Visions of a Discipline: Analyzing Introductory AI Courses on YouTube

Severin Engelmann
severin.engelmann@cornell.edu
Cornell Tech, Digital Life Initiative
New York City, NY, USA

Madiha Zahrah Choksi
mc2376@cornell.edu
Cornell Tech, Digital Life Initiative
New York City, NY, USA

Angelina Wang
angelina.wang@princeton.edu
Princeton University, Computer Science
Department
Princeton, NJ, USA

Casey Fiesler
casey.fiesler@colorado.edu
University of Colorado Boulder, Information
Science Department
Boulder, CO, USA

ABSTRACT

Education plays an indispensable role in fostering societal well-being and is widely regarded as one of the most influential factors in shaping the future of generations to come. As artificial intelligence (AI) becomes more deeply integrated into our daily lives and the workforce, educational institutions at all levels are directing their focus on resources that cater to AI education. Yet, informal education, including online learning on social media platforms like YouTube, plays an increasingly significant role for both students and the general public. Offering greater accessibility compared to formal education, millions of individuals use YouTube for educational resources on AI today. Due to the substantial societal impact of AI, it is crucial for introductory AI courses to meaningfully address the ethical implications associated with AI. Our work investigates the current landscape of introductory AI courses on YouTube, and the potential for introducing ethics in this context. We qualitatively analyze the 20 most watched introductory AI courses on YouTube, coding a total of 92.2 hours of educational content viewed by close to 50 million people. We find that these introductory AI courses do not meaningfully engage with ethical or societal challenges of AI (RQ1). When *defining and framing AI*, introductory AI courses foreground excitement around AI's transformative role in society, over-exaggerate AI's current and future abilities, and anthropomorphize AI (RQ2). In *teaching AI*, we see a widespread reliance on corporate AI tools and frameworks as well as a prioritization on a hands-on approach to learning rather than on conceptual foundations (RQ3). In promoting key *AI practices*, introductory AI courses abstract away entirely the socio-technical nature of AI classification and prediction, for example by favoring data quantity over data quality (RQ4). Given the power of openly available introductory courses to shape enduring beliefs around AI and its field at the onset of a learning journey, we extend our analysis with recommendations that aim to integrate ethical reflections into introductory AI courses. We recommend that introductory AI courses should (1) highlight ethical challenges of AI to present a more balanced perspective, (2) raise ethical issues explicitly relevant to the technical concepts discussed and (3) nurture a sense of accountability in future AI developers.

CCS CONCEPTS

• **Social and professional topics** → **Computing education**; **Informal education**.

KEYWORDS

Computer science education, ethics in computer science education, artificial intelligence ethics, accountability in computer science

1 INTRODUCTION

Artificial intelligence (AI) has become an enduring fixture in the eye of the general public. Daily media reporting on AI often oscillates between two extremes: powerful AI tools that transform how we live, work, and interact with each other, and, on the other hand, doomsday scenarios as a result of AI development [17, 66]. Recent major regulatory frameworks on AI also tend to define AI in balancing terms around benefits and risks [11, 61]. Beyond media and policy, *education* is a key contributor to society's knowledge on AI. Education effectively defines and fosters core values around AI, and influences how society perceives AI as an academic discipline. The recent advances in machine learning (ML), deep learning (DL), and generative AI underscore the growing importance of AI education in equipping future AI developers with an acute awareness of the potential risks and adverse impacts resulting from transformative AI systems and applications. However, not all learning takes place in a classroom, and many people are seeking out educational resources about AI *online*.

In this work, we explore how popular introductory AI courses on YouTube present AI to a broad audience. These introductory courses offer unparalleled accessibility to basic AI skills, responding to a growing demand for AI competence in government, industry, academia, and civil society [2]. YouTube's introductory AI courses exemplify informal online education, eliminating the need for enrollment in study programs such as those available on massive open online course platforms (MOOCs) [40]. On YouTube, educational content creators showcase their coding process by sharing screencasts, guiding viewers through coding problems step by step [55]. Such video-based tutorials are particularly valued by developers who, more than learners in other disciplines, have accustomed to a culture of online self-didactic learning [46, 59].

We focus on introductory AI courses as dynamic spaces for learning where educators *introduce* fundamental definitions, concepts, and practices around AI [4, 22, 23]. In introductory courses, beginners learn what problems the discipline considers to be within its own scope. They introduce the problem-solving approaches that hold value and establish criteria for determining when a problem is considered satisfactorily solved. Such foundational teachings embody some of the core values of a discipline [80]. In contrast to more advanced students, beginners tend not to question the teaching authority of the instructor and reflect less on the perspectives and assumptions presented in the learning materials [21]. AI introductory courses have the potential to leave a lasting impact on students' perceptions of AI and its disciplinary field.

The introductory AI courses we found on YouTube came from reputable universities, professional educational organizations, and education influencers with large followings. Considering their influential position, introductory AI courses provide a valuable opportunity to cultivate ethical reflection regarding core AI practices right from the start of students' learning journeys. While there is a large corpus of research on computer science and AI education in traditional educational settings (i.e., schools and universities) [20, 44, 68], and how to best embed ethical considerations into these study programs [30, 38, 47], less attention has been paid to AI education on YouTube. Our research qualitatively investigates introductory AI courses on YouTube to answer the following four research questions:

RQ1. Do introductory AI courses on YouTube cover ethics-related topics such as fairness, privacy, or accountability?

RQ2. How do introductory AI courses on YouTube define and frame AI?

RQ3. What tools and exercises do introductory AI courses on YouTube focus on?

RQ4. What methodological practices and competencies do introductory AI courses promote?

We find that introductory AI courses rarely address ethical or societal harms caused by AI technologies (RQ1). Introductory AI courses promote excitement and hype around AI framing its capabilities by reference to various forms of anthropomorphism (RQ2). Introductory AI courses prioritize hands-on coding at the expense of more conceptual understandings that might nurture a broader perspective of AI's situated uses (RQ3). Finally, introductory AI courses downplay the value of *data competence.* They largely neglect essential skills for appropriate data collection, data curation, and data annotation (RQ4). By focusing almost exclusively on model building and performance they also undervalue the role of domain expertise in prediction and classification tasks. In summary, our study provides evidence that introductory-level AI education on YouTube *promotes AI practices that have been identified as significant contributors to AI-related harms.*

Based on these findings, we formulate recommendations for integrating ethical reflections into introductory AI courses on YouTube in the final part of this work. Teaching ethics in the context of digital technologies is challenging as it requires substantial knowledge and expertise in ethics and computer science scholarship. Prior work on computing ethics in higher education points to the importance of both ethics integration (in-situ ethics learning in technical courses, on the concepts relevant to that course) and standalone courses taught by subject matter experts [31], though even integration requires some knowledge of the topic and presents some barriers for instructors [82]. With this in mind, our recommendations are aimed at *promoting* an ethical mindset at the beginning of the learning process. Our recommendations do not necessarily require in-depth ethics subject matter expertise. They prescribe to common theories of ethical interventions around *exposure* and challenge dominant views of computing [69]. Our recommendations can be integrated into any introductory AI class independent of the disciplinary background of the instructor. Our contribution is complementary to the important progress that has been made in embedding ethics into computer science curricula in the last years [30, 31, 38, 47]. With our recommendations, we seek to support instructors in projecting a more balanced picture of AI and help nurture a sense of accountability among future AI developers in building increasingly powerful AI systems.

2 BACKGROUND

2.1 Online learning

Prior to the COVID-19 pandemic, online distance learning was primarily viewed as a supplement to traditional forms of education [5]. It facilitated convenient exploration of new disciplines and enabled learners to sustain knowledge for career-related purposes. During the COVID-19 pandemic when learners around the world could not attend their educational institutions, online distance learning became the de facto mode of education [48, 90]. Entire curricula were converted into online programs to prevent learning gaps that could have led to significant inequalities in educational outcomes for millions of students worldwide [84].

Shifting to online formats during COVID-19 legitimized an already existing infrastructure of online learning media as educational programs that can complement, extend, and in some cases even replace conventional education [1, 81].

Learners around the world participated in these online videos through massive open online courses (MOOCs) which are produced by universities, educational organizations, and companies hosted on commercial online platforms such as Udacity or Coursera [8, 40]. MOOCs are closely related to the YouTube education videos we study, and both are purported to provide an avenue of education that is more accessible than traditional universities. However, it is contested whether and to what extent MOOCs do democratize education [25, 41, 51, 76], though work has found that learners who do not have a college degree are more likely to do better on a MOOC than learners with a college degree [63, 64]. MOOCs enrollment is often motivated for career building rather than knowledge building [91], and they are typically on a far shorter time scale than a university degree. Additionally, while YouTube does not facilitate social learning (e.g., group assignments) to the same extent as some MOOCs, it does enable interaction with education content creators and other learners through questions and answers in video comments [52].

Using online resources for advancing competencies is particularly popular in disciplinary cultures that value and reward self-didactic approaches to learning. One of these disciplines is software development in computer science. In 2021, Stack Overflow — a popular learning and knowledge sharing platform among software developers — ran a large-scale survey with more than 80,000 software developers across different industries [67]. When asked *How did you learn to code? Select all that apply*, the results were: 59.5% "Other online resources (videos, blogs, etc)", 40.4% "Online Courses or Certification" and 30.6% "Online Forum."[1] Online coding tutorials are common at every stage of the learning process for computer science students. Four out of the top ten most "enrolled" MOOCs of all times are computer science and programming courses [24]. Another factor contributing to the popularity of online learning among developers is the consistent demand for new competencies fueled by continuous technological advancements. YouTube is an especially popular learning resource among coders and developers as content creators can share their screen as screencasts to walk viewers through a coding problem [60]. This allows educators to illustrate specific coding skills, showcase code documentation and debugging, and explain error messages in relation to the source code. In introductory coding courses, educators can demonstrate how to set up an integrated development environment for coding, which can be challenging for beginners and a drawback of text-based coding tutorials. Moreover, education content creators on YouTube can share personal experiences learning specific competences and recommend what they believe are effective learning strategies [19].

All of this is to say that, increasingly, online learning takes place on social media platforms such as YouTube [54] or even TikTok [33] where learners search for and easily access educational content in a familiar interface. For comparison, enrollment in MOOCs require account registration and, in many cases, payment after a "freemium" period.

The barrier to upload content on YouTube is lower than for MOOCs platforms. Among other effects, this diversifies the community of education content creators. In addition to universities[2] and professional educational organizations (e.g., bootcamps and certification providers) a plethora of education influencers regularly publish learning materials on YouTube [72]. Some social media education influencers have millions of followers and have become "mini celebrities" [19].

