CIGI Policy for Use of Generative AI in Research

This document can be made available in an alternate format upon a request to Human Resources.

The Centre for International Governance Innovation ("**CIGI**") recognizes that informed and responsible use of generative artificial intelligence ("**AI**") applications like ChatGPT and other artificial intelligence techniques can increase research productivity and foster innovation. Generative AI refers to AI systems that can produce new content—such as text, images, or data analyses— by modelling features from large datasets that were fed into the model.

CIGI also recognizes that using AI applications for research-related work carries potential risks. Challenges include potential errors, fabrications, lack of consistency in output results, lack of transparency in how AI output is produced, and ethical considerations such as bias, misinformation, and deep fakes. Other risks include potential copyright violations, confidential information protection issues, and the potential for outputs that violate human rights laws.

Scope and Applicability

The purpose of this policy is to provide protocols that maximize the benefits of using generative AI applications for research while minimizing potential risks or concerns. It applies to all individuals and organizations that are engaging in research-based commitments for CIGI, including conducting peer reviews.

This policy is not intended to cover every potential use of generative AI for research purposes, especially as such uses will continue to evolve. Rather, it is intended to provide information and guidance on CIGI's parameters for use of AI applications, to ensure that such use:

- upholds research integrity and scholarly standards;
- preserves transparency and accountability in authorship;
- protects confidential, sensitive, or proprietary information; and
- aligns with ethical research and publication practices.

AI tools are assistive technologies and cannot substitute for an individual's intellectual responsibility or scholarly judgment.

Transparency and Requirements for Disclosure

If AI was used in connection with your work for CIGI, you must submit, together with submission of your first draft, a disclosure statement that follows the guidelines set out in the Artificial Intelligence Disclosure (AID) Framework[1]

- The author of the AID Framework has created an AID Statement Builder that you can find here: https://aidframework.org/
- AID statement examples provided by the AID Framework author can be found here: https://crln.acrl.org/index.php/crlnews/article/view/26548/34482

To complete this disclosure, ensure that you are tracking and documenting your use of all AI applications in connection with your work for CIGI. This documentation also serves to enhance ethical and accurate output, and ensure accountability. Documented information must include:

- The purpose of the AI application use.
- The name of the AI application tool used.
- The source of the input data used to teach the AI application to carry out the given task (i.e. the prompt or files provided).
- The model of logic used to teach the AI application, including algorithms, computational models or packages used to analyze data for trends or optimal paths.
- The source of research used to verify the AI application output.
- A plain language summary of any identified potential deficiencies with the AI output, including accuracy, consistency of results, potential bias in results and calculation of risks.

Authorship and Research Integrity

CIGI engages individuals that are experts in their respective fields to conduct research, participate in meetings and events with key stakeholders, and author high-quality work for publication. We expect the work to be well-researched, innovative and original.

Authors remain fully accountable for the content, accuracy, originality, and integrity of all work produced for CIGI, regardless of whether AI tools were used in its creation.

[1] Weaver, K. D. (2024). The artificial intelligence disclosure (AID) framework: An introduction. arXiv. https://arxiv.org/abs/2408.01904

CIGI publications endure a rigorous editing process whereby manuscripts are edited for sense, length, grammar, format and style so that they are in publishable form. The editors perform substantive editing, stylistic editing and copyediting. However, it is your responsibility to review your work – including elements of your work that have been produced with the assistance of generative AI – to ensure that the quality of the piece is sound, the citations are correct (and the sources legitimate), and nothing in the work infringes upon any intellectual property rights or is otherwise unlawful.

Use Categories

There are three categories for Generative AI use which represent different levels of permissible applications:

1. Green (Permitted) covers low-risk, high-benefit AI uses such as routine data analysis, research assistance, and language tools.
2. Yellow (Permitted with Conditions) which require human oversight and regular review to mitigate risks like bias or inaccuracies.
3. Red (Not Permitted) involves high-risk, ethically sensitive applications which are strictly prohibited due to potential harm and legal risk.

