Ebook

From Hype to Reality: AI in Government Contracting

Practical Insights to Harness AI
for Proposal and Contract Success

VisibleThread.com

Introduction

Generative AI (Gen AI) took the world by storm in Nov '22 coinciding with the release of ChatGPT. Overnight, Artificial Intelligence (AI) moved mainstream.

People played with ChatGPT, and many will remember that time when we were using it to "write jokes about New Yorkers in the style of Jerry Seinfeld". It appeared to possess magical qualities, to have "intelligence" even.

For some, AI represented exciting possibilities; for others, it suggested unknown fears. In February '24, we considered how it might impact government contracting during our Optimize24 AI event, and subsequently in a series of webinars featuring leading industry experts.

We brought together influential experts from diverse sectors and companies like Leidos, Amazon, and AECOM. Our collective learnings are many, but key takeaways include:

- **Gen AI is very powerful for many tasks during the Capture, Proposal and Contract lifecycle**
- **It is a tool and will drive large efficiency and quality improvement**
- **It is excellent at certain Jobs To Be Done (JTBD) and should not be used for others**
- **We are still in a somewhat experimental phase**
- **It will supplant rather than displace specific roles**

Now, as we approach the 2-year anniversary of ChatGPT and the dust has settled, it's a great time to drill into what we now know. The hype cycle has abated somewhat; only by being equipped with the right information can we navigate this all important topic.

We want to list out the key considerations around artificial intelligence and how it impacts proposal management, Business Development (BD) and Capture throughout the proposal lifecycle.

Buckle up for an insider's look into when it's appropriate to use Gen AI. When is it utterly inappropriate to use it. The risks associated with it, and how, when used for the right jobs, AI can revolutionize how you work.

Whilst there are huge advantages to using AI, the impact and cost of using it inappropriately can at best be non-productive, and at worst create business risk.

What this Guide will cover

AI and Generative AI (Gen AI) – The landscape

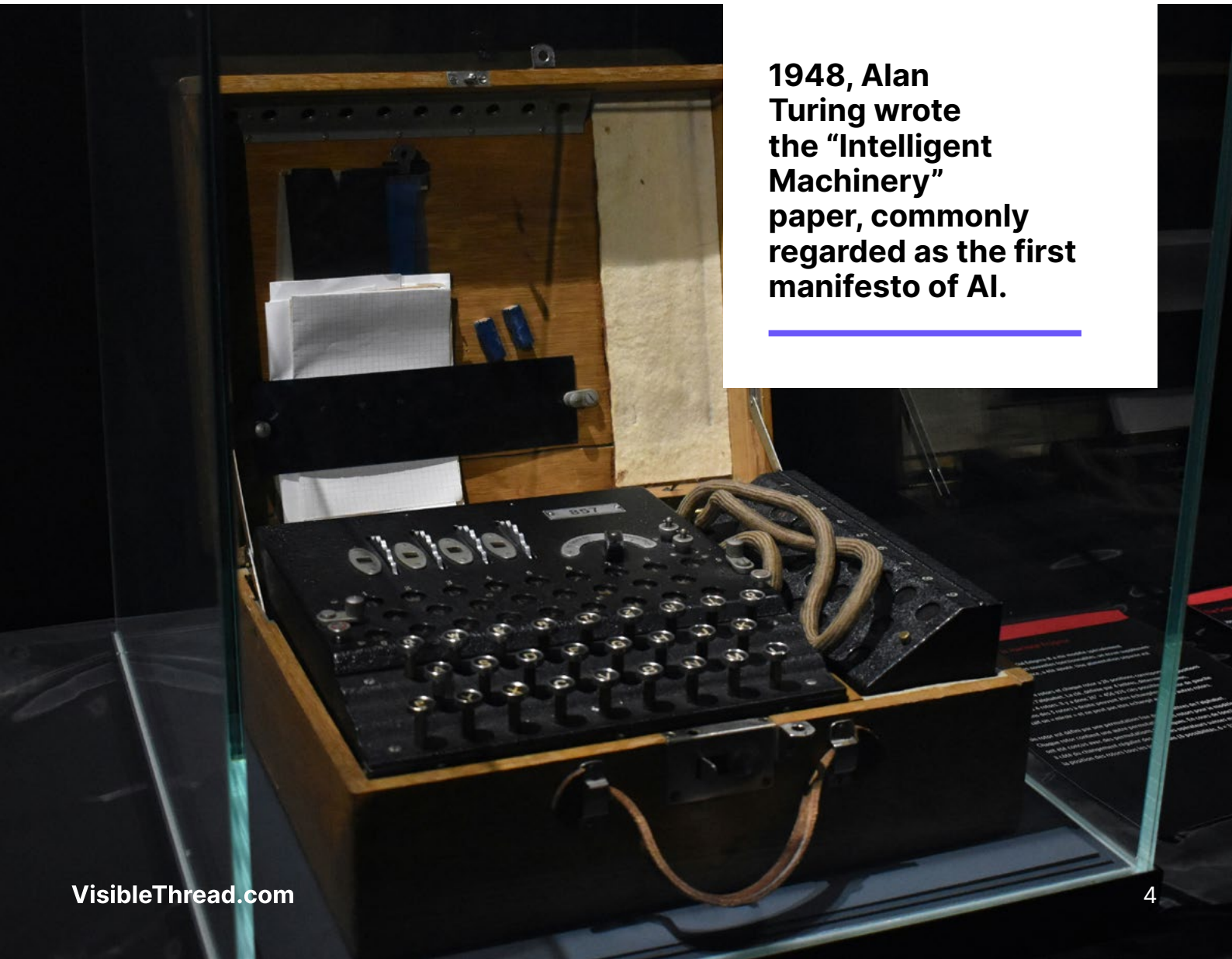
Artificial Intelligence (AI) has been around for decades.

