U.PORTO FACULDADE DE CIÊNCIAS UNIVERSIDADE DO PORTO U.PORTO FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO	**Master Degree in Artificial Intelligence** **Artificial Intelligence and Society**	**2024/2025** **1st Year** **1st Semester**
INSTRUCTOR: Miriam Seoane Santos		

Group Project

Deus Ex Machina: Power, Perception, and Prejudice

Theme

This project aims to explore the societal implications of Artificial Intelligence (AI) through the analysis of a real-world case study. Each group will choose or propose a significant case study where AI has impacted society and thoroughly analyze the case to identify and explain its main challenges, outcomes, and ethical dilemmas. During the analysis of the case study, one or more relevant topics should be identified (e.g., data quality, bias, fairness, explainability, transparency, privacy, etc.), and a technical experimentation on a preferred topic should be conducted to deepen the understanding of the concepts. **The deliverables of this project are a comprehensive report and a technical tutorial, which students will discuss in a final presentation.**

Project Phases and Deliverables

A. Project Planning [Checkpoint: 04/10]

You should start by finding or selecting a documented case study where AI has impacted society. You may choose one of the topics suggested below, provided that **there are no repeated topics among groups**. In case you decide to propose your own case study, make sure that it contains sufficient information to allow for a meaningful analysis and validate it with the teacher before you start.

- **Plan and schedule your project organization:** outline the main tasks required for each phase of the project, define tentative deadlines, and assign specific roles and responsabilities to each team member. Make an effort to *equally split the project load* among all members. Produce a *Gant* chart (or similar alternatives) to showcase your project organization. Identify the tools, and resources/bibliography (data, code, news, websites, blogs, research papers) needed for the project. Establish how the team will communicate and collaborate (e.g., shared documents, collaborative coding, etc.).

This should be <u>delivered</u> in class on Oct 4. Students should submit a PDF report <u>up to 2 pages</u> containing the following:

- Title of the chosen Case Study;
- Description of the Case Study;
- Team Members;
- Team Leader;
- Roles and Responsibilities of each team member;
- Draft of Project Organization: main task definition, responsible members, and tentative deadlines;
- Gant chart;
- Tools: Code Collaboration, Communication, Sharing Documents, Project Management
- Resources: data, code, videos, blog, open-source software, other tools;
- Bibliography: books, news, reports, documents, research papers

B. Case Study Analysis [Checkpoint: 25/10]

This task involves a comprehensive description of the chosen case study.

- Describe the context in which the AI was used and identify the issue that emerged, reflecting on AI's role and societal impact. Specify involved parties (institutions, stakeholders, companies, etc.).
- Discuss the technicalities of how the AI was developed and/or deployed, including relevant factors whenever you can (e.g., dataset quality, algorithmic choices, outcomes, examples).
- Focus on the outcomes of the AI application (positive or negative). Did the AI system improve efficiency or decision-making? Did it discriminate against certain groups? Was the logic behind the AI's decisions clear to its users? Was anyone accountable when (and if) the AI failed?
- Comment on the societal response (if there was any) or provide your own conclusions on the topic. How did society react? Were there legal consequences, public outrage, coverage on the news, political changes? What did the media report on the issue, how did stakeholders (government, public institutions, companies) react?

This should be <u>discussed</u> in class on Oct 25. For the final project, the following should be included:

- Context: an in-depth description of the case study;
- AI Systems details: how it was built, training, which technology was used (e.g., machine learning algorithms, etc.);
- Positive and/or Negative Outcomes;
- Societal Response.

C. Technical Component [Checkpoint: 15/11]

This task involves selecting a topic associated to your chosen case study (e.g., bias, fairness, explainability, accountability) and producing a technical tutorial on that topic. Read academic papers, articles, and reports to learn more about the topic and how it relates to your case study. Then, produce a hands-on tutorial that is reproducible. This technical component is open-ended and can comprise different approaches. Some examples are:

1. Producing a comprehensive analysis on a real-word dataset relevant to the selected topic;
2. Reproducing the techniques described on a research paper and validating its results;
3. Proposing and implementing new algorithms or machine learning approaches;
4. Testing and comparing different packages related to the topic;
5. (…)

This should be <u>discussed</u> in class on Nov 15. For the final project, the following should be included:

- A jupyter notebook (.ipynb) that explains your technical tutorial step-by-step. Make sure to include installation requirements if they exist (e.g., Python version, required packages and versions**). Your tutorial should be able to be reproduced and presented in class.**
- The notebook should include only a clear and easy to follow storyline, code snippets, and respective results (e.g., charts and a brief discussion of the main findings).The final report should include additional details about your dataset, used tools, methodology, and/or proposed approach, and a more in-depth discussion of the findings, when applicable.

