

How to evaluate AI research tools: A smart buyer's guide



Know it now.

Introduction

The world of consumer research is being reshaped by AI at speed. Every month brings a fresh wave of tools promising to transform how we understand consumers: surveys that write themselves, interviews run by digital moderators, insights served up in seconds. It's exciting. It's overwhelming. And it's changing the game for researchers.

But with so many tools flooding the market, and so few standards in place, separating what's genuinely useful from what's just well-marketed is becoming a full-time job.

It's easy to be swept up by shiny demos and big claims. Many vendors lean heavily on AI as a buzzword, but fail to demonstrate how their technology actually works or what value it truly adds.

And that's a problem. Because in research, the newest technology isn't necessarily the most effective. What matters is how responsibly that technology has been developed, how well it integrates into real workflows, and whether it can be trusted to deliver reliable, unbiased insights.

Where is the trust?

Trust, in this context, is everything. We're dealing with tools that shape how businesses understand their customers, how decisions are made, where budgets go. And yet, some vendors operate what we call 'AI black boxes': tools that won't – or can't – explain how their models are trained, where the data comes from, or how bias is detected and mitigated. That's not just opaque. It's risky.

As AI becomes more central to how we run research, transparency, accountability and methodological rigour must be non-negotiable. You deserve tools that show their workings.

That's why we've created this guide. To help you navigate the new AI research landscape with confidence.

We'll unpack the different types of AI tools on offer, explain what questions to ask potential vendors, and share a practical framework for evaluating them – not just on technical capability, but on integrity, transparency and quality. And we'll share what responsible AI development should look like, because in a market moving this fast, caution is just as important as curiosity.



Sam Killip
VP of Insights
Attest

Contents

- 03** A look at how AI is used in consumer research
- 07** What really matters: 3 core tenets for assessing AI tools
- 10** The 4-step framework for evaluating AI research tools
- 11** Documents checklist
- 12** Key questions to ask vendors
- 14** The Attest approach to AI: built for real researchers

A look at how AI is used in consumer research

AI is not one thing, it's a spectrum of capabilities. From models that simulate consumer reactions without speaking to a single person, to tools that analyse real responses in seconds, AI is reshaping every stage of the research process. But not all applications are equal in maturity, risk, or value.

To make sense of this complex landscape, we've broken down the key types of AI you're likely to encounter. Understanding what each one does will help you ask smarter questions and choose the tools that best serve your research goals.

Broadly, AI in consumer research is emerging in three key forms: synthetic offerings, which simulate human responses without collecting new data; conversational co-pilots, which support researchers in designing studies and interpreting results; and AI-moderated interviews, which automate and scale qualitative data collection. Every method presents unique benefits, as well as distinct implications for trust, transparency, and overall effectiveness.



Synthetic audiences: virtual people, real insight?

One of the most talked-about developments is the rise of synthetic research; tools that use large language models (LLMs) and generative techniques to simulate opinions, behaviours and even entire demographic groups. Rather than collecting data from real people, these platforms generate it using trained models.

What does that look like in practice? You might chat with a simulated Gen Z shopper about a new product, or run a concept test and get back a set of predictive scores, all without ever launching a survey. Synthetic models are being used to test creative ideas, stress-test messages, and fill in gaps where certain audiences are hard to recruit.

For fast-moving teams, the speed is a massive win. You can go from 30 ideas to a top 10 shortlist in hours, using synthetic feedback to narrow the field before investing in live research.

But speed shouldn't come at the cost of validity. The challenge with synthetic data is this: how do you know the outputs are representative of real consumer behaviour? If the training data is biased or limited, so are the results. As with any model, it's a case of garbage in, garbage out.

Many vendors are building or integrating synthetic models into their platforms. Some are developing proprietary systems, while others are white-labelling third-party technology. Either way, most will admit the technology is still in its early stages. It's promising, but not yet ready to consistently replace real people at scale.

Key challenge: Synthetic research offers valuable shortcuts, but only if you understand exactly what's being modelled, how the outputs are generated, and what data the models are trained on. Without this transparency, you risk basing decisions on assumptions rather than reality. Synthetic tools should not be viewed as a standalone replacement for real consumer data.



When to use synthetic audiences

Synthetic audiences are ideal for early-stage exploration: quickly narrowing down creative ideas, testing concepts with hard-to-reach segments, or validating hypotheses before committing to full-scale research. They offer speed and directional feedback when timelines are tight or budgets are limited. Use them to prioritise what to test with real consumers, not to replace live research entirely.

