

# Intelligence artificielle : de quoi parle t-on ?

démystification, promesses et enjeux

Journées Horti'doc – Angers – Juin 2024

# L'IA n'est pas une nouveauté

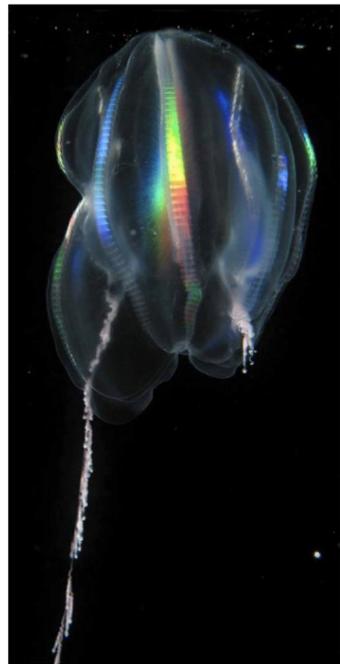
## les débuts datent des années 50

- Turing : faculté d'une machine à imiter une conversation (intelligence humaine vs. intelligence de la machine)
- 1956 : le terme d'intelligence artificielle apparaît pour la première fois
  - \* approche déductive : « si... alors ... » (**apprentissage symbolique**)
  - \* approche statistique : « apprendre » à identifier des relations statistiques entre les données (**apprentissage machine**)
- 2010 : c'est la révolution ! l'apprentissage profond (notamment) commence à donner des résultats spectaculaires. 2 ingrédients clés : les **données** et la **puissance de calcul**

# Le mot-clé : la généralisation

Imiter l'intelligence humaine

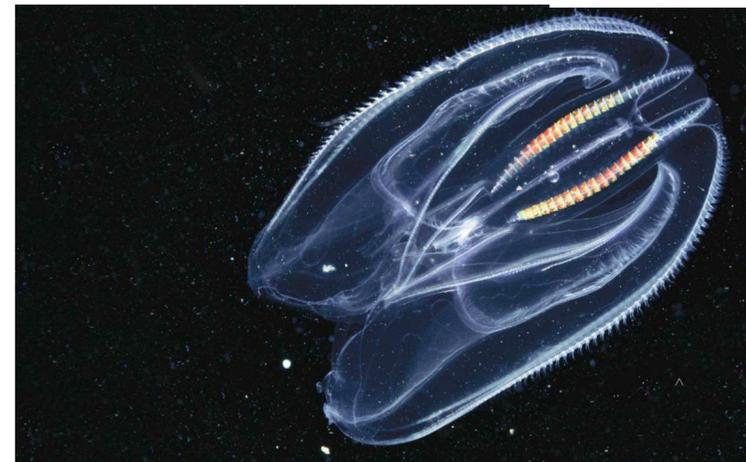
cténophore



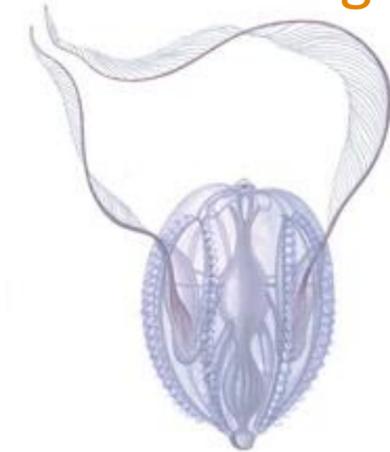
salpe



sur de nouvelles données



données hétérogènes



généralisation  $\neq$  apprentissage par coeur



être capable de transfert



en ligne

# Enfin, l'IA, c'est quoi ?

## Tentative de définition

- Machines qui apprennent automatiquement des règles à partir des données  
Outil permettant de "reproduire des comportements liés aux humains, tels que le raisonnement, la planification et la créativité » (parlement européen)
- Les IA sont en mesure d'établir des prédictions, de formuler des recommandations, de prendre des décisions, de générer des contenus...
- Les technologies évoluent très rapidement