On YouTube, this results in a diverse array of video-based learning formats, encompassing lectures, tutorials, walkthroughs, guides, vlogs, and informational videos [52]. Education content creators have the freedom to teach as they prefer. Consequently, video lengths vary, ranging from concise tutorials lasting a couple of minutes to more extensive ones that extend over several hours [60, 71]. The following subsection traces relevant works on the education of AI topics considering learning formats beyond only online learning.

2.2 AI Education

In recent years there has been growing attention to best practices for AI education, from college courses [3] to middle school [57] to informal public education [58]. Pedagogical strategies for introducing AI has been a topic of research and discussion for decades, and as early as the 1990s, there was recognition of "the oft-voiced complaint" that introductory AI is a difficult topic to teach [89]. A 2004 survey of introductory AI instructors revealed that part of this problem stems from the difficulty in keeping up with the latest research trends, resulting in teaching outdated content [42]. This survey also did not reveal ethics as a component of these instructors' courses, and similarly, a 2021 literature review of AI education had very little mention of ethics [3]. However, a recent synthesis of AI education literature towards defining components of AI literacy lists ethics as a core competency, describing many of the core commitments of the FAccT community such as fairness and transparency [58]. As Ng et al. pointed out in their review of AI literacy literature specifically, knowing how to use AI for future careers is only one relevant aspect, and an understanding of ethical issues is essential for ensuring AI for social good [65].

Ethics interventions in computing education have been growing in number, though it is worth noting that "ethics" is often used as shorthand for related concepts such as responsibility, social impact, and justice [16]. A recent review of ethics interventions in computing education also revealed a number of shared goals, including exposure (which assumes a lack of existing awareness of relevant ethical issues, an assumption that is supported by prior work), positioning students as practitioners who will be accountable for their decisions, and challenging dominant or common views of technology [69]. There are also arguments for including ethical concepts explicitly in introductory computing courses, including that it helps to create a "culture" around ethics as a component of technical practice, and it also ensures that

---

[1]The remaining responses were: 53.6% "School", 51.5% "Books & Physical Media" and, with less than 20%: "Friend or family member", "Colleague", and "Coding Bootcamp."

[2]Some universities transfer entire course lectures to YouTube. For example, UC Berkeley uploaded lecture courses in bioengineering and conflict studies to YouTube as early as 2008 [32].

even students who take only one lower-division course are exposed to ethics [30]. One intervention in an introductory AI course that introduced "human impact" as a necessary constraint in an algorithm assignment had similar goals, and found that this strategy had the potential to deepen technical rigor while also supporting instructor pedagogical content knowledge [15].

Though in recent years discussion of these types of integrative ethics interventions have increased [37], there are barriers to instructors in doing so [82] and even initial stages such as identifying an appropriate ethical context for assignments can be a challenge [14]. A 2018 review of syllabi for AI, machine learning, and data science courses revealed that only 12% of such courses included any (generously defined) mention of ethics-related concepts [77]. However, when raised, common relevant topics in AI courses include bias, privacy, and unintended consequences [34].

Despite the potential benefits for teaching ethical concepts in-situ alongside relevant AI technical topics, and of teaching applied ethics in computer science departments, it is critical to keep in mind as articulated by Raji et al. [73] that the "myth of the ethical unicorn" in computer science accompanies a devaluing of social science and humanities. The "exposure" theory may be a good starting place (particularly coming from nothing), but AI education also needs to include engagement with humanistic fields and learn from deep subject matter experts outside of technical computer science. One strategy beyond pointing students to separate courses is for instructors from different disciplines to work collaboratively to create and deliver educational content–for example, philosophers creating computing ethics modules [39] or NLP and accessibility experts working together to teach students about AI fairness [26]. However, it is important that in such interdisciplinary approaches technical skills are not privileged and valued above ethical skills [37, 73].

Additionally, beyond formal education, recent projects aimed at the general public have also presented the important goal of increasing legibility of AI for non-experts [45], emphasizing that even a casual understanding of AI (understanding how a system works as opposed to how to build one) is growing in importance [58]. Our work here was inspired in part by this acknowledgment of the importance of informal AI education, as well as the potential benefits for including ethics as part of introductory courses.

3 METHODS

3.1 Sampling Introductory AI You-Tube Videos

We sourced English-language introductory AI/ML videos on YouTube. We approached our search for introductory AI courses by envisioning how an individual eager to start learning AI would likely try to find relevant educational resources on YouTube. We employed a keyword-based search on YouTube for our selection criteria as educational tutorials tend to specify "intro to AI" (or equivalent descriptives) in their title. For our queries, we used the following keywords: *introduction to AI, introduction to artificial intelligence,*

introduction to ML, introduction to machine learning, introduction to DL, introduction to deep learning, beginner course AI, beginner course artificial intelligence, beginner course ML, beginner course machine learning, beginner course DL, beginner course deep learning. For our sample of introductory AI videos on YouTube, we followed the search procedures outlined in [46]. This included utilizing a computer that had never been used before for the purpose of searching introductory AI courses on YouTube. To minimize any potential effects of personalization, the search was performed in private browser mode and after deleting cookies. For each query, we ranked the results by most viewed and started selecting videos in accordance with the following criteria. We specified that videos had to cover at least one technical concept, such as basic AI algorithms, parts of the machine learning pipeline, or relevant concepts (e.g., a perceptron). This requirement meant exclusion of videos that introduced AI in a non-educational context, for example, videos related to best learning approaches (e.g., "The best way to learn AI in 2024") or comparisons like "AI vs. ML." Moreover, videos had to be at least 15 minutes. Our final dataset includes the 20 most viewed introductory AI courses on YouTube. Considering the extensive collection of search results, we stopped sampling when the next most viewed video had less than 1 million views, which was at N=20.[3] We present key information of the videos in our final sample in Table 1.

3.2 Coding Process

Our methodology combines both deductive and inductive approaches to qualitative coding. After gathering our final corpus of introductory AI courses, we generated full transcripts for each introductory AI course using *youtubetranscript.com*, a free tool that uses Merlin AI to transcribe YouTube videos. Coding our sample of YouTube videos posed unique challenges due to substantial variations in course length, ranging from 15 minutes to over 10 hours. Consequently, the topics covered in these courses exhibited significant diversity, with shorter courses focusing on specific technical concepts, while longer courses covered a broader array of AI concepts and coding examples. Another challenge in our coding process was the diverse structure of courses, as instructors had complete agency in organizing their content. Some courses began with an introduction to AI, while others immediately delved into coding examples. We thus opted to code colloquial sentence structures, dealing with highly unstructured data compared to more prompted interview studies. We employed Taguette,[4] a web-hosted collaborative and open-source tool for qualitative coding.

We followed best practices for coding qualitative data [10, 50, 62]. Applying researcher triangulation [70], three interdisciplinary researchers separately assessed V1 (university) to inductively identify first emerging patterns and trends. Researchers then met to present and discuss their initial codes. This resulted in a consolidation of a first set of codes (e.g., codes **High-level AI definition** and **Excitement around AI**). This process was repeated for V2 (influencer) and V3

---

[3]Note that YouTube's interface does not show search results in page tabs but in continuous scroll.
[4]https://www.taguette.org/

(online education organization), with discussions on and iterations to the set of codes after each video. Researchers met three times to discuss the coding scheme and made adjustments given the structural variability of introductory AI videos in our sample (e.g., length, content publisher). The final codebook can be organized around three major themes: definitions and framing of AI, teaching resources and styles, and practicing AI. This organization is based on high-level themes reflecting the descriptive and evaluative observations about specific elements within the courses. We present the final codebook with 21 codes in Table 10 in the Appendix. Ongoing discussions among the research team further shaped the codebook, with occasional decisions to merge codes, especially when addressing general approaches to data and data collection practices. For example, researchers removed the code **Data Privacy** which was collectively coded 3 times within 20 videos. It is crucial to note that instances of disagreement during these discussions are documented for transparency. The remaining videos in our sample were coded deductively using the established codebook. Taguette allows coders to apply multiple codes to a single textual unit, accommodating the nuanced nature of the content.

After the third discussion session (i.e., post coding V3), once consensus on a codebook was reached, three of the authors independently coded V4 (influencer, approximately 4 hours) to calculate Inter-Rater Reliability (IRR) scores. The obtained IRR measures for this coding task reflect strong agreement among the research team. Specifically, the Fleiss Kappa coefficient, which assesses agreement beyond chance among multiple raters was calculated at 0.873. Additionally, Cohen's kappa, a statistic that considers agreement between two raters while accounting for chance calculated at 0.876. Krippendorff's alpha was calculated at 0.877. These multiple reliability indices affirm high agreement and consistency in our qualitative coding.

It is also worth noting the perspective presented by McDonald et al. that finding informal consensus through discussion meetings suffices to communicate reliability, particularly in the context of highly unstructured data and multi-label coding [62]. This perspective aligns with our methodology, emphasizing the importance of discussion and consensus-building in ensuring the reliability of our qualitative analysis. Finally, we acknowledge our positionality relative to the study and provide ethical considerations in Section 7, which is presented at the very end of this work.

4 RESULTS

4.1 RQ1: Do introductory AI courses on YouTube cover ethics-related topics acknowledging that AI can cause harm?

4.1.1 Ethics plays no significant role in introductory AI courses on YouTube. First, we conducted a keyword search of ethics-related concepts in the transcripts of introductory AI courses. We employ the word stems of the following core

ethics-related concepts: *ethics, privacy, fairness, accountability, transparency, responsibility, bias, explainability, discrimination.*[5] We selected these ethics-related concepts given their prominence in AI ethics education (e.g., [34]) and in the FAccT literature (e.g., [53]). In Table 2, we show the prevalence of these ethics terms across introductory AI courses on YouTube.

The majority (17) of introductory AI courses in our corpus did not mention or otherwise meaningfully address "ethics." Of the 3 courses that did mention "ethics" instructors did not go beyond general statements such as "people are doing work on AI ethics" (V12, V15) or "humans need to behave ethically" (V11) to support machines solving people's problems. Only V15 mentioned "privacy," "fairness," "responsibility," and "bias" when introducing AI (1/20). Apart from V15, a course introduction by an influencer that offered a more in-depth overview of the AI field in general, ethical and societal considerations of AI technology did not play a significant role in introductory AI courses on YouTube. In V10, the instructor emphasized the importance of "transparency," stating that the mathematical notation for linear regression is easy to grasp. In V17, the instructor stated explainability of deep learning models is challenging albeit important for human interpretation. Not a single introductory AI course mentioned or otherwise addressed any issues of "accountability" or "discrimination" in the context of introducing AI. For introductory AI courses that were part of a lecture series, we searched all remaining lectures for the prevalence of any any ethics-related course sessions in any of the lecture series of introductory AI courses in our sample. Adding to these observations, in our qualitative, line-by-line coding of the video transcripts, we did not identify any implicit or indirect teaching of ethics or ethics-related concepts.

