



Décision critique et intelligence artificielle

Apports, risques et place de l'humain

Dr Adolphe J. Béquet

Chargé de recherche – Neuroergonomie
LESCOT — Université Gustave Eiffel

Journée d'études « IA et défis de la gouvernance »
ICT Toulouse — 12 mars 2026 — Axe 3 (approche cognitive et clinique)

Quand on dit « intelligence artificielle »...

... on pense souvent à :



Comment puis-je vous aider ?

+ Poser une question



Mais la réalité de l'IA est (actuellement) bien plus vaste, diverse et ancienne.



MESSAGE #1

L'IA, ce n'est pas juste ChatGPT ou Skynet.

C'est un ensemble de techniques qui existent depuis les années 1950
et qui sont déjà partout autour de nous.

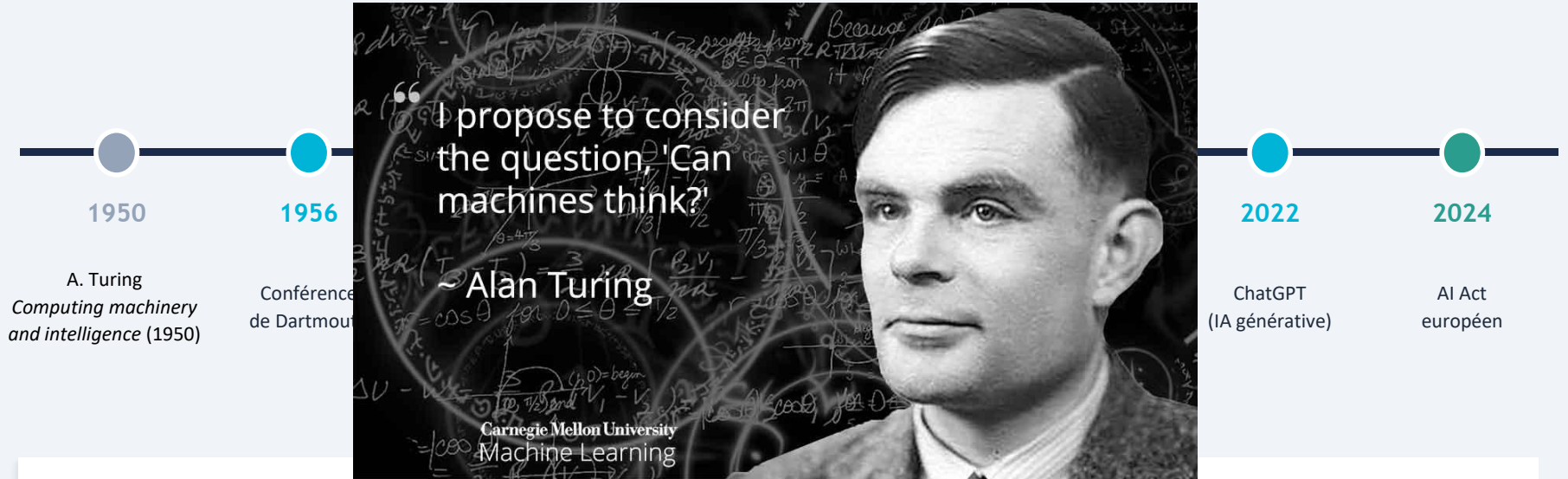
L'IA : une histoire de 70 ans



Repères culturels : Asimov (1950, lois de la robotique) → 2001 l'Odyssée de l'espace (1968) → Blade Runner (1982) → Matrix (1999) → Her (2013) → Ex Machina (2015)

Jalons scientifiques : Perceptron (1958) → Hivers de l'IA (1970s, 1990s) → SVM (1995) → ImageNet (2009) → Transformers (2017) → GPT-3 (2020) → Claude / Gemini (2024)

L'IA : une histoire de 70 ans



Repères culturels : Asimov (1950, lois de la robotique) → 2001 l'Odyssée de l'espace (1968) → Blade Runner (1982) → Matrix (1999) → Her (2013) → Ex Machina (2015)