Conversational AI and co-pilots: the new research assistant

A growing number of research platforms are integrating conversational AI; tools that allow users to interact with their data or workflow using natural language. These co-pilots are designed to help researchers streamline time-consuming tasks like building surveys, analysing results, and writing up findings, all by simply describing what they want to do.

For many teams, this is where AI becomes truly tangible. Instead of clicking through menus or building logic manually, users can ask the system to “create a concept test for a new beverage” or “summarise key findings by gender.” The result is a significant reduction in setup time, enabling researchers to spend more time on strategic thinking and less on repetitive tasks.

These tools vary widely in how they operate and how integrated they are into the research workflow. Some offer lightweight assistance for specific steps, while others aim to support the entire end-to-end process. The most effective co-pilots are trained on research-specific inputs and use built-in guardrails to encourage best practices and mitigate common errors. Others are little more than chat interfaces bolted onto generic LLMs, which may not be as finely tuned to the nuances of survey design or data interpretation.

There's also a growing interest in conversational tools that help researchers revisit and reuse past research; an area long under-leveraged in many organisations. These features help teams avoid duplication and unlock more value from historical data.

Key challenge: As with all AI tools, the output is only as good as the model and the design behind it. A co-pilot needs to be more than just a fancy chatbot; it should genuinely improve speed and quality, not simply repackage existing workflows with a conversational interface. It's important to evaluate whether these tools are truly solving for researcher pain points, or simply shifting where the work happens.



Meet Compass — Your research co-pilot

Compass is Attest's conversational AI that helps you design, edit and analyse research through natural conversation. Whether you're drafting questions or building a story from your results, Compass helps you move from intent to output in moments. Attest's model doesn't simply hack a generic LLM – it's specially built by researchers for researchers.

Need to write a survey about emerging snacking habits? Ask Compass. Want to know what's driving a shift in Gen Z shopping behaviour? Ask Compass. It's built for how you already work, accelerating and streamlining your process.

AI-moderated interviews: scaling qual without losing the human touch

One of the more intriguing developments in AI research is the rise of AI-moderated interviews. These platforms simulate the structure and flow of a traditional in-depth interview (IDI), prompting respondents, probing for detail, and recording responses at scale.

This opens up qualitative research to new possibilities. You can now run hundreds of interviews across global markets, at speed, without tying up your team for weeks. Some tools also analyse non-verbal cues like tone and body language, especially in video-based interviews.

That said, depth and nuance remain the biggest hurdles. Some tools still feel like a machine reading a script. Others lack the true interactivity that a skilled moderator brings; the ability to spot an unexpected insight and follow the thread.

AI is excellent at synthesising large volumes of qualitative data, summarising key themes and tagging recurring ideas. But the quality of the input still matters. If the conversation isn't rich, the analysis can only go so far.

Key challenge: Ensuring the AI doesn't just conduct interviews, but delivers insights with sufficient depth and nuance to inform decisions. This depends on how well the tool follows up on unexpected responses, avoids leading questions, and maintains respondent engagement. The real test is whether the output meaningfully approximates the value of a skilled human moderator.



When to Use AI-Moderated Interviews

AI-moderated interviews are best for scaling qualitative research quickly and consistently. They're ideal when you need rich, open-ended feedback across multiple markets, or want to automate elements like transcription and thematic coding. These tools are especially useful for exploratory research, pre-survey scoping, or when human moderation isn't feasible due to time or resource constraints.



What really matters: 3 core tenets for assessing AI tools

In the rush to build and adopt AI-powered tools, it's easy to get distracted by flashy demos or long lists of features. But the real test of any AI-powered research tool isn't what it claims to do – it's how it does it.

This section outlines the three most important criteria for assessing AI tools in consumer research: transparency & trust, data quality, and research & AI credentials. Use these as your North Star when considering any vendor's offer, especially when the tech is moving faster than the standards.

1. Transparency & trust: no black boxes

AI tools shouldn't be a mystery but some vendors still treat them like one. Transparency means more than listing a feature name or throwing around terms like "machine learning." It's about clearly articulating how the AI works, what it was trained on, and what safeguards are in place to prevent bias or error.

You should also pay close attention to how vendors present their capabilities. Are they clear about what their tools can genuinely do today versus what's still in development? Overpromising is common and frustrating. Buyers need to know what's ready for real use, not what's in the pipeline or being demoed with smoke and mirrors.

Ethical principles also matter here. Look for vendors who have a published stance on responsible AI, who are open about where AI is used within their product, and who communicate both the strengths and limits of their approach.

