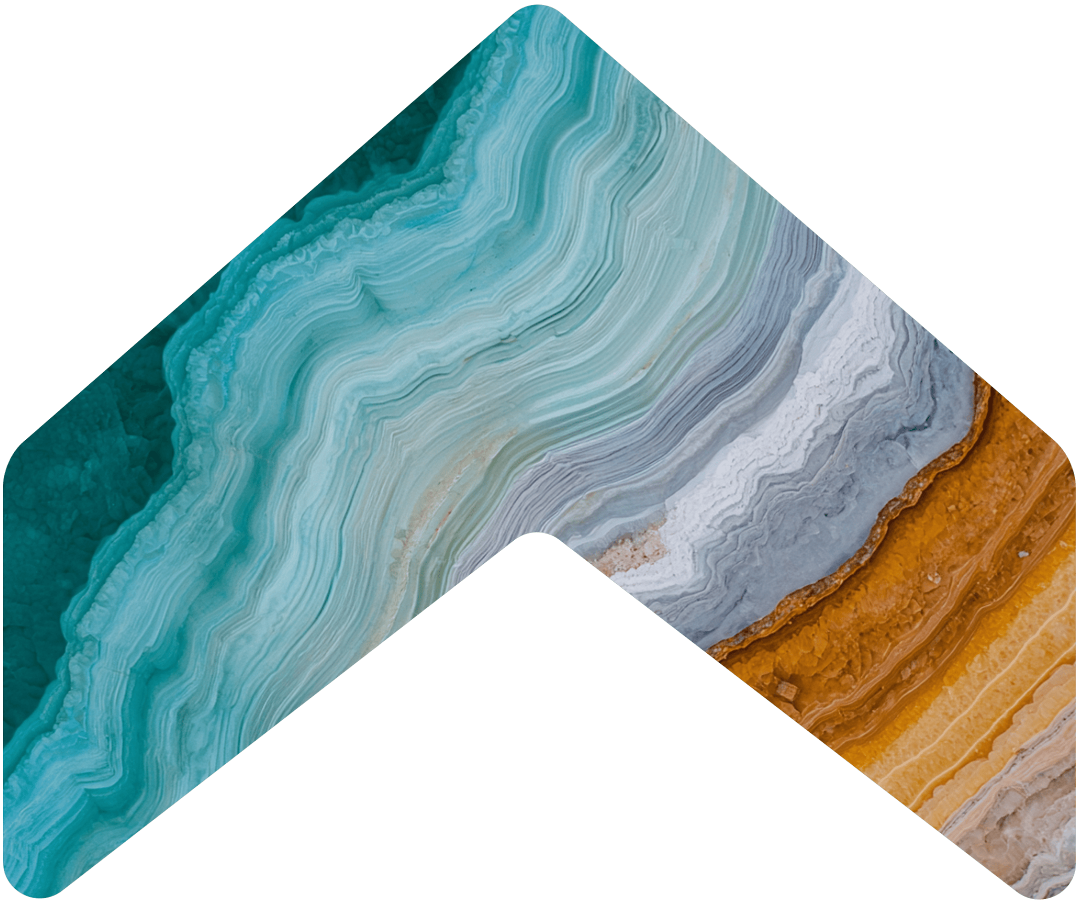

Evolution of Power and Energy Insurance Claims Handling in the Age of Artificial Intelligence

Andrew Bott | Audrey Wong

WOTTON KEARNEY

1. Introduction

Artificial Intelligence (AI) has become a popular buzzword in recent years. The AI technology which has attracted the most attention in recent years is generative AI ("**GenAI**"), specifically Large Language Models ("**LLMs**") like ChatGPT. The ability of LLMs to generate human-like text and content make them very useful across a wide range of applications, with many companies looking to integrate them into various products and services.

In this paper, we explore the potential use and impact of AI in insurers' management of losses in the power and energy sector. We begin by considering how LLMs work and their limitations and then discuss the current LLM use cases for insurers and insureds, finally discussing the applicability of common cyber exclusions in the context of the use of LLMs.

2. How do LLMs work?

Before we consider the impact of LLMs on power and energy claims, we first need to understand conceptually how LLMs are designed and operate.

When a user inputs text into an LLM, it is first broken down into tokens. Each token is a numerical representation of the meaning of words, or parts of words: the human equivalent of the letters of the alphabet.[1] These tokens are then processed through a multilayered neural network, or in other words, a complicated mathematical equation. This equation is trained through trial and error to recognise patterns and predict the most probable response to input text based on these patterns. The predicted response is then 'translated' back into a human-readable output text.

