

The Impact of AI on Future Education

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Jacob Clatterbuck
Spring, 2021

On my honor as a University Student, I have neither given nor received
unauthorized aid on this assignment as defined by the Honor Guidelines for
Thesis-Related Assignments



Signature _____ Date 5/8/2021
Jacob Clatterbuck

Approved _____ Date _____
Tsai-Hsuan Ku

Introduction

Recently, the field of natural language processing has made a huge leap thanks to the release of OpenAI's Generative Pre-trained Transformer 3 (GPT-3) in May of 2020. GPT-3 is a neural-net based language prediction model that can be used to generate text and essays that are difficult to distinguish from those that are human made. The potential benefit this AI can offer humanity is extremely promising, with individuals already having been able to guide it to generate Javascript code from basic instructions. With minimal human guidance, GPT-3 has also been used to generate entire blog posts and articles that are functionally indistinguishable from human writing made from scratch (Philip 2020). While this presents an incredible opportunity to industry, it also contains a potential risk for academic integrity. Basic high school essays can be generated in seconds with a prompt of a few sentences. Even as early as 2005, a far weaker language generator called SCIgen was used by MIT students to trick an academic journal into accepting a fake paper made by AI. The technological artifact of GPT-3 represents one of many advancements in AI that can challenge future educational systems. The main topic I wish to explore is the impact generative AIs like these will have on academic institutions.

While language generation technologies are still in their infancy, the educational world has already taken steps to find uses for it. Some schools and universities have installed automated teaching assistants to reduce the workload for teachers (Preston 2020). Others abroad in China have more recently used AI systems to ensure students are present and learning during online learning in the midst of the COVID-19 pandemic (Liu 2020). To further explore the impact of this technology on the relevant stakeholders, research was conducted through interviews and surveys of relevant stakeholders such as students and teachers. Interviews were held with students that made use of automated teaching assistant tools, and a survey of teachers was held to determine the efficacy of using text-generating AI to create essays for secondary school academic submission.

Literature Review

The existing literature on the topic of AI in education can be grouped into four major themes: overviews of the current status, papers analyzing the potential benefits, arguments for what the future of AIED will look like, and information that raises caution about potential risks. Current literature is stratified into these categories few take a holistic approach and with the rapidly advancing field of AI, there are gaps in the research.

In terms of an overview of modern uses, Lu, J. and Harris, L. (2018) offer the most succinct paper that addresses a variety of issues. The purpose of their article and organization is to provide a nonpartisan information source to members of Congress in an easily consumable manner, and it is therefore brief and focuses on American values and interests. It gives examples of current uses of AI in education such as Intelligent Tutoring Systems (ITS), Personalized Learning, Testing, and Task Automation. Policy considerations are also enumerated so that legislators may seek to take action on protecting student privacy and product procurement. This source is valuable for its legal analysis and ease of understanding, and is particularly useful for addressing research questions about the current use of AI in education. For a more in-depth overview of ITS, Beck, Stern, & Haugsjaa (1996) offer detailed descriptions of the different kinds of models in AI tutor systems and how they can be combined to create an effective ITS. This is an older source is invaluable for its breakdown on a commonly used application of AI in education, but does not, due to its age, provide information on newer technologies. For a more modern source with both breadth and depth covering the current status of AIED, Luckin, Holmes, Griffiths, & Forcier (2016) enumerate in their book the current uses, models, and benefits. In

their paper, they closely examine the efficacy of current AI models used to provide feedback to students and discuss the possibility of robots replacing teachers altogether.

Another theme of the academic literature that some of these sources fall into is that of benefits analysis research. Perotta and Selwyn (2019) give a detailed examination of scholarly work carried out on the effect of deep learning to predict educational performance. Their meta study analyses the existing research to find patterns in how much success students find with the help of AI in different educational topics, finding effectiveness in some subjects, especially those with memorization based needs. Perotta and Selwyn (2019) also look at demographic differences such as race and class to find differences in benefits gained. They conclude that there are different relations between different versions of learning and they question the scientific objectivity and neutrality of AIED systems. For a more positive outlook, Zhu & Zhang (2010), argue that ITS can help teachers by dynamically distinguishing in which concepts a student has shortcomings, finding that the immediate feedback AI can offer is unattainable through hand grading and other traditional methods of evaluation.

