

# Emerging AI Tools for Education and Research: Perspective and Policies for IISc

Sridharan Devarajan (CNS), Siddhartha Gadgil (MA), Kavita Isvaran (CES),  
Garima Jindal (OC), Arindam Khan (CSA), Viraj Kumar (KIAC), Y. Narahari (CSA),  
Danish Pruthi (CDC) Ramsharan Rangarajan (ME), and Suresh Sundaram (AE)

This Version: January 10, 2024

## Abstract

The Director, IISc, constituted a committee on August 21, 2023 with the following mandate: (a) Explore the challenges and benefits of emerging AI tools in the context of academic teaching and learning, (b) provide guidance and recommendations to faculty members on the use of AI tools, (c) share best practices from elsewhere, (d) explore innovative and ethical uses for better learning experiences and outcome, and (e) suggest methods to ensure the integrity of assessments. This committee (comprising the authors of this report) expanded the scope of this mandate by including other pertinent related topics, including research activities, and is pleased to share its findings through this report.

## 1 Executive Summary

Generative Artificial Intelligence tools and other such emerging tools (henceforth: *AI tools*) are evolving rapidly in their capabilities. These tools can generate complex digital artefacts in response to simple and meaningful natural-language prompts, and they are being rapidly integrated into professional workflows. This report examines the implications of easy, unconstrained access to such AI tools for education and research activities at IISc. The analysis and recommendations presented here are based on inputs from the IISc community (faculty and students), a review of global practices, and deliberations among the committee members.

### 1.1 Key Findings

IISc faculty and students have already recognised that the emergence of AI tools is a double-edged sword. On one hand, the ready availability of these tools can have positive implications for education, research, and learning. At the same time, the indiscriminate use of these tools is clearly detrimental to academic goals. It is the committee's view that there are well-justified and responsible ways to exploit AI tools that maximise the benefits and minimise the drawbacks.

### 1.1.1 Insights from Faculty and Student Surveys

#### Insights from Faculty Survey

A survey form was circulated among IISc faculty members and 97 responses were received. Many faculty members appear to be genuinely grappling with various challenges posed by the sudden appearance of AI tools. Some faculty members have switched from unproctored, take-home exams to in-class exams. Others have modified their assignments and tests so that AI tools are unlikely to affect assessment. Interestingly, many faculty members expect a positive impact of AI tools in learning. For example, more advanced assignments are feasible now that students have access to AI tools. Also, faculty members appear to find the use of AI tools acceptable for grammar and stylistic corrections to research writing. They agree that AI tools presently fall short in generating creative new ideas.

#### Insights from Student Survey

A survey form was circulated among IISc students and 525 responses were received. By and large, students appear to be interested in and highly supportive of using AI tools in academics and course work. They recommend conducting training programs on the responsible and ethical use of AI-tools, based on clear guidelines laid down by IISc. In general, paraphrasing and rewriting material for clarity were considered the major plus points, whereas logical reasoning, problem-solving, or mathematical proofs were considered the least helpful aspects of AI tools. An interesting, out-of-the-box suggestion was to explore the use of AI tools as a *virtual counsellor* to engage with anxious/depressed students.

### 1.1.2 Insights and Recommendations from the Committee

#### AI Tools are Error-Prone; Faculty should Track Improvements

The ability of AI tools to perform complex tasks in multiple domains (in the context of both coursework and research) is advancing rapidly. The content generated by these tools can contain flaws, some of which are subtle. Thus, anyone using AI tools for coursework, research, professional communication, etc. is responsible for verifying the accuracy and appropriateness of the generated content. Further, it is important for faculty members to keep a careful eye on developments in the quality and utility of AI tools in their respective areas of expertise.

#### AI Tool-Usage Policies

A high proportion of coursework at IISc is unproctored, including homeworks, projects, take-home exams, etc. Indiscriminate usage of AI tools by students on these tasks will be detrimental to their learning. Banning access to these tools is not prudent for many good reasons. Moreover, such a measure would be unpopular with students and difficult to enforce, given their ready availability. At present, automated tools for detecting AI-generated content are *unreliable*. Thus, there is an urgent need to articulate clear usage policies for AI tools. This report proposes institution-level policies and encourages programmes, divisions, departments, and course instructors to tailor these baseline policies as necessary. To limit unfair use of AI

tools, faculty may need to consider a combination of (a) new forms of “AI-proof” assessments (which may be difficult to create), and (b) manual processes such as proctored or oral examinations.

### **Positive Impacts on Teaching**

Many of the faculty members see positive uses of AI tools for coursework e.g., challenging students to solve more advanced problems using AI tools and better preparing them for professional contexts where such tools are being integrated into workflows. If such changes are made to assessments in a course, then the pedagogy and course learning outcomes may also need to be updated to reflect these shifts in emphasis. Further, AI tools can improve the productivity of both faculty (e.g., transforming notes into slides, or suggesting questions based on specified content) and students (e.g., providing personalised assistance by responding to each student’s specific needs). However, AI-generated content is far less authoritative than faculty-created content. Similarly, while students may seek some technical assistance from AI tools, TAs must continue to play a critical role in providing personal attention to students.

### **Impact on Research**

Researchers strive to use the best available tools for their work, and AI tools can be assistive in several ways. However, particularly in the context of research publications and theses, the usage of such tools must be declared as per clearly stated guidelines. These guidelines should be communicated to all members of the IISc community via appropriate channels. These guidelines must sensitise researchers about the risks of (a) entering sensitive content into AI tools (such content may be used to further train the underlying AI models, which is equivalent to posting the content in public), and (b) trusting the accuracy of AI-generated content. If any such content is included in a publication, the authors are responsible for its accuracy. It is important for faculty members to create an AI-responsible environment in their groups and it is important for the students to seek the approval of their advisors before using AI tools. Disclosure of use of AI tools should be an integral part of the theses submitted by the students.

## **1.2 Some Notes and Disclaimers**

World over, the level of current understanding of the capabilities of the *Large Language Models* (LLMs) that power the present generation of generative AI tools is limited. At the same time, the research community is actively exploring variants of LLMs and other models towards enhancing the capabilities of AI tools across a wider set of domains. Thus, it is prudent for IISc to conduct periodic surveys aimed at tracking how the use of these tools by the campus community is evolving. This data can be used to further refine or modify the specific recommendations made in this report (Sections 5.3 and 6).

This report presents what the committee members think would be appropriate practices towards a creative and constructive use of AI tools in education and research. What constitutes a best practice could vary based on the domain, purpose, process, and person. Our recommendations therefore are not universal. Individuals may adopt these best practices in a responsible manner that suits their situations optimally and correctly.

When we factor in the emerging AI tools, many (hitherto well-understood) notions such as plagiarism defy a crystal clear definition. There are conflicting notions advocated by different groups of individuals. We do not take any sides here and leave the final judgement to individual faculty members to decide on the best course of adopting what is recommended here.

### 1.3 Structure of the Report

Section 2 identifies global practices via a review of policies on the usage of AI tools put forth by premier academic institutions, journals, and publishers. These practices have informed this committee's recommendations.

Section 3 and Section 4 describe and analyse the faculty and student surveys respectively. This analysis has provided valuable inputs to the process of crafting recommendations. The insights from both surveys are summarised in the sections noted above, and further details are deferred to Appendix D.1 and Appendix D.2.

Section 5 and Section 6 summarise the Committee's views on practices for AI tool usage in the context of teaching and research respectively. Recommendations for changes to the Student Handbook and other Institute documents are specified in Section 7, and suggestions for using AI tools in research are summarised in Appendix D.6.

Section 8 focuses on expected advances in the capabilities of these tools that are of particular relevance to the IISc community.

Section 9 lists answers to frequently asked questions, and provides links to further resources. Many of these links were shared by faculty and student survey respondents, and the committee is grateful for these inputs.

The report provides an expansive collection of information in the appendices. Section A lists the questions that were included in the faculty survey while Section B lists the questions included in the student survey. For the benefit of readers, Section C provides an overview of the technology powering generative AI tools and examples illustrating their capabilities and limitations across domains. In particular, Section C.3 focuses on the inability (at present) of automated plagiarism detection tools to identify AI-generated content. Section D describes how the committee tested the AI tools for various use cases. Section D.1 reports a summary of faculty survey as produced by chatGPT while Section D.2 reports a summary of student survey as compiled by chatGPT. Section D.3 contains FAQs generated by ChatGPT while Section D.4 contains FAQs generated by BardAI. Section D.5 comprises 17 examples of mathematical reasoning that the committee members used, for eliciting responses from GPT4. Finally Section D.6 presents a summary of guidelines for use of AI tools in research as provided by ChatGPT in response to various questions.

## 2 Current Global Practices

Generative AI is viewed worldwide as a collection of tools with promise as well as perils and the consensus is in using these tools with clearly formulated guidelines to avoid their deleterious effects.

## 2.1 Some Representative Guidelines from Universities

Several universities have applied their mind in coming with guidelines for responsible use of AI tools in education and research. We provide a summary from a few selected universities.

### ETH Zurich

Regarding regulation of use, ETH Zurich in its FAQ on generative AI includes the following [1]:

*Any rapidly written overly specific or restrictive rules and regulations will run the danger of sounding like blind actionism, and apparent quick fixes might sound ridiculous a few months from now.*

*We also need to be careful about restricting tools where we cannot justify it. It makes sense to forbid the use of computers in chess tournaments, since computers could beat humans virtually anytime, which destroys the enjoyment of a wonderful game - this is similar to forbidding doping or "performance-enhancing drugs" in sports. Science, math and engineering never worked that way; we always worked with the best available tools, and in fact were expected to do so to stay competitive.*

*Having said that, simply copy/pasting the output of an AI-tool and submitting it as one's own work is unacceptable.*

### Cornell University

The committee on the use of AI tools at the Cornell University has the following to say about the opportunities and concerns surrounding GAI (Generative AI) [2]:

**Opportunity:** *GAI has been touted as a potential paradigm shift in education. Proposed benefits include providing a customized learning experience for all learners matching their individual needs; increasing accessibility for students with learning disabilities, anxiety, or language barriers; allowing instructors to scale constructive critiques for iterative learning and improvement in writing; and assisting in tasks in a number of domains including coding, creative composition, and more.*

**Concerns:** *Currently, GAI output can include inaccurate information, toxic output, biases embedded in the model through the training process, and infringement of copyrights on material and images. Students can use GAI to circumvent the process of learning and assessment in classes. In cases when GAI tools can serve learning outcomes, lack of affordable access for all students could exacerbate systemic inequalities. In addition, overreliance on these tools risks atrophying students' ability and willingness to interact with instructors and peers.*

## 2.2 Widely Followed Policies

We discuss next specific policies and suggestions that seem to be widely followed, emphasizing responsible use of AI tools with **disclosure, data and information protection, content re-**

**sponsibility, fair and ethical use**, and following **course policies**. The summary here closely follows that of Caltech [3].

### Disclosure and Attribution

When using AI tools, always disclose promptly, or reference the use of these tools and application plug-ins, as applicable. This transparent disclosure ensures that others are aware when AI tools were used to generate content and reduces misunderstandings regarding the source of information, potentially limiting claims of academic dishonesty or plagiarism. When using AI tools to write or publish, make sure to follow the guidance provided by the course instructor or journal or manuscript publisher/editor. For example, some require that the AI tools be included as an author and others may simply require acknowledgement.

### Private or Confidential Information

It is safest to assume data or queries uploaded into open AI tools will become public information, unless otherwise indicated. Do not enter, contribute, or otherwise input sensitive, confidential, or restricted information into open AI tools.

### Errors and Content Responsibility

Remember that AI tools are fallible. Responses can be inaccurate and misleading. Therefore, one should always review and assess all output generated by AI tools for accuracy before relying on them or distributing them publicly.

### Fair and Ethical Use

Ensure that the use of AI tools is fair and ethical. Ethical issues surrounding the use of AI tools include all concerns related to accuracy of the generated outputs and privacy of the inputs discussed earlier. In addition, there are important ethical concerns related to the data used to train the underlying AI models (bias as well as copyright issues) and the environmental costs of training and deploying these tools [4].

### Course Policies

Faculty members may choose to allow or prohibit the use of AI tools in their courses. Students must follow the policies set by the instructor for each course. However, any enforcement of policies prohibiting the use of AI tools in non-proctored settings should be done with caution, as it is difficult to detect the use of AI tools. To quote from the Cornell committee's report [2]:

*... attempts to identify text generated by GAI can only be done statistically. This method will likely continue to produce both false positives and false negatives, and cannot decisively provide evidence of academic integrity violations. Using these methods currently could lead to unfairly identifying academic integrity violations (for example, bias against non-native speakers), creating a lack of trust between the instructor and students, and damaging the learning environment.*

It is suggested that faculty members adapt their courses if they are affected by AI tools. For instance, MIT has the following suggestions, with more details at [5]:

- Refine your goals for learning
- Be clear about your expectations and the acceptable use of AI tools
- Rethink your assignments & assessments
- Consider alternate subject formats
- Design for equity, accessibility, and student privacy

## 2.3 Policies Followed by Journals and Conferences

### 2.3.1 Journals

According to policies of most (perhaps all) Journals, AI tools **cannot** be listed as authors as they cannot take responsibility for correctness. The use of AI tools is permitted with appropriate detailed attribution. We quote below policies of *Nature* and the *Association for Computing Machinery (ACM)*.

#### The Nature Magazine

The Nature magazine has the following policy on the use of AI tools [6]:

*First, no LLM tool will be accepted as a credited author on a research paper. That is because any attribution of authorship carries with it accountability for the work, and AI tools cannot take such responsibility.*

*Second, researchers using LLM tools should document this use in the methods or acknowledgements sections. If a paper does not include these sections, the introduction or another appropriate section can be used to document the use of the LLM.*

#### Association for Computing Machinery (ACM)

We quote from FAQs of the ACM [7]:

*Can I use generative AI software tools to prepare my manuscript?*

*Yes, the use of such tools is permitted, subject to the following requirements:*

- *That these systems do not plagiarize, misrepresent, or falsify content in ACM submissions.*
- *That the resulting Work in its totality is an accurate representation of the authors' underlying work and novel intellectual contributions and is not primarily the result of the tool's generative capabilities.*
- *That the authors accept responsibility for the veracity and correctness of all material in their Work, including any computer-generated material.*

*If I use generative AI software tools, under what conditions must I disclose their use in my Work?*

Today's generative AI software tools perform different tasks and it is reasonable to expect that their use and use cases will expand and become more robust over time. Whether disclosure in your published ACM Work is necessary depends on how you utilize these tools. The following guidelines should be followed when determining whether and to what extent disclosure must be provided in your Work.

