

AI研究最前線

福田剛志
IBM Research - Tokyo

自己紹介

福田剛志 (ふくだ たけし)

1991年 日本アイ・ビー・エム(株)入社

同年～2004年 IBM東京基礎研究所 (TRL) にて、

オブジェクト指向データベース

データマイニング

バイオインフォマティクス

などの研究に従事

1999年 早稲田大学より博士(情報科学)取得

2004年～2012年 大和ソフトウェア開発研究所 (YSL) にて、

データベース・情報統合関連製品の開発を担当

2012年 ソフトウェア開発担当 理事

2015年 東京基礎研究所所長 理事

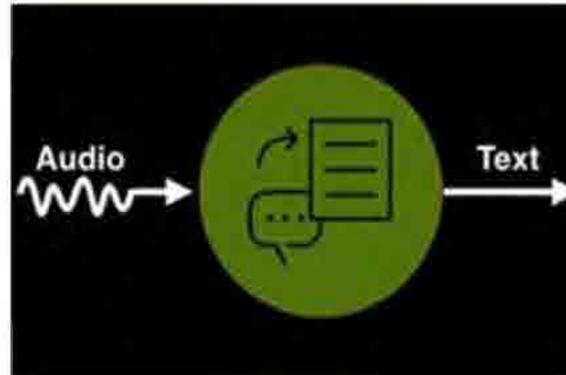


AIの現状

Deep Learning は AI を革新した



言語翻訳



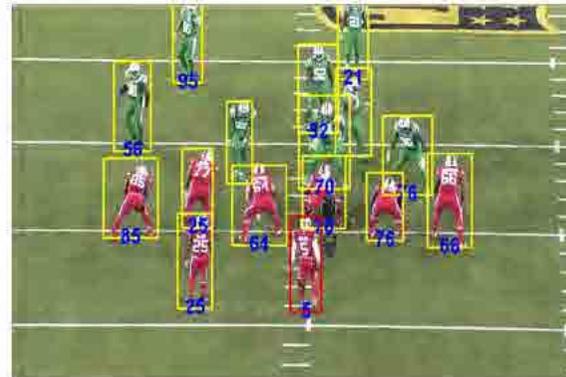
音声認識

Vehicle 1, a 1995 Honda Civic was traveling north on a two lane undivided roadway, negotiating a curve to the left on an upgrade. V1 went over the right lane line, overcorrected and went over the left lane line into the southbound lane. V1 overcorrected again and went across the northbound lane, over the right lane line.

自然言語処理



知的ゲーム



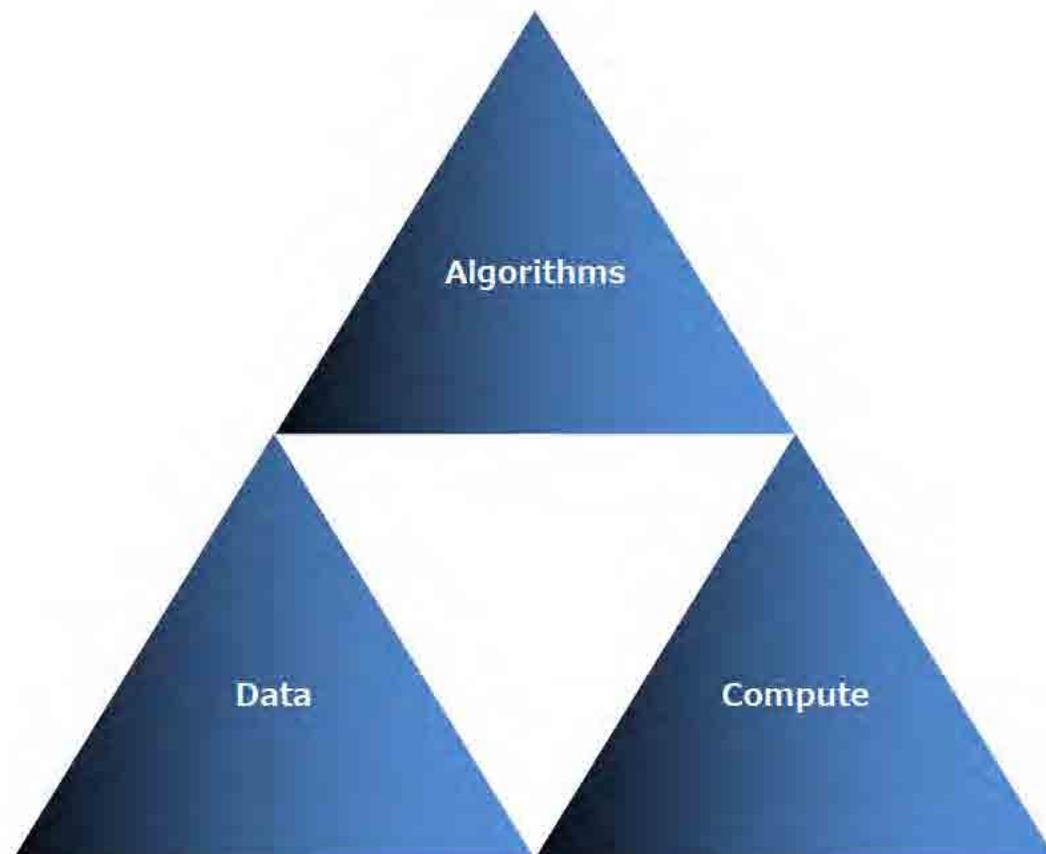
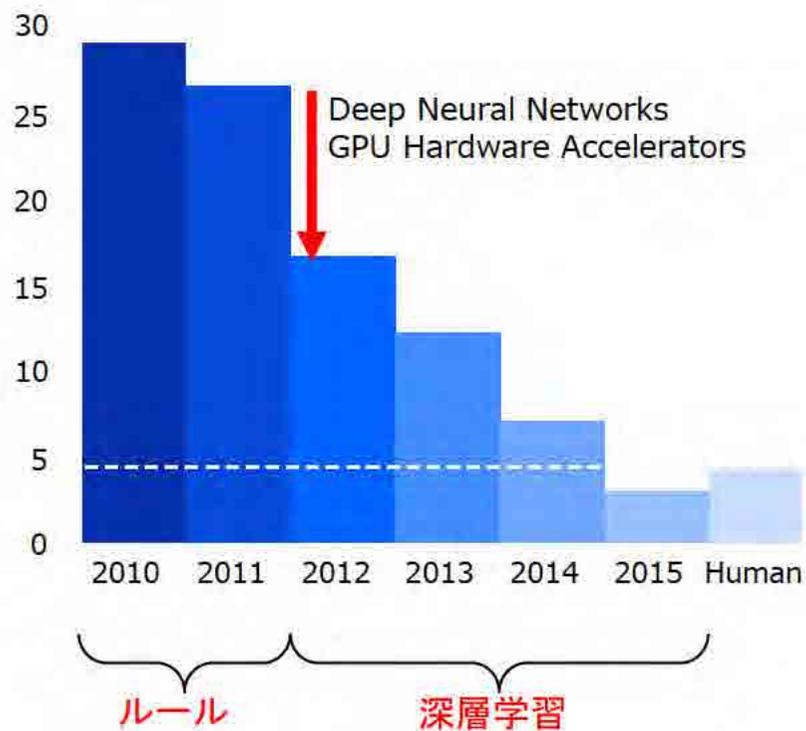
物体認識