Why it's critical: Researchers' credibility is dependent on the quality of their studies. If a vendor can't explain how their AI operates, what data it uses, or what its current limitations are, then you risk introducing uncertainty into your research (and possibly into the minds of your stakeholders).





2. Data Quality: AI is only as good as the inputs

We've all heard the term "garbage in, garbage out". This applies exponentially to AI. The effectiveness of any AI system depends entirely on the quality of the data it learns from and processes. For synthetic tools, that means understanding the training data. For analytics platforms, it's about the accuracy, structure and integrity of the source data being analysed.

A trustworthy vendor should be upfront about their data lineage. What's the origin of their data? What controls are in place to prevent survey fraud or bias? How do they ensure that samples are representative and statistically valid? What rules are in place for upholding significance when applying splits?

Quality and transparency are inseparable here. If a vendor can't explain why their data is robust or how their systems maintain accuracy, you have to question the integrity of the insights.

Why it's critical: With AI, flawed inputs aren't corrected; they're magnified. That's why poor data leads to more than poor outcomes – it leads to confidently wrong ones. Look for vendors who invest in quality controls and can clearly justify why their data – and your output – can be trusted.



Data quality at Attest

Attest's approach to data quality combines rigorous fraud detection, behavioural screening, and real-time response validation. We're actively developing advanced models that use LLMs to detect patterns consistent with LLM-generated survey responses, protecting against AI pollution in respondent data. Our system flags suspicious behaviours, including rapid completion times, inconsistent answers, and linguistic anomalies, ensuring datasets remain representative, human-authored, and decision-grade.



3. Research & AI credentials: the right people, the right foundation

In AI, having the tech isn't enough; you also need the expertise. Building responsible, high-performing AI for research means integrating data science with deep domain knowledge. This isn't a space for generalists. The most effective tools are designed by teams who understand both how research works and what AI can do responsibly.

Look for signs of that depth: Does the vendor have experienced researchers involved in product development? Do they demonstrate methodological rigour? Are they transparent about the sources of their training data, and do they disclose partnerships used to fill any gaps?

Now consider the company's background; are they an established research brand that's been steadily growing their AI capabilities, or are they an AI startup turning their hand to research? While time in the industry isn't the only indicator here, it should be evaluated if they have the necessary knowledge and resources in-house to do AI-assisted research well.

Why it's critical: The real promise of AI is speed, but if the researcher spends more time verifying the AI's output than they saved, the tool is a net loss. Make sure your vendor has the in-house data and expertise to deliver on their claims.

Looking beyond the features

One final note: it's just as important to evaluate what a vendor doesn't offer as what they do. A platform that jumps into synthetic data generation or automated interviews without first solving for quality and rigour might be ticking boxes, not solving real research problems.

That's where "AI washing" becomes a risk. Flashy features that sound impressive but deliver limited value can distract from what really matters: actionable, trustworthy insights.

The 4-step framework for evaluating AI research tools

Once you've defined what matters – transparency, data quality, and credentials – the next step is knowing how to interrogate a vendor's offer. What do you ask in a demo? What should you focus on when reviewing documentation? And how do you spot the difference between genuine capability and marketing gloss?

This 4-step framework is designed to help you go beyond the pitch deck. Use it to stress-test the maturity, ethics and utility of any AI-powered consumer research tool. You can find the key questions to ask for each tool type – whether it's synthetic, co-pilots or AI interviews – on page 12.

The 4 steps to follow

1. Clarify the claims

What does the tool actually do right now? Separate current functionality from future roadmap. Ask for clear use cases and demo what's available today, and not what's "coming soon."

2. Probe the foundations

What powers the tool? Understand the underlying data, training sources, and guardrails. Is it transparent about how outputs are generated and how it handles bias or errors?

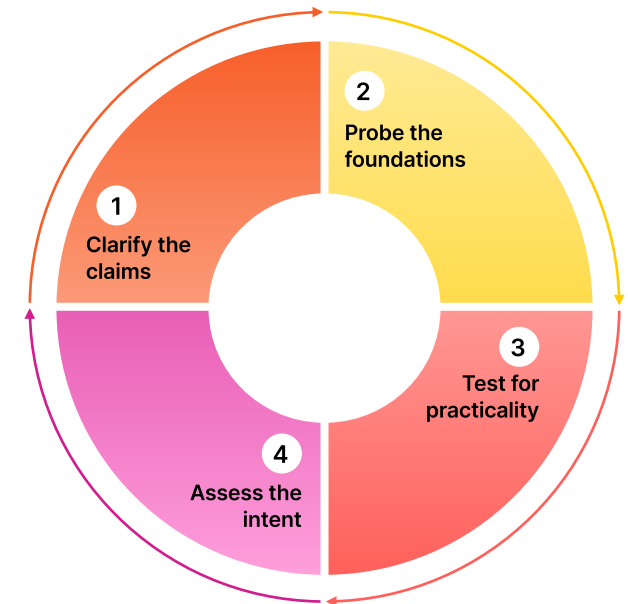
3. Test for practicality

Will it fit into your workflow? Look for real examples of time savings, accuracy improvements or process enhancement. Can users interact with and refine the output?