- *If one is using generative AI software tools such as ChatGPT, Jasper, AI-Writer, Lex, or other similar tools to generate new content such as text, images, tables, code, etc., you must disclose their use in either the acknowledgements section of the Work or elsewhere in the Work prominently. The level of disclosure should be commensurate with the proportion of new text or content generated by these tools.*
- *If entire sections of a Work, including tables, graphs, images, and other content were generated by one of these tools, you should disclose which sections and which tools and tool versions you used to generate those sections by preparing an Appendix or a Supplementary Material document that describes the use, including but not limited to the specific tools and versions, the text of the prompts provided as input, and any post-generation editing (such as rephrasing the generated text). Authors should also note that the amount or type of generated text allowable may vary depending on the type of the section or paper affected. For example, using such tools to generate portions of a Related Work section is fundamentally different than generating novel results or interpretations.*
- *If the amount of text being generated is small (limited to phrases or sentences), then it would be sufficient to add a footnote to the relevant section of the submission utilising the system(s) and include a general disclaimer in the Acknowledgements section.*
- *If one is using generative AI software tools to edit and improve the quality of your existing text in much the same way you would use a typing assistant like Grammarly to improve spelling, grammar, punctuation, clarity, engagement or to use a basic word processing system to correct spelling or grammar, it is not necessary to disclose such usage of these tools in your Work.*

*Can a generative AI tool be listed as an author?*

*No, generative AI software tools cannot be listed as authors on ACM Works under any conditions.*

### **2.3.2 Conferences**

Major computer science conferences have started specifying their policies regarding the use of AI tools including large image deep generative models (image DGMs such as Stable Diffusion).



While some conferences, like ICML (International Conference on Machine Learning) and AISTATS (International Conference of Artificial Intelligence and Statistics), take a conservative approach and restrict certain usage of AI tools, others, such as NeurIPS (Neural Information Processing Systems), ICLR (International Conference on Learning Representation), and CHI (International Conference on Human factors in Computing) adopt a more permissive approach, emphasising transparency and accountability. These variations reflect the ongoing exploration and current understanding of the impact of AI tools in scientific publishing.

We provide a brief highlight of different policies included at some of the major conference venues in computer science.

### NeurIPS 2023

*Use of Large Language Models (LLMs): We welcome authors to use any tool that is suitable for preparing high-quality papers and research. However, we ask authors to keep in mind two important criteria. First, we expect papers to fully describe their methodology, and any tool that is important to that methodology, including the use of LLMs, should be described also. For example, authors should mention tools (including LLMs) that were used for data processing or filtering, visualization, facilitating or running experiments, and proving theorems. It may also be advisable to describe the use of LLMs in implementing the method (if this corresponds to an important, original, or non-standard component of the approach). Second, authors are responsible for the entire content of the paper, including all text and figures, so while authors are welcome to use any tool they wish for writing the paper, they must ensure that all text is correct and original.*

### ICML 2023

*Papers that include text generated from a large-scale language model (LLM) such as ChatGPT are prohibited unless the produced text is presented as a part of the paper's experimental analysis. . . . (It) prohibits text produced entirely by LLMs (i.e., "generated"). This does not prohibit authors from using LLMs for editing or polishing author-written text. The LLM policy is largely predicated on the principle of being conservative with respect to guarding against potential issues of using LLMs, including plagiarism. The LLM policy applies to ICML 2023. We expect this policy may evolve in future conferences as we understand LLMs and their impacts on scientific publishing better.*

### AISTATS 2023

*LLMs and image DGMs are not allowed for the following use cases: Fully automatically generating text of more than one page, unless the produced text is presented as a part of the paper's experimental analysis. Generating quantitative figures (such as learning curves), unless the produced images are presented as a part of the paper's experimental analysis. Other potential use cases of LLMs such as polishing text*

(e.g., paragraph-wise, prompted by a manually-written paragraph of content) are not banned.

Even with the usage of LLMs and DGMs, it is still the authors' responsibility to ensure the quality, correctness, and originality of their submission(s). We ask the authors to respect the academic publishing process and ensure their submission(s) do not constitute scientific misconduct (e.g., plagiarism, deceptive figures, dual submissions, etc.).

If a paper with LLM-generated text and/or large DGM-generated images (except for experimental analysis purposes) has been accepted to the conference, we will require the authors to disclose this information to the Program Chairs before the camera-ready submission. If the Program Chairs are in doubt about potential scientific misconduct, then the submission(s) in question will be tested through checks for e.g., plagiarism and/or other forms of misconduct. Those submissions violating AISTATS submission policies will be rejected from publication even after the reviewing process.

## ICLR 2024

The use of LLMs is allowed as a general-purpose writing assist tool. Authors should understand that they take full responsibility for the contents of their papers, including content generated by LLMs that could be construed as plagiarism or scientific misconduct (e.g., fabrication of facts). LLMs are not eligible for authorship.

## CHI 2024

Text generated from a large-scale language model (LLM) such as ChatGPT must be clearly marked where such tools are used for purposes beyond editing the author's own text. While we will not be using tools to detect LLM-generated text, we will investigate submissions brought to our attention and will desk reject papers where LLM use is not clearly marked.

## 3 IISc Faculty Survey: A Summary

The questions from the Faculty Survey are reproduced in Appendix A.

### 3.1 A Summary of Faculty Responses

- 97 faculty members, spanning multiple science and engineering departments, responded to the survey.
- Overall, faculty appear to be most familiar with the use of AI tools for rewriting text to improve readability (27%), for summarising complex passages (18%), or for writing and debugging code (18%).

- Unproctored work contributes a significant fraction to course grades at IISc. Thirty-four percent of faculty provide 20% to 50% and 16% more than 50% of grade weightage to unproctored work; whereas 48% provide less than 20% weightage to such work.
- Faculty are roughly evenly split on the extent of the impact of AI tools on student learning. A large fraction (28%) believe that AI tools have mixed impact – both beneficial and detrimental – whereas many others either feel that it has no impact (25%) or are uncertain (25%). Only a relatively small fraction (12%) felt that it is only detrimental to student learning.
- Based on their textual responses, many appear to be genuinely grappling with the challenge posed by AI tools on student learning and modes of evaluation. Some faculty have switched from unproctored to in-class exams, others have modified their assignments and tests so that AI tools are unlikely to affect assessment, and some feel that AI tools are unlikely to be useful for the work that they set. Interestingly, some see a positive use of AI tools in learning; more advanced assignments could be set now that students have access to AI tools.
- Consistent with this view of prevalent confusion, when asked about their preferred policy for AI tools, the highest proportion (39%) had “not yet thought about this” whereas 35% would encourage students to use AI tools “sometimes”. Only 12% seek to strongly disallow the use of AI tools.
- Again, reinforcing these mixed views, while 31% of faculty are fine with the use of AI tools to replace spelling and grammar checking tools, and 21% are fine with its use for composing emails, only a small fraction are fine with using it for homework (11%), research proposals (12%), research publications (10%) or paper reviews (10%).
- While the vast majority of faculty (95%) feel the need for IISc to maintain an up-to-date resource page for faculty, only about 20% are willing to contribute to it.

### 3.2 A Summary of Observations from Open-ended Textual Responses

- Overall, it appears that many faculty prefer in-class exams as the best safeguard against AI tools.
- Yet, such exams prevent students from engaging with more ambitious and deeper problems. So, take-home exams and assignments cannot be entirely eliminated.
- Some have suggested supplementing these take-home exams with an oral presentation in class to gauge the student’s level of understanding.
- Although this may work for small classes, it may not be a good solution for large (50-100+) class sizes.
- One compromise solution is proctored open-book exams. In this case, we would allow textbooks or other source material but no internet access in the exam hall. The exams can be 4-8 hours long, allowing the faculty to pose challenging questions or assignments.

- Several faculty felt that AI tools are unlikely to affect the unproctored work they set because these tools are unlikely to help the type of tasks and questions involved.
- Most appear to be happy for students to use AI tools for improving the clarity of their writing but not for research or solutions for exams.
- Those who have said AI tools can be used for research mention that due acknowledgment must be provided indicating the use of AI tools.
- When it comes to research, faculty opine that there is no foolproof way to monitor and prevent the use of AI tools. In this case, we may have no option but to go with an honour code.
- The overall consensus appears to be: Let us not completely restrict the use of these tools, as that would be counterproductive. Instead, let us hold training sessions for students on productive and ethical use versus inappropriate use of these tools.
- We can also alert students to the danger of unintended “idea” plagiarism with AI tools because these tools can (and will) derive their source material from already published academic papers.
- We need a clear institute-level policy for thesis content. For courses, instructors should provide clear, practical, and enforceable guidelines for the use of AI tools.

In addition to this human summary, we provide in Appendix D.1 a summary generated by ChatGPT-3.5 to specific, open-ended questions addressed by the faculty, including guidelines on modifying assessments considering the advent of AI tools, as well as use and misuse of these tools. As can be seen from these responses, ChatGPT not only summarised the large volume of responses elegantly (albeit somewhat boringly) but, upon due prompting, was also able to generate LaTeX code for visually formatting these responses.

## 4 IISc Student Survey: A Summary

The questions from the Student Survey are reproduced in Appendix B.

### 4.1 A Summary of Student Responses

- A total of 525 students participated in the survey. The largest proportion of students were from the Mechanical Sciences division (26%) followed by Electrical Sciences (19%). Other divisions had 10% or less representation. Interestingly, 17% of the respondents were UG students.
- Students from Engineering departments, and Mechanical Engineering, in particular, were most strongly represented compared to students from Science departments.

- Among the students, the vast majority (53%) were familiar with ChatGPT and had already used it in coursework. Interestingly, comparatively fewer students (14%) had used Github Copilot, perhaps because of explicit proscription by the instructors. Only about 6% had used image generation tools like DALL-E.
- In terms of utilizing these tools, nearly half the students were familiar with their use in the context of writing/language tasks, such as rephrasing text (28%) or summarizing research articles (19%). Around 22% had used these tools for coding. Relatively few sought to use these tools for mathematical proofs (12%) or problem-solving (9%).
- Consistent with this trend, when asked to rate AI tools in terms of their capabilities the highest rating (3.9/5) was awarded for improving readability, with progressively lower ratings (3.3) for both article summarising and code debugging. The lowest rating (2.5) was given to the problem-solving abilities of AI tools, reflecting the well-known weakness of these tools in terms of their logical vulnerabilities.
- Similarly, most students opined that coding and technical writing were key areas in which they were likely to benefit from the capabilities of AI tools. Conversely, mathematics and problem-solving were identified as the areas where AI tools were likely to be least helpful.
- Most students (over 60%) sought to use AI tools to improve the quality of their written submissions. The majority (36%) were happy to leverage such tools, whereas the remaining sought not to become dependent on them (25%). Only about 30% of the students felt that AI tools were not very useful (or useful at all) for improving written content.
- While the majority of the students (54%) felt that clear guidelines should be prescribed by the Institute for the use of AI tools, interestingly, a small but significant fraction (17%) felt such guidelines were unnecessary. The remaining (28%) were ambivalent.
- Among those who felt that such guidelines should be prescribed, many (33%) sought to make this the responsibility of the specific course or instructor. Around 20% each of the student respondents felt that this could be specified in IISc's website, the Student Handbook, or during orientation, respectively.
- When asked about their preferred Institute policy for AI tool usage in coursework or research papers, the vast majority (over 95%) favoured allowing its usage.
- Yet, for coursework, students were nearly evenly divided between allowing its unrestricted usage (26%) versus for only specific tasks (e.g., writing, 26%). Interestingly, the remaining significant fraction (40%) sought to allow free usage but with attribution.
- By contrast, these proportions were different when asked about AI tool usage in research papers. Now, the majority opinion (39%) sought to restrict its usage for specific tasks (e.g., writing), whereas a smaller proportion (33%) sought usage with attribution. Only a minority (21%) sought to allow unrestricted usage.

## 4.2 A Summary of Observations from Open-ended Textual Responses

- Generally, students appear to be highly supportive of using AI tools in academics and coursework. A helpful recommendation was to conduct training programs on the responsible and ethical use of AI-based on guidelines laid down by the University.
- Many students viewed AI as beneficial for language, creative, and scientific writing courses. Debugging code was also considered a reliable and useful application for AI tools, rendering them useful for programming courses.
- Providing summaries of course material or long papers in reading courses was also considered a major benefit. In general, paraphrasing and rewriting material for clarity were considered the major plus points, especially given that English is not always the first language for most Indian students.
- By contrast, students were highly sceptical of using AI tools for logical reasoning, problem-solving, or mathematical proofs. One may infer that many of them may have tried to use AI tools for courses of this flavour and found the results wanting.
- Some pointed to the fragility of AI tools in frequently producing egregious errors in reasoning, often ignoring simple rules of mathematical logic. Some also opined that over-reliance on AI tools in writing may create challenges for students who seek to learn to write independently.
- On balance, students consider that the benefits of AI tools outweigh their drawbacks as long as the responsible use of AI tools is encouraged. In fact, some opined that it could also save faculty time! Students may prefer to seek answers to course-related questions from AI tools rather than from faculty, and research students may benefit from assistance with repetitive and routine tasks like spelling or grammar checks in written research output .
- Overall, most were not in favour of banning or, in other ways, severely limiting the use of AI tools. Given their potential benefits, a general opinion is that even if such bans were imposed, students may find ways around it.
- Moreover, this may place us at a disadvantage over other institutions that permit the liberal use of such tools to improve the quality of academic output. In fact, some recommended that the University have a standing policy to purchase and make accessible the most recent AI tools.
- It was felt that free use and open acknowledgment of these tools was the best way forward. Students also advocated periodic reviews and revisions of usage policies (by the University) as AI tools become more efficient and capable in the future.
- Beyond these specific comments, one student opined about using AI tools as a virtual friend to relieve anxiety and stress. This is an important, potentially beneficial use case that could be overlooked often.

- It would be critical to consider this application also (“virtual counsellor”) in terms of efficacy, risks, and ethical issues, especially given that mental health issues are frequently encountered on academic campuses.

In addition to this human-generated summary, we provide in Appendix D.2 a summary generated by ChatGPT-3.5 to specific, open-ended questions addressed by the students on positive and negative aspects of AI tools and suggestions on AI tool usage in education. Again, ChatGPT summarised the large volume of responses elegantly.

## 5 Guidelines and Recommendations for Teaching

This discussion on the role of AI tools in teaching acknowledges that there is now a new entity in the classroom, and that students are likely to have easy access to increasingly powerful AI tools [34]. Changes in education are generally gradual and cautiously executed. Experiments are deliberate, and evaluating outcomes takes time. In this setting, the disruptive potential of AI tools is without any parallel. New forms of teaching and learning beckon, but known and unknown risks are at play. Hence, as learners and educators, we are at an inflection point, and both parties are eager to strike the right balance. As universities and organisations worldwide grapple with the advent of new tools and technologies, let us remind ourselves that no one knows the right way forward [8, 9, 10, 11]. Yet nearly all seem to be aware that AI tools have the potential to transform the classroom. Based on the student and faculty surveys, the perception of AI tools at IISc spans the full spectrum, from being an active deterrent to learning to having a personal teaching assistant. Faculty members are concerned that students are too excited and vice-versa.