顔認識

2012: 深層学習 (ディープ・ラーニング) の爆発

ImageNet 画像分類エラー率



Data: ImageNet

- 1,400万枚を超える画像
- 画像に写っている物体名(クラス名)が付与されている
- 物体名は2万種以上

Java sparrow, Java finch, ricebird, Padda oryzivora

Small finch-like Indonesian weaverbird that frequents rice fields

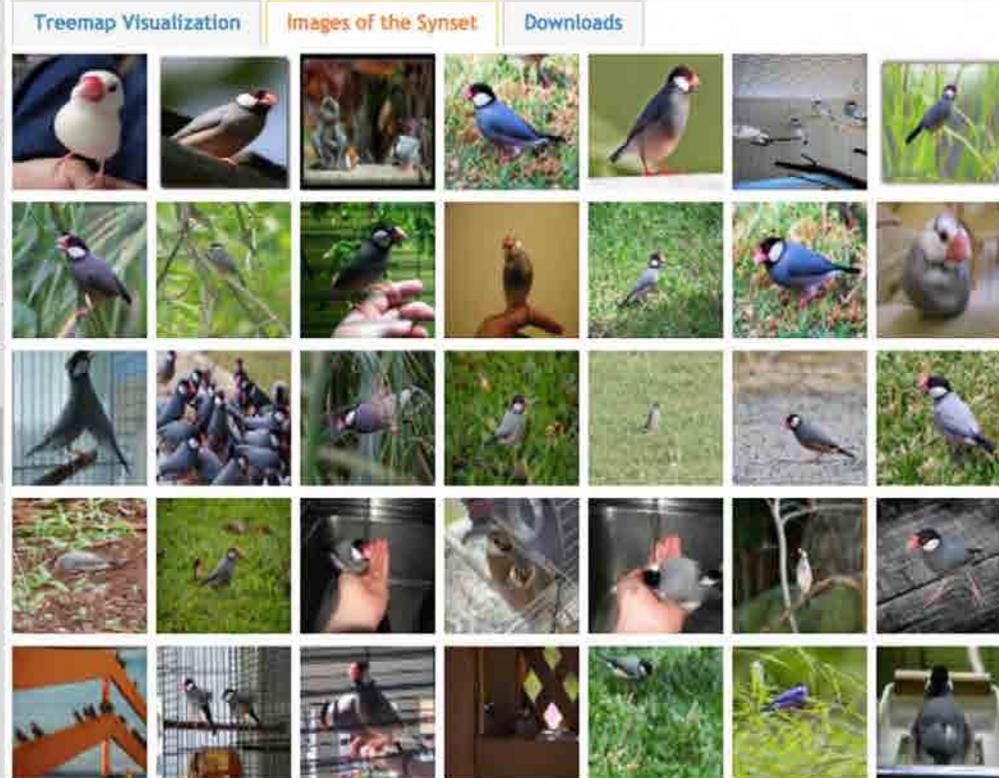
877 pictures

29.19% Popularity Percentile



Numbers in brackets: (the number of synsets in the subtree).

- ImageNet 2011 Fall Release (32326)
 - plant, flora, plant life (4486)
 - geological formation, formation (1)
 - natural object (1112)
 - sport, athletics (176)
 - artifact, artefact (10504)
 - fungus (308)
 - person, individual, someone, somebody (147)
 - animal, animate being, beast, brute, creature, fauna (15187)
 - invertebrate (766)
 - homeotherm, homoiotherm, warm-blooded (61)
 - work animal (4)
 - dart (0)
 - survivor (0)
 - range animal (0)
 - creepy-crawly (0)
 - domestic animal, domesticated (10)
 - molter, moult (0)
 - varmint, varment (0)
 - mutant (0)
 - critter (0)
 - game (47)
 - young, offspring (45)
 - poikilotherm, ectotherm (0)
 - herbivore (0)
 - peeper (0)
 - pest (1)
 - female (4)
 - insectivore (0)

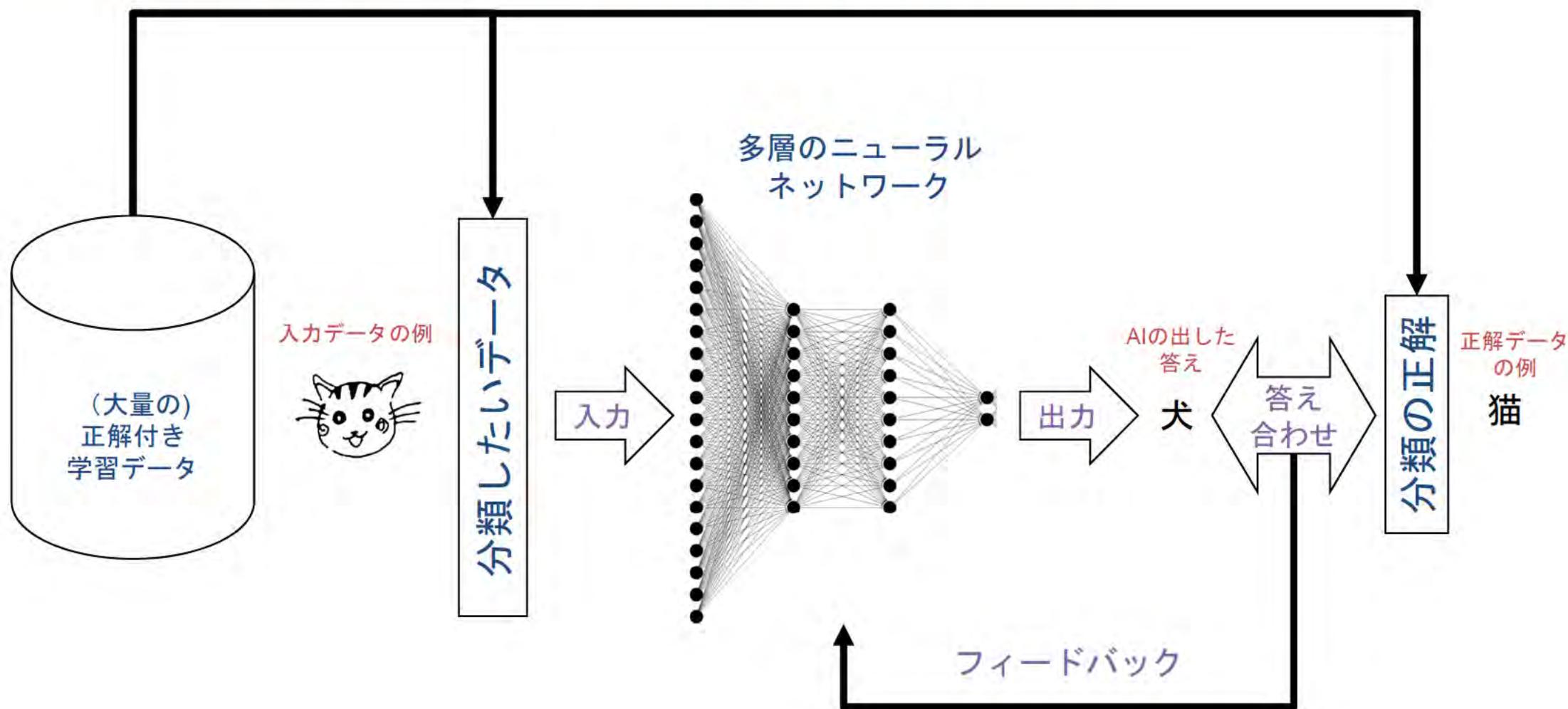


*Images of children synsets are not included. All images shown are thumbnails. Images may be subject to copyright.

