1. [**Guidelines for Using ChatGPT and other Generative AI tools at Harvard**](https://provost.harvard.edu/guidelines-using-chatgpt-and-other-generative-ai-tools-harvard)

We write today with initial guidelines on the use and procurement of generative artificial intelligence (AI) tools, such as OpenAI’s ChatGPT and Google Bard. The University supports responsible experimentation with generative AI tools, but there are important considerations to keep in mind when using these tools, including information security and data privacy, compliance, copyright, and academic integrity.

Generative AI is a rapidly evolving technology, and the University will continue to monitor developments and incorporate feedback from the Harvard community to update our guidelines accordingly.

**Initial guidelines for use of generative AI tools:**

* + **Protect confidential data:** You should not enter data [classified as confidential](https://security.harvard.edu/data-classification-table) (Level 2 and above), including non-public research data, into publicly-available generative AI tools, in accordance with the University’s [Information Security Policy.](https://policy.security.harvard.edu/) Information shared with generative AI tools using default settings is not private and could expose proprietary or sensitive information to unauthorized parties.
  + **You are responsible for any content that you produce or publish that includes AIgenerated material:** AI-generated content can be inaccurate, misleading, or entirely fabricated (sometimes called “hallucinations”), or may contain copyrighted material. Review your AI-generated content before publication.
  + **Adhere to current policies on academic integrity:** Review your School’s student and faculty handbooks and policies. We expect that Schools will be developing and updating their policies as we better understand the implications of using generative AI tools. In the meantime, faculty should be clear with students they’re teaching and advising about their policies on permitted uses, if any, of generative AI in classes and on academic work. Students are also encouraged to ask their instructors for clarification about these policies as needed.
  + **Be alert for AI-enabled phishing:** Generative AI has made it easier for malicious actors to create sophisticated scams at a far greater scale. Continue to [follow security best practices](https://security.harvard.edu/click-wisely) and report suspicious messages to phishing@harvard.edu.
  + **Connect with HUIT before procuring generative AI tools:** The University is working to ensure that tools procured on behalf of Harvard have the appropriate privacy and security protections and provide the best use of Harvard funds.

○ If you have procured or are considering procuring generative AI tools or have questions, contact HUIT at ithelp@harvard.edu.

○ Vendor generative AI tools must be [assessed for risk by Harvard’s Information Security and Data Privacy office prior to use.](https://harvard.service-now.com/ithelp?id=sc_cat_item&sys_id=63cf445adb33e41060c0d9fcd3961900)

It is important to note that these guidelines are not new University policy; rather, they leverage existing University policies. You can find more information about generative AI, including a survey to collect data on its potential use, [on the HUIT website,](https://huit.harvard.edu/ai) which will be updated as new information becomes available.

1. [**Generative AI Policy Guidance**](https://communitystandards.stanford.edu/generative-ai-policy-guidance)

##### Honor Code Implications of Generative AI Tools

The Board on Conduct Affairs (BCA) has been asked to address the Honor Code implications of generative AI tools such as ChatGPT, Bard, DALL-E, and Stable Diffusion. These are novel tools, and both students and instructors have been experimenting with their use in academic settings. While these tools have applications that foster student learning and understanding, these tools can also be used in ways that bypass key learning objectives.

To give sufficient space for instructors to explore uses of generative AI tools in their courses, and to set clear guidelines to students about what uses are and are not consistent with the Stanford Honor Code, the BCA has set forth the following policy guidance regarding generative AI in the context of coursework:

Absent a clear statement from a course instructor, use of or consultation with generative AI shall be treated analogously to assistance from another person. In particular, using generative AI tools to substantially complete an assignment or exam (e.g. by entering exam or assignment questions) is not permitted. Students should acknowledge the use of generative AI (other than incidental use) and default to disclosing such assistance when in doubt.

Individual course instructors are free to set their own policies regulating the use of generative AI tools in their courses, including allowing or disallowing some or all uses of such tools. Course instructors should set such policies in their course syllabi and clearly communicate such policies to students. Students who are unsure of policies regarding generative AI tools are encouraged to ask their instructors for clarification.

The BCA will continue to monitor developments in these tools and their use in academic settings and may update this guidance. Members of the community are encouraged to contact the BCA to provide input, suggestions, and comments on this policy.

NOTE: As part of the BCA’s guidance on clear communication of a course’s generative AI policy, OCS recommends course instructors provide clear advance notice that they may use detection software to review work submitted for use of generative AI. Other helpful information for faculty and course assistants can be found [HERE.](https://communitystandards.stanford.edu/resources/faculty-tas/tips-faculty-teaching-assistants)

If you are in doubt about whether a generative AI source (or any source) is permitted aid in the context of a particular assignment, talk with the instructor.

**13.** [**Getting Started with AI-Enhanced Teaching**](https://mitsloanedtech.mit.edu/ai/teach/getting-started/)

Welcome to our guide to leveraging generative AI for teaching at MIT Sloan. The fundamentals of great teaching haven’t changed with the emergence of new AI tools. However, if you’re struggling to find the time to implement certain research-backed teaching strategies, these new technologies could be just what you need. Here are just a few of the many ways you can use AI in your teaching:

* Do you want to provide students with concrete examples that help illustrate abstract concepts? AI can generate examples on demand.
* Looking to create low-stakes quizzes for comprehension checks? AI can instantly generate practice questions tailored to your needs.
* Want your students to teach new concepts to an inquisitive partner? Consider asking them to converse with an AI model.

This guide will equip you with foundational knowledge, MIT policies, curated tools, ethical considerations, suggested use cases, and avenues to get support when teaching with generative AI tools. Our goal is to provide you with the knowledge and resources to smoothly incorporate these technologies into your teaching.

##### The Basics

Generative AI is an artificial intelligence subset that learns from data to produce new, unique outputs on a vast scale, ranging from educational content to software code and more. Central to this are foundational AI models trained on massive datasets. Generative AI models are essentially advanced language prediction tools.

There’s a lot of jargon involved in discussing generative AI systems. Learn more about generative AI terminology in the [**AI Glossary**.](https://mitsloanedtech.mit.edu/ai/basics/glossary/)

The following video is the first in [Wharton Interactive’s five-part course on Practical AI for Instructors and Students.](https://youtube.com/playlist?list=PL0EdWFC9ZZrUAirFa2amE4Hg05KqCWhoq&si=jAJ4EgRKRzWeIECx) In these videos, MIT Sloan alum and Wharton Associate Professor Ethan Mollick, along with Lilach Mollick, Director of Pedagogy at Wharton Interactive, provide an accessible overview of large language models and their potential for enhancing teaching and learning.

In this first video, you can learn about the following:

* Why AI is now accessible to everyone and how students are using it
* What we mean by AI, specifically large language models and generative AI
* How models like ChatGPT work and their surprising capabilities
* The potentially outsized impact of AI on educators and creative professionals
* Ethical considerations and risks related to generative AI

You can watch the other four videos in the Mollicks’ [Practical AI for Instructors and Students Course](https://youtube.com/playlist?list=PL0EdWFC9ZZrUAirFa2amE4Hg05KqCWhoq&si=jAJ4EgRKRzWeIECx) to learn more about large language models, prompting AI, using AI to enhance your teaching, and how students can use AI to support their learning.

##### Generative AI Tools

We encourage you to spend some time exploring the generative AI tools in this resource hub. It’s important to get a sense of any technology’s capabilities and limitations before you integrate it into your teaching. Also, trying these technologies yourself may help you get a sense of how your students are using generative AI.

Before you start using AI tools in your teaching, make sure to review MIT Sloan’s [**Guiding Principles for the Use of Generative AI in Courses**.](https://mysloan.mit.edu/academics/tl/directlinks/AI%20Guiding%20Principles.pdf)

The tools we’ve curated in this resource hub fall into these categories:

* [**AI Writing and Content Creation Tools:**](https://mitsloanedtech.mit.edu/ai/tools/writing/) Large language models accessed through tools like ChatGPT and Claude can help generate written content, provide grammar suggestions, summarize texts, and more. Our overview of AI writing assistants covers the types of support they can provide along with important ethical considerations. While not a substitute for human writing, these tools can help accelerate drafting and revision.
* [**AI Data Analysis and Quantitative Tools:**](https://mitsloanedtech.mit.edu/ai/tools/data-analysis/) Complex data sets are now more understandable thanks to AI analytics and visualization platforms. Explore options like IBM Watson and ToolsAI to see how algorithms can help process, interpret, and generate insights based on quantitative data. Consider use cases for statistical modeling, data visualization, and other applications while keeping key limitations in mind.
* [**AI Image Generation Tools:**](https://mitsloanedtech.mit.edu/ai/tools/images/) Models like DALL-E 3 and Stable Diffusion enable the creation of original images, videos, and other multimedia just by describing desired outputs. With experimentation, they may enable you to transform your visual media workflows.

While you explore each platform’s potential, make sure to closely monitor for quality, bias, and responsible usage.

##### Ethical Considerations

The emergence of powerful generative AI systems presents exciting possibilities for enhancing teaching and learning. However, integrating these technologies into teaching also raises important ethical questions. Three key areas of concern are data privacy, AI-generated falsehoods, and bias in AI systems.

## Data Privacy

Make sure to treat unsecured AI systems like public platforms. As a general rule, and in accordance with [MIT’s Written Information Security Policy,](https://infoprotect.mit.edu/) **you should never enter any data or input that is confidential or sensitive into publicly accessible generative AI tools**. This includes (but is not limited to) individual names, physical or email addresses, identification numbers, and specific medical, HR, financial records, as well as proprietary company details and any research or organizational data that are not publicly available. If in doubt, please consult with MIT Sloan Technology Services Office of Information Security.

Note that some of this data is also governed by [FERPA](https://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html) (Family Educational Rights and Privacy Act), the federal law in the United States that mandates the protection of students’ educational records (U.S. Department of Education), as well as various international privacy regulations including the [European GDPR](https://gdpr-info.eu/) and [Chinese PIPL.](http://en.npc.gov.cn.cdurl.cn/2021-12/29/c_694559.htm)

Microsoft Copilot provides the MIT Sloan community with **data-protected** access to AI tools

GPT-4 and DALLE-3. Chat data is not shared with Microsoft or used to train their AI models. Access Microsoft Copilot by logging in with your MIT Kerberos account at [**https://copilot.microsoft.com/**.](https://copilot.microsoft.com/) To learn more, see [**What is Microsoft Copilot (AI Chat)?**](https://mitsloan.service-now.com/sloanservice?id=kb_article_view&sysparm_article=KB0010943) Beyond never sharing sensitive data with publicly available AI systems, we recommend that you remove or change any details that can identify you or someone else in any documents or text that you upload or provide as input. If there’s something you wouldn’t want others to know or see, it’s best to keep it out of the AI system altogether (Nield, 2023). This is not just about personal details, but also proprietary information (including ideas, algorithms or code), unpublished research, or sensitive communications.

It’s also essential to recognize that once data is entered into most AI systems, it’s challenging—if not impossible—to remove it (Heikkilä, 2023). Always exercise caution and make sure any information you provide aligns with your comfort level and understanding of its potential longterm presence in the AI system, as well as with [MIT’s privacy and security requirements.](https://infoprotect.mit.edu/)

## Falsehoods and Bias

There are well-documented issues around AI systems generating content that includes falsehoods (“hallucinations”) and harmful bias (Germain, 2023; Nicoletti & Bass, 2023). Educators have a responsibility to monitor AI output, address problems promptly, and encourage critical thinking about AI’s limitations.

We encourage you to review our resources on protecting privacy, integrating AI responsibly into your course, and mitigating AI’s issues with hallucinations and bias:

* [**Navigating Data Privacy:**](https://mitsloanedtech.mit.edu/ai/policy/navigating-data-privacy/) Using generative AI tools to enhance your teaching requires a strong commitment to data privacy. This article outlines considerations for protecting your and students’ privacy when using publicly available generative AI tools for teaching and learning. These include avoiding sharing sensitive data, treating AI inputs carefully, and customizing privacy settings.
* [**Teaching Responsibly with Generative AI:**](https://mitsloanedtech.mit.edu/ai/teach/teaching-responsibly-with-generative-ai/) This guide offers strategies for harnessing AI tools to augment education while addressing AI biases and hallucinations, guiding student engagement with AI tools, and developing AI literacy.
* [**When AI Gets It Wrong: Addressing AI Hallucinations and Bias:**](https://mitsloanedtech.mit.edu/ai/basics/addressing-ai-hallucinations-and-bias/) This article provides an overview of the biases and inaccuracies currently common in generative AI outputs. It outlines strategies for identifying and mitigating the impact of problematic AI content.

By proactively addressing ethical considerations and AI’s limitations, we can realize the promise of generative AI while upholding principles of fairness, accuracy, and transparency.

### AI-Powered Teaching Strategies

Thinking about using generative AI in your teaching but not sure where to start? In this section, we’ll walk through several simple strategies for implementing research-based teaching best practices with the help of generative AI tools. These approaches are grounded in the principles of [Universal Design for Learning (UDL)](https://udlguidelines.cast.org/) and insights from the [learning sciences.](https://www.learningscientists.org/) You can use the strategies as-is or think about creative ways to adapt them to your own courses.

#### 1. Use AI to Generate Concrete Examples

Teaching often involves explaining abstract concepts or theories. While these are essential for academic understanding, they can sometimes be challenging for students to grasp without realworld context. You can use generative AI tools to come up with many concrete examples to make abstract ideas more relatable and understandable for students.

**How to Implement This Strategy:**

**Identify an abstract concept.** Select one abstract concept or theory that you’ll be covering in your lesson.

**Choose a generative AI tool.** Select one or several [AI Writing and Content Creation Tools](https://mitsloanedtech.mit.edu/ai/tools/writing/) that you’ll use for this task.

**Teach the AI.** Prompt your chosen AI tool to engage with the concept you’ve selected. If the tool is connected to the internet, you can ask it to look up and summarize the concept. If the tool is not connected to the internet, provide it with open-source content describing the concept and ask it to summarize that information.

**Prompt the AI.** Ask your chosen chatbot for examples or applications of the chosen concept. You can use a prompt like this one created by Ethan Mollick and Lilach Mollick:

“I would like you to act as an example generator for students. When confronted with new and complex concepts, adding many and varied examples helps students better understand those concepts. I would like you to ask what concept I would like examples of, and what level of students I am teaching. You will look up the concept, and then provide me with four different and varied accurate examples of the concept in action” (Mollick & Mollick, 2023-b).

**Review and select examples.** From the generated examples, select the most relevant and clear examples that align with the lesson’s objectives. Always verify the accuracy of the examples provided by the AI using trusted sources. Make sure to address and eliminate any harmful bias in AI-generated examples.

**Integrate the examples into lessons.** Incorporate these examples into your lectures, discussions, or assignments.

**What’s the research?** Concrete examples help bridge the gap between abstract theories and realworld applications. Research shows that exploring tangible instances can help students better relate to and understand complex concepts, activating their background knowledge and making learning experiences more meaningful (Smith & Weinstein, n.d.-a; CAST, n.d.-b).

#### 2. Use AI to Create Practice Quizzes

Frequent low-stakes quizzes are a great way to help students test their knowledge and reinforce their understanding. However, creating quizzes can be time-consuming for faculty. With the rise of generative AI tools like ChatGPT, though, it’s now possible to streamline the quiz creation process. You can use AI to generate practice quizzes tailored to specific topics. Moreover, these AI-generated quizzes can be adapted to fit various teaching approaches and course requirements, offering a flexible solution for assessment needs.

**How to Implement This Strategy:**

**Choose your topics.** Identify the topics or concepts for which you want to create practice quizzes.

**Select a generative AI tool.** Identify one or several [AI Writing and Content Creation Tools](https://mitsloanedtech.mit.edu/ai/tools/writing/) that you’ll use for this task.

**Teach the AI.** Prompt your chosen AI tool to engage with the concept you’ve selected. If the tool is connected to the internet, you can ask it to look up and summarize the concept. If the tool is not connected to the internet, provide it with open-source content describing the concept and ask it to summarize that information.

**Prompt the AI.** Ask the AI tool to generate quiz questions related to these topics. Use a quiz-question generating prompt like this one created by Ethan Mollick and Lilach Mollick: “You are a quiz creator of highly diagnostic quizzes. You will look up how to do good low-stakes tests and diagnostics. You will then ask me two questions. (1) First, what, specifically, should the quiz test. (2) Second, for which audience is the quiz. Once you have my answers you will look up the topic and construct several multiple choice questions

to quiz the audience on that topic. The questions should be highly relevant and go beyond just facts. Multiple choice questions should include plausible, competitive alternate responses and should not include an ‘all of the above’ option. At the end of the quiz, you will provide an answer key and explain the right answer” (Mollick & Mollick, 2023-b).

**Review and refine the results.** Examine the generated questions for relevance and accuracy. Remove any content that perpetuates harmful biases. Modify or refine as necessary.

**Distribute the quizzes to students.** Share the practice quizzes with students. You may want to incorporate the questions into a [Canvas quiz.](https://community.canvaslms.com/t5/Instructor-Guide/How-do-I-create-a-quiz-with-individual-questions/ta-p/1248)

**What’s the research?** Retrieval practice, or the act of recalling information from memory, strengthens memory retention (Smith & Weinstein, n.d.-d). Practice quizzes offer students an opportunity to test their understanding and reinforce their learning, making the information more retrievable in the future.

#### 3. Assign Students to Generate Visual Summaries

Visual aids have always been a cornerstone in effective teaching, aiding in comprehension and retention. With the rise of image-generating AI models, we now have new tools in hand to help create these visual aids. In this use case, you’ll ask students to craft visual summaries of specific topics, blending both verbal descriptions and AI-generated imagery. This not only deepens their understanding but also fosters creativity and critical thinking as they evaluate and refine the visuals produced by AI tools.

**How to Implement This Strategy:**

**Assign topics.** Provide students with specific topics for which they should create visual summaries.

**Guide students to explore** [**AI Image Generation Tools.**](https://mitsloanedtech.mit.edu/ai/tools/images/) Follow the tips in our article [Teaching Responsibly with AI](https://mitsloanedtech.mit.edu/ai/teach/teaching-responsibly-with-generative-ai/) to set your students up for success with their chosen AI tool. Make sure they are aware of generative AI’s [limitations](https://mitsloanedtech.mit.edu/ai/basics/addressing-ai-hallucinations-and-bias/) and [privacy implications.](https://mitsloanedtech.mit.edu/ai/policy/navigating-data-privacy/)

**Have students create visual summaries.** Ask students to find or generate images that they can use to create visual aids for the assigned topics. Encourage students to combine text and visual information to summarize the topic’s main points.

**Review and discuss students’ work.** Examine the visual summaries in class, discussing the concepts and clarifying any misconceptions. If this assignment is graded, make sure to grade based on conceptual understanding rather than image quality.

**What’s the research?** Dual coding is when learners interact with content through both verbal and visual information, enhancing memory and understanding (Smith & Weinstein, n.d.-b). This research-backed study strategy aligns with the Universal Design for Learning checkpoint

“Illustrate through multiple media” (CAST, n.d.-a). Visual summaries allow students to integrate two forms of information, deepening their comprehension and making the learning experience more engaging.

#### 4. Ask Students to Teach the AI

Deep understanding often comes from the act of explaining. In the realm of education, having students articulate their understanding of a concept can solidify their grasp and highlight areas needing further clarification. With the advent of AI tools like ChatGPT, students now have an interactive platform where they can practice this act of elaboration. By engaging in detailed conversations with the AI, students can receive instant feedback, refine their understanding, and practice the art of explanation.

To see what this strategy can look like in action, check out our blog post: [**Harnessing AI in**](https://mitsloanedtech.mit.edu/2023/09/01/harnessing-ai-in-finance-eric-sos-innovative-take-on-teaching-value-investing/)

[**Finance: Eric So’s Innovative Take on Teaching Value Investing**.](https://mitsloanedtech.mit.edu/2023/09/01/harnessing-ai-in-finance-eric-sos-innovative-take-on-teaching-value-investing/) **How to Implement This Strategy:**

**Choose a generative AI tool.** Select a free [AI Writing and Content Creation Tool](https://mitsloanedtech.mit.edu/ai/tools/writing/) that students can use for this activity.

**Introduce your chosen platform to students.** Follow the tips in our article [Teaching Responsibly with AI](https://mitsloanedtech.mit.edu/ai/teach/teaching-responsibly-with-generative-ai/) to set your students up for success with their chosen AI tool. Make sure they are aware of generative AI’s [limitations](https://mitsloanedtech.mit.edu/ai/basics/addressing-ai-hallucinations-and-bias/) and [privacy implications.](https://mitsloanedtech.mit.edu/ai/policy/navigating-data-privacy/)

**Assign topics.** Provide students with specific topics or concepts they should explain to the AI.

**Invite students to interact with the AI.** Encourage students to have detailed conversations with the AI, explaining concepts and receiving feedback.

**Reflect and discuss.** Ask students to reflect on their conversation with the AI and identify areas for improvement.

**What’s the research?** Elaborative interrogation is a research-backed study strategy in which students deepen their understanding by asking questions and explaining concepts (Smith & Weinstein, n.d.-c). By interacting with AI, students can practice this strategy, enhancing their comprehension and reinforcing their learning.

### Get Support

As you consider how to best use generative AI in your course, questions will arise. [Contact us](https://mitsloan.service-now.com/sloanservice?id=sc_cat_item&sys_id=00c98afadbf03510fe942d8913961920) today for a personalized consultation. We’re here to be your thought partner during your development and implementation process.

### Conclusion

Integrating artificial intelligence into your teaching offers both opportunities and challenges. In this guide, we’ve provided an initial roadmap to begin exploring this new space. We’ve covered the basics of what generative AI is, considered its potential benefits, and explored practical use cases to incorporate generative AI tools into teaching. We’ve also emphasized the importance of ethical considerations like prioritizing student privacy and addressing potential biases.

While AI offers powerful tools to augment our teaching methods, the human touch remains irreplaceable. The goal is not to replace educators but to empower them with additional resources. By combining the strengths of AI with the expertise of skilled instructors, we can create richer, more effective learning experiences for our students.

As you move forward, remember that you’re not alone on this journey. Our team is here to support you, answer questions, and provide guidance. We’re excited to see how you’ll harness the potential of AI in your classrooms and look forward to hearing about your experiences. Let’s explore, learn, and innovate together.

**14.** [**Generative AI Guidance**](https://mcgraw.princeton.edu/faculty/generative-ai-guidance)

We encourage faculty to try out or experiment with generative AI (GAI) tools, which can be used to generate ideas, summarize articles, develop computer code, create images, and compose music.

We also encourage faculty to have explicit conversations with students about the permissibility of their use in your courses and their independent work. If you allow students to use GAI tools, we suggest that you set clear and explicit guidelines and help your students understand the risks associated with these programs. In large language models, such as ChatGPT, risks include inaccuracies, fabrications (“hallucinations”), and amplified biases (see more below). We also take seriously the risk that use of these tools may short-circuit learning.

**Syllabus Language**

Faculty have the discretion to set their own GAI policies (see Jill Dolan’s [August 2023 memo)](https://odoc.princeton.edu/about/official-deans-communications/2023/generative-ai-teaching-and-learning-fall-2023-faculty-and#:~:text=Faculty%20members%20have%20the%20discretion,website%20for%20more%20pedagogical%20guidance).). This means that students likely will encounter different rules about AI use in different courses. For this reason, we strongly encourage faculty to articulate a clear policy on AI use in your syllabus.

Make clear:

* Whether AI tools are prohibited or permitted in your courses
* Whether AI tools are permitted for certain tasks
* Whether AI tools are permitted for assignments or for certain stages of an assignment
* Which AI tools are permitted in your courses

○ AI tools include Chat GPT, Bard, Grammarly, Github Copilot, Google Translate, Adobe Firefly, etc.

* The the use of Generative AI should be acknowledged; see citation guidance from [MLA,](https://style.mla.org/citing-generative-ai/) [APA,](https://apastyle.apa.org/blog/how-to-cite-chatgpt) and [Chicago](https://www.chicagomanualofstyle.org/qanda/data/faq/topics/Documentation/faq0422.html)

Keep in mind that a course policy that allows some use of GAI may introduce complexity and open up the possibility of students using GAI in ways you don’t intend.

Below you will find Princeton-specific examples of syllabus language, grouped by category. You may also find this [Chronicle article](https://www.chronicle.com/article/should-you-add-an-ai-policy-to-your-syllabus?sra=true&cid=gen_sign_in) useful as you develop a policy statement. We also highly recommend [the guidance](https://cndls.georgetown.edu/ai/policies/#syllabus-policies) and examples our colleagues at Georgetown have curated.

### Ethical and Other Risks

We recommend that faculty consider and discuss with students the significant ethical considerations and risks of using generative AI. The most important concerns are:

**Equity and Access**

Students’ varying levels of AI literacy coupled with unequal access to technology and lack of exposure to AI tools exacerbates existing [digital divides in education.](https://www.brookings.edu/articles/ai-and-the-next-digital-divide-in-education/) Safer, more accurate AI tools are often locked behind paywalls, giving rise to concerns regarding [affordability and equitable access.](https://www.insidehighered.com/news/tech-innovation/artificial-intelligence/2023/06/05/how-ai-tools-both-help-and-hinder-equity) Even though today’s student population is often well-versed in digital technology, disparities in [digital literacy education and skill acquisition](https://login.ezproxy.princeton.edu/login?url=https://www.proquest.com/scholarly-journals/ict-use-digital-skills-students-academic/docview/2642491639/se-2?accountid=13314) can affect students’ performance in college.

**Student data and privacy**

When students create an account in a program, they share personally identifiable information like their email address and phone number. Large language models such as ChatGPT or Bard can [store conversations](https://blog.allenai.org/using-large-language-models-with-care-eeb17b0aed27) and uploaded content, which they might repurpose as training data. Princeton's Information Security Office has written the following position paper on the [Prohibition of University Data in Artificial Intelligence (AI) Solutions.](https://informationsecurity.princeton.edu/posts/2023/prohibition-university-data-artificial-intelligence-ai-solutions-new)

If you elect to use AI tools that require students to create accounts, we suggest that you highlight these risks and review the data usage policies with your students. Consider, in fact, [making this a classroom exercise.](https://autumm.edtech.fm/2023/01/18/prior-to-or-instead-of-using-chatgpt-with-your-students/) You might also offer alternative options for students who are not comfortable creating their own accounts.

**Inaccuracies and fabrication**

Generative AI fabricates data, invents facts, and produces persuasive but completely inaccurate arguments, [according to researchers at Stanford.](https://hai.stanford.edu/news/generative-search-engines-beware-facade-trustworthiness) When used as a research aid, these programs can concoct citations. ChatGPT, for instance, incorrectly stated that Princeton’s Hal Foster had written an article called “The Case Against Art History” in October. The citation included volume number, year, and page references[—all a fabrication.](https://artandarchaeology.princeton.edu/whats/news/news-chatgpt-goes-rogue-fabricating-citations-hal-foster-and-carolyn-yerkes) Making students aware of this tendency toward inaccuracy might help to deter them from relying on these tools.

**Cognitive Offloading**

[Cognitive offloading](https://doi.org/10.1016/j.tics.2016.07.002) involves delegating the mental demands of a task to a technology or tool, such as relying on a calculator or smartphone reminders instead of one’s own knowledge and abilities. People may offload a task when they think the technology is more capable, they have a high degree of trust in the tools, and the tools are easily accessible. Offloading may improve a student’s short-term performance (i.e., getting good grades on an assignment) but [diminish their long-term learning and cognition.](https://www.frontiersin.org/articles/10.3389/frai.2022.908261/full) We suggest that faculty encourage students to use AI to enhance their learning, not as a replacement for their own cognition.

**Bias and stereotypes**

Generative AI is fed and trained on data that can be biased and inaccurate, or geographically and racially skewed. It has a tendency to reproduce stereotypes. If prompted to depict a “Native American,” for instance, image-making software like DALL-E 2 and Stable Diffusion tend to produce images of people with traditional headdresses. Or, if asked to illustrate a profession using an adjective like “emotional” or “sensitive,” the program is more likely to produce an image of a woman as this article by [the MIT Technology Review demonstrates.](https://www.technologyreview.com/2023/03/22/1070167/these-news-tool-let-you-see-for-yourself-how-biased-ai-image-models-are/)

**Labor concerns with how AI tools are trained**

Companies like OpenAI have relied on labor from the Global South to train their models, requiring workers to read and categorize graphic texts to identify hate speech, violence, and sexual abuse.

[This source](https://time.com/6247678/openai-chatgpt-kenya-workers/) offers a fuller account.

**Environmental Impact**

The [computational requirements](https://www.techtarget.com/searchenterpriseai/feature/Energy-consumption-of-AI-poses-environmental-problems) associated with large language models like ChatGPT contribute to high rates of energy consumption, carbon emissions, and electronic waste. [Researchers at the University of Massachusetts](https://arxiv.org/abs/1906.02243) found that training large AI models can produce nearly five times the lifetime emissions of an average car (including fuel). As AI datasets and models grow in complexity, so do their environmental impact.

### Implications for Assignments

Generative AI requires us to be very intentional about assignment design -- to maximize students’ opportunity to engage critically with course material and minimize their risk of overusing GAI.

Regardless of whether you permit the use of GAI tools, we encourage you to:

* [Define your course learning goals,](https://mcgraw.princeton.edu/faculty/teaching-princeton/developing-learning-goals) and share them with students.

○ Explain to students what they will learn by completing your assignments. In what ways will it help them develop the skills or master the content of your discipline?

* [Include a generative AI policy on your syllabus.](https://mcgraw.princeton.edu/faculty/generative-ai-guidance#tab-0-0) 
  1. You might also work with students to set a class policy, as Associate Professor

Molly Crockett does in their Psychology courses. See McGraw’s [Faculty Resource Library](https://mcgraw.princeton.edu/faculty/teaching-princeton/faculty-resource-library) for guidance on [Creating Community Agreements.](https://mcgraw.princeton.edu/faculty/teaching-princeton/faculty-resource-library/creating-community-agreements)

* Test your prompts.
  1. To understand more about the strengths and limitations of GAI tools, experiment with your assignment prompts and evaluate the results. For guidance on how to effectively prompt Chat GPT, see [Open AI’s resource on Prompt Engineering.](https://platform.openai.com/docs/guides/prompt-engineering/strategy-write-clear-instructions)
* Scaffold assignments.
  1. Scaffold students’ work with draft and revision deadlines that offer you opportunities to give feedback.
* [Incorporate reflection into assignments.](https://mcgraw.princeton.edu/faculty/teaching-princeton/cultivating-active-learning#tab-0-4)
  1. Ask students to demonstrate their thought processes and reflect on their work. For example, they might annotate their solution to a problem, write an artist’s statement to accompany a submission, or write a cover letter for an essay.
* Assign “creative critical” assignments.
  1. Design assignments that ask students to engage creatively as well as critically with course material. This might take multiple forms, including [digital assignments](https://mcgraw.princeton.edu/faculty/digital-assignments) like digital exhibitions, podcasts, or story maps. Even without the use of digital technology, consider assignments that ask students to riff on, mix up, or playfully and purposefully engage course material. We have many ideas to share with you; feel free to consult with us.
* Try oral assignments, especially if you do not permit the use of GAI.
  1. Devise oral assignments such as presentations, simulations, or role plays. These can be low-stakes activities—for example, asking a student to talk through their response to a problem or share ideas as part of a “fishbowl” discussion—or higherstakes activities that require advanced planning and preparation.
* Make an appointment for a consultation with us.
  1. We’re very happy to help you think through how generative AI may affect your teaching. We offer consultations in person and over Zoom; be in touch with us at mcgraw@princeton.edu.

**Assigning Generative AI**

If navigating AI is a skill you think is important for students to develop, you might design activities and assignments that embrace it. If you do ask students to use GAI tools, be mindful of the [ethical concerns and other risks](https://mcgraw.princeton.edu/faculty/generative-ai-guidance#tab-0-1) that they present. Remember that some students may have access to subscription-based tools like Chat GPTPlus, while others will only have access to the less powerful free versions. Be prepared to offer alternative assignments or other workarounds to students who don’t feel comfortable using these tools–which often require students to create an account– themselves.