There are also associated risks with the increased usage of AI systems in education, both in the form of privacy and plagiarism concerns. In their paper, Arai & Matsuzaki (2014), argue that AI systems in 2016 were capable of passing the entrance examinations of more than half the universities in Japan. They raise concerns over what this means in terms of the value of what is being tested, implying that education will have to move away from the sorts of tasks that artificial intelligence can do with ease. Nils & Mossink (2020) performed an experiment with a state-of-the-art Natural Language Generation algorithm that showed humans are unable to tell AI made poetry apart from that written by humans. They argue that the implications of this study are numerous, including raising questions about academic fraud and how much the developer and training data authors should be credited. For non-poetry essays, Philip (2020) offers an anecdotal account of GPT-3 being used on the social media site, Reddit, and going undetected for some time. This source offers an example of how modern AI uses can

be used for academic fraud to generate, at times, passable papers and short stories that are untraceable with normal plagiarism detection algorithms. The earlier mentioned congressional research (Lu & Harris, 2018) offers an account of privacy concerns for students that may be in conflict with existing laws.

A last categorical theme found in the retrieved sources are those articles that speculate and hypothesize about what the next phase of AIED should or might look like. Andriessen and Sandberg (1999) argue that the role of AI in education is due a fundamental shift away from an ITS-based paradigm and towards a model that helps students “learn how to learn”. They present the compelling evidence that ITS helps with rote memorization, but argue for AI systems that can help with scenarios that deal with non-fixated learning goals. Neil Selwyn (2019) in his book makes the case for human teachers and claims that AI is a “double-edged sword”. While there are some fields that AI excels in, such as immediate feedback, there are downsides that include loss of jobs, heightened differences between different types of learners such as visual learners and social learners, and data privacy concerns. The existing literature is in agreement that current technology is no where near replacing human teachers, but AI aids can be an invaluable resource.

There was at the time of the beginning of research for this thesis, no source that looked into the grades that AI generated essays could receive in an academic setting. This research seeks to correct that, filling a hole in human knowledge that is now relevant thanks to the increasing ability of generative AI. Perhaps because of the stake so many institutions might have in the potential answer to this question, similar research was performed and released by the education resource site EduRef with a panel of professors as opposed to high school teachers.

STS Framework and Method

Of the STS frameworks used to analyze this topic, three in particular stood out as highly relevant, Winner's politics of artifacts framework, Hughes' LTS framework, and Pinch and Bijker's SCOT framework. Winner's framework is particularly useful for addressing the research questions focused on the deployment and availability of the technology for individual AI systems such as GPT-3. These deployments have relevance to macro politics through the lens of resource distribution of different school systems, creating potential differentiation between classes and locations. Hughes' large technical system framework is well suited to addressing the research questions pertaining to the technological momentum of AI's adoption in education and learning institutions. It is also useful for investigation into reverse salients that may impede the development and incorporation of these technologies. LTS analysis can also provide clarity into why some countries are being faster to adopt AIED than others through its reverse salient concept. Perhaps most relevant to the topic at hand is the framework of SCOT. With SCOT, the technological artifacts can be interpreted through the different lenses of stakeholders in the education system such as parents, teachers, and of course students.

Politics of artifacts:

Here the GPT-3 technology will be examined through the politics of artifacts framework of Langdon Winner. GPT-3 was made following the success of GPT-2 to serve as an experiment for the scalability of the Transformer model of neural networks first published in 2017. GPT-2 was able to successfully incorporate the Transformer model, with benefits such as eliminating the need for human oversight during the training of the model. After the concept was proven, the creators at OpenAI started work on a version that would be over 100 times larger by number of parameters, requiring significantly more processing power and time to develop.

