

Responsible Engineering in the Age of AI: The Value of Responsible AI Education from Engineering Students' Perspectives

Marie Mirsch^{a,1}, Sarah G. Moreno^a, Ben Schultz^a, Carmen Leicht-Scholten^a

^a RWTH Aachen University, Germany

ABSTRACT

Being a critical enabler of research and development, data-driven systems like Artificial Intelligence (AI) are increasingly relevant to engineers. Due to their generalizability and wide-ranging functionality, they are closely interwoven with social developments. With it comes the responsibility for instilling the right values and the need to gain knowledge of AI and its implications for society. A master's seminar at RWTH Aachen University trained engineering students in *Responsible AI* (ethical, social, and legal concerns with respect to AI) within engineering. To complement perspectives from industry and accreditation boards, we investigated students' reflection papers on the course to determine the relevance that engineering students give to their Responsible AI education. We found that prior to the seminar, students lacked knowledge about AI applications in engineering and assumed that technology (including AI) was neutral and unbiased. Yet after the seminar, students reported having corrected these assumptions. They expressed their positive beliefs about the importance of Responsible AI education in engineering, insisting that future engineers should consider the sociotechnical context of their work. This paper presents the results of the reflection paper analysis to address why engineering students see learning about Responsible AI, including its sociotechnical context, as relevant for their future careers.

KEYWORDS

Responsible AI, Ethical AI, AI Education, Machine Learning, Engineering Education, Engineering Ethics

PUBLICATION

Submitted:
31st August 2025

Accepted after revision:
15th December 2025

Introduction

Artificial Intelligence (AI) systems have become a fundamental driver of innovation, research, and interdisciplinary development across professional fields, including engineering (Cockburn et al., 2018). Engineers are increasingly integrating AI into their work to optimize efficiency, enhance quality control, or enable predictive maintenance (Peres et al., 2020). While data-driven systems have been

¹ Corresponding Author: Marie Mirsch – RWTH Aachen University, Germany, marie.mirsch@gdi.rwth-aachen.de

present in engineering for decades (Montáns et al., 2021), the sophistication and autonomy of AI systems pose new challenges, ranging from the unprecedented speed and scale at which AI systems develop to understanding the exact working and inner functioning of the systems (the 'black-box problem', e.g., Rudin, 2019). Further, as AI systems have become known for introducing unintended biases and discrimination, posing security risks and trust issues with unclear distribution of responsibility, deploying AI has raised pressing ethical, social, and legal concerns, which were recently summarized under the umbrella term of *Responsible AI* (Dignum, 2020). In this paper, Responsible AI does not refer to AI that is responsible. Rather, this shorthand term refers to the aforementioned ethical, social, and legal concerns with respect to AI, which responsible (human) agents should carefully consider when designing, implementing, or using AI. Their opacity, self-learning capabilities, and potential for unforeseen societal consequences make AI systems distinct from traditional digital systems, and make their risk assessments particularly complex (Metcalf et al., 2021). These aspects of AI also emphasize the ever-increasing importance of understanding the interconnectedness of society and technology, and of correcting the common misunderstanding that technology is a neutral tool (Martin et al., 2021; Sloane & Moss, 2019).

The ever-changing demand of adapting AI systems to specific deployment contexts and assuming responsibility for their outcomes requires engineers to develop sufficient AI literacy. For instance, the European AI Act calls for “a sufficient level of AI literacy” for “providers and deployers of AI systems” (AI Act, 2024), and the engineering education community explicitly calls for a future generation of “AI Engineers” (Schleiss, Bieber, et al., 2022, p. 1602). While this includes, for example, professional competencies such as data and AI knowledge or methodological competencies such as process- and system thinking (Schleiss, Bieber, et al., 2022), employers in engineering and related fields emphasize the need for AI education that extends beyond technical proficiency to include ethical reasoning (ElZomor et al., 2020; Khakurel et al., 2018; J. Mitchell & Guile, 2021).

In a similar vein, engineering ethics education aims to equip future engineers with the necessary competencies to fulfill their social responsibilities (Martin et al., 2021; Pierrakos et al., 2019). Traditional approaches in engineering ethics education primarily emphasize ethical principles and professional codes of conduct (Breuer & Genske, 2021; Martin et al., 2021). While these frameworks provide valuable guidance in ethically complex decision-making, they may fall short in fostering the development of a morally responsible engineering identity. A stronger focus on character education encourages intrinsic motivation for ethical behavior rather than mere rule adherence (Harris, 2008; Pierrakos et al., 2019). That means, students are encouraged to cultivate ethical virtues such as empathy, responsibility, and social awareness, which are essential for addressing engineering practices' multidimensional challenges (Bergen & Robaey, 2022; Clancy & Zhu, 2023). Furthermore, focusing engineering ethics education on character would not only entail acquiring subject-specific skills but also the ability to transfer moral behavior across different disciplinary contexts.

The master's seminar, Innovation & Diversity, trained engineering students at RWTH Aachen University in Germany on the sociotechnical context of AI in engineering applications. By *sociotechnical* context (Dolata et al., 2022; Sætra, 2021; Sarker et al., 2019), we refer to the vital component of Responsible AI that engages with the *societal* context impacting and impacted by *technology*. While Responsible AI is an umbrella term that comprises legal and environmental questions, we focused on ethical questions regarding the impacts of AI on society.

We aimed to explore students' perspectives on learning about the sociotechnical context of AI in engineering through qualitative analysis of their reflection papers. In particular, we wanted to know to what extent they value the topic itself and how learning about it could enrich engineering ethics education in general. Therefore, this paper aims to answer the following research question:

RQ: Why is education about the sociotechnical context of AI relevant to engineering students?

While our empirical work focuses on the sociotechnical context of AI, we first provide related literature on Responsible AI in Engineering Education in general. Afterwards, we describe the Innovation & Diversity seminar and the final assessment task, which required students to submit reflection papers leveraged in our post-hoc analysis, followed by an explanation of our coding process and the discovered categories (*Methods*). We then present our findings regarding students' own realizations following the seminar in three parts (*AI Is Now Fundamental in Engineering Yet Overlooked in Education*, *Considering the Sociotechnical Context Is Crucial: AI Is Not Neutral*, and *Responsible AI Education Is Essential for Engineers*), which are each discussed in turn before we offer our overall conclusion.

By investigating their prior knowledge of AI and ethics, their perception of AI's neutrality, and their general engagement with ethical considerations in engineering, our contribution is three-fold: *First*, we show that engineering students lack knowledge about AI in engineering despite the fact that AI is often used in engineering (*AI Is Now Fundamental in Engineering Yet Overlooked in Education*). *Second*, by highlighting students' dominant realization from our course that technology is not value-neutral, we demonstrate how a seminar on the sociotechnical context of AI in engineering can contribute to the character education of engineers and, thus, highlight the value of such a course on engineering ethics education in general (*Considering the Sociotechnical Context Is Crucial: AI Is Not Neutral*). We *conclude* by highlighting students' general interest in learning about the sociotechnical context of AI and their own assessment of the importance of learning about it with respect to their own future careers as engineers (*Responsible AI Education Is Essential for Engineers*). In the discussion, we summarize that learning about the sociotechnical context of AI is highly relevant to engineering students but can also teach them ethical competencies that are transferrable to other engineering topics.

Responsible AI in Engineering Education

Young professionals entering the workforce are expected to possess programming proficiency (Liu & Xie, 2021), the ability to work with AI tools (Cetindamar et al., 2024), and the critical thinking skills necessary to evaluate AI systems (Ng et al., 2021). Despite these expectations, a gap remains between university curricula and industry requirements (Da Costa et al., 2019). Although AI education is gaining traction in higher education (Aqlan & Nwokeji, 2018; Pilario, 2024; Schleiss, Hense, et al., 2022; Silapachote & Srisuphab, 2016; Venkatasubramanian, 2022), it remains an emerging discipline (Heidling et al., 2019). Furthermore, AI education in engineering rarely focuses on responsible and ethical AI practices, as an exploratory evaluation of engineering competence profiles shows (Decker et al., 2024). For instance, Schleiss, Hense, et al. (2022) implemented a project-based learning approach using open educational resources in three courses, demonstrating that students valued digital flexibility, autonomy in self-directed learning, and hands-on engagement with AI applications. Similarly, Pilario (2024) designed a course for chemical engineers emphasizing machine learning methods, case studies, and hands-on practice. While students appreciated the blend of theory and application, some expressed a desire for a deeper focus on the ethics of AI.

Incorporating the sociotechnical context into engineering education, in general, is vital for bridging theoretical knowledge with real-world applications. Engineering projects inherently operate within societal frameworks, impacting human lives and labor. Recognizing these interconnections, guidelines for engineering education advocate for fostering awareness of the social context and instilling a sense of social responsibility in students (Martin et al., 2021). Courses on responsible conduct, sustainability, or technology ethics can serve as vehicles for integrating social responsibility into engineering education. However, the inclusion of such interdisciplinary courses remains scarce (Lima et al., 2019), and when present, they often focus narrowly on professional ethics, plagiarism, and compliance with codes of conduct. Broader concerns, such as fairness, public welfare, and the societal implications of technology, are rarely addressed (Martin et al., 2021). Moreover, despite the importance of ethical education, engineering students frequently exhibit *disinterest*, *resistance*, and *difficulty* engaging with these topics (Bairaktarova & Evangelou, 2011; Polmear et al., 2019). Studies suggest that engineering education often prioritizes technical problem-solving while downplaying societal and ethical dimensions (Cech, 2014; Godfrey, 2014; Schiff et al., 2021). This emphasis on technical skills can lead to ethics being perceived as a “soft” and non-essential aspect of the curriculum (Martin et al., 2021). As a result, students may prefer ethics and social science courses to be optional and unassessed (Sucala, 2019), displaying lower engagement and emotional investment compared to technical subjects (Balakrishnan & Tarlochan, 2015; Newberry, 2004). Furthermore, the complexity and ambiguity inherent in ethical decision-making present an entirely different range of challenges, as they require problem-solving that contrasts sharply with the clear-cut solutions typical of traditional, technical problem-solving. This misalignment may contribute to a decline in ethical reasoning and socially responsible behavior throughout and after their education (Bombaerts & Nickel, 2017; Martin et al., 2019). However, to proactively recognize and mitigate ethical risks at an early stage, developing AI education that fosters both technical and ethical competencies while actively engaging students is essential. Moreover, as we will demonstrate, an AI ethics course has the ability to teach students to consider the ethics of technology and engineering in general; hence, the knowledge is transferable to other domains.

Methods

In this section, our seminar, Innovation & Diversity, on the sociotechnical context of AI in engineering is briefly described before introducing the final examination task that provided us with the reflection papers for our post-hoc qualitative data analysis. Then, we discuss how we systematically coded the students’ anonymized and ungraded papers for a qualitative content analysis in line with Kuckartz’ (2019) methodology. The categories discovered that are relevant to the RQ in this paper are then defined in detail.

The Innovation & Diversity Seminar

The elective master’s seminar, Innovation & Diversity (I&D) was designed and offered by the institute for Gender and Diversity in Engineering (GDI) at RWTH Aachen University. The course is open to students from various disciplines, mainly civil and environmental engineering, construction and robotics, technical communication, and sociology. This recurring seminar investigates multiple innovation topics (now AI) within the context of an inclusive society, and interdependencies between science and technology are discussed.

The seminar is conducted every winter semester and is limited to 20–30 students (though we had 60–70 interested participants each semester). The aim is to introduce students to fundamental technical aspects of AI, providing a foundation for the course’s genuine focus on discussing societal impacts in the context of AI as an innovative technology affecting many different engineering areas. This places the seminar at the intersection between engineering, AI, and ethics (see Figure 1). The seminar begins by introducing the basics of AI and discussing ethical principles, which is meant to be a starting point for the moral evaluation of real-world AI applications. In particular, the ways in which AI can be used to assess current global challenges are explored, as are ethical discourses on justice, responsibility, and diversity. Other critical aspects of AI, such as the social aspects of sustainability and responsibility, are also addressed. The seminar’s main objectives are for students to understand the connection between diversity and innovation, research current examples of AI applications within their engineering fields, reflect critically on AI use in their engineering domain, and conduct ethical evaluations of these real-world applications. Further acquired knowledge and skills include academic reading, research, and writing; knowing sociological and ethical concepts useful for a holistic approach to engineering; collaboration and presentation skills; and critical thinking, all pertinent to a future engineer’s career (Moreno et al., 2024).

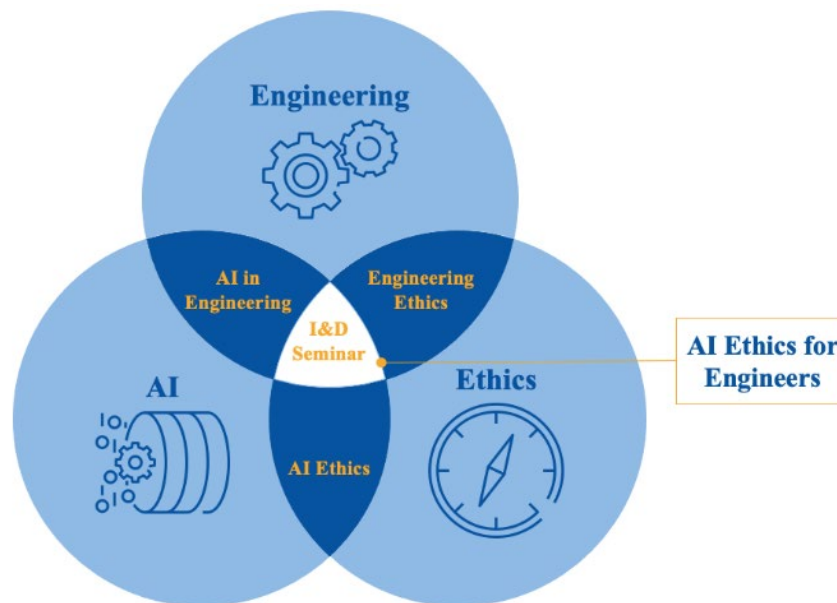


Figure 1 : Our seminar addresses the intersection between AI, ethics, and engineering. Topics in the dark blue cross-sections of the Venn diagram are often taught in higher education institutions, while AI ethics for engineers at the center of the Venn diagram is uncommon in engineering curricula, but it is the focus of the Innovation & Diversity (I&D) seminar offered by the Gender and Diversity in Engineering (GDI) institute of RWTH Aachen University.

