Ethics and Responsible Business leadership
Final Examination
Tuesday, May 7, 9am

Instructions – read carefully before you begin, and read again before submitting to ensure you are in compliance with all stated conditions

Ground rules
This exam is the exam will be open-book and open-notes but strictly individual. You may not use the Internet (including ChatGPT and equivalents), email, chat software, or any other medium of communication with others during the exam. The sole exception is you may use the Internet only to receive and submit your exam and access course material on bCourses.

The number of points given to each question is stated for each question. You may want to consider these points when allocating your time and space.

Format
There is no strict page limit. You do not need to rewrite the questions in your answers file. Allocate your time to allow for proofreading and final editing. Margins must be at least 1 inch—top, bottom, and sides—and font must be 12-point, Times New Roman.

Please title your exam file thus: [last name_first name_cohort.doc], where cohort is Blue/Gold/Oski/Axe.

Submission
Unless you have a special accommodation, you must upload your answers as an assignment submission through bCourses by 11:00am or, if you have difficulty uploading, through email to the course's GSI Sebastian (sebastian_arechaga@haas.berkeley.edu).

Help
If you need to ask (clarifying only) questions during the exam, you may ask the proctors. Professor Guo Xu will be present as will GSI Sebastian Arechaga, both of whom can answer questions.

MBA207 Business Ethics Exam 2024

The ethics of training generative AI models

Generative AI is disrupting business and ushering in large potential productivity gains, as writers, coders, and artists learn to deploy AI-based tools to speed up and enhance the quality of their work. Although evidence of productivity enhancements today has been modest,[1] the rapid pace of development and industry leaders' success in reducing errors—ChatGPT4 only hallucinates 3% of the time[2]—has led to widespread belief in the transformational power of this technology.

But has this growth been fueled by expropriating the property of content producers?

A growing chorus of disgruntled writers, artists, and other rights-holders thinks so. In December, the *New York Times* sued OpenAI for allegedly training ChatGPT using millions of proprietary articles.[3] Universal Music Group and Getty Images have recently sued other generative AI companies. Generative AI tools were "*built with stolen goods,*" the CEO of Condé Nast argued during a recent US Senate hearing.[4]

Generative AI companies need vast amounts of data to train algorithms predicting the next most likely token in a sentence, pixel in an image, or chord in a song. Where should they find this data? Most AI companies have primarily harvested data from the internet using web crawlers that scrape content from countless web pages, including news media, image banks, and music libraries. At times, AI companies may not even know exactly where these bots collected the data used to train their algorithms. Although today, many websites now include web crawler blockers, the plaintiffs noted above argue that before these blockers were developed the web crawlers plundered troves of proprietary content. Many now seek compensation for these violations of their copyright protections. Other content producers, such as the Associated Press and *Le Monde*, have chosen to make licensing deals with AI companies, albeit for relatively modest sums.[5]

Defenders of generative AI point out that human content producers also consume others' past content, which through memories or subconscious influence shapes whatever new content they produce.[6] Algorithms do this analogously but with a much larger database of "memories" upon which to draw. Traditional copyright laws, some argue, are too restrictive and outdated in today's world where content is produced in many different

[1] https://www.economist.com/business/2024/02/29/how-businesses-are-actually-using-generative-ai

[2] https://www.nytimes.com/2023/11/06/technology/chatbots-hallucination-rates.html

[3] https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html

[4] https://www.economist.com/business/2024/04/14/generative-ai-is-a-marvel-is-it-also-built-on-theft

[5] https://www.economist.com/business/2024/04/14/generative-ai-is-a-marvel-is-it-also-built-on-theft

[6] https://www.newyorker.com/magazine/2024/01/22/who-owns-this-sentence-a-history-of-copyrights-and-wrongs-david-bellos-alexandre-montagu-book-review

ways[7] "Because copyright today covers virtually every sort of human expression – including blogposts, photographs, forum posts, scraps of software of code, and government documents – it would be impossible to train today's leading AI models without using copyrighted materials."