The technology is owned by OpenAI, a for-profit artificial intelligence research laboratory based in San Francisco. Individuals who wish to use GPT-3 may contact OpenAI to request access to

the language model for use in research or project development, but it is not currently available and a waitlist currently exists preventing immediate access. Instead of open-sourcing the models as the company name might suggest is the usual intention, OpenAI has only released access to an API, claiming that this decision is based on ensuring funding and for reserving the ability to restrict access if harmful applications are discovered.

All decisions are completely controlled by the company that owns the model, OpenAI. Early adopters and researchers were included as they requested access before GPT-3 made headlines. For businesses seeking greater bandwidth and more processing through the API can pay more to have the access they seek, assuming they can afford it. An average, everyday person is currently totally excluded from using this artifact of technology themselves as they can't pay for direct access and are too late to get on the waitlist.

Large Technical System:

Modern natural language processing (NLP) models and their potential usage for malfeasance, academic fraud, and plagiarism can be viewed through Hughes' large technical system framework. The technology has a messy and incomplete relationship with society as it hasn't yet been fully deployed for societal use and the technology is advancing faster than laws and civil rules can keep up.

Elaborating on the sociotechnical complex through Hughes' system will hopefully shed some light on the relation between advancing NLP technology and humanity as a whole.

The system builders are the inventors and researchers of AI, along with the teams, companies, and governments that support them such as managers, investors, and nations with military interests in AI. The field of natural language processing began in the 1950's with Alan Turing's idea of the Turing test to measure machine intelligence. Since his time, the system builders changed to mathematicians and philosophers who defined the problem of understanding language in discrete terms. As

computational power advanced, system builders became computer scientists beginning in the 90s. In the near future, legislators may have a hand in being system builders for the laws that might govern usage of NLP systems and limit their potential harm to humanity.

The idea of technological momentum is highly present when dealing with computer-based technology and NLP is no exception. Moore's Law, which states that the computing power per area of computer chip doubles every two years, would seem to suggest that there is an inevitable advance of NLP capabilities in modern technology. However, historically, this has not been the case and the path of evolution has been sporadic, contingent on breakthroughs with AI models such as neural nets and reaching a critical state of computing power. There was competition through different models being used to tackle NLP problems such as statistical methods like Markov Chains. Neural networks proved superior for a variety of reasons and the field consolidated into research on the most effective neural nets.

Reverse salients can be seen both in the history of natural language processing and in the weaknesses of modern day NLP algorithms. Certain tasks have been hurdles that AI struggled with more than others. For instance, optical character recognition, or extracting text data from pictures, is a mostly solved problem, with progress made as early as the 1970s. Machine translation, the automatic translation between languages, is still a difficult problem that companies are paying millions to try and solve. For the era before the 90s, the major reverse salient was the lack of computing power. Time and Moore's Law eventually provided a solution through stronger and cheaper computers. A modern reverse salient seen in GPT-3 is the underperformance of its ability to recognize humor such as puns or sound such as rhyming. There is also a reverse salient in the social system as laws lag behind technological development.

Social Construction of Technology:

Relevant social groups in this topic include students, teachers, minorities that aren't fluent in English (in the case of American classrooms), parents of students, and to a lesser degree, workers at sites that write and sell essays to students looking to get out of work. Among these groups, teachers and educational professionals hold the most influence over what is allowed and acceptable for AI use in schools.

Each of these social groups can see AI use in the classroom in different ways, and this can be examined through the lens of interpretational flexibility. Teachers may understand the use of AI as a time saver through automated tools and as a potential source of cheating. A good design to them may be one that focuses on AI tools that can offer quick, accurate answers to simple questions to save the teacher the effort. Students may see it as a cheating resource in the worst cases and as a tool for learning faster in the best cases. They may prefer designs that offer longer answers of variable complexity to match their education level, and those looking to illicitly generate essays may wish for text generation that is undetectable. The goal of parents of students is likely to be focused on the increased learning potential that this technology can offer their children. Students in ESL programs may see text generating AIs as a way to bridge the gap between their command of English and a native speaker's, preferring a design that offers translation services and well controlled grammar. Illicit sites that sell essays to students wishing to cheat could see this technology as a business opportunity while their workers see it as a threat to their jobs.

