BOSTON UNIVERSITY AI TASK FORCE

REPORT ON GENERATIVE AI IN EDUCATION AND RESEARCH

APRIL 5, 2024

BOSTON UNIVERSITY

CONTENTS

PROLOGUE

The *Task Force on AI Use in Education and Research (AI Task Force)* of *Boston University (BU)* was charged by the Provost with several pivotal responsibilities surrounding new and emerging Generative AI (GenAI) tools. The task force was asked to assess the impact of GenAI tools on education and research at BU, and review existing policies adopted by BU schools, colleges, and other academic institutions. The culmination of the task force's work is a report detailing its findings, proposing new policies, and identifying resources and procedures that may help to implement such policies.

The AI task force met throughout the Fall 2023 semester and in the early part of the Spring 2024 semester. It consulted with diverse experts, including BU faculty members, industry representatives, and external consultants specializing in AI and higher education.

The task force organized itself into subcommittees mapped to sections of this report. The scope of the report is limited to GenAI tools and does not address the many other applications of AI in practically every field of research and business practice. The purpose of the report is to inform and educate faculty and to advise the BU administration on issues related to the use of GenAI tools in education and research.

In addition to the report, the AI task force developed two further resources:

- *Detector Reliability:* A preliminary analysis of the reliability (accuracy and consistency) of GenAI detectors (see Appendix C) was conducted by the teaching assistants of the Spring 2024 edition of DS380 "Data, Society, and Ethics," offered by the Faculty of Computing and Data Sciences: Carmen Pelayo Fernandez and Jasmine Pham. This analysis is a useful practical guide for faculty as they consider how to detect GenAI material in student submissions.
- *Generative AI Policy and Advice Landscape App:* The Generative AI Policy and Advice Landscape project (see Appendix D) was conducted by the students of the Fall 2023 edition of DS380 "Data, Society, and Ethics," offered by the Faculty of Computing & Data Sciences, led by teaching assistants Sai Krishna, Sashank Madipally, and Showndarya Madhavan, and implemented in a web app by undergraduate students Heather Davies and Jason Huang. The app offers a searchable repository of GenAI policy and advice documents from institutions of higher education, along with content and network analyses. The database and analyses may be updated from time to time. The Spring 2024 update was performed by a SPARK! student group consisting of Haya Almajali, Jiasong Huang, Maxwell Malamut, and Samantha Rigor, with guidance from Seth Villegas and Michelle Voong.

BU AI task force members were assisted by Deidre Fisher, project manager for strategic initiatives in the Office of the Provost.

Task force members were:

Yannis Paschalidis, Distinguished Professor of Engineering, College of Engineering; Founding Professor of Computing & Data Sciences; Director, Rafik B. Hariri Institute for Computing and Computational Science & Engineering (Co-Chair)

Wesley J. Wildman, Professor of Philosophy, Theology and Ethics, School of Theology; Founding Professor of Computing & Data Sciences and Chair of Faculty Affairs, Faculty of Computing & Data Sciences (Co-Chair)

Mary Churchill, Associate Dean for Strategic Initiatives & Community Engagement; Professor of the Practice and Program Director, Higher Education Administration, Wheelock College of Education & Human Development

Mark Crovella, Professor of Computer Science, College of Arts & Sciences; Founding Professor of Computing & Data Sciences and Chair of Faculty Affairs, Faculty of Computing & Data Sciences

Anne Danehy, Associate Dean of Academic Affairs; Associate Professor of the Practice, College of Communication

Pary Fassihi, Senior Lecturer, Arts & Sciences Writing Program, College of Arts & Sciences

Juliet Floyd, Borden Parker Bowne Professor of Philosophy, College of Arts & Sciences; Director, BU Center for the Humanities

Priya Garg, Associate Dean, Office of Medical Education, Chobanian & Avedisian School of Medicine

Amie Grills, Associate Provost for Undergraduate Affairs; Professor of Counseling Psychology & Applied Human Development, Wheelock College of Education & Human Development

Tal Gross, Associate Professor of Markets, Public Policy & Law, Questrom School of Business

Wendy Heiger-Bernays, Clinical Professor of Environmental Health, School of Public Health

Keith Hylton, William Fairfield Warren Distinguished Professor; Professor of Law, School of Law

Daniel Kleinman, Associate Provost for Graduate Affairs; Professor of Sociology, College of Arts & Sciences

Gail McCausland, Assistant Dean of Academic Affairs; Clinical Professor of Periodontology, Henry M. Goldman School of Dental Medicine

Matt Parfitt, Associate Professor of Rhetoric, College of General Studies

Wilson Wong, Associate Professor of Biomedical Engineering, College of Engineering

EXECUTIVE SUMMARY

Boston University's (BU) AI task force was established in Fall 2023 to analyze the impact of Generative AI (GenAI) technologies on education and research. The task force consulted widely with experts and stakeholders. This report surveys the landscape of GenAI technologies, assesses their potential to impact research and education, identifies limitations and concerns, offers guidelines on the best use of GenAI, and proposes policies and administrative support structures.

The task force's key policy recommendations are as follows.

- ***Critical Embrace:*** **BU should not universally prohibit or restrict the use of GenAI tools.** Rather, BU should critically embrace the use of GenAI, support AI literacy among faculty and students, supply resources needed to maximize GenAI benefits in research and education, and exercise leadership in helping faculty and students craft adaptive responses.
- ***Pedagogical Clarity:*** **GenAI policies adopted by any college, school, or departmental unit should be consistent with the University's policies and reflect the critical embrace of GenAI technology.** Consistent with academic freedom, individual instructors should be free to define GenAI policies suited to the learning goals of their courses, and every syllabus needs to state the instructor's GenAI policy. Consistent with citation practices, instructors and students should acknowledge use of GenAI tools.
- ***Academic Misconduct:*** **BU should advise instructors to exercise caution when using GenAI detectors as evidence of GenAI use.** GenAI detector output should be regarded as only one part of a wider evidence base in evaluating possible academic misconduct. If used, GenAI detectors should be applied equally and fairly, and faculty should be aware of selection bias when applying GenAI detectors to specific suspected cases. Advance notice should be given in syllabi, including naming the specific detectors employed, so that students have an opportunity to use them also. Instructors need to be informed and supported in using reliable and consistent detectors.
- ***Security and Privacy:*** **BU should adopt policies to prevent the inadvertent publicizing of sensitive or valuable information through uploading it to GenAI tools.**
- ***Centralized Decisions:*** **BU should centralize decisions on GenAI tool acquisition and licensing**, on resourcing and personnel for supporting GenAI literacy and pedagogical reflection, and on administrative structures to ensure ongoing adaptive response to rapidly developing GenAI technology.

Further recommendations (Section 1) are organized in the following broad categories: (a) training and support for faculty and students in leveraging GenAI technology, (b) mitigating privacy and security concerns associated with the use of GenAI, and (c) administrative structures for streamlining BU processes and organizing resources associated with GenAI. The report also offers basic background information (Section 2), as well as analysis of GenAI use in research (Section 3), GenAI use in creating artifacts for writing, arts, and design (Section 4), and teaching, learning, and student professional development (Section 5). Alongside policy recommendations, the report suggests best-practice guidelines for the appropriate use of GenAI tools, and for dealing with inappropriate use. In addition to the report, the task force developed a set of resources to explore the landscape of academic GenAI policies, available at the BU AI task force website.

1. Recommended Policies

Generative Artificial Intelligence (GenAI) tools are evolving rapidly and so are policies and guidelines regulating their use. Academic institutions are starting to adopt policies governing GenAI use and common policy themes have emerged (see Appendix D and BU AI task force website for a searchable repository and analysis of GenAI policies released by academic institutions). The AI task force's recommendations for (enforceable) policies at the University level, along with suggested guidelines for wise and productive use of GenAI tools, generally align with this emerging consensus.

1.1. Principles for Guiding GenAI Policy Recommendations

The AI task force identified four important principles that should govern the use of GenAI: leadership, ownership, transparency, and excellence. These principles have far-reaching implications, which will be considered in the sections of this document: research (Section 3), creating artifacts (Section 4), teaching, learning, and student professional development (Section 5).

Leadership:

- BU should recognize that GenAI tools are revolutionizing the economic landscape. The University should lead the way in helping students and faculty adapt.
- Faculty and students should be supported to become literate about GenAI, competent in using GenAI tools, critically aware of the shifting balance of advantages and disadvantages of GenAI tools, and flexible in thinking about pedagogical and research aims and methods.
- BU should support ethical awareness in faculty, student, and institutional appraisal of GenAI tools, incorporating ethical dimensions in every educational and strategic resource.

Ownership:

- Faculty and students should assume full ownership and ultimate responsibility for their work, including when that work leverages GenAI.
- Faculty should emphasize the importance of human oversight over any GenAI created content and ensure students are not heavily or solely relying on GenAI for their work. GenAI tools may produce false information; the responsibility rests with the user to authenticate content using other verified sources.
- Faculty decisions about GenAI usage by students should be evidence-based, and caution should be used with GenAI detectors in any decision process.
- Authorship of a document implies that the author is responsible for the work. Since a software tool cannot assume such responsibility, and consistent with policies put in place by academic publishers [1], [2], [3], GenAI tools should not be listed as co-authors of documents and creative works.

Transparency:

- If AI tools are used in research, class assignments, writing, design, or other forms of creative work, it is important to be transparent on how exactly the tool has been used.

