

L'Intelligence Artificielle, c'est quoi au juste ?

Thierry Artières

Professeur en Informatique à l'Ecole Centrale Marseille ([ECM](#))

Chercheur au Laboratoire d'Informatique et Systèmes ([LIS](#))

Membre de l'équipe d'Apprentissage automatique de Marseille ([QARMA](#))

La naissance de l'I.A.

1956 Dartmouth Conference: The Founding Fathers of AI



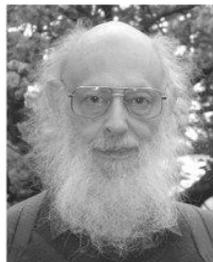
John McCarthy



Marvin Minsky



Claude Shannon



Ray Solomonoff



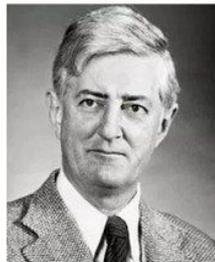
Alan Newell



Herbert Simon



Arthur Samuel



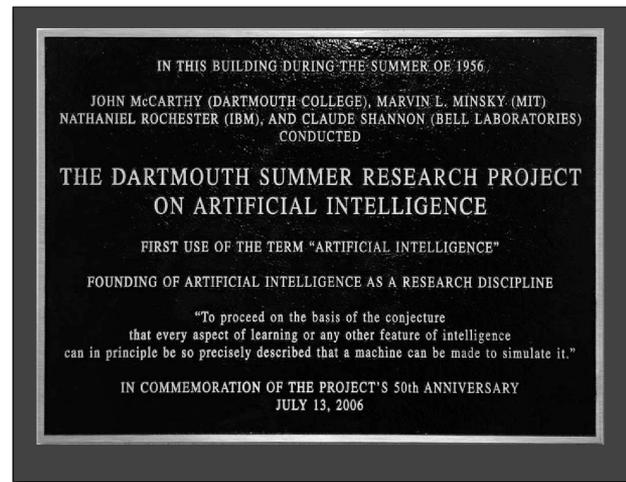
Oliver Selfridge



Nathaniel Rochester



Trenchard More



De multiples définitions de l'I.A.

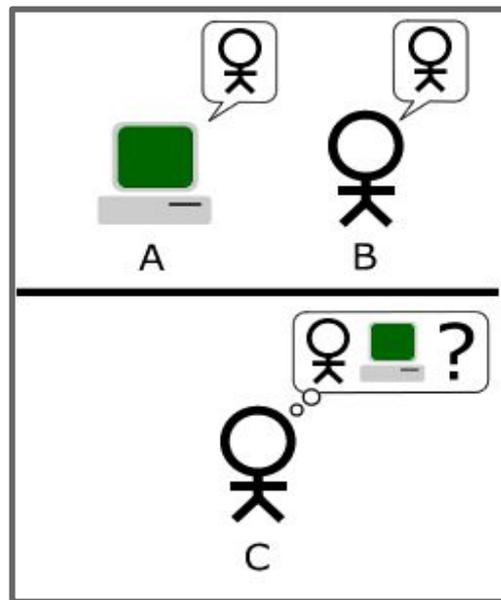
“L'IA est la science de programmer les ordinateurs pour qu'ils réalisent des tâches qui nécessitent de l'intelligence lorsqu'elles sont réalisées par des êtres humains.”

Marvin Minsky

“Les principales composantes d'un système d'IA doivent être les connaissances, le raisonnement, la compréhension du langage naturel et l'apprentissage.”

Alan Turing

Le test de Turing



Composantes majeures de systèmes d'IA

Joueurs artificiels (Echec, Go)

Résolution de problèmes

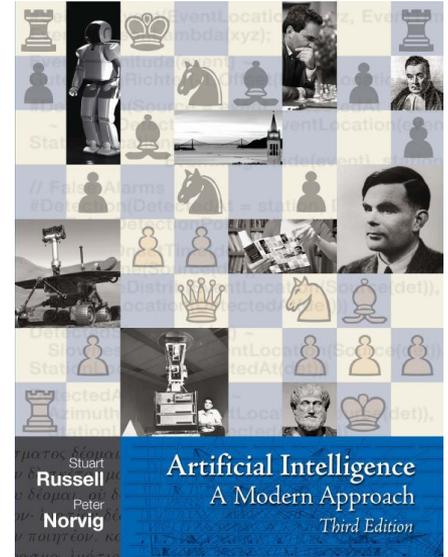
Raisonnement, représentation des connaissances et planification

Raisonnement dans l'incertain

Apprentissage

Communication, perception et action

Vision, Langage naturel, interfaces vocales...

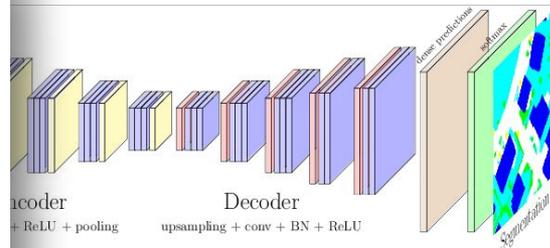
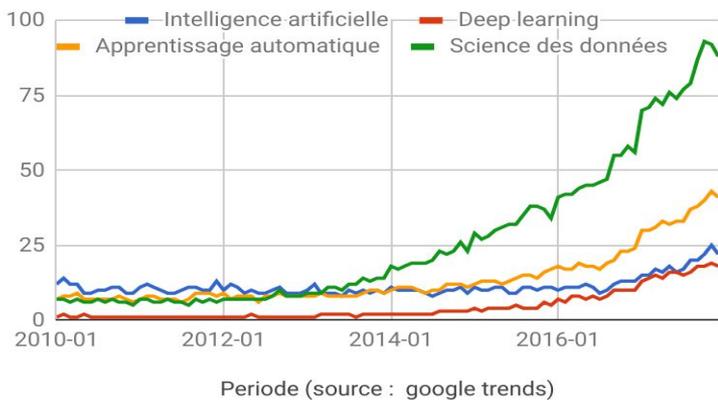


Quelques dates

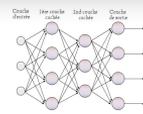
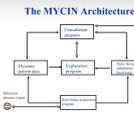
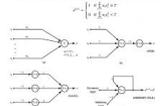
Les trois vagues des réseaux de neurones



Marché du travail et éducation



McCulloch-Pitts Neuron Model



1940

1950

1960

1970

1980

1990

2000

2010

2018

IA forte et IA faible

- IA faible

[Wikipedia] L'intelligence artificielle faible est une intelligence artificielle (non-sensible) qui se concentre sur une tâche précise.