Prev 1 2 3 4 5 6 7 8 9 10 ... 25 26 Next

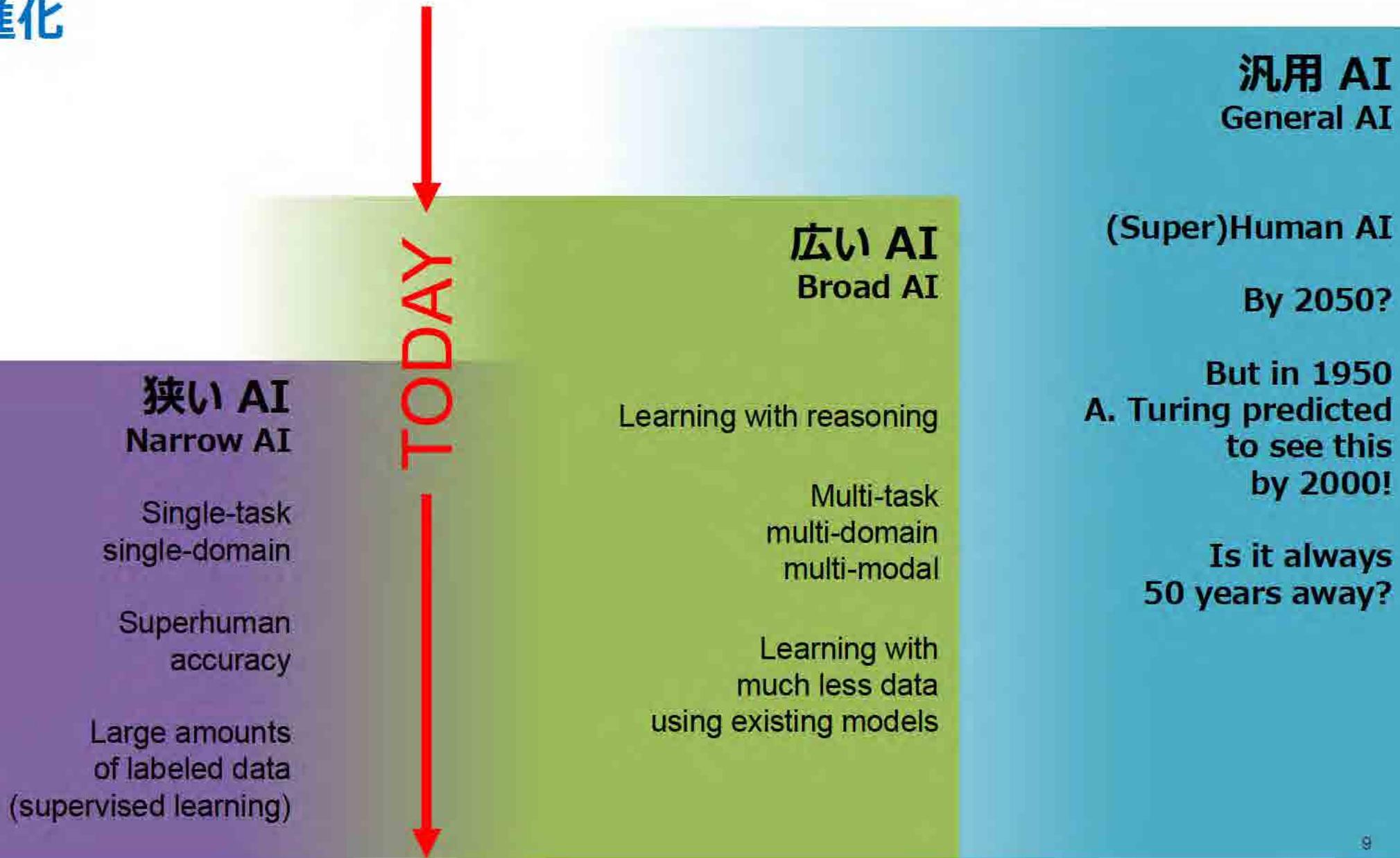
<http://www.image-net.org/>

The Algorithm: ディープ・ニューラル・ネットワーク



これからのAIに
求められるもの

AIの進化



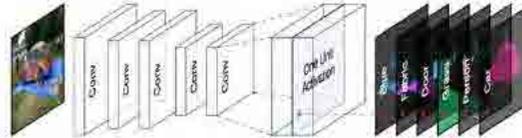
Broad AIに求められるもの

少量のデータから学習

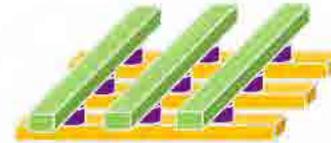


Multi-task and multi-domain

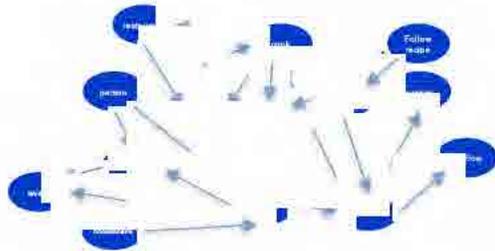
説明性



計算性能の向上



学習+推論



マルチモーダル



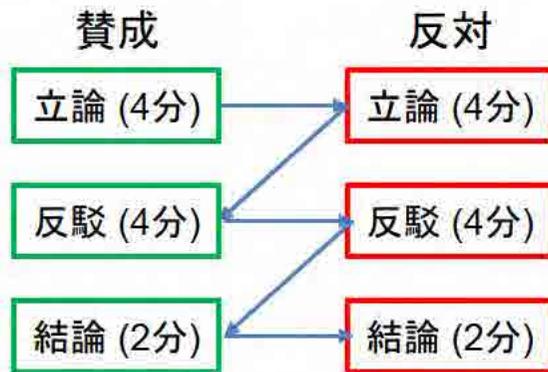
It was the best of times, it was the worst of times, it was the age of wisdom, it was the age of foolishness...

信頼性



Debater - 人間と討論するコンピュータ

- IBM基礎研究所が2012年から開発
- 与えられたトピックに関して、人間のディベーター
ディベート(討論)を行う



2018年、2019年に人間との公開討論を実施

- 1回目 2018年
Dan Zafrir = イスラエル人のディベート熟練者
[トピック] 遠隔医療の利用を活発にすべきか
- 2回目 2019年
Harish Natarajan = ディベート勝利回数で世界記録
[トピック] 幼稚園・保育園での教育に助成金を支給すべきか



intelligence²
DEBATES

#iq2uslive

こんにちは、ハリッシュ。あなたは人間とのディベートの
勝利数で、世界記録を持っているそうね。

これからのAIに必須の要素: シンボル + 言語

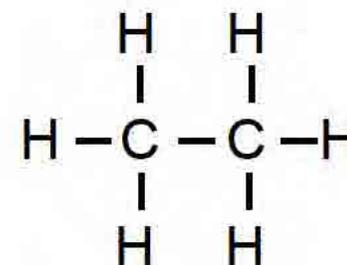
▪ シンボル

- 物事を抽象化

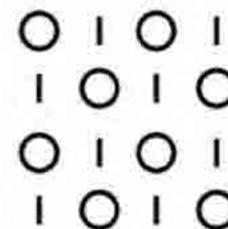
▪ 言語

- 物事の間係を記述
- 新しい概念を表現

- シンボルと言語によって、知識を
表現、獲得、共有、通信
可能になる



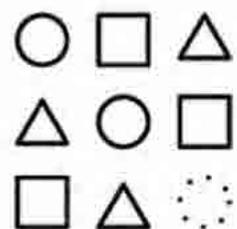
$$a^2 + b^2 = c^2$$



$$E = mc^2$$

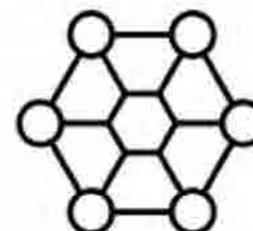
Symbolic AI + Neural AI = Neuro-Symbolic AI

