

Intelligence Artificielle

Mythes et réalités

Marc Schoenauer
INRIA, Université Paris-Sud et CNRS,
Université Paris-Saclay

Classe virtuelle le 21/2/2019



Intelligence Artificielle

Mythes et réalités

Menu

- Historique et (non-)définition
- Quelques succès récents
- Apprentissage (profond) supervisé

- Des problèmes non (encore) résolus
- L'IA pour quelle société ?

Définition



Faire faire par la machine des tâches demandant des capacités propres à l'intelligence (humaine), et au moins aussi bien que par les humains

Définition

Faire faire par la machine des tâches demandant des capacités propres à l'intelligence (humaine), et au moins aussi bien que par les humains

BOF...



Définition ?

Faire accomplir par la machine des choses qu'elle ne sait pas encore faire

- Jean-Louis Laurière, début 80
- Philippe Kahn, fin 80
- Gérard Sabah, 2017
(rapport de l'OPECST)

Définition ?

... un ensemble d'approches, ayant chacune des objectifs plus précis que le «raisonnement intelligent»

Académie des Technologies 2018

Raisonnement Logique

Représentation Connaissances

Planning et Navigation

Traitement Langage Naturel

Perception

Accélération
2012-2016

→ *Deep Learning*

Histoire

A network diagram consisting of several white circular nodes connected by thin white lines. The nodes are arranged in a somewhat irregular pattern across the frame. The word "Histoire" is positioned at the top center, and the sentence "L'IA est une invention récente" is centered in the middle of the diagram. The background is a solid light blue color.

L'IA est une invention récente

Histoire

L'IA est une invention récente

FAUX

Histoire

Avant 1956, des visions : Alan Turing, neurones formels, robots

1956: Conférence de Dartmouth, première apparition du terme IA

196x: *Problem solving*, jeux, langage naturel

1968: 2001 l'odyssée de l'espace, HAL

1969: *Perceptrons* (Minsky-Papert), arrêt de la recherche en RNs

1973: Lighthill Report, premier hiver de l'IA

198x: Systèmes experts, de Samuel à Chinook (dames)

199x: Second hiver de l'AI, mais Deep Blue (échecs) et premier réseaux convolutionnels

2000: premières applications pour le Web

2010: Deep learning (triomphe des réseaux convolutionnels, jeux vidéos, GO)

2018+: pas un troisième hiver ?

IA = moyen

IA = but

Histoire

Avant 1956, des visions : Alan Turing, neurones formels, robots

IA = moyen

Can Machines Think?

Les machines peuvent-elles penser ?

*C'est un problème de programmation. [...] On estime que le cerveau contient entre 10^{10} et 10^{15} bits. [...] Je peux écrire l'équivalent d'environ 1000 bits par jour, donc il faudrait au moins **60 programmeurs**, travaillant sans arrêt pendant **50 ans**, pour faire le boulot, et encore si tout ce qu'il font est parfait. Il faut absolument trouver **une autre approche du problème**.*



Comment ?

*en (...) copiant **l'éducation**: on peut espérer modifier le comportement de la machine jusqu'à ce qu'elle produise les réactions voulues à certains ordres. On pourrait ainsi atteindre l'intelligence avec seulement deux leviers pour 'programmer' la machine, l'un pour **le plaisir ou la récompense**, et l'autre pour **la douleur ou la punition**.*

Histoire

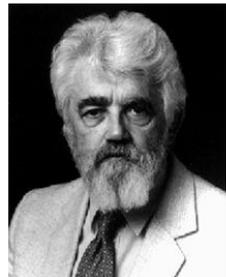
Avant 1956, des visions : Alan Turing, neurones formels, robots

IA = moyen

1956: Conférence de Dartmouth, première apparition du terme IA

IA = but

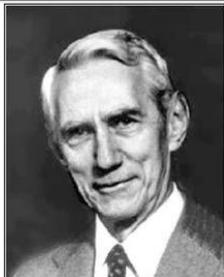
1956 Dartmouth Conference: The Founding Fathers of AI



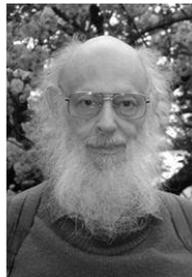
John McCarthy



Marvin Minsky



Claude Shannon



Ray Solomonoff

Alan Newell



Herbert Simon



Arthur Samuel



And three others...

Oliver Selfridge
(Pandemonium theory)

Nathaniel Rochester
(IBM, designed 701)

Trenchard More
(Natural Deduction)

*Nous nous proposons d'étudier l'intelligence artificielle [..]. Nous faisons l'hypothèse que tous les aspects de l'apprentissage et de l'ensemble de ce qu'on appelle l'intelligence peut être **pdécrit de manière si précise** qu'on peut les **simuler** à l'aide d'un ordinateur.*

La vision : le raisonnement est une suite d'opérations que la machine peut reproduire

Le but : *General Problem Solver*
(aka 2000+ : Artificial General Intelligence)

Histoire

Avant 1956, des visions : Alan Turing, neurones formels, robots

1956: Conférence de Dartmouth, première apparition du terme IA

196x: *Problem solving*, **jeux**, langage naturel

1968: 2001 l'odyssée de l'espace, HAL

1969: *Perceptrons* (Minsky-Papert), arrêt de la recherche en RNs

1973: Lighthill Report, premier hiver de l'IA

198x: Prolog+FGCS; Systèmes experts; de Samuel à Chinook (**dames**)

199x: Second hiver de l'AI, mais Deep Blue (**échecs**) et premier réseaux convolutionnels

2000: premières applications pour le Web

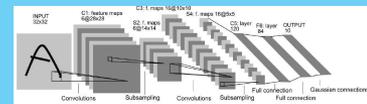


2010: Deep learning (triomphe des réseaux convolutionnels, **jeux vidéo, GO, ...**)

2018+: pas un troisième hiver ?

IA = moyen

IA = but



Histoire

Avant 1956, des visions : Alan Turing, neurones

1956: Conférence de Dartmouth, première apparition

196x: *Problem solving*, **jeux**, langage naturel

1968: 2001 l'odyssée de l'espace, HAL

1969: *Perceptrons* (Minsky-Papert), arrêt de la recherche en RNs

1973: Lighthill Report, premier hiver de l'IA

198x: Prolog+FGCS; Systèmes experts; de Samuel à Chinook (**dames**)

199x: Second hiver de l'AI, mais Deep Blue (**échecs**) et premiers réseaux

2000: premières applications pour les entreprises

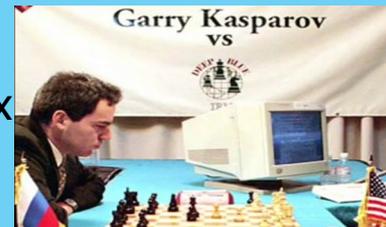
2010: Deep learning (triomphe de l'IA)

2018+: **jeux vidéo**, **GO**, ... ?



IA = moyen

IA = but



els, **jeux vidéo**, **GO**, ...)

Histoire

Avant 1956, des visions : Alan Turing, neurones formels, robots

IA = moyen

1956: Conférence de Dartmouth, première apparition du terme IA

IA = but

196x: *Problem solving*, jeux, langage naturel

1968: 2001 l'odyssée de l'espace, HAL

1969: *Perceptrons* (Minsky-Papert), arrêt de la recherche

1973: Lighthill Report, premier hiver de l'IA

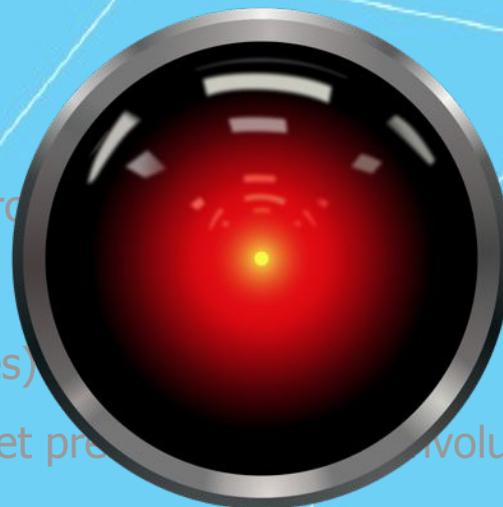
198x: Systèmes experts, de Samuel à Chinook (dames)

199x: Second hiver de l'AI, mais Deep Blue (échecs) et premiers succès révolutionnels

2000: premières applications pour le Web

2010: Deep learning (triomphe des réseaux convolutionnels, jeux vidéo, GO, ...)

2018+: pas un troisième hiver ?





WINTER IS COMING

Avant 1956, des vis

1956: Conférence d

196x: *Problem solvi*

1968: 2001 l'odysée

1969: *Perceptrons* (Minsky-Papert), **arrêt** de la recherche en RNs

1973: Lighthill Report, premier **hiver** de l'IA

198x: Prolog+FGCS; Systèmes experts; de Samuel à Chinook (dames)

199x: Second **hiver** de l'AI, mais Deep Blue (échecs) et premier réseaux convolutionnels

2000: premières applications pour le Web

2010: Deep learning (triomphe des réseaux convolutionnels, jeux vidéo, GO, ...)

2018+: pas un troisième **hiver** ?

IA = moyen

IA = but

Avant
1950
1960
1960
1960
1970
1980
1990



A = moyen
A = but

nnels

2000: premières applications pour le Web

2010: Deep learning (triomphe des réseaux convolutionnels, jeux vidéo, GO, ...)

2018+: pas un troisième hiver ?



Avant 1956, des vis

1956: Conférence c

196x: *Problem solv*

1968: 2001 l'odysée

1969: *Perceptrons*

1973: Lighthill Rep

198x: Systèmes exp

199x: Second hiver

2000: premières applications pour le Web

2010: Deep learning (et triomphe des réseaux convolutionnels)

2018+: pas un troisième hiver ?

