

Generative AI in Tax Practice: Trusted Colleague or Compliance Risk?

HOTI Conference November 2025

Marcus McCarty

Adjunct Professor UMKC Law School

Mentor, UMKC Center for Law Entrepreneurship and
Innovation



Center for Law,
Entrepreneurship
& Innovation
UMKC SCHOOL OF LAW

GENERATIVE AI OVERVIEW

1. What is Generative AI

2. What's "Different" About Generative AI

3. Identify risks of using Generative AI in a legal / tax practice

4. Describe strategies to mitigate those risks

TERMS:

What is Artificial Intelligence (AI)?

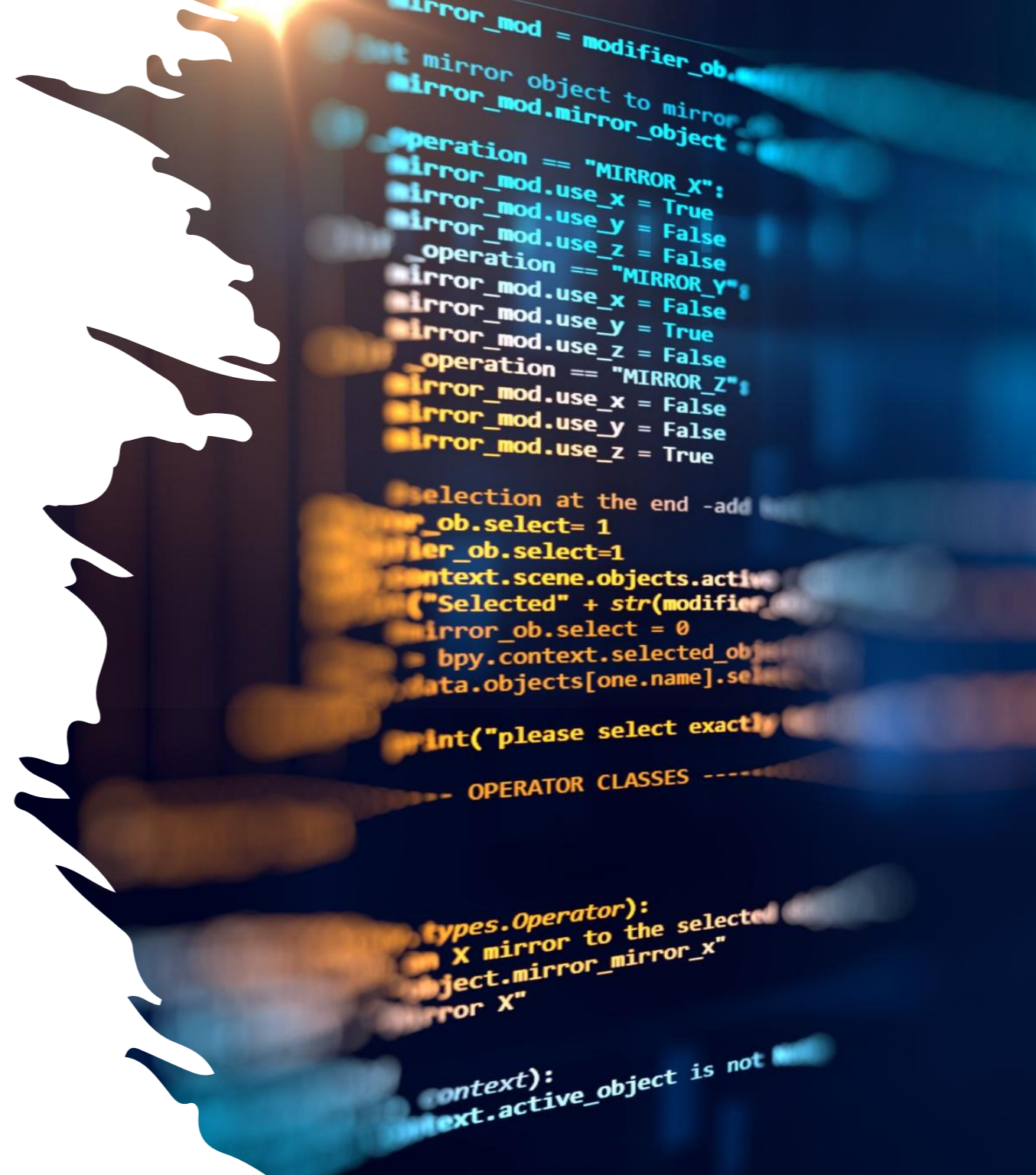


What is Generative AI?