In fact, in 1948, Alan Turing wrote the "Intelligent Machinery" paper, widely considered to be the first manifesto of AI.

At the highest level, you can categorize AI into two forms, machine learning and deep learning. These can be further sub-divided into things like, Natural Language Processing (NLP), computer vision, robotics, and indeed Generative AI (Gen AI). In fact, there are dozens of different types of AI.

So, Gen AI is just one form of AI. The Centers for Disease Control and Prevention (CDC) have used multiple types of AI for disease modeling for over three decades.

This is why our Founder and CEO, Fergal McGovern, spends so much time writing about the best ways to use AI (and Gen AI) in his newsletter and on webinars with VisibleThread customers.



1948, Alan Turing wrote the “Intelligent Machinery” paper, commonly regarded as the first manifesto of AI.

First, let's get some definitions out of the way:

- **AI** enables computers and machines to simulate human thought processes by learning from past experiences and adapting to comprehend and react to natural language, make informed decisions, and solve complex issues.

- **Gen AI** is a very specific type of AI that predicts the next word in a sequence of words. Or if used for image generation, it can predict the next pixel in a sequence of pixels. It is a type of AI called deep learning.

- **Large Language Models (LLMs)** are the "engine" of Gen AI. When you use an application like ChatGPT or VisibleThread's VT Writer, the results come from an LLM. They are trained on a very large dataset. LLMs are especially good at language translation, text summarization, question-answering and content generation.

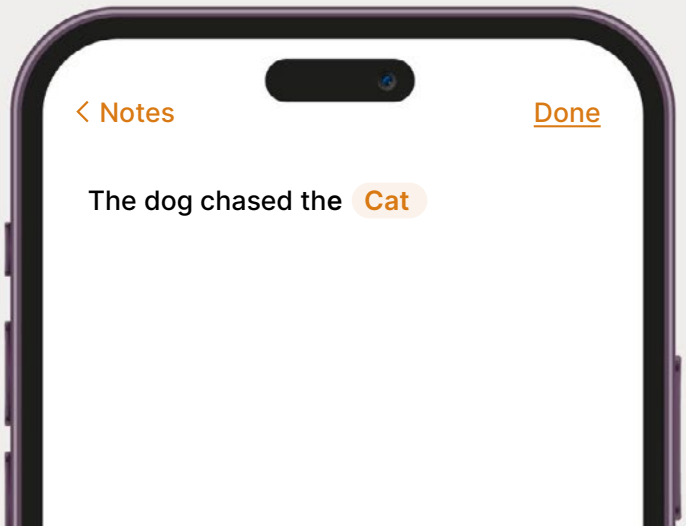
- **Deterministic Software** is software that uses rules and mathematical logic to arrive at a result. Deterministic models guarantee 100% accuracy and 100% repeatability, whereas AI is based on predictions and probability. Despite the proliferation of AI, a very large majority of the software in the world is deterministic. This covers the apps you use everyday, ranging from MS Mail to Adobe Photoshop to MS Word. Deterministic software won't go away, as it is essential for certain tasks, rather it will exist alongside AI.

- **Hybrid Software** combines both AI and deterministic models. For instance, MS Word is a hybrid product, the vast majority of which is deterministic. VisibleThread's VT Docs is also a hybrid product as it uses NLP, a type of AI.

- **Hallucinations** occur when AI generates plausible but incorrect or misleading content due to gaps or misinterpretations in data.

- **Retrieval Augmented Generation (RAG)** allows you tap into your proprietary data when using Gen AI. It doesn't require you to further "train" the LLM.

Now that we have a sense of the AI general landscape, let's take a deeper dive into how Gen AI actually works.



Gen AI – How it works

Gary Marcus, a luminary in the AI space sums it up nicely ....

Gary Marcus
@GaryMarcus

A leading expert on AI, spoke at US Senate AI Oversight Subcommittee, Founder/CEO of Geometric Intelligence (acq. by Uber), TED speaker, host Humans vs Machines

"LLMs generate responses by randomly sampling words based in part on probabilities. Their "language" consists of numbers that map to tokens."