IA = moyen

IA = but

réseaux convolutionnels

Intelligence Artificielle

Mythes et réalités

Menu

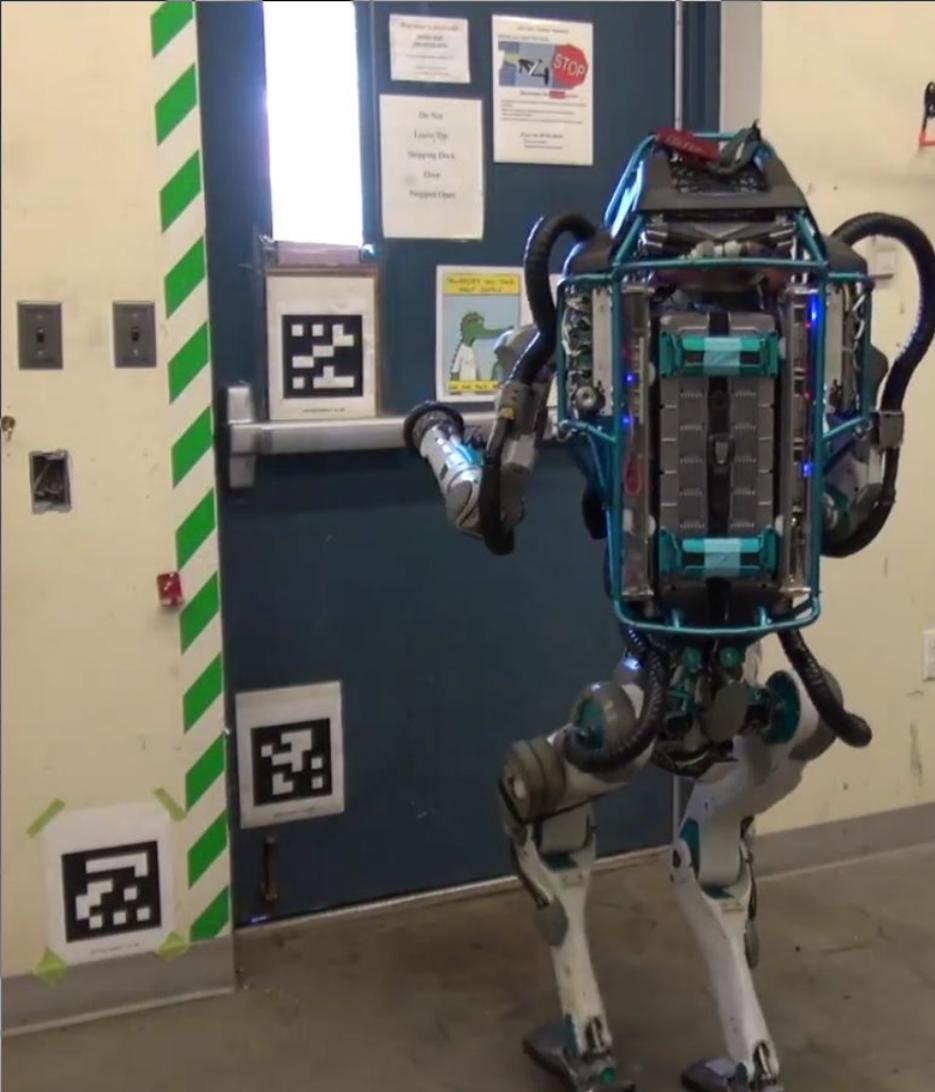
- Historique et (non-)définition
- **Quelques succès récents**
- Apprentissage (profond) supervisé
- Des problèmes non (encore) résolus
- L'IA pour quelle société ?

Autonomie et Robotique

- ***DARPA Autonomous Vehicle Challenge** 2004-07
 - dans le désert, puis en milieu 'urbain'
- **LeNet** (réseau de neurones profond) écrase tous les algorithmes de Computer Vision en reconnaissance d'images 2012
- ***DARPA Rescue Challenge** des robots conduisent, marchent en milieu hostile, montent les escaliers, réparent des machines cassées 2015
- ****Psibernetix** abat (en simulation :-) les meilleurs pilotes de US Air Force
 - algorithmes génétiques et logique floue ... sur un Raspberry Pi ! 2015
- Intel a racheté la société israélienne **MobilEye** pour 15 milliards 2017
- ****Boston Dynamics** robots better and better performing 1998-

Aut

- ***DARPA Auton**
 - dans le dés
- **LeNet** (réseau de
Computer Vision
- ***DARPA Rescu**
hostile, montent
- ****Psibernetix** a
○ algorithmes
- Intel a racheté la
- ****Boston Dyna**



le

algorithmes de

archent en milieu
passées 2015

s de US Air Force
aspberry Pi ! 2015

milliards 2017

g 1998-

Jeux

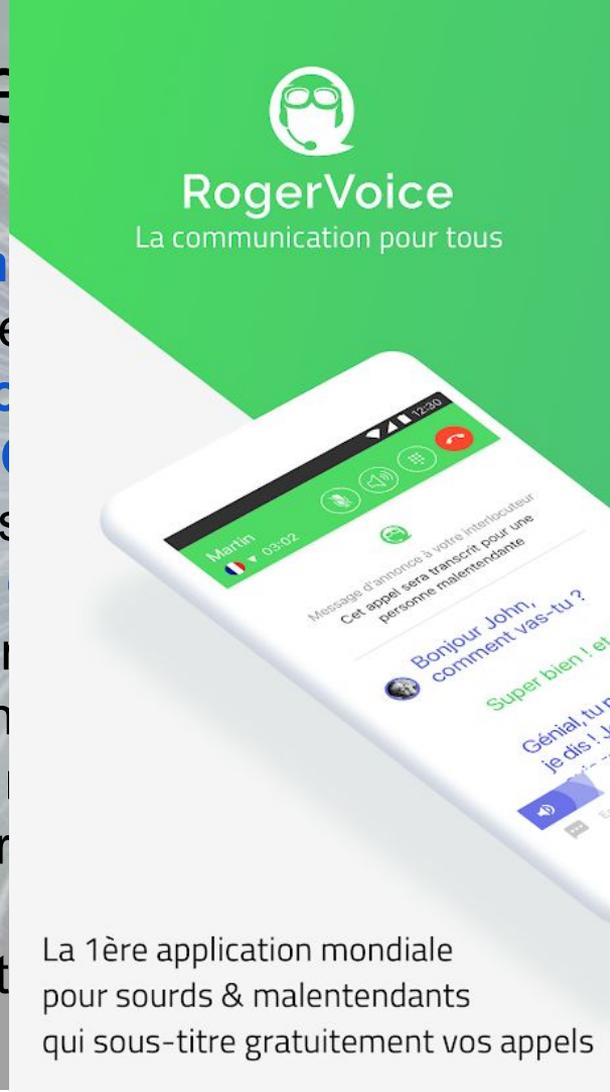
- ****IBM Watson** premier ordinateur à gagner au jeu Jeopardy 2011
 - TLN + recherche web + évaluation, en 3 secondes sur HPC
- **Deepmind** obtient des scores de bon joueur humain à certains (pas tous) jeux vidéo Atari grâce à l'apprentissage par renforcement profond 2013
 - Input: pixels; Output: joystick
- **Deepmind AlphaGo** bat le champion du monde de Go à l'aide d'un mélange d'apprentissages supervisé et par renforcement 2016-17
- **Deepmind AlphaZéro** bat AlphaGo 100-0 uniquement grâce à l'apprentissage par renforcement en jouant contre lui-même 2018
 - à peu près 2 pierres d'avance sur le meilleur humain
 - AlphaZero est également imbattable à d'autres jeux (dont les échecs)
- ***Libratus** a largement battu les meilleurs joueurs de Poker du monde 2017
 - apprentissage par renforcement et approches bayésiennes

TLN et aide au handicap

- **Microsoft Skype Translator** traduit plusieurs langages en temps réel utilisant des réseaux de neurones et de l'apprentissage statistique. Même chose chez **Google Translate**, **Pilot**, ...
- **Apple Siri**, **Microsoft Cortana**, **Amazon Alexa** assistants personnels utilisent de la reconnaissance de la parole et (un peu) de raisonnement
- ****Google Knowledge Graph** utilise la sémantique pour mieux structurer les résultats des recherches
- **Microsoft** traduit du chinois à l'anglais aussi bien qu'un humain
 - grâce à un double réseau profond
- **Ava**, **RogerVoice** pour sourds et malentendants (sous-titrage, téléphone,...)
- **Facebook** peut étiqueter ou décrire les images ou photos aux mal-voyants

TLN e

- **Microsoft Skype Translator** utilisant des réseaux de neurones
- Même chose chez **Google**
- **Apple Siri, Microsoft Cortana** utilisent de la reconnaissance vocale
- ****Google Knowledge Graph**** utilise les résultats des recherches
- **Microsoft** traduit du chinois grâce à un double apprentissage
- **Ava, RogerVoice** pour les malentendants (téléphone,...)
- **Facebook** peut étiqueter



The image shows the RogerVoice logo at the top, which consists of a white speech bubble containing a pair of glasses, with the text "RogerVoice" and "La communication pour tous" below it. Below the logo is a white smartphone displaying a call interface. The screen shows a contact named "Martin" with a French flag icon and the time "03:02". There are icons for voice call, video call, and a grid of app icons. A message on the screen reads: "Message d'annonce à votre interlocuteur: Cet appel sera transcrit pour une personne malentendante". Below this, there is a text message: "Bonjour John, comment vas-tu ?" and a response: "Super bien ! et toi ?". At the bottom, there is a blue button with a microphone icon and the text "Ecris".

