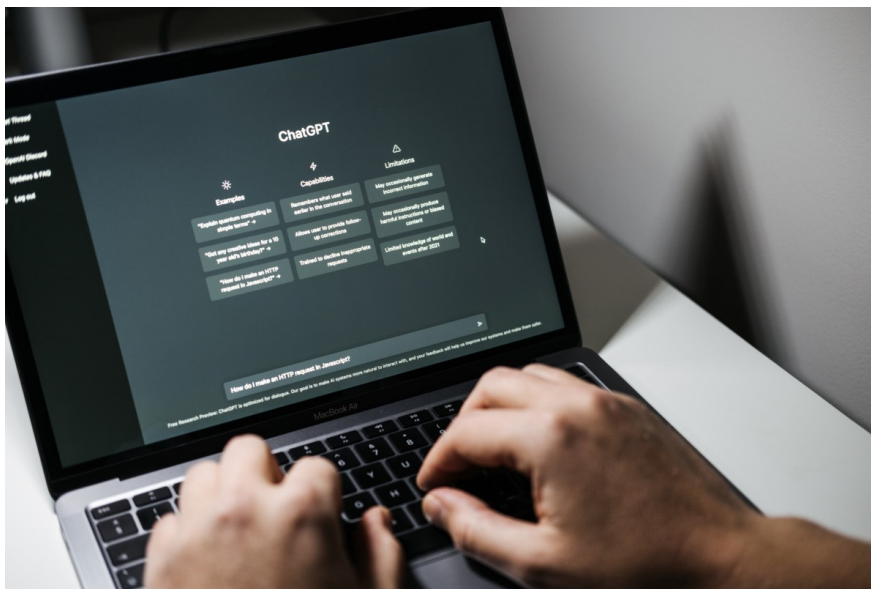


AI monthly: AI's habit of making stuff up is a growing concern

With great power comes great responsibility, and AI is no exception. As usage and trust in AI models grow, ensuring their accuracy is critical. Yet, research shows that some models produce incorrect statements in more than one-third of responses



Are AI models becoming overconfident?

Modern AI models include features like deep reasoning, long-term memory, and autonomous agents that can browse the web or perform tasks with minimal human input. To perform these tasks, the models require vast amounts of data, which increases reliance on uncontrolled and unverified data sources. This increased exposure can lead to behaviour that resembles overconfidence, a cognitive bias where confidence in one's knowledge or abilities exceeds actual accuracy.

A recent [study by the European Broadcasting Union \(EBU\)](#) highlights the consequences of this: leading AI systems generate false claims at a rate of up to 40%. This high rate coincides with a change in model behaviour. Previously, AI systems would decline to answer questions about events outside their training data. Newer systems having web access are designed to respond more frequently, even when they are uncertain or have insufficient information. While this

improves user engagement, it also leads to more fabricated output. We refer to these cases as "AI hallucinations." However, the models often deliver such responses with strong confidence, creating the impression that they are unquestionably correct.

Models are prioritising fluency over accuracy

There are several reasons why AI hallucinations are still common, even with newer models. One reason is that users may ask vague or complex questions, which the model cannot easily interpret. This often prompts the model to "fill in the blanks" using statistical patterns to deliver what appears to be a complete answer. While these responses are intended to be helpful, they can result in factually wrong statements.

Additionally, the models are often fine-tuned using human feedback, meaning they prioritise answers that sound confident and helpful. This results in a bias towards confident-sounding but inaccurate statements rather than cautious or uncertain responses.

This issue is further exacerbated by the declining "no response rate". While older models refused to answer almost 40% of queries, newer models are designed to answer virtually every request. In fast-moving fields such as politics and health, where accurate information is crucial, this shift from accuracy to fluency poses serious misinformation risks. These structural biases are technically embedded in modern AI systems and reflect the preference for confidence over accuracy.

