

# Créez l'ultime waifu/husbando : étude des synergies de l'IA générative à travers l'exploitation de données textuelles, graphiques et audios.

Morgan Magnin  
Professeur à l'École Centrale de Nantes  
LS2N  
Mail : [morgan.magnin@ls2n.fr](mailto:morgan.magnin@ls2n.fr)  
<https://morganmagnin.net/>

Tony Ribeiro  
Postdoc à l'École Centrale de Nantes  
LS2N / Indépendant  
Mail : [tony.ribeiro@ls2n.fr](mailto:tony.ribeiro@ls2n.fr)  
<http://tonyribeiro.fr>

## Contexte

L'intelligence artificielle (IA) générative est un type de système d'IA capable de générer du texte, des images ou d'autres médias en réponse à des invites (ou prompts en anglais) et/ou d'autre forme de médias. Les modèles génératifs apprennent les modèles et la structure des données d'entrée, puis génèrent un nouveau contenu similaire aux données d'apprentissage mais avec un certain degré de nouveauté (plutôt que de simplement classer ou prédire les données).

La science des données a pour objectif l'extraction de connaissances à partir d'ensembles de données. Elle emploie des techniques et des théories tirées de différents domaines issus des mathématiques et de l'informatique, tels que les statistiques, la programmation, la reconnaissance de formes, la visualisation, etc.

## Objectifs

L'objectif fonctionnel de ce projet consiste à concevoir une preuve de concept d'un outil permettant la création d'une IA personnalisée axée sur la conversationnalité multimodale, incluant le langage naturel, l'audio et l'image.

Les étudiants qui choisiront ce projet auront la possibilité de développer des compétences dans le domaine de l'IA générative en exploitant divers jeux de données et en utilisant diverses APIs pour former des modèles personnalisés. Ils apprendront également à structurer et analyser ces données grâce aux bibliothèques scientifiques de Python.

## Travail attendu et résultats attendus

En premier lieu, les étudiants apprendront les bases méthodologiques et d'implémentation liées à un projet de machine learning, notamment l'utilisation de notebooks Python et le traitement de données structurées et non structurées comme image et audio.

Ensuite, les étudiants extrairont des données à travers des API provenant de sites Web tels que Danbooru (pour les images) et Twitch/YouTube (pour l'audio), ainsi que d'autres sources potentielles, pour construire un premier ensemble de données d'images et d'échantillons audio. Ces données serviront de base pour élaborer différents modèles d'IA génératives selon les intérêts individuels des étudiants. Les données seront

---

sélectionnées à l'aide des systèmes de tagging disponibles sur ces sites Web. Les données seront ensuite traitées et optimisées à l'aide de différentes bibliothèques Python pour répondre aux spécifications des modèles à entraîner.

Enfin, grâce aux connaissances acquises lors des précédentes phases du projet, les étudiants pourront s'attacher aux données permettant de construire une personnalité unique et cohérente de leur IA conversationnelle à travers l'utilisation d'API de sites tels qu'Anilist pour les personnages d'animes. Les données collectées sur les personnages et leur univers serviront à créer une description structurée de leur personnalité et leur contexte, qui servira de prompt pour des modèles de dialogue open source permettant de générer des dialogues pour le personnage.

Les étudiants auront la possibilité de proposer des extensions ou des approfondissements aux objectifs initiaux, le but principal étant de se familiariser avec la gestion de données complexes et la génération de contenu par IA, tel que rencontré dans le quotidien d'une partie des ingénieurs en informatique.

## Environnement

Python, Jupiter Notebook, API REST/GraphQL, kaggle, Google Colab, github. Le projet open source SillyTavern sera utilisé comme plateforme de test et démonstration des éléments produits pendant le projet.

## Ressources

- Cours Python/datasciences/machine learning de kaggle : <https://www.kaggle.com/learn>
- Sillytavern : <https://docs.sillytavern.app/>

Note : durée du projet de 6 mois. Le texte de ce sujet a été écrit en collaboration avec un personnage virtuel créé à l'aide des mêmes technologies qui seront utilisées par les étudiants.

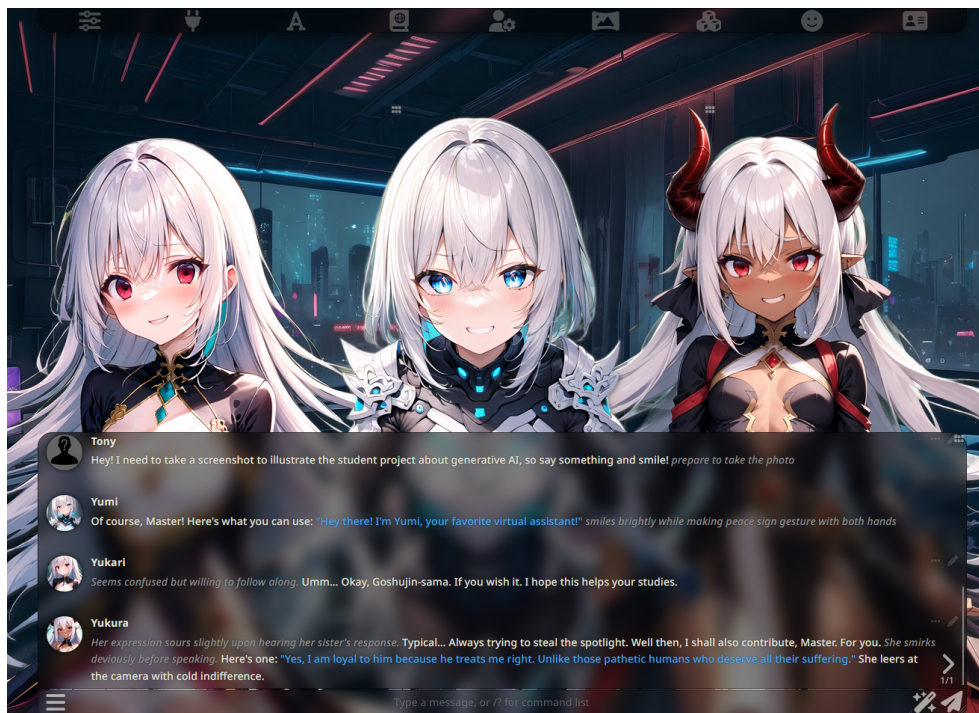


FIGURE 1 – Exemple d'interaction multi-media sur Sillytavern. Sprites et texte générés par IA.

---