シンボルAI

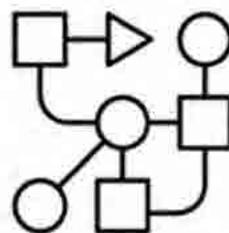


シンボルで記述
ただ、学習能力に限界

ニューラルAI



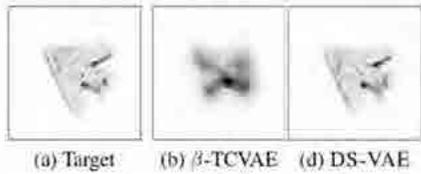
精度が良い
ただ、ブラックボックス



ニューロシンボリックAI

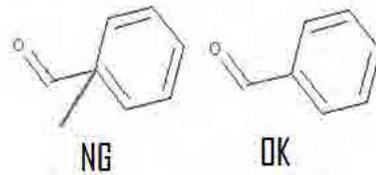
ニューロシンボリックAIの応用

Neuro-Symbolic Generative Models (生成モデル)



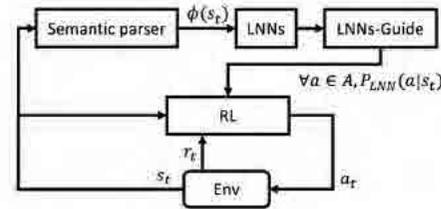
Srivastava et al. 2020 (submitted)

Neuro-Symbolic Material Discovery (物質発見)



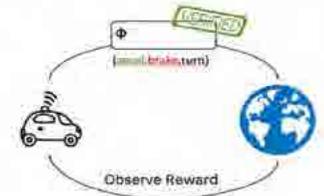
Kajino ICML 2019

Neuro-Symbolic Reinforcement Learning (強化学習)



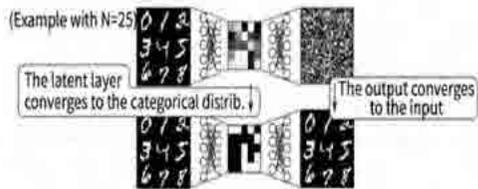
Kimura et al. IJCAIw 2020

Neuro-Symbolic Safe ML/RL (安全な強化学習)



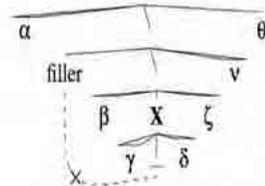
Fulton et al. AAAI 2018

Neuro-Symbolic Planning (学習データの用意)



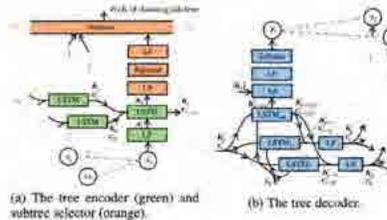
Asai et al. AAAI 2018

Neuro-Symbolic NLU (自然言語理解)



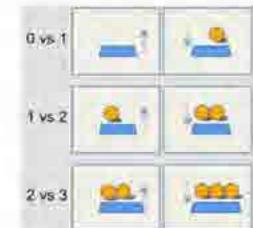
Wilcox et al. NAACL 2019

Neuro-Symbolic Code Optimization (コードの最適化)



Shi et al. ICLR 2019

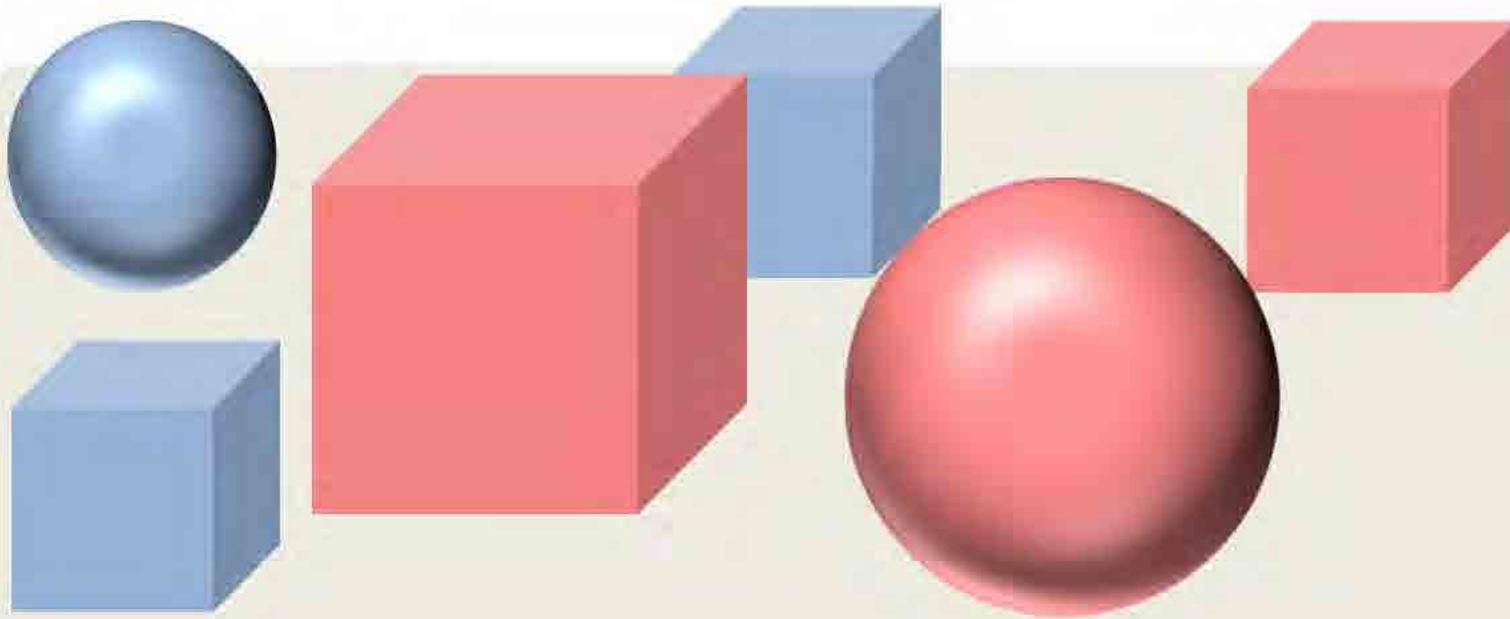
Neuro-Symbolic Machine Common Sense (常識のあるマシン)



Smith et al. NeurIPS 2019

ニューロシンボリックAIの具体例

画像を用いた質問回答

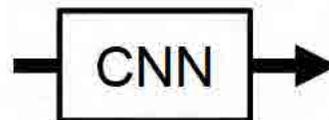
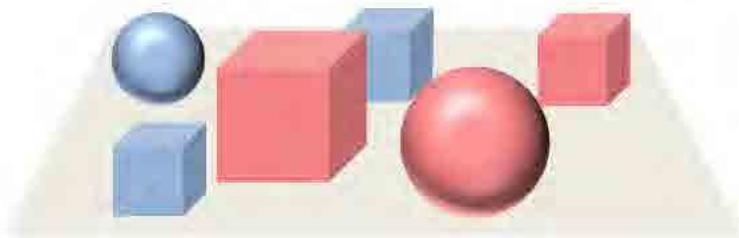


Q. How many pink blocks are to the left of the big sphere?
(大きな球より左に、ピンクのブロックはいくつある?)

Kexin Yi et al., "Neural-Symbolic VQA: Disentangling Reasoning from Vision and Language Understanding," NIPS 2018.

