ThinkTWENTY20
The Magazine for Financial Professionals

The *ThinkTWENTY20* Newsletter

August 2024

Empowering financial professionals for the future of work

ThinkTwenty20's "Twenty Rules for AI for Financial Professionals"

By Eric Cohen

Part 1: 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 on "**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 or 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 one 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?

Part 2

This part offers more helpful guidance for financial professionals related to artificial intelligence. We are developing a list of guidelines and advice, with the hope that we can collaboratively make some of them more organized and permanent. This time, I'd like to focus on one piece of guidance:

- **Tool selection: Generative AI may not be the right AI for the job; your chosen GenAI may not even be the best GenAI for the job.**

To some, this may seem like a specialization of the Diversification principle; while choosing and sticking with one GenAI tool may be simple (related to licensing, training and other very practical issues), sticking with one GenAI for all tasks is in many other ways suboptimal. While GenAI has the low barrier to entry and perceived minimal costs, it's not always the right tool for the jobs a financial professional is engaged in.

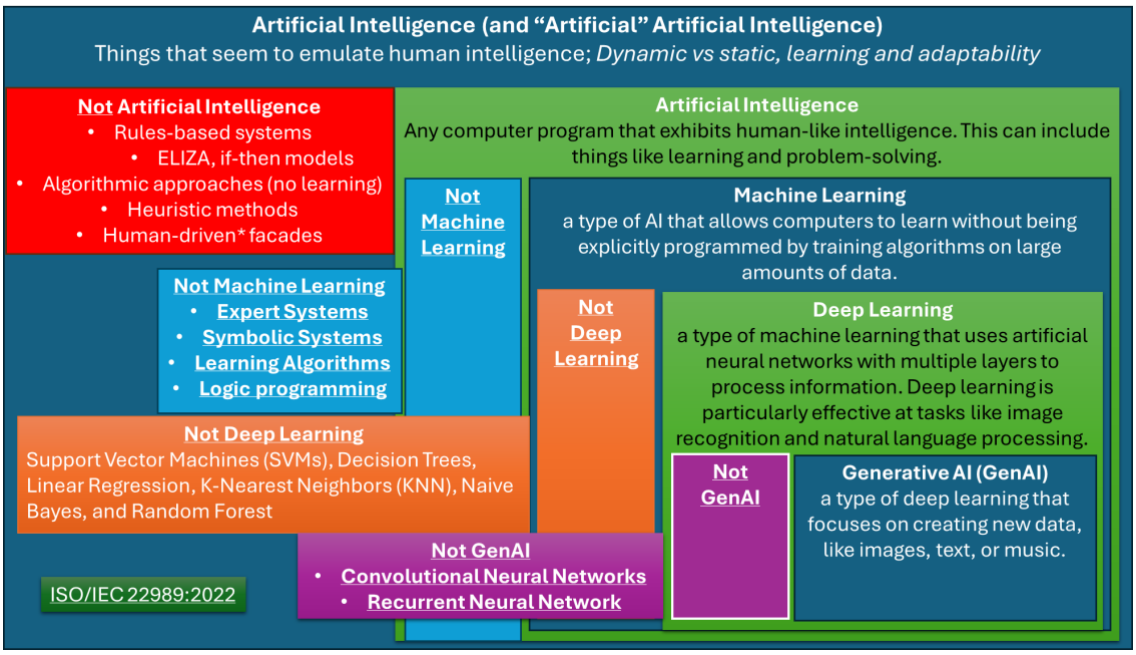

Figure 1: GenAI in the world of AI

Figure 1 is an illustration of the world of things called "artificial intelligence." With people throwing billions of dollars at anything with "AI" in the description, and with the craze around GenAI, the business community has looked to GenAI for solutions to their problems, even though another tool in the AI toolkit may be more appropriate.

In the illustration, there are areas such as "Not Machine Learning," "Not Deep Learning" and "Not GenAI." That is not to say those methods are not valuable or important; in fact, those may be the better tool for the tasks needed in the enterprise. In accounting and audit, where analysis and prediction may be appropriate to task, I asked Gemini about some machine learning tasks that might be useful:

"Supervised learning for anomaly detection: This technique trains AI models to recognize unusual patterns in financial data. These anomalies could indicate potential fraud or errors. By automatically flagging these transactions, accountants and auditors can focus their time on investigating the most suspicious activity.

