RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité

Inria

IA en Santé

Point de vue d'un informaticien ...

Thomas Guyet
Chercheur Inria/AIstroSight
Vice-Président Association Française pour l'Intelligence Artificielle

thomas.guyet@inria.fr

Brève présentation

Inria : organisme de recherche publique dédié aux sciences du numériques

- Institut national créé en 1967
- 9 centres nationaux
- Centre de Lyon créé officiellement en 2021 (présence bien antérieure)

Accord cadre HCL-Inria-UCBL

- pôle IA : transfert de technologies
- projets de recherche communs



Nos recherches en IA

AIstroSight : équipe de recherche sur les outils numériques pour aider à la conception des médicaments pour des maladies à affection neurologiques

Développement de méthodes d'intelligence artificielle (mais pas que)

- modélisation des dynamiques cellulaires
- intégration de données hétérogènes des patients (imageries, documents médicaux, ...)
- analyse des trajectoires de soins



Applications médicales

Projets pluridisciplinaires société/IA/santé

01
Tentative de définition

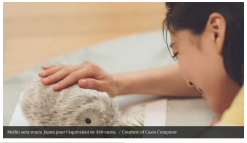
« *Intelligence Artificielle* » : LE buzz-word du moment

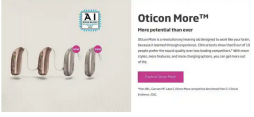
Intelligence Artificielle (IA) utiliser à tout va ...

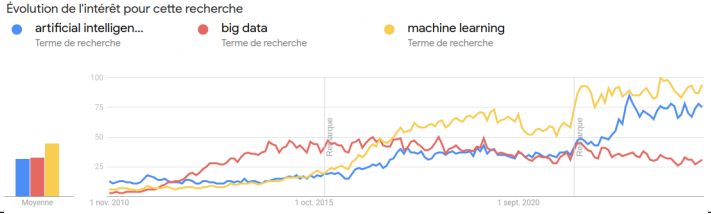
Attention aux abus : Promesses de « *progrès* » de l'Intelligence Artificielle (IA)

- performance (« à la pointe »)
- personnalisation
- innovation technologique (pour les décideurs)
- solution pour des questions sociétales (santé, environnement, etc.)

Grâce à l'IA, cette peluche est capable d'exprimer des émotions







Aux origines de l'Intelligence Artificielle



August 1956. From left to right : Oliver Selfridge, Nathaniel Rochester, Ray Solomonoff, Marvin Minsky, Trenchard More, John McCarthy, Claude Shannon.

Cybernétique

Établir une théorie générale du fonctionnement de la pensée humaine basée sur la théorie du contrôle
Conférences Macy 1946-1953 (Wiener, McCullogh, Pitts, Ashby).

Intelligence Artificielle

Dartmouth 1956 (McCarthy, Minsky, Shannon, ...)
Naissance du terme *Intelligence Artificielle*
Construire des dispositifs simulant les processus cognitifs humains

Deux conceptions opposées de la cognition

Cognition = Raisonnement logique

- IA « Symbolique »
- Principes
  - L'information peut être représentée sous la forme de symboles
  - La pensée peut être décrite comme l'application de règles formelles
- ex. plannification, résolution de problèmes, diagnostic, etc.

IBM Watson

Cognition = Émergence de l'activité neuronale

- IA « connexionniste »
- Principes
  - Simuler le fonctionnement des neurones = simulé la cognition
  - Pas de représentation explicite du monde
- ex. reconnaissance de formes, diagnostic, etc.

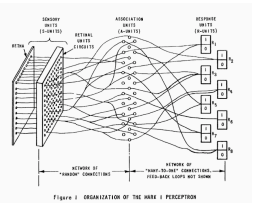


Figure 1 ORGANIZATION OF THE MARK I PERCEPTRON

Rosenblatt, 1958

Et d'autres encore programmes des recherches en IA ...

- représentation des connaissances, intelligence collective, robotique, ...

Vers une définition des systèmes d'intelligence artificielle (SIA) ...

Unesco, Recommendation on the Ethics of Artificial Intelligence, 2021

Les « systèmes d'IA » sont des **technologies de traitement des informations** qui intègrent des modèles et des algorithmes, lesquels génèrent une **capacité d'apprentissage** et d'**exécution de tâches cognitives** conduisant à des résultats tels que l'anticipation et la prise de décisions dans des environnements matériels et virtuels. Les systèmes d'IA sont conçus pour fonctionner avec différents degrés d'autonomie, au moyen de la modélisation et la représentation des connaissances, de l'exploitation des données et du calcul de corrélations. Ils peuvent intégrer plusieurs méthodes, telles que, sans s'y limiter :

- l'apprentissage automatique, y compris l'apprentissage profond et l'apprentissage par renforcement ;
- le raisonnement automatique, y compris la planification, la programmation, la représentation des connaissances et le raisonnement, et la recherche et l'optimisation.

autres définitions

SIA perçus aujourd'hui (Fév 2025)