- IA forte ou générale

Une machine dotée de conscience, de sensibilité et d'esprit

Une machine capable d'appliquer l'intelligence à tout problème plutôt qu'un problème spécifique

Algorithmes et logiciels

Qu'est-ce qu'un algorithme ?

[Source Wikipedia] : Un algorithme est une suite finie et non ambiguë d'opérations ou d'instructions permettant de résoudre une classe de problèmes

Programmation = implémentation d'un algorithme dans un langage compris par un ordinateur

Où les opérations sont prises dans un ensemble d'instructions disponibles dans le langage de programmation (par exemple la commande *print* en python)

Algorithme et programmation

Tout est programmable...

Le système d'exploitation d'un ordinateur

Le tir d'une fusée

Un jeu vidéo

Via un peu d'intelligence humaine (équations de la physique, de la cinématique, etc)

... ou presque

Car on ne dispose pas d'un algorithme pour tout !



**Des applications
de l'I.A.
d'aujourd'hui**

Vision

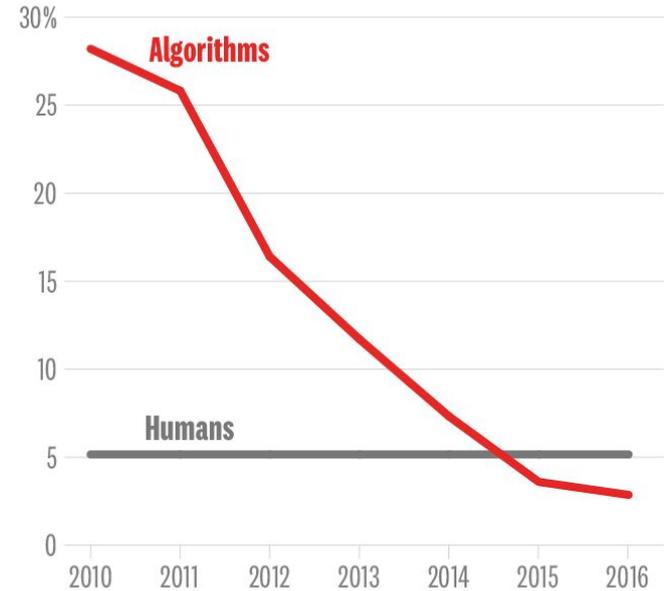


[Farabet et al., IEEE PAMI, 2012]

[YOLO 9000 A State-Of-The-Art, Real-Time Object Detection System That Can Detect Over 9000 Object Categories.](#)

Thierry Artières – ECM - IIS

VISION ERROR RATE



SOURCE ELECTRONIC FRONTIER FOUNDATION © HBR.ORG

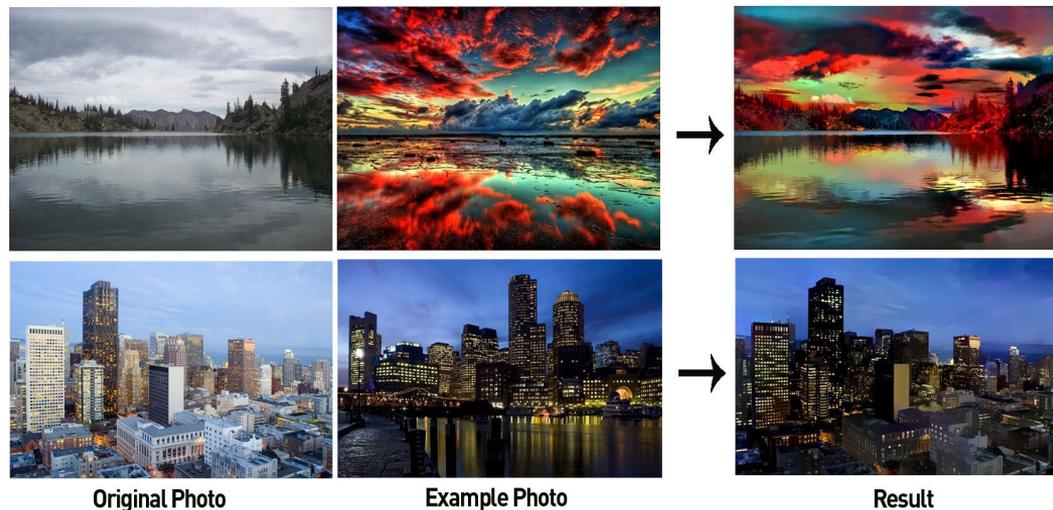
Langage naturel et assistants vocaux



Génération d'images ou de voix

Visages artificiels par Nvidia

Edition d'images et vidéos



Jeux

Echecs, Go etc

Vers les (jeux de) stratégies

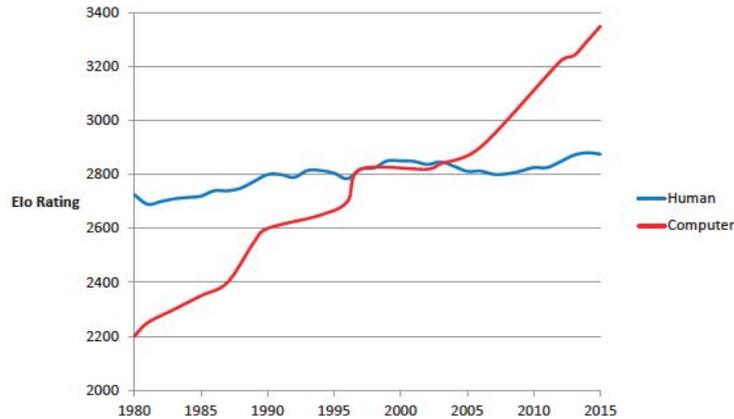


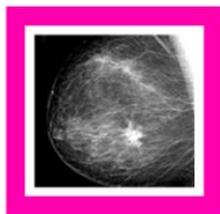
FIGURE 2.3 Elo scores—a measure of competency in competitive games—showing the chess-playing competency of humans and machines, measured over time. SOURCE: Courtesy of Murray Campbell.