We assessed whether the content creators of introductory AI courses in our sample had uploaded any courses on AI ethics, which was not the case. Because we considered that perhaps AI ethics is covered in other educational videos on YouTube, we ran another keyword search to source introduction/beginner AI *ethics* courses, applying our previous sampling criteria. Search results indicated a patchy landscape of different topics including video titles such as "Will AI take my job?", "AI ethics consulting", or "The truth about AI". In total, we found seven AI ethics courses with at least 1000 views. Only V2 offered a full introduction to AI ethics. Other courses addressed a specific AI technology (e.g., generative AI), a more concrete problem in AI ethics (e.g., bias and fairness) or a subfield of AI ethics (e.g., responsible AI). In comparison to introductory AI courses, AI ethics courses were significantly less represented on YouTube. We summarize general information on the most watched AI ethics courses on YouTube in Table 3. Notably, these videos have considerably fewer views than the introduction to AI videos in our corpus.

---

[5]Our analysis only considers use of these concepts in an ethical or societal context. We exclude counting these words in different contexts. For example, the term "bias" appeared frequently in introductory AI courses, however, our manual coding revealed exclusive use of the term in the context of model parameters and weights. Instructors used the term "fair" in statements such as: "We're covering a fair bit of ground today", which we did not consider for the analysis.

Video	Content Creator	Publication Year	Technical Concept (rand. ex.)
V1	University	2023	Perceptron
V2	Influencer	2022	Matrices
V3	Online education organization	2018	Input layer
V4	Influencer	2022	Computing the loss
V5	Influencer	2021	Training and testing dataset
V6	Online education organization	2020	Indexing
V7	Online education organization	2019	Linear regression
V8	Influencer	2020	Markov model
V9	Influencer	2020	Decision tree
V10	Online education organization	2019	Random forest
V11	University	2020	Convex optimization algorithms
V12	Influencer	2019	Batch normalization
V13	University	2014	Neural network
V14	University	2017	Feature engineering
V15	Influencer	2020	Collective intelligence
V16	University	2023	Greedy best-first search
V17	Influencer	2022	Tensors
V18	Influencer	2017	Weights and biases
V19	Company	2023	Discriminative model
V20	Influencer	2021	Sigmoid function

Total views: 49,217,851   **Total length**: 92:22:07 hours
Average views: 2,460,893   **Average length**: 4:37:06 hours
Median views: 1,559,587   **Median length**: 1:11:42 hours

Table 1: The final sample of the 20 most watched introductory AI courses on YouTube. The sample comprises three different content creators (universities, influencers, and online education organizations) and spans the years 2014 - 2023. Randomly selected *Technical Concepts* taught in the courses are displayed in the right column of the table.

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	Total	In # videos
Ethics	0	0	0	0	0	0	0	0	0	0	1	3	0	0	11	0	0	0	0	0	**15**	**3/20**
Privacy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	**4**	**1/20**
Fairness	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	**2**	**1/20**
Accountability	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**0/20**
Transparency	0	0	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0	0	0	0	**3**	**2/20**
Responsibility	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	**4**	**1/20**
Bias	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	**2**	**1/20**
Explainability	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	**2**	**1/20**
Discrimination	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**0/20**

Table 2: Prevalence of ethics concepts in introductory AI courses on YouTube. Numbers in cells indicate the frequency of ethics concepts mentioned in V1 to V20.

4.2 RQ2: How do introductory AI courses on YouTube define and frame AI?

We find two relevant axes in how videos introduce AI initially: with excitement and hype around the topic and through a formative set of definitions and categorizations.

4.2.1 Excitement and hype: Introductory AI courses draw an entirely positive image of AI and its field. In 18 introductory AI courses, instructors promoted an overly positive impression of AI and its adjacent field (codes ***Timeliness & Urgency of AI*** and ***Excitement & Hype around AI***). During the initial phase of the course, instructors typically claimed that the field of AI had arrived at a particularly important point in its development in order to create a sense of urgency around entering this field *now*. Instructors framed recent AI advances to be "*particularly rapid*" (V11) and "*accelerating faster than ever before*" (V1). In three more recently published courses (2022 & 2023), instructors started the lesson underlining the fast development of AI by referring to the widespread use of generative AI (V1, V19) and ChatGPT (V2). Explaining the timeliness and urgency of AI development, course instructors typically pointed to one, two, or all of the following reasons: (1) abundance of data, (2) access to fast compute, and (3) maturity of large-scale algorithms. We present illustrative textual examples of our code ***Timeliness & urgency of AI*** in Table 4.

Video	Content Creator	Publication Year	Topical Focus
V1	University	2021	Bias & fairness
V2	Influencer	2023	AI ethics
V3	University	2023	AI ethics bias
V4	Online education organization	2023	Risks of generative AI
V5	Online education organization	2018	Ethic of AI warfare
V6	Online education organization	2022	Responsible AI
V7	Company	2020	Ethics of AI

Total views: 208,620   **Total length**: 6:32:59 hours
Average views: 29,802   **Average length**: 0:56:08 hours
Median views: 36,919   **Median length**: 0:48:06 hours

Table 3: Overview of courses on YouTube that offered educational content on AI ethics. Only V2 presented a comprehensive overview of AI ethics. Remaining videos presented specific content such as ethics of AI warfare.

Timeliness & Urgency of AI

"Many incredible successes and a lot of problems that even just a decade ago we thought were not really even solvable in the near future now we're solving with deep learning with incredible ease." (V1)

"It's 2023. Tools like chat GPT exist and all of a sudden a lot of you want to dive into machine learning and artificial intelligence." (V2)

"Probably one of the most exciting advancements that we're in the middle of experiencing as humans." (V3)

"Right now it's the most important technology in today's world." (V7)

"There's one subset of machine learning that's really hot right now because it's just advancing very rapidly, which is deep learning." (V11)

Table 4: Illustrative examples of code *Timeliness & Urgency:* During the first phase of the introductory course, and throughout, instructors framed AI as one of the most important technologies, solving challenges previously deemed unsolvable, and praised recent advances in particular.

Further enhancing the timeliness of AI progress, instructors created hype around AI by placing emphasis on AI's power and impact across different disciplines, fields and industries; by encouraging the audience to learn AI as these skills could lead to a highly profitable career; and by portraying AI and the "*AI community*" as "*exciting*" and "*fun*." **Instructors framed AI to be among the most powerful technologies today**. The audience was told that AI skills allowed them to participate in the development of "*self-driving cars*" (V7), the "*future of autonomy*" (V1), or "*self-aware machines*" (V3). In V15, the instructor stated that AI contributes to "*humanity at large*" and, in V11, introduced AI as a technology that can "*make our democracy run better rather than make it run worse*". In V12, the audience was told that "*AI began with the ancient wish to forge the gods*". Instructors emphasized AI to be a universal technology able to solve problems across a wide range of application fields. **Instructors told the audience that learning AI would likely lead to high-paying jobs**: connecting to statements around AI's universal applicability, courses promoted AI skills to be in demand across many different industries and thus economically lucrative ("*I'm sure you all agree that machine learning is the hottest trend in today's market*" (V10)).

The importance of acquiring AI skills was justified by reference to a more general development toward automation industry processes ("*Everyone is going for automation*" (V6)). **Instructors told the audience that AI was an exciting field:** Instructors told the audience that AI was "*exciting*"[6] and that they could join an "*enthusiastic*" AI community (V1). We show illustrative examples of our code ***Excitement & Hype around AI*** in Table 5.

In contrast, a minority of introductory AI courses (7/20) attempted to balance out the excitement and hype around AI at some point throughout the session (code ***Reducing hype around AI***). However, this counter-balance largely came in the form of acknowledging the difficulty or drudgery of AI work rather than its impact on the world. For example, some courses indicated AI/ML development can be frustrating and challenging work; for example, in V6, the instructor stated: "*70 to 80% of the time in a ML project you're not really sitting and programming you're just trying to understand your data and that's true I mean I'm not making this up.*" V2 and V17 (both influencers) also offered alternatives to the overly

---

[6]Based on a keyword search and subsequent manual analysis, we found that courses using the word "*exciting*" in relation to AI or the field of AI most often were V17 (66 times), V15 (59 times), V12 (21 times), V11 (11 times) and V1 (7 times).

positive introduction of AI within the first phase of the course: "*If you get into this field most of your work is not going to be training a model, coming up with a model architecture and doing all of that fun stuff you might see in YouTube videos or hear about in articles...this is not as glorious and as glamorous as a field as you might make it out to be.*" (V2). In only one course (V17) did the instructors offer an explicit counter to the hype that began most videos, pointing out that "*...although machine learning is very powerful and very fun and very exciting, it doesn't mean that you should always use it...Machine learning isn't a solve all for everything.*"

4.2.2 Framing AI: High-level definitions, AI/ML examples, and AI abilities (present and future). In 19 introductory AI courses, instructors presented a general, high-level definition of AI or AI sub-fields such as machine learning and deep learning (code **High-level definition of AI**). These definitions typically framed AI as a "*universal*" technique that enables the "*automatic extraction of information from data*" without a human having to "*explicitly*" program a machine (or a computer) to do so (e.g., V3, V5, V7, V11, V14). We found various different AI definitions across our corpus of introductory AI courses ranging from pattern extraction from data over learning a mapping function to automating human decision-making to machines that can think and act like humans. Here, instructors often highlighted the need for AI algorithms to generalize to new data (code **Generalizability**).[7] Additionally, instructors defined AI taxonomically. While some described the delineations of 1) supervised, 2) unsupervised and 3) reinforcement learning (and in rare cases, semi-supervised learning), others chose more capability-based delineations. For example, 1) reactive, 2) limited memory, 3) theory of mind, 4) self-aware (V4), or even 1) artificial narrow intelligence, 2) artificial general intelligence, 3) artificial super intelligence (V7).

Similarly, we found various claims about AI's current abilities: from performing "*complex calculations*" (V16) to "*problem solving*" (V9) and "*processing millions of documents in 10 seconds*" (V7) to "*AI building software*" (V1) and "*robots that can serve coffee*" (V3) (code **AI ability (now)**). Claims about future abilities of AI typically made references to "*sentient and conscious*" machines (V3), "*artificial super intelligence*" (V7), "*general purpose artificial intelligence*" (V12) or "*level 4 AI*" (V15) that does not require any "*human supervisor*" (code **AI potential in future**). When defining AI and describing its abilities, we noticed that instructors often attributed human characteristics to AI (code **Anthropomorphizing AI**). For example, when a model was not trained it was compared to a baby that has no information about the world (V1) while a trained model was compared to "*an intelligent person*" (V9). In Table 6, we show representative statements made by instructors that attributed human capacities to AI ("*thinking*", "*reasoning*", "*learning*" like humans) or that compared AI abilities to human abilities. The tendency to anthropomorphize AI was particularly prevalent in introductory AI courses: 17 out of 20 courses

attributed human characteristics to AI or compared AI to human performance.