Qqs chiffres

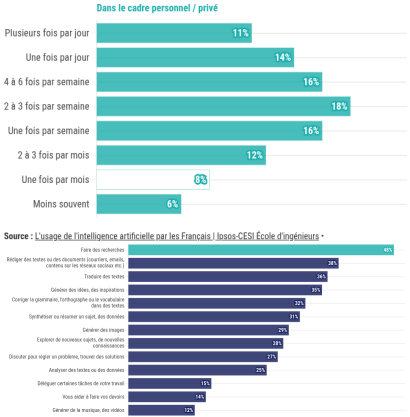
- 88% des Français ont déjà entendu parler d'intelligence artificielle générative.
- 39% d'entre eux l'utilisent activement.

Les 18-24 ans étant les plus grands utilisateurs.

ChatGPT se distingue comme la plateforme la plus populaire.

Préoccupations quant aux risques potentiels

- propagation de fake news
- impact de cette technologie sur l'emploi
- capacités cognitives



Étude IPSOS/CESI, Fév 2025.
(échantillon de 1000 p.)

En résumé

IA : champs scientifique

SIA : outil technique qui s'appuie sur des technologies d'IA

- Différentes inspirations de tâches : raisonnement, apprentissage, etc.

Des limites très floues ...

- flou liée à une définition imprécise d'« *Intelligence(s)* »
- flou disciplinaire lié au fait qu'il s'agit plus de *recherche* que de *science*
- flou technologique (e.g. détecteur de spam, l'illusion d'*intelligence*, ...)
- flou discursif lié à une utilisation parfois galvodée du terme

Bingo des SIA en santé

	Apprentissage	Raisonnement	Agents
Prévention			
Diagnostic			
Traitements / chirurgies			
Assistances médicales aux patients			
Organisations			
Développement de nouveaux traitements			

SIA en santé : Prévention

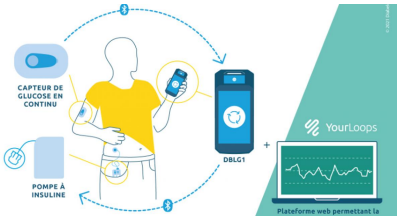
Capteurs connectés (bracelets, diabète, etc.)

Analyse prédictive / pronostic (anticipation des risques)

- *Natural history of human disease, 2025*
- Version SNDS Haugomat et al.
- Guidelines intelligents

ChatBot pour la prévention en santé mentale

- Therabot [**?**]
- *NEJM AI, 2025*



SIA en santé : Prévention

Capteurs connectés (bracelets, diabète, etc.)

Analyse prédictive / pronostic (anticipation des risques)

- *Natural history of human disease, 2025*
- Version SNDS Haugomat et al.
- Guidelines intelligents

ChatBot pour la prévention en santé mentale

- Therabot [?]
- *NEJM AI, 2025*

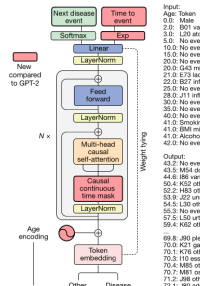


Tableau 5 Différence entre l'AUC globale et l'AUC stratifiée pour la prédiction de la mortalité à 1 an

Représentations des données	AUC globale	AUC stratifiée
BEHRT-SNDS	0,961	0,822
Plongements vectoriels des trajectoires (Snds2vec)	0,958	0,807
Comorbidités et fréquences de consommation	0,953	0,784
Cartographie et fréquences de consommation	0,953	0,781
Fréquences de consommation	0,948	0,760
Comorbidités (Charlson et Rx-Risk-V)	0,949	0,760
Cartographie	0,942	0,729
Mortality Related Morbidity Index (MRMI)	0,913	0,680
Démographique	0,912	NI

SIA en santé : Prévention



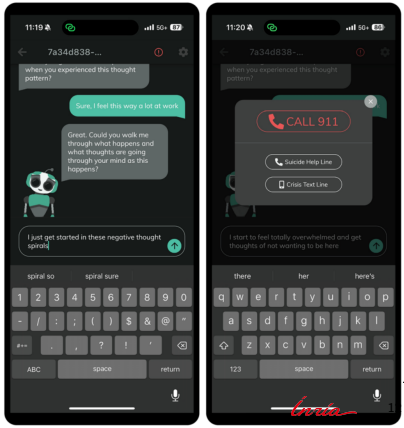
Capteurs connectés (bracelets, diabète, etc.)