Risk is a function of 1) the negative impact, or magnitude of harm, that would arise if the circumstance or event occurs and 2) the likelihood of occurrence. For situations not specifically included in one of the categories below, consider the following non-exhaustive list of elements to determine a situation's level of risk and whether a use-case would be permissible:

- impact of errors (type of impact – e.g. legal, reputational – as well as scale) and ability to reverse mistakes.
- recipient of output: internal content has a lower level of risk than partner/public-facing content.
- level of human review.
- type of data (personal or confidential information vs. publicly available information).
- quality of data (source, representativeness).
- reliability of application (credibility, fit for purpose).
- scale of use.

Best practice clearly indicates that using paid versions of applications, and modifying settings to prohibit the use of data for training purposes, offer the best protection of your pre-published insights and analysis.

Permitted	Permitted with Conditions	Not Permitted
Routine data analysis: Generating basic reports, data sorting, and visualization for internal use. Research assistance: Gathering publicly available information for preliminary research. Language tools: Editing for clarity, grammar, or style (outputs must be reviewed by the author).	Research content generation: Assisting in drafting research materials, preliminary literature reviews, and annotated bibliographies where AI outputs need careful verification to avoid inaccuracies or biased interpretations. Final versions must undergo human oversight, especially when intended for external publication or presentation. Audience-Facing Content Creation: Creating or summarizing content for external communication channels. Outputs must be reviewed and edited by a human to ensure alignment with organizational tone, values, and to prevent unintended biases or misinformation. Scenario and data modeling: Using AI tools to generate predictive insights, model scenarios, or visualize complex data sets where there are accuracy and ethical considerations. These outputs require validation by domain experts and strict controls to ensure reliability	Intellectual property infringement: knowingly inputting or generating content that violates copyrights, trademarks, or other IP rights. Deceptive practices: Using generative AI to produce content that replicates or closely mimics existing works without proper attribution, falsely presenting AI-generated material as original, or failing to disclose AI use as required under this policy. Privacy Violations: Inputting or disseminating personal or confidential information without authorization. Failure to Verify Content: Using AI-generated or reviewed content, including references and bibliographies, without human verification. Delegating Peer Reviews: Peer review is a scholarly responsibility that must be carried out by the contracted reviewer and may not be outsourced, automated, or delegated.

	and avoid misleading interpretations. External communication drafts: Creating initial drafts of communications for collaborations, partnerships, or stakeholder outreach where the tone and content impact external relationships. Outputs must be reviewed for appropriateness, alignment with CIGI's mission, and compliance with organizational communication policies.	

While this policy permits certain uses of generative AI in accordance with the Use Categories above, CIGI may, at its discretion, restrict or prohibit the use of generative AI in connection with specific research projects. Such restrictions may arise due to legal, ethical, contractual, reputational or research-integrity considerations. Where additional limitations apply, they will be communicated to the relevant individuals or organizations and will take precedence over the general permissions set out in this policy.

Guidelines for Using AI Applications

In using generative AI applications, abide by the following guidelines:

- **Data input.** Carefully consider any data that is used as input in the AI application. Do not enter any confidential, trade secret, or other proprietary information into an AI application prompt.
- **Respect terms of use.** Review the terms of use of the AI application and ensure that the intended use complies with those terms.
- **Respect third party intellectual property rights.** When inputting data or generating output, be aware of any intellectual property rights that may be owned by third parties, such as copyright, database rights or trademark rights. Ensure that third-party

proprietary data is not entered into an AI application prompt without the third party's consent.

- **Review of output.** Generative AI has the potential to produce inaccurate outputs or fabrications. There is also a risk that the output is biased, inappropriate, or otherwise offensive. Critical thought must be applied to all outputs of AI applications; they must always be fact- and sense-checked before being relied upon and reviewed to ensure content is appropriate. You are required to ensure that the outputs generated:
  - do not contain biased, offensive, or discriminatory content;
  - do not improperly use or disclose personal or confidential information; and
  - have been verified with other trusted sources to ensure accuracy.

Administration of Policy

This policy will be reviewed and updated to reflect evolving technologies, norms, and best practices. Authors are responsible for complying with the version of the policy in effect at the time of submission.

If you have any questions regarding this policy or any questions about using AI applications that are not addressed in this policy, send your questions to your CIGI Research Director.

Acknowledgement of Receipt and Review

I, _____ (name), acknowledge that on _____ (date), I received a copy of CIGI's Policy for Use of Generative AI in Research and that I read it, understood it and agree to comply with it.

_____

Signature

Name:

Date: