CADWALADER

AI and Bank Operations and Supervision

Presented by:

Mercedes Kelley Tunstall, Partner

and

Jay Agarwal, CEO, ThriveAI360

January 29, 2025

What We Will Be Discussing

1. Artificial Intelligence Tool Basics

2. Examples of AI Tools In Financial Services

3. Managing AI Risks

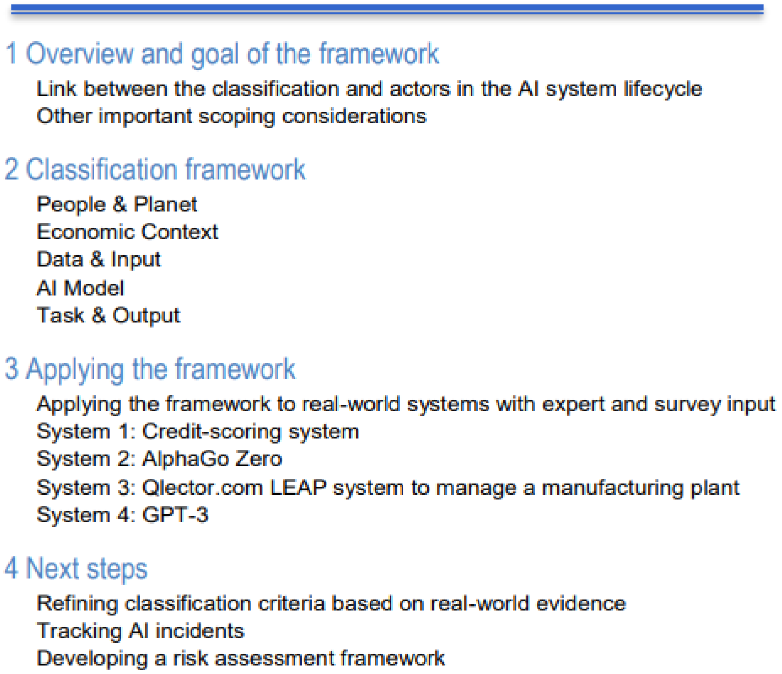
4. Legal Ethical Considerations

Artificial Intelligence Basics

Defining Artificial Intelligence

From the National Artificial Intelligence Initiative
15 U.S.C. 9401(3)

The term "artificial intelligence" means a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. Artificial intelligence systems use machine and human-based inputs to-
(A) perceive real and virtual environments;
(B) abstract such perceptions into models through analysis in an automated manner; and
(C) use model inference to formulate options for information or action.

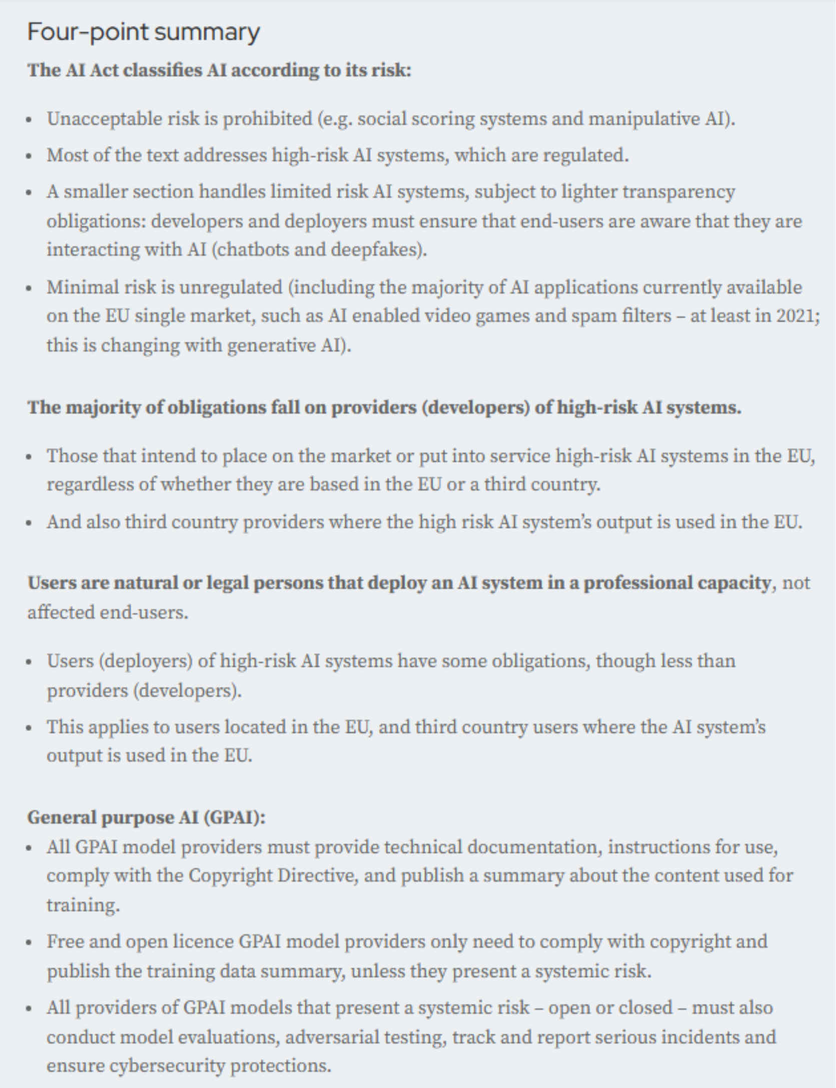


1 Overview and goal of the framework
- Link between the classification and actors in the AI system lifecycle
- Other important scoping considerations

2 Classification framework
- People & Planet
- Economic Context
- Data & Input
- AI Model
- Task & Output

3 Applying the framework
- Applying the framework to real-world systems with expert and survey input
- System 1: Credit-scoring system
- System 2: AlphaGo Zero
- System 3: Qlector.com LEAP system to manage a manufacturing plant
- System 4: GPT-3

4 Next steps
- Refining classification criteria based on real-world evidence
- Tracking AI incidents
- Developing a risk assessment framework



Four-point summary

The AI Act classifies AI according to its risk:

- Unacceptable risk is prohibited (e.g. social scoring systems and manipulative AI).
- Most of the text addresses high-risk AI systems, which are regulated.
- A smaller section handles limited risk AI systems, subject to lighter transparency obligations: developers and deployers must ensure that end-users are aware that they are interacting with AI (chatbots and deepfakes).
- Minimal risk is unregulated (including the majority of AI applications currently available on the EU single market, such as AI enabled video games and spam filters – at least in 2021; this is changing with generative AI).