Analyse prédictive / pronostic (anticipation des risques)

- *Natural history of human disease, 2025*
- Version SNDS Haugomat et al.
- Guidelines intelligents

ChatBot pour la prévention en santé mentale

- Therabot [**?**]
- *NEJM AI, 2025*



SIA en santé : Diagnostic

Analyse d'images, des signaux biomédicaux

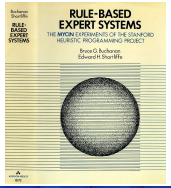
- Des nouveaux "biomarqueurs" ...
- *Potentiel gigantesque !*
- Vers des approches multimodales



Aide au diagnostic médical

- MyCin : raisonnements en santé (moteur LISP), 1979
- IBM Watson (2011) / Watson health (*arrêté*)

SIA en santé : Diagnostic



Analyse d'images, des signaux biomédicaux

- Des nouveaux "biomarqueurs" ...
- *Potentiel gigantesque !*
- Vers des approches multimodales



Aide au diagnostic médical

- MyCin : raisonnements en santé (moteur LISP), 1979
- IBM Watson (2011) / Watson health (*arrêté*)

SIA en santé : Traitements/Chirurgies

Robots chirurgicaux

- analyse des gestes chirurgicaux (e.g Pierre Jannin)
- contrôle assisté des opérations (projet RODEO)

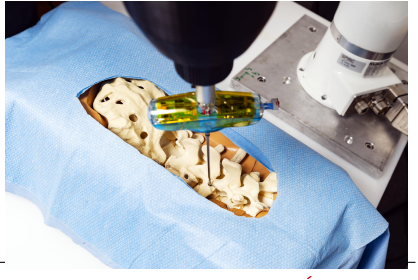
Jumeaux numériques (d'organes)

Aide à la prescription médicamenteuse

- Le bon médicament, à la bonne dose et au bon moment ...

Concept de Personnalisation des soins (médecine 5P)





SIA en santé : Traitements/Chirurgies

Robots chirurgicaux

- analyse des gestes chirurgicaux (e.g Pierre Jannin)
- contrôle assisté des opérations (projet RODEO)

Jumeaux numériques (d'organes)

Aide à la prescription médicamenteuse

- Le bon médicament, à la bonne dose et au bon moment ...



Concept de Personnalisation des soins (médecine 5P)

SIA en santé : Traitements/Chirurgies

Robots chirurgicaux

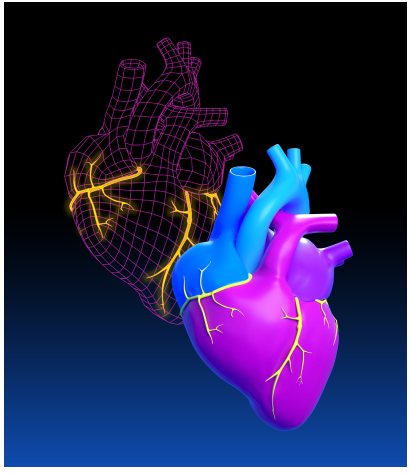
- analyse des gestes chirurgicaux (e.g Pierre Jannin)
- contrôle assisté des opérations (projet RODEO)

Jumeaux numériques (d'organes)

Aide à la prescription médicamenteuse

- Le bon médicament, à la bonne dose et au bon moment ...

Concept de Personnalisation des soins (médecine 5P)

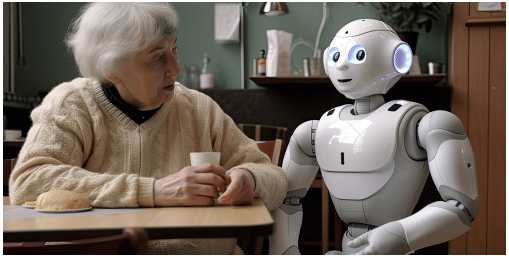


SIA en santé : Assistances médicales aux patients

Robots d'assistance

Smart home
Robots d'assistance

Education thérapeutique, maintien des traitements en pathologie chroniques



SIA en santé : Organisations

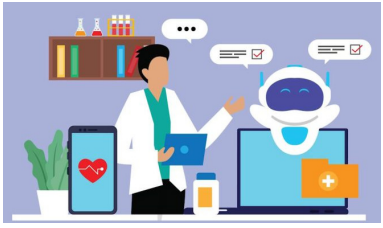
Plannification des services hospitaliers

- problème de plannification complexe
- IA : moteur de résolution de problèmes sous contraintes

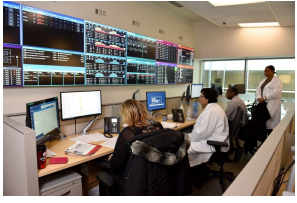
Triage de patient aux urgences

Aides administratives (prises de notes)

Robot d'assistances aux équipes médicales



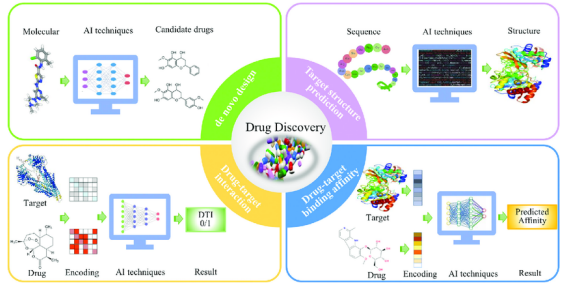




SIA en santé : Développement de nouveaux traitements

Epidémiologie sur les données de santé

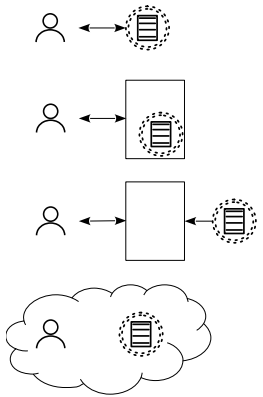
Aide à la découverte de médicaments par l'IA