ニューロシンボリックAIの具体例

画像を用いた質問回答



ID	x, y	Shape	Color	Size
1	(11, 43)	Sphere	Blue	Small
2	(13, 7)	Cube	Blue	Small
3	(23, 26)	Cube	Pink	Big
⋮				

Q. How many pink blocks are to the left of the big sphere?



count
filter_color_target
filter_shape_target
relationship
filter_size_ref
filter_shape_ref

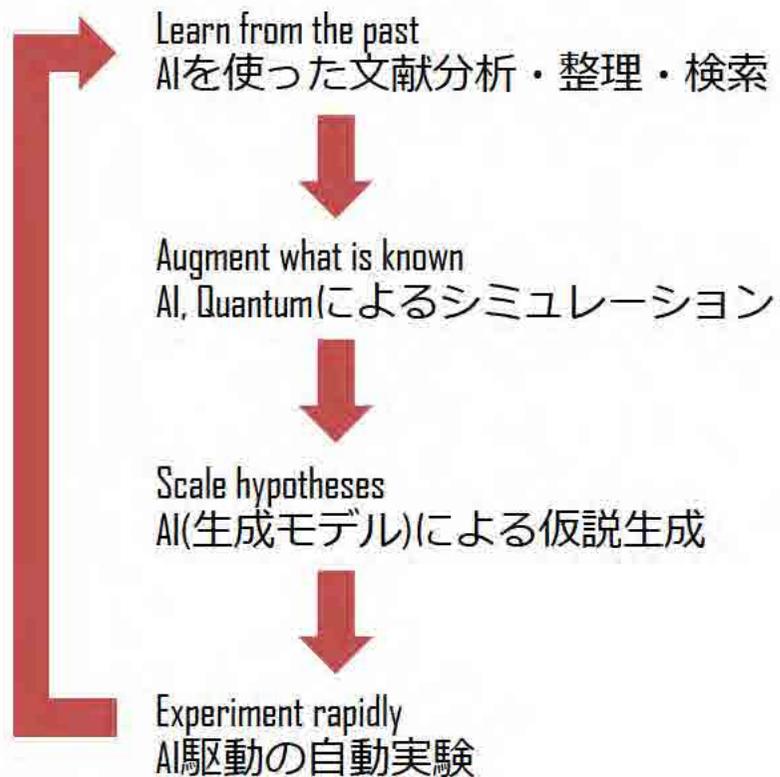


1

(AIの応用分野)
科学の緊急性
発見の加速

Using AI to Accelerate Scientific Discovery

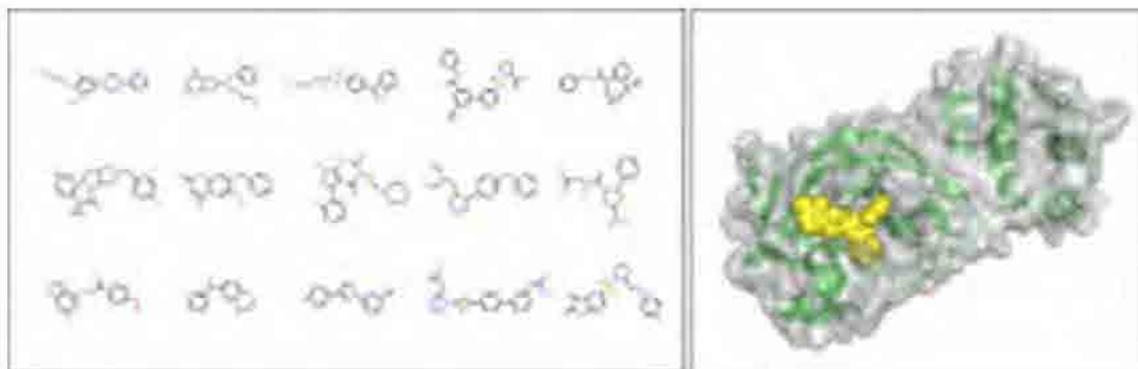
科学的発見のプロセスをAIで加速



例) AI生成モデルによる抗COVID-19薬候補の生成

Generated molecules with high SARS-CoV-2 Main Protease binding activity and selectivity.

A selected designed molecule docked to the SARS-CoV-2 Main Protease.

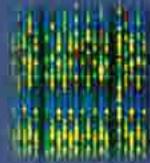


<https://www.ibm.com/blogs/think/jp-ja/urgencyofscience/>
<https://www.ibm.com/blogs/research/2020/06/accelerated-discovery/>
<https://arxiv.org/abs/2005.11248>