Jalons scientifiques : Perceptron (1958) → Hivers de l'IA (1970s, 1990s) → SVM (1995) → ImageNet (2009) → Transformers (2017) → GPT-3 (2020) → Claude / Gemini (2024)

L'IA : une histoire à 70 ans

1950

A. Turing
*Computing machinery
and intelligence* (1950)

1956

Conférence
de Dartmouth

Repères culturels : /

(1999) → Her (2013) → E

Jalons scientifiques

(2017) → GPT-3 (2020) →



Photographer: Joe Mehling

Trenchard More, John McCarthy, Marvin Minsky, Oliver Selfridge et Ray Solomonoff
faisaient partie de ceux qui ont assisté à la conférence de Dartmouth sur l'intelligence
artificielle en 1956. [Joe Mehling, CC BY](#)

2022

ChatGPT
(générative)

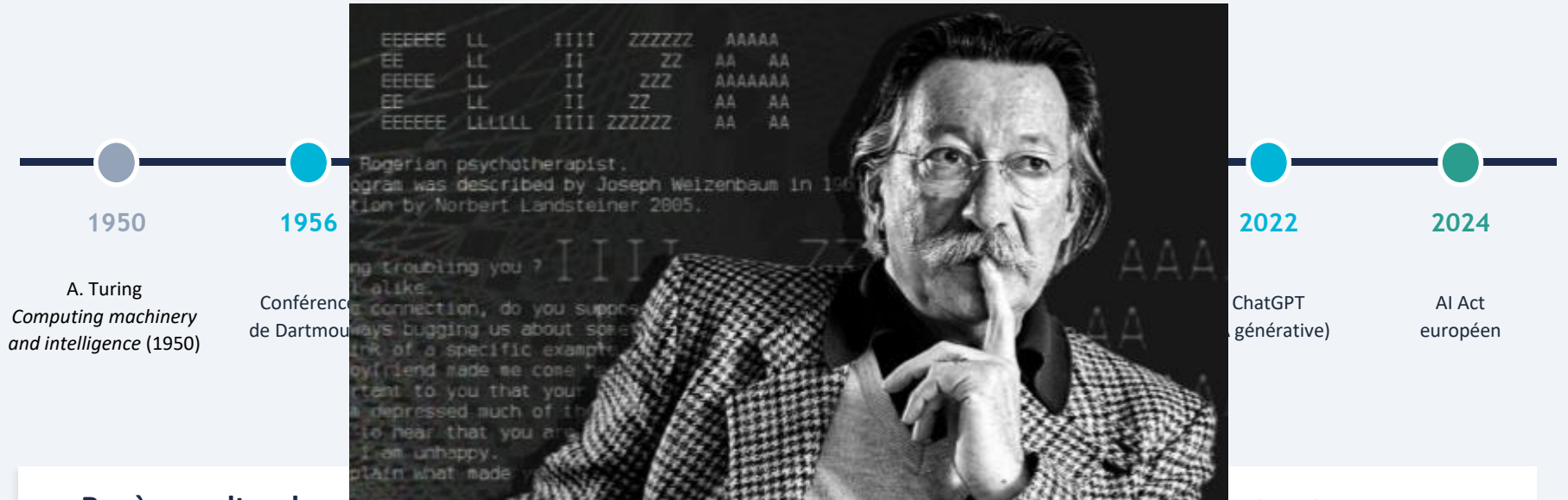
2024

AI Act
européen

r (1982) → Matrix

→ Transformers

L'IA : une histoire de 70 ans



Repères culturels : Asimov (1950, lois de la robotique) → 2001 l'Odyssee de l'espace (1968) → Blade Runner (1982) → Matrix (1999) → Her (2013) → Ex Machina (2015)

Jalons scientifiques : Perceptron (1958) → Hivers de l'IA (1970s, 1990s) → SVM (1995) → ImageNet (2009) → Transformers (2017) → GPT-3 (2020) → Claude / Gemini (2024)

