

# Intelligence artificielle et savoirs musicaux relevant de l'oralité

*Marc Chemillier (EHESS, CAMS)*  
*12 novembre 2025*

## Intelligence artificielle et savoirs musicaux

Algorithmes et données

Qu'est-ce que l'intelligence artificielle ?

Apprentissage automatique : l'IA « apprend » quoi ?



European Research Council (ERC) REACH project

## Algorithmes et données

*Solo de l'IA Djazz sur « Save The Earth » du guitariste malgache Charles Kely Zana-Rotsy*

<https://www.youtube.com/watch?v=tsTI2M0OBWg&t=217s>



*Le logiciel Djazz improvise avec des **données** de solos pré-enregistrés. Ici l'apprentissage est basé sur une transcription du pianiste Brad Mehldau :*

<https://www.tiktok.com/@digitaljazz/video/7043331222591835398>

*L'**algorithme** de Djazz permet de recombiner ces données et de les adapter au contexte (rythme, harmonie).*

**Algorithme** = suite d'opérations reproductibles quelles que soient les données (ex : la division)

*Jeu de dés de Mozart : système similaire de composition musicale avec deux dés, des cartes (mesures pré-composées = **données**), et des tables (règles d'enchaînement = **algorithme**)*



KV 516f, quelques mesures écrites par Mozart en marge de son manuscrit de l'Adagio du Quintette à cordes K 516.

**Démo du jeu de dés de Mozart :**  
<https://www.buschs.de/Mozart/index.html>

*Les deux tables indiquent pour chacune des 8 mesures du menuet (première et deuxième parties) quelle carte il faut choisir en fonction du résultat du lancer des dés. Les connaissances musicales sont donc à la fois dans les données et dans l'algorithme (règles) :*

- données = cartes contenant les mesures pré-composées
- règles = tables définissant quelle carte peut succéder à quelle autre

*Avec Djazz, les données peuvent être captées en direct : c'est de l'**apprentissage automatique**.*

*Elles sont recombinées par Djazz pour improviser en tenant compte du rythme et des harmonies.*

<https://www.youtube.com/watch?v=J52SEoHvx1o>



## Qu'est-ce que l'intelligence artificielle ?

1956 : John McCarthy invente le terme "intelligence artificielle" = réaliser des programmes informatiques capables de simuler l'intelligence humaine dans l'exécution de tâches complexes.

