

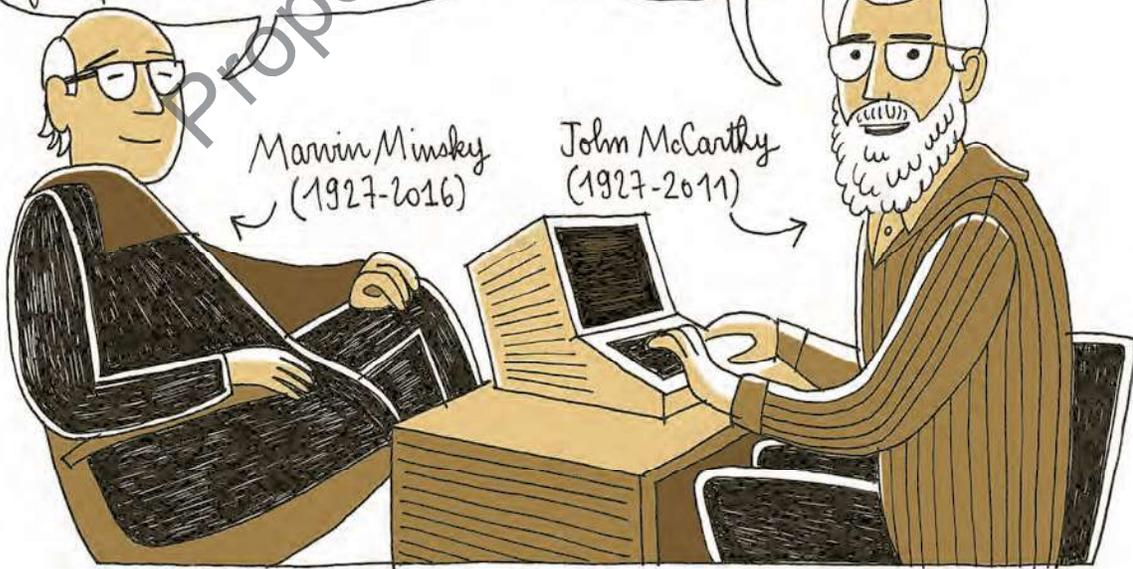
Alan Turing, l'un des fondateurs de l'informatique, met en place un test pour déterminer si l'ordinateur est capable de penser comme un humain.



Les premiers chercheurs vont suivre cette piste pour créer de l'intelligence artificielle. C'est au cours de la conférence de Dartmouth en 1956 que le nom de cette discipline est pour la première fois utilisé. Et dans la bouche des fondateurs, ça donne ça :

L'IA, c'est la science de faire faire à des machines des choses qui demanderaient de l'intelligence si elles étaient faites par des humains.

Je dirais même plus : l'IA c'est la science et l'ingénierie de créer des machines intelligentes, en particulier des programmes informatiques.





Et pour décrire tout ça, il faut...

...des algorithmes!

Une liste très précise de tâches à effectuer les unes après les autres.

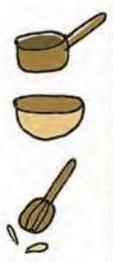


```
func somme(nombre_a, nombre_b)
for i in 1, 2, ..., taille_nombre;
sum = nombre_a[i] + nombre_b[i];
resultat[i] = sum;
ret [i+1] = sum [i+1];
return resultat;
```



Ça marche comme une recette de cuisine :

- étape 1** faire fondre dans une casserole 200 g de chocolat et 50 g de beurre
- étape 2** dans un saladier, placer 100 g de sucre, 3 œufs et 50 g de farine
- étape 3** mélanger



Les pionniers avaient des idées très ambitieuses. Très vite, ils ont reçu de gros financements.

Mais les résultats n'ont pas été au rendez-vous.

La technologie s'est retrouvée limitée pour certaines tâches: les chatbots ou la reconnaissance d'images par exemple.

Les investisseurs ont coupé court.

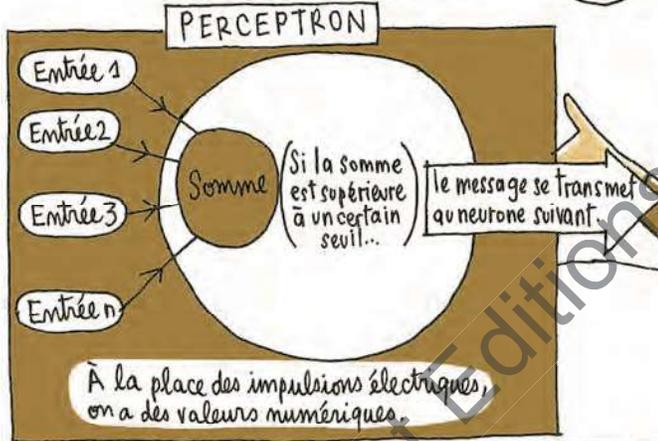
On rentre? Fait un peu froid, non?

