Values in Education: Exploration of Artificial Intelligence Ethics Syllabi Using Natural Language Processing Analyses

Kerrie Hooper*, Stephanie J. Lunn†
Florida International University
Miami, FL 33199, USA
Email: *khooper@fiu.edu, †sjlunn@fiu.edu

***Abstract*—With new technologies come additional responsibilities. Examining Artificial Intelligence (AI) through an ethical lens has become increasingly important and significant. Advancements in AI have led numerous organizations, such as IEEE, to develop AI ethics guidelines for consideration in academia and industry. Additionally, higher education has an essential role in fostering innovation and developing skilled professionals who will work on topics that span social, philosophical, scientific, and technical spheres. To assess the content being covered in tertiary classrooms, we utilized a Natural Language Processing (NLP) approach for analysis. This study examines (n = 45) AI ethics syllabi that were publicly available online. The course description, topics, department, and year were some important features captured from each syllabus. Using various NLP tools for analysis, a general exploration of AI ethics curricula was conducted. Through supervised clustering, k-means clustering, and latent Dirichlet Allocation (LDA), various patterns in the contents of the AI ethics syllabus were found. Some of these include trends and patterns from syllabi across various academic departments, years, and the pre-post Chat-GPT era. Cluster evaluation was also done on the unsupervised clusters using various metrics to determine the viability of the clusters. The LDA analysis enabled a review of topics that are consistent among the clusters, which helped highlight salient areas of focus in AI ethics syllabi. The findings from this study can serve to inform administrators and educators, acting as a baseline for including language around AI ethics topics and uncovering potential topical gaps in the contents of AI ethics syllabi. They can also provide insight into how different academic departments, like computer science and philosophy, may approach the topic. Such understanding is critical to ensuring the next generation of graduates not only considers how to utilize AI but also promotes doing so responsibly and with regard to its societal implications**

***Index Terms*—artificial intelligence, AI, ethics, natural language processing, clustering**

I. Introduction

Artificial Intelligence (AI) refers to the use of machines to perform tasks typically requiring human intelligence, such as pattern recognition and problem-solving [1]. It offers opportunities for technological development and applications across various domains (e.g., finance, law, education) and in society more broadly [2], [3]. Under the AI umbrella is **generative AI**, which describes computational approaches to the development of new material using technology, like text, images, or audio [4]. This includes Large-Language Models (LLMs), such as OpenAI's ChatGPT [5].

The pervasive integration and growth of AI and its capabilities have impacted varying aspects of people's lives. It has led to the emergence of new research, education, and job positions in the space, as well as novel use cases within existing roles [6]. As a result, scholars have highlighted the need to recognize the challenges and risks associated with such tools and models and to examine AI through an ethical lens [5]. These changes have also led job requirements and governing policies to adapt accordingly [7]–[9]. Efforts towards regulation and oversight may be undertaken locally by smaller agencies or at the national level [10], [11].

The Institute of Electrical and Electronics Engineers (IEEE) is self-described as the "world's largest technical professional organization dedicated to advancing technology for the benefit of humanity" [12]. IEEE, as well as many other similar organizations, have put together various guidelines and standards that give rise to suggestions of what ethical designs may look like around AI and autonomous systems [13]. IEEE published eight ethically designed (EAD) general principles, as illustrated in Figure 1. These standards can guide and govern practices of ethical AI.

Fig. 1: IEEE EAD principles, adopted from [13]

Academia, in particular, is crucial for driving innovation and producing talented individuals who serve in both the technical and ethical spaces [14]. It strives to produce responsible developers, innovators, critical thinkers, business leaders, and policymakers well-equipped with a wide array of knowledge based in multiple disciplinary domains. As such, it is fitting for post-secondary institutions to prepare students to be culturally responsive, have a collaborative mindset with interdisciplinary skills, and be aware of equity issues and other social problems in society. Much onus is on computing education (especially since AI is categorized as a computing field [15]) to prepare students so that they are instilled with the balance of technological innovation and ethical concerns. Therefore, it is suggested that students in computer science (CS) and other related computing fields should be trained to communicate across disciplines and appreciate non-technical concerns that may be posed from a social science context [16].

Stanford University Cyberlaw recently discussed some actions and/or roles universities can play in supporting AI ethics [17]. These included: 1) Identify, frame, and inform key issues such as highlighting questions, answers, and voices often excluded; 2) Develop and communicate an affirmative vision for ethical AI, which entails embracing opportunities to mitigate risks and foster interdisciplinary collaborations; 3) Model or enact the vision through the work of the university which allows institutions to fully embrace the responsible use of AI in teaching and research. As such, it can be valuable to examine ways in which universities can contribute to AI ethics research and development. As a first step, evaluating curricula on AI ethics can provide context and language around how academia may approach AI ethics in pedagogy. This can serve to understand patterns towards topics addressed as well as identify potential gaps in AI ethics or areas that may need improvement or expansion.