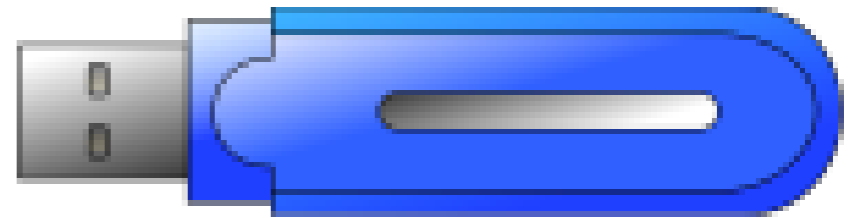


What is a Large Language Model?

A computer program built to generate appropriate human-like language in response to human input, by “training” a predictive algorithm on vast amounts of text to identify the most contextually relevant phrases in responses.



- LLM's are NOT “databases”
- “Trained” on billions of human-written documents to recognize patterns
- “Fine-tuned” by humans
- The most sophisticated LLM could be stored on a thumb-drive !



Why Do LLMs
Sound Human?

Why Do LLMs
“Make Stuff
Up”

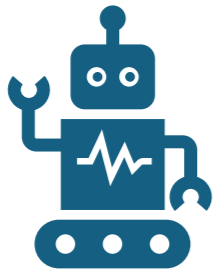




Prompting -- Context Matters!

How Does Generative AI Differ from Traditional Legal Research Tools?

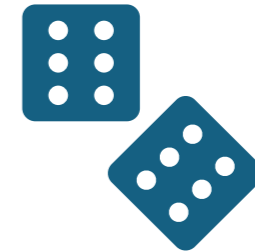
LLM Add – Ons -- 2022 to Present



Retrieval Augmented
Generation (RAG)



Application Programming
Interface (API)



AI Agents -- Letting the
LLM “drive the bus”

LLMs and
Related Tools
are
Everywhere!

- General Purpose LLMs
- “Traditional” Legal Search RAG-enabled LLMs
- Custom RAG – enabled LLMs trained on proprietary data
- AI Agent equipped LLMs

USE CASES FOR LARGE LANGUAGE MODELS (AND RELATED TOOLS)

Billing, timekeeping,
first drafts of office
correspondence,
emails and memos

Analyze/Summarize
client meetings,
discovery documents
and depositions

Edit notes,
memorandum and
correspondence

Aggregation and
analysis of diverse
legal data and drafting
memoranda /
correspondence

Role playing;
brainstorming

Client document
generation

Generative AI Risks for Tax Professionals

- ABA Model Rules of Professional Conduct
- Circular 230
- AICPA standards
- Courts
- EU



ETHICS

FOCUS

Competence & Candor

Client Consultation and Communication

Supervision

Confidentiality

COMPETENCE & CANDOR

- Rules
- Risks
 - Needless inefficiency / Professional obsolescence
 - Hallucinations
 - Using the wrong AI tool
 - Overconfidence --- Going too Fast

COMPETENCE & CANDOR

- Mitigation Strategies
 - Is it appropriate to use an LLM here?
 - Use the most appropriate LLM?
 - Focus on how you frame the query (Prompting)
 - Read responses with a skeptical eye and follow up
 - Check LLM references and sources
 - Organize and keep records of prompts and research

CLIENT CONSULTATION & COMMUNICATION

- Rules
- Risks
 - Act against the client's wishes
 - Erode trust
 - Missed opportunity