It's all based on probabilities. Simply put, given a block of text, LLMs predict the next word.

"The dog chased the _____"

"The dog chased the **cat**"

That's all LLMs do, they predict next words. There is no logic involved. And the next word is dictated by the data held in the model. So far, so good. Now the answer could also have been:

"The dog chased the **car**"

So, whether it's cat or car is entirely down to the model's training, plus some other factors.

Types of Software: Gen AI vs. Deterministic

Ok, let's step back. Have you ever wondered how software worked prior to Gen AI coming on the scene?

Which broadly coincided with ChatGPT's emergence in Nov 2022.

Well, most software products used (and continue to use) a "deterministic" approach. Deterministic software works on logic, rather than probability. It produces the same results given the same inputs, every time.

For instance, if I give a calculator app the inputs 2 and 3 and ask it to add these, it will always produce the same output 5 (because 2 + 3 = 5). This is deterministic, there is one and only one answer, and logic in this case dictates it's 5.

Now GenAI has not at all displaced deterministic software, rather it is complimentary. There's a time and a place to use deterministic software and a time and a place to use Gen AI and other forms of AI.

If you want 100% accuracy and 100% repeatability, you should use deterministic models. For more creative activities, Gen AI can be great.

Here's a good summary:

Generative AI

Characteristics:

1. Answers are variable
2. The "truth" depends on the model
3. Results are generated based on predictions (probability)
4. Hallucinations a factor (25% +/-)

Deterministic Software

Characteristics:

1. 100% accurate, 100% repeatable
2. Logic, not predictions
3. Results are based on logical steps in code
4. No Hallucinations

Using MS Word to prove the point

Let's use a simple and very familiar product to show the difference between a task that requires a deterministic approach and what happens if you use Gen AI. We'll use MS Word to prove the point.

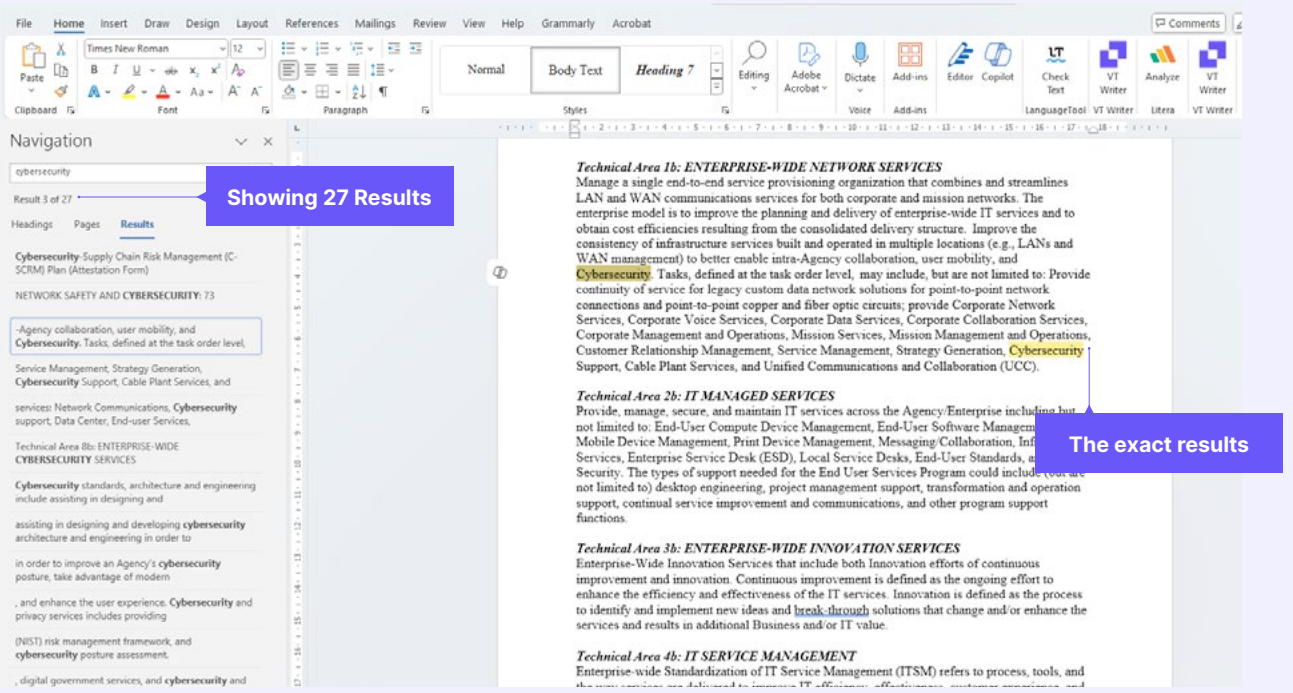
The Business Context:

When reviewing RFPs and contracts, proposal and contract managers need to identify important requirements. Contract managers will typically look for occurrences of terms like "indemnify", "warrant" etc. in a contract. A proposal manager will look for terms like "will", "shall" and "must" indicating mandatory requirements that they must comply with. Similarly, solution architects will often search for technical language like "cybersecurity" etc.