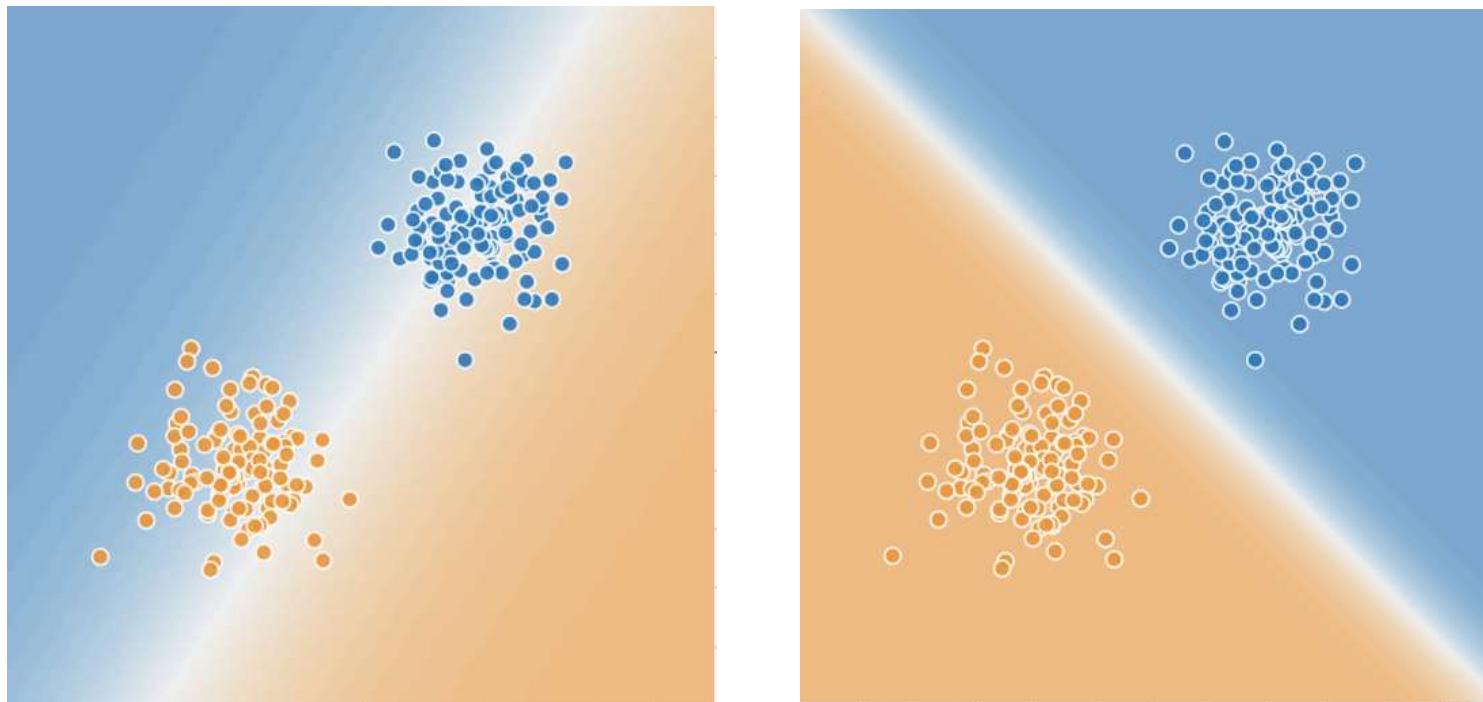
- 1950-1980 : règles (« systèmes-experts »)
- 1990-2000 : apprentissage automatique (données d'apprentissage étiquetées à la main)
- 2010-2020 : apprentissage profond (big data, données brutes et abondantes)

« IA générative » = apprentissage profond, mais il y a d'autres types d'IA (Djazz...)

*En 1990-2000, c'est l'augmentation de la taille mémoire des ordinateurs qui crée l'apprentissage automatique (machine learning) grâce au stockage de données.*

*L'un des modèles d'apprentissage automatique est le réseau de neurones basé sur la régression logistique classant des données de façon linéaire. Ex : grouper des points en deux classes séparées par une ligne droite. On ajuste progressivement les paramètres de la droite ( $a$  et  $b$  dans  $y = ax + b$ ) par entraînement, jusqu'à obtenir la bonne droite qui sépare les points orange et bleu.*

Démo sur les réseaux de neurones:  
<https://playground.tensorflow.org>



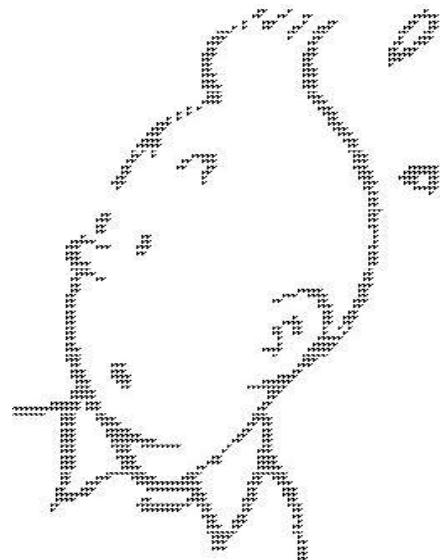
*Chaque petite unité de calcul sépare les points selon une droite. Un algorithme ajuste les paramètres selon que les données de test sont séparées correctement ou non.*

*Si les points ne sont pas séparables linéairement, un réseau de neurones avec plusieurs couches d'unités de calcul permet de faire la séparation selon des courbes plus complexes.*

