INSIGHTS · TOOLS AND THE FUTURE

Artificial Intelligence for Insight Generation

What artificial intelligence, including machine learning and large language models, can and cannot do in insight work, from coding text and finding patterns to simulating respondents, with the evidence on its benefits and risks and guidance on using it responsibly.

ModQu research article · 816 words · 9 references (APA 7th edition)

Artificial intelligence is changing how insight is produced. Machine learning models find patterns in large datasets, and large language models can summarise interviews, code open-ended survey responses, draft reports and even answer survey questions as if they were consumers. These tools can make insight work faster and cheaper. They also introduce new ways to be wrong. This article summarises what AI does well in insight work, where caution is needed, and how to use it responsibly.

What AI does well

Agrawal et al. (2018) argued that the main economic effect of AI is to make prediction cheap, and that this raises the value of the human judgement needed to decide what to predict and what to do with the results. Ma and Sun (2020) reviewed the use of machine learning in marketing and described its strengths in handling unstructured data such as text, images and voice, and in finding complex patterns in large datasets.

In insight work, these strengths apply to tasks such as:

- Classifying and summarising large volumes of open-ended text, reviews and transcripts.
- Detecting patterns and segments in behavioural data.
- Forecasting demand and customer behaviour.
- Drafting summaries, discussion guides and questionnaires for human review.

Evidence on productivity is encouraging for suitable tasks. Noy and Zhang (2023) found in an experiment that professionals given access to a large language model completed writing tasks substantially faster and produced work rated as higher quality. Dell'Acqua et al. (2023) studied consultants using AI and found large gains in speed and quality on tasks within the AI's capabilities, but worse performance on a task outside them, where consultants using AI were more likely to reach incorrect conclusions than those working without it.

Where AI helps in insight work, and what to check

Five common tasks

Task	How AI helps	What to check
Coding open text	Fast and consistent	Check samples against human coding
Summarising interviews	Quick first drafts	Check quotes and claims in the source
Finding patterns	Scales to many variables	Check that patterns hold in new data
Simulating respondents	Cheap early exploration	Check against real respondents
Drafting reports	Saves time	Check every number and claim

Diagram by ModQu ModQu.com

Figure 1. Five insight tasks where AI helps, and what to check. Diagram by ModQu.

Where caution is needed

Fabrication. Language models can produce fluent text that is not supported by their sources. Ji et al. (2023) reviewed research on this problem, often called "hallucination", and found it across many kinds of text generation. A summary of interviews may contain plausible quotes no participant said, or a report may contain statistics that do not appear in the data.

Simulated respondents. Argyle et al. (2023) found that a language model, prompted with the characteristics of real survey respondents, produced answers that matched the patterns of political attitudes in some human samples closely. Brand et al. (2023) found that a language model's responses to questions about products reproduced some well-known patterns of consumer behaviour, such as lower willingness to pay at higher prices. These results suggest that simulated respondents may be useful for early exploration. But they reproduce patterns in the data the models were trained on, and may miss what is new, local or specific to a brand's own customers, which is often what insight is for. Bisbee et al. (2024) also found that simulated responses varied less than real ones, which can make opinions seem more uniform than they are.

Over-trust. Logg et al. (2019) found that people often gave more weight to advice when told it came from an algorithm than from a person. Combined with fluent output, this can lead teams to accept AI conclusions without the scrutiny they would apply to a colleague's work.

An example

Consider a hypothetical retailer that receives 40,000 open-ended survey comments a year. Previously, analysts coded a sample of 1,000 by hand, which took two weeks. The insight team now uses a language model to code all 40,000 comments into themes in a few hours. To check the results, two analysts independently code a random sample of 500 comments, and the team compares their codes with the model's. Agreement is high for most themes but low for one, "value for money", which the model confuses with complaints about price. The team refines the instructions to the model, checks again, and documents the agreement rates in its report. The model's speed allows the team to track themes monthly rather than annually, while the human check ensures that the themes mean what the report says they mean.

Using AI responsibly

- **Keep humans responsible for conclusions.** AI can propose; people should decide what is true and what to do.
- **Check against the source.** Verify quotes, numbers and claims in AI-generated summaries.
- **Validate patterns and predictions** on data not used to find them.
- **Treat simulated respondents as hypotheses**, to be tested with real people.
- **Protect data.** Ensure that customer data shared with AI tools is handled in line with consent and privacy commitments.
- **Record how AI was used**, so that others can judge the results.

Key takeaways

AI makes some insight tasks much faster, especially those involving large volumes of text and data. It also introduces risks: fabricated content, simulated respondents that miss what is new, and a tendency to over-trust fluent outputs. Used with human judgement, checks against sources and validation on real data, it can extend what insight teams can do.

References

Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction machines: The simple economics of artificial intelligence*. Harvard Business Review Press.

Argyle, L. P., Busby, E. C., Fulda, N., Gubler, J. R., Rytting, C., & Wingate, D. (2023). Out of one, many: Using language models to simulate human samples. *Political Analysis, 31*(3), 337–351. https://doi.org/10.1017/pan.2023.2

Bisbee, J., Clinton, J. D., Dorff, C., Kenkel, B., & Larson, J. M. (2024). Synthetic replacements for human survey data? The perils of large language models. *Political Analysis, 32*(4), 401–416. https://doi.org/10.1017/pan.2024.5

Brand, J., Israeli, A., & Ngwe, D. (2023). *Using GPT for market research* (Working Paper No. 23-062). Harvard Business School.

Dell'Acqua, F., McFowland, E., III, Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023). *Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality* (Working Paper No. 24-013). Harvard Business School. https://doi.org/10.2139/ssrn.4573321

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys, 55*(12), Article 248. https://doi.org/10.1145/3571730

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes, 151*, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

Ma, L., & Sun, B. (2020). Machine learning and AI in marketing: Connecting computing power to human insights. *International Journal of Research in Marketing, 37*(3), 481–504. https://doi.org/10.1016/j.ijresmar.2020.04.005

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science, 381*(6654), 187–192. https://doi.org/10.1126/science.adh2586