This study sought to understand how educators may approach AI Ethics topics through their syllabi. In particular, we aimed to address the following research questions (RQs):

- **RQ1:** *What topics do AI Ethics syllabi address?*
- **RQ2:** *How well do AI Ethics syllabi topics align with the eight IEEE standards for Ethical Design?*
- **RQ3:** *How has the emergence of ChatGPT and other Large-Language Models (LLMs) impacted content in AI Ethics syllabi?*

To answer these questions, we employed Natural Language Processing (NLP) clustering algorithms and other NLP techniques. In the work that follows, we first describe the Related work in (Section II). Then, in the methods (Section III), we described the steps we took in developing the dataset as well as the rationale for the various NLP techniques we utilized. Next, we demonstrated our results (Section IV) using charts and figures and discussed the patterns and findings (Section V), limitations (Section VI), and conclusion (Section VII).

II. Related Work

A. Teaching AI and Ethics

The expansion of AI has resulted in the need for teaching and learning around its development and applications, as well as addressing the effects of such tools [18]. A joint task force of the Association for Computing Machinery (ACM), an international organization centered around computing, and IEEE worked together to define a guiding Software Engineering Code of Ethics [19]. This code offers a useful starting foundation for approaching ethical topics and "responsible innovation" [20].

AI can impact academia in areas ranging from help with academic writing for students to potential opportunities for evaluation and assessment for instructors [21], [22]. In practice, educators have applied different approaches to incorporate AI into lessons and to address its potential usage [23], [24]. As an example, it has been applied in the context of science fiction as a way of encouraging students to consider its current state and possible ramifications on society [25]. It has also been used to analyze curricula, which can have implications at the classroom and administrative levels at universities. This is useful as there can be misalignments between learning outcomes and teaching methods, which may result in the need to reevaluate teaching practices and develop more integrated interdisciplinary approaches to AI ethics education [26]. Discovering patterns through syllabi analysis also revealed that ethics can be incorporated into technical CS courses and not limited to non-technical courses [27]. Dialogue approaches and case studies have been suggested as additional ways to address such material in classrooms. Additionally, sociotechnical models could enhance students' abilities to address ethical issues when learning AI ethics in the classroom [28]. For instance, a combination of Saltz et al.'s framework [29], which integrates ethics into technical modules but lacks real-world context, and Krakowski et al.'s framework [30], which prompts broader ethical questions but is less applicable to university courses, could be applied to improve AI ethics curricula [28].

B. Generative AI and Ethics

Given that generative AI, including LLMs, are based on the data fed, it is important to consider not only their accuracy but also how items produced can be used responsibly and safely, as well as how they may perpetuate bias [31]. There is a need to implement robust ethical frameworks in higher education, as the advent of LLMs can raise concerns about preserving academic integrity, intellectual property rights, and plagiarism [32]. As a result, many scholars call for more transparency on various ethical issues that can undermine academic quality and diversity [32]. Academia continues to work towards this goal, analyzing AI's current status through syllabi, policies, research, and practice to develop innovative ways that ensure academic trustworthiness and that support ethical guidelines and frameworks. In the work that follows, we reconcile these concerns through analyses of AI ethics syllabi.

III. Method

Our methods are presented in two main subsections: Dataset Description and NLP Methods. In Dataset Description (Section III-A), we elaborate on our approach to search for AI ethics syllabi; then, we describe the inclusion and exclusion criteria and present our dataset. In the NLP Methods subsection (Section III-B), we elucidate our data preprocessing steps and text embedding tool, and lastly, we describe the specific algorithm analysis methods that help us to answer our RQs.

A. Dataset Description

1) Search Terms: Since we aimed to gather data from publicly available syllabi, we used the Google database and relevant keywords to build our dataset. The search terms applied to identify relevant AI ethics syllabi were: "AI Ethics Syllabus," "AI Ethics Curriculum," and "AI Ethics Course." Results from the first 15 pages of each search term were included because after testing the search terms and results returned, we observed that the results were not rendered relevant to our study in a pilot. Moreover, additional syllabi related to AI ethics were collected from the open source "Tech Ethics Spreadsheet" gathered by Dr. Casey Fiesler [33].

Each syllabus found was collected and saved in PDF format. Then, the variables of importance to this study were collected and organized in a comma-separated value (CSV) format. Each row represented a unique syllabus, and each column tracked relevant features and contents.

2) Inclusion and Exclusion Criteria: The primary inclusion criteria were that the syllabus must cover ethical dimensions of AI in the course description, learning objectives, and/or topics described. The exclusion criteria were that the syllabus must not: 1) be from a Massive Open Online Course (MOOC); or 2) have missing data (e.g., the university name, the department offering the course, and/or the year offered).

3) Dataset: The final dataset included a total of *n = 45 syllabi*, each from a different course. The data collection process involved retaining the most recent syllabi in cases where multiple versions existed, and a CSV file was manually created and organized with data from 45 syllabi. To handle missing values, corresponding topics replaced any missing course descriptions. We want to note that some universities, like Massachusetts of Technology, University of Pennsylvania, and DePaul University, had more than one course in AI ethics, which were also included in the dataset. Although the dataset gathered more than 20 features for each syllabus, some of which are course title, course code, United States of America (USA)/International, and graduate-level/undergraduate-level, we focus on the following features to answer our RQs: Department, Year, Public/Private, Course Description, and Topics.

