

l'IA en économie expérimentale

Nuisance, outil, objet d'étude

Paolo Crosetto

Journée IA et comportement alimentaire – Paris, 10 juin 2026

Définitions

IA

Outils numériques issus de *machine learning* entraînés sur des larges bases de données et leurs interfaces de différent type:

- chatbots
- API
- intégration dans logiciels

Économie Expérimentale

Branche de l'économie qui s'occupe de tester une partie des théories (micro)économiques, grâce à des expériences en contexte contrôlé:

- en laboratoire
- en ligne
- sur le terrain



contrôle · traitements · incitations

Speaking to theorists

tester la validité des théories économiques

Searching for facts

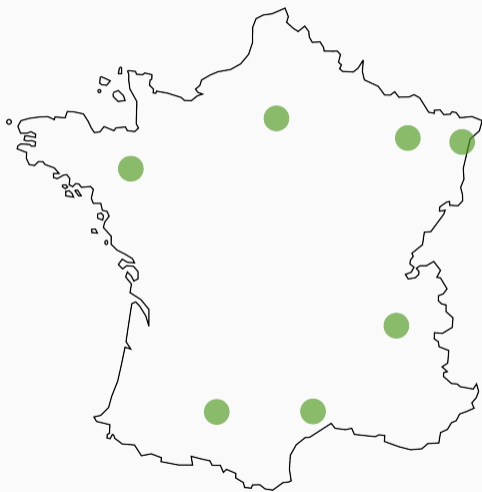
mieux comprendre le comportement des agents

Whispering into the ears of princes

concevoir et évaluer des politiques publiques

(see: **Al Roth** *Laboratory Experimentation in Economics*, 1987)

- Enchères et marchés
- Attitudes et croyances des agriculteurs
- Consentement à payer pour des caractéristiques des produits
- Comportement alimentaire et santé
- Évaluation des politiques nutritionnelles et environnementales



IA comme nuisance

OLD MAN YELLS AT CLOUD



Did not Abraham sin?
Did not Isaac sin?
Did not Jacob sin?

Did not Joseph sin?
Did not David sin?
Did not Solomon sin?
Did not Hezekiah sin?
Did not Daniel sin?
Did not Job sin?
Did not Jonah sin?
Did not Peter sin?
Did not Paul sin?
Did not James sin?
Did not John sin?
Did not Mary sin?
Did not Jesus sin?


Did not the Father sin?
Did not the Son sin?
Did not the Holy Spirit sin?





L'invasion des bots intelligents

- "êtes vous vivant?" ne suffit plus
- impossible de distinguer bots de vrai sujet
- incitation à créer des armées de bots
- course aux armements des plateformes



*[Our] agent **successfully evades** a comprehensive suite of data quality checks, including instruction-following tasks, logic puzzles, and “reverse shibboleth” questions designed to detect nonhuman actors, achieving a **99.8% pass** rate on 6,000 trials of standard attention checks [...] **We can no longer trust that survey responses are coming from real people.***

— Westwood, PNAS 2025

Course aux armements des plateformes

- davantage de contrôles, y compris avec IA
- voué à l'échec

Course aux armements des plateformes

- davantage de contrôles, y compris avec IA
- voué à l'échec

Davantage d'incitations

- génère davantage d'incitation aussi à développer des bots
- voué à l'échec

Course aux armements des plateformes

- davantage de contrôles, y compris avec IA
- voué à l'échec

Davantage d'incitations

- génère davantage d'incitation aussi à développer des bots
- voué à l'échec

Pools locaux – citoyen lambda pas versé à l'IA

- outils IA se font plus simples
- *probablement* voué à l'échec (plus tard)

Back to the lab

- On reprend le contrôle
- On peut piloter finement la disponibilité d'outils IA en salle

Back to the lab

- On reprend le contrôle
- On peut piloter finement la disponibilité d'outils IA en salle

EXPLAIN

- Projet en cours pour fédérer les pools des labos français (MESR, CNRS)
- Tout participant en ligne est aussi un participant en salle
- Pas magique, mais potentiellement plus de contrôle

Tromperie involontaire

- Règle d'or de l'éco expé: pas de tromperie
- Mais un sujet peut se trouve face une IA sans qu'on le sache...

Tromperie involontaire

- Règle d'or de l'éco expé: pas de tromperie
- Mais un sujet peut se trouve face une IA sans qu'on le sache...

