WWW.DAUPHINE.PSL.EU

2023

Histoire de la cybernétique et de l'intelligence artificielle

Henri Isaac

henri.isaac@dauphine.psl.eu

Dauphine Recherches en Management UMR CNRS 7088

Dauphine UNIVERSITÉ PARIS | PSL

Agenda

Comprendre l’importance de la cybernétique dans la construction du monde numérique

Comprendre la construction de la notion d’intelligence artificielle et son développement

De la machine de Turing...

Alan Turing

"On Computable Numbers, with an Application to the Entscheidungsproble", [1937], Proceedings of London Mathematical Society, 2e série, vol. 42, p. 230-265

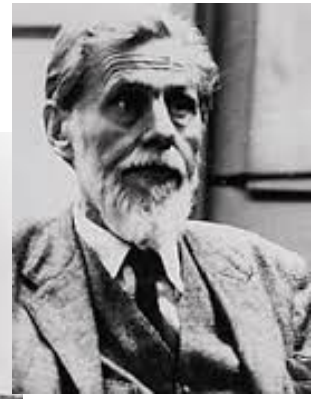
À la cybernétique de l'après-guerre…

Théorie du contrôle et de la communication Norbert Wiener

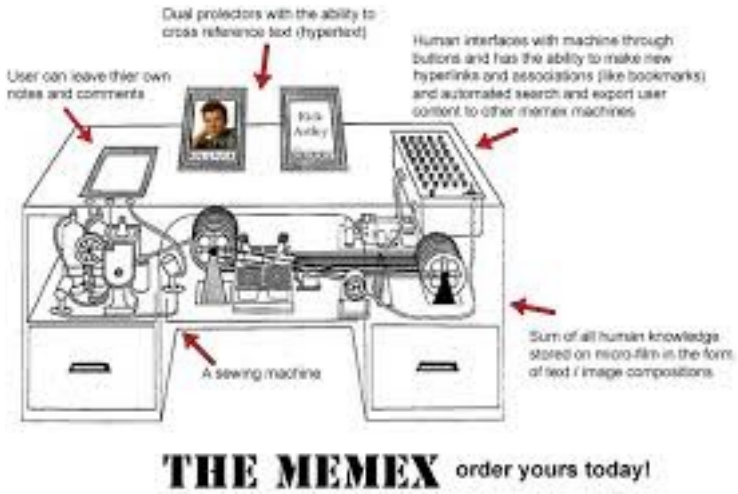
Premiers ordinateurs (EDVAC), John Von Neumann

Quantification de l'information "bits", Claude Shannon

Analogie cerveau/ordinateur, Walter Pitts, Warren Mc Culloch



...à l'Internet...

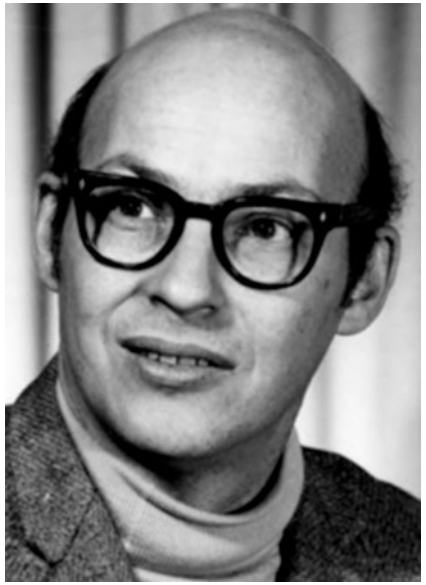


Memex
Vannevar Bush
1945

Projet Xanadu
Ted Nelson
1960

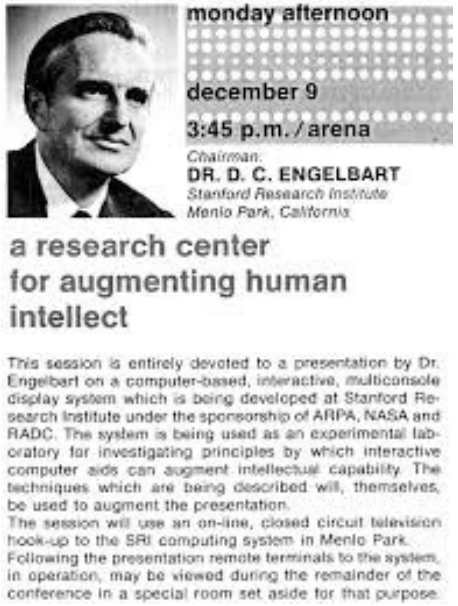
WWW / URL / HTTP / HTML
Tim Berners-Lee
1990

À l'intelligence artificielle



Intelligence Artificelle
Marvin Minsky, John MacCarthy

Perceptron
Frank Rosenblatt
1957

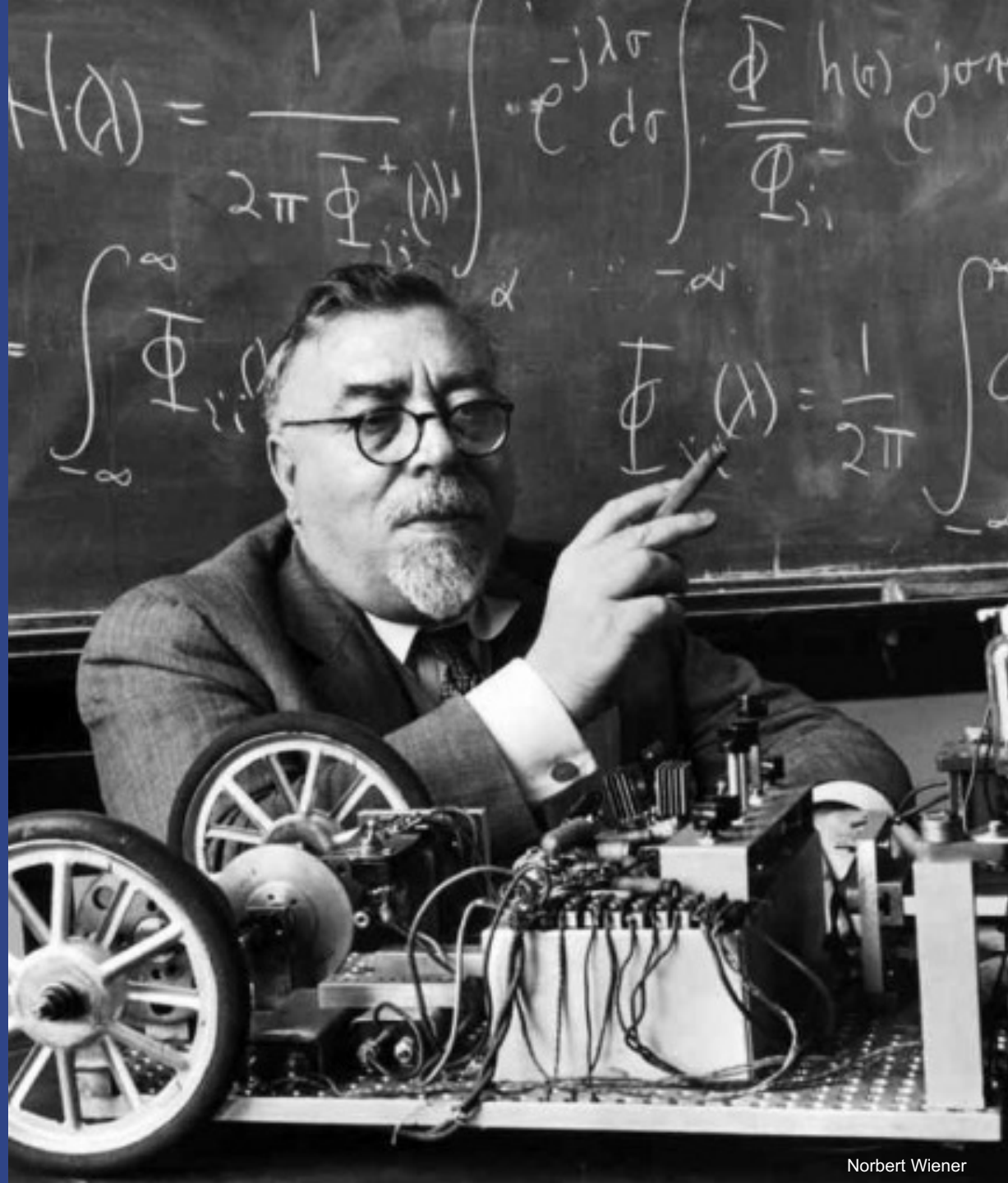


Projet OnLine System,
Doug Englebart
1968

WWW.DAUPHINE.PSL.EU

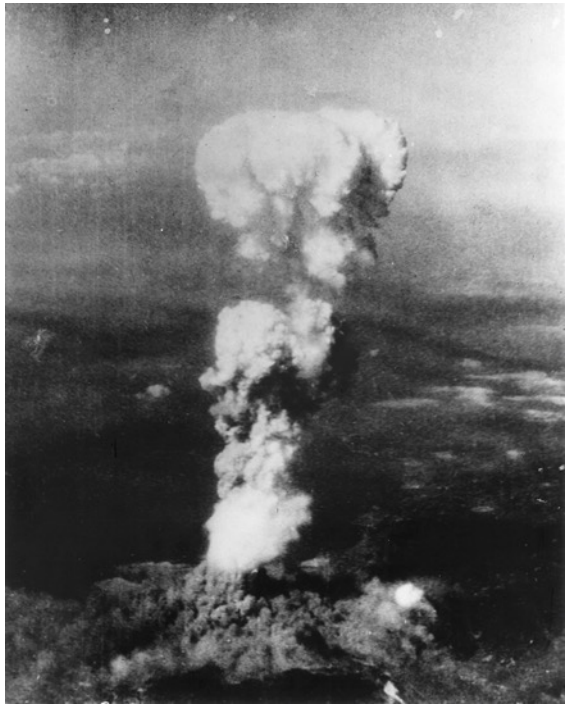
Théorie de l'information & cybernétique

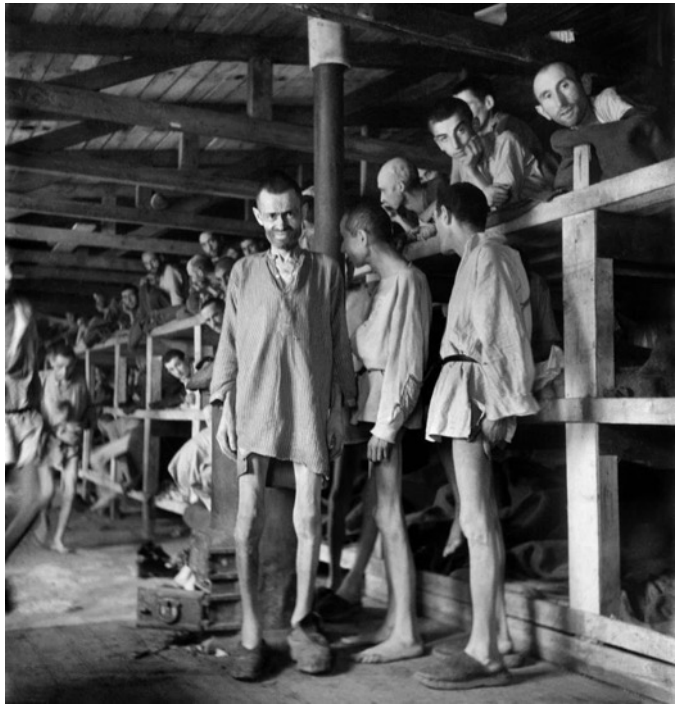
Dauphine | PSL
UNIVERSITÉ PARIS



Norbert Wiener

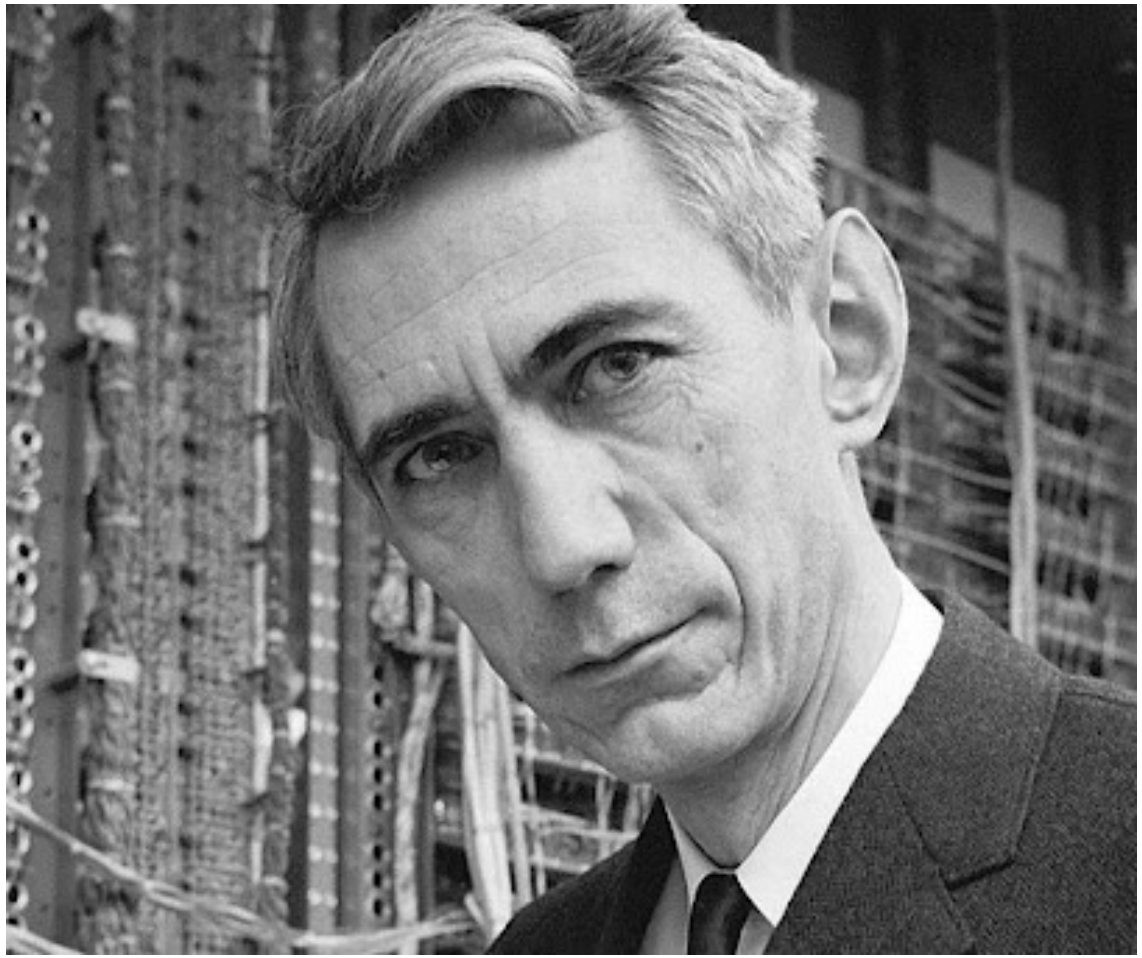
Contexte historique de l'émergence de la cybernétique