On a appelé ça: **L'HIVER DE L'IA** (il y en a eu plusieurs)

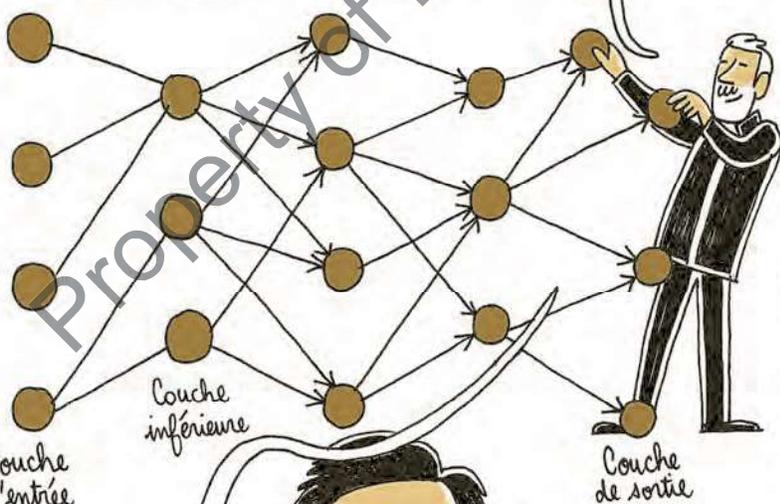


Pour autant, la recherche continue !

En 1957, Frank Rosenblatt, un chercheur américain, modélise le perceptron, une version mathématique très simplifiée du fonctionnement d'un neurone.



Pendant très longtemps, on a eu du mal à faire fonctionner ces réseaux de neurones pour que les valeurs se propagent proprement.



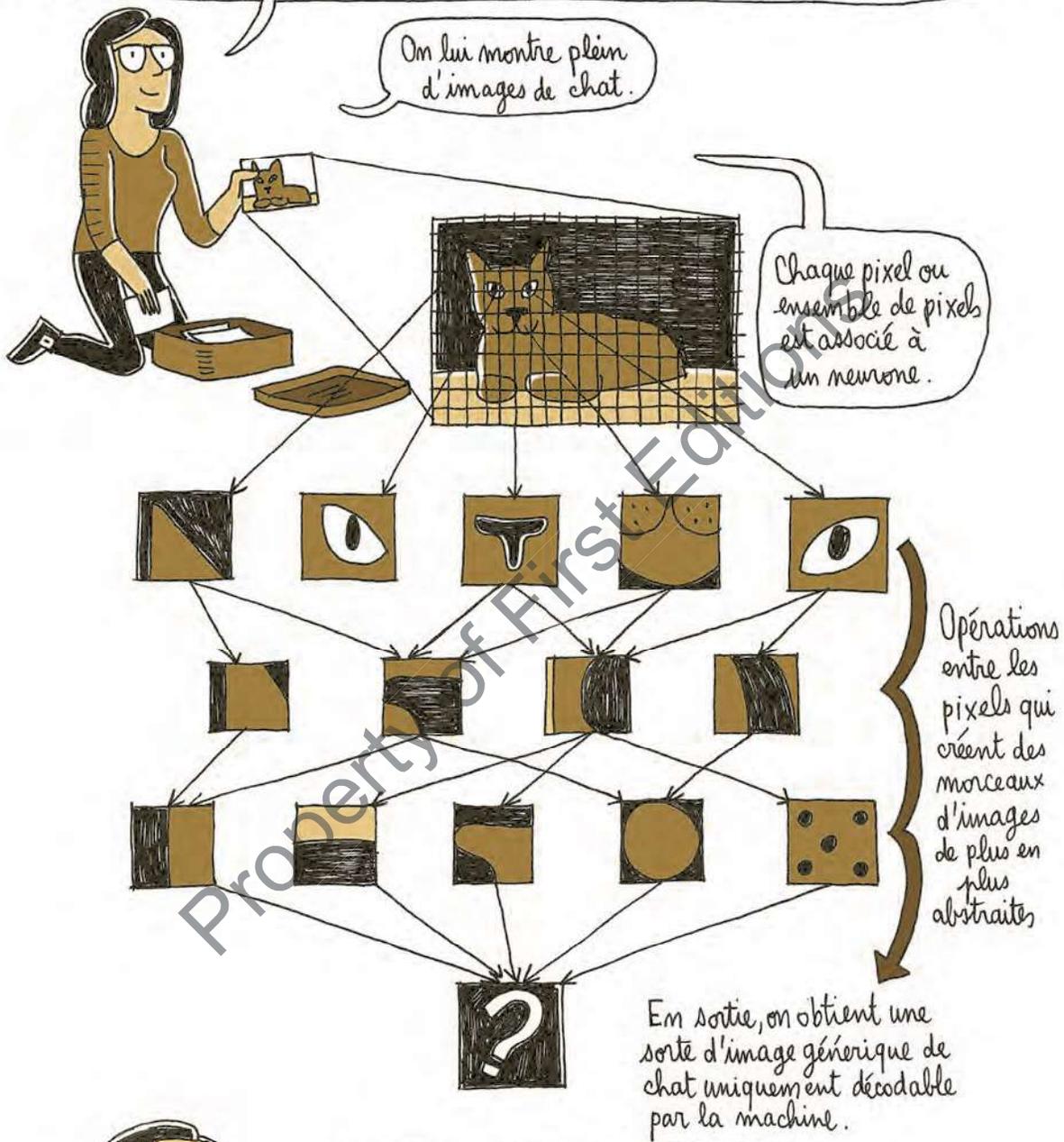
Jusqu'à ce qu'un chercheur français, Yann Le Cun, réussisse à les agencer correctement dans les années 90!