The majority of obligations fall on providers (developers) of high-risk AI systems.

- Those that intend to place on the market or put into service high-risk AI systems in the EU, regardless of whether they are based in the EU or a third country.
- And also third country providers where the high risk AI system's output is used in the EU.

Users are natural or legal persons that deploy an AI system in a professional capacity, not affected end-users.

- Users (deployers) of high-risk AI systems have some obligations, though less than providers (developers).
- This applies to users located in the EU, and third country users where the AI system's output is used in the EU.

General purpose AI (GPAI):
- All GPAI model providers must provide technical documentation, instructions for use, comply with the Copyright Directive, and publish a summary about the content used for training.
- Free and open licence GPAI model providers only need to comply with copyright and publish the training data summary, unless they present a systemic risk.
- All providers of GPAI models that present a systemic risk – open or closed – must also conduct model evaluations, adversarial testing, track and report serious incidents and ensure cybersecurity protections.

Types of AI Models

- Symbolic or Knowledge-Based – Uses expert human knowledge

- Statistical – Identifies patterns via algorithms

- Discriminative – Predicts labels based on distinguishing between datasets

- Generative – Creates new content

- Agentic – AI agents take actions autonomously and without human oversight

Why Do We Care About Identifying the Type of AI Model?

- Most AI Tools in the marketplace are composites of one or more types of models and some AI tools consist of an “ensemble” of models, meaning that a group of models act together in parallel and thereby cooperate on a single task or decision.

- Once we know what AI Model or Models are involved, then we can ask questions that will help us know what kinds of risks the AI Models pose and that we need to control for

- Additional information about models that is useful to understand:

  - ✓ Supervised AI Models are those where the data used to train the model has been labeled in advance. Supervised AI models are useful for text recognition and classification tasks.

  - ✓ Unsupervised AI Models are those where the data used to train the model has NOT been labeled. Unsupervised AI models are useful for pattern recognition tasks.

  - ✓ Reinforcement AI Models are those where the model is built such that humans can evaluate the output and either reward the model or penalize the model, in order to guide the model to better results over time.

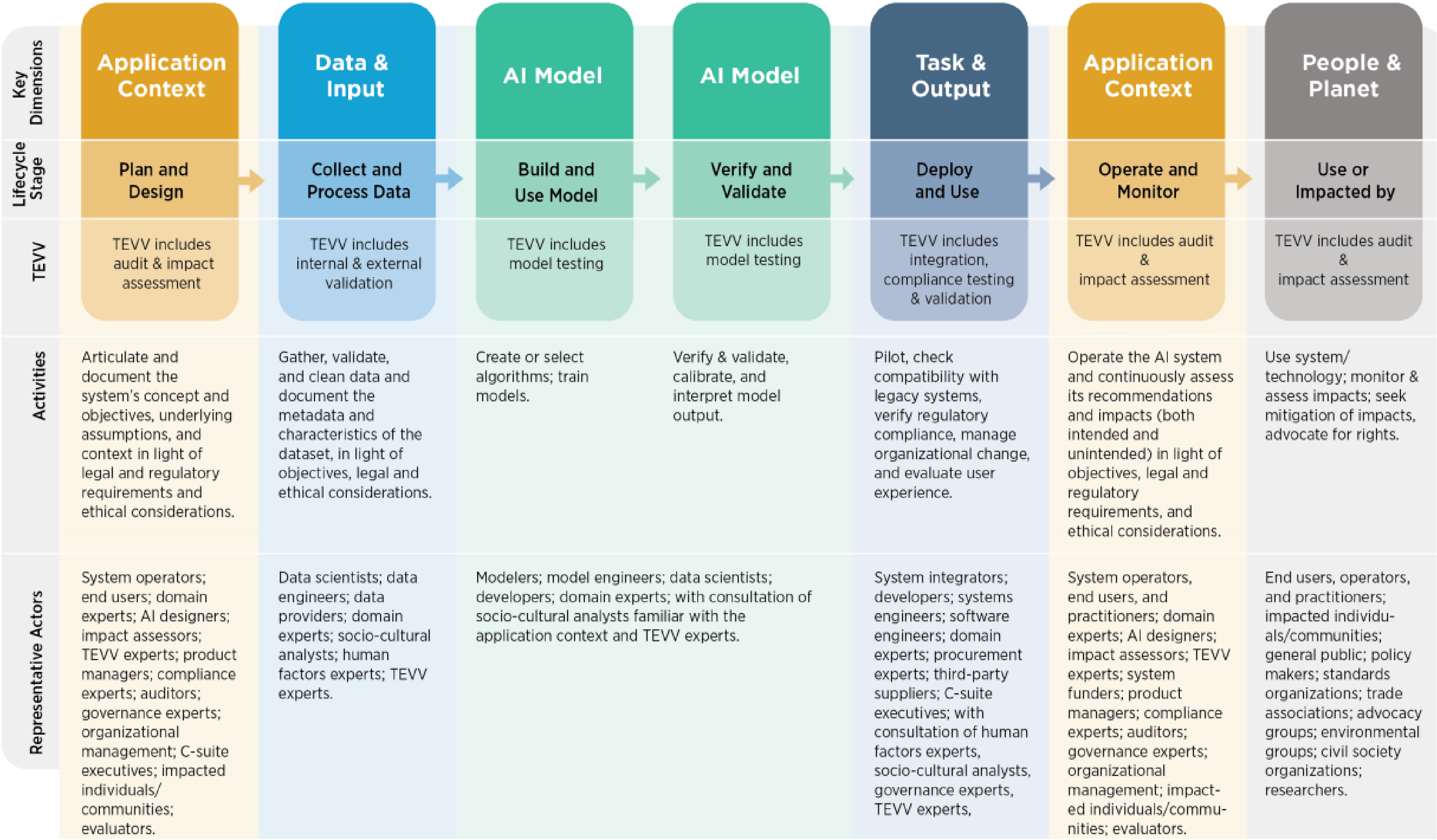
  - ✓ Semi-supervised AI Models are most useful when it is challenging to create a large enough labeled data set, such as when it is necessary to analyze images and speech at the same time