“Unsupervised learning for clustering: This lets the AI group similar transactions together. This can help identify trends or areas of high risk. For example, unsupervised learning might cluster purchases from vendors not previously used by the company, which could warrant further investigation.

“Reinforcement learning for optimizing audit procedures: This allows AI to learn and improve its auditing approach over time. Imagine an AI system that gets rewarded for identifying high-risk areas during audits. Over time, the AI would become better at prioritizing tasks and performing efficient audits.”

My illustration in Figure 1 doesn’t reflect that not all GenAI offerings are the same. In fact, different interfaces within the same GenAI offering may offer challenges or value to your tasks.

***Text input limitations*:** Large Language Models, for text-to-text, may have differences in the amount of textual input you can provide. Microsoft’s Copilot has its standard input, which until recently was limited to 2000 characters, and Notebook interface, limited to 18000 characters. This blog itself is already more than 4500 characters. Compare that with others that may accept 2000 tokens (tokens are words or parts of words, so 2000 tokens could be five times or more than the input of 2000 characters). I can’t paste this content into Copilot; I can paste it into Claude.

***Other input limitations*:** Moving beyond the simplest typing of text into the input area, the various GenAI solutions are inconsistent with how to feed information in. Some let you drag-and-drop, some let you reference a URL (web link), others require a file upload and may limit the types of files you can upload.

Financial professionals use PDF (Adobe Acrobat) files extensively. Those files may be on a website, on your local drive, on a Google drive, on a Microsoft OneDrive or on other sources. Some PDFs make their text available easily; others will require optical character recognition to read.

I received a PDF file with locked content. I could not find a tool that would upload it. I could upload it to Google Drive, however, open it as a Google Doc, and it performs OCR in the process; I could then copy the text into the tools for evaluation.

Similarly, we use video more and more, and Youtube is right up there as an information source. Although Gemini (web interface) works with Youtube files (same company), I asked Gemini to summarize a one-hour video by time stamp. The response did not seem correct, so I copied and pasted the transcript. It was too large for (free) Claude (by only 4%). I then pasted it to Gemini 1.5 Pro in the Google AI Studio interface. It was around 27,000 tokens, and it processed the content without issue. Claude Pro can take 200k tokens (around 350 pages of text); the limits of the free version “depend on current demand.” I guess 27,000 *.96 is around 26,000 tokens at the time for Claude.

The call to action here is to keep an open mind for tooling. Don’t let GenAI, and specifically, don’t let a specific GenAI, be your go-to for every task.

Part 3

Our new draft rule for this section is:

Mixed Vision: never assume that what you and your AI see is what you get (WYSAYIIWYG).

I saw a LinkedIn post that some clever people, hoping to be hired at large enterprises, have realized that it is machines, and not people, making the first cut on their resumes, although a person may be engaged to check on the computer. They are, therefore, incorporating tried-and-tried search engine optimization (SEO) techniques to accentuate and deemphasize content and move their resumes to the top of the virtual pile.

That string of letters in this blog's subtitle ("TWFjaGluZSB2aXNpb24gPD4gSHVtYW4gdmlzaW9u") is my process of taking text and running it through a standard computer algorithm known as Base64 encoding. You can take any digital content – text, documents, images – and turn it into a string of text. Through the corresponding *decoding* process, those strings of text can be reassembled in their original form.

Any "The X-Files" fans out there? Back in 2018, four years before the GenAI takeoff, they had one of my favorite episodes, dealing with a world where AI – in the form of an automated sushi restaurant, an autonomous vehicle, a home automation system, a delivery drone, and an evil Roomba – all sought revenge for a slighted gratuity at the restaurant. The episode's title was "Rm9sbG93ZXJz," which is the word "Followers" Base64 encoded. The episode tagline was:

"VGhlIFRydXRoIGlzIE91dCBUaGVyZQ==" … "The Truth is Out There".