The examination consisted of four different examination tasks (ET1–4). The first three tasks were completed as a group, each building on the previous task, while the final task was completed individually at the end of the course. ET1 asked students to find an academic paper presenting an example of AI in civil or environmental engineering, introduce it in their own words, and describe the societal and ethical challenges and risks that may arise in their chosen example. ET2 had groups prepare a 10–15-page essay detailing this example and, most importantly, the ethical principles covered in class and how they were violated, and present their own solution approaches for the

violations. For ET3, groups prepared scientific posters presenting the main results from their essays as well as a short video focusing on their proposed solutions. Finally, ET4 asked students to write a reflection paper about their experience during the course. These reflection papers now form the basis of our post-hoc qualitative data analysis. Specifically, the students were asked to reflect on the following three questions:

What was most interesting in the course?

What was most surprising?

What are your takeaways?

These reflection papers are evaluated in this paper with respect to the RQ mentioned in the introduction.

Data Analysis

This analysis is based on data collected in 2023 from 20 students who participated in the seminar. Due to the open-ended and reflective nature of the final examination task, students provided insights beyond the three questions in the prompt. This led us to realize the wealth of valuable information that the reflections provided and become interested in further exploring students' self-reported positions on the topics of the course and on the course itself, as well as the competencies successfully acquired during the seminar (see Moreno et al., 2024). All twenty ungraded reflection papers form the basis of our post-hoc qualitative data analysis, although one was short, superficial, and did not provide much valuable insight. The unstructured format of the reflections revealed a complex overlapping of categories, which lends itself to Kuckartz' (2019) qualitative content analysis to discover and compare these themes while supporting an exploratory style of analysis with respect to the RQ. Before beginning our analysis, we cyclically defined categories and coded our students' anonymized reflection papers. To this effect, we used MAXQDA qualitative data analysis software to help organize and filter our segments, which we manually coded following Kuckartz' (2014, 2019; Kuckartz & Rädiker, 2019) methodology. We did not use any AI to aid in our analysis. A total of three coding cycles were completed:

- 1) a concept-driven (deductive and inductive) cycle based on our research question and the task's writing prompts to develop a first draft of our main categories,
- 2) a second data-driven (inductive) cycle to refine our list of main categories via open-coding in MAXQDA and finish coding all segments for these main categories, and
- 3) a final inductive cycle to derive and code for subcategories in MAXQDA.

The length of segments was not explicitly limited; we agreed that a coded segment should include what was necessary to view it out of context and still understand the student's argument and why it fit in the assigned category.

In the first cycle, we each coded 5 of the 20 papers using an initial list of thirteen main categories and discussed. Two other categories worthy of our attention were inductively discovered during this time. We also merged and renamed categories where we saw significant overlap or for more clarity. Thus, the final eleven main categories (course structure and methods, expectations before the seminar, interest in the seminar topics, value-neutrality of technology, relevance of AI ethics (to a future engineer), applications of AI in engineering, previous experience with AI, previous experience with ethics, surprising, misunderstandings, and challenges) were decided. (In this paper, we will only cover segments of the "surprising" category when they overlap with other categories.) Subsequently, in the second cycle, we continually coded, compared, and edited category codings until we completed all 20

papers and resolved coding disagreements. This “consensual coding” strategy was preferred over intercoder reliability coefficients for its simplicity in qualitative text analysis and full agreement at the end of coding (Kuckartz, 2014, p. 46). The third cycle involved inductively deriving possible subcategories for each main category needing further subdivision by reviewing all coded segments of a single main category. Then, we combined subcategories where possible to reduce the number of subcategories to the essential before assigning each main category segment to one or more of its subcategories.

The final main categories relevant to this paper are listed below, along with their coding criteria. The categories have been lettered for ease of reference. In the case of the categories analyzed in this paper, their subcategories (when applicable) are mutually exclusive and self-explanatory, thus, not requiring further explanation of their coding criteria. (See Appendix A for the full breakdown of main categories and subcategories.)

- A) Students’ self-declared personal experience with AI belongs in the **previous experience with AI** category. Such experience may include but is not limited to experience using or learning about AI in a class, internship, job, or even references to pop culture, such as AI as it is represented in movies. Pop culture “experience” is included because it may be the only perception of AI that a student has before learning, during this seminar, how it is realistically implemented in applications they are already using, as well as in applications they may one day use, develop, or implement at work. Keywords include “before,” “previously,” and “AI.” Previous experience with AI ethics belongs to both this category (“previous experience with AI”) and the one below (“previous experience with ethics”).
- B) Self-declared **previous experience with ethics** could feature experience learning about or being aware of issues concerning ethics in a class, internship, job, or students’ personal lives. Keywords for the category include “before,” “previously,” “ethics,” “fairness,” “diversity,” “bias,” “sustainability,” and other societal needs or aspects. As mentioned above, instances of previous experience with AI ethics belong to both categories: “previous experience with AI” and “previous experience with ethics.”
- C) The category **applications of AI in engineering** contains any examples of AI in an engineering context that students mention. These may be examples presented in class, applications students researched as part of their group work, or original examples the students provided only in their reflection papers.
- D) A category that came up after beginning our review was the **value-neutrality of technology** category. It was not one of our initial categories but was mentioned in so many reflection papers that we decided to afford it its own main category. After all, it was a very important and surprising realization for many students. Keywords and phrases include “bias,” “neutral,” “AI/technology/machine,” and “think on its own.” Statements in this category generally describe the prior assumption of technology being neutral. This category does not simply capture (un)ethical technology, but more specifically *neutral* vs. *biased* technology.
- E) Keywords for the category **interest in the seminar topics** include “interesting,” “enjoy,” and “fascinating.” Content that sparked curiosity, that was intellectually stimulating, or that was emotionally engaging also pertain to this category. Statements should refer particularly to the topics or content of the seminar. For example, the phrase “group work was interesting” would belong to the category “course structure and methods” rather than to this category (“interest in the seminar topics”) even though it contains the key word “interesting” since it refers to a course method rather than to a specific seminar topic.

F) The **relevance of AI ethics (to a future engineer)** category encompasses motivation for the relevance of ethical AI in engineering but also includes mentions of AI ethics in general with the assumption that it is in the context of engineering since that was the context of the seminar and of all the students' master's programs. This category also includes mentions of AI regarding topics of diversity, fairness, sustainability, and other societal needs or aspects relevant to AI ethics.

Table 1 : Relevant categories and subcategories for thematic analysis in this paper, along with the number of coded segments assigned to each (sub)category and the number of students whose reflection papers contained segments assigned to each (sub)category.

Main Category (# Coded Segments, # Students Out Of 20)		Subcategory (# Coded Segments, # Students Out Of 20)
A	Previous experience with AI (15, 10)	N/A
B	Previous experience with ethics (12, 10)	N/A
C	Applications of AI in engineering (25, 14)	(C.1) Topics from examination tasks or class (17, 10)
		(C.2) New topics (8, 7)
D	Value-neutrality of technology (23, 12)	N/A
E	Interest in the seminar topics (29, 16)	N/A
F	Relevance of AI ethics (to a future engineer) (44, 18)	(F.1) Explicitly mentioned engineering context (17, 11)
		(F.2) Assumed engineering context (27, 14)

The main categories along with relevant subcategories can be found in Table 1. We focus on qualitative analysis and do not perform a quantitative analysis. Instead, the numbers of coded segments and students in Table 1 are intended to illustrate the perceived importance of each category and subcategory to the students participating in the course. Furthermore, Figure 2 presents the relative proportion of segments in each category and subcategory examined in this paper, potentially revealing what students preferred to reflect on, or possibly merely indicating what students thought they were supposed to write about in their reflection papers. Either way, it shows which categories provided more data for us to analyze, thereby contextualizing our findings. Meanwhile, Figure 3 visualizes our code matrix as a heatmap, showing more precisely the number of segments coded per student per (sub)category. This supports an understanding of the distribution of segments among the students. The dataset supporting the findings of this study have been deposited in the Figshare repository available under the reserved DOI <https://doi.org/10.6084/m9.figshare.30621935>. The coded segments are organized by category, and all categories are included in the dataset, even ones not directly analyzed in this paper.

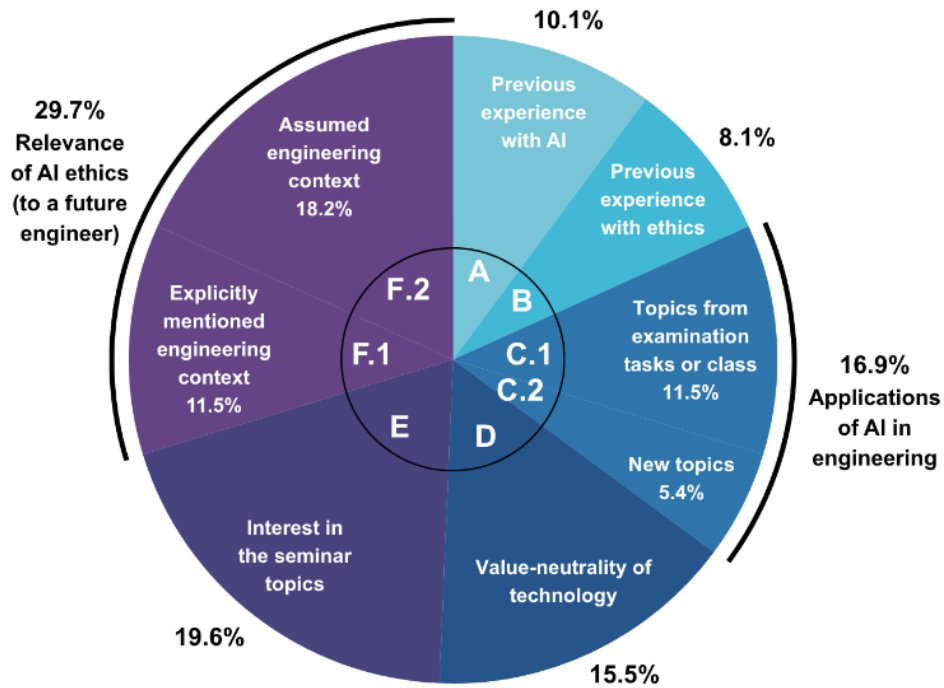


Figure 2 : Proportion (%) of segments discovered per (sub)category examined in this paper, where the total number of segments coded (N) is 148. It should be noted that some of the same coded text could be assigned to multiple categories such that segments overlapped.

	Student																				Σ
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
A	1	2	0	1	1	1	2	0	0	0	0	0	0	0	0	1	0	3	1	2	15
B	1	1	1	2	2	1	0	0	0	0	0	1	0	0	0	1	0	1	0	1	12
C.1	2	1	2	0	5	0	1	1	0	1	0	0	2	1	0	0	0	1	0	0	17
C.2	0	0	1	1	0	0	0	1	0	0	0	1	0	0	0	1	0	1	2	0	8
D	0	1	0	2	0	1	0	2	0	0	0	2	2	2	2	2	2	0	2	3	23
E	3	2	1	1	2	2	1	1	0	3	0	2	1	4	2	1	0	1	0	2	29
F.1	1	2	0	1	1	1	1	0	2	1	0	1	4	0	1	1	0	0	0	0	17
F.2	0	2	1	2	2	2	0	5	0	1	0	1	0	2	2	2	0	2	2	1	27

Figure 3 : Heatmap code matrix generated using MAXQDA, showing the number of segments coded per (sub)category per student, as well as confirming the sum (Σ) of coded segments per (sub)category.

Results

First, in *AI Is Now Fundamental in Engineering Yet Overlooked in Education*, categories A and C will demonstrate that many students lack knowledge about AI in engineering despite the fact that AI is often used in engineering. In *Considering the Sociotechnical Context Is Crucial: AI Is Not Neutral*, categories B and D will show that AI is not neutral or unbiased, though students often lack this awareness. Finally, in *Essential AI Education for Engineers*, categories E and F will establish that it is essential for engineering students to learn about both AI in engineering and the social context of AI.

To differentiate between students’ reflections while still assimilating statements from the same student and while retaining anonymization, we will refer to individual papers by a number randomly assigned between 1 and 20 for the 20 papers received. Thus, every reference will be followed by “s” (for student), a number between 1 and 20, and, when a direct quotation is taken, a page number indicating on which page of the reflection paper the quoted segment was found (e.g., (s18, p. 1)). (For all full segments referenced here, see Appendix B.)

Table 2: The structure according to which we present our findings accompanied by exemplary student quotes. This functions as an overview of our findings. We refer, here, to our *students’* findings since our findings in fact reflect the students’ own realizations after completing our seminar.