There is both agreement and disagreement on how this technology is interpreted by the relevant social groups. Both students and teachers have an interest in AI as a way to generate quick answers to questions. Students with lacking academic integrity have a shared interest with "essay farms" in AI-generated essays. Parents, teachers, and existing essay farms all have a shared interest in limiting the potential for nefarious uses of the technology, as it could interfere with students' learning or pose a threat to job security. There's disagreement between parties that are interested in this as a tool make

learning easier and those that are interested in having it do the work for them. As these social groups are not even in power, teachers have a clear authority over students, negotiation between groups may be more likely to fall on the side of the educational professionals getting what they want.

As this technology is still developing and the potential uses, both positive and negative, are becoming more clear, some pathways to closure can be theorized. Laws or commercial practices that limit the availability of generative AIs can provide a top down mechanism to guide the interpretation of this technology as a tool for learning as opposed to a tool for cheating. Teachers can also use their authority to force closure as they are ultimately the ones who decide what is allowed inside their classrooms.

Method:

For my research I wish to make use of several methods that may help provide information of the current state of AI in education. Intelligent Tutoring Systems are the most broadly used example of AI in education. Interviews with those who have interacted with them can help shed some light on how helpful they were, along with any problems in the adoption of the technology. Ideally, I want the input of multiple stakeholders such as students, teachers, and teaching assistants. Retrieving and analyzing the syllabus for classes taught with AI assistance and comparing them to a normally taught class's syllabus may also provide some insight. Additionally I aim to perform document analysis on existing educational laws and standards of education that may affect the adoption of AI in the classroom. While there have been studies showing that certain modern Natural Language Processing algorithms cannot be distinguished from some human writing under some conditions, there is no data as to how these AI-generated essays might be graded if used for academic fraud. A survey of middle and high school teachers would therefore be an appropriate way to determine the efficacy of such a scheme, gathering data on the letter grade an AI-made essay might receive at different grade levels.

Data Analysis

Through collection of existing documents and research in addition to original research and data collection, much progress has been made to addressing the research questions. Interviews and document inspection have shed light on the current state and value of existing AI education methods and technologies. A survey of secondary school teachers at Lake Braddock and Robinson high schools has provided insight into addressing the research questions pertaining to potential risk of AI usage in schools, especially AI-generated essays used for plagiarism. Additionally, by coincidence, research has been conducted and released after the creation of the thesis prospectus on the same topic. Document and data analysis will also be conducted on the results of this parallel research.

In an interview with a Georgia Tech grad student who attends a class that uses an AI-based teaching assistant, information was gained pertaining to the current usage and efficacy of automated aid in modern learning environments. Georgia Tech makes use of an artificially intelligent social agent called Jill Watson. This social agent is trained by professors and uses over 4 years of responses and feedback to cater its chat system to specific classes. Jill is made to be capable of answering both predictable questions such as, “What percentage of our final grade is determined by this next test?” and more open ended questions such as, “What is a sigmoid function?” The interviewee concluded that while their work quality was not affected, there was significant time saved by using the AI teaching assistant for common questions pertaining to scheduling and grading.

The survey of high school teachers found that an AI generated essay can receive a passable grade in secondary schools. Grades ranged from C- to A-, with the essay generally receiving higher grades at lower grade levels. The concern that AI can be used for plagiarism and in essay farms is therefore not entirely unfounded. Feedback from the teachers often commented on the lack of cohesiveness between passages, implying that AI text generation is still imperfect and incapable of

replicating exactly realistic essays that mimic human writing. Cohesiveness in this case meaning that passages and paragraphs at times did not have smooth transitions and were often jarring to the reader. Coincidentally, in the time between the submission of the thesis prospectus and the data collection, an independent study was conducted on this same topic. EduRef conducted a survey on college professors being given more tailored and curated GPT-3 made essays on several subjects. The results were similar in that it was found that AI generated school work can very likely receive a passing grade and is unable to be found by typical plagiarism detection software. However, the college professors gave generally lower grades to the submitted essay than the high school teachers that this thesis was concerned with.