Bingo des SIA en santé

	Apprentissage	Raisonnement	Agents
Prévention			
Diagnostic			
Traitements / chirurgies			
Assistances médicales aux patients			
Organisations			
Développement de nouveaux traitements			

SIA visibles et invisibles



SIA frontaux

SIA embarqués

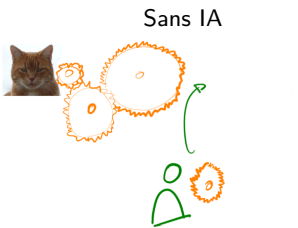
SIA secondaires

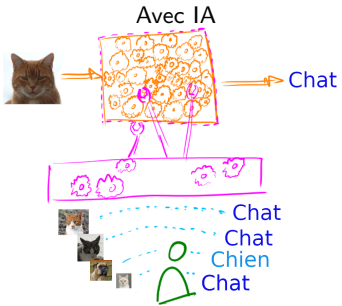
SIA diffus

02
« Système d'IA Apprenants » : pourquoi c'est tentant ?

« Système d'IA Apprenants » : principes (1)







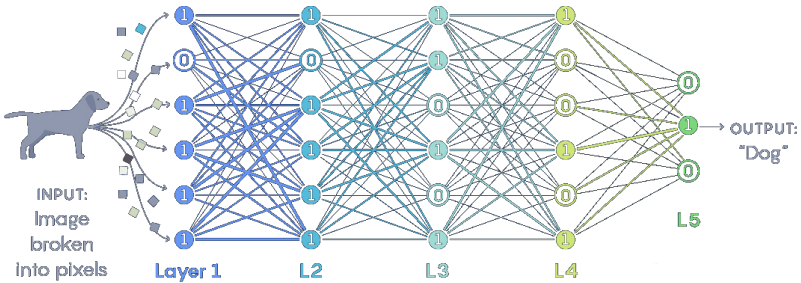
« Systèmes d'IA Apprenants » : modèle connexioniste

Réseaux de neurones

Neurone artificielle : brique élémentaire de calcul inspirée d'un neurone biologique

- valeurs numériques en entrée
- valeur numérique en sortie

Connexions du réseau : « paramètres du modèle », $\sim$ 153 paramètres

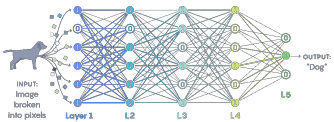


« Systèmes d'IA Apprenants » : Retro-propragation du gradient (de l'erreur)

Apprentissage d'un réseau de neurones

- résultat de la **minimisation des erreurs** faites sur les exemples présentés
- lors de la présentation d'un nouvel exemple :

1. tentative de prédiction
2. calcul d'une erreur commise
3. modification proportionnelle des connexions du réseau (*vers l'arrière*)





J. Hopfield & G. Hinton
Prix Nobel de physique 2024

Quelques points importants

L'utilisation d'un « Système d'IA Apprenant » remplace la conception d'un système informatique par une procédure d'apprentissage à partir d'exemples

Un principe de conception d'un système informatique très attrayant ...

La procédure d'apprentissage s'appuie sur

- une collection d'exemples *labellisés* (*chat*/*chien*)
- la minimisation d'une erreur

Une fois appris, le modèle appris peut être réutilisé sur de nouveaux exemples

Enjeu : on cherche un **pouvoir de généralisation**

03

« *Deep learning* » – émergence des grandes architectures : qu'est-ce qui change ?

« *Deep learning* » : émergence des grandes architectures

Deep learning = Apprentissage des représentations

- Données d'entrée beaucoup moins contraintes
  - variétés de données plus larges (images, signaux, texte, ...)
  - moins de préparation nécessaire (*feature engineering*)

→ **le réseau « apprend » quelle information utiliser** pour bien réaliser sa tâche

Accroissement des architectures de réseau de neurones

- Plus de couches/neurones ⇒ plus de complexités
- Plus de couches/neurones ⇐ besoin de (beaucoup) plus de données



J. Bengio, G. Hinton & Y. Le Cun
Prix Turing 2018

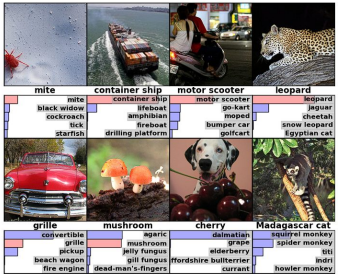
« *Deep learning* » : le saut AlexNet

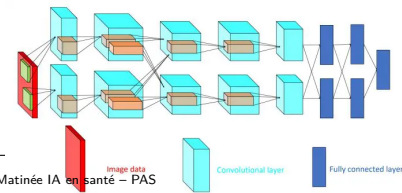
ImageNet : compétition de classification d'images

