

October 2023

The Value and Future of AI Language Models in the Global Insurance Industry

John Keddy and Mitch Wein



Prepared for:



The Value and Future of AI Language Models in the Global Insurance Industry



John Keddy and Mitch Wein

Table of Contents

Introduction	2
Methodology	2
Key Concepts and Definitions	3
Paths for Insurer Language Model Utilization	6
Decision Table	9
Conclusion	11

List of Tables

Table A: Language Model Decision Paths.....	9
---	---

Introduction

Throughout 2023, the conversation around artificial intelligence (AI) has permeated all aspects of our society. Insurers have not been immune from these discussions. Datos Insights' direct experience is that the understanding of AI topics among C-level insurer executives has matured dramatically in recent months. Insurers are increasingly trying to get more specifics on how AI will impact the broader insurance industry and their specific lines of business.

Large language models (LLMs) have been a source of particular interest and confusion. In conversations with Datos Insights, insurance technology executives express a spectrum of views on these tools, ranging from questioning whether an insurer should build LLMs to questioning if LLM capabilities have any value whatsoever.

This report, sponsored by DXC and written by Datos Insights, provides basic definitions of various language models and explores how insurers will most likely utilize LLMs. Additionally, the report presents a decision table outlining factors that insurers can consider as they contemplate various options to determine reasonable courses of action while avoiding expensive and unrealistic choices that do not create value.

Methodology

This report is based on practitioner experience, solution provider engagement, ongoing Datos Insights research on AI, and the authors' experience and engagement with AI.

Key Concepts and Definitions

A few short months ago, LLMs were not a topic of widespread conversation among insurers. Insurers' understanding of these tools has expanded dramatically over the course of 2023. Due to the amount of conversation and occasional confusion, it is useful to define several key terms upfront:

- **LLM:** "Large" here means models with billions of parameters. The largest models are now moving to trillions of parameters. Parameters are the factors that models use to generate predictions. Insurers should also take "large" to imply "large scale" in terms of computing power, data volumes, costs, and intellectual capital of data scientists. All of these elements are necessary to build and test an LLM. The second component of a basic definition is that LLMs have a connection with language. Here, language can extend to computer code, images, and audio. Examples of LLMs are numerous, but some of the most commonly discussed include Anthropic's Claude, Hugging Face's BLOOM, Meta's LLaMA, and OpenAI's GPT.
- **Foundational models:** Stanford defined the term "foundational model" in August 2021 as "any model that is trained on broad data ... that can be adapted (e.g., fine-tuned) to a wide range of downstream tasks." These models can be considered foundational to technologies such as LLMs but can also be viewed as a broader technology approach. For example, a foundational model might be used to train for capabilities that will require computer vision. Somewhat confusingly, OpenAI GPT versions are sometimes viewed and discussed both as foundational models and as LLMs. A detailed understanding of the intricacies of LLMs versus foundational models is not required to leverage the insights from this paper. Examples of foundational models include Google's BERT and AWS' Titan.
- **Generative AI:** Generative AI produces assets that did not previously exist by relying on statistical prediction of the next word, pixel, or note. Those predictions are based on extensive model training that allows for a more complex version of end-user applications that simply predict the next word a user might type when composing a text message or email. The content, or corpus, that serves as the basis for training is critical to the capabilities of the generative AI product, which includes language, computer code, images, synthesized voices, music, audio, video, and other assets. Generative AI can invoke search tools, such as Google, but generative output is predictive. It may provide wrong answers based on a bad prediction, which is why these tools should be understood less as "truth-telling" than "really good storytelling."

The speed and common utilization of ChatGPT have impressed many with generative AI's capabilities but have also caused a lot of confusion as people conflate generative AI with other aspects of AI. There are many subdomains within AI, including computer vision, analytical AI, and machine learning. It is critical to remember that generative AI is just one component of AI. In addition to ChatGPT, commonly used generative AI tools include DALL-E-2, Jasper, and GitHub Copilot.