L'IA : une histoire de 70 ans



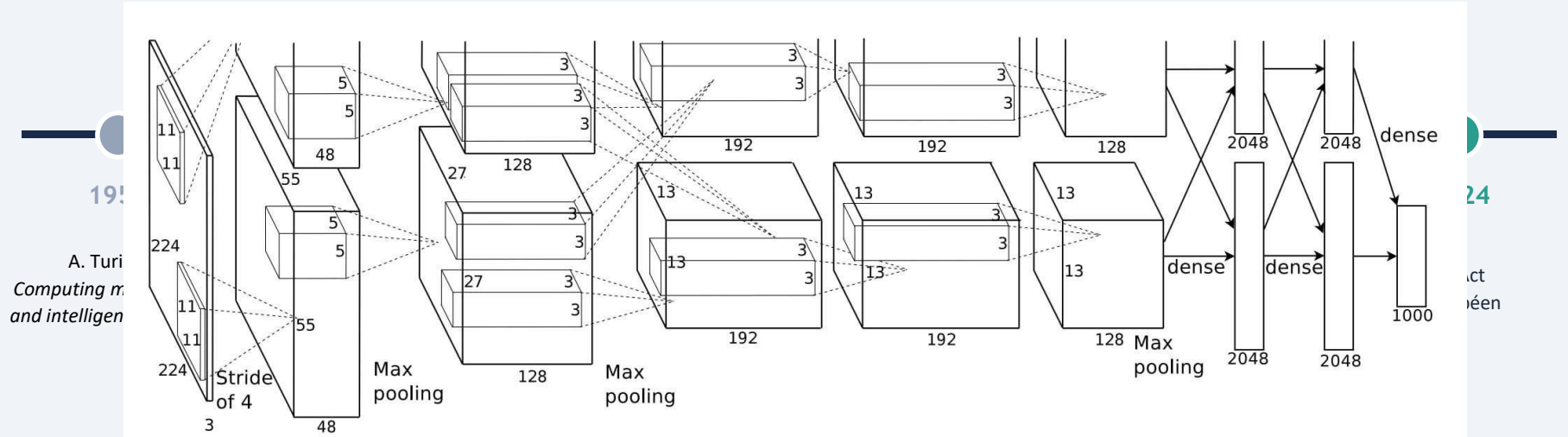
Repères culturels : Asimov (1950, lois de la robotique) → 2001 l'Odyssée de l'espace (1968) → Blade Runner (1982) → Matrix (1999) → Her (2013) → Ex Machina (2015)

Jalons scientifiques : Perceptron (1958) → Hivers de l'IA (1970s, 1990s) → SVM (1995) → ImageNet (2009) → Transformers (2017) → GPT-3 (2020) → Claude / Gemini (2024)

L'IA : une histoire de 70 ans



L'IA : une histoire de 70 ans



Repères culturels : Asimov (1950, lois de la robotique) → 2001 l'Odyssée de l'espace (1968) → Blade Runner (1982) → Matrix (1999) → Her (2013) → Ex Machina (2015)

Jalons scientifiques : Perceptron (1958) → Hivers de l'IA (1970s, 1990s) → SVM (1995) → ImageNet (2009) → Transformers (2017) → GPT-3 (2020) → Claude / Gemini (2024)