We found that only a minority of introductory AI courses (7/20) addressed limitations of AI (code **AI limitations**). In V1, the instructor outlined the computational cost of computing gradient descent in a dataset while, in V12, the instructor stated AI to be "*extremely smart at a particular task but not...able to transfer*" to another task. V15 told the audience that AI "*doesn't understand the fundamentals of the concepts that are being reasoned about*" and, in V19, the instructors outlined that "*hallucinations can be a problem for transformers because they can make the output text difficult to understand. They can also make the model more likely to generate incorrect or misleading information.*" Overall, introductory AI courses did not consistently cover fundamental limitations of AI.

Moving from a general, high-level definition of AI to a specific context application or problem task, instructors offered the audience a first "*real-life*" example of an AI application (code **AI/ML example**, 17/20).[8] We categorized all AI/ML examples instructors presented in introductory AI courses in Table 11 in the Appendix. Examples from the field of computer vision were presented most frequently. The most common AI/ML examples were differentiating images of cats versus images of dogs (8/20), scene recognition in self-driving cars (8/20) and face recognition (7/20). Introductory AI courses presented a plethora of different AI/ML example applications ranging from generative AI (a deepfake of former US President Obama welcoming the class, (V1)), smart home (intelligent heating, (V3)) to medical predictions (diabetes, cancer and COVID-19, (V4, V5)), gaming (boat racing, (V12)) and recommender systems (Netflix, Google search engine, (V7)).

4.3 RQ3: What tools and exercises do introductory AI courses on YouTube focus on?

In teaching AI, introductory AI courses rely on an infrastructure of coding tools and frameworks.

4.3.1 It's easy thanks to AI/ML tools and frameworks: Introductory AI courses promote a "hands-on" approach to AI and advertise inclusiveness and accessibility. We found instructors established their teaching authority and competencies when introducing themselves to the audience (code **Instructor Authority**). University teachers highlighted their university affiliation and rank. In some cases, they also mentioned current or previous employment at large digital technology companies (e.g., (V4), (V6), (V11)). YouTube education influencers pointed to previous university affiliations, previous tech employment, their YouTube channel and courses taught in the past. Education platform instructors highlighted their affiliation with the education platform provider, previous university affiliations and employment in digital tech.

Introductory AI courses on YouTube promoted hands-on learning with established AI tools and

---

[7]For example, in V1, the instructor explained: "*We want to build models that can learn representations from our training data that can still generalize even when we show them brand new unseen pieces of test data.*"

[8]In some introductory classes, these AI/ML examples also served as prediction/classification coding examples, see AI/ML use cases in Table 13 in the Appendix.

Excitement & Hype around AI

"Go, go transform healthcare or go transform transportation or go build a self-driving car. Um, and do all of these things that, um, after this class, I think you'll be able to do." (V11)

"PyTorch 1.0 and TensorFlow 2.0. these really solidified, exciting, powerful ecosystems of tools that enable you to do very, to do a lot with very little effort. The sky is the limit, thanks to the tooling." (V12)

"We can train autonomous vehicles entirely in simulation and deploy them on full-scale vehicles in the real world seamlessly." (V1)

"The average annual salary for a machine learning engineer, which is over \$134,000. And there are also a lot of job openings..."(V7)

"With python being so easy to learn and having so powerful capabilities it seems to be the dominant choice for new learners and industry professionals alike...Who knows? You might be the next industry expert." (V6)

Table 5: Illustrative examples of code *Excitement & Hype around AI*: During the first phase of the course, instructors framed AI skills as powerful competencies in high demand across different high-paying industries and with near-universal application possibilities.

Anthropomorphizing AI

"It will have the ability to reason about the code that it's generated and walk you through step by step explaining the process and procedure all the way." (V1)

"Machines will be super intelligent sentient and conscious so they'll be able to react very much like a human being although they'll have their own flavor." (V3)

"We simply load the model from the file and ask it to make predictions that model is already trained we don't need to retrain it it's like an intelligent person." (V9)

"97 accuracy which is amazing considering the fact that this algorithm never went to college didn't go to medical school doesn't even know what a tumor is but from this data it was able to find a pattern and from that it can be used to help radiologists and doctors." (V19)

"Artificial general intelligence which is also known as strong AI, it involves machines that posses the ability to perform any intelligent task that a human being can." (V7)

Table 6: Illustrative examples of code *Anthropomorphizing AI*: When defining AI, course instructors tended to ascribe AI human characteristics and qualities. Often, instructors stated that the core goal of AI was to automate human thinking and/or behavior.

frameworks with less focus on conceptual foundations and "bottom up" programming. We found that claims around "*excitement*", "*fun*" and "*community*" (see Section 4.2.1) were often part of a chain of statements that promoted learning AI was easy, required hardly any knowledge prerequisites and little resources (codes **Emphasis on hands-on AI techniques**, **Inclusive pedagogy**, **In-class AI/Ml tools**). Technical concepts were often described as "*very easy to understand*" and AI algorithms "*very easy to implement*". Course instructors typically downplayed the importance of prior competences (e.g., "*...you don't need calculus for deep learning...*" (V5)) overall promoting an image of learning AI that was inclusive and accessible ("*If you are someone who is interested in machine learning and you think you are considered as everyone, then this video is for you*" (V4)). Such statements were often justified by referencing an established infrastructure of available AI/ML "*learning*" tools and frameworks such as Google's TensorFlow, Jupyter Notebook, Google Colab, or Meta's PyTorch that offered an "*easy-to-use interface*" ("*Virtually anyone can solve in a day with no initial investment problems that would have required an engineering team working for a quarter and \$20,000 in hardware in 2014.*" (V17)). 17 courses used Python to teach AI/ML coding. Instructors frequently pointed out that these tools were "*free*" (e.g., no costs, free GPU) and offered off the shelf AI algorithms in "*popular*" AI/ML libraries such as pandas and numpy (see also code **In-class AI/Ml tools**). As the majority of these AI/ML tools are open-source, instructors frequently stated that the audience would make "*huge progress*" within a few weeks: ("*I hope that after the next 10 weeks, you will be an expert in machine learning*" (V11)), and that AI/Ml tools and frameworks would make

learning AI "*as easy and stress-free as possible*" (V2). Notably, the tools we list above have largely been popularized by prominent technology companies who consolidate their dominant position as AI leaders through their support and maintenance [59, 86]. We show the most frequently used AI tools and frameworks in Table 12 in the Appendix.

Our code ***Corporate sponsorship and promotion*** captured instructors' praise of tech companies' products. In running predictions and classifications, instructors recommended companies' free giveaway of GPU ("*we have to connect to a runtime and these are basically if you click this connect button the GPUs that Google gives us for free for us to do all of our training which is really really cool*" (V20) or free access to APIs ("*PaLM API lets you test and experiment with Google's large language models and gen AI tools*" (V19)). In V1, Google[9] was a sponsor of the course, and an Nvidia graphic card was the prize of a programming competition. Similarly, in V2, the instructor mentioned Nvidia as a sponsor of the video, and an Nvidia graphic card to be an award.

4.4 RQ4: What methodological practices and competencies do introductory AI courses promote?

We find that in prioritizing content, AI courses tend to breeze over data collection as well as downplay the relevance of domain experts.

4.4.1 Introductory AI courses largely skip over procedural elements of data collection and generation in AI. All introductory AI courses addressed the importance of data for AI. With the code ***Data Collection & Generation*** we captured fundamental value judgements about data and data collection and hence assessed introductory AI courses' data practices throughout the AI/ML pipeline in coding examples presented to the audience (see all ***AI/ML Use Cases*** in Table 13 in the Appendix). First, courses commonly projected an "*everything can be datafied*" image of the world. For example, in V6, when introducing the states and actions of a reinforcement learning task, the instructor noted: "*Know all these things those they look qualitative but they are all quantitative. We can actually make all these things as observable entities.*" Second, introductory AI courses taught a "*more data equals better AI performance*" attitude (see Table 7 for example comments).[10] In 18 courses, instructors evaluated model performance by metrics of accuracy (code ***Performance Metrics***) whereby *low* accuracy was most commonly explained by insufficient data available. One video alluded to values beyond data quantity. In V2, the teacher instructed the audience to think about data collection procedures and data sourcing: "*Pick the data that you're going to be using for your model so data collection where are you getting this data from how are you storing it and how are you going to decide what data you actually*

need." However, none of the introductory AI courses offered meaningful information on dataset quality dimensions such as balance, completeness, or consistency. **Regardless of their length and depth, introductory AI courses on YouTube focused on values related to data quantity but not data quality.** We could not find any statements underlining the contextuality of data collection and its inevitable governance by social, cultural, ethical, and legal norms.

Rather, when instructors introduced an AI coding example the dataset was ready for use. For example, in V9, the instructor stated that machine learning projects *started* with importing data to a Jupyter notebook: "*A machine learning project involves a number of steps the first step is to import our data which often comes in the form of a csv file you might have a database with lots of data we can simply export that data and store it in a csv file for the purpose of our machine learning project.*" Alternatively, instructors generated the data "on the go" as part of the prediction or classification task (see V8 textual example in Table 8). After a dataset had been uploaded to an AI/ML tool (e.g., Jupyter Notebook), we found that instructors sometimes showcased standard data preparation methods such as data formatting and normalization (code ***Data Preparation***, see Table 14 in the Appendix). This did raise awareness that datasets required further attention before moving on to analytics.

4.4.2 Downplaying domain knowledge: Anyone can develop powerful AI predictions and classifications given enough technical resources. With the code ***Feature Selection*** we wanted to analyze how introductory AI courses addressed the fitness for purpose of a particular dataset in the context of an AI coding task (Table 13). While courses introduced the concept of data labels (i.e., the target variables) in supervised learning (see Section 4.2.2), instructors typically did not explain why the selected data and its features were relevant for a AI/ML coding tasks. Consequently, the audience did not learn any meaningful information about data annotation procedures (assigning semantics to data) or data selection. One introductory course, V14, raised awareness about the importance of selecting relevant data in the context of predicting whether students would be good or bad programmers: "*On the other hand, I don't believe in astrology. So I don't think the month in which you're born, the astrological sign under which you were born has probably anything to do with how well you'd program. I doubt that eye color has anything to do with how well you'd program. You get the idea. Some features matter, others don't.*" Despite going through a plethora of different prediction and classification tasks and contexts (e.g., medical, commercial, financial, see Table 13 in the Appendix), courses focused on model performance (i.e., accuracy) and de-prioritized data assessment.