### 5.1 Balancing Excitement with Concerns

As with most new adventures in teaching and learning, faculty members are justifiably worried that AI tools could disrupt the existing system. It is certainly the case that indiscriminate use of such tools to solve homework problems circumvents the learning process. Printer-ready outputs from AI tools often bear the reassurance of material found in well-researched textbooks despite misleading students with inaccuracies (factual, logical, etc.). Well-packaged AI tools and convenient interfaces distract students from privacy and intellectual property issues. Students could easily become over-reliant on the technology. The realistic danger is that only the machines learn, not the students. At the same time, such apprehension should be contrasted with the fact that university education in science and engineering disciplines are at peace with web browsers replacing library stacks, despite gloomy predictions of the impact search engines could have on learning. Simulation tools such as MATLAB and Mathematica have enhanced learning experiences with data analysis and visualisation tools without discouraging pen-and-paper techniques.

Largely, the excitement surrounding the adoption of AI tools in teaching is fuelled by the fact that reality does not reflect what we know about learning. Students gain immensely from personalisation of course content, individual attention from instructors, and interactive

classroom discussions [61]. AI tools may provide efficient and scalable means to achieve these goals.

## 5.2 Plausible Use Cases for AI Tools in Teaching

### Solving Complex Problems using AI Tools

For certain types of problems, the ability of students to “get the answer” may simply indicate their ability to effectively prompt an AI tool rather than reflecting their grasp of the relevant concepts (e.g., see Figures 2 and 3). Knowing that students *can* use such tools for coursework (and perhaps *will* use such tools in their professional careers), it is incumbent on faculty to question whether such problems remain an effective way to promote student learning. In some domains, faculty may see an opportunity to pose more complex questions, where students must learn to deconstruct an unstructured problem into a sequence of simpler questions (some of which could be answered by AI tools). Asking the right questions in a manner that helps solve complex problems is an essential skill that students will take away.

Further, as AI tools are integrated into professional workflows, students will benefit from having firsthand experience of using these tools before they leave the campus. Companies are adopting AI at a breakneck pace in what may be the next iteration of an industrial revolution, and students are eager to be an integral part of this transformation. To this end, the Institute and its faculty must consider that exposure to solving complex problems using AI tools may be helpful in preparing students for an AI-infused world.

### Personalised and Immersive Learning Experiences

Our courses should cater to students with varying learning speeds, giving them sufficient time to grasp and exercise new concepts. Students come to IISc from diverse backgrounds and will greatly benefit from such adaptivity. Instead, those who do not get sufficient assistance settle down in long tails of grading curves in courses with pre-planned trajectories. Such expectations, however unrealistic, should guide teaching policies.

AI tools offer the possibility of personalised assistance to students. These tools can serve as a tutor by answering questions students may hesitate to ask in class or as a source of practice problems before an exam. Spending an hour with ChatGPT or AI tools, prompting dozens of questions without worry of judgement ensures active engagement that is difficult to achieve otherwise in classes of more than a few. A showcase example of learning of this sort comes from tools providing programming assistance. Coding skills vary widely among students at IISc. Nothing short of personal assistance is effective in getting a student with little prior experience to become proficient.

A handful of examples highlight the possibilities of immersive and personalised learning experiences that can be created using AI tools. As specific examples, the addition of a tutor bot in a freshman computer science course [12], the project to create mobile applications for personalised learning [13], the virtual AI/VR museum project with hands-on experiments such as earthquake simulators [14], Andrew Ng’s popular MOOC on AI/ML, and student simulators for real-time teaching feedback.



The use cases highlighted may not be feasible to emulate but serve as a source of inspiration. They all involve significant effort from teams of faculty members and administrators with the single-minded purpose of improving student engagement. Online programs could be a place to start.

### **Language Assistance**

Many students at IISc require English language assistance. Lack of language proficiency can stand in the way of communicating their ideas clearly in oral presentations, manuscripts, and theses. Tools such as Grammarly, widely used on campus, help address some of this need. In addition, the semester-long course currently offered to students could include the use of AI tools. ETH Zurich's short-term course on using AI tools for scientific writing could be considered as an example in this context [15].

### **Course Administration**

The impact of AI on automating tedious tasks extends to teaching as well. AI tools can improve the grading experience, giving faculty more time to provide meaningful feedback on homework and exams. AI-enhanced lecture transcription can ensure that students focus on the lecture during class and not scramble with notes. Such utilities will also aid students in catching up if they miss a lecture or two [16].

## **5.3 Recommendations for Teaching**

### **Technology is No Substitute for Faculty**

Let there be no doubt that faculty will remain in charge in the classroom and that AI will not generate meaningful course content. As the recent pandemic reminded us, students need instructors to learn from and peers to learn with. At the same time, the pandemic also reminded us that our teaching methods must evolve to become more impactful. AI tools can further spur the necessity for this evolution and provide a means towards that goal.

### **AI Tools will not be Virtual Teaching Assistants**

Teaching assistants are an integral part of our courses and will remain so. While AI tools can be assistive, they cannot play all the roles expected of teaching assistants. For our courses to remain student-centric, instructors and TAs should continue to develop close interactions with students and devote personal attention to their progress.

### **Update Learning Outcomes, Pedagogy, and Assessments on Need Basis**

The institute policy should encourage incorporating aspects of AI tools in courses as instructors deem appropriate so that students learn to engage with the technology constructively. In many courses, a key learning outcome is that students will be able to create certain types of *non-trivial artefacts*. These artefacts can take the form of code, proofs, designs, etc. The

ability to create such artefacts has thus far been recognised as cognitively demanding for students. Today, as AI tools are increasingly able to create such artefacts, faculty must bear the responsibility for carefully tracking the capabilities and limitations of AI tools in their respective domains. On this basis, faculty must consider whether it is necessary to update course learning outcomes to better prepare students for the impact of these technologies on their respective professional communities. For instance, instead of primarily focusing on the ability of students to *create* such artefacts, faculty could consider emphasising the ability of students to *critique* artefacts created by AI tools. Modest changes to pedagogical and assessment strategies would be necessary to achieve these revised learning outcomes, such as including examples of plausible-but-incorrect artefacts created by AI tools in lectures, and asking quiz/exam questions where students must identify errors in a given artefact [31]. (Such examples may help students appreciate the capabilities and limitations of the technology as it matures.)

### Transparent and Simple Course Policies

An extensive collection of course policies related to the use of AI can be found at [17]. These policies fall into one of four broad categories: *prohibited*, *allowed with attribution*, *encouraged*, or *mandatory*. We recommend *allowed with attribution* as the default institutional policy. This default policy should be articulated via an appropriate channel, such as the Student Handbook. (Suggested changes to the language are specified in Section 7.1).

Furthermore, this default policy can be tailored at each level in the hierarchy: individual course instructor, department, division, or programme. (The policy specified at the most granular level will apply.) Faculty should clearly communicate the relevant policy on the permitted usage of AI tools to students at the beginning of each course. Irrespective of the specific policy, students must be made aware that they are wholly responsible for the content of their submissions.

The roles and capabilities of AI are guaranteed to change rapidly. Therefore, it is essential that the policies at each level be re-evaluated periodically, while avoiding reactionary/impulsive stances [18].

### Assume that Students will use AI Tools

It is not easy to AI-proof courses or assignments [19] (also see Section C.2), nor is it currently possible to reliably detect the use of AI tools (see Section C.3). Course policies should adopt proactive approaches during evaluation. For instance, a take-home exam should not prevent using AI resources the way in-class exams can.

### Infrastructure and Administrative Support

AI is changing fast. Students and faculty need the Institute's support to help them keep up. Training workshops will improve AI literacy on campus. Technology demonstrations by centres such as IKEN will improve technology absorption and encourage constructive use. Access to lectures on the ethical use of AI by experts in the field will remind users to exploit the technology responsibly [20]. The Institute can conduct periodic surveys to determine how the campus community uses AI tools so that its policy remains relevant and informed.

## 6 Guidelines and Recommendations for Research

Generative AI is a powerful tool for research and is likely to grow more powerful in the future (though there are some indications that their performance could degrade as well). In Science, Engineering, and Mathematics we always use the best available tools to enable us to do our best work. However, generative AI must be used responsibly. It is crucial to remember that while AI tools can be a valuable resource, they should complement, not replace, traditional academic guidance and mentorship. Students should always critically assess the information and suggestions provided by AI tools and consult with their professors, advisors, and peers for expert guidance in their specific fields of study.

We provide a few recommendations to follow, in using AI tools in research. More suggestions on capabilities and limitations of generative AI are provided in Appendix D.6.

### Disclosure and Attribution

Use of generative AI should be fully disclosed, with attribution that is precise, transparent, and with adequate details. In particular, when one is using generative AI in publications, one should follow the guidelines of the Journal, Conference, or Publisher. For thesis and other Institute documents, one should follow the guidelines for publications given below (these are similar in spirit to the ACM guidelines). It is recommended that one follows the guidelines of the Institute for all publications, including those not published by the Institute.

- If one is using generative AI software tools to generate new content such as text, images, tables, code, etc. one must disclose their use either in the **Methods** or in the **Acknowledgements** section of the Work. One must disclose details of the extent to which the tools were used.
- If entire sections of a Work, including tables, graphs, images, and other content were generated by AI tools, one should disclose which sections and which tools and tool versions were used to generate those sections, the text of the prompts provided as input, and any post-generation editing (such as rephrasing the generated text). Such tools may only be used to generate portions of a Related Work or similar section, and not for generating novel results or interpretations.
- If the amount of text being generated is small (limited to phrases or sentences), then it would be sufficient to add a footnote to the relevant section of the submission utilising the system(s) and include a general disclaimer in the Acknowledgements section.
- If one is using generative AI software tools to edit and improve the quality of existing text in much the same way one would use a typing assistant like Grammarly to improve spelling, grammar, punctuation, clarity, engagement or to use a basic word processing system to correct spelling or grammar, it is not necessary to disclose such usage of these tools in one's Work.

## Private and Confidential Information

Note that AI tools often use queries for their training, and hence may reveal their content in subsequent queries by other users. Hence, entering data into an AI tool should be regarded as equivalent to posting it in public (unless there are strong reasons to trust that a specific tool will not use your queries). In particular, do not enter, contribute, or otherwise input sensitive, confidential, or restricted information into open generative AI tools.

## Responsibility for Content

Generative AI systems are prone to errors and inaccuracies. If you use such systems in your work, as an author, you are **wholly** responsible for the correctness of the content generated by the system. You must review and assess all output generated by AI tools for accuracy before relying on them or distributing them publicly.

Another danger when using an AI tool is that its output depends on its training data, which includes content that has been written by others who need to be credited. So using the output could amount to plagiarism, especially if you use the output of open-ended questions, possibly with some paraphrasing. The use of AI tools to improve your own writing or to extract summaries from works that you are citing is likely to be safer. Once again, you are **wholly** responsible that the content of your work is either novel or has adequate attribution.

## Recommendations Specific to Programs, Departments, Research Groups, and Courses

Programs, departments, and courses may have specific rules for the use of AI tools. One must follow these rules in addition to the above rules. It is important that faculty members create an AI-responsible environment in their research groups.

# 7 Modifications to the Institute Documents

The following modifications are recommended to the Institute documents.

## 7.1 Student Handbook

In the following, suggested changes to the wording in the Student Handbook [21] (pp. 56–59) are shown in *italics*.

1. It would be appropriate to mention AI tools in principle (i) of the “principles of academic integrity” section.
  - (i) properly acknowledges and cites the use of the ideas, results, material or words of others, *where ‘others’ includes both web sources and AI tools (e.g., ChatGPT)*.
2. The definition of Plagiarism (Point 1 in the section on ‘Violations of this policy’, pg. 57) can be modified as follows:

Plagiarism: It includes the use of material, ideas, figures, code or data as one's own, without appropriately acknowledging the original source. This may involve the submission of material, verbatim or paraphrased, that is authored by another person, *by AI tools*, or published earlier by oneself.

Further, the examples can include:

*Using AI tools to assist in course work in violation of policies specified by the Course, Department, Division, or Institute (in that order of priority).*

3. The Guidelines (Point 3 in the section on 'Violations of this policy', pg. 58) may be modified as follows:
  - (a) Use proper methodology for experiments and computational work. *Any use of AI tools for tasks other than word processing (see below) must be disclosed in the **Methods** or **Acknowledgements** section, including details of precisely which tools, their usage, and the extent to which these tools were used. Accurately describe and compile data. AI tools must not be used for generating novel results or interpretations.*
  - (b) Carefully record and save primary and secondary data... *Do not input sensitive, confidential, or restricted information into open generative AI tools.*
  - (e) Write clearly in your own words. It is necessary to resist the temptation to "copy and paste" from the Internet or other sources for class assignments, manuscripts and thesis. *AI tools may be used freely, without acknowledgement, for word-processing tasks such as improving grammar, punctuation, clarity, etc.*
4. The following may be added to Point 1 on 'Student Roles' in Section 13.14 of the Student Handbook (pg. 59): *Any student who uses AI tools must do so within the permitted framework. Further, the student is **wholly** responsible for the correctness of the content generated by such tools.*

## 7.2 Other Documents

- The letter sent out to the thesis referees should include briefly the guidelines for the use of AI tools in theses.
- The institute website should have the guidelines for the use of AI tools in research and education, and also maintain useful resources.

## 8 What is in Store for the Future?

While much is uncertain about what the future holds, seven major corporations—OpenAI, Microsoft, Google, Meta, Amazon, Anthropic, and Inflection—have committed to **watermark their models** [54]. Watermarking has emerged as an effective solution for distinguishing human-written and LLM-generated text, amidst growing concerns of LLMs being misused for generating misinformation or completing homework assignments.

As model capabilities continue to improve, particularly for tasks that require reasoning and multimodal understanding, we speculate that these tools would find increasing use in education, including interactive tutoring, learning assistance (e.g., by providing feedback, answering questions, explaining concepts), and grading exams. We certainly hope that future AI tools would support a diverse set of natural languages, including major Indian languages. It is worth seeking to influence the development and usage of AI tools to bring in more diversity and fairness, thereby making education more accessible and inclusive. A key unknown is: What effect will easy access to these powerful AI tools have on the learning process? The answer is likely to depend in part on the educational context, which includes both the academic domain and the student's level of expertise. For example, while some research finds that programmers with limited experience are able to leverage GitHub Copilot more effectively than their more experienced peers [53], this contradicts the experience in the introductory programming course at IISc (UENG-101).

To prepare for the future and to actively shape constructive use of AI in education and research, we recommend that the Institute identify and support marquee AI-based projects. For example, MOOC courses with AI-customised delivery on topics of contemporary interest will appeal to and benefit varied audiences. To this end, the Institute can encourage collective efforts and provide necessary resources and infrastructural support.

We conclude the report on a cautionary note. The recent events at OpenAI [22] are a timely reminder of the fickle and dynamic nature of the technology. AI has ebbed and flowed in the past and will see highs and lows in the future. The tools we use today may not exist tomorrow. The applications available for free today may become unaffordable tomorrow. The use of data sets supporting the exciting AI tools we see now may be regulated in the future. We must be aware that AI technology's evolution and availability will be shaped not just by researchers, engineers, and technocrats but also by businesses, governments, and perhaps even the courts. With these issues in mind, and given the fast-paced nature of the field, it is crucial to frequently revisit the Institute's approach to research in AI and its policies on using AI tools in research and education. In fact, an annual review is warranted.