  - ✓ Large Language Models (LLMs) are a type of semi-supervised AI Models

Artificial Intelligence In Financial Services

Managing AI Risks

NIST AI Risk Management Framework



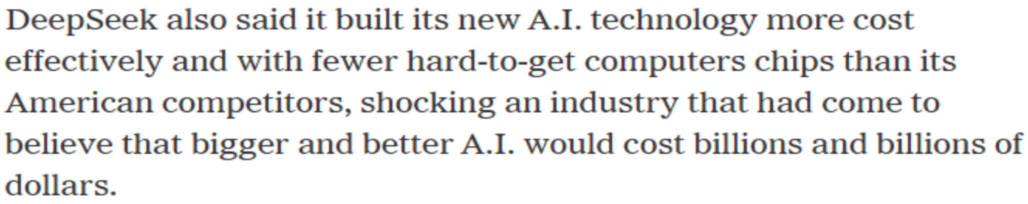
Key Dimensions	Application Context	Data & Input	AI Model	AI Model	Task & Output	Application Context	People & Planet
Lifecycle Stage	Plan and Design	Collect and Process Data	Build and Use Model	Verify and Validate	Deploy and Use	Operate and Monitor	Use or Impacted by
TEVV	TEVV includes audit & impact assessment	TEVV includes internal & external validation	TEVV includes model testing	TEVV includes model testing	TEVV includes integration, compliance testing & validation	TEVV includes audit & impact assessment	TEVV includes audit & impact assessment
Activities	Articulate and document the system's concept and objectives, underlying assumptions, and context in light of legal and regulatory requirements and ethical considerations.	Gather, validate, and clean data and document the metadata and characteristics of the dataset, in light of objectives, legal and ethical considerations.	Create or select algorithms; train models.	Verify & validate, calibrate, and interpret model output.	Pilot, check compatibility with legacy systems, verify regulatory compliance, manage organizational change, and evaluate user experience.	Operate the AI system and continuously assess its recommendations and impacts (both intended and unintended) in light of objectives, legal and regulatory requirements, and ethical considerations.	Use system/ technology; monitor & assess impacts; seek mitigation of impacts, advocate for rights.
Representative Actors	System operators; end users; domain experts; AI designers; impact assessors; TEVV experts; product managers; compliance experts; auditors; governance experts; organizational management; C-suite executives; impacted individuals/ communities; evaluators.	Data scientists; data engineers; data providers; domain experts; socio-cultural analysts; human factors experts; TEVV experts.	Modelers; model engineers; data scientists; developers; domain experts; with consultation of socio-cultural analysts familiar with the application context and TEVV experts.		System integrators; developers; systems engineers; software engineers; domain experts; procurement experts; third-party suppliers; C-suite executives; with consultation of human factors experts, socio-cultural analysts, governance experts, TEVV experts.	System operators, end users, and practitioners; domain experts; AI designers; impact assessors; TEVV experts; system funders; product managers; compliance experts; auditors; governance experts; organizational management; impacted individuals/communities; evaluators.	End users, operators, and practitioners; impacted individuals/communities; general public; policy makers; standards organizations; trade associations; advocacy groups; environmental groups; civil society organizations; researchers.

It’s All About the Data

- What kind of data was used to train the model, and what kind of data is being generated by the model?
- For data used to train the model what is the provenance of that data?
  - Expert data?
  - Provided data?
  - Observed data?
  - Synthetic data that was generated algorithmically?
  - Derived data (data taken from other data to become a new data element)
    - Proprietary data is most likely derived.
    - Examples of derived data include computational data (like a credit score); categorical data (such as age groups of borrowers); inferred data (like a fraud score); or aggregated data (data abstracted from more fine-grained data)

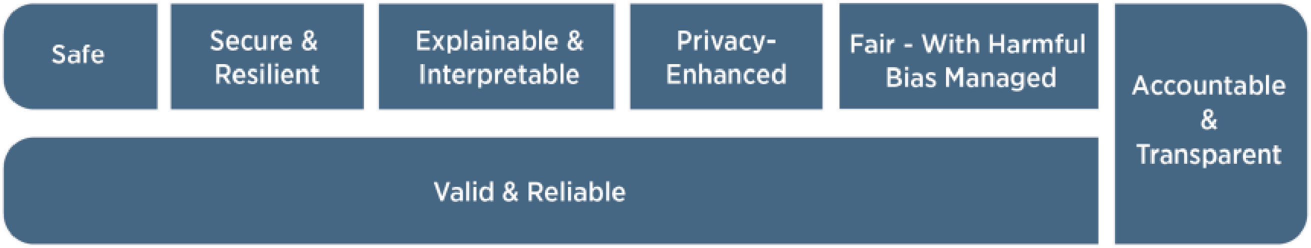
Problems That Can Arise With AI Tools

- Consider whether the data is static, dynamic and updated from time-to-time or dynamic and updated real time.
- Consider the scale of the data, which can impact the effectiveness of the AI Tool