Yann Le Cun
prix Turing 2018

On appelle ça des réseaux de neurones profonds. On passe de l'IA symbolique à l'IA numérique.

Avec ces réseaux de neurones profonds, on va apprendre à la machine à réaliser des tâches à partir de beaucoup de données. Exemple : reconnaître un chat sur une image.

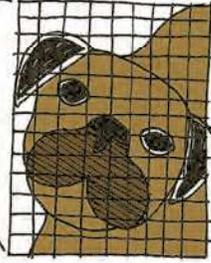


C'est la méthode la plus utilisée aujourd'hui. On l'appelle l'apprentissage profond ou deep learning.

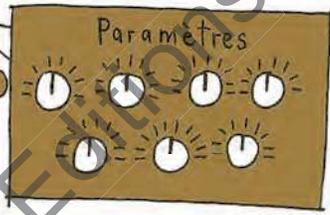
Il y a trois types d'apprentissage.

① L'apprentissage supervisé

On montre à la machine plein d'images de chien et de chat préalablement étiquetées



En sortie, elle doit dire si l'image est un chien ou un chat.



En fonction, elle ajuste ses paramètres.



Au début, elle se trompe.

Elle compare sa réponse avec la réponse attendue pour déterminer la mesure d'erreur.

Une fois que le réseau de neurones est suffisamment entraîné, on passe à la phase de prédiction avec des images inconnues.



CLAP!
CLAP! CLAP!
CLAP!

Mais que se passe-t-il si on entraîne le réseau de neurones avec une base de données déséquilibrée : beaucoup d'images de chiens, peu de chats ? C'est ce qu'on appelle les biais.

Aux États-Unis, des chercheurs avaient mis en place un logiciel capable de prédire le risque de récidive chez les prisonniers.

Le problème, c'est qu'il présageait pour tous les Noirs un fort taux de récidive.

La raison? Beaucoup plus de Noirs dans les prisons américaines.

Le logiciel a reproduit un biais raciste!

Le deuxième type d'apprentissage, c'est l'apprentissage par renforcement.

Le système ne reçoit aucune règle, uniquement un retour en sortie qui lui dit:

GAGNÉ ou PERDU

En fonction, il ajuste ses paramètres.

C'est très utilisé pour les jeux vidéo, typiquement le jeu de go.

Mais aussi en robotique, pour apprendre au robot à réaliser certaines tâches qui ne sont pas programmées. Marcher par exemple.

Et le troisième type, c'est l'apprentissage non-supervisé.

Les systèmes de recommandation.

Tu donnes au système plein de données et tu lui demandes de faire des groupes en fonction des similarités.

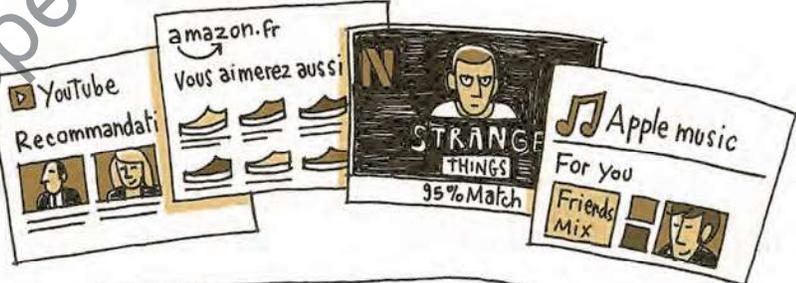
Prenons comme critère le style vestimentaire.



Tu ajoutes un deuxième algorithme pour repérer les groupes les plus proches et proposer des articles similaires.