4) Data Description: The syllabi collected ranged across seven departments, with the majority of syllabi from the Computer Science department; see Table I. The years of the syllabi range from 2014 to 2023. After 2022 was considered the post-ChatGPT era since the tool gained traction in late November 2022 [34]. Table II illustrates the total number of syllabi obtained pre and post-ChatGPT. The ratio of syllabi from private and public universities was 46% to 53%, respectively, and the ratio of syllabi from USA and international universities was 71% to 29%, respectively.

Department	**Total**
Media and Design	2
Business and Law	3
Philosophy	14
Computer Science	22
Public Policy	2
Political Science	1
Religion and Gender Studies	1

TABLE I: Total number of syllabi from each department in the dataset

Year	**Total**
Before ChatGPT (2014 - 2022)	35
After ChatGPT (2023 -)	6
No year stated	4

TABLE II: Total syllabi from the pre-and post-ChatGPT era

B. NLP Methods

1) Data Preprocessing: The preprocessor package in Python (tweet-preprocessor) was used to conduct text preprocessing; it took care of tokenization, parsing, and the elimination of reserved words and URLs. Additionally, English stopwords were eliminated using the Natural Language Toolkit package, such as "AI" and "course." The preprocessed text data was then ready for embedding.

2) Text Embedding: Two text embedding techniques were examined to determine which one was optimal. Specifically, we compared SentenceBERT (S-BERT) against Term Frequency Inverse Document Frequency (TF-IDF). S-BERT is a modern technology, while TF-IDF is domain-specific within the context of the document or data. S-BERT works by training neural networks to convert input sentences into high-dimensional vectors in such a way that similar sentences have closer vector representations, facilitating semantic similarity comparisons in vector space [35]. Conversely, TF-IDF works by assigning numerical weights to words in a document based on their frequency within that document and inversely proportional to their frequency across a collection of documents, aiming to highlight the importance of terms in capturing the content-specific relevance of words. Choosing the better-performing embedding is a good practice and has been described as facilitating credibility and more accurate results in NLP analysis.

3) Similarity Correlations: Course descriptions and course topics were two data points that could answer our RQs. In an effort to reduce the complexity of our analysis, we elected to choose one of those data points. Based on the similarity correlation comparison, there were greater similarities or overlaps among topics than in course descriptions. Therefore, data was grouped by topics to uncover uniqueness. Going forward, the analysis is focused on course topics unless otherwise specified (see Figure 2).

4) Methods: To answer our first research question, three kinds of clustering techniques were used. They are k-means clustering, Latent Dirichlet Allocation (LDA) analysis, and supervised clustering by various categories of the data. K-means clustering is an unsupervised clustering method that was used to find patterns across the syllabi topics. A key component of k-means clustering is to determine the best value for k. To do so, values from k = [2 - 10] were tested, and the silhouette score metric was used to determine the best k.

Silhouette score measures how similar the syllabi are in their cluster (cohesion) compared to other clusters (separation). A score closer to 1 indicates that the clustering is good, while a score closer to -1 indicates a bad cluster. As demonstrated in Table III, k = 2 has the best Silhouette Score.

Cluster	Silhouette Score
2	0.14
3	0.11
4	0.08
5	0.06
6	0.05
7	0.06
8	0.04
9	0.05
10	0.04

TABLE III: Determination of k based on Silhouette Score

We then compared the S-BERT and TF-IDF embeddings (Section III-B2) on k = 2 to determine which embedding method would serve to be most effective in this case. Based on the silhouette scores and inter and intra-cluster distances, the S-BERT model performed better. This implied that S-BERT embeddings display more semantic strength because, with S-BERT, the silhouette scores were workable compared to the TF-IDF. Therefore, moving forward, the S-BERT embeddings were used in this study. See Tables IV and V.

Metric	Cluster 1	Cluster 2
Silhouette Score	0.00	0.04
Intra-cluster Distance	1.28	1.23
Inter-cluster Distance	1.07	1.07
Cluster Size	22	23

TABLE IV: TF-IDF Cluster Performance

Metric	Cluster 1	Cluster 2
Silhouette Score	0.16	0.13
Intra-cluster Distance	11.35	10.54
Inter-cluster Distance	8.72	8.72
Cluster Size	37	8

TABLE V: S-BERT Cluster Performance

Next, we utilized LDA (Latent Dirichlet Allocation), which is a probabilistic model used to determine the top words covered across a collection of documents. It assumes that each word is attributed to a document topic and that a document contains a mixture of topics. LDA is used here to look for patterns across clusters to highlight the most frequent topics in the dataset. The Python LDAVis Library was used to conduct this analysis. We also made use of supervised clustering techniques to cluster topics based on individual features like department or year. This approach allowed for the organization and analysis of data in a structured manner, providing insights into how departments operate across various types of institutions. In addition, we examined topics that were least covered across the dataset. To accomplish this, the first step was to devise a method for extracting unique words from the "topics" corpus that were mentioned only once. Once the unique words were obtained, the next stage involved manually scanning through the list to identify and disregard any typographical errors or related issues. This process required careful attention to detail and a systematic approach to ensure accuracy.

