

# For B2B Generative AI Apps, Is Less More?

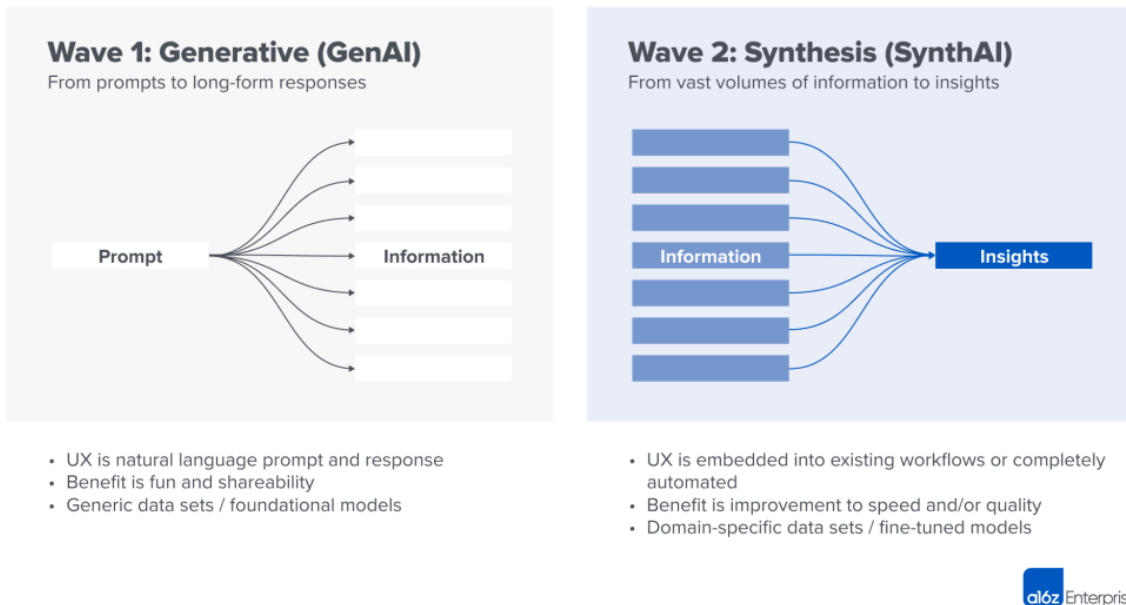
by Zeya Yang and Kristina Shen

We've watched large language models (LLMs) become mainstream over the past few years and have studied the implementations in the context of B2B applications. Despite some enormous technological advances and the presence of LLMs in the general zeitgeist, we believe we're still only in the first wave of generative AI applications for B2B use cases. As companies nail down use cases and seek to build moats around their products, we expect a shift in approach and objectives from the current "Wave 1" to a more focused "Wave 2."

Here's what we mean: To date, generative AI applications have overwhelmingly focused on the *divergence* of information. That is, they create new content based on a set of instructions. **In Wave 2, we believe we will see more applications of AI to converge information.** That is, they will show us *less* content by *synthesizing* the information available. Aptly, we refer to Wave 2 as **synthesis AI** ("SynthAI") to contrast with Wave 1. While Wave 1 has created some value at the application layer, we believe Wave 2 will bring a step function change.

Ultimately, as we explain below, the battle among B2B solutions will be less focused on dazzling AI capabilities, and more focused on how these capabilities will help companies own (or redefine) valuable enterprise workflows.