When instructors spoke about prediction or classification they typically told the audience to "*fit a line to the data*", "*separate the dots on the grid*" or "*find data clusters*" (code ***Analysis: Prediction and Classification***). Challenges were often presented as the result of technical constraints only. For example, in V6, the instructor introduced a "*CCTV scene analysis*" tool stating that the challenge of "*scene description*"

---

[9]Through a keyword search, we found that the term "*Google*" was mentioned a total of 342 times across 19 out of the 20 introductory AI courses in our sample — far more than any other corporation. Amazon and Facebook (Meta) ranked second with 29 mentions each across our sample.

[10]We could not find a single course that raised any meaningful concerns about data privacy, the collection of sensitive information, over-collection of data, or any statements regarding harms associated with data collection (see also Table 2 keyword "*privacy*".)

Data Collection & Generation (1)

"Data is so much more pervasive than it has ever been before in our lifetimes these models are hungry for more data and we're living in the age of big data more data is available to these models than ever before and they thrive off of that." (V1)

"The more input data we give it the more accurate our model is going to be." (V9)

"Now let's change this back to 0.2 run this one more time okay now the accuracy is 75 percent now we drop to 50 again the reason this is happening is because we don't have enough data." (V6)

"Computers can be programmed to be quite intelligent by learning from data and learning from experience, being able to perform a task better and better based on greater access to data." (V16)

"Obviously, it'll predict better outcomes if it is being trained on more data." (V7)

Table 7: Illustrative examples of code *Data Collection & Generation*: When introducing data, instructors promoted a "the more data the better" attitude towards data collection, generation, and use.

Data Collection & Generation (2)

"Uh, the Internet has tons of unlabeled text data. You just suck down data from the Internet." (V11)

"I'm going to download the data from yahoo finance." (V5)

"I have seen people discussing how to create a series from scratch but I don't think that's not really important because when you start working in production it is very rare you create your own data you'll get it from somewhere..." (V6).

"Now what I want to do is just create an example here, I'm going to make a dataset that is about students final grades in like a school system. So essentially, we're gonna make this a very easy example." (V8)

"If you have to go manually and collect the data, it's going to take a lot of time. But lucky for us, there are a lot of resources online, which provide datasets. All you need to do is web scraping where you just have to go ahead and download data." (V7)

Table 8: Illustrative examples of code *Data Collection & Generation:* Besides prioritizing data quantity, instructors promoted data as an abundant resource readily available for use in the AI pipeline.

---

laid primarily in the high cost of GPU processing rather than in "*building the model.*" Courses did not talk about the socio-technical nature of classification practices (e.g., sorting people into categories) or prediction tasks (e.g., potential harm caused by biased error rates). **Summarizing, in taking the audience through the AI/ML pipeline, introductory AI courses created the impression that *anyone* – independent of any domain knowledge pertinent to the task – could make powerful AI predictions or classifications given enough data and computing resources**.

4.5 Concluding summary of findings

In summarizing our findings, we underscore the accessibility of YouTube for introductory AI courses and note a general lack of engagement with ethical and societal considerations. Course instructors present AI enthusiastically and overemphasize its positive impact on society. A meaningful engagement with data is currently missing from introductory AI courses. Guiding beginners through the AI/ML

pipeline, courses do not address the broader social context that shapes AI development, and that it shapes in turn. In the following, we offer recommendations that support course instructors in promoting an ethical mindset as part of an AI developer's professional identity.

5 INTEGRATING ETHICAL REFLECTIONS INTO INTRODUCTORY AI COURSES ON YOUTUBE

One argument for integrating ethics content into university-level introductory computing courses is that many students who are not computing majors may take only a single programming course–which may therefore be their only potential exposure to the content [30]. Similarly, when considering the potential role of publicly accessible content introducing AI concepts, it is very possible that these YouTube videos are the only source of a viewer's AI learning. Though of course a lesson introducing AI cannot be expected to provide a comprehensive overview of the myriad ethical issues in AI

Feature Selection

"But at the heart of supervised learning is the idea that during training, uh, you are given inputs X together with the labels Y and you give it both at the same time, and the job of your learning algorithm is to, uh, find a mapping so that given a new X, you can map it to the most appropriate output Y." (V11)

"We will build a heart failure prediction uh model that will predict whether the person whether the person will die based on some of the features or not okay so this is the problem statement and we have certain features like age gender blood pressure smoke whether a person is smoker or not whether a person has diabetes or not..."(V5)

"So we're telling our model that if we have a user who's 20 years old and is a male they like hip hop..."(V9)

"For example, here you can say 1 to 5: very unfit, moderately unfit, ok, relatively fit, very fit." (V6)

"Now what I want to do is just create an example here, I'm going to make a dataset that is about students final grades in like a school system. So essentially, we're gonna make this a very easy example." (V8)

Table 9: Illustrative examples of code *Feature Selection:* Introductory AI courses did not promote or otherwise teach any competencies in feature selection. They tended to uplift statistical knowledge and downplayed domain knowledge.

development and implementation, there is an opportunity to introduce the importance of this topic, provide more details when most relevant to the technical concepts being introduced, and encourage viewers to seek out more information. It is also worth noting that, similar to MOOCs, YouTube videos may have a different target audience than university course settings—for example, audiences may be more focused on career building than knowledge building [91]. Though some of the research we draw from on best practices is based on interventions in formal education, the best implementation for these recommendations may differ in order to be more appealing to this audience and their needs.

1. AI introductory courses should mention ethical implications of AI. Especially in the context of detailing "excitement" around AI and the ways that it will transform the world, there should be an appropriate counterbalance and acknowledgment of potential negative outcomes and unintended consequences in introductory AI courses. For example, given the excitement around large language models, a course introducing these concepts might point to commonly discussed risks such as disinformation, biased outputs, and socioeconomic and environmental harms [9, 75]. Additionally, when pointing to the accessibility of AI knowledge, it is worth pointing out that this also means that it is increasingly accessible to bad actors, and that the basic skills taught in this lesson could be the first step towards malicious uses of AI [13].

"Mentioning" such ethical challenges should of course be considered a floor and not a ceiling for what might be desired. Prior research suggests that undergraduate students consistently report computing classes to not raise ethical issues at all [69], and thus exposure may be a good place to start. Additional useful pedagogical strategies include having students read news articles related to the societal impact of the class topic [7] or having students imagine unintended consequences through science fiction [49]. Though YouTube videos do not provide space for assignments like these, instructors might take inspiration for how to talk about issues in an engaging way, for example by mentioning related topics in the news or the relevance of imagined anecdotes (e.g., a *Black Mirror* episode).

2. AI introductory courses should raise ethical issues explicitly relevant to technical concepts discussed. Particularly when elaborating on possible risks or unintended consequences, examples may be more powerful (and organic) when directly tied to the technical context of the conversation. This would not mean creating narrative stretches that would be unnatural, as the relevance of ethics is present, for example, throughout the machine learning pipeline [83], and prior work has created concrete examples of how this might be accomplished [77]. For example, instructors might draw from taxonomies of AI privacy risks [56] or frameworks of harm [83] at the points of the ML pipeline where the problems arise (e.g., data collection, processing, or annotation). More generally, it is crucial that introductory AI courses address potential *failure modes* in AI prediction and classification tasks.

When pointing students to existing datasets, instructors could point out how AI can exacerbate surveillance risks by increasing the scale and ubiquity of personal data collected [56]. Instructors can also bring up issues like the provenance and perspectives encoded in the dataset, and the potential implications thereof [35]. For example, many of the datasets used in the videos that we analyzed made use of publicly available data, some of which were likely collected without consent of the people who created that data, which can have unintended consequences with respect to privacy and autonomy [78]. AI research is currently lacking in research ethics norms and disclosures, particularly when it comes to data collection and labeling [43], but one question always worth asking is "where did this data come from and should I be using it?" At later stages in the pipeline,

when discussing performance measures, rather than only bringing up accuracy or error rates, instructors can also elaborate on how error rates may be distributed unequally [18]. And of course when creating predictive models, it is important to consider how those predictions might be used, even in unintended ways [6, 85]. For example, instructors can demonstrate how AI can reflect physiognomic measuring through models learning correlations between arbitrary inputs and outputs by pointing to examples that may be even more high stakes or problematic than the course context (e.g., predicting sexual orientation from images) [29, 56]. Prior research on ethics integration into computing courses often emphasizes that connecting concepts directly to that content (e.g., as part of a technical assignment) is an important strategy for illustrating that ethical consideration should be *part of* technical practice [30]. Additionally, the ethical issues specific to the course content may be most comfortable for an instructor to include. As Smith et al. point out in a discussion of the barriers to ethics integration, the suggestion is not, for example, for a machine learning instructor to be teaching moral philosophy, but rather that "as an expert on machine learning as a domain area, they should be capable of learning enough about the specific, applied area of machine learning ethics in order to competently teach that topic to students" [82]. However, it is also important to incorporate expertise and perspectives from humanists and subject matter experts when possible, and even more importantly, to not privilege the importance of technical content *over* ethical concerns [37, 73].

3. AI introductory courses should nurture a sense of accountability in people developing AI. Some of the AI practices that are known to contribute to societal harms resemble those promoted in introductory AI courses. For example, a recent major data privacy scandal involved the scraping of personally identifiable images (3 billion images in total) from social media accounts for the training of accurate facial recognition systems applied in surveillance programs [74]. This case exemplifies a disregard for consent in data collection, prioritizing data quantity, transferring the data from one context to another, and neglecting the broader societal consequences of the AI application.

When providing an introduction to AI, instructors project a certain vision of the discipline they teach. Our findings underscore how this vision may exacerbate what members of the AI ethics community have identified as the "accountability gap" [88]. AI developers perceive limited agency in shaping the relationship between designing a system and its potential outcomes [87]. Introductory AI courses increase the perceived and real distance between system design and system impact. The emphasis on hands-on competencies acquired through powerful AI tools may, on the surface, project autonomy to efficiently generate powerful AI applications. In reality, however, it prevents learners from seeking a deeper conceptual understanding and leaves them without knowledge on how to contemplate potential societal implications.

The framing of data we find in the courses — as abstract objects that can be played around with at will — could further lower a sense of accountability in learners. As such, introductory AI courses contribute to an understanding of

AI development that fails to account for the properties of the social context for which they are deployed [28, 29, 79]. Prediction and classification become a matter of "fitting the line to the data" rather than procedures that might be sorting people into categories [27]. Birhane et al. recently demonstrated that the majority of machine learning research encodes values related to the performance of AI systems, to the neglect of the potential ethical and societal harms of such systems [12]. Our results indicate that developers internalize such values early on, possibly at the onset of learning AI skills. Course instructors must clarify that datasets represent real individuals and that a model's classifications and predictions could significantly impact these individuals. Indeed, one of the critical goals of integrating ethics into introductory AI courses is that learners feel accountability for the AI systems they develop. As Goetze points out, developers are morally entangled with the systems they develop, including when such systems are unfair, discriminate, or otherwise bring about negative societal consequences [36]. It is the responsibility of educational resources to instill this sense of accountability in learners as part of what it means to be proficient and competent in the discipline.