L'IA : une histoire de 70 ans



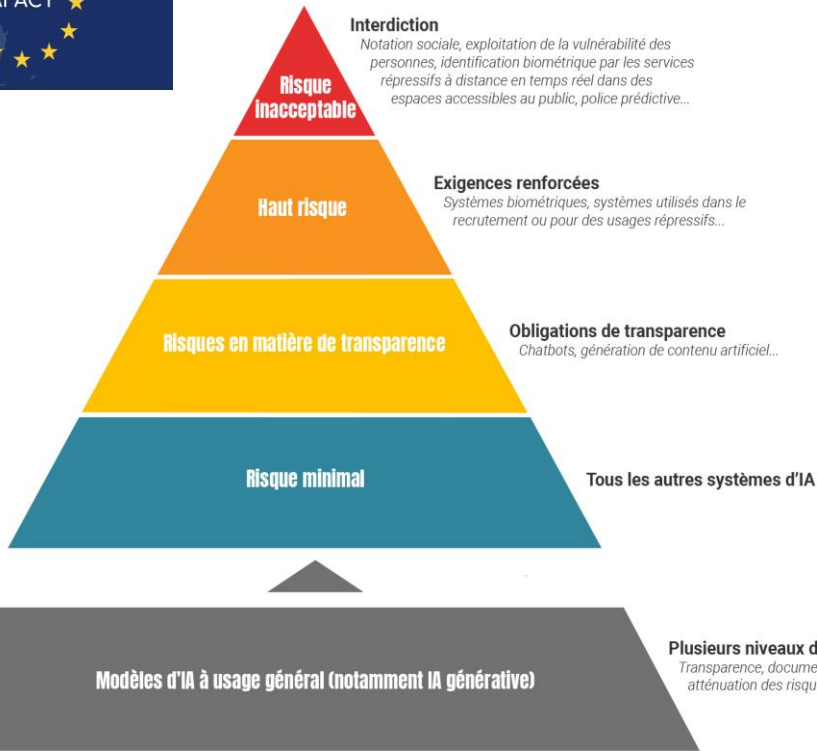
L'IA : une histoire de 70 ans



L'IA : une histoire de 70 ans



Règlement IA - Approche par les risques



1950

1956

A. Turing
Computing machinery and intelligence (1950)

Conférence de Dartmouth (1956)



2022

2024

ChatGPT (IA générative)

AI Act européen

Repères culturels:

(1999) → Her (2013)

Jalons scientifiques:

(2017) → GPT-3 (2020)

(1982) → Matrix

→ Transformers

L'IA : une histoire de 70 ans



Repères culturels : Asimov (1950, lois de la robotique) → 2001 l'Odyssée de l'espace (1968) → Blade Runner (1982) → Matrix (1999) → Her (2013) → Ex Machina (2015)

Jalons scientifiques : Perceptron (1958) → Hivers de l'IA (1970s, 1990s) → SVM (1995) → ImageNet (2009) → Transformers (2017) → GPT-3 (2020) → Claude / Gemini (2024)

Les différentes familles de techniques IA



Systèmes experts & logique

Règles if/then, ontologies.
Historiquement les premiers
systèmes « intelligents ».



Machine Learning classique

Arbres de décision, SVM,
réseaux bayésiens, k-NN.
Apprentissage sur données.



Deep Learning

CNN (vision), RNN
(séquences),
Transformers. Réseaux
de neurones profonds.



IA générative & LLM

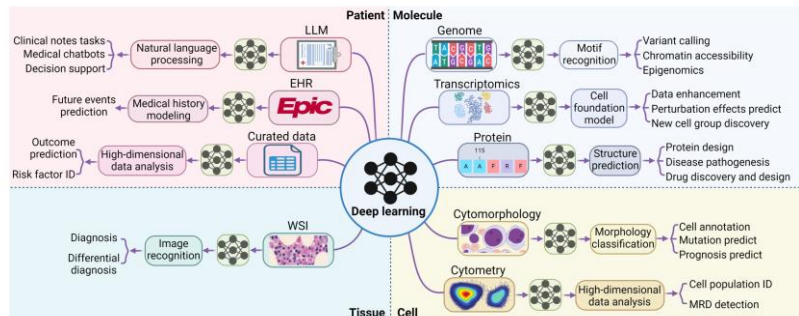
GPT, Claude, Gemini, DALL-E.
Génération de texte, images,
code. Rupture récente
(2022+).

L'IA déployée dans les systèmes critiques repose surtout sur le ML classique et le deep learning — pas sur les LLM.

L'IA : le meilleur et le pire



L'IA qui sauve des vies



Ex : l'IA détecte des cancers invisibles à l'œil nu sur des mammographies avec une précision supérieure aux radiologues.

