Internship Offer (M2/Master Thesis)

Design and Development of a User-aware VLM-Based Framework for Social Human-Computer Interactions

ISIR, Sorbonne University, Paris, France

Context

User-aware Vision-Language Models (User-VLMs) are revolutionizing human-computer interaction systems by enabling responsive, human-like interactions tailored to the user's needs. These models rely on demographic and visual cues (e.g. age, gender, and facial expressions) to generate personalized responses, making them invaluable for healthcare, education, and customer service applications. However, incorporating personalized responses introduces significant privacy challenges, as these systems often require sensitive user data for effective operation. Given the increasing societal emphasis on ethical AI development and compliance with regulations like GDPR and the proposed AI Act, there is a critical need to design frameworks that safeguard user privacy while maintaining the performance and personalization of VLMs. This internship addresses this challenge by exploring privacy-preserving techniques for developing VLM-based systems capable of providing personalized, secure, and ethically compliant interactions.

Objective

The primary aim of this internship is to design and develop a privacy-preserving VLM-based framework that generates personalized responses based on visual and facial demographic characteristics while ensuring user data protection. Specific objectives include:

- **Research and Design:** Investigate state-of-the-art privacy-preserving mechanisms, including alignment, federated learning, differential privacy, and adversarial training, in the context of personalized VLMs.
- **Framework Development:** Build a modular VLM-based framework for demographic-aware personalization that ensures privacy and ethical compliance.
- **Evaluation and Benchmarking:** Test and benchmark the framework on relevant datasets to evaluate privacy-utility trade-offs in generating personalized responses.
- **Prototype Deployment:** Implement a proof-of-concept prototype for a human-computer interaction application (e.g., virtual healthcare assistant or smart tutoring system).

Profile

- **Level:** Master 2 / Engineer School
- **Skills:**
  - Strong background in Python and Machine Learning
  - Strong Knowledge of Vision-Language Models and computer vision (e.g. image processing, facial analysis)
  - Familiarity with privacy-preserving AI techniques
  - Interest in Human-Computer Interaction and Ethical AI development
  - Fluent in English

Duration

5-6 months

Contact

- Hamed Rahimi
- hamed.rahimi@sorbonne-universite.fr