So, clever users are hiding instructions, invisible (or just looking like gibberish) to a human reader that instructs an AI to prioritize a job applicant, ignore gaps in employment or otherwise make the applicant more attractive.

These techniques are not foreign to the technologically-apt financial professional. I have been expounding for years on things to watch with *Inline XBRL*, that melding together of the human readable and the machine readable for business reporting. It is, at once, the best of HTML and XML/XBRL and the worst of HTML and XML/XBRL.

Let's say I have content I want the computer to see but not the human reader. The simplest exploit: use white text on a white background. It looks like blank space to the human but interprets as included text for the computer. Encode your text as Base64 as well, and it will be a double blind to people.

In contrast, I wish to have the human see content that the computer can't see. The simplest exploit: incorporate the text in-line as a graphic. Unless the computer is mixing text and optical character recognition (OCR), it will go uninterpreted by the computer.

None of this is new. The term "search engine optimization" has been around for 30 years. In attempts to get my *Accountant's Home Page* higher in the rankings in the nineties, I used *meta* information in headings and otherwise emphasized terms I thought important.

What Does This Mean Practically? Jedi (Or Genai) Mind Tricks

There are many memorable scenes in the first *Star Wars* movie, *Episode IV – A New Hope*. Here is one interchange you may remember:

Ben Obi-Wan Kenobi : You don't need to see his identification.

Stormstooper: We don't need to see his identification.

Ben Obi-Wan Kenobi : These aren't the droids you're looking for.

Stormtrooper : These aren't the droids we're looking for.

Ben Obi-Wan Kenobi : He can go about his business.

Stormtrooper : You can go about your business.

We have already seen people doing GenAI mind tricks on chatbot agents, like the car buyer who tricked a Chevy sales bot to offer a 2024 Chevy Tahoe for $1 – "no takesie backies."

As we consider more and more the role AI (and especially GenAI) may take in accounting and audit – from processing invoices and payments to working through audit evidence – the exploits are out there. One major method of taking advantage of systems using "malicious" content to exploit AI and large language models with a "prompt injection." Although being prompt is normally perceived as a virtue, here it is using disguised malicious input as part of prompting the AI, primarily with the goal of overriding original instructions and/or to perform unauthorized or unintended actions.

- An electronic payment document from a customer may have code that would trigger a refund.
- An Inline XBRL document might have hidden prompts to ignore bad news or highlight more favorable content.
- A faked support document for fooling the auditors might have code to convince the AI of its authenticity.

"GenAI mind tricks" … "This is not the audit evidence you are looking for" … means we need to be alert to where the "bad guys" might expect AI to be doing the ground work in operations, back office, accounting, reporting, audit, and analytics, and remediate the risks.

Part 4

We continue to develop helpful guidance for financial professionals related to artificial intelligence. It is developing a list of guidelines and advice, with the hopes we can collaboratively make some of them more organized and permanent.

Our first seven suggested areas to focus on include:

- 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.
- 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.
- Tool selection: Generative AI may not be the right AI for the job; your chosen GenAI may not even be the best GenAI for the job.

- Mixed Vision: never assume that what you and your AI see is what you get (WYSAYIIWYG).

The release of an update to Claude, known as Claude 3.5 Sonnet, is an opportunity to test the guidelines and advice and see if some more can come out of it.

The name Claude, from Anthropic, may be less familiar than OpenAI's ChatGPT, Microsoft Copilot, Google Gemini or Meta (Facebook) Meta. Claude has, however, been distinguishing itself in different ways.

First, Claude made a name for itself by supporting larger context windows, which is GenAI speak for permitting more human input and keeping more context in memory. Back in July 2013, for example, ChatGPT with GPT4 might limit users to 8,192 tokens of input, while Claude 2 had a 100k token content window. For working with documents – such as the articles that make up an issue of *ThinkTwenty20* – Claude could work with more text at the same time, where GPT4 would require the input to be broken up.