* Ask students to analyze its output. For example, after they complete their own drafts of an assigned essay, you might ask students to request a draft of the assignment from a generative AI tool and analyze and/or critique the work it produces. Jacob Shapiro, Professor of Politics and International Affairs, [requires](https://www.dailyprincetonian.com/article/2023/11/princeton-features-professors-creative-approaches-to-chatgpt-2023) students to prompt Chat GPT and then share the responses with classmates to revise them. Associate Professor Alexander Glaser from Mechanical and Aerospace Engineering asks students to compare their answers to those composed by ChatGPT and reflect on the differences between responses from a human and those from a machine. Steven Strauss, Visiting Professor in SPIA, asks graduate students to grade ChatGPT’s response to a prompt and then reverses the process, asking students to submit their draft answers to ChatGPT (with the appropriate context) so it can give them feedback.
* Allow students to use the tool for one part of a larger assignment. For example, Heather Thieringer, University Lecturer in Molecular Biology, allows students to use ChatGPT to create a potential introduction to their lab report, which they include as an appendix. The students critique and correct the response as part of the assignment.
* Emphasize the skill of prompt engineering. Assign students to use the tool and to turn in the prompts they use to get their responses. Ask them to write a short paper reflecting on how altering their prompt changed the output.
* Use the tools to enhance students' creativity. In his *Storytelling* course, Professor of Slavic Languages and Literatures Yuri Leving asks students to illustrate their writing projects with images produced by an AI generator. He also asks students to write stories inspired by images he has generated using the tool, giving them experience both creating images from text and generating text from images.
* Ask students to analyze the benefits and drawbacks of generative AI for certain tasks in class discussions, debates, or written assignments. For example, Steven Strauss, Visiting Professor in SPIA, devotes class time to what he calls “[GAI housekeeping”](https://steven-strauss.medium.com/embracing-generative-ai-gai-in-education-some-personal-reflections-2372e6036efb) before requiring students to use the tools, addressing topics such as student accountability, algorithmic bias, the potential for hallucinations, ethical dilemmas, and replicability concerns. Once students understand the challenges and limitations inherent in GAI technology, Strauss asks them to make and support an argument about how ChatGPT and similar tools might be used to improve productivity on an everyday task.

Faculty have expressed interest in hearing about how colleagues are using AI tools in their courses. If you are assigning GAI in your course and would be willing to share your assignment, please reach out to us at mcgraw@princeton.edu.

### Detection Software and Academic Integrity

Though companies like Turnitin, ZeroGPT, and OpenAI have all developed AI detection capabilities, we do not recommend you use such software to attempt to determine if student work is AI-generated. Our recommendation against using these tools is based both on Princeton’s standards for academic integrity and the practical limits of these tools. Detection tools seem [unreliable](https://www.usatoday.com/story/news/education/2023/04/12/how-ai-detection-tool-spawned-false-cheating-case-uc-davis/11600777002/) at best and biased at worst. The creators of these tools have [warned against](https://arxiv.org/abs/2304.02819?ref=aiweirdness.com) using them to make decisions about academic honesty. [Research](https://arstechnica.com/information-technology/2023/07/why-ai-detectors-think-the-us-constitution-was-written-by-ai/#:~:text=If%20you%20feed%20America's%20most,can't%20be%20the%20case.) has also demonstrated that the software consistently misclassifies writing samples by non-native English writing as AI-generated. Instead, we encourage you to emphasize your learning goals, [consider our guidance on assignment design,](https://mcgraw.princeton.edu/faculty/generative-ai-guidance#tab-0-2) and [include a clearly stated GAI policy on your syllabus.](https://mcgraw.princeton.edu/faculty/generative-ai-guidance#tab-0-0)

If you suspect a student has used an unauthorized GAI tool in an assignment, please contact Joyce Chen at jgchen@princeton.edu or 609-258-3054. If you suspect a student has used an unauthorized GAI tool in an exam, please contact the Honor Committee at honor@princeton.edu.

**Teaching guidance and case reports on teaching with GAI**

1. [Assigning AI: Seven Approaches for Students, with Prompts](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4475995)**(Social Science Research Network)**
2. [ChatGPT and the Rise of Generative AI: Threat to Academic Integrity?](https://doi.org/10.1016/j.jrt.2023.100060)**(Science Direct)**
3. [Artificial Intelligence and the Future of Teaching and Learning](https://www2.ed.gov/documents/ai-report/ai-report.pdf)**(U.S. Department of Education)**
4. [Artificial Intelligence and Education: A Reading List](https://daily.jstor.org/artificial-intelligence-and-education-a-reading-list/)**(JSTOR Daily)**
5. [Reactions: Princeton faculty discuss ChatGPT in the classroom](https://www.dailyprincetonian.com/article/2023/02/reactions-chatgpt-ai-princeton-professors-faculty-writing-assignments-skills-tools-thinking-analysis-education)**(The Princetonian)**
6. [The AI Pedagogy Project](https://aipedagogy.org/)**(Harvard University)**
7. [Learn With AI Toolkit](https://umaine.edu/learnwithai/)**(University of Maine)**
8. [Teaching in the Age of AI](https://cft.vanderbilt.edu/guides-sub-pages/teaching-in-the-age-of-ai/)**(Vanderbilt University)**
9. [Artificial Intelligence Teaching Guide](https://teachingcommons.stanford.edu/teaching-guides/artificial-intelligence-teaching-guide)**(Stanford University)**
10. [Intentional Pedagogy with AI Technology](https://www.brown.edu/sheridan/teaching-learning-resources/teaching-resources/intentional-pedagogy-ai-technology)**(Brown University)**
11. [Why I’m Encouraging my Students to use Generative AI (e.g., ChatGPT ) When Writing Their Assignments](https://steven-strauss.medium.com/why-im-encouraging-my-students-to-use-generative-ai-e-g-eaed72e0cc3d)**(Medium)**
12. [Embracing Generative AI (GAI) in Education: Some Personal Reflections](https://steven-strauss.medium.com/embracing-generative-ai-gai-in-education-some-personal-reflections-2372e6036efb)**(Medium)**
13. [Teaching CS50 with AI](https://cs.harvard.edu/malan/publications/V1fp0567-liu.pdf)**(SIGCSE)**
14. [Developing AI Standards of Conduct as a Class](https://exploringaipedagogy.hcommons.org/2023/11/14/developing-ai-standards-of-conduct-as-a-class/)**(Exploring AI Pedagogy)**
15. [What I Learned From an Experiment to Apply Generative AI to My Data Course](https://www.edsurge.com/news/2023-11-03-what-i-learned-from-an-experiment-to-apply-generative-ai-to-my-data-course)**(EdSurge)**

**Regularly updated sources on GAI: channels; podcasts; listservs; substacks**

1. [Arvind Narayanan](https://substack.com/@arvindnarayanan)
2. [Dr Philippa Hardman](https://substack.com/@learningfuturesdigest)
3. [One Useful Thing](https://www.oneusefulthing.org/)
4. [Exploring AI Pedagogy](https://exploringaipedagogy.hcommons.org/)
5. [AI+EDU=Simplified](https://aiedusimplified.substack.com/)
6. [AutomatED: Teaching Better With Tech](https://automatedteach.com/)
7. [AI in Education](https://groups.google.com/g/ai-in-education/about)

**Links to PU faculty research and initiatives on GAI**

1. [Princeton Language and Intelligence](https://pli.princeton.edu/)
2. [Future Values Initiative](https://uchv.princeton.edu/academic-programs/future-values-initiative)
3. [Princeton Center for Information Technology Policy](https://citp.princeton.edu/)
4. [Princeton Precision Health](https://pph.princeton.edu/)
5. [Princeton Dialogues on AI and Ethics](https://aiethics.princeton.edu/)

**University Report and Resources**

1. [Generative AI Working Group Report](https://princetonu.sharepoint.com/sites/mcgraw/gen-ai/Shared%20Documents/GenAI-working-group-report-9-15-23.pdf?csf=1&web=1.&cid=0e26eae0-4c3e-4f5d-bbc1-eeb891553f18) (Office of the Dean of the College) 2. [Generative AI](https://libguides.princeton.edu/generativeAI) (Princeton University Library)

**15.** [**Guidance for Syllabus Statements on the Use of AI Tools**](https://teaching.uchicago.edu/sites/default/files/2023-09/CCTL_AI%20Syllabus%20Statements.pdf)

Guidance for Syllabus Statements on the Use of AI Tools As Autumn Quarter approaches, the University teaching community is thinking through how to address the use of AI tools, such as ChatGPT, in their courses. Not unlike the disruptions posed by the COVID-19 pandemic, one of the essential strategies for responding to this development is clear, transparent communication with students about our expectations. With the rapid onset of availability and sophistication of AI tools over the past year, students have only had a short amount of time to discover the utility and limits of these tools for their learning as well as expectations of their use in their courses of study. Moreover, the expectations around AI tools will likely vary from one course to the next, making clear and consistent communication of expectations for your course particularly important. This document provides considerations that might shape how instructors craft a syllabus statement about AI tools, followed by example statements covering a range of approaches. The staff in the Chicago Center for Teaching and Learning (CCTL) are always available for one-on-one consultations on using AI tools in your courses, communicating to students about their use of these tools, and writing syllabus statements. To schedule a consultation please email: teaching@uchicago.edu.

### Considerations and Examples of Syllabus Statements on the Use of AI Tools

Below are some examples of syllabus statements prepared by the CCTL to help guide your writing of your own statement on the use of AI tools in your courses. This is not an exhaustive list and the statements are generic in nature, with the idea that these examples will provide some ideas and starting points for thinking about how you might want to communicate your own approach on the role of AI tools in your courses. Please feel free to adapt and modify these statements to your particular context and needs. As you draft your own syllabus statement, here are a few considerations to keep in mind: • Communicate clearly and specifically when AI tools are and are not allowed, and what uses constitute a violation of academic integrity. • When AI tools are permitted, communicate when and how they should be correctly attributed. o APA guide for citing ChatGPT o Chicago Manual of Style citation guidelines for AI tools o MLA guide for citing AI tools • As with any course policy, providing reasoning that connects it to supporting the learning process helps students to understand the pedagogical rationale for the policy. Communicate to students how the use of AI tools do or do not support the learning goals for the course. In many courses, the development of foundational skills and knowledge that are integral to a student’s academic and personal growth will need to be developed without the use of AI tools. In other cases, the development of skills and knowledge around the use of AI tools in particular fields may be an important learning goal. • Consider linking to or referencing the student policy on Academic Honesty and Plagiarism. Every new student at the University is provided with the University’s Student Manual of University Policies & Regulations. Under the policy, AI tools that are not explicitly allowed by the instructor(s) of a course, will be considered a violation of academic integrity. While it is the responsibility of each student to make themselves familiar with these policies, it can be helpful to remind them of the specific policy on Academic Honesty and

Plagiarism. • Consider how word choice and the tone of your statement will communicate and foster trust in your students as they navigate the availability and use of these new technologies. In general, most instructors will allow or limit the use of AI tools in ways that fall under four broad categories: (1) use prohibited in all situations; (2) use with prior permission; (3) use with proper citation; and (4) free use with no citation required. Below are examples of general statements in each of these categories. These examples are best used as starting points to adapt to your teaching, assignments, course design, and style of communicating with your students.

1. **Use prohibited in all situations:**

“In this course, we will be developing skills and knowledge that are important to discover and practice on your own. Because use of AI tools inhibits development of these skills and knowledge, students are not allowed to use any AI tools, such as ChatGPT or DallE 2, in this course. Students are expected to present work that is their own without assistance from others, including automated tools. If you are unclear if something is an AI tool, please check with your instructor. Using AI tools for any purposes in this course will violate the University’s academic integrity policy. I treat potential academic integrity violations by […]”

1. **Use with prior permission:**

“Students are only allowed to use AI tools, such as ChatGPT or Dall-E 2, on assignments in this course when advance permission is given by the instructor. Students must submit a written request with an explanation of how they will use a particular tool in their assignment. Students are not permitted to use these tools until permission is granted in writing. The instructor may encourage and give permission to students to use AI tools during class activities and in other contexts when it is considered in support of the course learning goals. Unless given permission to use those tools, each student is expected to complete each course assignment without substantive assistance from others, including AI tools. If you are unclear if something is an AI tool, please check with your instructor. Unauthorized use of AI tools for any purposes in this course will violate the University’s academic integrity policy. I treat potential academic integrity violations by […]

1. **Use only with proper citation:**

“The use of AI tools, such as ChatGPT or Dall-E 2, for this course is allowed for specific assignments only when determined to be in support of the course learning goals. Assignments in which AI tools are permitted will be clearly identified by the instructor and noted in the assignment directions. You are not required to use AI tools, but if you choose to use them for any part of the assignment (from brainstorming to text editing), you must use proper citation (please use APA citation format). Failure to properly cite AI tools is considered a violation of the University of

Chicago’s Academic Honesty and Plagiarism policy. If you are unclear if something is an AI Tool, please check with your instructor. I treat potential academic integrity violations by […]”

1. **Free use with no citation required:**

“In this course, students are allowed to use AI tools (such as ChatGPT) on all assignments. No citation is required.”

### Additional Resources

To schedule a one-on-one consultation with a member of the CCTL staff, please email: teaching@uchicago.edu If you are interested in reading more, below are some curated resources to learn more about AI tools and how others in higher education are approaching their use in their courses: • “Should You Add an AI Policy to Your Syllabus? What to consider in drafting your own course policy on students’ use of tools like ChatGPT.” by Kevin Gannon • “Sentient Syllabus Project” • “My Assessments Next Semester – Modified for Avoiding & Embracing AI” by Maha

Bali • Crowdsource Classroom Policies for AI Generative Tools

**16.** [**Considerations for AI Tools in the Classroom**](https://ctl.columbia.edu/resources-and-technology/resources/ai-tools/)

#### Considerations for AI Tools in the Classroom

Given the rapid pace of technological innovation and development, higher education, like nearly all industries, is continuously called upon to consider creative approaches to teaching and learning. The following resource offers instructors a brief introduction to Artificial Intelligence (AI) Tools, specifically ChatGPT, along with several strategies they might consider for navigating or engaging with these tools in their courses.

**For instructors: share your thoughts with us!**

Have additional questions about AI Tools in your classroom? Have examples of how you are integrating these tools in your course or redesigning your assignments to address them? Email CTLFaculty@columbia.edu to share your questions and innovations!

**Cite this resource:** Columbia Center for Teaching and Learning (2023). Considerations for AI

Tools <https://ctl.columbia.edu/resources-and-technology/resources/ai-tools/>

in

the

Classroom.

Columbia

University.

Retrieved

[

today’s

date]

from

**What are AI Tools?**

While instructors will continue to encounter new tools and technological innovations, which can sometimes feel overwhelming, the science of teaching and learning offers support for navigating and responding to innovative tools, expanded opportunities, and ongoing shifts.

Most recently, higher education has grappled with the seemingly-overnight introduction of ChatGPT, which first became available in November 2022 and has since captivated the discourse in higher education with its ongoing evolution and iterations. The following section offers an overview of ChatGPT, what it does, and its current limitations. It’s important to recognize, however, that because the technology is self-learning, it can improve and evolve quickly; thus, what is a limitation today could be addressed in future iterations or versions. This was perhaps made most evident with the release of GPT-4 during the Spring 2023 semester, which introduced a series of new capabilities for the generative AI tool. Furthermore, since the release of ChatGPT hundreds of AI apps have exploded into the scene, some of which assist in research (e.g. consensus.app and scite.ai), brainstorming (e.g. mymap.ai), and student presentations (e.g. lenovo.ai, twelvelabs.io, and invideo.io).

#### ChatGPT

ChatGPT, created by OpenAI, is a “model […] which interacts in a conversational way” and has the ability to “answer followup [sic] questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests” ([OpenAI Blog,“ChatGPT”)](https://openai.com/blog/chatgpt/). In respect to the classroom, ChatGPT can produce written responses to input prompts, write essays and poems, assist with computer code, provide feedback on student-written text, and more.

The most recent version, GPT-4, is described as “more creative and collaborative” with the ability to “generate, edit, and iterate with users on creative and technical writing tasks” ([OpenAI, GPT4)](https://openai.com/product/gpt-4). Unlike its predecessors, GPT-4 is capable of producing longer form responses, with more complex analyses in response to particular prompts. It has wider capabilities such as image analysis, and the addition of plug-ins that connect ChatGPT to other third party services, including information from the web. It is important to note, however, that GPT-4 and its capabilities are only available through a *paid subscription* to GPT Plus; these more advanced features are not built into the free uses of ChatGPT.

Despite these expanded capabilities, like many AI tools, ChatGPT does have a variety of limitations. Some of its current limitations include:

* Sharing incorrect information: The paid GPT-4 version of ChatGPT has access to the internet, while GPT-3.5 (the free version) does not and can only access information prior to 2021. Since ChatGPT is a predictive text model, it can at times make up information, especially if the information is not easily accessible. This makes ChatGPT prone to making factual mistakes, often confusing similar information, and even making up citations when asked to produce them.
* Generating personal reflections: ChatGPT can generate responses to questions and prompts, but it is still an AI bot; thus, it is unable to respond to a prompt that requires a student’s personal experience or reflections.
* Producing non-text based responses: Despite the availability of other image-based AI generators such as Dall-E 2 and MidJourney, ChatGPT responses are strictly text-based. Therefore, any response produced can only offer text. With its most recent updates, ChatGPT can, however, accept image input prompts to generate captions and produce image analyses.

An awareness of ChatGPT’s capabilities and limitations can help instructors talk openly with their students about the potential role of the tool in the course and what is expected in terms of student engagement with AI tools. Additionally, knowing what ChatGPT can and cannot do can also help instructors make decisions around course and assignment design. The following section offers several approaches for instructors to consider when developing their own approaches to ChatGPT in their own courses. Instructors can explore OpenAI’s documentation [Educator considerations for ChatGPT](https://platform.openai.com/docs/chatgpt-education) as they consider their own approach to teaching and learning with ChatGPT.

#### ChatGPT in the Classroom

There is no one correct way to navigate the use of ChatGPT in the classroom. The role it plays (or doesn’t) in a course will depend on several factors including course objectives and goals, disciplinary skills that students need to practice, and instructors’ own comfort level with engaging with the tool. No matter how instructors respond to the evolution of AI tools, it’s important to have explicit expectations, and provide opportunities for clear conversations with students about those expectations. Read on to learn more about several possible approaches and considerations to help instructors start navigating, leveraging, and possibly even engaging ChatGPT (and future AI tools) in their course.

*View recording of the “Teaching and ChatGPT” forum held on February 13, 2023. Columbia contributors shared resources and facilitated an informal conversation about AI tools, specifically ChatGPT, and its implications in the classroom.*

#### Develop Course Policies that Include Digital Transparency

It is important to be explicit with students about the expectations around the usage of ChatGPT and other AI tools in your course. For example, ChatGPT is capable of reading a student’s essay and providing meaningful feedback that can then be used by the student to make edits. As the instructor, be clear about these expectations: can students use the tool for feedback on their own writing? If so, how should they disclose their use? As with all course policies, especially those around [academic integrity,](https://ctl.columbia.edu/resources-and-technology/resources/academic-integrity/) it is essential for instructors to be explicit and transparent with their expectations, and to have frank conversations with their students. Some colleagues are collecting and maintaining [an open-source repository](https://docs.google.com/document/d/1RMVwzjc1o0Mi8Blw_-JUTcXv02b2WRH86vw7mi16W3U/edit) of sample digital transparency language and policies from higher education. While not affiliated with Columbia, this collection can offer some insight into the different approaches institutions and individual instructors are taking when it comes to addressing AI tools in the classroom.

Since ChatGPT’s introduction, there has been a parallel rise in tools claiming accurate detection of AI-generated work. As with any form of detection software, there are risks of misidentification, which can have consequences in the classroom. These products are best used with careful consideration, and as one of many ways to work with students. It is also important to include the use of these tools in any discussion with students around course policies, making clear why and how such services may be used in the course.

When developing usage policies in a course, instructors might also consider [partnering with their students](https://ctl.columbia.edu/faculty/sapp/ai-tools/#:~:text=Create%20space%20for%20student%2Dinstructor%20dialogue%20on%20academic%20integrity.) to develop the policies. This partnership can create opportunities for instructors and students to talk in detail about the evolution of particular tools, their potential benefits in specific disciplines, and their limitations. Instructors should be explicit about the course objectives and how the use of these tools might interfere with students’ learning and their achievement of particular learning goals.

#### Scaffold Activities and Assignments

Regardless of AI tool innovation or evolution, one important approach for instructors is to leverage scaffolded activities and assignments. Scaffolding is the process of breaking down a larger assignment into subtasks, which create opportunities for students to check-in and receive [feedback.](https://ctl.columbia.edu/resources-and-technology/resources/feedback-for-learning/)

At the same time, scaffolding can help instructors become more familiar with students’ work as the semester progresses. This cyclical process of feedback and revision makes the use of tools like ChatGPT challenging and perhaps even unlikely, as students will provide drafts incrementally, and engage in the process of drafting and revision. More importantly, though, this breaking down of a large project into incremental parts helps students to more deeply engage with the different skills and component parts, while also creating valuable time for feedback and reflection throughout the process. For ideas on how to scaffold student work, view the [recording](https://columbiauniversity.zoom.us/rec/play/FgguV5NOcc4LLHRDPDx24LMY_aA5qLBPoo9N13YssERjp23IIFc7xTPSazg3d85x3LWt_2Fz1bZYmxzV.Lmq7IpirNES7tcMp?canPlayFromShare=true&from=share_recording_detail&startTime=1683043386000&componentName=rec-play&originRequestUrl=https%3A%2F%2Fcolumbiauniversity.zoom.us%2Frec%2Fshare%2FpAEefeb6AsMDKNhiqn6FZIYfIqnaeQKu58MwiOYcay60tbNhrXOrTlUc2u04Uszr.C0YLf4zxCJQg5gK9%3FstartTime%3D1683043386000) of “Using

AI writing tools in your scientific writing process” (May 2, 2023) in which Tim Requarth shares how he uses AI tools to streamline and support the writing process.

#### Design Authentic Assessments for Learning

[Authentic assessments centered on student learning](https://ctl.columbia.edu/resources-and-technology/teaching-with-technology/teaching-online/designing-assignments/) can help instructors and learners make intentional choices about the integration of AI tools into the writing process. These kinds of authentic assessments ask students to apply the course concepts they have learned to a “real world” situation or problem. In doing so, authentic assessments can enhance student learning by engaging students in “doing” a particular subject and practicing specific disciplinary skills that will help prepare them for their professional lives outside the classroom. Additionally, these kinds of assignments engage students in higher order thinking and require that they grapple with real world problems and challenges. Designing authentic assessments can ask students to draw from and engage with specific course materials, explore their local community, make connections between course concepts, or also ask students to incorporate their personal experiences or reflections. In this way, authentic assessments can help prevent the use of tools like ChatGPT in that their very design and objectives are rooted within specific course concepts, while also asking students to infuse their own experiences and reflections.

#### Incorporate AI Tools into Assignment Design

For some courses, depending on the goals and objectives, instructors might consider ways to incorporate AI tools in their assignment design; in doing so, instructors can provide students with opportunities to practice and foster the digital literacy skills they will need for the future. These kinds of [creative assignments](https://ctl.columbia.edu/resources-and-technology/resources/creative-assignments/) might ask students to produce AI-written texts as a way to develop awareness of voice, authorship, and accuracy. Additionally, students could apply a rubric and offer feedback on AI-produced texts to build deeper awareness of a course prompt. Lastly, in some instances, ChatGPT might be called upon as a learning support tool, where students ask for feedback on their own texts, have readings summarized, create personalized study materials, or brainstorm for ideas. No matter the assignment design approach taken, instructors should offer opportunities to discuss the assignment with students, asking them to [reflect](https://ctl.columbia.edu/resources-and-technology/resources/metacognition/) on the experience and analyze their engagement with the tool.

#### Conclusion

Perhaps unsurprisingly, higher education, like all industries, will continue to feel the impacts of technological evolution and growth. As such, and like they always have, classrooms will continue to remain flexible and responsive to this evolution. With digital innovation and developments, the capabilities (and limitations) of today’s AI tools, including ChatGPT, will shift and evolve. For that reason, trying to completely ignore or shut out these tools, or even adopting an approach of complete disengagement, will not serve instructors and their students in the long-term. Instead, instructors have an opportunity to rethink and focus on the elements of their course over which they have the most control, including transparent course policies, explicit communication, partnerships with students, and course and assignment design. Leveraging these aspects of teaching and learning can better serve instructors and their students no matter the digital innovations of the future.

**The CTL is here to help!**

Looking to develop scaffolded authentic assignments for your course? Want to know more about designing assignments that incorporate AI Tools? Have questions about digital transparency in your course policies? The CTL is here to help – email CTLFaculty@columbia.edu to schedule a 1-1 consultation!

#### Columbia CTL Related Resources

* [Designing Assignments for Learning](https://ctl.columbia.edu/resources-and-technology/teaching-with-technology/teaching-online/designing-assignments/)
* [Feedback for Learning](https://ctl.columbia.edu/resources-and-technology/resources/feedback-for-learning/)
* [Getting Started with Creative Assignments](https://ctl.columbia.edu/resources-and-technology/resources/creative-assignments/)
* [Learner Perspectives on AI Tools: Digital Literacy, Academic Integrity, and Student Engagement](https://ctl.columbia.edu/faculty/sapp/ai-tools/)
* [Metacognition](https://ctl.columbia.edu/resources-and-technology/resources/metacognition/)

### Additional Related Resources

* Barnard Center for Engaged Pedagogy (2023). [Generative AI & the College Classroom.](https://cep.barnard.edu/generative-ai-college-classroom)
* Cornell University Center for Teaching Innovation (2023). [Promoting Academic Integrity in Your Course.](https://teaching.cornell.edu/promoting-academic-integrity-your-course)
* Digital Futures Institute, Teachers College, Columbia University (2023). [ChatGPT and Other Artificial Intelligence (AI) in the Classroom.](https://www.tc.columbia.edu/digitalfuturesinstitute/future-of-education/chatgpt-and-other-ai-in-the-classroom/)
* Digital Futures Institute, Teachers College, Columbia University (2023). [Thinking About Assessment in the Time of Generative Artificial Intelligence.](https://www.tc.columbia.edu/digitalfuturesinstitute/future-of-education/thinking-about-assessment-in-the-time-of-generative-artificial-intelligence/)
* Montclair State Office for Faculty Excellence. (2023). [Practical Responses to ChatGPT.](https://www.montclair.edu/faculty-excellence/practical-responses-to-chat-gpt/)
* Penn Center for Teaching and Learning (2023). [ChatGPT and Its Implications for Your Teaching.](https://www.ctl.upenn.edu/resources/tech/chatgpt/)
* University of Central Florida Faculty Center. (2023). [Artificial Intelligence Writing.](https://fctl.ucf.edu/teaching-resources/promoting-academic-integrity/artificial-intelligence-writing/)
* Yale Poorvu Center for Teaching and Learning (2023). [AI Guidance.](https://poorvucenter.yale.edu/AIguidance)

1. [**Guidance on the Use of Generative AI and Large Language Model Tools**](https://www.imss.caltech.edu/services/ai)

**Guidance on the Use of Generative AI and Large Language Model Tools**

Caltech provides this initial guidance to encourage the responsible use of generative artificial intelligence (GenAI) and large language model (LLM) tools and technologies, such as OpenAI's ChatGPT and Dall-E and Google's Bard in research, education, and administrative work at Caltech. As a research and education institute, committed to advancing the frontiers of science and engineering and expanding knowledge, we support a responsible, measured experimentation with and use of new technologies. While doing this, however, Caltech requires that you follow all existing applicable regulations and Institute policies. These include, but are not limited to, ensuring protection of confidential, personal, or business information and intellectual property, and adherence to the honor code, course requirements, research integrity, and publication ethics.

GenAI and LLM technologies have evolved rapidly in their use and application this past year and are expected to continue to evolve in ways society cannot predict. Likewise, our guidance for the appropriate use of these tools is written for the present moment and will likely evolve alongside the technology. In the interim, however, as you use GenAI and LLM technologies in your work at Caltech, we ask that you apply these four guiding principles to your practice: disclosure, data and information protection, content responsibility, and Caltech's honor code.

* 1. Disclosure: When using GenAI, always disclose promptly, or reference the use of GenAI tools and application plug-ins, as applicable. This transparent disclosure ensures that others are aware when GenAI was used to generate content and reduces misunderstandings regarding the source of information, potentially limiting claims of academic dishonesty or plagiarism. When using GenAI to write or publish, please make sure to follow the guidance provided by the course instructor or journal or manuscript publisher/editor. For example, some may require that the GenAI be included as an author, others may simply require acknowledgement.
  2. Data and Information Protection: Federal, state, and local laws as well as Caltech policies may limit data that can be disclosed. Unless you are using a GenAI application that ensures separation of your entry from other entries and confidentiality (usually a paid service), uploading content into GenAI (Open GenAI), is a public disclosure. It is safest to assume data or queries uploaded into Open GenAI tools will become public information, unless otherwise indicated.

In order to protect Caltech data and information, do not enter, contribute, or otherwise input sensitive, confidential, or restricted information into open GenAI tools. This includes, but is not limited to, data covered by regulations such as FERPA and HIPAA, any intellectual property or unpublished research data, export-controlled data, and other sensitive HR, business, or administrative data. Caltech is considering a subscription to a restricted GenAI and will keep you apprised of its progress in securing such a service. In the meantime, Caltech has and continues to reserve the right to disable or limit access to AI companion tools in enterprise business software and applications, such as Zoom and Microsoft Office suites.

* 1. Content Responsibility: Remember that GenAI systems are fallible. Responses can be inaccurate, misleading, and even entirely fabricated. Therefore, you should always review and assess all output generated by GenAI tools for accuracy before relying on them or distributing them publicly.
  2. Honor Code: Caltech's honor code underscores the importance of ethical conduct and fairness and extends to the use of GenAI tools and is stated as follows: "No member of the Caltech community shall take unfair advantage of any other member of the Caltech community." Please note that we offer these guidelines in addition to the teaching resources that have already been provided by the [Center for Teaching and Learning and Outreach.](https://ctlo.caltech.edu/universityteaching/resources/resources-for-teaching-in-the-age-of-ai)

The Institute's guidance promotes responsible and ethical use of GenAI and LLM tools at Caltech, and fosters a community that values transparency, integrity, privacy, accuracy, and fairness. Caltech may update its guidance as the technology and regulatory and commercial landscapes evolve. Thank you for adhering to these guidelines. If you have any questions about the use of this dynamic technology, please email gen\_ai@caltech.edu.

1. [**Appropriate use of ChatGPT and Similar AI Tools**](https://ethics.berkeley.edu/privacy/appropriate-use-chatgpt-and-similar-ai-tools)

With the emergence of ChatGPT and other AI tools, many members of our community are eager to explore their use in the university context. This advisory provides guidance on how to use these tools safely, without putting institutional, personal, or proprietary information at risk. Additional guidance may be forthcoming as circumstances evolve.

**Allowable Use:**

* + Publicly-available information ([Protection Level P1](https://security.berkeley.edu/data-classification-standard#plminimal)
  + [(link is external)](https://security.berkeley.edu/data-classification-standard#plminimal)
  + ) can be used freely in ChatGPT.
  + In all cases, use should be consistent with [UC Berkeley’s Principles of Community](https://diversity.berkeley.edu/principles-community) ● [(link is external)](https://diversity.berkeley.edu/principles-community)

**Prohibited Use:**

* + At present, any use of ChatGPT should be with the assumption that no personal, confidential, proprietary, or otherwise sensitive information may be used with it. In general, [student records subject to FERPA](https://registrar.berkeley.edu/ferpa-privacy-disclosure/)
  + [(link is external)](https://registrar.berkeley.edu/ferpa-privacy-disclosure/)
  + , and any other information classified as [Protection Level P2, P3, or P4](https://security.berkeley.edu/data-classification-standard#plclassification)
  + [(link is external)](https://security.berkeley.edu/data-classification-standard#plclassification)
  + should not be used.
  + Similarly, ChatGPT should not be used to generate output that would be considered nonpublic. Examples include, but are not limited to, proprietary or unpublished research; legal analysis or advice; recruitment, personnel or disciplinary decision making; completion of academic work in a manner not allowed by the instructor; creation of non-public instructional materials; and grading.
  + Please also note that OpenAI explicitly forbids the use of ChatGPT and their other products for certain categories of activity, including fraud and illegal activities. This list of items can be found in their [usage policy document](https://openai.com/policies/usage-policies)
  + [(link is external)](https://openai.com/policies/usage-policies) **Additional Guidance:**

For further guidance on the use of ChatGPT for teaching and learning, please see [Understanding](https://teaching.berkeley.edu/understanding-ai-writing-tools-and-their-uses-teaching-and-learning-uc-berkeley)

[AI Writing Tools and their Uses for Teaching and Learning at UC Berkeley](https://teaching.berkeley.edu/understanding-ai-writing-tools-and-their-uses-teaching-and-learning-uc-berkeley)

[(link is external)](https://teaching.berkeley.edu/understanding-ai-writing-tools-and-their-uses-teaching-and-learning-uc-berkeley)  from Research, Teaching & Learning.

