Deep Analysis

Vendor Vignette

Vertesia

Vertesia is an early-stage start-up looking to provide a platform for organizations to build and maintain their own in-house large language model (LLM) and artificial intelligence (AI) applications. While Vertesia is very new, its leaders come from Nuxeo and have a solid track record of selling into technically adept enterprises and understanding their specific concerns.

Founded 2024 | Fully remote | 10 employees (approx.) January 2025

The Company

Vertesia is an early-stage start-up looking to provide a platform for organizations to build and maintain their own in-house large language model (LLM) and artificial intelligence (AI) applications. Founded in 2024 primarily by alumni of content management vendor Nuxeo (acquired by Hyland in April 2021), the company is led by CEO Eric Barroca (replicating his position at Nuxeo when it was acquired). Vertesia is at present a fully remote organization with its team of around 10 people spread across the US, Europe, and Japan. The company has some active pilot projects in progress, which it hopes to convert into its first paying customers during 2024. A $4 million seed funding round was completed in February 2024 led by Elaia Partners and Illuminate Financial, with participation from Motier Ventures, Kima Ventures, m.ai club, Super Capital, and a few unnamed angel investors.

The Technology

Vertesia's central proposition is that even for technically adept organizations looking to develop applications using LLMs, the overheads in managing multiple model developers, model types, and inference providers (which provide access to multiple models as a service) quickly becomes unsustainable. The explosion in models – both commercial and open source – along with the Pandora's box of potential use cases, coupled with disparate technical challenges set against rapid change, inhibits organizations nearing the point where their experiments can be safely considered for production.

Research commissioned by Vertesia suggests that 94% of organizations polled (500 US-based organizations with over 1,000 employees) anticipated utilizing two or more LLMs in the next two years, with 54% believing that number would be over six. 70% of that same group believed themselves to be at best only somewhat prepared to run the multiple LLM projects they anticipated would be required in the next two years.

To address these emerging needs, Vertesia has developed an architecture that allows organizations to develop, test, and deploy LLM-derived applications from a wide range of model developers and inference providers (see Figure 1). Within this, Vertesia provides two primary applications: Studio and Platform.

- **Studio** is concerned with the development and testing of applications using the multi-LLM setup that Vertesia provides.
- **Platform** contains the engineering-level setup that enables those connections to the model/inference providers, as well as access to add-on functions such as its own retrieval augmented generation (RAG) setup, caching, and search services called "RAG Services."