## 9 Frequently Asked Questions

1. **What is generative AI?** Generative AI is a type of Artificial Intelligence that generates new content, including text, images, and code. It refers to large language models that train on immense data sets, use machine learning algorithms to identify statistically probable patterns and structures, and draw on this training together with user input to create new content based on user prompts.

You can learn more about these AI tools here: [Google Techspert](#)

2. **What are the different AI tools available, and are they free?** There are many generative AI tools available, including those that allow generating text, for coding, images and graphic design, speech, and music. Because many of these tools were developed in academic research settings, many of them are free to use and some are even open-sourced!

An elaborate and frequently updated list can be found here: [Awesome Generative AI](#)

- 3. How are generative AI tools useful in teaching?** AI tools are useful in teaching to both teachers and students. For students, AI tools can help with coding, mathematical reasoning, improve writing and presentation of graphs and slides, explore topics being covered in the course, clarify doubts, provide practice problems, provide a personalised learning experience, and summarise and synthesize literature. Of course, if you use these tools in formal coursework, you should ensure that you abide by the usage policy for your course. For teachers, AI tools can help with tasks such as generating course plans, syllabi, and course policies, facilitating grading, and providing assistance to students outside class.

There are limitations and disadvantages too, to the use of these tools. So, for both students and teachers, it is very important to familiarise yourself with the positive and negative impacts of these tools on an individual's learning.

For further details, see the section on [Guidelines and Recommendations for Teaching](#). Further reading on an interesting perspective: [Why teachers should explore ChatGPT's potential — despite the risks](#).

- 4. How are generative AI tools useful in research?** In research, AI tools can help with various components, including generating research ideas and questions, reviewing literature, guiding data analysis, writing rough drafts of research proposals, and improving writing and presentation in research papers, thesis chapters, conference presentations, and so on.

It is important to be aware of and adhere to the privacy and scientific integrity policies of IISc, journals, conferences, or other relevant bodies when using these tools.

For further details, see the section on [Guidelines and Recommendations for Research](#). For an intriguing read, see: [Living guidelines for generative AI — why scientists must oversee its use](#).

- 5. What are the limitations of AI tools?** As you may have experienced first-hand, AI tools have several serious limitations for both research and learning. AI tools can provide factually incorrect or outdated information. The answers may vary depending on how the question is phrased. Furthermore, the output from AI tools may reflect biases in the training data and may also be out of context. There also may be privacy, plagiarism, and security concerns. Over-reliance on AI tools can reduce critical thinking, problem-solving skills, creativity, and a deep understanding of complex concepts, as well as pose problems related to academic integrity.

For more details, see the section on [Capabilities and Limitations of LLMs](#).

- 6. Why is the use of AI tools under discussion? Isn't it similar to using other online resources, like Grammarly or search engines?** Generative AI represents a huge leap in natural language interactions with computers. Based on prompts from a user, AI tools can draw from an immense training dataset to create new content, such as text, images, code, music, and more. Students can use these tools to answer homework assignments

and write essays. Researchers can use these tools to summarise and synthesise literature, carry out data analysis and so on. However, there are serious problems when relying solely on these tools. These include using incorrect, incomplete, and/or biased information. In addition, learning goals, like developing critical thinking, problem-solving skills, creativity, and depth in understanding, could be negatively affected. So, much of the discussion on AI tools revolves around using them productively while being aware of and avoiding inappropriate use.

7. **What are the ethical concerns with using AI tools in research?** Generating content without proper attribution leads to plagiarism. The use of generative AI should always be accompanied by the appropriate attribution.

Using private and confidential information in queries may make this information available widely. So, being aware of privacy and security concerns when using these tools is important. Moreover, the output from AI tools depends on the training data. Biases in the training data can propagate and have negative consequences for research outputs and recommendations.

Generative AI systems are prone to errors and fabrication. If you use such systems in your work, as the author, you will be wholly responsible for the correctness of the content generated by the system. You must review and assess all output generated by generative AI tools for accuracy before relying on them or distributing them publicly.

8. **Is the use of AI tools like ChatGPT considered cheating?** It depends. Consult IISc's policies on using AI tools in coursework and research. Some of the policies of institutes and universities across the world are summarised in the section on [Current Global Practices](#).
9. **Is it ethical to use AI tools for writing reports, assignments, etc., in courses?** There are appropriate and inappropriate ways to use AI tools. Do consult your course instructor's and IISc's policies. Do also be aware of the positive and negative impacts of AI tools on learning and educational goals. For more details, see the section on [Recommendations for Teaching](#).
10. **If I use AI tools, can it be detected by plagiarism software such as Turnitin?** As of now, standard plagiarism detection software is unreliable and can make errors when trying to detect the use of AI. But powerful tools are evolving day by day, and very soon, AI-generated content could be detected reliably. (see the section on [Detecting Plagiarism](#)).
11. **If course instructors cannot detect the use of AI tools in assignments/reports, do I need to bother about plagiarism?** Yes, because academic integrity is essential to doing science. If plagiarism is detected, the academic consequences can be serious – these are discussed in the Student Information Handbook.
12. **Does IISc have any policies on using AI in homework assignments, thesis reports, etc.?** Yes. Institutional policies are listed in the Student Information Handbook and also



in IISc's website. Instructors are free to formulate policies on the use of AI tools depending on the requirements of their specific courses and will inform students accordingly.

13. **Can faculty make their assignments AI-proof?** AI tools are evolving fast, making this task difficult. Some guidelines are to think of the learning objectives, design assessments that address these objectives, and communicate this to students. An instructor may want to focus on a student's process and steps in addressing an assignment or solving a problem rather than just the end product. The instructor could incorporate the use of AI tools into the assignment to make it more challenging.

14. **What are best practices for the use of AI, and how do we make students aware of them?** AI tools can provide a personalised learning experience, help with the exploration of a topic, provide help with problem-solving, programming, and data analysis, and improve writing. However, there are potential negative impacts, including plagiarism, lack of originality and depth in work, reduced critical thinking and understanding, privacy concerns, and incorrect and biased information.

To mitigate these negative impacts, it is essential that students and educators use AI tools responsibly and ethically. Emphasising the importance of originality, proper citation, and maintaining academic integrity is crucial. Educators and institutions should clearly communicate guidelines on the ethical use of AI tools and thoughtfully integrate them into the learning process; these can serve as supplements rather than replacements for traditional teaching and research methods. These topics can be discussed in courses, in workshops on AI tools, and in institutional documents.

For more details, see the section on [Current Global Practices](#).

15. **Can I use AI tools for writing parts of my thesis, like the introduction?** AI tools can help with improving writing and language. AI tools can also help structure the thesis better and explore the literature. However, these tools have important limitations, including providing incorrect and/or biased information. Do ensure that you are aware of these limitations when using AI tools and critically evaluate the information and responses the tools provide. It is advisable to cross-check information from reliable sources and use AI tools as a supplementary tool rather than a sole or definitive source of information or knowledge. Also, make sure to discuss the use of these tools in writing your thesis with your supervisor/committee and to adhere to your institution's policies.

16. **If I have used AI for my research, can the tool be an author?** Journals and conferences typically do not allow this because the tool cannot take responsibility for content. Instead, correct and comprehensive attribution of AI tools is advised. Make sure to be familiar with what constitutes ethical and responsible use of AI tools. Further details can be found in the section on [Policies Followed by Journals and Conferences](#).

17. **Do I need to cite AI tools when I use them in my coursework or research?** Yes, clear and comprehensive attribution is essential. How to attribute these tools is explained in the section on Coursework [Recommendations for Teaching](#) and also in the section on [Guidelines and Recommendations for Research](#).

18. **Was AI used to prepare this page or any policies of IISc?** Not this document itself, but we did test the generative capabilities of multiple large language models (LLMs) to both generate and answer some of these questions! See the Appendix on [ChatGPT's FAQs on Generative AI](#).

## Acknowledgments

We wish to thank Director, IISc, for providing us an opportunity and privilege to work on this important initiative. We place on record our thanks to the 97 faculty members who responded to the faculty survey, 525 students who responded to the survey, and to all the faculty members who provided feedback during the meeting of the faculty of science, the meeting of the faculty of engineering, and through the feedback survey. Our thanks to Prof Jayant Haritsa and Prof Aditya Gopalan for helpful suggestions.

## Appendices

### A Faculty Survey: Questionnaire

1. Generative AI tools such as ChatGPT, GitHub Copilot, etc. can generate plausible (but not always correct) responses for a variety of academic tasks.  
For which of these tasks are you familiar with the quality (good or bad) of responses produced by Generative AI?
2. In your courses, approximately what percentage of the course grade is based on unproctored work (homeworks, projects, take-home exams)?
3. For your courses, does the fact that students have access to Generative AI tools for unproctored work impact the way these courses are offered?
4. If your answer to Question 3 is Yes or Mixed impact, have you thought of how to modify your assessments (homework, reports, programming tasks, etc.)? If so, please describe briefly.
5. If your answer to Question 3 is No impact, how would Generative AI have to improve in order to change your answer?
6. For your courses, what would be your preferred policy for Generative AI tools?
7. To help stay abreast with rapid advances in this space, should IISc maintain an up-to-date resource page for faculty?  
For example: <https://teaching.pitt.edu/resources/chatgpt-resources-for-faculty>
8. If your student has good ideas but struggles to convey them in written English, should that student use Generative AI?

9. What specific guidelines should be provided on Generative AI use/misuse?
10. Please provide any other suggestions. In particular, please do point us to any interesting articles/blogs/institutional policies on this topic which you think would be useful.

## **B Student Survey: Questionnaire**

1. Please enter the Division under which your department falls.
2. Please enter your department (major in case of undergraduates)
3. What Generative AI tools have you used in your course/research work (check all that apply)?
4. Generative AI tools such as ChatGPT, GitHub Copilot, etc. can generate plausible (but not always correct) responses for a variety of academic tasks.  
For which of these tasks are you familiar with the quality (good or bad) of responses produced by Generative AI?
5. How would you rate generative AI for solving homework problems, writing logical arguments, including mathematical proofs/derivations?
6. How would you rate generative AI in writing/debugging/optimizing code?
7. How would you rate generative AI in summarising research articles?
8. How would you rate generative AI in rephrasing content to improve readability?
9. For courses in what subjects/topics do you feel generative AI will have a positive impact on your learning?
10. For courses in what subjects/topics do you feel generative AI will have a negative impact on your learning?
11. Would generative AI help in improving the language of your written submissions?
12. In what way, if any, do expect generative AI to have an impact on your research?
13. Do you feel that the institute should have clear guidelines on the use of generative AI in courses and in research?
14. What is the best way to communicate institute guidelines on generative AI to students?
15. What would you prefer as the Institute policy on generative AI for use in course work?
16. What would you prefer as the Institute policy on generative AI for use in research submissions such as papers and thesis?
17. Please provide any other suggestions. In particular, please do point us to any interesting articles/blogs/institutional policies on this topic which you think would be useful.

## C Background on Generative AI and State-of-the-Art

The release of ChatGPT has captured the interest of researchers, educators, students, industry professionals, startup leaders, and the general public. Despite the overall attention that ChatGPT and other AI tools have received, our understanding of the Large Language Models (LLMs) that underly these tools is still limited. Thanks largely to open-source efforts, the recipe for building these models is becoming standardised (see Section C.1). On the other hand, our understanding about model capabilities—and how these capabilities emerge in the first place—is an area of active research. In Section C.2, we shed light on some of the known capabilities relevant for faculty and students at IISc. Given the pressing need, we also discuss methods to distinguish model-generated content from human-written text in Section C.3.

### C.1 Overview of Large Language Models (LLMs)

#### C.1.1 Development Process of LLMs

The overall development process can be (roughly) divided into four steps.

##### A. Gathering Data and Computation

Current large language models are trained on about *one trillion words* (for comparison, all of English Wikipedia contains about one billion words). For instance, LLaMA [58] was trained using CommonCrawl and C4, which are wholesale dumps of the internet webpages (this makes up for over 80% the training data), Wikipedia (4.5% of the training data), GitHub (4.5%), books (4.5%), ArXiv (2.5%) and Stack Exchange (2%). This text data is predominantly English. These models also require massive compute: LLaMA 65 B model required over 1 million hours of A100 GPUs.

##### B. Pretraining

Pretraining is the first phase of training, which is an unsupervised learning process. That is, the data is not labeled for any specific task. Instead, within a specified “window” of text, the model blanks out (or masks) some of the text and learns to predict the blanked-out text. This task is similar to “fill-in-the-blanks” questions common in early school education. Optimising the model to fill-in these blanks at this scale equips the model with a vast amount of world knowledge and (to a certain extent) reasoning capabilities. There is much ongoing research on optimising pretraining methods. For instance, increasing the length of the “window” increases the cost of pretraining, but it also enables these models to discover patterns between elements that are related but spatially separated in the data.

##### C. Instruction Tuning

Many large language models are developed on top of the transformers architecture [59]. These models excel at predicting the most likely word that follows a given piece of text (known as the ‘prompt’). By repeatedly generating the next word and appending it to the prompt, the

model can perform a ‘prompt completion’ task. The goal of instruction tuning is to transform a language model from a mere ‘prompt completer’ to a more useful ‘instruction follower’. Instruction tuning is achieved through supervised learning i.e., by fine-tuning the (pretrained) model on a set of instructions and the desired outputs.

#### D. Reinforcement Learning from Human Feedback

To ensure that models are safe, correct, and aligned with human preferences, another layer of feedback is included in the form of thumbs up/down ratings by human raters. Such a feedback is crucial as it provides feedback to align the model’s internal knowledge rather than forcing it to generate desirable outputs. Such feedback is incorporated through reinforcement learning.

#### C.1.2 LLM Tools

While some sophisticated LLMs are available as open-source in their “raw” form, it is important to note that most users will access these models via specialised tools. The use of these tools is expected to grow in a variety of professional contexts, and it is likely that students will continue to have access to “free” versions of these AI tools [34]. Executing LLM queries is compute-intensive (and therefore expensive), but the “free” version of an tool often collects detailed usage data including the prompt, the generated response, and whether the user accepted or rejected this response. This data can be extremely valuable for refining the LLM. For this reason, users (especially researchers) need to be particularly careful about entering sensitive data into AI tools. This point is elaborated further in Section 6.