**Rationale for the Above Guidance:**

UPDATE 7/1/2023: The University of California recently renegotiated the UC systemwide agreement with Microsoft to include Microsoft Azure OpenAI. UC Berkeley is currently working with Microsoft to establish how to use this service under UC’s agreement. [Note: For questions regarding the approved use of Microsoft's Azure's Open AI service, please consult with the Privacy Office at privacyoffice@berkeley.edu

(link sends e-mail)

As of September 2023, the University of California's agreements with the parent companies of

ChatGPT and other generative AI services, which include our terms and conditions, Appendix Data Security and other privacy protections, do not cover the use of ChatGPT or other similar generative AI services. The UC Office of the President is working on this issue. We hope to see this addressed in the near future and will update this guidance when additional information is available.

Personal Liability: ChatGPT uses a click-through agreement. Click-through agreements, including OpenAI and ChatGPT terms of use, are contracts. Individuals who accept click-through agreements without delegated signature authority may face personal consequences, including responsibility for compliance with terms and conditions. [1]

#### Guidance on Appropriate Use

For questions regarding the appropriate use of ChatGPT and other AI tools, please contact privacyoffice@berkeley.edu (link sends e-mail)

#### References

* [Educator considerations for ChatGPT](https://platform.openai.com/docs/chatgpt-education)
* [(link is external)](https://platform.openai.com/docs/chatgpt-education)
* [OpenAI sharing & publication policy](https://openai.com/policies/sharing-publication-policy)
* [(link is external)](https://openai.com/policies/sharing-publication-policy)
* [OpenAI usage policies](https://openai.com/policies/usage-policies)
* [(link is external)](https://openai.com/policies/usage-policies)
* [OpenAI privacy policy](https://openai.com/policies/privacy-policy)
* [(link is external)](https://openai.com/policies/privacy-policy)
* [OpenAI terms & policies](https://openai.com/policies)
* [(link is external)](https://openai.com/policies)
* [[1] Delegations of Authority: To find out who has signature authority at Berkeley to “sign” a click-through agreement, go to the Delegations of Authority webpage in the Office of the Chancellor Compliance Services website.](https://ethics.berkeley.edu/privacy/appropriate-use-chatgpt-and-similar-ai-tools#Anchor)

#### Additional Readings

* [Governor Newsom Signs Executive Order to Prepare California for the Progress of](https://www.gov.ca.gov/2023/09/06/governor-newsom-signs-executive-order-to-prepare-california-for-the-progress-of-artificial-intelligence/)

[Artificial Intelligence](https://www.gov.ca.gov/2023/09/06/governor-newsom-signs-executive-order-to-prepare-california-for-the-progress-of-artificial-intelligence/)

* [(link is external)](https://www.gov.ca.gov/2023/09/06/governor-newsom-signs-executive-order-to-prepare-california-for-the-progress-of-artificial-intelligence/)

●

* [Understanding AI Writing Tools and their Uses for Teaching and Learning at UC Berkeley](https://teaching.berkeley.edu/understanding-ai-writing-tools-and-their-uses-teaching-and-learning-uc-berkeley)
* [(link is external)](https://teaching.berkeley.edu/understanding-ai-writing-tools-and-their-uses-teaching-and-learning-uc-berkeley)
* , from UC Berkeley Research, Teaching & Learning
* [Quantamagazine: The Unpredictable Abilities Emerging From Large AI Models](https://www.quantamagazine.org/the-unpredictable-abilities-emerging-from-large-ai-models-20230316/)
* [(link is external)](https://www.quantamagazine.org/the-unpredictable-abilities-emerging-from-large-ai-models-20230316/)
* [The Atlantic: Don’t Be Misled by GPT-4’s Gift of Gab](https://www.theatlantic.com/newsletters/archive/2023/03/dont-be-misled-by-gpt-4s-gift-of-gab/673411/)
* [(link is external)](https://www.theatlantic.com/newsletters/archive/2023/03/dont-be-misled-by-gpt-4s-gift-of-gab/673411/)
* [University of California Presidential Working Group on Artificial Intelligence](https://www.ucop.edu/ethics-compliance-audit-services/compliance/presidential-working-group-on-artificial-intelligence.html)
* [(link is external)](https://www.ucop.edu/ethics-compliance-audit-services/compliance/presidential-working-group-on-artificial-intelligence.html)
* [Inclusive Intelligence: Artificial Intelligence in the Service of Science, Work, and the Public Good](https://vcresearch.berkeley.edu/signature-initiatives/inclusive-intelligence)
* [(link is external)](https://vcresearch.berkeley.edu/signature-initiatives/inclusive-intelligence)
* , from UC Berkeley Research

**19.** [**Guidelines for the Use of Generative AI Tools**](https://provost.yale.edu/news/guidelines-use-generative-ai-tools)

Dear Members of the Yale Community,

Publicly available generative Artificial Intelligence (AI) tools such as ChatGPT, Bing, Midjourney, and Bard have garnered tremendous attention in the past year. The field of generative AI is developing rapidly. While its precise impacts are unknown, this technology will transform how we learn, teach, conduct research, and carry out daily tasks. We encourage you to experiment with AI tools. As you explore AI’s potential, please adhere to the following general guidelines, which align with existing university policies and uphold our institutional commitment to safety, security, and academic integrity.

1. **Protect Yale’s confidential information and your own.** Do not enter confidential or legally restricted data or any data that Yale’s data classification policy identifies as moderate or high-risk into an AI tool. If you are not sure whether you should share certain data, please review [Yale’s](https://click.message.yale.edu/?qs=843aa8b121129d67d076be4d618cdc732f3a204fc7258d35e43facc86a5ff100e8b62aa8b3f92e2fca9c4920da28821dc44ef0e5afc30f3b2aafe94a4e3ab2d8) data classification polic[y.](https://click.message.yale.edu/?qs=843aa8b121129d67d076be4d618cdc732f3a204fc7258d35e43facc86a5ff100e8b62aa8b3f92e2fca9c4920da28821dc44ef0e5afc30f3b2aafe94a4e3ab2d8)

1. **Assume all information shared will be made public.** Treat all information shared with an AI tool as if it will become public. Do not share information that is personal or sensitive, and be

mindful that the information you input into an AI tool may be retained.

1. **Always follow academic integrity guidelines and institutional standards of conduct.** All students and faculty are expected to know and adhere to their school’s academic integrity policies. Faculty members are expected to provide clear instructions on the permitted use of generative AI tools for academic work and requirements for attribution. Likewise, students are expected to follow

their instructors’ guidelines about permitted use of AI for coursework.

1. **Be alert for bias and inaccuracy.** AI-generated responses can be biased, inaccurate, inappropriate, or may contain unauthorized copyrighted information. We are each responsible for the content of our work product. Always review and verify outputs generated by AI tools, especially before publication.

1. **Protect yourself and your credentials.** Never share your Yale NetID and password with AI tools, and always be aware of phishing schemes. For information, tips, and toolkits on cybersafe practices, visit [Yale’s Cybersecurity website,](https://click.message.yale.edu/?qs=843aa8b121129d675dab98f5abf9201354a8b88d7b01f7493c84e83b0862cc340403806a9473235c0f04d6e71de7de082615ad2e9fe1c2cccf760cffef5762c4) which also includes information about security policies and standards.

1. **Seek support.** The university is working to support procurement practices that coordinate shared interests and minimize institutional risk. If you are considering acquiring an AI product, please conduct an initial review of the tool to ensure that it conforms to institutional security requirements. Use [Yale’s purchasing intake portal](https://click.message.yale.edu/?qs=843aa8b121129d67fc2531d889846a91cc339baf2c0459c4e58ef79886f7a64938a31ea2bd0ed8e8f99a8bc82b7273e4f8f72fe4990e992502ee6f13ed01d4f1) or contact purchasing.helpdesk@yale.edu.

Generative AI is evolving quickly, as is institutional support for utilizing these tools. For up-todate recommendations on teaching and learning, please visit the [Poorvu Center’s webpage on “AI](https://click.message.yale.edu/?qs=843aa8b121129d675ac29c5810144872c769b85eedad651984fd683b06e049431e0a1d81121a844ead42252b5b527c9f77df0df784b1478187a99d5db3d5d7e9)

[Guidance.”](https://click.message.yale.edu/?qs=843aa8b121129d675ac29c5810144872c769b85eedad651984fd683b06e049431e0a1d81121a844ead42252b5b527c9f77df0df784b1478187a99d5db3d5d7e9) Detailed AI guidelines for staff are available on Yale’s [data governance website.](https://click.message.yale.edu/?qs=843aa8b121129d67dc346a0366a62e0176c97e4f887ff7743c5c12d690e4b20d7447e641fd3ba8af425a1a033fac3ff2d89ca8b3840c7cc1296341d82b350d49) As a university community dedicated to exchanging ideas, disseminating knowledge, and fostering a climate for breakthrough discoveries, we will embrace technological tools and harness their power for innovation. We must do this safely and responsibly. We thank you for your efforts in adhering to these guidelines.

**20.** [**Statement on Guidance for the University of Pennsylvania Community**](https://www.isc.upenn.edu/security/AI-guidance#:~:text=In%20the%20absence%20of%20other,AI%20to%20complete%20the%20task.)

Penn embraces innovations like generative artificial intelligence (“AI”) models in teaching, learning, research, and the effective stewardship of Penn’s resources. To this end, this document provides guidelines for members of the Penn community who are using, or interested in using, AI in pursuit of Penn’s mission.

#### Scope

This document is scoped to generative AI using large language models provided by third parties. Generative AI describes algorithms, such as ChatGPT and other large language models, that can be used to create new content, including text, code, and simulations.

This statement is not intended as legal advice or an exhaustive set of best practices and should not be viewed as a final policy. The AI field is rapidly evolving in terms of technology, deployment models, third-party relationships, terms of service, regulatory landscape, and academic-industry partnership structures. It is anticipated that this document will be updated regularly and interact with other sources of policy, ethics, and governing legal authority.

#### General Guidance for Penn Community (Educators, Staff, Researchers, and Students)

Transparency

Be transparent about the use of AI. Disclose when a work product was created wholly or partially using an AI tool and, if appropriate, how AI was used to create the work product.

Accountability

The user of AI should endeavor to validate the accuracy of created content with trusted first party sources and monitor the reliability of that content. Users are accountable for their use of content created by AI and should be wary of misinformation or “hallucinations” by AI tools (e.g., citations to publications or source materials that do not exist or references that otherwise distort the truth).

Bias

When using AI, keep in mind that these tools are often trained on large, unmoderated bodies of text, such as text posted to the internet. This can result in the production of biased and other unintended content. The ability to avoid such biased content is still in the early stages of development. Privacy & Contracts

|  |  |
| --- | --- |
| Most AI tools and services use input and data from users of the tool to train the model. Additionally, existing tools may incorporate AI features in their service offerings. For this reason, users of AI should avoid sharing personal or sensitive data with the tool and should not input moderate or high-risk Penn data as defined by the [Penn Data Risk Classification,](https://www.isc.upenn.edu/security/data-classification) or intellectual property, without: | |
|  | 1. Careful consideration and understanding of the tool’s use of Penn data and the service provider’s stated rights to the data, including, but not limited to whether the service provider offers the option to opt-out of using customer’s data to train the AI; 2. A contract in place to protect Penn data; and 3. Review by Penn’s Privacy Office and consultation with the Office of Information   Security as coordinated by Procurement when moderate or high-risk data is involved. |
| Consultation with the Penn Center for Innovation, where intellectual property is involved. | |

Patient Privacy Protection

It is not permissible under the Health Information Portability & Accountability Act (HIPAA) or Penn Medicine policy to share patient or research participant information in connection with open or public AI tools and services, such as ChatGPT. This is because, as currently configured, such open or public tools and services can use and share any data without regard to HIPAA restrictions and other protections. Therefore, individual patient data and patient data sets (even if deidentified) may not be exposed to open or public AI tools or services, absent institutional approval.

Security

When using AI to write computer code or when creating new technology that leverages AI it is important to be aware of the new kinds of cyberattacks that are being used against AI users.

Review the [Office of Information Security guidance](https://www.isc.upenn.edu/security/LLM-guide) on these risks or consult with the Office of Information Security if in doubt.

Data Scraping

The rise of AI models has led to a significant increase in individuals and organizations scraping (i.e., copying) information posted on the internet for the purpose of training new AI models. Be aware that any data posted publicly will likely be scraped and used in this way by third parties. Similarly, while these practices are common, their legality and the potential consequences of these actions are currently being developed but remain unresolved at the time this guidance was issued.

Intellectual Property

Members of the Penn community should adhere to established principles of respect for intellectual property, particularly copyrights when considering the creation of new data sets for training AI models. Avoid uploading confidential and/or proprietary information to AI platforms prior to seeking patent or copyright protection, as doing so could jeopardize IP rights.

University Business Processes

While automating tasks using AI may improve operational efficiency for University Business processes, oversight and review of the use of AI and verification of its outputs for these University business processes should be in place to ensure reliability, consistency, and accuracy.

**Additional Guidance**

Guidance for Educators

|  |  |
| --- | --- |
|  | Use of AI should always be in alignment with Penn’s [Principles of Responsible Conduct.](https://oacp.upenn.edu/oacp-principles/)  As expectations may vary between classes and instructors, it is important for instructors to provide students with clear guidelines similar to the [guidelines on collaboration,](https://ctl.upenn.edu/resources/academic-integrity/promoting-integrity/collaboration/) on the use of AI within coursework, and when and how the use of AI within a course should be cited.  Disclose to students when course materials have been created with the use of AI and when AI detection software will be used in the course.  Be aware of how AI tools may impact assignments and exams in your discipline. [Penn’s Center for Teaching and Learning](https://ctl.upenn.edu/resources/tech/chatgpt/) provides useful guidance in this space. |
| Guidance for Students | |
|  | All use of AI should be in line with Penn’s [Code of Student Conduct](https://catalog.upenn.edu/pennbook/code-of-student-conduct/) and the [Code of Academic Integrity.](https://catalog.upenn.edu/pennbook/code-of-academic-integrity/)  Educators may have requirements and guidance for citing the use of generative AI output and for attributing AI created content to the specific AI tool and parameters used.  Individual courses may have different or more narrow guidance on the use of AI that should be adhered to within the context of that course.  In the absence of other guidance, treat the use of AI as you would treat assistance from another person. For example, this means if it is unacceptable to have another person substantially complete a task like writing an essay, it is also unacceptable to have AI to complete the task.  Keep in mind that having access to data is not the same as having permission to scrape the data or use it to train an AI model. |
| Guidance for Researchers | |
|  | Consult with your department leadership and your discipline’s publishing standards to determine how the use of AI should be accounted for with regard to authorship in publications.  Researchers should adhere to federal or international requirements on obtaining informed consent, and Institutional Review Board approvals should be obtained prior to exposing research participant data to AI tools.  Caution should be adopted when research involves the examination of high-risk data, including Personally Identifiable Information (PII) and research participant health information (both identifiable and non-identifiable) exposed to AI. |
|  | Intellectual property law may limit what AI can be copyrighted or patented depending on the specifics of their creation, and using data to train AI may implicate copyright laws and/or the patentability of an idea or discovery. Researchers should avoid uploading confidential and/or proprietary information to AI platforms, at least prior to seeking patent or copyright protection, as doing so could jeopardize IP rights. AI users should consult further with the Penn Center for Innovation for additional guidance, including determinations as to whether potential IP rights should be protected prior to the inclusion of data and findings in AI models.  Researchers should reinforce with their mentees (undergraduate students, predoctoral students, and postdoctoral trainees) on the appropriate use of AI. |
| **FAQs AI Guidance** | |

What kinds of AI does this guidance apply to?

This AI Guidance is largely focused on large language models, such as ChatGPT and Google Bard; but these principles apply generally to other machine learning and artificial intelligence technologies.

Who does this AI Guidance apply to?

This Guidance applies to Penn faculty, staff, undergraduate students, pre-doctoral students, and post-doctoral trainees

Who developed this Guidance?

The Offices of Information Security and University Privacy developed this Guidance in collaboration with the Office of the Vice Provost for Education, the Office of the Vice Provost for Faculty, the Office of the Senior Vice Provost for Research, the Office of Clinical Research, the Office of General Counsel, Procurement, and the Penn Center of Innovation.

Can Schools, Centers or departments publish more specific guidance around the use of AI? Yes, Schools, Centers, and other departments may publish more specific guidance on the use of AI to address certain situations or use cases.

|  |  |
| --- | --- |
| Who do I contact if I have additional questions about this AI Guidance? | |
|  | For questions about educational policies, please contact the Vice Provost for Education at provost-ed@upenn.edu.  For questions regarding clinical research, please contact the Office of Clinical Research at lauraee@upenn.edu. |

For questions about research use, please contact the [Office of the Senior Vice Provost for Research](https://research.upenn.edu/about/contact/) at vpr@upenn.edu.

For questions about procurement, including contracting, sourcing, and supplier management, please contact [Procurement Services](https://cms.business-services.upenn.edu/purchasing/about/contact-information.html) at procure@upenn.edu.

For questions regarding intellectual property and AI, contact the Penn Center for Innovation at pciinfo@pci.upenn.edu.

For questions about privacy, please contact the University Privacy Office at privacy@upenn.edu.

For questions about security, please contact the Office of Information Security at security@isc.upenn.edu.

For legal questions related to this guidance document, contact the Office of General Counsel at [Contact Us | OGC (upenn.edu).](https://ogc.upenn.edu/contact-us)

**21.** [**Guidance for the use of generative AI**](https://teaching.ucla.edu/resources/ai_guidance/)

#### About this guide

Students and instructors are embracing ChatGPT (GPT-4 as of March 2023) and similar artificial intelligence (AI) technologies across disciplines for different learning goals. There is no one-sizefits-all best practice for their use. This document is meant as a guideline for instructors on what to consider as these tools evolve. We will provide strategies for adopting AI technologies in a responsible, ethical manner, and innovating within each discipline, major, and course. Exploring and communicating about the opportunities and limitations to using these tools will allow instructors and students to critically think about how knowledge is created.

**What is ChatGPT?**

ChatGPT is a “chatbot” developed by a private company called OpenAI. Users can enter question prompts and within seconds ChatGPT will produce text-based responses in the form of poems, essays, articles, letters and more. It can also create structured responses like tables, bulleted lists and quizzes. ChatGPT can provide translation and copy language style and structure. It can also be used to develop and debug programming code. New and expanded uses continue to be developed and launched. A similar tool called DALL-E uses AI to create art pieces, and other AI tools have been created or are in rapid development to do even more – music, animation, multimedia video, powerpoints – and the list goes on.

**Are students using ChatGPT?**

Yes, students have already been exploring and using it to support completion of their coursework. Some courses explicitly encourage the use of ChatGPT for assignments. However, one concern is that students may be using ChatGPT to draft responses to homework responses without learning the material, and this presents the challenges and opportunities for reflection on teaching and learning at this time.

**Explore how ChatGPT works**

CEILS Education Research Talk with Jess Gregg. Demo portion starts just after the 5 minute mark.

Revolutionizing Education with ChatGPT: How AI is Transforming the Way We Learn

–

Learning

and Technology with Frank

Dig Deeper (article):

[How ChatGPT Work](https://towardsdatascience.com/how-chatgpt-works-the-models-behind-the-bot-1ce5fca96286)

[s](https://towardsdatascience.com/how-chatgpt-works-the-models-behind-the-bot-1ce5fca96286)

**Try out ChatGPT and Reflect: What are the opportunities?**

**Recommendation**

**Try out ChatGPT (or other AI tech) with your own course materials and assignments**

Instructors can sign up for a free account at [ChatGPT (or GPT-4) on Open AI’s platform.](https://chat.openai.com/) A first step in exploring the tool may be to enter some of your assignment prompts and assess the accuracy of the output. Then reflect on how you might embrace the tool or implement strategies that make use of the tool unnecessary.

PRO TIP: Including your students in the reflective process is also a learning opportunity to help them understand the benefits and limitations of the tool.

A few things to try:

* **Ask ChatGPT a question** – it could be a homework assignment or any question.
* How would you evaluate the response provided by ChatGPT?
* Try modifying the prompt and see how that changes the response.
* **Ask Chat GPT to synthesize text from large documents.** For example, enter a 3500 word paper as a prompt, and ask ChatGPT to create an 18 slide PowerPoint presentation, with headings and bullet points, making a persuasive case for action.
* P**rompt for writing samples specific to your area of expertise.** For example, ask ChatGPT to generate a nurse practitioner note for a 53 year old male with hypertension presenting with shortness of breath and dizziness. Another example is to ask for an email introducing your upcoming course to enrolled students.
* **Ask Chat GPT to translate something.** Together with your students examine the translation to see how well it did. Ask for improvements, or consider when such translation capabilities might be
* **If you teach students how to code**, ask ChatGPT to correct incorrect code (debug code). Consider ways this might help students who are learning to code in your course.
* **If you teach a writing intensive course,** try asking ChatGPT to respond to a writing prompt in a specific style (like a popular author) or create a poem on a specific topic.
* Reflect on the potential for ChatGPT to support student writing. Which writing skills do you feel are fundamental for students to do independent of artificial intelligence? How might ChatGPT facilitate the development of writing or problem solving skills?
* Have students use ChatGPT to write a draft and then have them edit what it produces or check for errors

*Note: If you do not wish to create an account, reach out to your local teaching support to set up a consultation and explore your assignments together. There are also many demonstrations of ChatGPT available online that you can search and watch.*

#### Ideas for Updating Your Course Activities with AI in Mind

**Adjusting assignments and activities**

* **Consider utilizing ChatGPT and other AI tools explicitly:** After experimenting with ChatGPT you may decide that you want to incorporate exercises where students are explicitly encouraged to interact with ChatGPT.

*Example:*

* Ask students to use ChatGPT and “fact check” the response provided by finding primary and secondary sources to back up the information provided.
* Ask students to generate a first draft using ChatGPT then keep track changes in a document to refine/edit.
* Reflecting upon prompt engineering – -use prompting logic used by students to generate information and then provide a different prompt to help guide revision. Showcase that small changes can lead to major differences in output!
* The University of Wisconsin, Madison provides some examples for [how to integrate AI into the writing process in your classroom](https://idc.ls.wisc.edu/guides/using-artificial-intelligence-in-the-classroom/)

*View this collaborative* [*Google Doc: AI Examples and Resources*](https://docs.google.com/document/d/1luwax_ps5tqRGBL4XWyr_Y_ab8h7Y7k6bsp4v1-pIRc/edit) *to see examples and resources curated by UCLA’s teaching and learning community.*

* **Adapt or create assignments that are not easily completed using AI:** Be more explicit about having students provide references for assignments, use a social annotation tool like [Hypothes.is](https://ucla.service-now.com/support?id=its_browse_service&sys_id=61c33e621b41999092f9a641b24bcb92&table=u_service_definitions) or [Perusall,](https://ucla.service-now.com/support?id=its_browse_service&sys_id=e041171a1b8219109ca4a602b24bcb72&table=u_service_definitions) utilize comments, Microsoft Word Track Changes or Google Docs Suggesting mode for individual or group annotation, have students complete written assignments in class, ask students to connect learning to their personal experiences and/or current events. Support students in developing oral communication skills by providing more opportunities for in-class presentations (or during discussion sections).
* Montclair State has created a guide that include [“Practical Suggestions to](https://www.montclair.edu/faculty-excellence/practical-responses-to-chat-gpt/)

[Mitigate Non-Learning/Cheating”](https://www.montclair.edu/faculty-excellence/practical-responses-to-chat-gpt/)

* **Explore** [**Gradescope**](https://ucla.service-now.com/support?id=its_browse_service&sys_id=038400d91b7c91109ca4a602b24bcbc2) **for implementing additional hand-written assignments:** UCLA has a campus-wide license for Gradescope that [integrates with Bruin Learn.](https://ucla.service-now.com/support?id=kb_article&sys_id=KB0013878) One feature of this tool is that students can upload photos or PDFs of written work, and the system easily allows for streamlined grading (by question, by page) and digital commenting and rubrics.
* You can [watch](https://youtu.be/IIy0SpIDRQo?t=1652) UCLA’s Will Conley, Department of Mathematics, provide a recorded overview of how to use Gradescope (approx. 56 mins). While we have now shifted to BruinLearn so the initial steps for linking to the gradebook will differ slightly, the within Gradescope interface is the same. [At the 27 minute mark](https://www.youtube.com/watch?v=IIy0SpIDRQo&t=1652s) you can view what it looks like to see handwritten work submitted to Gradescope for streamlined digital grading by instructors and TAs.
* **Require students make a connection to class discussions:** Prompt students to explicitly reference in-class discussions, lecture material and course readings in their homework assignments. (Example: Share three takeaways from our in-class discussion on the issues with how Covid-19 testing was implemented across the US at the onset of the pandemic.) **Communicate: Talk With Your Students About AI**

**Discuss the potential:** Many of our students will go on to become leaders at organizations that utilize and/or develop new AI technologies. How will these tools support advancements in your field (medicine, science, art, music, humanities, health, and more)?

**Prepare students for the future when they will work and interact with AI:** This technology is likely to develop and become embedded in many parts of our lives. Preparing students to thoughtfully engage with it, co-create with it and be curious about and know how to interact with other technological developments as they occur.

**Seize the opportunity to center the importance of critical thinking and digital literacy.** Students will have the opportunity in the future to break the cycle of spreading disinformation, lack of journalistic integrity in news, and elevating accurate and factual research and scholarship.

Emphasize the importance of digital literacy, research, and writing skills with students; connect students to library resources for research and writing. As educators, we have an obligation to help guide our students through many types of literacy, including digital media and AI literacy. [UCLA’s WI+RE](https://uclalibrary.github.io/research-tips/workshops/) has created the [Understanding Misinformation: A Lesson Plan Toolkit,](https://uclalibrary.github.io/research-tips/misinfo/) that is geared towards educators who want to prepare students to learn about misinformation. We can get students to [vet information like experts.](https://www.edutopia.org/article/what-fact-checkers-know-about-media-literacy-and-students-should-too)

**Lean in to talking to your students about how learning happens:** Learning happens when actively engaging with the course material, through conversations and dialogue leading to deepening conceptual understanding.

**Discuss Academic Integrity with Your Students:**

One of the main concerns instructors have expressed is how to uphold academic integrity and prevent the misuse of tools like ChatGPT (intentional or not). These concerns include:

* Plagiarism (copying and pasting the response that the tool provides; running material through multiple AI generators to avoid detection)
* Lack of proper citation of sources
* Inaccurate, misleading, biased, false, or limited information in responses to question prompts. While some AI detectors have been recently developed, it is unclear how effective they will be long-term and early reports indicate that individuals can easily avoid detection through simple modifications to produced text.

The [UCLA Student Conduct Code](https://deanofstudents.ucla.edu/individual-student-code#:~:text=102.01c%3A%20Plagiarism&text=Unless%20otherwise%20specified%20by%20the,must%20clearly%20acknowledge%20the%20source.) states, “Unless otherwise specified by the faculty member, all submissions, whether in draft or final form, to meet course requirements (including a paper, project, exam, computer program, oral presentation, or other work) must either be the Student’s own work, or must clearly acknowledge the source.” Unless an instructor indicates otherwise, the use of ChatGPT or other AI tools for course assignments is akin to receiving assistance from another person and raises the same concern that work is not the student’s own. Please communicate this to your students, and consider incorporating this language into your syllabus.

Teaching Assistants will be seeking guidance on how to discuss ChatGPT with students and what to do if they suspect submitted work may be AI generated. In addition to talking with your students, make space for conversations with your TAs and other instructional team members to explore this topic and co-construct guidelines.

***The ultimate decision and responsibility for how to teach about AI and the establishment of or revision of course policies related to its use lies with the instructor.***

View this collaborative Google Doc: [AI Examples and Resources](https://docs.google.com/document/d/1luwax_ps5tqRGBL4XWyr_Y_ab8h7Y7k6bsp4v1-pIRc/edit) to see examples (**including Syllabus**

**Language and Policy Language**) and resources curated by UCLA’s teaching and

## Discuss the ethical issues and limitations of AI

Facilitate discussions with your students on the impacts of spreading disinformation or biased information, lack of regulation of companies that develop these technologies, and other dangers. While students will likely still continue to use ChatGPT and other tools like it, it is crucial that our community has this shared understanding of both dangers and opportunities.

These technologies are not infallible and their accuracy is subject to a variety of factors, some listed below:

* Prone to filling in replies with incorrect data if there is not enough information available on a subject.
* Lack the ability to understand the context of a particular situation, which can result in inaccurate outputs.
* Large, uncurated datasets scraped from the internet are full of biased data that then informs the models.
* Data is collected from the past, it tends to have a regressive bias that fails to reflect the progress of social movements.

Our nation has yet to catch up to the regulation needed to prevent the potential for tremendous harm when false or biased information is taken as fact. Our community must continue to explore the value and innovation that can come from AI while simultaneously contributing to the dialog about these potential harms.

**Share current examples of scholarly discussion on this topic.**

As our own UCLA experts have shared with our instructional community during the recent [UCLA](https://www.youtube.com/watch?v=j2WVsxXKa0s)

[Virtual Town Hall: What is ChatGPT and How Does it Relate to UCLA’s Academic Mission,](https://www.youtube.com/watch?v=j2WVsxXKa0s) there are concerns about the ethics and practices around tools like ChatGPT. Algorithms can and do replicate and produce biased, racist, sexist, etc. outputs, along with incorrect and/or misleading information.

Additional Examples:

* [Educause – Special Report on Artificial Intelligence](https://er.educause.edu/-/media/files/articles/2022/6/er2225w.pdf?la=en&hash=5CDDC654009BB896FC1B15AE1B6753FF0C05C128)
* [Noam Chomsky: The False Implications of ChatGPT](https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html)
* [Schools Must Embrace the Looming Disruption of ChatGPT](https://www.the74million.org/article/schools-must-embrace-the-looming-disruption-of-chatgpt/)

Be proactive in discussing concerns around privacy and intellectual property that students may

have

**On Requiring the Use of ChatGPT:** Creating an account to use ChatGPT requires sharing of personal information. Depending on context, the use of ChatGPT may also mean sharing student intellectual property or student education records with ChatGPT under their terms and conditions of use. Individual students may have legitimate concerns and therefore may be unwilling to create an account. Discuss these concerns and consider alternatives.

* If you will be requiring use of ChatGPT, consider making this explicit in the syllabus (for a related example of considerations related to privacy, see [Privacy Tips for Your Syllabus)](https://cio.ucop.edu/privacy-tips-for-your-syllabus/).

**Protecting student privacy as required by FERPA**: Academic records, such as examinations and course assignments, are considered a student record and protected by FERPA. For example, ChatGPT should not be used to draft initial feedback on a student’s submitted essay that included their identifying information. Asking ChatGPT to respond to question prompts would not be a FERPA violation, as no student information is provided to ChatGPT.

ChatGPT is currently in the process of review through UCLA’s [Third Party Risk Management](https://ociso.ucla.edu/third-party-risk-management-tprm) to understand where there will be gaps in accessibility and security.

Ensure equity and accessibility concerns are

addressed

**As with any emerging technology, ChatGPT may not always be accessible by individuals with disabilities.** Open a conversation with the Center for Accessible Education (CAE) for ideas on exploring accessible alternatives. As always, we encourage faculty to use [this sample syllabus language](https://cae.ucla.edu/faculty-handbook/suggested-syllabus-statement) to direct a student toward CAE to discuss their options for accommodations and support.

Requests for support should be directed to caeintake@saonet.ucla.edu or the student’s listed Disability Specialist on their accommodation letter.

As the technology evolves, there may be a cost to using it, so continuing to revisit your learning goals and activities with respect to access is a critical equity issue.