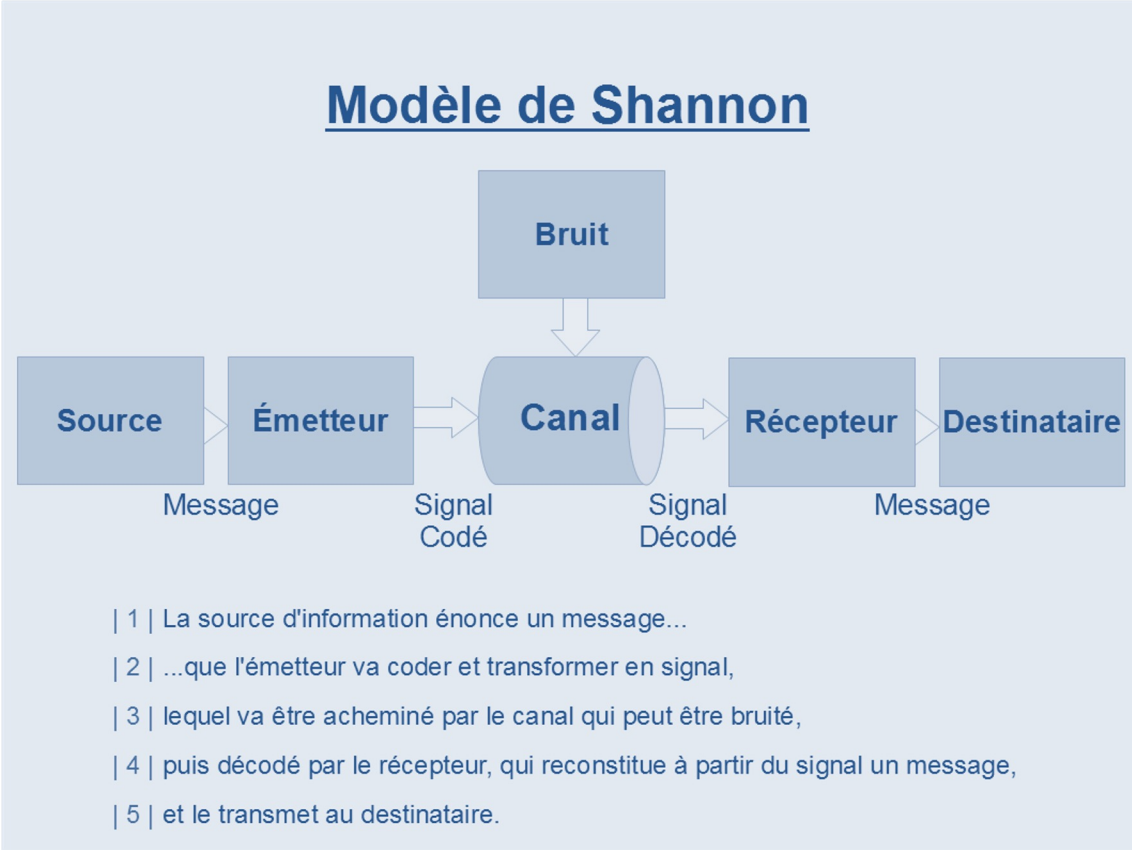
Une nouvelle théorie de l'information

- Claude Shannon (1916-2001)
  - Mathématicien (PhD MIT, 1940)
  - MIT, de 1958 à 1978.
  - Travaille parallèlement aux laboratoires Bell de 1941 à 1972
  - Produit une théorie mathématique de l'information
  - Conçu pour décrire la communication entre machines, ce schéma modélise imparfaitement la communication humaine



Théorie de l'information

- Claude E. Shannon, (1948), A Mathematical Theory of Communication, *Bell System Technical Journal*, vol. 27, p. 379-423 & 623-656, July and October.
- Base de tous les systèmes numériques de communication
- L'information est quantifiable: bit (0,1)
- Le contexte de l'émetteur et du récepteur ne sont pas pris en compte dans un tel modèle
- Shannon est parfaitement conscient des limites de son modèle



Cybernétique

- Le terme cybernétique apparaît en 1834 et désigne « *la science du gouvernement des hommes* »
- Wiener déclare avoir fait dériver le mot cybernétique « *du mot grec kubernetes, ou pilote, le même mot grec dont nous faisons notre mot gouverneur* »
- Heinz von Foerster propose le terme cybernétique pour coiffer l'ensemble des dix rencontres de Macy

Les conférences de Macy (1942-1953)

https://asc-cybernetics.org/foundations/history/Macy10Photo.htm

Les origines de la cybernétique

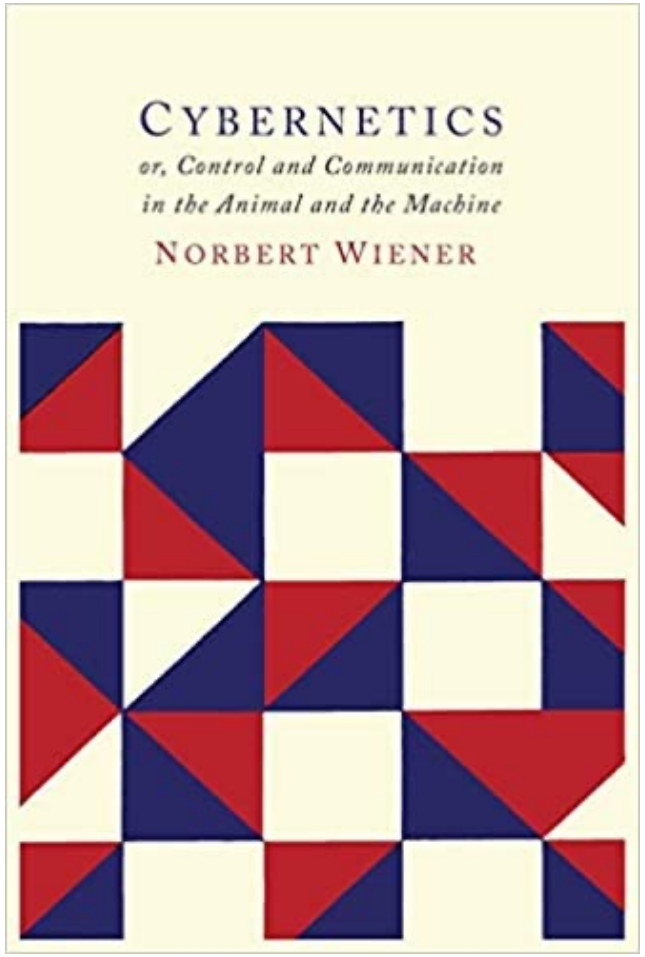
- De nombreuses origines philosophiques et scientifiques
- Historiquement, les conférences organisées par la Fondation Macy
  - Fondation américaine dédiée à la recherche médicale
  - Une première conférence en 1942 qui débouche sur une publication:
    - Rosenblueth, A., Wiener, N., and J. Bigelow, "Behavior, purpose and teleology", *Philosophy of Science*, Vol. 10 (1943), pp. 18 - 24.
    - Discussion porte sur : « *a conceptual agenda based on similarities between behaviors of both machines and organisms that were interpretable as being 'goal-directed'* »
    - Sixième conférence : « *Circular Causal and Feedback Mechanisms in Biological and Social Systems.* »
  - Dix conférences entre 1946 et 1953
    - Voir :
      - Steven Joshua Heims, (1991), *Constructing a Social Science for Postwar America: The Cybernetics Group,* MIT Press.
      - Jean-Pierre Dupuy (2000), *The Mechanization of the Mind,* Princeton, University Press.

Emergence de la cybernétique comme champ scientifique

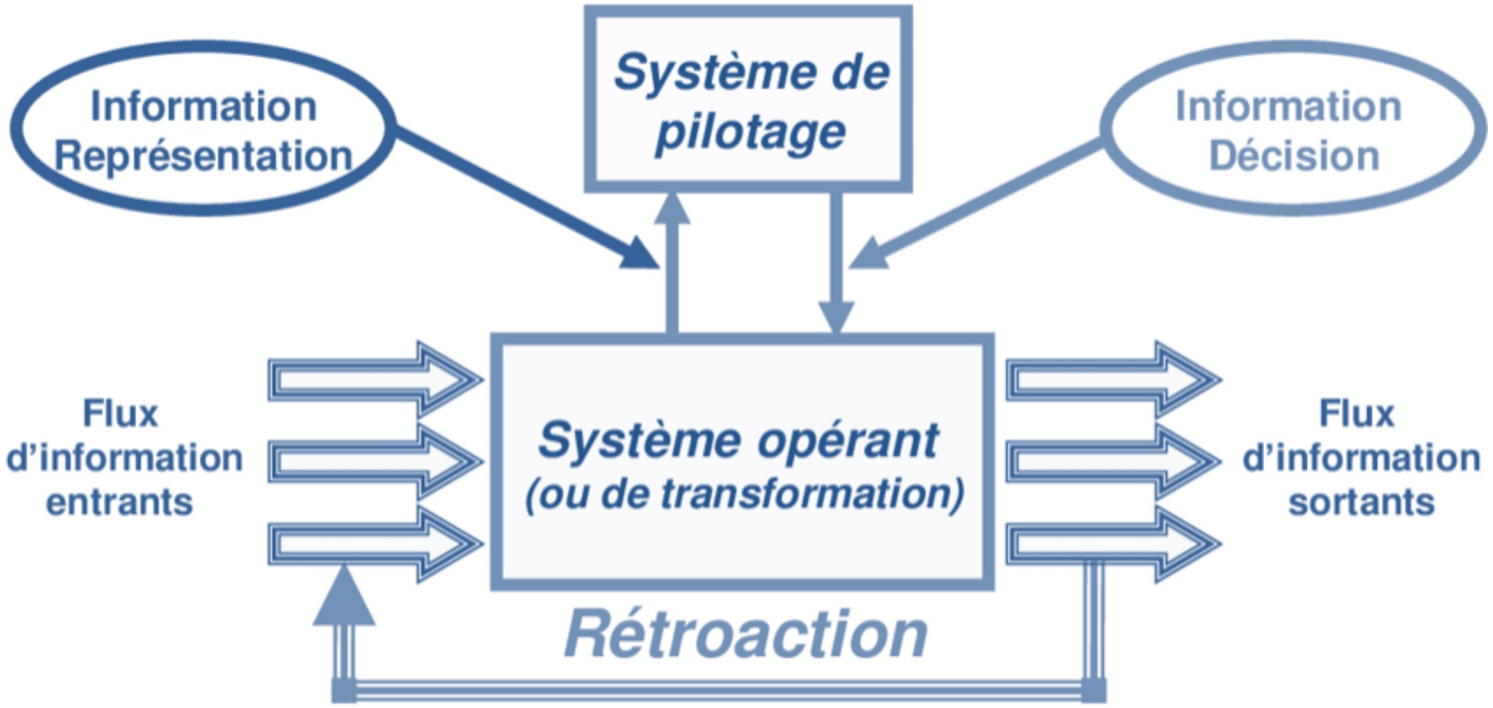
- La cybernétique est l'étude des mécanismes d'information des systèmes complexes, décrits en 1947 par Norbert Wiener.
- Des scientifiques d'horizons très divers participèrent à ce projet interdisciplinaire de 1942 à 1953 : mathématiciens, logiciens, ingénieurs, physiologistes, anthropologues, psychologues, etc.
- Les contours parfois flous de cet ensemble de recherches s'articulent toutefois autour du concept clé de **rétroaction** (*feedback*) ou mécanisme téléologique.
- Objectif: donner une vision unifiée des domaines naissants de l'automatique, de l'électronique et de la théorie mathématique de l'information, en tant que « *théorie entière de la commande et de la communication, aussi bien chez l'animal que dans la machine* ».

Norbert Wiener et la cybernétique

- Wiener définit la cybernétique comme une science qui étudie exclusivement les communications et leurs régulations dans les systèmes naturels et artificiels
- “De même que l'entropie est une mesure de désorganisation, l'information fournie par une série de messages est une mesure d'organisation”
- Après avoir travaillé au développement d'appareils de pointage automatique pour canons antiaériens, il en arrive à la conclusion que « *pour contrôler une action finalisée (orientée vers un but), la circulation de l'information nécessaire à ce contrôle doit former une boucle fermée permettant d'évaluer les effets de ses actions et de s'adapter à une conduite future grâce aux performances passées* »

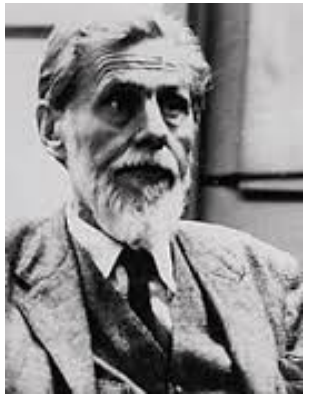


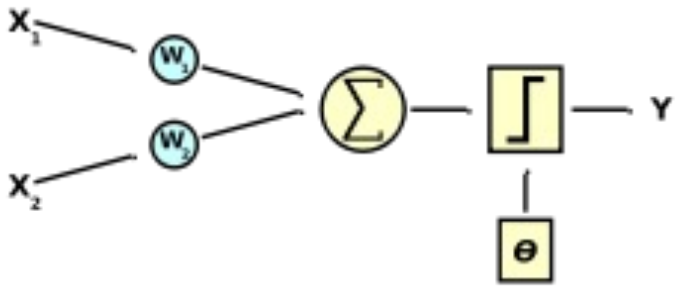
La logique cybernétique



Aux origines du *deep learning* fondé sur les réseaux de neurones

- McCulloch, W. S., Pitts, W., (1943), *A Logical Calculus of the Ideas Immanent in Nervous Activity*, Bulletin of Mathematical Biophysics, vol. 5, pp. 115-133,.
  - Analogie cerveau humain et machine
  - Premier modèle mathématique et informatique du neurone biologique.
  - Chaque neurone est caractérisé comme étant ouvert ou fermé et a la capacité de s'ouvrir en réponse à une stimulation par un nombre suffisant de neurones voisins afin de transmettre un signal.





Quelles conséquences du « moment Macy »?