To provide an example of an essay generated by GPT-3 used in these research experiments, the following section contains one such submission:

Example Essay:

In this paper, I will be addressing the issue of American exceptionalism. I will first define American exceptionalism and then give examples of what it silences. Finally, I will discuss the possibilities of creating productive communities and identities that defy a tradition of subjugation.

American exceptionalism can be defined as “the idea that America is different from other nations and superior to them” (Rogers, 2006). This idea is based on two main ideas: one being that the U.S. is a place where people can achieve their dreams through hard work; and two being that there are certain principles that make America great such as liberty, democracy, freedom, and equality (Rogers, 2006). These ideas are not necessarily true; however they are still held by many Americans today.

In regards to what American exceptionalism silences, it silences those who do not fit into the narrative of America being a land of opportunity for all. It silences those who do not have power or who do not have access to power. For example, in 1789 when Congress passed legislation to allow African Americans to become citizens of the United States it was seen as a great step forward in American history; however African Americans were still treated differently than white Americans.

They were not allowed to vote or serve on juries in most states until 1870 and even then they could only vote if they were literate. In addition African Americans were discriminated against in employment opportunities and lived under Jim Crow laws. They were also segregated in schools, restaurants, and public places. In 1857 the Supreme Court ruled that African Americans were not citizens of the United States (Rodgers, 2006). This was a huge blow to African Americans who had been fighting for their rights for many years.

Other groups of people silenced by American exceptionalism are Native Americans and immigrants. The United States has a long history of treating Native Americans unfairly. In fact Native Americans were forced off their land and forced to live on reservations. In addition they were forced to send their children to boarding schools where they were punished if they spoke their native language or practiced their culture (Meltzer & McNamara, 2011). Immigrants come to America from all over the world in search of better opportunities for themselves and their families. However these opportunities do not always exist for them because they face discrimination in many different areas such as employment and housing (Rodgers, 2006).

However there are also possibilities for creating productive communities and identities that defy a tradition of subjugation. The first way to create a productive community is to educate people about the history of the United States and about the different groups of people who have been silenced by American exceptionalism. This will help people understand why certain groups are still being silenced today. It will also help them understand how certain groups have been treated differently than others in the past and why they continue to be treated differently today.

Another way to create a productive community is to make sure that everyone has equal rights and opportunities. This can be done by making sure that everyone has access to things such as education, housing, healthcare, employment, etc. This can also be done by making sure that everyone has equal rights when it comes to voting and serving on juries. We can also take action against

discrimination wherever it exists. Discrimination can take many forms such as racism, sexism, homophobia, etc. It can take place in any aspect of life including employment, housing, education, healthcare, voting rights and more. When discrimination takes place it silences those who are being discriminated against because they are denied equal rights with others.

Finally, there is one more way to create a productive community: this is by having open dialogue about American exceptionalism. Open dialogue allows people from all walks of life to come together and discuss their differences without judgment or prejudice. This helps people become more tolerant of each other's beliefs

End of Example Essay.

Conclusion

Artificial intelligence is developing faster than the education system can adapt to it. There are spaces that have been able to make use of this new cutting-edge technology, but society currently exists in a technological overhang where the capacity for change is not met by its actualization. Through original research and analysis of existing research, it can be seen that the potential for academic fraud with the use of language generation agents exists and can be used for passable grades at the highest levels of education. The use of these AIs is often totally undetectable by commonly used plagiarism detection software. While this may not yet be a concern for a society as the most advanced natural language generators aren't available for public use, it is unknown how long that may last. However, there are noticeable benefits that honest students and teachers can reap from these advancements. Intelligent Tutoring Systems show measurable advantages for memorization based learning and provide invaluable feedback to teachers. Automated TA systems can save students precious time.