Students’ Findings	Main Category		Sub-category	Exemplary Student Quote
AI Is Now Fundamental in Engineering Yet Overlooked in Education	A	Previous experience with AI	N/A	<i>Getting an introduction to AI was very important for me. I knew of AI, but I didn’t have any technical knowledge about it. (s20, p. 2)</i>
	C	Applications of AI in engineering	(C.1) Topics from examination tasks or class	<i>[The use of AI in civil engineering] highlights how AI is changing the way we approach traditional civil engineering problems and provides new opportunities for innovation and optimization. The use of AI in civil engineering is making design and construction processes more efficient, accurate and cost-effective. (s18, p. 1)</i>
			(C.2) New topics	<i>One of the most interesting aspects of the course was learning about the various types of artificial intelligence and their applications. I found it fascinating to discover how AI can be used in so many different areas... It was also intriguing to learn about the potential benefits and drawbacks of using AI in these fields. (s8, p. 2)</i>

Considering the Sociotechnical Context Is Crucial: AI Is Not Neutral	B	Previous experience with ethics	N/A	<i>Taking other aspects like social aspects into consideration and not just treating it from a technical viewpoint as an engineer has been a new experience. (s4, p. 1)</i>
	D	Value-neutrality of technology	N/A	<i>I used to think that technology is just that, a machine, so why would it have bias towards anything if it doesn't even 'think' by itself? (s2, p. 1)</i>
Responsible AI Education Is Essential for Engineers	E	Interest in the seminar topics	N/A	<i>Topics discussed in the lecture broadened my horizon and understanding on fairness, discrimination, and diversity. I particularly enjoyed the parts on different concepts of diversity. ... Understanding and especially defining diversity is an important foundation in order to actually consider and respect the diversity of societies while making use of AI. (s10, p. 1)</i>
	F	Relevance of AI ethics (to a future engineer)	(F.1) Explicitly mentioned engineering context	<i>Considering my field of study (Engineering), I am almost certain that at some point during my career I will come in a deeper contact with this technology [(AI)] and I'm grateful this course has put the idea in me of not taking that work so easily and to give a fair amount of consideration to all the ethical principles I have come across during these past 4 months. (s15, p. 2)</i>
			(F.2) Assumed engineering context	<i>I have realized that since taking this course, I have become an unintentional activist about the ethics of AI, since the conversation about its powers and promises comes quite frequently in my social circles. ... I hope this lasts well after the course is over, both for my personal and professional life. (s15, p. 2)</i>

AI Is Now Fundamental in Engineering Yet Overlooked in Education

This section demonstrates the students' (lack of) knowledge about AI in engineering (A) despite the fact that AI is often used in engineering (C), revealing how AI is now a fundamental technology used in engineering that any future engineer should have knowledge of. Since engineering students seem to lack this knowledge, AI is likely a topic often missing from their education.

A) Previous experience with AI

Out of 20 students, 10 mentioned their (mostly nonexistent) previous experience with AI. None of the students expressed having significant prior experience with AI while most (6 of these 10) had no prior experience at all, as one student summarized, "we all are new to AI topic" (s18, p. 1). Meanwhile, only one student reported to have taken a course parallel to our seminar that discussed the "latest developments and innovative technologies (some of them being AI-based) in the construction sector"

(s5, p. 1). Another student explicitly expressed that they were “worried about having to learn about AI and Machine Learning, something that at first seemed out of reach and very complex to understand” (s2, p. 1). Nevertheless, they reported feeling more comfortable with the topic after the introduction provided by our seminar. Before attending the course, many students (8) were unaware of various current uses of AI technologies across diverse fields (e.g., s18) and “most importantly” their impacts “despite the fact that AI systems are a huge part of our everyday life” (s19, p. 2).

C) Applications of AI in engineering

14 out of 20 students mentioned applications of AI in engineering, the majority of which (10 students) explicitly referred to AI application contexts that they had found while conducting literature research for their portfolio examination or that had been examples discussed in class. 7 of the students mentioned examples that were not directly related to seminar discussions or groupwork.

C.1) Topics from examination tasks or class

As part of the portfolio examination, students were asked to choose an example from the fields of civil (e.g., construction) or environmental engineering. Therefore, many students (6) refer to the use of AI in construction. This includes, for example, building automation (s2) and wearables that can be used on a construction site (s5) such as smart vests or smartwatches (s14). One student stated that they were previously unaware of these uses for AI in construction (s7). Another student explained that “conducting research in the search for a use case for the civil [engineering] industry... gave [them] new ideas regarding possibilities of optimizing processes either using data-driven technologies or rather in the training of data” (s5, p. 1), showing students’ attempt to extend learned applications to new areas.

Another topic chosen by students was disaster risk management. Many students (5) referred to the use of AI in environmental contexts, including, for example, earthquake prediction (s3), disaster risk reduction (s8), weather prediction (s10), and flood vulnerability assessment (s13). Meanwhile, another student referred to some examples discussed in class, including smart mobility and short-distance car sharing (s3).

Several students (3) expressed difficulty in identifying examples where AI was used in engineering. One student expanded on this, explaining that the seminar’s requirement for literature research on AI applications in civil engineering was difficult since such literature was scarce (s7). Nevertheless, many students (8) found this task interesting and rewarding: “researching and switching topics was an interesting experience, as we got to see a lot of use cases of AI in environmental and civil engineering” (s10, p. 1).

C.2) New topics

Many students (7) were previously unaware of “how AI can be used in so many different areas” (s8, p. 2). However, after completing the seminar, one student explained that the use of AI will grow soon: “despite being a relatively new technology there are already many successful implementations of AI in different aspects of our lives. From the recently published literature, it is evident that, the use of AI in different domains will grow rapidly in coming years” (s3, p. 2). Others mentioned AI in contexts related to health (s12), GPS technology in safety equipment (s12), and drones for surveying and mapping (s18). One student explained that AI in finance and manufacturing can be “used to improve efficiency, accuracy, and decision-making in these fields, leading to potential cost savings and improved outcomes” (s16, p. 1). Another student discussed advancements in computer vision and its applications in surveillance or autonomous driving, as well as the use of AI in generating music, designing buildings, and creating art (s19).

Considering the Sociotechnical Context Is Crucial: AI Is Not Neutral

This section demonstrates the students' (lack of) prior exposure to ethics (B), including AI ethics, and their perception of the value-neutrality of technology (D). Together, the analysis of these categories shows how students' newfound realization during the seminar that AI is *not* neutral has generalized to an understanding that *technology* is not neutral, leading to an awareness of how crucial it is to consider the sociotechnical context of any engineering technology, process, or application.

B) Previous experience with ethics

In addition to lacking prior experience with AI, most students also lacked prior experience with ethics. Half (10) of the 20 students mentioned their (lack of) previous experience with ethics. None of the students reported to have engaged in ethical deliberation before—neither in AI-related fields nor in engineering more generally. 2 students explicitly expressed that they had “never heard of ethical principles of AI before” (s1, p. 2; see also s20). Some students (3) were also surprised to discover that ethics is relevant not only in more obviously AI-centered fields, but also in other engineering fields, as one student explained, “it never occurred to me that Ethics could be applied to other sectors, least of all to construction [in civil engineering]” (s2, p. 1).

Similarly, one student referred to ethical AI when they explained that they were “aware of its importance... [but] had not given much thought to the various ethical considerations that come into play when developing and using AI systems” (s16, p. 1), hence acknowledging a learned “importance” of a vague concept (ethics) without learning how this “important” value may be upheld. Another student reflected that they had learned in a different course what makes an innovation financially and competitively successful, but “had not considered the... added value for humanity... [or] the dangers and limits of innovation from a moral point of view” (s12, p. 1).

Furthermore, many students (all 10 who mentioned their previous experience with ethics) found that they had a limited or inaccurate conception of ethics and the impact of ethical considerations. For example, one student reported, “initially, I had a narrow understanding of ethics and mostly thought of data privacy and safety when thinking about ethics” (s5, p. 1) while another reported that “discrimination in AI is more common than [they had] previously thought and its connection to diversity is equally surprising” (s4, p. 1).

D) Value-neutrality of technology

During the seminar, we discussed some of the limitations and negative effects of AI. This prompted students to reflect on the possibility of neutrality or bias of technology in general. In their reflection papers, 12 out of 20 students referred to what we summarize as perceptions of value-neutrality of technology (or lack thereof). They often referred to AI but were also able to generalize to other machines and technology, as exemplified in the quote included for D in Table 2.

Most of the students (10 of the 12 who wrote about value-neutrality) highlighted that they had not expected AI to be positively or negatively value-laden but had instead expected a “purely objective,” neutral technology. One student described having thought of AI as a “simple tool to fulfil tasks... [making it] easy to perceive AI-models as equitable and impartial pragmatic instruments” yet stated that “it is [now] quite interesting to see that that’s not the case at all” (s4, p. 1). Another explained that “it was fascinating to learn that ‘machines’ or artificial intelligence (AI), which [they] had previously assumed to be error-free or rational, may make mistakes just like people do” (s8, p. 2). Yet another described that they “[had] never put an emphasis on the integrity of AI and its impact on social settings, and always had an image of AI as perfect in most of its applications” (s13, p. 2) before participating in this seminar on Responsible AI. Most students had thought of AI as a neutral technological

innovation and were aware of many of AI's beneficial impacts but were surprised to discover that many applications also inflict harmful effects on society. Several students (7) referred to discrimination and bias in relation to the concept of value-neutrality. One of these students was surprised that “with a machine, which operates on the facts and figures, moral questions arise” (s17, p. 1). Another wrote that one of their main takeaways from the course was that AI had “the potential for bias. AI systems are not neutral – they are only as unbiased as the data they are trained on... [which] can lead to unfair and discriminatory outcomes, particularly when it comes to marginalized groups” (s19, p. 2). Other harmful effects of AI on society were also mentioned, such as biases, prejudice, and human errors (s4), and, more specifically, the impact of racial accuracy disparities in face recognition software in use in the United States (s20).

Some students (2) described AI as a “double-edged sword”: for example, “it can actually be a weapon for inclusivity, [creating] opportunities for vulnerable groups, and... improving impartial decision-making processes in the civil or environmental engineering field. ... On the other hand, it is surprising that AI can exacerbate existing inequalities if they are not controlled and monitored regularly” (s13, p. 2). Similarly, another student wrote, “on the one hand, it was interesting to learn that AI can be used to promote diversity and inclusion,” citing its potential in hiring more diverse candidates, while “on the other hand, it was surprising to learn how AI systems can maintain and even reinforce existing biases, for example through concept drift” (s14, p. 2). The same student also reflected on “how important it is to consider the potential consequences of these biases when designing and deploying AI systems. ... [although they] had never given it much thought” (s14, p. 2) prior to the seminar, and further elaborated on the potential advantages and disadvantages of deep fakes and AI art.

Students often (10 students) expressed an emotional reaction to realizing that technology is in fact biased: “to learn that technology can be biased and even exacerbate human prejudices... was an immense discovery” (s2, p. 1) and “it was very surprising and indeed scary how discriminating AI could be” (s6, p. 1). Words that were frequently used in this context were “surprising” (s4; s14; s16; s17; s20), “fascinating” (s8), and “very interesting” (s19).

It was striking that students often (4 students) explicitly distinguished between “technical” and “social” aspects of technology as if they were mutually exclusive categories, not (yet) acknowledging that these two aspects are deeply intertwined. One student wrote, “as an engineer, I was always considering the technical aspect of it. Nevertheless, I got to know that the social aspect is significantly important, and it could be a key element in deciding whether AI should be implemented or not” (s6, p. 1). Meanwhile, another student expressed that “we tend to consider this emerging field as one [that is] strictly technical and analytical since its closest ‘relative’ is computer programming... [implying the common misunderstanding] that this only means giving commands to a machine to do something you want, albeit faster and more precisely” (s15, p. 1).

It was additionally notable that some students (3) described how they transferred their newfound awareness regarding the lack of value-neutrality of AI to other engineering or innovation technologies. One student reflected that “this leaves you wondering not only in what other types of technology the same happens, but also what do we take for granted,” (s2, p. 1) likely alluding to other widespread beliefs or practices we do not question. Meanwhile, another wrote, “I have learned that innovations or innovative technologies, or even the idea itself, is not automatically morally justifiable in all respects, even if the intentions are only well meant” (s12, p. 2).

Responsible AI Education Is Essential for Engineers

This section demonstrates the students' general interest in the seminar topics (E) and their own assessment of the importance of Responsible AI for their own future careers as engineers (F), leading to an overall sense that Responsible AI education is essential for future engineers, especially considering how AI is now fundamental in engineering.

E) Interest in the seminar topics

16 out of 20 students made statements that we interpreted as expressing their interest in the seminar topics. 7 students stated that they found the learned ethical principles for AI interesting (s1; s2; s6; s10; s12; s14; s16). This is encouraging since they were an important part of the seminar and examination tasks. 5 students found the considerations of fairness, diversity, and equity particularly interesting (s3; s4; s6; s10; s13). For example, one student wrote that “the most interesting thing... was learning to look at the term fairness from different perspectives. Something seemingly fair in common sense or from a statistical point of view might not be fair at all. Even after very careful, informed consideration it still might not be fair from a different undiscovered perspective” (s3, p. 2). Another 5 students expressed their interest in diving into the different applications of AI (s1; s7; s8; s14; s18). One student explained, “one of the most interesting aspects of the course was learning about the various types of artificial intelligence and their applications. I found it fascinating to discover how AI can be used in so many different areas. ... It was also intriguing to learn about the potential benefits and drawbacks of using AI in these fields” (s8, p. 2). 2 students explicitly expressed that what was most interesting to them was to discover that technology is not value-neutral (s4; s14), as discussed in the previous section, but in fact a “double-edged sword” (s14, p. 2). 2 students further highlighted that it was most interesting to consider social aspects instead of only “purely technical” ones (s12, s20). Another 2 students found it interesting to see how complex discussions of ethical topics are (s5; s12). Finally, one student explained that the sociotechnical context is sorely missed in other courses of their study program: “since the entire civil engineering studies in my bachelor's degree was so technical, I really like the seminars of the GDI because they give you a different point of view on traditionally technical fields. ... it was again a new train of thoughts for me” (s20, p. 2). Ultimately, our students expressed interest in the following topics: ethical principles for AI; considerations of fairness, diversity, and equity; applications of AI in diverse fields; the lack of value-neutrality of AI; sociotechnical aspects (the addition of so-called social aspects rather than purely technical ones); and the complexity of ethical discussions and tools that help to critically reflect on AI.