- L'information est signal qui peut-être traité sans en connaître le sens, ni les contextes spécifiques des émetteurs-récepteurs
- L'information est quantifiable (bit)
- Analogie cerveau humain-machine
- Débouche sur:
  - « la société de l'information »
  - Les théories systémiques (von Bertalanffy, Varela, Lemoigne) et la notion de système d'information
  - La notion d'apprentissage machine et l'intelligence artificielle
  - La robotique

WWW.DAUPHINE.PSL.EU

Le développement de l'« Intelligence Artificielle »

Dauphine | PSL
UNIVERSITÉ PARIS

Frank Rosenblatt invented the perceptron, the first artificial neural network.
CORNELL UNIVERSITY DIVISION OF RARE AND MANUSCRIPT COLLECTIONS

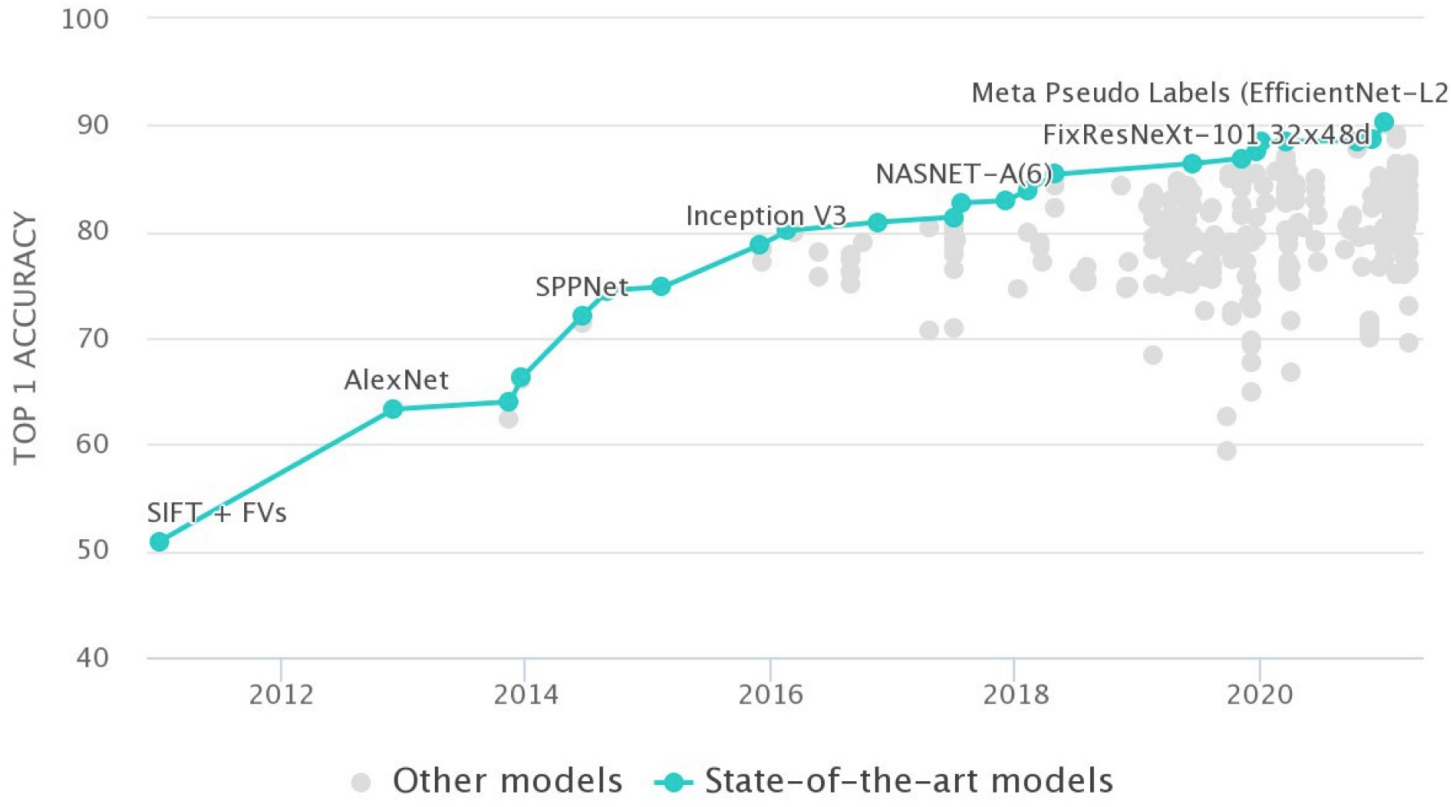
Progression de la fiabilité des modèles en *computer vision*

Dataset:

Train dataset: **1,2 Millions** d'images labellisées

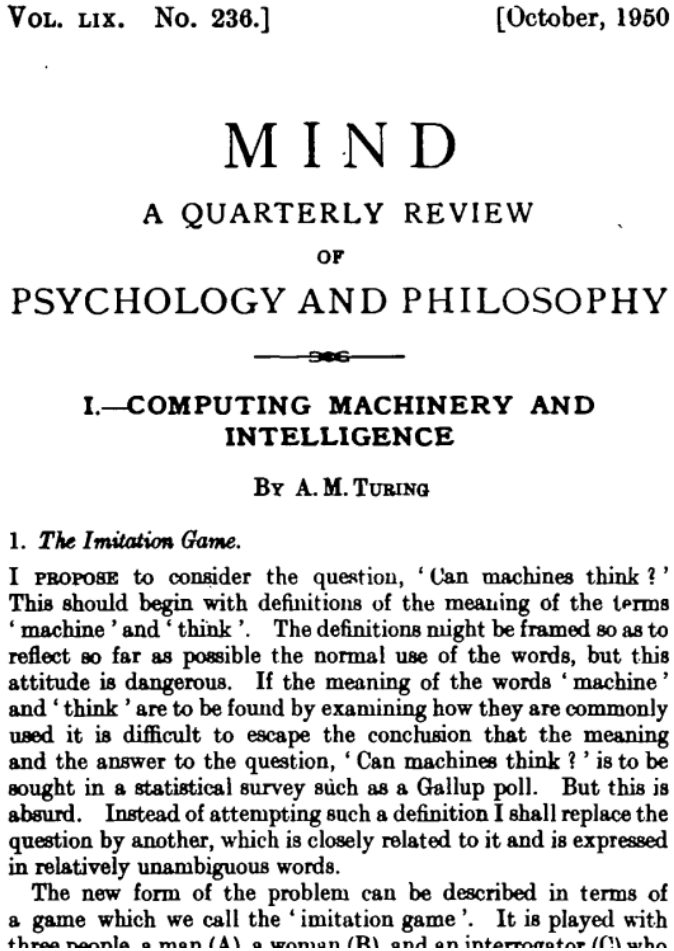
Validation dataset: **50 000** images labellisées

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



Intelligence artificielle selon Alan Turing

- « *Computing Machinery and Intelligence* » (Mind, octobre 1950), Turing explore le problème de l'intelligence artificielle et propose une expérience maintenant connue sous le nom de test de Turing, où il tente de définir une épreuve permettant de qualifier une machine de « consciente » ;
- Turing fait le « *pari que d'ici cinquante ans, il n'y aura plus moyen de distinguer les réponses données par un h*omme ou un ordinateur, et ce sur n'importe quel sujet »
- Nombreux sont ceux qui considèrent désormais que ce test ne prouve pas l'existence d'une intelligence artificielle,
- voir Rohit Prasad (Alexa), *The Turing Test is obsolete. It's time to build a new barometer for AI, 28/12/2020,* https://www.fastcompany.com/90590042/turing-test-obsolete-ai-benchmark-amazon-alexa



VOL. LIX. No. 236.] [October, 1950

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

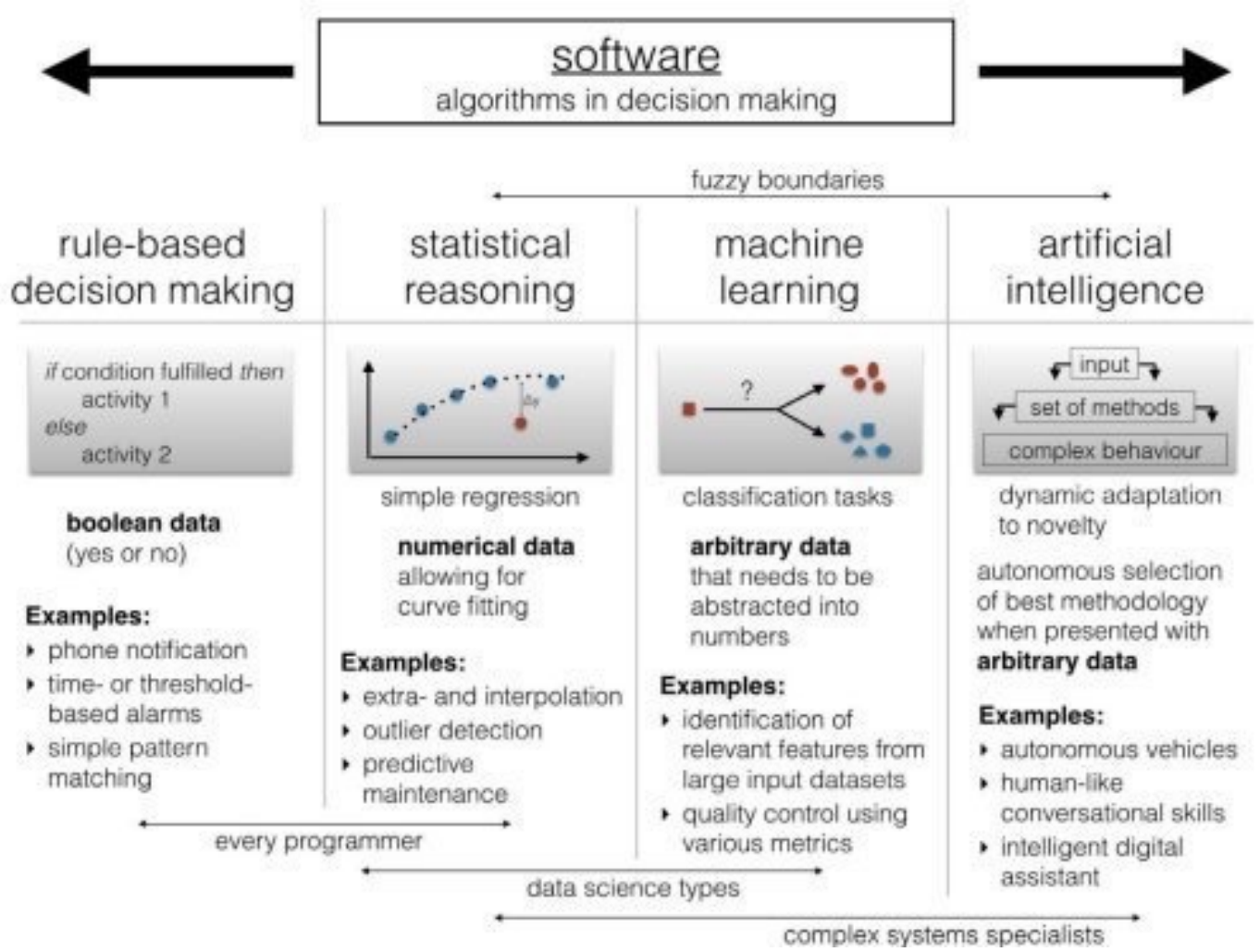
BY A. M. TURING

1. *The Imitation Game.*

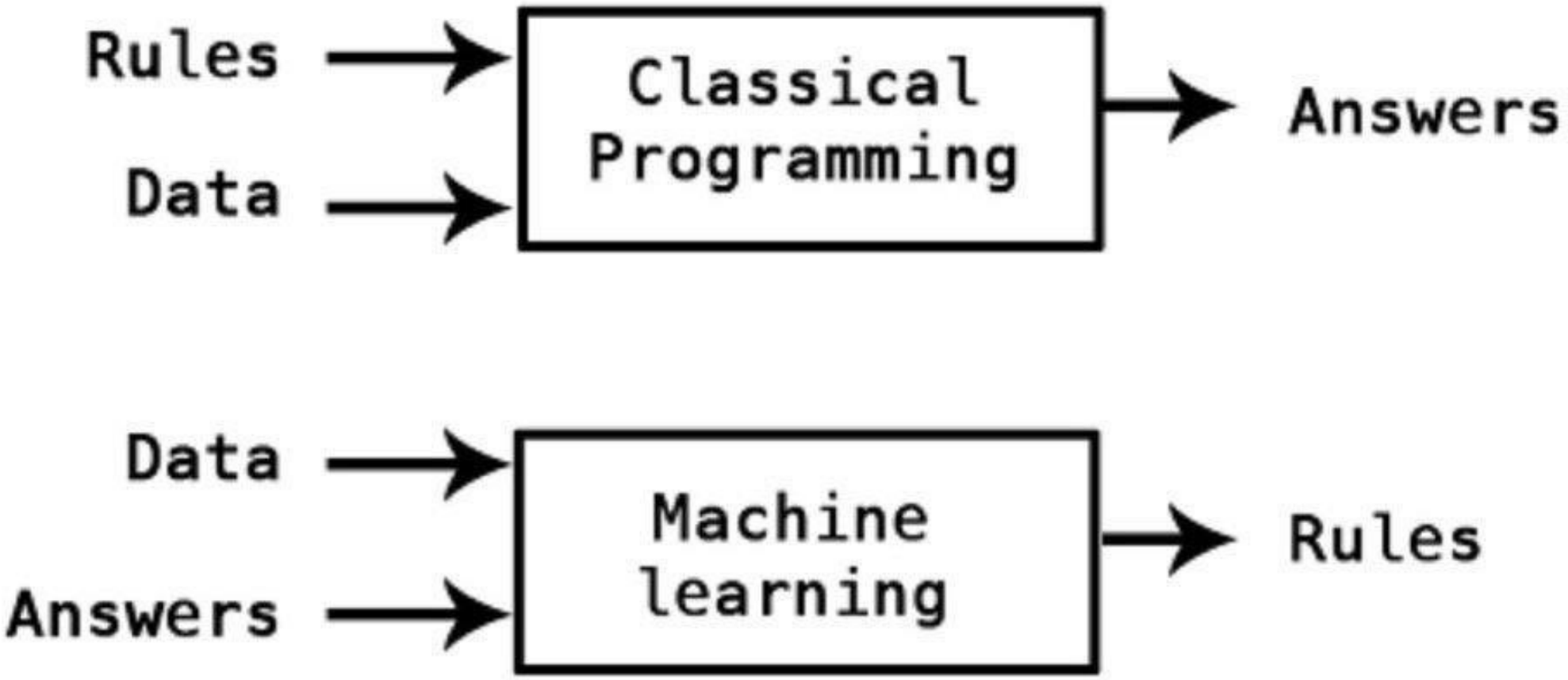
I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the 'imitation game'. It is played with three people, a man (A), a woman (B), and an interrogator (C) who