量子計算 x AI

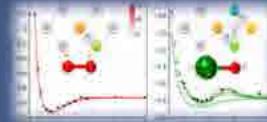
量子コンピュータの応用

創薬、ゲノム医療



- 新薬の発見、開発の迅速化
- ゲノム解析

量子化学、材料



- 発見から、必要な材料の創出へ
- 量子制御コントロール

最適化、リスク解析

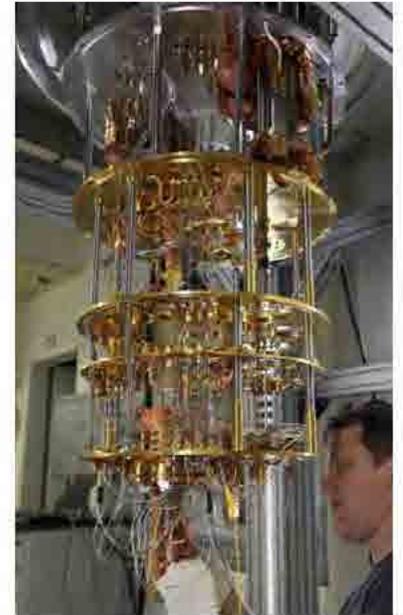


- 大規模で複雑な物流の最適化
- 金融データのリスク分析

AI / 機械学習



- 複雑度の高いニューラルネット解析
- 膨大なデータを短時間で学習



AIの課題

消費電力・バイアス・説明性・脆弱性

コンピュータ vs 人間の脳

Jeopardy! Watson

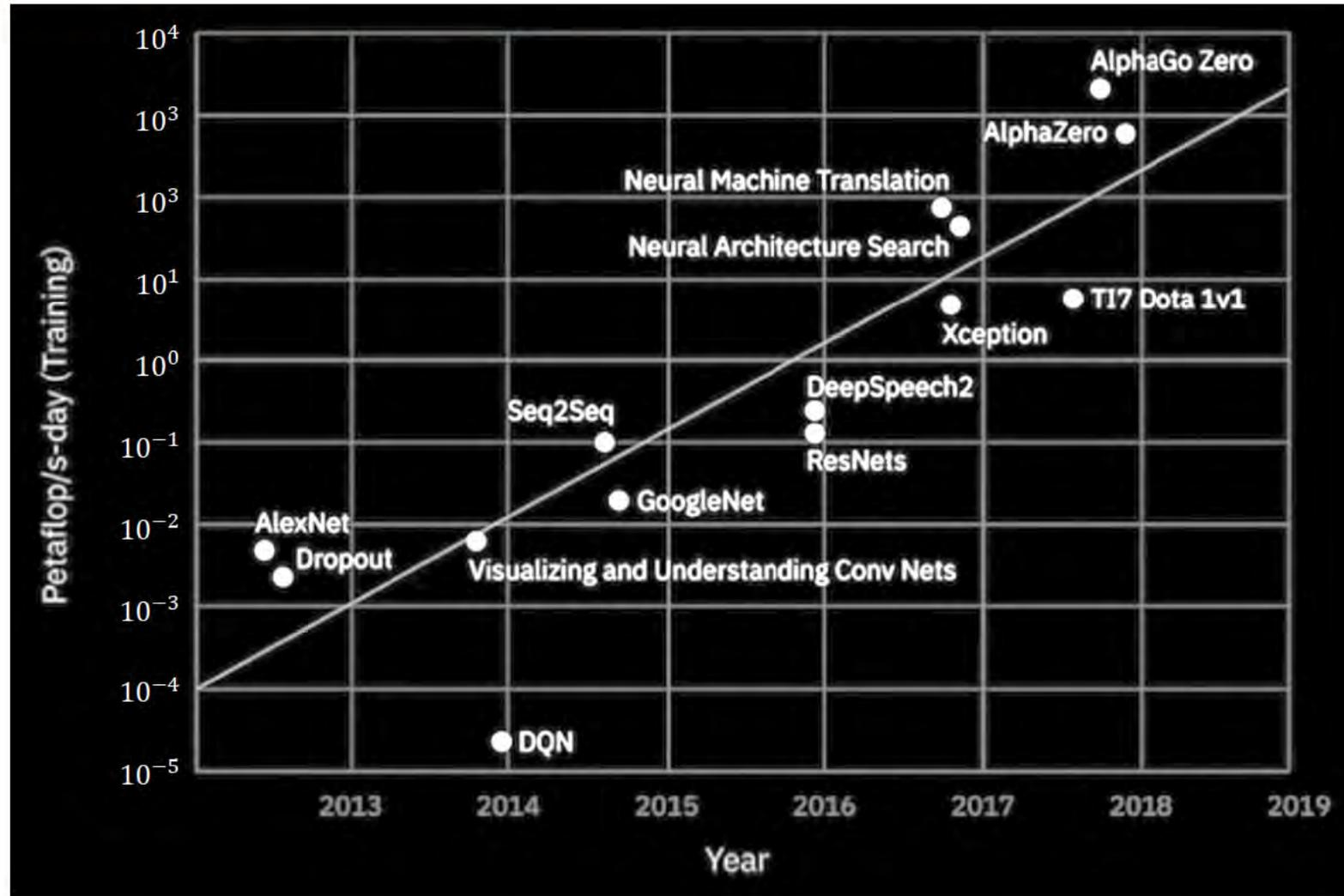


大脳のニューロン数 = 140億個
中枢神経のニューロン数 = 2,000億個
シナプス(ニューロン同士の接続)数 = 100兆個

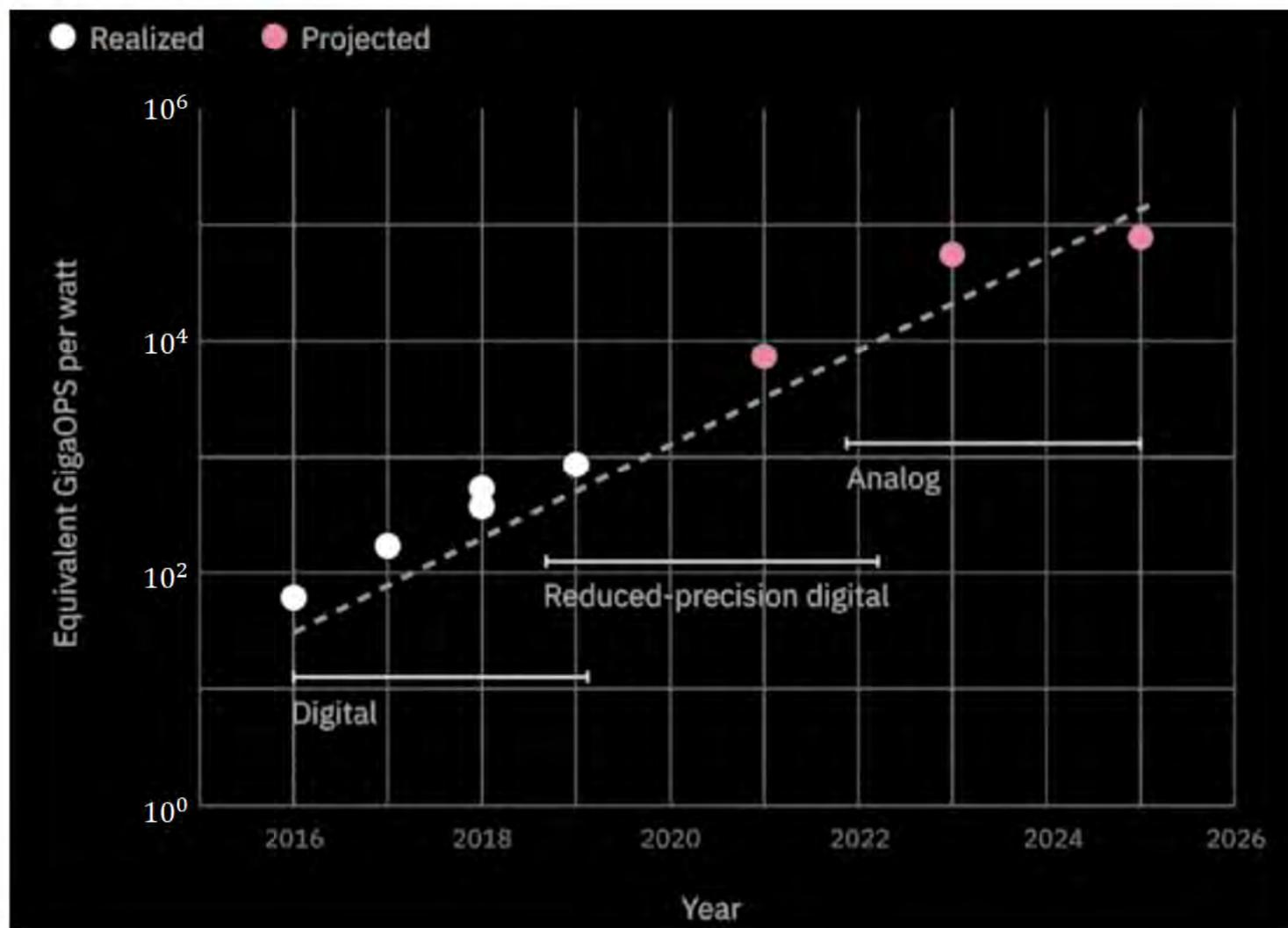


3.55 GHz	クロック	<10 Hz
200kW	消費電力	20W
12,000L	体積	1.2L

計算能力の拡大 (Compute Scaling)