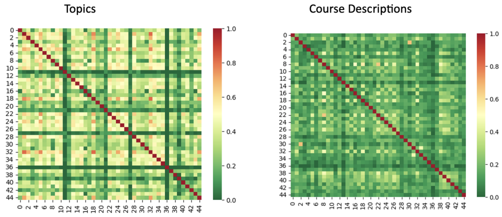


Fig. 2: Similarity correlations of topics and course descriptions of the syllabi

For our second research question centered around IEEE's EAD principles, we calculated the Similarity score for each principle against each syllabus. This method calculates the Word Similarity score by leveraging the Python SpaCy library, which employs the cosine similarity metric. First, it utilizes SpaCy to compute the similarity between words. Then, Seaborn, a visualization library, was used to create heatmaps, providing a visual representation of the similarity scores. This process offered a means of visualization to analyze and interpret the relationships between words based on their similarity scores.

The third research question was answered using the supervised clustering method as well, where 2023 and onwards are grouped as the post-chatGPT era, while syllabi before 2023 were grouped as the pre-chatGPT era. This method aided in observing how syllabi topics may have changed during the era preceding the advent and traction of ChatGPT and generative AI technologies. Then, we examined the differences between the topics covered.

IV. Results

The results are presented according to the three guiding research questions.

A. RQ1: What topics do AI Ethics syllabi address?

1) K-means Clustering: Here, we find that when clustering syllabi topics with k = 2 in a semantic manner, they tend to cluster around whether topics are technical or less technical (general). Given the textual nature of these results, they are also open to reader interpretation. [explain more and signpost]

Topics	
Cluster 1 (15 words) ***Classification: Philosophical or General***	***Cluster 2 (15 words)*** ***Classification: Technical or Specific***
moral, autonomous, technology, robot, history, work, future, superintelligence, new, safety, economic, human, politics, LLM, good	machine learning, privacy, transparency, algorithmic, ethical, autonomy, design, discrimination, accountability, society, introduction, value, issues, emerging, overview

TABLE VI: Unique Words in Clusters (k=2)

2) Supervised Clustering: Department: Supervised clustering aided in depicting the unique topics covered in each department. Table VII highlights the topical vastness of AI ethics as well as how it is distributed across multiple departments that offer an AI ethics-related course.

Media and Design	ownership, trade, agency, credit, predictive policing, liability
Business and Law	accuracy, manipulation, collection, prediction
Philosophy	surveillance, superintelligence, rights, consent
Computer Science	autonomous, accountability, application, vs, value
Public Policy	government, bureaucracy, police, tech
Political Science	health care, malicious, uses, governance
Religion and Gender Studies	race, pop culture, posthuman, nanotechnology, cyborg, critiques

TABLE VII: Unique topics across the departments.

Public and Private Universities: Table VIII shows that Public and Private universities have some unique topics. However, the silhouette scores were low (see Table IX). The low silhouette score indicates a bad cluster and that there is not much of a difference in topics covered in public and private universities.

Private	machine, algorithmic, social, design, discrimination, regulation, emerging
Public	society, accountability, human, safety, autonomous, superintelligence, theory

TABLE VIII: Unique topics from Public and Private Universities

Public/ Private	Silhouette Score	Intra-cluster Distance	Inter-cluster Distance	Cluster Size
Public	-0.0	1.19	0.64	24
Private	0.0	1.18	0.64	21

TABLE IX: Public and Private University cluster evaluation.

3) Latent Dirichlet Allocation (LDA): For this analysis we examined the top 5 words occurring frequently in each cluster. We conducted this analysis on k = 2 because it was the best k in k-means clustering. The LDA analysis helped us to observe where the topics converge across all the syllabi. We can see that Data, Bias, Privacy, and Fairness are the top converging topics across the syllabi (see Table X). This means that the majority of the syllabi focused on these topics as a priority in AI ethics.

Cluster 1	Cluster 2
Data	Bias
Bias	Data
Fairness	Privacy
Privacy	Fairness
Transparency	Emerging

TABLE X: Top five (5) converging topics across the syllabi.

4) Scarce Topics: After observing topics that converge, we decided to examine topics that were scarce or mentioned only once across the dataset. We find that the topics less covered are related to religion/ culture, controversial topics and various philosophical terms (see Table XI).

Deontological, Kantian, Ubuntu, Buddhist, Islamic, Indian, Capitalism, Consequentialism, Counterfactual, Storyteller, Anthropomorphism, Utopianism, Faith

TABLE XI: Topics mentioned only once across the syllabi dataset

B. RQ2: How well do AI Ethics syllabi topics align with the eight IEEE standards for Ethical Design?

As we study AI ethics curricula, it is vital to have a reference point by which we understand the contents and breadth of the topics covered in the courses. If we recall the eight EAD principles outlined by IEEE, these were tested against each syllabus. First, the computer science and philosophy syllabi were compared since AI is rooted in both of those fields (see Figure 3), and then all the courses were compared together (see Figure 4). A similar pattern was seen throughout computer science and philosophy courses, where there appears to be a lesser similarity score of syllabi topics covering areas of human rights, well-being, and data agency. In comparison to other courses, all courses exhibit lower similarity levels to "Human Rights," "Well-being," "Data Agency," and, to a lesser extent, "Misuse." This suggests that there are distinct differences in the content and focus of these courses compared to the overarching themes encapsulated by these key principles.