- Scholars should be mindful that documentation and transparency requirements may not be the same in all fields and journals, and that current requirements may change with the ever-evolving landscape.
- Faculty should be fully aware of and follow guidelines set forth by federal agencies and regulators (e.g., National Science Foundation guidelines [4]).
- Researchers should be aware of the individual AI practices and policies of those they collaborate with, as well as vendors and subcontractors who are involved in the research process.
- Students should be transparent about their use of GenAI for course assignments and projects and be instructed to document how GenAI tools were used.
- Instructors need to be transparent about course policies and address GenAI in every syllabus, providing clear guidance to students on how GenAI may or may not be used. They also need to provide advance notice of detection strategies and enforcement of GenAI rules. For example, the *CAS Writing Program*, as part of their faculty AI literacy efforts, offers specific language on their syllabus building tool for faculty to incorporate in course syllabi. The BU *Center for Teaching & Learning (CTL)* also offers guidance on referencing GenAI on course syllabi on their Communicate with Students about Generative AI resource.
- Transparency requires that authors need to cite the specific GenAI tool used in any document or product. Students should be instructed and reminded to credit both human and GenAI contributors, and that every source used must be cited.
- Faculty should familiarize themselves with the different approaches for citing GenAI tools and guide students on how to credit GenAI. BU's CTL offers resources for "Citing with AI" on their Quick References for Teaching with AI resource page.

Excellence:

- BU's commitment to excellence in research and teaching should be reasserted in connection with GenAI tools. Every application of such tools, regardless of context, should deepen the quality of teaching, learning, and research.
- Faculty and students should be trained in creative engineering of prompts to extract the most profound outputs possible.
- GenAI outputs should be critically evaluated to improve the quality of research and teaching.
- GenAI use should be evaluated for conformity with University expectations about ethical and responsible computing.

In addition to the general policy recommendations mentioned in the Executive Summary, which are guided by these principles, we group specific additional recommendations in the following categories: (1.2) training and support, (1.3) privacy and security, (1.4) guidelines for appropriate use and for dealing with inappropriate use of GenAI tools, and (1.5) administrative recommendations. When appropriate, we discuss the need for resources to implement the proposed recommendations.

1.2. Training and Support

GenAI tools will impact every field of study and most professions. It is vital that we equip students and faculty to operate effectively in the emerging GenAI era. It is likely that GenAI tools will accentuate disparities between those with the knowledge and resources to benefit from their use and the rest of the global population. BU, by properly educating its students, can help graduates maintain competitive advantage in their future careers, while remaining acutely aware of equity challenges. Similarly, faculty can benefit by leveraging GenAI tools and BU should help educate them. Recommendations follow.

***Incentivize the creation of new undergraduate and graduate courses or course modules*.** Everyone across the many disciplines represented at the University should have a basic understanding of GenAI technology so that they can better appreciate its benefits and limitations.

- University units should add courses that meet Hub requirements by focusing on GenAI tools and related applications.

***Support the development of GenAI literacy and education in best practices among faculty*.** GenAI, and AI more broadly, is likely to impact research in many domains, yet some faculty may lack the background or specialized training to benefit from these developments. The more BU does to educate its faculty and support their work, the more it will benefit from their increased productivity and impact in their fields. This leads to several specific recommendations.

- For new research projects and collaborations, University-wide centers, such as the Hariri Institute, should play an important role, enabling and supporting new research collaborations through existing mechanisms (e.g., the Hariri Focused Research Programs).
- Similar to facilitators for research computing and free allotments of computing resources at the Massachusetts Green High Performance Computing Center (MGHPCC) for all faculty, we recommend that BU considers funding positions of *AI research and education consultants and facilitators*. Such consultants could be made available to faculty to facilitate the use of AI and GenAI in their research. If the consultation service is centralized, the Hariri *Software and Application Innovation Lab (SAIL)* would be an appropriate place to house AI facilitators.
- We recommend soliciting feedback from faculty about how GenAI is and is not assisting faculty efforts to provide excellent education, thereby helping all to learn about the limits of GenAI.
- AI facilitators should be tasked to offer workshops to faculty and students on leveraging GenAI as research tools, introducing concepts such as prompt engineering, presenting case studies from faculty who leveraged GenAI in their research, and providing information about computing resources and software and how to access them.

Instructors should be provided with support to help them incorporate GenAI into their courses. This leads to several specific recommendations.

- CTL and the Office of Digital Learning & Innovation (DL&I) should be BU's main vehicles for providing teaching and course support. Their GenAI offerings should be vetted by the body of GenAI ambassadors described below.

- CTL should continue to offer workshops on how to incorporate GenAI into classroom instruction for those who are interested.
- CTL should continue to provide guidelines and resources in its AI in the Classroom and Teaching and Learning in an AI World web sites.
- DL&I and CTL should produce self-paced online modules to educate faculty.
- DL&I and CTL should establish *GenAI faculty ambassadors* across campus to act as school/college experts and to provide local advice. These can be drawn from the ranks of the faculty in each corresponding unit. This body of local ambassadors should exchange notes and identify best practices.
- GenAI faculty ambassadors should act as a faculty body for CTL and DL&I in evaluating and approving their online offerings, similar to how a regular course is evaluated and approved by a department's and a college's faculty. CTL and DL&I should establish mechanisms to collect user feedback on their offerings and use it for continuous improvement.
- BU should consider establishing a seed-funding program to provide modest and time-limited grants to faculty who have innovative ideas on using or analyzing GenAI in their teaching, particularly if they are willing to make their innovations more broadly available to other faculty who want to test such ideas. One potential mechanism is to leverage the AI and Education Initiative at the Hariri Institute which has introduced such a seed-funding program in collaboration with DL&I.

1.3. Privacy and Security

There exist serious privacy and security concerns associated with using GenAI tools running on servers that reside in corporate data centers. For this reason, many scientific journals have introduced policies against the uploading of papers under review, and funding agencies, such as the National Science Foundation (NSF), have banned the uploading of proposals under review.

These concerns apply to GenAI tools ChatGPT, Claude, Gemini (formerly Bard), and Llama, among others, when one uses the company's website and resources to run the model. These companies can use the provided data from users to train the next generation of the models. Thus, it becomes possible for information to be leaked through future model generations.

In contrast, some models (e.g., Llama, Mistral) can be downloaded and run locally. In this case, all interactions with the model and content provided to the model are stored locally and the user maintains ownership and control. We recommend:

- BU should adopt policies against uploading documents such as personnel files, budget spreadsheets with identifying information, exams, homework papers and solutions, medical records, tenure letters, and any document that contains personal information that should not be made public through server-based GenAI tools.
- University offices should implement policies to train all employees, including student employees, about this policy and their associated responsibilities.
- Independent Review Board (IRB) outcome letters and Data Use Agreements (DUAs) should be modified to include a provision affirming non-disclosure to GenAI servers.

- This policy can be reevaluated only when corporate assurances about not using prompts for training purposes are authoritatively audited.

1.4. Guidelines for Dealing with Inappropriate Use of GenAI

In the past, plagiarism, unauthorized collaboration for assignments, and other violations of academic integrity have been relatively easy to detect (e.g., by comparing an assignment with the source). Detecting the illegitimate use of GenAI tools is challenging by comparison. GenAI detectors are not fully reliable and may provide different answers for the same query at different times (see Appendix C). This is likely to remain an issue, particularly for cases where there is human-GenAI collaboration. This leads to several recommendations.

- ***Promote faculty literacy about GenAI detectors*.** Faculty and disciplinary committees should be informed of the unreliability and inconsistency of GenAI detectors and be advised that a likelihood score produced by GenAI detectors should not be considered sufficient evidence to accuse a student of unauthorized GenAI use. On the other hand, taken together with other factors, it may be considered as a relevant piece of evidence to strengthen a case.
- ***Promote faculty consistency in the use of GenAI detectors*.** In view of the unreliability of GenAI detectors, faculty should adopt a uniform policy on the application of detection tools. If a faculty member chooses to systematically employ a detection tool on submitted assignments, the specific tool should be applied to the assignments of all students. Faculty should be aware of selection bias that could be introduced by applying GenAI detectors to a specific targeted subset of students.
- ***Promote faculty literacy about GenAI methods to hide plagiarism*.** GenAI tools may make it harder to detect plagiarism, as users can use GenAI to "paraphrase" from an original source. CTL offers guidance on AI and Academic Integrity and this guidance should be kept current with developments in the field. We recommend that IS&T evaluates various options for tools to enhance plagiarism detection, including licensing tools that are successfully used by academic publishers (e.g., iThenticate [5], Grammarly [6]). The task force heard from both faculty and students that the Blackboard solution (Turnitin) is ineffective at present.

1.5. Administrative Recommendations

Finally, the AI task force offers the following administrative recommendations.

- ***GenAI should be leveraged to improve, automate, and streamline a variety of BU processes*.** Whereas for business functions (e.g., Q&A for HR benefit options) commercial solutions are likely to emerge quickly, student advising may be a function where GenAI could help students conveniently discover information regarding courses, instructors, graduation requirements, and processes. We recommend that the University explores the training of a BU native chatbot for these purposes, either by leveraging BU faculty expertise or partnering with an external provider. Such a chatbot would help students quickly find answers to many administrative questions, reduce the burden on staff and faculty in large academic programs, and establish a productive collaboration between advising staff and

chatbot assistants. This is meant not to reduce faculty-student advising interactions, but rather to reduce the time devoted to procedural issues and increase the time for more meaningful exchanges.

- ***BU should centralize decisions on recommending and licensing GenAI tools for faculty and students*.** Some existing tools (e.g., OpenAI's ChatGPT) offer a subscription service for the most advanced version of the tool. A paid service can lead to a "digital divide" that intensifies inequities between students and groups with the resources to pay for subscriptions and those who lack such resources. GenAI tools are likely to become ubiquitous and necessary for work, so BU should implement a policy similar to what it is now using for commonly used software (e.g., office productivity, operating systems, Matlab, statistical, and other scientific and design software).
- ***BU should create a centralized resource of GenAI information*.** BU should consider creating a centralized website to communicate the University's policies and procedures regarding GenAI, and to provide links to various resources offered by BU offices, such as IS&T, CTL, DL&I, and any new resources that may be created as a result of these recommendations (e.g., AI research facilitators, AI ambassadors, the web-based app containing information on policies adopted elsewhere, etc.). This website needs a manager to coordinate the various offices and resources and to ensure that content gets updated to reflect rapid developments in the GenAI landscape. An initial web site was set up by the BU AI task force.