- **Natural language processing (NLP):** NLP capabilities enable intuitive human interaction with complex AI tools. NLP tools allow input questions to be more open-ended and output dialogue to sound more natural. True NLP differs from tools like chatbots and interactive voice response (IVR), both widely used in insurance, which are not inherently natural; they give predictable answers to a predictable or known set of questions. NLP capabilities are tremendously powerful and drive much of the hype around generative AI (e.g., "ask ChatGPT to write a haiku"), but they are not the right solution for every use case. For example, if a consumer requests an account value, a chatbot or IVR with a fixed path to the correct answer is the appropriate solution. Chatbots will be enhanced with generative AI solutions to make the engagement more natural, but generative AI shouldn't statistically "predict" an account value. There is one factual answer to an account's current value, and the implementation must support that reality.
- **Open-source AI models:** It is well known that OpenAI's investment in developing ChatGPT was massive. Google, AWS, Microsoft, and many other for-profit entities have made similar investments in generative AI, and these entities have to monetize those investments. While ChatGPT has dominated the AI conversation in 2023, insurers shouldn't overly fixate on OpenAI. Instead, they should understand that there are many other players in generative AI, and open-source LLMs may be more cost-effective (though, notably, not "free") for some insurers. Meta, for example, has released an LLM, leading developers to report LLM success with an open-source approach. Subsequently, on June 18, 2023, Microsoft and Meta announced a go-to-market strategy. Examples of LLM that are available on an open-source basis include Google PaLM 2, Meta LLaMA, Falcon, and OpenLLaMA. The introduction of open-source, or near open-source, is an important consideration for insurers as costs and skill set availability will be key challenges.
- **Small language models (SLMs):** SLMs are similar to LLMs, but they are smaller in size, perhaps half the size of LLMs, and may be more cost-effective. At the time of writing, LLMs are more powerful than SLMs. For many insurers and real-life use cases, however,

the best solution may be an SLM as opposed to an LLM due to cost considerations, the breadth of the model's coverage, and the time it takes to train or retrain the model. Insurers should monitor SLMs closely as current models progress and more SLMs are introduced. Their smallness may have benefits in terms of costs, scale, time invested, and ability to customize training to specific domains. LLMs can be used to train SLMs, and some SLMs are open-source. Examples of SLMs include Chinchilla and Gopher.

Paths for Insurer Language Model Utilization

There are five major language model paths that insurers could pursue for the balance of 2023 and into 2024.

- 1. Build your own LLM:** This approach involves an insurer deciding that the LLM work done to date is insufficient—or, more likely, wanting to develop and control their own LLM. Though it is always dangerous to speak in absolutes, Insurers building LLMs will be rare. While exact data is very difficult to obtain, credible cost estimates for building a true LLM start in the tens of millions of dollars and go higher, a prohibitive expense profile for most insurers and reinsurers. The most obvious example in Financial Services of an attempt to build an LLM may be BloombergGPT. Bloomberg had over 100 years of existing data and deep domain knowledge. It partnered with third parties to refine and implement the model. Costs were not publicly disclosed but are thought to be significant. For most insurers, this will not be possible, and the value of developing a unique LLM won't outweigh the cost.
 - **Value:** Insurers considering this path are likely global insurers or reinsurers with board buy-in and a unique opportunity where the billions of dollars in past LLM investments cannot be leveraged. This is likely under 1% of insurers.
 - **Future:** At the time of writing, various SLMs are posting impressive results. The era of large-scale investments in LLMs, where size is viewed as the most desirable factor, may be ending. As opposed to many new LLMs being developed, we may see increasing leverage of existing LLMs. Some of the largest leverage points may prove to be training SLMs. Leveraging weights from LLaMA, Stanford built a model called Alpaca for US\$600 vs. the estimated US\$100 million for the full GPT-4 LLM. Moreover, biases and flaws that result in incorrect answers will be more quickly identified in SLMs. There are legal issues that will have to be addressed, depending on which LLM an SLM is trained on, but to ignore the potential impact is poor judgment. Insurers that believe they should not ignore all of these factors and build an LLM from scratch will be incredibly rare.
- 2. Fine-tune LLMs:** This option involves taking an LLM and refining the model with partners to create an offshoot model focused on a specific area. Like a college major builds upon past educational experiences, fine-tuning an LLM creates an area of focus.

In insurance, we will see fine-tuning, but it is unlikely that each insurer will directly fine-tune an LLM, as we will discuss in future options.

