E3B Generative AI Guidelines

Duncan Menge, on behalf of the E3B department
Originally drafted October 12, 2026
Distributed November 19, 2026

Like most of the world, we in E3B recognize that there are both concerns about AI and benefits of AI. The non-research- and non-education-related concerns, including environmental issues, data security, and societal collapse, need to be taken seriously. Our focus for this document, however, is on issues related to research and education.

There is substantial overlap between research and education. Research often takes place as part of an educational program, both in classes and as part of degree programs. For the purposes of this document, "education" includes everything related to classes and degree programs, including all student research for classes and dissertations, whereas "research" includes all other research.

Our overarching educational framework is a brain-first approach. AI is beneficial where it enhances learning, but detrimental where it obstructs or replaces learning. Humans have invented machines that can lift heavy weights, but there is no point in going to a gym and having a machine lift weights for you. Similarly, offloading the intellectual work needed to develop our brains is antithetical to education. On the contrary, machines that facilitate weightlifting can make you stronger, just like some uses of AI can facilitate learning. This analogy can also help with the distinction between research and education. Whereas the focus of education is always to develop our brains, the focus of research is to discover new knowledge, which sometimes means lifting weights by the most efficient means necessary. AI is useful as a tool, but not as a crutch.

Generative AI can offer a number of benefits. Just like calculators have made arithmetic more efficient, AI can make some tasks more efficient. We encourage the use of AI in contexts where it makes non-brain-related tasks more efficient. In many cases, such as data processing, AI enables tasks that would otherwise be impossible, so in some research and in some classes, certain uses of AI are essential. We also acknowledge the potential equity benefits of AI, such as leveling the playing field for non-native English speakers. While acknowledging the benefits of AI, we also respect the rights of E3B members not to use AI if that is their desire.

Our departmental AI guidelines sits within a hierarchy of AI policies. Above the departmental level, Columbia University and Arts & Sciences (including GSAS) have AI policies, which we have to follow. Journals and funding agencies have AI policies, which we have to follow when we submit manuscripts and proposals. Below the departmental level, our departmental guidelines give substantial agency to individual advisers and instructors.

The intent of this document is to provide a framework, but we fully intend for instructors and advisers to provide more detailed guidance for classes and research programs.

For each use case, we offer three categories.

- "**Never**" means that generative AI should never be used, typically because it compromises academic integrity.

- "**Usually not**" means that generative AI should not be used unless an instructor or adviser gives explicit guidance to the contrary. AI use for any of these purposes must be disclosed in writing in submitted products.
- "**Usually fine**" means that generative AI can be used unless an instructor or adviser gives explicit guidelines that it cannot.

In cases where policies overlap, the less permissive option is the more relevant. For example, if the education category for a task is **Usually not** but the research category is **Usually fine**, a research project that is part of a class gets the **Usually not** classification. We note that future uses of many tasks are uncertain. A piece of text might initially be used for a thesis development class, then for a dissertation, then for a paper, which might be submitted to multiple journals with different policies. If AI is allowed for an initial use but not a later use, ethical problems could arise. We urge all E3B members to consider these uncertainties as they make decisions.

Although our guiding principles are laid out above, we thought it would be useful to give guidance for specific tasks. Below we list some of these tasks, grouped into categories.

A few notes:
- For all tasks, we emphasize the importance of checking the output of AI. Anyone using AI to code, for example, should verify that the AI-written or AI-edited code is doing the intended task. (This has always been true for coding.) As another example, even if using AI to format a reference list is fine ethically, it is currently a task where AI makes many mistakes, so it is essential to check the output.
- It is forbidden by multiple groups (funding agencies, Columbia) to upload personal, proprietary, or otherwise sensitive information, including unpublished research, into open (non-sandbox) LLMs. This includes your drafts and data, students/colleagues drafts and data, and so forth. Do not upload these to open LLMs. Columbia provides sandbox versions of many LLMs.

Writing:

Task	Education	Research
Outlining:	**Usually not**	**Usually not**
Writing initial drafts:	**Never**	**Never**
Major editing:	**Usually not**	**Usually not**

- Outlining:
  - Faculty opinion was split, with some nevers, some usually nots, and a couple usually fines. For split opinions, the departmental guideline is the median opinion. The "never" opinion shared the same motivation as writing initial drafts: outlining is where the thought process of writing begins, so it should be done by the author.
- Writing initial drafts:
  - All faculty agreed with "never" for education, and nearly all agreed with "never" for research. For education, our current understanding is that "never" is Columbia and GSAS policy for writing initial drafts. Writing occupies a special place in academia. To many people, writing *is* thinking. The process of putting ideas on paper is the way we organize our thoughts and the way we communicate.
- Major editing:
  - Faculty opinion was split: half nevers, some usually nots, and one usually fine. The challenge here is that there is a wide gray area between writing the initial

draft and minor editing. In some cases, major editing is more similar to writing the initial draft, whereas in others it is more similar to minor editing.