C. RQ3: How has the emergence of ChatGPT and other Large-Language Models (LLMs) impacted content in AI Ethics syllabi?

Through supervised clustering, we were able to demonstrate the uniqueness of topics covered pre-post Chat-GPT era. Post Chat-GPT era is defined as the year 2023 and onwards. As new technology develops, we can see shifts in topical foci across the syllabi. We can clearly see topics like explainability being a new area of focus in the field (see Figure XII).

Pre-ChatGPT	machine, morals, history, society, discrimination, emerging, value, autonomy, moral
Post-ChatGPT	machine-learning, analysis, explanation, decision-making, models, policy, practice, risk, LLM, explainability

TABLE XII: Unique words in syllabi pre and post-ChatGPT era.

V. Discussion

Our approach allowed us to study patterns across AI ethics syllabi through NLP techniques. Examining course "topics" as a data point demonstrated more distinctness because they involved lists of words or phrases as opposed to course descriptions, which in technical terms may contain more noise (see Figure 2). The k = 2 clustering in k-means showed some polarity in the concepts based on the sentence-transformers embeddings. Using supervised clustering, we observed differences and patterns across various departments, but there were no significant differences when considering whether institutions were public or private (see Table IX). For departments,

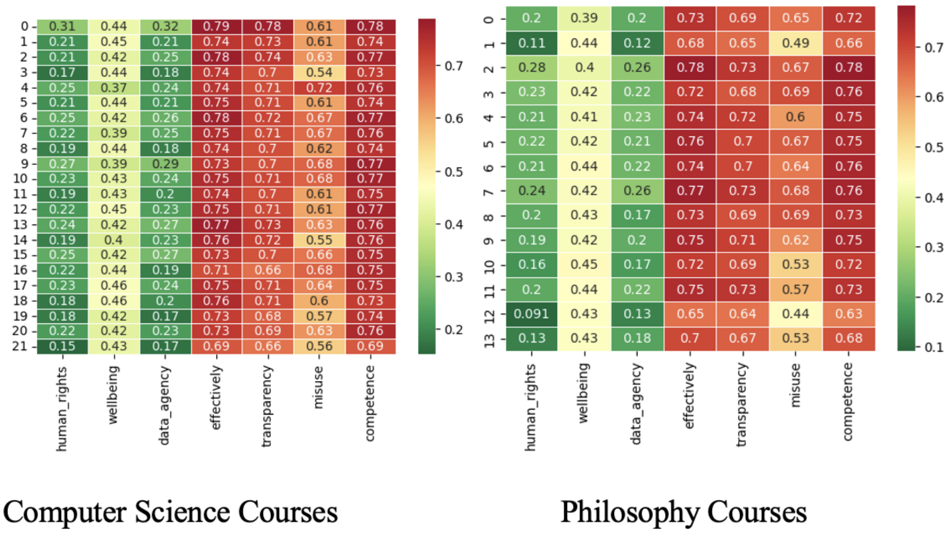
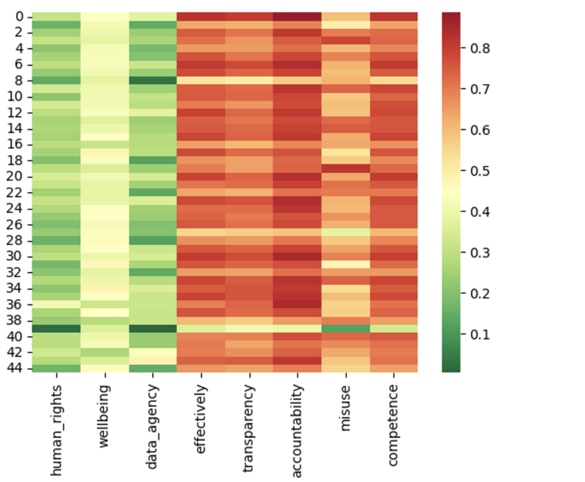


Fig. 3: Heatmap of similarity score of syllabi topics from computer science and philosophy departments with IEEE's EAD principles.



Heatmap of AI Ethics syllabi similarity to IEEE standards for ethical AI

Fig. 4: Heatmap of AI Ethics syllabi similarity to IEEE's EAD principles.

the range of unique words per cluster reinforces the interdisciplinary nature of AI ethics [36] and the need for cross-disciplinary collaboration both in academia and industry.

Through the topic modeling technique generated from the LDA analysis, we noted that privacy, bias, fairness, and data were key issues of focus across AI ethics syllabi. Our findings align with the converging topics and AI ethics principles previously described by other scholars [37]. Overall, the NLP methodology proved to be one that other researchers, administrators, and instructors can use to analyze, build, and refine syllabi.