4. Assess the intent

Is this tool solving real problems, or ticking an AI box? A thoughtful, phased approach to feature rollout is a sign of a vendor focused on utility, and not simply jumping on the bandwagon.

The 4-step research evaluation framework



Documents checklist

Don't stop at a demo, push the vendor to provide written documentation. This list outlines the key materials every buyer should request to verify a vendor's claims and assess their readiness. Used alongside the evaluation framework, these documents give you the visibility to make informed, confident decisions about AI in research.

- Product overview or feature guide**
A detailed breakdown of the tool's current capabilities, use cases, and technical functionality.
- AI methodology statement**
An explanation of how AI is used within the platform, including model types, training data sources, and intended outputs.
- Data quality assurance protocols**
Documentation on how data integrity is maintained, covering respondent validation, fraud prevention, and any use of synthetic data.
- Privacy & data handling policy**
Clear information on data collection, storage, retention, and whether customer data is ever used for training models.
- Bias detection & mitigation report**
Evidence of steps taken to identify and reduce bias in AI-generated outputs.
- Model transparency overview**
A plain-language guide explaining how key models work and how users can interpret, interrogate, or audit outputs.
- Security & compliance certifications**
Relevant standards met (e.g. ISO 27001, SOC 2, GDPR compliance) to ensure safe handling of sensitive research data.
- Product roadmap (if available)**
A forward-looking view of planned AI developments, clearly distinguishing between near-term launches and future concepts.
- Case studies or benchmarks**
Real-world examples of how the AI tool has improved efficiency, accuracy, or outcomes for other users.
- Terms of use & SLAs**
Legal documentation outlining what's guaranteed, what's experimental, and what happens if things go wrong.

Key questions to ask vendors

When it comes to AI research tools, asking the right questions is your best defence against hype. It's how you uncover what's real, what's useful, and what's still aspirational. The right questions don't just validate features; they reveal intent, rigour and long-term thinking.

Questions to ask when evaluating synthetic offerings

Training data & source

- What dataset(s) was your foundational model trained on?
- What is the recency and geographic/demographic balance of this data?
- Do you include any primary data collection to ground the synthetic output in current reality, or is it purely simulation?

Validity & accuracy

- What external benchmarks (e.g. election results, market share data) can you provide to demonstrate the accuracy of your synthetic populations?
- How do you handle low-incidence or highly niche audiences where training data may be sparse or outdated?

Methodological transparency

- Can we audit the 'reasoning path' of a synthetic answer, similar to a prompt chain in an LLM?
- What safeguards are in place to prevent repetitive or overly plausible "template" responses that sound good but reveal little?

✅ Look for answers that reflect both ambition and caution. Vendors who can speak confidently about their capabilities but also acknowledge the limits of synthetic data today.

Questions to ask when evaluating conversational co-pilots

Model design & training

- Is your co-pilot built on a general-purpose LLM, or is it fine-tuned for research-specific tasks and language?
- What steps have you taken to ensure the co-pilot follows research best practices (e.g. unbiased question phrasing, appropriate methodologies)?

User experience & flexibility

- Can users still manually interact with the survey design or data analysis, or are they locked into the co-pilot's output?
- Does the co-pilot learn from user context (e.g. past projects or brand preferences), or is it a one-size-fits-all tool?

Impact & validation

- What evidence can you share to show this tool actually saves time or improves outcomes?
- Have you tested it across a range of use cases – from survey creation to thematic analysis – and what did you learn?

✅ Look for co-pilots that are embedded in real workflows, not standalone features. They should understand research logic, support editing and iteration, and offer explainable outputs you can trust.

Questions to ask when evaluating AI-moderated interview tools

Respondent experience & engagement

- What's the average drop-off rate compared to a traditional human-led interview?
- How does the AI handle unexpected respondent behaviour, emotional responses, or requests for clarification in a way that feels human, not robotic?