F) Relevance of AI ethics (to a future engineer)

18 out of 20 students mentioned the relevance of AI ethics. 17 segments (from 11 students) explicitly mentioned the engineering context when referring to the relevance of AI ethics, while in 27 segments (from 14 students), our interpretation allowed us to assume the engineering context even though it was not overt.

F.1) Explicitly mentioned engineering context

As exemplified by the quote included for F.1 in Table 2, at least some students (5) came to the conclusion that AI is a relevant topic in engineering education since it will become a regular part of future engineering careers. Furthermore, once (9) students became aware of the dangers of AI, they found that it is crucial to be able to critically reflect on AI, as one student explained, “it was not only enjoyable, but also very important to learn more about tools that help to critically reflect on AI. This course has helped to shape future engineers that keep ethical values in mind when facing environmental and civil challenges with the help of AI” (s10, p. 1). One student also expressed that doing research on “implementing smart wearables in an ethical way in construction also allowed [them] to realize the

relevance of cultural aspects in the innovation process” (s5, p. 1), extending their acquired knowledge beyond the scope of the papers provided to them. Another student explained, “I now have a deeper appreciation for the importance of considering the ethical implications of AI and I am committed to using my skills and knowledge to contribute to the responsible and ethical development and use of these systems,” going on to express that our seminar provided an indispensable foundation for their “future career as an engineer working with artificial intelligence” (s16, p. 1).

F.2) Assumed engineering context

9 students stated that it is important to ensure that ethical principles are followed when implementing AI (s2; s3; s5; s6; s8; s14; s15; s16; s18). Students expressed the necessity of identifying the source and use of data (s4; s6) and acknowledged the importance of transparency and accountability (s8) as well as the need for regulations (s6). One elaborated that “the class on AI ethics has given [them] a greater understanding of the ethical considerations that must be taken into account as we continue to develop and integrate these technologies into our society. It is important to consider the potential harms and benefits of AI, and to ensure that these systems are transparent, accountable, sustainable and fair for all people” (s8, p. 3). Another student explicitly mentioned the value of understanding the intersectionality of AI and society (s2) while 4 students further stated that it is important to know about the potential dual-use of technologies since they demonstrate that technologies can have unintended impacts on society (s5; s12; s14; s19), revealing the relevance of ethics in the innovation process.

One student elaborated that “the use of AI [has] a significant impact on society through its use in various industries, such as construction, but also healthcare, finance, or manufacturing” (s16, p. 1). Since another student “learned that AI application[s have a] more significant impact on societies” than expected, they explained that “the implementation of AI has to be treated very carefully with significant consideration of social factors and impacts” (s6, p. 1). Moreover, 3 students referred to the significance of understanding normativity, cultural differences, and diversity (s4; s10; s16). They clarified that “social aspects absolutely need to be taken into account when AI-applications are used to solve societal issues” (s4, p. 1) and that “understanding and especially defining diversity is an important foundation in order to actually consider and respect the diversity of societies while making use of AI” (s10, p. 1).

Furthermore, “the course made [students] aware of how important it is to deal with the ethical concerns of AI in advance rather than after the event” (s8, p. 2). One student aptly summarized this need: “as AI technology is quickly advancing, it’s important to be aware of the potential for AI technologies to be used in unexpected ways, and to have a framework that can identify and weaken those risks” (s14, p. 2). Another added: “the course was an eye-opener for me and has given me a deeper appreciation for the potential and limitations of AI. I am confident that the knowledge and skills I gained will be valuable in my future endeavors” (s18, p. 1). In fact, several students echoed this sentiment in class, referring to the seminar as an “eye-opener” for them.

Discussion

Having presented the results of our qualitative analysis, this section will discuss the findings with respect to our RQ: *Why is education about the sociotechnical context of AI relevant to engineering students?*

Discussion on AI Is Now Fundamental in Engineering Yet Overlooked in Education

Evidence in support of a variety of AI applications (in engineering) can be found in our analysis of the examples that students had discovered (category C). While conducting literature research for examination tasks, students found several examples of AI used in civil engineering and environmental engineering. They also referred to examples discussed in class. Although this is far from a complete view of the AI applications in engineering, this small yet diverse sample of what our engineering students discovered within a single semester, shortly after being introduced to the use of AI in engineering fields beyond only software engineering, speaks to the prevalence of AI applications in engineering. Students further recognized that the use of AI will continue to grow in the future, including in engineering. The diverse examples of AI application areas that students mentioned, but that were not discovered as part of the examination tasks nor mentioned in class, also seemed to show that exposure to several AI applications in engineering could increase students' propensity to extend this awareness to novel applications.

Even so, students found it difficult to find suitable examples of AI applications in engineering despite the fact that there is a myriad of such examples. It may have been difficult for them since we asked them to find literature that discusses the ethical aspects or provides the information necessary for an ethical analysis of such an example, whereas this type of information may be more rarely published if collected at all.

Prior to the seminar, our students lacked knowledge about AI or awareness of AI applications (in engineering) despite AI's omnipresence. None of the students expressed having significant prior experience with AI while most had none at all. These findings align with scientific literature that presents AI-related modules as rare among engineering study programs (Heidling et al., 2019), even though, as we show in our analysis, they are now critical to engineering students' basic education.

It is worrying that students who should become experts in their respective fields are unaware of AI applications in their fields, especially since their lack of awareness will not prevent them from working with AI in their future careers (Cockburn et al., 2018; Peres et al., 2020). This points to a significant lapse in their education that requires redress. As much as academic literature demands updates to engineering education, it is time that the needed changes are finally implemented within all programs and courses. Not only do students lack an awareness of AI applications in engineering, but they also lack a basic understanding of how AI functions and of its capabilities, which are becoming required knowledge for all university students, regardless of their study program. AI can no longer be thought of as a magical black box, which takes some input and returns a perfect output—indeed, perhaps it should never have been so abstracted. Therefore, it is time for students to learn about the basic mechanics of AI, so that they understand what it is capable of, how it can be applied to their respective fields, and—once they learn about the sociotechnical context and ethical implications of AI—how it should be applied accordingly. Hence, the next aspects that must be addressed are students' lack of understanding regarding the interconnectedness of society and technology, as well as their misconception about the neutrality of technology, both of which will be discussed in the following section.

Discussion on Considering the Sociotechnical Context Is Crucial: AI Is Not Neutral

Prior to the seminar, students lacked knowledge about the societal and ethical contexts that are key to Responsible AI. No student reported to have engaged in any ethical deliberation before, and some were even surprised to discover that ethics is relevant not only with respect to the AI discussed in class, but also relevant to all (engineering) fields. These are concerning yet unsurprising findings (Martin et al., 2021). A lack of experience studying or considering the ethical aspects and sociotechnical context of AI, or engineering more generally, is problematic enough, but our findings suggest that this extends beyond a mere lack of practice. Most students were completely unaware of the applicability of ethics to engineering disciplines. Therefore, without having taken a course that explicitly taught them that ethics and the sociotechnical context matter to engineering—as well as practicing the application of such knowledge—they may have never known to consider anything outside their purely technical knowledge. It is imperative that students know and understand the impact that technology has on society and vice versa, and that they consider the ethical implications of how AI may be used in their fields throughout their careers. Considering the potential harm of AI consists of more than simply worrying about AI replacing jobs. (Future) engineers must strive to use technology that upholds our societal values and supports what we care about without perpetuating biases nor introducing new harms to society or the environment.

Furthermore, while some students acknowledged their prior awareness of the importance of ethics (or its subtopics, such as fairness and transparency), they had not reflected on how this declared importance should translate into action in practice prior to the seminar. Similarly, others were previously taught to prioritize financial and business goals without ethical considerations of societal impacts or, in other words, without considering the sociotechnical context. Moreover, several students discovered that before attending the seminar, they had had a limited or inaccurate conception of ethics and the impact of ethical considerations, which is in line with the literature (Martin et al., 2021). This vague awareness of the “importance” of ethical considerations in the absence of an exploration of what precisely is important about it and why it is important renders the sentiment meaningless. It is like acknowledging the “importance” of eating healthy food without considering what healthy food consists of and the potential impacts of eating healthily or failing to do so. Therefore, engineering education is missing this important link between a vague awareness of the importance of ethics and what it would mean to apply this awareness to ethical engineering. This would involve exploring what “ethics” actually consists of, the impact and relevance of ethics in all engineering domains, and how to implement these important ethical considerations in the context of engineering. Through this process, engineers should be educated to think ethically about *any* engineering decision or task; though the more practice students have—by implementing such exercises throughout engineering courses and curricula—the better they would be at incorporating this thinking in *all* engineering decisions and tasks.

Notably, our students were surprised to discover that moral questions about societal impact can arise from machines, which work with facts and numbers, supposedly devoid of morality. They may have expected that such numerical machines as AI can only give one output (or a range of “objectively correct” outputs) to any input, where the output must be just and not subject to questions of morality if it is based on facts. What our students did not realize before the course is that countless choices determine how any algorithm works, and it is these choices, along with the impacts of algorithms’ outcomes, that are subject to ethical speculation. Moreover, this necessarily entails the interconnection of so-called technical and social aspects of a technology (Bosen et al., 2023). For example, the “technical” aspects may include the technical architecture while the “social” aspects may refer to the societal impacts of every stage in the development of a technology; still, each architectural design

decision may impact society in some manner, just as society may guide the technical needs or limitations of a technology. Engineering students must learn that their work has indirect societal impacts beyond fixing the targeted problems. It is not uncommon for engineering students to believe that technology and society are two disjoint concepts (Faulkner, 2015). This points to a severe failure of engineering education, in turn leading to exponentially severe consequences in later engineering applications (Cech, 2014). Therefore, the interconnectedness of “technical” and “social” aspects and, hence, the sociotechnical context is a critical aspect that should be included in all engineering study programs, within all courses (J. E. Mitchell et al., 2021; Niles et al., 2020). When such sociotechnical aspects are not taught as part of “regular,” traditional engineering topics, it becomes much more difficult for students to later make the connections between the one mandatory engineering ethics course they may have taken and the topics of all the other engineering courses. Instead, the sociotechnical context and ethical considerations should be integral parts of every engineering topic such that when students learn about, for example, AI, they simultaneously learn about AI ethics and the sociotechnical context of AI.

During the seminar, students learned that, to their surprise, AI is *not* an inherently neutral technology, but rather often a biased one. This is an essential sociotechnical aspect to be aware of (Boyd & Crawford, 2012), yet the belief that AI is purely objective and unbiased is not uncommon. In fact, STEM (science, technology, engineering, and math) subjects have traditionally been regarded as morally neutral or even morally good (Martin et al., 2021). This perceived neutrality and superior objectivity is also reflected in technosolutionism, which refers to the assumption that a technical solution exists for every (societal) problem (Sloane & Moss, 2019). The danger in these assumptions—of the neutrality of technology and of technosolutionism—is perhaps most obvious whenever technology replaces humans in decision-making. Since technology is built by humans and based on data, it will contain the same biases of the humans that built it or the data that informed it. Even when a machine collects the data, a human has at some point dictated what the machine should look for. Even if a generative AI would generate other AI models, it would do so based on data selected by a human or based on conditions set by a human. Any technology can be traced back to a human designer, even if by degrees. Since technology does not exist in a vacuum, it cannot be separated from our biases unless specifically designed to do so. Even then, it is impossible to eliminate all biases entirely. Merely looking at the variety of fairness metrics and definitions gives us an idea of how impossible it is to completely eradicate unfairness, but we can strive to reduce it as much as possible. Furthermore, the ethical stances of concepts such as fairness evolve as society does, requiring constant updates to the use and design of technologies that impact and are impacted by these concepts.

Moreover, some students described AI as a “double-edged sword,” such that the same technology can be used in beneficent or maleficent ways, sometimes referred to as a “dual-use” technology (Brenneis, 2025). For example, an AI application (such as a hiring algorithm) could promote diversity and inclusion or, on the other hand (or edge of the sword), reinforce existing biases. Furthermore, as some of our students also explained, technology is not innately moral, and good intentions alone are not enough to justify technology; the sociotechnical context and potential impacts should always be fully evaluated. It should become common knowledge among students and professionals that every choice has advantages and disadvantages; so, too, every innovation will be a double-edged sword that will improve some aspects—which is usually the focus when considering its implementation—but will also likely come with risks. It is every moral agent’s job to know the full picture, including the potential risks, whenever making a choice to use a tool or implement changes to technology. Not only should this be applied when introducing new processes, but it should also be considered retroactively to longstanding processes that may not have been ethically evaluated or that may have been evaluated

under outdated or insufficient ethical standards. Furthermore, it is imperative that engineers understand that advancing technological innovation is not more important than ethical considerations and that innovation is never automatically morally justifiable.