The neural network technology behind LLMs is not new and has been in development since the 1940s.[2] Recent advancements, particularly the availability of massive amounts of data, and improvements in computing power, have made it possible to build much larger and more complex models capable of handling sophisticated language tasks.[3] It is these advancements, and their potential, that have brought LLMs back to the attention of the public in recent years.

3. Limitations of LLMs

As described above, LLMs understand language in a fundamentally different way from humans. LLMs operate purely on mathematical representations of language, and generate responses based on statistical predictions, using the data they are trained on.[4] While LLMs can mimic human understanding, it does not grasp meaning and nuances the way humans do. As such, LLMs can sometimes get things very wrong and not realise or correct the error.

LLMs also do not "learn" in the way humans do. Once the model is trained, it is fixed in time and does not adapt to new user interactions.[5] This means that unless the model is re-trained and fine-tuned with new data, it will repeat the same mistake again and again.

---

[1] See Lee, T.B. and Trott, S. (2023) *Large language models, explained with a minimum of math and jargon*. Available at: https://www.understandingai.org/p/large-language-models-explained-with (Accessed: 16 April 2025).
[2] Hardesty, L. (2017) *Explained: Neural networks*, *Massachusetts Institute of Technology*. Available at: https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414 (Accessed: 16 April 2025).
[3] See fn 2.
[4] See fn 1.
[5] The LLM's knowledge will need to be fine-tuned and refined post-training in order to improve performance: see Sanderink, U. (2025) *LLM post-training: A deep dive into reasoning large language models*, *Medium*. Available at: https://medium.com/@sanderink.ursina/llm-post-training-a-deep-dive-into-reasoning-large-language-models-b910786275b5 (Accessed: 16 April 2025).

As such, the quantity and quality of the training datasets used to train the LLM is crucial, and directly affects the accuracy and reliability of the response generated by the model. Using high quality and comprehensive datasets will enable the LLM to perform more reliably and ethically, at least according to human standards.

Another point to consider is that LLMs use billions[6] of parameters to calculate a response. The complexity of this internal process makes it challenging to trace the exact reasoning followed by an LLM to produce its response. Indeed, it is not currently possible to fully understand the internal logic followed by LLMs, and how or why a particular response is generated.[7] As such, it may be difficult to know whether the LLM has considered all the relevant factors in its response, and it remains unclear if any actual reasoning is occurring within an LLM.

These limitations underscore the importance of providing people with the opportunity to intervene before AI-generated responses are actioned, especially in novel or ambiguous situations.

4. Use of LLMs in Insurers' management of Power & Energy claims

The ability of current LLMs to quickly extract and analyse large volumes of data can, in theory, be particularly useful in the management of insurance claims.

Indeed, some insurers have reportedly been able to successfully leverage LLMs to achieve no touch claims handling in personal lines of business, such as motor and homeowners insurance, wherein the entire process from claim submission to payment is automated and can be completed within seconds.[8]

The effectiveness of LLMs in these relatively high-volume low-value claims environment seems obvious. The range of possible claim scenarios and claims documents to be considered are relatively uniform and predictable. The policy terms are also largely consistent across the insurer's portfolio. It is believed that the use of LLMs would result in improved claims handling and fraud detection and lower claims cost, albeit there is limited publicly available information on the true effectiveness of these AI systems (such as the error rate in the claims handled by the AI system, and the extent to which claims leakage is reduced).[9]

However, the current limitations of LLMs means that they are likely to play a more limited role in power and energy claims, such as summarising reports, providing general information, extrapolating claim insights and patterns from claim records. No or low touch claims handling is unlikely to be achievable nor practical due to the relatively low volume and bespoke nature of these types of claims.

There is a scarcity of comprehensive high-quality datasets available for each insurer to independently train LLMs capable of automating power and energy claim handling. As we noted in the above sections,