## The SynthAI Revolution

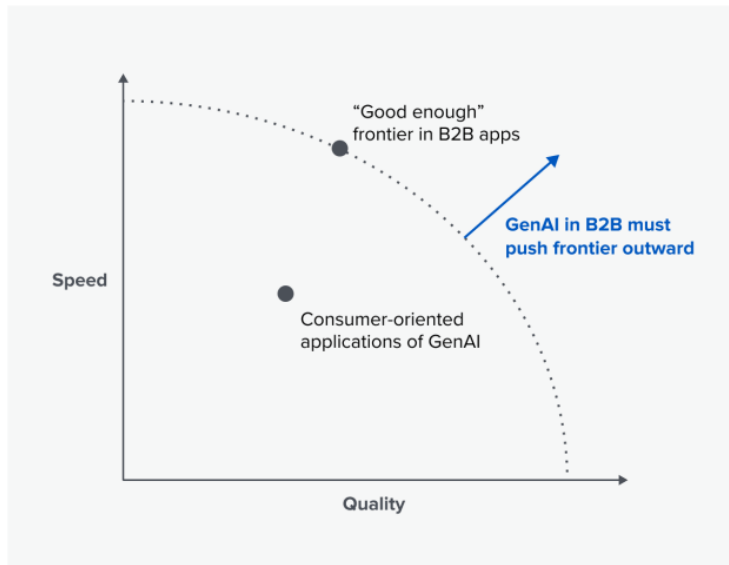


### Wave 1: Crossing the bridge from consumer to enterprise

To analyze Wave 1, it's helpful to first draw the distinction between B2C and B2B applications. When we use generative AI as consumers, our objectives are oriented toward *having fun* and *having something to share*. In this world, quality or correctness are not high priorities: It's fun to have an AI model generate art or music you can share in a Discord channel, before you quickly forget about it. We also have a psychological tendency to believe more = productive = good, and so we are drawn to automated creation. The rise of ChatGPT is a great example of this: we tolerate the shortcomings in quality because having something longer to share is more impressive.

When it comes to B2B applications, the objectives are different. Primarily, there is a cost-benefit assessment around *time* and *quality*. You either want to be able to generate *better* quality with the *same* amount of time, or generate the *same* quality but *faster*. This is where the initial translation from B2C to B2B has broken down.

## Speed and Quality are Critical Factors in B2B



We use B2B applications in workplace settings, where quality matters. However, the content generated by AI today is passable largely for repetitive and low-stakes work. For example, generative AI is good for writing short copy for ads or product descriptions; we have seen many B2B applications demonstrate impressive growth in this area. But we've subsequently seen that generative AI is less reliable for writing opinions or arguments (even when AI-generated content is compelling or confident, it's often inaccurate), which are more valuable when it comes to innovation and collaboration in a B2B setting. A model might be able to generate usable SEO spam, but a blog post announcing a new product for software developers, for example, would require a fair amount of human refinement to ensure it's accurate and that the message will resonate with the target audience.

Another increasingly common example of this is for writing outbound sales emails. Generative AI is useful for a generic, cold outbound email, but less reliable for accurate personalization. From the perspective of a good sales rep, generative AI may help write more emails in less time, but to write emails that increase response rates and ultimately lead to booked meetings (which is what a rep is evaluated on), the rep still needs to do research and use their judgment about what that prospect wants to hear.

In essence, Wave 1 has been successful for more-substantive writing in the brainstorming and drafting stages, but, ultimately, the more creativity and domain expertise are required, the more human refinement is required.

## What's the cost (or benefit) of disrupting the workflow?

Even in cases where generative AI is useful for longer blog posts, the prompt must be precise and prescriptive. That is, before the AI can express them in long form, the authors must already have a clear understanding of the concepts that represent the substance of the blog post. Then, to get to an acceptable end result, the author must review the output, iterate on the prompts, and potentially re-write entire sections.

An extreme example here is using ChatGPT to generate legal documents. While it's possible to do so, the prompt requires a human who is familiar with the law to provide all the required clauses, which ChatGPT can then use to generate a draft of the longer-form document. Consider the analogy of going from term sheets to closing docs. An AI can't perform the negotiation process between the principal parties, but once all the key terms are set, generative AI could write a preliminary draft of the longer closing docs. Still, a trained lawyer needs to review and edit the outputs to get the docs to a final state that the parties can sign.

This is why the cost-benefit assessment breaks down in the B2B context. As knowledge workers, we are evaluating whether it's worth our time to add an additional AI-powered step to our workflows, or if we should just do it ourselves. **Today, with Wave 1 applications, the answer is frequently that we're better off doing it ourselves.**

## Wave 2: Converging information for improved decision making

As we move into the next wave of generative AI applications, we expect to see a shift in focus from the generation of information to the synthesis of information. In knowledge work, there is huge value in decision-making. Employees are paid to make decisions based on imperfect information, and not necessarily the quantity of content generated to execute or explain these decisions. In many cases, longer is not better, it's just longer.