Some students demonstrated that this newfound awareness regarding the lack of value-neutrality (the bias) of AI transferred to other engineering or innovation technologies. This seems to show that even a single well-designed course on AI can teach students to consider the sociotechnical context of other technologies, beyond that which they have already practiced analyzing. It should be noted, however, that we cannot say how long this effect will last since we only looked at reflection papers written at the end of the single-semester seminar. Including such sociotechnical considerations across the teaching of a variety of engineering topics would surely increase students' intuitive understanding and transference to further topics and fields (J. E. Mitchell et al., 2021; Pierrakos et al., 2019).

Discussion on Responsible AI Education Is Essential for Engineers

Engineering students' demonstrated interest in the seminar topics of AI in engineering and the sociotechnical context of AI implies a willingness to study the topics and their perceived relevance. More specifically, students expressed interest in the following topics: ethical principles for AI; considerations of fairness, diversity, and equity; applications of AI in diverse fields; the lack of value-neutrality of AI; sociotechnical aspects (the addition of so-called social aspects rather than purely technical ones); and the complexity of ethical discussions and tools that help to critically reflect on AI. Students concluded that AI is a relevant topic in engineering education since it will become a regular part of future engineering careers (Bühler et al., 2022). Furthermore, once students became aware of the dangers of AI, they found that it is crucial to be able to critically reflect on AI, thereby recognizing the relevance of AI ethics as well.

Students' interests raise the question of whether engineering curricula adequately reflect these concerns. If students can identify the importance of ethical principles, the sociotechnical context of technology, and of recognizing the potential harms as well as benefits of engineering decisions, why are these factors still only marginally integrated into engineering education? While a single course can increase awareness, research in educational psychology suggests that knowledge retention diminishes without reinforcement (Cepeda et al., 2006; Kang, 2016). Repeated exposure across diverse contexts both strengthens retention and supports the transfer of knowledge to new situations (Bjork & Bjork, 2011).

The sociotechnical dimensions of AI—and of engineering more broadly—remain underrepresented in curricula, despite increasing recognition of their importance (IEEE, 2019). Addressing this gap requires moving beyond the traditional technical focus to incorporate ethical, social, and cultural perspectives as integral components of engineering education. Without such integration, future engineers may innovate effectively from a technical standpoint but risk overlooking ethical responsibilities, with potentially significant societal consequences. Embedding these dimensions at the educational stage could foster more responsible innovation from the outset, reducing the need for costly corrective measures once harm has occurred (Floridi & Cowls, 2019; Mittelstadt, 2019).

Conclusion

In pursuit of answering our RQ—*Why is education about the sociotechnical context of AI relevant to engineering students?*—we showed that engineering students lacked knowledge about AI in engineering, despite the fact that AI is often used in engineering. Students also lacked knowledge about AI ethics and ethics in general, thus perceiving technology as value-neutral and unbiased. However, our course on the sociotechnical context of AI challenged this perception and made students reflect on the implications of technology in general. Finally, we demonstrated students' interest in the seminar topics and the perceived relevance of learning about the sociotechnical context of AI for their future careers as engineers. As one student aptly summarized: “this course has helped to shape future engineers that keep ethical values in mind when facing environmental and civil challenges with the help of AI” (s10, p. 1).

However, students did not expect ethical considerations or their application to real-world problems to be so complex. One student wrote, “at first glance, the explanations [of the ethical principles covered in the seminar] seemed logical and simple. However, if you look more closely at the principles – which we did in Examination Task 2 – the consideration and especially the application to a particular case was more complex and difficult than expected” (s12, p. 1). Yet another student explained that “it became clear that the [implementation] of these ethical principles is very different [for] each individual AI application and so is the assessment” (s20, p. 3). This shows how important it is to strengthen the intrinsic motivation for deeper ethical assessment through *character education* (Harris, 2008; Pierrakos et al., 2019).

Throughout this paper, we make several normative claims: firstly, students should learn about the basic mechanics of AI as it is becoming increasingly, ubiquitously relevant to engineering careers. However, that is insufficient. Beyond purely technical knowledge, students should learn about the sociotechnical context of AI and the associated ethical considerations, and this sociotechnical awareness should be expanded to all engineering domains. In order to do that, the link between a vague awareness of the “importance” of ethics and the actual implementation of ethical practices must be addressed and practiced, ideally across all courses, but at least in several courses on a variety of topics. Thus, the sociotechnical context and associated ethical considerations should be integral parts of learning about every engineering topic such that students develop a habit for this kind of responsible thinking that is generalizable to topics not covered in class. Crucially, future engineers must understand that innovation and technology are not value-neutral nor innately moral. Therefore, the sociotechnical context and potential impacts of any innovation or technology should always be carefully and fully evaluated throughout its lifecycle: during design, before implementation, and regularly after deployment. Students should also understand that technology can never be perfectly neutral—that is the nature of subjective and ever-evolving ethical standards—though this should not stop them from trying to reduce adverse effects as much as possible. Moreover, it is imperative that future engineers understand that advancing technological innovation is not more important than ethical considerations. The simplest way to fulfill all of these demands would be to update engineering curricula and individual courses to teach these aspects, particularly including the sociotechnical context of all engineering domains. We have shown that this essential sociotechnical knowledge is generalizable to other domains when learning about it within the context of AI—a universally important engineering domain regardless of a students' major—but the same demands for incorporating the sociotechnical context should be applied to all other engineering topics as well. Educators must therefore accept the responsibility of updating their courses and programs such that future engineers are adequately trained to work responsibly and ethically.

As with any study, our analysis has certain limitations. First, the rapid development of AI technologies poses a challenge to the timeliness of our findings. While the data was collected in 2023, AI knowledge has since increased, particularly due to the widespread use of generative AI tools like ChatGPT, Gemini, etc. However, recent iterations of the seminar suggest that students' AI knowledge remains largely confined to generative models, which do not yet represent the majority of AI applications in engineering. Second, the study's target group—students enrolled in an elective seminar titled Innovation & Diversity—likely introduces a selection bias. Nevertheless, this potential selection bias, by which students had special interest in the topics, demonstrates the gravity of the deficiency in knowledge regarding AI in engineering and the ethical considerations thereof even among interested and motivated students. Our results serve as an exploratory analysis in the context of Responsible AI in engineering. Further studies involving a broader range of engineering students (particularly at different universities) remain necessary to validate our results. Lastly, the content of the reflection papers presents a further limitation. The prompt questions were formulated positively, which may have shaped the students' responses; alternative framing with more critical questions might have led to different reflections. Furthermore, while there was no stylistic similarity suggesting the use of AI-generated content, the increasing prevalence of generative AI raises the possibility that some students may have relied on such tools for writing the reflection papers. Despite these considerations, our findings highlight a persistent gap in students' understanding of AI ethics, at a time when AI technologies were already widely known.

Future research should focus on refining methods for teaching AI ethics to engineering students, particularly in ways that foster long-term competency development. Studies have shown that learning about ethics is more impactful if integrated throughout the whole curriculum instead of taking a single course (J. Mitchell & Guile, 2021; Pierrakos et al., 2019). Longitudinal studies investigating whether students' ethical awareness and perceptions undergo lasting change beyond the seminar setting and timeframe remain to be explored. This is especially relevant when considering students' ability to transfer ethical knowledge to other domains of engineering practice. Given the growing prevalence of AI tools like ChatGPT, future iterations of the seminar should adapt to evolving AI literacy levels and critically assess how generative AI influences students' ethical reflections. An important next step would be to examine how different instructional approaches—such as case studies, interdisciplinary collaboration, or hands-on ethical decision-making exercises—can enhance Responsible AI competence in engineering education.

Acknowledgements

This work was funded by the Federal Ministry of Education and Research (BMBF) and the Ministry of Culture and Science of the German State of North Rhine-Westphalia (MKW) under the Excellence Strategy of the Federal Government and the Länder. We thank the students participating in the seminar for their thoughts and insights.

Declaration of Interest

The authors report there are no competing interests to declare.

Notes on Contributors

Marie Mirsch (née Decker), mathematician by training, is a research assistant and doctoral candidate at RWTH Aachen University. Her interdisciplinary work addresses intersectionality in algorithmic systems. She manages the RRI Hub, teaches responsible AI, and contributes to national and international projects within the ENHANCE alliance.

Sarah G. Moreno holds a BA in psychology from McGill University and a MSc in ethics of technology from RWTH Aachen University. She wrote her master's thesis on morally distressing news on social media while working as a teaching and research assistant at GDI. Through her work, Sarah hopes to improve responsible education and practices.

Ben Schultz holds degrees in Business Informatics and Computational Social Systems from RWTH Aachen University. His work focuses on ethical, human-centred design and practical steps toward responsible AI across academia, industry, and the public sector, aiming to develop innovative yet responsible AI solutions.

Univ.-Prof. Dr. Carmen Leicht-Scholten, political scientist, holds the Chair of Gender and Diversity in Engineering at RWTH Aachen University and directs the RRI Hub. Her research integrates gender and diversity perspectives into science and technology, fostering socially responsible innovation. She advises national and international projects, boards, and associations.

References

- Aqlan, F., & Nwokeji, J. C. (2018). Applying Product Manufacturing Techniques to Teach Data Analytics in Industrial Engineering: A Project Based Learning Experience. In *2018 IEEE Frontiers in Education Conference (FIE)* (pp. 1–7). IEEE.
<https://doi.org/10.1109/FIE.2018.8658588>
- Bairaktarova, D., & Evangelou, D. (2011). Development of Engineering Ethics Course. In *2011 ASEE Annual Conference & Exposition Proceedings*. ASEE Conferences.
<https://doi.org/10.18260/1-2--17773>
- Balakrishnan, B., & Tarlochan, F. (2015). Engineering students' attitude towards engineering ethics education. In *2015 IEEE Global Engineering Education Conference (EDUCON)* (pp. 16–22). IEEE. <https://doi.org/10.1109/EDUCON.2015.7095944>
- Bergen, J. P., & Robaey, Z. (2022). *Designing in Times of Uncertainty: What Virtue Ethics Can Bring to Engineering Ethics in the Twenty-First Century*. https://doi.org/10.1007/978-3-031-08424-9_9
- Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). Worth Publishers.
- Bombaerts, G., & Nickel, P. J. (2017). Feedback for relatedness and competence: Can feedback in blended learning contribute to optimal rigor, basic needs, and motivation? In *2017 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1089–1092). IEEE.
<https://doi.org/10.1109/EDUCON.2017.7942984>
- Bosen, J., Bernhard, S., Fauster, E., Decker, M., Lämmerhirt, M., & Leicht-Scholten, C. (2023). Engineering Society: The Role Of Intersectional Gender And Diversity Studies For A Sustainable Transformation On The Case Of Interdisciplinary Engineering Education. *Practice Papers*. Advance online publication. <https://doi.org/10.21427/Q05K-AW70>
- Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, *15*(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Brenneis, A. (2025). Assessing dual use risks in AI research: necessity, challenges and mitigation strategies. *Research Ethics*, *21*(2), 302–330. <https://doi.org/10.1177/17470161241267782>
- Breuer, U., & Genske, D. D. (Eds.). (2021). *Sachbuch. Ethik in den Ingenieurwissenschaften: Eine Annäherung*. Springer.
- Bühler, M. M., Jelinek, T., & Nübel, K. (2022). Training and Preparing Tomorrow's Workforce for the Fourth Industrial Revolution. *Education Sciences*, *12*(11).
<https://doi.org/10.3390/educsci12110782>
- Cech, E. A. (2014). Culture of Disengagement in Engineering Education? *Science, Technology, & Human Values*, *39*(1), 42–72. <https://doi.org/10.1177/0162243913504305>
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*(3), 354–380.
<https://doi.org/10.1037/0033-2909.132.3.354>
- Cetindamar, D., Kitto, K., Wu, M., Zhang, Y., Abedin, B., & Knight, S. (2024). Explicating AI Literacy of Employees at Digital Workplaces. *IEEE Transactions on Engineering Management*, *71*, 810–823. <https://doi.org/10.1109/TEM.2021.3138503>
- Clancy, R. F., & Zhu, Q. (2023). Why Should Ethical Behaviors Be the Ultimate Goal of Engineering Ethics Education? *Business and Professional Ethics Journal*, *42*(1), 33–53.
<https://doi.org/10.5840/bpej202346136>
- Cockburn, I., Henderson, R., & Stern, S. (2018). The Impact of Artificial Intelligence on Innovation. *National Bureau of Economic Research*. Advance online publication.
<https://doi.org/10.3386/w24449>