Often, these tools help users leverage the underlying models to their fullest for common tasks, but they prohibit access to the models for tasks deemed off-topic. As an illustrative example, consider GitHub Copilot, an AI tool used by an increasing number of software developers. It provides access to two types of LLMs: one for generating code (see Section C.2) and one for chatting about the generated code. The LLM underlying the Chat feature (Figure 1) is similar to ChatGPT, but it is typically accessed via a menu of prompts for common tasks (fixing code, explaining code, and writing test cases). The tool internally unpacks these prompts into a series of carefully tuned sub-prompts, in order to optimise the quality of the model’s responses. The tool also allows users to input open-ended queries. While the underlying model is extremely versatile and is capable of responding to queries such as *Write a poem*, the Chat tool restricts the usage of the LLM to “software development topics.” In contrast, the user cannot directly prompt the code-generation model. Instead, the GitHub Copilot tool sends snippets of the code along with carefully tuned prompts [23] to that model, and displays the responses via an interface that is more suitable for code development (Figure 2). *A few undergraduate and*

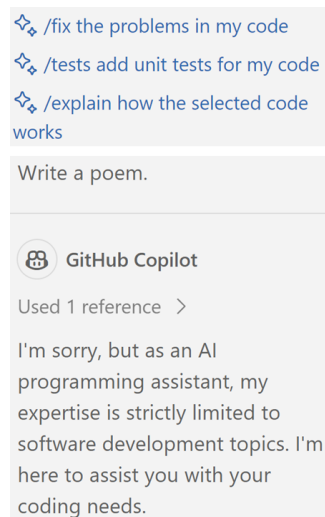


Figure 1: Key features of GitHub Copilot Chat.

graduate-level courses at IISc have adapted their curriculum to include exposure to this tool and its effective usage.

## C.2 Capabilities and Limitations of LLMs

Large language models are versatile, and can perform (to varying degrees) a range of tasks: from creative writing, solving mathematical puzzles, to assisting in programming. Here, we focus on tasks that we believe are relevant for IISc faculty and students.

### Writing Assistance

Arguably, writing assistance is one of the most sought after ChatGPT functionality. A recent study by MIT economists [50] evaluates the impact of ChatGPT use in writing quality and productivity. The study recruited 453 marketers, data analysts, and college-educated professionals to complete writing tasks typical of the writing required in their work, such as writing press releases, short reports, or analysis plans. Half of the participants could use ChatGPT to help them. A group of other professionals then evaluated their work, grading the writing on a scale of 1 to 7, with 7 being the best. The study concludes that the writers who chose to use ChatGPT took 40% less time to complete their tasks, and produced work that the raters scored 18% higher in quality. The writers who were already skilled at writing were able to reduce the amount of time they spent on their work, while those who were assessed as being weaker writers produced higher-quality work upon using ChatGPT.

Additionally, large language models can be used for “style-transfer”, e.g., rewriting text more formally; or rewriting the content for a certain audience. However, a recent study notes that **models’ responses could introduce factual errors or hallucinate** while adapting a certain style [55]. For all writing-related tasks, it is highly recommended that users verify the model responses for correctness and appropriateness.

Above and beyond writing, these tools (including text-to-image generation models) could aid faculty and students in improving their **presentations**, e.g., by generating illustrative figures for talks, for making powerpoint presentations, or by helping with code to draw TikZ figures for their papers and reports.

### Coding Assistance

There are several LLMs for generating code, including Codex (OpenAI), AlphaCode (DeepMind), and CodeLlama (Meta). These LLMs have been trained on large volumes of code and the accompanying natural-language descriptions of tasks achieved by specific code segments. Thus, when these models are given an appropriate prompt that describes the task in natural language, they are capable generating *plausible* code.

The accuracy of these code-generation models has improved significantly, and several powerful models are freely accessible for students. As an illustrative example, consider the following task: *Write a Python function to count the number of unique words in a*

```
1 def num_unique(words):  
2     return len(set(words))
```

Figure 2: A student has translated the given task into a suitable prompt in Python (Line 1), and GitHub Copilot has generated an efficient solution (Line 2).

```

1 def num_unique(words):
2     """Count the number of unique words of
3     length at least three, ignoring differences
4     between upper-case and lower-case letters."""
5     unique = set()
        for word in words:
            if len(word) >= 3:
                unique.add(word.lower())
        return len(unique)

```

Figure 3: A student has copy-pasted the more complex task specification into the prompt (Lines 2-4), and GitHub Copilot has suggested an appropriate solution (starting from Line 5).

given list of words. While this task seeks to develop the ability of students to express simple algorithms as code, a student can “get the answer” by merely knowing how to prompt an LLM-powered tool such as GitHub Copilot, as shown in Figure 2.

To thwart code generation models and ensure that students solve the task themselves, faculty may attempt to make the task more complex, for instance by increasing the number of basic algorithmic steps to be performed. Consider this variant: *Count the number of unique words of length at least three, ignoring differences between upper-case and lower-case letters.* Even if this exact problem is not in the model’s training set, it has seen each sub-problem (strings of a certain minimum length, upper vs. lower case, etc.) and several instances of combining solutions to sub-problems into an overall solution. Thus, the LLM can also generate an efficient solution to this variant (Figure 3). A recent study found that GitHub Copilot can solve about 50% of typical introductory programming tasks by merely copy-pasting the given task as a comment, and that this percentage rises to about 80% when a more detailed comment is provided that manually spells out the steps involved [39].

Code-generation models often fail when the steps involved are not obvious. For example, Figure 4 shows a clearly incorrect solution generated by GitHub Copilot for a variant of the problem that is best solved by querying an external source for US vs. UK spellings. (If details about this source are specified, the tool is likely to generate a correct solution.) Unfortunately, such problems may be poorly aligned to the learning outcomes of an introductory programming course. Developing viable alternatives to simple code-writing tasks for the era of LLMs is an active area of research [40], with some calling for a rethink of traditional learning outcomes in programming courses [56, 60, 34].

## Mathematical Reasoning

Multiple studies have tested the capabilities of LLMs across a range of aspects regarding mathematical comprehension and creativity. A study by Microsoft Research [36] found that while GPT-4’s abilities certainly go beyond memorisation or simple pattern matching (it sometimes

```

1 def num_unique(words):
2     """Count the number of unique
3     words in a list of words. Ignore
4     differences in US vs UK spelling."""
5     words = [word.lower() for word in words]
        return len(set(words))

```

Figure 4: GitHub Copilot generates an incorrect suggestion for a variant where the steps to be performed are not obvious.

demonstrates a profound understanding of undergraduate-level mathematical concepts as well as non-trivial creativity), it is still quite far from the level of an expert carrying out creative mathematical research. GPT-4’s mathematical understanding is context-dependent. In some cases, it is able to apply quantitative reasoning while encountering problems that require mathematical thinking, abstraction, and model-building. Researchers at MIT [41], NYU Abu Dhabi [43], and UT Austin [24] have found that GPT-4 generated answers have achieved similar or higher grades than the median grades in many undergraduate-level courses.

Another study by Frieder et al. [42] found that GPT-4 can be used effectively as a mathematical assistant for querying facts, acting as a mathematical search engine and database interface. It can also be used to assist in problem-solving for undergraduate-level courses. For graduate-level courses, the present level of GPT-4 can be rated as inferior to that of an excellent graduate student. It is notable, however, that GPT-4 has successfully cleared almost all (barring one or two) courses from the mathematics dataset GHOSTS1.

At IISc, GPT-4 was tested on problems from two graduate-level courses. In *Randomized Algorithms* (a graduate-level course in the Department of Computer Science and Automation), the median mark obtained by students was 39/60. ChatGPT scored 20/60, GPT-4 scored 39/60 without any hints, and 48/60 with hints. Some of these problems were actually created by the instructors and were not directly taken from any textbook or other available materials on the internet. It was observed that the answers are “stochastic,” and thus if one repeats the same question (for which GPT-4 has once returned the correct answer), one might get an incorrect answer the next time. Also, for GPT-4, it seems to be easier to prove a given mathematical equality than to solve open-ended problems.

In *Random Processes* (a graduate-level course in the Department of Electrical Communication Engineering), the instructor similarly tested GPT-4 on three problems. The questions were created by the instructor, and it was impressive that GPT-4 was able to apply the right concepts and step-by-step procedures to answer them.

In some cases, AI tools have revealed new possibilities and achieved better bounds in mathematical research in problems such as (a) matrix multiplication [25], (b) a geometric conjecture related to plane coloring [26], and (c) a central problem in extremal graph theory [49].

The transformer architecture [48] has also been used to generate symbolic, closed-form solutions to integrals and first and second-order differential equations, which outperformed other usual solvers, such as Mathematica, MATLAB, and Maple. Additional examples of GPT-4’s mathematical reasoning are provided in Appendix D.5.



## Scientific Discovery

Several recent scientific breakthroughs have leveraged machine learning for the statistical modeling of massive observational datasets. More recently, researchers at Microsoft Research are exploring the use of machine learning emulators to accelerate the simulation of physical processes: the so-called fifth paradigm of scientific discovery [27]. This research group has also found evidence that LLMs can be used for scientific discovery by “handling complex problem-solving and knowledge integration tasks” [32]. Researchers at CMU have suggested that the versatility of LLMs “across domains, including natural language, biology, chemistry, and computer programming, can foster interdisciplinary collaboration” [35]. National labs in the United States such as Argonne National Laboratory [28] have initiated a major *Autonomous Discovery* program that envisages to use AI tools, robotics, and simulation to assist in the planning, execution, and analysis of large scale scientific experiments towards new discoveries.

AI systems specialised to scientific domains have already had a significant impact and promise to have much more of an impact. The most notable such system is AlphaFold [45] which predicts protein structures as accurately as experimental methods. As a result the number of proteins with known structures has exploded from mere thousands to several millions. This has the potential to revolutionize drug discovery and other areas of biology.

A very recently developed important AI system is GraphCast [47], which predicts weather up to 10 days in advance better than any other prediction system, including giant numerical models running on supercomputers. This has been trained and the model code and weights are publicly available. Given these, weather prediction can be done in just 5 minutes with modest computing resources. This has the potential to revolutionize weather prediction and climate science.

Nuclear fusion, if it becomes practical, would be an energy source with huge benefits and few costs or dangers. A key difficulty in making such a system is the design of the variable magnetic field that confines the plasma. An AI system [37] has been developed that can design such magnetic fields that were shown, for the first time, to be able to control plasma in fusion reactors.

### C.3 Detecting Plagiarism

As AI tools gain popularity among students, there is a pressing need for instructor tools to distinguish model-generated content from human-written text. One approach to detect plagiarism is to train post-hoc classifiers that exploit the distributional differences between LLM generations and human writing. However, these classifiers often suffer from large false-positive rates and have been the cause for false accusations of plagiarism [52]. OpenAI, citing low detection accuracy, has terminated their detection service [51]. Additionally, as LLMs continue to improve and these distributional differences narrow, the effectiveness of such approaches may wane. Many available solutions on the web, including the popular tool Turnitin, are based on post-hoc detection approaches. These tools can have high false positive rates, particularly when the length of the generated content is short [29]. For this reason, the Teaching Center at the University of Pittsburgh states the following [30]:

*Based on our professional judgment, the Teaching Center has concluded that current AI detection software is not yet reliable enough to be deployed without a substantial risk of false positives and the consequential issues such accusations imply for both students and faculty.*

*At this time, we concur with this assessment.* We note that a promising alternative is to intentionally embed a ‘watermark’ signal into LLM-generated text that is imperceptible to an unsuspecting reader but can be algorithmically detected [33, 38, 57, 44]. One prominent watermarking strategy for LLMs works by pseudo-randomly partitioning the model’s vocabulary tokens into ‘green’ and ‘red’ lists at every generation step [46]. To implant a watermark, the probabilities associated with tokens in the green list are artificially boosted before decoding a token. The watermark can later be verified by assessing the proportion of green tokens in a given piece of text (as content with a large fraction of green tokens is likely to be generated). Given the pressing need to detect generated content, seven major corporations including OpenAI, Microsoft, Google, Meta, Amazon, Anthropic, and Inflection have recently committed to watermark their models [54]. Therefore, future tools for detecting AI-generated content may be more reliable than existing ones.

## D How the Committee Tested AI Tools

### D.1 Generative AI Summarizes the Faculty Survey

#### D.1.1 *ChatGPT’s summary of responses to the question on “modifying assessments considering AI tools”*

##### **Summary:** Educators’ Responses on Generative AI Tools in Academia

The responses indicate various strategies, concerns, and adaptations regarding the integration of Generative AI tools in academic settings. The overarching themes are:

- 1. Assessment Modifications:** Educators are considering or implementing adjustments in their assessments. This includes increased proctored exams, modifications in assignment structures, and the shift towards in-class tests to mitigate AI tool usage for unmonitored work.
- 2. Balancing Benefits and Challenges:** There’s recognition of the benefits and challenges of AI tools. While they can aid in writing, they might deter critical thinking or the need for understanding. Educators are attempting to strike a balance, allowing for AI assistance while maintaining the necessity for original thought.
- 3. Focus on Problem-Solving and Authenticity:** The emphasis is on problem-solving tasks that demand understanding and reasoning. Some educators are relying on in-person discussions, oral assessments, and real-world problem assignments to ensure students engage authentically with the content.
- 4. Challenge of Subject Adaptation:** Educators in fields like mathematics, program-

ming, and design express difficulties in aligning AI tools with their subjects due to the nature of problems or the need for hands-on, creative approaches.

5. **Ethical Considerations and Vigilance:** There are concerns about the ethical use of AI tools and the potential for academic dishonesty. Some are relying on vigilance, in-class exams, or oral assessments to mitigate the misuse of AI for assignments.
6. **Exploring Pedagogical Shifts:** Some educators are considering changes in teaching methodologies to prevent the ease of AI-generated solutions, such as more focus on class interaction, conceptual learning, or case studies.
7. **Continual Evaluation and Learning:** Many educators are in the process of adapting and continuously evaluating their methods to ensure fairness, authentic learning, and the prevention of misuse of AI tools.

In summary, the responses indicate a collective effort among educators to navigate the integration of AI tools into education while ensuring that learning, critical thinking, and authentic engagement remain at the forefront. There is an ongoing exploration of various assessment methods and teaching strategies to strike a balance between the benefits and challenges posed by AI tools.

Even the LaTeX code for this box was generated by ChatGPT, with the following prompts:

- “Can you summarize the following responses (cut-paste)?”
- “Can you send me a Latex version of the previous response?”
- “Can you put these responses within a box?”
- “Can you have the font be red and the box background be light yellow?”

#### D.1.2 *ChatGPT’s Summary of Responses to the Question on “What Guidelines should be provided on AI Use/Misuse”*

These responses expand on sentiments and concerns surrounding the utilization of Generative AI tools in academic settings. The collective themes are as follows:

**Awareness and Accountability:** Many emphasize the importance of educating students about the advantages and disadvantages of AI use. They recommend clear acknowledgments if AI tools are employed in academic work.

**Ethical Considerations:** There’s a significant call for ethical guidelines. Respondents stress the need for accountability and stringent policies to prevent AI tool misuse. Declarations or acknowledgments are viewed as crucial, particularly in professional or academic submissions.

**Learning and Self-Discipline:** Some opinions highlight the importance of self-discipline in using AI tools, encouraging students to utilize them for learning and improvement rather than circumventing understanding. Suggestions involve using AI for specific tasks like writing improvements rather than generating complete solutions or content.

**Specific Usage and Restrictions:** Recommendations involve employing AI tools for particular tasks, such as coding or grammar correction. However, there's a general agreement against using AI to complete assignments or tasks that hinder the learning process.

**Balancing Use and Learning:** The challenge is to find a balance where AI tools are used as aids to enhance learning and understanding rather than acting as substitutes for critical thinking and original work.