- Minor editing (grammar, spell check, word choice): **Usually fine** **Usually fine**
  - Unanimous.
- Generate and/or edit reference lists: **Usually not** **Usually fine**
  - Faculty opinion was split: one never, some usually nots, and some usually fines. Most faculty agreed that using generative AI like Zotero or EndNote is ethically fine, but that it is an area where AI currently makes a lot of mistakes, which has major consequences. Errors in reference lists can signal that references have not been read or checked properly. Mistakes in reference lists can be viewed as academic misconduct. Of course, it is never acceptable to submit a reference list without reading the references themselves.
- Translating to or from another language: **Usually fine** **Usually fine**
  - Near unanimous for education, unanimous for research.
- Uploading sensitive text into open LLMs **Never** **Never**

Ideas:

Task	Education	Research
Generating/brainstorming ideas:	**Usually not**	**Usually not**
Organizing ideas:	**Usually not**	**Usually not**
Revising ideas:	**Usually not**	**Usually not**

- Faculty opinion for all three was split between never, usually not, and usually fine. Generating, organizing, and revising ideas are a major part of the intellectual process of science.

Experimental design:

Task	Education	Research
Designing experiments:	**Usually not**	**Usually fine**

- Faculty opinion was split for education, with one never, many usually nots, and a couple usually fines. Faculty opinion was also split for research, with many usually fines and some usually nots.

Getting feedback on experimental design:	**Usually fine**	**Usually fine**

- Faculty opinion was split with many usually fines and some usually nots.

Coding:

Task	Education	Research
Writing code:	**Usually fine**	**Usually fine**

- Nearly unanimous

Editing code:	**Usually fine**	**Usually fine**

- Unanimous

Troubleshooting/debugging:	**Usually fine**	**Usually fine**

- Unanimous

Interpreting code:	**Usually fine**	**Usually fine**

- Unanimous

Learning about packages/functions:	**Usually fine**	**Usually fine**

- Nearly unanimous

- Uploading sensitive data to a non-sandbox LLM **Never** **Never**

Creating figures/display items:

Task	Education	Research
Generating figures:	**Usually not**	**Usually fine**
○ Split between usually not and usually fine for both		
Modifying figures:	**Usually fine**	**Usually fine**
○ Nearly unanimous		
Uploading data to a non-sandbox LLM	**Never**	**Never**

Search engine:

Task	Education	Research
Searching for general information	**Usually fine**	**Usually fine**
○ Nearly unanimous.		
Searching for literature:	**Usually fine**	**Usually fine**
○ Nearly unanimous		
Searching for grants/funding/conferences:	**Usually fine**	**Usually fine**
○ Unanimous		

Statistics:

Task	Education	Research
Suggesting models:	**Usually fine**	**Usually fine**
○ Split for education, unanimous for research		
Checking output interpretation:	**Usually fine**	**Usually fine**
○ Split for both.		
Using machine learning or other standard statistics:	**Usually fine**	**Usually fine**
○ Nearly unanimous.		
Uploading sensitive data to a non-sandbox LLM	**Never**	**Never**

Reading/studying:

Task	Education	Research
Summarizing papers to screen for relevance:	**Usually fine**	**Usually fine**
○ Split between one never, one usually not (for education), and many usually fines		
Citing AI summaries without verifying source:	**Never**	**Never**
○ Unanimous		
Generating flash cards/study tools:	**Usually fine**	**Usually fine**
○ Nearly unanimous		
Uploading sensitive data to a non-sandbox LLM	**Never**	**Never**

Data processing:

Task	Education	Research
Data processing:	**Usually fine**	**Usually fine**
○ Split between usually not and usually fine		
Uploading sensitive data to a non-sandbox LLM	**Never**	**Never**

Grading/reviewing:

Task	Education	Research
- Using AI to assist with grading or giving feedback:	**Usually not**	**Never**
o Faculty opinion was mostly usually not for education, with one usually fine and one never, and split between never and usually not for research.		
- Uploading sensitive data to a non-sandbox LLM	**Never**	**Never**