

AI Policy for Teaching, Research and Administration in the Department of HSS, IIT Bombay

Artificial Intelligence (AI) involves use of advanced computing tools such as machine learning and algorithms to generate texts, images, videos, audios, codes/formulae, and other such content, based on human user inputs and prompts. Current GenAI (Generative Artificial Intelligence) output can be indistinguishable from creative output of humans. At the same time, it is important to keep in mind that guaranteed content accuracy is not a feature of AI and human user alone is responsible for the content. In particular, AI content could reproduce human bias, prejudice, and malice. Other concerns regarding the increasing use of AI include: (a) the problem of junk content, (b) inequalities of AI literacy skills, (c) disparities in the access to technology (digital divide), and (d) the environmental impact of energy-intensive AI. Therefore, prudent and responsible use of AI tools is strongly recommended. Recent developments in AI pose new challenges for the practice of academic integrity, ethical research, and shared responsibility among faculty, researchers, and students. For disciplines of the humanities and social sciences, which are reading and writing intensive and rooted in critical thinking and analysis, it becomes crucial to impart and preserve their core values and skill sets in this emerging situation. Rigorous training and academic output, reflecting original content of the researchers' own analysis, interpretation and ideas must continue alongside increasing prevalence of AI. Given this dynamic and constantly changing scenario, this policy is specifically meant to address the concerns of the Department of Humanities and Social Sciences (HSS), IIT Bombay and must be revisited every semester by the DPGC (Department Post-Graduate Committee) to incorporate revisions, if necessary.

1. This policy specifically addresses the use of AI for Teaching, Research and Administration in the Department of HSS, IIT Bombay.
2. The use of AI according to the strict demands of academic and professional integrity is the ultimate responsibility of the user/author/writer, whether one is faculty, researcher, student or staff.
3. Human verification of the accuracy of AI content by the user of AI is mandatory in all cases of permissible use.
4. Any information fed into AI platforms (such as the essays we write) will most likely be used by the same platforms in ways beyond our control. AI platforms are also prone to generating content that is inaccurate and hallucinatory. Awareness and application of tools or functions available to restrict or reduce the use of information fed into AI platforms is recommended. The department may consider conducting periodic training sessions to ensure healthy use of AI.
5. In the interest of intellectual property rights, privacy and confidentiality, we must not feed any information into AI platforms that must not be available in the public domain, including institutional and sensitive material.
6. Explicit prior permission from the instructor or supervisor is mandatory for the use of AI, except for the use of basic functions like grammar and spell-check. For the latter, Institute

subscribed and recommended (if any) AI platforms must be used. However, use of grammar and spell-checking tools must be acknowledged (see Point No.11).

7. Students and researchers will be held responsible for the use of AI in any work submitted in the department for evaluation and any violation of academic integrity with respect to the use of AI will be solely their responsibility.
8. It is mandatory to abide by the regulations of the Institute Ethics Committee (IEC) or Review Board (IRB) regarding collection of data from human respondents for maintaining confidentiality, anonymity and consent requirements. Therefore, use of AI for transcription of interviews and survey data is not permissible. Any data containing personally identifiable information must not be uploaded in any platform, including AI platforms, in ways that will violate the Digital Personal Data Protection (DPDP) Act, 2023 of the Government of India (<https://www.meity.gov.in/documents/act-and-policies>).
9. If quantitative and qualitative data are analysed by researchers using AI tools under the supervision of their research supervisor, the user must be responsible and remain in control of the analytical process. In that case, AI use must only be supplementary to the user's own critical judgment and it must be duly acknowledged.
10. In cases where translation is a primary exercise of research, use of AI is not recommended, unless specifically permitted by the supervisor, and if used, it should be duly acknowledged and must be in accordance with the DPDP Act.
11. Use of AI within permissible limits should be duly acknowledged with proper citation in all types of teaching materials and research outputs. Citation must include: (a) prompts used, (b) AI tools used, and (c) clear demarcation of the content/output generated by or with the aid of AI. If required, the user should be able to produce relevant interactions with AI for verification. Methods and examples of citation style can be found in the links below:

<https://apastyle.apa.org/blog/cite-generative-ai-references>
<https://www.chicagomanualofstyle.org/qanda/data/faq/topics/Documentation/faq0422.html>
12. Use of AI without human verification is not permissible for any task that involves evaluation or assessment of a student, staff or researcher, including (but not limited to) assessment of student submissions by faculty, of staff by Department functionaries, and screening of applications for admission, recruitment and promotion.
13. When there is any case of inappropriate AI use by students/research scholars, instructors/supervisors must report the academic malpractice to D-ADAC (Departmental Academic Disciplinary Action Committee) or ADAC ([Institute-level] Academic Disciplinary Action Committee) rather than resolving the issue at a personal or individual level. The malpractice or misconduct will be verified by D-ADAC/ADAC. Disciplinary measures will be decided as per gravity of the malpractice and as per prevalent Institute rules and regulations. [see: https://acad.iitb.ac.in/files/2024-04/punishments_201521July.pdf]
14. Along with specific citations (see point No. 11), all dissertations and research reports (such as Annual Progress Seminar [APS], Pre-Synopsis Seminar, MA/MA.Res. project

and other seminar reports) submitted for evaluation in the Department of HSS must include a declaration of AI use/non-use by the student/researcher as per IIT Bombay requirement (see <https://acad.iitb.ac.in/files/2024-04/Dissertation17june09-10.pdf>; rule No.3.10). Instructors can also ask for such a declaration for course assignments submitted for evaluation.

15. AI detection tools (such as Turnitin) cannot be used as the sole criterion for disciplinary action without human verification.
16. In the absence of the AI protocol in academic publishing houses, this policy is applicable for the publication of research output of concerned members of the Department.