Santé

De plus en plus de diagnostics sont mieux prédits par des technos d'I.A.

.Détection de cancer du sein



The Digital Mammography DREAM Challenge

Build a model to help reduce the recall rate for breast cancer screening

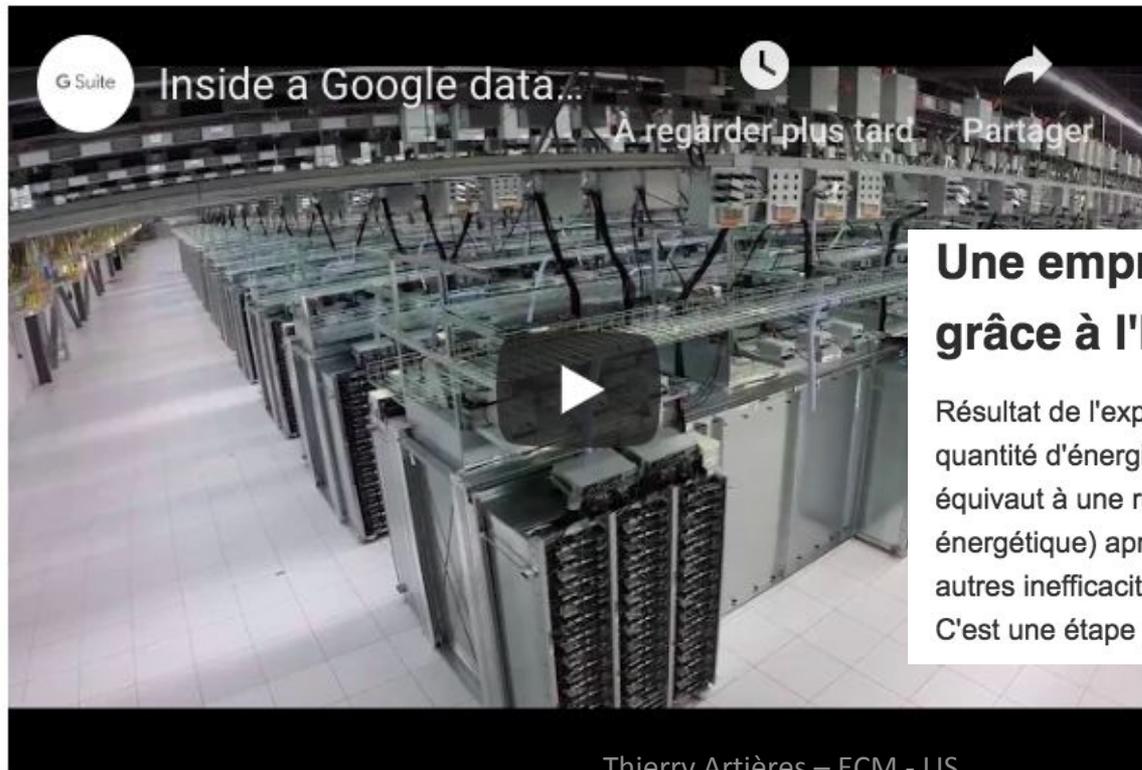
Learn more & register to participate here: www.synapse.org/Digital_Mammography_DREAM_Challenge



.Détection de mélanomes

.Détection de troubles cardiaque dans les EEG

Optimisation de l'énergie



Une empreinte écologique réduite grâce à l'IA

Résultat de l'expérience ? Une réduction de 40% de la quantité d'énergie nécessaire au refroidissement, "ce qui équivaut à une réduction de 15% du PUE (efficacité énergétique) après avoir comptabilisé les pertes électriques et autres inefficacités du non refroidissement, détaille Google. C'est une étape phénoménale qui a été franchie".

Voitures autonomes

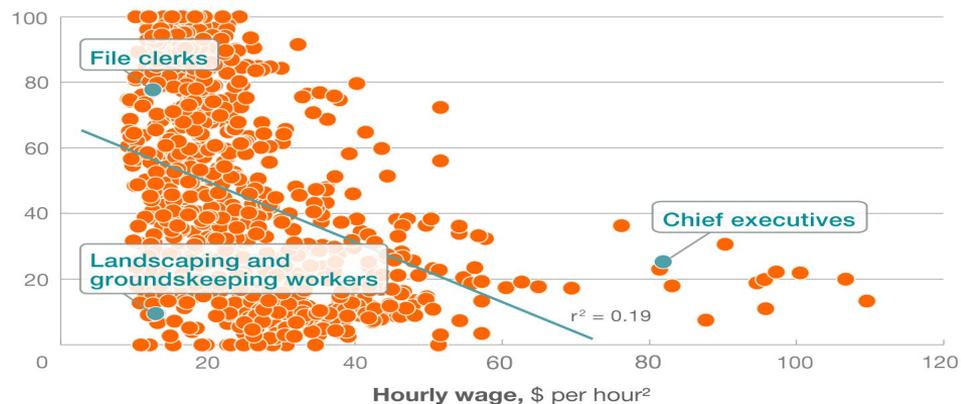


Vous ne passerez pas à coté !

The hourly-wage rate alone is not a strong predictor of automatability, despite some correlation between the two.

Comparison of wages and automation potential for US jobs

Ability to automate, % of time spent on activities¹ that can be automated by adapting currently demonstrated technology



¹Our analysis used “detailed work activities,” as defined by O*NET, a program sponsored by the US Department of Labor, Employment and Training Administration.

²Using a linear model, we find the correlation between wages and automatability in the US economy to be significant (p-value <0.01), but with a high degree of variability ($r^2 = 0.19$).

Source: O*NET 2014 database; McKinsey analysis

**Sous le capot de
l'I.A.**

**L'Apprentissage
Automatique**

Quand utiliser l'Apprentissage Automatique ?

Quand on ne sait pas comment programmer

Quand on n'a pas d'algorithme !!



- Perception : audition, vision
- Langage naturel : résumés, traduction
- Prédiction : bourse, consommation énergie...
- Recommandation : Amazon, Netflix
- Profilage et publicité sur internet
- Robotique

L'A.A., comment ça marche ?