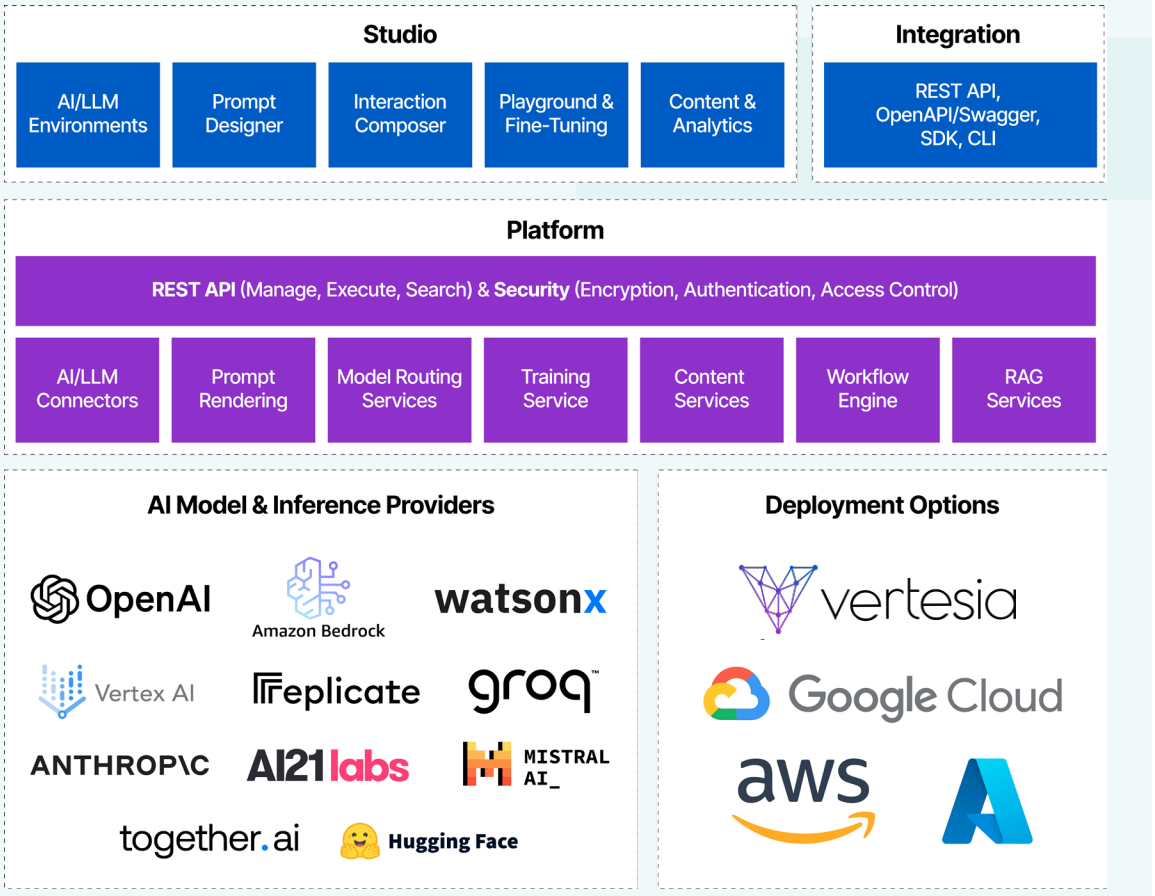
It's Studio that application developers will need to become quickly familiar with, as it is where most of their time will be spent. Within Studio, prompts can be developed and tested against a variety of models, with the application largely taking care of transposing the model-specific syntaxes. This means that prompts with the same intent can then be tested across a variety of models, not only to ensure that the right responses are generated, but also to perform on-the-fly testing for model suitability for that specific task.

From those outputs, developers can develop

prompts and perform fine-tuning on the models themselves. This combination of functionality enables developers to manage prompts centrally, without having to worry about which model will ultimately process them, as the configuration for each individual project managed within Studio contains the logic for model use. This model use also allows for targeted splits across models, so that loads can be tested across prompts on a specified variety and the results analyzed. This also makes it possible to create synthesized models ("Synthetic LLMs" in Vertesia terms) which can be built from different models and are executed depending on developer choices, potentially as a result of that prior split testing (using input and output data schemas as required).

Another important area that Studio can add to application planning is estimating the token use of each prompt across the range of possible models. This is an important emerging challenge for organizations planning to utilize LLMs in production, especially where multiple models for multiple providers might have different ways of calculating the cost of each prompt. Of note here is that Vertesia works on a bring your own model (BYOM) basis for model licensing; it does not provide bundled model access as part of its platform, so customers will have to enter into their own agreements with model developers and inference providers.

Figure 1
Vertesia Architecture



At present, the Vertesia platform is set up to utilize models and inference services from OpenAI, Amazon Bedrock, watsonX, Google Vertex, Groq, Together AI, Replicate, Hugging Face, and Mistral AI. It offers deployment of the platform to AWS, Azure, Google Cloud, and its own SaaS environment.

As those familiar with the founders' previous enterprise, Nuxeo, will attest, Vertesia is primarily built to serve technically savvy organizations; predominantly large enterprises already familiar with LLMs and the challenges in developing applications using them. The intent is for it to become a tool that is deployed for strategic use of LLM-derived applications once an organization's tactical experiments have been exhausted. Given that, Vertesia expects technical buyers, but where the financial sponsor often will be the line-of-business that is commissioning the ultimate application.

Initially, Vertesia is managing its early pilots and subsequent sales directly, in-house. However, it sees that to scale up in the short term, a combination of reseller agreements and partner relationships – especially with the large systems integrators – will be required, and both are high on its current go-to-market to-do list.

