Getting a lay of the land: Insights from compiling AI Ethics courses across Higher Education in the EU

Introduction

The responsible development of Artificial Intelligence is a clear goal of the European Union, spearheaded by the High-Level Expert Group's Guidelines (HLEG) on Trustworthy AI (EUCommission, 2018) which outline the necessary requirements for trustworthy AI development. This goal is underlined in the EU's digital strategy, which emphasises the need to train professionals that can "shape technology in a way that respects European values" (EU, 2020). The need for education into the technical, legal, ethical and economic elements of AI is not only relegated to developers: a digitally literate population is essential to enable social debate on the public policies and legal frameworks that are to govern the use of AI. Incorporating all aspects of Trustworthy AI in Higher Education is therefore becoming a necessity to train future members of the artificial intelligence community as a whole -developers, deployers, policymakers, users- and, more widely, informed citizens.

To this end, calls have arisen for the inclusion of artificial intelligence ethics into universities' curricula, taking into account matters that are not exclusively technical (Raji et al., 2021). A key consideration is the importance of having interdisciplinary teams involved in AI ethics education and governance: the challenges emerging in relation to AI cross over disciplinary lines and are often too complex for any single type of expertise to handle. Insights from lawyers, sociologists, policy scholars, philosophers, together with scientists and engineers are all necessary for a rounded perspective on the ethical, legal, social and economic implications of AI. This not only means that educators of different backgrounds should be involved, but also that AI ethics education should not be relegated to STEM degrees but rather should be included across disciplines.

Given these acknowledged needs, a key question arises: to which extent is AI ethics education currently included in Higher Education in the EU, across disciplines? Indeed, getting a *lay of the land* on the state of AI ethics education is a first step to understanding *how*, *where* and *to whom* AI ethics is currently taught, as well as in identifying current gaps and needs. To this aim, the AI4EU Working Group on AI Ethics Education has undertaken the goal of compiling a list of the Higher Education courses within the EU that include education on AI ethics and providing an accessible repository with this information. In this work we detail the challenges faced in the building of this repository, as well as the insights and gaps that we have preliminarily identified.

Defining and finding AI Ethics courses

The initial stages of creating the course repository revolved around a single question: what does it mean for a course to be about, or to include, AI Ethics? Indeed, wide definitions of what constitutes AI and what ethics encompasses make a simple definition hard to reach. For this reason, the Working Group decided to anchor the idea of AI ethics in the requirements set

down by the High-Level Expert Group's Guidelines on Trustworthy AI (EUCommission, 2018). As such, the decision was made to include any course that considers at least one topic from the seven requirements: Human agency and oversight, Technical robustness and safety, Accountability, Privacy and data governance, Transparency, Societal and environmental well-being, Diversity, non-discrimination and fairness. To warrant inclusion, these topics should be approached with AI applications as a focus or an example, where AI is construed broadly (including data, learning, reasoning and autonomous applications).

Once this definition was established, the work was divided between the members of the Working Group to collect AI ethics courses by country. In the collection effort, there were two main challenges encountered: the difficulty of accessing detailed course syllabi, and the lack of explicitness connecting course contents with the HLEG Requirements. When it comes to course syllabi, the information made available openly often consists on course objectives, whereas detailed content covered in each course is not necessarily made available to non-enrolled students. This constitutes an obstacle in determining whether any Requirements are tackled in a course, as well as understanding whether they are covered with AI applications in mind. Additionally, although topics related to Requirements are often covered (e.g. good data practices, debiasing, auditing), they are not explicitly related to the concepts of AI ethics or trustworthy AI in the syllabus.

Overall, the collection effort has yielded a list of 111 courses across 19 countries. A repository of such courses will be made available at the AI4EU observatory[1], enabling access for both students and researchers. The repository will indicate, among others features, country, university/degree to which they belong, language and official links. In addition, it will be possible to request a course to be added to the repository: this feature is fundamental to alleviate the shortcomings of the Working Group's web search, where courses whose syllabus was not made available online can be added on request if relevant. If you want to contribute to the repository by suggesting additional material, a public form is available.

Findings and discussion

Overall, we have found that AI ethics is a growing topic in education across countries. All across disciplines and languages, concepts and methodologies to tackle the HLEG's Requirements are taught in both undergraduate and postgraduate programs. Although the topics are covered, an important shortcoming is the lack of relating specific course content to the HLEG Guidelines or to Trustworthy AI. This lack of explicitness not only constitutes a hurdle to searching for courses, but can also prevent students from understanding that they are in fact being taught about Trustworthy AI.

Overall, the final repository of courses constitutes a first step in understanding the state of the art in AI ethics education in the EU, which can inform on policy initiatives destined to boost

[1] A new section with the education repository will be added at https://www.ai4europe.eu/ethics

Trustworthy AI education. Although drafting a comprehensive resource has proved challenging, we hope that the publication of this repository will encourage higher education institutions to add their missing courses to it, as well as being more explicit on how AI ethics is included in their education.

References

EU. (2020). WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust).

EUCommission. (2018). ETHICS GUIDELINES FOR TRUSTWORTHY AI. *High-Level Expert Group on Artificial Intelligence*, (December). Retrieved from https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

Raji, I. D., Scheuerman, M. K., & Amironesei, R. (2021). "you can't sit with us": Exclusionary pedagogy in AI ethics education. In *FAccT 2021 - Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 515–525). New York, NY, USA: ACM. https://doi.org/10.1145/3442188.3445914