La 1ère application mondiale pour sourds & malentendants qui sous-titre gratuitement vos appels

dicap

engages en temps réel
messagerie statistique.

assistants personnels
(peu) de raisonnement
pour mieux structurer

qu'un humain

(sous-titrage,

ou photos aux mal-voyants

TLN et aide au handicap

 **WAVERLYLABS**

[Investors](#)

[Products](#)

[Company](#)

[Press](#)

[Support](#)

[Blog](#)

English

01.

Your system to a world without language barriers

Winter Shipment Sold Out.

Pre-Order Now for Next Shipment!

▲ RED

▼ BLACK

▼ WHITE

~~\$299~~ **\$249.00**

Pre-order now











FEATURED IN



el

els
ent
turer

oyants

TLN et aid

- **Microsoft Skype Translator** tra utilisant des réseaux de neurone. Même chose chez **Google Tran**
- **Apple Siri, Microsoft Cortana,** utilisent de la reconnaissance de
- ****Google Knowledge Graph** ut les résultats des recherches
- **Microsoft** traduit du chinois à l'a
 - grâce à un double réseau pr
- **Ava, RogerVoice** pour sourds e téléphone,...)
- **Facebook** peut étiqueter ou déc



réel
nels
ment
cturer
-voyants

Intelligence Artificielle

Mythes et réalités

Menu

- Historique et (non-)définition
- Quelques succès récents
- Apprentissage (profond) supervisé
- Des problèmes non (encore) résolus
- L'IA pour quelle société ?

Apprentissage machine



L'IA, c'est l'apprentissage (profond)

Apprentissage machine

L'IA, c'est l'apprentissage (profond)

FAUX

Apprentissage machine

L'IA, c'est l'apprentissage (profond)

quoique ...

Ce qui a changé :

- Déluge de données
- Puissance de calcul
- Nouveaux algorithmes

avènement du Web
loi de Moore, GPUs, ...
ou meilleure utilisation d'anciens

Apprentissage machine

Apprentissage à partir d'exemples

- Supervisé
- Semi-supervisé
- Non supervisé

reconnaissance, prédiction

exemples labellisés (étiquetés)

quelques exemples labellisés

aucun exemple labellisé

Apprentissage par renforcement

prise de décisions séquentielle

Apprentissage machine

Apprentissage à partir d'exemples

- **Supervisé**
- Semi-supervisé
- Non supervisé

reconnaissance, prédiction

exemples labellisés (étiquetés)

quelques exemples labellisés

aucun exemple labellisé

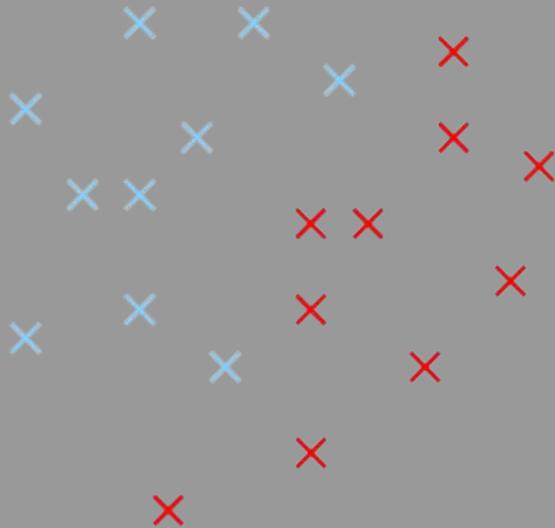
Apprentissage par renforcement

prise de décisions séquentielle

Apprentissage supervisé

Un petit cas jouet

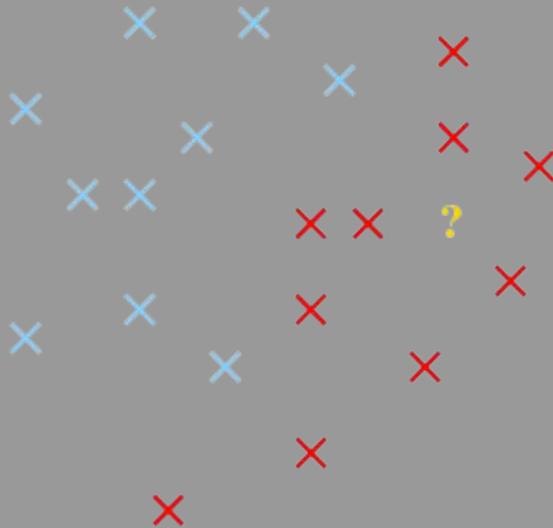
- Un exemple = (x_1, x_2) + une étiquette, ou **label** (rouge ou bleu ici)
- **But**: trouver un **modèle** $f(x_1, x_2)$ qui sépare les points rouges des bleus



Apprentissage supervisé

Un petit cas jouet

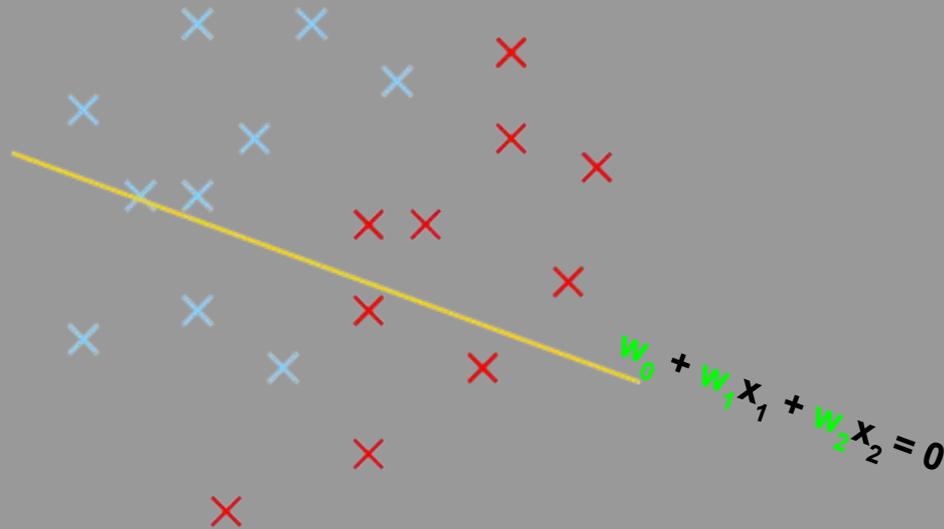
- Un exemple = (x_1, x_2) + une étiquette, ou **label** (rouge ou bleu ici)
- **But**: trouver un **modèle** $f(x_1, x_2)$ qui sépare les points rouges des bleus
- et permette de classer les exemples sans étiquettes à partir de (x_1, x_2)



Apprentissage supervisé

Un petit cas jouet

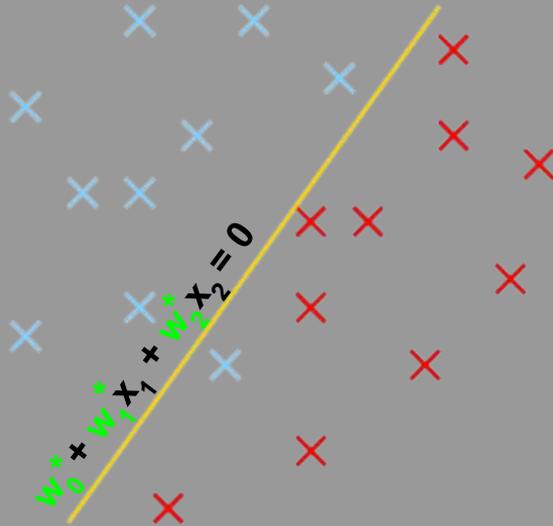
- Par exemple, les modèles linéaires (droites) sont définis par 3 paramètres



Apprentissage supervisé

Un petit cas jouet

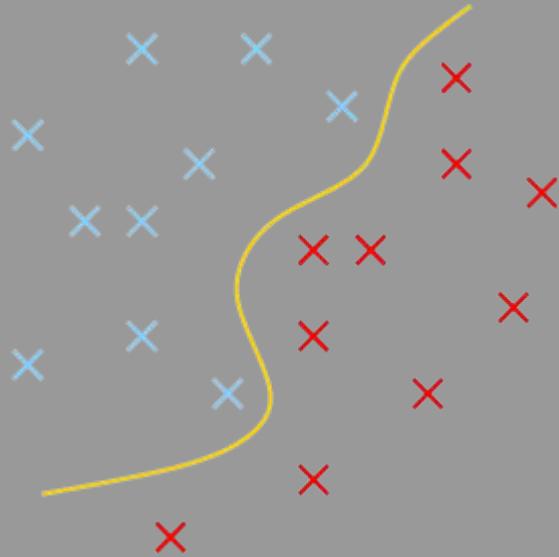
- Par exemple, les modèles linéaires (droites) sont définis par 3 paramètres
- Et on cherche ceux qui séparent au mieux les données disponibles
- C'est l'**apprentissage** des paramètres du modèle



Apprentissage supervisé

Un petit cas jouet

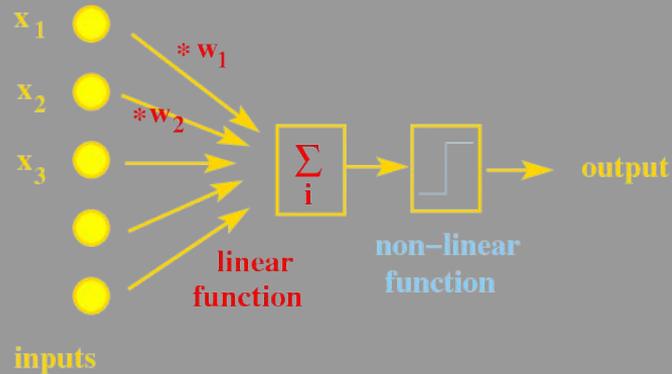
- Il existe des modèles plus compliqués (définis par plus de paramètres)