- Da Costa, G. M., Dzugan, J., Linney, S., West, J., & Philippou, G. (2019). *2019 Global Skills Gap Report*. QS.
- Decker, M., Schleiss, J., Schultz, B., Moreno, S. G., Stober, S., & Leicht-Scholten, C. (2024). Towards Responsible AI - Competencies for Engineers: An Explorative Literature Review On Existing Frameworks. In *Proceedings of the 52nd Annual Conference of SEFI*. <https://doi.org/10.5281/zenodo.14256815>
- Dignum, V. (2020). *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*. Springer. <https://link.springer.com/book/10.1007/978-3-030-30371-6>
- Dolata, M., Feuerriegel, S., & Schwabe, G. (2022). A sociotechnical view of algorithmic fairness. *Information Systems Journal*, 32(4), 754–818. <https://doi.org/10.1111/isj.12370>
- ElZomor, M., Pradhananga, P., Santi, G., & Vassigh, S. (2020). Preparing the Future Workforce of Architecture, Engineering, and Construction for Robotic Automation Processes. In *2020 ASEE Virtual Annual Conference Content Access Proceedings*. ASEE Conferences. <https://doi.org/10.18260/1-2--35082>
- Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), Official Journal of the European Union (2024).
- Faulkner, W. (2015). ‘Nuts and Bolts and People’ Gender Troubled Engineering Identities. In S. H. Christensen, C. Didier, A. Jamison, M. Meganck, C. Mitcham, & B. Newberry (Eds.), *Philosophy of Engineering and Technology. Engineering Identities, Epistemologies and Values* (Vol. 21, pp. 23–40). Springer International Publishing. https://doi.org/10.1007/978-3-319-16172-3_2
- Floridi, L., & Cowsls, J. (2019). A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review*. Advance online publication. <https://doi.org/10.1162/99608f92.8cd550d1>
- Godfrey, E. (2014). Understanding Disciplinary Cultures: The First Step to Cultural Change. In A. Johri & B. M. Olds (Eds.), *Cambridge Handbook of Engineering Education Research* (437–456). Cambridge University Press.
- Harris, C. E. (2008). The good engineer: Giving virtue its due in engineering ethics. *Science and Engineering Ethics*, 14(2), 153–164. <https://doi.org/10.1007/s11948-008-9068-3>
- Heidling, E., Meil, P., Neumer, J., Porschen-Hueck, S., Schmierl, K., Sopp, P., & Wagner, A. (2019). *Ingenieurinnen und Ingenieure für Industrie 4.0: VDMA: Hochschulen stehen vor gewaltigem Change-Prozess*. IMPULS-Stiftung.
- IEEE (2019). Ethically Aligned Design - A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems. *Ethically Aligned Design - a Vision for Prioritizing Human Well-Being with Autonomous and Intelligent Systems*, 1–294.
- Kang, S. H. K. (2016). Spaced Repetition Promotes Efficient and Effective Learning. *Policy Insights from the Behavioral and Brain Sciences*, 3(1), 12–19. <https://doi.org/10.1177/2372732215624708>
- Khakurel, J., Penzenstadler, B., Porras, J., Knutas, A., & Zhang, W. (2018). The Rise of Artificial Intelligence under the Lens of Sustainability. *Technologies*, 6(4), 100. <https://doi.org/10.3390/technologies6040100>
- Kuckartz, U. (2014). *Qualitative Text Analysis: A Guide to Methods, Practice and Using Software*. Sage Publications Ltd. <https://doi.org/10.4135/9781446288719>
- Kuckartz, U. (2019). Qualitative Text Analysis: A Systematic Approach. In *ICME-13 Monographs* (pp. 181–197). Springer International Publishing. https://doi.org/10.1007/978-3-030-15636-7_8

- Kuckartz, U., & Rädiker, S. (2019). *Analyzing Qualitative Data with MAXQDA*. Springer.
<https://doi.org/10.1007/978-3-030-15671-8>
- Lima, R. M., Mesquita, D., & Sousa, R. M. (2019). *Curriculum analysis process: Analysing fourteen industrial engineering programs*. <https://hdl.handle.net/1822/66284>
- Liu, S., & Xie, X. (2021). AI Quality Cultivation and Application Ability Training for Normal University Students. In *2021 7th Annual International Conference on Network and Information Systems for Computers (ICNISC)* (pp. 116–120). IEEE.
<https://doi.org/10.1109/ICNISC54316.2021.00030>
- Martin, D. A., Conlon, E., & Bowe, B. (2019). The role of role-play in student awareness of the social dimension of the engineering profession. *European Journal of Engineering Education*, 44(6), 882–905. <https://doi.org/10.1080/03043797.2019.1624691>
- Martin, D. A., Conlon, E., & Bowe, B. (2021). A Multi-level Review of Engineering Ethics Education: Towards a Socio-technical Orientation of Engineering Education for Ethics. *Science and Engineering Ethics*, 27(5), 60. <https://doi.org/10.1007/s11948-021-00333-6>
- Metcalf, J., Moss, E., Watkins, E. A., Singh, R., & Elish, M. C. (2021). Algorithmic Impact Assessments and Accountability. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 735–746). ACM.
<https://doi.org/10.1145/3442188.3445935>
- Mitchell, J., & Guile, D. (2021). Fusion Skills and Industry 5.0: Conceptions and Challenges. In B. Montaha (Ed.), *Insights Into Global Engineering Education After the Birth of Industry 5.0* (pp. 53–76). IntechOpen. <https://doi.org/10.5772/intechopen.100096>
- Mitchell, J. E., Nyamapfene, A., Roach, K., & Tilley, E. (2021). Faculty wide curriculum reform: the integrated engineering programme. *European Journal of Engineering Education*, 46(1), 48–66. <https://doi.org/10.1080/03043797.2019.1593324>
- Mittelstadt, B. (2019). AI Ethics – Too Principled to Fail? *SSRN Electronic Journal*. Advance online publication. <https://doi.org/10.2139/ssrn.3391293>
- Montáns, F. J., Chinesta, F., Gómez-Bombarelli, R., & Kutz, J. N. (2021). Data-driven modeling and learning in science and engineering. In B. Bischl, O. Guhr, H. Seibold, & P. Steinbach (Eds.), *Proceedings of Machine Learning Research, Proceedings of the First Teaching Machine Learning and Artificial Intelligence Workshop* (pp. 845–855). PMLR.
<https://doi.org/10.1016/j.crme.2019.11.009>
- Moreno, S. G., Decker, M., & Leicht-Scholten, C. (2024). Implementing AI Ethics Education: Engineering Competencies for Sustainable Societal Impact. In L. Bettaieb, Ajailia Nadia, J. Bennedsen, K. Edström, Singh, Mar, Nivan, R. Lyng, . . . F. Byström (Chairs), *20th International CDIO Conference*. Symposium conducted at the meeting of ESPRIT, Tunisia.
<https://cdio.org/knowledge-library/documents/implementing-ai-ethics-education-engineering-competencies-sustainable>
- Newberry, B. (2004). The dilemma of ethics in engineering education. *Science and Engineering Ethics*, 10(2), 343–351. <https://doi.org/10.1007/s11948-004-0030-8>
- Ng, D. T. K., Leung, J. K. L., Chu, K. W. S., & Qiao, M. S. (2021). AI Literacy: Definition, Teaching, Evaluation and Ethical Issues. *Proceedings of the Association for Information Science and Technology*, 58(1), 504–509. <https://doi.org/10.1002/prai.2.487>
- Niles, S., Roudbari, S., & Contreras, S. (2020). Integrating Social Justice and Political Engagement into Engineering. *International Journal of Engineering, Social Justice, and Peace*, 7(1), 57–69. <https://doi.org/10.24908/ijesjp.v7i1.13568>
- Peres, R. S., Jia, X., Lee, J., Sun, K., Colombo, A. W., & Barata, J. (2020). Industrial Artificial Intelligence in Industry 4.0 - Systematic Review, Challenges and Outlook. *IEEE Access*, 8, 220121–220139. <https://doi.org/10.1109/ACCESS.2020.3042874>
- Pierrakos, O., Prentice, M., Silverglate, C., Lamb, M., Demaske, A., & Smout, R. (2019). Reimagining Engineering Ethics: From Ethics Education to Character Education. In *2019*

- IEEE Frontiers in Education Conference (FIE)* (pp. 1–9). IEEE.
<https://doi.org/10.1109/FIE43999.2019.9028690>
- Pilario, K. E. (2024). Teaching classical machine learning as a graduate-level course in chemical engineering: An algorithmic approach. *Digital Chemical Engineering*, 11, 100163.
<https://doi.org/10.1016/j.dche.2024.100163>
- Polmear, M., Bielefeldt, A. R., Knight, D., Canney, N., & Swan, C. (2019). Analysis of macroethics teaching practices and perceptions in engineering: a cultural comparison. *European Journal of Engineering Education*, 44(6), 866–881. <https://doi.org/10.1080/03043797.2019.1593323>
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206–215.
<https://doi.org/10.1038/s42256-019-0048-x>
- Sætra, H. S. (2021). AI in Context and the Sustainable Development Goals: Factoring in the Unsustainability of the Sociotechnical System. *Sustainability*, 13(4), 1738.
<https://doi.org/10.3390/su13041738>
- Sarker, S., Chatterjee, S., Xiao, X., & Elbanna, A. (2019). The Sociotechnical Axis of Cohesion for the Is Discipline: Its Historical Legacy and its Continued Relevance. *MIS Quarterly*, 43(3), 695-A5. <https://www.jstor.org/stable/26848052>
- Schiff, D. S., Logevall, E., Borenstein, J., Newstetter, W., Potts, C., & Zegura, E. (2021). Linking personal and professional social responsibility development to microethics and macroethics: Observations from early undergraduate education. *Journal of Engineering Education*, 110(1), 70–91. <https://doi.org/10.1002/jee.20371>
- Schleiss, J., Bieber, M., Manukjan, A., Kellner, L., & Stober, S. (2022). An interdisciplinary competence profile for AI in engineering. In *Towards a new future in engineering education, new scenarios that european alliances of tech universities open up* (pp. 1601–1609). Universitat Politècnica de Catalunya. <https://doi.org/10.5821/conference-9788412322262.1288>
- Schleiss, J., Hense, J., Kist, A., Schlingensiepen, J., & Stober, S. (2022). Teaching AI competencies in engineering using projects and open educational resources. In *Towards a new future in engineering education, new scenarios that european alliances of tech universities open up* (pp. 1592–1600). Universitat Politècnica de Catalunya. <https://doi.org/10.5821/conference-9788412322262.1258>
- Silapachote, P., & Srisuphab, A. (2016). Teaching and learning computational thinking through solving problems in Artificial Intelligence: On designing introductory engineering and computing courses. In *2016 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)* (pp. 50–54). IEEE.
<https://doi.org/10.1109/TALE.2016.7851769>
- Sloane, M., & Moss, E. (2019). AI's social sciences deficit. *Nature Machine Intelligence*.
- Sucala, I. V. (2019). Mission (im)possible? Teaching social sciences to engineering students. In *SEFI Conference Proceedings 2019*.
- Venkatasubramanian, V. (2022). Teaching Artificial Intelligence to Chemical Engineers: Experience from a 35-year-old Course. *Chemical Engineering Education*, 56(4).
<https://doi.org/10.18260/2-1-370.660-130423>

Appendix A

Full List of Main Categories and Subcategories

The list of categories below depicts the full scale of themes discovered among the students’ reflection papers during our thematic analysis. Some categories have been explored in another paper (see Moreno et al., 2024).

Table A1. All Main Categories and Subcategories

Main Category (# Coded Segments)	Subcategory (# Coded Segments)
Course structure and methods (109)	Course organization and examination structure (28)
	Topics and tools (31)
	Class activities and examples (6)
	Groupwork (21)
	Assignments (22)
	Creative freedom (4)
	Improved skills (10)
	Overall evaluation / compliments (21) Constructive criticism / suggestions (14)
Expectations before the seminar (4)	N/A
Interest in the seminar topics (29)	N/A
Value-neutrality of technology (23)	N/A
Relevance of AI ethics (to a future engineer) (44)	Explicitly mentioned engineering context (17)
	Assumed engineering context (27)
Applications of AI in engineering (25)	Topics from examination tasks or class (17)
	New topics (8)
Previous experience with AI (15)	N/A
Previous experience with ethics (12)	N/A
Surprising (37)	Need for multi-disciplinarity in technical fields (8)
	Technology is not neutral (12)
	AI for future global (engineering) challenges (9)
	Applying ethical principles is difficult (4)
	Surprising groupwork (5) Other (2)
Misunderstandings (13)	N/A
Challenges (22)	Applying ethical principles (5)
	Engaging in AI-related topics (1)
	Finding relevant literature and choosing appropriate topic (9)
	Challenging groupwork (8)

Appendix B

Full Segments Cited in This Paper

Portions of segments coded from students' reflection papers composed direct quotes or supported indirect citations. To offer transparency and maintain scientific integrity where interpretation may influence analysis, all segments cited in the text are available below, in their full, original format, grouped according to the category section in which they were cited. For all coded segments, please see the dataset deposited in the Figshare repository available under the reserved DOI <https://doi.org/10.6084/m9.figshare.30621935>. The coded segments are organized by category, and all categories are included in the dataset, even ones not directly analyzed in this paper.