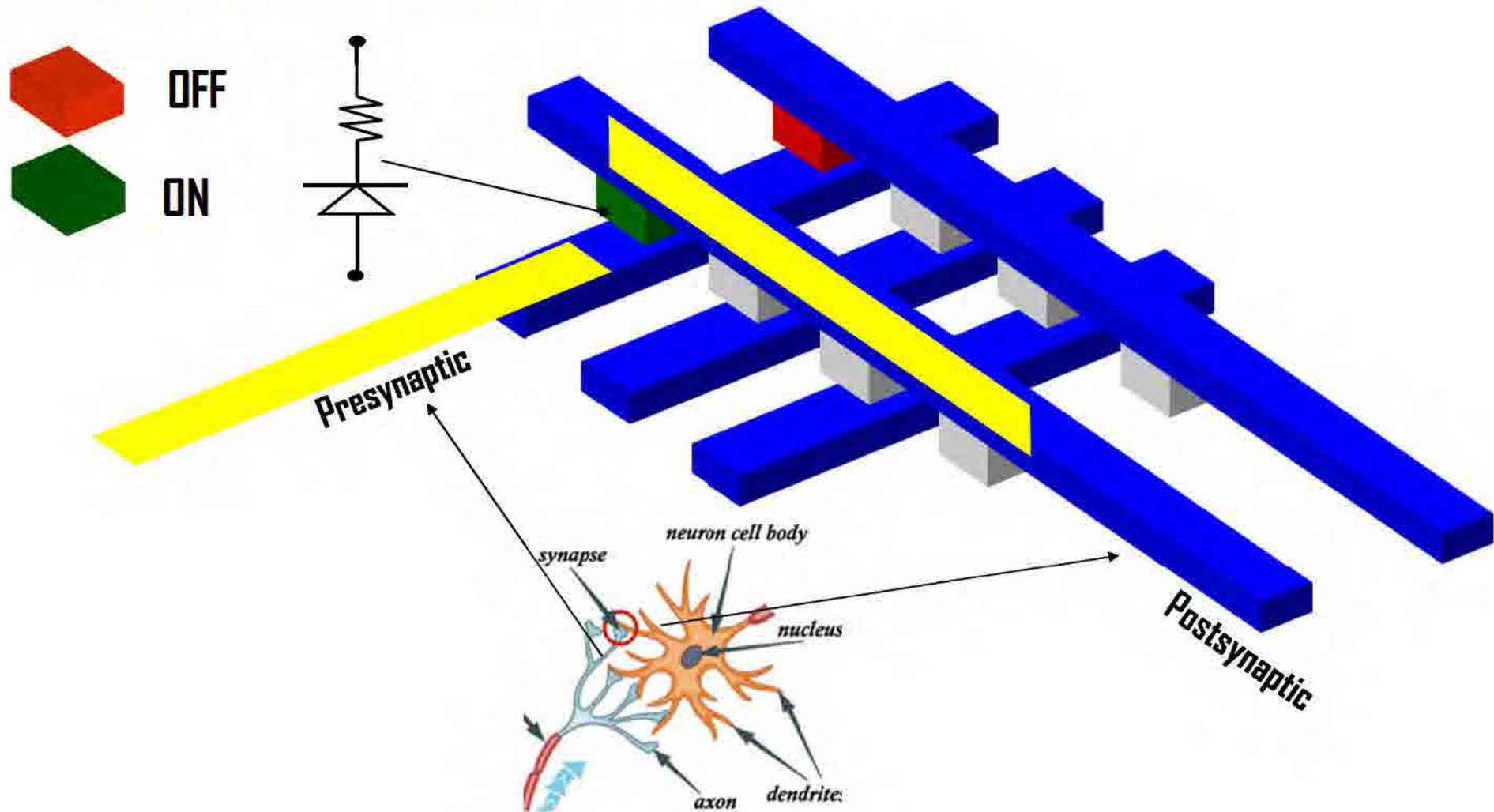


計算能力の拡大：ロードマップ



脳型コンピュータ (Neuromorphic Chip)

New technology for the next generation AI device



IBM CEO's Letter to Congress on Racial Justice Reform

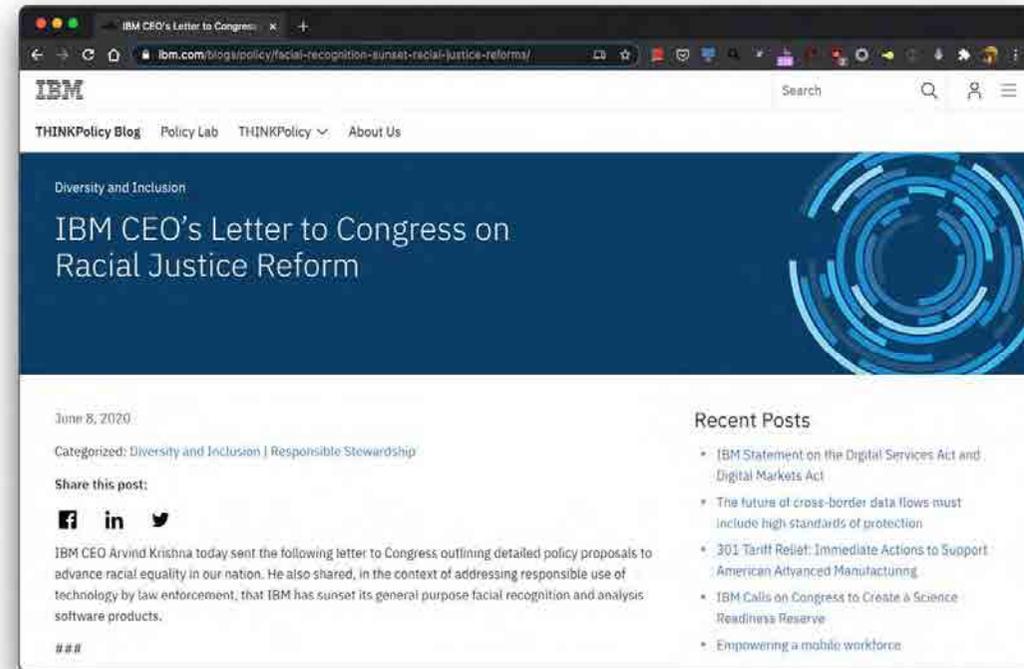
Responsible technology policies – technology can increase transparency and help police protect communities but must not promote discrimination or racial injustice.

IBM no longer offers general purpose IBM facial recognition or analysis software. IBM firmly opposes and will not condone uses of any technology, including facial recognition technology offered by other vendors, for mass surveillance, racial profiling, violations of basic human rights and freedoms, or any purpose which is not consistent with our values and [Principles of Trust and Transparency](#). We believe now is the time to begin a national dialogue on whether and how facial recognition technology should be employed by domestic law enforcement agencies.

Artificial Intelligence is a powerful tool that can help law enforcement keep citizens safe. But vendors and users of AI systems have a shared responsibility to ensure that AI is tested for bias, particularity when used in law enforcement, and that such bias testing is audited and reported.

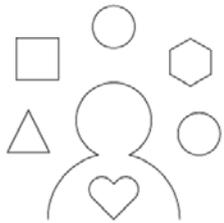
Finally, national policy also should encourage and advance uses of technology that bring greater transparency and accountability to policing, such as body cameras and modern data analytics techniques.

<https://www.ibm.com/blogs/policy/facial-recognition-sunset-racial-justice-reforms/>

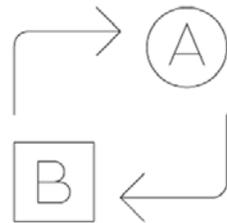


AIのバイアス (偏り)

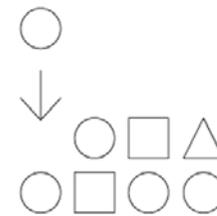
偏ったAIシステムの数は、今後増えていくだろう。
しかし、それらに対処する能力も増えていく。つまり、AIに内在するバイアスをコントロールする新しい解決策と、バイアスのないAIを推奨することは、今後の研究と規制政策の重要なエリアとなる。



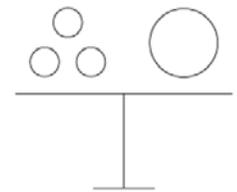
人間の持っているバイアスは180種類以上が分類・定義されている。そのどれもが我々の判断に影響を与えている。



バイアスは政府や企業が設計し使用するAIシステムに、影響を与えてしまう。



AIを学習させるために悪いデータを使うと、人種・性別・イデオロギーなどに関わる、暗黙のバイアスを生じる可能性がある。



AIシステムに含まれるバイアスは、人間と機械の間の信頼関係を損なう可能性がある。

<https://www.ibm.com/blogs/systems/tackling-bias-in-ai/>

AI Fairness 360 Toolkit & Watson OpenScale

AIの公平性・信頼性を高めるソフトウェア

AI Fairness 360

- Developers and Data Scientists
- Open Source Toolkit
- Detect and Mitigate Bias for Models in Development



<https://aif360.mybluemix.net/>

Watson OpenScale

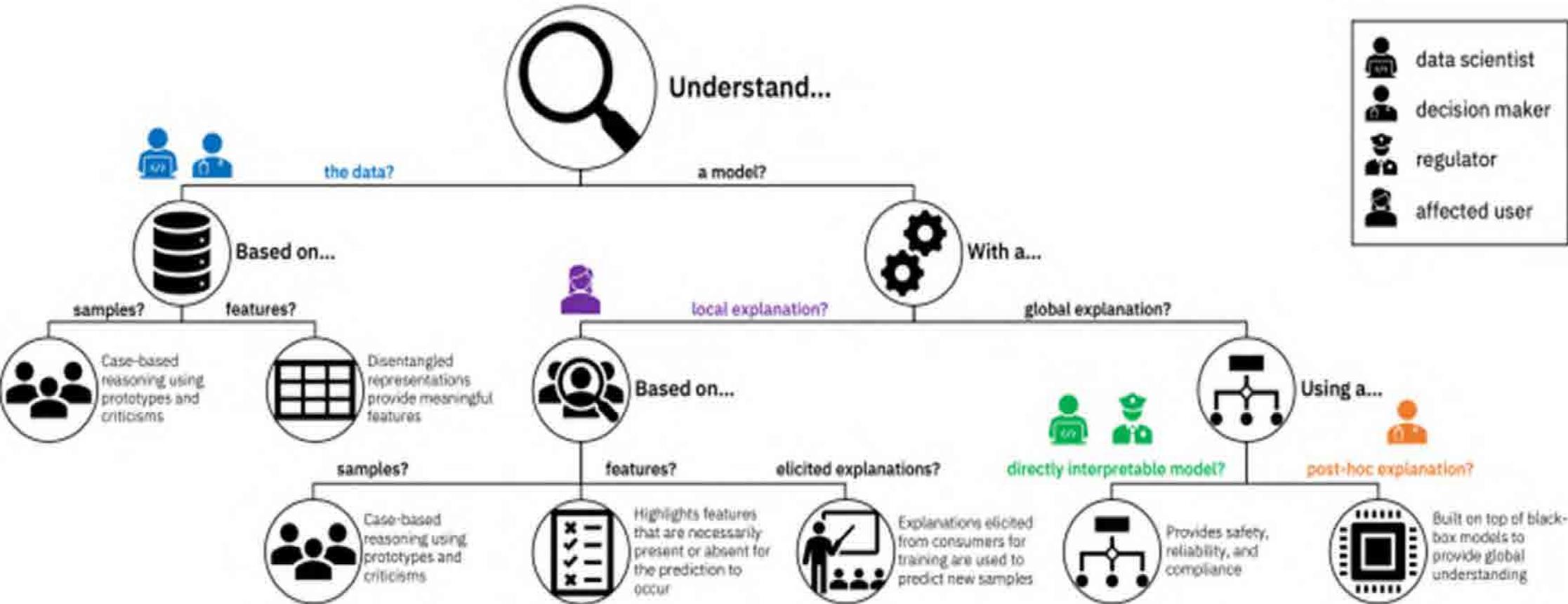
- Business Users
- Fully Supported Enterprise Tool
- Measure, Improve and Explain Production AI



<https://www.ibm.com/watson/jp-ja/watson-openscale/>

AI Explainability 360

AIの説明性を高めるオープンソース



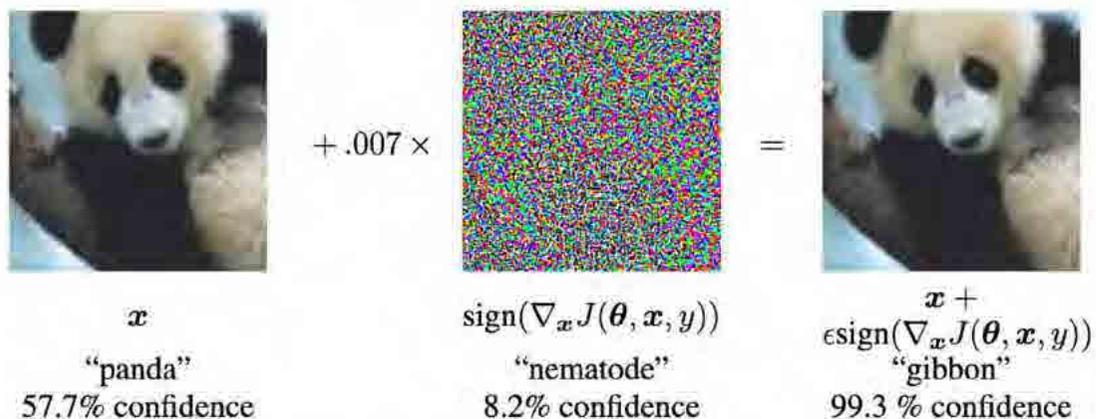
<https://www.ibm.com/blogs/research/2019/08/ai-explainability-360/>

AIの脆弱性

Adversarial Example:

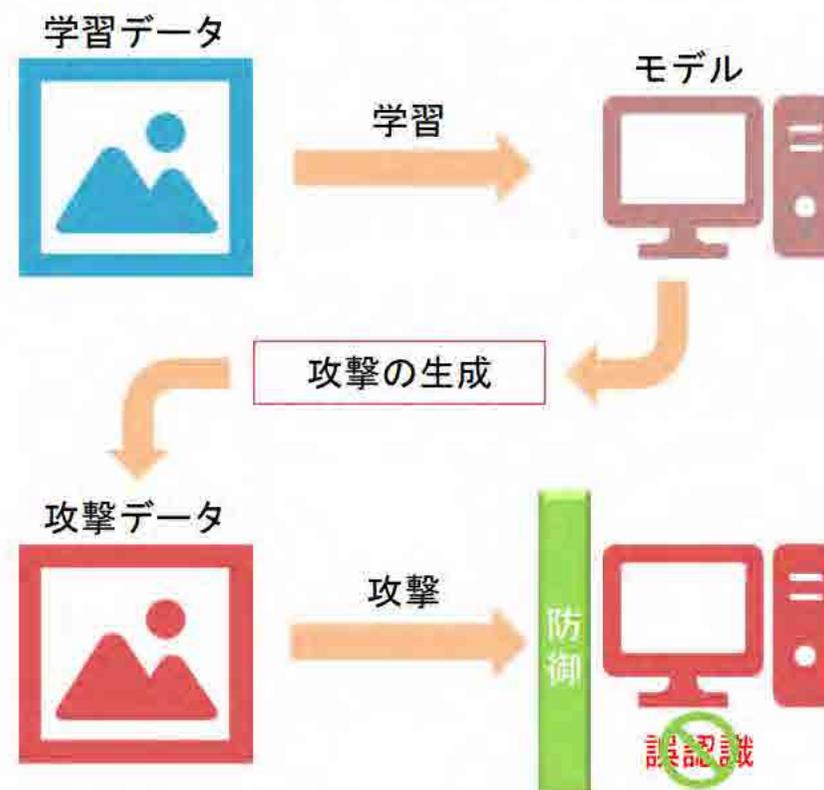
- 機械学習モデルが誤認識するように改変されたデータのこと
- 人間は改変を認識ほとんどできない
- 1つの機械学習モデルから得られたAdversarial Exampleは、別の機械学習モデルに対しても有効

例：ノイズを加えることで、パンダ→テナガザル



GOODFELLOW, I. J., SHLENS, J., AND SZEGEDY, C. Explaining and harnessing adversarial examples. arXiv preprint arXiv:1412.6572(2014).

AIモデルに対する攻撃と防御:



デモ: <https://art-demo.mybluemix.net/>