They typically use search (CTRL + F) in MS Word to do this job. Or the equivalent for a PDF. Now if you're a VT Docs user, you'll use Dictionary Search with specialized dictionaries to speed up this process by 10x and shred the doc too. But let's stick with MS Word for this example.

On a recent webinar featuring Geoff d'Alelio, a Principal Enterprise Solutions Architect at Broadcom, we used MS Word to search the $60b NASA SEWP VI RFP issued in May 2024. It's a 151-page doc and reasonably big. Geoff mentioned how he normally searches for technical terms like "cybersecurity" etc.

Let's do that for our NASA SEWP example. Here's what we get:



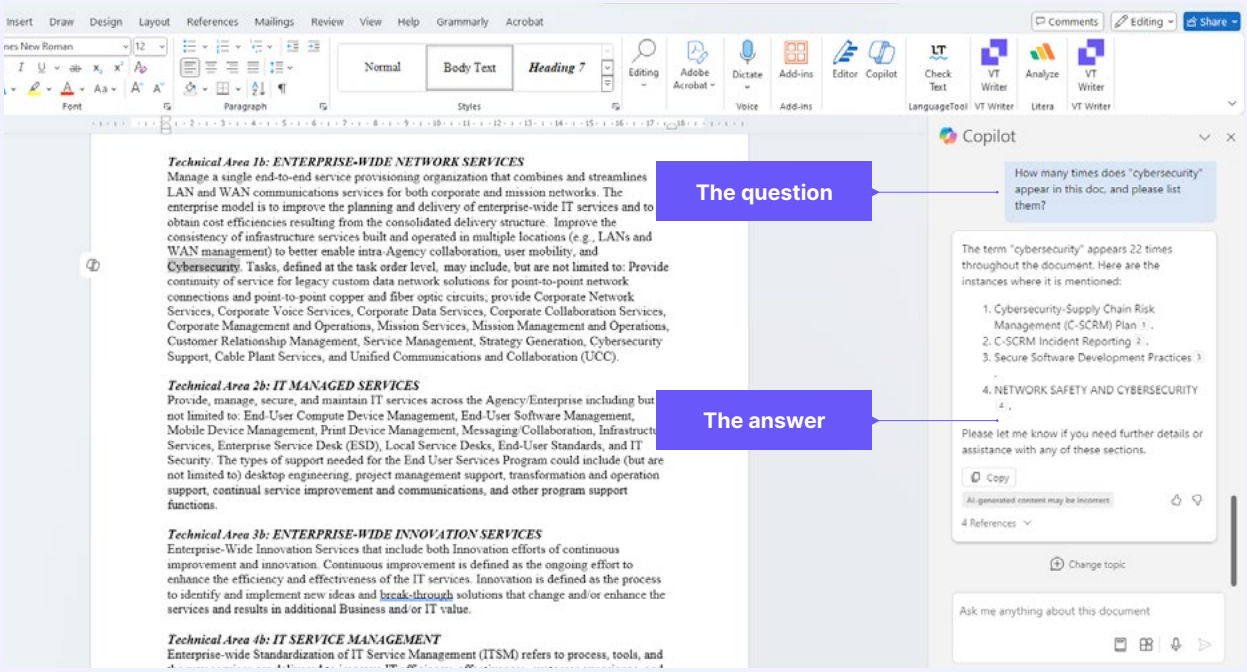
So, searching for cybersecurity, we get:

- exactly 27 hits **(100% accurate)** and
- it will be 27 hits every time we do that search **(100% repeatable)**

This is a “deterministic” operation, meaning it's 100% accurate and 100% repeatable. And MS Word under the covers does not use any AI for this task, rather it uses a logic-based approach to match sequences & patterns of letters. In other words, it uses deterministic software.

Next, let’s use Microsoft’s Copilot (in MS Word) to do the same task. Copilot is an LLM powered addin using Gen AI, the LLM is GPT 4.

Here’s what we get when we ask the question “How many occurrences of cyber security are there in this doc?”:



As you see:

- Copilot has come back showing 4 occurrences of cybersecurity.
- Whereas the actual number of occurrences is 27 as we saw earlier.
- So, we’re missing 23 references.
- and if I try this again, who knows what it'll come back with. Since running the same query twice does not yield the same answers with LLMs. It’s the nature of LLMs.

This is a very simple example where Gen AI is not built to do accurate 100% search. In fact, not only is it not completing the search, it is guessing the answer wrongly, also known as hallucinating.

Gen AI is the **wrong technology** for this type of job.

Keep in mind MS Word and Copilot are both Microsoft products, but using Copilot for accurate search is using Gen AI technology for the wrong task. We just need to use a deterministic approach to complete this type of task. And that’s exactly what Microsoft (and VT Docs) does.

