



GenAI and Insurance: What OpenAI's o1 'reasoning' model could mean for insurers

By Jeff Heaton

OpenAI released a major update to ChatGPT on September 12. This release, named o1 and codenamed Strawberry, is already available on the ChatGPT service.

Key takeaways

- OpenAI released its o1 model, codenamed Strawberry, on September 12, 2024.
- The new release adds some fairly advanced reasoning capabilities to the large language model (LLM), allowing it to go through a process that OpenAI calls "thinking."
- This new capability gives ChatGPT some self-checking ability that limits one of its most common weaknesses – hallucinations.

What's new?

The o1 engine boosted ChatGPT's scores particularly in the areas of PhD-level physics questions, LSAT, and mathematics. It also shows great promise to mitigate one of the greatest risks of LLM usage – hallucinations.

OpenAI added additional reasoning capabilities to achieve these goals. But what, exactly, is meant by "reasoning?"

Miriam-Webster defines reason as "the power of comprehending, inferring, or thinking, especially in orderly rational ways." OpenAI does not claim to have achieved human-level intelligence; however, this is one of the most significant advances to the ChatGPT platform to date.

Think back to the Miriam-Webster definition, specifically "in orderly rational ways." The most significant change with o1 is a process that OpenAI calls "thinking." This involves ChatGPT forming a plan to answer a question by creating additional questions to help answer the main query.



Learn how you can put RGA's data and technology experience and insights to work for you.

[Start exploring today →](#)

The o1 model's codename of Strawberry is a reference to a simple question that cannot be answered without reasoning: "If you place a strawberry in a cup, then turn the cup upside down and place it on a table, then place the cup in the microwave, where is the strawberry?"

Most AI systems would simply say the strawberry is in the cup, which is now in the microwave. They failed to work through the problem and see that the strawberry fell out and is now on the table.

The o1 upgrade in action

Insurers might consider the last time they were asked a similar question that took multiple steps to think through. For example:

Jane is a 32-year-old female who has recently been diagnosed with hypertension and an anxiety disorder. She has not yet been prescribed any medications for either. What additional information might a life underwriter want to know and who would it be evaluated?

When asked to walk through the reasoning process it would follow to answer this query, the o1-upgraded ChatGPT platform provided the following response:

Identifying key factors

- I'm beginning to piece together what a life underwriter is considering: Jane's age, medical history, lifestyle, occupation, and how these elements influence mortality risk.

Assessing Jane's circumstances

- I'm gathering info on Jane's 32-year-old age, hypertension, and anxiety diagnosis. No medication details yet, but underwriters need more info to evaluate her risk.

Gathering pertinent data

- I'm collecting detailed medical history, treatment plans, lifestyle factors, family medical history, occupation details, driving record, and other risk factors to comprehensively understand the patient's situation.

Evaluating risk

- Life underwriters assess risk by considering hypertension control, organ damage, cardiovascular risk levels, anxiety disorder severity, stability, compliance, and functional impairments to assign a relevant risk class.

Gathering insights

- I'm assembling detailed medical records, focusing on hypertension and anxiety indicators. This involves recent readings, stages, associated conditions, and treatment plans. Additionally, assessing lifestyle factors like smoking, alcohol use, and stress levels.

Assessing risk factors

- I'm mapping out the family medical history to gauge potential health risks. Plus, additional medical tests might come into the picture to ensure a comprehensive risk assessment.

After then working through this question with the new o1 engine (presumably following a process similar to the one outlined above), ChatGPT was particularly concerned with control of the hypertension and anxiety and wished to monitor the drugs prescribed, as these two conditions frequently have drugs that cannot be simultaneously prescribed. ChatGPT was also interested in Jane's weight and build, including any recent changes.

The primary change with the new o1 engine is that ChatGPT will now form/research additional questions based on the initial request – similar to the human reasoning process. Because it is now running additional queries based on the first query, this can decrease hallucinations, which occur when an LLM essentially fills in a gap by offering a best guess that may or may not be true. In other words, ChatGPT is now checking itself.

As OpenAI marches toward the much-anticipated GPT-5, we are seeing significant advances in the interim releases – first GPT-4o and now o1. Based on the extremely promising features of this latest offering, we should prepare for a transformative road ahead.

Learn how you can put RGA's data and technology experience and insights to work for you. [Start exploring today.](#)



AUTHOR

Jeff Heaton

Vice President,
Data Science, Data Strategy and Infrastructure