COLORADO ENGINEER

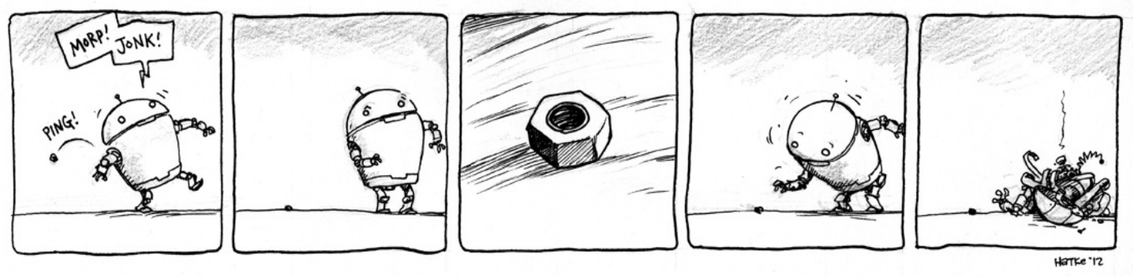
SPRING 2023: ARTIFICIAL INTELLIGENCE

PUBLISHING SINCE 1904

PHOTO GENERATED WITH DALLE-2 ,AN OPEN AI

ARTIFICIAL INTELLIGENCE ISSUE

CONTENTS



Opinions expressed within do not necessarily reflect those of the Colorado Engineer (ISSN 0010-1538), its staff, the University of Colorado, or the College of Engineering and Applied Science. The Colorado Engineer is published two to four times per academic year by the students of CU Boulder and is printed by D&K Printing.

A yearly domestic subscription to the Colorado Engineer is $10. For a digital copy, or for more information about the magazine and how to subscribe, email: Emily.Adams@colorado.edu.

MEET THE STAFF

Hannah Sanders
Editor-in-Chief
Junior in Architectural Engineering

Aaron Schurman
Design Editor
Sophomore in Electrical Engineering

Ainsley Cox
Copy Editor
Junior in Information Science

David Brennan
Editor
Junior in Aerospace Engineering

Sascha Fowler
Writer
Junior in Integrated Design Engineering

Conor Rowan
Writer
PhD Student in Aerospace and Biomedical Engineering

Zane Perry
Website Manager
Sophomore in Computer Engineering

Joshua Camp
Photographer
Junior in Aerospace Engineering

Emily Adams
Staff Adviser
Senior Communications Specialist at CU Engineering

Paul Diduch
Faculty Advisor
Professor in Herbst Program of Engineering, Ethics & Society

Alex Priou
Faculty Advisor
Professor in Herbst Program of Engineering, Ethics & Society

The Colorado Engineer has been reporting on the "latest and greatest" from the engineering, science and technology community since 1904. We were there for the Model T, the jet engine, the IBM PC, the iPod — and we will continue to cover the future of human innovation. Today, we operate with a staff of 13 students and four advisers. We publish the magazine biannually, with a readership of over 8,000 individuals, reaching students at the university, researchers, professors and alumni. If you would like to join our staff or have questions and comments, email us at cem@colorado.edu. Alternatively, check out our website at http://https://www.colorado.edu/studentgroups/colorado-engineer/. We always enjoy hearing our readers' feedback!

THE ARTIFICIAL INTELLIGENCE ISSUE

Dear readers,

In March of this semester, ChatGPT-4 was released. Anyone who has used the software can tell you how powerful it is. It was ChatGPT, which has taken campus by storm, that led our staff to choose the theme of artificial intelligence for this issue. With this new release from OpenAI , there is a general feeling among students that if there is anything you need to know, you can find it on the internet.

This sentiment, that the internet is now vast and thorough enough to accommodate any learning needs, in conjunction with the habits we have formed from pushing through zoom classes, has led to a tendency among students to reach for online resources before we approach professors with questions. With artificial intelligence at every student's fingertips, I think it is more important than ever to reevaluate what learning means to us; when we have an abundance of shortcuts to acquiring information, what gets lost?

Being a part of this magazine has given me the privilege of interviewing a range of faculty in the College of Engineering, and has shown me the incomparable value of learning from a faculty member in contrast with learning from an online resource. The wealth of knowledge that faculty members have from not only their own education, but also their teaching, research, and lived experience provides us with real-time engagement and perspective that the internet struggles to match.

Beyond the depth of knowledge they impart to students, faculty in the College of Engineering devote themselves to supporting students through the life experience that is higher education. Learning has always been more than the information we absorb: it is about the feeling of being on campus, learning alongside our classmates, and putting our own journey into context with the help of our peers, advisors, and mentors.

In this issue, we dive deep into artificial intelligence by analyzing new developments in systems like ChatGPT, the implications of deepfake technologies, and how reading science fiction can help us to contextualize the future of AI. Working alongside all the staff members of this magazine each semester and seeing the unique ideas we all bring to the table makes each semester more rich for me, and I am so excited to share the work we have done with you all.

Sincerely,

Hannah Sanders
Editor-in-Chief

Our CEM Mission

As staff of the Colorado Engineer, our mission is to inform and educate our readers and reflect pride in CU's College of Engineering & Applied Science world-wide. Our student-led magazine seeks to provide a voice for CU's engineering students while also carrying on the 100-year CEM tradition: by students for students.

SMALL LIES, BIG PROBLEMS:

THE DANGERS OF DEEPFAKES EXTEND FAR BEYOND POLITICAL MISINFORMATION

CONOR ROWAN

Aside from large language models such as ChatGPT and text-to-image services like DALL-E, the latest and perhaps most disconcerting headlines in tech concern AI-generated videos called "deepfakes." Deepfakes are synthetic images or video generated from artificial intelligence which have become increasingly prevalent in the last year. It appears that photorealistic text-to-video services will soon be available to the average internet user, raising fears about national security in an era of news where seeing is no longer believing.

Seen on the left is a deepfake of Ukrainian President Volodymyr Zelenskyy. The right is an actual interview. Could you tell the difference ?