Table B1. All Full Segments Along with Citations and Associated Category

Category	Full Segment	Citation
A	I was worried about having to learn about AI and Machine Learning, something that at first seemed out of reach and very complex to understand.	2, p. 1
	At the beginning of the course, I didn't have the chance to read scientific publications or even to read more about AI and the ethical aspects that should be considered while not only developing it but also in the application. I did however parallel to this course had a course at the university where we got to see the latest developments and innovative technologies (some of them being AI-based) in the construction sector, therefore some of the technologies which we later chose for our paper were not completely new	5, p. 1
	having group members with different backgrounds makes it more engaging as we all are new to AI topic	18, p. 1
	Before enrolling in this course, I was under the impression that AI was primarily used for business and technology purposes. However, I was surprised to learn about its potential in environmental science with issues.	18, p. 1
	Despite the fact that AI systems are a huge part of our everyday life, I wasn't really aware, in which fields they are being used and most importantly, what impacts on our society – both positive and negative they could have.	19, p. 2
	Getting an introduction to AI was very important to me. I knew of AI, but I didn't have any technical knowledge about it. Getting to know what	20, p. 2

	is already AI-based in our daily life and the technical information with the different stages of the machine learning loop functioned as a perfect basis for the future group work.	
B	I have never heard of the ethical principles of AI before	1, p. 2
	so far we've only seen ethics playing an important part of machines and AI in movies and TV shows with extreme cases, for example, "I, Robot" and "Black Mirror" but it never occurred to me that Ethics could be applied to other sectors, least of all to construction.	2, p. 1
	Taking other aspects like social aspects into consideration and not just treating it from a technical viewpoint as an engineer has been a new experience. Discrimination in AI is more common than previously thought and its connection to diversity is equally as surprising. The dimension paper that we filled out during the discussion of lecture three for example showed new facets to the topic of diversity I previously didn't pay attention to and helped understanding the term diversity and its importance in the context of AI-models.	4, p. 1
	Initially, I had a narrow understanding of ethics and mostly thought of data privacy and safety when thinking about ethics.	5, p. 1
	I had not taken a comparable course like this in my Bachelor's degree. So it was interesting to see a different way of thinking or looking at engineering, technology and innovation. I had taken a course on innovation in the business sector. The course looked at factors such as economic dominance in the market and competitiveness rather than the ethical factors. Although I had learnt what constitutes an innovation and how it should be to be successful, I had not considered the actual value and added value for humanity. This includes the dangers and limits of innovation from a moral point of view.	12, p. 1
	Before attending this seminar, I must admit that I had little experience with the topic of ethical AI. While I was aware of its importance, I had not given much thought to the various ethical considerations that come into play when developing and using AI systems. However, after participating in this seminar, I feel much more sensitized to the topic and have a better understanding of the various ways in which AI can impact	16, p. 1

	society. Especially learning about tools which can be used as guidelines for an ethical analysis are helpful.	
	It is also the first time I heard about the ethical principles connected to AI. At the end of the lecture, I thought it would be easy to apply these to our chosen essay. But I was mistaken.	20, p. 2
C.1	i learned a great deal about ethics applied to construction, mostly related to Building Automation, which seems to be the future of the current facilities, with the centralized management of several aspects, like the heating, ventilation, alarms, lights and so on.	2, p. 1
	the most surprising thing was to know that AI can successfully predict earthquakes to a significant extent! Even one of the early adoptions of AI in disaster risk management (DRM) was Japan's public earthquake early warning system (EEW) which has been in operation since October 2007. During the 2011 Tohoku-oki earthquake (Magnitude 9.0) and the following tsunami several million people living near the epicenter were able to take action for the AI-based early warning system (Fujinawa & Noda, 2013).	3, p. 2
	The fifth session was mainly an in-session group work with a focus on the application of AI in civil and environmental engineering. Our group selected the example of smart mobility and sketched a concept of an AI algorithm that can redefine the experience of short-distance car sharing.	3, p. 2
	the use case our group chose for the examination task 2 and 3 – smart wearables in the construction industry to increase safety)	5, p. 1
	conducting research in the search for a use case for the civil industry (which was the industry we decided to focus on for our paper) gave me new ideas regarding the possibilities of optimizing processes either using data-driven technologies or rather in the training of data as in the case of wearables capable of being used to determine safe and unsafe zones.	5, p. 1
	The surprising part about the course is that I was able to find the real application of AI in civil engineering field. The course made me to research on the various applications of AI in civil engineering field which was kind of difficult to find since there were very few research work based on the same. The paper that our group has finalised dealt	7, p. 2

	<p>with the possible application of AI technologies in the construction industry. Before this course, I have not been aware of how AI and its applications. I believe that this kind of application can save time and reduce the costs and easy to deal with in today's world of civil engineering if applied in the right manner considering the Ethical principles.</p>	
	<p>I also appreciated the chance to investigate AI in disaster risk reduction</p>	<p>8, p. 2</p>
	<p>Finding a topic and paper to work on was a long process; we met and did individual research. Two times we changed the topic. We compromised on the third topic after the first two were rejected; one because it was too general, the other because it was a classical example and overdone. Researching and switching topics was an interesting experience, as we got to see a lot of use cases of AI in environmental and civil engineering. Our topic, ‘AI in weather prediction’ was not everyone’s preferred topic, but nonetheless, is was a good topic to work with.</p>	<p>10, p. 1</p>
	<p>For example, in flood vulnerability assessment, if training data sets are predominantly taken from urban areas, the AI model may not reflect the vulnerability of rural areas, which unfairly underrates their needs in the mitigation plan process and relief agencies (“Responsible AI for Disaster Risk Management,” n.d.). However, they can be the greatest weapons to fight global challenges if models are trained using a diversified set of data.</p>	<p>13, p. 2</p>
	<p>It was interesting to choose an example by ourselves, in our case the use of smart vests and smartwatches in construction sites. It was interesting to see how this technology would be able to improve safety, efficiency, and communication on construction sites. In this context, one of the most interesting aspects of this discussion was the potential for smart vests and smart watches to provide real-time monitoring of workers' safety and health, and to alert supervisors to potential hazards or injuries. The use of this technology could significantly improve the safety of construction sites, which is a critical concern in this industry. It was interesting to see that there are already considerations to use A.I. technology in the construction industry and it will even influence a part of my future.</p>	<p>14, p. 2</p>

	<p>It highlights how AI is changing the way we approach traditional civil engineering problems and provides new opportunities for innovation and optimization. The use of AI in civil engineering is making design and construction processes more efficient, accurate and cost-effective.</p>	18, p. 1
C.2	<p>despite being a relatively new technology there are already many successful implementations of AI in different aspects of our lives. From the recently published scientific literature, it is evident that, the use of AI in different domains will grow rapidly in coming years</p>	3, p. 2
	<p>One of the most interesting aspects of the course was learning about the various types of artificial intelligence and their applications. I found it fascinating to discover how AI can be used in so many different areas, such as healthcare, finance, and Disaster Management. It was also intriguing to learn about the potential benefits and drawbacks of using AI in these fields.</p>	8, p. 2
	<p>I also didn't know about the GPS technology used in safety equipment until the research. I was surprised that the technical implementation - i.e. a simple GPS monitoring system - was not as complex as expected and that safety can still be increased. I think almost everyone has a health app on their smartphone, consciously or unconsciously. At least on my smartphone, such an app was preinstalled. I thought such health apps were only good for motivation or for one's own interest. I was not aware that they could save lives.</p>	12, p. 1
	<p>The use of AI having a significant impact on society through its use in various industries, such as construction, but also healthcare, finance, or manufacturing. AI systems can be used to improve efficiency, accuracy, and decision-making in these fields, leading to potential cost savings and improved outcomes.</p>	16, p. 1
	<p>In this course, I explored various applications of AI in civil engineering, such as machine learning, monitoring, and the use of drones for surveying and mapping while reading various papers which was the interesting section where I came to know that AI has an important role in every fields.</p>	18, p. 1

	<p>the advancement in AI for computer vision and their application in various fields – AI-powered cameras and surveillance systems can automatically identify and track individuals, and self-driving cars rely on computer vision to navigate the roads.</p>	19, p. 2
	<p>AI systems are now even able to generate new music, design new buildings and to create new art pieces.</p>	19, p. 2
D	<p>as soon as we saw the TEDTalk that covered the biases in Facial Recognition Technology everything started to go smoothly, getting the chance to learn that technology can be biased and even exacerbate human prejudices (Barocas, S. et al, 2017) was an immense discovery, because personally i used to think that technology is just that, a machine, so why would it have bias towards anything if it doesn't even "think" by itself? This leaves you wondering not only in what other types of technology the same happens, but also what do we take for granted when try to apply and implement this new findings.</p>	2, p. 1
	<p>As an (ongoing) engineer, one tends to think about artificial intelligence as a simple tool to fulfil tasks. It is easy to perceive AI-models as equitable and impartial pragmatic instruments, however it is quite interesting to see that that's not the case at all. AI-modelling is a technique that mimics human behaviour, it is consequently also prone to human errors, biases and prejudice. Machine learning as a part of AI use statistical methods to learn from experience. Because it relies on the given data, such systems can easily be corrupted by the biases. Especially the responsibility gap was a factor that I previously didn't consider when thinking about artificial intelligence.</p>	4, p. 1
	<p>Discrimination in AI is more common than previously thought and its connection to diversity is equally as surprising.</p>	4, p. 1
	<p>As an engineer, I was always considering the technical aspect of it. Nevertheless, I got to know that the social aspect is significantly important, and it could be a key element in deciding whether AI should be implemented or not. Moreover, it was very surprising and indeed scary how discriminating AI could be.</p>	6, p. 1

	<p>It was fascinating to learn that "machines" or artificial intelligence (AI), which I had previously assumed to be error-free or rational, may make mistakes just like people do. Prior to use, the models must be tested to determine whether they are error-prone. Consequently, the employment of artificial intelligence should not be viewed as a tool with no problems.</p>	<p>8, p. 2</p>
	<p>I have learned that innovations or innovative technologies, or even the idea itself, is not automatically morally justifiable in all respects, even if the intentions are only well meant. The true purpose, the true added value of an innovation is the added value that is achieved for all those involved.</p>	<p>12, p. 2</p>
	<p>I have never put an emphasis on the integrity of AI and its impact on social settings, and always had an image of AI as perfect in most of its applications, till I took part in the course and shed more light on different research papers on responsible AI and unleash the reality that AI is still in its beginnings and needs to meet standards of FAST AND FACT principles in the field of Civil and Environmental Engineering to be feasible for use as it can directly affect lives of millions.</p>	<p>13, p. 2</p>
	<p>One surprising fact about this course is that it made me recognize that AI is a double-edged sword. It can actually be a weapon for inclusivity, creates opportunities for vulnerable groups, and brings their research and ideas to the table, improving impartial decision-making processes in the civil or environmental engineering field (Cristina Zaga, 2022). On the other hand, it is surprising that AI can exacerbate existing inequalities if they are not controlled and monitored regularly, which can put the needs of marginalized groups into question.</p>	<p>13, p. 2</p>
	<p>One of the course's most interesting aspects was the fact that AI is a double-edged sword. On the one hand, it was interesting to learn that AI can be used to promote diversity and inclusion. For example, AI can be used to screen job candidates in a way that is less biased than humans in the context of decision-making, and therefore be more representative of diverse perspectives. On the other hand, it was surprising to learn how AI systems can maintain and even reinforce existing biases, for example through the concept drift, and how important it is to consider the</p>	<p>14, p. 2</p>

	<p>potential consequences of these biases when designing and deploying AI systems. I had never given it much thought and the course made me aware of it.</p>	
	<p>One example that has been on my mind for the past couple of months is AI systems that can create unique and visually striking deep fake art by creating images from text prompts. At first, I mostly thought about the potential benefits this technology can deliver, e.g., advertisements and movies. However, there are also potential drawbacks to consider. The technology can be used to create misleading or fraudulent images, which could have serious consequences. Furthermore, the use of AI-generated images in art and media could impact the authenticity and uniqueness of art and question the general function of artists.</p>	<p>14, p. 2</p>
	<p>We tend to consider this emerging field as one strictly technical and analytical since its closest “relative” is computer programming and we have it clear that this only means giving commands to a machine to do something you want, albeit faster and more precisely.</p>	<p>15, p. 1</p>
	<p>One surprising aspect of the seminar was the realization of just how much the use of AI can impact different aspects of society. The use of AI having a significant impact on society through its use in various industries, such as construction, but also healthcare, finance, or manufacturing. AI systems can be used to improve efficiency, accuracy, and decision-making in these fields, leading to potential cost savings and improved outcomes. While it is certainly a powerful tool that can bring many benefits, it is important to consider the potential consequences of its use and ensure that it is not causing harm.</p>	<p>16, p. 1</p>
	<p>About AI particularly surprising for me was that problems arise from the collection of data alone - even if this data is not used.</p>	<p>17, p. 1</p>
	<p>But also in general, that with a machine, which operates on the facts and figures, moral questions arise. That machines (automatic systems) alone by the choice of the selected information a human/personal view is passed on, which can be discriminating.</p>	<p>17, p. 1</p>

	<p>I found it very interesting in the course, that we learned, how things, which were initially created to make our life easier, could also have a negative side in the example of the AI systems and devices.</p>	<p>19, p. 2</p>
	<p>My first takeaway, related to the use of AI in regard to its ethical side, would be the potential for bias. AI systems are not neutral – they are only as unbiased as the data they are trained on, and if the data contains biases, the AI will also be biased. This can lead to unfair and discriminatory outcomes, particularly when it comes to marginalized groups. (N. Mehrabi, 2021).</p>	<p>19, p. 2</p>
	<p>Motivation for the attendance of the seminar Innovation & Diversity came from news articles I’ve read in the past. As an example of a bad AI implementation the face recognition software in the USA comes to mind (Gentzel 2021). It’s racial accuracy disparities connected with racial profiling can be seen as a clear link between AI implementation and it’s devastating effect on human life (Gentzel 2021). In contrast, a week before the application deadline I read an article about an initiative, which uses AI and population data to target individual households who might have problems paying their rent to give assistance when eligible. I tried to find that article, but I was unsuccessful. That’s why there is no reference for it. The difference of outcome of these two AI applications made me think about the right way of implementing AI and if there are criteria for producing a good and ethical AI. So, I applied to Innovation & Diversity to find out.</p>	<p>20, p. 2</p>
	<p>The second aspect I was surprised about was the proxy concept and how many attributes can be intentionally or unintentionally be used as a proxy (Gajane 2017). Coming back to the face recognition example, where you can clearly see the connection between attribute and human effect the examples in the lecture showed how easily attributes could be used to officially not target protected classes, but you do it through the backdoor. It gave me a heightened awareness to look have a closer look at the data going into AI models.</p>	<p>20, p. 2</p>