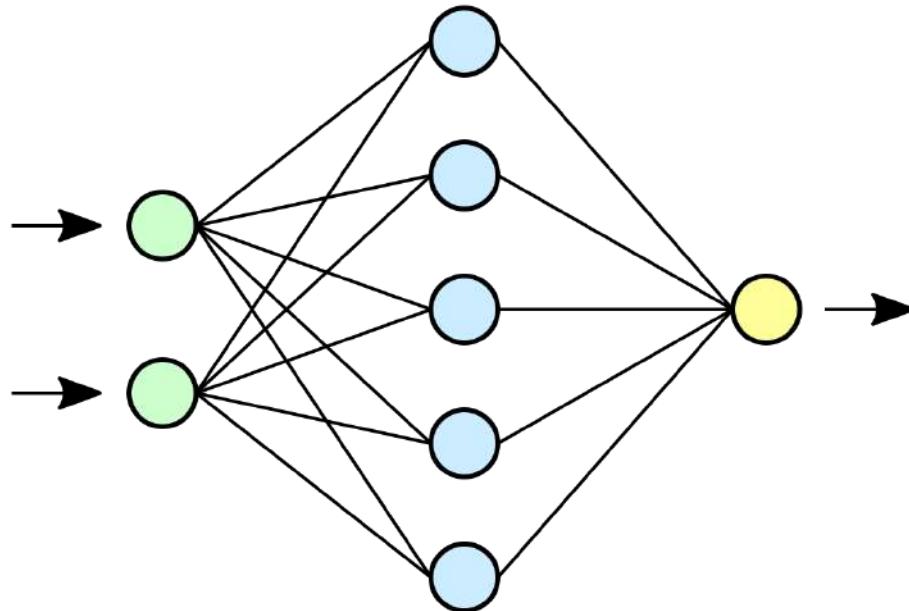
*= principe des pixels, formes simples (carré), mais un grand nombre de pixel permet de reproduire des formes complexes.*

*Si les points ne sont pas séparables linéairement, un réseau de neurones avec plusieurs **couches** d'unités de calcul permet de faire la séparation selon des courbes plus complexes.*

*= principe des pixels, **formes simples** (carré), mais un grand nombre de pixel permet de reproduire des **formes complexes**.*



*Dans un réseau de neurones simple, on a 3 couches : entrée, couche cachée, sortie*

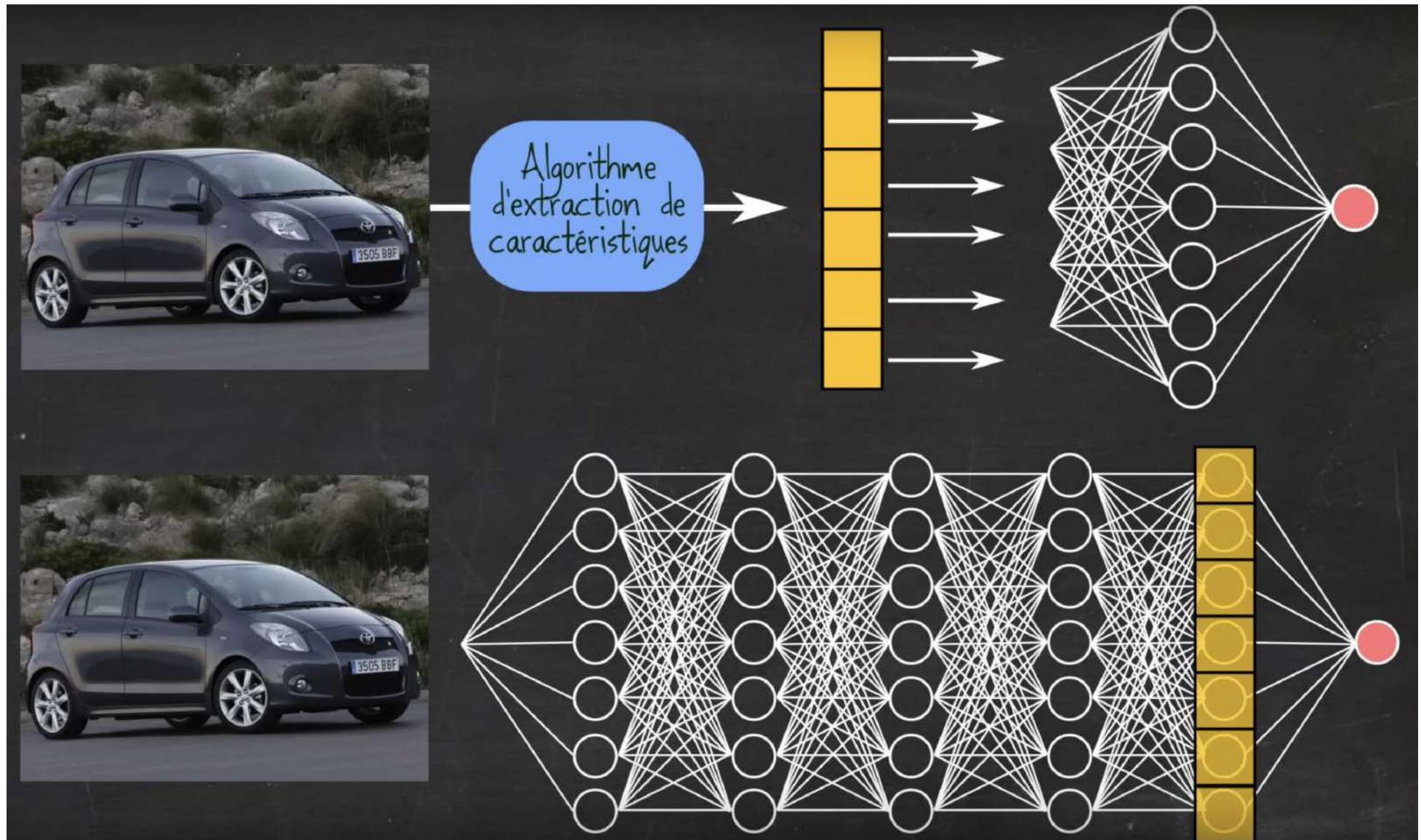


*Pour des images numérisées, il faut simplifier les données par extraction préalable de caractéristiques (abstractions d'images) qu'on donne en entrée au réseau (et non les pixels).*

2012 : succès spectaculaire de SuperVision (Yann Le Cun)

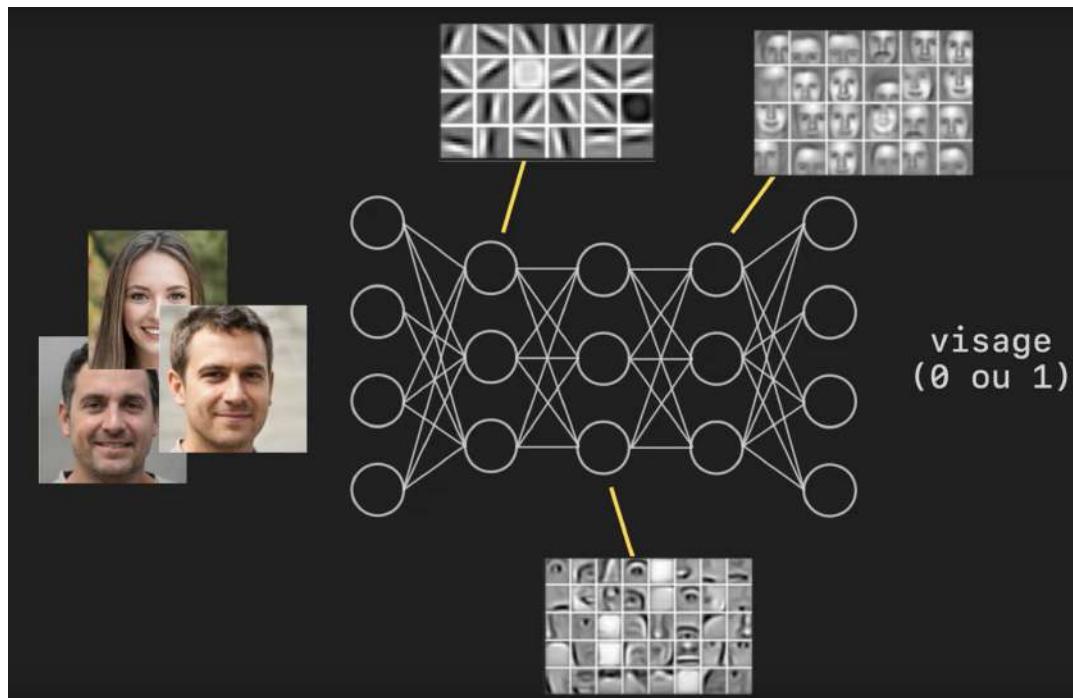
= pas d'extraction préalable de caractéristiques, il traite directement les pixels, mais avec un très grand nombre de couches. Ce nouveau modèle est appelé **apprentissage profond** (= profondeur de couches).