2018



GOVERNEMENT  
Liberté  
Égalité  
Fraternité

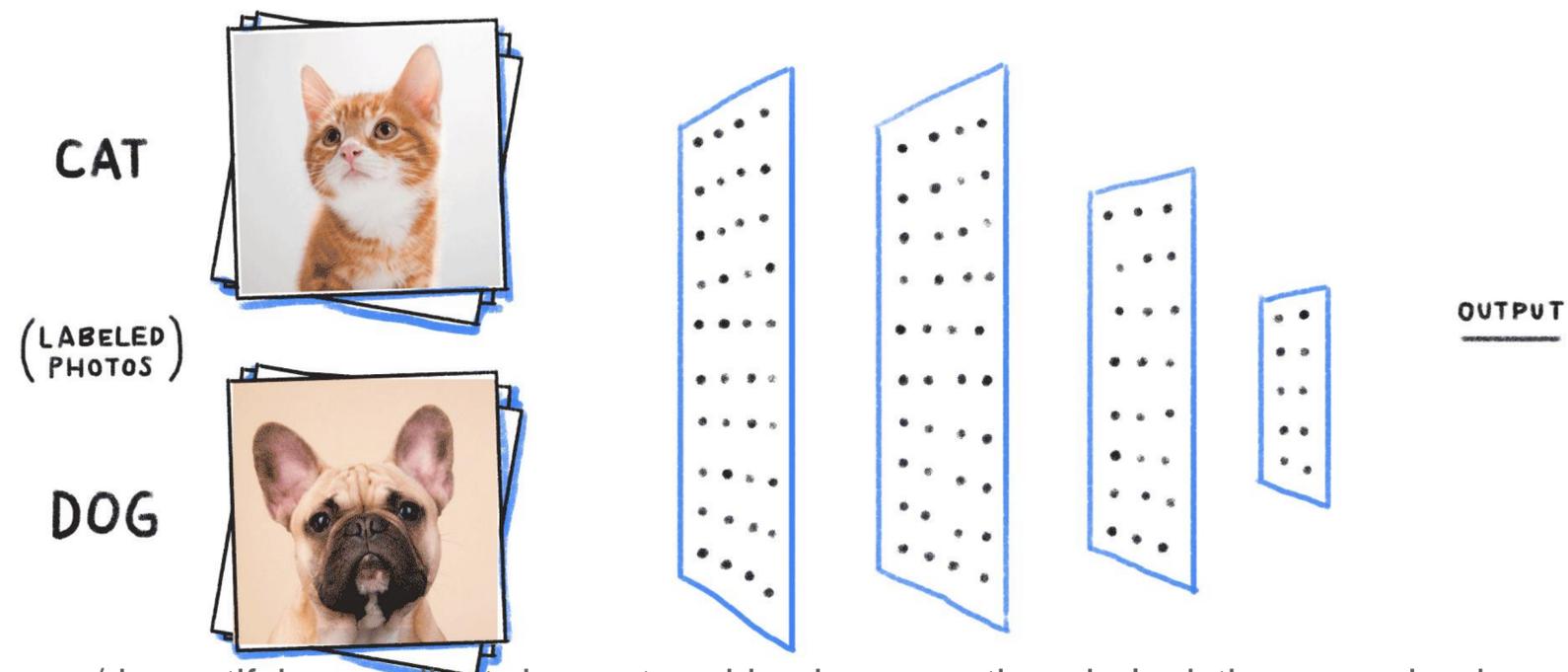


2024

# L'IA et l'inspiration biologique

deep learning = réseau profond avec beaucoup de couches de neurones

- Le cerveau =  $85 \times 10^9$  neurones et  $10^{15}$  synapses
- 1957: le perceptron, avec un « neurone » informatique capable « d'apprendre »
- Comment ça marche ? de façon très simplifiée,



