

THE RESEARCH UNIVERSITY PRESS

International
Computational Law Journal

VOLUME 2 / ISSUE 2 / CURATED SOURCE EDITION

Task Force on Responsible Use of Generative AI for Law

Dazza Greenwood

MIT Computational Law Report · Originally published 2023-02-28

Independent reader edition under CC BY 4.0. This attribution cover was added; original source PDF pages follow unchanged. Original authorship and source identity are retained. MIT, Stanford University, Stanford Law School and CodeX have not sponsored or endorsed this republication.

Source, license and version evidence accompany the online article.

MIT Computational Law Report

Task Force on Responsible Use of Generative AI for Law

Dazza Greenwood

Published on: Feb 28, 2023

URL: https://law.mit.edu/pub/generative-ai-responsible-use-for-law

License: Creative Commons Attribution 4.0 International License (CC-BY 4.0)

The research arm of law.MIT.edu is convening an experts and stakeholders group to examine and report upon principles and guidelines for applying due diligence and legal assurance applicable to Generative AI for law and legal processes.

The purpose of this Task Force is to develop principles and guidelines on ensuring factual accuracy, accurate sources, valid legal reasoning, alignment with professional ethics, due diligence, and responsible use of Generative AI for law and legal processes.

The work of the Task Force has served as the basis for law reform efforts around the country and the world, such as the State Bar of California, who wrote in their recently promulgated Guidelines for the Use of Generative Artificial Intelligence in the Practice of Law:

> "This document is based on the principles and guidelines prepared by MIT's Task Force on Responsible Use of Generative AI for Law, and addresses some of the initial concerns surrounding lawyer use of generative AI, as well as use of other applications of AI."

This Task Force has also helped catalyze related work in other countries, such as Argentina's Guidelines for the use of ChatGPT and text generative AI in Justice, where co-author Mariana Sánchez Caparrós wrote:

> "For us, the UBA IALAB research team, the work of the Task Force was a source of inspiration that made us realize that we needed to work locally on some guide to the use of generative AI, because we noticed that it was beginning to be used in the justice system with little caution and knowledge. So, inspired by the work of the Task Force, we prepared a Guide of recommendations."

Task Force on Responsible Use of Generative AI for Law

Version 0.2 June 2, 2023; law.MIT.edu/AI

Task Force Members

Dazza Greenwood, Chair

Shawnna Hoffman, Co-Chair

Olga V. Mack, LexisNexis / CodeX Fellow (Stanford) / Berkeley Law Lecturer

Jeff Saviano, EY / MIT Connection Science Fellow

Megan Ma, Stanford / MIT

Aileen Schultz, MIT Computational Law Report

This Task Force is convened to examine and report upon principles and guidelines for applying due diligence and legal assurance applicable to Generative AI for law and legal processes. Below, for comment, is an initial public draft version of the principles with some scenarios showing consistent and inconsistent examples in the context of law practice.

Visit the web version of this article to view interactive content.

law.MIT.edu Task Force on Responsible Use of Generative AI for Law - Discussion Forum, Summer 2023

The purpose of this Task Force is to develop principles and guidelines on ensuring factual accuracy, accurate sources, valid legal reasoning, alignment with professional ethics, due diligence, and responsible use of Generative AI for law and legal processes. The Task Force believes this technology provides powerfully useful capabilities for law and law practice and, at the same time, requires some informed caution for its use in practice.

In light of the Mata vs. Avianca Airlines., Inc case, where an attorney filed citations and cases fabricated by ChatGPT to a court, the need for responsible use guidelines have arisen more pointedly than before. Recently, a federal judge on the Northern District of Texas promulgated a rule requiring more explicit and specific certification than what Rule 11 provides to confirm clearly "...any language drafted by generative artificial intelligence…will be checked for accuracy, using [authoritative legal sources], by a human being before it is submitted to the Court." At this point in history, we think it's appropriate to encourage the experimentation and use of generative AI as part of law practice, but caution is clearly needed given the limits and flaws inherent with current widely deployed implementations. Eventually, we suspect every lawyer will be well aware of the beneficial uses and also the limitations of this technology, but today it is still new. We would like to see an end date attached to technology-specific rules such as the certification mentioned above, but for the present moment, it does appear reasonable and proportional to ensure attorneys practicing before this court are explicitly and specifically aware of and attest to the best practice of human review and approval for contents sourcing from generative AI.

The Task Force is publicly releasing this early version of draft principles in an open and iterative spirit, not because we believe they are complete or perfect, but in order to engage a broader set of views and expertise in order to improve the end result. In future versions, we intend to include guidelines and commentary to apply and augment the principles.

The Task Force invites your feedback on this draft and asks, in particular, the following questions:

- The first guideline we intend to include focuses on the need to apply and dynamically adapt standards and best practices for data governance and information security in all the usage of AI applications. We seek feedback on this and other needed guidelines that may be useful to include as part of the work of this Task Force.
- Are the existing duties we identified correct and complete, or would you recommend any deletions, additions, or modifications?
- What can we learn from, and what are the implications for legal malpractice insurance?
- Does your firm or other organizations have a policy, procedures, or other guidelines for using generative AI? We invite you to share a copy of this document (including redacted and/or anonymous versions) through our feedback form.
- Do you have input on jurisdiction-specific approaches that may exist elsewhere in the world beyond the United States?
- Are you aware of and can you identify for us any other teams or task forces etc, working on these same issues?
- Use the following form to provide your feedback: https://forms.gle/W2G8d419eFSqUqtLA

Future versions of these principles will be followed by additional guidelines and materials. We invite your contributions and suggestions on guidelines, best practices, examples, forms, policies, and any other inputs, questions, and ideas you may have.