*Du point de vue de l'anthropologie de la connaissance, on a supprimé l'étape d'extraction de caractéristiques simplement en ajoutant des couches supplémentaires. Mais ces « caractéristiques » sont recréées par le réseau lui-même dans son avant-dernière couche.*



**David Louapre (ScienceEtonnante): Le deep learning**  
<https://www.youtube.com/watch?v=trWrEWfhTVg>

*Détection de visage : 1ère couche = traits simples, avant-dernière couche = formes élaborées (yeux, nez,...). Le réseau redécouvre par lui-même ces **concepts** : si une entrée ne contient que le concept, on voit quel neurone est activé.*



**Alexandre TL : Comprendre les réseaux de neurones**  
<https://www.youtube.com/watch?v=bkoNI7ImPBU>

## Apprentissage automatique : l'IA « apprend » quoi ?

*L'apprentissage profond utilise des milliards de neurones en couches. Il faut des milliards de données d'entraînement (années 2010 bases d'images classifiées : ImageNet de Stanford). Le **big data** est l'utilisation (pas très contrôlée...) des données disponibles en ligne.*

**Générateur de chanson Suno (big data) :** <https://www.suno.ai/>

**Utilisation de Suno pour l'arrangement (pub Facebook):**  
<https://www.facebook.com/reel/786557050763355>

*Djazz utilise une IA maigre. Les données d'entraînement du modèle ne proviennent que*

- *du musicien interagissant avec Djazz*
- *d'un corpus de transcriptions sélectionnées*

*Le modèle d'apprentissage de Djazz (oracle des facteurs) n'est pas un réseau de neurones .*

*Donc Djazz ne fait pas d'apprentissage profond.*

*Avantages :*

- *écologique : calcul léger (stockage, calcul)*
- *esthétique : intention artistique dans le choix des données*

Concernant l'anthropologie de la connaissance, on a vu que l'apprentissage profond n'utilise **pas de caractéristiques extraites au préalable**.

**agnosticisme** : il n'y a **pas de connaissance a priori hors de l'expérience** (métaphysique, religion)

Dans Djazz, il y a étiquetage préalable (pulsations, harmonie). Mais **pas de connaissance musicale a priori sur ces étiquettes** (= simples marques). Donc Djazz est agnostique, d'où l'**adaptabilité à différents contextes culturels**.

*Critique de l'IA vue comme modèle de domination occidentale : **décolonisation des savoirs**.*

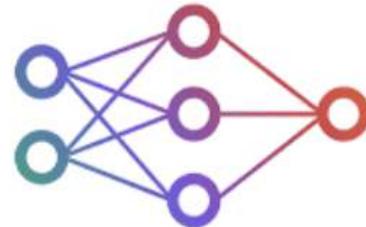
*Autre question : l'IA remplace-t-elle l'humain ?  
Transhumanisme, vie digitale post-mortem,  
musique artificielle (Suno)...*

*Critique autour de l'idée de **post-réalité** (Bronner 2025). Djazz n'est pas conçu comme substitution, mais comme augmentation = **cocréativité** (interagir **avec des humains** dans l'improvisation collective).*

*Paradoxe de ces deux critiques de l'IA **opposées** :*

- épistémologies relativistes (décoloniales)
- rationalisme universaliste (post-réalité)

IA



algorithme  
(objectif)

données  
(subjectives)

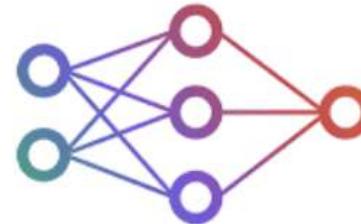
biais racistes

post-réalité

épistémologies  
relativistes  
(décoloniales)

rationalisme  
universaliste

IA



algorithme  
(objectif)

données  
(subjectives)

biais racistes

post-réalité

agnos-  
tique

cocréa-  
tivité

épistémologies  
relativistes  
(décoloniales)

rationalisme  
universaliste