C'est très utilisé par des plateformes comme :



Plus tu les utilises, plus l'algorithme apprend à te connaître et te fait des propositions adaptées.



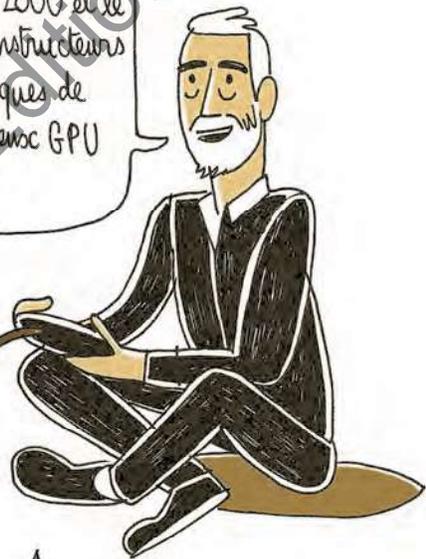
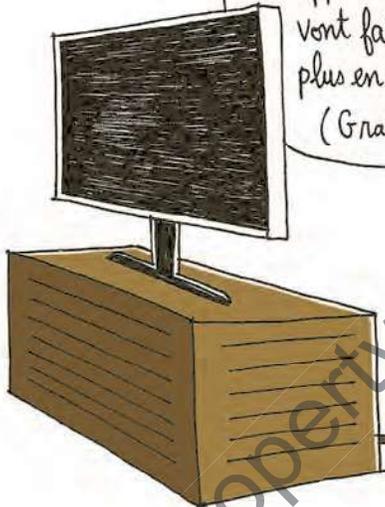
Une photo, c'est à peu près 10 millions de pixels, ça fait beaucoup de données à traiter!

Plus la tâche est complexe, plus tu ajoutes des couches, plus ça crée des opérations.

Pour tout ça, tu as besoin d'une grosse puissance de calcul !!!

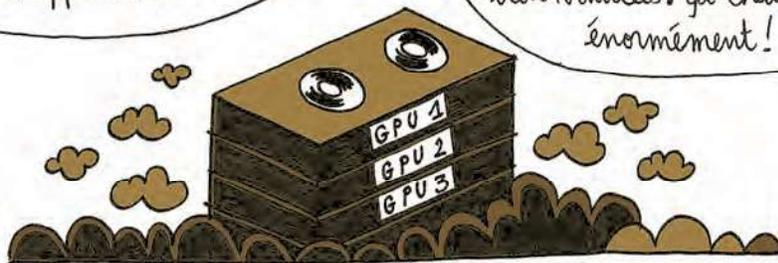


Ça tombe bien! À partir des années 2000 et le développement des jeux vidéo, les constructeurs vont fabriquer des cartes graphiques de plus en plus puissantes, les fameux GPU (Graphic Processing Unit)



Les chercheurs commencent à les utiliser pour entraîner leurs réseaux de neurones à apprendre.

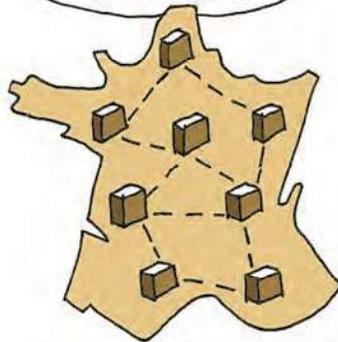
On place ça dans des grandes salles qu'on appelle des fermes, des "clusters" ou des "data centers" bien ventilées. Ça chauffe énormément!



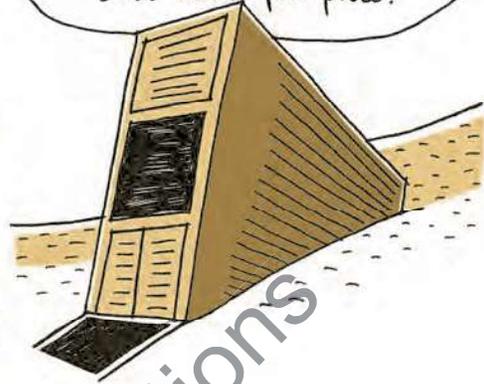
Notre première ferme avait 250 processeurs.



Puis en regroupant toutes les fermes des centres de recherche de France, on est arrivés à 5000 GPU.



Aujourd'hui, Google a un bâtiment entier en Alaska avec à peu près 5000 unités par pièce!



Le problème de cette technologie c'est qu'elle est très polluante!

Pour apprendre à reconnaître un chat, un bébé a besoin de 2-3 exemples. Une machine s'entraîne sur des millions d'exemples!



On a la méthode, on a le calcul.

Avec la démocratisation du web, on va pouvoir se servir d'énormes bases de données!



C'est comme ça que la discipline va exploser dans les années 2000!