- **Value:** The overall cost, time, and skill set involved in fine-tuning an LLM will be significantly lower than building a new LLM. However, the skill sets and dollars required still represent a sizable commitment. For other than the largest insurers and reinsurers, costs and skill sets will prove to be barriers.
- **Future:** More insurers are likely to fine-tune an LLM than build a new model from the ground up (perhaps 5% of insurers).

3. Leverage language model capabilities from a proven insurance solution provider: In this approach, the insurer avoids the intricacies of building or directly fine-tuning LLMs by leveraging the work of proven solution providers. The solution provider will have implemented language models into its own tooling or have stand-alone language model products optimized for insurance. Insurers will bear licensing and any implementation costs. Insurers will fine-tune the model with their own documents, data, and business rules but will be guided by a solution provider who has productized and machined the approach. In this model, insurers can't outsource ownership or view current solutions as mature, off-the-shelf products; generative AI is simply too embryonic. Insurers will still need to review, test, and be highly involved during the implementation of this approach. Commonly, in these models, insurer-specific data may be used to fine-tune a vendor-provided product. Insurer-provided data may be augmented with data from a third party. For example, a solution provider might offer a generic customer support tool trained on insurance terms, and an insurer may require specific fine-tuning with their internal data, due perhaps to unique products or customer profiles. In a model like this, where the language model tool is calibrated to the insurer's needs, implementation and total cost of ownership will be higher. Insurers will also need to validate all responses and make sure the model is retrained as needed.

- **Value:** This category is likely to bring greater value to insurers due to faster implementation times and lower investment. Because insurers can access their vendor partner's data science expertise, the insurer is freed to focus on developing a deep understanding of their specific data and how best to use the vendor's solution within the insurer's production environment.
- **Future:** This is likely to be a common path. Insurers should be pressing their solution providers now for clarity on the vendor's roadmap and investments, as well

as how the insurer can leverage those investments. Still, insurers should cultivate an understanding of LLMs and other generative AI tools, as being an educated consumer will prevent misunderstandings and expensive missteps. We expect this to be a very common solution scenario.

- 4. Leverage insuretech/innovators for LLMs or SLMs:** This path will be similar to working with large established insurance vendors but with two caveats that insurers should note. The first is innovators' lack of broad domain knowledge. Typically, startups and innovators focus on a narrow portion of the insurance value chain; solutions may be narrower in scope, point or niche solutions rather than broad tools. The second caveat is that innovators may not have the resources required to invest in computing expenses, storage expenses, or other resources at a large scale. Obviously, cloud capabilities offer more cost-effective solutions, but to assume that "cloud is cheap" creates dangerous blind spots. Neither caveat is a barrier that can't be overcome, but when partnering with innovators, insurers may have to take on more in terms of dollar investment and time. Nonetheless, this path will be used by many insurers due to the skill set and abilities that many innovators and insuretechs possess.
 - **Value:** Solution provider costs may be lower in this category, but insurer time investment may be higher. For an insurer to pursue this option, the innovator should demonstrate data science excellence that the insurer can complement with the insurer's deep domain expertise. In the right situation, the value proposition could be strong.
 - **Future:** This is likely to be a common path, and many insurers are actively engaged with startups on generative AI proof of concepts. As with working with large providers, the insurer must be an educated consumer but may also have to be ready to fill in any expertise gaps and determine how the solution fits holistically within the insurer's environment.
- 5. Do nothing:** Despite the current euphoria around AI, many insurers are not engaging with the technology. Some may be focusing on existing planned initiatives; some may believe AI is overhyped; some may lack clarity on where to start or have other logical reasons for staying on the sidelines. By the end of 2023, perhaps as many as 30% of insurers have not done much exploration in this domain at all. However, insurers who are not investing any effort or mindshare into these topics do risk placing themselves at a competitive disadvantage. While most insurers are not likely to be building their own

data science foundation, understanding the concepts and options around AI will be necessary for insurers to leverage, invest in, or effectively work with external entities.

In summary, Datos Insights expects fewer than 1% of insurers to build LLMs and around 5% of insurers to fine-tune an LLM. Most insurers will leverage a solution provider’s AI capabilities or partner with insurerechs. Some will choose to do nothing.

