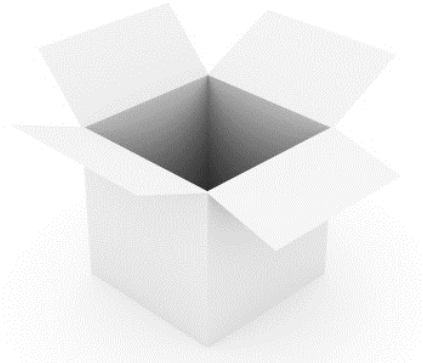
All AI that has impact on human lives, will need to be explainable and accountable,

the interpretability of machine learning based models will be key for the usage of these models.

Black box vs white box

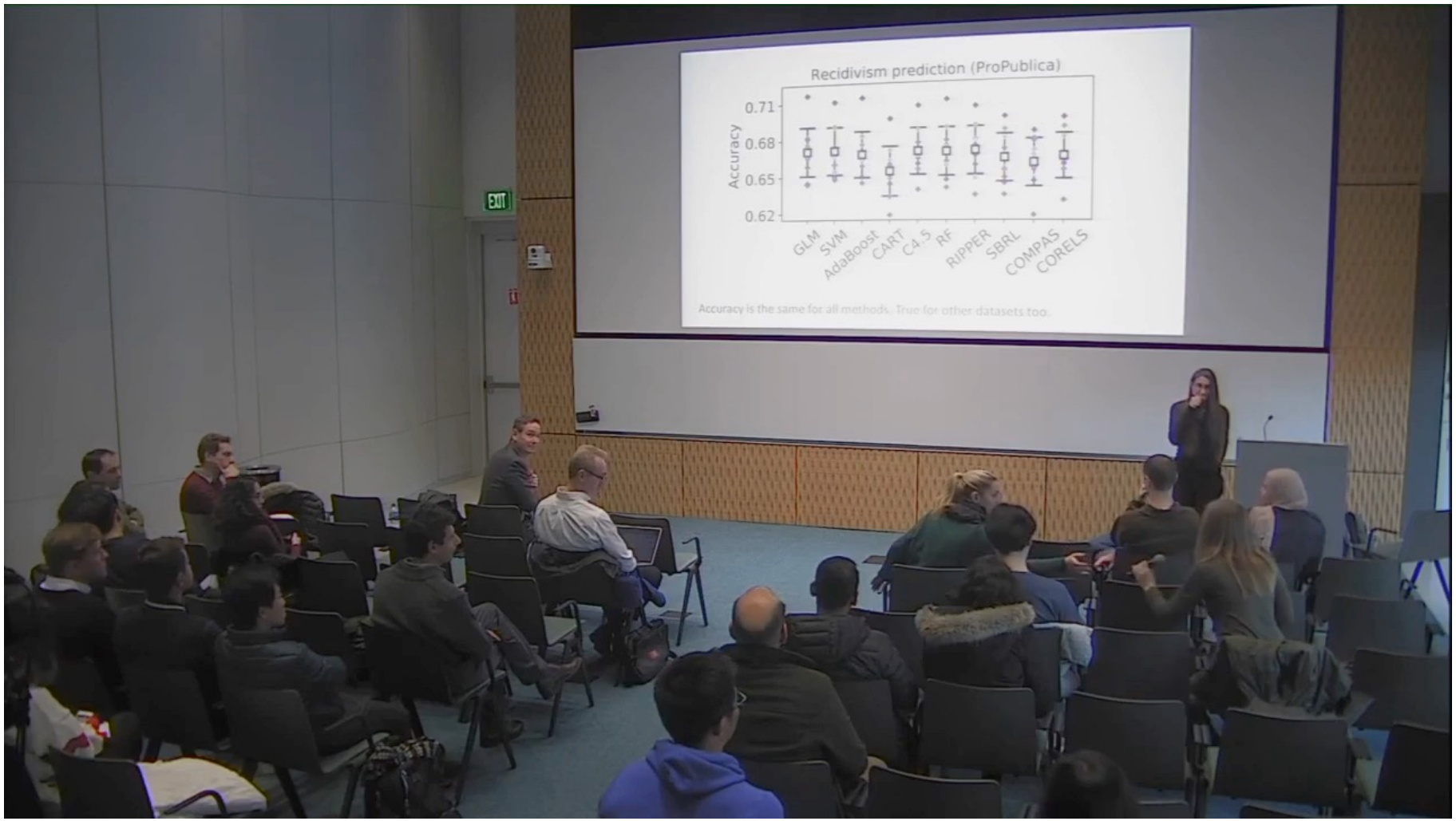


Black box model
Often very performant
But can not explain itself



White box
Interpretable model
Transparent model
Robust model

The uggly



Prof. Cynthia Rudin

Ann Nowé

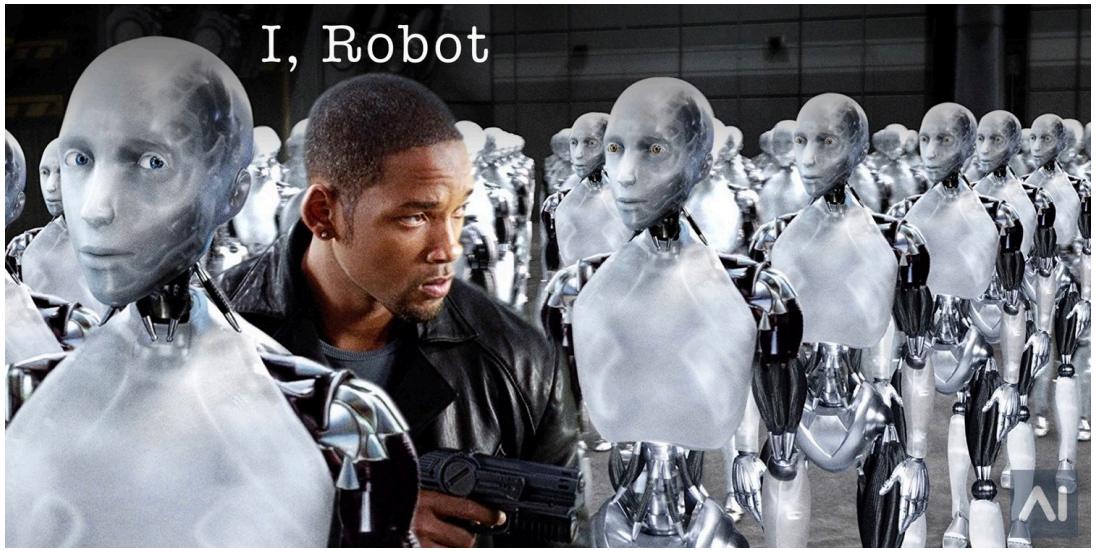
Artificial Intelligence versus consciousness

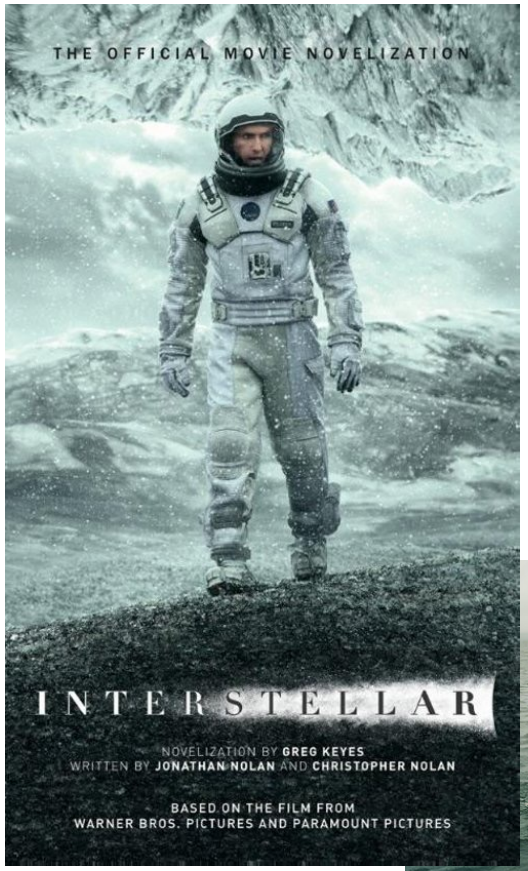
Yuval Noah Harari, Antwerpen, 27/1/2020

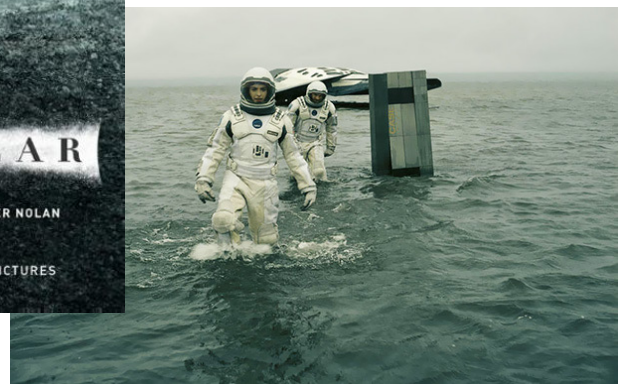
ARTIFICIAL INTELLIGENCE LAB

Ann Nowé

I, robot *versus* Interstellar







Vlaams AI-onderzoeksprogramma

LUIK 1: STRATEGISCH BASISONDERZOEK VERSTERKEN - VLAAMS AI-ONDERZOEKSPROGRAMMA

LUIK 2: GEBRUIK VAN AI DOOR BEDRIJVEN STIMULEREN

LUIK 3: ONDERSTEUNING BIEDEN MET BEWUSTMAKING, OPLEIDING EN ETHISCHE OMKADERING

Kenniscentrum Data & Maatschappij

Kenniscentrum voor juridische, ethische en maatschappelijke aspecten van artificiële intelligentie en datatoepassingen.

CONSORTIUM

FLANDERS MAKE | imec | VUB | Universiteit Antwerpen
vito | UHASSELT | KU LEUVEN | sirris
VIB | GHENT UNIVERSITY | Lees meer over Consortium

VUB ARTIFICIAL INTELLIGENCE LAB

Ann Nowé

AI experience center