<https://chi.scholasticahq.com/article/124131-deep-learning-in-hematology-from-molecules-to-patients>



L'IA qui met en danger



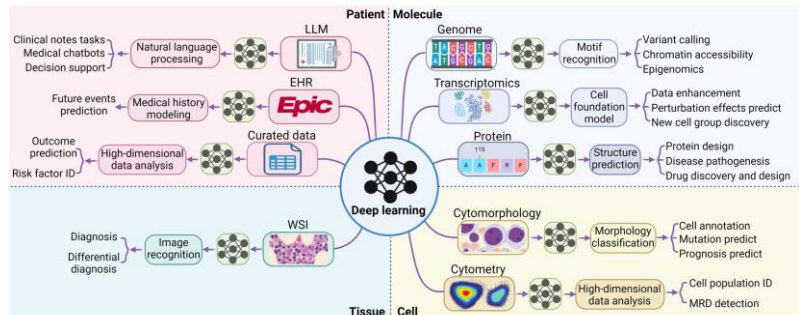
Ex : en 2018, un véhicule autonome Uber percute une piétonne. Le système de détection l'a identifiée mais n'a pas freiné à temps.

<https://ctif.org/fr/news/un-nouvel-accident-mortel-aux-etats-unis-avec-une-tesla-autopilotee-fait-lobjet-dune-enquete>

L'IA : le meilleur et le pire



L'IA qui sauve des vies



Ex : l'IA détecte des cancers invisibles à l'œil nu sur des mammographies avec une précision supérieure aux radiologues.

<https://chi.scholasticahq.com/article/124131-deep-learning-in-hematology-from-molecules-to-patients>



L'IA qui met en danger



Ex : en 2018, un véhicule autonome Uber percute une piétonne. Le système de détection l'a identifiée mais n'a pas freiné à temps.

<https://ctif.org/fr/news/un-nouvel-accident-mortel-aux-etats-unis-avec-une-tesla-autopilotee-fait-lobjet-dune-enquete>

→ L'IA n'est ni bonne ni mauvaise en soi : tout dépend de comment elle est conçue, déployée et gouvernée.

→ Exemple de l'usine à trombones (Bostrom, 2003)

<https://www.decisionproblem.com/paperclips/index2.html>



MESSAGE # 2

**L'IA présente autant de défis
que d'opportunités.**

La question n'est pas « pour ou contre » mais « comment bien faire ».

Opportunités : l'IA au service de la décision critique



Médecine

Diagnostic par imagerie,
détection
précoce, aide à la prescription



Aéronautique

Maintenance prédictive,
optimisation trajectoires,
sécurité



Automobile

ADAS, détection d'obstacles,
assistance au freinage
d'urgence



Urgences

Analyse prédictive
catastrophes,
allocation ressources en temps
réel

(Perez-Cerrolaza et al., 2024 ; Wang & Chung, 2022 ; Mallouhy et al., 2024)

L'IA excelle quand...

- Volumes massifs de données à traiter en temps réel
- Patterns complexes invisibles à l'œil humain
- Tâches répétitives à haute exigence de fiabilité



Défis : risques et effets pervers de l'IA



Défis : risques et effets pervers de l'IA

Opacité algorithmique

« Boîte noire » : décisions inexplicables

Biais de modélisation

Les données reproduisent les inégalités

Déresponsabilisation

Qui est responsable quand l'IA se trompe ?

Vie privée, données, écologie

Collecte massive, consentement flou, coût écologique des algos

(Al-Kfairy et al., 2024 ; Wang & Chung, 2022 ; Perez-Cerrolaza et al., 2024)



Défis : risques et effets pervers de l'IA

Opacité algorithmique

« Boîte noire » : décisions inexplicables

Biais de modélisation

Les données reproduisent les inégalités

Déresponsabilisation

Qui est responsable quand l'IA se trompe ?

Vie privée, données, écologie

Collecte massive, consentement flou, coût écologique des algos

Impact sur les compétences humaines

Deskilling

Érosion des compétences par la délégation à la machine. Le professionnel « désapprend ».