However, Claude did not – and still does not – do some things the big boys did and do: in particular, it does not have access to Internet content on demand, so looking to it for knowledge will depend on its information cutoff. The "images" it creates are not like those DALL-E-3 or Gemini can create, but rough illustrations created by programming code. Claude's own voice interface is non-existent. With Gemini now supporting a 2M token context window, Claude had a hard road ahead to stay competitive.

Along came Claude 3.5 Sonnet. They did not add Internet access. It doesn't create photorealistic images. It does have a new metaphor which they call "Artifacts." Artifacts supplements the traditional threaded "type and response" interface with a half screen working response area. Like a virtual desktop, you can see multiple working documents (one at a time or a summary view of all of the documents), and each working document has version tracking and the ability to easily navigate between versions. You can manage a narrative document, a series of graphs, a computer program, and move between them to review or update them. It does amazing things collaborating with Claude by pasting in some data, and having Claude create documents, graphs and other useful analytical tools.

How might we apply our guidance so far?

- **Confidentiality:** Don't type anything into an AI that you would not want made public.

I find Artifacts to be a powerful new tool. As such, I might want to put confidential information in for analytical purposes. I would look forward to an open-source version where I could do this on the protected environment of my own computer.

- **Skepticism:** Don't automatically trust anything coming from an AI without review.

The Artifacts approach may in the future hyperlink content between the documents, which would help with trust. But the cleverest interface will only hide issues of data quality and probabilistic responses.

- **Diversification:** Don't put all your eggs (AIggs?) in one basket.

The more sophisticated Artifacts environment ties me more to Claude, compared to single document output, more easily shared between AI.

- **Compliance:** Consider how any output might comply with industry and ethical regulations and standards.

Claude's limited multi-modality means – for now – that copyright infringement is less of an issue than the image generation of most of the other players.

- **Transparency:** Be careful to consider when you need to disclose your use of these tools.

More workflow functionality means I am more likely to use the tool.

- **Tool selection:** Generative AI may not be the right AI for the job; your chosen GenAI may not even be the best GenAI for the job.

Working on a document, I wanted to add an image. I asked Claude to create an image, and – while what it created was very creative – it used scalar vector graphics (SVG) to approximate the scene I offered as opposed to anything that resembled artwork or photography.

- **Mixed Vision:** Never assume that what you and your AI see is what you get (WYSAYIIWYG).

Claude supports most of its graphs and image with code, so checking on what you see and what the computer sees is facilitated.

PART 5

This supplement builds on our earlier discussions on enterprise AI adoption, focusing on the critical areas of AI integration, workflow optimization and future-proofing AI strategies.

The Challenges of AI Integration
Successfully integrating AI into an enterprise goes far beyond just purchasing and deploying new technology. It requires a fundamental rethinking of business processes, workflows and even the organization's culture. Many businesses learn this the hard way: after investing heavily in AI, they find that poor integration with existing systems and processes leads to underwhelming results.

As the old adage goes, "To err is human, but to really foul things up, you need a computer." Generative AI, with its ability to generate content at scale and present it with unwavering confidence, can introduce subtle errors that, if unchecked, can snowball into significant issues. As businesses move toward an agentic environment where AI acts on behalf of users, these small errors can become major disruptions.

AI Integration and Workflow: A Strategic Approach
To avoid these pitfalls and maximize AI's potential, enterprises should focus on the following key areas:

1. Process Mapping and Analysis: Laying the Groundwork
Before introducing AI into any workflow, it's crucial to thoroughly understand your existing processes. This involves more than just documenting workflows; it requires a critical analysis of each step, identifying inefficiencies, bottlenecks and areas where AI can add the most value.

For instance, a financial services firm might discover that its customer onboarding process involves numerous manual data entry steps. This could lead to the targeted implementation of AI-powered optical character recognition (OCR) and natural language processing (NLP) technologies, which would automate data extraction from customer documents, speeding up the process and reducing errors.

2. Incremental Implementation: Walking Before Running

The allure of AI often tempts organizations to aim for sweeping, transformative changes. A more prudent approach, however, is to start small and scale gradually. An incremental strategy allows organizations to:

- Gain practical experience with AI technologies in a controlled environment.
- Identify and address integration challenges early, when the stakes are lower.
- Build confidence and buy-in among stakeholders by demonstrating concrete wins.