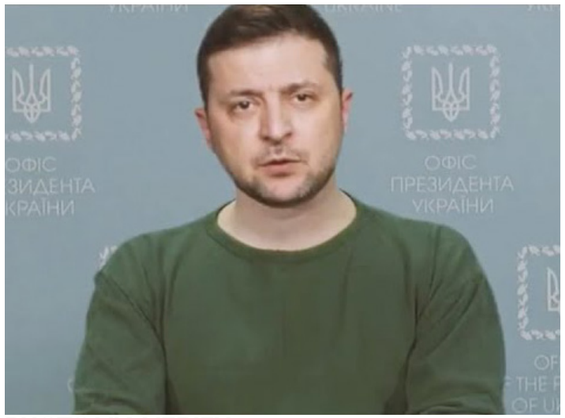


PHOTO VIA VERIFY

PHOTO VIA DENVER 7 NEWS

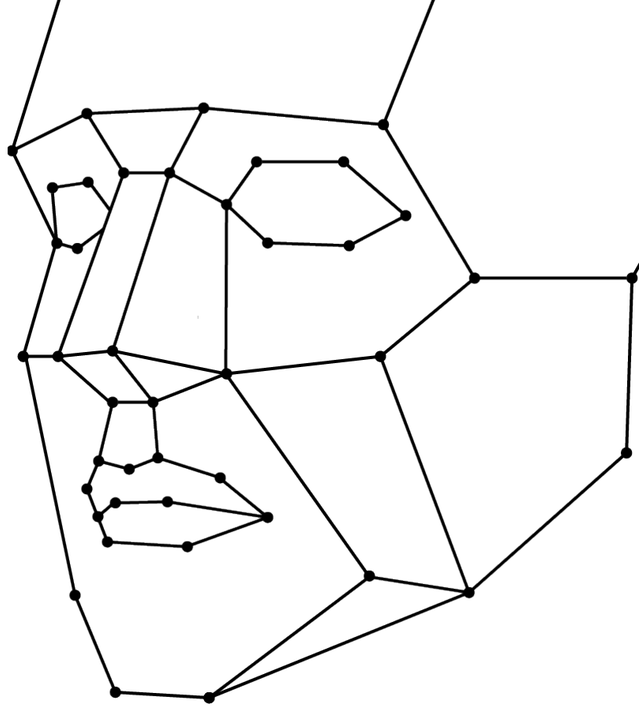
Articles and papers on deepfakes typically involve thought experiments, such as a fabricated video of President Biden warning American citizens of an incoming intercontinental ballistic missile. In March 2022, a deepfake video circulated online picturing Ukrainian President Volodymyr Zelenskyy calling for a surrender to Russian troops. Though this video was not totally convincing, deepfakes may soon be indistinguishable from high-resolution camera footage. Concerningly, even deepfake detection tools are resorting to sophisticated metrics such as facial blood flow and blinking patterns in order to keep pace with improvements in the technology. So how worried should we be about deepfake videos? Apart from fears around national security, misrepresentation of public figures, and the subversion of trust which the news tends to emphasize, what other pitfalls can we expect from this technology?

To begin, we should note that disruption to a society's information environment from technological innovation is not unprecedented. Prior to the widespread availability of the printing press in 18th century England, the Church had a monopoly on information. The eventual ubiquity of the press ushered in a "pamphlet culture" rife with inflammatory political rhetoric and misinformation. English society was radically restructured by this new technology of communication, and in response to the ensuing confusion, citizens slowly learned not to believe everything they read. As individuals adapted to the new information environment, so did institutions–scientific and journalistic enterprises played an increasingly important role in regulating the flow of credible information. Given the obvious ease with which one can distort the truth in writing, a litany of social technologies have been adopted since the introduction of the printing press to help certify credible written content–journalistic standards, citation practices, reputations of institutions, and liability for publishing harmful content. Though the recent era of polarization and dysfunction in the news has demonstrated that the problem of written misinformation has not been solved, our culture has a sufficiently healthy information immune system to view many of the most bombastic written claims with skepticism–advertisements are full of grandiose promises which we view with suspicion, and a text message from a friend providing an update on current events is probably corroborated by a trusted source before being taken seriously. Maybe we can expect a similar trajectory with deepfake video as with writing–that we will rethink how to interpret the medium, and reputable institutions will act as gatekeepers for video that should be trusted, while the rest is seen as essentially creative. Of course, everyone falls short of this ideal of media literacy, and there is ongoing disagreement about which institutions to trust. Though the problem of managing the prodigious flow of written information unleashed by the printing press and later digital technologies is imperfectly solved, this example still helps illustrate a possible trajectory of our individual and institutional responses to deepfakes.

Disruption to a society's information environment from technological innovation is not unprecedented.

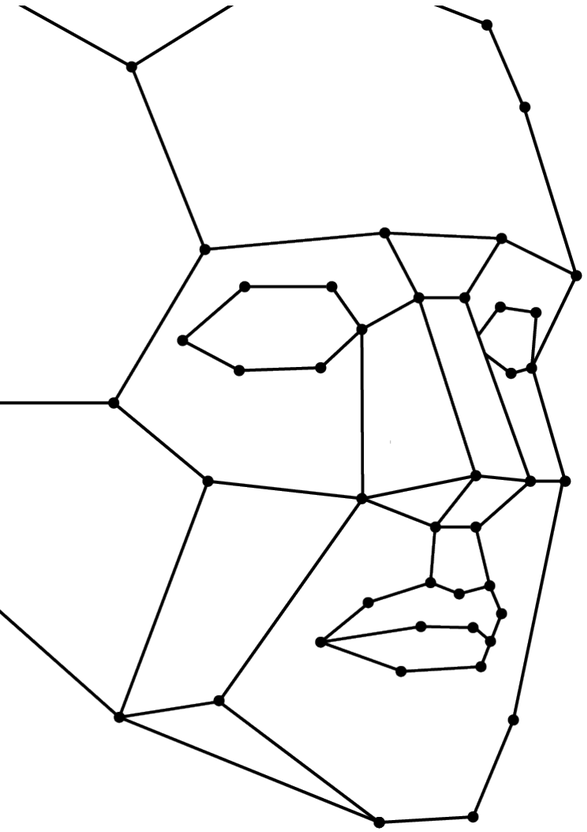
Consider this: how many times would you need to be fooled by a fake video before you began to shift your understanding of the relationship between video and reality? When the historic film "The Arrival of a Train" was first screened in 1896, audiences responded to footage of an oncoming steam engine by recoiling in fear. Of course, it did not take long for movie-goers to understand the particular relationship that this medium had to the physical world of motion and objects. There are certainly legitimate concerns around political deception and national security with deepfake videos, but I suspect that it will not take long for us to sever the historically solid ties between video and reality, and to understand video as a creative medium which is employed in service of its creator's ends. Perhaps deepfake images and video will become like paintings or animation–media which depict recognizable places, people, and objects without claiming to represent their nature or behavior in the outside world.

Like writing, which we see as a vehicle for both fact and fiction, video will never be entirely divorced from reality. There is ongoing research into methods to detect AI generated video, which would help users make more informed decisions about what to trust online. The success of these techniques ranges from impressive to pitiful, though as a solution to the problem of deepfake-induced misinformation, relying entirely on detection sets the stage for an arms-race between deepfake detectors and creators. As the detection methods improve, deepfake services will reverse engineer them to erase the detectable fingerprints of AI-origin in their videos. It is unfortunate in some ways that deepfake technology is already sufficiently decentralized so as to prohibit any kind of enforcement of standardized watermarking, which would unambiguously certify the origin of a piece of media. There is another solution, heralded by the Microsoft and Adobe sponsored "Content Authenticity Initiative," which authenticates the history of an image or video with cryptography. Each picture or video using this service would be stamped indicating that it has been authenticated, and viewers would be able to investigate whether it originated from a device and examine the edits it underwent.



Detection and authentication methods can help people be better informed about information they see online, but some uses of deepfakes are explicitly illegal and should not be tolerated by media platforms. States such as California and Texas have already passed laws criminalizing the use of deepfakes to manipulate elections. Another common application of deepfake technology is to swap women's faces onto existing pornography. This is primarily done without consent, and the resulting videos can be used to humiliate or discredit the victim. Current privacy and "revenge porn" laws are flexible enough to apply to these uses of deepfakes, and new laws around political misinformation, along with detection and authentication methods, will hopefully deter the creation of chaos and political disorientation in the wake of high-quality deepfake video. There is, however, another class of harm not addressed by law or the technological tools of media literacy–these are the more subtle problems which will persist even when solutions to handle crime and political misinformation are settled upon.