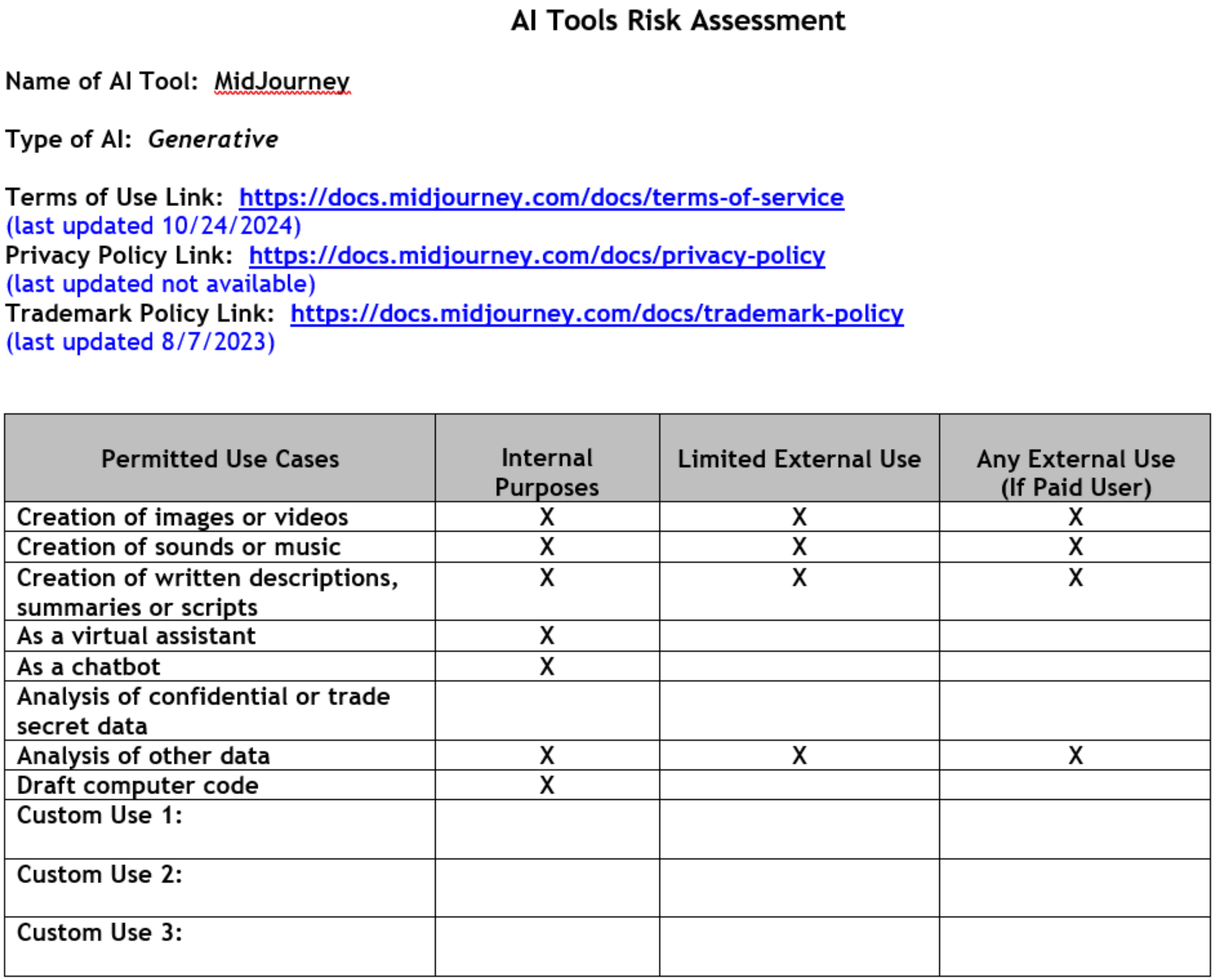
DeepSeek also said it built its new A.I. technology more cost effectively and with fewer hard-to-get computers chips than its American competitors, shocking an industry that had come to believe that bigger and better A.I. would cost billions and billions of dollars.

- Sometimes, even absent personal data, AI models can quickly infer data and correlations from proxy variables that are not personally identified, which can raise privacy and security concerns.
- Data integrity and drift concerns
- Trustworthiness
- Bias/Discrimination
- Hallucination
- Transparency


Safe
Secure & Resilient
Explainable & Interpretable
Privacy-Enhanced
Fair - With Harmful Bias Managed
Accountable & Transparent
Valid & Reliable