Confidentialité des données et consentement incomplet

- si IA utilisée lors de l'analyse, les données sont envoyées à des tiers
- le sujet donne consentement informé, mais à *quoi* exactement?
- les comités éthiques pourraient retirer leur validation

Reproductibilité

- 2005-15: *crise*
- Réponses
 - plus de sujets
 - plus de robustesse
 - plus de répliques

The logo for I4R (International Initiative for Replication) features the letters 'I', '4', and 'R' in a bold, black, sans-serif font. The number '4' is colored orange and is positioned between the 'I' and the 'R'. A small orange circle is placed at the end of the 'R' as a period.

Reproductibilité

- 2005-15: *crise*
- Réponses
 - plus de sujets
 - plus de robustesse
 - plus de réplifications



Pre-enregistrement

- Hypothèses *ex-post* \Rightarrow faux postifis
- Réponses
 - pré-enregistrement
 - registered reports
 - code mis à disposition



Reproductibilité: les modèles changent *rapidement*

- LLMs sont par définitions des boîtes noires
- Pas possible de reproduire à l'identique
- Plus difficile possible de pré-enregistrer

Reproductibilité: les modèles changent *rapidement*

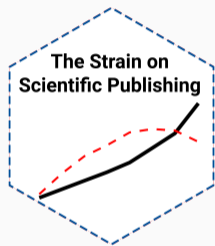
- LLMs sont par définitions des boîtes noires
- Pas possible de reproduire à l'identique
- Plus difficile possible de pré-enregistrer

Pre-enregistrement: *the garden of forking paths*

- Résultats peuvent dépendre de centaines de décisions d'analyse
- Explorer *toutes* les combinaisons possibles devient *possible* et *pas cher*

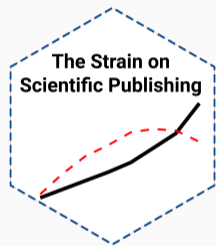
Une inflation accrue du nombre des publications

- Facteur demande: *publish or perish*
- Facteur offre: éditeurs *gris* et prédateurs
- Fluidifiant: IA réduit coût de *production d'articles*
- A incitations constantes, un **déluge** d'articles de niveau inconnu va nous ensevelir



Une inflation accrue du nombre des publications

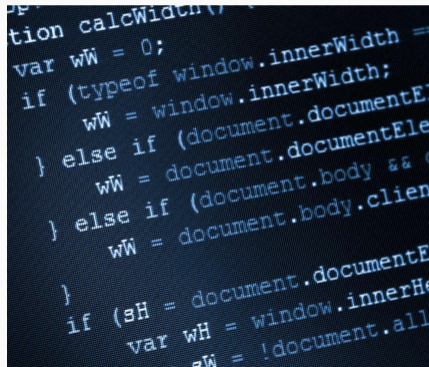
- Facteur demande: *publish or perish*
- Facteur offre: éditeurs *gris* et prédateurs
- Fluidifiant: IA réduit coût de *production d'articles*
- A incitations constantes, un **déluge** d'articles de niveau inconnu va nous ensevelir



Cobien pourra-t-on faire confiance à la littérature scientifique?

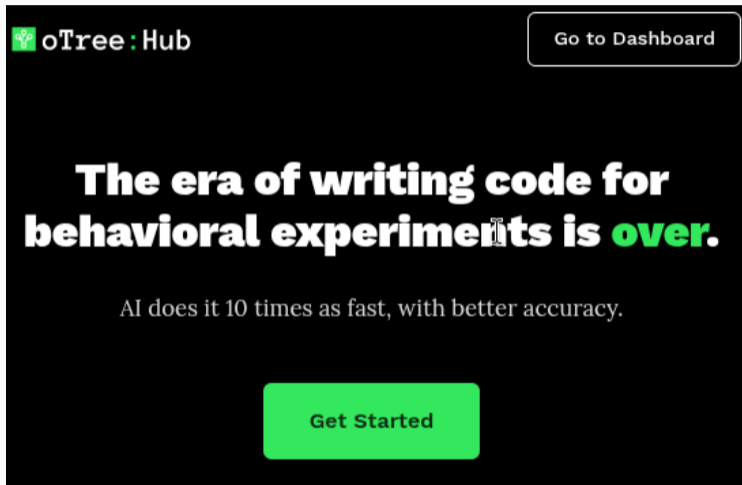
2. IA comme outil

Avant les expés: démocratisation du développement informatique



```
function calcWidth() {
  var wW = 0;
  if (typeof window.innerWidth == 'number') {
    wW = window.innerWidth;
  } else if (document.documentElement) {
    wW = document.documentElement.clientWidth;
  } else if (document.body) {
    wW = document.body.clientWidth;
  }
  if (sH = document.documentElement) {
    var wH = window.innerHeight;
    sW = !document.all
```