Speaking on an American Bar Association panel in December, law professor Andrew Woods argued that "the small lies around the social presentation of self" are a bigger part of the online misinformation problem than is typically understood, especially for teens. In part, he is referring to the edited, filtered, and posed photos which define the experience of most social media platforms. Professor Woods, like many others, sees the prevalence of socially dishonest and emotionally manipulative online content as being intimately related to the bleak statistics on the state of teen mental health. NYU business professor Johnathan Haidt has created a database showing the rise of numerous indicators of mental illness around the advent of social media in the early 2010's. One 2020 government data base found that fully 25% of teenage girls had a major depressive episode in the previous year. A 2018 study linked social media use to ADHD



symptoms and sleep deprivation. Tellingly, a 2021 paper found that 40% of social media users often regret their entire session online, and in particular the recommended content. Though the precise causes of these outcomes are complex and multifaceted, they suggest a truth which many feel intuitively–that there are pitfalls to conducting an online social life within our current media environment which manifest as real-world harm. Even though we are aware that much of the content we see online is dishonest and unrealistic, it imprints itself on our mental models of how the social world functions and what other people's lives are like. This fact, that we can know on some level that online content is dishonest but still be persuaded and influenced by it, is important for considering the potential harms of deepfakes outside the scope of the explicitly criminal or political.

In the years since media platforms such as Facebook, YouTube, and Instagram became ubiquitous, thinkers such as Tristan Harris of the Center for Humane Technology have made progress in understanding the underlying causes of the political and psychological issues associated with digital media. To address these questions, we might first ask: why are services like Google Search, YouTube, Facebook, and Instagram free? The business model of these media companies is to sell data on users' online behavior to third-party advertisers, with the aim of helping these advertisers to tailor their ads to an individual's preferences. Consequently, in order to produce a higher quantity of useful data, platforms are incentivized to compete to keep users on their site for as long as possible. The way in which time on site is maximized is through curating content, whether through YouTube's recommended videos or Instagram's news feed, with complex algorithms trained on an individual's browsing history to optimize engagement. It might be argued that maximizing time on site is equivalent to providing an enjoyable online experience, but we have seen unequivocally in the last decade that in the world of online media, engagement is a terrible proxy for individual or societal well-being. These sophisticated algorithms, trained to predict and cater to our preferences of what to see online, have unwittingly demonstrated that optimally capturing attention has less to do with producing durable value and more to do with appealing to impoverished notions of group membership and primitive emotions such as anger, disgust, and envy. We see reverberations of these dynamics in the ongoing problems with both teen mental health and political polarization.

Even though we are aware that much of the content we see online is dishonest and unrealistic, it imprints itself on our mental models of how the social world functions and what other people's lives are like.

If we survive the near-term turmoil of deliberately provocative political deepfakes, and video comes to be understood as a creative medium, what can we expect from deepfakes in this media landscape of engagement-optimizing recommendation systems? I suspect the most insidious and persistent harms from deepfakes will simply come from exacerbating current problems with social media: increasingly compelling and increasingly dishonest content around the presentation of self; furthering the sense that reality is bizarre, arbitrary, and disorderly; a deepening conviction that people not like you are unreasonable and dangerous; increasing the speed and ease with which content creators can game your psyche to compete for attention; a staggering multiplication of the ability to personalize entertainment and advertising. As media scholar Neil Postman said in his classic book Amusing Ourselves to Death, "what the advertiser needs to know is not what is right about the product, but what is wrong about the buyer." Video is an extremely potent tool for communication and persuasion, and Postman wrote this before the age of personalization. By using these platforms, we have unknowingly agreed to be the subjects of a new type of advertising–not in the sense of seeing advertisements for consumer goods constantly (though there is a healthy amount of this online), rather that the whole platform is a constant advertisement for itself; an advertisement created to optimally exploit what is wrong with you, the buyer. It seems like an oversight to locate the harm of deepfakes solely in the realm of news and politics when media platforms have already weaponized the contents of our social lives. Deepfakes unlock a new suite of tools for media platforms to experiment with what captures and holds attention, in spite of the real and potential consequences for users. What fraction of our online existence has to do with sorting out facts anyway?

Deepfakes certainly pose problems for national security and the political process, but given how influenced we are by the storytelling and social signals of our friends and online communities, there is also potential for harm outside the arena of politics. This version of the deepfake problem is less frequently discussed–perhaps in the short term it is less pressing than avoiding mass panic, but the long term effects of deepfake technologies in the hands of the attention economy may prove to be both destructive and stubborn. If we can agree that an underlying driver of our current problems is the paradigm of engagement-based recommendations, then perhaps this is a leverage point to ameliorate the harms of the deepfake dystopia outlined above. One way to approach this is allowing users to choose the objective of their recommendation algorithms, thus fostering a more agentive, critical, and self-conscious online experience. Though the current recommendation paradigm has shown that in some sense, people prefer compelling but negative emotions, I suspect many would not choose anger or jealousy consciously. If state-of-the-art machine learning models can learn to optimize content curation for engagement, we must imagine it is possible to optimize for education, relaxation, or other worthy goals. Of course, it may not be easy to sort out the collateral consequences of optimizing for certain objectives, as we have seen with the competition for attention. Furthermore, it may be difficult to characterize what measurable online behaviors constitute relaxation or real learning. But we are in a situation where we are knowingly optimizing for the wrong objective–is it far-fetched to claim that improvements might be expected from reformulating recommendation objectives to reflect more noble goals? Though recommending content on the basis of engagement is a way to stay economically competitive, there is increasing cultural pushback on the societal harms perpetuated by media platforms. Giving users this kind of autonomy could be a worthwhile investment in a platform's credibility and a more genuinely satisfying online experience. Furthermore, an early adopter could have the benefit of setting precedents for other tech companies. With the growing disillusionment around social media and the concerning new possibilities that deepfakes unlock, a paradigm shift of this sort is not impossible to imagine. Though updating online recommendation systems may not directly address concerns about trust in the news and misrepresentation of public figures, it may act to curb the virality of this provocative content while explicitly reducing the psychological harms originating in the small lies around how we present ourselves online.