### Examples and Resources

WEBSITE: [UCLA Online Teaching and Learning – Information on Chat GPT and AI](https://online.ucla.edu/chatgpt-ai/) – Resources from UCLA Online Teaching and Learning

CURATED RESOURCE LIST: [Generative AI Tools and Resources](https://docs.google.com/document/d/1G3JxQpsnyhD3QFbnrg0fnWbycpLzp0m_3D9xE_CUyeQ/edit) – From Dr. Kim DeBacco, Senior Instructional Designer UCLA Online Teaching and Learning

SUBSTACK POST: [“Carving out time to learn: A conversation with ChatGPT”](https://docs.google.com/document/d/1G3JxQpsnyhD3QFbnrg0fnWbycpLzp0m_3D9xE_CUyeQ/edit) – From Caroline

Kong, Instructional Designer and Technologist at the Center for the Advancement of Teaching

RECORDED WEBINAR: [“What’s All the Buzz About ChatGPT and AI in Higher Ed?”](https://youtu.be/kQG39meoaFw) RECORDED WEBINAR: [UCLA Virtual Town Hall: What is ChatGPT and How Does it Relate to UCLA’s Academic Mission](https://www.youtube.com/watch?v=j2WVsxXKa0s)

RECORDINGS AND RESOURCES: [AI in Action](https://sites.google.com/g.ucla.edu/aiinaction) – Events brought to you by UCLA’s Center for the Advancement of Teaching (CAT), the Center for Education, Innovation, and Learning in the Sciences (CEILS), the Excellence in Pedagogy and Innovative Classrooms program (EPIC), Online Teaching and Learning (OTL), the Bruin Learn Center of Excellence (CoE), the Writing

Programs, and Humanities Technology (HumTech)

UCLA NEWSROOM ARTICLE: [“Can AI and creativity coexist?”](https://newsroom.ucla.edu/stories/artificial-ingelligence-and-creativity-coexist-foster-snelson) – In a joint interview, UCLA professors Jacob Foster and Danny Snelson discuss how chatbots could be used in teaching, offer historic analogs for the current AI explosion and opine about whether technology is actually capable of creativity.

**Example Syllabi Language and Activity Ideas**

View this collaborative [Google Doc: AI Examples and Resources t](https://docs.google.com/document/d/1luwax_ps5tqRGBL4XWyr_Y_ab8h7Y7k6bsp4v1-pIRc/edit)o see examples and resources curated by UCLA’s teaching and learning community. This includes syllabus language, assignment ideas, and other strategies shared by instructors from UCLA and across the US.

**22.** [**Cornell Gudelines for artificial intelligence**](https://it.cornell.edu/ai)

Generative artificial intelligence (AI), offered through tools such as ChatGPT, Claude, Bard, Copilot (Microsoft, GitHub, etc.), and DALL-E, is a subset of AI that uses machine learning models to create new, original content, such as images, text, code, or music, based on patterns and structures learned from existing data.

### Guidelines

Cornell’s guidelines seek to balance the exciting new possibilities offered by these tools with awareness of their limitations and the need for rigorous attention to accuracy, intellectual property, security, privacy, and ethical issues. These guidelines are upheld by [existing university policies.](https://policy.cornell.edu/policy-library) A key to exploring AI tools centers on important choices about which tools we use and the privacy and protection of an individual’s personal information and institutional data. Free AI tools that are not offered by Cornell do not provide any material protection of data and should not be used to share or process institutional academic or administrative information.

**Accountability**

You are accountable for your work, regardless of the tools you use to produce it. When using generative AI tools, always verify the information for errors and biases and exercise caution to avoid copyright infringement. Generative AI excels at applying predictions and patterns to create

|  |  |  |
| --- | --- | --- |
|  | new content, but since it cannot understand what it produces, the results are sometimes misleading, outdated, or false.  **Confidentiality and Privacy**  If you are using public generative AI tools, you cannot enter any Cornell information, or another person's information, that is confidential, proprietary, subject to federal or state regulations, or otherwise considered sensitive or restricted. Any information you provide to public generative AI tools is considered public and may be stored and used by anyone else.  As noted in the [University Privacy Statement,](https://privacy.cornell.edu/privacy-cornell/university-privacy-statement) Cornell strives to honor the [Privacy Principles:](https://privacy.cornell.edu/privacy-cornell/privacy-principles-cornell) Notice, Choice, Accountability for Onward Transfer, Security, Data Integrity and Purpose Limitation, Access, and Recourse.  **Use for Education and Pedagogy**  Cornell encourages a flexible framework in which faculty and instructors can choose to prohibit, to allow with attribution, or to encourage generative AI use. In addition to the [CU Committee Report: Generative Artificial Intelligence for Education and Pedagogy](https://teaching.cornell.edu/generative-artificial-intelligence/cu-committee-report-generative-artificial-intelligence-education) delivered in July 2023 and [resources from the Center for Teaching Innovation,](https://teaching.cornell.edu/generative-artificial-intelligence) check with your college, department, or instructor for specific guidance.  **Use for Research**  The widespread availability of generative AI tools offers new opportunities of creativity and efficiency and, as with any new tool, depends on humans for responsible and ethical deployment in research and society. The [Cornell University Task Force Report, Generative AI in Academic Research: Perspectives and Cultural Norms](https://research-and-innovation.cornell.edu/generative-ai-in-academic-research/) (December 2023), offers perspectives and practical guidelines to the Cornell community on the use of generative AI in the practice and dissemination of academic research.  **Use for Administration and Other Purposes**  By the end of 2023, Cornell is aiming to offer or recommend a set of generative AI tools that will meet the needs of staff doing administrative work, while providing sufficient risk, security, and privacy protections. The use of generative AI for administration purposes must comply with the guidelines of the [Cornell Generative AI in Administration Task Force Report](https://it.cornell.edu/sites/default/files/itc-drupal10-files/Cornell%20Generative%20AI%20in%20Administration%20Task%20Force%20Report-Jan2024_1.pdf) (January 2024). |  |
| **Cornell AI Services and Pilots** | | |
|  | Cornell researchers and other community members have long been involved in AI and machine learning in their scholarly, teaching, and technical work. With the rapid growth of generative AI tools, the Cornell community has more opportunities to enhance our understanding, practice, policy, and technology to use AI in new ways.  Cornell faculty, research, and administrative communities across Ithaca, New York City, and Doha have worked diligently to explore the emerging AI needs, concerns, and potential use cases and lay the groundwork for a deeper understanding of Cornell's near-term institutional needs and potential responses. The resulting university-wide reports are informing AI explorations to meet needs in the teaching and learning, research, and administration spaces.  Throughout spring and summer 2024, we will be piloting a range of services in order to determine where to best invest Cornell’s time and talent to meet the university’s evolving needs.  For those seeking to purchase generative AI tools or subscriptions outside of those plans, the [IT Statement of Need process](https://it.cornell.edu/itgovernance/checklist-purchasing-it-applications-software-and-services) is required. |  |
| **Recommended Resources** | | |
|  | [Cornell AI Initiative](https://ai.cornell.edu/)  [Cornell University AI for Science Institute](https://science.ai.cornell.edu/)  [Center for Teaching Innovation: Generative AI](https://teaching.cornell.edu/generative-artificial-intelligence)  **University Task Force Reports**  [CU Committee Report: Generative Artificial Intelligence for Education and Pedagogy](https://teaching.cornell.edu/generative-artificial-intelligence/cu-committee-report-generative-artificial-intelligence-education) (June 2023)  [Cornell University Task Force Report - Generative AI in Academic Research: Perspectives and Cultural Norms](https://research-and-innovation.cornell.edu/generative-ai-in-academic-research/) (December 2023)  [Cornell Generative AI in Administration Task Force Report](https://it.cornell.edu/sites/default/files/itc-drupal10-files/Cornell%20Generative%20AI%20in%20Administration%20Task%20Force%20Report-Jan2024.pdf) (January 2024) **AI Questions, Concerns, Needs, or Ideas?**  [Contact AI at Cornell](https://it.cornell.edu/ai-question-feedback)  **Comments?**  To share feedback about this page or request support, [log in](https://it.cornell.edu/user/login?destination=/node/10263) with your NetID |  |

**23.** [**Generative Artificial Intelligence in the classroom**](https://www.viceprovostundergrad.utoronto.ca/strategic-priorities/digital-learning/special-initiative-artificial-intelligence/) The latest generation of Artificial Intelligence (AI) systems will impact teaching and learning in many ways, presenting both opportunities and challenges for the ways our course instructors and students engage in learning. At the University of Toronto, we remain committed to providing students with transformative learning experiences and to supporting instructors as they adapt their pedagogy in response to this emerging technology.

While many generative AI systems have recently become available, [~~ChatGPT~~](https://chat.openai.com/auth/login) is currently the most prominent, garnering worldwide media attention. This is an AI tool that uses predictive technology to create or revise written products of all kinds, including essays, computer code, lesson plans, poems, reports, and letters. The products that the tool creates are generally of good quality, although they can have inaccuracies. We encourage you to try the system to test its capabilities and limitations.

In this FAQ, ChatGPT refers to the free, online AI chat system that utilizes the OpenAI GPT technology. Please note that this is only one of a variety of generative AI tools currently available.

### Sample Syllabus Statements

April 2023: The University has created sample statements for instructors to include in course syllabi and course assignments to help shape the message to students about what AI technology is, or is not, allowed. These statements may be used for both graduate and undergraduate level courses.

### SGS Guidelines

July 2023: The School of Graduate Studies (SGS) announced new [Guidance on the Appropriate Use of Generative Artificial Intelligence in Graduate Theses](https://www.sgs.utoronto.ca/about/guidance-on-the-use-of-generative-artificial-intelligence/) which will be of interest to graduate students, supervisors, supervisory committee members, Graduate Chairs and Graduate Units.

### Copyright Considerations

September 2023: There remains significant legal uncertainty concerning the use of generative AI tools in regard to copyright. This is an evolving area, and our understanding will develop as new policies, regulations, and case law become settled. Some of the concerns surrounding generative AI and copyright include:

* Input: The legality of the content used to train AI models is unknown in some cases. There are a number of lawsuits originating from the US that allege Generative AI tools infringe on copyright and it remains unclear if and how the fair use doctrine can be applied. In Canada, there also remains uncertainty regarding the extent to which existing exceptions in the copyright framework, such as fair dealing, apply to this activity.
* Output: Authorship and ownership of works created by AI is unclear. Traditionally, Canadian law has indicated that an author must be a natural person (human) who exercises skill and judgement in the creation of a work. As there are likely to be varying degrees of human input in generated content, it is unclear in Canada how it will be determined who the appropriate author and owner of works are. More recently, the US Copyright Office has published the following guide addressing these issues: [Copyright Registration Guidance for Works Containing AI-Generated Materials.](https://copyright.gov/ai/ai_policy_guidance.pdf)

If you have further questions about copyright, please view the U of T Libraries webpage, [Generative AI tools and Copyright Considerations](https://onesearch.library.utoronto.ca/copyright/generative-ai-tools-and-copyright-considerations) for the latest information.

If you are an instructor who is interested in in using generative AI to develop course materials, review the FAQ below for considerations.

### Frequently Asked Questions

How can I test out ChatGPT to see its capability?

Instructors are welcome and encouraged to test ChatGPT, use of which is currently free upon registration. You can also test other similar AI tools to assess their capability, for instance to see if they can respond to the assignments used in your courses, or the way in which they improve the readability and grammar of a paragraph. Experimentation is also useful to assess the limits of the tool. However, confidential information should never be entered into an AI tool such as ChatGPT.

All content entered may become part of the tool’s dataset and may inadvertently resurface in response to other prompts.

Please note that due to high demand, access to ChatGPT is, at times, unavailable.

Is ChatGPT accurate and reliable?

Large Language Models, like ChatGPT, are trained to predict the next word in a sentence, given the text that has already been written. Early attempts at addressing this task (such as the next-word prediction on a smartphone keyboard) are only coherent within a few words, but as the sentence continues, these earlier systems quickly digress. A major innovation of models such as GPT is their ability to pay attention to words and phrases which were written much earlier in the text, allowing them to maintain context for much longer and in a sense remember the topic of conversation. This capacity is combined with a training phase that involves looking at billions of pages of text. As a result, models like ChatGPT, and its underlying technology GPT-3 (and now, GPT-4), are good at predicting what words are most likely to come next in a sentence, which results in generally coherent text.

One area where generative AI tools sometimes struggle is in stating facts or quotations accurately. This means that models like GPT-4 sometimes generate claims that sound real, but to an expert are clearly wrong.

A related area where ChatGPT seems to struggle is in the discussion of any event or concept that has received relatively little attention in online discourse. To assess these limitations, you could try asking the system to generate your biography. Unless there are numerous accurate biographies of yourself online, ChatGPT is unlikely to generate a comprehensively correct biography.

What are the ethical considerations regarding the use of generative AI systems?

This is a threshold question that instructors may want to consider. Mainstream media has been covering this issue extensively, and alternate viewpoints are widely available.

Given that generative AI systems are trained on materials that are available online, it is possible that they will repeat biases present online. OpenAI has invested substantial effort into addressing this problem, but it remains a danger with these types of systems. You may also want to familiarize yourself regarding questions about the way the technology was developed and trained (e.g., who were the people who trained it?), the way we use the responses it provides, and the long-term impacts of these technologies on the world.

The Provost is consulting with faculty and staff experts on these larger questions involving ChatGPT and other generative AI systems, and welcomes debate and discussion on these issues.

As an instructor, can I use generative AI to generate content for my courses?

The question of copyright ownership remains one of the biggest unknowns when using generative AI tools. The ownership of outputs produced by generative AI is unsettled in law at the current time. If, as an instructor, you would like to use generative AI tools for content generation in your course, consider the following before doing so:

Have an understanding that while you can use these tools to create content, you may not own or hold copyright in the works generated.

Be mindful of what you input into tools: never input confidential information or intellectual property you do not have the rights or permissions to (e.g., do not submit student work or questions without their permission). All content entered may become part of the tool’s dataset and may inadvertently resurface in response to other prompts.

Review the terms of service of each tool, which will establish terms of use and ownership of inputs and outputs (for example, view the Terms of Use for OpenAI). Note that terms of use are subject to change without notice.

Be explicit in how you have used these tools in the creation of your work.

View the U of T Libraries, Generative AI tools and Copyright Considerations for more information.

Can I use generative AI tools for pedagogical purposes in my classroom?

Yes. Some instructors may wish to use the technology to demonstrate how it can be used productively, or what its limitations are. The U of T Teaching Centres have developed more information and advice about how you might use generative AI as part of your learning experience design.

**24.** [**Generative AI – Academic Integrity at UBC**](https://www.google.com/search?q=university+of+british+columbia+policy+on+artificial+intelligence&sca_esv=8ae9a4f0a7e64b77&rlz=1C1UEAD_enUS1024US1024&ei=gbjKZfTtDNKnmtkPj9uyiAk&ved=0ahUKEwi0ssauiKeEAxXSkyYFHY-tDJEQ4dUDCBA&uact=5&oq=university+of+british+columbia+policy+on+artificial+intelligence&gs_lp=Egxnd3Mtd2l6LXNlcnAiQHVuaXZlcnNpdHkgb2YgYnJpdGlzaCBjb2x1bWJpYSBwb2xpY3kgb24gYXJ0aWZpY2lhbCBpbnRlbGxpZ2VuY2UyCBAhGKABGMMEMggQIRigARjDBDIIECEYoAEYwwQyCBAhGKABGMMEMggQIRigARjDBEiEJVAAWK8kcAJ4AJABAJgBaqAB0wOqAQMzLjK4AQPIAQD4AQHiAwQYACBB&sclient=gws-wiz-serp)

**Thinking about ChatGPT?**

Conversations around the impacts of artificial intelligence (AI) tools are ongoing as their capabilities continue to evolve. AI tools have the potential to change the way we teach, learn and work at UBC.

This list brings together important things to know about ChatGPT and generative artificial intelligence in the classroom for instructors and students at UBC. Generative AI technology is evolving quickly and this list will be updated as new developments arise. If you have a question that is not answered here, we invite you to share it through the website [feedback form.](https://academicintegrity.ubc.ca/contact/)

**What is ChatGPT and what is generative artificial intelligence?**

ChatGPT (GPT stands for Generative Pre-trained Transformer) is a tool developed by OpenAI that is capable of producing human-like responses to prompts. This AI system is a large language model that has been trained on a dataset to interact with users in a conversational way. The current version of ChatGPT bases its output on its training data of internet content up to September 2021. While ChatGPT is widely known, it is part of the broader category of generative AI, a form of artificial intelligence that generates new content based on the data it has been trained on. The new content that is generated can be text, images, code, videos, etc. Other examples of generative artificial intelligence tools (“AI tools”) include Bing, Bard and Dalle-2.

Some generative AI tools are available for free, while others such as ChatGPT Plus have a paid subscription version with further features. Some tools are connected and others offer plugins to connect to the internet and others are not connected and are limited to the information they are trained on which may not be up to the present day.

**Can instructors test AI tools to see what they can do, and what they cannot?**

Instructors are welcome to test AI tools to assess their capabilities and limitations and experiment with how the tools respond to particular course assignments or prompts. For example, by using prompts from their own assignments and assessments, instructors can gain a sense of how the tool could potentially be used by learners as well as its limitations. Keep in mind that some students may be sophisticated in prompt engineering and could be able to prime GenAI tools to return results that are superior to those obtained by entering an assignment question into a fresh chat.

Additionally, in trying out tools, be sure not to share in prompts any personal or sensitive information, or any information you wish to be kept private, as this content may become part of the dataset the models train on.

**Is the use of AI tools considered to be academic misconduct at UBC?**

The use of ChatGPT or other generative AI tools does not automatically equate to academic misconduct at UBC. At this time, whether the use of AI tools in courses is or is not allowed is a course or program-level decision and there is no overall, UBC-wide ban on its use in teaching and learning. Individual instructors should clarify expectations with their students early in the term, such as on the syllabus. If instructors have questions about any Department of program level policies on artificial intelligence tools, they should reach out to their Department or program. Further information is available on [generative AI syllabus language.](https://academicintegrity.ubc.ca/generative-ai-syllabus/)

* If using ChatGPT and/or generative AI tools on coursework has been **prohibited** by the instructor, then using these tools would be considered to be academic misconduct.
* If using ChatGPT and/or generative AI tools has been **permitted** by the instructor, then instructors should make sure to convey the limitations of use and how it should be acknowledged and use should stay within those bounds.
* If the use of ChatGPT and/or generative AI tools has **not been discussed or specified** by the instructor, then it is likely to be considered as prohibited as an example of the “use or facilitation of unauthorized means to complete an examination or coursework” and more specifically as “accessing websites or other online resources not specifically permitted by the instructor or examiner” (Discipline for Academic Misconduct, [**Vancouver**](https://www.calendar.ubc.ca/vancouver/index.cfm?tree=3,54,111,0) and [**Okanagan**](https://www.calendar.ubc.ca/okanagan/index.cfm?tree=3,54,111,1547) 3.1.b.iv), and potentially plagiarism (3.1.e).

Students should not assume that all available technologies are permitted. If students are not sure about whether AI tools are allowed, as with any tool, they must ask their instructor for clarity and guidance.

UBC’s Academic Calendar provides guidance on what is considered academic misconduct (Academic Misconduct by UBC Students, [**Vancouver**](https://www.calendar.ubc.ca/vancouver/index.cfm?tree=3,54,111,1745) and [**Okanagan**)](https://www.calendar.ubc.ca/okanagan/index.cfm?tree=3,54,111,1547). Academic misconduct is any conduct by which a student gains or attempts to gain an unfair academic advantage or benefit, thereby compromising the integrity of the academic process. Artificial intelligence tools like

ChatGPT are not expressly named in the academic misconduct regulation but their use could be considered an attempt to gain an unfair academic advantage, as well constituting unauthorized means to complete an assignment or assessment, the accessing of a website that is not permitted, or other, depending on the specific case.

It is important for instructors to address academic integrity throughout the semester. Create opportunities to discuss expectations around academic integrity ahead of assignments and exams and throughout the term. Refer to [**academicintegrity.ubc.ca**](https://academicintegrity.ubc.ca/) fo[r **tools to teach academic integrity and respond to academic misconduct.**](https://academicintegrity.ubc.ca/faculty-start/)

**How can instructors use AI tools in their classroom?**

There are many innovative and creative ways that an instructor might use generative artificial intelligence technologies in their courses. From open conversations around the ethical and societal implications of artificial intelligence to dialogues with chatbots and critical engagement with their outputs, there is potential both for learning about and for learning to use the technology. If instructors choose to integrate ChatGPT or other GenAI tools into course activities, they should consider the privacy implications of doing so. Using ChatGPT, for example, requires a login that asks for personal information, and it is important to offer students an alternative option if they do not wish to provide this information to engage with the tool themselves. One option could be for the instructor to generate conversations and share them with students to engage with outside of the tool.

Suggestions and examples around how generative AI can be integrated into teaching and learning, as well as how to design assignments and assessments that make it harder to use AI, have been developed by the CTLT. Instructors may wish to review the resource from the CTLT [Teaching and Learning in an Era of Generative AI](https://ctlt.ubc.ca/resources/assessment-design-in-an-era-of-generative-ai/?login) or [AI in Teaching and Learning.](https://ai.ctlt.ubc.ca/)

**Can students use AI tools to complete assignments?**

A key expectation of academic integrity for students is completing their own work. Besides producing essays in seconds, generative AI has proven itself capable of completing multiplechoice exams and short answer questions, generating code, and producing creative output.

While artificial intelligence technologies should not be used to complete academic work, there may be times where instructors and students engage with it as part of student learning. If this is the case, instructors should provide clear guidance to students around how they are allowed to engage with the tools.

**Can or should instructors use AI detectors to detect the use of AI tools on assignments or assessments?**

UBC discourages the use of artificial intelligence detectors on student work, and is not at this time

planning to purchase or support any such tools at the institutional level.

There are several AI detectors currently in existence, such as GPTZero, Turnitin, and AI Content Detector. Despite the availability of such tools, it is important to remember that they might not be fully tested and that the technology to potentially outwit them continues to evolve. The detectors are not foolproof and can produce false negatives and false positives. It may also be possible for the user to modify content to avoid detection. Instructors might wish to consider [UBC’s response](https://lthub.ubc.ca/2023/04/04/ubc-not-enabling-turnitins-ai-detection/) to Turnitin’s AI detection feature and the concerns that were raised.

If instructors still choose to use AI detectors, they should be aware and understand their limitations and issues they can raise. None of the detectors has undergone a UBC Privacy Impact Assessment, and as such there may be privacy and security concerns with submitting student work to them, particularly without their knowledge or consent. Instructors should not use these tools to evaluate any student work that contains the name of the student or any other personal information of the student or third parties. If student work may be submitted through one or more AI detectors, instructors should be transparent with students and let them know at the beginning of the term, such as in the syllabus.

Finally, it is recommended that AI detection tools not be used as the sole factor in decision-making around an allegation of academic misconduct. If an instructor suspects that an assignment or assessment has been completed with unauthorized use of AI tools, they should proceed as they would for any other potential allegation of academic misconduct. An [overview of the academic misconduct process for instructors](https://academicintegrity.ubc.ca/regulation-process/faculty/) is available on the academic integrity website.

**Can I use Turnitin’s new AI-detection feature in my courses?**

No, currently Turnitin’s new AI-detection functionality is not available for use in any UBC course.

[**UBC has reaffirmed their decision to not enable Turnitin’s new AI-detection feature**.](https://lthub.ubc.ca/2023/08/28/ubc-affirms-decision-to-not-enable-turnitin-ai-detection/)

On April 4 2023, Turnitin activated [**a new and separate feature**](https://www.turnitin.com/solutions/ai-writing) that attempts to identify AIgenerated text . The LT Hub Leadership group, with the support of the Provosts at both UBC Vancouver and UBC Okanagan, made the decision not to enable this feature at that time ([April 2023)](https://lthub.ubc.ca/2023/04/04/ubc-not-enabling-turnitins-ai-detection/) and has recently reaffirmed their decision ([August 2023)](https://lthub.ubc.ca/2023/08/28/ubc-affirms-decision-to-not-enable-turnitins-ai-detection/) for the following reasons:

**Effectiveness of the feature is still unclear**

* Testing for accuracy in the AI-detection feature is in early stages.
* Testing for potential bias in the feature continues to be in early stages.
* Ability of the feature to keep up with rapidly evolving AI is unknown.

**It is not possible to double-check or review the results**

* Instructors cannot double-check the feature results.
* Results from the feature are not available for students to review.

You can read more about UBC’s [**decision to not enable Turnitin’s new AI-detection**](https://lthub.ubc.ca/2023/08/28/ubc-affirms-decision-to-not-enable-turnitins-ai-detection/) feature on

UBC’s Learning Technology Hub website.

UBC is continuing to wait before deploying any AI-detection features, including the one by Turnitin, as more information is needed about the feature’s effectiveness, accuracy and bias mitigation, and about the ability to employ an AI detector as a robust component of talking to students about suspected academic misconduct. The use of other AI detection tools is also not recommended, due to similar issues as those noted above, as well as privacy and security concerns; to date, no AI detection tool has undergone a UBC Privacy Impact Assessment process.

**What should I do if I suspect the unauthorized use of AI tools?**

Instructors who suspect that a student has used generative artificial intelligence tools contrary to expectations should follow the [standard academic misconduct process.](https://academicintegrity.ubc.ca/regulation-process/faculty/) If an instructor has a suspicion based on the student’s work, they should follow the procedure as they would for any misconduct allegation. Instructors should not rely on AI detectors to form the basis of an allegation of academic misconduct . If students have concerns about any allegations of academic misconduct against them, they can reach out to the [Office of the Ombudsperson for Students,](https://ombudsoffice.ubc.ca/contact-us/) [AMS Advocacy](https://www.ams.ubc.ca/support-services/student-services/advocacy/) or [SUO Advocacy.](https://www.suo.ca/suo-advocacy-office/)

**What are the recommendations for citing content developed by Generative AI if its use is permitted?**

There has been much conversation about how generative AI should be cited if its use is permitted in assignments and academic publications. The [**American Psychological Association**](https://apastyle.apa.org/blog/how-to-cite-chatgpt) (APA), [**Modern Language Association**](https://style.mla.org/citing-generative-ai/) (MLA) and the [**Chicago Manual of Style**](https://www.chicagomanualofstyle.org/qanda/data/faq/topics/Documentation/faq0422.html) have all provided recommendations in this area. Further information on this topic can be found in the [Generative AI and ChatGPT LibGuide](https://guides.library.ubc.ca/GenAI) from the UBC Library.

**Is there funding available for teaching and learning projects on generative AI?**

Funding is available for projects that explore the intersection of generative artificial intelligence and teaching and learning, as well as other teaching and learning topics ([**Teaching and Learning Enhancement Fund**,](https://tlef.ubc.ca/) [**Aspire 2040 Learning Transformations Fund,**](https://alt-2040.ok.ubc.ca/about/) [**SOTL Seed Program**)](https://isotl.ctlt.ubc.ca/services/sotl-seed-program/).

Explore teaching and learning topics and further support through the [**CTL**](https://ctl.ok.ubc.ca/) and the [**CTLT.**](https://ctlt.ubc.ca/)

1. [**Principles on Generative AI in Teaching and Learning at McGill**](https://www.mcgill.ca/provost/files/provost/principles_on_generative_ai_in_teaching_and_learning_at_mcgill.pdf)

Generative Artificial Intelligence (AI) tools have broad applications across society, including for institutions of higher education. The Subcommittee on Teaching and Learning (STL) of McGill’s Academic Policy Committee (APC) established a working group in January of 2023 to develop an approach regarding Generative AI and its impacts on Teaching and Learning. Their final report, available online, was discussed at APC on 26 October 2023, and the recommendations were endorsed by McGill’s Senate on 15 November 2023. The five principles emerging from the report are reproduced below. These provide a framework for ongoing conversations about Generative AI at the University in the context of teaching and learning, and can be used as a guide for instructors, students, and staff, and also for Faculties when they consider their own internal guidelines. Some of the principles contain within them areas of ongoing focus and attention, notably around the importance of education and awareness. Members of the community can also consult the Teaching and Learning Knowledge base for additional information about Generative AI

1. [**AI-Squared – Artificial Intelligence and Academic Integrity**](https://www.ualberta.ca/centre-for-teaching-and-learning/teaching-toolkit/teaching-in-the-context-of-ai/artificial-intelligence-academic-integrity.html#:~:text=Let%20students%20know%20that%20although,proper%20attribution%2C%20instructors%20can%20consider)

At the outset, a number of universities around the world elected to ban the student use of ChatGPT (as well as other Generative AI Tools) in academic work based on the news that ChatGPT was capable of generating outputs of unexpected quality and sophistication. Academic integrity concerns, both real and imagined, were raised over how students might use ChatGPT inappropriately in their academic assessments (Shiri 2023). The majority of these concerns are premised on the view that Generative AI-tools have the potential to do more harm than good, and, that, when wielded by uninformed students, the technology’s impact ranges from quick ‘workaround’ to weapon (Sawahel 2023, Weisman 2023). **To AI or Not to AI?**

While some individuals and institutions may view Generative AI with suspicion and as a threat to higher education, there are others who want to weigh the options from more nuanced perspectives. Many universities, for example, have put together task forces to make teaching in the context of AI a priority (Baucom, nd.; Black, 2023; Fox, 2023). When used ethically and in pedagogically sound ways, AI-tools can offer academics the chance to reconsider and reimagine an educational focus, not on deliverables and summative end-products (such as written assignments and standard exams) as measures of learning, but instead on process-driven and evaluated assessment. Stated another way, learning is not only about the product; learning is also about the process of acquiring new knowledge or learning ways to think and reason. This gives the instructor a window through which to focus on what students are ‘doing’ in their classes to develop the requisite disciplinary knowledge and allied critical-thinking abilities. While tools like ChatGPT are prone to fabrication (factually inaccurate outputs) and generating biases found in its data lakes, they can also help to deepen student engagement and enhance teaching, learning, and assessment (Mollick, 2023). Importantly, these technologies should be recognized as potential tools enabling increased accessibility to learning that will support a wider diversity of student needs than previously possible.

Therein lies another important consideration: AI and its various tools are well on their way to becoming omnipresent in our lives. Learning to adapt to AIs presence in our academic spaces is part of teaching today’s learners. To simply ban or avoid AI is to avoid the reality that U of A students are engaging and experimenting with (i.e. yes, like right now, as you read this) with and experiment with AI tools in our courses.

In order to do so, instructors are encouraged to experiment with different approaches, and to find ways to adapt and improve teaching and assessment to embrace the new reality of working and studying in a world where these emerging technologies are freely and widely available.

Guidance - University of Alberta’s Provost’s Task Force on Artificial Intelligence and the

Learning Environment

The suggestions provided in this section are in alignment with the initial [Guidance](https://www.ualberta.ca/the-quad/2023/03/generative-ai-in-the-learning-environment.html) proposed by the University of Alberta’s Provost’s Task Force on Artificial Intelligence and the Learning Environment (March 2, 2023).

Given the diversity of learning environments across our campuses, the general guidance that we can give includes the following:

1. Have conversations with your students about your expectations regarding the use of Generative AI, particularly in your course assignments. If students are using Generative AI, how would you like them to indicate that to you (e.g. in the sources cited page, methodology section, prefatory comments, or in-text citation)? Please make sure that you also summarize these conversations in a written format and include them in eClass in a place where students will find them for those who may not have been in class. This also gives students a place to refer back to when completing assignments. Your Department or Faculty may also have specific guidance for you.
2. Identify creative uses for Generative AI in your course (idea generation; code samples; creative application of course concepts; study assistance; language practice). Discuss the limitations of tools like ChatGPT in the topics covered by your course, including the limitation of data used (prior to 2021), factually inaccurate information, biases and discrimination in the data used to generate text and in the output, and the use of culturally inappropriate language and sources.
3. Remind students that the Code of Student Behaviour states: “No Student shall represent another’s substantial editorial or compositional assistance on an assignment as the

Student’s own work.” Submitting work created by generative AI and not indicating such would constitute cheating as defined above.

1. Stress to students the value of building their own voice, writing skills, and so on. Motivating students to share their ideas, perspectives, and voice may make generative AI less appealing. Similarly, asking students to share their reflections (reflective writing) can help reinforce student investment in the learning process. If instructors are equipped to do so, they can even show how generative AI can be used as a tool to aid in work as opposed to replacing student work.
2. Remind students that AI tools such as ChatGPT gather significant personal data from users to share with third parties.