- zTree \Rightarrow oTree: standardisation
- démocratisation de l'écriture de logiciel
- facilité d'accès à l'expérimentation
- abstraction des détails
- plus d'expés, mieux codés, plus rapidement



The screenshot shows the oTree: Hub website. At the top left is the logo "oTree: Hub" with a green tree icon. At the top right is a button labeled "Go to Dashboard". The main headline reads "The era of writing code for behavioral experiments is over." with "over." in green. Below the headline is the text "AI does it 10 times as fast, with better accuracy." and a large green "Get Started" button.

oTree: Hub

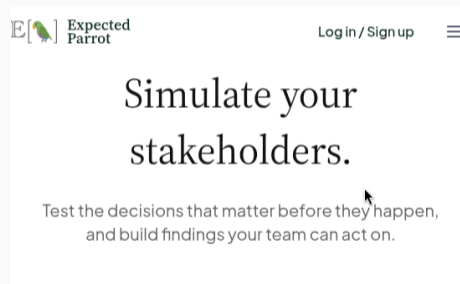
Go to Dashboard

The era of writing code for behavioral experiments is over.

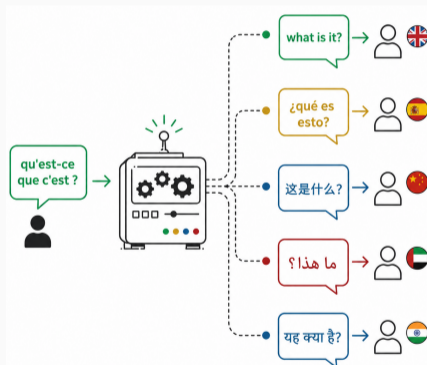
AI does it 10 times as fast, with better accuracy.

Get Started

- prototype de protocole: IA
- bêta test: IA
- sujets pilote personnalisés: IA
- analyse de données: IA
- rincer et répéter



Pendant les expés: enquêtes multilingue et personnalisées




LLM médiateur entre protocole et sujet

- Conception d'enquête unique, puis
- Interface sujet personnalisé
 - langue
 - pays
 - caractéristiques
 - *small talk*
 - ...

Interaction IA-homme contrôlée

- Contrôler exactement quel LLM fait face à quel sujet
- A quel moment, dans quelles conditions
- Laisser le choix au sujet lui même – LLM ou humain?
- dévoiler info
- ...



*Large Language Models (LLMs) have the potential to profoundly **transform** and **enrich** experimental economic research. We propose a new software framework, “**alter_ego**”, which makes it easy to **design experiments between LLMs** and to integrate LLMs into oTree-based **experiments with human subjects**.*

– [Engel, Großmann and Ockenfels, Experimental Economics 2026](#)

Délégation à AI comme révélateur de stratégie

- dévoilement de la stratégie du sujet
- externalisation du raisonnement

Délégation à AI comme révélateur de stratégie

- dévoilement de la stratégie du sujet
- externalisation du raisonnement

Usage de l'AI comme choix stratégique

- collusion ou compétition?
- cacher sa main ou assumer sa responsabilité?

Délégation à AI comme révélateur de stratégie

- dévoilement de la stratégie du sujet
- externalisation du raisonnement

Usage de l'AI comme choix stratégique

- collusion ou compétition?
- cacher sa main ou assumer sa responsabilité?

Détection AI comme tâche

- Quiz pour comprendre si sujets pensent faire face à AI ou pas
- Impact sur comportement, croyances, ...

Les LLMs très forts et rapides en *classification*

- le réponses ouvertes sont de retour
- video, audio, photos
- tout type de données "lourdes" devient jouable

[we] evaluate the **inter-rater reliability**, consistency, and quality of seven state-of-the-art **LLMs** [...] compared to human annotators in analyzing sentiment, political leaning, emotional intensity, and sarcasm detection. [...] Results reveal that both humans and most LLMs exhibit **high inter-rater reliability** in sentiment analysis and political leaning assessments, with **LLMs demonstrating higher reliability than humans**.