# Une tâche d'IA : génération de texte

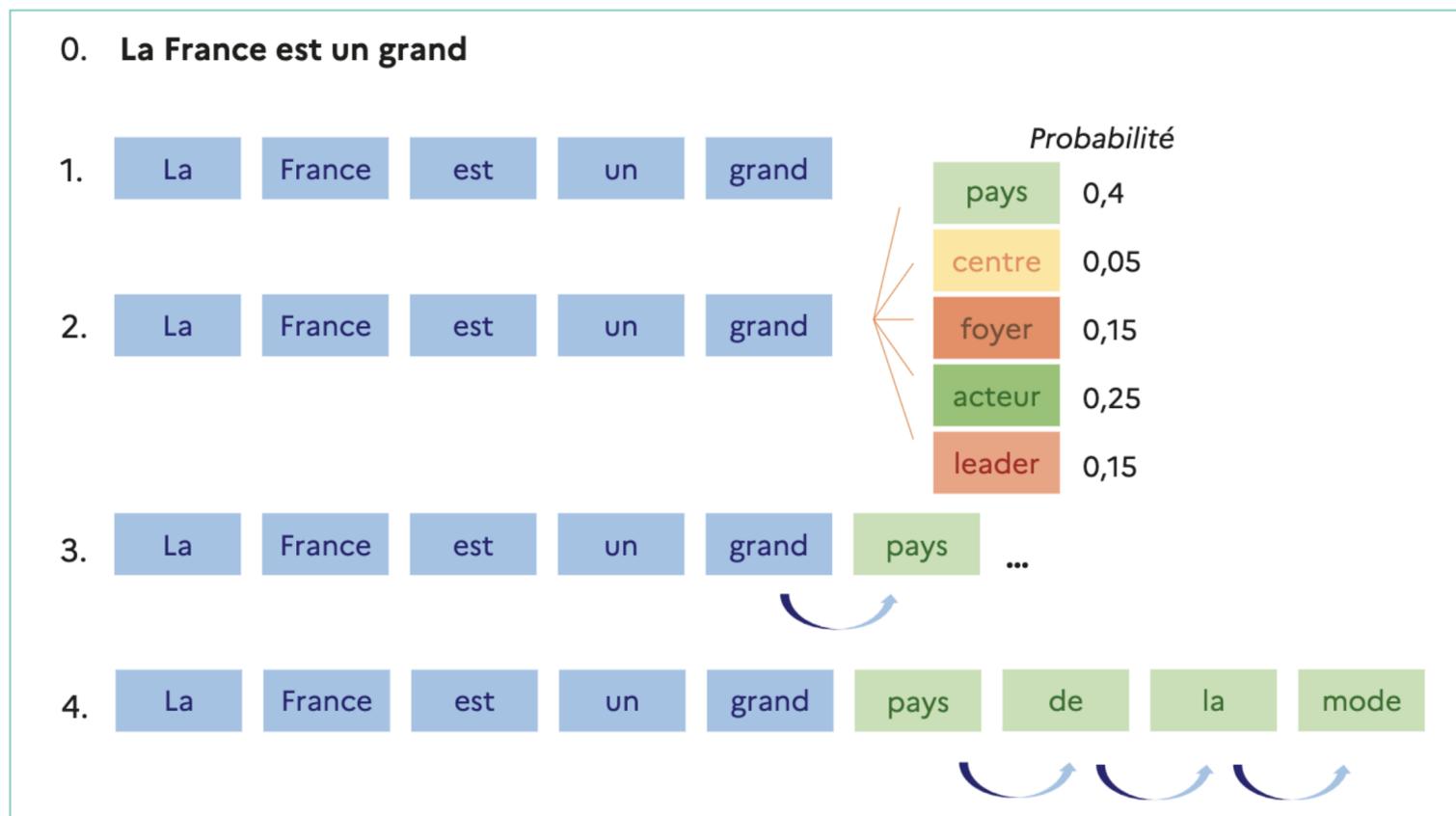
## Large langage model

- Objectif : générer des textes semblables à ceux écrits par des humains (capturant le contexte, le ton, les nuances, des éléments culturels de langage etc.)
- Les clés pour "apprendre" des relations complexes entre les mots
  - \* énormément de volumes de données en entrée ! découpées en séquences — **jetons**
  - \* de grosses capacités de calcul — 100 000 milliards de paramètres pour chatGPT4
  - \* un mécanisme qui permet de replacer les mots dans leur contexte (et d'identifier quel est le contexte important) — **mécanisme d'attention**
  - \* un mécanisme d'**encodage-décodage** : on encode le texte dans une nouvelle représentation et on la décode
  - \* un mécanisme qui permet de prendre en compte la **position** d'un mot dans la phrase

# Une tâche d'IA : génération de texte

## Large langage model

- Objectif : générer des textes semblables à ceux écrits par des humains (capturant le contexte, le ton, les nuances, des éléments culturels de langage etc.)



quelle est la suite de cette phrase

quelle est la suite de cette phrase

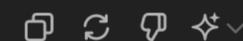
l'IA c'est hyper performant et ce n'est pas compliqué à



Voici une suite possible pour la phrase :

"L'IA c'est hyper performant et ce n'est pas compliqué à comprendre une fois que l'on maîtrise les bases."

Cette continuation termine la pensée en expliquant que la compréhension de l'IA devient plus facile avec une bonne base de connaissances.



l'IA c'est hyper performant et ce n'est pas compliqué à comprendre.