---

[6] Alarcon, N. (2020) *OpenAI presents GPT-3, a 175 billion parameters language model*, *OpenAI Presents GPT-3, a 175 Billion Parameters Language Model*. Available at: https://developer.nvidia.com/blog/openai-presents-gpt-3-a-175-billion-parameters-language-model/ (Accessed: 16 April 2025).
[7] Hazra, R. *et al.* (2024) *Can large language models reason? A characterization via 3-sat*, *arXiv.org*. Available at: https://arxiv.org/abs/2408.07215 (Accessed: 16 April 2025).
[8] Willard, J. (2023) *Lemonade shatters record by using AI to settle a claim in two seconds - reinsurance news*, *Lemonade shatters record by using AI to settle a claim in two seconds*. Available at: https://www.reinsurancene.ws/lemonade-shatters-record-by-using-ai-to-settle-a-claim-in-two-seconds/?gad_source=1&gclid=EAIaIQobChMIhsuQicWYjAMVay2DAx0L5jKPEAAYASAAEgIfmPD_BwE (Accessed: 20 March 2025); *Ping An takes lead in Fortune Fintech Innovator Asia list, ranked 1st in Insurtech* (2024) *PingAn*. Available at: https://group.pingan.com/media/news/2024/pingan-takes-lead-in-fortune-fintech-innovator-asia-list-ranked-1st-in-insurtech.html#:~:text=(hereafter%20%E2%80%9CPing%20An%E2%80%9D%2C,retail%20consumers%20across%20the%20country (Accessed: 20 March 2025).
[9] The Geneva Association (September 2023) *Regulation of Artificial Intelligence in Insurance: Balancing consumer protection and innovation*. rep. Available at: https://www.genevaassociation.org/sites/default/files/2023-09/Regulation%20of%20AI%20in%20insurance.pdf.

the quality of the datasets used to train the model directly affects the reliability and accuracy of the response produced.

There is a relatively small number of power and energy claims in the world, and not all insurers will have access to all the data for these claims. In many cases, proprietary information is also used relied upon when assessing a loss, and may have to be excluded from the LLM's training dataset. As such, the available training dataset may be incomplete and/or does not capture the full range of claims patterns.

According to Open AI, the creator of Chat GPT, a minimum of 10 well-crafted relevant examples is required to fine tune an existing model, and clear improvements typically require 50-100 training examples.[10] Therefore, the limited dataset available for similar power and energy claims is likely to impair the reliability and performance of the LLM.

Further, policies issued in the power and energy sector are often manuscript policies. As such, the LLM will need to be constantly fine-tuned to ensure that it has the correct data to provide guidance on the relevant policy. This constant refining of the LLM would invariably increase the operational cost to maintain the LLM. If this fine-tuning process is not carefully managed, the frequent re-calibration of training datasets could compromise the balance and quality of the dataset and undermine the reliability of the LLM.

These limitations mean that no or low touch claims handling is unlikely to be realistic at the current stage of AI development. Indeed, given that the claims seen in the power and energy industry are generally more complex and require more in-depth investigations, often requiring technical experts to assess the full circumstances of the loss, it would also be unwise to rely on current LLMs to handle them without considerable human oversight.

5. Considerations for Insurers when deploying LLMs

Having regard to the limitations of LLMs, Insurers deciding to utilise LLMs as part of their business should design appropriate protocols and guidelines to manage the limitations.

First, there should be robust data governance measures to manage and audit LLM training datasets in order to ensure that the LLMs remain reliable and up to date.

Second, Insurers have legal and/or ethical duties to handle claims fairly. As LLMs are not infallible, it is important to ensure that LLM-generated responses are reviewed by human claims professionals to ensure that they are reasonable and defensible having regard to the circumstances of the claim and the policy terms before they are relied upon. This aligns with the standard practice of Insurers reviewing reports and recommendations from adjusters and other service providers before making a claims decision based upon them.

6. Evolving Insured Risk Profiles

Like insurers, the power and energy industry is also exploring ways to improve their operations utilising LLMs to manage ever more complex tasks. As such, Insurers may also observe a shift in the types of losses they encounter.

While it has been possible to automate and remotely control industrial systems for decades, such systems traditionally operate under many restrictions and/or have limited capabilities.

---

[10] *Fine-tuning* (no date) *OpenAI*. Available at: https://openai.com/form/custom-models (Accessed: 16 April 2025).

Technological advances in AI and LLMs are transforming the manner in which power assets are operated. For example, LLMs are seen as particularly useful in optimising plant performance[11] and predicting maintenance needs.[12]

According to a 2016 Marsh Report, 30% of the power generation losses between 2001 and 2015 (by value) were caused by human errors, and 5% of losses were caused by age- and maintenance-related issues.[13] As such, when appropriately implemented, LLMs may mitigate these traditional risks of loss.

That said, as LLMs are increasingly relied on to make system-critical decisions, new types of losses could arise. Unlike Insurers use of LLMs, where its decision can be reversed, decisions on plant operations have to be made in real-time. There may be limited opportunities for human intervention before a system-critical decision guided by an LLM is put into effect. Any error by the LLM could also have significant operational and/or safety consequences, resulting in larger quantum losses.