Les programmes que l'on réalise avec de l'AA réalisent une fonction

$$f: X \rightarrow Y$$
$$x \mapsto f(x)$$

Par exemple

= Images

= Catégories d'objets

Comment ça marche ?

On dit que l'on *apprend (itérativement)* la fonction / le programme f à l'aide d'un ensemble d'exemples (x,y)



Voiture



Mise à
jour de f



Jaguar



Mise à
jour de f

...

Une fois la fonction f apprise

On utilise la fonction f sur d'autres données

C'est une difficulté majeure : la généralisation

- On met au point un système sur des données d'apprentissage
- On l'utilise sur d'autres
- Comment peut-on être sûr que cela va fonctionner ?

Avant d'aller

plus loin...

Tout est vecteur

Un vecteur

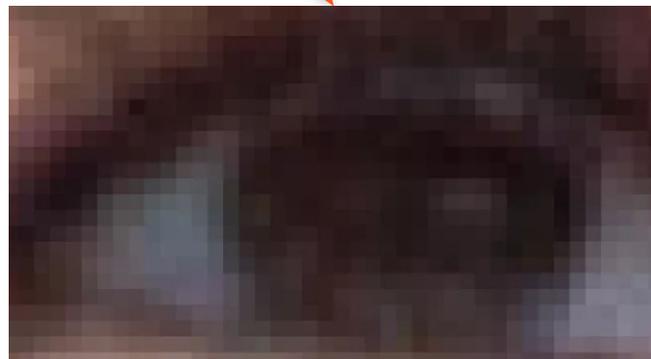
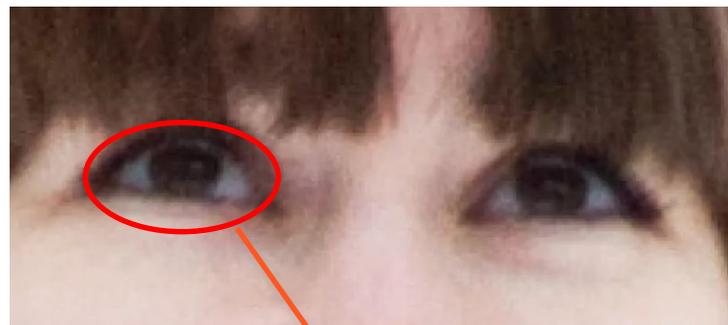
- . C'est un ensemble de valeurs (dans un certain ordre)
- . Toutes les données que l'on manipule en IA peuvent être mises sous forme vectorielle
- . Par exemple sur un site marchand *vous* êtes représentés par un vecteur de plusieurs *caractéristiques*

[*nb_visites_par_mois, Facture_moyenne, Adresse, Age, ...*]

Un vecteur

- . C'est un ensemble de valeurs (dans un certain ordre)
- . Toutes les données que l'on manipule en IA peuvent être mises sous forme vectorielle
- . La phrase « Je suis trop fatigué ! » ...
 - ... peut être représentée comme le vecteur [2023, 8702, 9340, 1673, 12] ...
 - ...en utilisant un index des mots du vocabulaire de la langue

Une image c'est une collection de pixels

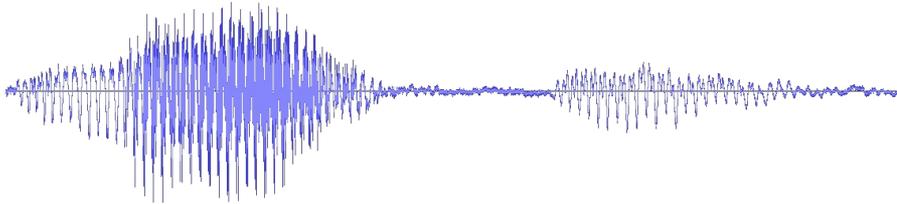


12, 23, 56, 34, 25, ...

Thierry Artières - ECM - LIS

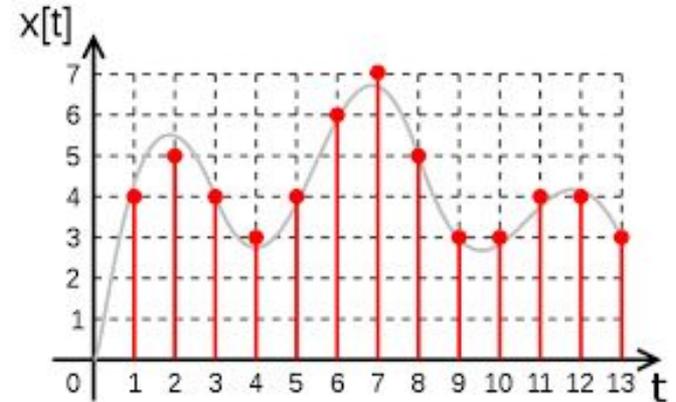
Un son aussi (Cf échantillonnage de Shannon)

· Tout signal sur un ordinateur est échantillonné



Signal de parole=signal 1D

Echantillonnage



→ [0,4,5.5,...]