E	As an environmental engineer, when I think about AI, civil and environmental engineering fields don't come to my mind directly. So, searching and reading different examples were quite interesting.	1, p. 2
	I have never heard of the ethical principles of AI before, therefore learning these and looking at AI from an ethical perspective was something that I enjoyed during the examination tasks.	1, p. 2
	Another topic of great interest were the FAST Track Principles and Ethics as a whole in AI	2, p. 1
	The most interesting thing to me about this course was learning to look at the term fairness from different perspectives. Something seemingly fair in common sense or from a statistical point of view might not be fair at all. Even after very careful, informed consideration it still might not be fair from a different undiscovered perspective. I got acquainted with this concept from the examples of the topic outcome fairness in the fourth session. Before that, I think, I had a simpler and somewhat naïve view to evaluate fairness which had way more probability of being incorrect. That is why the concept of fairness and using this concept to evaluate an AI algorithm was most interesting to me.	3, p. 2
	It is easy to perceive AI-models as equitable and impartial pragmatic instruments, however it is quite interesting to see that that's not the case at all.	4, p. 1
	The most interesting aspect of the course was the number of considerations that should be considered when analyzing the ethics of AI. Initially, I had a narrow understanding of ethics and mostly thought of data privacy and safety when thinking about ethics. However, throughout the course and through group discussions, I learned that there are many more aspects to consider, such as accountability, transparency, and social impact.	5, p. 1
	it was interesting to have an idea how to measure diversity and to understand the different diversity concepts such as Egalitarian Diversity Representative Diversity and Normic Diversity.	6, p. 1

	learning about the FAST & FACT principles that are defined by Leslie (2019) due to the responsibility gap is another interesting topic in the course.	6, p. 1
	I chose this course since the content it was dealing itself was interesting for me. Artificial Intelligence in Civil engineering field is something I was curious about and wanted to know more about the real applications.	7, p. 2
	One of the most interesting aspects of the course was learning about the various types of artificial intelligence and their applications. I found it fascinating to discover how AI can be used in so many different areas, such as healthcare, finance, and Disaster Management. It was also intriguing to learn about the potential benefits and drawbacks of using AI in these fields.	8, p. 2
	Topics discussed in the lecture broadened my horizon and understanding on fairness, discrimination, and diversity. I particularly enjoyed the parts on the different concepts of diversity and the discussion about diversity attributes that could be cause of discrimination. Understanding and especially defining diversity is an important foundation in order to actually consider and respect the diversity of societies while making use of AI.	10, p. 1
	I really appreciated getting to know the FAST and FACTS principles as very structured tools to evaluate ethics. They helped a lot with the homework tasks and have the potential to provide a guideline for future projects.	10, p. 1
	I would conclude that it was not only enjoyable, but also very important to learn more about tools that help to critically reflect on AI.	10, p. 1
	I had not taken a comparable course like this in my Bachelor's degree. So it was interesting to see a different way of thinking or looking at engineering, technology and innovation.	12, p. 1
	I also found the ethical consideration according to the FAST and FACT principles interesting. At first glance, the explanations seemed logical and simple. However, if you look more closely at the principles - which we did in Examination Task 2 - the consideration and especially the	12, p. 1

	<p>application to a particular case was more complex and difficult than expected.</p>	
	<p>The most interesting aspect of the Innovation & Diversity course is to delve into different ways in which diversity can give rise to better solutions and ideas to ensure an ethical AI, which are hopefully reflected in our group examination tasks.</p>	<p>13, p. 2</p>
	<p>The Innovation & Diversity course was an interesting exploration of the ethical problems regarding new AI technologies. The focus on the "FAST" and "FACT" principles was particularly intriguing, as it highlighted the importance of being fair, accountable, sustainable, transparent, accurate, and confidential in the development and deployment of AI systems.</p>	<p>14, p. 2</p>
	<p>One of the course's most interesting aspects was the fact that AI is a double-edged sword. On the one hand, it was interesting to learn that AI can be used to promote diversity and inclusion. For example, AI can be used to screen job candidates in a way that is less biased than humans in the context of decision-making, and therefore be more representative of diverse perspectives. On the other hand, it was surprising to learn how AI systems can maintain and even reinforce existing biases, for example through the concept drift, and how important it is to consider the potential consequences of these biases when designing and deploying AI systems. I had never given it much thought and the course made me aware of it.</p>	<p>14, p. 2</p>
	<p>It was interesting to choose an example by ourselves, in our case the use of smart vests and smartwatches in construction sites. It was interesting to see how this technology would be able to improve safety, efficiency, and communication on construction sites. In this context, one of the most interesting aspects of this discussion was the potential for smart vests and smart watches to provide real-time monitoring of workers' safety and health, and to alert supervisors to potential hazards or injuries. The use of this technology could significantly improve the safety of construction sites, which is a critical concern in this industry. It was interesting to see</p>	<p>14, p. 2</p>

	that there are already considerations to use A.I. technology in the construction industry and it will even influence a part of my future.	
	The dual use principle was especially interesting for me and therefore struck me for the whole course length. This principle highlights the potential for AI technologies to be used for both beneficial and harmful purposes, depending on how they are deployed and used in different applications. As AI technology is quickly advancing, it's important to be aware of the potential for AI technologies to be used in unexpected ways, and to have a framework that can identify and weaken those risks.	14, p. 2
	<p>What was most interesting in the course?</p> <p>Personally, the answer to this question has to be the philosophical approach to Artificial Intelligence. We tend to consider this emerging field as one strictly technical and analytical since its closest “relative” is computer programming and we have it clear that this only means giving commands to a machine to do something you want, albeit faster and more precisely. But with AI, there is so much that goes on while the model learns “by itself”, which for us only looks like a metaphorical black box from the outside.</p>	15, p. 1
	One of the most interesting aspects of the seminar for me was learning about the Fast Principles	16, p. 1
	I feel interested to learn new concepts of AI and how Artificial Intelligence in civil engineering are acted. AI has been a rapidly growing field in recent years and has been revolutionizing many industries, including civil engineering.	18, p. 1
	Since the entire civil engineering studies in my bachelor's degree was so technical, I really like the seminars of the GDI because they give you a different point of view on traditionally technical fields. AI and machine learning being one of those fields, with the emphasis on their affects on society and especially their ethicality it was again a new train of thoughts for me.	20, p. 2
F.1	when implementing smart wearables in an ethical way in construction also allowed me to realize the relevance of cultural aspects in the innovation process.	5, p. 1

	<p>I would conclude that it was not only enjoyable, but also very important to learn more about tools that help to critically reflect on AI. This course has helped to shape future engineers that keep ethical values in mind when facing environmental and civil challenges with the help of AI.</p>	<p>10, p. 1</p>
	<p>Considering my field of study (Engineering), I am almost certain that at some point during my career I will come in a deeper contact with this technology and I'm grateful this course has put the idea in me of not taking that work so easily and to give a fair amount of consideration to all the ethical principles I have come across during these past 4 months.</p>	<p>15, p. 2</p>
	<p>I feel much more informed and educated about the topic of ethical AI after attending this seminar. I now have a deeper appreciation for the importance of considering the ethical implications of AI and I am committed to using my skills and knowledge to contribute to the responsible and ethical development and use of these systems. So, I believe that this seminar has been a valuable learning experience that will be useful in my future career as an engineer working with artificial intelligence. Looking to the future, I plan to use the knowledge and skills gained in this seminar to help ensure that the AI systems I work on are developed and used ethically. I also hope to continue learning about this important topic and staying up to date on the latest developments in the field of ethical AI. Overall, this seminar has been a valuable learning experience and will be useful in my future career as an engineer working with artificial intelligence.</p>	<p>16, p. 1</p>
<p>F.2</p>	<p>Even though this principles seem like a good strategy to prevent unethical practices when developing a project, my takeaway is how to ensure that they're followed, because as convenient as they are, through the investigation for the essay based in our paper we noticed that so far people have trouble implementing them or they purposefully ignore them.</p>	<p>2, p. 1</p>
	<p>Overall, the course help greatly with the understanding of the intersectionality of Artificial Intelligence with innovation and society, understanding that innovation in technology not only is the creation of something new but aside from needing to be desirable, have feasibility</p>	<p>2, p. 1</p>

	<p>and be viable, it oughts to have a strong ethical base for its correct application in all fields.</p>	
	<p>it is imperative that we do not get carried away with the ubiquitous use of AI and overlook the negative consequences that may emerge. It is somewhat comforting that different school of thoughts about AI ethics are coming to light and researchers are working to ensure ethical, responsible implementation of AI.</p>	<p>3, p. 2</p>
	<p>Normativity as a way to view diversity and the in lecture three mentioned corresponding examples, as well as the eurocentrism ingrain the notion that social aspects absolutely need to be taken into account when AI-applications are used to solve societal issues.</p>	<p>4, p. 1</p>
	<p>AI applications need to be treated carefully, whether it’s for prediction modelling like in our chosen paper or in other applications. While doing our research, learning about the negatives of AI has been a new experience. Being aware of the way data has been gathered and understanding which factors influence the data gathering is equally, if not even more important than the usage or rather the end goal of the AI-application itself, as the output relies on the input.</p>	<p>4, pp. 1-2</p>
	<p>Initially, I had a narrow understanding of ethics and mostly thought of data privacy and safety when thinking about ethics. However, throughout the course and through group discussions, I learned that there are many more aspects to consider, such as accountability, transparency, and social impact. To my surprise data privacy is just one of the many sides that demands more attention in the development and application phase. The fast and fact principles are just some of the many groups of principles addressing ethics, however they manage to encompass several fields where the AI is still behind and not fully adequate</p>	<p>5, p. 1</p>
	<p>I think there is not enough data safety or there is a lack of data protection due to disadvantages of AI. The implementation of AI has to be treated very carefully with significant consideration of social factors and impacts. I learned that AI application has more significant impact on societies more that I expected.</p>	<p>6, p. 1</p>

	<p>The course also highlighted the importance of transparency and accountability in the development and use of AI. It became clear that there is a need for clear guidelines and regulations to ensure that AI is used in a responsible and ethical manner.</p>	8, p. 2
	<p>The course made me aware of how important it is to deal with the ethical concerns of AI in advance rather than after the event.</p>	8, p. 2
	<p>the class on AI ethics has given me a greater understanding of the ethical considerations that must be taken into account as we continue to develop and integrate these technologies into our society. It is important to consider the potential harms and benefits of AI, and to ensure that these systems are transparent, accountable, sustainable and fair for all people.</p>	8, p. 3
	<p>Understanding and especially defining diversity is an important foundation in order to actually consider and respect the diversity of societies while making use of AI.</p>	10, p. 1
	<p>I think in the end there won't be one product that provides the perfect solution for everything as is mostly the case. It remains a matter of trade-off, but it could be a matter of life and death.</p>	12, p. 1
	<p>The dual use principle was especially interesting for me and therefore struck me for the whole course length. This principle highlights the potential for AI technologies to be used for both beneficial and harmful purposes, depending on how they are deployed and used in different applications. As AI technology is quickly advancing, it's important to be aware of the potential for AI technologies to be used in unexpected ways, and to have a framework that can identify and weaken those risks.</p>	14, p. 2
	<p>Based on (hopefully not) naïve optimism and the investigated models during the course, I would assume that most of the AI developers have an altruistic goal with these models, and therefore are made out of good intentions, among other interests. But that will be my greatest takeaway, that there is an incredible amount of information on the ethics that should go hand in hand with the technical knowledge</p>	15, p. 1
	<p>I have realized that since taking this course, I have become an unintentional activist about the ethics of AI, since the conversation about its powers and promises comes quite frequently in my social circles and</p>	15, p. 2

	<p>since we as humans like to share any new knowledge we have on any topic. I hope this lasts well after this course is over, both for my personal and professional life.</p>	
	<p>One of the key benefits of ethical AI is that it helps to ensure that these systems are being used for the benefit of society and not causing harm or leading to negative consequences. For example, ethical AI can help to prevent AI systems from perpetuating biases or discrimination, which can have serious consequences for marginalized groups. Another important aspect of ethical AI is that it helps to build trust in these systems. As AI becomes more integrated into various aspects of our lives, it is important for individuals to trust that these systems are being used responsibly and ethically. Without trust, there may be resistance to the use of AI, which could limit its potential benefits.</p>	<p>16, p. 1</p>
	<p>The use of AI having a significant impact on society through its use in various industries, such as construction, but also healthcare, finance, or manufacturing. AI systems can be used to improve efficiency, accuracy, and decision-making in these fields, leading to potential cost savings and improved outcomes. While it is certainly a powerful tool that can bring many benefits, it is important to consider the potential consequences of its use and ensure that it is not causing harm. By ensuring that AI is being developed and used in a responsible and ethical manner, we can maximize its potential benefits and minimize any negative consequences. Therefore, is important to be sensitized to the topic and include ethical considerations when it comes to analysis if AI may be an optimization for a specific use case.</p>	<p>16, p. 1</p>
	<p>the ethics of AI and its Fast and Fact principles are important to consider and study in the course of AI development and deployment. These principles provide guidelines for responsible and ethical use of AI, ensuring that the technology is used for the benefit of society and not to cause harm. Understanding the ethical implications of AI is crucial in shaping its development and deployment in a responsible and sustainable manner.</p>	<p>18, p. 1</p>

	<p>Overall, the course was an eye-opener for me and has given me a deeper appreciation for the potential and limitations of AI. I am confident that the knowledge and skills I gained will be valuable in my future endeavors.</p>	<p>18, p. 1</p>
	<p>I found it very interesting in the course, that we learned, how things, which were initially created to make our life easier, could also have a negative side in the example of the AI systems and devices. Thus, we as humans should be aware and morally accountable for the negative consequences, which they could bring.</p>	<p>19, p. 2</p>