7. Cover for AI-related losses

In the power and energy sector, it is not uncommon to see cyber exclusions, such as LMA 5400 or LMA 5426, inserted in operational policies excluding most if not all cyber-related losses.

In general, LMA 5400 and LMA 5426 are drafted in broad terms, and seek to exclude all losses and damage caused by a 'Cyber Incident', subject to write-backs for physical loss or damage caused by fire or explosion (or additionally under LMA 5426, machinery breakdown). A 'Cyber Incident' is defined to include non-malicious errors and omissions involving processing or operation of any software or server.

At first glance, these exclusions appear to apply to LLMs (which is technically 'software'), such that cover for 'errors' made by LLMs are potentially limited to physical damage caused by fire, explosion or machinery breakdown.

As discussed above, there is currently limited understanding of how LLMs make 'decisions'. As such, it may not always be clear whether the model has made an 'error'. As a LLM is trained to predict the most likely outcome based on patterns in the training dataset, an unexpected 'decision' may simply be a valid alternative interpretation of the user input and/or the training data itself. Arguably this is not an error, but an unexpected outcome.

This means the scope of application of cyber exclusions may be more limited than initially apparent. As such, in assessing cover for AI-related losses, claims professionals will need to carefully consider LLM operation to determine if cyber exclusions are properly triggered.

8. Conclusion

Ultimately, AI is a computer program, albeit an exciting new program that functions in a more human-like manner when compared to traditional computer programs. Whilst AI has the potential to disrupt the power and energy insurance space, it is unlikely to fully replace claims professionals in the near to medium future. Until a significant advancement in the capability of LLMs occurs, their role will remain adjunctive – useful but not essential.

---

[11] For example, Aboitiz is developing digital twins for its circulating fluidized bed (CFB) coal-fired power plants which would help to optimize plant performance and enhance operational efficiency and resilience: see Akella, S. (2024) Aboitiz power to introduce first smart power plants in Philippines, Power Technology. Available at: https://www.power-technology.com/news/aboitiz-power-smart-power-plants-philippines/?cf-view (Accessed: 16 April 2025).
[12] See e.g. Cooper, K. (2024) *AI 'revolutionising' predictive maintenance in renewables*, *Power Technology*. Available at: https://www.power-technology.com/news/ai-revolutionising-predictive-maintenance-as-it-gains-traction-in-renewables/ (Accessed: 16 April 2025).
[13] Marsh (2016) *Global Loss Trends: Analysing the Causes of Power Generation Claims*. Available at: https://www.marsh.com/content/dam/marsh/Documents/PDF/UK-en/Global Loss Trends Analysing The Causes of Power Generation Claims.pdf (Accessed: 16 April 2025).

As LLMs and other AI systems become more widely adopted in the power and energy industry, claims professionals will need to adapt to a shifting profile of losses. While the cyber exclusions such as LMA 5400 and LMA 5426 are drafted in technology-agnostic terms and remain broadly applicable, claims professionals will need to carefully consider LLM operation to determine if these exclusions are properly triggered.

So long as the power and energy industry (including Insurers) continue to deploy LLMs responsibly having regard to their limitations, it will be well placed to take advantage of the operational and productivity benefits, while effectively managing the risks of errors occurring.

Authors

Andrew Bott

Executive Loss Adjuster
Chubb Insurance (Australia)
+61 457 833 802
andrew.bott@chubb.com

Audrey Wong

Senior Associate
Wotton Kearney
+65 6967 6467
audrey.wong@wottonkearney.com

Australian offices

Adelaide

Lvl 1, 25 Grenfell St
Adelaide, SA 5000
+61 8 8473 8000

Brisbane

Level 21, 71 Eagle Street
Brisbane, QLD 4000
+61 7 3236 8700

Canberra

Ste 4.01, 17 Moore St
Canberra, ACT 2601
+61 2 5114 2300

Melbourne

Lvl 30, 500 Bourke St
Melbourne, VIC 3000
+61 3 9604 7900

Perth

Lvl 49, 108 St Georges Tce
Perth, WA 6000
+61 8 9222 6900

Sydney

Lvl 9, Grosvenor Plc,
225 George St
Sydney, NSW 2000
+61 2 8273 9900

New Zealand offices

Auckland

Lvl 8, 21 Queen St
Auckland 1010
+64 9 377 1854

Christchurch

203/237 High Street
Christchurch 8011
+64 3 667 4003

Wellington

Lvl 12, 342 Lambton Qy
Wellington 6011
+64 4 499 5589

Asia office

Singapore

138 Market St,
05-01, CapitaGreen
Singapore, 048956
+65 6967 6460