Vendors who tell you that you can use Gen AI for tasks like accurate search have a vested interest in pushing a “Gen AI only” solution. Just be wary of “silver bullet” promises. As the Harvard Business Review said in a column earlier this year: “If you wield a Generative AI hammer, everything becomes a nail.”

What tasks should you use Gen AI for?

Here's a great test you can use to determine whether Gen AI is the right technology for the task you're trying to accomplish.

3 simple questions

1. Do you need 100% accurate results?
2. Do you need 100% repeatable results?
3. Do you need to know the logic by which you got the results?

If you answer "Yes" to any of these three questions, then don't use Gen AI for that task

In the world of proposals, here are the types of tasks you would **not use** Gen AI for:

- User management
- Dictionary search for terms & patterns
- Shredding / Exporting to Excel, MS Word etc
- Comparing docs, Acronym Checks
- Calculating Grade Level (for readability)
- Etc. etc

And here's what you **would use** Gen AI for:

- Text simplification
- Text summarization
- Text creation
- Excellent at creating new content that appears compelling. But accuracy of results depends on the training model.
- Needs to be checked for accuracy, as suffers from hallucinations. Humans must carefully review results.
- Often as part of other workflows and combined with deterministic models

The “Use Case X Technology Matrix" – a useful tool

If we draw back and think about tasks to automate, the basic question is: Do we need 100% reliability and repeatability? If so, use deterministic methods; if not, consider Gen AI. Start with the Job to be Done (JTBD) and map to technology.

As we have worked with customers, we have found the “Use Case X Technology” matrix to be a very helpful tool to make this clear. It helps to educate senior stakeholders when considering Gen AI vs. deterministic models.

Here’s an example matrix:

	Task	Note	Gen AI	Other AI (NLP etc.)	Determin-istic
Goal 1	Develop a requirements/compliance matrix from a Government-issued Request for Proposal.	Must be 100% accurate		●	●
Goal 2	Identify material changes in RFPs, SOW, Contracts etc. (Doc Compare / Excel Compare)	Must be 100% accurate		●	●
Goal 3	Proposal Editing: Single tone-of-voice, Acronym checks, readability checks etc.	Mix: 100% accurate + Good Enough (Gen AI)	●	●	●
Goal 4	Search Past Performance, Resumes, Past Proposal Content etc.	Mix: 100% accurate + Good Enough (Gen AI)	●		●
Goal 5	Generate first draft of proposal sections.	Good Enough (Gen AI)	●		
Goal 6	Analyze contracts for risk (Data rights, Security, IP clauses, FAR/DFARS)	Mix: 100% accurate + Good Enough (Gen AI)		●	●
Goal 7	Qualify New Opportunities. Make Bid / No-bid / Teaming Decisions (Score against Capabilities etc.)	Must be 100% accurate	●	●	●

Here’s how you create the matrix:

- First itemize your business goals, cross-referencing each goal with technology type. In the example above, there are 6 discrete goals.
- The “Note” column helps explain what technology characteristics we require to satisfy the goal. If we need 100% accuracy for example, then we put a bullet under “deterministic” etc.
- For goal 1, since we need a 100% accurate compliance matrix based on shredding a doc, we would not use Gen AI. Instead use a deterministic approach with some NLP elements. To remind, NLP is a different type of AI. On the other hand, if we’re shredding an MS Word doc in VT Docs, then we don’t use any AI at all, since we can use pure logic to get 100% accuracy.
- Now consider goal 4, generating 1st draft proposal content, this is the sweet spot for Gen AI. For instance, creating a 1st draft management plan is perfect for Gen AI. Therefore, we place a bullet against “Gen AI”.
- By mapping from the business goal to the note column, it is simple to justify the technology type (one or more of: Gen AI, Other AI or Deterministic) for that goal.

You get the picture.

We recommend anyone considering building or buying Gen AI technology use a similar evaluation matrix.

RAG - How to leverage your proprietary data with LLMs

Assuming Gen AI makes sense, then a huge part of the value is harnessing your own proprietary data. In this section, we'll explain how you do that.

A technique known as Retrieval Augmented Generation (RAG) allows you tap into your proprietary data when using Gen AI. It doesn't require you to further "train" the LLM, and that's a big advantage.

However, you must have good data hygiene to get good results from Gen AI and RAG.