**Guidelines and Implementation:** Suggestions include clearer guidelines, declarations, and limitations on AI tool usage, especially in academic submissions or research works.

**Caution and Awareness:** Concerns are raised regarding the impact on learning and the importance of personal responsibility and effort in academic work. Encouraging acknowledgment of AI tool use is seen as a step towards transparency and responsibility.

In summary, the responses cover various opinions about integrating Generative AI tools into education, emphasizing ethical usage, responsibility, and the importance of maintaining learning and critical thinking rather than solely relying on AI-generated content.

**Note.** When prompted slightly differently, *ChatGPT* used a different kind of textbox (*tcolorbox* instead of *mdframed*) to generate these responses; this demonstrates the tool's versatility with providing different, acceptable solutions to the same (or similar) questions.

## D.2 Generative AI Summarizes the Student Survey

### D.2.1 *ChatGPT's* Summary of Responses to the Question on *Positive Impacts of AI Tools*

#### Summary of Students' Responses on the Impact of Generative AI on Learning

##### 1. Positive Impact Areas:

- **Programming and Coding:** Particularly helpful for coding-intensive courses, debugging, code understanding, and implementation. Also for coding-related tasks in various subjects.
- **Machine Learning and AI:** Seen as beneficial for courses related to AI, ML, deep learning, and tasks involving research papers, survey, and summarization.
- **Mathematics:** For conceptualization, simplification of content, and finding multiple explanations.
- **Technical Writing:** Aiding in improving grammar, syntax, academic writing, and summarizing articles.
- **Research-Oriented Courses:** Support in understanding dense papers, abstractive summarization, and helping with queries.
- **Theoretical Subjects:** Useful for understanding theoretical concepts and logical thinking.

##### 2. Areas with Mixed or Limited Impact:

- **Humanities and Literature:** Some find it helpful for writing tasks or composing emails, while others find limited utility.
- **Subjects with Extensive Available Resources:** Helpful in summarizing and consolidating available information, especially for courses with abundant material on the internet.

### 3. Areas with Scepticism or No Apparent Use:

- **Specific Mathematical Topics:** Views are divided; while some find it useful for algorithms or coding, others express scepticism regarding math-oriented courses.
- **Courses Not Centered on Programming or Theoretical Logic:** Some express doubts about its usefulness, especially in subjects that involve logical reasoning.

### 4. Varied Personal Experiences:

- Individuals report varied experiences, such as using AI for summaries, coding help, understanding new topics, improving writing, and searching for relevant information.

Overall, the responses highlight the diverse ways generative AI can aid in coding, technical writing, research-oriented tasks, and certain theoretical subjects. However, scepticism exists regarding its applicability in certain areas, indicating that its effectiveness can vary based on individual experiences and the nature of the course material.

## D.2.2 *ChatGPT's Summary of Responses to the Question on Negative Impacts of AI Tools*

### Summary of Negative Impacts of Generative AI in Learning

1. **Mathematics:** Relying on generative AI might hinder the understanding and learning of rigorous mathematical concepts, mathematical proof writing, courses heavily based on mathematics, and problem-solving in mathematics.
2. **Science and Research:** Subjects involving research, scientific thinking, scientific writing, theoretical concepts, and fields like physics, biology, or complex problem-solving in various scientific domains might suffer due to inaccuracies or limitations in AI-generated responses.
3. **Programming and Computer Science Courses:** Coding-related subjects, courses in data structures, algorithms, and computer science that demand original coding, problem-solving, or writing technical codes might be negatively impacted. Also, the concern is raised about the misuse of AI to simply copy and paste code without genuine understanding.

4. **Humanities and Writing Courses:** Subjects involving creativity, writing skills, literature, or fields that require subjective interpretation, creative thinking, and original expression might be negatively affected by the use of generative AI.
5. **Complex, Evolving, and Cutting-Edge Topics:** Fields requiring original or cutting-edge ideas, evolving sciences, or subjects with recent advancements where AI may not be equipped with the latest data or understanding might be negatively impacted.
6. **Critical Thinking and Logical Reasoning:** Courses that demand logical arguments, critical thinking, and deep analysis may suffer if students heavily rely on AI-generated responses.
7. **Education and Learning Process:** Overreliance on AI could impede the development of problem-solving skills, thinking abilities, creativity, and the joy of learning itself.
8. **Responsibility and Misuse:** Misuse or overdependence on AI-generated responses might hinder the development of students' own skills and create a dependency that restricts their growth and understanding.

In summary, while generative AI has its advantages, there are concerns about its use potentially undermining the development of critical skills, original thinking, and deep understanding across various subjects, particularly in cases where genuine learning, problem-solving, or creativity is essential. The consensus leans toward the importance of using AI as a tool while fostering one's own intellectual development and understanding.

### D.2.3 *ChatGPT's Summary of Responses to the Question on Any Other Suggestions on Generative AI*

#### Summary of Suggestions on Generative AI in Education

1. **Positive Impact on Communication and Stress Relief:** Some users highlighted the positive impacts of generative AI on correcting literature, acting as a friendly conversational tool, and alleviating stress among introverted individuals. It's noted that, for some, chatting with AI in a relaxed manner can be beneficial.
2. **Positive Use and Constructive Implementation:** It's emphasized that technology should be utilized for the betterment of life. The suggestion is to use generative AI constructively, acknowledging its continuous development and improvement, rather than imposing a ban.
3. **AI as an Aid in Learning:** Generative AI is seen as a valuable tool that helps in saving time, particularly by reducing search time and providing guidance in learning. Emphasizing that discouraging its use might not be beneficial.
4. **Necessity and Inevitability of AI:** It's noted that generative AI is perceived as the

need of the century, like oxygen for various fields. There's a suggestion that the curriculum should adapt to incorporate this need.

5. **Course Structure and Syllabus Enhancement:** Suggestions include restructuring courses to reduce the necessity for generative AI, making courses less difficult unless they are specifically intended to be advanced. Encouraging a greater number of well-defined basic courses in relevant subjects is advised.
6. **Usage for Grammar and Writing Improvement:** Generative AI is recommended for grammar and writing improvement, particularly highlighting its efficiency compared to other paid platforms. There's an argument that using such tools does not harm the academic integrity of the work.
7. **Language Barriers in Science Communication:** Recognizing that English is predominantly used in the scientific community, suggestions point out the need to effectively communicate science, advocating for the use of tools like AI for improving language proficiency.
8. **Caution in Use:** Several suggestions emphasize that while AI should be used, it should not replace the student's efforts entirely. AI should be viewed as an assistance tool and not be solely relied upon, especially for critical thinking, logical reasoning, research problems, and coursework assignments/exams.
9. **Policy References and Ethical Guidelines:** References to AI policies, guidelines, and articles from various sources, including institutions and respected organizations like UNESCO, OpenAI's Terms of Use, and Harvard Business School Publications, are provided.
10. **Concerns About Blind Trust and Misuse:** Caution is advised against blindly trusting generative AI. Suggestions urge the incorporation of awareness courses or sessions to promote proper and ethical usage of AI tools and prevent misuse.
11. **Focus on Workshops, Tutorials, and Education:** A recurring suggestion is to introduce workshops, tutorials, and educational programs that teach students and faculty how to use AI tools effectively, responsibly, and ethically.
12. **Discussion on Usage and Integration:** The need for discussions within the academic community on integrative use of generative AI in courses and research is highlighted. Suggestions include incorporating its use into the curriculum, but with clear guidelines.

This extensive summary showcases the varied viewpoints and recommendations provided concerning the use, implications, and guidelines for generative AI in educational settings.

### D.3 ChatGPT's FAQs on Generative AI

To generate FAQs, we tried an interesting experiment. We prompted ChatGPT with the following questions:

- “Can you generate a 25-point FAQ on the use of generative AI in coursework and research for academic institutions?” *At this point, it generated an FAQ of questions without the answers! A separate prompt (next) had to be provided to generate answers to the questions.*
- “Can you provide brief answers along with each question? Each answer should be between 50 and 100 words long?” *This prompt failed; it did generate answers, but all answers were short phrases or sentences . . . then, after some more prompting.*
- “Can you generate even more detailed responses to the above questions?”

The following FAQ was generated:

#### **1. What is generative AI, and how does it differ from other forms of AI?**

Generative AI is a subset of artificial intelligence that focuses on creating new content such as images, text, or music. It differs from other AI forms like discriminative models by generating original data rather than analyzing or recognizing patterns in existing data. It utilizes advanced models such as Generative Adversarial Networks (GANs) or Variational Autoencoders (VAEs) to create content.

#### **2. How can generative AI be applied to various academic disciplines and research fields?**

Generative AI finds application in numerous academic fields. In biology, it assists in image generation for cellular structures and pathology. In language studies, it helps in text synthesis and translation. Physics and chemistry benefit from simulations, while various fields use it for data analysis and visualization in research.

#### **3. What are some specific examples of generative AI applications that have made an impact in academic settings?**

Generative AI applications, such as OpenAI's GPT (Generative Pre-trained Transformer) models for text generation, StyleGAN for creating realistic images, and molecular generation for drug discovery in pharmaceutical research, have significantly influenced academic settings.

#### **4. What ethical considerations are essential when employing generative AI in coursework and research?**

Ethical considerations include ensuring authenticity and ownership of generated content, addressing biases in AI models and outputs, being transparent about the technology's limitations, and addressing ethical use and responsible sharing of AI-generated material.

#### **5. What benefits does the integration of generative AI offer to educational programs?**



Integrating generative AI encourages creativity, problem-solving, and innovation. It offers students tools for practical application in research and fosters critical thinking by encouraging exploration of new technologies and concepts.

**6. Are there recommended tools, platforms, or software for using generative AI in academic research and teaching?**

Tools such as TensorFlow, PyTorch, OpenAI's GPT models, NVIDIA's StyleGAN, and various other libraries and frameworks provide a robust foundation for implementing generative AI in academic research and teaching.

**7. In what ways can generative AI enhance the learning experience for students and aid researchers?**

Generative AI enables students to engage in hands-on learning experiences, assists researchers in visualizing and analyzing complex datasets, facilitates the creation of diverse content, and allows access to cutting-edge technology for exploring and experimenting with new ideas.

**8. What challenges or limitations might one encounter when utilizing generative AI in academic settings?**

Challenges include addressing biases in AI-generated content, managing ethical considerations, navigating the technical complexities of AI models, and the requirement for substantial computational resources, potentially limiting access for some academic institutions.

**9. How can educators and researchers ensure the responsible and ethical use of generative AI in coursework?**

By establishing clear guidelines and policies, fostering discussions about ethical considerations, providing adequate training on the technology, and implementing oversight and review mechanisms to ensure the responsible and ethical use of generative AI.

**10. Are there established guidelines or best practices for integrating generative AI into academic curricula?**

While formal standards are evolving, best practices involve emphasizing ethical considerations, offering interdisciplinary learning opportunities, providing access to diverse tools, and fostering collaborative and open learning environments.

**11. Can generative AI assist in the generation of original content for research papers or academic projects?**

Yes, generative AI can assist in creating original content. However, strict verification procedures and ethical guidelines are necessary to ensure the authenticity, validity, and ethical use of AI-generated materials in academic work.

**12. What security measures should academic institutions implement when working with generative AI?**

To ensure data security, academic institutions should enforce robust data protection policies, regulate access to sensitive AI models, implement secure storage practices, and manage generated data to safeguard against potential security breaches.

**13. How might the incorporation of generative AI impact traditional research methodologies in academia?**

Generative AI has the potential to accelerate experimentation, automate data analysis, and introduce new avenues for data-driven exploration, potentially transforming traditional research methodologies in various academic disciplines.

**14. Are there specific fields or subjects that could significantly benefit from integrating generative AI?**

Various fields such as healthcare (medical imaging analysis), computer science (language generation), art (creative content generation), linguistics (translation), and material science (simulations) stand to gain significant benefits from integrating generative AI.

**15. What skill sets or knowledge are necessary for students and researchers to effectively use generative AI in their work?**

Proficiency in programming languages, strong data analysis skills, a deep understanding of machine learning principles, and a heightened awareness of ethical considerations are crucial for effectively utilizing generative AI tools in academic work.

**16. Are there ongoing developments or research in generative AI likely to impact academia in the future?**

Ongoing research focuses on improving AI fairness, interpretability, and developing more sophisticated generative models, promising significant advancements that will likely influence the applications of generative AI in academia.

**17. How can academic institutions create a supportive environment for the use of generative AI in coursework and research?**

By providing access to resources, encouraging interdisciplinary collaborations, establishing clear guidelines, fostering ethical considerations, and promoting an inclusive and supportive environment for exploring generative AI applications.

**18. Can you provide examples or case studies demonstrating successful use of generative AI in academia?**

Projects showcasing AI-generated artwork, language translation models, and AI-driven simulations in scientific research highlight successful applications of generative AI within academic contexts.

**19. How does the use of generative AI align with current curriculum standards and academic objectives?**

Generative AI aligns with educational objectives by fostering innovation, critical

thinking, and technological literacy, contributing to various academic goals across multiple disciplines.

**20. What are the potential future implications of generative AI in shaping academic research and learning?**

The integration of generative AI could revolutionize research methodologies, foster personalized learning experiences, and encourage interdisciplinary collaborations, shaping the future landscape of academia.

**21. What are the implications of bias and fairness when using generative AI in academic contexts?**

Bias in training data and model outputs can significantly impact the integrity and fairness of research, urging the need for addressing biases in generative AI applications used in academic settings.

**22. How can generative AI assist in data analysis and interpretation for academic research?**

Generative AI aids in processing and interpreting large datasets, recognizing patterns, and providing insights, empowering researchers across various disciplines to derive meaningful conclusions from complex data.

**23. Are there any regulatory or compliance considerations related to the use of generative AI in academia?**

Regulatory concerns revolve around data privacy, intellectual property rights, and the ethical use of AI-generated content, necessitating evolving compliance measures within academic institutions.

**24. How might generative AI impact the role of educators and researchers within academic institutions?**

Generative AI could augment educators' and researchers' capabilities, demanding a focus on interpreting AI-generated outputs, adapting teaching methodologies, and overseeing ethical considerations in academic work.

**25. What are the potential collaborative opportunities for academic institutions in utilizing generative AI for interdisciplinary research and learning?**

Collaborations across disciplines could result in diverse perspectives, new breakthroughs, and innovative solutions by leveraging generative AI technologies, fostering interdisciplinary collaborations across academic fields.

#### **D.4 BardAI's FAQs on Generative AI**

Repeating the experiment described in Section D.3 with BardAI failed when asked to generate a 25- or even 15-point FAQ. The experiment worked for a 10-point FAQ, though – even then note that it generated only a 7-point FAQ. It also appears to have conflated the summary (last point) with one of the FAQ items. Interestingly, Bard did not have to be prompted once to generate the FAQ and once to generate the answers.

## 10-Point FAQ on the Use of Generative AI in Coursework and Research for Academic Institutions

### 1. What is generative AI?

Generative AI is a type of artificial intelligence that can create new text, images, music, and other content. It is trained on a massive dataset of existing content, and then learns to generate new content that is similar to the training data.