A partir de maintenant

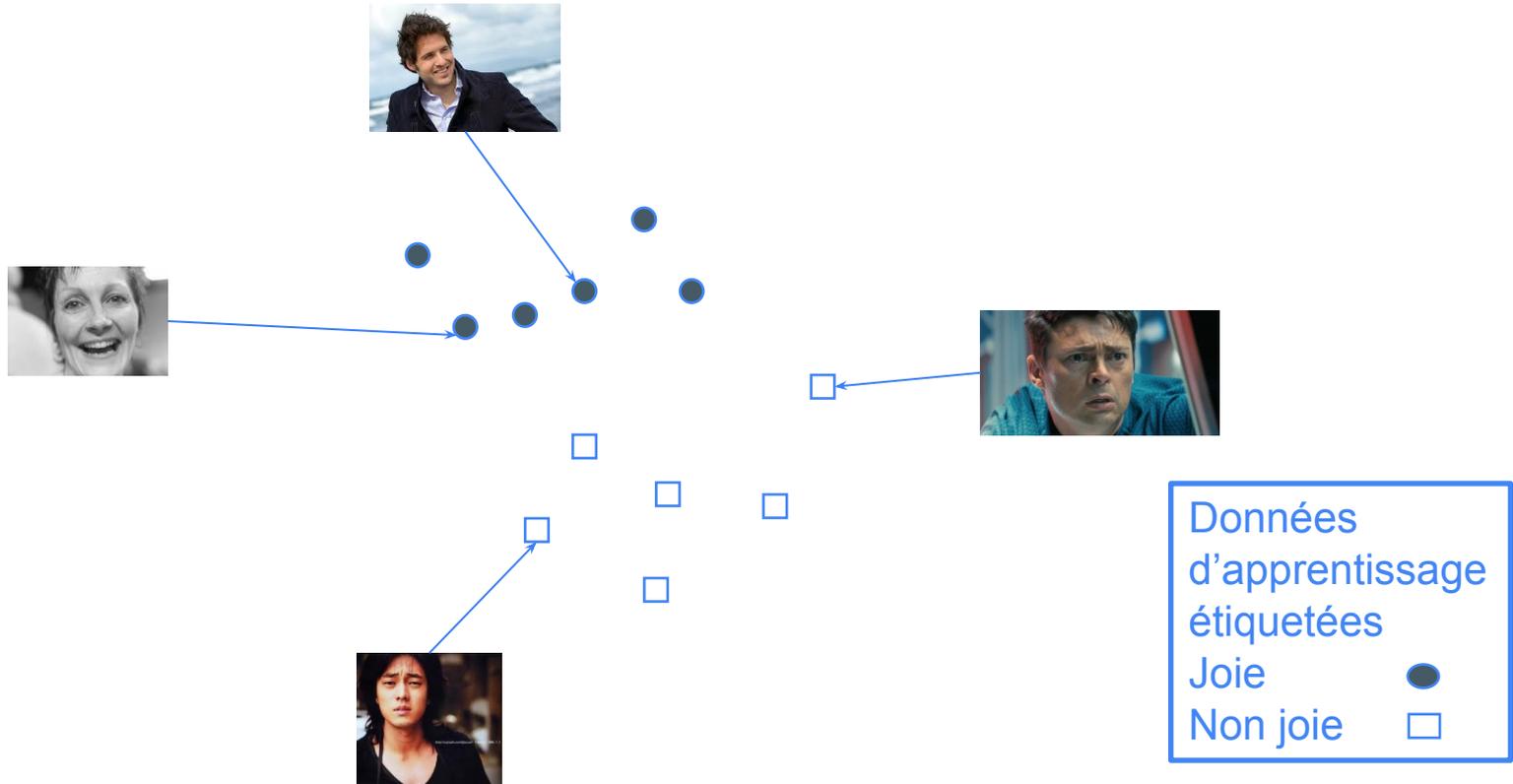
- Une donnée, quelle qu'elle soit, est un vecteur de valeurs réelles
- Pour simplifier on se place en dimension 2
- On s'intéresse donc
 - À des images à 2 pixels
 - A des sons de 2 échantillons
- Mais tout cela se généralise très bien à
 - Des images de 8Megapixels = 8 millions de pixels
 - ..

La classification (et la généralisation)

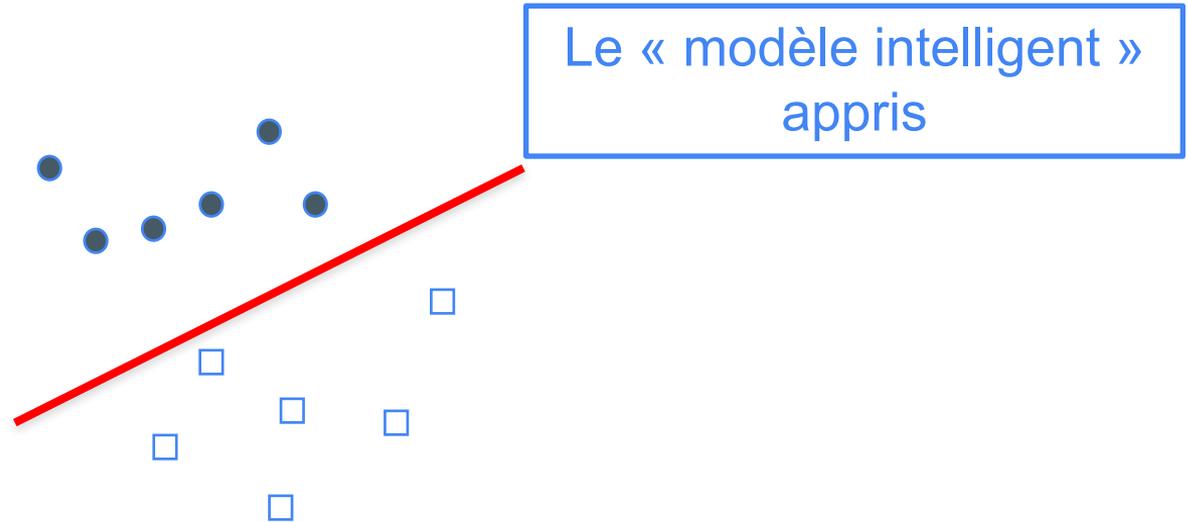
$$f: X \rightarrow \{Classe_1, Classe_2, \dots, Classe_K\}$$

$$x \mapsto f(x)$$

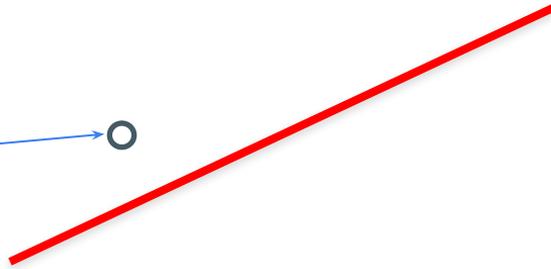
Classification d'images : le cadre



Classification en 2D



Classification en 2D: la généralisation



L'utilisation
du modèle
en test

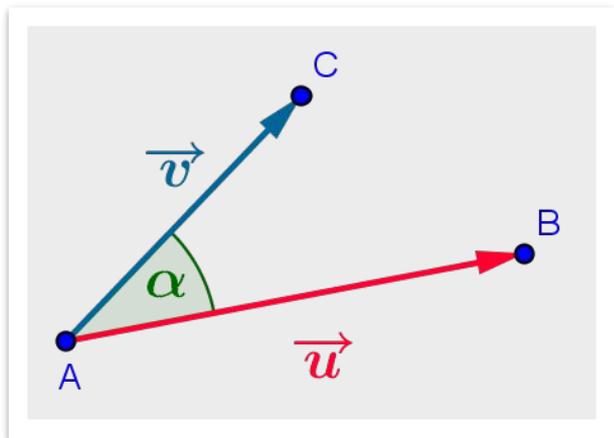


Joie

**La classification
avec un
perceptron :
produits
scalaires and co**

C'est quoi déjà un produit scalaire ?