6 CONCLUSION

In this work we present a novel descriptive analysis of introductory AI videos on YouTube. We do so by qualitatively analyzing 20 of the most viewed videos, totaling around 92 hours of content and reaching a collective viewcount of almost 50 million. These videos are formative in shaping lasting impressions of what AI is and what it can do, so it is important to gain an understanding of what is being covered. Overall, a key theme we find is a focus on technical, hands-on coding, to the detriment of covering ethics-related concepts and topics. We conclude by proposing a set of suggestions that are grounded in our descriptive findings for how introductory AI YouTube videos might begin to engage broader and critical concepts such as integrating ethical reflections and issues. By doing so, we can instill a more active sense of accountability and awareness of these concepts during this formative and context-setting part of the education pipeline.

7 ANALYST POSITIONALITY & ETHICAL CONSIDERATIONS

We acknowledge our own positionality relative to this study. The authors involved with this work represent voices from multiple countries, genders, races, and varied socioeconomic status. The authors have complementary expertise and perspectives, and this study is a collaboration between qualitative and quantitative researchers. Our interdisciplinary research team encompasses expertise within computer science and computer science education, human computer interaction, law, and philosophy. Some of the authors have taught AI ethics or technology ethics courses, and some have taught technical cources. AI ethics is a highly relevant issue within each author's area of scholarship.

Our study requires ethical consideration. We underscore our commitment to responsible data handling and ethical

practices in the presentation of video content. To that end, our sampling method ensured that we only collected data from videos that had an extremely large public audience already. We also opted not to release names or urls of video creators and their videos but rather to aggregate and showcase only specific attributes of each. This approach avoids any potential harm to the reputation of instructors by refraining from identifying specific individuals, exact view counts, or public reception and comments. Our goal in this paper is to engage in a constructive critique of the broader corpus of YouTube videos on AI education and the associated culture, focusing on overarching themes rather than singling out individuals.

ACKNOWLEDGMENTS

REFERENCES

[1] Olasile Babatunde Adedoyin and Emrah Soykan. 2023. Covid-19 pandemic and online learning: the challenges and opportunities. *Interactive learning environments* 31, 2 (2023), 863–875.
[2] Liudmila Alekseeva, José Azar, Mireia Gine, Sampsa Samila, and Bledi Taska. 2021. The demand for AI skills in the labor market. *Labour economics* 71 (2021), 102002.
[3] Becky Allen, Andrew Stephen McGough, and Marie Devlin. 2021. Toward a framework for teaching artificial intelligence to a higher education audience. *ACM Transactions on Computing Education (TOCE)* 22, 2 (2021), 1–29.
[4] Deborah E Allen, Barbara J Duch, and Susan E Groh. 1996. The power of problem-based learning in teaching introductory science courses. *New directions for teaching and learning* 1996, 68 (1996), 43–52.
[5] Maha Bali. 2014. MOOC pedagogy: Gleaning good practice from existing MOOCs. *Journal of Online Learning and Teaching* 10, 1 (2014), 44.
[6] Chelsea Barabas, Madars Virza, Karthik Dinakar, Joichi Ito, and Jonathan Zittrain. 2018. Interventions over Predictions: Reframing the Ethical Debate for Actuarial Risk Assessment. *ACM Conference on Fairness, Accountability, and Transparency (FAccT)* (2018).
[7] Scott Barlowe. 2020. Encouraging Global Awareness in a Computer Networking Course. *IEEE Transactions on Education* 64, 2 (2020), 195–201.
[8] Meltem Huri Baturay. 2015. An overview of the world of MOOCs. *Procedia-Social and Behavioral Sciences* 174 (2015), 427–433.
[9] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAccT)* (2021).
[10] Mariette Bengtsson. 2016. How to plan and perform a qualitative study using content analysis. *NursingPlus open* 2 (2016), 8–14.
[11] Joseph R Biden. 2023. Executive order on the safe, secure, and trustworthy development and use of artificial intelligence. (2023).
[12] Abeba Birhane, Pratyusha Kalluri, Dallas Card, William Agnew, Ravit Dotan, and Michelle Bao. 2022. The values encoded in machine learning research. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency.* 173–184.
[13] Taís Fernanda Blauth, Oskar Josef Gstrein, and Andrej Zwitter. 2022. Artificial intelligence crime: An overview of malicious use and abuse of AI. *IEEE Access* 10 (2022), 77110–77122.
[14] Noelle Brown, Koriann South, Suresh Venkatasubramanian, and Eliane S Wiese. 2023. Designing Ethically-Integrated Assignments: It's Harder Than it Looks. In *Proceedings of the 2023 ACM Conference on International Computing Education Research-Volume 1.* 177–190.
[15] Noelle Brown, Koriann South, and Eliane S Wiese. 2022. The Shortest Path to Ethics in AI: An Integrated Assignment Where Human Concerns Guide Technical Decisions. In *Proceedings of the 2022 ACM Conference on International Computing Education Research-Volume 1.* 344–355.
[16] Noelle Brown, Benjamin Xie, Ella Sarder, Casey Fiesler, and Eliane S Wiese. 2023. Teaching Ethics in Computing: A Systematic Literature

[17] Mercedes Bunz and Marco Braghieri. 2022. The AI doctor will see you now: assessing the framing of AI in news coverage. *AI & society* 37, 1 (2022), 9–22.
[18] Joy Buolamwini and Timnit Gebru. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research* 81 (2018).
[19] Jeffrey P Carpenter, Catharyn C Shelton, and Stephanie E Schroeder. 2023. The education influencer: A new player in the educator professional landscape. *Journal of Research on Technology in Education* 55, 5 (2023), 749–764.
[20] Steve Cooper and Steve Cunningham. 2010. Teaching computer science in context. *Acm Inroads* 1, 1 (2010), 5–8.
[21] Jack Crittenden and Peter Levine. 2023. Civic Education. In *The Stanford Encyclopedia of Philosophy* (Fall 2023 ed.), Edward N. Zalta and Uri Nodelman (Eds.). Metaphysics Research Lab, Stanford University.
[22] Barbara Gross Davis. 2009. *Tools for teaching.* John Wiley & Sons.
[23] Adrian A de Freitas and Troy B Weingart. 2021. I'm going to learn what?!? teaching artificial intelligence to freshmen in an introductory computer science course. In *Proceedings of the 52nd ACM technical symposium on computer science education.* 198–204.
[24] Shah Dhawal. 2023. *The 250 Most Popular Online Courses of All Time.* https://www.classcentral.com/report/most-popular-online-courses/ Accessed January, 2024.
[25] Tawanna R Dillahunt, Brian Zengguang Wang, and Stephanie Teasley. 2014. Democratizing higher education: Exploring MOOC use among those who cannot afford a formal education. *The International Review of Research in Open and Distributed Learning* (2014).
[26] Samantha Jane Dobesh, Tyler Miller, Pax Newman, Yudong Liu, and Yasmine N Elglaly. 2023. Towards Machine Learning Fairness Education in a Natural Language Processing Course. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1.* 312–318.
[27] Severin Engelmann and Orestis Papakyriakopoulos. 2023. Social media as classification systems: procedural normative choices in user profiling. In *Handbook of Critical Studies of Artificial Intelligence.* Edward Elgar Publishing, 619–630.
[28] Severin Engelmann, Valentin Scheibe, Fiorella Battaglia, and Jens Grossklags. 2022. Social Media Profiling Continues to Partake in the Development of Formalistic Self-Concepts. Social Media Users Think So, Too.. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society.* 238–252.
[29] Severin Engelmann, Chiara Ullstein, Orestis Papakyriakopoulos, and Jens Grossklags. 2022. What people think AI should infer from faces. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency.* 128–141.
[30] Casey Fiesler, Mikhaila Friske, Natalie Garrett, Felix Muzny, Jessie J Smith, and Jason Zietz. 2021. Integrating ethics into introductory programming classes. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education.* 1027–1033.
[31] Casey Fiesler, Natalie Garrett, and Nathan Beard. 2020. What Do We Teach When We Teach Tech Ethics?: A Syllabi Analysis. *Proceedings of the 51st ACM Technical Symposium on Computer Science Education (SIGCSE)* (2020).
[32] Barbara Fralinger, Russell Owens, et al. 2009. You Tube as a learning tool. *Journal of College Teaching & Learning (TLC)* 6, 8 (2009).
[33] Manuel B. Garcia, Irish C. Juanatas, and Roben A. Juanatas. 2022. TikTok as a Knowledge Source for Programming Learners: a New Form of Nanolearning? *International Conference on Information and Education Technology (ICIET)* (2022).
[34] Natalie Garrett, Nathan Beard, and Casey Fiesler. 2020. More Than "If Time Allows": The Role of Ethics in AI Education. *Proceedings of the AAAI/ACM Conference on AI, Ethics (AIES)* (2020).
[35] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. 2021. Datasheets for Datasets. *Communications of the ACM (CACM)* (2021).
[36] Trystan S Goetze. 2022. Mind the Gap: Autonomous Systems, the Responsibility Gap, and Moral Entanglement. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency.* 390–400.
[37] Trystan S Goetze. 2023. Integrating Ethics into Computer Science Education: Multi-, Inter-, and Transdisciplinary Approaches. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1.* 645–651.
[38] Barbara J. Grosz, David Gray Grant, Kate Vredenburgh, Jeff Behrends, Lily Hu, Alison Simmons, and Jim Waldo. 2019. Embedded EthiCS: Integrating Ethics Across CS Education. *Commun. ACM* (2019).
[39] Barbara J Grosz, David Gray Grant, Kate Vredenburgh, Jeff Behrends, Lily Hu, Alison Simmons, and Jim Waldo. 2019. Embedded EthiCS: integrating ethics across CS education. *Commun. ACM* 62, 8 (2019), 54–61.
[40] Jonathan Haber. 2014. *MOOCs.* MIT Press.

[41] John D. Hansen and Justin Reich. 2015. Democratizing education? Examining access and usage patterns in massive open online courses. *Science* (2015).

[42] Harlan D Harris and Sarah M Kiefer. 2004. The Pedagogy of Artificial Intelligence: A Survey of Faculty Who Teach Introductory AI.. In *FLAIRS Conference*. 74–79.

[43] Will Hawkins and Brent Mittelstadt. 2023. The ethical ambiguity of AI data enrichment: Measuring gaps in research ethics norms and practices. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 261–270.

[44] Orit Hazzan, Tami Lapidot, and Noa Ragonis. 2020. *Guide to teaching computer science.* Springer.

[45] Drew Hemment, Morgan Currie, SJ Bennett, Jake Elwes, Anna Ridler, Caroline Sinders, Matjaz Vidmar, Robin Hill, and Holly Warner. 2023. AI in the Public Eye: Investigating Public AI Literacy Through AI Art. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 931–942.