**In order to help you think through the various options available to you, CTL suggests you start with the following questions as part of your decision-making process:**

1. What are your discipline's conventions and assumptions? How might students use AI to support their academic work in your discipline?
2. What role, if any, do AI driven technologies in the course/classroom play in your personal teaching philosophy?
3. Is assessment task redesign needed? How significant is this redesign and development? How do the new assessments fit and align with the course learning outcomes?
4. What do you want your students to know about your expectations regarding AI and academic integrity?
5. Which University resources would you like to direct your students to for further guidance if necessary?
6. What kind of classroom environment would your students like to see? How might you include them in the conversation about AI use in academic work?

Where did you land with your responses to the above questions? Are you leaning towards experimenting with/integrating AI tools in your teaching? Or, do you plan not to use AI in your classroom but still allow students to use it for specific purposes in the course of their learning? Whatever your final decision, it is important that you be transparent and share this information with your students. If you are a YES to AI use:

Stress to your students that if they use generative AI in their academic work, it is important they do so honestly, transparently, and according to the expectations you set for them. The substance of these conversations should match the language of expectations spelled out in any reference documentation you provide to students, such as the course syllabus or a [Statement of Expectations for AI Use.](https://www.ualberta.ca/centre-for-teaching-and-learning/teaching-toolkit/teaching-in-the-context-of-ai/statements-of-expectations.html) Another issue for you to consider, it is very possible that you may encounter students in your courses who do not want to use AI tools. In such cases – and assuming AI is not integral to course contents and learning outcomes – instructors can offer alternative assessment tasks.

If you are a NO to AI use:

Let students know that although the University of Alberta's most recent (November 2022) [Code of Student Behaviour](https://www.ualberta.ca/governance/media-library/documents/resources/policies-standards-and-codes-of-conduct/cosb-updated-december-05-2022.pdf) does not explicitly reference Generative AI and its use, if a student submits academic work, including text, images, code, and designs, generated by AI without proper attribution, instructors can consider this an act of plagiarism under the Code, which states: “No Student shall represent another’s substantial editorial or compositional assistance on an assignment as the Student’s own work. If you are a NO, it is important for you to consider that not every use of AI tools by students may qualify as cheating – students might use the tools in ways that support and deepen their learning.

**Tenets of Postplagiarism: Writing in the Age of Artificial Intelligence**

Responding to concerns about AI and writing, plagiarism, and academic integrity, Sarah E. Eaton contemplates a future in which we enter an era marked for “postplagiarism.” This will be a time when archaic print-based notions of copyright are set aside to make way for human-AI partnerships and new definitions of authorship and originality.

Although Eaton first put forward these ideas in her book, [*Plagiarism in Higher Education: Tackling Tough Topics in Academic Integrity (2021)*,](https://www.abc-clio.com/products/a6237p/) she has recently revised her thinking and put forward [*6 Tenets of Postplagiarism: Writing in the Age of Artificial Intelligence*.](https://drsaraheaton.wordpress.com/2023/02/25/6-tenets-of-postplagiarism-writing-in-the-age-of-artificial-intelligence/)

Eaton identifies six notions that will likely come to characterize the age of postplagiarism:

1. Hybrid Human-AI Writing Will Become Normal
2. Human Creativity is Enhanced
3. Language Barriers Disappear
4. Humans can Relinquish Control, but not Responsibility
5. Attribution Remains Important
6. Historical Definitions of Plagiarism No Longer Apply

*For more information, please also see her article,* [*“Artificial intelligence and academic integrity, post-plagiarism”*](https://www.universityworldnews.com/post.php?story=20230228133041549) *(2023)*.

### Dialogue with students

Begin with a conversation, in-person or synchronously, if possible, so you have the best opportunity to openly dialogue about your expectations and gauge your students’ responses. Talk to your students. The purpose of this initial dialogue is to share with them your expectations, and explore together in two-way conversation the possibilities and limitations of using Generative AI tools in the context of your course(s) and their academic work in your discipline. Speak to them about the academic integrity concerns that have been raised at the U of A and elsewhere in higher education. Where appropriate, encourage your students to ask questions, provide inputs, and offer suggestions. You might be surprised to discover instructors sometimes need to explore AI basics with their students too. Not all students will be up to date on the AI.

**A few key questions to guide your conversation with your students include:**

1. What do you know about artificial intelligence and AI tools such as GPT-4, Midjourney, and Microsoft’s (GPT-powered) search engine, Bing?
2. Have you used any of them before? Why?
3. Have you used an AI tool for learning (specifically, in your academic work)?
4. If so, how did you use them?
5. How do you think you can ethically use AI tools to support your learning? (Adapted from, Eaton, 2023)

This conversation is a great opportunity for you to discuss (November 2022) [Code of Student Behaviour](https://www.ualberta.ca/governance/media-library/documents/resources/policies-standards-and-codes-of-conduct/cosb-updated-december-05-2022.pdf) and [Academic Misconduct](https://www.ualberta.ca/dean-of-students/policies/student-conduct-and-accountability/index.html) with students, so, together, you can all consider the ethical implications and responsibilities.

**If you plan to integrate AI into your in-person or hybrid courses, here are a couple of options you can use to continue the conversation:**

1. Create an AI-based Discussion forum to share in (and monitor) your students’ experiences and conversations about their use of AI tools.
2. Create a Journal activity, and request that your students transparently track and reflect on their use of AI-tools as part of their learning process during your class.

**27.** [**Artificial Intelligence and ChatGPT – Academic Integrity**](https://uwaterloo.ca/academic-integrity/artificial-intelligence-and-chatgpt)

Quick Links

● [Instructor resources for Artificial Intelligence & ChatGPT tools](https://uwaterloo.ca/academic-integrity/artificial-intelligence-and-chatgpt#quicklink1) ● [Tips for students](https://uwaterloo.ca/academic-integrity/artificial-intelligence-and-chatgpt#quicklink2)

Artificial intelligence is everywhere and with it comes new opportunities and challenges.

As these new platforms evolve so must the way instructors, TAs and students use them. Given the rapid pace of change and growth in artificial intelligence - this issue will be ongoing for the foreseeable future.

At present, **it is important for instructors to be explicit about whether artificial intelligence or tools like ChatGPT are allowed to be used to complete assignments, tests or exams, and if so, the extent to which it is allowed, and if it should be cited and how to cite it**. A student who does not comply with the instructors rules about the use of such tools will be subject to Policy 71 and an investigation into academic misconduct.

A memo from the Associate Vice-President, Academic about Chat GPT (and other AI tools) can be found [here.](https://uwaterloo.ca/associate-vice-president-academic/news/avpa-memo-chat-gpt-and-other-ai-tools-university-waterloo)

FAQ for Instructors and TAs can be found [here.](https://uwaterloo.ca/associate-vice-president-academic/frequently-asked-questions-chatgpt-and-generative-ai) Wiley has also put out a publication on *AI in Higher Education: Current Uses and Future Applications* which can be found [here.](https://universityservices.wiley.com/wp-content/uploads/2020/12/201811-AI-in-Higher-Education-TLH-with-new-bage.pdf)

Thinking about Generative Artifical Intelligence in the Classroom (Video): An overview for University of Waterloo instructors on considerations for generative artificial intelligence in teaching and learning. Includes recommendations instructors may wish to implement in their teaching practice.

### INSTRUCTOR RESOURCES FOR ARTIFICIAL INTELLIGENCE & CHATGPT TOOLS

* [AIO Guidance Regarding Student Use of AI in Your Course](https://www.torontomu.ca/academicintegrity/ai/) (Toronto Metropolitan University)
* [Teaching and Learning with Artificial Intelligence Apps](https://taylorinstitute.ucalgary.ca/teaching-with-AI-apps) (University of Calgary)
* [Artificial intelligence and Assessment Design (](https://staff.flinders.edu.au/learning-teaching/good-practice-guides/good-practice-guide---designing-assessment-for-artificial-intell)Flinders University)
* [Artificial Intelligence in Teaching & Learning](https://act.utoronto.ca/ai-readings/) (University of Toronto)
* [Conversations with Students about Generative Artificial Intelligence Tools](https://uwaterloo.ca/centre-for-teaching-excellence/catalogs/tip-sheets/conversations-students-about-generative-artificial) (Centre for Teaching Excellence)

### TIPS FOR STUDENTS

**Article:** [5 things students need to know before using GenAI](https://uwaterloo.ca/academic-integrity/sites/ca.academic-integrity/files/uploads/files/5_things_students_need_to_know_before_using_ai_to_generate_assignments_2.pdf)

**Interactive Lesson** - [The Assignment Dilemma: Using GenAI Productively](https://lib.uwaterloo.ca/online_learning/assignment-dilemma-using-genai-productively)

This interactive lesson offers an engaging experience that delves into the ethical considerations, challenges, and applications of Generative Artificial Intelligence (AI). By participating in this interactive lesson, you can better understand the practical aspects of handling Generative AI responsibly.

**Before beginning a project/assignment/test:**

* Review the instructions and ensure you are clear on your instructors expectations. If you are unsure if you can use AI tools, ask your instructor **before** starting the project/assignment/test.
* If AI tools are permitted in a project or written assignment, ensure you are aware of how to cite the use of such tools. If you are not sure which citation style to use or how to cite AI tools, ask your instructor. A helpful guide to citing AI tools in APA style can be found [here.](https://subjectguides.uwaterloo.ca/chatgpt_generative_ai)
* Check out resources from the Writing and Communication Centre [here](https://uwaterloo.ca/writing-and-communication-centre/Resources-AI-Overview) that instruct you how to use GenAI tools in the writing process.
* Develop a plan working backwards from the due date.
* Create a schedule to stay on track. Find schedule templates in the SSO’s section on [‘Time](https://uwaterloo.ca/student-success/resources)

[Management’.](https://uwaterloo.ca/student-success/resources)

* If you find you are falling behind and might not make the due date reach out to your instructor immediately.
* Reach out to campus resources like the [Writing and Communication Centre o](https://uwaterloo.ca/writing-and-communication-centre/)r [SSO](https://uwaterloo.ca/student-success/learning-support) for support and guidance.

The University of Queensland Australia has developed a brief module for students on ways to incorporate AI tools into their studies. To review the module click [here.](https://uq.pressbooks.pub/digital-essentials-artificial-intelligence/chapter/module-overview/#3)

**Remember you only get out what you put into your academic courses! To graduate with the best skills for your future - ensure you are learning the content and can demonstrate your knowledge.**

**28.** [**Montreal Declaration on Responsible AI**](https://montrealdeclaration-responsibleai.com/about/#:~:text=The%20Montreal%20Declaration%20for%20a,%2C%20knowledge%2C%20democracy%20and%20responsibility.)

On November 3, 2017, the Université de Montréal launched the coconstruction process for the Montréal Declaration for a Responsible Development of Artificial Intelligence (Montréal Declaration) . A year later, the results of these citizen deliberations are public. Dozens of events were organized to stimulate discussion on social issues that arise with artificial intelligence (AI), and 15 deliberation workshops were held over three months, involving over 500 citizens, experts and stakeholders from all backgrounds.

The Montréal Declaration is a collective work that aims to put AI development at the service of the well-being of all people, and to guide social change by developing recommendations with strong democratic legitimacy.

The selected citizen co-construction method is based on a preliminary declaration of general ethical principles structured around seven (7) fundamental values: well-being, autonomy, justice, privacy, knowledge, democracy and responsibility. Following the process, the Declaration was enriched and now presents 10 principles based on the following values: well-being, autonomy, intimacy and privacy, solidarity, democracy, equity, inclusion, caution, responsibility and environmental sustainability. **Our process for responsible artificial intelligence**

The **Montreal Declaration for a Responsible Development of Artificial Intelligence** is based on a declaration of ethical principles built around 7 core values: well-being, autonomy, justice, privacy, knowledge, democracy and responsibility. These values, suggested by a group of ethics, law, public policy and artificial intelligence experts, have been informed by a deliberation process. This deliberation occurred through consultations held over three months, in **15 different public spaces**, and sparked exchanges between **over 500 citizens, experts and stakeholders** from every horizon.

Educational, ethical and methodological introduction

What is artificial intelligence? What are the ethical issues raised by AI? What is coconstruction, and most importantly, what is expected of citizens? The scientific co-directors of the Declaration answer these questions and set the stage for the ensuing discussions.

**Debating**

Education, Health, Smart Cities, Justice and the Workforce are the 5 sectors around which prospective scenarios were developed. Using these scenarios set in 2025, groups of 5 to 8 people, with the help of a facilitator, discuss ethical issues.

**Suggesting**

Using the issues developed for 2025, participants must now imagine recommendations to allow a responsible rollout and use of AI in Quebec.

**29.** [**Provisional Guidelines on the Use of Generative AI in Teaching and Learning**](https://provost.mcmaster.ca/office-of-the-provost-2/generative-artificial-intelligence/task-force-on-generative-ai-in-teaching-and-learning/provisional-guidelines-on-the-use-of-generative-ai-in-teaching-and-learning/#:~:text=Unless%20otherwise%20stated%2C%20students%20should,McMaster's%20citation%20guide%20for%20examples.)

### Provisional Principles

These overarching provisional principles have guided the work of the Task Force on Generative AI in Teaching and Learning.

* Students want to learn, and instructors want to support their learning.
* Participatory learning – learning which happens in relationships and community – continues to be a valuable and vital way for students to learn.
* Assessments that require students to document the process of learning continue to be meaningful for student learning.
* Generative AI poses risks, as well as opportunities. Individuals will have different reactions and different expectations for the technology.

**Provisional Guidelines on the Use of Generative AI in Teaching and Learning**

[**General Guidelines**](https://provost.mcmaster.ca/office-of-the-provost-2/generative-artificial-intelligence/task-force-on-generative-ai-in-teaching-and-learning/provisional-guidelines-on-the-use-of-generative-ai-in-teaching-and-learning/#accordion-General-Guidelines-1)

[Academic Integrity](https://provost.mcmaster.ca/office-of-the-provost-2/generative-artificial-intelligence/task-force-on-generative-ai-in-teaching-and-learning/provisional-guidelines-on-the-use-of-generative-ai-in-teaching-and-learning/#accordion-Academic-Integrity-1)

McMaster’s existing [academic integrity policy](https://secretariat.mcmaster.ca/app/uploads/Academic-Integrity-Policy-1-1.pdf) applies when using generative AI. Its overall definition of academic dishonesty, which is to knowingly act or fail to act in a way that results or could result in unearned academic credit or advantage, allows for allegations related to generative

AI. The policy states under item 18(c) that “It shall be an offence knowingly to … submit academic work for assessment that was purchased or acquired from another source”.

* 1. Unless otherwise stated, students should assume use of generative AI is prohibited.
  2. Instructors who incorporate generative AI into courses should explain to students in writing and verbally in-class how generative AI material should be acknowledged or cited. [See McMaster’s citation guide](https://libguides.mcmaster.ca/cite-gen-ai/mla) for examples.

2. Generative AI plagiarism detection software is currently unavailable or not recommended at McMaster. This software will continue to be reviewed and may be used in the future.

* 1. These detectors will produce false positives and are not approved for use through the University’s policy. Students have not consented to the sharing of their intellectual work through these tools. It is also unclear how the material submitted to the third-party detectors is retained or used.
  2. Until more is understood about generative AI detection tools, instructors should not submit student work to generative AI detection tools.
  3. McMaster has an institutional membership to Turnitin, a plagiarism detection software. Turnitin [announced an update](https://www.turnitin.com/blog/the-launch-of-turnitins-ai-writing-detector-and-the-road-ahead) aimed at detecting writing produced by generative AI. McMaster, like many other institutions, has not yet turned on this feature as there is a need to understand the functionality of the tool, assess the security and privacy considerations for student work and determine whether it aligns with existing policies.
  4. If you do suspect student work may have violated the academic integrity policy, [please review the steps to take.](https://www.mcmaster.ca/academicintegrity/instructors/processdishonesty/index.html) [**Instructor or TA Use of Generative AI**](https://provost.mcmaster.ca/office-of-the-provost-2/generative-artificial-intelligence/task-force-on-generative-ai-in-teaching-and-learning/provisional-guidelines-on-the-use-of-generative-ai-in-teaching-and-learning/#accordion-Instructor-or-TA-Use-of-Generative-AI--1)

1. If instructors use generative AI in their teaching materials instructors should explain in the course outline the extent to which generative AI has been, or will be, used.
   1. Instructors should fact-check any generative AI produced materials.
   2. Instructors should not submit student work to generative AI tools for feedback without students’ consent and ability to opt-out.
2. Course instructors have three options for directing teaching assistant use of generative AI
   1. Permitting teaching assistants to use generative AI for any aspect of teaching assistant work, with the exception of summative evaluation, with no expectation that they use generative AI and no training specific to generative AI required. TAs must inform the instructor of the intended use of generative AI, and receive approval, before implementation. Summative evaluations are those which significantly impact a student’s grade or progress in a course. This includes providing a quantitative grade (number or letter grade).
   2. Requiring teaching assistants to use generative AI for specified teaching tasks as outlined in the hours of work form and with training provided.

■ In the instance of required use: As directed by the course instructor explicitly in the hours of work form, teaching assistants will use generative AI for the specific teaching tasks. Course instructors will provide teaching assistants with the necessary training to use generative AI for the specified teaching purpose(s) with this training included in the hours of work. Teaching assistants will evaluate all teaching materials/formative feedback developed with generative AI for accuracy before use with students. Any planned use of generative AI by teaching assistants will be shared with students in the course outline.

* 1. Prohibiting teaching assistants from using generative AI for teaching tasks

1. Generative AI tools can be used to provide formative feedback on student work; generative AI tools cannot be used to provide summative evaluation of student work.
   1. AI-generated formative feedback is intended to guide learning and improve understanding, by pointing out strengths and areas for improvement in student work.
   2. Summative evaluations are those which significantly impact a student’s grade or progress in a course. This includes providing a quantitative grade (number or letter grade).
   3. Instructors, or teaching assistants when directed, should review AI-generated formative feedback to ensure it aligns with the learning objectives and course materials, and add their own insights where necessary. Formative feedback that uses AI should not be given a quantitative grade by the AI tool. A “pass/fail” or

“completion” may be applied.

* 1. Instructors, or teaching assistants when directed, are responsible for summative evaluations to ensure appropriateness and accuracy.
  2. Data collection should be turned off on generative AI tools when used for providing formative feedback.
  3. When providing AI-generated formative feedback, students should be made aware that it is generated by AI explicitly in the course syllabus.

1. Instructors who include assessments that incorporate generative AI should:
   1. Consider including reflective components that invite students to comment on the use of/experience with generative AI in the assessment
   2. Explicitly review criteria and/or rubrics in ways that demonstrate how the use of generative AI is being assessed.

[**Privacy, Security and Selection of Tools**](https://provost.mcmaster.ca/office-of-the-provost-2/generative-artificial-intelligence/task-force-on-generative-ai-in-teaching-and-learning/provisional-guidelines-on-the-use-of-generative-ai-in-teaching-and-learning/#accordion-Privacy-Security-and-Selection-of-Tools-1)

1. Instructors incorporating generative AI should be aware of the privacy policies and user agreements of each generative AI tool and alert students to these policies in the course outline.
2. Where possible, courses that incorporate generative AI should rely on free versions of generative AI tools (e.g. Microsoft Copilot, ChatGPT 3) for student use.

a. Alternatives should be provided for Generative AI tools that are restricted to users 18+ (e.g. ChatGPT).

[**Student Assessments**](https://provost.mcmaster.ca/office-of-the-provost-2/generative-artificial-intelligence/task-force-on-generative-ai-in-teaching-and-learning/provisional-guidelines-on-the-use-of-generative-ai-in-teaching-and-learning/#accordion-Student-Assessments-1)

1. Assessment alternatives that may be less susceptible to the use of generative AI include oral exams, presentations followed by a Q and A, invigilated/in-class assessments, practical tests, assessments that incorporate class discussion/activities, and process-based work.
2. Instructors may consider adding an honour pledge to assessments.
3. Students may opt-out of assessments that require the use of generative AI only in exceptional circumstances as approved by the course instructor. If approved to opt-out of an assessment that requires the use of generative AI based on an exceptional circumstance, students will not face academic penalty, but will be required to provide alternative and equivalent evidence of their learning as proposed to, and agreed to by, the course instructor.

[Ongoing Support and Continued Work](https://provost.mcmaster.ca/office-of-the-provost-2/generative-artificial-intelligence/task-force-on-generative-ai-in-teaching-and-learning/provisional-guidelines-on-the-use-of-generative-ai-in-teaching-and-learning/#accordion-Ongoing-Support-and-Continued-Work-1)

1. The MacPherson Institute will continue to provide training and resources for instructors and students on how to use generative AI effectively. See [the MacPherson Institute website](https://mi.mcmaster.ca/) for current workshops, resources and to schedule a consultation.
2. McMaster will explore an annual donation to carbon offsetting programs to address the environmental impact of training large AI models.
3. The MacPherson Institute will collect feedback from instructors and students this fall on their experiences, questions and concerns about using generative AI in teaching and learning in an effort to update and improve these guidelines.
4. These guidelines will be regularly reviewed and revised.

### Sample Syllabus Statements

These sample syllabus statements may be included on a course syllabus to communicate with students the expectations around generative AI in a course. Instructors may adapt or modify these statements according to their individual teaching goals and course learning outcomes.

[**Use Prohibited**](https://provost.mcmaster.ca/office-of-the-provost-2/generative-artificial-intelligence/task-force-on-generative-ai-in-teaching-and-learning/provisional-guidelines-on-the-use-of-generative-ai-in-teaching-and-learning/#accordion-Use-Prohibited-2)

Students are not permitted to use generative AI in this course. In alignment with [McMaster academic integrity policy,](https://secretariat.mcmaster.ca/app/uploads/Academic-Integrity-Policy-1-1.pdf) it “shall be an offence knowingly to … submit academic work for assessment that was purchased or acquired from another source”. This includes work created by generative AI tools. Also state in the policy is the following, “Contract Cheating is the act of “outsourcing of student work to third parties” (Lancaster & Clarke, 2016, p. 639) with or without payment.” Using Generative AI tools is a form of contract cheating. Charges of academic dishonesty will be brought forward to the Office of Academic Integrity.

[**Some Use Permitted**](https://provost.mcmaster.ca/office-of-the-provost-2/generative-artificial-intelligence/task-force-on-generative-ai-in-teaching-and-learning/provisional-guidelines-on-the-use-of-generative-ai-in-teaching-and-learning/#accordion-Some-Use-Permitted-2)

Example One

Students may use generative AI in this course in accordance with the guidelines outlined for each assessment, and so long as the use of generative AI is referenced and cited following citation instructions given in the syllabus. Use of generative AI outside assessment guidelines or without citation will constitute academic dishonesty. It is the student’s responsibility to be clear on the limitations for use for each assessment and to be clear on the expectations for citation and reference and to do so appropriately.

Example Two

Students may use generative AI for [editing/translating/outlining/brainstorming/revising/etc] their work throughout the course so long as the use of generative AI is referenced and cited following citation instructions given in the syllabus. Use of generative AI outside the stated use of [editing/translating/outling/brainstorming/revising/etc] without citation will constitute academic dishonesty. It is the student’s responsibility to be clear on the limitations for use and to be clear on the expectations for citation and reference and to do so appropriately.

Example Three

Students may freely use generative AI in this course so long as the use of generative AI is referenced and cited following citation instructions given in the syllabus. Use of generative AI outside assessment guidelines or without citation will constitute academic dishonesty. It is the student’s responsibility to be clear on the expectations for citation and reference and to do so appropriately.

[Unrestricted Use](https://provost.mcmaster.ca/office-of-the-provost-2/generative-artificial-intelligence/task-force-on-generative-ai-in-teaching-and-learning/provisional-guidelines-on-the-use-of-generative-ai-in-teaching-and-learning/#accordion-Unrestricted-Use-2)

Students may use generative AI throughout this course in whatever way enhances their learning; no special documentation or citation is required.

**30.** [**Artificial Intelligence and the Future of Teaching and Learning**](https://www2.ed.gov/documents/ai-report/ai-report.pdf)

### Introduction

The U.S. Department of Education (Department) is committed to supporting the use of technology to improve teaching and learning and to support innovation throughout educational systems. This report addresses the clear need for sharing knowledge and developing policies for “Artificial Intelligence,” a rapidly advancing class of foundational capabilities which are increasingly embedded in all types of educational technology systems and are also available to the public. We will consider “educational technology” (edtech) to include both (a) technologies specifically designed for educational use, as well as (b) general technologies that are widely used in educational settings. Recommendations in this report seek to engage teachers, educational leaders, policy makers, researchers, and educational technology innovators and providers as they work together on pressing policy issues that arise as Artificial Intelligence (AI) is used in education. AI can be defined as “automation based on associations.” When computers automate reasoning based on associations in data (or associations deduced from expert knowledge), two shifts fundamental to AI occur and shift computing beyond conventional edtech: (1) from capturing data to detecting patterns in data and (2) from providing access to instructional resources to automating decisions about instruction and other educational processes. Detecting patterns and automating decisions are leaps in the level of responsibilities that can be delegated to a computer system. The process of developing an AI system may lead to bias in how patterns are detected and unfairness in how decisions are automated. Thus, educational systems must govern their use of AI systems. This report describes opportunities for using AI to improve education, recognizes challenges that will arise, and develops recommendations to guide further policy development.

### Rising Interest in AI in Education

Today, many priorities for improvements to teaching and learning are unmet. Educators seek technology-enhanced approaches addressing these priorities that would be safe, effective, and scalable. Naturally, educators wonder if the rapid advances in technology in everyday lives could help. Like all of us, educators use AI-powered services in their everyday lives, such as voice assistants in their homes; tools that can correct grammar, complete sentences, and write essays; and automated trip planning on their phones. Many educators are actively exploring AI tools as they are newly released to the public1 . Educators see opportunities to use AI-powered capabilities like speech recognition to increase the support available to students with disabilities, multilingual learners, and others who could benefit from greater adaptivity and personalization in digital tools for learning. They are exploring how AI can enable writing or improving lessons, as well as their process for finding, choosing, and adapting material for use in their lessons. Educators are also aware of new risks. Useful, powerful functionality can also be accompanied with new data privacy and security risks. Educators recognize that AI can automatically produce output that is inappropriate or wrong. They are wary that the associations or automations created by AI may amplify unwanted biases. They have noted new ways in which students may represent others’ work as their own. They are well-aware of “teachable moments” and pedagogical strategies that a human teacher can address but are undetected or misunderstood by AI models. They worry whether recommendations suggested by an algorithm would be fair. Educators’ concerns are manifold. Everyone in education has a responsibility to harness the good to serve educational priorities while also protecting against the dangers that may arise as a result of AI being integrated in edtech. To develop guidance for edtech, the Department works closely with educational constituents. These constituents include educational leaders—teachers, faculty, support staff, and other educators— researchers; policymakers; advocates and funders; technology developers; community members and organizations; and, above all, learners and their families/caregivers. Recently, through its activities with constituents, the Department noticed a sharp rise in interest and concern about AI. For example, a 2021 field scan found that developers of all kinds of technology systems—for student information, classroom instruction, school logistics, parent teacher communication, and more—expect to add AI capabilities to their systems. Through a series of four listening sessions conducted in June and August 2022 and attended by more than 700 attendees, it became clear that constituents believe that action is required now in order to get ahead of the expected increase of AI in education technology—and they want to roll up their sleeves and start working together. In late 2022 and early 2023, the public became aware of new generative AI chatbots and began to explore how AI could be used to write essays, create lesson plans, produce images, create personalized assignments for students, and more. From public expression in social media, at conferences, and in news media, the Department learned more about risks and benefits of AIenabled chatbots. And yet this report will not focus on a specific AI tool, service, or announcement, because AI-enabled systems evolve rapidly. Finally, the Department engaged the educational policy expertise available internally and in its relationships with AI policy experts to shape the findings and recommendations in this report.

### Three Reasons to Address AI in Education Now

“I strongly believe in the need for stakeholders to understand the cyclical effects of AI and education. By understanding how different activities accrue, we have the ability to support virtuous cycles. Otherwise, we will likely allow vicious cycles to perpetuate.” —Lydia Liu During the listening sessions, constituents articulated three reasons to address AI now: First, AI may enable achieving educational priorities in better ways, at scale, and with lower costs. Addressing varied unfinished learning of students due to the pandemic is a policy priority, and AI may improve the adaptivity of learning resources to students’ strengths and needs. Improving teaching jobs is a priority, and via automated assistants or other tools, AI may provide teachers greater support. AI may also enable teachers to extend the support they offer to individual students when they run out of time. Developing resources that are responsive to the knowledge and experiences students bring to their learning—their community and cultural assets—is a priority, and AI may enable greater customizability of curricular resources to meet local needs.

As seen in voice assistants, mapping tools, shopping recommendations, essay-writing capabilities, and other familiar applications, AI may enhance educational services. Second, urgency and importance arise through awareness of system-level risks and anxiety about potential future risks. For example, students may become subject to greater surveillance. Some teachers worry that they may be replaced—to the contrary, the Department firmly rejects the idea that AI could replace teachers. Examples of discrimination from algorithmic bias are on the public’s mind, such as a voice recognition system that doesn’t work as well with regional dialects, or an exam monitoring system that may unfairly identify some groups of students for disciplinary action. Some uses of AI may be infrastructural and invisible, which creates concerns about transparency and trust. AI often arrives in new applications with the aura of magic, but educators and procurement policies require that edtech show efficacy. AI may provide information that appears authentic, but actually is inaccurate or lacking a basis in reality. Of the highest importance, AI brings new risks in addition to the well-known data privacy and data security risks, such as the risk of scaling pattern detectors and automations that result in “algorithmic discrimination” (e.g., systematic unfairness in the learning opportunities or resources recommended to some populations of students). Third, urgency arises because of the scale of possible unintended or unexpected consequences. When AI enables instructional decisions to be automated at scale, educators may discover unwanted consequences. In a simple example, if AI adapts by speeding curricular pace for some students and by slowing the pace for other students (based on incomplete data, poor theories, or biased assumptions about learning), achievement gaps could widen. In some cases, the quality of available data may produce unexpected results. For example, an AI-enabled teacher hiring system might be assumed to be more objective than human-based résumé scoring. Yet, if the AI system relies on poor quality historical data, it might de-prioritize candidates who could bring both diversity and talent to a school’s teaching workforce. In summary, it is imperative to address AI in education now to realize key opportunities, prevent and mitigate emergent risks, and tackle unintended consequences.

### Toward Policies for AI in Education

The 2023 AI Index Report from the Stanford Institute for Human-Centered AI has documented notable acceleration of investment in AI as well as an increase of research on ethics, including issues of fairness and transparency.2 Of course, research on topics like ethics is increasing because problems are observed. Ethical problems will occur in education, too.3 The report found a striking interest in 25 countries in the number of legislative proposals that specifically include AI. In the United States, multiple executive orders are focused on ensuring AI is trustworthy and equitable, and the White House Office of Science and Technology Policy has introduced a Blueprint for an AI Bill of Rights (Blueprint)4 that provides principles and practices that help achieve this goal. These initiatives, along with other AI-related policy activities occurring in both the executive and legislative branches, will guide the use of AI throughout all sectors of society. In Europe, the European Commission recently released Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators. 5 AI is moving fast and heralding societal changes that require a national policy response. In addition to broad policies for all sectors of society, education-specific policies are needed to address new opportunities and challenges within existing frameworks that take into consideration federal student privacy laws (such as the Family Educational Rights and Privacy Act, or FERPA), as well as similar state related laws. AI also makes recommendations and takes actions automatically in support of student learning, and thus educators will need to consider how such recommendations and actions can comply with laws such as the Individuals with Disabilities Education Act (IDEA). We discuss specific policies in the concluding section.