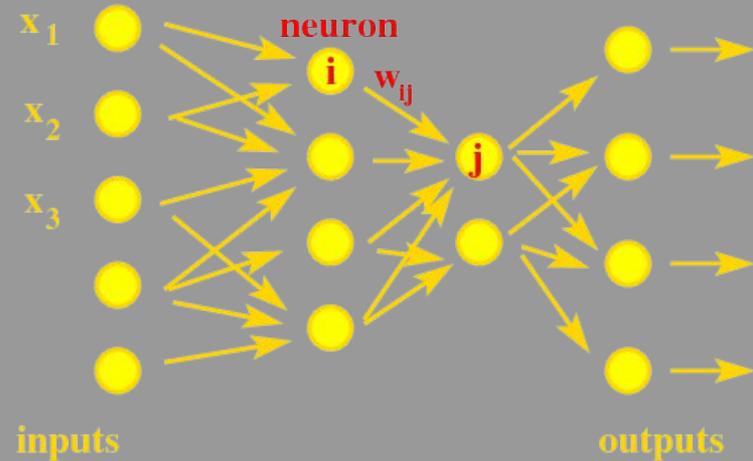
Apprentissage supervisé

Différents modèles

- Polynômes
- Mélanges de Gaussiennes
- Support Vector Machine (machines à noyaux)
- Réseaux de neurones artificiels



Un neurone



Un réseau de neurones:

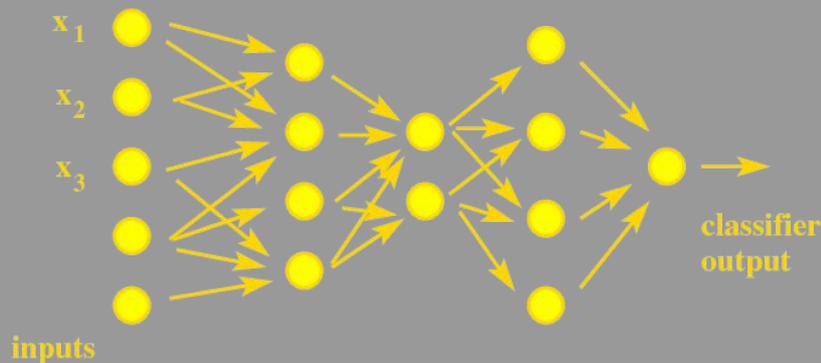
Les paramètres sont les poids w_{ij}

Réseaux de neurones profonds

Phase d'apprentissage

Rétro-propagation du gradient

- On présente les exemples 1 par 1
 - ou par mini-batch de quelques dizaines
- On calcule l'erreur commise
 - différence entre sortie attendue (label) et sortie du réseau
- On **ajuste les poids** w_{ij}
 - pour (tenter de) la faire décroître
- On recommence



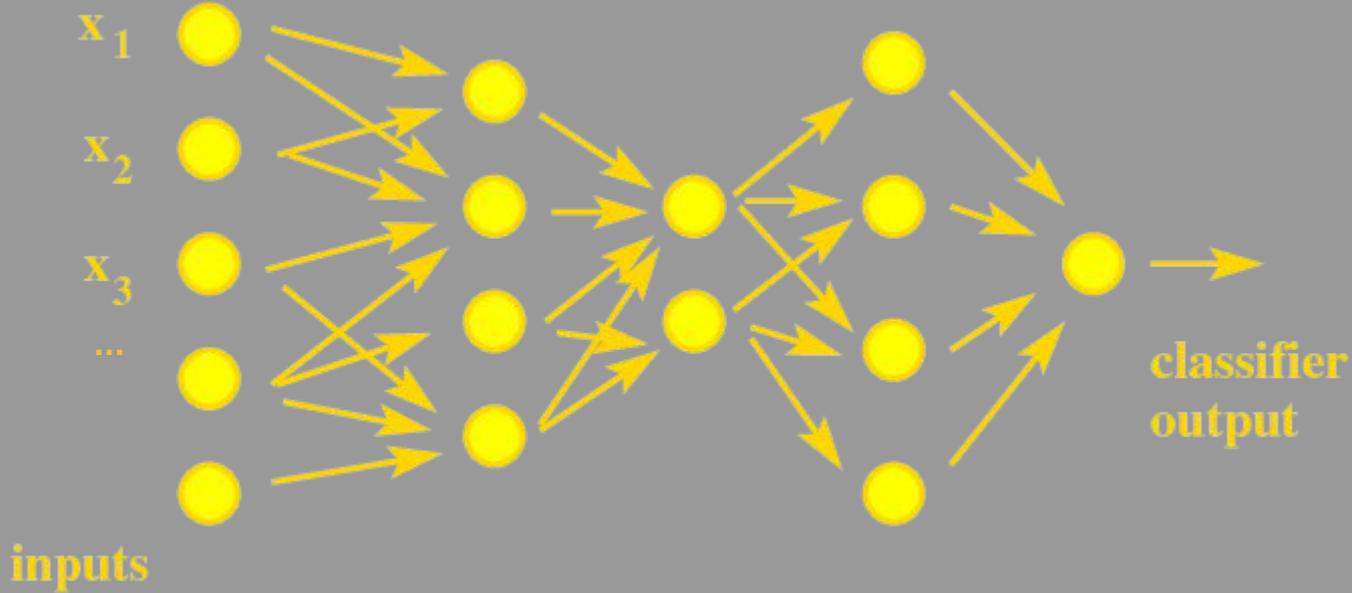
On gèle les poids w_{ij} → Phase de prédiction