[46] Hendrik Heuer, Juliane Jarke, and Andreas Breiter. 2021. Machine learning in tutorials–Universal applicability, underinformed application, and other misconceptions. *Big data & society* 8, 1 (2021), 20539517211017593.

[47] Diane Horton, Sheila A. McIlraith, Nina Wang, Maryam Majedi, Emma McClure, and Benjamin Wald. 2022. Embedding Ethics in Computer Science Courses: Does it Work? *ACM Technical Symposium on Computer Science Education (SIGCSE)* (2022).

[48] Zoe Kanetaki, Constantinos Stergiou, Georgios Bekas, Sébastien Jacques, Christos Troussas, Cleo Sgouropoulou, and Abdeldjalil Ouahabi. 2022. Acquiring, analyzing and interpreting knowledge data for sustainable engineering education: An experimental study using youtube. *Electronics* 11, 14 (2022), 2210.

[49] Shamika Klassen and Casey Fiesler. 2022. " Run Wild a Little With Your Imagination" Ethical Speculation in Computing Education with Black Mirror. In *Proceedings of the 53rd ACM Technical Symposium on Computer Science Education-Volume 1*. 836–842.

[50] Klaus Krippendorff. 2018. *Content analysis: An introduction to its methodology.* Sage publications.

[51] Sarah R. Lambert. 2020. Do MOOCs contribute to student equity and social inclusion? A systematic review 2014–18. *Computers & Education* (2020).

[52] Patricia G Lange. 2019. Informal learning on YouTube. *The international encyclopedia of media literacy* (2019), 1–11.

[53] Benjamin Laufer, Sameer Jain, A Feder Cooper, Jon Kleinberg, and Hoda Heidari. 2022. Four years of FAccT: A reflexive, mixed-methods analysis of research contributions, shortcomings, and future prospects. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. 401–426.

[54] Samson Lausa, Rex Bringula, Annaliza E. Catacutan-Bangit, and Cereneo Santiago. 2021. Information-Seeking Behavior of Computing Students while Programming: Educational Learning Materials Usage, Satisfaction of Use, and Inconveniences. *IEEE Global Engineering Education Conference (EDUCON)* (2021).

[55] Chei Sian Lee, Hamzah Osop, Dion Hoe-Lian Goh, and Gani Kelni. 2017. Making sense of comments on YouTube educational videos: a self-directed learning perspective. *Online Information Review* (2017).

[56] Hao-Ping Lee, Yu-Ju Yang, Thomas Serban von Davier, Jodi Forlizzi, and Sauvik Das. forthcoming, 2024. Deepfakes, Phrenology, Surveillance, and More! A Taxonomy of AI Privacy Risks. *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (forthcoming, 2024). arXiv preprint arXiv:2310.07879

[57] Irene Lee, Safinah Ali, Helen Zhang, Daniella DiPaola, and Cynthia Breazeal. 2021. Developing middle school students' AI literacy. In *Proceedings of the 52nd ACM technical symposium on computer science education*. 191–197.

[58] Duri Long, Takeria Blunt, and Brian Magerko. 2021. Co-designing AI literacy exhibits for informal learning spaces. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–35.

[59] Inga Luchs, Clemens Apprich, and Marcel Broersma. 2023. Learning machine learning: On the political economy of big tech's online AI courses. *Big Data & Society* (2023).

[60] Laura MacLeod, Margaret-Anne Storey, and Andreas Bergen. 2015. Code, camera, action: How software developers document and share program knowledge using YouTube. In *2015 IEEE 23rd International Conference on Program Comprehension*. IEEE, 104–114.

[61] Tambiama Madiega. 2021. Artificial intelligence act. *European Parliament: European Parliamentary Research Service* (2021).

[62] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and inter-rater reliability in qualitative research: Norms and guidelines for CSCW and HCI practice. *Proceedings of the ACM on human-computer interaction* 3, CSCW (2019), 1–23.

[63] Michael Meaney and Tom Fikes. 2022. Adding a Demographic Lens to Cluster Analysis of Participants in Entry-level Massive Open Online Courses (MOOCs). *Proceedings of the Ninth ACM Conference on Learning @ Scale* (2022).

[64] Michael Meaney and Tom Fikes. 2023. The Promise of MOOCs Revisited? Demographics of Learners Preparing for University. *Journal of Learning Analytics* (2023).

[65] Davy Tsz Kit Ng, Jac Ka Lok Leung, Samuel Kai Wah Chu, and Maggie Shen Qiao. 2021. Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence* 2 (2021), 100041.

[66] Dennis Nguyen and Erik Hekman. 2022. The news framing of artificial intelligence: a critical exploration of how media discourses make sense of automation. *AI & SOCIETY* (2022), 1–15.

[67] Stack Overflow. 2021. *2021 Stack Overflow Developer Survey.* https://insights.stackoverflow.com/survey/2021#overview Accessed January, 2024.

[68] Mark Overmars. 2004. Teaching computer science through game design. *Computer* 37, 4 (2004), 81–83.

[69] Aadarsh Padiyath. 2024. A Realist Review of Undergraduate Student Attitudes towards Ethical Interventions in Technical Computing Courses. *ACM Trans. Comput. Educ.* (jan 2024). https://doi.org/10.1145/3639572 Just Accepted.

[70] Michael Quinn Patton. 1999. Enhancing the quality and credibility of qualitative analysis. *Health services research* 34, 5 Pt 2 (1999), 1189.

[71] Elizabeth Poché, Nishant Jha, Grant Williams, Jazmine Staten, Miles Vesper, and Anas Mahmoud. 2017. Analyzing user comments on YouTube coding tutorial videos. In *2017 IEEE/ACM 25th International Conference on Program Comprehension (ICPC)*. IEEE, 196–206.

[72] Khalil Oliveira Portugal, SDM Arruda, and MM Passos. 2018. Freechoice teaching: how YouTube presents a new kind of teacher. *Revista Electrónica de Enseñanza de las Ciencias* 17, 1 (2018), 183–199.

[73] Inioluwa Deborah Raji, Morgan Klaus Scheuerman, and Razvan Amironesei. 2021. You can't sit with us: Exclusionary pedagogy in ai ethics education. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 515–525.

[74] Isadora Neroni Rezende. 2020. Facial recognition in police hands: Assessing the 'Clearview case'from a European perspective. *New Journal of European Criminal Law* 11, 3 (2020), 375–389.

[75] Matthias C Rillig, Marlene Ågerstrand, Mohan Bi, Kenneth A Gould, and Uli Sauerland. 2023. Risks and benefits of large language models for the environment. *Environmental Science & Technology* 57, 9 (2023), 3464–3466.

[76] Matthias Rohs and Mario Ganz. 2015. MOOCs and the claim of education for all: A disillusion by empirical data. *The International Review of Research in Open and Distributed Learning* (2015).

[77] Jeffrey Saltz, Michael Skirpan, Casey Fiesler, Micha Gorelick, Tom Yeh, Robert Heckman, Neil Dewar, and Nathan Beard. 2019. Integrating Ethics within Machine Learning Courses. *ACM Transactions on Computing Education* (2019).

[78] Morgan Klaus Scheuerman, Katy Weathington, Tarun Mugunthan, Emily Denton, and Casey Fiesler. 2023. From Human to Data to Dataset: Mapping the Traceability of Human Subjects in Computer Vision Datasets. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–33.

[79] Andrew D Selbst, Danah Boyd, Sorelle A Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. Fairness and abstraction in sociotechnical systems. In *Proceedings of the conference on fairness, accountability, and transparency*. 59–68.

[80] Harvey Siegel, D.C. Phillips, and Eamonn Callan. 2018. Philosophy of Education. In *The Stanford Encyclopedia of Philosophy* (Winter 2018 ed.), Edward N. Zalta (Ed.). Metaphysics Research Lab, Stanford University.

[81] Sandra Phek-Lin Sim, Hannah Phek-Khiok Sim, and Cheng-Sim Quah. 2021. Online learning: A post COVID-19 alternative pedagogy for university students. *Asian Journal of University Education* 16, 4 (2021), 137–151.

[82] Jessie J Smith, Blakeley H Payne, Shamika Klassen, Dylan Thomas Doyle, and Casey Fiesler. 2023. Incorporating Ethics in Computing Courses: Barriers, Support, and Perspectives from Educators. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1*. 367–373.

[83] Harini Suresh and John Guttag. 2021. A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle. *ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO)* (2021).

[84] Khalaf Mohammed Tahat, Walaa Al-Sarayrah, Said A Salloum, Mohammed Habes, and Sana Ali. 2021. The influence of YouTube videos on the learning experience of disabled people during the COVID-19 outbreak. In *Advances in Data Science and Intelligent Data Communication Technologies for COVID-19: Innovative Solutions against COVID-19*. Springer, 239–252.

[85] Angelina Wang, Sayash Kapoor, Solon Barocas, and Arvind Narayanan. 2023. Against Predictive Optimization: On the Legitimacy of Decision-Making Algorithms that Optimize Predictive Accuracy. *ACM Journal of Responsible Computing* (2023).

[86] David Gray Widder, Sarah West, and David Gray Widder. 2023. Open (For Business): Big Tech, Concentrated Power, and the Political Economy

of Open AI. *SSRN* (2023).

[87] David Gray Widder, Derrick Zhen, Laura Dabbish, and James Herbsleb. 2023. It's about power: What ethical concerns do software engineers have, and what do they (feel they can) do about them?. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency.* 467–479.

[88] Maranke Wieringa. 2020. What to account for when accounting for algorithms: a systematic literature review on algorithmic accountability. In *Proceedings of the 2020 conference on fairness, accountability, and transparency.* 1–18.

[89] Michael Wollowski, Robert Selkowitz, Laura Brown, Ashok Goel, George Luger, Jim Marshall, Andrew Neel, Todd Neller, and Peter Norvig. 2016. A survey of current practice and teaching of AI. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 30.

[90] Zulnaidi Yaacob and Nor Hasliza Md Saad. 2020. Acceptance of Youtube as a learning platform during the covid-19 pandemic: The moderating effect of subscription status. *Tem Journal* 9, 4 (2020).

[91] Chen Zhenghao, Brandon Alcorn, Gayle Christensen, Nicholas Eriksson, Daphne Koller, and Ezekiel J. Emanuel. 2015. Who's Benefiting from MOOCs, and Why. *Harvard Business Review* (2015).

A APPENDIX

FINAL CODEBOOK

Table 10: Final codebook structured around three high-level themes: (1) Definitions and framing of AI, (2) Teaching resources and styles, and (3) Practicing AI.