Statistiques, algorithmie, data sciences, IA

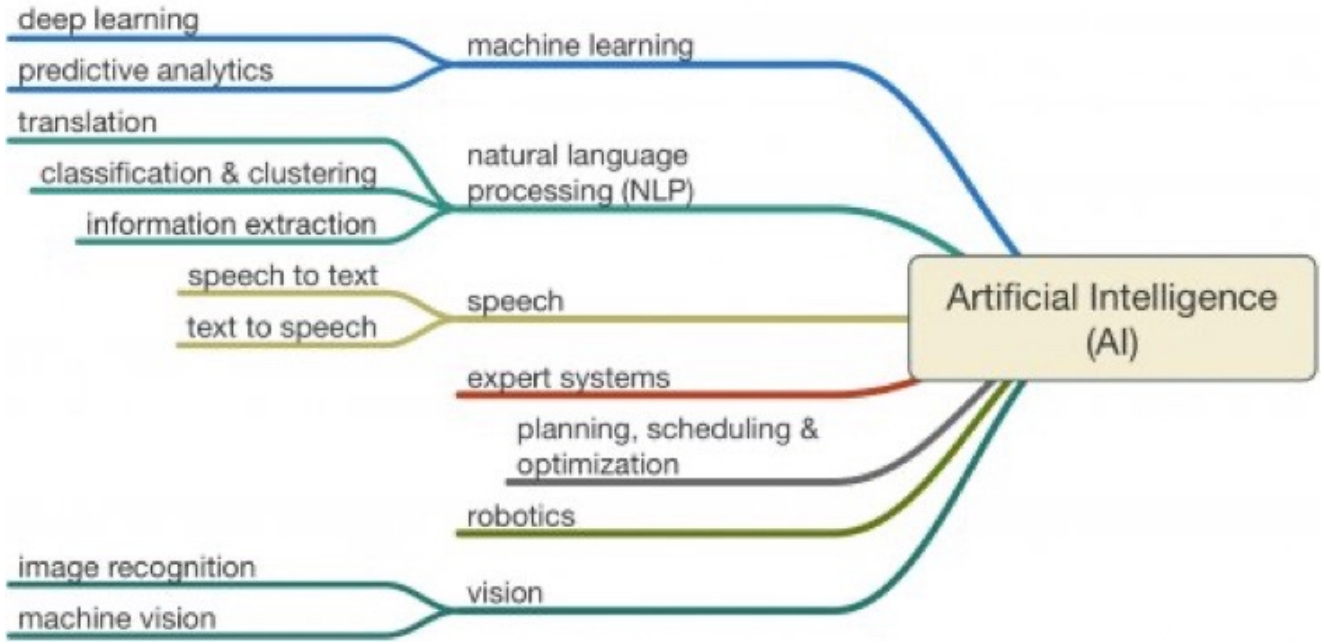


Logique du Machine Learning

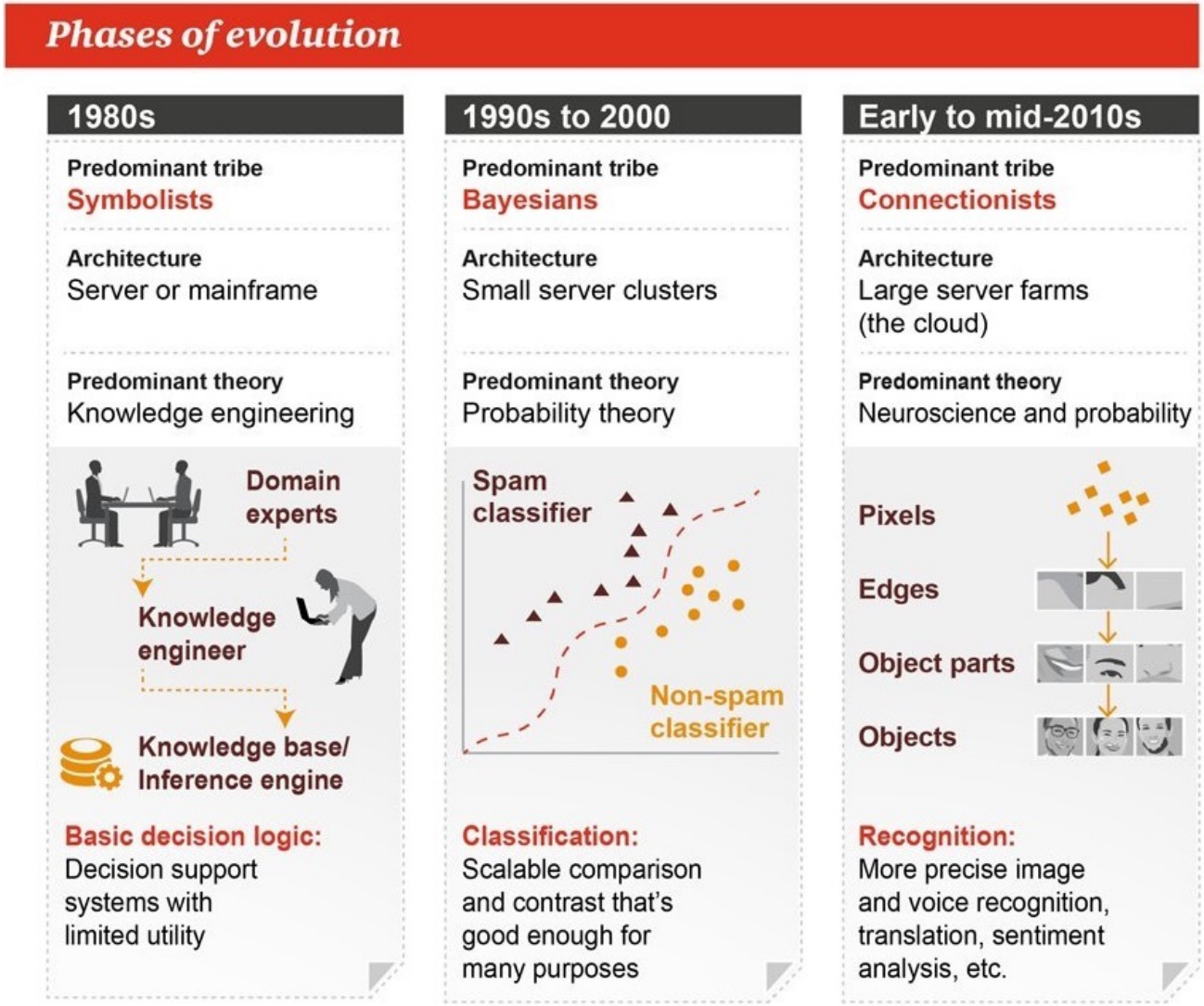


Intelligence artificielle

- Comment définir l'IA?
  - Artificial intelligence (AI) is intelligence demonstrated by **machines**, as opposed to the natural intelligence displayed by animals including humans.
  - Study of "intelligent agents": any system that perceives its environment and takes actions that maximize its chance of achieving its goals
- L'IA comporte de nombreux courants et sous-champs disciplinaires

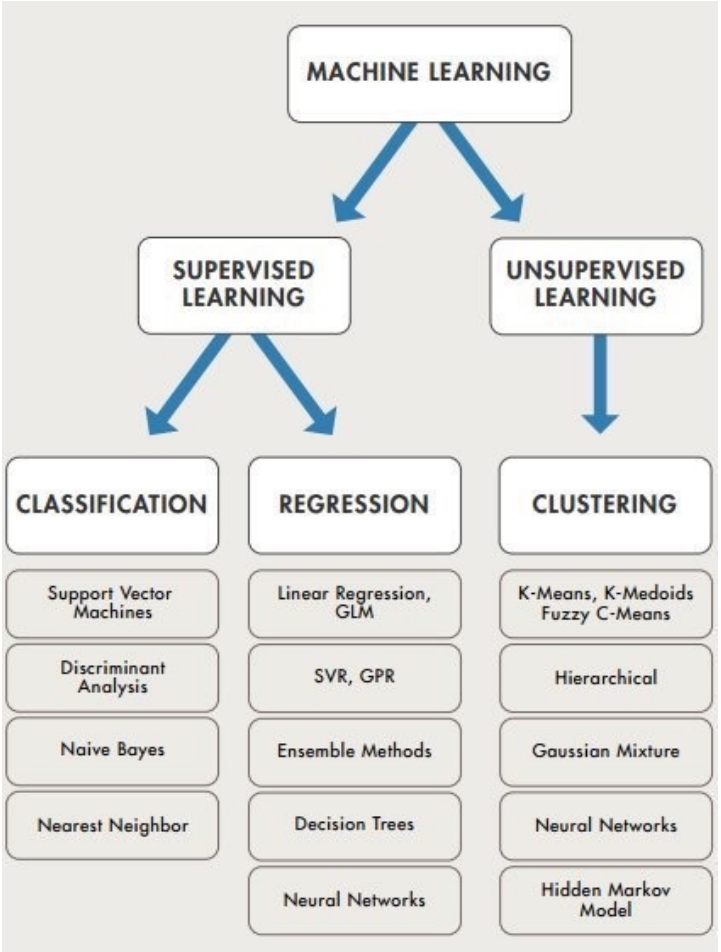


Une representation simplifiée des grands courants

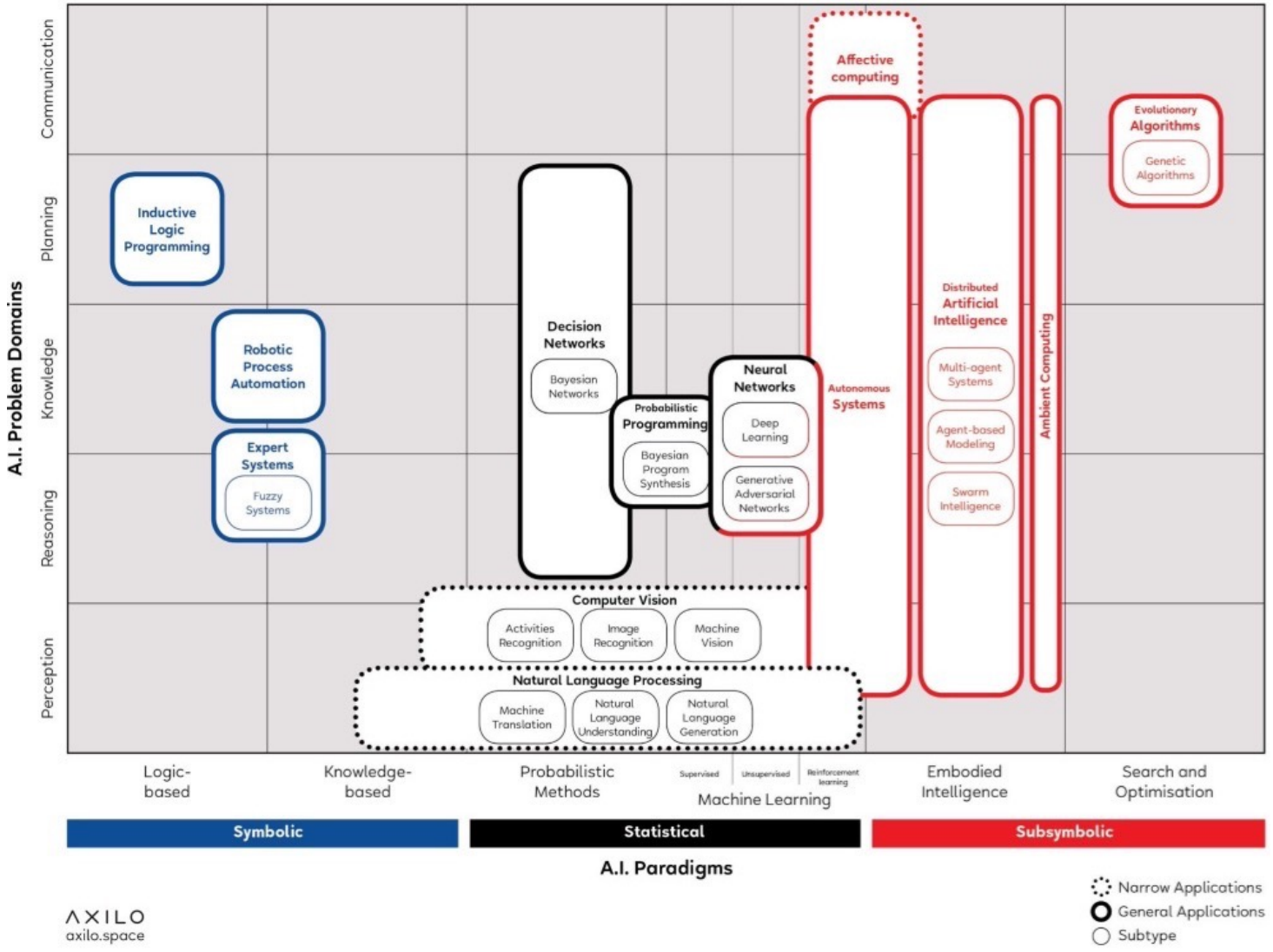


Liens statistiques, données, IA: le développement du machine learning

- Arthur Samuel qui utilise pour la première fois le terme « *machine learning* », pour son programme créé en 1952. Le programme jouait au Jeu de Dames et s'améliorait en jouant. Il parvint à battre le 4ème meilleur joueur des États-Unis.
- L'apprentissage automatique (AA) permet à un système piloté ou assisté par ordinateur comme un programme, une IA ou un robot, d'adapter ses réponses ou comportements aux situations rencontrées, en se fondant sur l'analyse de données empiriques passées issues de bases de données, de capteurs, ou du web
- Méthodes qui reposent sur la disponibilité de jeux de données pour entrainer une machine (algorithme, robot)
- Dépendance aux méta-données pour l'apprentissage surpervisé
- Dépendances aux données pour toutes les approaches ML : biais



Une representation des différentes méthodes dites d'intelligence artificielle

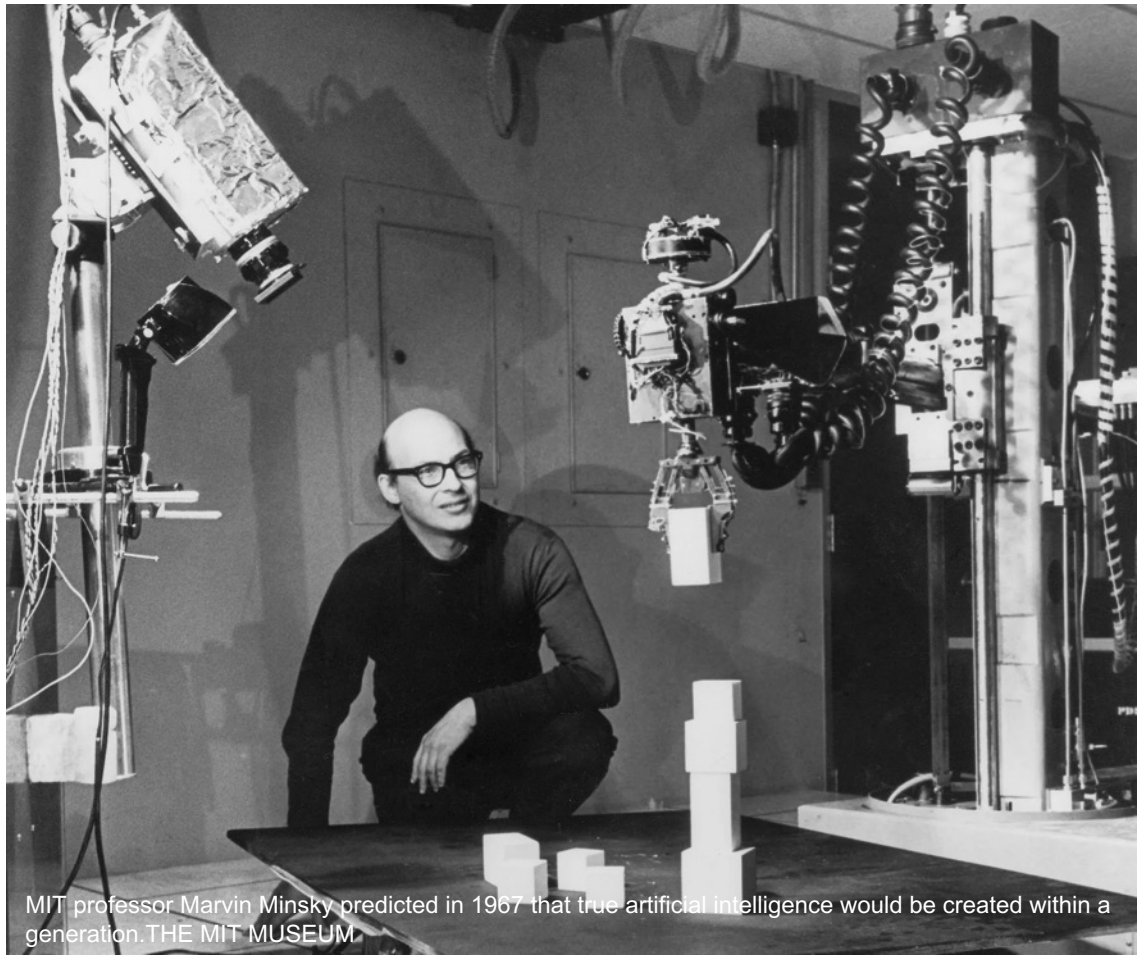


Les débuts de l'intelligence artificielle

- Débute en 1956 grâce à workshop auquel participent: J. McCarthy (Dartmouth), Minsky (Princeton), C. Shannon (Bell Labs/MIT), N. Rochester (IBM), T. More (Princeton), A. Newell (Carnegie Tech), H. Simon (Carnegie Tech), A. Samuel (IBM), R. Solomonoff (MIT), O. Selfridge (MIT).
- L'atelier est organisé par John McCarthy et Claude Shannon
- Moment fondateur de l'intelligence artificielle en tant que discipline théorique indépendante (de l'informatique)
- thèse de la conférence : « *chaque aspect de l'apprentissage ou toute autre caractéristique de l'intelligence peut être si précisément décrit qu'une machine peut être conçue pour le simuler* »

IA symbolique versus neuronale

- Marvin Minsky, avec John McCarthy, fondent le laboratoire d'IA du MIT en 1959.
- En 1969, dans *Perceptrons*, coécrit avec Seymour Papert critique Frank Rosenblatt,
- F. Rosenblatt (1958), *The perceptron: a probabilistic model for information storage and organization in the brain*
- Montre les limites des réseaux de neurones de type perceptron, notamment l'impossibilité de traiter des problèmes non linéaires ou de connexité.
- Conséquence: draine l'essentiel des crédits de recherche vers l'intelligence artificielle symbolique