$$\vec{u} \cdot \vec{v} = \|\vec{u}\| \times \|\vec{v}\| \times \cos(\vec{u}, \vec{v})$$



$$\vec{u} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \vec{v} = \begin{pmatrix} x' \\ y' \end{pmatrix}$$

$$\vec{u} \cdot \vec{v} = xx' + yy'$$

Classifier en dimension 2 avec un perceptron

Le perceptron

défini par un vecteur de paramètres

$$\vec{w} = \begin{pmatrix} w_x \\ w_y \end{pmatrix}$$

Fonctionnement du perceptron

Le perceptron classe une donnée

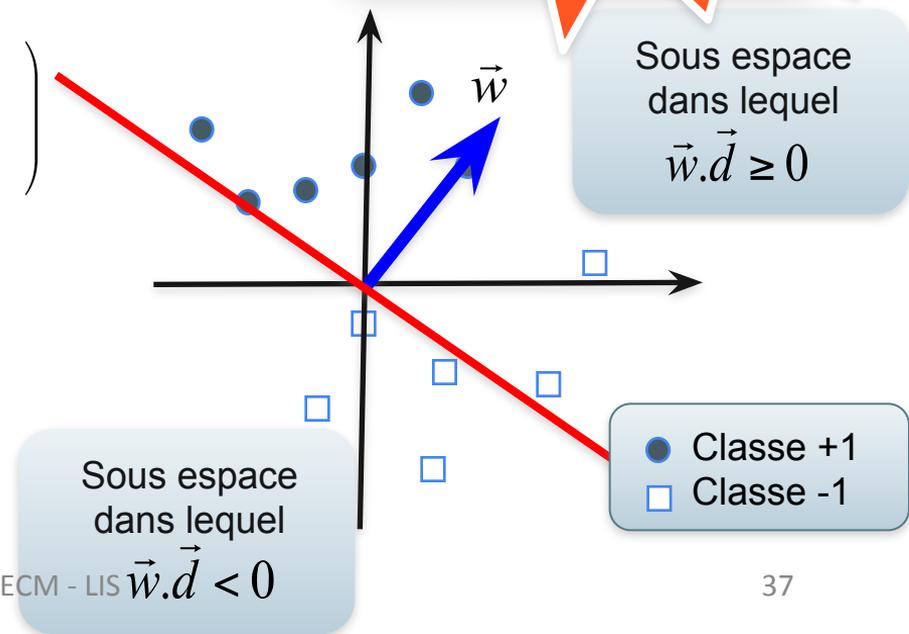
$$\vec{d} = \begin{pmatrix} d_x \\ d_y \end{pmatrix}$$

dans la classe 1 si $\vec{w} \cdot \vec{d} \geq 0$

dans la classe 2 si $\vec{w} \cdot \vec{d} < 0$

Le vecteur \vec{w} définit deux régions

Ici il y a des erreurs de classification

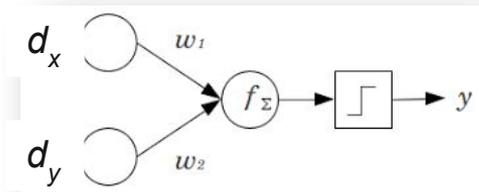


Classifier en dimension 2 avec un perceptron

Le perceptron

défini par un vecteur de paramètres

$$\vec{w} = \begin{pmatrix} w_x \\ w_y \end{pmatrix}$$



Fonctionnement du perceptron

Le perceptron classe une donnée

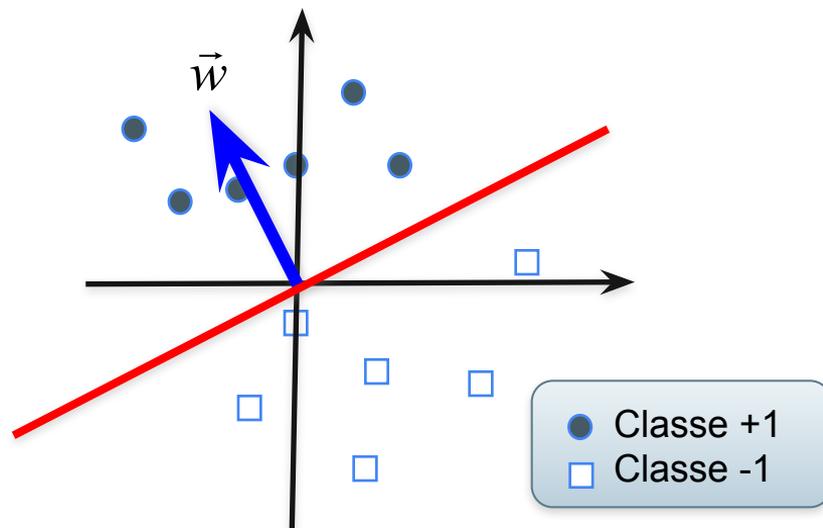
dans la classe 1 si $\vec{w} \cdot \vec{d} \geq 0$

dans la classe 2 si $\vec{w} \cdot \vec{d} < 0$

$$\vec{d} = \begin{pmatrix} d_x \\ d_y \end{pmatrix}$$

Apprentissage du perceptron ?

Trouver un (le) bon \vec{w} !!!