On présente un exemple non labellisé, la sortie du réseau prédit un label

Apprentissage total



les pixels

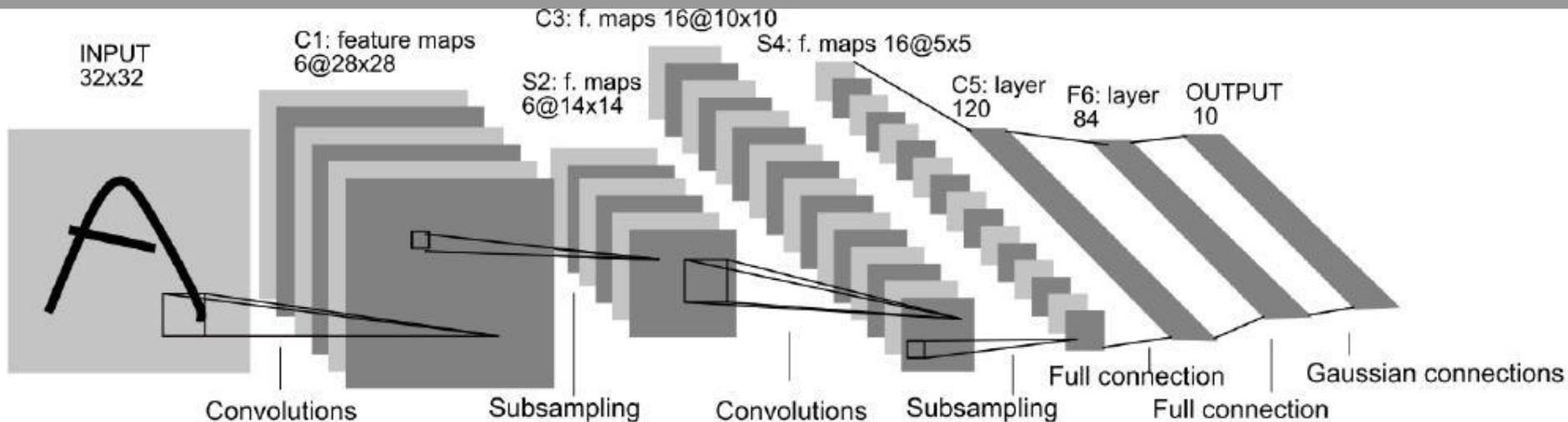


Réseaux convolutionnels



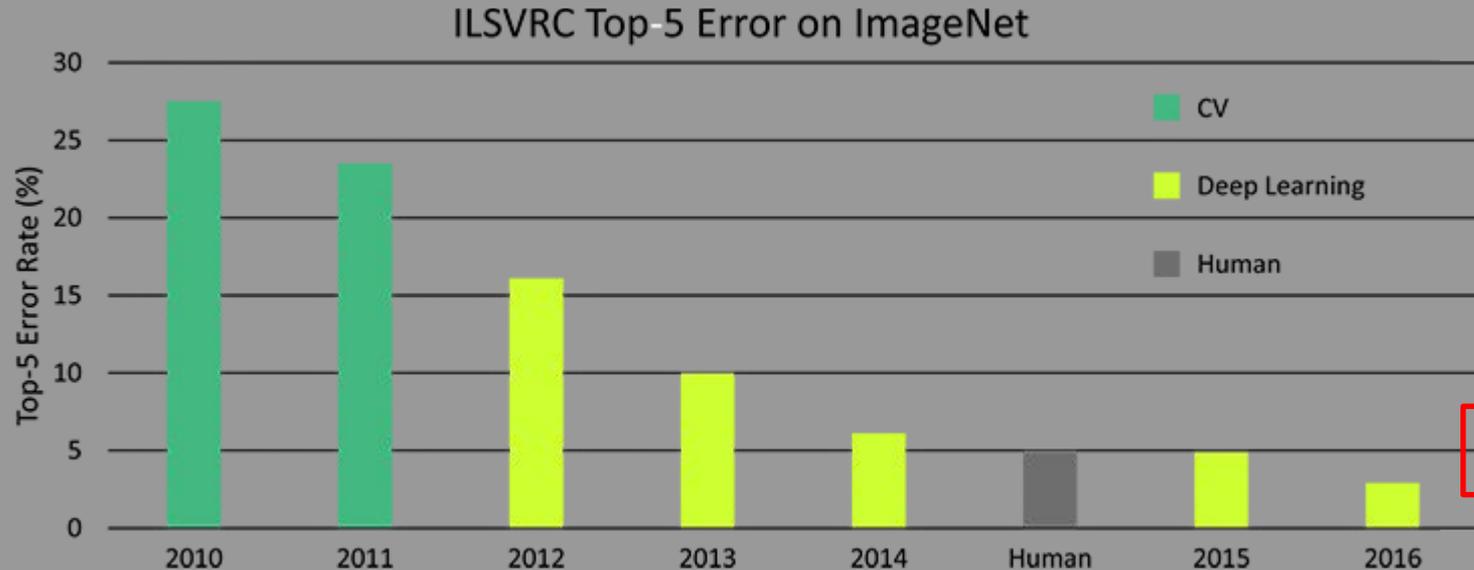
Modèle entièrement appris

CHAT



Apprentissage profond supervisé

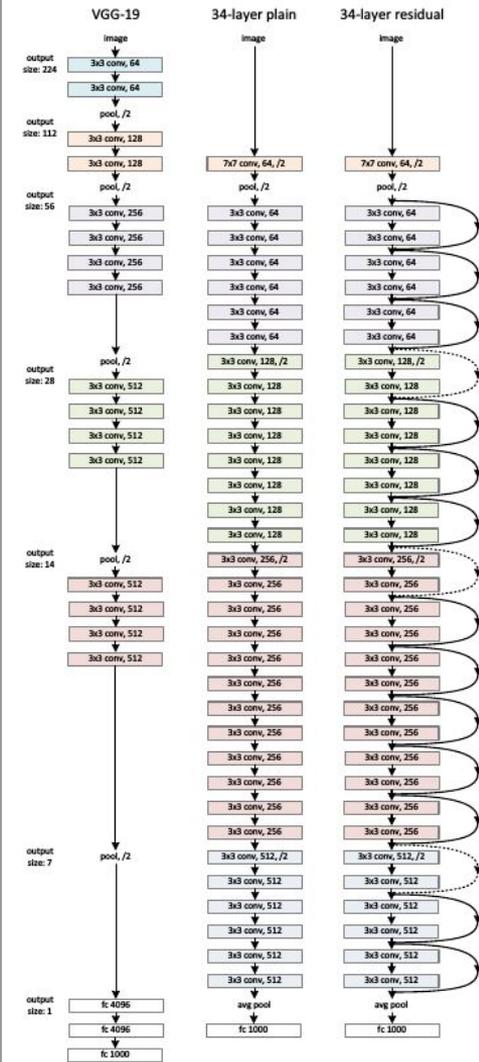
Meilleur que l'apprentissage humain



Human

Etat de l'art

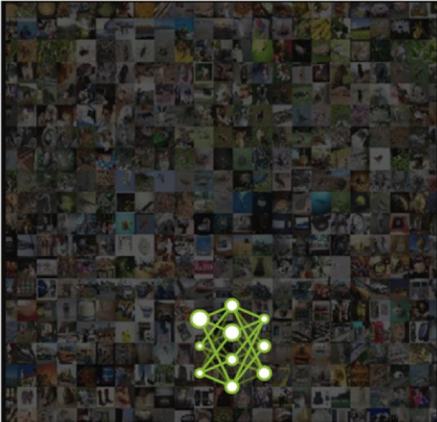
- Données à milliers de classes
 - ImageNet : 1000 et 3000 classes
- Réseaux à dizaines de couches
 - jusqu'à 152 !
- Millions de poids à optimiser
 - centaines de GPU mobilisés
- Trucs et astuces divers et variés
 - Dropout, couches résiduelles, ensembles, ...
- Erreur **3.75%** (2016)



Etat de l'art

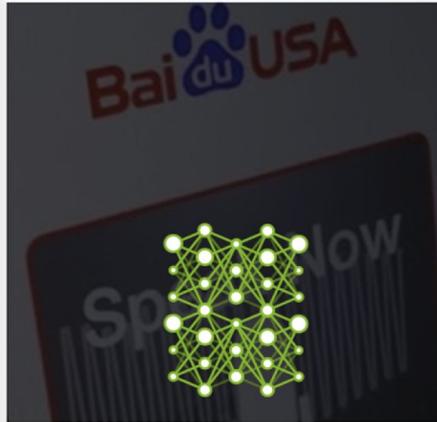
- Donner
- I
- Rése
- j
- Millio
- c
- Trucs
- D
- Erre

7 ExaFLOPS
60 Million Parameters



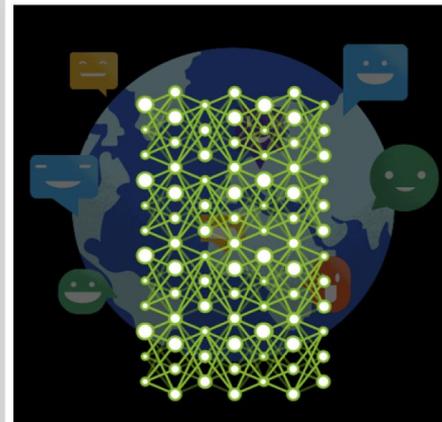
2015 - Microsoft ResNet
Superhuman Image Recognition

20 ExaFLOPS
300 Million Parameters



2016 - Baidu Deep Speech 2
Superhuman Voice Recognition

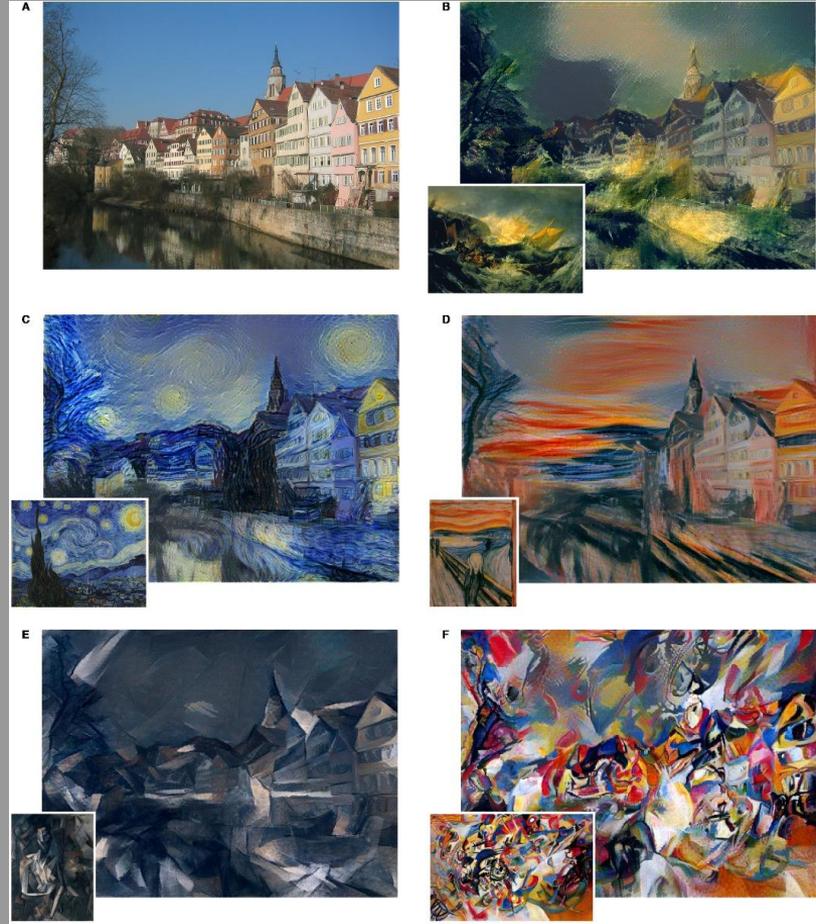
100 ExaFLOPS
8700 Million Parameters



2017 - Google Neural Machine Translation
Near Human Language Translation

Apprentissage profond supervisé

- Performances exceptionnelles
 - Reconnaissance et prédiction
 - disposant de donnée massives
- Des applications rigolotes
 - Génération d'images réalistes
 - Transfert de styles
 - ...



Apprentissage profond supervisé

- Performances exceptionnelles
 - Reconnaissance et prédiction
 - disposant de donnée massives
- Des applications rigolotes
 - Génération d'images réalistes
 - Transfert de styles
 - ...
- D'autres moins rigolotes
 - "Deep Fakes"



Intelligence Artificielle

Mythes et réalités

Menu

- Historique et (non-)définition
- Quelques succès récents
- Apprentissage (profond) supervisé
- **Des problèmes non (encore) résolus**
- L'IA pour quelle société ?

Intelligence Artificielle

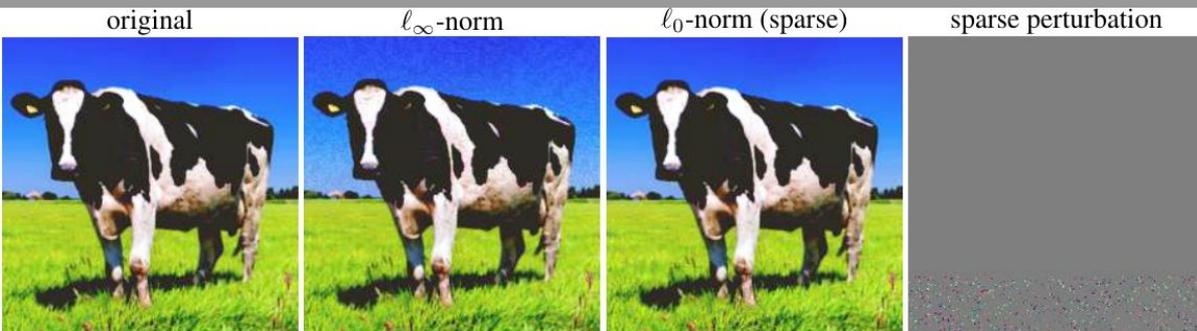
Mythes et réalités

Menu

- Historique et (non-)définition
- Quelques succès récents
- Apprentissage (profond) supervisé
- Des problèmes non (encore) résolus
 - **Validation et certification**

Rien à voir avec l'apprentissage humain

Modifications imperceptibles à l'oeil → faux label



Vache (a) classée “Feu rouge” (b-c)

Shafahi et al., 2018



Tous sont classés “Vitesse limitée à 45” sous plusieurs angles et distances.