NIST AI Risk Mgmt. Framework Characteristics of trustworthy AI systems, Fig. 4

Risk Assessment Example



AI Tools Risk Assessment

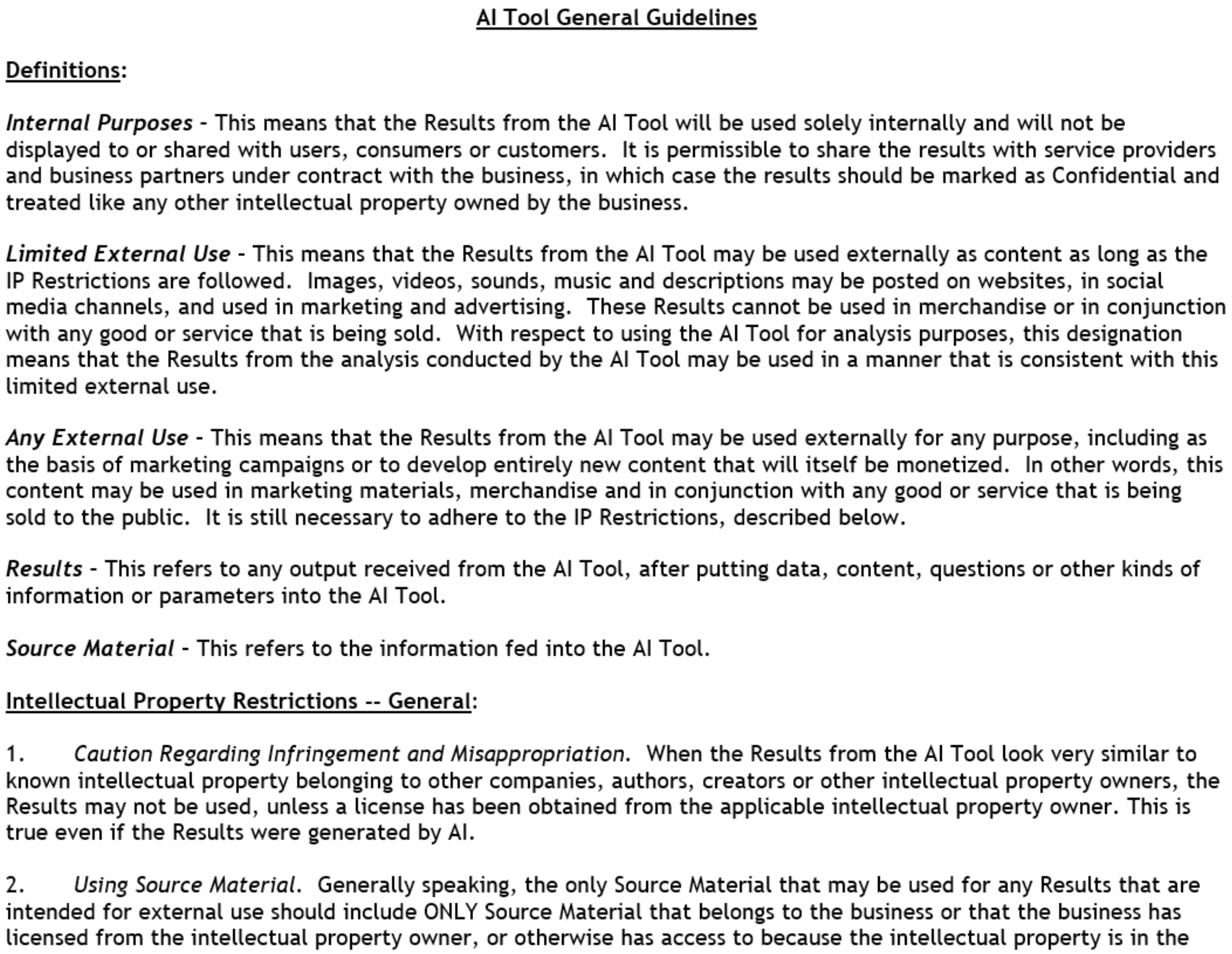
Name of AI Tool: MidJourney

Type of AI: *Generative*

Terms of Use Link: https://docs.midjourney.com/docs/terms-of-service
(last updated 10/24/2024)
Privacy Policy Link: https://docs.midjourney.com/docs/privacy-policy
(last updated not available)
Trademark Policy Link: https://docs.midjourney.com/docs/trademark-policy
(last updated 8/7/2023)

Permitted Use Cases	Internal Purposes	Limited External Use	Any External Use (If Paid User)
Creation of images or videos	X	X	X
Creation of sounds or music	X	X	X
Creation of written descriptions, summaries or scripts	X	X	X
As a virtual assistant	X		
As a chatbot	X		
Analysis of confidential or trade secret data			
Analysis of other data	X	X	X
Draft computer code	X		
Custom Use 1:			
Custom Use 2:			
Custom Use 3:			

Risk Assessment Example



AI Tool General Guidelines

Definitions:

Internal Purposes - This means that the Results from the AI Tool will be used solely internally and will not be displayed to or shared with users, consumers or customers. It is permissible to share the results with service providers and business partners under contract with the business, in which case the results should be marked as Confidential and treated like any other intellectual property owned by the business.

Limited External Use - This means that the Results from the AI Tool may be used externally as content as long as the IP Restrictions are followed. Images, videos, sounds, music and descriptions may be posted on websites, in social media channels, and used in marketing and advertising. These Results cannot be used in merchandise or in conjunction with any good or service that is being sold. With respect to using the AI Tool for analysis purposes, this designation means that the Results from the analysis conducted by the AI Tool may be used in a manner that is consistent with this limited external use.

Any External Use - This means that the Results from the AI Tool may be used externally for any purpose, including as the basis of marketing campaigns or to develop entirely new content that will itself be monetized. In other words, this content may be used in marketing materials, merchandise and in conjunction with any good or service that is being sold to the public. It is still necessary to adhere to the IP Restrictions, described below.

Results - This refers to any output received from the AI Tool, after putting data, content, questions or other kinds of information or parameters into the AI Tool.

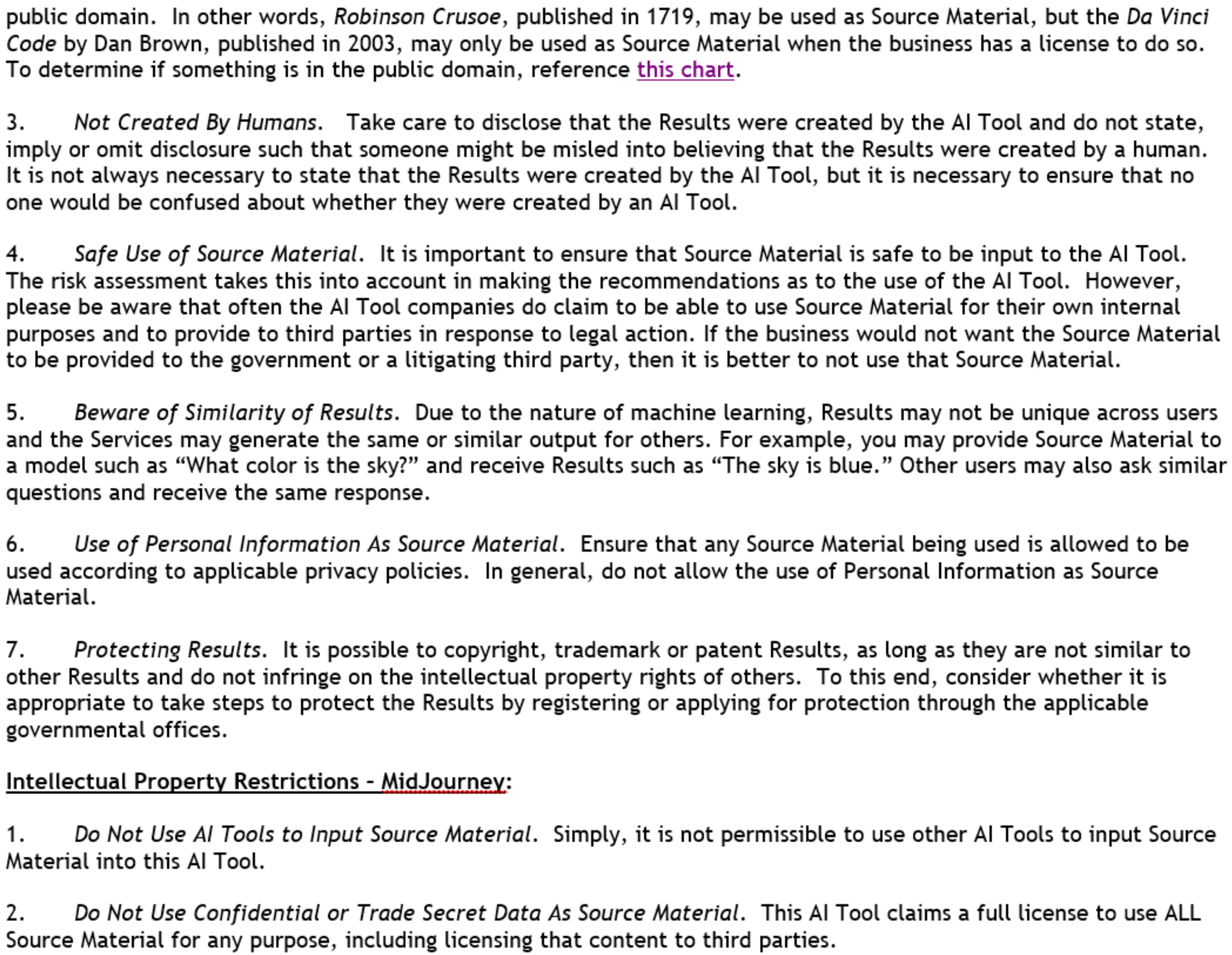
Source Material - This refers to the information fed into the AI Tool.