AI is advancing exponentially (see Figure 1), with powerful new AI features for generating images and text becoming available to the public, and leading to changes in how people create text and images6 . The advances in AI are not only happening in research labs but also are making news in mainstream media and in educational-specific publications. Researchers have articulated a range of concepts and frameworks for ethical AI7 , as well as for related concepts such as equitable, responsible, and human-centered AI. Listening session participants called for building on these concepts and frameworks but also recognized the need to do more; participants noted a pressing need for guardrails and guidelines that make educational use of AI advances safe, especially given this accelerating pace of incorporation of AI into mainstream technologies. As policy development takes time, policy makers and educational constituents together need to start now to specify the requirements, disclosures, regulations, and other structures that can shape a positive and safe future for all constituents—especially students and teachers. Policies are urgently needed to implement the following: 1. leverage automation to advance learning outcomes while protecting human decision making and judgment; 2. interrogate the underlying data quality in AI models to ensure fair and unbiased pattern recognition and decision making in educational applications, based on accurate information appropriate to the pedagogical situation; 3. enable examination of how particular AI technologies, as part of larger edtech or educational systems, may increase or undermine equity for students; and 4. take steps to safeguard and advance equity, including providing for human checks and balances and limiting any AI systems and tools that undermine equity.

**Building Ethical, Equitable Policies Together**

In this report, we aim to build on the listening sessions the Department hosted to engage and inform all constituents involved in making educational decisions so they can prepare for and make better decisions about the role of AI in teaching and learning. AI is a complex and broad topic, and we are not able to cover everything nor resolve issues that still require more constituent engagement.

This report is intended to be a starting point. The opportunities and issues of AI in education are equally important in K-12, higher education, and workforce learning. Due to scope limitations, the examples in this report will focus on K-12 education. The implications are similar at all levels of education, and the Department intends further activities in 2023 to engage constituents beyond K12 schools.

### Guiding Questions

Understanding that AI increases automation and allows machines to do some tasks that only people did in the past leads us to a pair of bold, overarching questions: 1. What is our collective vision of a desirable and achievable educational system that leverages automation to advance learning while protecting and centering human agency? 2. How and on what timeline will we be ready with necessary guidelines and guardrails, as well as convincing evidence of positive impacts, so that constituents can ethically and equitably implement this vision widely? In the Learning, Teaching, and Assessment sections of this report, we elaborate on elements of an educational vision grounded in what today’s learners, teachers, and educational systems need, and we describe key insights and next steps required. Below, we articulate four key foundations for framing these themes. These foundations arise from what we know about the effective use of educational technology to improve opportunity, equity, and outcomes for students and also relate to the new Blueprint.

### Foundation 1: Center people (parents, educators, and students)

Education-focused AI policies at the federal, state, and district levels will be needed to guide and empower local and individual decisions about which technologies to adopt and use in schools and classrooms. Consider what is happening in everyday lives. Many of us use AI-enabled products because they are often better and more convenient. For example, few people want to use paper maps anymore; people find that technology helps us plan the best route to a destination more efficiently and conveniently. And yet, people often do not realize how much privacy they are giving up when they accept AI-enabled systems into their lives. AI will bring privacy and other risks that are hard to address only via individual decision making; additional There should be clear limits on the ability to collect, use, transfer, and maintain our personal data, including limits on targeted advertising. These limits should put the burden on platforms to minimize how much information they collect, rather than burdening Americans with reading fine print.8 As protections are developed, we recommend that policies center people, not machines. To this end, a first recommendation in this document (in the next section) is an emphasis on AI with humans in the loop. Teachers, learners, and others need to retain their agency to decide what patterns mean and to choose courses of action. The idea of humans in the loop builds on the concept of “Human Alternatives, Consideration, and Fallback” in the Blueprint and ethical concepts used more broadly in evaluating AI, such as preserving human dignity. A top policy priority must be establishing human in the loop as a requirement in educational applications, despite contrary pressures to use AI as an alternative to human decision making. Policies should not hinder innovation and improvement, nor should they be burdensome to implement. Society needs an education-focused AI policy that protects civil rights and promotes democratic values in the building, deployment, and governance of automated systems to be used across the many decentralized levels of the American educational system.

### Foundation 2: Advance Equity

“AI brings educational technology to an inflection point. We can either increase disparities or shrink them, depending on what we do now.” —Dr. Russell Shilling A recent Executive Order9 issued by President Biden sought to strengthen the connection among racial equity, education and AI, stating that “members of underserved communities—many of whom have endured generations of discrimination and disinvestment—still confront significant barriers to realizing the full promise of our great Nation, and the Federal Government has a responsibility to remove these barriers” and that the Federal Government shall both “pursue educational equity so that our Nation’s schools put every student on a path to success” and also “root out bias in the design and use of new technologies, such as artificial intelligence.” A specific vision of equity, such as described in the Department’s recent report, Advancing Digital Equity for All10 is essential to policy discussion about AI in education. This report defines digital equity as “the condition in which individuals and communities have the information technology capacity that is needed for full participation in the society and economy of the United States.” Issues related to racial equity and unfair bias were at the heart of every listening session we held. In particular, we heard a conversation that was increasingly attuned to issues of data quality and the consequences of using poor or inappropriate data in AI systems for education. Datasets are used to develop AI, and when they are nonrepresentative or contain undesired associations or patterns, resulting AI models may act unfairly in how they detect patterns or automate decisions. Systematic, unwanted unfairness in how a computer detects patterns or automates decisions is called “algorithmic bias.” Algorithmic bias could diminish equity at scale with unintended discrimination. As this document discussed in the Formative Assessment section, this is not a new conversation. For decades, constituents have rightly probed whether assessments are unbiased and fair. Just as with assessments, whether an AI model exhibits algorithmic bias or is judged to be fair and trustworthy is critical as local school leaders make adoption decisions about using AI to achieve their equity goals. We highlight the concept of “algorithmic discrimination” in the Blueprint. Bias is intrinsic to how AI algorithms are developed using historical data, and it can be difficult to anticipate all impacts of biased data and algorithms during system design. The Department holds that biases in AI algorithms must be addressed when they introduce or sustain unjust discriminatory practices in education. For example, in postsecondary education, algorithms that make enrollment decisions, identify students for early intervention, or flag possible student cheating on exams must be interrogated for evidence of unfair discriminatory bias—and not only when systems are designed, but also later, as systems become widely used.

### Foundation 3: Ensure Safety, Ethics, and Effectiveness

A central safety argument in the Department’s policies is the need for data privacy and security in the systems used by teachers, students, and others in educational institutions. The development and deployment of AI requires access to detailed data. This data goes beyond conventional student records (roster and gradebook information) to detailed information about what students do as they learn with technology and what teachers do as they use technology to teach. AI’s dependence on data requires renewed and strengthened attention to data privacy, security, and governance (as also indicated in the Blueprint). As AI models are not generally developed in consideration of educational usage or student privacy, the educational application of these models may not be aligned with the educational institution’s efforts to comply with federal student privacy laws, such as FERPA, or state privacy laws Further, educational leaders are committed to basing their decisions about the adoption of educational technology on evidence of effectiveness—a central foundation of the Department’s policy. For example, the requirement to base decisions on evidence also arises in the Elementary and Secondary Education Act (ESEA), as amended, which introduced four tiers of evidence (see Figure 2). Our nation’s research agencies, including the Institute of Education Sciences, are essential to producing the needed evidence. The Blueprint calls for evidence of effectiveness, but the education sector is ahead of that game: we need to insist that AIenhanced edtech rises to meet ESEA standards as well.

### Foundation 4: Promote Transparency

The central role of complex AI models in a technology’s detection of patterns and implementation of automation is an important way in which AI-enabled applications, products, and services will be different from conventional edtech. The Blueprint introduces the need for transparency about AI models in terms of disclosure (“notice”) and explanation. In education, decision makers will need more than notice—they will need to understand how AI models work in a range of general educational use cases, so they can better anticipate limitations, problems, and risks. AI models in edtech will be approximations of reality and, thus, constituents can always ask these questions: How precise are the AI models? Do they accurately capture what is most important? How well do the recommendations made by an AI model fit educational goals? What are the broader implications of using AI models at scale in educational processes? Building on what was heard from constituents, the sections of this report develop the theme of evaluating the quality of AI systems and tools using multiple dimensions as follows: ● About AI: AI systems and tools must respect data privacy and security. Humans must be in the loop. ● Learning: AI systems and tools must align to our collective vision for high-quality learning, including equity. ● Teaching: AI systems and tools must be inspectable, explainable, and provide human alternatives to AI-based suggestions; educators will need support to exercise professional Formative Assessment: AI systems and tools must minimize bias, promote fairness, and avoid additional testing time and burden for students and teachers. ● Research and Development: AI systems and tools must account for the context of teaching and learning and must work well in educational practice, given variability in students, teachers, and settings. ● Recommendations: Use of AI systems and tools must be safe and effective for students. They must include algorithmic discrimination protections, protect data privacy, provide notice and explanation, and provide a recourse to humans when problems arise. The people most affected by the use of AI in education must be part of the development of the AI model, system, or tool, even if this slows the pace of adoption. We return to the idea that these considerations fit together in a comprehensive perspective on the quality of AI models in the Recommendations section.

### Overview of Document

We begin in the next section by elaborating a definition of AI, followed by addressing learning, teaching, assessment, and research and development. Organizing key insights by these topics keeps us focused on exploring implications for improving educational opportunity and outcomes for students throughout the report. Within these topics, three important themes are explored: 1. Opportunities and Risks. Policies should focus on the most valuable educational advances while mitigating risks. 2. Trust and Trustworthiness. Trust and safeguarding are particularly important in education because we have an obligation to keep students out of harm’s way and safeguard their learning experiences. 3. Quality of AI Models. The process of developing and then applying a model is at the heart of any AI system. Policies need to support evaluation of the qualities of AI models and their alignment to goals for teaching and learning during the processes of educational adoption and use.

**What is AI?**

Our preliminary definition of AI as automation based on associations requires elaboration. Below we address three additional perspectives on what constitutes AI. Educators will find these different perspectives arise in the marketing of AI functionality and are important to understand when evaluating edtech systems that incorporate AI. One useful glossary of AI for Education terms is the CIRCLS Glossary of Artificial Intelligence Terms for Educators. 11 AI is not one thing but an umbrella term for a growing set of modeling capabilities, as visualized in Figure 3.

### Perspective: Human-Like Reasoning

“The theory and development of computer systems able to perform tasks normally requiring human intelligence such as, visual perception, speech recognition, learning, decision-making, and natural language processing.” 13 Broad cultural awareness of AI may be traced to the landmark 1968 film “2001: A Space Odyssey”—in which the “Heuristically-programmed ALgorithmic” computer, or “HAL,” converses with astronaut Frank. HAL helps Frank pilot the journey through space, a job that Frank could not do on his own. However, Frank eventually goes outside the spacecraft, HAL takes over control, and this does not end well for Frank. HAL exhibits humanlike behaviors, such as reasoning, talking, and acting. Like all applications of AI, HAL can help humans but also introduces unanticipated risks—especially since AI reasons in different ways and with different limitations than people do. The idea of “human-like” is helpful because it can be a shorthand for the idea that computers now have capabilities that are very different from the capabilities of early edtech applications. Educational applications will be able to converse with students and teachers, co-pilot how activities unfold in classrooms, and take actions that impact students and teachers more broadly. There will be both opportunities to do things much better than we do today and risks that must be anticipated and addressed. The “human-like” shorthand is not always useful, however, because AI processes information differently from how people process information. When we gloss over the differences between people and computers, we may frame policies for AI in education that miss the mark.

### Perspective: An Algorithm that Pursues a Goal

“Any computational method that is made to act independently towards a goal based on inferences from theory or patterns in data.” 14 This second definition emphasizes that AI systems and tools identify patterns and choose actions to achieve a given goal. These pattern recognition capabilities and automated recommendations will be used in ways that impact the educational process, including student learning and teacher instructional decision making. For example, today’s personalized learning systems may recognize signs that a student is struggling and may recommend an alternative instructional sequence. The scope of pattern recognition and automated recommendations will expand Correspondingly, humans must determine the types and degree of responsibility we will grant to technology within educational processes, which is not a new dilemma. For decades, the lines between the role of teachers and computers have been discussed in education, for example, in debates using terms such as “’computer-aided instruction,” “blended instruction,” and “personalized learning.” Yet, how are instructional choices made in systems that include both humans and algorithms? Today, AI systems and tools are already enabling the adaptation of instructional sequences to student needs to give students feedback and hints, for example, during mathematics problem solving or foreign language learning. This discussion about the use of AI in classroom pedagogy and student learning will be renewed and intensify as AIenabled systems and tools advance in capability and become more ubiquitous. Let’s start with another simple example. When a teacher says, “Display a map of ancient Greece on the classroom screen,” an AI system may choose among hundreds of maps by noting the lesson objectives, what has worked well in similar classrooms, or which maps have desirable features for student learning. In this case, when an AI system suggests an instructional resource or provides a choice among a few options, the instructor may save time and may focus on more important goals. However, there are also forms of AI-enabled automation that the classroom instructor may reject, for example, enabling an AI system or tool to select the most appropriate and relevant readings for students associated with a historical event. In this case, an educator may choose not to utilize AI-enabled systems or tools given the risk of AI creating false facts (“hallucinating”) or steering students toward inaccurate depictions of historical events found on the internet. Educators will be weighing benefits and risks like these daily. Computers process theory and data differently than humans.

AI’s success depends on associations or relationships found in the data provided to an algorithm during the AI model development process. Although some associations may be useful, others may be biased or inappropriate. Finding bad associations in data is a major risk, possibly leading to algorithmic discrimination. Every guardian is familiar with the problem: A person or computer may say, “Our data suggests your student should be placed in this class,” and the guardian may well argue, “No, you are using the wrong data. I know my child better, and they should instead be placed in another class.” This problem is not limited exclusively to AI systems and tools, but the use of AI models can amplify the problem when a computer uses data to make a recommendation because it may appear to be more objective and authoritative, even if it is not. Although this perspective can be useful, it can be misleading. A human view of agency, pursuing goals, and reasoning includes our human abilities to make sense of multiple contexts. For example, a teacher may see three students each make the same mathematical error but recognize that one student has an Individualized Education Program to address vision issues, another misunderstands a mathematical concept, and a third just experienced a frustrating interaction on the playground; the same instructional decision is therefore not appropriate. However, AI systems often lack data and judgement to appropriately include context as they detect patterns and automate decisions. Further, case studies show that technology has the potential to quickly derail from safe to unsafe or from effective to ineffective when the context shifts even slightly. For this and other reasons, people must be involved in goal setting, pattern analysis, and decision-making.15

### Perspective: Intelligence Augmentation

Foundation #1 (above) keeps humans in the loop and positions AI systems and tools to support human reasoning. “Intelligence Augmentation” (IA)17 centers “intelligence” and “decision making” in humans but recognizes that people sometimes are overburdened and benefit from assistive tools. AI may help teachers make better decisions because computers notice patterns that teachers can miss. For example, when a teacher and student agree that the student needs reminders, an AI system may provide reminders in whatever form a student likes without adding to the teacher’s workload. Intelligence Automation (IA) uses the same basic capabilities of AI, employing associations in data to notice patterns, and, through automation, takes actions based on those patterns. However, IA squarely focuses on helping people in human activities of teaching and learning, whereas AI tends to focus attention on what computers can do.

### Definition of “Model”

The above perspectives open a door to making sense of AI. Yet, to assess AI meaningfully, constituents must consider specific models and how they are developed. In everyday usage, the term “model” has multiple definitions. We clarify our intended meaning, which is a meaning similar to “mathematical model,” below. (Conversely, note that “model” as used in “AI model” is unlike the usage in “model school” or “instructional model” as AI model is not a singular case created by experts to serve as an exemplar.) AI models are like financial models: an approximation of reality that is useful for identifying patterns, making predictions, or analyzing alternative decisions. In a typical middle school math curriculum, students use a mathematical model to analyze which of two cell phone plans is better. Financial planners use this type of model to provide guidance on a retirement portfolio. At its heart, AI is a highly advanced mathematical toolkit for building and using models. Indeed, in well-known chatbots, complex essays are written one word at a time. The underlying AI model predicts which next words would likely follow the text written so far; AI chatbots use a very large statistical model to add one likely word at a time, thereby writing surprisingly coherent essays. When we ask about the model at the heart of AI, we begin to get answers about “what aspects of reality does the model approximate well?” and “how appropriate is it to the decision to be made?” One could similarly ask about algorithms—the specific decision-making processes that an AI model uses to go from inputs to outputs. One could also ask about the quality of the data used to build the model—for example, how representative is that data? Switching among three terms— models, algorithms, and data—will become confusing.

Because the terms are closely related, we’ve chosen to focus on the concept of AI models. We want to bring to the fore the idea that every AI model is incomplete, and it's important to know how well the AI model fits the reality we care about, where the model will break down, and how. Sometimes people avoid talking about the specifics of models to create a mystique. Talking as though AI is unbounded in its potential capabilities and a nearly perfect approximation to reality can convey an excitement about the possibilities of the future. The future, however, can be oversold. Similarly, sometimes people stop calling a model AI when its use becomes commonplace, yet such systems are still AI models with all of the risks discussed here. We need to know exactly when and where AI models fail to align to visions for teaching and learning.

### Insight: AI Systems Enable New Forms of Interaction

AI models allow computational processes to make recommendations or plans and also enable them to support forms of interaction that are more natural, such as speaking to an assistant. AIenabled educational systems will be desirable in part due to their ability to support more natural interactions during teaching and learning. In classic edtech platforms, the ways in which teachers and students interact with edtech are limited. Teachers and students may choose items from a menu or in a multiple-choice question. They may type short answers. They may drag objects on the screen or use touch gestures. The computer provides outputs to students and teachers through text, graphics, and multimedia. Although these forms of inputs and outputs are versatile, no one would mistake this style of interaction with the way two people interact with one another; it is specific to humancomputer interaction. With AI, interactions with computers are likely to become more like humanto-human interactions (see Figure 4). A teacher may speak to an AI assistant, and it may speak back. A student may make a drawing, and the computer may highlight a portion of the drawing. A teacher or student may start to write something, and the computer may finish their sentence—as when today’s email programs can complete thoughts faster than we can type them. Additionally, the possibilities for automated actions that can be executed by AI tools are expanding. Current personalization tools may automatically adjust the sequence, pace, hints, or trajectory through learning experiences.18 Actions in the future might look like an AI system or tool that helps a student with homework19 or a teaching assistant that reduces a teacher’s workload by recommending lesson plans that fit a teacher’s needs and are similar to lesson plans a teacher previously liked.20 Further, an AI-enabled assistant may appear as an additional “partner” in a small group of students who are working together on a collaborative assignment.21 An AI-enabled tool may also help teachers with complex classroom routines.22 For example, a tool may help teachers with orchestrating23 the movement of students from a full class discussion into small groups and making sure each group has the materials needed to start their work.

### Key Recommendation: Human in the Loop AI

Many have experienced a moment where technology surprised them with an uncanny ability to recommend what feels like a precisely personalized product, song, or even phrase to complete a sentence in a word processor such as the one being used to draft this document. Throughout this supplement, we talk about specific, focused applications where AI systems may bring value (or risks) into education. At no point do we intend to imply that AI can replace a teacher, a guardian, or an educational leader as the custodian of their students’ learning. We talk about the limitations of models in AI and the conversations that educational constituents need to have about what qualities they want AI models to have and how they should be used. These limitations lead to our first recommendation: that we pursue a vision of AI where humans are in the loop. That means that people are part of the process of noticing patterns in an educational system and assigning meaning to those patterns. It also means that teachers remain at the helm of major instructional decisions. It means that formative assessments involve teacher input and decision making, too. One loop is the cycle of recognizing patterns in what students do and selecting next steps or resources that could support their learning. Other loops involve teachers planning and reflecting on lessons. Response to Intervention is another well-known type of loop. The idea of humans in the loop is part of our broader discussions happening about AI and society, not just AI in education. Interested readers could look for more on human-centered AI, responsible AI, value-sensitive AI,

AI for social good, and other similar terms that ally with humans in the loop, such as “humancentered AI.” Exercising judgement and control in the use of AI systems and tools is an essential part of providing the best opportunity to learn for all students—especially when educational decisions carry consequence. AI does not have the broad qualities of contextual judgment that people do. Therefore, people must remain responsible for the health and safety of our children, for all students’ educational success and preparation for their futures, and for creating a more equitable and just society.

#### Learning

The Department’s long-standing edtech vision sees students as active learners; students participate in discussions that advance their understanding, use visualizations and simulations to explain concepts as they relate to the real world, and leverage helpful scaffolding and timely feedback as they learn. Constituents want technology to align to and build on these and other research-based understandings of how people learn. Educators can draw upon two books titled How People Learn and How People Learn II by the National Academies of Sciences, Engineering, and Medicine for a broad synthesis of what we know about learning.24 As we shape AI-enhanced edtech around research-based principles, a key goal must be to strengthen and support learning for those who have experienced unfavorable circumstances for learning, such as caused by the COVID-19 pandemic or by broader inequities. And we must keep a firm eye toward the forms of learning that will most benefit learners in their future lives in communities and workplaces. Examples of AI supporting learning principles in this section include the following: AI-based tutoring for students as they solve math problems (based on cognitive learning theories), adapting to learners with special needs (based on the Universal Design for Learning framework and related theories), and

AI support for effective student teamwork (based on theories in the field called “Computer

Supported Collaborative Learning”).

### Insight: AI Enables Adaptivity in Learning

Adaptivity has been recognized as a key way in which technology can improve learning.25 AI can be a toolset for improving the adaptivity of edtech. AI may improve a technology’s ability to meet students where they are, build on their strengths, and grow their knowledge and skills. Because of AI’s powers of work with natural forms of input and the foundational strengths of AI models (as discussed in the What is AI? section), AI can be an especially strong toolkit for expanding the adaptivity provided to students. And yet, especially with AI, adaptivity is always more specific and limited than what a broad phrase like “meet students where they are” might suggest. Core limits arise from the nature of the model at the heart of any specific AI-enabled system. Models are approximations of reality. When important parts of human learning are left out of the model or less fully developed, the resulting adaptivity will also be limited, and the resulting supports for learning may be brittle or narrow. Consequently, this section on Learning focuses on one key concept: Work toward AI models that fit the fullness of visions for learning—and avoid limiting learning to what AI can currently model well. AI models are demonstrating greater skills because of advances in what are called “large language models” or sometimes “foundational models.” These very general models still have limits. For example, generative AI models discussed in the mainstream news can quickly generate convincing essays about a wide variety of topics while other models can draw credible images based on just a few prompts. Despite the excitement about foundational models, experts in our listening sessions warned that AI models are narrower than visions for human learning and that designing learning environments with these limits in mind remains very important. The models are also brittle and can’t perform well when contexts change.

In addition, they don’t have the same “common sense” judgment that people have, often responding in ways that are unnatural or incorrect.26 Given the unexpected ways in which foundational models miss the mark, keeping humans in the loop remains highly important.

### Intelligent Tutoring Systems: An Example of AI Models

One long-standing type of AI-enabled technology is an Intelligent Tutoring System (ITS).27 In an early success, scientists were able to build accurate models of how human experts solve mathematical problems. The resulting model was incorporated into a system that would observe student problem solving as they worked on mathematical problems on a computer. Researchers who studied human tutors found that feedback on specific steps (and not just right or wrong solutions) is a likely key to why tutoring is so effective.28 For example, when a student diverged from the expert model, the system gave feedback to help the student get back on track.29 Importantly, this feedback went beyond right or wrong, and instead, the model was able to provide feedback on specific steps of a solution process. A significant advancement of AI, therefore, can be its ability to provide adaptivity at the step-by-step level and its ability to do so at scale with modest cost. As a research and development (R&D) field emerged to advance ITS, the work has gone beyond mathematics problems to additional important issues beyond step-by-step problem solving. In the early work, some limitations can be observed. The kinds of problems that an ITS could support were logical or mathematical, and they were closed tasks, with clear expectations for what a solution and solution process should look like. Also, the “approximation of reality” in early AI models related to cognition and not to other elements of human learning, for example, social or motivational aspects. Over time, these early limitations have been addressed in two ways: by expanding the AI models and by involving humans in the loop, a perspective that is also important now. Today, for example, if an ITS specializes in feedback as a student practices, a human teacher could still be responsible for motivating student engagement and self-regulation along with other aspects of instruction. In other contemporary examples, the computer ITS might focus on problem solving practice, while teachers work with students in small groups. Further, students can be in the loop with AI, as is the case with “open learner models”—a type of AIenabled system that provides information to support student self-monitoring and reflection.Although R&D along the lines of an ITS should not limit the view of what’s possible, such an example is useful because so much research and evaluation has been done on the ITS approach. Researchers have looked across all the available high-quality studies in a meta-analysis and concluded that ITS approaches are effective.31 Right now, many school systems are looking at high-intensity human tutoring to help students with unfinished learning. Human tutoring is very expensive, and it is hard to find enough high-quality human tutors. With regard to large-scale needs, if it is possible for an ITS to supplement what human tutors do, it might be possible to extend beyond the amount of tutoring that people can provide to students.

### Important Directions for Expanding AI-Based Adaptivity

Adaptivity is sometimes referred to as “personalization.” Although this is a convenient term, many observers have noted how imprecise it is.32 For some educators, personalization means giving learners “voice and choice,” and for others it means that a learning management system recommends an individual “playlist” of activities to each student. Hidden in that imprecision is the reality that many edtech products that personalize do so in limited ways. Adjusting the difficulty and the order of lesson materials are among the two most common ways that edtech products adapt. And yet, any teacher knows there is more to supporting learning than adjusting the difficulty and sequence of materials. For example, a good teacher can find ways to engage a student by connecting to their own past experiences and can shape explanations until they really connect in an “aha!” moment for that student. When we say, “meet the learner where they are,” human teachers bring a much more complete picture of each learner than most available edtech. The teacher is also not likely to “over personalize” (by performing like an algorithm that only presents material for which the learner has expressed interest), thereby limiting the student’s exposure to new topics. The nature of “teachable moments” that a human teacher can grasp is broader than the teachable moments today’s AI models grasp. In our listening sessions, we heard many ways in which the core models in an AI system must be expanded. We discuss these below. 1. From deficitbased to asset-oriented. Listening session attendees noted that the rhetoric around adaptivity has often been deficit-based; technology tries to pinpoint what a student is lacking and then provides instruction to fill that specific gap. Teachers also orient to students' strengths; they find competencies or “assets” a student has and use those to build up the students’ knowledge. AI models cannot be fully equitable while failing to recognize or build upon each student’s sources of competency. AI models that are more asset-oriented would be an advance. 2. From individual cognition to including social and other aspects of learning. The existing adaptivity rhetoric has also tended to focus on individualized learning and mostly on cognitive elements of learning, with motivational and other elements only brought in to support the cognitive learning goals. Attendees observe that their vision for learning is broader than cognition. Social learning is important, for example, especially for students to learn to reason, explain, and justify. For students who are learning English, customized and adaptive support for improving language skills while learning curricular content is clearly important. Developing self-regulation skills is also important. A modern vision of learning is not individualistic; it recognizes that students learn in groups and communities too. 3. From neurotypical to neurodiverse learners. AI models could help in including neurodiverse learners (students who access, process, and interact with the world in less common ways than “neurotypical” students) who could benefit from different learning paths and from forms of display and input that fit their strengths. Constituents want AI models that can support learning for neurodiverse learners and learners with disabilities. Thus, they want AI models that can work with multiple paths to learning and multiple modalities of interaction. Such models should be tested for efficacy, to guard against the possibility that some students could be assigned a “personalized” but inadequate learning resource. In addition, some systems for neurodiverse students are presently underutilized, so designs that support intended use will also be important. 4. From fixed tasks to active, open, and creative tasks. As mentioned above, AI models are historically better at closed tasks like solving a math problem or logical tasks like playing a game. In terms of life-wide and lifelong opportunities, we value learning how to succeed at open-ended and creative tasks that require extended engagement from the learner, and these are often not purely mathematical or logical. We want students to learn to invent and create innovative approaches. We want AI models that enable progress on open, creative tasks. 5. From correct answers to additional goals. At the heart of many adaptivity approaches now on the market, the model inside the technology counts students' wrong answers and decides whether to speed up, slow down, or offer a different type of learning support. Yet, right and wrong answers are not the only learning goals. We want students to learn how to self-regulate when they experience difficulties in learning, for example, such as being able to persist in working on a difficult problem or knowing how and when to ask for help. We want learners to become skilled in teamwork and in leading teams. As students grow, we want them to develop more agency and to be able to act on their own to advance toward their own learning goals. Listing every dimension of expansion that we heard in our listening sessions is beyond the scope of this report. Some additional dimensions are presented in the following sections on Teaching, Assessment, and Research. For example, in Research, we discuss all the ways in which AI systems have trouble with context—context that humans readily grasp and consider. Overall, constituents in the listening sessions realized we need an ambitious outlook on learning to respond to the future today’s learners face. Constituents were concerned about ways in which AI might narrow learning. For example, if the incorporation of AI into education slowed attention to students’ skills on creative, open-ended tasks and their ability to lead and collaborate in teams, then school districts may be less able to realize their students’ progress in relation to a Portrait of a Graduate who excels in communication and other skills valued in communities and careers. Constituents reminded us that as we conceptualize what we want AI in edtech to accomplish, we must start and constantly revisit a human-centered vision of learning.

### A Duality: Learning With and About AI

As AI is brought into schools, two broad perspectives about AI in education arise: (1) AI in support of student learning; and (2) support for learning about AI and related technologies. So far, we’ve discussed AI systems and tools to support student learning and mastery of subjects like mathematics and writing. Yet, it is also important that students learn about AI, critically examine its presence in education and society, and determine its role and value in their own lives and careers. We discuss risks across each section in this report. Here, it is important for students to become more aware of and savvy to the risks of AI—including risks of bias and surveillance—as they appear in all elements of their lives. In the recent past, schools have supported students’ understanding of cybersecurity, for example. AI will bring new risks, and students need to learn about them. We are encouraged by efforts we’ve seen underway that would give students opportunities to learn about how AI works while also giving them opportunities to discuss relevant topics like privacy and security. 33 Other learning goals are noted in the K-12 Computer Science Framework. We’ve seen that students can begin learning about AI in elementary, middle, and high school. They can use AI to design simulations and products that they find exciting. And we’ve seen that students want to talk about the ethics of products they experience in their everyday lives and have much to say about the kinds of products they’d like to see or not see in school. (And later, in the Research section, we note the desire for co-design processes that involve students in creating the next generation of AI-enabled edtech). Overall, it’s important to balance attention to using AI to support learning and giving students opportunities to learn about AI.

### A Challenge: Systems Thinking About AI in Education

As AI expands into the educational system, our listening session attendees reminded us that it will be entering parts or locations of the system that are presently dysfunctional. AI is certainly not a fix for broken systems, and instead, must be used with even more care when the systems’ context is unstable or uncertain. As discussed previously, because AI systems and tools do not fully align with goals for learning, we have to design educational settings to situate AI in the right place, where educators and other adults can make effective use of these tools for teaching and learning. Within the ITS example, we saw that AI could make learning by practicing math problems more effective, and a whole curricular approach might include roles for teachers that emphasize mathematical practices like argumentation and modeling. Further, small-group work is likely to remain important: Students might work in small groups to use mathematics to predict or justify as they work on responding to a realistic challenge. At the present, one “right place” for people, and not AI, is understanding how learning can be culturally responsive and culturally sustaining, as AI is not even close to being ready to connect learning to the unique strengths in a student’s community and family

### Open Questions About AI for Learning

With advances occurring in the foundations for AI, opportunities to use AI in support of learning are rapidly expanding. As we explore these opportunities, the open questions below deserve ongoing attention: ● To what extent is AI enabling adaptation to students’ strengths and not just deficits? Is AI enabling improved support for learners with disabilities and English language learners? ● How are youth voices involved in choosing and using AI for learning? ● Is AI leading to narrower student activities (e.g., procedural math problems), or the fuller range of activities highlighted in the National Educational Technology Plan (NETP), which emphasizes features such as personalized learning, project-based learning, learning from visualizations, simulations, and virtual reality, as well as learning across school, community, and familial settings? ● Is AI supporting the whole learner, including social dimensions of learning such as enabling students to be active participants in small group and collaborative learning? For example, does AI contribute to aspects of student collaboration we value like shared attention, mutual engagement, peer help, self-regulation, and building on each other’s contributions? ● When AI is used, are students’ privacy and data protected? Are students and their guardians informed about what happens with their data? ● How strong are the processes or systems for monitoring student use of AI for barriers, bias, or other undesirable consequences of AI use by learners? How are emergent issues addressed? ● Is high-quality research or evaluations about the impacts of using the AI system for student learning available? Do we know not only whether the system works but for whom and under what conditions?