### 2. How can generative AI be used in coursework and research?

Generative AI can be used in a variety of ways in coursework and research. For example, it can be used to:

- (a) Generate new questions for students to answer,
- (b) Create practice problems for students to solve,
- (c) Generate creative text formats, such as poems, code, scripts, musical pieces, email, and letters,
- (d) Help students learn new concepts by providing them with interactive simulations,
- (e) Help researchers generate new hypotheses and ideas,
- (f) Help researchers analyze large datasets more efficiently,
- (g) Create new tools and resources for students and researchers.

### 3. What are the benefits of using generative AI in coursework and research?

Generative AI can offer a number of benefits for coursework and research, including:

- (a) It can help students learn more effectively by providing them with personalized and engaging content.
- (b) It can help researchers be more productive by automating tasks and freeing up time for more creative work.
- (c) It can help academic institutions to provide better educational experiences for their students and researchers.

### 4. What are the challenges of using generative AI in coursework and research?

There are a number of challenges that need to be addressed before generative AI can be widely adopted in coursework and research. These challenges include:

- (a) Generative AI models can be biased, which can lead to the generation of inaccurate or harmful content.
- (b) Generative AI models can be computationally expensive to train and run.

- (c) It can be difficult to assess the quality of the content generated by generative AI models.
  - (d) There is a need for guidelines and best practices for the use of generative AI in coursework and research.
5. How can academic institutions mitigate the challenges of using generative AI?  
Academic institutions can mitigate the challenges of using generative AI by:
- (a) Carefully selecting and evaluating generative AI models.
  - (b) Using generative AI models in conjunction with human oversight.
  - (c) Educating students and researchers about the potential biases and pitfalls of generative AI.
  - (d) Developing guidelines and best practices for the use of generative AI in coursework and research.
6. What are the future trends in the use of generative AI in coursework and research?  
The use of generative AI in coursework and research is likely to continue to grow in the coming years. As generative AI models become more sophisticated and less biased, they will be able to be used in a wider range of applications.
7. Here are some specific examples of how generative AI is being used in coursework and research today:
- (a) Generative AI is being used to create new educational games and simulations. For example, the University of Southern California is using generative AI to create a virtual world where students can learn about climate change.
  - (b) Generative AI is being used to develop new tools for scientific research. For example, the University of Cambridge is using generative AI to develop new ways to analyze large datasets of medical images.
  - (c) Generative AI is being used to create new forms of creative expression. For example, the University of Edinburgh is using generative AI to develop new ways to write poetry and music.
8. Overall, generative AI has the potential to revolutionise the way that coursework and research are conducted. By carefully addressing the challenges and implementing best practices, academic institutions can ensure that generative AI is used in a safe and responsible manner.

## D.5 Extended Examples of Mathematical Reasoning

The following are results of experiments with GPT-4 for harder mathematical reasoning tasks. In each case three answers were sought.

## Easy Linear Algebra

We begin with some non-trivial (but not too hard) Linear algebra question.

**Question 1.** *Let  $A$  be a square matrix. Prove that if  $A^2$  is invertible then  $A$  is invertible.*

**Question 2.** *Let  $A$  be a square matrix. Prove that if  $A^2 = I$  then  $A$  is diagonalizable.*

Both the above questions were answered correctly by GPT-4. The following question was accidentally asked to GPT-4.

**Question 3.** *Let  $A$  be a square matrix. Prove that if  $A^2$  is diagonalizable then  $A$  is diagonalizable.*

GPT-4 correctly answered that this was false.

## Ph.D. Interview Questions

As a better challenge the following questions were asked to GPT-4. These were asked in the IISc Ph.D. interview and are non-standard questions. The only students who answered these were once who were selected for the Ph.D. program.

**Question 4.** *Find all  $2 \times 2$  matrices with rational entries such that  $A^5 = I$ , where  $I$  is the  $2 \times 2$  identity matrix.*

All three proofs given by GPT-4 were correct. They were based on characterizing the minimal polynomial of  $A$ , using a non-trivial fact. So this was asked to GPT-4 as another question.

**Question 5.** *Prove that the polynomial  $x^4 + x^3 + x^2 + x + 1$  is irreducible over the rationals.*

The first of the proofs given by GPT-4 was nonsense. The second and third were correct, and were based on substituting  $y = x + 1$ , expanding and then applying Eisenstein's criterion (which cannot be directly applied).

## Analysis Questions

GPT4 was asked the following Ph.D. interview-level analysis question.

**Question 6.** *Suppose  $f : \mathbb{R} \rightarrow \mathbb{R}$  is a continuous function such that  $f''(x) > 0$  for all  $x \in \mathbb{R}$ . Prove that  $f$  is unbounded.*

GPT-4 repeatedly gave wrong proofs, invariably claiming that any increasing sequence is unbounded, which is false. When asked whether an increasing sequence was unbounded, GPT-4 correctly answered that it was not, and gave a counterexample.

To explore further, an easier (but non-trivial) question was asked.

**Question 7.** *Prove or disprove: Suppose  $f : \mathbb{R} \rightarrow \mathbb{R}$  is a continuous function such that  $f'(x) > 0$  for all  $x \in \mathbb{R}$ . Then  $f$  is unbounded.*

The responses were surreal.

- GPT-4 claimed that  $f$  was unbounded, which is false. It used the same incorrect reasoning as before, claiming increasing sequences are unbounded.
- When asked to disprove the result, GPT-4 correctly disproved it with a counterexample.
- When asked to check its (incorrect) proof line-by-line with details, GPT-4 claimed that the proof was correct.
- When told that a specific line was incorrect, GPT-4 correctly explained why it was incorrect.

### Probability Questions

GPT4 was asked the following questions related to probability, in Random Processes (a graduate-level course in the Department of Electrical Communication Engineering) and Randomized Algorithms (a graduate-level course in the Department of Computer Science and Automation). Many of these problems are not from any standard textbooks or web, but made by the instructors. However, GPT4 was often able to apply the right concepts and step-by-step procedure to answer them.

**Question 8.** *Suppose that you are an expert in probability and stochastic processes. Can you answer the following question?*

*The random variable  $M$  is uniformly distributed between 1 and 9. A student draws the line with equation  $y = Mx - M^{3/2}$  on the  $xy$  plane. What is the probability that the line cuts the  $x$ -axis after 2?*

GPT4 was able to solve this question by following a step by step systematic approach.

**Question 9.** *Take the sample space of outcomes for two successive coin tosses. Let the function  $X$  be defined as  $\alpha \cdot \text{indicator}\{HT\} + \beta \cdot \text{indicator}\{HH\} + \gamma \cdot \text{indicator}\{TT\}$ . Find real number  $\alpha$ ,  $\beta$ ,  $\gamma$  such that the smallest sigma-algebra with respect to which  $X$  is measurable is  $\{\text{nullset}, \text{sample space}, \{HH, HT\}, \{TH, TT\}\}$ .*

GPT4 was able to solve this problem using concepts from  $\sigma$ -algebra.

**Question 10.** *A fair coin is tossed independently for an infinite number of times. If  $X_1, X_2, \dots$  denote the binary sequence of tosses obtained, find the probability of the event that  $\lim_{n \rightarrow \infty} \max_{i=1}^n X_i = 1$ .*

GPT4 was able to give a correct answer, however the initial proof was not very precise. However, when requested it to be argue rigorously using the axioms of probability measure, it did answer correctly.

**Question 11.** *Let  $A$  and  $B$  be two Bernoulli random variables with means  $\mu_1$  and  $\mu_2$ . Suppose,  $\mu_1 - \mu_2 = \theta$  and we sample  $A$  and  $B$  both  $K/\theta^2$  times to obtain empirical means  $\hat{\mu}_1, \hat{\mu}_2$ . Show the following:*

$$\mathbb{P}(\hat{\mu}_1 \leq \hat{\mu}_2) \leq 2 \cdot e^{-K/2}.$$

GPT4 was able to answer this question correctly by using Hoeffding's inequality.

**Question 12.** Let  $M$  be an array containing the integers  $\{1, 2, \dots, n\}$  in random order, with all  $n!$  permutations of the integers in  $M$  to be equally likely. Call  $M[k]$ , where  $1 \leq k \leq n$ , a hero if  $M[j] < M[k]$  for all  $j < k$ . Call  $M[k]$ , where  $2 \leq k \leq n$ , a superhero if both  $M[k-1]$  and  $M[k]$  are heroes. Show that the expected number of superheroes in  $M$  is  $(1 - 1/n)$ .

GPT4 was able to answer this question correctly by appropriately defining the indicator random variables, using linearity of expectations, and then applying a telescoping sum argument.

**Question 13.** Assume we have a fair coin, i.e., flipping the coin generates head and tail with equal probability. We want a six-sided die that generates all the values in  $\{1, 2, 3, 4, 5, 6\}$  with equal probability. Present an algorithm that does so using at most 4 coin-flips in expectation.

This was a standard exercise problem in the textbook. GPT4 was able to answer this correctly.

**Question 14.** Use the Lovász Local Lemma to show that, if

$$4 \binom{k}{2} \binom{n}{k-2} 2^{1-\binom{k}{2}} \leq 1$$

then it is possible to color the edges of  $K_n$  with two colors so that it has no monochromatic  $K_k$  subgraph.

GPT4 was able to correctly apply Lovász Local Lemma to solve this problem.

**Question 15.** Show that the VC dimension of the following range space is at least 4:  $S := (X, \mathcal{R})$ , where  $X = \mathbb{R}^2$  and  $\mathcal{R}$  is the set of all arbitrary squares (may not be axis-aligned).

This was a standard problem from computational geometry. GPT4 was able to answer this as well.

**Question 16.** Consider the randomized min-cut algorithm presented in the class. We showed that for any graph  $G$  with  $n$  vertices, the probability that the algorithm finds a specific min cut  $C$  of  $G$  is at least  $2/n(n-1)$ .

- What can we say about the maximum number of distinct min cuts that a graph  $G$  can have?
- Give an example of a graph (with  $n$  vertices) with the maximum number of distinct min cuts.
- Use the randomized edge contraction algorithm to find all the global min cuts in any graph  $G$ .

GPT4 could answer the first part of the problem. However, it could not answer other parts, despite several hints.



**Question 17.** Let  $n$  equidistant points  $\{0, 1, \dots, n - 1\}$  be marked on a circle. Initially, a wolf begins at 0, and there is one sheep at each of the remaining  $n - 1$  points. The wolf takes a random walk on the circle. For each step, it moves with probability  $1/2$  to one neighboring point and with probability  $1/2$  to the other neighboring point. On the first visit to a point, the wolf eats a sheep if there is still one there. What is the probability that the sheep at position  $i$  is the last eaten sheep?

GPT4 was able to answer the correct probability, but the answer was logically incorrect.

**Question 18.** Consider the following variation of the coupon collector's problem. Each box of cereal contains one of the  $2n$  different coupons. The coupons are organized into  $n$  pairs:  $(1, 2), (3, 4), \dots, (2n - 1, 2n)$ . Once you obtain one coupon from every pair, you win a prize. Assuming that the coupon in each box is chosen independently and uniformly at random from the  $2n$  possibilities, what is the expected number of boxes you must buy before you can claim the prize?

This is again a standard problem. However, despite several hints GPT4 was not able to answer this problem.

## D.6 Guidelines from AI Tools on Using AI Tools in Research

The content in this section has been generated using the output produced by chatGPT for multiple questions.

### Use cases for AI Tools in Research

AI Tools are useful in many ways for the Research Process.

- Literature Review: Help streamline and automate the initial stages of a literature review. Provide keywords and questions to generate summaries and relevant research papers
- Idea Generation: One can brainstorm research ideas, hypotheses, and potential research questions in any given field.
- First Approximation Research Proposals: Create a rough draft of your research proposal or grant application with assistance from AI tools. The tools can help structure your thoughts and suggest relevant content.
- Writing Support: Get help with writing and proofreading research papers, thesis chapters, or conference submissions. AI tools can provide suggestions for clarity, coherence, and grammar.
- Data Analysis Guidance: Helps formulating data analysis plans, selecting appropriate statistical methods, and interpreting results
- Prepare presentations: AI tools can generate presentation slides and content for seminars, conferences, or research meetings.

- Interview and Survey Question and FAQ Development: One can generate interview or survey questions and get feedback on the quality and clarity of your questions. One can generate an excellent first version of an FAQ.
- Time Management and Productivity Improvement: One can create personalized time management plans, schedules, and strategies for staying organized and efficient.
- Language Learning and Translation: Learn a new language or assist in translating research materials in languages you are less familiar with.
- Coding and Programming Assistance: Get coding help for data analysis, simulations, or software development. AI tools can provide code snippets, explanations, and debugging assistance.
- Problem Solving: Conversations with the tool can help stimulate critical thinking, debate, and deeper understanding of your research topic. One can make initial progress on tackling challenging research problems and brainstorm innovative solutions
- Stress Management: Get advice on managing stress, maintaining work-life balance, and promoting overall well-being during your research journey.

It is important to remember that while AI tools can be a valuable resource, it should complement, not replace, traditional academic guidance and mentorship. Always critically assess the information and suggestions provided by AI tools and consult with your professors, advisors, and peers for expert guidance in your specific field of study.

### Be Aware of Limitations

One should be extremely cautious about many limitations of AI Tools. The tools may not have up-to-date information, and the responses can vary considerably in quality, so exercise caution in relying on it for critical tasks and for some specific tasks. Specifically:

- Lack of Common Sense: AI tools often struggles with common-sense reasoning, and their responses may lack context or be based on outdated or incorrect information.
- Inaccurate Information: AI tools can provide information that is factually incorrect or misleading, which can be disastrous for research, education, and decision-making.
- Sensitivity to Input Wording: The model's responses can vary based on how a question or prompt is phrased, leading to inconsistencies in answers
- Hallucinations: AI tools tend to provide answers even when they do not fully understand the question, which can lead to incorrect or irrelevant responses
- Bias and Offensive Content: AI tools can sometimes generate biased, offensive, or inappropriate responses, as it may inadvertently reflect biases present in the training data

- **Inability to Ask Clarifying Questions:** AI tools do not ask clarifying questions when faced with ambiguous or incomplete queries, which human interlocutors often do to resolve misunderstandings
- **Long-Form Coherence:** Generating coherent and contextually relevant long-form text can be challenging for the model, resulting in responses that drift off-topic or lose coherence over extended interactions.
- **Privacy Concerns:** Using AI tools may require sharing sensitive or personal information, which raises privacy and security concerns if not handled properly
- **No Independent Verification:** Most AI tools lack the ability to independently verify the accuracy of the information they provide, which means users need to fact-check the responses
- **Limited Domain Knowledge:** The knowledge is limited to what the tool was trained on up until its last update. For example, ChatGPT is trained only until April 2023. The tools may not have access to recent developments in various fields.
- **Overuse and Dependency:** Over-reliance on AI tools can lead to a reduction in critical thinking and problem-solving skills, as well as issues related to academic integrity
- **Potential for Misinformation:** In some cases, AI tools might generate information that appears authoritative but is entirely fictional, potentially contributing to the spread of misinformation.