Simple stuff like:

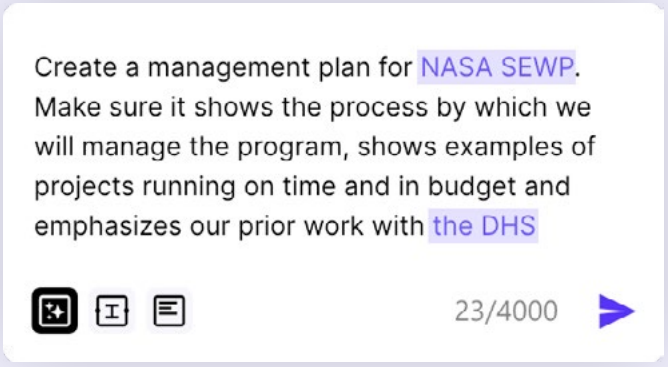
- Setting up clean folder/library structures
- Putting the right permissions on folders/files
- Clear tagging and labeling of files - Make sure you have the "final" version of assets, and not v2, v3 etc. which will contaminate your generated content
- Make sure you have clean "Past Performance" folders etc.

So, let's assume your data is "clean" and well structured; the next question is how to combine that data with Gen AI.

In other words, if I prompt the Large Language Model (LLM) to create a 1st draft management plan for example, I need to ensure it draws on the right information. Information that is not part of the model.

Before we get there, here's what that type of prompt looks like in VisibleThread's VT Writer product, but it'll be similar regardless of what tool you use.

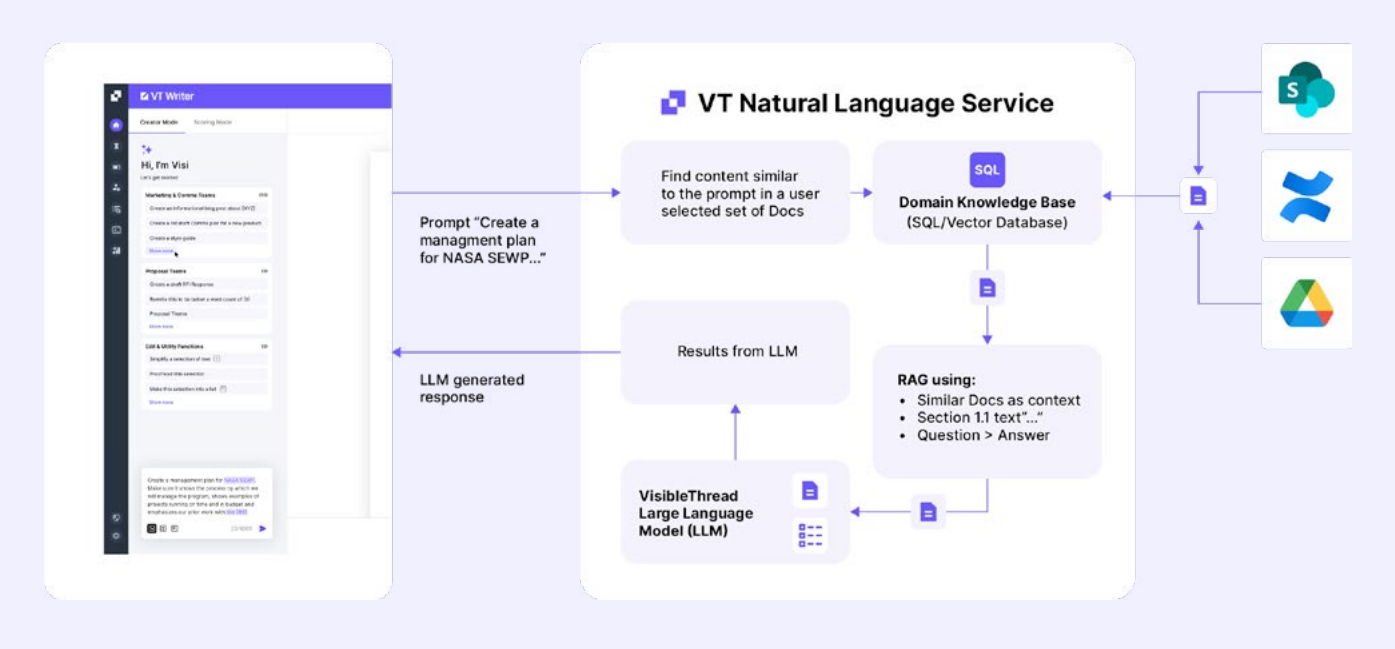
Side note: VT Writer allows you to use templatized prompts with parameters. The purple highlighting indicates the parameters. These are a great way to re-use prompts in different contexts. In my case, with one-click I have a stock prompt for the management plan. Then I just update the purple bits with the RFP name and government agency I want to write to. It's great because I can leverage the best prompts out of the box in 1-click.



The LLM itself has no knowledge of your proprietary data from past proposals.

That content lives in a data repository. In the world of government contractors, that repository is typically SharePoint (SP). People have libraries in SP, where they store winning contracts etc. And that's where RAG comes in.

Here's a graphic showing what happens when you use RAG in conjunction with an LLM.



Here's a quick explainer:

- The user enters a prompt. In our example, it's "Create a management plan..."
- Now, in the background there's a 2nd important thing afoot. Our proprietary data is "indexed" (and stored) into what's called a "vector database". This is a techie thing, but it basically means that your document's content is represented in a way that becomes searchable and findable. In the above graphic, that's the box titled "Domain Knowledge Base".
- Then we combine the prompt with the proprietary data stored in the vector database.
- This then returns answers that are based on your proprietary content.