Decision Table

Table A summarizes the five paths for insurers looking to engage with language models discussed in this paper. The columns capture four different strategic parameters: budget, insurer data science expertise, insurance domain knowledge, and competitive advantage.

Table A: Language Model Decision Paths

Path	Budget	Data science expertise	Domain knowledge	Competitive advantage
Build your own LLM	Tens of millions of dollars and up	Expert level	Deep	Seen as unique and current LLM cannot be leveraged
Fine-tune LLM	Very large	Very deep	Deep	Seen as unique
Leverage LM capabilities from proven provider	Moderate; serious consideration for most insurers	Some; not very deep	Moderate	Utilization of capabilities of greater value than unique LLM
Leverage insurerech/ innovator	Limited; serious consideration for most insurers; innovators may take more time and budget	Some; not very deep	Very deep	Utilization of capabilities and partnership of greater value than unique LLM
Do nothing	Zero cost today but risk falling behind	N/A	N/A	No strategic benefits for any engagement

Source: Datos Insights

Insurers should start with their budget. Cost is a significant matter that is sometimes overlooked. Although public data is not perfect, there is general agreement that LLM development is very expensive. SLMs may mitigate this. All AI endeavors must have consideration for the financial resources an insurer has to invest.

Available data science expertise will determine which options an insurer can effectively execute. The range of data science expertise in insurance ranges tremendously. Some very large insurers have invested significantly in data science teams with expertise that rivals organizations in any vertical. Other companies are far more limited, and their data investments to date may be more centered on operational reporting, dashboards, and other tactical investments. As insurers consider these options, they must be candid about the current state of their data science capabilities.

Insurance domain knowledge is also a key decision parameter. Of course, insurers commonly (though not universally) have deep domain knowledge about their own data. However, whether those resources can be freed up for significant time investments is an issue. Even if the resources are available, AI requires new skills and approaches, and resources with deep domain knowledge may not be able to learn the new skillset.

Conclusion

Insurers of all sizes are talking about language models—and they should be. Insurance leaders must understand these concepts to set technology direction, participate in industry conversations, understand regulatory impacts, and engage across their organizations, including with board members.

Most insurers, but not all, are likely to leverage the power of language models through partners. Those partners may implement capabilities via LLMs or SLMs. While OpenAI's ChatGPT has dominated the AI conversation in 2023, insurers should understand that there are many other players in generative AI. Meta, for example, has released an LLM, leading developers to report LLM success at open-source prices; subsequently, on June 18, 2023, Microsoft and Meta announced a go-to-market strategy.

Insurers should continue to closely monitor developments in generative AI, especially in SLMs, as strategic decisions that hinge on cost and time spent training LLMs may be evaluated differently as SLM capabilities improve and become more widely available.

About Datos Insights

Datos Insights is an advisory firm providing mission-critical insights on technology, regulations, strategy, and operations to hundreds of banks, insurers, payments providers, and investment firms—as well as the technology and service providers that support them. Comprising former senior technology, strategy, and operations executives as well as experienced researchers and consultants, our experts provide actionable advice to our client base, leveraging deep insights developed via our extensive network of clients and other industry contacts.

Contact

Research, consulting, and events:

sales@datos-insights.com

Press inquiries:

pr@datos-insights.com

All other inquiries:

info@datos-insights.com

Global headquarters:

6 Liberty Square #2779
Boston, MA 02109

www.datos-insights.com

Author information

John Keddy

jkeddy@datos-insights.com

Mitch Wein

mwein@datos-insights.com

Contributing author:

Caitlin Simmons

csimmons@datos-insights.com

© 2023 Datos Insights or its affiliates. All rights reserved. This publication may not be reproduced or distributed in any form without Datos Insights' prior written permission. It consists of information collected by and the opinions of Datos Insights' research organization, which should not be construed as statements of fact. While we endeavor to provide the most accurate information, Datos Insights' recommendations are advisory only, and we disclaim all warranties as to the accuracy, completeness, adequacy, or fitness of such information. Datos Insights does not provide legal or investment advice, and its research should not be construed or used as such. Your access and use of this publication are further governed by Datos Insights' Terms of Use.