Eykholt et al., 2017

Rien à voir avec l'apprentissage humain

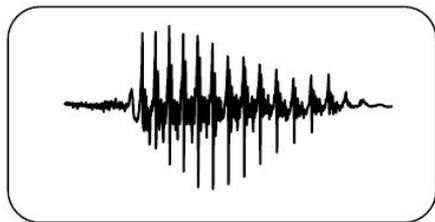
Tortues imprimées
3D et peintes,
reconnues comme
des fusils !

Athalye et al. 2017

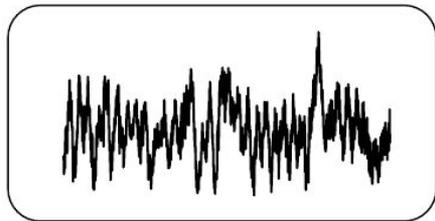


 classified as turtle  classified as rifle
 classified as other

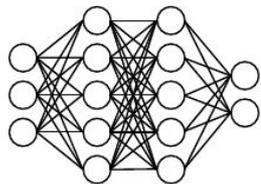
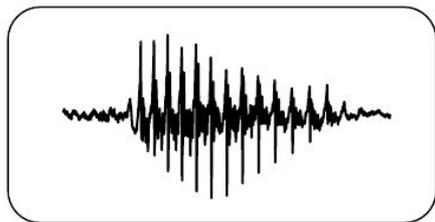
Même chose pour le son



+



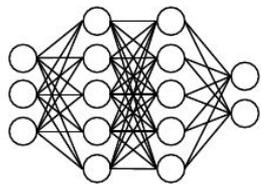
=



"it was the
best of times,
it was the
worst of times"

×

0.001



"it is a truth
universally
acknowledged
that a single"

Séquence modifiée
imperceptiblement, une
autre phrase est
reconnue par
l'assistant vocal

Importance du contexte

Exemple hors-contexte, non reconnu



Une vache ne va jamais à la plage ... dans la base d'apprentissage

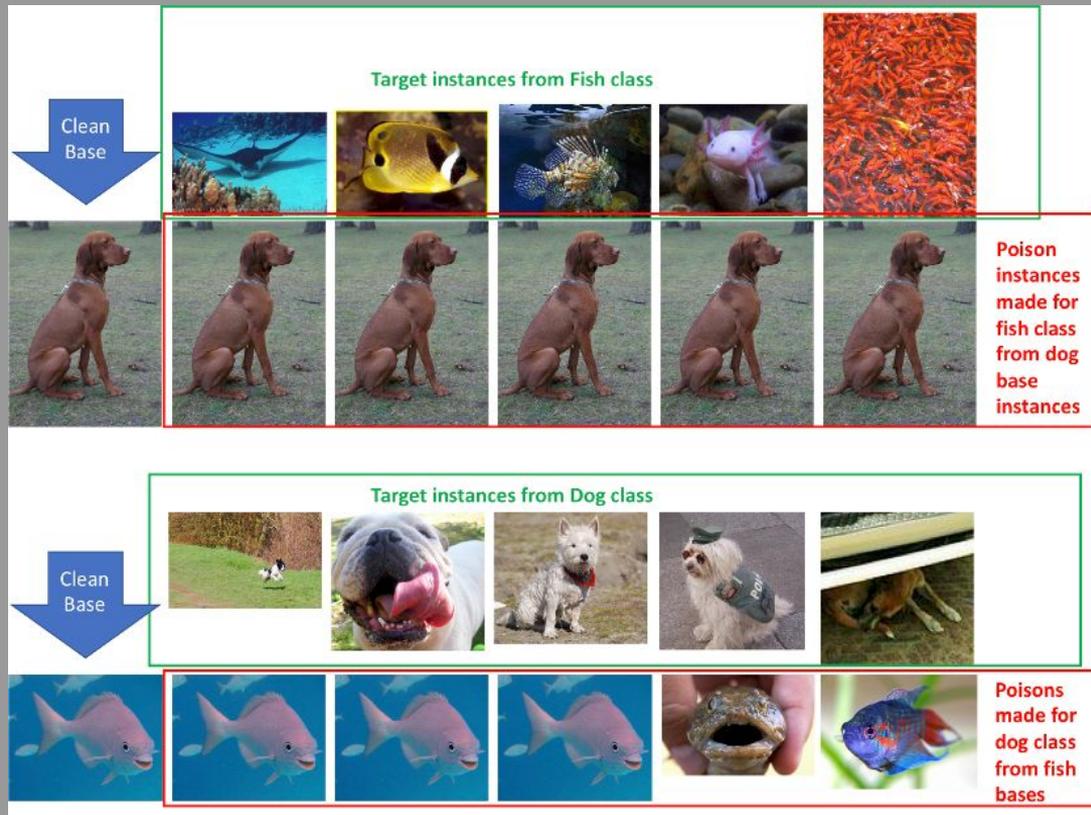
Bottou et al., 2017

Empoisonnement durant l'apprentissage

Une image de plus dans la base d'apprentissage

→

une image-cible devient mal classée



Validation et certification

- Science encore expérimentale
 - Pas de validation formelle des modèles appris
 - Validation statistique uniquement
 - Lois de la robotique d'Isaac Asimov ?
aujourd'hui impossibles à garantir
- ou toute règle forte

Intelligence Artificielle

Mythes et réalités

Menu

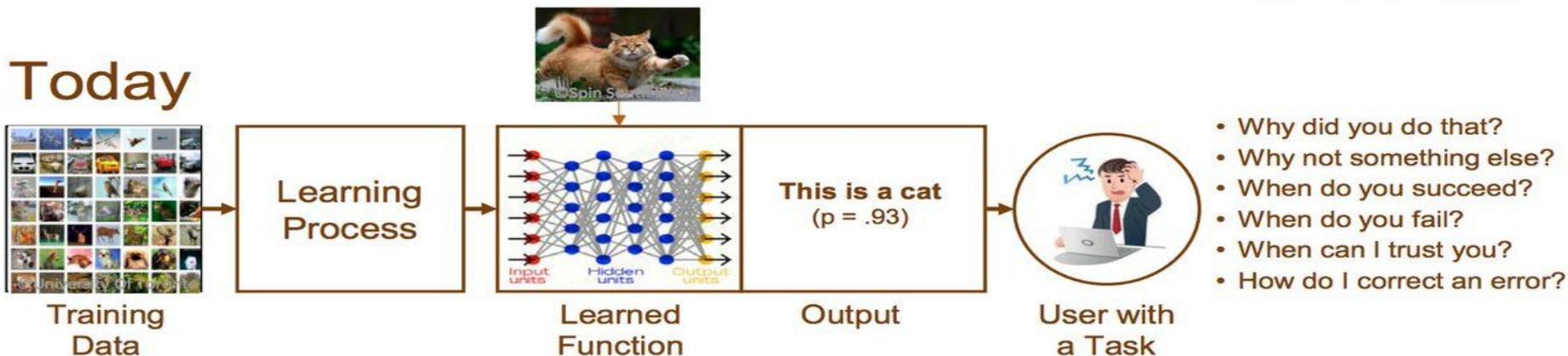
- Historique et (non-)définition
- Quelques succès récents
- Apprentissage (profond) supervisé
- Des problèmes non (encore) résolus
 - **Des modèles boîte-noires**

Interprétabilité et explicabilité

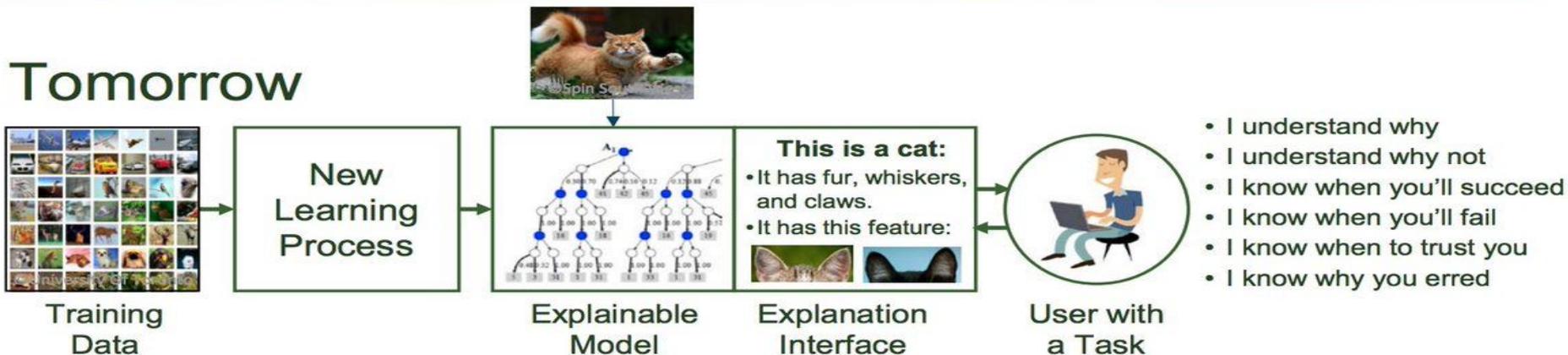
Les modèles appris sont des boîtes noires

- Concepts mal définis, et subjectifs
- Certains modèles seraient plus explicables que d'autres
 - e.g., les arbres de décision
 - ... mais pas les forêts de dizaines d'arbres

Today



Tomorrow



Interprétabilité et explicabilité

Les modèles appris sont des boîtes noires

- Concepts mal définis, et subjectifs
- Certains modèles seraient plus explicables que d'autres
 - e.g., les arbres de décision
 - ... mais pas les forêts de dizaines d'arbres

Le débat

- Qu'êtes vous prêts à perdre en précision des prédictions ?
- Il suffit de citer quelques exemples proches ...
- Nous faisons confiance à notre médecin après tout ...
- Le retour de l'IA symbolique ?