In the US, legal ambiguity rests on the interpretation of what constitutes "fair use", which allows the use of copyrighted content without the owner's explicit permission in certain circumstances, such as when the use is deemed sufficiently "transformative". For instance, Google Search shows an excerpt from each website—a use of content that was found sufficiently transformative in what it enabled that it satisfied the conditions for "fair use." [8] As OpenAI wrote in defense of their practice, *Training AI models using publicly available internet materials is fair use, as supported by long-standing and widely accepted precedents. We view this principle as fair to creators, necessary for innovators, and critical for US competitiveness."*

The following questions ask you to consider this case through the lens of the frameworks we discussed in class. The case description contains all the information you need to answer the questions. Please be explicit when making any additional assumptions.

---

[7] https://arstechnica.com/information-technology/2024/01/openai-says-its-impossible-to-create-useful-ai-models-without-copyrighted-material/

[8] https://www.newyorker.com/magazine/2024/01/22/who-owns-this-sentence-a-history-of-copyrights-and-wrongs-david-bellos-alexandre-montagu-book-review

Questions:

1. (10 points in total) Would "Justice as Entitlements" agree with the statement that Generative AI tools are "built with stolen goods"? Explain why or why not.

2. (10 points in total) Now consider the Duty Ethics framework and the individual maxim: "I will use unlicensed proprietary content as an input to produce and sell new content and thus improve my wellbeing." Does this maxim universalize? Specifically, assess whether this individual maxim:

    a. Leads to a contradiction in logic.

    b. Leads to a contradiction in will.

3. (10 points in total) An alternative application of the Duty Ethics framework is the Formula of Humanity. Please explain whether the use of scraped web data to train one's AI model violates the Formula of Humanity. Be sure to clarify whose "ends" you are considering and how they are, or are not, affected.

4. (10 points in total) Consider the Act Utilitarian framework. In a statement released by OpenAI, the company argues that it would be "impossible" to train effective AI models without copyrighted materials. What would be the strongest Act Utilitarian case you could make to ignore existing copyright protections for training AI models? Please be specific about the stakeholders, their individual welfare impacts, and any additional assumptions you make about the net aggregate impact.

5. (10 points in total) A key argument in favor of copyright is that it incentivizes content producers to generate high-quality content. Would your answer to the previous question change in a Rule Utilitarian thought experiment in which all companies ignored existing copyright protections when training AI models? Please specify any assumptions you make.

6. (20 points in total) Now consider international competition for the AI sector. Imagine two countries choosing whether to regulate the sector and require AI companies to use licensed content when training their algorithms. Both countries choose their action simultaneously. Imagine they face the following strategies and payoffs.

		Country B	
		Regulate	Not Regulate
Country A	Regulate	10, 10	4, 12
	Not Regulate	12, 4	5, 5

Example: when Country A chooses not to regulate and Country B regulates, the payoff to Country A is 12, and the payoff to Country B is 4.

    a. (5 points) What is the outcome if both countries seek to maximize their own payoff? Is this a Nash Equilibrium? Why?

    b. (5 points) In class, we emphasized the importance of recognizing underlying incentive structures. What underlying "archetype" does this game reflect?

    c. (5 points) Assume both countries behaved like Act Utilitarians – would the outcome differ from your answer in a.?

    d. (5 points) Would your previous answer (to c) change if both countries behaved like Rule Utilitarians?

7. (10 points) In your opinion, would "Justice as Fairness" support legislation requiring AI companies to only train their models with licensed content? Explain why or why not. Please specify any assumptions you are making.

8. (10 points) Copyright laws vary across countries. Imagine that country A legally requires AI companies to obtain licenses for all content they use to train their algorithm, while country B has no such legal requirements. Are the following firms practicing corporate social responsibility according to the definition from class? Please explain your reasoning.

    a. (2.5 points) A firm in Country A that, in the name of an ethical objective, only trains its algorithm with licensed content.
    b. (2.5 points) A firm in Country A that trains its algorithm with licensed and unlicensed content.
    c. (2.5 points) A firm in Country B that, in the name of an ethical objective, only trains its algorithm with licensed content.
    d. (2.5 points) A firm in Country B that trains its algorithm with licensed and unlicensed content.

9. (10 points) A newly founded NGO, *Fairly Trained*, now offers certification for companies that trained their AI models exclusively on content that they own, have licensed, or that is in the public domain.[9] Would your conclusion in Question 1 change when considering these certified companies? Why or why not?

---

[9] https://www.wired.com/story/proof-you-can-train-ai-without-slurping-copyrighted-content/