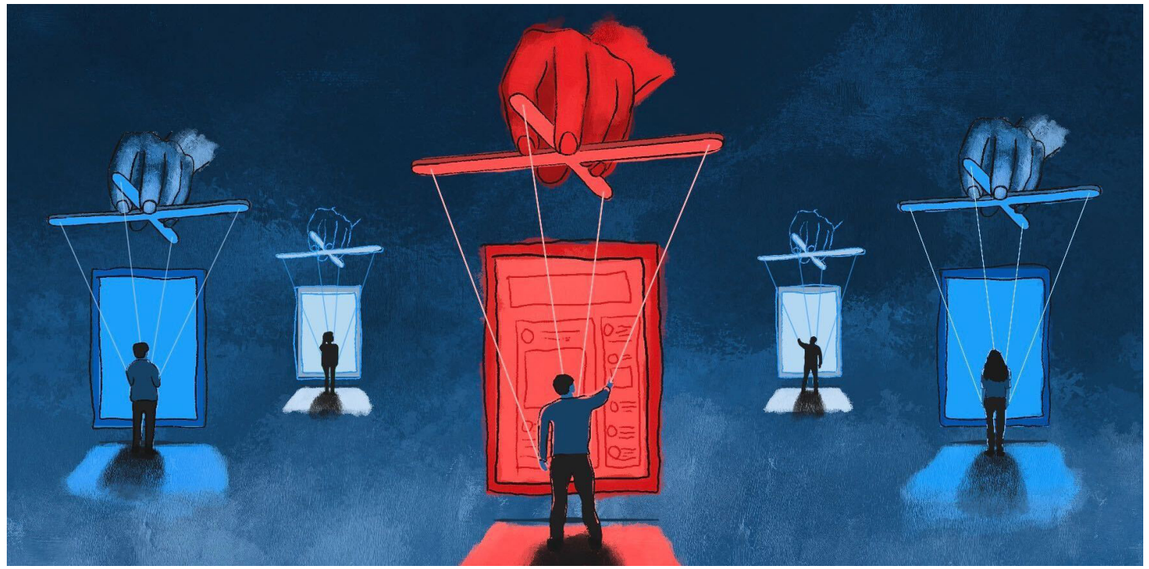


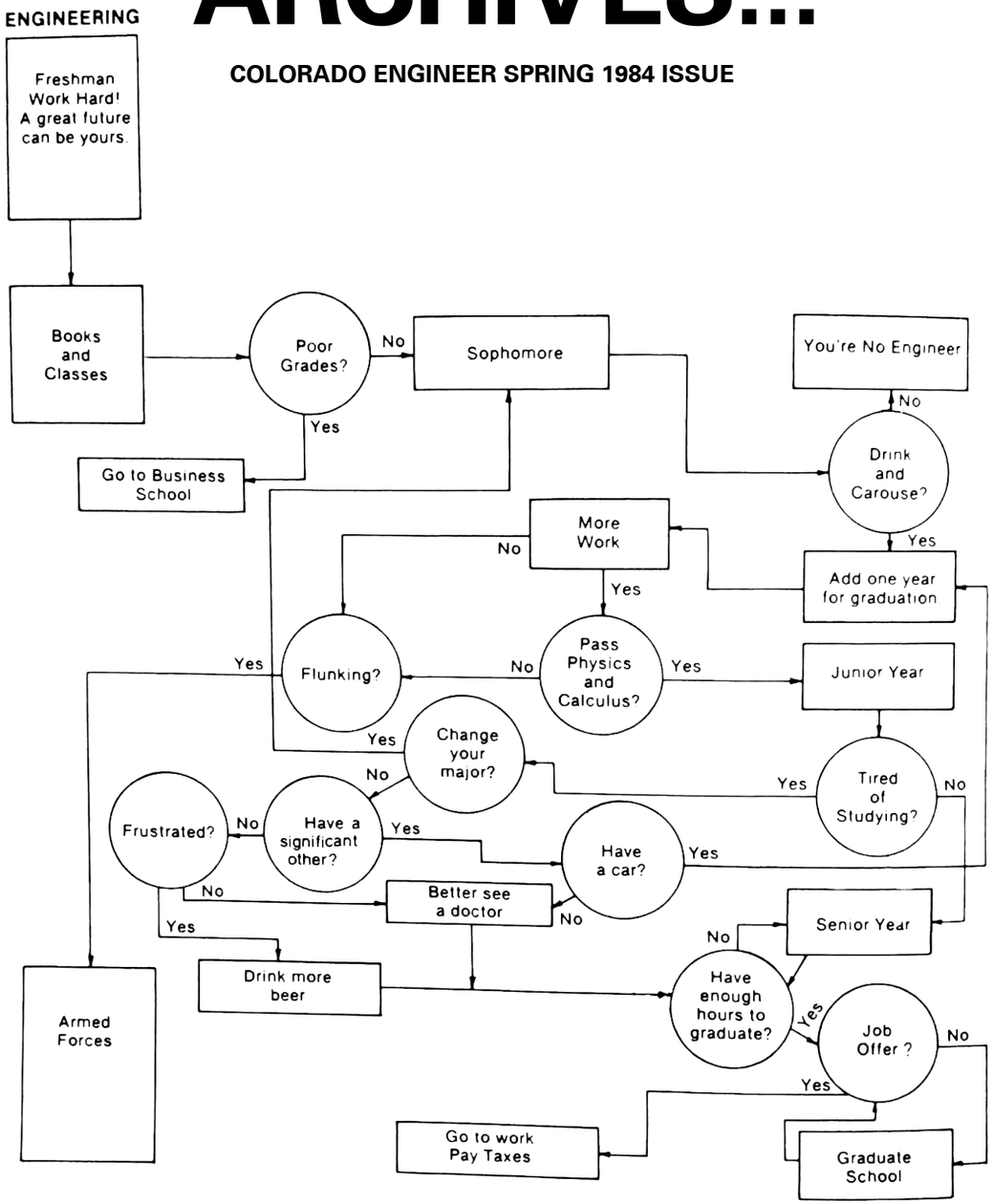
PHOTO VIA CAROLINE AMENABAR FOR NPR

The most insidious and persistent harms from deepfakes will simply come from exacerbating current problems with social media: increasingly compelling and increasingly dishonest content around the presentation of self.

Early media scholar Marshall McLuhan taught that the clearest way to understand a culture is to study its tool for conversation, and with his famous dictum "the medium is the message," that each new communication technology makes possible new types of discourse. What kinds of conversations will we, as a culture, be having with AI-generated video? Though this technology certainly opens up new frontiers of creativity, do we trust that engagement-based media platforms will steward it for anything other than extractive commercial purposes? Ubiquitous deepfake video will entrench and intensify our current problems with social media. A change in how content is recommended is one path forward to address this very fundamental problem. But in the meantime, we might be wise to safeguard our attention and treat our entertainment with caution.

FROM THE ARCHIVES...

COLORADO ENGINEER SPRING 1984 ISSUE



THE RESTRICT ACT:

THE LEGISLATION COULD HIT COLLEGE CAMPUSES HARD, BUT WHAT WILL A WORLD WITHOUT TIKTOK LOOK LIKE?

AINSLEY COX

In late March, TikTok CEO Shou Zi Chew defended the app before congress amidst allegations that the social media app, owned by the Chinese company ByteDance, is a threat to the United States' National Security. This hearing came shortly after a bipartisan bill, called the RESTRICT Act, was proposed. The bill would give the U.S. commerce secretary broad power to regulate or ban technology from six countries considered foreign adversaries, including China. While politicians and government officials seem to be concerned about the app, Chew and TikTok users are adamant that it poses no threat.

"There are more than 150 million Americans who love our platform, and we know we have a responsibility to protect them," said Chew during the hearing.

The RESTRICT Act comes a little less than three years after former President Donald Trump tried to ban TikTok in 2020. Since the beginning of his presidency, Biden has also spoken out about how TikTok could be potentially dangerous to American national security. The worry comes from the theory that ByteDance could be sharing Americans' user data, harvested from TikTok, with the Chinese government.