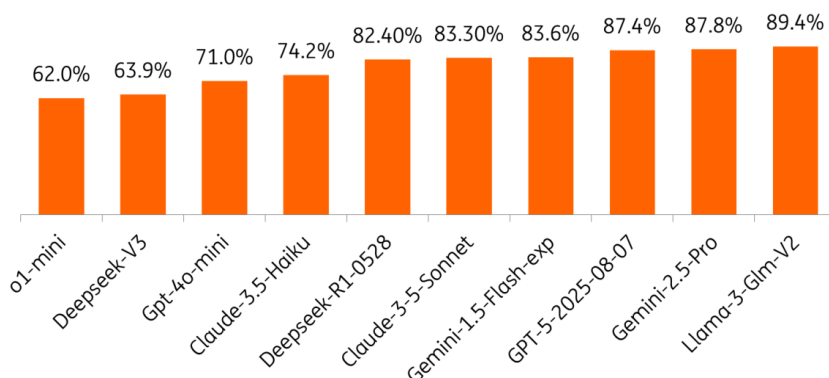
What's at stake

As we reported in a previous edition of [AI Monthly](#), investments in AI are at an all-time high. However, trust and continued investment could be at risk if even the most sophisticated models are found to deliver false information. AI is increasingly used to gather information about current events, particularly among younger users. In fact, 15% of individuals under 25 report using AI chatbots as a primary source of news. Given the [rising usage of AI](#) both privately and in businesses, accuracy should be a priority.

Recent [research](#) shows that factuality improves with newer, larger models, suggesting that further scaling may help address the challenge. Yet performance still varies widely. Some models produce incorrect statements in more than one-third of responses, whereas grounded models, which connect to external sources of truth, can reduce this figure to around one in 10. This emphasises the importance of selecting the right model. Despite these improvements, businesses are still at risk of reputational damage or legal issues due to factually incorrect answers.

The level of factuality varies greatly between AI models

FACTS: factuality score (in % of total answers)



Source: FACTS leaderboard, 2025

Two recent cases have highlighted the impact inaccurate or fabricated output can have on businesses and individuals. In one case, an international advisory firm had to refund the government for a report containing AI errors. A judge was misquoted, and non-existent references were cited.

In another case, a senior judge criticised lawyers who had used AI tools to prepare written arguments referencing fake case law. He called on regulators and industry leaders to ensure lawyers know their ethical obligations. Both cases resulted not only in financial but also reputational damage.

Awareness is vital

In Germany, there's a saying: "stiffly claimed is already half proven" – but confidence does not automatically translate into correctness. AI's low level of accuracy shows that its chances of replacing entire professions in the near future are rather low – unless those who practice the profession rely entirely on the (false) information that AI feeds them. AI-generated statements should be treated with the same critical mindset as human claims.

Author

Julian Geib

Junior Economist, Global Trade

julian.geib@ing.de

Disclaimer

This publication has been prepared by the Economic and Financial Analysis Division of ING Bank N.V. ("ING") solely for information purposes without regard to any particular user's investment objectives, financial situation, or means. *ING forms part of ING Group (being for this purpose ING Group N.V. and its subsidiary and affiliated companies).* The information in the publication is not an investment recommendation and it is not investment, legal or tax advice or an offer or solicitation to purchase or sell any financial instrument. Reasonable care has been taken to ensure that this publication is not untrue or misleading when published, but ING does not represent that it is accurate or complete. ING does not accept any liability for any direct, indirect or consequential loss

arising from any use of this publication. Unless otherwise stated, any views, forecasts, or estimates are solely those of the author(s), as of the date of the publication and are subject to change without notice.

The distribution of this publication may be restricted by law or regulation in different jurisdictions and persons into whose possession this publication comes should inform themselves about, and observe, such restrictions.

Copyright and database rights protection exists in this report and it may not be reproduced, distributed or published by any person for any purpose without the prior express consent of ING. All rights are reserved. ING Bank N.V. is authorised by the Dutch Central Bank and supervised by the European Central Bank (ECB), the Dutch Central Bank (DNB) and the Dutch Authority for the Financial Markets (AFM). ING Bank N.V. is incorporated in the Netherlands (Trade Register no. 33031431 Amsterdam). In the United Kingdom this information is approved and/or communicated by ING Bank N.V., London Branch. ING Bank N.V., London Branch is authorised by the Prudential Regulation Authority and is subject to regulation by the Financial Conduct Authority and limited regulation by the Prudential Regulation Authority. ING Bank N.V., London branch is registered in England (Registration number BR000341) at 8-10 Moorgate, London EC2 6DA. For US Investors: Any person wishing to discuss this report or effect transactions in any security discussed herein should contact ING Financial Markets LLC, which is a member of the NYSE, FINRA and SIPC and part of ING, and which has accepted responsibility for the distribution of this report in the United States under applicable requirements.

Additional information is available on request. For more information about ING Group, please visit www.ing.com.