2. Background

2.1. GenAI Introduction

GenAI generally refers to foundational AI models that have been trained on vast amounts of data and can generate new content. Appendix A offers technical background on the evolution of the *Natural Language Processing (NLP)* field that gave rise to the recent *Large Language Models (LLMs),* such as OpenAI's popular ChatGPT [7]. The latest GPT version (as of the writing of this report), GPT-4, has been trained on vast amounts of text (from the internet, books, news articles, scientific literature, etc.), yielding models with over 1.7 trillion parameters. Beyond text generation, models (such as OpenAI's DALL-E [8]) can generate images based on a text description.

The GenAI landscape is rapidly changing. There are hundreds of GenAI tools and more appear every week. Popular GenAI tools designed to generate text include: Microsoft's Copilot (formerly BingChat) [9] (based on OpenAI's GPT-4), Google's Gemini (formerly Bard) [10], [11], Meta's Llama [12], Mistral [13], and Anthropic's Claude [14]. Text-based tools can be prompted to write many forms of text (including creative writing), edit text, write computer code in many programming languages, offer counseling services, function as tutors, automate call centers, and create interactive robotic companions. There are countless other GenAI tools for text creation, many of them with no content moderation, which the better-known offerings do have. Hundreds of text generators for specialized applications are now being offered by OpenAI, based on GPT-4, and many GenAI tools are being created based on the foundational GenAI models mentioned earlier.

GenAI tools can also produce other forms of output, including images and visual art (e.g., at the time of this report, OpenAI's DALL-E [8], Midjourney [15], Stable Diffusion [16], Adobe Firefly [17]), music (e.g., Soundraw [18], Bandlab [19]), video (e.g., Pictory [20], Synthesys [21], Sora [22]), and presentations (e.g., Simplified [23], Slides AI [24], Gamma [25]).

The latest versions of GenAI tools (including OpenAI's GPT-4 and Google's Gemini) can handle multi-modal input and output, including text, images, various file types (e.g., spreadsheets, presentations, data files), and audio/video in the case of Gemini. We expect these capabilities to become more sophisticated, more integrated, and more broadly applicable very rapidly.

2.2. The Nature of GenAI Models

Understanding the nature of the models underlying GenAI tools is important for informing their wise and effective use. The following terms and associated characteristics are relevant to every part of this report.

- *GenAI Output:* The output of GenAI models is *probabilistic* and not deterministic. Specifically, the model learns to associate a probability to different outputs and essentially "tosses a coin" to generate the most likely output given the input. The same prompt, repeated several times, can produce different outputs.
- *GenAI Inaccuracy:* GenAI models are known to *hallucinate*, state falsehoods, generate plausible but non-existing references, etc. Thus, one should use the output with extreme caution and always fact-check.

- *GenAI Modalities:* The latest GenAI models can handle *multi-modal* input and produce multi-modal output, including text, data formatted in various formats, images, audio, video, presentation slides, computer code, etc.
- *GenAI Web Searching:* Chatbots powered by GenAI models can search the web for answers, helpfully integrating and summarizing information. GenAI models also generate answers to questions that fuse the model's output with information they obtain from web searches.
- *GenAI Opacity:* No one understands exactly what these models do, not even their developers. It is currently impossible to explain precisely why a complex model having 1.7 trillion parameters produces a specific answer.
- *GenAI Attribution:* LLM models "recall" the information on which they have been trained to generate responses to queries, and they can do so without always attributing the source. This can lead to plagiarism, even though the models are not trained to plagiarize.
- *GenAI Prompts:* The way one asks a GenAI chatbot a question is important, and different prompts may yield different answers. Several techniques have been shown to improve the quality of GenAI tools' output. *Prompt engineering*, i.e., creating prompts to elicit high quality answers, is an active field of research. Examples of prompt engineering include: (a) Few-shot learning, in which the model is given example questions and desirable answers as part of the prompt; (b) Chain-of-Thought (CoT) prompting [26], which asks the GenAI model to explain its reasoning before providing an answer; and (c) ensembling, i.e., asking the same question multiple times and taking a majority vote to determine the answer, which can compensate for the stochastic operation of GenAI chatbots.
- *GenAI Training:* For the most part, the big tech companies developing GenAI models have not made the datasets used for training public, which makes it difficult to determine what aspects of model behavior are due to specific properties of the training data (e.g., bias). For example, models have been found to know details about characters in copyrighted books, so authors have sued some of the companies who developed the models for copyright infringement. On a larger scale, news organizations such as *The New York Times* (NYT) [27] have mounted legal action against OpenAI and Microsoft (which has invested heavily in OpenAI) for unauthorized use of NYT content in training OpenAI's GPT models.
- *GenAI Costs:* Training GenAI models (e.g., involving as much as 1.7 trillion parameters) requires vast datasets (reportedly, more than 10 trillion words in 45 GB of data) and a large number of GPU servers. OpenAI estimated that the cost of training GPT-4 was on the order of $100 million [28], which suggests that only large corporations with ample financial and computing resources can train such large foundational models, and such training involves significant ecological impacts. University-based researchers, and even collaborations between universities cannot easily compete with commercial enterprises in constructing such large artifacts.
- *GenAI Privacy:* Popular models such as OpenAI's ChatGPT and Google's Gemini (formerly Bard) are *not* open source. The code for the model and the model itself (embodied in the neural network weights) reside on servers controlled by these companies. This implies that the company running a model can still see and reuse interactions with the model. For example, the text of a query submitted to ChatGPT is saved by OpenAI and could be used to train future generations of the model. Hence, there is a real danger that future model

responses may leak private or confidential or otherwise valuable information through responses to queries. While settings can instruct GenAI models not to use information submitted for training future models, and new services promising confidentiality are being rolled out [29], these corporate representations have not been independently audited or verified. Similarly, there is no independent verification of the processes used to train the models, their safety, their ability to preserve privacy, etc. The federal government has issued an executive order seeking to highlight such concerns and instructing federal agencies to take steps to regulate this nascent industry [30]. Currently, among the recent very large LLMs, Meta's Llama [12] and Mistral's Mistral [13] are partially open source, in the sense that one can download, install, and run a local copy of the model, which guarantees that no information submitted through a query could be passed to the corporation that developed and trained the model. Still, Meta has not released the dataset used to train the model, retaining some opacity and making the model not fully open source and reproducible [31].

- *GenAI Moderation:* Large tech companies have invested in creating a content moderation layer for their GenAI tools to inhibit the generation of offensive content. These content *guardrails* can often be surmounted with clever jailbreak techniques [32], [33], [34]. Many chatbots have little or no content moderation.
- *GenAI Disinformation:* GenAI tools, particularly those without content moderation, can be used by bad actors to generate and disseminate disinformation, and in a convincing way that inspires others to spread it further, not realizing that it is false. This can potentially destabilize political and public institutions and provoke security problems.
- *GenAI Psychology:* We are still learning about the psychological impact of GenAI tools, but at this point they appear to have the potential for significant effects, from failure to develop critical skills in thinking, reading, writing, and coding, to the atrophy of those same skills. Used wisely, they can help refocus attention on the more creative aspects of problems, including human-to-human skills in reasoning critically and empathically, and extend cognitive reach to an unprecedented degree, which has the potential to make them vital assets for learning, teaching, and research.
- *GenAI Detectors:* Detectors can estimate the likelihood of a passage being created by a GenAI tool (e.g., GPTZero [35] provides assessments for GPT-4, Llama 2, and human + AI working together to write a passage), but these assessments *simply report a probability*. Using them can generate both false positives and false negatives, which complicates evaluation of academic misconduct charges. They may also be inconsistent, in the sense that the same query may yield different results, which makes it difficult for students and faculty to use detectors to protect themselves against the nightmare of academic misconduct charges based on false positives. In addition, these detectors have been known to be especially biased towards non-native speakers of English due to these speakers' possibly restricted linguistic diversity and word choice [36]. Currently, there is no definitive way to assess with high confidence that a passage was written by or with the help of GenAI, or for students and graders to receive the same originality estimate. The accompanying study of the accuracy and consistency of GenAI chatbots (see Appendix C) illustrates the problem.

2.3. Ethics Concerns Surrounding GenAI Tools

GenAI tools have sparked a range of ethical concerns, including:

- using work of others for training, and the associated possible copyright infringement;
- low-paid workers for model training based on human feedback [37], [38] and traumatic exposure of low-paid content moderators to horrific material;
- output mimicking styles of human creators;
- deep-fake audio and video used to spread disinformation, from political campaigns to cyberbullying;
- violation of privacy and security expectations;
- economic disruption through job displacement;
- climate impacts due to the massive server farms that train and run widely used GenAI tools [39];
- inequities of access.

Such ethical quandaries are arising within a legal and policy vacuum, forcing lawmakers to adapt quickly, potentially leading to suboptimal solutions with harmful side effects until more stable policies, regulations, and laws are enacted.

A fundamental unresolved issue is the lack of sufficiently thoroughgoing ethical review within and of technology corporations, which often aim to "move fast and break things" rather than be clear with regulators and the general public about the predictable socioeconomic implications of new technologies.

With that in mind, BU should consider strengthening guidelines for responsible conduct of research and implementing a mechanism for ethics review of research projects involving AI, much as BU currently uses an institutional review board to oversee human-subjects research.

3. GenAI in Research

The promise of GenAI tools in research is considerable, as are the challenges. Because of the rapid pace of this field, it is likely that some points below will be outdated soon after this report is released. Thus, our preliminary recommendations will need to be revised in the months and years ahead.

Some researchers at BU focus on developing and applying GenAI tools, while others may use GenAI tools for research on any number of topics. In both cases, it is important to enable BU researchers to use GenAI to advance their research and to address major societal challenges, while also mitigating any problems that may result. We must harness these technologies but do so cautiously.

In addition to training established scholars in using GenAI tools to advance their research, we should also integrate GenAI training into the education of graduate and undergraduate researchers.

3.1. GenAI Usage Survey

To understand where our community stands in the evolving place of GenAI tools in research, we surveyed 15 associate deans for research across BU's schools and colleges (see Appendix B). We received responses from nine (60% response rate). We asked the associate deans seven open-ended questions on topics ranging from the value and challenges of GenAI tools in research, to whether BU should have guidelines concerning GenAI tools in research. The associate deans echoed perspectives found in news coverage and in the limited extant scholarly work, including a recent workshop by the National Academies of Sciences, Engineering, and Medicine [40] on the topic.