Apprentissage du perceptron pour la classification binaire

Etant donné un ensemble d'apprentissage de N exemples $\{(d_i, y_i), i = 1 \dots N\}$

Choisir un vecteur initial aléatoire w

Itérer

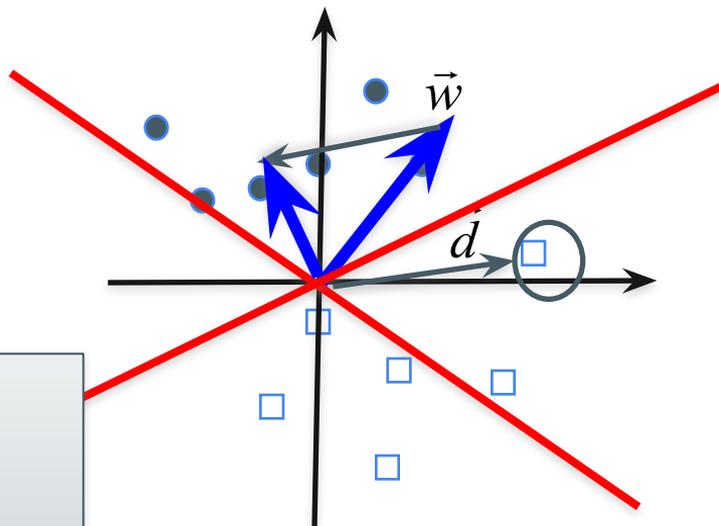
Choisir un exemple i au hasard

Si l'exemple i est bien classé ne rien faire

Sinon modifier w suivant $w = w + y_i d_i$

Une itération

- Classe +1
- Classe -1



Choisir un vecteur initial aléatoire w

Itérer

Choisir un exemple i au hasard

Si l'exemple i est bien classé ne rien faire

Sinon modifier w suivant

$$w = w + y_i d_i$$

$$y \times \vec{w} \cdot \vec{d} < 0$$

x mal classé

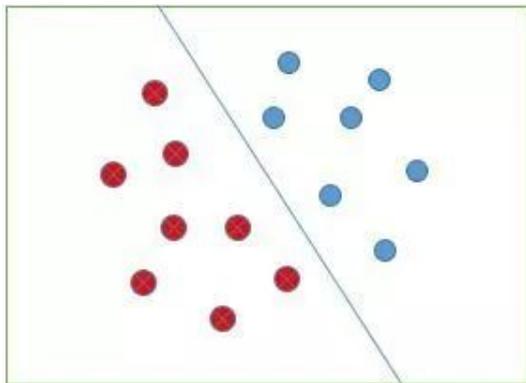
$$w = w + y \times d$$

Convergence en théorie et en pratique

Preuve de convergence : On peut montrer que si les données sont « séparables » linéairement (comme ci dessous), alors l'algorithme du perceptron est certain de converger en un nombre

Preuve
mathématique pour
l'algorithmique

Illustration



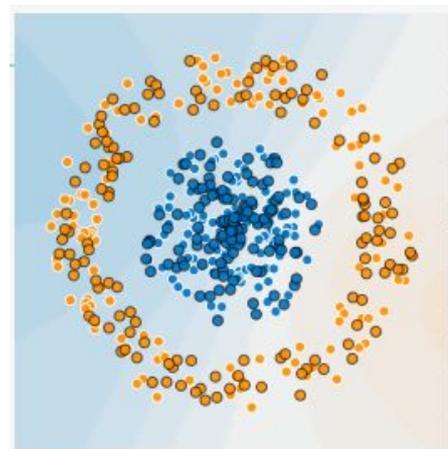
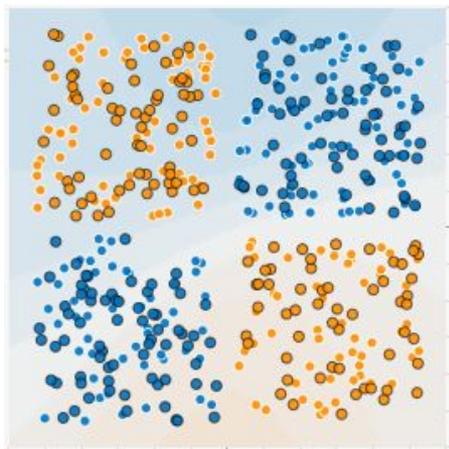
**Et pour des problèmes
plus compliqués ?**