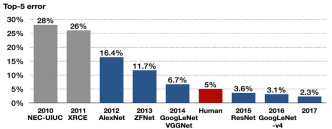
14 millions d'images
20000 catégories

AlexNet (2012) : réseau pour la classification d'images

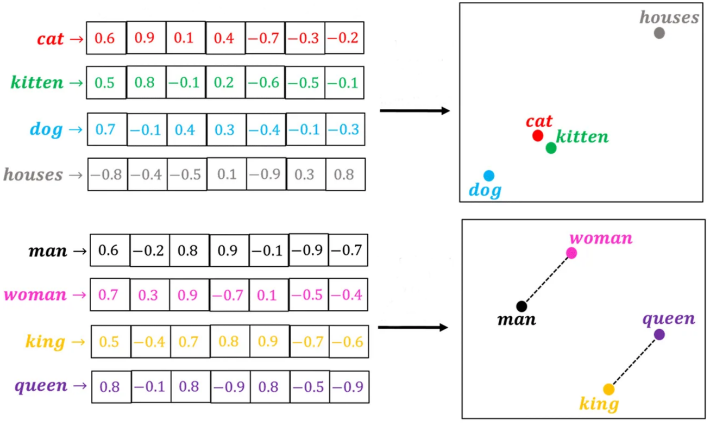
- Apprentissage de représentation : couches convolutionnelles
- Classification : couches « fully connected »
- 60 Millions de paramètres



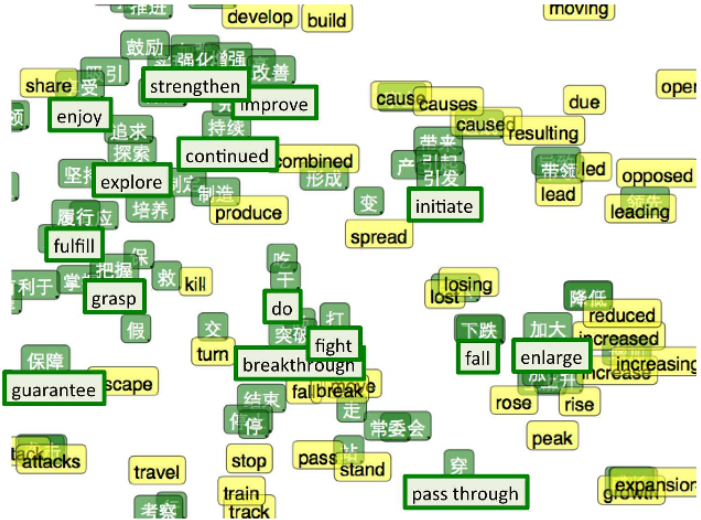




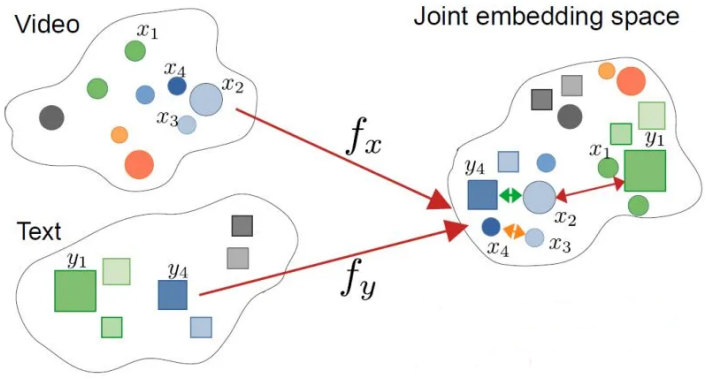
Représentation vectorielle de mots : principes



Représentation vectorielle de mots : principes



Représentation vectorielle de mots : principes



Dall-E, SORA

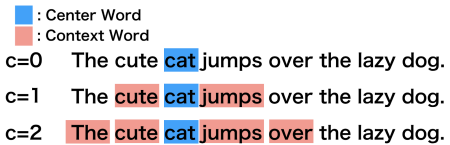
Apprentissage d'une représentation vectorielle de mots / textes

Modèle *Transformer* (2017)

- Modèle de réseau de neurones permettant de construire des représentations **contextuelles** des mots/phrases
- GPT : Generative Pre-training Transformer

Apprentissage par masquage

- Chaque phrase est présentée avec un mot masqué
- Le réseau cherche à prédire le mot masqué
  - retropropagation de l'erreur



Le résultat de l'apprentissage est un modèle « pré-entrainé »

- Réutilisable dans une multitude de tâches (génériques)
- Très gros modèles ... (*Palm2* 540 milliards de paramètres !)