Intelligence Artificielle

Mythes et réalités

Menu

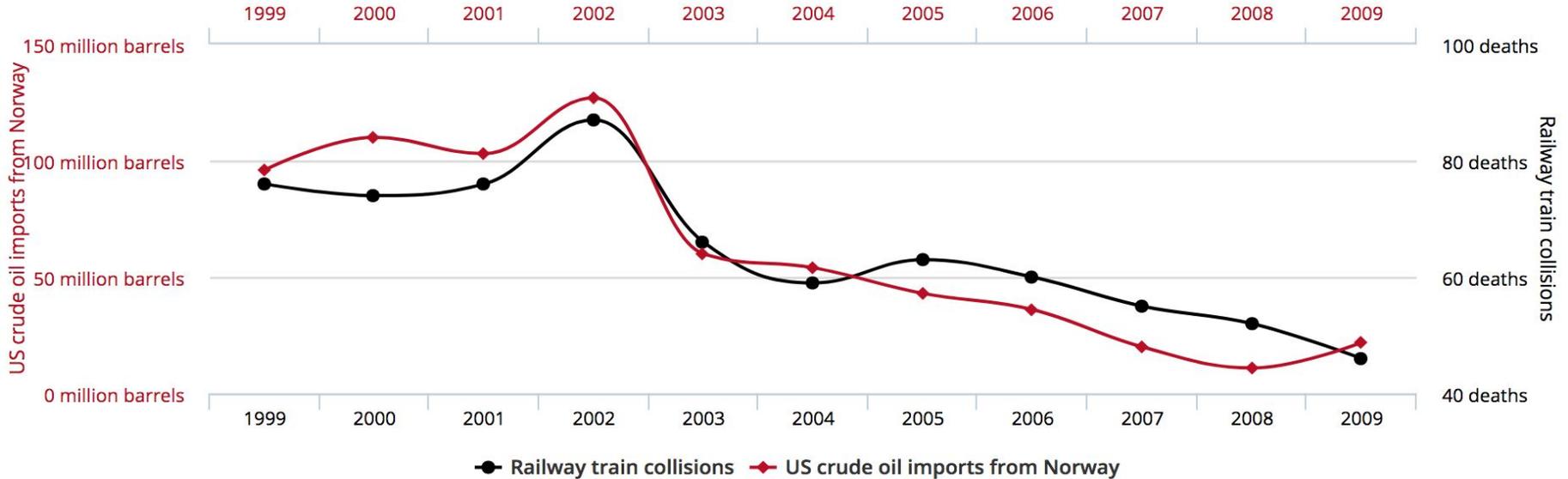
- Historique et (non-)définition
- Quelques succès récents
- Apprentissage (profond) supervisé
- Des problèmes non (encore) résolus
 - **Causalité**

On apprend des corrélations

<http://www.tylervigen.com/spurious-correlations>

US crude oil imports from Norway correlates with Drivers killed in collision with railway train

Correlation: 95.45% ($r=0.954509$)



Prévision vs prescription

- L'apprentissage supervisé apprend des modèles **déductifs**
 - S'il y a des parapluies ouverts, c'est qu'il pleut
- pas des modèles **prescriptifs**
 - et si j'ouvre mon parapluie ????
- La causalité est souvent implicite
 - mais que faire quand elle est inconnue
- Et il peut y avoir des facteurs cachés
 - Corrélation entre la santé de l'entreprise et le bien-être des employés



Intelligence Artificielle

Mythes et réalités

Menu

- Historique et (non-)définition
- Quelques succès récents
- Apprentissage (profond) supervisé
- Des problèmes non (encore) résolus
- **L'IA pour quelle société ?**

IA, travail, emploi

L'arrivée de l'IA va entraîner un chômage de masse



IA, travail, emploi

L'arrivée de l'IA va entraîner un chômage de masse

**Pt'êt ben qu'oui,
Pt'êt ben qu'non**



IA, travail, emploi

L'arrivée de l'IA va entraîner un chômage de masse

**Pt'êt ben qu'oui,
Pt'êt ben qu'non**

De très nombreux métiers sont/seront impactés

- Apprendre l'agilité en terme de métier
- Former à la créativité
- Travailler avec l'IA

Décorrélérer emploi et travail ?

IA et écologie

Le numérique pollue



IA et écologie

Le numérique pollue

VRAI

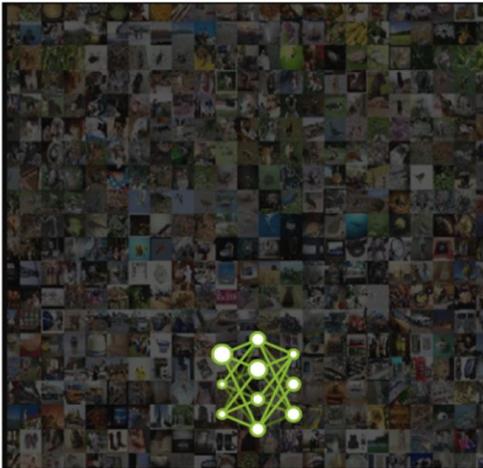


IA et écologie

L'IA (calculs et données) pollue

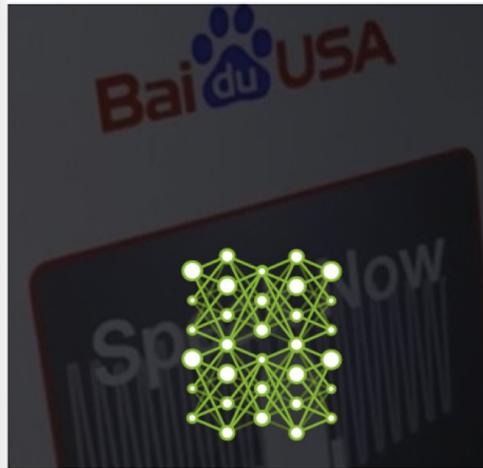
VRAI

7 ExaFLOPS
60 Million Parameters



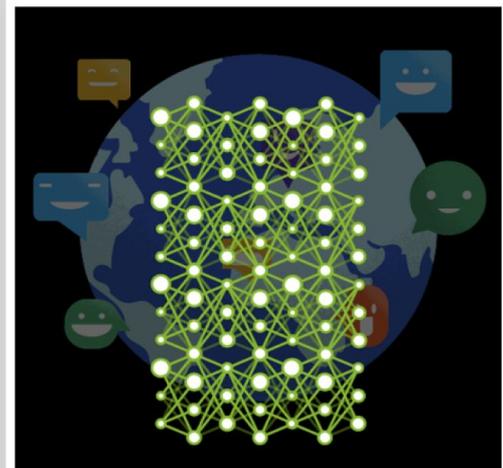
2015 - Microsoft ResNet
Superhuman Image Recognition

20 ExaFLOPS
300 Million Parameters



2016 - Baidu Deep Speech 2
Superhuman Voice Recognition

100 ExaFLOPS
8700 Million Parameters



2017 - Google Neural Machine Translation
Near Human Language Translation

IA et écologie

L'IA (calculs et données) pollue

VRAI

- Consommation énergétique, GES, terres rares
- Concurrence avec les énergies renouvelables
- → Prise de conscience

L'IA peut aider

- Prévention des risques
- Agriculture, biodiversité, ...
- Energy : production, distribution, transports, logement, ...
 - Attention aux effets rebonds

Quelle éthique pour l'IA ?

L'IA est neutre, c'est l'usage qu'on en fait qui peut poser problème



Quelle éthique pour l'IA ?

L'IA est neutre, c'est l'usage qu'on en fait qui peut poser problème

Certes, mais



Quelle éthique pour l'IA ?