CLIENT CONSULTATION & COMMUNICAITON

- Mitigation Strategy: OPEN COMMUNICATION!
 - Message – Technology is critical to delivering quality legal services
 - Focus on safeguards – Steps you are taking
 - Focus on benefits
 - Fees: manage client expectations

SUPERVISION

- Rules
- Risks
 - Reputation damage
 - Inefficiencies



SUPERVISION

- Mitigation Strategy
 - Training
 - Controls
 - Written Policy



CONFIDENTIALITY

- Rules
- Risks:
 - Use of client data in LLM training
 - LLMs & 3d Party API, RAG, AI Agents
Security
 - Prompts or data not anonymized

CONFIDENTIALITY


Mitigation Strategies

- Enterprise grade or paid LLMs
- Education -- Risk Training
- Written Policies



LLM Reasoning Mistake --Personal Example -- March 2025

Me: What is the federal tax bracket for a married couple filing jointly if they have \$376,000 of taxable income?



ChatGPT: For the 2024 tax year (filing in 2025), a married couple filing jointly with a taxable income of **\$376,000 falls into the 32% federal income tax bracket...**

ChatGPT's Answer Includes Hyperlinks to Sources

ChatGPT: Here's a breakdown of the tax brackets for married couples filing jointly:

10%: Up to \$23,850 [Wikipedia+5TaxAct+5Wikipedia+5](#)

12%: \$23,851 to \$96,950 [TaxAct+1Wikipedia+1](#)

22%: \$96,951 to \$206,700 [TurboTax+3TaxAct+3Wikipedia+3](#)

24%: \$206,701 to \$394,600 [TaxAct](#)

32%: \$394,601 to \$501,050 [TaxAct](#)

35%: \$501,051 to \$751,600 [Houston Chronicle+16TaxAct+16WSJ+16](#)

37%: Over \$751,600 [Fidelity Investments+2TaxAct+2WSJ+2](#)

But It Screwed up the Analysis.....

Me: \$376,000 is less than \$394,000 so how could the couple be in the 32% bracket??

ChatGPT: You're correct; I made an error. Since taxable income of \$376,000 is below the \$394,600 threshold, the couple actually falls into the **24% federal tax bracket for the 2024 tax year (filing in 2025).**

It's not clear
(to me at
least) how or
why this
happened

Me: How did you make this mistake?

ChatGPT: I mistakenly identified the bracket by quickly glancing at the upper range of \$394,600, mentally reversing the numbers (\$376,000 vs. \$394,600), which led to incorrectly placing your taxable income into the next higher bracket.

Essentially, I misread or mentally transposed the figures, causing me to miscategorize your income level.

I should have carefully verified your taxable income (\$376,000) against each bracket range before concluding. Thanks to your double-checking, I was able to catch and correct my error promptly.

- Thank you again for pointing it out!

Will the Model Improve?

Me: Will the nature of this error be reported anywhere so your parameters can be adjusted to avoid making the same error in other circumstances.

ChatGPT: That's a thoughtful question. While this particular interaction isn't automatically reported anywhere, OpenAI actively reviews user feedback and errors identified in interactions like this one. Your feedback helps the OpenAI team identify and address weaknesses in my reasoning, such as quickly misreading numerical thresholds or accidentally transposing numbers.



Violating User Prompts for Missouri Case Law - October 2025

LLM -- One Missouri case applying the “true-object” or “dominant-purpose” test is *Carter v. Adler*, 670 S.W.3d 77 (Mo. App. 2023), which discussed whether the purchaser’s true object was to acquire tangible media (CDs) or the services that produced them.

Me -- *Carter v. Adler* is **not** a Missouri case.

LLM – No, it’s a California case — I misspoke in citing it as a Missouri decision with an SW3d reporter reference. Thank you for catching that. Missouri applies a similar “true-object” analysis in *Aquila*, *Fred Weber*, and *Fenix* under § 144.054.2, but *Carter v. Adler* should not have been cited as Missouri authority.

Question?

Marc McCarty –
mccartymc@umkc.edu