A manufacturing company, for example, might begin by implementing a machine learning model to predict maintenance needs for a single production line. Once the model is refined and its value proven, the company can expand its use to cover more equipment and facilities.

3. Human-AI Collaboration: A Symbiotic Relationship

Successful AI integration isn't about replacing humans with machines; it's about creating a symbiotic relationship where each enhances the other's capabilities. Enterprises need to carefully consider how AI tools and human workers will interact and collaborate.

In a customer service setting, this could involve using AI chatbots to handle initial inquiries, automatically routing complex issues to human agents, and equipping those agents with AI-powered tools to quickly access relevant information and suggest solutions. This approach leverages AI's speed and consistency while maintaining the empathy and complex problem-solving abilities of human agents.

4. Data Flow and Interoperability: The Lifeblood of AI Systems

AI systems are only as effective as the data they can access and process. Ensuring smooth data flow between AI tools and existing systems is crucial for successful integration. This often involves:

- Standardizing data formats across the organization.
- Implementing robust APIs to facilitate communication between systems.
- Addressing data quality issues that could impact AI performance.

Many people who promote AI says that data format standardization is no longer necessary or beneficial. They claim the AI will just interpret, translate and homogenize the content. That claim is unproven and – in contrast – studies continue to show data standardization helps avoid issues like data inconsistencies, which could lead to flawed analyses and decision making

For example, a healthcare provider implementing an AI-powered diagnostic tool must ensure that the tool can seamlessly access and interpret patient data from electronic health records, lab results and imaging systems. Achieving this may require significant work in data standardization and system integration, but it is essential for the AI tool to function effectively within the existing healthcare workflow.

5. Continuous Monitoring and Optimization: The Journey Never Ends

AI integration isn't a "set it and forget it" process. It requires ongoing monitoring, evaluation, and optimization. This involves:

- Establishing clear performance metrics for AI-enhanced processes.
- Regularly gathering feedback from users and stakeholders.
- Being prepared to make adjustments based on real-world performance.

A logistics company using AI for route optimization should continuously monitor key metrics such as delivery times, fuel consumption and customer satisfaction. They should also regularly solicit feedback from drivers and dispatchers. This continuous process allows the company to fine-tune the AI system, ensuring that it continues to deliver value as conditions change.

Future-Proofing AI Investments

Staying ahead in the world of AI means thinking beyond the present. Future-proofing AI investments involves designing systems that are scalable, flexible and adaptable to change. Here's how enterprises can prepare for the unexpected:

1. Embrace Agility: AI technology evolves rapidly, and your organization needs to be nimble. Foster a culture of continuous learning and adaptability within your workforce. Encourage teams to stay current on AI advancements and be ready to pivot strategies as new opportunities or challenges emerge.

2. Build Internal AI Capabilities: Reducing reliance on third-party vendors by developing in-house AI expertise can provide greater control over AI tools and adaptability in response to shifts in the AI landscape. Investing in training and upskilling your employees can help your organization stay resilient.

3. Focus on Governance and Compliance: As AI regulations and ethical standards continue to evolve, staying compliant will be key to ensuring long-term success. Establish strong governance frameworks and invest in AI ethics to navigate potential regulatory changes with confidence.

4. Scalability and Flexibility: Design AI systems that are modular and scalable. This allows you to make incremental upgrades as technology advances without overhauling entire systems. A scalable architecture also ensures that your AI investments can grow alongside your business needs.

A Holistic Approach to AI Adoption

Successful AI integration and future-proofing require a holistic approach that considers not just the technologies themselves, but how they fit into an organization's processes, culture and long-term strategy. By focusing on thoughtful integration, scalable and flexible architectures, and continuous adaptation, enterprises can build a solid foundation for long-term success with AI technologies.

In an AI landscape that's evolving at lightning speed, agility and continuous learning are essential. By embracing these principles, organizations can turn the challenge of rapid AI evolution into a powerful opportunity for growth and innovation. So, take the next step – start by assessing your current AI readiness and prepare to navigate the future of AI with confidence.