The topics least associated with IEEE standards topics were more human-centered and less social justice-related (see Figures 3 and 4). A potential issue of concern can arise when AI discourse, such as syllabi, misconstrues AI as strictly autonomous and attributes its "intelligence" to its autonomy rather than attributing its intelligence to its being social and relational, especially to humans [38]. This suggests a potential gap in the syllabi on critical topics and principles that play a crucial role in AI ethics. This result was prevalent in the majority of the courses. With computer science being the most frequent department in the dataset, we found that more technical aspects of AI ethics, such as transparency, security, privacy, and mitigating bias, were taught, focusing more on the implementation and development of models. Accordingly, there may be a need for more inclusiveness of human-centered ideals such as human rights around topics like surveillance, well-being (e.g., addiction that may manifest from efforts to increase user engagement with technology), and oversight for AI ethics to be more present in syllabi content. By deliberately calling attention to the fact that humans are responsible for what is being presented, it can empower students to take an active role in mitigating risk and reconciling potential concerns for individuals, groups, and society more broadly, as other scholars have described [39].

The unique words (pertaining to least covered topics, Table XI) further described that other ethics-centric topics are not often covered, although these cultural and spiritual factors may influence ethical decisions. The role of culture in establishing AI ethics cannot be stressed enough. Teaching this at the classroom level serves as a vital foundation for involving students of AI ethics in these discussions. Examples of pedagogical approaches include case studies, role plays, and examples [40]. Ethical plurality and cultural diversity are important in the construction of ethics in AI; otherwise, AI could easily become an instrument of domination and modern imperialism [41]. For instance, religion and faith have an undeniable impact on influencing students' ethical decision-making process due to various moral codes instilled by religion [42]. If students are presented with scenarios that conflict with their beliefs, it can affect how they handle ethical dilemmas from their religious and cultural viewpoints [42]. Therefore, it is vital for educators to be mindful and considerate of human factors such as religion, which can impact AI ethics education in the classroom and beyond.

The pre- and post-ChatGPT comparison (see Table XII) was also insightful, as we observed a shift in salient topics in the syllabi. The pre-ChatGPT era illustrated a higher emphasis on topics such as “history,” “morals,” “emerging,” and “autonomy,” which tend to suggest that topics were more focused on the historical and theoretical foresight of AI ethics. However, since the spurt in generative AI technologies, such as ChatGPT, we have observed a shift to topics like “explainability,” “decision-making,” and “LLMs,” which represents a timely change in the topical focus to being more hands-on and addressing more practical concerns regarding AI ethics. This represents how quickly the field of AI ethics is evolving and how it may impact what students and educators may need to know to address current trends in the field. While this urges for constant refinement of curricula in AI ethics across disciplines, it can also be worth recognizing that changing institutional curricula may also be constrained by various organizational and accreditation expectations. However, as applications and technology change, we must continue to understand how AI may have a global impact across society and consider sustainability.

All-in-all, this work has implications for university professors and curriculum developers who can benefit from the study by identifying potential gap areas in AI ethics syllabi as well as areas that are critical and need to be addressed at the policy and industry levels to ensure that academia can adequately prepare students for the next steps in their professional trajectories. This can include refining the curriculum constantly and allocating space to further open up AI ethics syllabi discourse to notions of ethical plurality regarding culture and societal impacts. There may also be different ways of doing this, like through the incorporation of case studies, news stories, or projects that can raise awareness and promote ethical thinking [27]. Policy also plays a crucial role in this regard, as it can provide an avenue for advocacy, aid in guiding the responsible development of AI, and suggest various areas for risk mitigation. Broader areas of impact include addressing the understanding of the socio-technical system of AI ethics and the interdisciplinary nature of the field. AI ethics requires all hands on deck, regardless of discipline and professional background, because it does not only involve technical applications. It entails public perception and establishing trust, thereby making it necessary to consider what technology can do and anticipate how it may be used. Additionally, this study can potentially bridge gaps between academia and industry by understanding what students in these fields are learning so as to equip them better for industry and policy expectations.

VI. Limitations and Future Directions

There are several limitations we would like to acknowledge. Only syllabi available in English were used, which may neglect additional perspectives. Going forward, we suggest expanding to be more inclusive of institutions from additional countries that are posted in other languages.

Moreover, publicly available syllabi focused specifically on AI ethics were limited, and often, such content was embedded within courses rather than serving as a standalone option. Hence, the sample size of 45 syllabi could be additionally expanded to consider how it may be applied, even if it is not the primary focus. While other syllabi were found, they ignored critical information such as the institution’s name and /or department that offered the course. In the future, we plan to expand our initial analysis.

Finally, we only applied a subset of possible analytical approaches, and further analyses could yield additional insight. K-means clustering forces a specification of k, and when there are outliers, k-means values tend to skew [43]. Accordingly, assessing the clusters can be difficult. Furthermore, LDA analysis can be difficult to interpret since it merely shows the distribution of topics across different clusters [44]. Including other analytical approaches, like agglomerate clustering, may emphasize different features and could yield unique insights.

VII. Conclusion

Our approach highlighted the utility of NLP clustering for analyzing AI Ethics syllabi and revealed distinct patterns. K-means clustering demonstrated a polarization between general/philosophical and specific/technical topics. Supervised clustering uncovered additional patterns that suggested implications for interdisciplinary collaboration and technological advancement. LDA topic modeling findings further aligned with globally converging critical areas of AI ethics such as data, privacy, bias, fairness, and transparency. Although it is encouraging to see strides in AI ethics syllabi, teaching, and learning in recognition of its importance, we must also continue to expand our efforts if we intend to prepare the next generation to not only know what AI can do but to use it responsibly and with regard to its potential impact.