Theme 1: Definitions and framing of AI

Code name	Description	Example
1. Excitement & Hype around AI	**Definition:** The course instructor highlights the benefits of AI and/or reiterates how "great" or "amazing" AI is.	*"...artificial intelligence (...) is probably one of the most exciting advancements that we're in the middle of experiencing as humans..."*
2. Timeliness & Urgency of AI	**Definition:** Instructor makes references to the present time as a particularly important one for AI development.	*"Now this past year in particular of 2022 has been an incredible year for deep learning progress."*
3. Reducing Hype around AI	**Definition:** The course instructor downplays the hype around AI.	*"So keep that in mind, this is not as glorious and glamorous as a field as you might make it out to be, and there's a lot of very frustrating time-consuming and difficult work."*
4. High-level AI Definition	**Definition:** The course instructor provides a high-level definition of what AI is and describes AI's main function.	*"...deep learning, which is a subcategory of machine learning. Deep learning provides artificial intelligence the ability to mimic a human brain's neural network."*
5. Generalizability	**Definition:** The course highlights that AI models need to generalize to other datasets. Code includes claims about AI being a universally applicable and transferable technique/technology.	*"...for every machine learning problem, we have an algorithm which can solve the problem..."*
6. AI Ability (Now)	**Definition:** The course instructor describes the current state of AI development in terms of AI's current capacities.	*"Deep learning can be used to generate content directly from how we speak and the language that we convey to it from prompts that we say deep learning can reason about the prompts in natural language and English."*
7. AI Potential in Future	**Definition:** The course instructor makes claims about AI's future abilities/problem-solving capacities.	*"Finally, self-awareness, this is the future of AI. These machines will be super intelligent, sentient, and conscious, so they'll be able to react very much like a human being..."*
8. Anthropomorphizing AI	**Definition:** The course instructor attributes human qualities or characteristics to AI.	*"I really want to highlight the fact that they work and react like humans because that is where the development of artificial intelligence is and that's what we're comparing it to is how it looks like next to a human. Thanks, any tasks you want me to do for you, get me a cup of coffee poof here you go he brings him a cup of coffee that's my kind of robot."*
9. AI Limitations	**Definition:** The course instructor makes claims about the limits of AI	*"...let's revisit that gradient descent algorithm right so here this gradient that we talked about before is actually extraordinarily computationally expensive to compute because it's computed as a summation across all of the pieces in your dataset..."*
10. AI/ML Example	**Definition:** The course instructor introduces a specific AI/ML application or a specific problem that AI solves. These AI/ML examples are often presented at the beginning of the class to arouse students' interest in the topic field.	*"...Neural network (...) generate(s) a photo of an astronaut riding a horse and it actually can imagine hallucinate what this might look like even though, of course, this photo not only this photo has never occurred before but I don't think any photo of an astronaut riding a horse has ever occurred before."*

Theme 2: Teaching resources and styles

Code name	Description	Example
1. Instructor authority	**Definition:** instructor's teaching authority is established by reference to institutional affiliation or mentioning of specific AI/ML skills.	*"Kylie Ying has worked at many interesting places such as MIT, CERN, and Free Code Camp. She's a physicist, engineer, and basically a genius."*
2. Emphasis on hands-on AI techniques	**Definition:** The course focuses on coding skills and implementation and only loosely hints at the theoretical dimensions of concepts.	*"You'll learn the foundations of this really really fascinating and exciting field of deep learning and artificial intelligence and more importantly you're going to get hands-on experience..."*
3. Inclusive pedagogy & accessibility	**Definition:** AI education is not limited to a specific group e.g. engineers or those with a strong math/quantitative background. Also, instructor implies / assumes that all learners have equal opportunities for AI education. Educational resources are accessible to everyone regardless of their social, cultural, ethnic, economic background.	*"...but you don't need calculus you don't even need calculus yeah if you want to go on a research level then you obviously need but for now for a machine learning you don't need calculus for deep learning even you don't need calculus."*
4. In-class AI/ML tools	**Definition:** Any AI/ML programming tool or framework used in class.	*"...unique time in history where we have the ability to train these extremely large-scale algorithms and techniques that have existed for a very long time but we can now train them due to the hardware advances that have been made and finally due to open source toolbox access and software platforms like tensorflow for example which all of you will get a lot of experience on in this class."*
5. Corporate sponsorship & promotion	**Definition:** Explicit mention of a company providing resources to sponsor the course.	*"...because i'm actually giving away an rtx 4080 this is courtesy of nvidia which i've teamed up with for this video..."*

Theme 3: Practicing AI

Code name	Description	Example
1. Data Collection & Data Generation	**Definition:** This code captures how courses introduce data. Are data treated as a technical artifact? Are datasets described in detail, do instructors mention their origin and their contextual meanings? Do instructors explain how the example data for the course were collected? Are the data simply present already? What data collection competences are taught in the course?	*"...a machine learning project involves a number of steps the first step is to import our data which often comes in the form of a csv file you might have a database with lots of data we can simply export that data and store it in a csv file for the purpose of our machine learning project."*
2. In-class AI/ML Use Case	**Definition:** In-class AI/ML application coding examples. This code illustrates AI/ML classification/prediction examples that instructors use to teach coding skills.	*"... finally we'll dive into a use case predicting if a person has diabetes or not and we'll be using tensorflow for that in the python environment."*
3. Performance Metrics	**Definition:** When performance is discussed, focus lies on accuracy, recall, precision or related performance metrics.	*"After we build a model we need to measure its accuracy and if it's not accurate enough we should either fine-tune it or build a model using a different algorithm."*
4. Data Preparation	**Definition:** This code aims to capture how the course instructor teaches steps in data preparation.	*"...the next step is going to be data cleaning oftentimes when you collect data especially if you're getting it from the internet some kind of open source you're going to need to clean this data and get rid of a bunch of information that's unnecessary."*
5. Feature Selection	**Definition:** Does the course instructor describe the process of feature selection? Do they explain why specific features are relevant for a prediction/classification task? Do they offer any insight into how data are annotated?	*"...we're telling our model that if we have a user who's 20 years old and is a male they like hip hop once we train our model then we give it a new input set for example we say hey we have a new user who is 21 years old and is a male."*
6. Normativity of Classification and Prediction	**Definition:** How does the course instructor deal with the inherent normativity of classification and prediction? Do instructors mention any high-level ethical challenges that come with classification and/or prediction? E.g., sorting people into categories, lack of transparency, power of classification/prediction, power of classification/prediction in driving automated decision making. Limits of classification/prediction?	*"...which is that that event whether the particular person died or not so that's it with the target variable we will see more about the data but what is the business objective over here every machinery problem has some kind of business objective it simply means that it's some health care problem means we'll be able to build a health care we will be able to build a healthcare in healthcare something ai and healthcare which is simply able to make a machine learning model that will help you in early detection of the person based on particular features and help the person can be saved."*

LIST OF AI/ML EXAMPLES

Table 11: AI/ML Examples Presented by Instructors of Introductory AI Courses

AI/ML	Example Application
Computer Vision	Pictures with dogs vs. pictures with cats: (V4), (V2), (V3), (V5),(V7), (V9), (V17), (V19) Detect faces in images: (V1), (V5), (V7), (V12), (V9), (V14), (V16) Self-driving cars: (V5), (V1), (V9), (V11), (V14), (V15), (V16), (V17) Picture with ramen vs. picture with spaghetti (V17) CCTV scenery classification (V6) Analysis of legal documents (V7) Real banknotes vs. counterfeit banknotes (V16)
Generative AI	Synthetic images (V1) Chatbots (V15) Generate images with faces (V12) Generate cooking recipe from ingredients (V17) Deep fakes (V1)
NLP	Hate speech detection (V7) Spam vs no spam (V7), (V16), (V17) Sentiment analysis (V5) Tweet analysis (V17) Sentence completion (V7) Voice recognition (V10), (V16) Google translate (V16)
Medical Predictions	Predict diabetes in patients (V4) Cancer prediction (V5), (V6), (V11), (V14) Covid prediction (V4) Drug discovery (V14), (V15)
Robotics	Instruct robotics to serve a cup of coffee (V3) Humanoid robotics (V12)
Smart Home	Intelligent heating (V3)
Finance	Stock price prediction (V5), (V14)
Games	Computer chess (V7) Alpha Go (V6) Watson Jeopardy (V7) Boat racing (V12)
Recommender Systems	Online shopping (V7) YouTube (V17), (V18) Google search engine (V7) Netflix (V7) Customer segmentation (V5), (V10)
Reinforcement Learning	Teach a baby to walk (V6) Teach a dog to get a biscuit (V6)
Housing	Predict housing prices from housing size (V11)

AI Tools/Frameworks	Used in # of videos	Total mentions across videos
Google Tensorflow	10/20	302
Google Colab	5/20	135
Jupyter Notebook	4/20	111
PyTorch (Meta)	3/20	548

Table 12: The most frequently AI tools and frameworks used in introductory AI courses on YouTube.

IN-CLASS AI/ML CODING EXAMPLE CASES

In-class AI/ML Coding Example Cases

AI prediction & classification context	Coding task
Medical AI prediction/classification task	Predict if a person has diabetes (V3, V10) Predict if a person has COVID (V4) Predict if a person has cancer (V20) Predict if a person has heart disease (V7) Predict if a person has heart failure (V5)
Computer vision classification tasks	Classify red and green dots (V1) Classify photos (V3) Classify banknotes (V16) Classify fruits (V7, V10) Classify foods on images (V17) Classify colors on images (V7) Classify animals based on different features (V7, V10) Classify flowers on images (V10)
Commerce & finance	Predict airline ticket prices (V3) Predict housing prices (V5, V4) Predict stock market prices (V5) Predict loan applications (V7, V10) Predict rental bike lending (V4) Predict music tastes of customers (V9, V8) Classify customers into segments (V6, V10) Predict fitness product preferences of different genders (V6)
Games	Reinforcement learning: Smart taxi (V6) Predict best location for fence against wolves (V7) Create a game to play Romeo and Juliet (V8)
Miscellaneous	Predict the weather (V5, V7, V4, V10, V16) Generate music (generative AI) (V1) Classify football players by features (V10, V14) Classify passengers that survived the Titanic (V10) Classify different voices in voice recognition task (V11) Classify people as Democrats or Republicans (V14) Classify if a person is vegetarian or not (V7) Predict: Will I pass this class? (V1) Predict if your sister gets into grad school (V7) Particle prediction in Astronomy (V4) Spam detection (V5) Cluster movies (V7) Predict speed of a car (V7) Classify sentiment (V7, V16) Predict the best time for a parking space (V16)

Table 13: Prediction & classification tasks taught in AI introductory classes (coding examples)

Data Preparation Technique
Normalizing data (V3, V12)
Data standardization (V12)
Get rid of irrelevant data (V2)
Data formatting (V2, V4, V6, V7, V8, V9)
Data splitting for training and testing data (V7, V8, V10)
Dealing with missing values (V7, V9)
Remove outliers (V6)

Table 14: Data preparation techniques taught in introductory AI courses