Intellectual Property Restrictions -- General:

1. *Caution Regarding Infringement and Misappropriation.* When the Results from the AI Tool look very similar to known intellectual property belonging to other companies, authors, creators or other intellectual property owners, the Results may not be used, unless a license has been obtained from the applicable intellectual property owner. This is true even if the Results were generated by AI.

2. *Using Source Material.* Generally speaking, the only Source Material that may be used for any Results that are intended for external use should include ONLY Source Material that belongs to the business or that the business has licensed from the intellectual property owner, or otherwise has access to because the intellectual property is in the

Risk Assessment Example



public domain. In other words, *Robinson Crusoe*, published in 1719, may be used as Source Material, but the *Da Vinci Code* by Dan Brown, published in 2003, may only be used as Source Material when the business has a license to do so. To determine if something is in the public domain, reference this chart.

3. *Not Created By Humans*. Take care to disclose that the Results were created by the AI Tool and do not state, imply or omit disclosure such that someone might be misled into believing that the Results were created by a human. It is not always necessary to state that the Results were created by the AI Tool, but it is necessary to ensure that no one would be confused about whether they were created by an AI Tool.

4. *Safe Use of Source Material*. It is important to ensure that Source Material is safe to be input to the AI Tool. The risk assessment takes this into account in making the recommendations as to the use of the AI Tool. However, please be aware that often the AI Tool companies do claim to be able to use Source Material for their own internal purposes and to provide to third parties in response to legal action. If the business would not want the Source Material to be provided to the government or a litigating third party, then it is better to not use that Source Material.

5. *Beware of Similarity of Results*. Due to the nature of machine learning, Results may not be unique across users and the Services may generate the same or similar output for others. For example, you may provide Source Material to a model such as "What color is the sky?" and receive Results such as "The sky is blue." Other users may also ask similar questions and receive the same response.

6. *Use of Personal Information As Source Material*. Ensure that any Source Material being used is allowed to be used according to applicable privacy policies. In general, do not allow the use of Personal Information as Source Material.

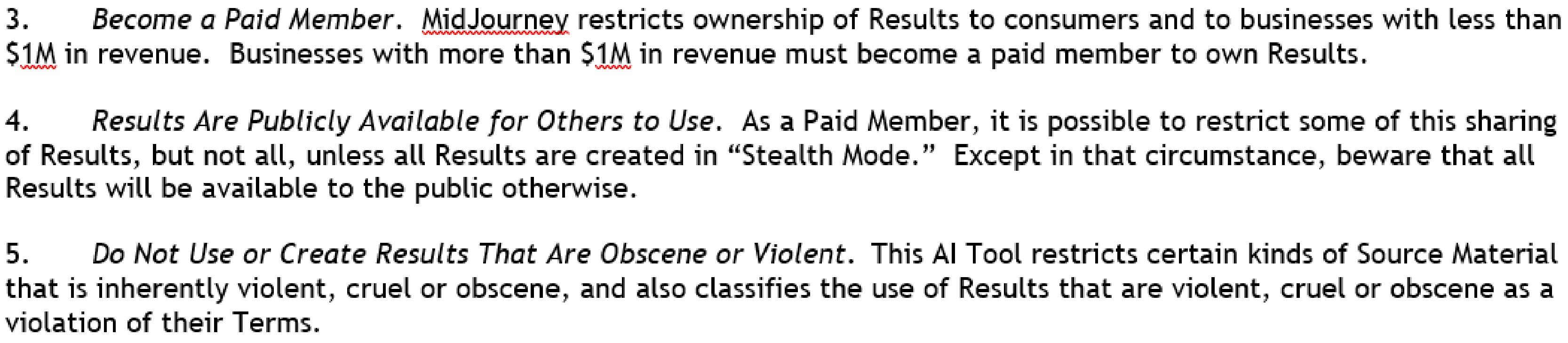
7. *Protecting Results*. It is possible to copyright, trademark or patent Results, as long as they are not similar to other Results and do not infringe on the intellectual property rights of others. To this end, consider whether it is appropriate to take steps to protect the Results by registering or applying for protection through the applicable governmental offices.

Intellectual Property Restrictions – MidJourney:

1. *Do Not Use AI Tools to Input Source Material*. Simply, it is not permissible to use other AI Tools to input Source Material into this AI Tool.

2. *Do Not Use Confidential or Trade Secret Data As Source Material*. This AI Tool claims a full license to use ALL Source Material for any purpose, including licensing that content to third parties.