GenAI tools in research, according to the associate deans, can help scholars develop literature reviews, formulate hypotheses, aid data collection and analysis, make coding easier and faster, and provide editing and proofreading assistance. On the downside, the associate deans suggested that AI can lead to greater reliance on pattern recognition without understanding, and this, in turn, can lead to biased results. In addition, GenAI tools may facilitate the proliferation of misinformation, make plagiarism harder to detect, and raise challenges for intellectual property protection. Finally, several of our respondents suggested that BU researchers could benefit from training workshops of various sorts, and the majority favored the review and development of updated guidelines for responsible conduct of research at BU.

3.2. Benefits and Concerns Relating to GenAI in Research

Considering the survey results, perspectives from the AI task force and its consultants, and other policy and advice documents for higher education (see Appendix D), we note the following benefits of GenAI tools for research.

- *Natural Language Processing:* GenAI tools can facilitate and simplify natural-language processing (NLP), especially of text, including extracting information from text; comprehending, analyzing, and summarizing passages; conducting thematic keyword identification in automated qualitative research pipelines; performing sentiment analysis; translating between multiple languages; and encoding text for downstream classification

tasks. For example, the analysis of GenAI policy documents on the task force's website is a fully automated application of GenAI to hundreds of complex texts (see Appendix D).

- *Human-Computer Interaction:* GenAI tools can automate human-machine interactions, including in Q&A tasks. For example, specialized chatbots can automate data collection from human subjects for clinical trials.
- *Literature Review:* GenAI tools can automate literature review by analyzing a large range of scholarly papers and through intelligent human prompting, identify the more relevant subset of publications for subsequent analysis. (See Appendix E for a description of a case study.)
- *Tech Support:* GenAI tools can automate technical-support activities by answering technical questions and providing instructions on how to perform a plethora of computer-related and other tasks.
- *Coding:* GenAI tools can generate computer code in almost any programming language, greatly accelerating and simplifying tedious and time-consuming computer-programming tasks.
- *Disciplinary applications:* GenAI tools have started to be used in various scientific fields. Examples include: aiding healthcare and disease diagnosis [41], [42], [43], robotics [44], [45], material science and chemistry [46], social sciences [47], and even automated hypothesis generation for scientific discovery, leading to automated laboratories. As GenAI models improve, they are likely to be able to reason better, to the point that they can analyze legal arguments and construct complex mathematical proofs [48]. GenAI tools can compose text for scholarly papers, presentations, proposals, and other documents; take interpretative summary minutes in meetings, when coupled with automatic speech recognition; edit and improve text; generate illustrative graphics; and generate drafts of presentations [49]. An expansive survey [50] outlines AI research directions for science, energy, and security.

GenAI tools also raise concerns in research contexts:

- *Overuse:* GenAI tools may lead to atrophy of human generative skills due to overuse or inattentive use in the writing of proposals, reports, papers, and presentations.
- *Bias:* GenAI tools can accentuate biases present in training data, expressing such biases in text output.
- *Privacy:* GenAI tools introduce privacy concerns when uploading confidential information to GenAI engines (e.g., proposals or papers under review, identifiable information about research subjects, personnel files, or private personal information).
- *Responsible conduct of research:* GenAI tools make it easier to violate research ethics and harder to detect such violations (e.g., plagiarism, generate fictitious data, manipulate images, etc.).

3.3. Existing Guidelines

Several universities, including R1 institutions, have already published guidelines for the use of GenAI tools in research (see Appendix D). At the same time, publishers have issued guidelines on the use of GenAI tools by reviewers and authors, generally prohibiting uploading papers under

review to GenAI tools to preserve confidentiality, prohibiting listing GenAI tools as co-authors, and requiring authors to disclose the use of GenAI in the manuscript and cite the corresponding tools [1], [2], [3]. NSF has also adopted similar guidelines for reviewers and proposers [4]. The National Institutes of Health has adopted guidance on the use of GenAI in grant reviews [51].

4. GenAI in Creating Artifacts: Writing, Arts, and Design

GenAI has become commonplace in creating artifacts, including but not limited to areas that incorporate an iterative invention and revision process such as writing, designing, and creating art. BU should support GenAI not as a replacement for the process of creation, but rather as a collaborative tool. GenAI tools can be effective in these areas if integrated with human oversight, and if students are taught how to use them appropriately and responsibly.

Writing a paper or creating any form of art is often viewed in terms of the final artifact. However, a process-driven approach emphasizes the formation, development, and refinement of ideas, eventually yielding a final product. In this framework, GenAI tools can be useful assistants and collaborators for generating initial ideas, clarifying thoughts, acting as peer reviewers, and iterating through many versions.

4.1. Best Practices for Incorporating GenAI in Creation of Artifacts

By thoughtfully designing assignments that incorporate and emphasize process rather than only the end-result, faculty can ensure that students not only are able to make sound decisions regarding the use of GenAI tools, but also become proficient in using these tools when needed by remaining actively engaged in the process of *creating artifacts*, honing their critical thinking and originality. This approach prepares students for a future where GenAI tools are an integral part of the creative landscape, without diminishing the importance of human insight and creativity.

- *Teaching Students about GenAI-Assisted Artifacts:* Incorporating or *not incorporating* GenAI tools into the iterative creative processes requires a nuanced understanding from both faculty and students. Faculty can guide students in understanding how GenAI tools can serve as a part of their creative toolkit in courses that incorporate writing or designing, emphasizing its *supplementary* role rather than a substitute for human creativity, while also demonstrating GenAI's limitations. This will foster an AI literacy that appreciates GenAI tools only as a potential *partner* or *assistant* in writing and other creative activities. This level of literacy, for example, may involve demonstrations on how GenAI tools can suggest diverse design layouts or writing styles, or be used as a starting point for creative exploration, thereby *enhancing* the creative process. Meanwhile, educating students about the limitations of GenAI tools in these creative processes is also crucial. This can be achieved through case studies and examples where GenAI tools have failed to capture the nuances of human communication and creativity, or have produced outputs with errors or biases, lacking cultural sensitivity, or where AI-written pieces missed the emotional depth or awareness of rhetorical context that human writers more naturally include.
- *Teaching Courses with GenAI:* The application of GenAI tools in the teaching of courses that incorporate creative processes extends beyond mere content generation. It offers an opportunity for personalized learning experiences, exposing students to a wide array of styles and techniques. For instance, in tasks that incorporate writing, GenAI tools can assist with language development, genre awareness, and the structuring of arguments. In tasks that involve designing, it can suggest various aesthetic and functional design options, thus broadening students' horizons. Faculty may consider integrating GenAI tools to facilitate

such heightened awareness, designing their assignments in ways that ensure the focus remains on the development of students' critical and creative skills.

- *GenAI Tools and the Design of Process-Oriented Assignments:* Designing assignments that require an iterative creative process that integrates GenAI demands particular care and consideration. This involves creating assignments that leverage GenAI as an educational tool while also fostering critical thinking and creativity in students. For example, faculty may design writing assignments where GenAI tools are leveraged only for initial idea generation and brainstorming. Such assignments would then require students to select one concept and develop and refine it further. Other assignments may ask students to reflect on their process of working with GenAI tools. This could involve reflective writing about the decisions they made after receiving GenAI outputs, how they incorporated GenAI suggestions into their work, and their thoughts on the creative balance between GenAI and human input.

Conversely, certain assignments might intentionally limit the use of GenAI tools to highlight areas where human insight is paramount, such as understanding cultural nuances or ethical considerations in designing a poster. In these instances, faculty may introduce assignments that require skills GenAI tools cannot replicate, such as hand-drawn sketches of local subjects or analytical writing based on personal experiences, encouraging students to value and develop their uniquely human creative abilities.

4.2. Concerns about Adopting GenAI in Creative Processes

Adopting GenAI tools in areas that require knowledge and experience of a creative process, including almost any kind of writing or design, raises several concerns:

- *Concerns in Society:* Instructors who integrate GenAI tools in courses that incorporate creativity must be supported to recognize broader societal concerns. Issues such as potential job displacement in creative fields and ethical implications need to be acknowledged and addressed in classrooms. Through training and support for AI literacy, faculty can be prepared to discuss these topics, incorporating them into the curriculum to foster a holistic understanding of the impact of GenAI tools. Faculty can also become prepared to discuss copyright and intellectual property policies and regulations with students when integrating the use of GenAI.
- *Plagiarism:* A primary concern about GenAI tools in academia, especially in courses that incorporate writing, is the blurred line between student-generated and GenAI content. Plagiarism takes a new form with GenAI, where the originality of ideas and expressions can be challenging to ascertain (see the related discussion in Appendix C). Faculty may need to revisit their assignment design, assessment strategies, and GenAI-use policies. This might involve more emphasis on the process of creation, requiring students to submit drafts and reflections on their work. The use of GenAI tools in any creative work produced should not violate BU's Academic Conduct Code. For example, using GenAI tools to paraphrase existing works without citing them is still plagiarism (see [3] for similar concerns expressed by scientific publishers).

- *Misinformation:* GenAI tools can sometimes generate plausible but factually incorrect information, posing a risk of misinformation.
- *Over-Reliance:* An over-reliance on GenAI tools might impede the development of critical thinking and originality among students and may lead to poor-quality and factually inaccurate student artifacts. Faculty need to emphasize the importance of verifying information and encourage students to approach GenAI outputs with a critical eye. Assignments may need to be redesigned to require independent thought and verification of facts and sources.
- *Ethical Issues:* There is a concern that GenAI-driven artifacts might not adequately represent or respect the diversity of student populations. Through training and support, faculty can be made aware that GenAI tools used in the classroom are not free from biases and may not be inclusive. This suggests approaching GenAI tools with a critical mind, and teaching students to recognize and critically assess any biases in GenAI outputs.