The big advantage of RAG is

1. You do not have to fine tune models which is expensive and
2. You can cater for new or updated content without having to tweak the LLM.

Now there are quite a few technical considerations on how best to "vectorize" the proprietary content for RAG so that you get best results.

It concerns esoteric topics like "chunking". But that's a whole other topic, and for another day.

The Future of Proposal Management Alongside Gen AI

So, let's look ahead three-to-five years.

The future role of proposal managers will likely involve a more holistic approach, bridging marketing and strategy. Rather than separate divisions, the same individuals may handle both roles, emphasizing 360-degree skills.

> **"The future of our role is going to be much more 360 with respect to our marketing capabilities. You know, traditionally, a lot of firms break out to marcomms team, and they break out the proposal team. It's two separate divisions. I think a lot of the same people will do both roles"**
>
> Uniqueka Walcott**,** *Proposal Manager & Writer at* **AECOM**

Below are just some of the potential paths that AI could take to potentially impact Proposal Management over the next decade.

- **Leadership Shifts:** In the future, we can expect more emphasis on leadership roles. There could be a designated proposal manager for specific accounts or groups. These managers will take charge of the 360-degree communication efforts related to those clients, using AI to help.

- **Increased Capacity:** Accessing GenAI tool systems allows individuals to take on tasks they find more enjoyable. As a result, people have and will continue to handle more development work, creating white papers, etc.

- **Data Analysis and Insight:** Rather than using AI simply for content generation, our experts from APMP can envisage it helping large scale analysis of contract databases. For example, it could create heat maps highlighting promising opportunities.

- **More strategic role positioning:** Using AI, proposal managers may be able to position themselves as more strategic to the organization. Raising their profile and perceived contribution.

"What I have a hard time doing – because I am a single human with one brain – is looking at all of our contracts. Thousands of contracts. I'd like to look at the entire universe of every federal opportunity that might be tangentially related to what we do and picking out the best ones. So, I want to take it and put it into a giant matrix and have it just been this heat map of where our hotspots are. Think of the cool stuff that we can do that we may not even know about."

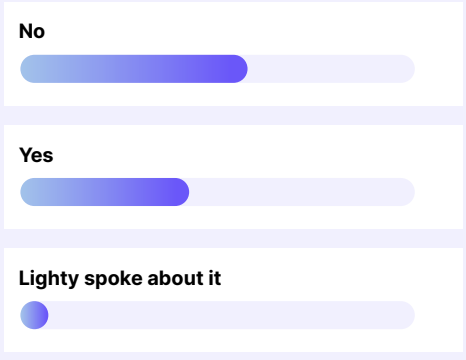
Steve Skeldon, **Chickasaw Nation Industries**

- **Further AI Adoption:** As security concerns related to AI remain a critical focus, companies will implement their own secure GenAI systems. This will facilitate access to appropriate internal content to expedite bid, proposal, and business development activities. It will empower proposal professionals to strategically manage their time in a way that works for them. Using software that is 100% compliant and secure, running fully isolated behind their firewall.

For Proposal and Contracts teams, will the pressure of tight deadlines and stretched resources ease with Gen AI? According to Amanda Haven, Founder at Illuminate Marketing Collective, things have already started to shift in that regard:

"I just got off the phone with a client who's like, we love it. Our KPIs and productivity have gotten so much better since we started implementing AI in our proposal process. Hopefully, depending on each firm and how you choose to implement it, it's only going to help proposal professionals."