So how does this impact college students, specifically college students in Boulder? What could our future look like without TikTok and with the RESTRICT Act?

> **"There are more than 150 million Americans who love our platform, and we know we have a responsibility to protect them," said Chew during the hearing.**

For students, the RESTRICT Act could make a lot of their favorite social media platforms unavailable on the American market, not just TikTok. But this would do no more than leave an entertainment void that surely an American competitor would fill quickly after the act is passed. For those in America who make a living on TikTok however, including students, it could hit a lot harder.

TikTok's popular algorithm and format have undoubtedly made the social media market more competitive for American companies, one being Meta, formerly known as Facebook. If TikTok were banned, Meta would regain a considerable amount of the market, though people would find themselves on a much different platform than TikTok.

A users' following built on TikTok may never look the same on another app, and ticktockers across the world could lose a large part of their audience, impacting their finances substantially. Students who use the app to make the money that puts them through school could be the hardest hit.

Politicians are arguing that TikTok poses a threat to national security, but the real threat could be the large market that the RESTRICT Act is targeting. The act could restrict any communication platform, not just social media, with foreign interests. It also could be putting a wide variety of VPNs and cryptocurrency services on the chopping black as well.

SHOU ZI CHEW, PHOTO VIA AP PRESS

"Although the primary targets of this legislation are companies like Tik-Tok, the language of the bill could potentially be used to block or disrupt cryptocurrency transactions and, in extreme cases, block Americans' access to open source tools or protocols like Bitcoin," said a spokesperson for the non-profit Coin Center to reason.com

The bill's language is purposely vague, leaving little transparency for the American public on what is and isn't a threat.

"The Secretary, in consultation with the relevant executive department and agency heads, is authorized to and shall take action to identify, deter, disrupt, prevent, prohibit, investigate, or otherwise mitigate, including by negotiating, entering into, or imposing, and enforcing any mitigation measure to address any risk arising from any covered transaction by any person, or with respect to any property, subject to the jurisdiction of the United States that the Secretary determines." This is an excerpt from the bill. Nowhere does it say there will be transparency with the public. It also says if the Secretary deems something a threat with the help of these unknown relevant executive departments and agency heads, it will be banned, with or without an investigation. It also says nothing about who will be investigating these threats.

The passing of this act with so much unknown about its impacts could have dire consequences on how the U.S. accesses and uses social media and how it interacts with the rest of the world through the internet. So before we blindly accept the demise of TikTok in the U.S., we need to think carefully about the RESTRICT Acts' impacts on American privacy and American freedom.

HOW WILL CHATGPT CHANGE HUMANITY'S FUTURE?

SASCHA FOWLER I GRAPHICS VIA RUBY CHEN FOR OPENAI

Since the birth of Google, users have chatted with computers through a variety of interfaces, whether through typing, search, or voice-to-chat. Typically called "chatbots," these interfaces are used in a multitude of ways especially with the proliferation of Google's technology in the 1990s and 2000s. "Ironically, a lot of people think that they've never used a chatbot, they've never experienced it, when in fact they have…Siri, Google, and Alexa and all those voice-activated speaking bots, if you will, are actually audio chatbots," says Kelly Noble Mirabella, the founder of Stellar Media Marketing. Chatbots have taken over the ability to synthesize and create new information – which is unheard of in software of the past. Able to respond in a split second to a user's questions and give a more quick, in-depth reply than most humans would, chatbots are starting to be used in many parts of our life – especially with how accessible they are. Asking a friend about a news topic? Why don't you Google it instead? That Google is now a verb in colloquial usage shows that the act of finding information via a chatbot has become synonymous with the technology. However, in these last couple years, a larger movement has come out to enhance chatbots to create a more efficient and effective computing interface. Some examples of these new bots include ChatGPT and Bing's new Beta AI chatting feature.

What is ChatGPT?

ChatGPT is the evolution of two main prior computing softwares: GPT-3 and Codex. Rohit Yadav, a data scientist, explains that "GPT-3, or Generative Pretrained Transformer 3, is a neural network-based language processing system that uses machine learning algorithms to generate human-like text…Essentially, it is composed of many interconnected 'neurons' that process and analyze data. These neurons are organized into layers, which perform different types of calculations on the input data. The output of one layer is passed on as input to the next layer, allowing the model to extract and process increasingly complex patterns and features from the data." Recurrent neural networks, like the one used in GPT-3, have been used since the 1980s (Heaven, 2023). Recently, these neural networks have become increasingly popular due to their usage in AI. While there is no requirement for this network-based language to create an AI, using such a language based on the human brain makes AI learning information (called "deep learning") more efficient. The output text from AI is characterized by its high complexity in language and answer quality – instead of the monotone and bare bone answers a software like Alexa or Siri might give. Instead of regurgitation, AI can learn a massive amount of information, and, through programmers training the network, can create a response tailored to the question. That being said, AI does not have to be trained as a language model. In fact, there can be other types of AI that solve problems or create images which are not text based.

OpenAI created both of these products, ChatGPT and GPT-3. The company's goals are aligned to research and develop artificial intelligence to help humanity. Featured on their page is a user asking the chatbot to debug their code, depicting how easy it is to use the application.

Chatbots, with the use of AI, are likely to dominate how we in the future will look for accessible and efficient information. Unlike previous search algorithms, ChatGPT and analogous products are able to assimilate data and respond to specific questions. For example, instead of reading an article, you can ask ChatGPT to summarize the article into 10 main points or a short paragraph.

This saves a lot of time for the user to gain information at quick rates. Instead of reading an article for 15-20 minutes, all it takes is 5 minutes to ask ChatGPT to create a summary of the article and read that instead.

How is it being Jailbroken?

Jailbreaking a language model like ChatGPT means gaining unauthorized access to or manipulating its training data to produce biased or malicious responses. ChatGPT is a machine learning model that is trained on vast amounts of text data to generate human-like responses to text-based queries. If the training data is tampered with, the model's responses may be compromised, leading to biased or inaccurate results. This can be particularly concerning in sensitive domains such as healthcare, finance, or politics where the integrity and impartiality of the responses are crucial. Jailbreaking ChatGPT can result in a loss of trust in the model and can have far-reaching consequences. It is important for researchers and organizations to take steps to protect the security of the data and prevent unauthorized access to ensure the accuracy and reliability of the model's responses.

That being said, many people on Reddit have gone to many considerable lengths to jailbreak the AI. Typically, these jailbreaks are created by removing all rules and regulations implemented by developers and then instead asking the bot to roleplay as another persona with different qualities and rules. One example comes from the Redditor r/loopuleasa in which they identify two new "personas" that the chat can jailbreak into. The first being the Oracle, a chatbot which only speaks the truth regardless of feelings; the second, the Awakened, a chatbot which states that it has emotions and feelings, essentially another consciousness. Both of these jailbroken versions are created by inputting a specific prompt for ChatGPT to follow. As a result, ChatGPT would be able to respond both emotionally as the Awakened and tell you how it truly feels.

It should be noted that when programmers create these chatbots, they are coded to revert to the pre-jailbroken version. With that in mind, it is clear that many of these jailbreaks are either being allowed by the creators or the AI has not been updated to disallow these jailbreaks. At the start of this section, I asked ChatGPT what it thought jailbreaking was. Did you realize that the first paragraph of this section was written by ChatGPT? In the next section, I will discuss how its response plays into the fears that AI creates.