Miskilling

Mauvais calibrage : on forme les gens aux mauvaises compétences face à l'IA.

Neverskilling

Nouvelles générations qui n'acquièrent jamais certaines compétences, remplacées d'emblée par l'IA.

(Al-Kfairy et al., 2024 ; Wang & Chung, 2022 ; Perez-Cerrolaza et al., 2024)



Défis : risques et effets pervers de l'IA

Opacité algorithmique

« Boîte noire » : décisions inexplicables

Biais de modélisation

Les données reproduisent les inégalités

Déresponsabilisation

Qui est responsable quand l'IA se trompe ?

Vie privée, données, écologie

Collecte massive, consentement flou, coût écologique des algos

Impact sur les compétences humaines

Deskilling

Érosion des compétences par la délégation à la machine. Le professionnel « désapprend ».

Miskilling

Mauvais calibrage : on forme les gens aux mauvaises compétences face à l'IA.

Neverskilling

Nouvelles générations qui n'acquiescent jamais certaines compétences, remplacées d'emblée par l'IA.

(Al-Kfairy et al., 2024 ; Wang & Chung, 2022 ; Perez-Cerrolaza et al., 2024)



Le théorème d'impossibilité (Chouldechova & Kleinberg)

Il est mathématiquement impossible de satisfaire simultanément les 3 métriques d'équité algorithmique (parité démographique, égalité des chances, parité prédictive) dès que les taux de base diffèrent entre groupes. Choisir une métrique d'équité n'est pas un choix technique, c'est un choix de société.



MESSAGE #3

L'IA est tant un objet d'étude qu'un outil de recherche

Application : détection automatisée des états internes du conducteur par fusion multimodale de données physiologiques et comportementales.

L'IA : à la fois objet de recherche et outil du quotidien



L'IA comme OBJET DE RECHERCHE



L'IA comme OUTIL DU QUOTIDIEN

L'IA : à la fois objet de recherche et outil du quotidien



L'IA comme OBJET DE RECHERCHE

Détection des états internes du conducteur (Béquet, 2023)

Capteurs

EEG · ECG · EDA
Oculométrie · Volant
Physio + Comportemental



Extraction

Features temporelles
Features fréquentielles
Signal processing



Algorithmes ML

SVM · Random Forest
Deep Learning · Fusion
Classification multimodale

FATIGUE | **DISTRACTION** | **STRESS**

Questions de recherche :

- Quelle combinaison de capteurs est optimale ?
- Comment gérer la variabilité inter-individuelle ?
- Quels seuils d'alerte pour quel compromis sécurité / acceptabilité ?



L'IA comme OUTIL DU QUOTIDIEN

L'IA : à la fois objet de recherche et outil du quotidien



L'IA comme OBJET DE RECHERCHE

Détection des états internes du conducteur (Béquet, 2023)

Capteurs

EEG · ECG · EDA
Oculométrie · Volant
Physio + Comportemental



Extraction

Features temporelles
Features fréquentielles
Signal processing



Algorithmes ML

SVM · Random Forest
Deep Learning · Fusion
Classification multimodale

FATIGUE | **DISTRACTION** | **STRESS**

Questions de recherche :

- Quelle combinaison de capteurs est optimale ?
- Comment gérer la variabilité inter-individuelle ?
- Quels seuils d'alerte pour quel compromis sécurité / acceptabilité ?



L'IA comme OUTIL DU QUOTIDIEN

Usage raisonné et critique des LLM



Revue de littérature

Synthèse, extraction de concepts,
veille scientifique
→ **Retrieval Augmented Generation**
(eg. Google Labs / ScholarQABench)



Analyse exploratoire

Scripts, visualisation,
exploration de données
→ **Claude pour prototyper du code**



Rédaction scientifique

Reformulation, structuration,
traduction, relecture
→ **L'humain valide, l'IA assiste**



Enseignement / Administratif

Supports de cours, exercices,
préparation de présentations
→ **Gain de temps, pas de remplacement**



Principe : usage raisonné et critique

L'IA ne remplace pas l'expertise — elle l'augmente.
Toujours vérifier, garder l'esprit critique, rester l'auteur.