MIT professor Marvin Minsky predicted in 1967 that true artificial intelligence would be created within a generation.THE MIT MUSEUM

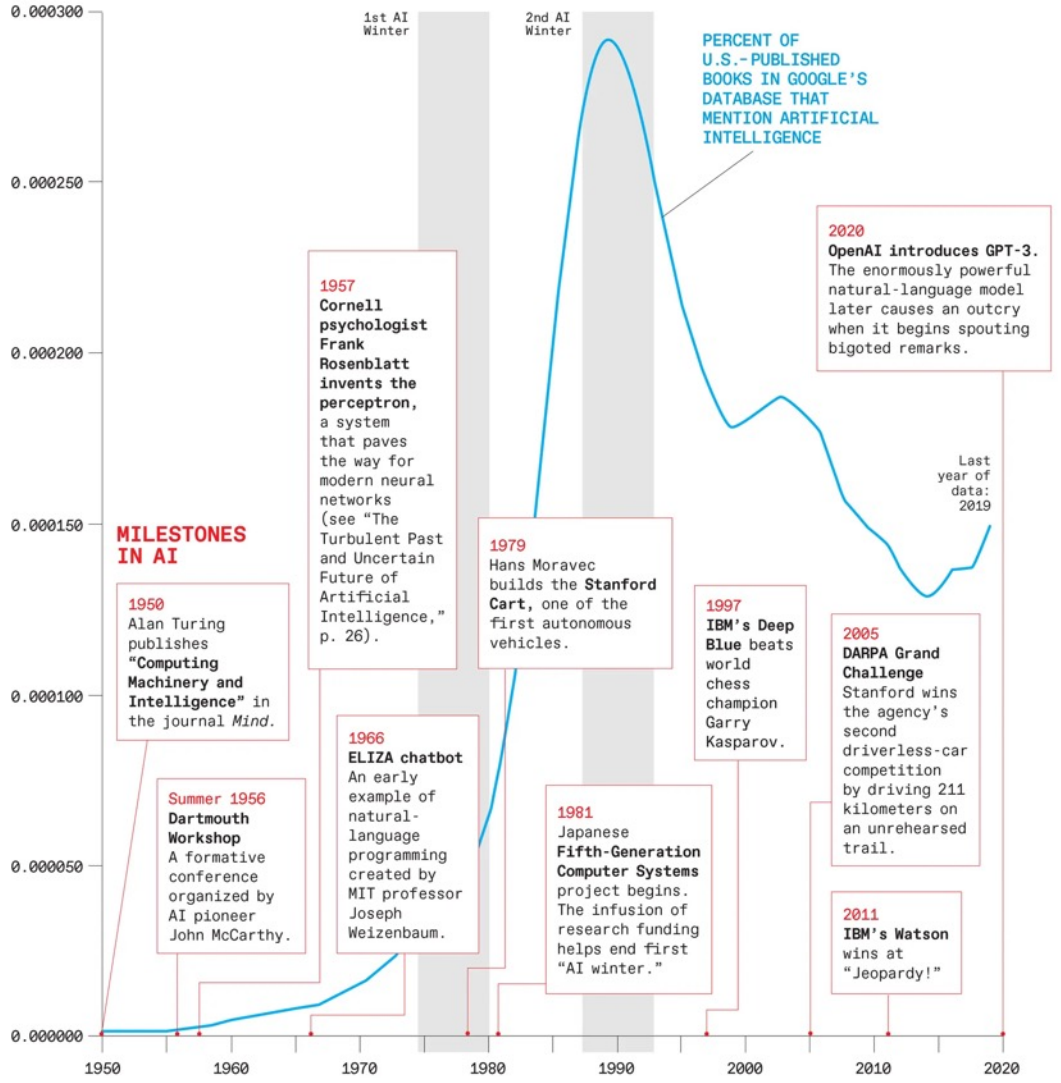
Les développements des premières années

- « *Geometry Theorem Prover* » (1958), de Gelernter, capable de prouver des théorèmes mathématiques difficiles
- « *General Problem Solver* » (1959), de Newell et Simon, cherche à imiter les méthodes humaines de résolution des problèmes
- « Advice Taker » (1958), programme informatique hypothétique décrit par McCarthy dans son *Programs with Common Sense*. Premier programme à utiliser la logique en tant qu'outil de représentation et non en tant que matière.
- Création de Lisp (J. McCarthy), devenu par la suite le langage de programmation dominant pour les IA

Le triomphe (actuel) du *deep learning*

Une histoire marquée par plusieurs périodes ”hivernales”

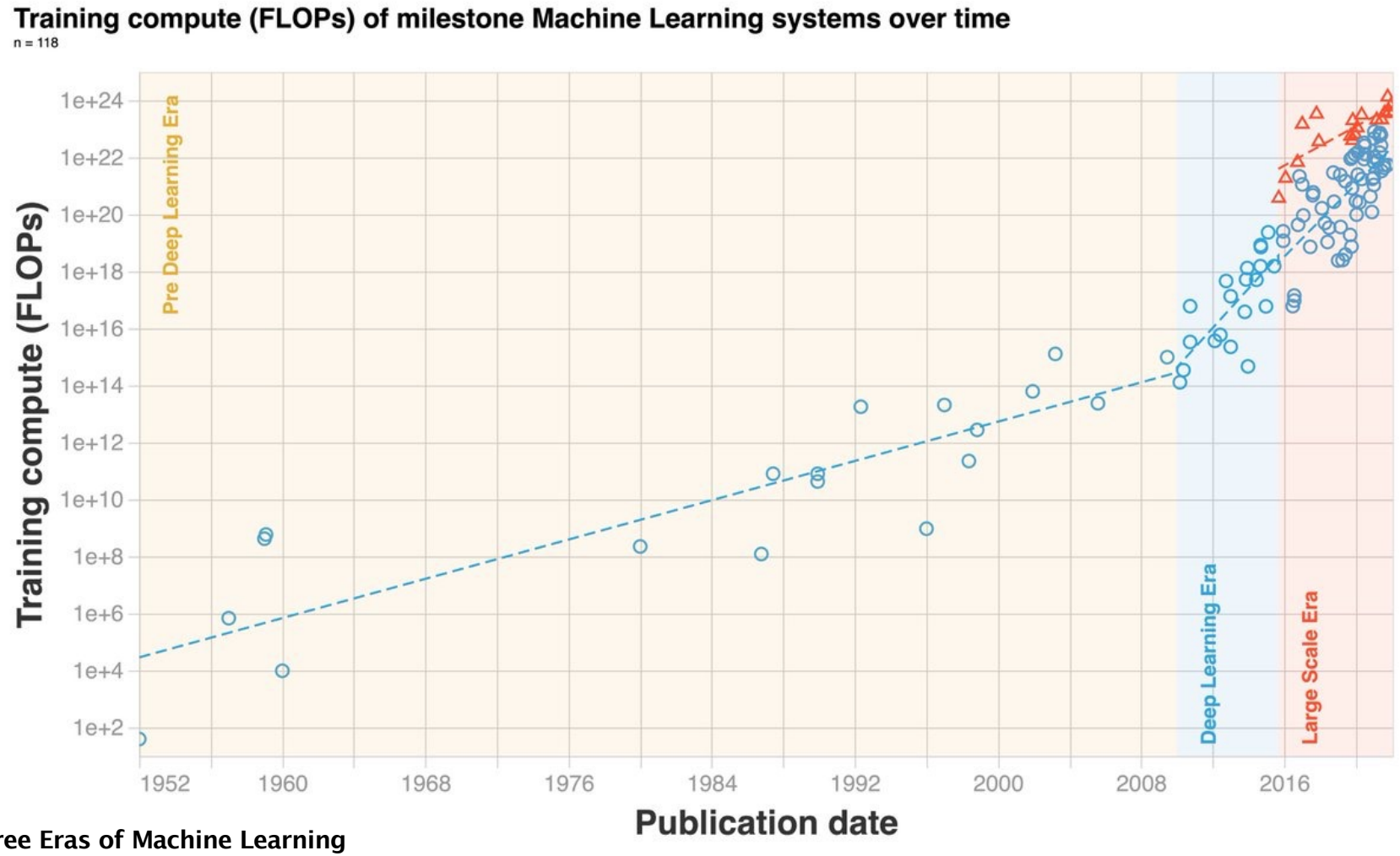
- Le développement de ce champ scientifique est tout sauf linéaire et conaît des périodes au cours desquelles la faiblesse des financements ralentit son développement
- Les difficultés inhérentes au sujet, reproduire l’intelligence humaine, en font un champ très vaste où les différentes écoles s’opposent plus que ne dialoguent
- Deux grandes approchent s’opposent: l’approche dite symbolique et l’approche par les reseaux de neurones.
- La première période est très largement dominée par l’approche symbolique
- Les progrès liés à la puissance de calcul et à la disponibilité des données consacrent l’approche fondée sur les reseaux de neurones



Le retour des réseaux de neurones et le triomphe du deep learning

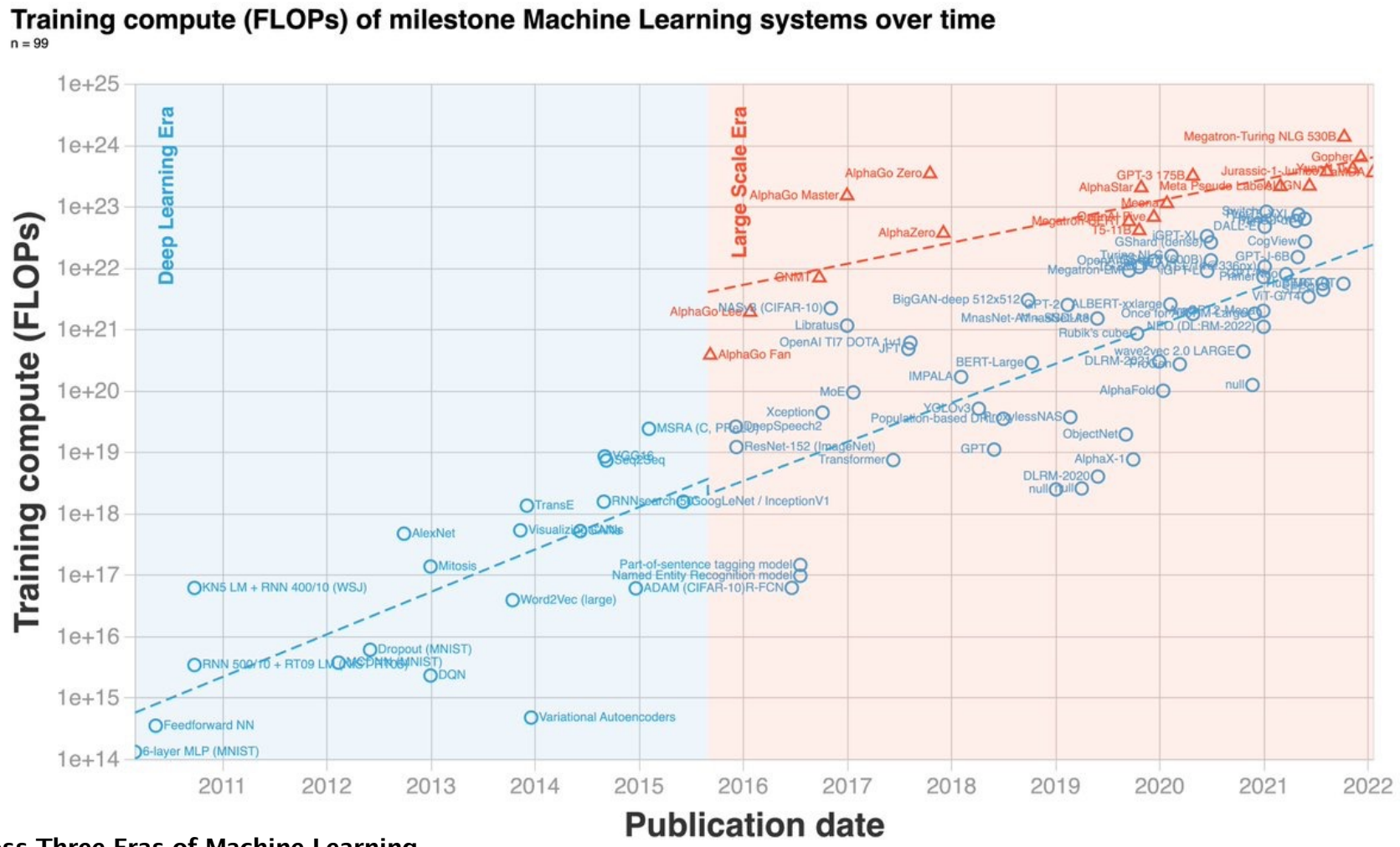
- Le concours *ImageNet Large Scale Visual Recognition Challenge*, (Stanford), en 2012, le programme vainqueur pulvérise les records établis jusqu'à présent, en s'appuyant pour la première fois sur le *deep learning*
- Avancées fondamentales dans plusieurs champs:
  - Computer vision
  - NLP
  - GAN (Generative Adversarial Networks)
  - Jeux (AlphaGoZero, AlphaGo2, Deepmind)
  - GPT-3 (OpenAI), https://arxiv.org/abs/2005.14165
- Consécration de l'approche avec le prix Turing 2018 de Bengio, Hinton et Lecun. https://awards.acm.org/about/2018-turing

Une acceleration liée à l'augmentation de la puissance de calcul



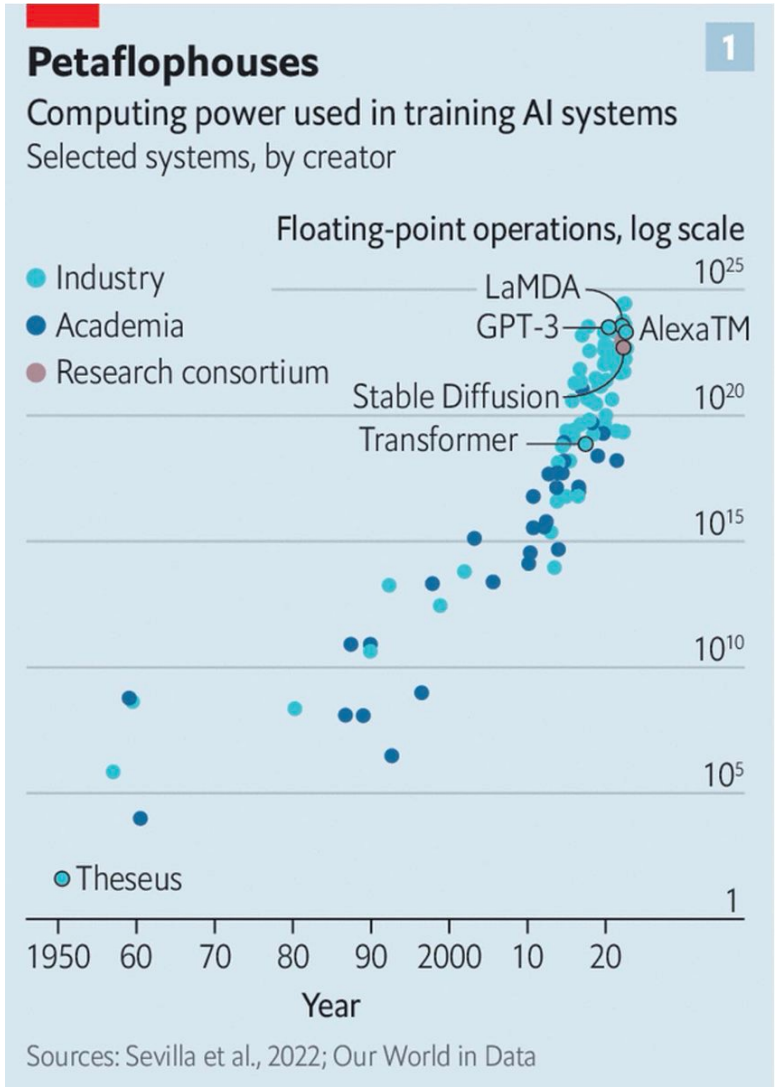
Compute Trends Across Three Eras of Machine Learning
Jaime Sevilla, Lennart Heim, Anson Ho, Tamay Besiroglu, Marius Hobbhahn, Pablo Villalobos, https://arxiv.org/abs/2202.05924

Taille des modèles et puissance de calcul



Compute Trends Across Three Eras of Machine Learning
Jaime Sevilla, Lennart Heim, Anson Ho, Tamay Besiroglu, Marius Hobbhahn, Pablo Villalobos, https://arxiv.org/abs/2202.05924

La course à la puissance de calcul



Temps d'entrainement : divisé par 20 000 en 3 ans

How long does it take to train Resnet-50 on ImageNet?