It is important for users to be aware of these limitations when using AI tools and to critically evaluate the information and responses the tools provide. It is advisable to cross-check information from reliable sources and use AI tools as a supplementary tool rather than a sole or definitive source of information or knowledge. Developers and researchers are continually working to address these limitations and improve the capabilities and reliability of these AI tools.

### **Use AI Tools Responsibly and Ethically**

AI tools can have negative impacts on research and learning if not used responsibly. Here are some ways in which they can contribute to corner-cutting and hinder the educational process.

- **Plagiarism:** Students/researchers may misuse AI tools to generate content for assignments, research papers, or reports without proper attribution. This leads to plagiarism and undermines the development of critical thinking and research skills.
- **Ethics:** Students/researchers may not be aware of the ethical considerations of using AI tools in their academic work, which can lead to ethical dilemmas and academic misconduct.
- **Reduced Critical Thinking:** Overreliance on AI tools can lead to a lack of independent critical thinking and problem-solving skills. Students/researchers may become passive consumers of information rather than active researchers and learners.

- **Incomplete Understanding:** Using AI Tools for complex concepts or topics without fully comprehending them can result in surface-level knowledge that doesn't serve the long-term educational goals of students
- **Lack of Originality:** Relying heavily on AI-generated content can result in work that lacks originality and depth, which is a crucial aspect of academic research and learning.
- **Reduced Engagement:** Depending solely on AI tools can lead to reduced engagement in class discussions, reduced interaction with peers, and a lack of in-depth understanding of the subject matter.

To mitigate these negative impacts, it is essential for students and educators to use AI tools responsibly and ethically. Emphasizing the importance of originality, proper citation, and maintaining academic integrity is crucial. Educators and institutions should also provide guidance on the ethical use of AI tools and integrate them into the learning process thoughtfully, as supplements rather than replacements for traditional teaching and research methods. Responsible use of AI tools can help students develop skills and knowledge while avoiding shortcuts and corner-cutting.

## References

- [1] <https://ethz.ch/en/the-eth-zurich/education/educational-development/ai-in-education/chatgpt.html>. Accessed: October 31, 2023.
- [2] <https://teaching.cornell.edu/generative-artificial-intelligence/cu-committee-report-generative-artificial-intelligence-education>. Accessed: October 31, 2023.
- [3] <https://www.caltech.edu/about/news/caltech-announces-new-ai-ethics-fellowship-pr>. Accessed: October 31, 2023.
- [4] <https://www.scientificamerican.com/article/a-computer-scientist-breaks-down-gene>. Accessed: December 2, 2023.
- [5] <https://t11.mit.edu/teaching-resources/course-design/gen-ai-your-course/>. Accessed: October 31, 2023.
- [6] <https://www.nature.com/articles/d41586-023-00191-1>. Accessed: October 31, 2023.
- [7] <https://www.acm.org/publications/policies/frequently-asked-questions>. Accessed: October 31, 2023.
- [8] <https://www2.ed.gov/documents/ai-report/ai-report.pdf>. Accessed: October 31, 2023.

- [9] <https://hai.stanford.edu/news/ai-will-transform-teaching-and-learning-lets-get-i>  
Accessed: October 31, 2023.
- [10] [https://teaching.cornell.edu/sites/default/files/2023-08/Cornell-GenerativeAIForEducation-Report\\_2.pdf](https://teaching.cornell.edu/sites/default/files/2023-08/Cornell-GenerativeAIForEducation-Report_2.pdf). Accessed: October 31, 2023.
- [11] <https://www.iit.edu/sites/default/files/2023-08/Faculty%20Guide%20on%20Teaching%20and%20Generative%20AI.pdf>.
- [12] <https://www.pcmag.com/news/harvards-new-computer-science-teacher-is-a-chatbot>.  
Accessed: October 31, 2023.
- [13] <https://www.cmu.edu/news/stories/archives/2022/october/hcii-researchers-awarded-2m-grant-to-test-ai-based-mobile-tutoring-software>.  
Accessed: October 31, 2023.
- [14] <https://www.cs.cmu.edu/news/2022/norilla-spain>. Accessed: October 31, 2023.
- [15] <https://library.ethz.ch/en/news-and-courses/courses/wissenschaftliches-schreiben--chatgpt-effektiv-und-verantwortung.html>. Accessed: October 31, 2023.
- [16] <https://www.microsoft.com/en-us/translator/APPS/PRESENTATION-TRANSLATOR/>.
- [17] [https://docs.google.com/document/d/1RMVwzjc1o0Mi8Blw\\_-JUTcXv02b2WRH86vw7mi16W3U/edit](https://docs.google.com/document/d/1RMVwzjc1o0Mi8Blw_-JUTcXv02b2WRH86vw7mi16W3U/edit). Accessed: October 31, 2023.
- [18] <https://www.viceprovostundergrad.utoronto.ca/strategic-priorities/digital-learning/special-initiative-artificial-intelligence/>. Accessed: October 31, 2023.
- [19] [https://medium.com/\spacefactor\@m{}rwatkins\\_7167/updating-your-course-syllabus-for-chatgpt-965f4b57b003](https://medium.com/\spacefactor\@m{}rwatkins_7167/updating-your-course-syllabus-for-chatgpt-965f4b57b003). Accessed: October 31, 2023.
- [20] <https://onlinedegrees.sandiego.edu/classes/ethics-in-artificial-intelligence-maa>  
Accessed: October 31, 2023.
- [21] <https://iisc.ac.in/wp-content/uploads/2023/05/Student-Information-Handbook-2023.pdf>. Accessed: December 2, 2023.
- [22] <https://www.nytimes.com/2023/11/21/briefing/open-ai-sam-altman-microsoft.html>. Accessed: December 1, 2023.
- [23] <https://github.blog/2023-07-17-prompt-engineering-guide-generative-ai-llms/#the-prompt-engineering-pipeline-for-github-copilot>. Accessed: December 2, 2023.

- [24] <https://scottaaronson.blog/?p=7209>. Accessed: October 31, 2023.
- [25] <https://www.quantamagazine.org/ai-reveals-new-possibilities-in-matrix-multiplication>. Accessed: October 31, 2023.
- [26] <https://www.quantamagazine.org/mathematicians-break-bounds-in-coloring-problem-2>. Accessed: October 31, 2023.
- [27] <https://www.microsoft.com/en-us/research/video/plenary-the-fifth-paradigm-of-scientific-discovery/>. Accessed: December 2, 2023.
- [28] <https://www.anl.gov/autonomous-discovery>. Accessed: December 2, 2023.
- [29] <https://www.turnitin.com/blog/understanding-false-positives-within-our-ai-writing>. Accessed: December 2, 2023.
- [30] <https://teaching.pitt.edu/resources/encouraging-academic-integrity/>. Accessed: December 2, 2023.
- [31] Nimisha Agarwal, Viraj Kumar, Arun Raman, and Amey Karkare. A bug’s new life: Creating refute questions from filtered cs1 student code snapshots. In *Proceedings of the ACM Conference on Global Computing Education Vol 1*, CompEd 2023, page 7–14, New York, NY, USA, 2023. Association for Computing Machinery.
- [32] Microsoft Research AI4Science and Microsoft Azure Quantum. The impact of large language models on scientific discovery: a preliminary study using GPT-4, 2023.
- [33] Mikhail J Atallah, Victor Raskin, Michael Crogan, Christian Hempelmann, Florian Kerschbaum, Dina Mohamed, and Sanket Naik. Natural language watermarking: Design, analysis, and a proof-of-concept implementation. In *Information Hiding: 4th International Workshop, IH 2001 Pittsburgh, PA, USA, April 25–27, 2001 Proceedings 4*, pages 185–200. Springer, 2001.
- [34] Brett A. Becker, Paul Denny, James Finnie-Ansley, Andrew Luxton-Reilly, James Prather, and Eddie Antonio Santos. Programming is hard - or at least it used to be: Educational opportunities and challenges of ai code generation. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1*, SIGCSE 2023, page 500–506, New York, NY, USA, 2023. Association for Computing Machinery.
- [35] Daniil A. Boiko, Robert MacKnight, and Gabe Gomes. Emergent autonomous scientific research capabilities of large language models, 2023.
- [36] Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott M. Lundberg, Harsha Nori, Hamid Palangi, Marco Túlio Ribeiro, and Yi Zhang. Sparks of artificial general intelligence: Early experiments with GPT-4. *CoRR*, abs/2303.12712, 2023.

- [37] Jonas Buchli and Federico Felici. Magnetic control of tokamak plasmas through deep reinforcement learning. *Nature*, 602(7871):414–419, 2021.
- [38] Yuei-Lin Chiang, Lu-Ping Chang, Wen-Tai Hsieh, and Wen-Chih Chen. Natural language watermarking using semantic substitution for chinese text. In *Digital Watermarking: Second International Workshop, IWDW 2003, Seoul, Korea, October 20-22, 2003. Revised Papers 2*, pages 129–140. Springer, 2004.
- [39] Paul Denny, Viraj Kumar, and Nasser Giacaman. Conversing with Copilot: Exploring Prompt Engineering for solving CS1 problems using natural language. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1, SIGCSE 2023*, page 1136–1142, New York, NY, USA, 2023. Association for Computing Machinery.
- [40] Paul Denny, Juho Leinonen, James Prather, Andrew Luxton-Reilly, Thezyrie Amarouche, Brett A. Becker, and Brent N. Reeves. Prompt problems: A new programming exercise for the Generative AI era, 2023.
- [41] Iddo Drori, Sarah Zhang, Reece Shuttleworth, Leonard Tang, Albert Lu, Elizabeth Ke, Kevin Liu, Linda Chen, Sunny Tran, Newman Cheng, et al. A neural network solves, explains, and generates university math problems by program synthesis and few-shot learning at human level. *Proceedings of the National Academy of Sciences*, 119(32):e2123433119, 2022.
- [42] Simon Frieder, Luca Pinchetti, Ryan-Rhys Griffiths, Tommaso Salvatori, Thomas Lukasiewicz, Philipp Christian Petersen, Alexis Chevalier, and Julius Berner. Mathematical capabilities of chatgpt. *CoRR*, abs/2301.13867, 2023.
- [43] Hazem Ibrahim, Fengyuan Liu, Rohail Asim, Balaraju Battu, Sidahmed Benabderrahmane, Bashar Alhafni, Wifag Adnan, Tuka Alhanai, Bedoor K. AlShebli, Riyadh Baghdadadi, Jocelyn J. Bélanger, Elena Beretta, Kemal Celik, Moumena Chaqfeh, Mohammed F. Daqaq, Zaynab El Bernoussi, Daryl Fougny, Borja Garcia de Soto, Alberto Gandolfi, András György, Nizar Habash, J. Andrew Harris, Aaron Kaufman, Lefteris Kirousis, Korhan Kocak, Kangsan Lee, Seungah S. Lee, Samreen Malik, Michail Maniatakos, David Melcher, Azzam Mourad, Minsu Park, Mahmoud Rasras, Alicja Reuben, Dania Zantout, Nancy W. Gleason, Kinga Makovi, Talal Rahwan, and Yasir Zaki. Perception, performance, and detectability of conversational artificial intelligence across 32 university courses. *CoRR*, abs/2305.13934, 2023.
- [44] Zunera Jalil and Anwar M Mirza. A review of digital watermarking techniques for text documents. In *2009 International Conference on Information and Multimedia Technology*, pages 230–234. IEEE, 2009.
- [45] John Jumper, Ricard Evans, Adrian Pritzel, Jonas Ovchinnikov, Martin Lampel, Miles Senior, Kathryn Read, Julian Bateman, Google Research, and al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7871):587–595, 2021.

- [46] John Kirchenbauer, Jonas Geiping, Yuxin Wen, Jonathan Katz, Ian Miers, and Tom Goldstein. A watermark for large language models. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett, editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 17061–17084. PMLR, 23–29 Jul 2023.
- [47] Remi Lam, Alvaro Sanchez-Gonzalez, Matthew Willson, Peter Wirnsberger, Meire Fortunato, Ferran Alet, Suman Ravuri, Timo Ewalds, Zach Eaton-Rosen, Weihua Hu, Alexander Merose, Stephan Hoyer, George Holland, Oriol Vinyals, Jacklynn Stott, Alexander Pritzel, Shakir Mohamed, and Peter Battaglia. Graphcast: Learning skillful medium-range global weather forecasting. *Science*, 2023.
- [48] Guillaume Lample and François Charton. Deep learning for symbolic mathematics. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net, 2020.
- [49] Abbas Mehrabian, Ankit Anand, Hyunjik Kim, Nicolas Sonnerat, Matej Balog, Gheorghe Comanici, Tudor Berariu, Andrew Lee, Anian Ruoss, Anna Bulanova, Daniel Toyama, Sam Blackwell, Bernardino Romera-Paredes, Petar Velickovic, Laurent Orseau, Joonkyung Lee, Anurag Murty Naredla, Doina Precup, and Adam Zsolt Wagner. Finding increasingly large extremal graphs with alphazero and tabu search. *CoRR*, abs/2311.03583, 2023.
- [50] Shakked Noy and Whitney Zhang. Experimental evidence on the productivity effects of generative artificial intelligence. *Available at SSRN 4375283*, 2023.
- [51] OpenAI. New ai classifier for indicating ai-written text. 2023.
- [52] Natalie O’Neill. Texas professor flunked whole class after chatgpt wrongly claimed it wrote their papers, May 2023.
- [53] Sida Peng, Eirini Kalliamvakou, Peter Cihon, and Mert Demirer. The impact of ai on developer productivity: Evidence from github copilot, 2023.
- [54] WH Press. Biden harris administration secures voluntary commitments from leading artificial intelligence companies to manage the risks posed by ai. *White House Press Statement*, 2023.
- [55] Dongqi Pu and Vera Demberg. Chatgpt vs human-authored text: Insights into controllable text summarization and sentence style transfer. *arXiv preprint arXiv:2306.07799*, 2023.
- [56] Arun Raman and Viraj Kumar. Programming pedagogy and assessment in the era of AI/ML: A position paper. In *Proceedings of the 15th Annual ACM India Compute Conference, COMPUTE ’22*, page 29–34, New York, NY, USA, 2022. Association for Computing Machinery.



- [57] Mercan Topkara, Giuseppe Riccardi, Dilek Z. Hakkani-Tür, and Mikhail J. Atallah. Natural language watermarking: challenges in building a practical system. In *Electronic imaging*, 2006.
- [58] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*, 2023.
- [59] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- [60] Matt Welsh. The end of programming. *Commun. ACM*, 66(1):34–35, Dec 2022.
- [61] Nesra Yannier, Scott E Hudson, Kenneth R Koedinger, Kathy Hirsh-Pasek, Roberta Michnick Golinkoff, Yuko Munakata, Sabine Doebel, Daniel L Schwartz, Louis Deslauriers, Logan McCarty, et al. Active learning: "hands-on" meets "minds-on". *Science*, 374(6563):26–30, 2021.