Intelligent Interactive Systems – Final Design Project Instructions (2025–26)

General Instructions for the Project

In this project, you'll create an original **human–AI interaction system** in teams of two. You can pick any kind of system, e.g., a recommender system, creativity support system, or decision-support tool.

Graduate students: If you'd like to do a research project instead, see instructions [here](#).

Your project will include several milestones:

1. Project proposal

2. Prototype

3. Final presentation (demo)

4. Final report + evaluation

High-level grading expectations for final project:

Since you are allowed to use LLMs, we expect very high-quality projects! (A fully LLM-generated project provides a baseline for grading)

96–100 – Outstanding: Original, theory-grounded idea with exceptional implementation or design insight; goes well beyond course expectations.
90–95 – Excellent: Strong, well-executed concept with solid theoretical grounding and polished prototype; minor issues only.
85–89 – Very good: Clear, coherent, and technically sound project showing good grasp of principles but lacking top-tier originality or depth.
80–84 – Good: Competent and functional work that meets requirements but is conventional or uneven in quality.
70–79 – Adequate: Working system with notable flaws in concept, design, or evaluation; limited theoretical connection.
60–69 – Major issues: Partially functional or conceptually weak, relying heavily on generic outputs or showing gaps in understanding.

55–59 – Minimum standard: Barely meets requirements; little originality or human contribution—an LLM could have produced it.

Policy on Large Language Models (LLMs)

LLMs (e.g., ChatGPT, Claude, Copilot) may be used in this course **with restrictions**:

- **Allowed use:** brainstorming ideas, debugging, generating/revising code snippets, exploring interface designs, or as part of your system.
- **Not allowed:** using an LLM to write your **reflections** or other personal evaluation components.

Guidelines for Project Ideas

Projects are open-ended, but must satisfy:

- A **real AI technique** (algorithmic or LLM-based) must run in the background.

- A **user interface** must exist (e.g., simple web tool, GitHub Pages, Streamlit, etc.).

- The system design should be **theory-grounded** (identify relevant theories specific to your application that can inform the design of the system)

- The prototype should focus on the **core human–AI interaction concept**, not full infrastructure (e.g., login, database).

Examples of high-quality past projects (demo videos):

Note - this year we added the theoretical grounding of your app, which is not seen in these examples

- "[The future of typing](#)" - an LLM-based interface for augmented reality that does not require a mouse (using hand gestures instead) or a keyboard (text suggestions by LLMs)
- "[CustomTales](#)"/"[What story would you like to hear today](#)?" - LLM-based systems for generating children's stories
- "[MyMoir](#)" - an app for generating videos summarizing a trip/experience

Submission instructions for all components:
- One team member submits each milestone
- Documents should be written in English, Arial font, size 11, 1.5 spacing
- Note the length expectations described below for each component of the project
- Code should be documented and accompanied by a README file in a GitHub repository.

Milestone 1: Project Proposal (10%)

Deadline: 9/12/25

Your first task is to identify a problem and brainstorm solutions. Your proposal (up to 2 pages) should include:

- **Problem & motivation** – what problem are you tackling? Why is it interesting/important?

- **System objectives** – at least 2 preliminary objectives your system should support.

- **Existing technology and theoretical review** – what tools exist, how yours differ, inspirations, theories you might draw on.

- **Approach (Intelligence Design)** – Describe how "intelligence" will appear in your system. This can involve:

  - Algorithmic components you might implement (e.g., clustering, search, and personalization).

  - LLM-based components, including prompting strategies, orchestration, or hybrid designs.

  - How the system's intelligence will enable meaningful human interaction.

- **Plan** – major tasks and timeline.

Grading criteria:
- Problem & motivation (25%)
- System objectives (20%)
- Existing technology review (15%)
- Approach: Intelligence design (25%)
- Innovativeness (5%)
- Plan (10%)

Milestone 2: Initial Mockup(15%)

Deadline: 23/12/25

Submit an initial design mockup and describe possible backend approaches.

Deliverables:

1. **Project summary** (problem, solution, unique approach) - essentially the same as in the proposal, but should be revised based on updates to your project. (~0.5 page)
2. **Design principles** – guided by at least 3 HAI design guidelines (e.g., [PAIR Guidebook](#)). Use an LLM of your choice to brainstorm design ideas for the guidelines (~0.5 page)
3. **Interface design and instruction** – screenshots/mockups of at least 3 user tasks + video walkthrough (up to 3 minutes).
4. **Theoretical basis -** describe the theory or theories (support with academic references) that inform the design of your system (~0.5 page)
5. **Intelligence design approach** – describe how the "intelligence" of your system will work (~0.5 page)

    ○ If algorithmic: approaches, libraries, or tools you might implement.

    ○ If LLM-based: prompting strategies, role configurations, evaluation/guardrails, or orchestration logic.

Grading rubric:

- Project summary (10%)
- Design principles (20%)
- Interface instruction (20%)
- Theoretical grounding (25%)
- Intelligent design approach (25%)

Milestone 3: Demo (10%)

Deadline: 25/1/26

Prepare a **3-minute demo video** showcasing your project.

In addition to the demo video, there will be an in-class session with mandatory attendance in the last week of the semester (all team members must attend to receive the grade for the demo).

Milestone 4: Final Prototype + Evaluation (65%)

Deadline: 29/1/26

Submit a functional prototype and final report.

Final Report Sections:

1. **Project Summary** (problem, solution, unique approach). (up to 0.5 page)

2. **System Description – Link to Your Working Prototype**. (up to 2.5 pages)

   ○ **Interface** – describe core components, supported tasks, and how at least 3 HAI guidelines are addressed. Describe changes you made to the prototype from the prior submission based on the feedback from us/users/LLM
   ○ **Theory** - briefly describe the concrete theory(ies) you built on and how they informed your design
   ○ **Intelligent design** – describe how the "intelligence" of your system works. This may include algorithmic methods, LLM prompting/orchestration, or hybrid approaches.

3. **User evaluation (human users)** – at least 5 people test your system; describe tasks, feedback, and lessons learned. up to 1 page)

4. **LLM-as-User Evaluation**

   ○ Think of a way you could use an LLM to provide feedback on your prototype - the feedback should be about the design approach in general and not low-level aspects such as colors etc. Describe the strategies you've tried (~0.5 page)
   ○ Submit the **prompts and raw responses** as an appendix.
   ○ In the report, write a **critique**: Were the LLM's insights useful, biased, or unrealistic? Compare briefly with human users' feedback. (~0.5 page)

5. **Reflection on LLM use** – write (in your own words) about how you used LLMs during the project. Discuss what helped, what hindered, and what you learned. **LLMs may not be used to generate this section.**

6. **Summary** – describe what you learned overall, and what you would extend with more time.

Grading rubric (see also high-level expectations at the top of the document):

- Project summary (5%)
- Interface (12%)
- Theory (11%)
- Intelligence design (12%)

- Prototype functionality & documentation (15%)
- User evaluation (15%)
- LLM-as-user evaluation (10%)
- Reflection on LLM use (5%)
- Summary & future directions (5%)
- Innovativeness/effort (10%)