

Using AI Superpowers for Good

An Update on AI's Quest to Conquer the Insights Industry

Fall 2024

The speed of AI acceptance is unmatched by any innovation in history, and the pace of active adoption is not far behind. We're all scrambling to figure out what AI is good for and what we can safely do with it – but the dizzying rate of advance makes it hard to take stock. A scant nine months ago, NAXION published a late-breaking article describing some of AI's limitations for insights work based on empirical observations and some tire-kicking experiments. Today, nearly every one of those cautionary notes has been qualified or contradicted by rapid advances in the models and improvements in the products they are spawning. *The only point that remains indisputable – presumably for the indefinite future – is the premise that safe AI always needs a human in the room.* Even today's third- and fourth-gen LLMs retain the potential to hallucinate, and they can certainly make mistakes. No matter how good they get, they must be guided, and they need to be checked.

Recognizing that the shelf life on any discussion of AI is short, we'd like to offer a brief update on where AI is seeing the greatest and most successful use by insights-centered consulting organizations – recognizing that this will be the first of many installments. *Our focus here is on the implications it may have for clients, who care more about the quality of insights served than what happens in the kitchen so long as the outcome is nutritional and safe to consume.*

This article briefly highlights four angles on AI application:

- AI-powered increases in efficiency
- The state of AI qualitative research
- The role for synthetic data
- AI-guided data harmonization



In day-to-day workings of the insights field, AI's major contributions remain efficiency-related and, thus, largely imperceptible to clients.

One of AI's chief contributions to market research takes the form of added efficiency in programming. LLMs adapt nicely to code writing and can now call up a large reservoir of GitHub tools, freeing humans to perform more complex analytic tasks that are not so easily delegated. This use of AI is invisible and immaterial to clients despite the ubiquity of marketing terms like "AI-powered." (True fact: In a business that has touted innovation for years, everyone is afraid to be caught looking like a slow adopter.)

Another remarkably successful application has come in the form of qualitative analysis and report drafting, where even early products have shown themselves to be very good at thematic analysis and basic data summaries. Although AI tools are ill-equipped to plumb psychological depth and intuit motivation or resolve contradictions, the best of the AI-drafted topline summaries free up humans to do other things (besides supervise their AI tools) including spend more time on higher-level analysis. The setup required to guide them through a study can also have the added benefit of forcing teams to think harder about hypotheses, goals, and proof standards, potentially enhancing the outcome. And on occasion, AI can spot something a human analyst has missed while combing through reams of transcripts.

Improvements in AI will also facilitate basic survey data analysis by calling up prescribed statistical actions and converting data outputs into basic "table talk." Both the value and limitations remain to be seen – especially for complex analyses and challenging marketing questions. Again, humans will still have lots to do here, both in auditing the processes and interpreting the data contextually. AI recognizes and organizes far better than it thinks and concludes.

So far, none of this has brought down the price tag for quality work, in part because industry margins are already tight. More to the point, there is still lots of human oversight needed to avoid error and supply nuanced interpretation. AI may increase the speed of the work sooner than it brings down the cost.

While LLM widgets can be trained to follow up on open-ended survey responses with a few shallow probes, they cannot yet be trusted to replace human moderators any more than they can be trusted to replace human respondents.

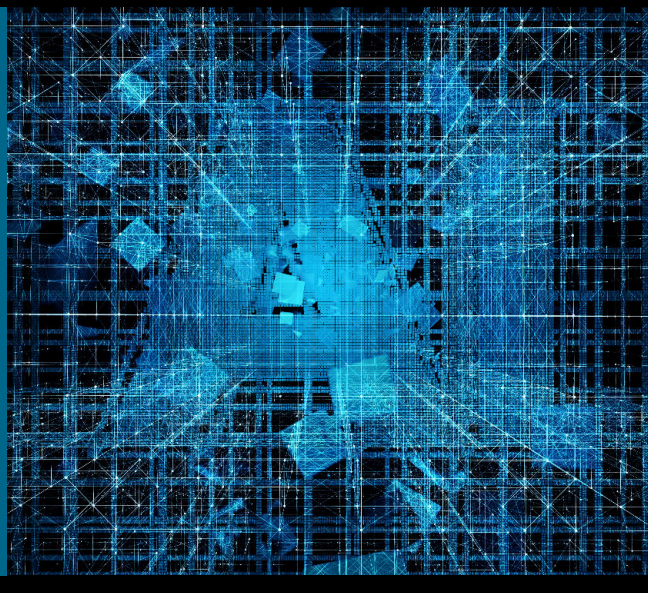


It is possible to use an AI widget to probe open-ended survey responses, but the results we've seen so far have not been impressive. They don't necessarily improve on what we can accomplish by anticipating possible responses and creating our own pre-programmed follow-ups. We continue to experiment with these products, so there's more to come, but by its nature, this application is certainly not game-changing – it's more likely just benign.

