

What synthetic data can and cannot do for a brand decision, and how we use it at One Inch Whale

Ask a synthetic consumer whether they are sure, and they almost always are. In Pew Research Center's September 2026 test, AI respondents chose "unsure" 4% of the time. Real people chose it 16% of the time. That gap captures the problem with synthetic data in one number: it sounds more certain than consumers are.

Synthetic data here means survey or interview answers generated by AI to stand in for consumers, from simple demographic personas to "digital twins" built on what real individuals have said before. The promise is obvious: a thousand respondents in minutes, no recruitment, a fraction of the cost. The question that matters is a different one. Would you put a launch budget behind those answers?

Fluent, confident and twelve points off

Pew's test is the most rigorous so far. Researchers built a digital twin of each panelist from that person's own earlier survey answers, then put nearly 300 questions to the twins. Even with the best-performing model, the synthetic results missed the human results by about 12 percentage points on average.

The twins also overestimated what people know. On factual questions, real panelists were right about half the time; their AI counterparts about 80% of the time. And the outcome depended on the model: switch from one leading model to another and the answers shift. Change the model, and you change your "consumer".

Pew measured public opinion, and brand research shows the same pattern. A 2026 study by the Nuremberg Institute for Market Decisions (NIM) found that synthetic respondents captured broad brand preferences but rated brands more positively than real people did, most of all well-known brands. They also varied far less from one respondent to the next.

The average answer, at scale

Put those findings together and a clear picture appears. Synthetic samples converge. They show less variance, more consensus, more certainty and more goodwill towards big brands than real consumers do. They produce the average answer, and they produce it at scale. Generating more responses only multiplies the effect.

For a brand, the cost is concrete. Inflated positivity can wave through a weak launch. Missing variance can hide the niche worth building a proposition around. And when results move with the prompt and the model, a researcher's assumptions can come back dressed up as consumer evidence.

Decisions are rarely won on the average. They are won at the edges: the ambivalent shopper, the minority motive, the reason nobody says out loud. Those are exactly the signals synthetic samples smooth away, and exactly where our work starts.

How we use synthetic data at One Inch Whale

We give synthetic data one clearly defined job: exploration. It has a fixed place on our Evidence Ladder, the framework we use to decide how much certainty a decision needs and which evidence delivers it.

Level 1, AI intelligence: what do we already know?

Level 2, Synthetic exploration: what do we expect?

Level 3, Human directional: does it hold with real people?

Level 4, Human validation: is it strong enough to put serious money behind?

Synthetic data lives on Level 2, and it earns its place there. It helps us stress-test hypotheses, anticipate objections and sharpen the questions we take into fieldwork. NIM's guidance on synthetic samples points to the same use: refining research materials before they reach real respondents. Simulated objections to a new pack can make an interview guide sharper. They cannot tell us what a shopper will do at the shelf.

That leads to four commitments:

Every output we deliver states its evidence level. Clients always know how much weight a finding can bear.

Synthetic responses are never counted in a sample and never quoted as if they came from consumers.

Synthetic data is never the only basis for a go or no-go decision.

Client data stays confidential whenever it informs a synthetic exercise.

Four questions to ask any synthetic data provider

1. What human data was this benchmarked against, and in which category?
2. How much variation does it show compared with real respondents?
3. Does it ever say "I don't know"?
4. What happens to the results when you change the model or the prompt?

If the answers are vague, so is the evidence.

AI makes our work faster, and we use it every day. Speed only matters when the answer holds. That is what we mean by Always Sharp, Never Shallow.



Sources

Pew Research Center (30 September 2026). How well synthetic samples replicate public opinion; How synthetic polling results change based on the AI model. pewresearch.org/data-labs

Kaiser, C., Kaiser, J., Manewitsch, V., Schallner, R. & Rau, L. (2026). Leaving Insight to Digital Twins? Promise, Progress and Limits of Synthetic Respondents. NIM Marketing Intelligence Review, 18(1), 48-53.

Sarstedt, M., Adler, S. J., Rau, L. & Schmitt, B. (2026). Your Next Respondent Might Be an LLM: Guidelines for Using Silicon Samples in Marketing Research. NIM Marketing Intelligence Review, 18(1), 24-29.