Risk Assessment Example

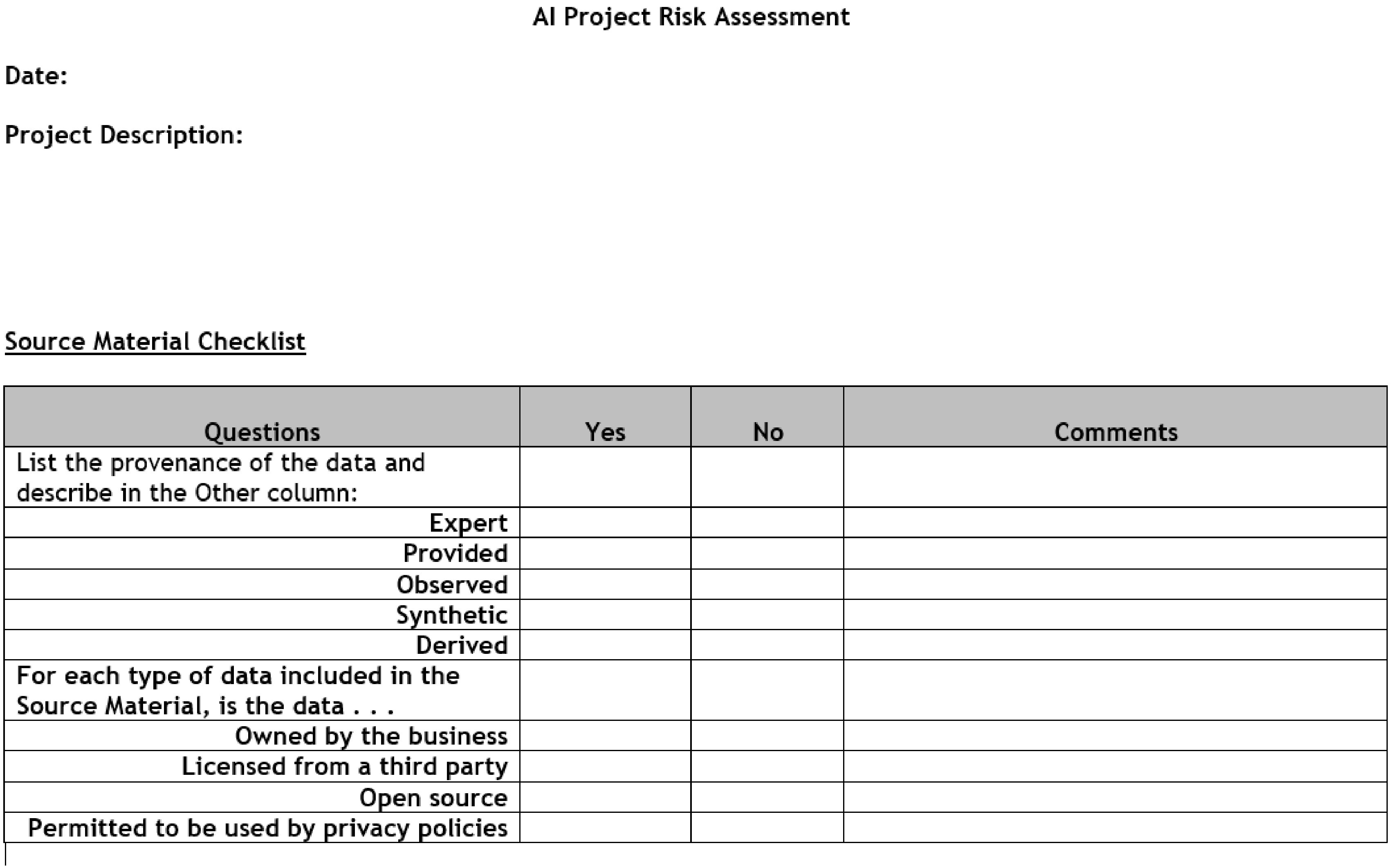


3. *Become a Paid Member*. MidJourney restricts ownership of Results to consumers and to businesses with less than $1M in revenue. Businesses with more than $1M in revenue must become a paid member to own Results.

4. *Results Are Publicly Available for Others to Use*. As a Paid Member, it is possible to restrict some of this sharing of Results, but not all, unless all Results are created in "Stealth Mode." Except in that circumstance, beware that all Results will be available to the public otherwise.

5. *Do Not Use or Create Results That Are Obscene or Violent*. This AI Tool restricts certain kinds of Source Material that is inherently violent, cruel or obscene, and also classifies the use of Results that are violent, cruel or obscene as a violation of their Terms.

Data Checklist



AI Project Risk Assessment

Date:

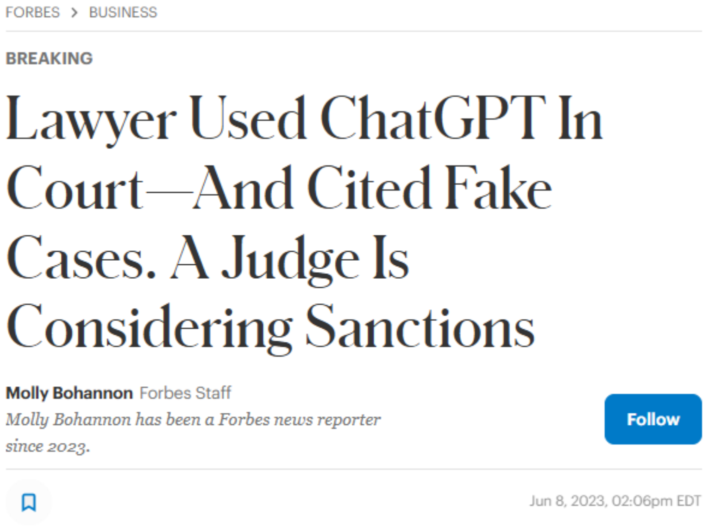
Project Description:

Source Material Checklist

Questions	Yes	No	Comments
List the provenance of the data and describe in the Other column:			
Expert			
Provided			
Observed			
Synthetic			
Derived			
For each type of data included in the Source Material, is the data . . .			
Owned by the business			
Licensed from a third party			
Open source			
Permitted to be used by privacy policies			

Legal Ethical Considerations

- Both the ABA and the NC Bar has issued guidance on lawyers using AI tools themselves, with particular focus on the use of Generative AI tools by lawyers.