Many axioms support this: lines of code written is not a good measure of engineering productivity; longer product specs do not necessarily provide more clarity on what needs to be built; and longer slide decks don't always provide more insights.

Barry McCardel, CEO and co-founder of [Hex](#), believes in human-computer symbiosis and highlights how LLMs can improve the way we work:

*"AI is here to augment and improve humans, not replace them. When it comes to understanding the world and making decisions, you want humans in the loop."*

*What AI can do is help us apply more of our brainwaves to valuable, creative work, so that we not only spend more hours in a day on the work that matters, but also free ourselves to do our best work.”*

How can AI improve human decision-making? **We believe LLMs will need to focus on synthesis and analysis — SynthAI — that improves the quality and/or speed of decision-making** (remember our B2B diagram above), if not make the actual decision itself. The most obvious application here is to summarize high volumes of information that humans could never digest themselves directly.

The real value of SynthAI in the future will be in helping humans make better decisions, faster. We are envisioning almost the opposite of the ChatGPT user interface: **Instead of writing long-form responses based on a concise prompt, what if we could reverse engineer from massive amounts of data the concise prompt that summarizes it?** We think there’s an opportunity to rethink the UX as one that conveys large amounts of information as efficiently as possible. For example, an AI-powered knowledge base like Mem that holds notes from every meeting in an organization could proactively suggest relevant decisions, projects, or people that someone should reference as they begin a new project, saving them hours (even days) of navigating prior institutional knowledge.

Returning to our outbound sales email example, one potential manifestation is for AI to identify when a target account is at its highest level of intent (based on news reports, earnings calls, talent migration, etc.) and alert the relevant sales rep. The AI model would then, based on the synthesized research, suggest the one or two most important issues to mention in the email, along with the product features most relevant to that target account. Ironically, these inputs could then be fed into a Wave 1 solution, but the value comes from the synthesis phase and saving a sales rep potentially hours of research into just a single prospect.

**A fundamental shift in ensuring this synthesis is sufficiently high quality will be a movement away from large-scale, generic models**, to architectures that leverage multiple models, including more fine-tuned models trained on domain- and use-case-specific data sets. For example, a company building a customer-support application may primarily use a support-centric model that has access to the company’s historical support tickets, but then fall back to GPT for corner cases. To the extent that the fine-tuned models and data sets are proprietary, there’s an opportunity for these components to be moats in the delivery of speed and quality.

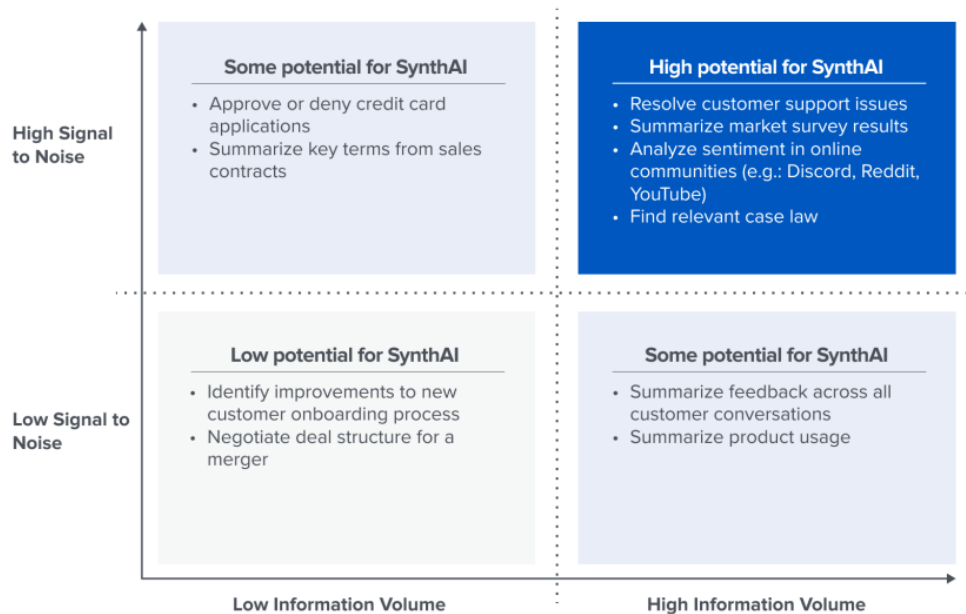
## Implementing SynthAI

As we think through what Wave 2 might look like, we believe the use cases that will benefit most from synthesis AI will be when there is both:

- A high volume of information, such that it's not pragmatic for a human to manually sift through all the information.
- A high signal-to-noise ratio, such that the themes or insights are obvious and consistent. In the name of accuracy, you don't want to task an AI model with deciphering nuance.

In the diagram below, we categorize examples of common analysis and synthesis by these dimensions to help bring this to life.

## SynthAI Effectiveness



This helps us think about the types of outcomes Wave 2 applications will deliver, and how they'll differ from Wave 1 outcomes. Below, we try to offer some examples to bring the comparisons to life, but they are by no means meant to be comprehensive.

## Comparing Wave 1 and Wave 2 Outcomes

Function	Wave 1 - GenAI	Wave 2 - SynthAI
Sales	Generate high volumes of somewhat personalized cold outbound emails	Identify which prospects to send a cold outbound email to, and when, by researching relevant reports Inform how cold outbound email is best personalized based on research on both account and individual levels
Marketing	Generate ad copy	Inform which email campaign a customer should get based on product usage Design ideal cadences for email campaigns based on customer segment and prior success with peer companies
Customer Success	Generate updates on product releases	Understand most important features for a CSM to highlight at the next QBR with a customer that will deepen their usage
Product	Generate specs	Evaluate tradeoffs between competing feature requests by summarizing options, pros/cons, and propose how to weigh criteria based on previous decisions
Legal	Draft legal documents	Summarize most important issues in a contract, as well as any irregular language Suggest research leads or reassure precedent case law doesn't exist when revising a litigation brief



## A battle to own the workflow

Naturally, there is a race between existing systems of record and workflow solutions trying to embed AI-augmented capabilities, and new solutions that are AI-native. We want to be clear what they are racing toward: **the prize is not about who can build the AI synthesis capability; rather, it's who can own the workflow.** For existing solutions, vendors are racing to entrench their existing workflows by improving them with AI. For challengers, vendors will use a best-in-class AI implementation as a wedge and seek to expand from there to redefine the workflow.

On the product feedback use case, Sprig has always used AI to analyze open-text responses and voice responses, and to summarize them into themes. Sprig founder and CEO Ryan Glasgow is excited about the potential for LLMs to improve their synthesis solution:

*“With LLMs, we can save our customers even more time than before. With our prior models, we had a human-in-the-loop review process before customers could see the themes; now, we’re comfortable presenting the themes right away, and doing the review process afterward. Additionally, we’re now able to add a*

*descriptor to each theme to provide more specificity, which makes the insights more actionable.*

*“In the future, we think there’s an opportunity to allow the user to ask follow-up questions if they want to dig further into a theme. At the end of the day, it’s about delivering the end-to-end workflow — from gathering data quickly to understanding it quickly — to help make decisions in real time.”*

At the same time, we’re already seeing new startups exclusively focused on using AI to summarize user feedback, by integrating with existing platforms that are collecting the raw feedback.

On the outbound sales use case, [ZoomInfo](#) recently [announced](#) that they are integrating GPT into their platform and shared a [demo video](#). Certain parts of the video are not far off from the Wave 2 examples we described. Similarly, we’re already seeing new startups exclusively focused on trying to automate as much of the outbound sales process as possible with an AI-first approach.

The potential for how AI may change the way we work is endless, but we are still in the early innings. Generative AI in B2B applications needs to evolve beyond creating more content, to synthesis AI that enables us to do our work better and faster. In B2B applications, it’s a constant dance around who can own the workflow, and AI-native applications will make this dance ever more interesting to watch.