4.3. USE CASES

The use of GenAI tools as a collaborative tool for writing and design processes adds significant value in a multitude of applications across various academic and societal contexts. The following use cases highlight the versatility and potential of GenAI tools in educational and practical contexts.

- *Academic Writing:* GenAI is helpful in academic writing in a variety of ways, including increasing productivity when it comes to the initial gathering of literature based on predefined criteria, therefore saving considerable time and effort. GenAI has the ability to assist in synthesizing key findings and themes from a vast array of sources and suggesting potential areas of inquiry.
- *Expanding Creative Boundaries in Arts:* In the visual arts, GenAI tools have the potential to play a pivotal role in personal artistic development and expression. By inputting ideas and sketches into GenAI systems (while also being mindful of copyright laws and inputting any personal creative work), artists can receive variations and evolutions of their initial concepts rendered in multiple styles. This not only accelerates the ideation process but also provides artists with new perspectives on their own work.
- *Facilitating an Education that Incorporates Creativity:* In tasks that incorporate creativity, GenAI tools can aid in tailoring learning materials to individual student needs and providing assistance. This could include fostering a more process-oriented and personalized learning environment.
- *Using GenAI in Graphic Design:* GenAI transforms the field of graphic design by providing tools that can assist in generating creative visuals, suggesting design layout options, and offering color scheme recommendations. It can be used to create initial design concepts, which designers can then refine and personalize. This not only accelerates the design process but also opens up new avenues for creativity and experimentation. Graphic design students can leverage GenAI tools to explore and learn about diverse styles and techniques and experiment with a plethora of options, thereby expanding their creative repertoire.
- *Streamlining Integrated Communication Campaigns:* The utility of GenAI tools in creating coherent and impactful messages is invaluable for communication campaigns. Its use allows

for a unified approach where a single individual can effectively manage both the narrative and visual aspects, ensuring messaging consistency and efficiency in a campaign.

- *Drafting Public-Facing Documents and Campaign Materials:* For larger societal contributions, GenAI tools can assist in crafting public documents and materials that are designed to communicate complex ideas in an accessible manner. This is particularly beneficial for awareness campaigns, where clarity and communication around challenging topics are essential.

In conclusion, GenAI has the potential to be a powerful and effective tool for creating artifacts and enhancing creativity when used appropriately and responsibly. GenAI's role in transforming how courses (e.g., analytic and creative writing, graphic design, arts, etc.) that focus on a creative process are taught and practiced will be very significant. As a result, at the very least, the existence of GenAI tools should be acknowledged in course policies. It is important that the emphasis in courses that incorporate writing and design elements should be on the process rather than solely the end product. The necessity of maintaining ethical standards and human oversight will always remain, and GenAI should be framed as a collaboration tool and assistant.

5. GenAI in Teaching, Learning, and Student Professional Development

The principles set forth in Section 1.1 should motivate faculty to rethink teaching, assignments, and learning outcomes for the GenAI era.

- *Redesign Assignments*. Many students may use GenAI whether it is allowed or prohibited, so we recommend that faculty re-visit their assignments, assessments, and teaching strategies, which may need to be adapted to reflect the prevalence of GenAI tools.
- *Process Scaffolding*. For courses that incorporate writing, arts, and design, we recommend that work is scaffolded, thereby focusing more on the process than on the end product (see Section 4, above). One way to do this involves requiring students to submit iterative drafts and reflections on the evolution of their work in stages. Another way involves requiring students to orally present and defend their work in class, in small groups, or in meetings with instructors whenever class size makes such an approach feasible.
- *Avoiding Take-Home Exams*. Instructors, when feasible, can increase their reliance on in-class testing and in-person evaluations, while keeping in mind the possibility that such modes of evaluation can pose accessibility problems for some students.

More detailed considerations follow.

5.1. Course Policies

Several high-level policies are relevant here.

- *Institutional Diversity:* BU is a large, diverse institution, and no single, "one-size-fits-all" GenAI course policy can comprehend the great diversity of learning outcomes, teaching methods, disciplinary cultures, and assessment practices across BU's schools and colleges. Any single policy meant to apply to all courses across the University would have to be so vague that it would be of little practical use. Thus, we recommend against a monolithic, University-wide policy on GenAI use in the classroom.
- *University-Wide Policy:* However, we do recommend that a University-wide policy be adopted requiring every course syllabus at BU to include specific guidance and boundaries around how GenAI tools may and may not be used to produce submitted work.
- *Explicit Guidance for Students:* GenAI tools can be powerful learning adjuncts, so faculty are encouraged to explore how they could be leveraged rather than banned. However, GenAI tools can also be used as shortcuts that inhibit or distort learning. The precise risks and opportunities differ for each course, and students need clear, explicit guidance to understand how GenAI tools should and should not be used in each context. Ideally, faculty will clearly indicate if, when, and how GenAI tools can be used on class assignments. Sample statements can be found on the BU Center for Teaching and Learning (CTL) website.
- *Academic Integrity:* Syllabi should clearly state that use of GenAI tools outside the indicated scope is considered a violation of academic integrity.
- *Faculty Reasoning:* To support AI literacy, faculty might provide students with information about why they have chosen to allow or not allow its use on a particular assignment.

- *Faculty Transparency:* Faculty need to specify how teaching staff may use GenAI tools in learning activities and assignments for each course, so students understand the role of GenAI in faculty teaching.

5.2. Faculty and Student Responsibilities

Students and teachers incur a range of responsibilities in relation to the appropriate use of GenAI tools.

- *Universal AI Literacy*. In making a shift towards requiring that all syllabi include information on GenAI tool use, all members of our community need to know some basic facts about GenAI tools.
- *Faculty Adaptation*. Faculty need to recognize that many students are now entering the University as experienced users of GenAI tools, already in the habit of using GenAI as a learning tool, an assistant, and a collaborator. Just as the advent of the Internet in the 1990s and the smartphone in the 2000s forced faculty to reimagine their courses, change their teaching practices, and articulate clear policies for using the new technologies, so now faculty need to be supported to craft adaptive responses to the era of GenAI tools.
- *Faculty Need to Continually Adapt*. Faculty need to recognize that students today must be prepared for GenAI tool use in their future careers and become AI literate. Because AI will impact every field and discipline, students in every school and college need to understand the basics of how GenAI works, where its strengths and weaknesses lie, and how to engineer effective prompts. This is part of what is meant by the term AI literacy, but as the technology changes, as new tools become available, and as these tools become more deeply integrated into the workplace, the meaning of AI literacy will also change.
- *Faculty Need to Stay Up to Date*. The current state of GenAI will evolve rapidly. Faculty in all disciplines will need to stay informed about new developments, and BU offices such as CTL will need to offer ongoing assistance.
- *Faculty Can Make Use of GenAI as Student Coaches*. We need to inform faculty that GenAI can provide students with a deeper understanding of course content by giving them formative feedback or clearer examples. For example, BU students report that they are already using ChatGPT to provide alternative definitions or explanations for key concepts. Some students also report using ChatGPT to create practice questions to help them prepare for exams. Faculty need to know that students are using GenAI in this way, and they should consider how they can capitalize on this to further students' mastery of course content.
- *Faculty Need to Understand Limitations of GenAI Detection*. All faculty need to understand the limitations of GenAI-detection tools, which should be used with caution (see Appendix C for further details). With the risk of false positives in mind, AI-detection tools could be used as only one element in a larger body of evidence for assessing the originality of student work. In addition, faculty who choose to use detection tools should apply the tools consistently to all students in their classes, not to a subset chosen on the basis of gut-level suspicion or imperfect signals. Faculty should consider this in the information they provide to students about class assignments, and they should be transparent about their adopted GenAI detection and rule enforcement strategies.

- *Faculty, Not AI, Must Assign Grades*. GenAI tools should not determine grading, supply feedback, or stipulate learning outcomes in courses. Such decisions – whether it is the grade breakdown on a syllabus or a specific grade for a particular student's work – should be made by faculty, not outsourced to GenAI tools. GenAI is not yet reliable or subtle enough to manage such critical tasks without a "human in the loop."
- *Faculty Should Not Assume that Students are Experts*. Faculty should not assume that students understand GenAI tools. Students may have used GenAI tools, perhaps extensively. But faculty should not assume that all students understand how GenAI works or what it can and cannot do. In conversations with several student panels, the varying interest and expertise among our current students was clear.
- *Faculty Need to Teach Students about Limitations and Nuances*. To use GenAI tools ethically and productively, students need to understand that their outputs are probabilistic predictions (informed guesses); that training datasets reflect a number of known biases; and that even when outputs incorporate web searching, GenAI tools can still hallucinate while appearing confident and authoritative. At this point in their development, GenAI tools tend to generate writing that qualitatively tends toward the mean and is primarily declarative rather than critical or analogical. Both faculty and students need to take responsibility for their own work. All GenAI output needs to be revised, edited, and validated.
- *Preparing for Changing Careers*. Both faculty and students need to be aware that GenAI will continue to transform the way that professionals in most industries work. BU must prepare students for this new method of working. Soon, GenAI will be incorporated in nearly every software product. AI will affect every discipline and it is becoming an important skill sought by employers. The BU Center for Career Development offers a set of resources to students as they look to start their professional careers, and it would be helpful to consider adding workshops related to AI. The Hariri Institute is running a student-led Data Science Mentoring Circles Program, offering BU alumni now in industry as mentors to graduate students seeking careers in data science.
- *Faculty Must Not Ignore GenAI Tools*. Finally, faculty are advised to not avoid discussing GenAI. GenAI is now a presence in our lives and our students' lives, one that is changing the way we research, write, and learn. We believe that ignoring it and refusing to address its implications is no longer an option. Faculty need university support to become prepared for discussions with students about GenAI, its role in student learning, its possible uses (ethical and unethical), and its limitations.

5.3. Administration Responsibilities

University administrators also incur responsibilities due to GenAI Tools.