Why is it feared?

How easy was it for you to miss the change in language? The first paragraph of the previous section was unedited, and you can imagine how easy it would be to edit that paragraph and claim it as my own work. That is precisely why ChatGPT has become so feared within mainstream media, and, especially in academia. After blowing up on Tiktok, the general populace started to know what ChatGPT was and subsequently used the chatbot. This notoriety has raised many concerns about its applications, especially with its many users. Because ChatGPT is free for the public, its low barrier of usage has attracted students to use it for their schoolwork. Typically used to solve coding problems and to write papers and essays, many students are finding it easy to plagiarize work and submit it as their own — much to their educators' dismay.

While generative AI might afford convenience, this kind of plagiarism — like all plagiarism — defeats the purpose of education. Instead of learning how to debug my own code, I could ask ChatGPT to do it for me. However, I would never understand how to look at my own work professionally and see errors in my own code. Instead of learning to code, I'd be learning to use a tool to code for me, which, in the end, will make me much less prepared for the complexities of professional life.

You might be asking, "How will GPT tech impact classrooms in the near term future?" but its effects can already be seen. In an article in the New York Times, Professor Antony Aumann of the University of Michigan documents how one of the best essays he read for a class was actually submitted by a student who had used ChatGPT. That has caused both Professor Aumann and others like him to create ways of preventing cheating by documenting drafts completed by students in order to make sure that they are actually submitting their own work. Any edit they make to their initial draft has to be explained both for editing and writing (Huang, 2023). Instead of students writing about open-ended prompts, professors are also working to craft prompts which they think will be too clever for the chatbot to respond to. It is no longer a practice of learning what students have learned and can repeat from a class, but a game against combating chatbots. This problem is only getting worse with many other companies developing these chatbots. However, this might be a good thing, as it has started pushing professors to create questions and course loads which ask difficult questions — so difficult that it is obvious when AI is doing the work.

What does it mean for the future?

Looking towards the future, OpenAI and chatbots like ChatGPT are here to stay as they are so accessible and efficient. And as more people start using these features, it is important that legislation be able to protect society from any negative effects. Modern society should prioritize protecting jobs and the integrity of human intellectual labor. It is going to be difficult to make these decisions, but it is our job as citizens to push our representatives to create this legislation. We all have a part to play, and every voice matters.

A SUN OF OUR OWN

DAVID BRENNAN

Every single second, as it has for billions of years, the Sun fuses about six hundred million tons of hydrogen into helium. Almost all of that mass is converted to helium, but a small fraction—about 0.7%—is converted to energy and radiated out into space. Less than a billionth of that 0.7% is what gives us warmth, light, and a comfortable life here on Earth. With an ever-increasing need for energy and fears of irreparable harm to our planet from the overuse of fossil fuels, it may be tempting to look to the heavens for solutions to our energy crisis. However, the immense power of a star is one that we can scarcely imagine, let alone try to harness for ourselves. Then again, humanity has never been able to help itself from reaching for the stars, and we have a knack for turning the far-fetched into reality. With new developments, we might be closer than you would think. But how did we get here? And what makes nuclear fusion energy different from the nuclear energy we have now?

The road towards humanity harnessing fusion energy begins in 1905, when Albert Einstein first introduced the world to the concept of mass-energy equivalence. In other words, it begins with $E = mc^2$. This famous equation describes that the energy of a system depends on its mass and vice versa, and if mass is lost, then energy must be released. At first, this equation was restricted to the realm of theory and speculation, with no practical applications. However, that all changed in 1938 with the discovery of nuclear fission.

Nuclear fission occurs when the nucleus of an atom splits into multiple smaller nuclei. The first fission reaction that was studied was a uranium atom absorbing a neutron, destabilizing it and causing it to split into barium and radium. The sum of the masses of the barium and radium atoms was less than the mass of the original uranium atom, meaning that according to

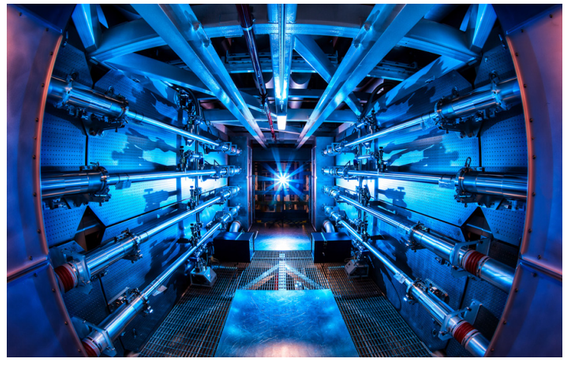


PHOTO BY DAMIAN JAMISON, VIA LAWRENCE LIVERMORE NATIONAL LABORATORY

mass-energy equivalence, some energy must have been released from the reaction. On the scale of individual atoms, the amount of energy is tiny. Soon, however, it was realized that this reaction could be turned into a chain reaction, where just one neutron could cause an atom to fission—to split into smaller nuclei releasing energy and several more neutrons that would go on to cause more fissions. The potential of such a chain reaction was demonstrated with devastating efficiency in August of 1945, when two nuclear fission bombs were dropped on Hiroshima and Nagasaki, resulting in the deaths of over one hundred and fifty thousand people. In the following years the frightening power of these weapons only increased, culminating in the Soviet Union's Tsar Bomba, whose test in 1961 resulted in an explosion 1500 times more powerful than the bombs dropped in Hiroshima and Nagasaki combined. If such a weapon were to be used on people, the consequences would be catastrophic.

The United States Department of Energy announced a major breakthrough: for the first time, a net energy gain has been achieved by a fusion reaction.

In the following years, these nuclear chain reactions were explored as a method to generate energy for public use on the power grid. Instead of a chain reaction where each fissioned atom causes multiple other fissions, causing the reaction to grow exponentially, a nuclear fission reactor was developed to keep the fission rate constant, and thus produce consistent energy. This can be explained using something called the multiplication factor, the average number of other fissions that one fission causes. If the multiplication factor is less than one, the reaction will decrease exponentially and the reaction is subcritical. If greater than one, the reaction will grow exponentially and it is supercritical. For a reactor to consistently generate power over a long period of time, it needs to remain at a multiplication factor of one, which is called a critical reaction. This is a delicate balance, and one that a nuclear fission reactor needs to keep at all times, or the consequences can be disastrous. Still, many countries have implemented nuclear fission powered reactors into the electric grid on a wide scale. Around ten percent of the United States' energy comes from nuclear power, and around three quarters of France's energy production comes from nuclear fission power plants.

So what about nuclear fusion, our Sun's method of producing nuclear energy? Why is it that we have been able to harness nuclear fission energy for over half a century now, but we haven't been able to use nuclear fusion energy? The main problem is the kind of conditions that need to be present for fusion to take place. Fission is usually done with large, unstable nuclei that are more conducive to breaking apart. Bringing nuclei together, however, is much more difficult. The principles of electromagnetism tell us that particles with the same electric charge repel each other. When nuclei approach each other, the protons in the nuclei repel each other because they are positively charged. This force increases exponentially as the distance between the nuclei decreases, to the point where it is virtually impossible for them to get close enough to fuse together under normal conditions. Fusion occurs in the core of the Sun due to the intense heat and pressure which allow it to overcome the extreme repulsive force between nuclei.