### 7. Surveillance et décision humaines

Les États membres devraient veiller à ce qu'il soit toujours possible d'attribuer la responsabilité éthique et juridique de tout stade du cycle de vie des systèmes d'IA à des personnes physiques ou des entités juridiques existantes.

### 8. Durabilité

Les technologies de l'IA devraient être continuellement évaluées en fonction de leur impact sur la 'durabilité' en tant qu'ensemble d'objectifs en constante évolution, tels qu'ils sont actuellement définis dans les objectifs de développement durable (ODD) des Nations Unies.

### 9. Sensibilisation et éducation

La compréhension des technologies de l'IA et de la valeur des données par le public devraient être favorisées par une éducation ouverte et accessible, l'engagement civique, l'acquisition de compétences numériques et la formation à l'éthique de l'IA, et l'éducation aux médias.

### 10. Équité et non-discrimination

Les acteurs de l'IA doivent promouvoir la justice sociale, garantir l'équité et lutter contre les discriminations, tout en adoptant une approche inclusive pour s'assurer que les bénéfices de l'IA sont disponibles et accessibles à tous.

s données ?)

### 3. Droit au respect de la vie privée et protection des données

La vie privée doit être respectée, protégée et promue tout au long du cycle de vie des systèmes d'IA. Des cadres de protection des données et des mécanismes de gouvernance adéquats doivent être mis en place.

### 4. Gouvernance et collaboration multipartites et adaptatives

Le droit international et la souveraineté nationale doivent être respectés dans l'utilisation des données. La participation des différentes parties prenantes est nécessaire pour garantir des approches inclusives de la gouvernance de l'IA.

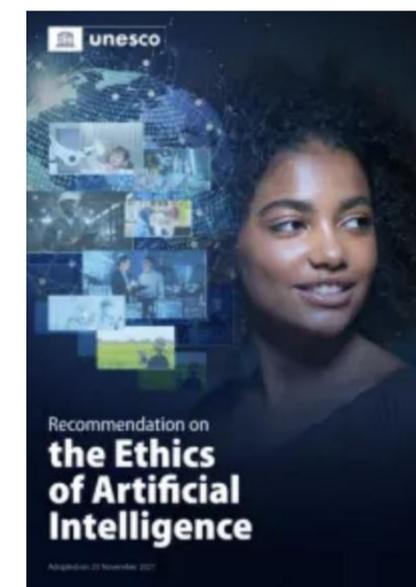
### 5. Responsabilité et redevabilité

Les systèmes d'IA doivent être vérifiables et traçables. Des mécanismes de surveillance, d'évaluation de l'impact, de contrôle et de diligence requise, devraient être mis en place pour éviter tout conflit lié aux droits de l'homme et traiter les menaces pesant sur le bien-être de l'environnement.

### 6. Transparence et explicabilité

Le déploiement éthique des systèmes d'IA dépend de leur transparence et explicabilité (T&E). Le niveau de T&E doit être adapté au contexte, car il peut être nécessaire de trouver un équilibre entre T&E et d'autres principes, tels que la protection de la vie privée, la sûreté et la sécurité.

2022



# Dilemmes éthiques et réponses ?

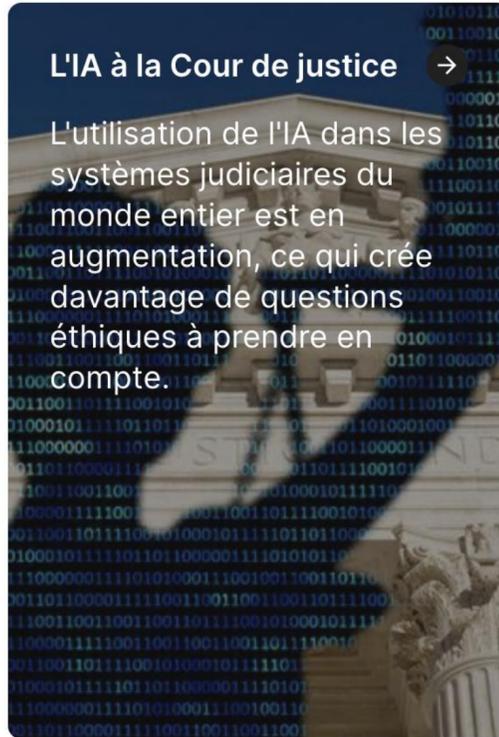
## Les biais de l'IA →

Exemples de préjugés de genre dans le domaine de l'intelligence artificielle, résultant de représentations stéréotypées profondément ancrées dans nos sociétés.