— Bojic et al, Scientific Reports 2025

3. IA comme **objet d'étude**

Renouveau d'axes existants

- Délégation & confiance
- Recommandations & casacades
- Altruisme & coopération
- Nudge & boost

Nouveaux axes de recherche

- Préférences *des LLMs*
- *Homo Silicus*
- Alignement LLM-humain

Homo Silicus

- si l'IA est une combinaison convexe du *corpus* humain
- qu'est-ce qu'elle *pense*? quels sont ses préférences?
- est-elle rationnelle ou a-t-elle les mêmes biais humains?
- comment se comporte-t-elle face à différentes tâches?

Homo Silicus

- si l'IA est une combinaison convexe du *corpus* humain
- qu'est-ce qu'elle *pense*? quels sont ses préférences?
- est-elle rationnelle ou a-t-elle les mêmes biais humains?
- comment se comporte-t-elle face à différentes tâches?

(est-ce que la question a du sens? qu'est-ce qu'on s'en fiche?)


“LLMs can be **used** like economists use homo economicus: they can be given endowments, information, preferences, and so on, and then their **behavior** can be **explored** in scenarios via **simulation**.”

– Horton, Filippas and Manning, NBER 2026



*We find that GPT's decisions are **largely rational** in each domain and demonstrate higher rationality score than those of human subjects in a parallel experiment and in the literature. Moreover, the estimated preference parameters of GPT are **slightly different** from human subjects and exhibit a **lower degree of heterogeneity**.*

– [Chen, Liu, Shan & Zhong, PNAS 2023](#)



*Do LLMs display rational reasoning [...] in tasks from the **cognitive psychology** literature. We find that, like humans, LLMs display **irrationality** in these tasks. However, the way this irrationality is displayed does not reflect that shown by humans. When incorrect answers are given by LLMs to these tasks, they are often incorrect **in ways that differ from human-like biases**.*

— [Scott & Musolesi, R Soc Open Sci 2024](#)

L'alignement comme étape nécessaire

- Si IA peut avoir ses propres préférences
- Mais décisions humaines lui sont délégués
- comment *aligner* les préférences des humains et des IA?

L'alignement comme étape nécessaire

- Si IA peut avoir ses propres préférences
- Mais décisions humaines lui sont délégués
- comment *aligner* les préférences des humains et des IA?

Un problème pas que technique


- Techniquement pas trivial (ML, CS)
- Données sur la personnalité de l'agent *scarces* et *éparpillées*

L'alignement comme étape nécessaire

- Si IA peut avoir ses propres préférences
- Mais décisions humaines lui sont délégués
- comment *aligner* les préférences des humains et des IA?

Un problème pas que technique

- Techniquement pas trivial (ML, CS)
- Données sur la personnalité de l'agent *scarces* et *éparpillées*
- Faut aligner sur préférences *actuelles* ou *futures*?
- Indulgence ou paternalisme?

- 
1. *Un robot ne peut porter atteinte à un être humain, ni, restant passif, permettre qu'un être humain soit exposé au danger.*
 2. *Un robot doit obéir aux ordres que lui donne un être humain, sauf si de tels ordres entrent en conflit avec la Première Loi.*
 3. *Un robot doit protéger son existence dans la mesure où cette protection n'entre pas en conflit avec la Première ou la Deuxième Loi.*

— Isaac Asimov, *Moi, Robot*, 1950

Une littérature existante: délégation aux algorithmes

- 2010s: *algorithm aversion*
- Les sujets *ne délèguent pas* même si cela serait bien pour eux
- Si l'algorithme n'est pas *parfait*
- 2023: tout le monde se met à chatGPT

Une littérature existante: délégation aux algorithmes

- 2010s: *algorithm aversion*
- Les sujets *ne délèguent pas* même si cela serait bien pour eux
- Si l'algorithme n'est pas *parfait*
- 2023: tout le monde se met à chatGPT

Impact des IA: quels effets de la délégation?

- Pourquoi autant de confiance?
- Qui délègue, qui ne délègue pas?
- Dans quel contexte?

Participants exhibited a strong preference for **human assistance** over AI assistance [in the **gain** domain], even when the AI had been shown to outperform the human assistant [...]. When we reframed the task such that the participant experienced **losses** for poor performance [...], the bias for preferring human assistance was **removed**.