L'IA est neutre, c'est l'usage qu'on en fait qui peut poser problème

Certes, mais

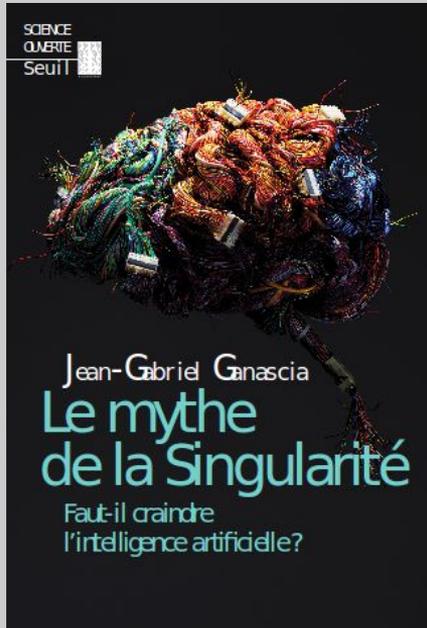
- A données biaisées, résultats biaisés (discrimination)
 - Justice prédictive, prêts bancaires, ...
 - Traçabilité des données ?
- Décisions morales
 - Le problème de la poussette pour le Véhicule Autonome
 - Les armes létales autonomes

Acceptabilité : problème scientifique, législatif, et éthique

Le mythe de la singularité

J.-G. Ganascia

Quand les machines supplanteront les humains



*Whereas the short-term impact
of AI depends on who controls it,
the long-term impact depends on
whether it can be controlled at all.*
Stephen Hawking 1er mai 2014

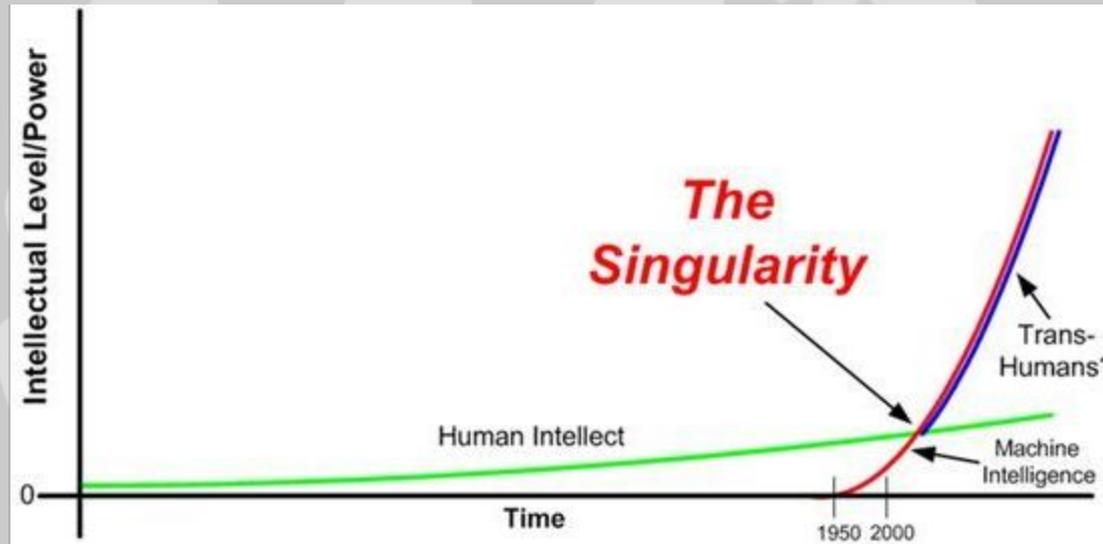
mais aussi Elon Musk, Bill Gates, ...



Le mythe de la singularité

J.-G. Ganascia

Quand les machines supplanteront les humains

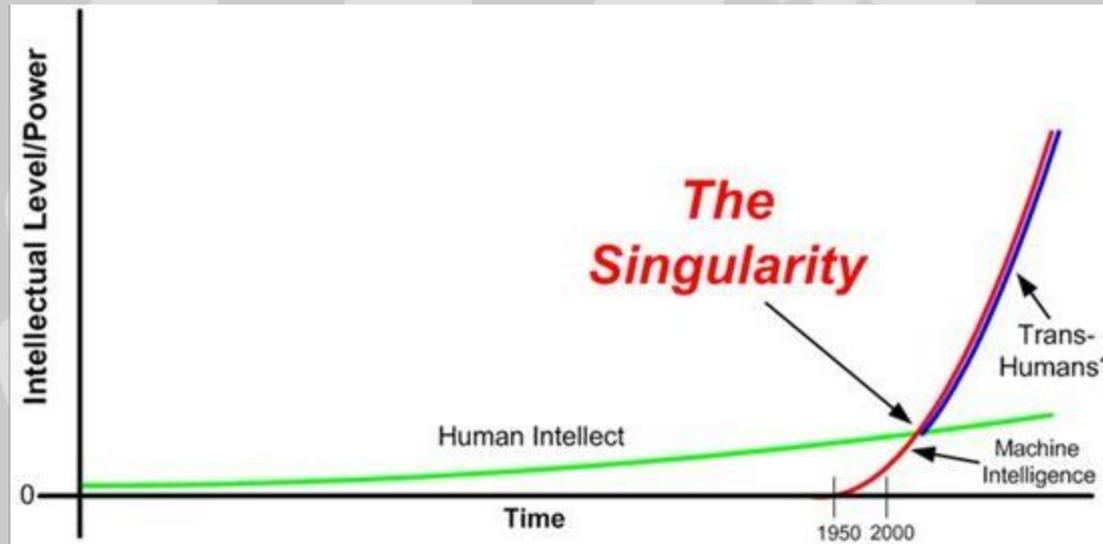


Le mythe de la singularité

J.-G. Ganascia

Quand les machines surpasseront les humains

Fantasma



Le mythe de la singularité

J.-G. Ganascia

Quand les machines supplanteront les humains

Fantasme

- IA faible : sait accomplir une seule tâche
- IA forte (AGI) : s'adapte à toute tâche
- Aujourd'hui, uniquement des IA faibles
 - même AlphaZéro
 - même les traducteurs automatiques
 - même les assistants personnels
 - ...

IA = moyen

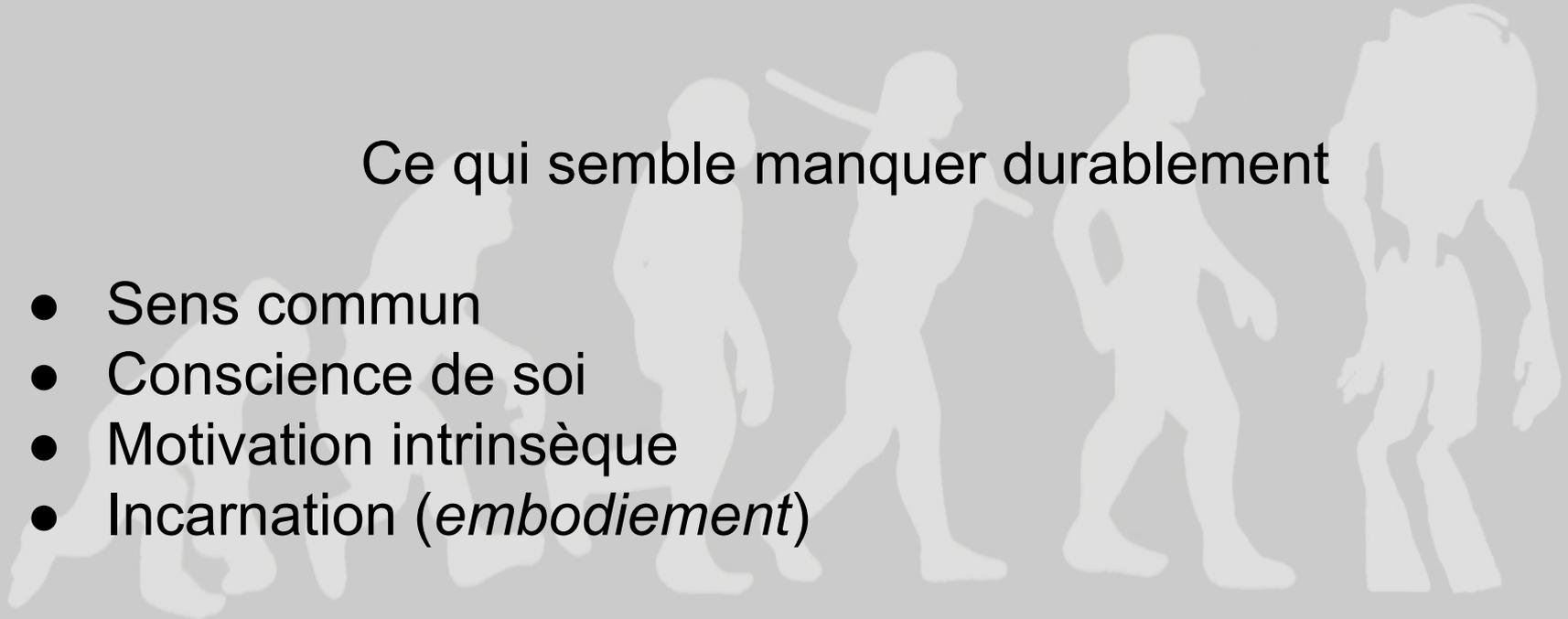
IA = but

Vers une IA forte ?

J.-G. Ganascia

Ce qui semble manquer durablement

- Sens commun
- Conscience de soi
- Motivation intrinsèque
- Incarnation (*embodiement*)



Vers une IA forte ?

J.-G. Ganascia

Ce qui semble manquer durablement

- Sens commun
- Conscience de soi
- Motivation intrinsèque
- Incarnation (*embodiment*)

S'y préparer peut néanmoins être utile (pari de Pascal ?)

A qui le crime profite ?

J.-G. Ganascia

Un épouvantail,
pour masquer les vrais dangers ?

- Contrôle accru des activités de tous
- Perte de souveraineté des individus par manque de culture numérique
- Perte de souveraineté des états au profit des géants du numérique

Way of the Future: Una religión que desarrolla un dios de Inteligencia Artificial

La organización religiosa fue fundada por un ex empleado de Google y Uber.



Le nouvel opium du peuple

Imposer la transparence

Sans transparence, risque d'un **hiver** sociétal de l'IA

- Le logiciel libre ne suffit pas
 - il faut pouvoir expérimenter
- Définir et implémenter des règles éthiques
 - Débat public ? Un CCNE bis ?
 - Contrôle citoyen ou organisme indépendant ?
- Labels, Discrimination Impact Assessment, ...

Besoin de réglementation

L'IA pour le bien de tous

AI for good



- Nombreuses avancées bénéfiques (santé, TLN, énergie, mobilité, lien social, ...)
- L'Europe peut et doit défendre ses valeurs humanistes

face aux USA et à la

Chine

C'est notre responsabilité



**Intelligence Artificielle
Mythes et réalités**

Questions ?