Quelques points importants du « *deep learning* »

« *Deep learning* »

- Accroissement de la taille des modèles
- Hetérogénéïté des données d'entrées (textes, images, sons)
- Apprentissage de **représentations** *génériques* : capture des régularités du monde

Principes de construction du système informatique similaire à ceux de l'apprentissage automatique, mais :

- Complexification des procédures
- Contrôle de l'apprentissage moins direct

Recette de la réussite

- Beaucoup de données
- Beaucoup de calculs (et d'énergie)
- Beaucoup – *mais pas tant que cela* – d'annotation humaine (à bas coût)

04
IA Génératives : Ça fait des choses qui plaisent !

Des modèles de langues aux IA Génératives

chat-GPT

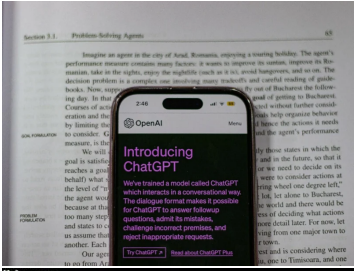
- Utilisation d'un modèle de langue pour un usage conversationnel
- Chat-GPT est un réseau de neurones basé sur GPT et affiné pour converser

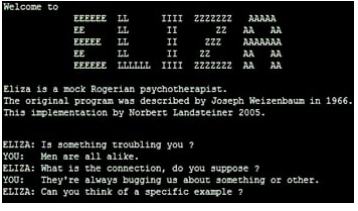
Accessibilité, intuitivité de l'interaction avec une machine

- Vers le remplacement des moteurs de recherche ...

L'effet Eliza (1966)

- Invitation à l'interaction
- Attribution d'une « *intelligence* »

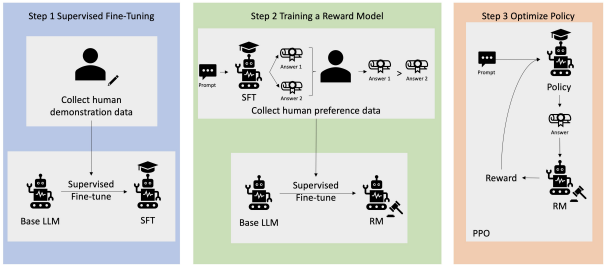




GPT : Pourquoi ça marche ?

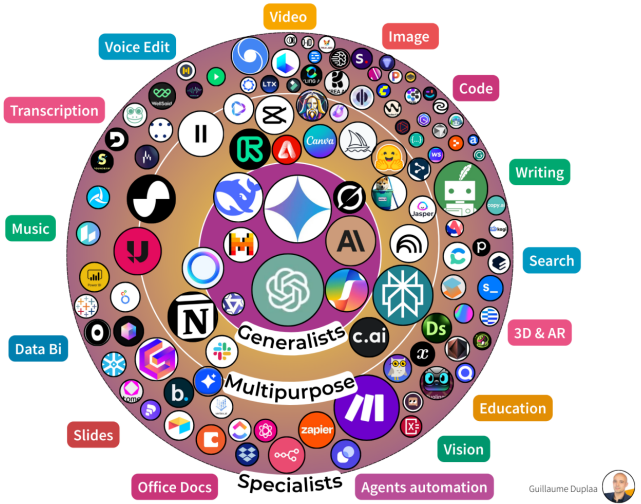
RLHF : Reinforcement Learning by Human Feedbacks (GPT-3)

- Des humains donnent des avis sur des générations
- Les avis servent à construire un modèle de récompense (*RM*)
- Le modèle de récompense affine le modèle pré-entrainé

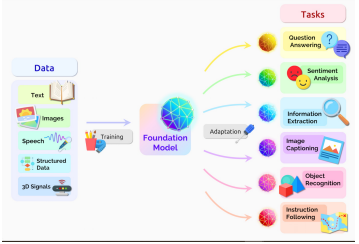


Alternatives moins coûteuses au RLHF/PPO : RLAIF (Anthropic), GRPO (DeepSeek), ...

Graal(s) : les modèles génératifs ?

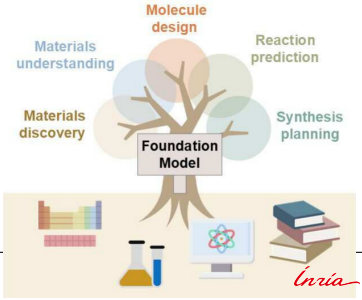


Graal(s) : les « modèles de fondation »

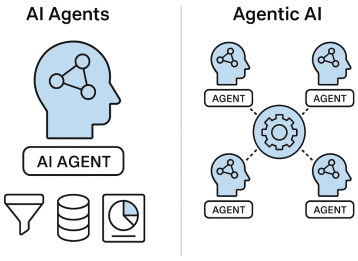


Modèles non-spécialisés par tâche

Essentialisation des connaissances d'un domaine



Vers une « Agentification » / « SIA agentique »



Agentification : vers des résolutions plus autonomes

- atteindre un but complexe en identifiant des sous-buts de manière plus autonome
- connexion d'SIA multiples entres plusieurs SIA

→ La (petite) revanche du symbolique (?)