Notes on the scope and application of these principles:

- These principles largely identify and apply, but in some cases extend existing professional rules and principles of conduct for the legal profession, such as the ABA Rules of Professional Conduct and the United Nations Principles for the Role of Lawyers. They have been extended or applied to encompass the use of AI within legal practice.

- These principles are intended to address oncoming issues and imminent concerns with regard to the use of generative AI, however, they may apply to more traditional applications of AI and data modeling as well e.g. for automating processes or analytics.
- Traditional AI ethics principles such as explainability or fairness principles were reviewed against existing rules of professional conduct within the legal sphere and it was determined that in most (if not all) cases these principles are nicely subsumed under existing professional principles e.g. the Fiduciary Duty of Care would include considerations of bias prevention and fair treatment.
- These principles are intended for use in governing best practices with the use of externally provided generative AI applications and services and do not necessarily cover the internal development and deployment of such models.

PRINCIPLES

Adhere to the following principles:

1 Duty of Confidentiality to the client in all usage of AI applications;

2 Duty of Fiduciary Care to the client in all usage of AI applications;

3 Duty of Client Notice and Consent* to the client in all usage of AI applications;

4 Duty of Competence in the usage and understanding of AI applications;

5 Duty of Fiduciary Loyalty to the client in all usage of AI applications;

6 Duty of Regulatory Compliance and respect for the rights of third parties, applicable to the usage of AI applications in your jurisdiction(s);

7 Duty of Accountability and Supervision to maintain human oversight over all usage and outputs of AI applications;

*Consent may not always be required - refer to existing best practices for guidance. We also seek feedback on whether or when consent may be advisable or required.

EXAMPLES

Principle	Example: Inconsistent	Example: Consistent
1 Duty of Confidentiality to the client in all usage of AI applications	You share confidential information of your client with a service provider through prompts that violate your duty because, for example, the terms and conditions of the service provider permit them to share the information with third parties or use the prompts as part of training their models.	Ensure you don't share confidential information in the first place, such as by adequately anonymizing the information in your prompts, or possibly ensure contractual and other safeguards are in place, including client consent.
2 Duty of Fiduciary Care to the client in all usage of AI applications	Not fact-checking or verifying citations from the outputs of generative AI. Using GAI for contract review or completion without regard for defensibility or accuracy.	Ensure diligence and prudence with respect to facts and law. Maintain existing practices for fact-checking.

3 Duty of Client Notice and Consent* to the client in all usage of AI applications;	The client would be surprised by and object to the way in which the attorney used GAI for their case. If the terms of an agreement with a client require disclosure of this type and the attorney fails to disclose.	The terms of the client engagement include the use of technology and specifically address the responsible use of GAI. If needed, an amended engagement letter is agreed upon with the client.
4 Duty of Competence in the usage and understanding of AI applications;	Making use of GAI and accepting outputs as fact without understanding how the technology works and or critically reviewing how outputs are generated.	Understanding and skillfully integrating generative AI with other relevant apps and tools in your workflow. Likewise, the currently required skill of adequately composing prompts to generate high-quality outputs that augment and improve upon existing human expertise.
5 Duty of Fiduciary Loyalty to the client in all usage of AI applications;	Accepting at face value the output of GAI that contains recommendations or decisions contrary to the client's best interests, e.g. potentially prioritizing the interests of a buyer when your client is the seller, or the employer when your client is the employee, etc.	Critically review, confirm, or correct the output of generative AI to ensure the content accurately reflects and supports the interests and priorities of the client in the matter at hand, including as part of advocacy for the client.
6 Duty of Regulatory Compliance and respect for the rights of third parties, applicable to the usage of AI applications in your jurisdiction(s);	Deploying GAI to employees and agents of your firm who practice in jurisdictions that have, for example, banned that technology.	Analyzing the relevant laws and regulations of each jurisdiction in which GAI is deployed in/by your firm and ensuring your compliance with such rules, e.g. by being able to turn off the tool for users in problematic jurisdictions.
7 Duty of Accountability and Supervision to maintain human oversight over all usage of AI applications.	Making use of GAI applications without adequate best practices and human oversight, evaluation, and accountability mechanisms in place.	Any language drafted by GAI is checked for accuracy, using authoritative legal sources by an accountable human being before submission to a Court. Responsible parties decide on use cases/tasks that GAI can and cannot perform and sign off on use on a client/matter basis.

Contributors to the Task Force

The Task Force is grateful for the generous support and input of the following invited contributors. However, the content, including errors and omissions, does not necessarily reflect the views of any given contributor.

Adrienne Valencia Garcia
Anessa Allen Santos, Esquire
Carmin Ballou
Damien Riehl
David Hobbie
Deepa Tharmaraj
Eli Nelson
Eric Ambinder
Eske Montoya Martinez van Egerschot
Iris Skornicki
John Nay
Massimiliano Nicotra
Peter Bilyk
Roberto Lopez-Davila
Sam Harden
Stephen Goldstein
Steve Goldstein
Una Kang

As with the law.MIT.edu Task Force on Privacy Principles for COVID-19 Contact Tracing and the Task Force on Computational Law for Combating Modern Slavery, the published materials of this Task Force are available under Creative Commons open license terms.

- Use this link for the most up-to-date version of this page.
- Use this form to request Task Force updates and to be notified when the final report is published.
- Find more information on the use of Generative AI for law and legal processes here.