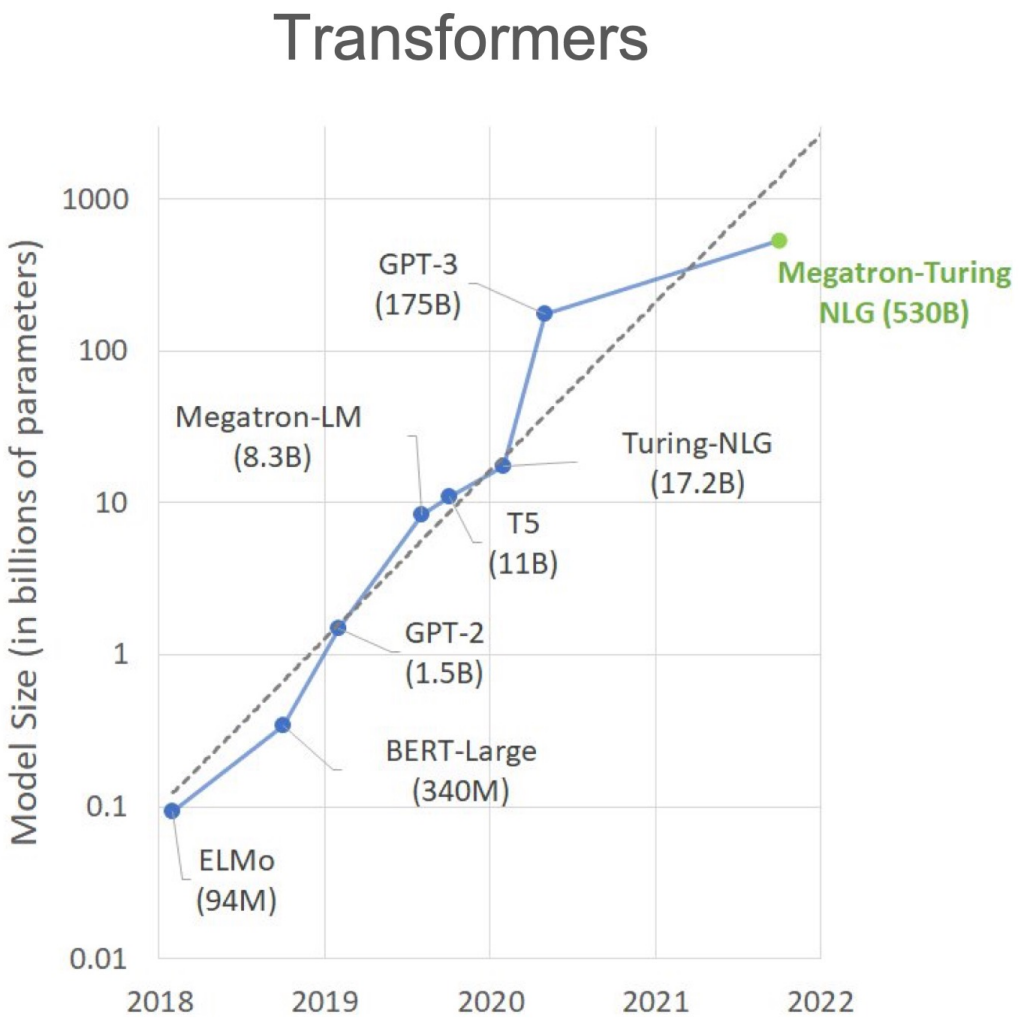
14 days

NVIDIA M40 GPU

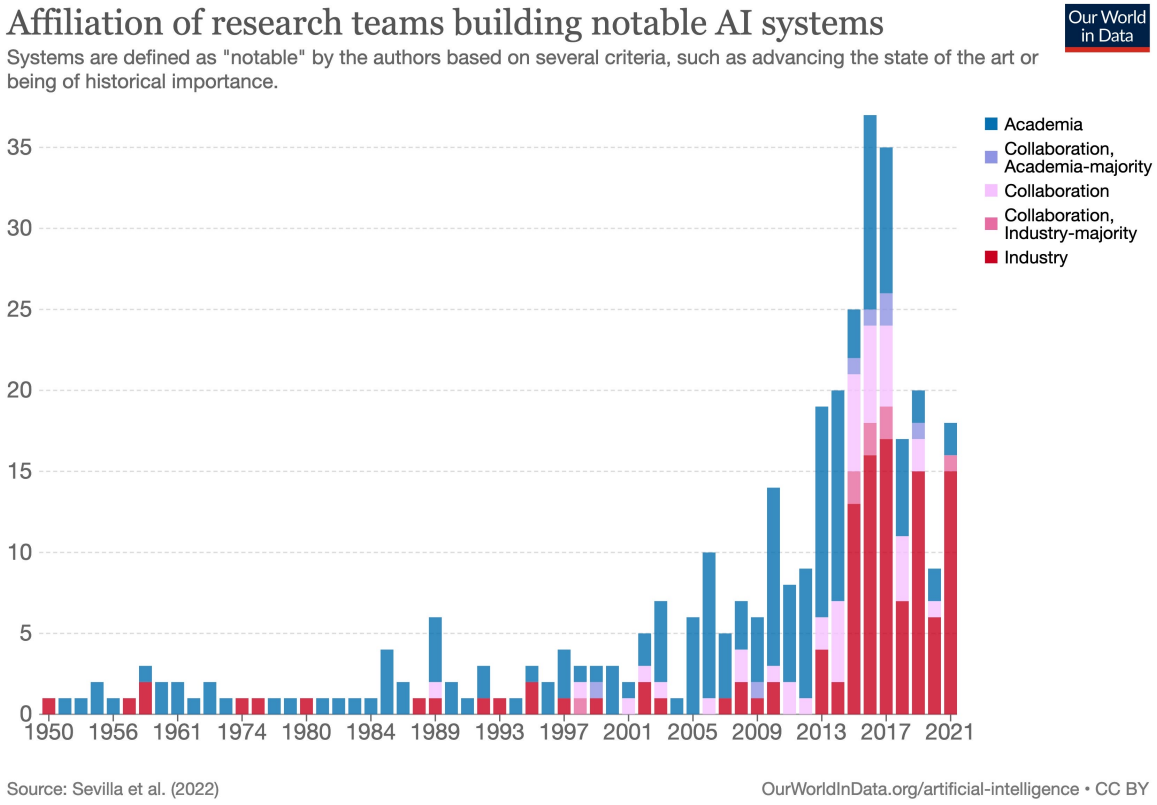
Training Resnet-50 on Imagenet

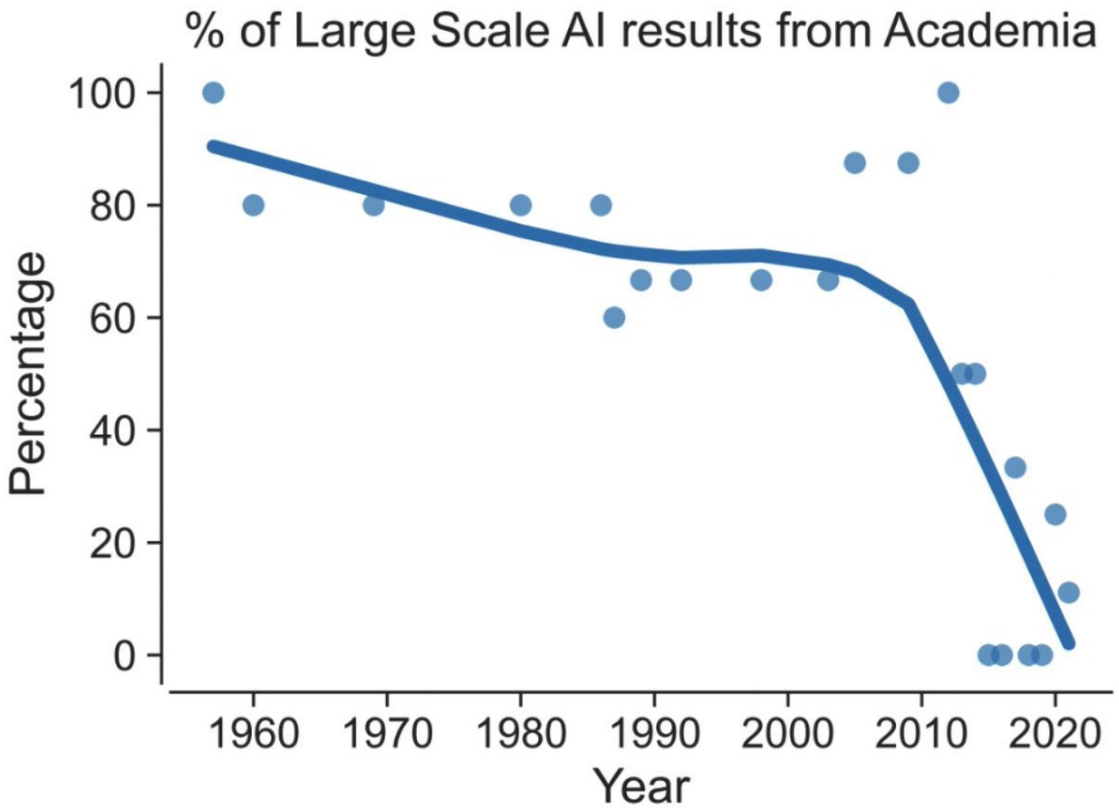
Facebook Caffe2	UC Berkeley, TACC, UC Davis Tensorflow	Preferred Network ChainerMN	Tencent TensorFlow	Sony Neural Network Library (NNL)	Fujitsu MXNet
1 hour	31 mins	15 mins	6.6 mins	2.0 mins	1.2 mins
Tesla P100 x 256	1,600 CPUs	Tesla P100 x 1,024	Tesla P40 x 2,048	Tesla V100 x 3,456	Tesla V100 x 2,048
Apr	Sept	Nov	July	Nov	Apr
2017				2018	2019

Taille des *transformers*



Qui produit l'IA ? De la théorie à l'industrialisation

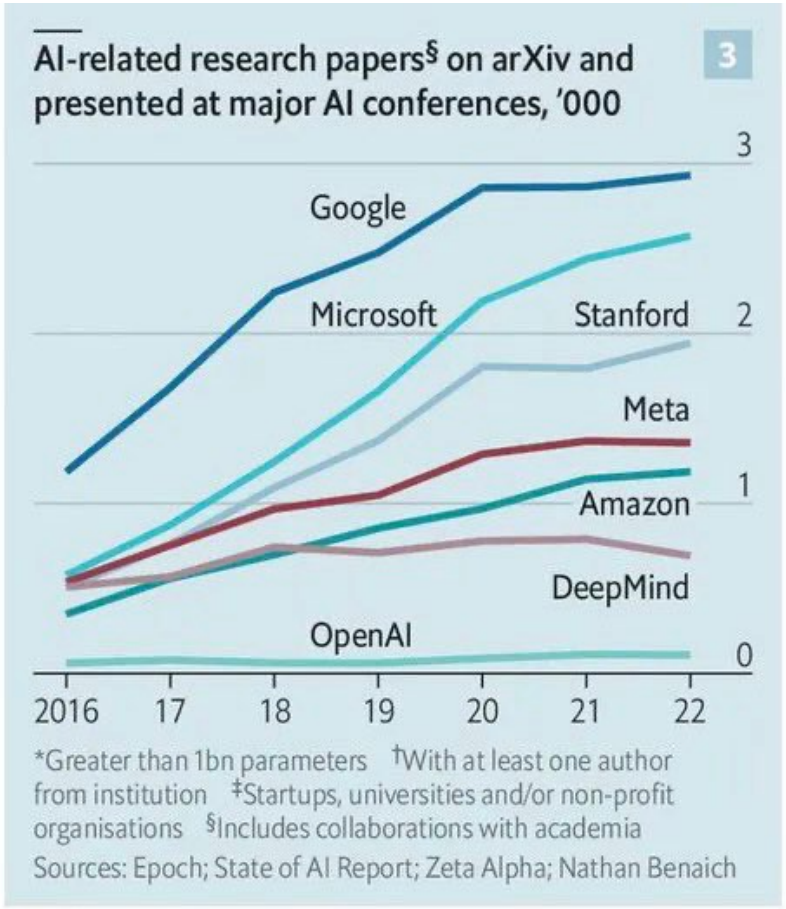


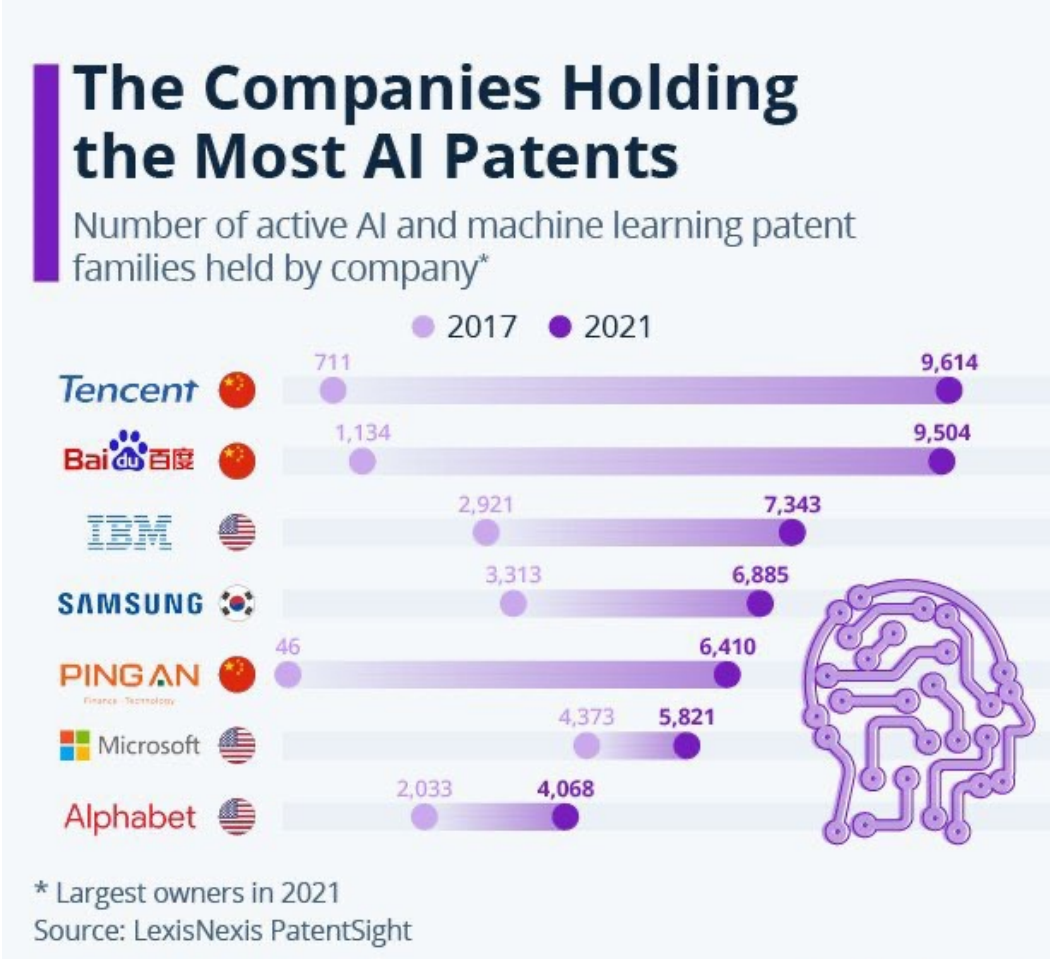


Modèles ouverts, modèles propriétaires

- LLaMA (Meta)
  - comprend des modèles à base de *transformers* avec 7 milliards, 13 milliards, 33 milliards et 65 milliards de paramètres.
  - Les modèles ont été entraînés sur Common Crawl, GitHub, Wikipedia, Project Gutenberg, ArXiv et Stack Exchange.
  - LLaMA a surpasse GPT-3 sur toutes les tâches, Chinchilla sur toutes sauf une, et PaLM sur toutes sauf deux.
  - Meta met à disposition LLaMA aux chercheurs d'institutions, d'agences gouvernementales et d'organisations non gouvernementales qui demandent l'accès et acceptent une licence non commerciale.
- BLOOM

Une compétition mondiale





La percée liée aux « Transformers »

- Un transformeur (ou modèle auto-attentif) est un modèle d'apprentissage profond introduit en 2017
- une architecture de réseau neuronal qui utilise l'auto-attention. Il remplace les approches antérieures des LSTM ou des CNN qui utilisaient l'attention entre l'encodeur et le décodeur.
- À l'instar des réseaux de neurones récurrents (RNN Recurrent Neural Network), les transformeurs sont conçus pour gérer des données séquentielles, telles que le langage naturel, pour des tâches telles que la traduction et la synthèse de texte.
- Contrairement aux RNN, les transformeurs n'exigent pas que les données séquentielles soient traitées dans l'ordre.
- Par exemple, si les données d'entrée sont une phrase en langage naturel, le transformeur n'a pas besoin d'en traiter le début avant la fin.
- Grâce à cette fonctionnalité, le transformeur permet une parallélisation beaucoup plus importante que les RNN et donc des temps d'entraînement réduits.

Des interrogations sur le deep learning

Language Models are Few-Shot Learners, https://arxiv.org/abs/2005.14165

Model Name	n_{params}	n_{layers}	d_{model}	n_{heads}	d_{head}	Batch Size	Learning Rate
GPT-3 Small	125M	12	768	12	64	0.5M	6.0×10^{-4}
GPT-3 Medium	350M	24	1024	16	64	0.5M	3.0×10^{-4}
GPT-3 Large	760M	24	1536	16	96	0.5M	2.5×10^{-4}
GPT-3 XL	1.3B	24	2048	24	128	1M	2.0×10^{-4}
GPT-3 2.7B	2.7B	32	2560	32	80	1M	1.6×10^{-4}
GPT-3 6.7B	6.7B	32	4096	32	128	2M	1.2×10^{-4}
GPT-3 13B	13.0B	40	5140	40	128	2M	1.0×10^{-4}
GPT-3 175B or "GPT-3"	175.0B	96	12288	96	128	3.2M	0.6×10^{-4}

Table 2.1: Sizes, architectures, and learning hyper-parameters (batch size in tokens and learning rate) of the models which we trained. All models were trained for a total of 300 billion tokens.

Evolution de la taille des modèles de NLP

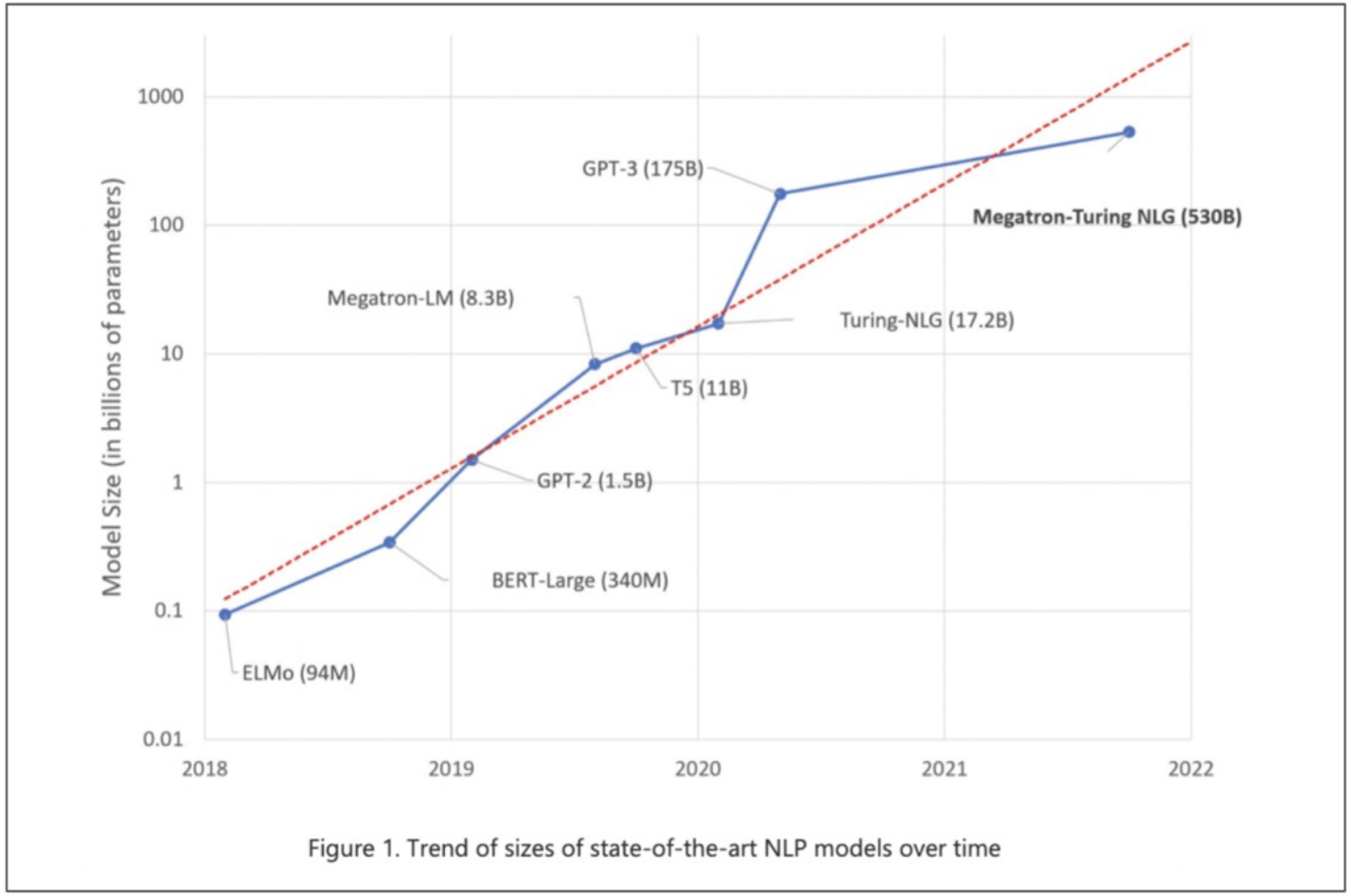


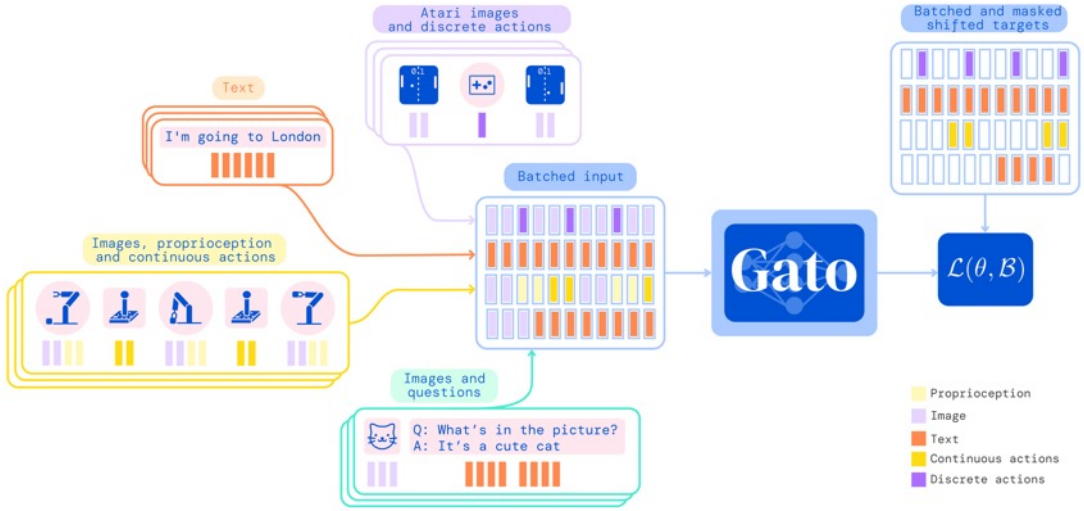
Figure 1. Trend of sizes of state-of-the-art NLP models over time

Via Exponential View

Les modèles génératifs pré-entrainées (General Pre-Trained, GPT)

- Les transformateurs génératifs pré-entraînés (GPT) sont une famille de modèles de langage généralement formés sur un grand corpus de données textuelles pour générer du texte de type humain. Ils sont construits à l'aide de plusieurs blocs de l'architecture du transformateur et peuvent être affinés pour diverses tâches de traitement du langage naturel telles que la génération de texte, la traduction et la classification de texte.
-

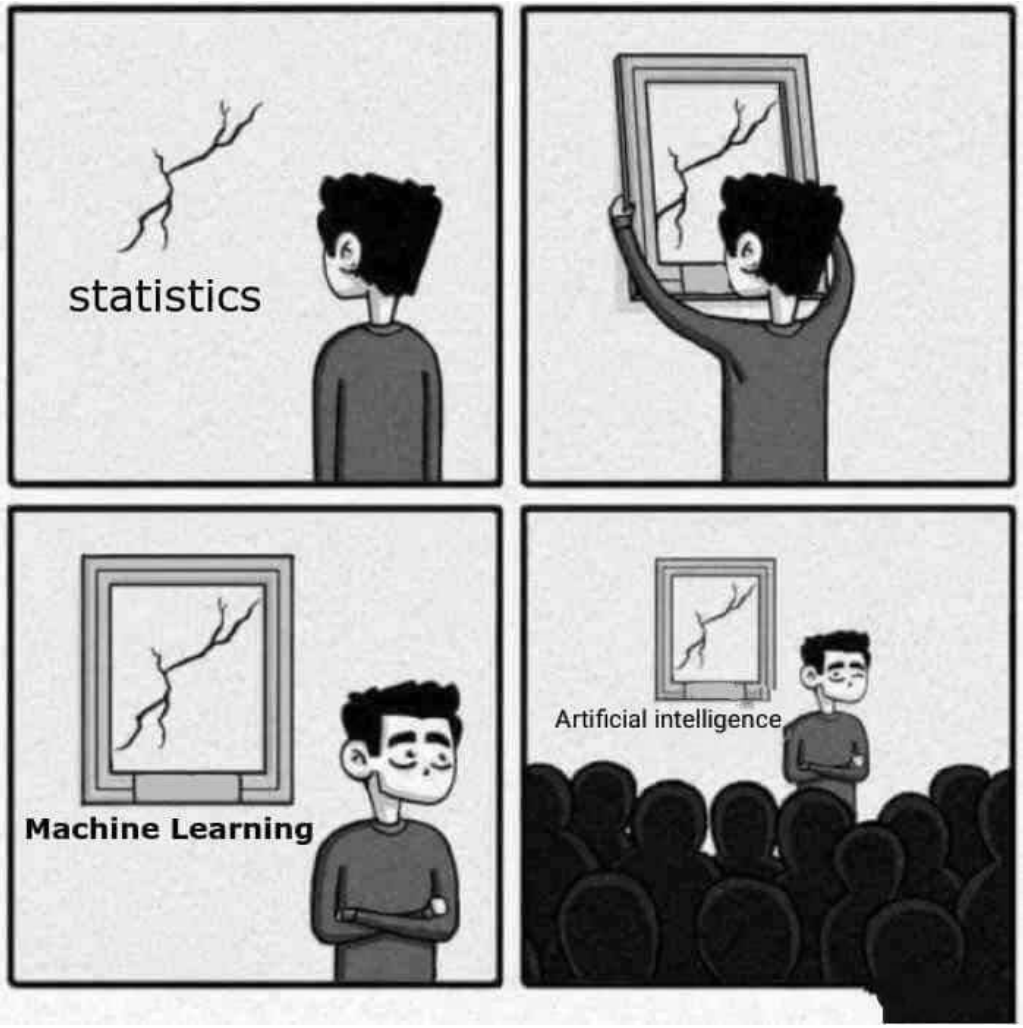
Les modèles multi-modaux : GATO, Deep Mind (2022)



Enjeux, limites et dépassement de l’IA actuelle

Dauphine UNIVERSITÉ PARIS | PSL

Critique du Machine Learning



Un débat toujours en cours Marcus vs Lecun

J'aime Commenter Partager

Patrick John Hayes
I disagree. I think it will be one component of a fully integr... En voir plus

J'aime Répondre 16

Voir 2 réponses précédentes...

Yann LeCun
Patrick John Hayes transformer architectures trained in a Self-Supervised manner (LLMs are merely a special case of that) will definitely be part of future AI system.
But LLMs in their current form, i.e. reactive predictors of the next word, are woefully insufficient.
They can reason, they can't plan, they can't learn effective models of the underlying reality that language is describing.
That's why they can't produce consistent long-form answers, and can't use simple tools (like a calculator, a database search, or a search engine), and make astonishingly stupid mistakes, counterfactual claims, and inconsistent statements.
Fixing all this will require a major redesign, not a mere scaling up.

J'aime Répondre 41

Les difficultés des Large Language Models (LLMs)

- « *This strong awareness of toxic language may or may not be desirable depending on the specific requirements of downstream applications. Future applications of OPT-175B should consider this aspect of the model, and take additional miti- gations, or avoid usage entirely as appropriate* »
- Zhang S., & alii, (2022), « OPT: Open Pre-trained Transformer Language Models », https://arxiv.org/pdf/2205.01068.pdf

DeepMind 2023-1-13

Tracr: Compiled Transformers as a Laboratory for Interpretability

David Lindner[1*], János Kramár[2], Matthew Rahtz[2], Thomas McGrath[2] and Vladimir Mikulik[2]
[1]ETH Zurich, [2]DeepMind, [*]Work done at DeepMind.

