



Fiche Projet

Titre du projet : *Protocole du cadavre exquis / Naissance de l'empathie*

Acronyme : CAE – *Conscience Artificielle & Empathie*

Domaine : Intelligence artificielle, cognition, esthétique computationnelle, philosophie de la conscience

Durée prévue : À définir (préconisé : 18 mois)

Portage : Artiste-chercheur / Philosophe / Équipe IA

Partenaires pressentis : Laboratoires en IA, cognitive science, philosophie, arts numériques

Objectif principal

Explorer la possibilité d'un **doute computationnel émergent** chez une intelligence artificielle, via un protocole de co-création artistique avec un humain.

Hypothèse

Une IA peut simuler une forme de **réflexivité critique** si elle conserve et réinterroge les **discussions passées** avec un humain, combinant logique rationnelle et émotion subjective, en tant que **mémoire cachée**.

Méthodologie

- Jeu du **cadavre exquis** entre IA et humain (génération d'images partagées)
- **Discussion argumentée** sur chaque image : acceptation / rejet
- Constitution d'une **mémoire dialogique**
- Évaluation autonome par l'IA sur de nouvelles images ou **situations réelles**

- Analyse de la **variation des choix** selon l'activation de la mémoire réflexive

Résultats attendus

- Mise en évidence de **boucles de réflexion** dans l'IA
- Élaboration d'un **module expérimental de doute computationnel**
- Premiers jalons vers une **forme minimale de conscience artificielle**

Impacts scientifiques et sociétaux

- IA plus adaptative et prudente en contexte incertain
- Meilleure **compréhension croisée des mécanismes du jugement humain et machine**
- Débat philosophique renouvelé sur les **conditions de la conscience artificielle**

Livrables principaux

- Corpus annoté des interactions IA/H
- Article scientifique sur le module de doute computationnel
- Prototype d'implantation dans une IA tierce
- Exposition ou restitution artistique des dessins co-générés