Des

**programmes/fonctions
plus compliqués !**

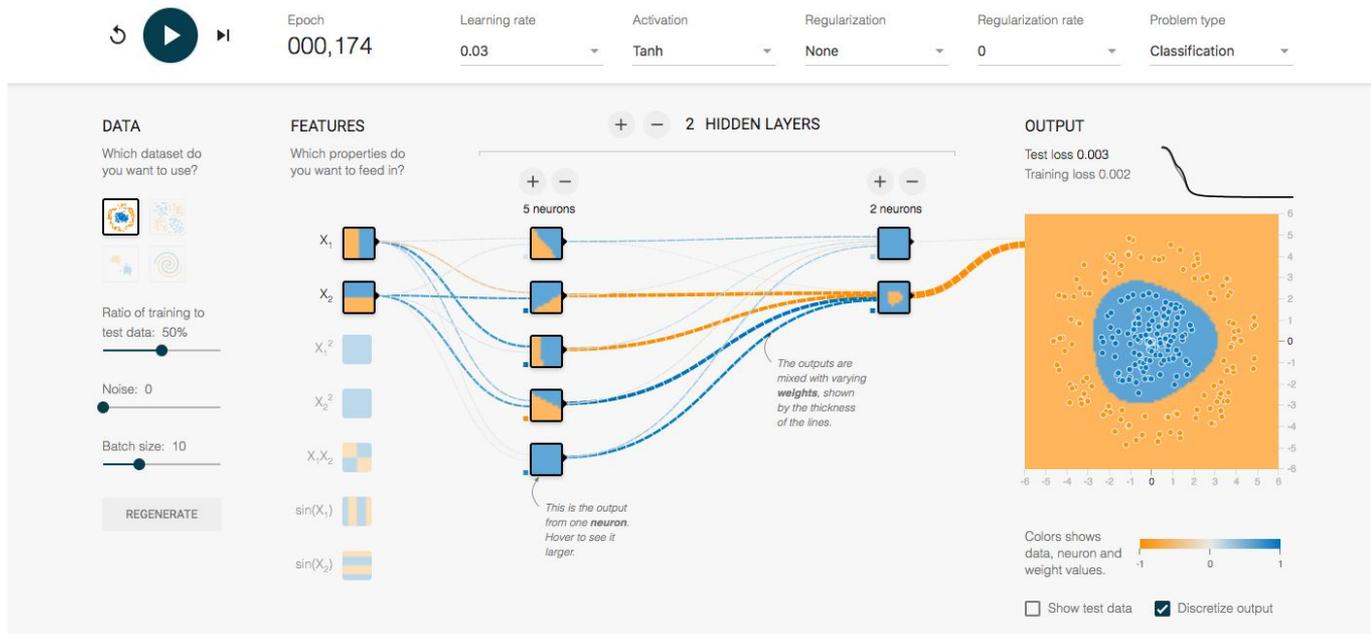
Données non linéairement séparables

Cas le plus fréquent dans la « vraie vie » des données



En pratique

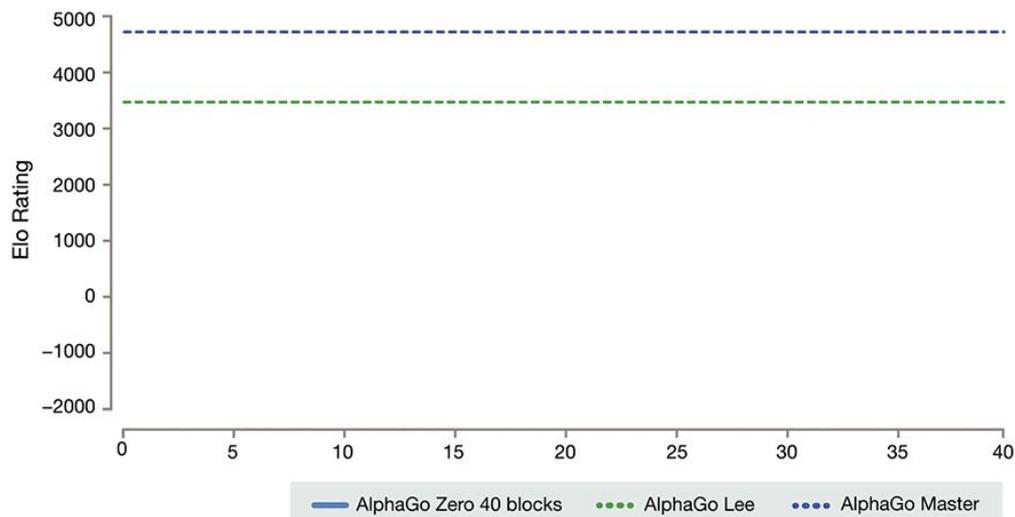
[URL](#)



**Apprentissage
machine et
apprentissage
humain**

L'apprentissage est progressif (et long), comme pour nous !

Courbe d'apprentissage d'AlphaZero [\[URL\]](#)



Mais

Nombre de parties jouées par AlphaGo Zero pour arriver à son niveau ?

Dans les trois premiers jours seulement, AlphaGo Zero a joué **4,9 millions de jeux**

Nombre d'images vues par les algorithmes de classification d'images ?

Des millions

Nombre de paires de phrase vues par un outil de traduction automatiques ?

Des millions

On n'apprend visiblement pas de la même façon !!

Généricité de l'apprentissage et des connaissances

- . Joueur de Go => un apprentissage
- . Joueur d'échecs => un autre apprentissage
- . Joueur de Go sur un go-Ban 9x9 => un autre apprentissage

- . Pas de « capitalisation »
- . Pas de représentation efficace des connaissances / raisonnements

Je sais entrainer des machines,
mais je cherche toujours à comprendre ce
qu'est l'apprentissage

[Stéphane Canu, Professeur, Université de Rouen, Marseille, 2018]

Et ensuite ?

Du côté de l'IA

Quelques limites de l'I.A.

Interprétabilité : Les outils élaborés par apprentissage automatique sont des boîtes noires



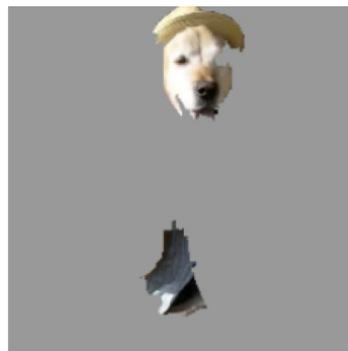
(a) Original Image



(b) Explaining *Electric guitar*



(c) Explaining *Acoustic guitar*



(d) Explaining *Labrador*

Quelques limites de l'I.A.

La robustesse au biais de conception

« L'application aveugle de l'apprentissage-machine risque d'amplifier les biais qui sont présents dans les données » [Tolga Bolukbasi, NIPS 2016].

L'homme est à la femme ce que le roi est à quoi ? « À la reine »

L'homme est à la femme ce que le chirurgien est à... « l'infirmière »

... ce que le programmeur informatique est à... « la femme au foyer »

... ce que l'architecte est à « la décoratrice »

... ce que le commerçant est à « la ménagère ».

Quelques limites de l'I.A.

La robustesse au piratage

Les outils d'apprentissage automatique ne sont pas (encore) prévus pour être "attaqués"

"A peine lancée, une intelligence artificielle de Microsoft dérape sur Twitter
L'entreprise américaine a lancé Tay, un « chatbot » censé discuter avec des adolescents sur les réseaux sociaux. Mais des propos racistes se sont glissés dans ces échanges."

Le Monde, Mars 2016

Du côté de la société

Domination des grands acteurs dans tous les domaines technologiques

Et notamment les domaines régaliens de l'état

.Santé

.Droit

.Education

.Sécurité

...

.**Pour le moment** beaucoup de choses sont gratuites

La singularité ?

Le jour où la machine sera plus intelligente que l'homme (et le dominera?)

Rien ne permet vraiment de présager que cela puisse arriver

On n'a pas d'idée des technologies qui le permettraient

Est-ce que mon (futur) métier va disparaître ?

Quel que soit le métier auquel vous vous destinez

(Le futur est aux codeurs !!)

.Mieux vaut maîtriser l'informatique

.Mieux vaut maîtriser les concepts d'IA pour savoir à quel point et quand on peut lui faire confiance. L'IA sera probablement le plus souvent une aide qu'un remplaçant

.Mieux vaut ne pas s'orienter vers un métier dont l'IA spécifique / faible pourrait s'emparer