### Key Recommendation: Seek AI Models Aligned to a Vision for Learning

We’ve called attention to how advances in AI are important to adaptivity but also to ways in which adaptivity is limited by the model’s inherent quality. We noted that a prior wave of edtech used the term “personalized” in differing ways, and it was often important to clarify what personalization meant for a particular product or service. Thus, our key recommendation is to tease out the strengths and limitations of AI models inside forthcoming edtech products and to focus on AI models that align closely to desired visions of learning. AI is now advancing rapidly, and we should differentiate between products that have simple AI-like features inside and products that have more sophisticated AI models. Looking at what’s happening in research and development, we can see significant effort and push toward overcoming these limitations. We noted that decision makers need to be careful about selecting AI models that might narrow their vision for learning, as general artificial intelligence does not exist. And because AI models will always be narrower than real world experience, we need to proceed with systems thinking in which humans are in the loop, with the strengths and weaknesses of the specific educational system considered. We hold that the full system for learning is broader than its AI component.

#### Teaching

Teachers have long envisioned many things that technology could make possible for teachers, their classrooms, and their students but not the changes wrought by the recent pandemic. Today, nearly all teachers have experienced uses of technologies for instruction that no one anticipated. Some of those experiences were positive, and others were not. All of the experiences provide an important context as we think further about teaching and technology. There is a critical need to focus on addressing the challenges teachers experience. It must become easier for teachers to do the amazing work they always do. We must also remember why people choose the teaching profession and ensure they can do the work that matters. This section discusses examples of AI supporting teachers and teaching including these concepts: AI assistants to reduce routine teaching burdens;

AI that provides teachers with recommendations for their students’ needs and extends their work with students; and AI that helps teachers to reflect, plan, and improve their practice.

### Always Center Educators in Instructional Loops

To succeed with AI as an enhancement to learning and teaching, we need to always center educators (ACE). Practically speaking, practicing “ACE in AI” means keeping a humanistic view of teaching front and center. ACE leads the Department to confidently respond “no” when asked “will AI replace teachers?” ACE is not just about making teachers’ jobs easier but also making it possible to do what most teachers want to do. That includes, for example, understanding their students more deeply and having more time to respond in creative ways to teachable moments. To bring more precision to how and where we should center educators, we return to our advocacy for human in the loop AI and ask, what are the loops in which teachers should be centered? Figure 5 suggests three key loops (inspired by research on adaptivity loops34): 1. The loop in which teachers make moment-to-moment decisions as they do the immediate work of teaching. 2. The loop in which teachers prepare for, plan, and reflect on teaching, which includes professional development. 3. The loop in which teachers participate in decisions about the design of AI-enabled technologies, participate in selecting the technologies, and shape the evaluation of technologies— thus setting a context for not only their own classroom but those of fellow teachers as well. Please note that in the next section, on Formative Assessment, we also discuss teachers’ important role in feedback loops that support students and enable school improvement. That section also includes a discussion of the concepts of “bias” and “fairness,” which are important to teachers.

### Insight: Using AI to Improve Teaching Jobs

The job of teaching is notoriously complex, with teachers making thousands of decisions each day. Teachers participate in classroom processes, in interactions with students beyond classrooms, in work with fellow teachers, and in administrative functions. They also are part of their communities and thus are expected to interact with families and caregivers alerts and notifications about events. Selecting music that we want to hear used to be a multistep process (even with digital music), and now we can speak the name of a song we want to hear, and it plays. Likewise, mapping a journey used to require a cumbersome study of maps, but now cell phones let us choose among several transportation options to reach a destination. Why can’t teachers be supported to notice changing student needs and provided with supports to enact a technology-rich lesson plan? Why can’t they more easily plan their students’ learning journeys? When things change in a classroom, as they always do, why don’t the tools of the classroom make it easier for teachers to adapt to student strengths and needs on the fly? A report by McKinsey36 first suggested that AI’s initial benefit could be to improve teaching jobs by reducing low-level burdens in administrative or clerical work (Figure 6). The report also suggests that recovered time from AI-enabled technology should be rededicated toward more effective instruction—particularly, outcomes such as reducing the average 11 hours of weekly preparation down to only six. We highlight these opportunities and two others below. 1. Handling low-level details to ease teaching burdens and increase focus on students. A good teacher must master all levels of details, big and small. When working with a particular student, the teacher may wish to later send that student a helpful learning resource. How will they remember to send it? A voice assistant or other forms of an AI assistant could make it easier to stay organized by categorizing simple voice notes for teachers to follow up on after a classroom session ends. We are beginning to see AIenabled voice assistants in the market, and they could do many simple tasks so that the teachers can stay focused on students. These tasks can include record-keeping, starting and stopping activities, controlling displays, speakers, and other technologies in the classroom, and providing reminders. Many workers may eventually use assistants to make their jobs easier, and teachers are the most deserving of efforts to ease their jobs now. 2. Extending beyond the teacher's availability with their students but continuing to deliver on the teacher’s intent. Teachers almost always want to do more with each student than they can, given the limited number of hours before the next school day. A teacher may wish to sit with the student as they practice 10 more math problems, giving them ongoing support and feedback. If the teacher can sit with the student for only three problems, perhaps they could delegate to an AIenabled learning system to help with the rest. Teachers cannot be at their best if on call at all hours to help with homework, but perhaps they can indicate what types of supports, hints, and feedback they want students to receive while studying after school hours. An AI assistant can ensure that students have that support wherever and whenever they do homework or practice skills on their own. Teachers may wish to provide more extensive personal notes to families/caregivers, and perhaps an AI assistant could help with drafts based on students’ recent classroom work. Then, the teacher could review the AI-generated comments and quickly edit where needed before returning it to the student for another draft. AI tools might also help teachers with language translation so they can work with all parents and caregivers of their students. AI tools might also help teachers with awareness. For example, in the next section, Formative Assessment, we note that teachers can’t always know what’s going on for each student and in each small group of students; emerging products might signal to the teacher when a student or teacher may need some more personal attention. 3. Making teacher professional development more productive and fruitful. Emerging products already enable a teacher to record her classroom and allow an AI algorithm to suggest highlights of the classroom discussion worth reviewing with a professional development coach.37 AI can compute metrics, such as whether students have been talking more or less, which are difficult for a teacher to calculate during a lesson.38 For teachers who want to increase student engagement, these metrics can be a valuable tool. Classroom simulation tools are also emerging and can enable teachers to practice their skills in realistic situations.39 Simulators can include examples of teaching from a real classroom while changing the faces and voices of the participants so that teaching situations can be shared and discussed among teachers without revealing identities. Note the emphasis above on what listening-session panelist Sarah Hampton said about the human touch. Teachers will feel that AI is helping them teach with a focus on their human connection to their students when the necessary (but less meaningful) burdens of teaching are lessened. In Figure 7, below, see concerns that teachers raised about AI during listening sessions.

### Preparing and Supporting Teachers in Planning and Reflecting

ACE also means preparing teachers to take advantage of possibilities like those listed above and more. In the Research section, we highlight how pre-service education still tends to compartmentalize and inadequately address the topic of technology. That section suggests a need to invest in research about how to deeply integrate technology in pre-service teacher training programs. In-service teachers, too, will need professional development to take advantage of opportunities that AI can provide, like those presented in the Teaching section. Professional development will need to be balanced not only to discuss opportunities but also to inform teachers of new risks, while providing them with tools to avoid the pitfalls of AI By nature, teaching requires significant time in planning as well to account for the breadth of needs across their rosters—especially for inclusive learning environments and students with IEPs and 504 plans. AI could help teachers with recommendations that are tuned to their situation and their ways of practicing teaching and support with adapting found materials to fit their exact classroom needs. For students with an IEP, AI could help with finding components to add to lesson plans to fully address standards and expectations and to meet each student’s unique requirements. Even beyond finding components, AI might help adapt standardized resources to better fit specific needs—for example, providing a voice assistant that allows a student with a visual difficulty to hear material and respond to it or permitting a group of students to present their project using American Sign Language (ASL) which could be audibly voiced for other students using an AI ASL-to-SpokenEnglish translation capability. Indeed, coordinating IEPs is time-consuming work that might benefit from supportive automation and customized interactivity that can be provided by AI. Reflection is important too. In the bustle of a classroom, it is sometimes difficult to fully understand what a student is expressing or what situations lead to certain positive or negative behaviors. Again, context is paramount. In the moment, teachers may not be aware of external events that could shape their understanding of how students are showing up in their classrooms. Tools that notice patterns and suggest ways to share information might help students and teachers communicate more fully about strengths and needs.

### Designing, Selecting, and Evaluating AI Tools

The broadest loop teachers should be part of is the loop that determines what classroom tools do and which tools are available. Today, teachers already play a role in designing and selecting technologies. Teachers can weigh in on usability and feasibility. Teachers examine evidence of efficacy and share their findings with other school leaders. Teachers already share insights on what is needed to implement technology well. While these concerns will continue, AI will raise new concerns too. For example, the following Formative Assessment section raises concerns about bias and fairness that can lead to algorithmic discrimination. Those concerns go beyond data privacy and security; they raise attention to how technologies may unfairly direct or limit some students’ opportunities to learn. A key takeaway here is that teachers will need time and support so they can stay abreast of both the well-known and the newer issues that are arising and so they can fully participate in design, selection, and evaluation processes that mitigate risks.

### Challenge: Balancing Human and Computer Decision-Making

One major new challenge with AI-enabled tools for teachers is that AI can enable autonomous activity by a computer, and thus when a teacher delegates work to an AI-enabled tool, it may carry on with that work somewhat independently. Professor Inge Molenaar40 has wondered about the challenges of control in a hybrid teaching scenario: When should a teacher be in control? What can be delegated to a computational system? How can a teacher monitor the AI system and override its decisions or take back control as necessary? Figure 8 expresses the tension around control. To the left, the teacher is fully in control, and there is no use of AI in the classroom. To the right, the technology is fully in control with no teacher involved—a scenario which is rarely desirable. The middle ground is not one dimensional and involves many choices. Molenaar analyzed products and suggests some possibilities: ● The technology only offers information and recommendations to the teacher. ● The teacher delegates specific types of tasks to the technology, for example, giving feedback on a particular math assignment or sending out reminders to students before an assignment is due. ● The teacher delegates more broadly to the technology, with clear protocols for alerts, for monitoring, and for when the teacher takes back control. These and other choices need to be debated openly. For example, we may want to define instructional decisions that have different kinds of consequences for a student and be very careful about delegating control over highly consequential decisions (for example, placement in a next course of study or disciplinary referrals). For human in the loop to become more fully realized, AI technologies must allow teacher monitoring, have protocols to signal a teacher when their judgment is needed, and allow for classroom, school, or district overrides when they disagree with an instructional choice for their students. We cannot forget that if a technology allows a teacher choice—which it should—it will take significant time for a teacher to think through and set up all the options, requiring greater time initially.

### Challenge: Making Teaching Jobs Easier While Avoiding Surveillance

We also recognize that the very technologies that make jobs easier might also introduce new possibilities for surveillance (Figure 9). In a familiar example, when we enable a voice assistant in the kitchen, it might help us with simple household tasks like setting a cooking timer. And yet the same voice assistant might hear things that we intended to be private. This kind of dilemma will occur in classrooms and for teachers. When they enable an AI-assistant to capture data about what they say, what teaching resources they search for, or other behaviors, the data could be used to personalize resources and recommendations for the teacher. Yet the same data might also be used to monitor the teacher, and that monitoring might have consequences for the teacher. Achieving trustworthy AI that makes teachers’ jobs better will be nearly impossible if teachers experience increased surveillance. A related tension is that asking teachers to be “in the loop” could create more work for teachers if not done well, and thus, being in the loop might be in tension with making teaching jobs easier. Also related is the tension between not trusting AI enough (to obtain assistance) or trusting it too much (and incurring surveillance or loss of privacy). For example, researchers have documented that people will follow instructions from a robot during a simulated fire emergency even when (a) they are told the robot is broken and (b) the advice is obviously wrong.41 We anticipate teachers will need training and support to understand how and when they will need to exercise human judgment

### Challenge: Responding to Students’ Strengths While Protecting Their Privacy

Educators seek to tackle inequities in learning, no matter how they manifest locally (e.g. in access to educational opportunities, resources, or supports). In culturally responsive42 and culturally sustaining43 approaches, educators design materials to build on the “assets”—individual, community, and cultural strengths that students bring to learning. Along with considering assets, of course, educators must meet students where they are, including both strengths and needs. AI could assist in this process by helping teachers with customizing curricular resources, for example. But to do so, the data inputted in an AI-enabled system would have to provide more information about the students. This information could be, but need not be, demographic details. It could also be information about students’ preferences, outside interests, relationships, or experiences.44 What happens to this data, how it is deleted, and who sees it is of huge concern to educators. As educators contemplate using AI-enabled technologies to assist in tackling educational inequities, they must consider whether the information about students shared with or stored in an AI-enabled system is subject to federal or state privacy laws, such as FERPA. Further, educators must consider whether interactions between students and AI systems create records that must be protected by law, such as when a chatbot or automated tutor generates conversational or written guidance to a student. Decisions made by AI technologies, along with explanations of those decisions that are generated by algorithms may also be records that must be protected by law. Therein, a third tension emerges, between more fully representing students and protecting their privacy (Figure 10).Further, representation would be just a start toward a solution. As discussed earlier in this report, AI can introduce algorithmic discrimination through bias in the data, code, or models within AIenhanced edtech. Engineers develop the pattern detection in AI models using existing data, and the data they use may not be representative or may contain associations that run counter to policy goals. Further, engineers shape the automations that AI implements when it recognizes patterns, and the automations may not meet the needs of each student group with a diverse population. The developers of AI are typically less diverse than the populations they serve, and as a consequence, they may not anticipate the ways in which pattern detection and automation may harm a community, group, or individual. AI could help teachers to customize and personalize materials for their students, leveraging the teacher’s understanding of student needs and strengths. It is time consuming to customize curricular resources, and teachers are already exploring how AI chatbots can help them design additional resources for their students. An elementary school teacher could gain powerful supports for changing the visuals in a storybook to engage their students or for adapting language that poorly fits local manners of speaking or even for modifying plots to incorporate other dimensions of a teacher’s lesson. In the Learning section, we noted that AI could help identify learner strengths. For example, a mathematics teacher may not be aware of ways in which a student is making great sense of graphs and tables about motions when they are in another teacher’s physics classroom and might not realize that using similar graphs about motion could help with their linear function lesson. AI might help teachers when they seek to reflect student strengths by creating or adapting instructional resources. Yet, the broad equity challenges of avoiding algorithmic discrimination while increasing community and cultural responsiveness must be approached within the four foundations we earlier outlined: human in the loop, equity, safety and effectiveness, and evaluation of AI models. We cannot expect AI models to respect cultural responsiveness. The Department is particularly concerned that equity is something that engaged educators and other responsive adults are in the best position to address and something that is never solely addressable as a computational problem.

### Questions Worth Asking About AI for Teaching

As leaders in both pre-service and post-service teacher education contemplate how AI can improve teaching (along with policymakers, developers, and researchers), we urge all in the ecosystem to spend more time asking these questions: • Is AI improving the quality of an educator’s day-to-day work? Are teachers experiencing less burden and more ability to focus and effectively teach their students? • As AI reduces one type of teaching burden, are we preventing new responsibilities or additional workloads being shifted and assigned to teachers in a manner that negates the potential benefits of AI? • Is classroom AI use providing teachers with more detailed insights into their students and their strengths while protecting their privacy? • Do teachers have oversight of AI systems used with their learners? Are they exercising control in the use of AI-enabled tools and systems appropriately or inappropriately yielding decision-making to these systems and tools? • When AI systems are being used to support teachers or to enhance instruction, are the protections against surveillance adequate? • To what extent are teachers able to exercise voice and decisionmaking to improve equity, reduce bias, and increase cultural responsiveness in the use of AIenabled tools and systems?

### Key Recommendation: Inspectable, Explainable, Overridable AI

In the Introduction, we discuss the notion that when AI is incorporated into a system, the core of the AI is a model. In the Learning section, we discuss that we need to be careful that models align to the learning we envision (e.g., that they aren’t too narrow). Now, based on the needs of teachers (as well as students and their families/caregivers), we add another layer to our criteria for good AI models: the need for explainability.45 Some AI models can recognize patterns in the world and do the right action, but they cannot explain why (e.g., how they arrived at the connection between the pattern and the action). This lack of explainability will not suffice for teaching; teachers will need to know how an AI model analyzed the work of one of their students and why the AI model recommended a particular tutorial, resource, or next step to the student. Thus, explainability of an AI system’s decision is key to a teacher’s ability to judge that automated decision. Such explainability helps teachers to develop appropriate levels of trust and distrust in AI, particularly to know where the AI model tends to make poor decisions. Explainability is also key to a teacher’s ability to monitor when an AI system may be unfairly acting on the wrong information (and thus may be biased. We discuss bias and fairness more in the Assessment section next). Surrounding the idea of explainability is the need for teachers to be able to inspect what an AI model is doing. For example, what kinds of instructional recommendations are being made and to which students? Which students are being assigned remedial work in a never ended loop? Which are making progress? Dashboards in current products present some of this information, but with AI, teachers may want to further explore which decisions are being made and for whom and know of the student-specific factors that an AI model had available (and possibly which factors were influential) when reaching a particular decision. For example, some of today’s adaptive classroom products use limited recommendation models that only consider student success on the last three mathematics problems and do not consider other variables that a teacher would know to consider, such as whether a student has an IEP Plan or other needs. Our call for attending to equity considerations as we evaluate AI models requires information about how discriminatory bias may arise in particular AI systems and what developers have done to address it. This can only be achieved with transparency for how the tools use datasets to achieve outcomes and what data they have available or that a teacher could include in her judgement but are not available to the system (IEP status is offered as an example above). Teachers will also need the ability to view and make their own judgement about automated decisions, such as decisions about which set of mathematics problems a student should work on next. They need to be able to intervene and override decisions when they disagree with the logic behind an instructional recommendation.46 Teachers need protection against adverse ramifications when they assert human judgement over an AI system’s decision.

#### Formative Assessment

Formative assessment is traditionally a key use of edtech because feedback loops are vital to improving teaching and learning.47 As we have emphasized throughout this report, a top priority with AI is to keep humans in the loop and in control, which includes focusing on the people engaged with formative assessments: students, teachers, school leaders, families/caregivers, and others who support learners. In the definition below, please note the overlap between definitions of AI and formative assessment; both have to do with detecting patterns and choosing a future course of action (that adapts to learner strengths and needs).

### Building on Best Practices

A number of dimensions hold potential for shaping the future of formative assessments, and many have ready extensions to the field of AI-enabled systems and tools. For example, the 2017 NETP discussed how technology can lead to improved formative assessments along seven dimensions, listed below: 1. Enabling Enhanced Question Types: to give students more ways to show what they know and can do. 2. Measurement of Complex Competencies: to better elicit growth in important skills that go beyond typical subject matter standards, for example, in measuring practices, social skills like teamwork, self-regulation, and work-relevant skills (e.g., making presentations or leading teams). 3. Providing Real-Time Feedback: to maintain and increase student engagement and to support effective learning, providing timely and helpful responses and suggestions to each learner. 4. Increasing Accessibility: to include neurodiverse learners and to engage learners’ best communication capabilities as they share what they know and can do.5. Adapting to Learner Ability and Knowledge: to make assessments more precise and efficient. 6.

Embedded Assessment in the Learning Process: to emphasize an assessment’s role in improving teaching and learning (this report does not focus on assessment for accountability purposes). 7.

Assess for Ongoing Learning: to reveal progress over time and not just predetermined milestones. AI models and AI-enabled systems may have potential to strengthen formative assessments. In one example, a question type that invites students to draw a graph or create a model can be analyzed with AI algorithms,49 and similar student models might be grouped for the teacher to interpret.

Enhanced formative assessment may enable teachers to better respond to students’ understanding of a concept like “rate of change” in a complex, real-world situation. AI can also give learners feedback on complex skills, such as learning American Sign Language50 or speaking a foreign language51 and in other practice situations where no person is available to provide immediate feedback. Generally, an AI assistant may be able to reduce the load for teachers related to grading simpler aspects of student responses, allowing the teacher to focus their specialized judgment on important qualities of a whole essay or a complex project. We also may be able to better provide feedback with accessibility. For example, an AI-enabled learning technology may be able to interact verbally with a student about their response to an essay prompt, asking questions that guide the student to clarify their argument without requiring the student to read a screen or type at a keyboard. In the examples shared earlier in the Learning section, we also see that AI can be embedded in the learning process, providing feedback to students as they work to solve a problem, rather than only later after the student has reached a wrong answer. When formative assessment is more embedded, it can better support learning, and timely feedback is critical.52 Although there are many points of connection like these between AI and formative assessments, our listening sessions also revealed attendees’ desire to tackle some existing shortcomings in the field of formative assessment; namely, the time-consuming and sometime onerous nature of taking tests, quizzes, or other assessments and the lack of perceived value in the feedback loop by teachers and students.

### Implications for Teaching and Learning

Real-time instructional feedback can be beneficial when it helps learners and teachers to improve. But common experience too often leaves students and teachers with unpleasant feelings toward assessment and thus poses a provocative conflict between the potential benefits of data collected through formative assessments and the practical implications of administering additional assessments in classrooms and schools. Some AI-enabled systems and tools seek to address this potential conflict. For example, one AIenabled reading tutor listens to students as they read aloud and provides on-the-spot feedback to improve their reading.53 Students reportedly enjoyed reading aloud, and the approach was effective. Researchers have also embedded formative assessments in games so that students can show how well they understand Newtonian physics as they play increasingly difficult levels of a game.54 If a student can more easily ask for and receive help when they feel frustrated or confused, reducing those feelings can feel encouraging. Student feelings of safety, confidence, and trust in the feedback generated by these AI-enabled systems and tools are essential to showcase their learning. That focus on learning growth and gains is optimal (absent negative consequences or a high-stakes environment).55 AI-enhanced formative assessments may have the potential to save teachers’ time (e.g., time spent on grading), allowing the instructor to spend more time engaged in helping students. AIenhanced assessments may also benefit teachers if they provide detailed insights about student strengths or needs that may not be visible and if they support instructional adaptation or improvement by suggesting a small set of evidence-based recommendations for helping students master content. Such assessments may also be helpful outside of the classroom if it can provide feedback when the teacher is not available, for example, in completing homework or practicing a concept during study hall. As we discussed in the Teaching section, an essential aspect of deploying AI-based formative assessment must be centering teachers in system design.

### Insight: AI Can Enhance Feedback Loops

The term “formative assessment” does not singularly mean a test or a measurement. Assessment becomes formative when it results in useful reflections and changes to the course of teaching, learning, or both.56 The term “feedback loops” emphasizes that measurement is only part of the process. Feedback loops that lead to instructional improvement—including adaptations in teaching and learning—yield the strongest outcomes for students. We also use “feedback loops” as a plural term because there are many types and levels of loops that are important. Students can benefit from feedback when they work individually, as a member of a small group, or in a classroom discussion.

Feedback loops are valuable “in the moment”—for example, as a student practices a skill. Further, feedback loops are valuable when they cover larger spans of effort and reflections, such as at the end of presenting a project or term paper. In addition, feedback loops can assist teachers, for example, helping them notice continuous improvement of products and the implementation of programs. Due to the importance of feedback loops, formative assessment could be a leading area for schools’ explorations of powerful uses of AI in teaching and learning. Educators can build upon alignments between their long-standing visions for formative assessment and the emerging capabilities that AI holds. Further, the professional assessment community brings a toolkit for asking and answering questions about topics like bias and fairness. The psychometric toolkit of methods is a strong start toward the questions that must be asked and answered because it already contains ways to measure bias and fairness and, more generally, to benchmark the quality of formative assessments. But as our discussion reveals, AI can only make feedback loops better if we keep a firm eye on the weaknesses of AI and how AI introduces new concerns.

### An Example: Automated Essay Scoring

One instructive example is Automated Essay Scoring (AES). To become strong writers, which is a valuable life skill, students need regular and specific feedback. However, reviewing and providing feedback on essays is very time consuming for humans. Hence, Ellis Page provided a first vision for computer programs that could review and provide feedback on student essays in 196657 , and much effort has gone into AES technologies in the intervening 56 years. Many research review articles are available to summarize the progress, which has been impressive.58

Further, some of today’s applications of AES technologies will be familiar to readers, such as Grammarly, Turnitin, and the various essay analysis engines used by publishers and assessment companies. Also note that while the traditional AES functionality emphasizes scoring or rating essays, newer AI-enabled products focus more on providing students with constructive criticism and developing their skills as writers. Writing is a life skill that is important to the pursuit of college and career ambitions, and developing writers require comprehensive feedback. If developers could inexpensively augment human feedback to developing writers with AI feedback, it’s possible that support for learning to write could become more equitable. And yet, AES is an instructive example because researchers have analyzed limitations, too.59 AES technologies in AI can analyze some features of student essays but can also be misled by the length of an essay, by a student who places appropriate keywords in sentences that don’t make sense, and other flaws that a human reader would easily notice. In a telling quote, one team that reviewed the state of the art wrote this: The authors further note that while human and AI judgements of essays may correlate, people and computers are not noticing the same things in student writing. Due to these limitations, we must continue to emphasize a human in the loop foundation for AI-enhanced formative assessment. AI may support but not replace high-quality, human-led processes and practices of formative assessment in schools.

### Key Opportunities for AI in Formative Assessment

Based on the listening sessions we held, we see three key areas of opportunity in supporting formative assessment using AI systems and models. First, we recommend a strong focus on measuring what matters61 and particularly those things that have not been easily measured before and that many constituents would like to include in feedback loops. The example above, AES, was chosen because writing remains a valuable academic, workplace, and life skill. Looking at community goals through the lens of their visions for their high school graduates, we see that families/caregivers, students, and community leaders want to nurture graduates who solve problems adaptively, who communicate and collaborate well, who persevere and self-regulate when they experience challenges. “What matters” today reaches beyond a sole focus on the core academic content measured by large-scale summative assessments, to support students and teachers with actionable feedback that nurtures the broader skills students need to succeed and thrive. Further, within core academic content, AI may help us to provide feedback on the more realistic and complex aspects of doing math, for example, or investigating scientific phenomena, understanding history, or discussing literature. Second, we’d like to see a strong focus on improving help-seeking and help-giving. 62 Asking for and giving help is crucial to learning63 and practicing a growth-mindset and central to the notion of human feedback loops. Students may not always know when they need help. In one example, computer algorithms can detect a student who is “wheel spinning” (working hard on mastering content but not making progress).64 A student who is working hard may not feel like they need help, and the teacher may not be aware that the student is struggling if he or she appears to be “on task.” AI may also be helpful by highlighting for students and teachers what forms of assistance have been most useful to the student in the recent past so that an educator can expand access to specific assistance that works for that individual student. Finally, educators may learn things from AI-enabled systems and tools that give feedback and hints during the completion of homework, utilizing that feedback to later reinforce concepts in direct instruction and strengthen the one-on-one support provided to students. 65 AI-enabled systems and tools can provide teachers with additional information about the students’ recent work, so their instructor has a greater contextual sense as they begin to provide help. Third, we advocate for teachers and students to be strongly involved in designing feedback loops as developers produce AI-enhanced formative assessments so they can directly voice what would make assessments less onerous and more convenient and valuable to them.66 Earlier in the Teaching section, we emphasized how important it is to involve teachers in designing, selecting, and evaluating AI-enhanced technologies. Students need to be centered, too. They are experiencing AI in their everyday lives, and they have strong opinions on what is valuable and safe. There are local and cultural variations in how people provide and receive feedback, so adjusting feedback to align with community norms is important.

### Key Recommendation: Harness Assessment Expertise to Reduce Bias

Bias and fairness are important issues in assessment design and administration, 67 and they hold relevance for the area of AI-enabled assessment. In traditional assessment, a test item might be biased if unnecessary details are included that differentially advantage some students (e.g., a storybased item that references a sport that only boys play regularly may be less helpful to girls). As discussed earlier, with AI, we now must worry about algorithmic discrimination which can arise due to the manner in which AI algorithms are developed and improved from large datasets of parameters and values that may not represent all cohorts of learners. Algorithmic discrimination is not just about the measurement side of formative assessment; it is also about the feedback loop and the instructional interventions and supports that may be undertaken in response to data collected by formative assessments. There is a question both about access to such interventions and the quality or appropriateness of such interventions or supports. When an algorithm suggests hints, next steps, or resources to a student, we have to check whether the help-giving is unfair because one group systematically does not get useful help which is discriminatory. Fairness goes beyond bias as well. In AI-enabled formative assessment, both the opportunity to learn through feedback loops, as well as the quality of learning in and outside of such loops, should be addressed. Issues of bias and fairness have arisen in traditional assessments, and the field of psychometrics has already developed valuable tools to challenge and address these issues.68 Assessment as a field may have a head start on tackling bias and fairness for AI in education. And yet the issues expand with AI, so the work is not done. Strong and deliberate attention to bias and fairness is needed as future formative assessments are developed.

### Related Questions

As indicated, formative assessment is an area in which AI is expanding along a continuum that can be guided by visions already in place, such as the 2017 NETP. It is an area in which AI is poised to grow, especially with capabilities that power more feedback loops in student learning.

As this growth takes place, we suggest ongoing attention to the following questions: ● Is formative assessment bringing benefits to the student learning experience and to the efficacy of classroom instruction? ● Are humans being centered in AI-enabled formative assessment and feedback loops? ● Are we providing empowering professional development to teachers so they can leverage feedback loops and safeguard against concerns? ● To what extent are the developers and implementers of AI-enabled systems and tools tackling new sources of algorithmic bias and continuing to make assessment fairer? ● Are governance policies regarding who owns, controls, and can view or use AI-enabled formative assessment data appropriate and adequate? ● Do we have sufficient guardrails against misuse of formative assessment data or automatically generated interpretations of student achievement and learning, such as on dashboards? ● Is trust in an AIenabled assessment system, feedback loops, and data generated by such assessments growing or diminishing?

#### Research and Development

Policy relies upon research-based knowledge; likewise, improving practice depends on feedback loops that analyze empirical evidence. Consequently, the 2010 NETP specified a series of “grand challenges” which were “R&D problems that might be funded and coordinated at a national level.” One 2010 NETP grand challenge was to create personalized learning systems that continuously improve as they are used: Since 2010, much R&D has addressed this challenge. Conferences about learning analytics, educational data mining, and learning at scale have blossomed. Developers have created platforms that use algorithms and the analysis of big data to tune learning experiences. The challenge has not been fully achieved, and further work on this challenge is still relevant today.

### Insight: Research Can Strengthen the Role of Context in AI

Despite the relevance of 2010’s grand challenges, it has become apparent that the R&D community is now looking to expand their attention. The 2010 challenges were stated as technical problems.

Today’s researchers want to more deeply investigate context, and today’s tech companies want to develop platforms that are responsive to the learners’ characteristics and situations more broadly— not just in terms of narrow cognitive attributes. We see a push to transform R&D to address context sensitivity. We look forward to new meanings of “adaptive” that broaden outward from what the term has meant in the past decade. For example, “adaptive” should not always be a synonym of “individualized” because people are social learners. Researchers therefore are broadening “adaptivity” to include support for what students do as they learn in groups, a form of learning that is prevalent in schools across the U.S. The focus on context is not an accident. Context is a traditional challenge in AI.70 Thus, researchers and developers are wise to prioritizing context. Unless we invest more in AI that is context-sensitive, it is quite likely that AI will break and fail to achieve educational goals. Agreeing to prioritize context won’t be easy. As illustrated above in Figure 12, there will be a tension between depth of context and pace of technological advances in AI R&D. On the one hand, AI is sometimes presented as a race to be the first to advance new techniques or scale new applications—innovation is sometimes portrayed as rapidly going to scale with a minimally viable product, failing fast, and only after failure, dealing with context. On the other hand, researchers and developers see that achieving good innovations with AI in education will clearly require bringing more context into the process early and often. For example, researchers highlight that humans must be continually adjusting the goals for technology and have noted that when we set forth goals, we often don’t yet fully understand context; and as we learn about context, the goals must change.71 This suggests that context must be prioritized early and habitually in R&D; we don’t want to win a race to the wrong finish line. Further, intensifying focus on context in this work will change the nature of the R&D. There won’t be just one type of change in R&D because context has multiple meanings. Attendees in our listening sessions described four types of context necessary for the future. We list these four types of context below and then expand on each one in its own section. These four types emerged as topics of provocations to think differently about R&D but certainly do not exhaust the important ways of investigating context. 1.