D. Final Presentation [Checkpoint: 29/11]

The final presentations of the project will occur on **Dec 6** and **Dec 13** (groups will be randomly selected, unless a consensus is established among Team Leaders). In the final presentation, **each group will have 25 minutes** at most to present their project. This involves the following:

- Presenting the chosen Case Study, according to the guidelines in **Section B**.
- Presenting the technical tutorial, according to the guidelines in **Section C.** The tutorial is to be replicated by the peers, so **it must be available online at the time of the presentation** (e.g., using GitHub, Google Colab, or other alternatives).
- **The group has complete freedom on how to present the final project** (e.g., using a website, playing a video, producing a poster). It is not mandatory to deliver any final presentation (e.g., Powerpoint). The only deliverables are the PDF report and technical tutorial.

All groups must attend the final presentations and take notes. Both the technical tutorial and the reports will be available to all students, and they are considered material subjected to Test (40%).

Final Delivery [Dec 8 (and Dec 22)]

By **Dec 8**, each group must submit in Moodle the following deliverables:

- **A PDF report (up to 8 pages using the specified templates).**
- **The technical tutorial**, submitted as a Jupyter Notebook.

Students will be allowed to improve their work, taking into account the instructor and peers feedback by submitting a revised version of their material by Dec 22, on Moodle.

Templates

Students should follow a double-column article format for the writing of the final reports:

- Template 1
- Template 2

Deadline Overview

- **October 4: (In-class + Moodle)** Checkpoint**.** Deliver a 2-page PDF with information on Case Study and Project Planning;
- **October 25: (In-class)** Checkpoint. Feedback of the Case Study;
- **November 15: (In-class)** Checkpoint. Feedback of the technical tutorial;
- **November 29: (In-class)** Checkpoint. Recommendations for the final presentation and reports;
- **December 6: (In-class)** Project Presentations (Part 1 – 5/6 groups);
- **December 8: (Moodle)** Deliver a 8-page PDF report and technical tutorials;
- **December 13: (In-class)** Project Presentations (Part 2 – 5/6 groups);
- **December 22 (Optional):** Final Delivery (*after presentation feedback*)

Evaluation

The project will be evaluated regarding the following expected outcomes:

- **Checkpoints (1 point):** Student involvement, progress, plan, and management of the project's timelines.
- **Final Report (2 points)**: Quality of the exploration of the case study, the produced insights, critical analysis, and conclusions.
- **Technical Tutorial (2 points)**: Quality of the technical implementation and documentation.
- **Presentation and Demo Quality (1 point)**: Quality of presentation delivered by the students.

All students should participate in the checkpoints and the final presentation, as they are graded according to each member's contribution.

Suggested Case Studies

CS1: COMPAS Machine Bias: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

CS2: Gender Shades: https://www.media.mit.edu/projects/gender-shades/overview/

CS3: Inside the Suspicious Machine: https://www.wired.com/story/welfare-algorithms-discrimination/

CS4: (De)Generative AI: https://www.bloomberg.com/graphics/2023-generative-ai-bias/

CS5: Netflix N'(o) Chill:

- https://knowledge.wharton.upenn.edu/article/data-privacy-real-dont-bet/
- https://arxiv.org/pdf/cs/0610105
- https://arxivblog.com/?p=142
- https://www.cs.utexas.edu/~shmat/netflix-faq.html

CS6: You Shall not Pass: https://gandalf.lakera.ai/baseline

CS7: GPT, the PhD:

- https://hdsr.mitpress.mit.edu/pub/pqiufdew/release/2
- https://github.com/The-CEAS-Library/A-Practical-Guide-to-Learning-to-Code-with-ChatGPT/blob/main/A_Practical_Guide_to_Learning_to_Code_with_ChatGPT.ipynb

CS8: Minority Report: https://themarkup.org/prediction-bias/2021/12/02/crime-prediction-software-promised-to-be-free-of-biases-new-data-shows-it-perpetuates-them

CS9: Premium Prejudice: https://themarkup.org/premium-penalty/2024/07/18/michigans-fair-and-reasonable-reforms-allowed-car-insurers-to-charge-more-in-black-neighborhoods

CS10: Unmasking Dropout: https://themarkup.org/machine-learning/2021/05/26/nycs-school-algorithms-cement-segregation-this-data-shows-how

CS11: Privacy Inspector: https://themarkup.org/blacklight/2020/09/22/blacklight-tracking-advertisers-digital-privacy-sensitive-websites

CS12: Home? Alone!: https://themarkup.org/investigation/2023/02/28/l-a-s-scoring-system-for-subsidized-housing-gives-black-and-latino-people-experiencing-homelessness-lower-priority-scores

DYI Case Study

Alternatively, **students are welcome to propose their own case study,** namely by visiting the AI Incident Report Database, MIT Technology Review, The Markup "Show Your Work", "Machine Learning" , or "Pixel Hunt" walls, and selecting one case study of their interest. There are also interesting GitHub repositories that might provide some inspiration: Failed Machine Learning and Awful AI.

Important guidelines:

- The case study should be a **relevant concern of AI in Society** and be **related with the topics discussed in the course** (data quality, bias/fairness, privacy, transparency, accountability, etc.).
- The case studies should **provide enough information for a comprehensive analysis**. This may include access to case details, used data, research papers, or reports (e.g., "one-shot" news in the media will be hard to map into a complete case study);
- If the case study is **not convincing enough** or it has **insufficient material**, it will not be accepted.