

Projet Tremplin Recherche 2025-2026

Titre du projet

Métafort-e ! Jeu de linguistique appliquée

Équipe encadrante

- Rose Moreau Raguenes (rose.moreau-raguenes@univ-eiffel.fr), co-porteuse du projet – Université Gustave Eiffel (LISAA) et CY Cergy Paris Université (AGORA / Institut des Humanités Numériques)
- Julien Longhi (julien.longhi@cyu.fr), co-porteur du projet – CY Cergy Paris Université (AGORA / Institut des Humanités Numériques)
- JérémY Demange (jeremy.demange@cyu.fr), chargé de développement d'outils numériques pour la valorisation – CY Cergy Paris Université (Pôle Universitaire d'Innovation CY Tranfer)

Filières visées

Informatique, Data-science et intelligence artificielle (E4/M1).

Le projet pourra être rattaché au Graduate program Digis (<https://digis.hypotheses.org/>), et dans ce cadre l'étudiant·e pourra postuler à une bourse de soutien mensuelle (auquel cas une poursuite en stage sera envisageable).

Présentation générale

Métafort-e ! est un projet Science et société en linguistique appliquée visant à créer un jeu pour sensibiliser/former aux usages et aux effets du langage en contexte. Un premier prototype (centré sur les effets des métaphores dans les discours politiques et médiatiques) a été testé avec succès lors de la Fête de la Science 2024, et *Métafort-e* est lauréat du programme Innovation et sociétés (Pôle Universitaire d'Innovation CY Transfer). La phase de prématuration, qui débute à l'automne 2025, vise à passer du prototype à un jeu autonome, diffusable et commercialisable et à préparer des déclinaisons futures (envisagées autour de phénomènes comme la violence verbale mais également pour former aux usages, compétences et normes discursives et interactionnelles propres à différentes sphères professionnelles).

Objectifs du projet

La contribution de l'étudiant·e en E4/M1 consistera à développer une interface, des outils d'extraction et de structuration de données linguistiques, afin d'alimenter le jeu et de stabiliser une méthodologie transférable aux éditions ultérieures du jeu.

Le projet sera formateur en termes de recherche avec et pour la société, notamment en termes de problématiques liées à l'application de l'IA aux projets en sciences humaines et sociales.

Les tâches concerneront les éléments suivants :

1. Collecte de données : mise en place d'un système de récupération de citations courtes issues de sources multiples [1] (par ex. presse en ligne, réseaux sociaux, discours écrit et oral). Selon les contraintes d'accès, cela pourra passer par des API existantes, par des techniques de scraping, ou par l'utilisation de corpus déjà constitués et diffusés en open access.
2. Prétraitement et structuration des données récoltées pour mettre en place une base structurée et exploitable.
3. Détection automatique de phénomènes linguistiques comme la métaphore [2] : revue des approches existantes (lexiques, règles heuristiques, modèles supervisés ou embeddings) ; expérimentation sur un sous-corpus de citations collectées ; comparaison de méthodes simples avec des modèles plus avancés (BERT, LLMs [3]).
4. Prototype d'interface : développement d'un outil minimal (web app) permettant de visualiser les citations sélectionnées et de générer semi-automatiquement des cartes de jeu (texte + métadonnées).

L'étudiant-e bénéficiera d'un suivi régulier ainsi que de soutien technique au sein du Pôle d'Innovation Universitaire CY Transfer. Modalité : distanciel, avec possibilité de points sur le campus de Marne-la-Vallée avec l'encadrante locale.

Bibliographie

[1] VIGOUROUX, Christian (2020). *Robotic Process Automation, Automatisez vos processus métier*. Éditions ENI.

[2] SU, Chang, HUANG, Shuman & CHEN, Yijang (2017). Automatic detection and interpretation of nominal metaphor based on the theory of meaning. *Neurocomputing*, vol. 219, 300-311.

[3] XIAO, Tong et ZHU, Jingbo (2025). Foundations of Large Language Models. <https://arxiv.org/abs/2501.09223>.