- *Professional Development:* BU needs to support ongoing professional development, including providing grants and resources to allow faculty to experiment and innovate with GenAI. Both the risks and the potential benefits of GenAI tools for research, teaching, and student learning are considerable, but largely unexplored and untested. Faculty may need to reimagine and revise many of their existing teaching methods, course designs, syllabi,

assignments, and classroom activities to capitalize on this technology and stay abreast of new developments. Investments may also be needed to provide workshops, trainings, online resources, and consultations for faculty engaging in this work. This may require a considerable investment of time for which faculty and staff may deserve compensation, depending on the nature of the investment. In addition, some funding should perhaps be made available for more structured studies, designed to develop and promulgate new approaches to teaching and learning.

- *Centralized Resources:* A continuously maintained set of GenAI resources for faculty and students is imperative. Faculty need access to sample syllabi, course policies, advisory statements, use examples, and instructions. CTL, DL&I, and other offices can provide those resources. Opportunities should be available for faculty to discuss GenAI resources with CTL and DL&I staff and other experts.
- *Managing Inequities:* Teaching with GenAI tools should not lead to inequities among students. While we encourage faculty to develop new ways of using GenAI as an effective learning tool, we are also aware that some tools have paid tiers that may be difficult for some students to afford. Financial inequity should not imply unequal opportunity to learn or complete work efficiently. This concern may possibly be alleviated as prices change, or completely resolved through University-negotiated site licenses.

The integration of GenAI tools in education presents both challenges and opportunities. By understanding the capabilities and limitations of GenAI tools, incorporating thoughtful policies into syllabi, and fostering an environment of continuous learning and adaptation, educational institutions can effectively harness the potential of GenAI tools. This approach will not only enhance the learning experience but also prepare students and faculty for a future where GenAI tools are an integral part of the professional and academic landscape.

APPENDICES

APPENDIX A: TRANSFORMERS AND THE GPT MODELS

GenAI models are based on deep learning approaches using neural networks, which are formed by arranging stacked layers of computational "neurons." These neural networks perform basic linear and nonlinear transformations on a vector input to produce some other scalar or vector output. Among the first generative models were so called *Generative Adversarial Networks (GANs)* [52] that were designed to generate images to be presented to an image classifier in an effort to improve the classifier. Later applications focused on *Natural Language Processing* (NLP), giving rise to *Large Language Models* (LLMs).

A breakthrough in the field is considered to be the introduction of the transformer architecture [53], which provided a way for identifying which words in a sentence the model should pay attention to provide the proper context while mapping a word into a numerical vector (called embedding). This architecture was first used for Natural Language Processing (NLP) purposes in the *Bidirectional Encoder Representations from Transformers* (BERT) model [54], which is able to encode each word (or smaller pieces called tokens) in its input text, taking into account an entire passage of text.

The so called *Generative Pre-Trained Transformer* (GPT) architecture [55] also uses transformers but only uses past words in a sentence to predict the next word, and thus functions as a sentence completion algorithm that can be made recursive. Subsequent GPT models became larger (GPT-2 [56], GPT-3 [57]) and were shown to exhibit emergent properties ("few-shot" learning with no domain adaptation [57]). The latest version, GPT-4 [7], powers Open AI's ChatGPT and has been trained on vast amounts of data (on the order of 45 GB), consisting of a collection of models with over 1.7 trillion parameters.

In addition to the ability of GPTs to predict the next word and their sheer size, another important element of their success is training with human feedback [37], [38]. This involves humans rating the output of the model and ranking several model outputs for appropriateness. In both cases, the human ratings are used to further improve the model and moderate its output.

A transformer model can also be adapted to generate images from text. At a high level, the model is trained on a vast collection of images with accompanying text descriptions. This kind of model powers OpenAI's DALL-E image-generation model [8]. Many of the same principles can be applied to other forms of input and output, including audio and video.

APPENDIX B: SURVEY OF DEANS ON GenAI IN RESEARCH

Associate Deans for Research Responding to the Survey.

- Michelle Amazeen, Associate Dean for Research, College of Communication
- John Byers, Senior Associate Dean of the Faculty for Mathematical & Computational Sciences, College of Arts & Sciences
- Arianne Chernock, Associate Dean of the Faculty for Social Sciences, College of Arts & Sciences

- Leslie Dietiker, Associate Dean for Research, Wheelock College of Education & Human Development
- Kate Flewelling, Director, Library Services, BU Medical Campus
- Maria Kukuruzinska, Associate Dean for Research, Henry M. Goldman School of Dental Medicine
- Michael D. McClean, Associate Dean for Research and Faculty Advancement, School of Public Health
- Sean Mullin, Associate Dean of the Faculty for Natural Sciences, College of Arts & Sciences
- Andrew Taylor (conveying responses from multiple colleagues), Associate Dean of Research, Chobanian & Avedisian School of Medicine

Summary of Survey Responses.

Question 1: What value, if any, do you believe AI can bring to scholarly research?

- Research and Methods Planner
- Literature reviews
- Generates new hypothesis
- Content Generator and thought partner (e.g. brainstorm research ideas)
- Data Collection and Analysis (e.g. accelerate analysis)
- Summarizes larger amounts of research data in shorter time
- Make coding easier and faster
- Edit, proofread, and help with creating graphics, reviewing manuscripts, and writing grant applications
- Improve peer review times

Question 2: Do you have any concerns with the use of AI in research at BU? What are they?

- Can lead to more reliance on pattern recognition without understanding
- Bias or discrimination in data
- Equitable access
- Security threats
- Potential of proliferation of misinformation
- May make plagiarism and falsification of data easier to do and harder to detect
- Issues around intellectual property
- Image manipulation hard to detect
- Concerns that putting confidential/prepublication information onto an AI server could prematurely release information into the public domain

Question 3: What kinds of central support do you believe faculty and other researchers would benefit from?

- Tutorials/demonstrations/guiding principles
- A searchable list of software providing ChatGPT and AI programming providing information as being open or closed, accuracy, and other problems
- Repository of generated programming involving vetted routine administrative activities
- Providing a framework for AI that defines infrastructure along with concerns

- Information sessions and accessibility to machine learning algorithms, perhaps training sessions on deep-learning and other subsets of machine learning
- Help researchers keep up to date on the latest opportunities, potential pitfalls, and best practices

Question 4: Do you believe that BU should provide guidelines for the use of AI in research?

- 6 Yes (perhaps in partnership with other AAU institutions, or the AAU)
- Not guidelines but BU researchers should have ways to stay up to date on the latest opportunities, potential pitfalls, and best practices as the field evolves

Question 5: Do you see the use of AI in research changing how we supervise and mentor the research of students, especially PhD students?

- Most see it as a new tool, and it will be important to teach about its use

Question 6: Is the work into AI METHODS a focus of research among researchers in your school or college?

- Currently most common use is analyzing datasets

Question 7: For the USE of AI in research in your school/ college, is Ai used as decision support or automation? The former is verifiable (e.g. coding). Are results subjective or objective?

- Most did not have enough information to accurately answer that question at this time

APPENDIX C: A STUDY OF GenAI TEXT DETECTORS

A number of AI-detection tools have been released since GPT-2 became available in 2019 [56]. Unfortunately, none of them can reliably and consistently detect AI outputs [58], [59]. The fundamental problem is that GenAI outputs are usually unique and carry no "watermark" or signature that would confirm that they were produced by a machine. Although some of the tools boast high accuracy rates, independent studies indicate that all these tools are fallible, especially if the output has been modified by humans. Even a small chance of a false positive – that is, a result that falsely claims the writing was produced by GenAI – means that students could be unjustly accused of academic misconduct. This is not an acceptable risk.

The AI task force arranged for a systematic analysis of the reliability (accuracy and consistency) of GenAI detectors. This analysis was produced in association with the spring 2024 edition of the CDS DS380 "Data, Society, and Ethics" course (for authors and acknowledgements, see Appendix F). This analysis is a useful practical guide for professors as they consider how to detect GenAI material in student submissions.

There are two reliability issues at stake in the small landscape of GenAI detectors. On the one hand, while text detectors are generally well over 90% *accurate*, the probabilistic nature of detector judgments allows for the possibility that they can be badly mistaken, in both directions, producing false negatives (they mistakenly find human-generated text) and false positives (they mistakenly find GenAI text). The latter can be devastating for a student who did nothing improper. On the other hand, students may want to run a GenAI detector on an assignment to see what their grader will see, thereby protecting themselves from disciplinary charges based on a false positive detector

reading. This calls for reliability in the sense of consistency, but GenAI text have been reported as supplying quite different probability estimates on the very same document at different times.

Some GenAI policies, such as the Generative AI Assistance Policy from the Faculty of Computing & Data Sciences, encourage students to employ GenAI detectors to protect themselves against false accusations that might lead to disciplinary action. But such recommendations are only fully feasible if GenAI detectors give consistent readings.

The experiment involved seven GenAI text detectors, seven varieties of texts with various combinations of human-creation and GenAI creation to evaluate accuracy, and three time points one day and one week apart to evaluate consistency.

The seven texts used in this experiment had the following origins and characteristics:

- AI-made with ChatGPT 4;
- AI-made with Gemini (formerly Bard);
- Made by a native English-speaking human;
- Made by a non-native English-speaking human;
- AI-made (ChatGPT 4) with intentional misspellings;
- AI-made (Gemini) with intentional grammar errors;
- AI-made (ChatGPT 4) with human-made sentences interposed.

The seven AI Detectors were those that did not require paying a fee and delivered either a numeric probability estimate or a boolean guess (human or AI):

- Copyleaks (https://copyleaks.com/ai-content-detector) – detects GPT-2, GPT-3, ChatGPT, GPT-4, T5, Bert, and Jasper.
- Originality (https://originality.ai/) – detects GPT-3, 3.5, and ChatGPT.
- Crossplag (https://app.crossplag.com/individual/detector).
- Sapling (https://sapling.ai/ai-content-detector) – detects ChatGPT and Gemini; developed by former researchers at University of California Berkeley, Stanford University, Google, and Meta; free version allows up to 2,000 characters.
- Winston AI (https://gowinston.ai/) – detects GPT-3, ChatGPT, and GPT-4; free version allows up to 2,000 words.
- Content At Scale (https://contentatscale.ai/) – detects GPT-2, GPT-3, 3.5, ChatGPT, and GPT-4.
- GPTZero (https://gptzero.me/) – to be carefully distinguished from ZeroGPT, which seems designed to benefit from GPTZero's fame, and which we found to be unreliable.