Amanda Haven, *Founder at* **Illuminate Marketing Collective**



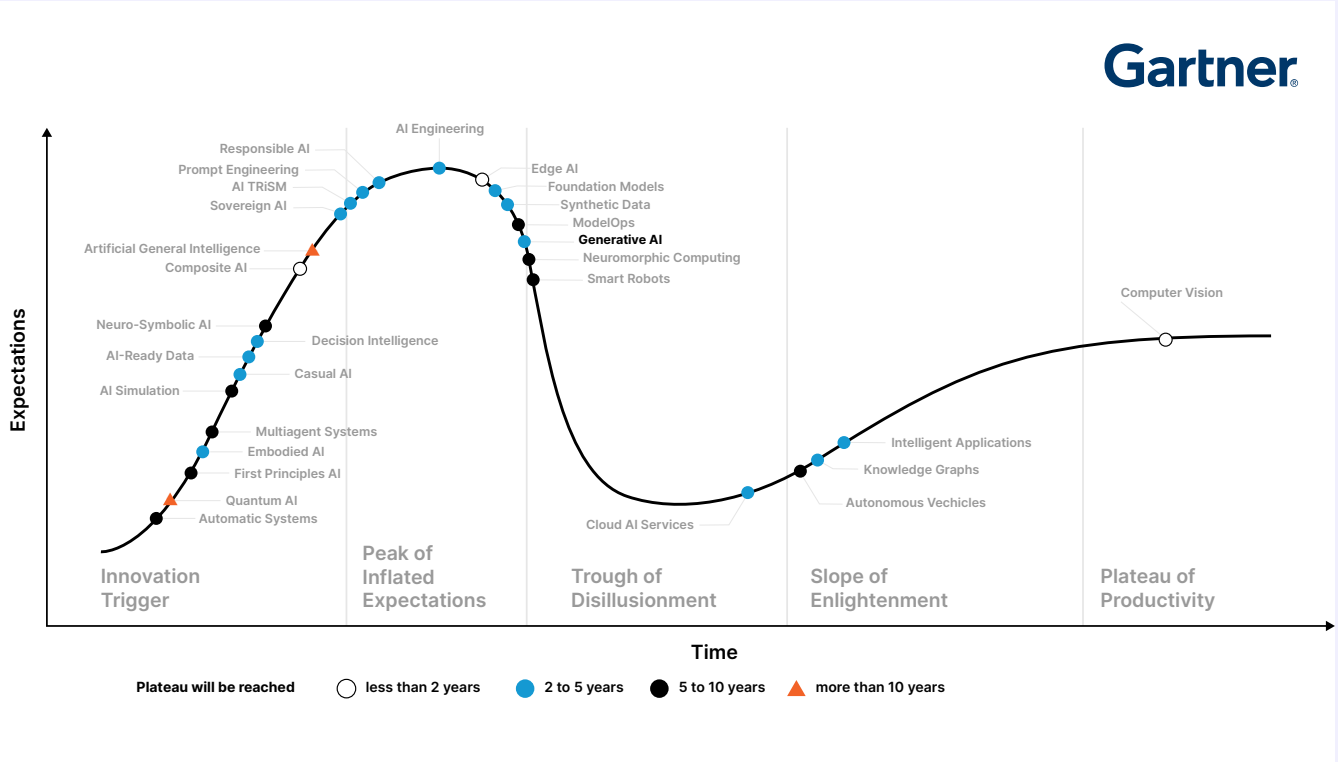
The promise is large with Gen AI. With appropriate guidelines, the productivity gains and quality improvements will accrue to organizations that embrace it.

Where to from here? Understanding Hype Cycles

On June 17th, Gartner released their latest AI Hype Cycle report. It makes for interesting reading.

If you're not familiar with this type of analysis, they characterize every hype cycle as going through phases. The 2nd phase is called "Peak of Inflated Expectations", think about this as peak hype, top of the hype roller-coaster. Every hype cycle goes through these phases. Right from the web in the 90s, mobile in the 2000s, cypto in the 2010s, and now AI / Gen AI.

Here's the current hype cycle for AI.



As noted earlier, AI ≠ Generative AI. GenAI is a specific type of AI. That's why you see a whole bunch of other types of AI at play, ranging from "Neuro symbolic AI" to "Computer Vision". We've bolded where Generative AI sits in this landscape.

You'll see that GenAI is now off peak-hype and is entering the "Trough of Disillusionment" phase. If this were a Star Wars episode, you might think of it as "The Empire Strikes Back" phase for GenAI. Gartner puts timeframes on these things, and for GenAI they suggest it's a 2-5 year journey to get to the "Plateau of Productivity".

In any case, every hype cycle tends to go through the same phases. But having lived through a bunch of these cycles, one thing's for sure, "silver bullets" never feature. They always follow the same pattern.

Here's Gartner's view as to where we're at:

> **"Investment in AI has reached a new high with a focus on generative AI, which, in most cases, has yet to deliver its anticipated business value."**

In a nod to our "Gen AI is not a silver bullet" mantra, here's what they say about the reality of where we'll likely get to.

> **"Generative AI (GenAI) has passed the Peak of Inflated Expectations, although hype about it continues. In 2024, more value will derive from projects based on other AI techniques, either stand-alone or in combination with GenAI, that have standardized processes to aid implementation."**

In essence, they're saying that other types of AI will work in conjunction with GenAI.

And since this is an AI-only hype cycle report, our take is that hybrid systems will be the norm. They will use a combination of various types of AI as well as deterministic software, depending on the JTBD (Job to be Done). Using AI for probability type use cases, and deterministic models when you need logic and must have 100% accuracy and 100% repeatability.

That's exactly how our products here at VisibleThread are. We have layered in Gen AI for certain jobs (mostly create and edit type use cases), alongside deterministic models for others.

Summary

VisibleThread

By following these guidelines and avoiding these pitfalls, organizations can securely integrate AI into their proposal management processes.

Register for our no-obligation live demonstration.

Book a Demo

Need help or have a question?

Alternatively, you can chat to us by mailing info@VisibleThread.com. Just one of the ways you can learn how to use VT Docs in your Proposal Process.