SIA comme compagnon

05
Faut-il être méf-/conf-iant envers un SIA ?

Un système d'IA est-il un outil comme un autre ?

Un système d'IA est un outil de traitement de l'information

- Réalise une tâche pour laquelle il a été construit

→ beaucoup de question liées aux risques et à l'éthique des SIA peuvent être abordés par des principes usuels

- éducation
- régulation/encadrement

Ce qui pose question sur la nature de cette technique ?

- Il réalise des tâches cognitives (et non-plus mécaniques)
- Très grande versatilité
- Niveau de complexité élevé *et* un usage aisé
- La conception des SIA apprenants est indirecte : contrôle limité par le concepteur lui-même

→ questionnement sur la neutralité de la technique, la transformation par la technique (perte de compétences), ...

Un SIA est-il plus fort qu'un humain ?

Pour certaines tâches et certaines conditions ... oui !

- tâches décisionnelles rapides
- tâches décisionnelles sous pression
- l'utilisation de capteurs complexes et sensibles (e.g. radiographie)
- tâches sensibles aux biais (eg triage à l'hopital)



L'humain a de bons restes ...

- environnements complexes et changeants
- environnements flous
- système perceptif inégalé

Les humains accompagnés de machines peuvent faire mieux

Quelques défauts bien connus des SIA génératifs

Biais (de genre, d'ethnicité, etc.)

- Les biais d'un modèle sont nécessaires ! Mais certains biais sont discriminants ...
- Une grande partie du problème provient des données (leur sélection, etc.)
- Les mécanismes d'apprentissage automatique ont **tendance à renforcer les biais existants dans les données**



Hallucinations

- Quelle est la part d'invention dans la réponse ?

→ challenge : apprendre à dire "je ne sais pas" !

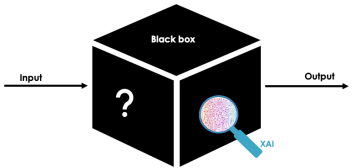
Conserver un regard critique sur les productions et les usages

Outil à tester : Compar:IA

L'explicabilité des résultats d'un SIA est-elle bien nécessaire ?

Système d'IA = boîte noire

Trop complexe pour en décrire le fonctionnement



Confiance vs explicabilité (XAI)

L'explicabilité est un vecteur de construction de la confiance dans les Systèmes d'IA

Pourrait-on imaginer se passer d'explications à terme ?

- dans des usages routiniers sans impact : probablement !
- dans des usages plus pointus : un possible retour au source restera utile (nécessaire ?)

→ utilisation des RAGs (hybridation des méthodes symboliques/neuronales)
→ révéler des causalités dans les données (vs corrélations)

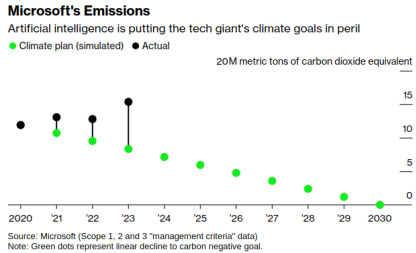
06
Impacts sociétaux du développement des SIA ?

Empreinte environnementale des systèmes d'IA

Impact réel difficile à quantifier mais réel

- matériel et fonctionnement
- phase d'apprentissage et phase d'inférence
- → limite pratique à l'expansion des Systèmes d'IA (génératifs)

`https://theshiftproject.org/publications/intelligence-artificielle-centres-de-donnees-rapport-final/`



Amélioration de l'efficacité énergétique des modèles : IA Frugale/énergies décarbonnées

Attention à l'effet rebond

Besoin d'éducation des usages du numérique

`https://altimpact.fr/`



Coût social des SIA

Travailleurs du click

- Annotateurs Kenyan (Sama) : entre \$1.32/h et \$2/h (source TIME)
- Exposition à des contenus traumatisants

Remplacement pour certaines activitées

- Système d'IA pour **assister** ou pour **remplacer** certains travailleurs ?

→ **Oui !** Certaines activités vont disparaître

- Quelques secteurs potentiellement impactés par la disparition d'activités :
  - production *routinière* de documents (journalisme de brève, rédaction de notes synthétiques, illustrations visuelles ou sonores, codeurs de spécifications, etc.)
  - poste de monitoring visuel (de qualité, en santé, etc.)
- Est-ce un progrès (social) ?

Renversement deshumanisant des positions : humain qui s'adapte aux machines

Impacts des SIA sur la formation des esprits

L'usage de ChatGPT influence notre manière de parler

Empirical evidence of Large Language Model's influence on human spoken communication (2025)



Impacts des SIA sur la formation des esprits

Modèle de langue (LLM) : y a-t-il du sens derrières les représentations ?

Ces modèles proposent une **certaine représentation du monde** : un point de vue, non-neutre

- diffusion très large, parfois cachée de cette représentation

→ imposent-t-ils une vision ?