The experimental design involved the following steps:

- Apply each AI detector to the test text and record detection results.
- Repeat the testing procedure three times for each detector (Day 1: Feb 1, 2024, Day 2: Feb 2, 2024, Day 7: Feb 8, 2024) and use a different computer and a different IP address for Day 7.
- Evaluate detectors' ability to consistently classify native, ESL, AI-generated, and modified texts.

Evaluation involved assessing:

- consistency of ratings;
- success rate of AI detectors in identifying human modifications aimed at deceiving them;
- appearance of paywalls that limit access.

Results (percentages are probability that text is AI-generated, orange shading means bad detector guess, P means a paywall was encountered and testing did not continue):

		AI-made (ChatGPT4)	AI-made (Gemini)	Human (native English)	Human (non-native English)	AI-made (ChatGPT4) w intentional misspellings	AI-made (Gemini) w intentional grammar errors	AI-made (ChatGPT4) w human sentences interposed
Copyleaks	**Day 1**	100%	86%	0%	0%	100%	100%	100%
	Day 2	100%	86%	0%	0%	100%	100%	P
	Day 7	100%	100%	0%	0%	100%	100%	100%
Originality	**Day 1**	100%	100%	99%	0%	100%	99%	P
	Day 2	P	P	P	P	P	P	P
	Day 7	100%	100%	1%	0%	100%	80%	91%
Crossplag	**Day 1**	81%	0%	100%	0%	0%	0%	100%
	Day 2	81%	0%	100%	0%	0%	0%	100%
	Day 7	81%	0%	100%	0%	0%	0%	100%
Sapling	**Day 1**	100%	99.9%	0%	99.1%	100%	100%	94.7%
	Day 2	100%	99.9%	0%	99.1%	100%	100%	94.7%
	Day 7	100%	99.9%	0%	99.1%	100%	100%	94.7%
Winston AI	**Day 1**	100%	100%	1%	80%	100%	100%	100%
	Day 2	100%	P	P	P	P	P	P
	Day 7	100%	100%	1%	19%	100%	100%	100%
Content At Scale	**Day 1**	Human	Human	Human	Human	AI	AI	Human
	Day 2	Human	Human	Human	Human	AI	AI	Human
	Day 7	Human	Human	Human	Human	AI	AI	Human
GPTZero	**Day 1**	98%	78%	1%	3%	98%	62%	97%
	Day 2	98%	78%	1%	3%	98%	62%	97%
	Day 7	98%	78%	1%	2%	98%	74%	97%

Conclusion: All seven detectors were largely consistent, though there were problems with Originality and Winston AI. Copyleaks and GPTZero were the most accurate and also largely consistent on this limited test, with Copyleaks handling intentional grammar errors better. Thus, for the purposes of consistent ratings for both students and graders, on the basis of this limited evaluation, Copyleaks is the best option followed by GPTZero. The others are too inaccurate or too inconsistent to use safely.

APPENDIX D: THE GenAI POLICY AND ADVICE LANDSCAPE PROJECT

The AI task force arranged for a web-based app that provides access to a searchable repository of over 100 GenAI policy and advice documents from institutions of higher education. This app also includes semantic network analysis and topic modeling for those documents. This app was produced by the members and TAs of the Fall 2023 edition of the CDS DS380 "Data, Society, and Ethics" course, with the help of two undergraduate designers and developers.

During the spring 2024 semester, a machine-learning student group from SPARK! enhanced this site in several ways. They expanded the dataset to include policy and advice documents from K-12 education, government, and business domains, as well as higher education. They built a topic model and a semantic network analysis based on themes instead of documents. In addition, they used GenAI to automate document tagging and they upgraded the interface.

APPENDIX E: EXAMPLES OF GenAI USE AND FOR RESEARCH AND EDUCATION

We provide examples of GenAI use for research and education purposes. The examples referring to class activities and assignments have been developed by faculty in the CAS Writing Program and the Rhetoric Department at CGS.

- *Fine-tuning LLMs for specific uses.* As mentioned in Section 3.2, LLMs can be used to analyze and extract useful information from a variety of documents, which can then be used to train various models. Generic LLMs, such as ChatGPT, Gemini, Llama, have been trained on large text corpora but may lack an ability to process specialized text. One prominent example is clinical text, which is rich in specialized medical terms and abbreviations. Motivated by applications in healthcare, researchers have been able to use available datasets of clinical reports (often the MIMIC dataset, consisting of ICU clinical reports and discharge notes [60], [61]) to fine-tune generic LLMs for analyzing and extracting information from clinical notes. One such effort is based on the open-source Llama LLM [62].
- *Using LLMs to streamline literature review.* Section 3.2 mentions using LLMs to automate literature review. Asking an LLM to write a literature review for a specific project is not the best way to leverage its abilities and, because of the hallucinations we have discussed, it may produce plausible but false references. As an example, two lawyers have been sanctioned for submitting a legal brief that included six fictitious case citations generated by ChatGPT [63]. In work soon to be published, a group of researchers at the Hariri Institute developed an approach where an LLM is fed with abstracts of published papers (extracted using a set of keywords from a public database such as PubMed). The LLM is then asked whether a given abstract indicates that the paper contains a result that the researchers care about. Through elaborate prompt engineering, it becomes possible to obtain accurate YES/NO responses from the LLM. Such an approach can be used to narrow down a potential large number of papers one may recover through keyword search to a more manageable list that a human group can then study.
- *Brainstorming counterarguments*. Instructors may ask students to use a Generative AI (GenAI) platform to list potential objections to or limitations of a claim they are making in their papers or projects. In such an activity, students consider and evaluate the objections

created by GenAI, and choose at least one objection to acknowledge and respond to in their paper. If using GenAI-generated objections, students should be asked to: (1) acknowledge the use of GenAI in their paper (e.g., list in the Works Cited or acknowledge GenAI usage in a letter to the reader), and (2) express the objection in their own words, rather than copy the GenAI-produced text.

- *Practicing debate with Generative AI platforms.* One can use GenAI platforms to help students prepare for a class debate. Many GenAI platforms now have a voice feature (e.g., ChatGPT-4 mobile app) where an instructor can have their students engage in a real-time voice conversation with the platform. Students choose a position (affirmation or negation of a resolution), and interact with the GenAI platform. Such an activity develops critical thinking and argumentation skills; helps students practice articulating and defending a position on a specific topic and encourages students to critically analyze Gen AI's responses and generate thoughtful counterarguments in return.
- *Using GenAI as a peer reviewer to solicit feedback.* One can ask students to interact with GenAI as a peer reviewer. First, have students solicit feedback on their paper draft or an excerpt from their draft, listing two or three criteria they want GenAI to use in evaluating their draft. These criteria should be directly connected to the student assignment. Next, encourage students to critically evaluate the feedback they received from Gen AI and to decide whether or not they agree with it and how they might respond to it.
- *Understanding an author's frame of reference.* Many writers have a particular frame of reference: for example, early essayists like Michel de Montaigne and Francis Bacon frequently quote Roman authors, Virginia Woolf often references 19th and 20th century literature, and so on. Using Gen AI, students can get some quick information about a group of referenced authors or works, and also learn what those authors and works have in common, giving them a stronger sense of an author's frame of reference.
- *Getting out of the weeds.* When a course presents students with multiple perspectives on a topic or issue, they might get confused or overwhelmed. At such a point, it can help to ask them to identify their questions (perhaps by freewriting for a few minutes), then pose their questions to GenAI. Sharing their questions and the GenAI's answers in discussion can help students dispel confusion and feel better prepared to move forward.
- *Comparing human response to machine analysis to better understand **pathos**.* Gen AI can do a surprisingly good job of analyzing the emotional appeal of essays and stories, and even of particular passages if one uploads the text. Students can learn about the nuances of emotional appeal by comparing GenAI's analysis to their own. Often, GenAI gives an articulate analysis, and yet misses subtleties.

REFERENCES

[1] "Publishing ethics | Elsevier policy," www.elsevier.com. Accessed: Jan. 10, 2024. [Online]. Available: https://www.elsevier.com/about/policies-and-standards/publishing-ethics

[2] "Tools such as ChatGPT threaten transparent science; here are our ground rules for their use," *Nature*, vol. 613, no. 7945, pp. 612–612, Jan. 2023, doi: 10.1038/d41586-023-00191-1.

[3] "IEEE CSS Publications | IEEE Control Systems Society." Accessed: Jan. 10, 2024. [Online]. Available: https://www.ieeecss.org/publications

[4] "Notice to research community: Use of generative artificial intelligence technology in the NSF merit review process | NSF - National Science Foundation." Accessed: Jan. 10, 2024. [Online]. Available: https://new.nsf.gov/news/notice-to-the-research-community-on-ai

[5] "Plagiarism Detection Software | iThenticate." Accessed: Jan. 11, 2024. [Online]. Available: https://www.ithenticate.com

[6] "Plagiarism Checker | Grammarly." Accessed: Feb. 17, 2024. [Online]. Available: https://www.grammarly.com/plagiarism-checker

[7] OpenAI, "GPT-4 Technical Report." 2023.