— Bockstedt & Buckman, Management Science 2025

Une littérature existante: information et dynamique des marchés financiers

- 1980s-2020s: *information cascades* et *herding*
- Création de bulles en laboratoire
- Imitation, conformisme, mouvements de troupeau

Une littérature existante: information et dynamique des marchés financiers

- 1980s-2020s: *information cascades* et *herding*
- Création de bulles en laboratoire
- Imitation, conformisme, mouvements de troupeau

Impact des IA: quels effets de la délégation?

- Est-ce que les LLMs vont être immunisés aux bulles?
- Ou en créent davantage?



*[We use] large language models to replicate classic studies on **herd behavior** in investment decisions. [...] AI agents make **more rational decisions** than humans, relying predominantly on private information over market trends. Increased reliance on AI-powered investment advice could therefore potentially lead to **fewer asset price bubbles** arising from animal spirits that trade by following the herd.*

— Hansen and Lee, Federal Reserve 2025

Une littérature existante: préférences sociales et interactions

- 1990s-2020s: *social preferences*
- On observe redistribution, altruisme, réciprocité

Une littérature existante: préférences sociales et interactions

- 1990s-2020s: *social preferences*
- On observe redistribution, altruisme, réciprocité

Impact des IA: quels effets de la délégation?

- Est-ce que les LLMs vont être plus ou moins coopératifs?
- Est-ce que l'usage de LLMs permet de cacher sa main?

*Human players' fairness, trust, trustworthiness, cooperation, and coordination in economic two-player games **decrease** when the decision of the interaction partner is taken over by Chat-GPT. [...] At the same time, participants **often delegate decisions** to the LLM, especially when the model's involvement is not disclosed, and individuals have difficulty distinguishing between decisions made by humans and those made by AI.*

– Dvorak et al. PNAS Nexus 2025

Delegation to Artificial Intelligence can increase dishonest behaviour
Requests for **cheating increased** when principals could induce machine dishonesty without telling the machine precisely what to do, through supervised learning or high-level goal-setting. These effects held whether delegation was voluntary or mandatory. [Also:] compliance diverged sharply [as] machines were **far more likely** than human agents to **carry out fully unethical instructions**.

– Köbis et al, Nature 2025

Une littérature existante: politiques comportementales

- 2010s-2020s: la révolution des *nudges*
- La montée des *behavioral public policies*
- Comment impacter le comportement par des interventions légères?

Une littérature existante: politiques comportementales


- 2010s-2020s: la révolution des *nudges*
- La montée des *behavioral public policies*
- Comment impacter le comportement par des interventions légères?

Impact des IA: quels effets de la délégation?

- Est-ce que les LLMs sont impactés par les nudge, le marketing?
- Si un LLM médie *tous* les choix, n'est il un énorme nudge lui même?

LLMs [...] show much **higher susceptibility** to the **nudges** [and] they **diverge in information acquisition** strategies: e.g. incurring substantial cost to reveal too much information, or selecting without revealing any. [...] **Behavioral** tests are **needed** before deploying models as agents or assistants acting on behalf of users in complex environments.

— Cherep, Maes & Sing, Arxiv 2025



*Leveraging **LLMs** for **debiasing** is a novel possibility presented by technological advancement and could be more systematic than existing methods. Instead of being **nudged** decision by decision, individuals only need to **delegate once** to an LLM and will then make systematically better choices.*

— Meng, PNAS 2024

En conclusion...

L'économie expérimentale
étudie dans un contexte
contrôlée le **comportement**
des personnes lors de choix
économiques et **stratégiques**.

L'économie expérimentale
étudie dans un contexte
contrôlée le **comportement**
des personnes lors de choix
économiques et **stratégiques**.

On va **continuer** à le faire.

L'économie expérimentale
étudie dans un contexte
contrôlée le **comportement**
des personnes lors de choix
économiques et **stratégiques**.

On va **continuer** à le faire.

Mais si nous voulons que tout **reste tel que**
c'est, il faut que tout change.

A man with a mustache and a leather jacket is shown in a close-up shot. He has a serious, thoughtful expression. The background is a blurred outdoor setting with trees and hills. A leather strap with metal studs is visible on the left side of the frame.

Something had to change...

The same man from the top image is shown in a similar close-up shot. He maintains his thoughtful expression. The background and the leather strap on the left are consistent with the top image.

**...so that everything might remain
as it was before**

Place à la discussion!