L'IA : à la fois objet de recherche et outil du quotidien



L'IA comme OBJET DE RECHERCHE

Détection des états internes du conducteur (Béquet, 2023)

Capteurs

EEG · ECG · EDA
Oculométrie · Volant
Physio + Comportemental



Extraction

Features temporelles
Features fréquentielles
Signal processing



Algorithmes ML

SVM · Random Forest
Deep Learning · Fusion
Classification multimodale

FATIGUE | **DISTRACTION** | **STRESS**

Questions de recherche :

- Quelle combinaison de capteurs est optimale ?
- Comment gérer la variabilité inter-individuelle ?
- Quels seuils d'alerte pour quel compromis sécurité / acceptabilité ?



L'IA comme OUTIL DU QUOTIDIEN

Usage raisonné et critique des LLM



Revue de littérature

Synthèse, extraction de concepts,
veille scientifique
→ **Retrieval Augmented Generation**
(eg. Google Labs / ScholarQABench)



Analyse exploratoire

Scripts, visualisation,
exploration de données
→ **Claude pour prototyper du code**



Rédaction scientifique

Reformulation, structuration,
traduction, relecture
→ **L'humain valide, l'IA assiste**



Enseignement / Administratif

Supports de cours, exercices,
préparation de présentations
→ **Gain de temps, pas de remplacement**



Principe : usage raisonné et critique

L'IA ne remplace pas l'expertise — elle l'augmente.
Toujours vérifier, garder l'esprit critique, rester l'auteur.

Les deux facettes se nourrissent mutuellement

La recherche sur l'IA éclaire les limites des outils qu'on utilise.
L'usage quotidien des outils IA alimente les questions de recherche.

(Cette présentation a d'ailleurs été préparée avec l'aide de Claude)

À retenir

1



L'IA ne se résume pas à ChatGPT.

C'est un ensemble de techniques qui existent depuis 70 ans et sont déjà déployées partout dans les systèmes critiques.

À retenir

1



L'IA ne se résume pas à ChatGPT.

C'est un ensemble de techniques qui existent depuis 70 ans et sont déjà déployées partout dans les systèmes critiques.

2



L'IA présente autant de défis que d'opportunités.

Biais, opacité, deskilling d'un côté ; diagnostic, sécurité, prédiction de l'autre. La gouvernance est clé.

À retenir

1



L'IA ne se résume pas à ChatGPT.

C'est un ensemble de techniques qui existent depuis 70 ans et sont déjà déployées partout dans les systèmes critiques.

2



L'IA présente autant de défis que d'opportunités.

Biais, opacité, deskilling d'un côté ; diagnostic, sécurité, prédiction de l'autre. La gouvernance est clé.

3



L'IA est un outil et un objet de recherche au quotidien.

Application concrète : détection des états internes du conducteur par fusion multimodale; assisté par l'IA. L'humain reste au centre.



Merci de votre attention

Adolphe J. Béquet

adolphe.bequet@univ-eiffel.fr

LESCOT — Université Gustave Eiffel

Références principales

- AI-Kfairy et al. (2024). Ethical challenges of generative AI. SSRN.
- Béquet, A. J. (2023). Étude du stress et de sa régulation en conduite autonome à l'aune des différences individuelles. Thèse, Lyon.
- Mallouhy et al. (2024). AI in emergency services. IJISA, 16(6).
- Perez-Cerrolaza et al. (2024). AI for safety-critical systems. ACM Comp. Surv., 56(7).
- Wang & Chung (2022). AI in safety-critical systems. IMDS, 122(2).

Quelques recommandations additionnelles :

- EGO (chaine youtube) : https://www.youtube.com/@ego_one
- Basilik de Roko (IA Générale) : <https://www.youtube.com/watch?v=JdxfhMbSwL0>
- IA generative d'Anthropic (largement meilleure que GPT pour l'instant) : <https://www.anthropic.com/>