Acknowledgment

The authors would also like to acknowledge Drs. Agoritsa Polyzou and Giri Narasimhan for their contributions to this project. We thank them for their support throughout its conceptualization and implementation.

References

[1] E. B. Hunt, *Artificial intelligence*. Academic Press, 2014.
[2] J. H. Korteling, G. C. van de Boer-Visschedijk, R. A. Blankendaal, R. C. Boonekamp, and A. R. Eikelboom, “Human-versus artificial intelligence,” *Frontiers in artificial intelligence*, vol. 4, p. 622364, 2021.
[3] M. Raza, “Artificial intelligence: What is it really?” 2019. [Online]. Available: https://datalchemy.ai/insights/posts/3-Artificial-Intelligence-Everything-You-Need-to-Know.html
[4] S. Feuerriegel, J. Hartmann, C. Janiesch, and P. Zschech, “Generative AI,” *Business & Information Systems Engineering*, vol. 66, no. 1, pp. 111–126, 2024.

[5] K. Kenthapadi, H. Lakkaraju, and N. Rajani, “Generative AI meets responsible AI: Practical challenges and opportunities,” in *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2023, pp. 5805–5806.

[6] V. Juho, “Ethics of ai technologies and organizational roles: Who is accountable for the ethical conduct,” in *Conference on Technology Ethics*, 2019.

[7] R. Strack, M. Carrasco, P. Kolo, N. Nouri, M. Priddis, and R. George, *The Future of Jobs in the Era of AI*, Mar 2021. [Online]. Available: https://web-assets.bcg.com/f5/e7/9aa9f81a446198ac5402aaf97a87/bcg-the-future-of-jobs-in-the-era-of-ai-mar-2021-r-r.pdf

[8] E. Brynjolfsson, D. Li, and L. R. Raymond, “Generative ai at work,” National Bureau of Economic Research, Tech. Rep., 2023.

[9] D. Oniani, J. Hilsman, Y. Peng, R. K. Poropatich, J. C. Pamplin, G. L. Legault, and Y. Wang, “Adopting and expanding ethical principles for generative artificial intelligence from military to healthcare,” *NPJ Digital Medicine*, vol. 6, no. 1, p. 225, 2023.

[10] L. Galindo, K. Perset, and F. Sheeka, “An overview of national ai strategies and policies,” 2021. [Online]. Available: https://www.oecd-ilibrary.org/science-and-technology/an-overview-of-national-ai-strategies-and-policies_c05140d9-en

[11] L. Schmitt, “Mapping global ai governance: a nascent regime in a fragmented landscape,” *AI and Ethics*, vol. 2, no. 2, pp. 303–314, 2022.

[12] IEEE, “About,” 2024. [Online]. Available: https://www.ieee.org/about/index.html

[13] ——, *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems First Edition*, 2019. [Online]. Available: https://standards.ieee.org/wp-content/uploads/import/documents/other/ead1e.pdf

[14] S. S. Bhullar, V. K. Nangia, and A. Batish, “The impact of academia-industry collaboration on core academic activities: Assessing the latent dimensions,” *Technological Forecasting and Social Change*, vol. 145, pp. 1–11, 2019.

[15] M. Khaleel and A. Jebrel, “Artificial intelligence in computer science,” vol. 2, pp. 1–21, 03 2024.

[16] R. Verdugo, M. Nussbaum, M. Claro, M. Sepulveda, B. Escobar, R. Rendich, and F. Riveros, “Preparing undergraduate computer science students to face intercultural and multidisciplinary scenarios,” *IEEE Transactions on Professional Communication*, vol. 56, no. 1, pp. 67–80, 2013.

[17] B. W. Smith, “An academic vision for ai ethics,” 2023. [Online]. Available: https://cyberlaw.stanford.edu/blog/2023/04/academic-vision-ai-ethics

[18] J. Antoniou and J. Antoniou, “Dealing with emerging ai technologies: Teaching and learning ethics for ai,” *Quality of Experience and Learning in Information Systems: Incorporating Learning and Ethics into Characterizations of Quality of Experience*, pp. 79–93, 2021.

[19] D. Gotterbarn, K. Miller, and S. Rogerson, “Computer society and ACM approve software engineering code of ethics,” *Computer*, vol. 32, no. 10, pp. 84–88, 1999.

[20] S. Spiekermann, “What to expect from ieee 7000: The first standard for building ethical systems,” *IEEE Technology and Society Magazine*, vol. 40, no. 3, pp. 99–100, 2021.

[21] M. D. Idris, X. Feng, and V. Dyo, “Revolutionising higher education: Unleashing the potential of large language models for strategic transformation,” *IEEE Access*, 2024.

[22] M. M. Rahman and Y. Watanobe, “Chatgpt for education and research: Opportunities, threats, and strategies,” *Applied Sciences*, vol. 13, no. 9, p. 5783, 2023.