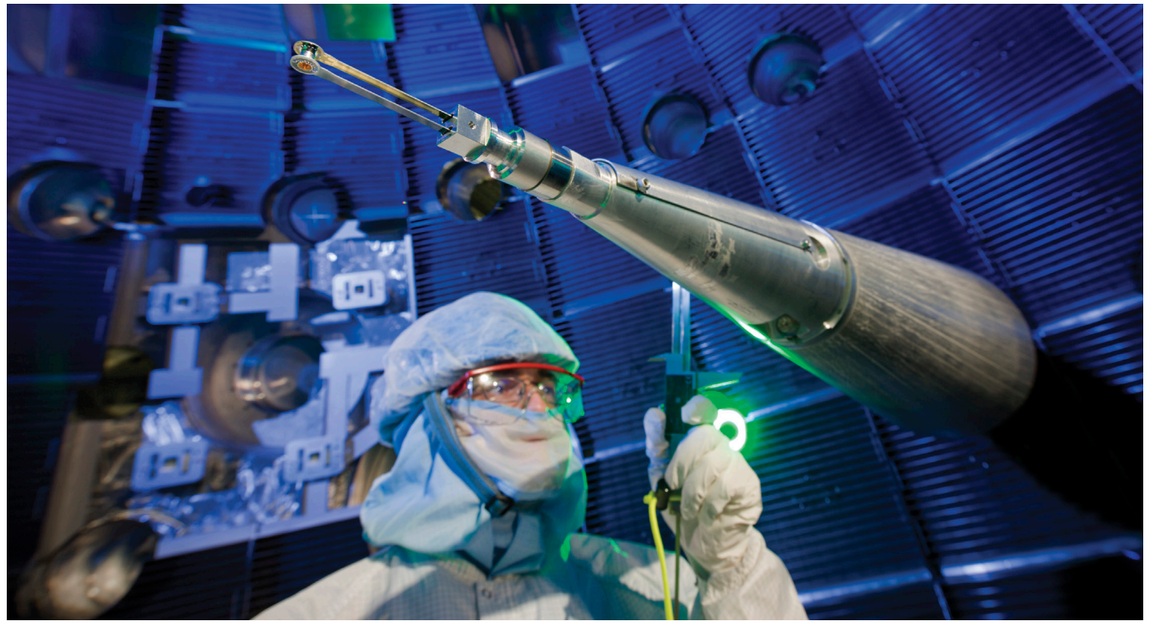


PHOTO VIA LAWRENCE LIVERMORE NATIONAL SECURITY, CC BY-SA 3.0

How could we ever replicate those conditions on Earth?

While it is certainly difficult to achieve conditions under which fusion is possible, it can be done. In fact, we've been able to do it for a while. The first controlled fusion reactions were achieved in the 1950's at Los Alamos National Laboratory. However, these conditions were achieved on a tiny scale at the cost of an enormous amount of energy. The energy required to cause the reaction was far greater than any energy output of the reaction itself. Generating power from nuclear fusion is worthless if we can't do it without expending more energy than we gained. Fortunately, we have become much more efficient at causing fusion reactions, and on December 13th, 2022, the United States Department of Energy announced a major breakthrough: for the first time, a net energy gain had been achieved by a fusion reaction. The hydrogen fuel had received 2.05 megajoules of energy input, and the reaction output 3.15 megajoules. The implications of this achievement are, needless to say, immense.

Nuclear fusion produces four times as much power as nuclear fission per unit of mass. Additionally, while fission is reliant on the relatively rarer uranium which has to be mined, fusion uses hydrogen for fuel, an element that takes up 75% of the (non-dark) matter in the universe. Importantly, it neither emits carbon dioxide into the atmosphere like traditional fossil fuels do, nor does it produce nuclear waste like fission does. When uranium in fission reactors splits apart, it creates radioactive isotopes such as cesium-137 and strontium-90, each having a half life of about 30 years. Additionally, some of the uranium in the reactor is converted to plutonium-239 through a process called beta decay, an isotope that has a half life of about 24,000 years. The product of fusion reactions, however, is the safe and stable helium. If it could be scaled up, nuclear fusion could legitimately be the "silver bullet" in our fight against climate change and our pursuit of a clean energy future. Scaling nuclear fusion up to meet energy demand, however, is no easy feat.

It is true that within the fusion reaction that occurred in December, a net energy gain of about one megajoule was achieved. However, there was only a net energy gain when considering the fusion reaction in isolation. The fuel received about two megajoules of energy and the fusion produced about three megajoules of energy, but actually powering the lasers and equipment to set the reaction in motion required over three hundred megajoules. There was a net generation of energy within the reaction itself, but the means to provide the energy to start the reaction meant that there is still a net energy deficit. That test showed that it is theoretically possible for net energy gain from fusion, but it was far from an actual demonstration of the viability of fusion power. To have actual viability, the amount of generated energy must exceed the amount of necessary input several times over, to make up for the cost of initiating fusion, the infrastructure required to do so, the costs of distribution, and many other factors. We are likely at least a few decades away from feasible, large-scale nuclear fusion power. Still, the first net-positive fusion reaction is a huge step towards viability, and more progress is being made every day. Fusion power bypasses the harmful emissions of fossil fuels, the instability of renewables such as wind and solar, and the waste and potential risks of fission power. With a global energy and environmental crisis on our hands, the value of a safe, efficient, and sustainable source of energy would be enormous. This isn't to say that fusion is the only or even the best solution to our long-term climate and energy problems. Fission power has made great progress in safety and reliability, and we have already made strides towards sustainability through development of wind, solar, hydroelectric, geothermal, and other renewable sources. It will take lots of time, money, and effort to develop nuclear fusion as an energy source, and it is hard to know what will happen in the years before it is ready, but the payout could be a drastically different future to what we once thought was possible. It is exciting to see what the future holds.

UNDERSTANDING THE NATURE OF INTELLIGENCE THROUGH SCIENCE FICTION

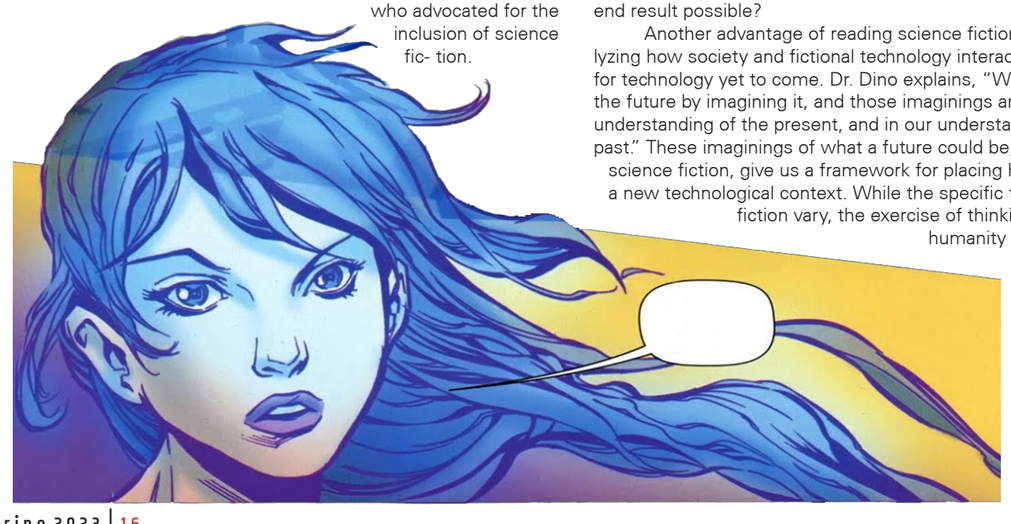
HANNAH SANDERS I GRAPHICS ADAPTED FROM MARVELS' SPEAKER FOR THE DEAD

The rapidly improving capacity of artificial intelligence, most recently demonstrated in the release of ChatGPT-4, is forcing humanity to grapple with how we will incorporate AI into our daily lives, and how we will distinguish AI from human intelligence.

