ThinkTwenty20's "Twenty Rules for AI for Financial Professionals": Alpha Version

Starting with some guidelines related to risk: guidelines 1-5

This column launches an ongoing series of postings to develop helpful guidance for financial professionals related to artificial intelligence. I don't know where it will go, but I will begin with a list of guidelines and advice, with the hopes we can collaboratively make some of them more permanent.

We are now 18 months into the ChatGPT era (November 30, 2022 – May 28, 2024). In that year and a half, generative AI – a branch of artificial intelligence that creates electronic content, primarily in response to prompts, based on their input training data – has evolved from a niche tool to one that is visible in every major business software product and social media tool. The pressure to keep up is so great that major companies are facing severe reputational risks as they try to compete.

I am hesitant to assign numbers to these rules, as I think priorities will change and future organization may benefit from reordering them. So I will simply assign a pithy name for now. In this series, I will concentrate on each one and seek your help in refining them, as well as adding to the list.

Let's begin with some risks, things you should know before you start (or keep) providing input to an AI.

- Confidentiality: Don't type anything into an AI that you would not want made public.
- Skepticism: Don't automatically trust anything coming from an AI without review
- Diversification: Don't put all your eggs (AIggs?) in one basket

… and some things you should know before users read/view and share the output

- Compliance: Consider how any output might comply with industry and ethical regulations and standards
- Transparency: Be careful to consider when you need to disclose your use of these tools

Confidentiality: Unless you are self-hosting an AI, type/attach nothing as a prompt and upload nothing for retrieval-augmented generation (RAG) that you would not want to be made public.

From the time of the first popular chatbot, ELIZA, in 1966 to the present, people have wanted to use the impassionate and judgement-free interaction with the computer as a confidant. When ELIZA creator Dr. Joseph Weizenbaum sought to review the interactions between users at MIT and his program, the reaction was quick and furious; people felt it was an invasion of privacy. Sixty years later, we have not learned that our input may be read by the developers, accidentally or purposefully leaked, or otherwise exposed. This will lead to a guideline "**Cybersecurity**".

Skepticism: "Trust, but verify". As financial professionals, we should always be curious and seeking to assess whether information presented to us makes sense, based on evidence, and invest in seeking that evidence based on the risks of information being incorrect. This is true of people and computers.

The topic popularly known as “hallucinations” – which I prefer to call broadly “undesirable results” – is well known in generative AI; as a word probability tool and not a database retrieval tool, generative AI is known for producing reasonable sounding but factually incorrect information. Efforts to minimize the impact of these undesirable results include the provision of information sources that can be verified.

In Russian, the phrase rhymes, “доверяй, но проверяй”, romanized, “*doveryay*, *no* proveryay” It is a Russian proverb made famous in the US by President Ronald Reagan; he learned the expression from Suzanne Massie, an American scholar of Russian history and his trusted advisor. My own interest in the Russian language was sparked by her daughter, with whom I went to high school. Her son, Bob, also attended my high school, and is known as a co-founder of the sustainability standards group, the Global Reporting Initiative (GRI).

Diversification: While many people have heard of ChatGPT, OpenAI’s web-based chatbot is one of many options available. There are free offerings, there are paid offerings, there are front-ends to some or all, you can load some on your PC, Mac, or iPad to run privately

. They will have different functions, excel at some tasks and fall back at others. The comparative strengths will change and evolve.

Do you know when to use ChatGPT, Claude , Copilot, Falcom, Gemini, Grok, HuggingFace Chat, Meta, Perplexity, Phi-3, Pi, Poe, You, ElevenLabs, HeyGen, Firefly from Adobe, Mistral, Llama … the list goes on and on and is ever changing.

With the ever-changing and expanding landscape, it’s easy to pick one – ChatGPT, Copilot, Gemini – and decide that’s enough. But we need an AI for our AI: want to get a great summary of a new Youtube video? Gemini works well directly with Youtube … but you may like the guidance from another tool where you will make the extra step of cutting and pasting the video transcript. Claude is great for text-to-text or image-to-text, but lacks other multi-modal functionality at the moment. So many tools, every changing capabilities, how do you leverage more than one for best of breed? This will connect to a guideline “**Stay up-to-date”**.

Compliance: This is a tough one in many ways, and certainly in the news. So many rules, expectations, and concerns.

Transparency: What disclosures do you need to provide on your use of AI?

Let me know in the comments what you think! Help me build common sense guidance to cope with and benefit from this disruptive phenomenon.