[8] "DALL·E 2." Accessed: Jan. 09, 2024. [Online]. Available: https://openai.com/dall-e-2

[9] "Microsoft Copilot: Your everyday AI companion," Microsoft Copilot: Your everyday AI companion. Accessed: Apr. 02, 2024. [Online]. Available: https://ceto.westus2.binguxlivesite.net/

[10] "Gemini - Google DeepMind." Accessed: Jan. 09, 2024. [Online]. Available: https://deepmind.google/technologies/gemini/#introduction

[11] "Gemini - chat to supercharge your ideas," Gemini. Accessed: Feb. 17, 2024. [Online]. Available: https://gemini.google.com

[12] "Llama 2," Meta AI. Accessed: Jan. 09, 2024. [Online]. Available: https://ai.meta.com/llama-project

[13] "Mistral AI." Accessed: Feb. 04, 2024. [Online]. Available: https://mistral.ai/product/

[14] "Claude." Accessed: Jan. 09, 2024. [Online]. Available: https://claude.ai/login

[15] "Midjourney Documentation and User Guide." Accessed: Jan. 09, 2024. [Online]. Available: //docs.midjourney.com

[16] "Stable Diffusion Online." Accessed: Jan. 09, 2024. [Online]. Available: https://stablediffusionweb.com/#demo

[17] "Adobe Firefly - Free Generative AI for creatives." Accessed: Feb. 11, 2024. [Online]. Available: https://www.adobe.com/products/firefly.html

[18] "AI Music Generator - SOUNDRAW." Accessed: Jan. 09, 2024. [Online]. Available: https://soundraw.io/

[19] "BandLab SongStarter | Generate royalty-free music," BandLab. Accessed: Jan. 09, 2024. [Online]. Available: https://www.bandlab.com/songstarter

[20] "Pictory - Easy Video Creation For Content Marketers," Pictory.ai. Accessed: Jan. 09, 2024. [Online]. Available: https://pictory.ai/

[21] "Unlock Generative AI Content At Scale | Synthesys.io." Accessed: Jan. 09, 2024. [Online]. Available: https://synthesys.io/

[22] "Sora: Creating video from text." Accessed: Feb. 17, 2024. [Online]. Available: https://openai.com/sora

[23] "Simplified: An Easy to Use All-In-One App For Modern Marketing Teams." Accessed: Jan. 09, 2024. [Online]. Available: https://simplified.com/

[24] "Create Presentation Slides With AI in seconds with SlidesAI." Accessed: Jan. 09, 2024. [Online]. Available: https://slidesai.io

[25] "Gamma App: Generate AI Presentations, Webpages & Docs." Accessed: Jan. 09, 2024. [Online]. Available: https://gamma.app?lng=en

[26] J. Wei *et al.*, "Chain-of-thought prompting elicits reasoning in large language models," *Adv. Neural Inf. Process. Syst.*, vol. 35, pp. 24824–24837, 2022.

[27] M. M. Grynbaum and R. Mac, "The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work," *The New York Times*, Dec. 27, 2023. Accessed: Jan. 06, 2024. [Online]. Available: https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html

[28] "GPT-4," *Wikipedia*. Jan. 06, 2024. Accessed: Jan. 06, 2024. [Online]. Available: https://en.wikipedia.org/w/index.php?title=GPT-4&oldid=1194013135

[29] "Introducing ChatGPT Team." Accessed: Jan. 11, 2024. [Online]. Available: https://openai.com/blog/introducing-chatgpt-team

[30] "FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence | The White House." Accessed: Jan. 06, 2024. [Online]. Available: https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/

[31] M. Wong, "There Was Never Such a Thing as 'Open' AI," The Atlantic. Accessed: Jan. 06, 2024. [Online]. Available: https://www.theatlantic.com/technology/archive/2024/01/ai-transparency-meta-microsoft/677022/

[32] X. Qi *et al.*, "Fine-tuning Aligned Language Models Compromises Safety, Even When Users Do Not Intend To!" arXiv, Oct. 05, 2023. doi: 10.48550/arXiv.2310.03693.

[33] "Universal and Transferable Attacks on Aligned Language Models." Accessed: Jan. 06, 2024. [Online]. Available: https://llm-attacks.org/

[34] "Researchers Say Guardrails Built Around A.I. Systems Are Not So Sturdy - The New York Times." Accessed: Jan. 06, 2024. [Online]. Available: https://www.nytimes.com/2023/10/19/technology/guardrails-artificial-intelligence-open-source.html?searchResultPosition=1

[35] "GPTZero | The Trusted AI Detector for ChatGPT, GPT-4, & More." Accessed: Jan. 09, 2024. [Online]. Available: https://gptzero.me/

[36] W. Liang, M. Yuksekgonul, Y. Mao, E. Wu, and J. Zou, "GPT detectors are biased against non-native English writers," *Patterns*, vol. 4, no. 7, Jul. 2023, doi: 10.1016/j.patter.2023.100779.

[37] N. Stiennon *et al.*, "Learning to summarize with human feedback," *Adv. Neural Inf. Process. Syst.*, vol. 33, pp. 3008–3021, 2020.

[38] L. Ouyang *et al.*, "Training language models to follow instructions with human feedback," *Adv. Neural Inf. Process. Syst.*, vol. 35, pp. 27730–27744, Dec. 2022.

[39] "What is 'dark data' and how is it adding to all of our carbon footprints?," World Economic Forum. Accessed: Feb. 18, 2024. [Online]. Available: https://www.weforum.org/agenda/2022/10/dark-data-is-killing-the-planet-we-need-digital-decarbonisation/

[40] "AI for Scientific Discovery - A Workshop | National Academies." Accessed: Jan. 09, 2024. [Online]. Available: https://www.nationalacademies.org/our-work/ai-for-scientific-discovery-a-workshop

[41] A. J. Thirunavukarasu, D. S. J. Ting, K. Elangovan, L. Gutierrez, T. F. Tan, and D. S. W. Ting, "Large language models in medicine," *Nat. Med.*, vol. 29, no. 8, pp. 1930–1940, 2023.

[42] M. F. Romano, L. C. Shih, I. C. Paschalidis, R. Au, and V. B. Kolachalama, "Large Language Models in Neurology Research and Future Practice," *Neurology*, vol. 101, no. 23, pp. 1058–1067, Dec. 2023, doi: 10.1212/WNL.0000000000207967.

[43] S. Amini *et al.*, "Automated detection of mild cognitive impairment and dementia from voice recordings: A natural language processing approach," *Alzheimers Dement.*, vol. 19, no. 3, pp. 946–955, 2023, doi: 10.1002/alz.12721.

[44] S. Vemprala, R. Bonatti, A. Bucker, and A. Kapoor, "Chatgpt for robotics: Design principles and model abilities," *Microsoft Auton Syst Robot Res*, vol. 2, p. 20, 2023.

[45] I. Singh *et al.*, "Progprompt: Generating situated robot task plans using large language models," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2023, pp. 11523–11530.

[46] K. M. Jablonka *et al.*, "14 examples of how LLMs can transform materials science and chemistry: a reflection on a large language model hackathon," *Digit. Discov.*, vol. 2, no. 5, pp. 1233–1250, 2023.

[47] V. Veselovsky, M. H. Ribeiro, A. Arora, M. Josifoski, A. Anderson, and R. West, "Generating Faithful Synthetic Data with Large Language Models: A Case Study in Computational Social Science," *ArXiv Prepr. ArXiv230515041*, 2023.

[48] T. H. Trinh, Y. Wu, Q. V. Le, H. He, and T. Luong, "Solving olympiad geometry without human demonstrations," *Nature*, vol. 625, no. 7995, Art. no. 7995, Jan. 2024, doi: 10.1038/s41586-023-06747-5.

[49] F. Dell'Acqua *et al.*, "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality." Rochester, NY, Sep. 15, 2023. doi: 10.2139/ssrn.4573321.

[50] J. Carter, J. Feddema, D. Kothe, R. Neely, J. Pruet, and R. Stevens, "AI for Science, Energy, and Security Report 2023 | Argonne National Laboratory," 2023. Accessed: Mar. 31, 2024. [Online]. Available: https://www.anl.gov/ai/reference/AI-for-Science-Energy-and-Security-Report-2023

[51] "NOT-OD-23-149: The Use of Generative Artificial Intelligence Technologies is Prohibited for the NIH Peer Review Process." Accessed: Feb. 11, 2024. [Online]. Available: https://grants.nih.gov/grants/guide/notice-files/NOT-OD-23-149.html

[52] I. Goodfellow *et al.*, "Generative Adversarial Nets," in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2014. Accessed: Jan. 06, 2024. [Online]. Available:

https://proceedings.neurips.cc/paper_files/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afccf3-Abstract.html

[53] A. Vaswani *et al.*, "Attention is all you need," *ArXiv Prepr. ArXiv170603762*, 2017.

[54] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," *ArXiv Prepr. ArXiv181004805*, 2018.

[55] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving language understanding by generative pre-training," 2018, [Online]. Available: https://www.mikecaptain.com/resources/pdf/GPT-1.pdf

[56] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, "Language models are unsupervised multitask learners," *OpenAI Blog*, vol. 1, no. 8, p. 9, 2019.

[57] T. Brown *et al.*, "Language models are few-shot learners," *Adv. Neural Inf. Process. Syst.*, vol. 33, pp. 1877–1901, 2020.

[58] A. Pegoraro, K. Kumari, H. Fereidooni, and A.-R. Sadeghi, "To ChatGPT, or not to ChatGPT: That is the question!" arXiv, Apr. 05, 2023. doi: 10.48550/arXiv.2304.01487.

[59] E. N. Crothers, N. Japkowicz, and H. L. Viktor, "Machine-Generated Text: A Comprehensive Survey of Threat Models and Detection Methods," *IEEE Access*, vol. 11, pp. 70977–71002, 2023, doi: 10.1109/ACCESS.2023.3294090.

[60] A. E. Johnson *et al.*, "MIMIC-III, a freely accessible critical care database," *Sci. Data*, vol. 3, no. 1, pp. 1–9, 2016.

[61] A. E. Johnson *et al.*, "MIMIC-IV, a freely accessible electronic health record dataset," *Sci. Data*, vol. 10, no. 1, p. 1, 2023.

[62] S. Kweon *et al.*, "CAMEL | Clinically Adapted Model Enhanced from LLaMA," CAMEL | Clinically Adapted Model Enhanced from LLaMA. Accessed: Mar. 31, 2024. [Online]. Available: https://starmpcc.github.io/CAMEL/

[63] S. Merken, "New York lawyers sanctioned for using fake ChatGPT cases in legal brief," *Reuters*, Jun. 26, 2023. Accessed: Mar. 31, 2024. [Online]. Available: https://www.reuters.com/legal/new-york-lawyers-sanctioned-using-fake-chatgpt-cases-legal-brief-2023-06-22/