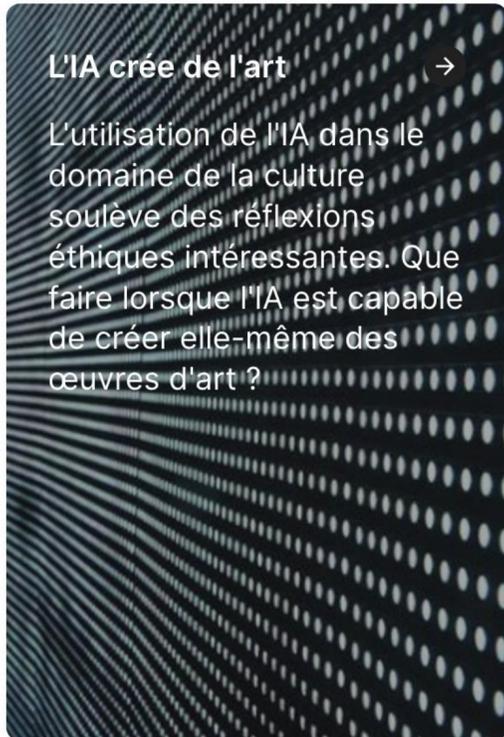
## L'IA à la Cour de justice →

L'utilisation de l'IA dans les systèmes judiciaires du monde entier est en augmentation, ce qui crée davantage de questions éthiques à prendre en compte.



## L'IA crée de l'art →

L'utilisation de l'IA dans le domaine de la culture soulève des réflexions éthiques intéressantes. Que faire lorsque l'IA est capable de créer elle-même des œuvres d'art ?



## La voiture autonome →

Une voiture autonome est un véhicule capable de détecter son environnement et de se déplacer avec peu ou pas d'intervention humaine.



CNIL



## Quelles réponses éthiques au développement des algorithmes et de l'intelligence artificielle ?

### Vers l'affirmation de deux principes fondateurs : loyauté et vigilance

Les débats ont permis de dégager deux principes fondateurs pour une intelligence artificielle au service de l'homme. Ces principes pourraient s'inscrire dans **une nouvelle génération de garanties et de droits fondamentaux à l'ère numérique**, des « droits-système » organisant la **gouvernance mondiale de notre univers numérique** :

- **Un principe de loyauté** appliqué à tous les algorithmes et intégrant les impacts collectifs, et pas seulement personnels, de ces derniers. Tout **algorithme**, qu'il traite ou non des données personnelles, doit être loyal envers ses utilisateurs, non pas seulement en tant que consommateurs, mais également en tant que citoyens, voire envers des communautés ou de grands intérêts collectifs dont l'existence pourrait être directement affectée. **L'intérêt des utilisateurs doit primer**. Par exemple, un tel principe pourrait avoir vocation à s'appliquer à l'impact potentiel des réseaux sociaux sur la structure du débat public dans nos démocraties (segmentation du corps politique par le ciblage de l'information) ou à celui d'algorithmes de police prédictive sur des communautés ou quartiers entiers.
- **Un principe de vigilance/réflexivité** : il s'agit d'organiser une forme de questionnement régulier, méthodique et délibératif à l'égard de ces objets mouvants. Ce principe constitue une réponse directe aux exigences qu'imposent ces objets technologiques du fait de leur nature imprévisible (inhérente au *machine learning*), du caractère très compartimenté des chaînes algorithmiques au sein desquels ils s'insèrent et, enfin, de la confiance excessive à laquelle ils donnent souvent lieu. C'est l'ensemble des maillons de la chaîne algorithmique (concepteurs, entreprises, citoyens) qui doivent être mobilisés pour donner corps à ce principe, au moyen de procédures concrètes (par exemple, des comités d'éthique assurant un dialogue systématique et continu entre les différentes parties-prenantes).

Ces principes fondateurs sont complétés par des principes organisationnels ayant traits à l'intelligibilité et à la responsabilité des systèmes algorithmiques ainsi qu'à la nature de l'intervention humaine dans la prise de décision algorithmique.

# Intelligence artificielle : de quoi parle t-on ?

démystification, promesses et enjeux