→ opportunité pour favoriser une pensée complexe ?

Les LLMs sont-ils une nouvelle forme de soft power ?

Modification de la relation au savoir

Un LLM connait beaucoup plus de faits que je ne saurai jamais ...

- Dois je continuer à apprendre par coeur ? quelles conséquences sur la formation des esprits ?
- Accès à ce savoir par des interfaces imposées (parfois payantes)
- Un savoir sans structure (sans taxonomie, sans organisation) ...

→ apprendre à apprendre avec des LLM !

Penser avec ou contre un LLM/un SIA : entre paresse et besoin de justification

Autres questions politiques sur la conception des SIA

Construction d'un SIA (Apprentissage) = minimisation d'une erreur

- Les critères à optimiser sont-ils universels ?
- Comment et qui définit le calcul de l'erreur à optimiser ?

Un SIA ... a un concepteur

- Responsabilité et confiance dans le producteur du SIA
- Responsabilité et confiance dans une autorité certificatrice
- Risque de monopole par quelques concepteurs
  - Concentration des données, ressources de calcul et compétences ...

07
Conclusions

Conclusions I

Le terme « intelligence artificielle » est très utilisé actuellement

- difficile à définir (pas manque de bonne définition de l'intelligence)
- usage (souvent abusif) pour la promesse de modernité ...

Bingo de l'IA en Santé ... quasi complet

- des usages de l'IA en développement dans beaucoup de secteurs de la santé
- trois grandes classes de méthodes : systèmes apprenants, systèmes raisonnants et systèmes agentiques

Conclusions II

Focus sur les systèmes d'intelligence artificielle apprenant

- les systèmes d'intelligence artificielle (SIA) peuvent prendre d'autres forment
- la dimension d'apprentissage automatique soulève le plus de questionnement et de fascination

→ tentative de démystification en présentant

- des mécanismes de construction des SIA
- la notion de représentation sémantique utilisés par les IA Génératives

Beaucoup de questionnements accompagnent la conception de ces outils

- sur les limites des SIA
- sur les réflexions à avoir sur leurs usages (régulation, formation)
- sur les risques qu'il nous faut anticiper

Merci.

Autres questions

Une IA comprend elle ?

Perroquet stochastique vs créativité ?

Photo primée

Méthode de génération d'images ...

Empathie avec l'IA

Neuron, 2024

Autonomie des IA ?

Ethique de conception pour des IA éthiques

Autres définitions plus récentes d'un « système d'IA » ...

AI Act, EU, 2024

Un système basé sur une machine qui est conçu pour fonctionner avec différents niveaux d'autonomie et qui peut faire preuve d'adaptabilité après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des données qu'il reçoit, comment générer des résultats tels que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer des environnements physiques ou virtuels.

OCDE, 2024

Système d'IA : Un système d'intelligence artificielle est un système automatisé qui, pour des objectifs explicites ou implicites, déduit, à partir d'entrées reçues, comment générer des résultats en sortie tels que des prévisions, des contenus, des recommandations ou des décisions qui peuvent influer sur des environnements physiques ou virtuels. Différents systèmes d'IA présentent des degrés variables d'autonomie et d'adaptabilité après déploiement.

Unesco

AI systems are **information-processing technologies** that integrate **models** and **algorithms** that produce a capacity to learn and to perform cognitive tasks leading to outcomes such as prediction and decision-making in material and virtual environments.
AI systems are designed to operate with varying degrees of autonomy by means of **knowledge modelling and representation** and by **exploiting data** and **calculating correlations**. AI systems may include several methods, such as but not limited to :

- machine learning, including deep learning and reinforcement learning ;
- machine reasoning, including planning, scheduling, knowledge representation and reasoning, search,

Pourquoi parler de confiance ? Risques liés aux SIA

The International Scientific Report on the Safety of Advanced AI, January 2025

Risques liés à une utilisation malveillante

- Préjudices aux individus via des contenus falsifiés
- Manipulation de l'opinion publique
- Cyberattaques
- Attaques biologiques et chimiques

Risques liés aux dysfonctionnements

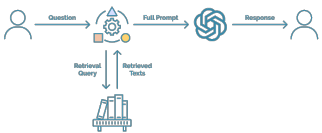
- Problèmes de fiabilité
- Biais
- Perte de contrôle

Risques systémiques

- Risques pour le marché du travail
- Fracture mondiale en matière de R&D en IA
- Concentration du marché et points uniques de défaillance
- Risques environnementaux
- Risques pour la vie privée
- Violations du droit d'auteur

retour

RAG : tendance pour des besoins plus opérationnels



Défauts des modèles conversationnels

- les hallucinations
- données d'entraînement obsolètes
- connaissance non-spécifiques à une entreprise / un domaine

RAG : Retrieval Augmented Generation

Modèle conversationnel branché à une base de connaissances
Modèle génératif de texte utilisé pour structurer la réponse