

# REPORT OUTLINE

*Federated RLHF Pipeline for Personalized Learning with ScholéAI*

## 1. Introduction

- Motivations : personalization, scalability, and alignment in educational AI.
- Problem : aligning generative models with student preferences in federated environments.
- High-level contributions and project scope.

## 2. Related Work

- Federated RLHF and educational recommendation systems.
- Overview of DPO, PPO, and their limitations.
- Positioning of FedBiscuit.

## 3. FedBiscuit Architecture

- End-to-end system design : clients, central server, and training loop.
- Client-side preference optimization and server-side aggregation.
- Advantages and current limitations of the approach.

## 4. Synthetic Dataset : Generation and Augmentation

- Prompt design using a knowledge graph for structured scenario generation.
- Integration of learning styles and diverse student profiles.
- Heuristic construction of preference pairs for DPO training.
- Implementation of a new dataloader in FedBiscuit.

## 4bis. New Evaluation Component

- PPE Benchmark ?

## 5. Experiment

- TL ;DR summarization task as a reference benchmark.
- Task : Assess the training dataset size required.
- Simulation design : client count, model size, number of rounds, hardware specs.
- Implementation details.

## **6. Results**

- Quantitative results : accuracy, alignment score, and system performance.
- Visualization with graphs and tables.
- Qualitative analysis and interpretation.

## **7. Discussion**

- Lessons learned and what worked well.
- Challenges in federated preference alignment (data distribution, model variance).
- Ideas for extending the work : personalization, real-world data, privacy constraints.

## **8. Conclusion and Future Work**

- Recap of objectives and core contributions.
- Next steps : Deployment with real student data and system improvements.