FORBES › BUSINESS

BREAKING

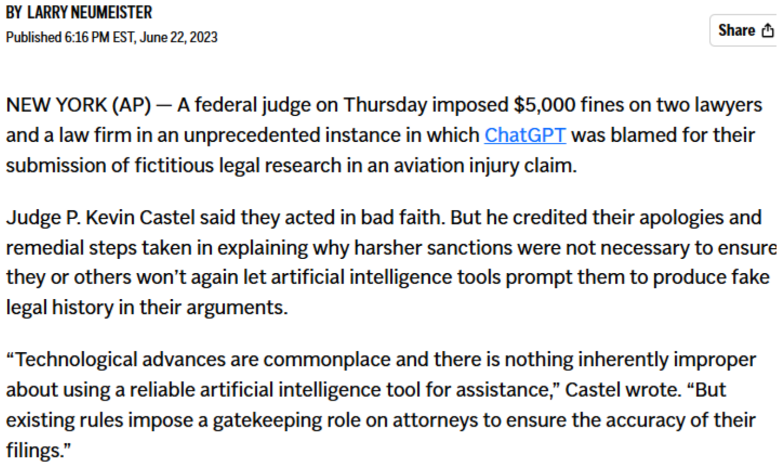
Lawyer Used ChatGPT In Court—And Cited Fake Cases. A Judge Is Considering Sanctions

Molly Bohannon Forbes Staff

Molly Bohannon has been a Forbes news reporter since 2023.

Jun 8, 2023, 02:06pm EDT

Updated Jun 8, 2023, 03:42pm EDT



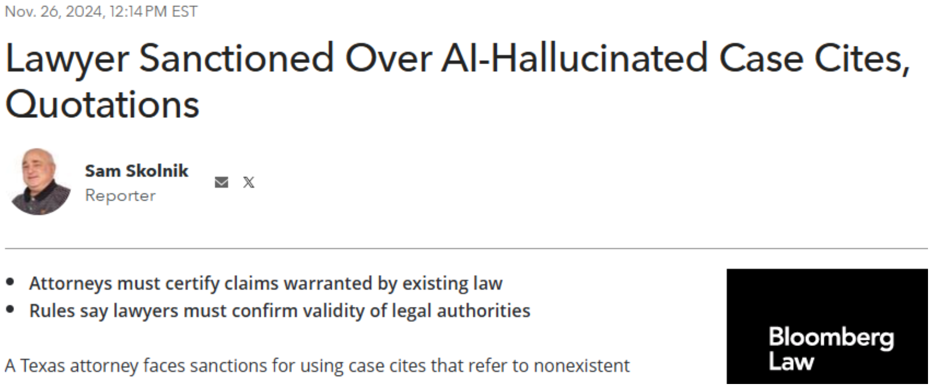
BY LARRY NEUMEISTER

Published 6:16 PM EST, June 22, 2023

NEW YORK (AP) — A federal judge on Thursday imposed $5,000 fines on two lawyers and a law firm in an unprecedented instance in which ChatGPT was blamed for their submission of fictitious legal research in an aviation injury claim.

Judge P. Kevin Castel said they acted in bad faith. But he credited their apologies and remedial steps taken in explaining why harsher sanctions were not necessary to ensure they or others won't again let artificial intelligence tools prompt them to produce fake legal history in their arguments.

"Technological advances are commonplace and there is nothing inherently improper about using a reliable artificial intelligence tool for assistance," Castel wrote. "But existing rules impose a gatekeeping role on attorneys to ensure the accuracy of their filings."



Nov. 26, 2024, 12:14 PM EST

Lawyer Sanctioned Over AI-Hallucinated Case Cites, Quotations

Sam Skolnik
Reporter

- Attorneys must certify claims warranted by existing law
- Rules say lawyers must confirm validity of legal authorities

A Texas attorney faces sanctions for using case cites that refer to nonexistent

AI Laws, Regulations and Guidance

OECD AI Framework

EU Artificial Intelligence Act

Trump AI Executive Order (Jan. 25, 2025)

NIST Artificial Intelligence Risk Management Framework 1.0

Bipartisan House Task Force Report on Artificial Intelligence

State Laws Addressing AI:

- California
  - AB 2013, Generative artificial intelligence: training data transparency
  - SB 942, California AI Transparency Act
- Colorado
  - SB 24-205, Consumer Protections for Artificial Intelligence
- Utah
  - SB 149, Artificial Intelligence Policy Act
- New York DFS
  - Industry Letter Re: Cybersecurity Risks Arising from AI (Oct. 16, 2024)
  - Insurance Circular Letter Re: Use of AI Systems and External Consumer Data . . in Insurance Underwriting and Pricing (July 11, 2024)

Notice and Disclaimer

Cadwalader, Wickersham & Taft LLP (Cadwalader) is a registered limited liability partnership under the laws of the State of New York, USA, with its principal place of business at 200 Liberty Street, New York, NY 10281. The Firm has additional offices in Washington, DC, Charlotte, NC and London, U.K.

This presentation has been prepared by Cadwalader for informational purposes only and does not constitute advertising or solicitation and should not be used or taken as legal advice. Those seeking legal advice should contact a member of the Firm or legal counsel licensed in their jurisdiction. Transmission of this information is not intended to create, and receipt does not constitute, an attorney-client relationship.

United States: The Firm is a registered limited liability partnership under the laws of the State of New York. The personal liability of our partners is limited to the extent provided in such laws. Additional information is available upon request made to our office in New York.

United Kingdom: The Firm is registered in England and Wales and is regulated by the Solicitors Regulation Authority. A list of our partners who are Solicitors or Registered Foreign Lawyers in England and Wales is available for inspection at our London office. Cadwalader lawyers resident in London who are admitted as members of the legal profession elsewhere may also be subject to regulation in their place of admission. Information on the regulations of The Solicitors Regulation Authority can be found in the Solicitors' Code of Conduct at www.sra.org.uk.

Ireland: The Firm operates in Ireland through Cadwalader, Wickersham & Taft Ireland, a general partnership formed under the laws of Ireland and regulated by the Irish Law Society. Cadwalader, Wickersham & Taft Ireland lawyers who are admitted as members of the legal profession elsewhere may also be subject to regulation in their additional jurisdictions of admission.