[23] K. Zhang and A. B. Aslan, “Ai technologies for education: Recent research & future directions,” *Computers and Education: Artificial Intelligence*, vol. 2, p. 100025, 2021.

[24] D. Lee, M. Arnold, A. Srivastava, K. Plastow, P. Strelan, F. Ploeckl, D. Lekkas, and E. Palmer, “The impact of generative ai on higher education learning and teaching: A study of educators’ perspectives,” *Computers and Education: Artificial Intelligence*, vol. 6, p. 100221, 2024.

[25] E. Burton, J. Goldsmith, and N. Mattei, “Teaching ai ethics using science fiction,” in *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.

[26] R. T. Javed, O. Nasir, M. Borit, L. Vanhée, E. Zea, S. Gupta, R. Vinuesa, and J. Qadir, “Get out of the bag! silos in ai ethics education: unsupervised topic modeling analysis of global ai curricula,” *Journal of Artificial Intelligence Research*, vol. 73, pp. 933–965, 2022.

[27] C. Fiesler, N. Garrett, and N. Beard, “What do we teach when we teach tech ethics? a syllabi analysis,” in *Proceedings of the 51st ACM technical symposium on computer science education*, 2020, pp. 289–295.

[28] A. Orchard and D. Radke, “An analysis of engineering students’ responses to an ai ethics scenario,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 13, 2023, pp. 15 834–15 842.

[29] J. Saltz, M. Skirpan, C. Fiesler, M. Gorelick, T. Yeh, R. Heckman, N. Dewar, and N. Beard, “Integrating ethics within machine learning courses,” *ACM Transactions on Computing Education (TOCE)*, vol. 19, no. 4, pp. 1–26, 2019.

[30] A. Krakowski, E. Greenwald, T. Hurt, B. Nonnecke, and M. Cannady, “Authentic integration of ethics and ai through sociotechnical, problem-based learning,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 11, 2022, pp. 12 774–12 782.

[31] A. Jo, “The promise and peril of generative ai,” *Nature*, vol. 614, no. 1, pp. 214–216, 2023.

[32] S. M. Qadhi, A. Alduais, Y. Chaaban, and M. Khraisheh, “Generative ai, research ethics, and higher education research: Insights from a scientometric analysis,” *Information*, vol. 15, no. 6, p. 325, Jun 2024. [Online]. Available: https://www.mdpi.com/2078-2489/15/6/325

[33] C. Fiesler, “Tech ethics curricula: A collection of syllabi,” 2018. [Online]. Available: https://cfiesler.medium.com/tech-ethics-curricula-a-collection-of-syllabi-3eedfb76be18

[34] K. Hines, “History of chatgpt: A timeline of the meteoric rise of generative ai chatbots,” Jun 2023. [Online]. Available: https://www.searchenginejournal.com/history-of-chatgpt-timeline/488370/

[35] N. Reimers and I. Gurevych, “Sentence-bert: Sentence embeddings using siamese bert-networks,” in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, 2019.

[36] R. Kusters, D. Misevic, H. Berry, A. Cully, Y. Le Cunff, L. Dandoy, N. Díaz-Rodríguez, M. Ficher, J. Grizou, A. Othmani *et al.*, “Interdisciplinary research in artificial intelligence: challenges and opportunities,” *Frontiers in big data*, vol. 3, p. 577974, 2020.

[37] A. Jobin, M. Ienca, and E. Vayena, “The global landscape of ai ethics guidelines,” *Nature machine intelligence*, vol. 1, no. 9, pp. 389–399, 2019.

[38] D. Siddarth, D. Acemoglu, D. Allen, K. Crawford, J. Evans, M. Jordan, and E. Weyl, *How AI Fails Us*, 2021. [Online]. Available: https://ethics.harvard.edu/files/center-for-ethics/files/howai_fails_us_2.pdf

[39] B. Shneiderman, “Bridging the gap between ethics and practice: guidelines for reliable, safe, and trustworthy human-centered AI systems,” *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 10, no. 4, pp. 1–31, 2020.

[40] L. Furze, “Teaching ai ethics,” Jan 2023. [Online]. Available: https://leonfurze.com/2023/01/26/teaching-ai-ethics/comment-page-1/

[41] E. R. Goffi, “The importance of cultural diversity in ai ethics* — beyond the horizon issg,” Jan 2021. [Online]. Available: https://behorizon.org/the-importance-of-cultural-diversity-in-ai-ethics/

[42] B. Balakrishnan, F. Tochinai, H. Kanemitsu, and A. Altalbe, “Engineering ethics education from the cultural and religious perspectives: a study among malaysian undergraduates,” *European Journal of Engineering Education*, vol. 46, no. 5, p. 707–717, Feb 2021.

[43] C. Piech, “Cs221,” 2013. [Online]. Available: https://stanford.edu/~cpiech/cs221/handouts/kmeans.html

[44] J. Tang, Z. Meng, X. Nguyen, X. Edu, Q. Mei, Q. Edu, and M. Zhang, *Understanding the Limiting Factors of Topic Modeling via Posterior Contraction Analysis*, 2014. [Online]. Available: https://proceedings.mlr.press/v32/tang14.pdf