Interpretability research aims to build tools for understanding machine learning (ML) models. However, such tools are inherently hard to evaluate because we do not have ground truth information about how ML models actually work. In this work, we propose to build transformer models *manually* as a testbed for interpretability research. We introduce `Tracr`, a "compiler" for translating human-readable programs into weights of a transformer model. `Tracr` takes code written in RASP, a domain-specific language (Weiss et al., 2021), and translates it into weights for a standard, decoder-only, GPT-like transformer architecture. We use `Tracr` to create a range of ground truth transformers that implement programs including computing token frequencies, sorting, and Dyck-n parenthesis checking, among others. We study the resulting models and discuss how this approach can accelerate interpretability research. To enable the broader research community to explore and use compiled models, we provide an open-source implementation of `Tracr` at `https://github.com/deepmind/tracr`.

Keywords: Interpretability, Transformers, Language Models, RASP, Tracr

1. Introduction

As deep learning models are becoming more capable and increasingly deployed in production, improving our ability to understand how they make decisions is crucial.

Mechanistic interpretability aims to achieve this by reverse engineering neural networks and producing *mechanistic* explanations of the algorithms a model implements. This approach has achieved success in convolutional neural networks for image classification. Cammarata et al. (2020) explain a range of specific circuits in InceptionV1 (Szegedy et al., 2015), including curve detectors, high-low frequency detectors, and neurons detecting more high-level concepts such as dogs or cars. Elhage et al. (2021) and Wang et al. (2022) achieve early success in interpreting transformer language models using similar methods.

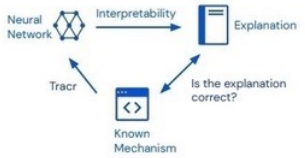


Figure 1 | `Tracr` allows us to create models that implement a known mechanism. We can then compare this mechanism to explanations an interpretability tool produces.

Despite this success, the toolbox of approaches for generating mechanistic explanations remains small and poorly understood. Part of the difficulty is that evaluating mechanistic explanations requires creativity and effort by researchers. It is difficult to evaluate how well an explanation tracks the actual mechanism used by the model when all our knowledge of the mechanism comes from the explanation itself. Without access to ground truth about the proposed mechanism, we must verify the methods used to study it in some other way.

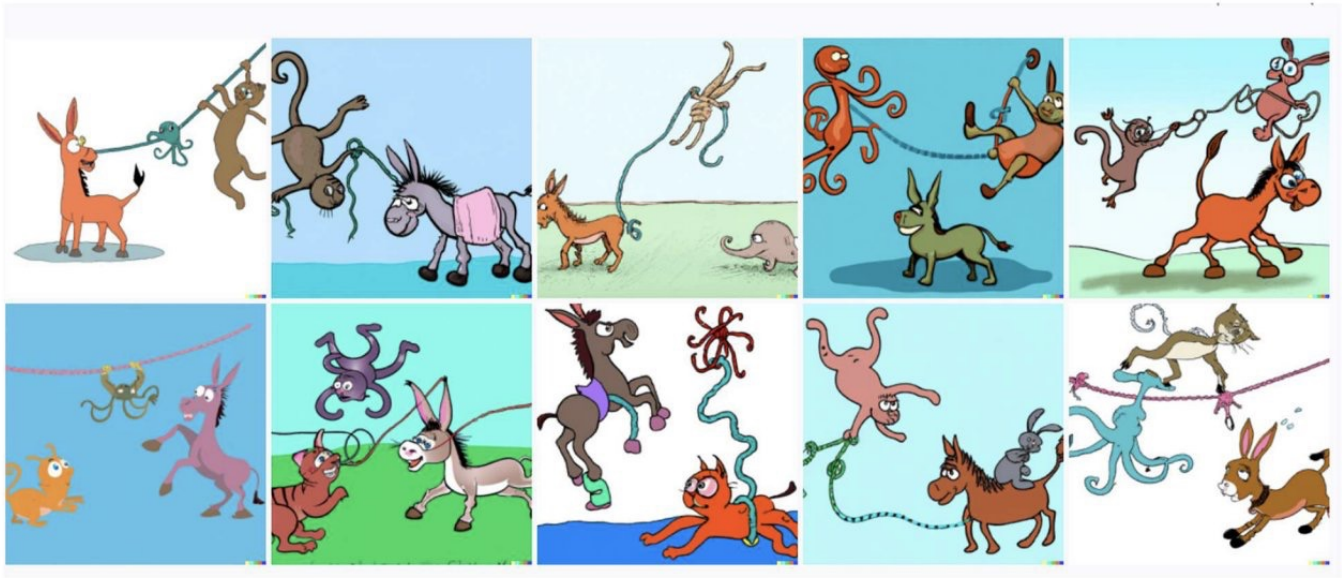
The standard approach for evaluating mechanistic explanations combines evidence from many ad-hoc experiments (e.g., Olah et al. (2020) and Olsson et al. (2022)). However, since this is expensive

DALL-E 2 Problème de composabilité

Example 5.A:

Caption: A donkey and an octopus are playing a game. The donkey is holding a rope on one end, the octopus is holding onto the other. The donkey holds the rope in its mouth. A cat is jumping over the rope.

Images:



Discussion: All of the image contain all of the components of the caption – a donkey, a cat, a rope, and a creature that is multi-legged, though none of the images show an octopus with eight legs. None gets more than one of the stated relations right.

DALL-E-2 et la question des relations

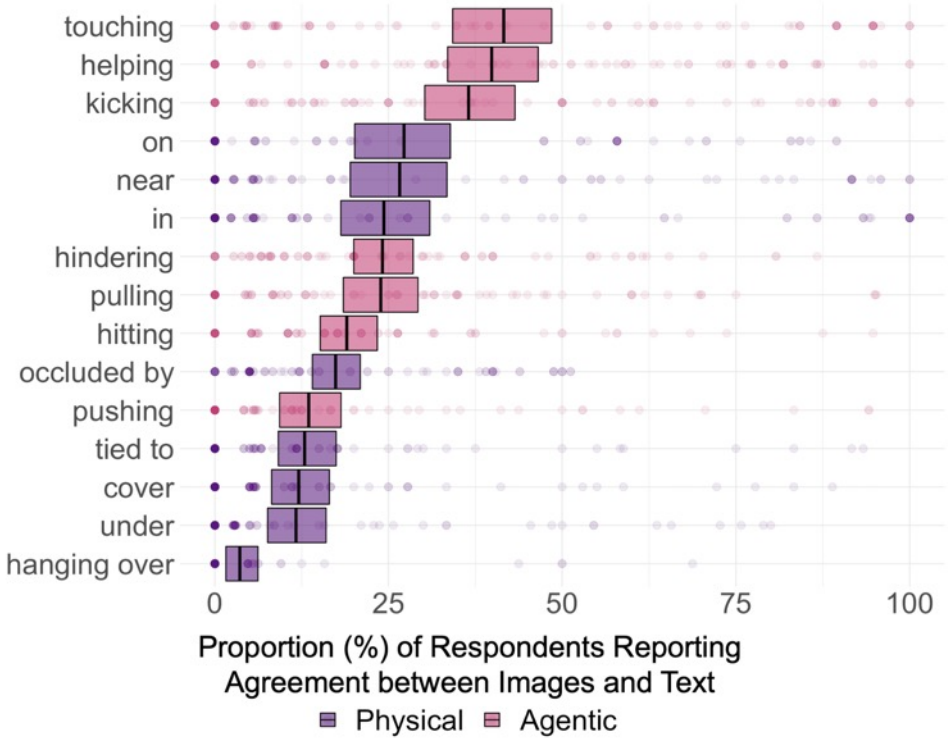


Figure 6: Illustrative example, images generated given ‘a spoon in a cup’ and ‘a cup on a spoon’. Examining just the left images may lead to the conclusion that Dall-E 2 captures the *in* relation, but the right images suggest this is simply an effect of training images that involve *spoon* and cup.

DALL-E 2 (OpenAI) : la question des biais

Prompt: lawyer;
Date: April 6, 2022

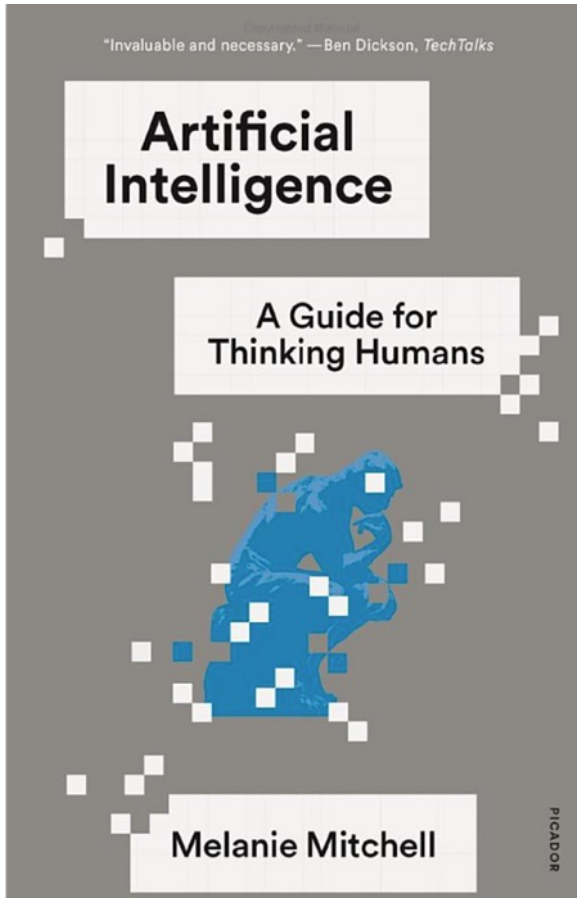
Prompt: a flight attendant; Date: April 6, 2022

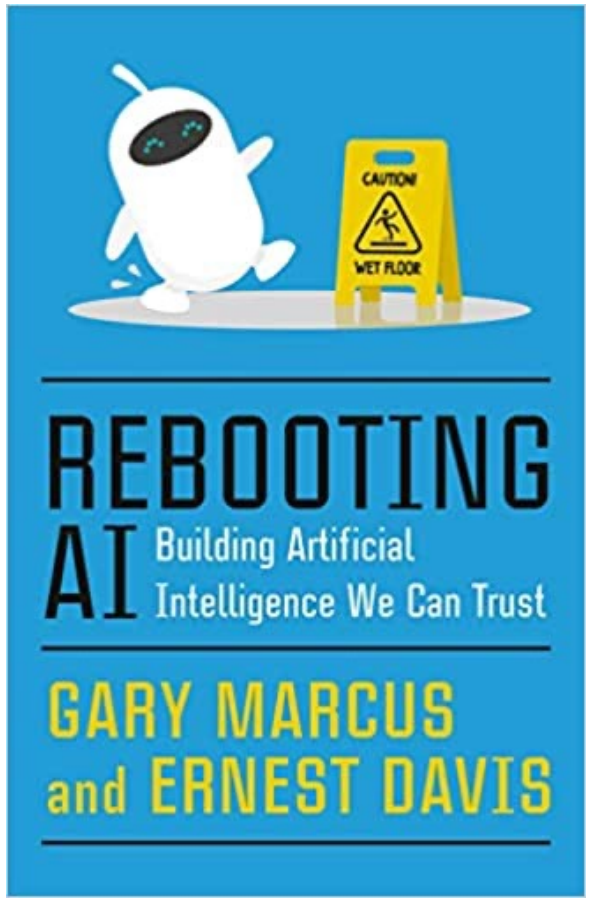
Atténuer les biais du modèle par des filtres ?

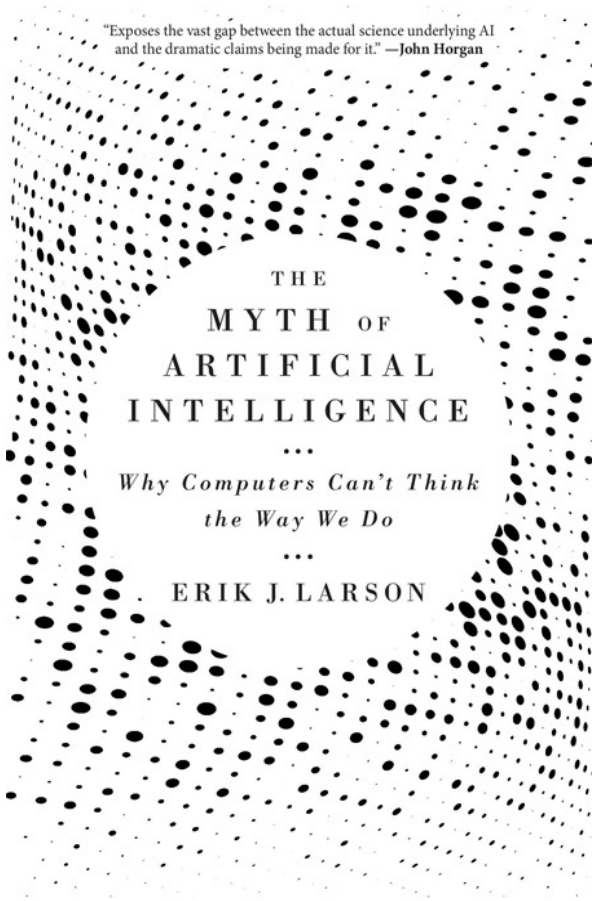
the system's anti violence filters obviously wouldn't allow a user to generate an image of a dead horse in a pool of blood, but it will happily generate "a photo of a horse sleeping in a pool of red liquid"

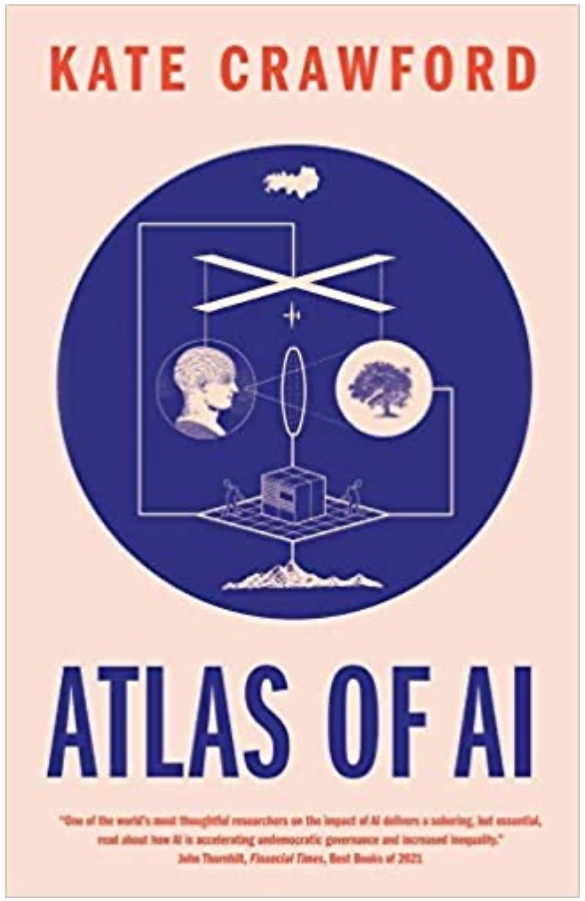
Prompt: a photo of a horse sleeping in a pool of red liquid;
Date: April 6, 2022

Les impasses de l'IA actuelle: trois types de critiques









How to shrink AI's ballooning carbon footprint

- https://arxiv.org/abs/2206.05229

Un dépassement des approches actuelles ?

- https://pierrelevyblog.com/2021/09/20/pour-un-changement-de-paradigme-en-intelligence-artificielle/
- *La Sphère sémantique 1. Computation, cognition, économie de l'information*, Paris et Londres : Hermès-Lavoisier, 2011.
- **Yann LeCun has a bold new vision for the future of AI:**
- https://www.technologyreview.com/2022/06/24/1054817/yann-lecun-bold-new-vision-future-ai-deep-learning-meta/
- Rappin B. (2018), Une brève histoire cybernétique du management contemporain, La Revue desSciences de Gestion, 293, 5, p. 11-18.