Our Opinion

Vertesia has developed a platform that is designed to provide a strategic response for large enterprises looking to rapidly build, evaluate, and deploy LLM-based tasks with enterprise-level standards and controls. It's anticipating a near future in which enterprise IT departments will both want and need to utilize multiple models from multiple suppliers in order to manage the requirements of their line-of-business customers. As such, allowing the separation of the management of prompts and model-specific syntax is sensible, as are the capabilities for mixed, synthetic LLMs for testing and potentially for redundancy planning in production. Vertesia is likely to find competition sparse only in the very short term, as the API management vendors are unlikely to want to cede any control they've gained over IT service coordination and likely alternative management panels from the inference providers themselves. Vertesia leadership's experience in winning the mind share of technically advanced organizations is likely to put it in a good initial position if it can continue to execute and turn those pilots into active, paying customers.

Advice to Buyers

Organizations that either find themselves having to manage the complexity of multiple LLMs or anticipate that they will soon be in that position should investigate whether Vertesia offers them the right tools to gain control over their situation. Where their primary issues are managing prompt creation and model tuning across a variety of models and inference providers, performing split testing, and developing synthetic LLMs approaches, Vertesia is likely to be a platform that should be evaluated. If managing the commercial relationships (i.e., licensing and token bundling/payment) for LLMs is the primary concern, Vertesia likely won't be a complete solution just yet, for the time being at least.

SOAR Analysis

Strengths

- Vertesia's leadership has a solid track record of selling into technically adept enterprises and understanding their specific concerns

Aspirations

- To be the go-to platform for enterprises developing and employing LLM-driven generative AI applications across multiple providers and models

Opportunities

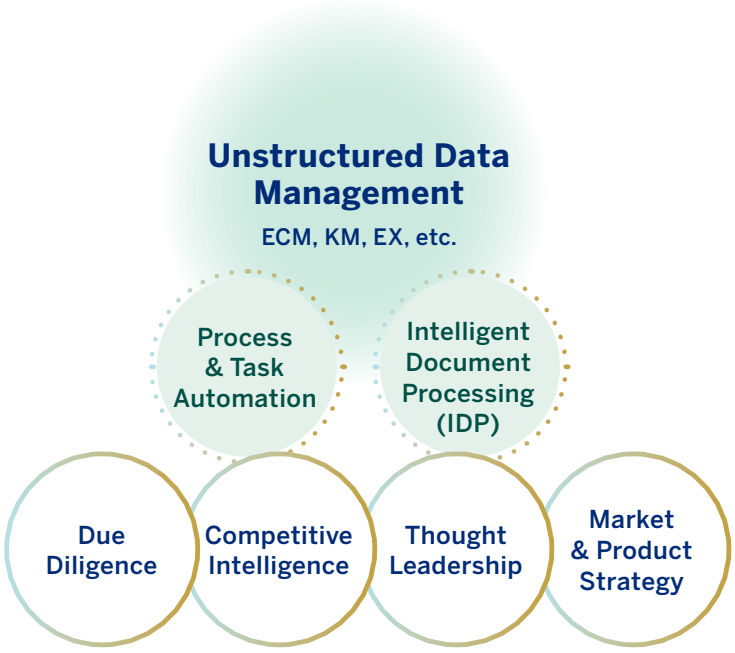
- In the short term, the proliferation of models and providers isn't going away; if Vertesia is able over time to add tighter integration on the commercial/financial metering side of the LLM management puzzle, it will get closer to a complete enterprise solution

Results

- Raising a strong seed round prior to landing any paying customers suggests that the company has a solid, defensible business plan for growth

About Deep Analysis

We Research Innovation



Our Services

Deep Analysis is an advisory firm that helps technology vendors, buyers, and investors understand and address the challenges of innovative and disruptive technologies in the enterprise software marketplace.

The firm's work is built on decades of experience advising and consulting to global technology firms large and small, from SAP, Oracle, and HP to countless start-ups.

Led by Alan Pelz-Sharpe, Deep Analysis works with technology vendors, buyers of enterprise technology, and investors in the ECM and enterprise automation market to improve their understanding of the information management technology landscape and provide actionable guidance.

Deep Analysis' timely book, "Practical Artificial Intelligence: An Enterprise Playbook," outlines strategies for organizations to avoid pitfalls and successfully deploy AI.

Contact us:

info@deep-analysis.net
+1 978 877 7915