Focus on the Long Tail: How could we use big data and AI to pay more attention to the “long tail” of edtech use—going beyond a few “most typical” ways of using emerging technology and instead solving for digital equity and inclusion? 2. Partnership in Design-Based Research: How can we change who is involved and influential in designing the future of AI in education to more centrally include students, teachers, and other educational constituents? 3. Connect with Public Policy: How can work on AI in education build on general advances in AI ethics, safety, and regulation and contribute additional advances specific to educational policy? 4. Rethink Teacher Professional Development: How can we solve for new systems of teacher professional development (both preservice and in-service) that align to the increasingly core role of technology in the teaching profession?

### Attention to the Long Tail of Learner Variability

At the core of R&D of AI in education, innovators will be building models that fit available data. The increasing scale and prevalence of technologies means that the data is coming from and including a wide range of different contexts and varied ways that people in those contexts engage in teaching and learning. Researchers in our listening sessions drew attention to the promise of AI for addressing “context” by reference to the long tail of learner variability As depicted in Figure 13, learners vary in their strengths and needs. The most frequently occurring mix of strength and needs (also known as “teaching to the middle”) is depicted leftmost, with less frequently occurring mixes spreading to the right. Rising upward, the figure depicts the number of learners who benefit from a particular learning design, pathway, or approach. We argue that AI can bring opportunities to address a wider spectrum of strengths and needs but only if developers and innovators focus on the long tail and not only “teaching to the middle.” For the sake of argument, the figure indicates three zones. In a first zone, curricular resources are mostly standardized, with perhaps a dimension or two of adaptivity. For example, many existing products adapt based on the correctness of student answers and may also provide options to read or hear text in a second language. However, the core of the instructional approach is highly standardized. In a second zone, there is greater balance between how much standardization and how much adaptivity students can access. Universal Design for Learning (UDL) is one set of recommendations for providing learning opportunities in multiple formats and for accommodating different learning progressions.72 UDL can enable accommodating more ways in which learners vary, and as teachers know, there are many more important ways to adapt to students than found in today’s edtech products. Students are neurodiverse. They bring different assets from their experiences at home, in their communities, and in their cultures. They have different interests and motivations. And they learn in varied settings—classrooms and schools differ, and at-home students learn in informal settings in ways that could complement school learning. These are all important dimensions of “context.” Zone 3 indicates highly adaptive learning, where standardization is less successful and where we need to discover a wider variety of approaches to engage learners and sustain powerful learning.

Researchers in our listening sessions noted the promise of Zone 3 because AI’s ability to recognize patterns in data can extend beyond the most common patterns and because AI's ability to generate customized content can extend beyond what people can reasonably generate on their own. Notice that although the Zone 1 bar appears to be the tallest, and thus tends to attract initial attention, there are more students in Zones 2 and 3, the regions where AI can provide more help. Thus, it’s important to ask where AI researchers and developers are directing their attention. When we say a model “fits,” are we saying it fits the most common and typical uses by teachers and learners? This sort of R&D is easier to do. However, machine learning and AI also can tailor a model to the less common and more culturally specific contexts, too. Therefore, how can constituents cultivate interdisciplinary expertise to direct attention among researchers and developers to focus on the long tail? If we do, the quality of what we do for those represented in that tail can be more adaptive and more context-sensitive. And to be most effective, it will require the integration of contextual, content, and technical expertise. Within the long-tail challenge, the community is wondering how we can get to research insights that are both general and specific enough. When research produces very general abstractions about learning, it often doesn’t give developers enough guidance on exactly how to adjust their learning environments. Conversely, when research produces a specific adaptive algorithm that works on one educational platform, it often remains hard to apply to additional platforms; research can be too detailed as well. The research community is also thinking about new partnerships that could bring more data and more diverse perspectives to the table, the topic of the next section. Focusing on the long tail of learner variability is particularly important to addressing a longstanding key research question: “Do new AI-enhanced approaches work to improve learning, and for whom and under what conditions?”

### Partnership in Design-Based Research

Of course, teachers must be included in rethinking their own professional development. This thought leads to another priority aspect of context: partnership in design-based research. With regard to inclusive design, attendees in our listening sessions brought up a variety of co-design73 and other participatory processes and goals that can be used in R&D.74 By co-design, they mean sharing power with non-researchers and non-developers through all the phases of design and development, which would result in more influence by teachers, students, and other constituents in the shape of AI-enabled edtech. The shift toward co-design was palpable throughout our listening sessions, but as researchers and developers have not standardized on one particular codesign method, we share some representative examples. ● Youth can powerfully participate in design when researcher methods include participant co-design. Such research can investigate how to improve edtech while educating students. A listening session attendee asked about developing students’ awareness of what data are being collected and how data are being used by developers. ● There is a near future need to go beyond representation so that co-designed solutions consider more generous contexts for broader possibilities, according to attendees. ● The shift of power dynamics is another research-worthy interest of the panel and attendees to understand the balance between a teacher’s agency and a machine’s suggestions. ● Likewise, such longitudinal research will require both the infrastructure and institutional support to fund necessary experimentation and requisite failures to elicit positive results and safe innovation. ● There is a desire for rapid cycle evaluations with inclusive feedback loops that return to the educators themselves as essential relative to traditional research approaches. ● Many researchers also mentioned a focus on explainable AI as essential to enable participation in the design and evaluation of emerging AI approaches in education. The conversations raised this question: how can co-design provide an empowering form of participation in design and thus achieve digital inclusion goals? Such digital inclusion can span many layers of design, including diverse representation in design of policies around data, design of adaptivity, and other user experiences in AI systems, design of plans for cultivating AI literacy for users of new platforms, and lastly, the design of plans to evaluate systems.

### Re-thinking Teacher Professional Development

With regard to teachers as professionals, both researchers and other educators attending our listening sessions were highly concerned about the disconnect between how teachers are prepared versus how they are expected to work with emerging technology. When we discuss learning, teachers are central actors, and thus the contexts in which they are prepared is centrally important to their ability to do great work in current and emerging technological environments. Teacher professional development, professional learning, and leadership (PD or PL) for emerging technologies was seen as an area needing intense re-thinking, and research could lead the way. Today, few who prepare to become a teacher in an established pre-service program learn about the effective use of educational technology in schools and classrooms; those who do have the opportunity to investigate technology rarely think about the structures that shape its use in the classroom and in educational leadership. Consequently, a troubling dichotomy arises between a small set of investigators who specifically consider educational technology in their research on teaching and a broader group of educators who see educational technology as a generic instructional resource. The challenge is high because teacher professional development will remain highly varied by local contexts. Yet insufficient attention to teachers as leaders in the use and further development of effective educational technology is widespread in teacher professional development research. One response can be in terms of investigating how to nurture greater AI literacy for all teachers. AI literacy is not only important to protect educators and students from possible dangers but also valuable to support teachers to harness the good and do so in innovative ways. A panelist reminded the group that this work implies how we prepare educators with a baseline AI literacy and understanding. More transparency and authentic dialogue can foster trust, which was mentioned by a researcher as a chief concern for all teachers and students. This is not to suggest that AI literacy is a complete or even a simple fix. Researchers want to ask fundamental questions about what it means for teachers to be professionals, especially as emerging technologies gain ground in schools and classrooms—our teachers’ professional workplaces. Researchers want to broadly reconceptualize teacher professionalism and to stop treating technology as an add-on element of professional development.

### Connecting with Public Policy

Defining human-centered AI for education requires the embrace of a human-centered principle and foundation for developing and formulating policies that govern the application and use of AI more generally throughout society. For example, power dynamics that arise between companies and consumers in society around issues like data ownership will also arise in the education-specific ecosystem. Further, the public discourse in which people are discussing ethics, bias, responsibility, and many other necessary concepts will be happening simultaneously in public policy and in educational ecosystems. One clear implication in our listening sessions was that efforts to improve AI literacy in education could be important and helpful to society more generally. For example, one panelist said that an overarching goal of improving AI literacy is necessary if they are to contribute to how those technologies are designed. Another researcher was interested in how edtech can provide environments where students can experience having difficult discussions across perspectives, an issue which is endemic to present society. A third researcher noted the insufficiencies of prior efforts to contend with algorithmic bias, ethics, and inclusion due to a classroom’s complex social dynamics. Researchers want to take a lead in going beyond checkbox approaches to take these issues seriously. And they also acknowledge that engaging with policy is often a new form of context for edtech and AI researchers, many of whom don’t have long experiences in policy arenas. Likewise, developers often do have experience with some policy issues, such as data privacy and security, but are now needing to become part of new conversations about ethics, bias, transparency, and more, a problem that the EdSAFE AI Alliance is addressing through multisector working groups and policy advocacy.75

### Key Recommendation: Focus R&D on Addressing Context

Attendees who have participated in listening sessions leading up to this report were exceptionally clear that their view of future R&D involved a shift from narrow technical questions to richer contextual questions. This expansive shift toward context, as detailed below, is the foundational orientation that the listening session attendees saw as being necessary to advancing R&D.

Attendees included these as dimensions of context: • learner variability, e.g., in disabilities, languages spoken, and other relevant characteristics; • interactions with peers, teachers, and others in the learning settings; • relationships across home, school, and community settings, including cultural assets; • instructional resources available while learning; • teacher preparation; and • policies and systems that structure teaching and learning. To more fully represent the context of teaching and learning, including these and other dimensions of text, researchers will have to work in partnership with others to understand which aspects of context are most relevant to teaching and learning and how they can be usefully incorporated into AI models.

### Ongoing Questions for Researchers

As mentioned earlier, people are good at context; AI—not so much. R&D investment in contextrich edtech thus could serve multiple national interests because finding ways to do a better job with context would be a fundamental advancement in AI. Indeed, questions like these reverberate across all applications of AI in society, and education is a centrally good context for investigating them: ● Are AI systems moving beyond the tall portions of the “long tail” to adapt to a greater range of conditions, factors, and variations in how people learn? ● To what extent are AI technologies enhancing rather than replacing human control and judgment of student learning?

● How will users understand the legal and ethical implications of sharing data with AI enabled technologies and how to mitigate privacy risks? ● To what extent does technology account for the complex social dynamics of how people work and learn together, or is technology leading humans to narrow or oversimplify? ● How can we more clearly define what we mean by a context-sensitive technology in terms that are both concrete and broad enough? How can we measure it? ● To what extent are technical indicators and human observations of bias or unfairness working together with human observations? How can concerns about ethics and equity in AI technologies become actionable both in R&D, and later, when AI is widely used? ● Are we learning for whom and under what conditions AI systems produce desired benefits and impacts and avoid undesirable discrimination, bias, or negative outcomes?

### Desired National R&D Objectives

Attendees sought immediate progress on some key R&D issues, such as these: • Clarifying and achieving a consensus on the terms that go beyond data privacy and data security, including ideas like human-centered, value-sensitive, responsible, ethical, and safe so constituents can advocate for their needs meaningfully and consistently • Creating and studying effective programs for AI literacy for students, teachers, and educational constituents in general, including literacy with regard to the ethics and equity issues specific to AI in educational settings • Advancing research and development to increase fairness, accountability, transparency, and safety in AI systems used in educational settings • Defining participatory or co-designed research processes that include educators in the development and conduct of research related to the development, use, and efficacy of AIenabled systems and tools • Highlighting and advancing R&D efforts that empower the participation and voices of youth regarding research, data, and design of AI applications for teaching and learning Longer term desires for a national R&D program include some of the following objectives: • Funding sustainable partnerships that uncover what context means and how it can be addressed over longer periods of time • Better connecting goals for “broadening participation” (for example, in STEM learning pathways) to strategies for addressing learner variability and diversity • Prioritizing research to revitalize support for instructors in light of the increasingly technological nature of K-12, higher education, and workplace learning settings • Creating infrastructure and new ways of working together beyond individual fieldinitiated grants so that R&D with big data and leveraging emerging AI capabilities becomes safer and more productive

#### Recommendations

Earlier, we asked two guiding questions: 1. What is our collective vision of a desirable and achievable educational system that leverages automation while protecting and centering human agency? 2. On what timeline will we be ready with necessary guidelines and guardrails along with convincing evidence of positive impacts, so that we can ethically and equitably implement this vision widely? Answers to the first question are provided throughout the Learning, Teaching, Assessment, and Research sections. This section turns to a call to action to education leaders and to recommendations. Core to the Department’s perspective is that education will need leadership specific to our sector. Leadership should recognize and build on prior accomplishments in edtech (such as strong prior work on student privacy and school data security) as well as broad frameworks for safe AI (such as the Blueprint for an AI Bill of Rights). Leadership must also reach beyond these accomplishments and frameworks to address emerging opportunities and risks that are specific to novel capabilities and uses of AI in education.

### Insight: Aligning AI to Policy Objectives

Individual sections of this policy report provided insights in each of four areas—learning, teaching, assessment, and research. These insights, synthesized from extensive stakeholder consultation and listening sessions, show that the advances in AI can bring opportunities to advance the Department’s policy objectives: ● In support of our objective of attracting and retaining teachers, our nation could focus on AI assistants that make teaching jobs better and provide teachers with the information they need to work closely and empathically with students. An emphasis on teachers in the loop could ensure that AI-enabled classroom technologies keep teachers in the know, in touch with their students, and in control of important instructional decisions. Keeping the teacher

in the loop is important to managing risks, as well. ● In support of equitable learning, especially for those most affected by the pandemic, AI could shift edtech from a current deficit-based model to a strengths-based alternative. In addition to finding student weaknesses and assigning fixes, edtech could make recommendations based on strengths that students bring to learning and how adapting to the whole student—a cognitive, social, and self-regulating person—could enable more powerful learning. Adapting to the whole student should include supporting students with disabilities as well as English learners. With regard to equity, we must remain highly attuned to the challenges of bias (which are inherent to how AI systems are developed) and take firm action to ensure fairness. ● With regard to growth trajectories to successful careers, AI-enabled assessments could provide students and teachers with formative guidance on a wider range of valuable skills, focusing on providing information that enhances learning. Aligned with the human-centric view, we should take a systems view of assessments where students, teachers, and others remain at the center of instructional decision making ● With regard to equity, as research advances and brings more context into AI, we will be better able to use AI to support goals that require customization of learning resources, such as enabling teachers to more easily transform materials to support neurodiverse learners and increase responsiveness to local communities and cultures. Going forward, educational leaders need to bring these and their own policy priorities to the table at every discussion about AI, driving the conversation around human priorities and not only their excitement about what new technology might do. Fundamentally, AI seeks to automate processes that achieve goals, and yet, AI should never set goals. The goals must come from educators’ vision of teaching and learning and educators’ understanding of students’ strengths and needs.

### Calling Education Leaders to Action

We summarize seven recommendations for policy action. These recommendations are for education leaders. In the introduction, we note the necessity of involving education constituents in determining policies for AI. We also observed throughout our listening sessions that people coming from many different roles in education all have passion, knowledge, and insights to contribute. In our view, all types of constituents can be education leaders. We are reluctant to suggest any constituent role is more important to advance any of the recommendations, but we call out specific needs for action within some of the recommendations where it is warranted.

### Recommendation #1: Emphasize Humans in the Loop

We start with a central recommendation throughout this report. This recommendation was a clear constituent favorite. Indeed, across more than 700 attendees in our listening sessions, the predominant discussion tackled how constituents can achieve a consensus vision for AI-enabled edtech where humans are firmly at the center. The Blueprint for an AI Bill of Rights similarly calls for “access to timely human consideration and remedy by a fallback and escalation process if an automated system fails, it produces an error, or you would like to appeal or contest its impacts…” Building on this consensus, we call upon all constituents to adopt “humans in the loop” as a key criterion for educational use of AI. We envision a technology-enhanced future more like an electric bike and less like robot vacuums. On an electric bike, the human is fully aware and fully in control, but their burden is less, and their effort is multiplied by a complementary technological enhancement. Robot vacuums do their job, freeing the human from involvement or oversight. Although teachers should not be the only humans involved in loops, Figure 5 provided examples of three types of teacher loops that are central to education and can be used to illustrate what

“human in the loop” means. Here, we use the example of an AI chatbot to elaborate on the meaning of the loops. First, as students become involved in extended interactions with AI chatbots, teachers will need to educate students about safe AI use, monitor their use, and provide human recourse when things go astray. Second, teachers are beginning to use chatbots to plan personalized instruction for their students; they will need to be involved in loops with other teachers to understand effective prompts, to know how to analyze AI-generated lesson plans for flaws, and to avoid the human tendency to overly trust AI systems and underapply human judgement. Third, teachers need to be involved in the design and evaluation of AI systems before they are used in classrooms and when needs for improvement are observed. In one example, to design AI-generated homework support for students, teachers’ in-depth understanding of the cognitive, motivational, and social supports their students need will provide much-needed guidance as a homework-support chatbot is designed. In framing AI in education, this report advances a key recommendation of

“human in the loop” AI because the phrase readily communicates a criterion that everyone can use as they determine which AI-enabled systems and tools are appropriate for use in teaching and learning. In a rather technical field, human in the loop is an approachable and humanistic criterion. Rather than suggesting that AI-enabled systems and tools should replace teachers, this term instead solidifies the central role of educators as instructors and instructional decision makers, while reinforcing the responsibility of teachers to exercise judgement and control over the use of AI in education. It resonates with the important idea of feedback loops, which are highly important to how people teach and learn. It also aligns with the ideas of inspectable, explainable, severable, and overridable AI. The Department agrees with listening session participants who argued that teachers should not be the only humans in the loop and calls upon parents, families, students, policy makers, and system leaders to likewise examine the “loops” for which they are responsible, critically analyze the increasing role of AI in those loops, and determine what they need to do to retain support for the primacy of human judgement in educational systems.

### Recommendation #2: Align AI Models to a Shared Vision for Education

As we have discussed across every section of this report, AI technologies are grounded in models, and these models are inevitably incomplete in some way. It is up to humans to name educational goals and measure the degree to which models fit and are useful—or don’t fit and might be harmful. Such an assessment of how well certain tools serve educational priorities may seem obvious, but the romance of technology can lead to a “let’s see what the tech can do'' attitude, which can weaken the focus on goals and cause us to adopt models that fit our priorities poorly. Here we call upon educational policy and decision makers at the local, state, and federal level to use their power to align priorities, educational strategies, and technology adoption decisions to place the educational needs of students ahead of the excitement about emerging AI capabilities. We want to strengthen their attention to existing state, district, and school-level policies that guide edtech adoption and use, such as the four levels of evidence in ESSA, the privacy requirements of FERPA, and enhanced policies to come. Local education leaders know best what their urgent educational priorities are. Every conversation about AI (or any emerging technology) should start with the educational needs and priorities of students front and center and conclude with a discussion about the evaluation of effectiveness re-centered on those needs and priorities. Equity, of course, is one of those priorities that requires constant attention, especially given the worrisome consequences of potentially biased AI models. We especially call upon leaders to avoid romancing the magic of AI or only focusing on promising applications or outcomes, but instead to interrogate with a critical eye how AI-enabled systems and tools function in the educational environment. We ask leaders to distrust broad claims and ask six types of questions, listed below. Throughout this report, we elaborated on which characteristics of AI model use in education are most important to evaluate for alignment to intended educational goals. To aid leaders, we summarize our insights about AI models and their use in educational tools and systems in Figure 14. In this figure, we center teaching and learning in all considerations about the suitability of an AI model for an educational use. Humans remain in the loop of defining, refining, and using AI models. We highlight the six desirable characteristics of AI models for education (elaborating from principles in the Blueprint for an AI Bill of Rights to fit the specifics of educational systems): 1. Alignment of the AI Model to Educators’ Vision for Learning: When choosing to use AI in educational systems, decision makers prioritize educational goals, the fit to all we know about how people learn, and alignment to evidence-based best practices in education. 2. Data Privacy: Ensuring security and privacy of student, teacher, and other human data in AI systems is essential. 3. Notice and Explanation: Educators can inspect edtech to determine whether and how AI is being incorporated within edtech systems. Educators’ push for AI models can explain the basis for detecting patterns and/or for making recommendations, and people retain control over these suggestions. 4. Algorithmic Discrimination Protections: Developers and implementers of AI in education take strong steps to minimizing bias and promoting fairness in AI models. 5. Safe and Effective Systems: The use of AI models in education is based on evidence of efficacy (using standards already established in education for this purpose) and work for diverse learners and in varied educational settings. 6. Human Alternatives, Consideration and Feedback: AI models that support transparent, accountable, and responsible use of AI in education by involving humans in the loop to ensure that educational values and principles are prioritized. Although we first address our recommendation to interrogate how educational systems use AI models to educational leaders who adopt technologies, other leaders also have integral roles to play. Teachers and students, as well as their families/caregivers, contribute significantly to adoption decisions also. And leaders and parents must support educators when they question or override an AI model based on their professional wisdom. Developers of technologies need to be forthcoming about the models they use, and we may need policymakers to create requirements for disclosure so that the marketplace can function on the basis of information about AI models and not only by the claims of their benefits. We also emphasize the need for a government role. AI models are made by people and are only an approximation to reality. Thus, we need policies that require transparency about the AI models that are embedded in educational systems, as well as models that are inspectable, explainable, and overridable. Our listening sessions featured constituent calls for government doing more to hold developers accountable for disclosing the types of AI models they employ in large-scale products and the safeguards included in their systems. Government leaders can make a positive contribution to market conditions that enable building trust as AI systems are procured and implemented in education. We discuss these guidelines more in recommendation #4, which is about building trust.

### Recommendation #3: Design Using Modern Learning Principles

We call for the R&D sector to ensure that product designs are based on best and most current principles of teaching and learning. The first decade of adaptivity in edtech drew upon many important principles, for example, around how to sequence learning experiences and how to give students feedback. And yet the underlying conception was often deficit-based. The system focused on what was wrong with the student and chose pre-existing learning resources that might fix that weakness. Going forward, we must harness AI’s ability to sense and build upon learner strengths. Likewise, the past decade of approaches was individualistic, and yet we know that humans are fundamentally social and that learning is powerfully social. Going forward, we must build on AI capabilities that connect with principles of collaborative and social learning and which respect the student not just for their cognition but also for the whole human skill set. Going forward, we also must seek to create AI systems that are culturally responsive and culturally sustaining, leveraging the growth of published techniques for doing so. Further, most early AI systems had few specific supports for students with disabilities and English learners. Going forward, we must ensure that AI-enabled learning resources are intentionally inclusive of these students. The field has yet to develop edtech that builds upon each student’s ability to make choices and to self-regulate in increasingly complex environments. We have to develop edtech that expands students’ abilities to learn in creative modes and to expand their ability to discuss, write, present, and lead. We also call upon educators to reject uses of AI that are based solely on machine learning from data—without triangulation based on learning theory and knowledge from practice. Achieving effective and equitable educational systems requires more than processing “big data,” and although we want to harness insights from data, human interpretation of data remains highly important. We reject a technological determinism in which patterns in data, on their own, tell us what to do. Applications of AI in education must be grounded in established, modern learning principles, the wisdom of educational practitioners, and should leverage the expertise in the educational assessment community around detecting bias and improving fairness.

### Recommendation #4: Prioritize Strengthening Trust

Technology can only help us to achieve educational objectives when we trust it. Yet, our listening sessions revealed the ways in which distrust of edtech and AI is commonplace. Constituents distrust emerging technologies for multiple reasons. They may have experienced privacy violations. The user experience may be more burdensome than anticipated. Promised increases in student learning may not be backed by efficacy research. They may have experienced unanticipated consequences. Unexpected costs may arise. Constituents may distrust complexity. Trust needs to incorporate safety, usability, and efficacy. The Department firmly takes the stance that constituents want AI that supports teachers and rejects AI visions that replace teachers. And yet, teachers, students, and their families/caregivers need support to build appropriate levels of trust in systems that affect their work. In the broader ecosystem, trustworthy AI is recognized as a multidimensional problem (including the dimensions of Figure 14, above). If every step forward does not include strong elements of trust building, we worry that distrust will distract from innovation serving the public good that AI could help realize. We expect that associations and societies have a key role in strengthening trust. Some important associations like the State Educational Technology Directors Association and the Consortium for School Network work with edtech leaders, and parallel organizations like EDUCAUSE work with postsecondary leaders.

Other associations and societies work with teachers, education leaders, and education staff developers. Industry networks, like the EdSAFE AI Alliance, can bring together industry leaders to work together to foster trust. Additional societies bring researchers together. These societies and associations have the reach necessary to bring all parts of the educational ecosystem into discussions about trust and also the ability to represent the views of their constituents in crosscutting policy discussions.

### Recommendation #5: Inform and Involve Educators

Our listening sessions also asked for more specific direction on the question of what education leaders should do (see Figure 15). The most frequent responses fit three clusters: the need for guidelines and guardrails, strengthening the role of teachers, and re-focusing research and development. These are activities that constituents are asking for and that could expand trust. The recommendations that follow respond to these requests.In particular, one concern that repeatedly arose in our listening sessions was the potential for AI to result in less respect for educators or less value for their skills. Across the nation, we are now responding to decreasing interest in entering or remaining in the teaching profession. Now is the time to show the respect and value we hold for educators by informing and involving them in every step of the process of designing, developing, testing, improving, adopting, and managing AI-enabled edtech. This includes involving educators in reviewing existing AI-enabled systems, tools, and data use in schools, designing new applications of AI based on teacher input, carrying out pilot evaluations of proposed new instructional tools, collaborating with developers to increase the trustworthiness of the deployed system, and raising issues about risks and unexpected consequences as the system is implemented. We have already seen educators rise to the challenge of creating overall guidelines, designing specific uses of available AI-enabled systems and tools, and ferreting out concerns. And yet, the influence of educators in the future of AI-enabled products cannot be assumed; instead, constituents need policies that put muscle behind it. Could we create a national corps of leading educators representing every state and region to provide leadership? Could we commit to developing necessary professional development supports? Can we find ways to compensate educators so they can be at the forefront of designing the future of education? Our policies should enable educators to be closely involved in design of AI-enabled educational systems. Although we know that the responsibility for informing and involving educators must be distributed at all levels of national and school governance, the Office of Educational Technologycan play a key role in informing and involving educators through its reports, events, outreach, and in a future NETP. Although examples above refer to K-12 teachers, higher education instructors must also be included. We also call on the edtech industry to involve educators throughout their design and development processes. For example, AI-enabled teaching assistants are only likely to help teachers do their job if teachers are thoroughly involved as the assistants are designed. We call upon institutions that prepare teachers to integrate technology more systematically into their programs; for example, the use of technology in teaching and learning should be a core theme across teacher preparation programs, not an issue that arises only in one course.

**Recommendation #6: Focus R&D on Addressing Context and Enhancing Trust and Safety** Research that focuses on how AI-enabled systems can adapt to context (including variability among learners) in instructional approaches and across educational settings is essential to answering the question of, “Do specific applications of AI work in education, and if so, for whom and under what conditions?” The italicized phrase points to variability among learners and diversity in the settings for learning. We call upon innovators in R&D to focus their efforts to advance AI on the long tail of learning variability, where large populations of students would benefit from customization of learning. We also call on R&D to lead by establishing how trust can be strengthened in AI-enabled systems, building on the Blueprint’s call for safe and effective systems yet also including education-specific requirements, such as how teachers can be meaningfully involved in design phases, not only in implementation and evaluation. Although many products today are adaptive, some adapt on just one or a few dimensions of variability, such as student’s accuracy in problem solving. As teachers know, there are many more important ways to adapt to students’ strengths and needs. Students are neurodiverse and may have specific disabilities. They bring different assets from their experiences at home, in communities, and in their cultures. They have different interests and motivations. They are in different places in their journeys to master the English language. And they learn in varied settings. Classrooms and schools are different, and at home, students learn in informal settings in ways that could complement school learning. We recommend attention to “context” as a means for expressing the multiple dimensions that must be considered when elaborating the phrase “for whom and under what conditions.” We also acknowledge the role of researchers in conducting evaluations, which must now consider not only efficacy but must also explore where harm may arise and the system problems that can occur through weak trust or over-trust in AI systems. R&D must take the lead in making AI models more context-sensitive and ensuring that they are effective, safe, and trustworthy for use with varied learners in diverse settings. Although AI has capabilities to find patterns beyond the limited number of variables that people normally think about, AI is not particularly good at understanding and working with context in the ways people do. Over time, we’ve seen learning sciences grow to be less about individualistic cognitive principles and more encompassing first of social learning and then of the many dimensions of context that matter in learning. Our use of AI needs to follow this trajectory toward context to support educational applications. To achieve human-centric vision, listening session attendees argued that teams will need time and freedom to explore how best to manage the tension between the pace of technological advancement and the need for broader contextual insights—for trust and for safety. They will need time and freedom to pioneer new processes that better involve teachers and students as codesigners, with attention to balancing power dynamics. And they will need to shift attention from older ways of framing priorities (such as achievement gaps) to new ways of prioritizing digital equity. We call on R&D funders to focus resources on the long tail of learner variability, the need for AI-enabled systems that better incorporate context, and time required to get contextual considerations right. We call upon researchers and developers to prioritize challenges of context, trust, and safety in their work to advance AI.

### Recommendation #7: Develop Education-Specific Guidelines and Guardrails

Our final recommendation is central to policymakers. A feature of the American educational system is the emphasis on local decision making. With technology growing in complexity at such a rapid pace, it is becoming difficult for local leaders to make informed decisions about the deployment of artificial intelligence. As we have discussed, the issues are not only data privacy and security but extend to new topics such as bias, transparency, and accountability. It will be harder to evaluate promising edtech platforms that rely on AI systems against this evolving, complex set of criteria. Regulations related to key student and family data privacy laws like the Family Educational Rights & Privacy Act (FERPA), the Children’s Internet Privacy Act (CIPA), and the Children’s Online Privacy Protection Act (COPPA) warrant review and further consideration in light of new and emerging technologies in schools. Laws such as the Individuals with Disabilities Education Act (IDEA) may likewise be considered as new situations arise in the use of AI-enabled learning technologies. As discussed throughout this document, the Blueprint for an AI Bill of Rights is an important framework throughout this work. The Department encourages parallel work by constituents in all levels of the educational system. In addition to the key federal laws cited immediately above, many states have also passed privacy laws that govern the use of educational technology and edtech platforms in classrooms. Further constituents can expect general frameworks for responsible AI in parallel sectors like health, safety, and consumer products to be informative but not sufficient for education’s specific needs. Leaders at every level need awareness of how this work reaches beyond implications for privacy and security (e.g., to include awareness of potential bias and unfairness), and they need preparation to effectively confront the next level of issues.

### Next Steps

We are heartened to see intensifying discussions throughout the educational ecosystem about the role of AI. We see progress that we can build upon occurring, as constituents discuss these three types of questions: What are the most significant opportunities and risks? How can we achieve trustworthy educational AI? How can we understand the models at the heart of applications of AI and ensure they have the qualities that align to educational aspirations? The Department developed this report with awareness of contributions arising from many types of organizations and collectives. Internationally, we recognize parallel efforts to consider AI in the European Union, at the United Nations, and indeed throughout the world. We are aware of progress being led by organizations such as UNESCO, the EdSAFE AI Alliance, and research organizations in many countries. We plan to continue cross-agency work, for example, by continuing to coordinate with the Office of Science and Technology Policy and other Federal agencies as agencies implement next steps guided by the Blueprint for an AI Bill of Rights. We see a broad and fertile context for necessary next steps: ● Working within this context and with others, the Department will consider specific policies and regulations so that educators can realize the opportunities of AI in edtech while minimizing risks. For example, the Department is developing a set of AI usage scenarios to strengthen the process of evaluating and enhancing policies and regulations. The principles and practices in the Blueprint for an AI Bill of Rights will be used to ensure the scenarios mitigate important risks and harms. ● Working with constituents (including education leaders; teachers, faculty, support staff, and other educators; researchers; policymakers; funders; technology developers; community members and organizations; and above all, learners and their families/caregivers), we will develop additional resources and events to increase understanding of

AI and to involve those who will be most affected by these new technologies. ● Working across sectors, such as education, innovation, research, and policy, we will revise and update the NETP to guide all constituents toward safe, equitable, and effective AI in education in the United States, in alignment with our overall educational priorities.