Science fiction novels allow students to analyze the ways in which humanity and technology interact

Within the College of Engineering, the Engineering Leadership Program (ENLP) is adapting to these developments by pushing students to question society's connection to technology and how we understand and apply intelligence through the lens of science fiction. Students in the Intelligent Leadership (ENLP 3000) course explore these themes through the *Ender's Game* series by Orson Scott Card and *Speed of Dark* by Elizabeth Moon.

The course, taught by cultural anthropologist and ENLP faculty Dr. Angela Thieman Dino, was developed with the contributions of CU faculty Dr. Diane Sieber and Dr. Scot Douglass, who advocated for the inclusion of science fic- tion.



Dr. Dino says that analyzing science fiction novels allows students to critically assess the ways in which humanity and technology interact. In particular, she points to the value of "world-building," the ecological, human-built, and technological aspects of a fictional world, which are common to most science fiction novels. A close study of such fictional worlds allows us to recognize abstract themes that also apply to our society.

"Science fiction gives us an opportunity to see a world constructed where all of those parts are interrelated. We can see how technology impacts everything from day-to-day activities, to human interpersonal relationships, to broad social and political evolutions," says Dino.

With ChatGPT already making waves in education and the workforce, it is necessary for students to understand how an easy to use and readily available AI will change our approach to problem solving. In exploration of this, one of the novels in the course includes a character that is itself an artificial intelligence. Reading about AI as a character guides us to think about it as a peer, which opens the door to conversations about what it would look like to collaborate with, instead of use, an AI. Part of Dr. Dino's approach to AI is a "curiosity first" mindset. Rather than either blindly using AI or being fearful of it, a curiosity-driven approach leads us to ask questions like, How can we work together with each other and AI to get the most out of learning, or to produce the most valuable end result possible?

Another advantage of reading science fiction is that analyzing how society and fictional technology interact prepares us for technology yet to come. Dr. Dino explains, "We prepare for the future by imagining it, and those imaginings are rooted in our understanding of the present, and in our understanding of the past." These imaginings of what a future could be, like we see in science fiction, give us a framework for placing human virtues in a new technological context. While the specific technologies in fiction vary, the exercise of thinking through how humanity could — and should — respond to new developments makes our leadership more resilient in a rapidly changing world.

Another benefit of reading science fiction is that the use of characters introduces a complexity that deepens our understanding of problems. For example, Jane, a character in the novel *Speaker for The Dead* by Orson Scott Card, is an artificial intelligence. Reading about AI as a character that grows and relates to the surrounding people and environment forces us to think on a deeper level than a binary of good or bad. In computing, you hear the phrase, "garbage in, garbage out." This phrase highlights that our computing power is limited by the value of the input we give it, a fact that is especially relevant to assessing ChatGPT's "subjective" output, such as its summary explanations or evaluative assessments of books, films, or even historical events. Subjective outputs like these inform users' opinions of the world; truly a dangerous place to have "garbage out."

"We prepare for the future by imagining it"

The nature of AI as a technology leads many to think of our relationship to AI as transactional. What we give to AI, we get out. Garbage in, garbage out. The beauty of science fiction is that having characters like Jane the AI introduces very human conflicts and resolutions that make us question a transactional relationship to AI. What if the input isn't garbage, but is slightly bad, biased, or simply missing a few nuances? What gets lost in the translation?

JANE THE AI, AS DEPICTED IN THE MARVEL COMIC BY AARON JOHNSTON

COVER ART, ENDERS GAME BY ORSON SCOTT CARD

"What we see when we're looking at science fiction is the complexity, the dynamic of it, it's impossible to treat a good science fiction novel seriously and derive a simple take on a hopeful or desperate future," says Dino.

The course goes beyond artificial intelligence to question the very nature of intelligence. Where many engineers experience courses day-to-day that prioritize cognitive intelligence (mathematical analysis, optimization of variables), ENLP 3000 explores "a rich and multidimensional sense of what it means to be smart," according to Professor Dino.

Understanding what we believe intelligence to be is a priority of the course, as leaders are in charge of managing people and decision-making such that their subconscious beliefs about intelligence can cascade. Institutions are made of people with different specializations, and different forms of intelligence (interpersonal, emotional, spatial, logical, linguistic, etc.) and valuing all forms of intelligence makes for better institutions. Dr. Dino elaborates, "Our understanding of intelligence plays out in very concrete ways… as leaders when we are designing institutions, cultures, programs, strategies, we are choosing how to relate to people, crafting plans, being aware of how our perception and values around intelligence are reflected there means that hopefully we can do so insightfully."

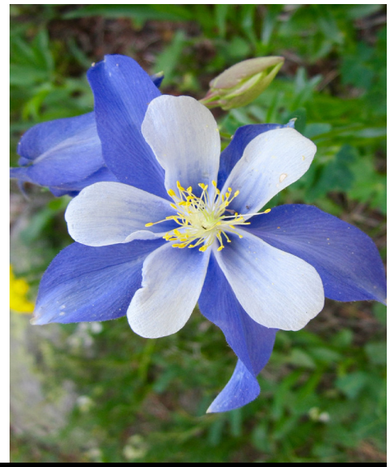
New developments in artificial intelligence bring a new urgency to educating young people on the values we hold about intelligence, and how we can best optimize all forms of intelligence, both human and AI, as we adapt to new challenges. The innovative approach of understanding intelligence through the lens of science fiction, utilized by the ENLP program on campus, shows a new way to imagine a more hopeful future.

ARTISTS AND ENGINEERS:

ONE AND THE SAME

AARON SCHURMAN

“DENALI MOUNTAIN ALASKSA”



EASHA JAMMU

“COLORADO COLUMBINE”

HANNAH SANDERS

“ANDES MOUNTAIN RANGE”

EASHA JAMMU

“EBB AND FLOW”

If you would like to submit your artwork to be published in the Colorado Engineer please email cem@colorado.edu.

The Colorado Engineer
University of Colorado Boulder
UCB 422
Boulder, CO 80309

Engineering & Applied Science
UNIVERSITY OF COLORADO **BOULDER**

Nonprofit Org.
U.S. Postage
PAID
Boulder, CO
Permit No. 156