Data integrity & non-verbal analysis

- If your tool includes video or audio sentiment analysis, what ethical frameworks and validation studies support your approach?
- Can human researchers step in during live interviews to re-route questioning or override the AI script when necessary?

Scalability & customisation

- How long does it take to deploy a complex new interview script within your platform?
- What language capabilities do you support, and how do you ensure local cultural nuance is reflected in both the conversation and the analysis?

✔ Prioritise platforms that blend automation with flexibility. AI should speed up execution without removing the researcher's ability to guide the process.

Questions to ask about the vendor's AI strategy

Beyond specific features, it's also important to understand the thinking behind a vendor's AI roadmap. Here are some deeper strategic questions to ask:

If the vendor offers only one AI feature (e.g. a co-pilot):

- Why have you chosen to focus here?
- What's your plan for expanding responsibly, and what safeguards are in place to ensure new tools meet the same standards of quality and transparency?

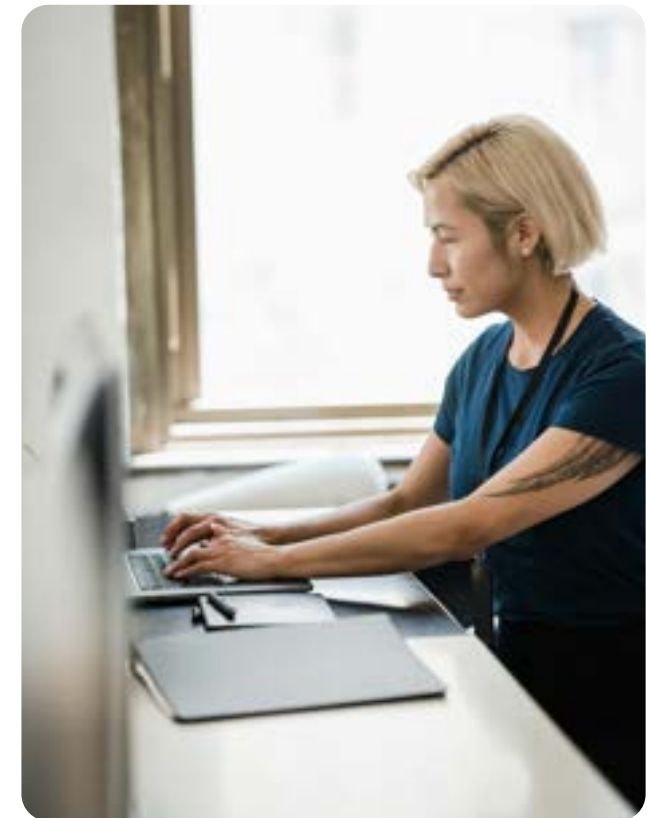
On complexity vs. usability:

- How do you balance sophisticated AI capabilities with practical, measurable utility for researchers?
- Can you share real examples of how your tools save time or improve output quality, not just hypothetical use cases?

On staged rollout:

- Do you have a clear roadmap for gradually introducing AI functionality in a way that integrates with researchers' workflows?
- How do you ensure new tools are tested, transparent and built with methodological integrity before release?

✔ Mature vendors won't promise everything at once. Choose one that shows you a deliberate, principled path toward a fuller AI ecosystem, starting with the features that solve real researcher problems today.



The Attest approach to AI: built for real researchers

The promise of AI in consumer research is immense, but so is the potential for missteps. As this guide has shown, the best tools are not the ones that launch the most features the fastest, but the ones that are built on a foundation of transparency, quality, and real-world utility.

That's the philosophy we follow at Attest.

We're actively developing AI across the full research journey – from design and analysis to synthesis and storytelling – but we're doing it with intention and caution. Rather than racing to tick every box, we're introducing AI where it solves real problems today, not just hypothetical ones tomorrow.

You won't find gimmicks. You'll find tools that are designed to be used, trusted, and built upon. Every new feature we launch is tested, transparent and aligned with research best practice. Our roadmap is shaped by the needs of real researchers, not just by what's technically possible.

How we use AI at Attest

AI at Attest is designed to make research faster, more intuitive and more accessible, without ever replacing human judgement. Our co-pilot, Compass, is already helping users create surveys, refine questions, and uncover insights more efficiently. And there's more to come.

At every stage, we prioritise:

- Transparency: you'll always know what the AI is doing, where it's used, and how it works.
- Data safety and privacy: we never use customer data to train AI models. Ever.
- Control: you remain in the driver's seat. AI helps, but it never takes over.



Want an introduction to Attest, the AI consumer insights engine?

[Book a demo](#)



askattest.com