More worrisome are some misguided assumptions about the value and scalability of "AI qual." LLMs are pretty good conversationalists, but AI-guided qual interviews do not equate to sophisticated qualitative research as most of us would (or should) think of it, and they are not going to be better by the pound, despite rumblings about *AI qual-at-scale*. In another piece, we'll discuss the limitations of even the best-trained bots as moderators and the serious deficiencies of AI qual-at-scale, which we think of as a lazy man's survey – neither fish nor fowl, neither richly insightful nor statistically informative.



“Synthetic” datasets that mimic real ones are now used extensively to train models and perform complex analysis of sensitive data to avoid privacy risk – but experiments with synthetic *opinion* data have raised significant flags.



Synthetic data are artificially derived datasets developed with the guidance of algorithms to mimic real ones. Current use of synthetic data is primarily in the service of (1) efficient model training (think of a ball machine to intensify practice) and (2) the development of de-identified alt-databases for circumstances in which confidentiality is a high priority. For instance, in healthcare, where highly sensitive and (thanks to AI) ever-larger datasets are available to model critical disease outcomes, synthetic datasets look just like the real thing but without PII attached. Financial databases, census information, and other highly sensitive data sources also come quickly to mind as ideal applications. Synthetic data can also be used successfully to uncover bias and retrain models to perform more *ethically* as well as more efficiently – a hugely important undertaking.

But as technical terms go, synthetic data is also a sexy one. It brings to mind AI-generated survey cyborg respondents who can do a better, cheaper job of conveying popular sentiment than real ones. The premise is that when AI has free run of the internet and the right prompts, it can avoid the sampling bias and validity concerns of traditional survey data. *But can it?*

As of today, this idea still looks more like wishful thinking than imminent reality. Recently published work on synthetic opinion data raises serious concerns about the way even newer LLM models perform when instructed to role-play as survey respondents guided by sophisticated, up-to-the-minute algorithms. Statistical discrepancies and anomalies abound when the data are modeled – and some of the problems may be fairly intractable.

Nor is AI able to invent small qualitative personas we can trust to add insight. Even with its now far-reaching access to web content and its more advanced training in role-play, LLMs dispatched to the internet often bring back uninteresting faux respondents whose “point of view” is not necessarily worth hearing (As a general rule, AI seems to do better at the mean than the margins with synthetic data, which makes modeling dicey.) And notwithstanding its superficial resemblance to real-world datasets, synthetic opinion data carries algorithmic bias that distorts the representation of critical subgroups. Moral: Don’t count on AI to write your love letters or invent your respondents *just yet*.



LLMs are playing a key role in harmonizing vast and highly diverse datasets to power up the models. but the ROI on AI-guided data integration for smaller data libraries like market research archives will vary case by case.

Data integration is not a new need or a new concept – just a decade ago, humans were still integrating datasets using relatively “primitive” computational tools. In recent years, exponential growth in the size of databases and our modeling aspirations have changed the game, not just the rules. To appreciate the utility of AI, think again of health outcomes modeling based on hundreds of different trials worldwide and millions of case records culled from many health systems that can only be harmonized with AI models. Clients seeking this same integration of market research archives (in hopes of spotting patterns not otherwise detectable) will see varying degrees of benefit based on factors like data quantity, variability, time frame, contextual information, and available documentation. Significant internal or external data science resources are required for an insights department to accomplish this, and a cautious assessment of likely ROI is advised. Feedback from clients thus far suggests a heavy lift, especially since access to internal resources can be scarce. An upcoming NAXION article will recount some of those experiences and practices.



We’re keeping a close and cautious eye on new developments, with the expectation that AI will continue to acquire new superpowers – and some may be accidentally misused. We’ll also be on the lookout for kryptonite. Watch this space.

About the Author



Susan Schwartz McDonald, Ph.D.
CEO, NAXION
215.496.6850
smcdonald@naxionthinking.com

Susan’s career focus has been on the development and protection of robust brands and the research methodologies needed to support them. She has contributed to the evolution of many standard research techniques, and she writes frequently on industry topics and issues of broader interest. Her commentary on language and culture can be found at www.smartmouth.blog. Susan holds M.A. and Ph.D. degrees from UPenn’s Annenberg School of Communication.

About NAXION

NAXION is a nimble, broadly resourced boutique that relies on advanced research methods, data integration, and sector-focused experience to guide strategic business decisions that shape the destiny of brands. Our century-long history of innovation has helped to propel the insights discipline and continues to inspire contributions to the development and effective application of emerging data science techniques. For information on what’s new at NAXION and how we might help you with your marketing challenges, please visit www.naxionthinking.com