INSIGHTS · TOOLS AND THE FUTURE

The Future of the Insights Industry

How the insights industry is changing and is likely to change, covering the shift from designed to organic data, online panels and data quality, privacy regulation, experimentation, artificial intelligence and synthetic respondents, and the skills and roles that will matter.

ModQu research article · 2,101 words · 24 references (APA 7th edition)

The insights industry, the businesses and teams that produce knowledge about customers and markets, has changed repeatedly over the past century. Door-to-door interviewing gave way to the telephone, the telephone to the internet, and periodic surveys to continuous streams of behavioural data. Another period of change is now under way, driven by artificial intelligence, new privacy rules and the declining willingness of people to answer surveys. Predicting the future is uncertain, and many confident forecasts about the industry have proved wrong. But several trends are clear enough to examine. This article describes them, considers what they mean for the work and the people who do it, and identifies what is likely to remain constant.

Three eras, and a fourth

Groves (2011) described three eras of survey research. In the first, roughly from the 1930s to the 1960s, the basic methods of sampling and questionnaire design were invented. In the second, from the 1960s to the 1990s, surveys expanded rapidly, helped by the telephone and computers, and response rates were high. In the third, beginning in the 1990s, response rates fell, costs rose and new sources of data emerged. Groves distinguished "designed data", collected for a specific research purpose, from "organic data", generated as a by-product of other activities such as transactions, web browsing and social media, and argued that the future would involve combining the two.

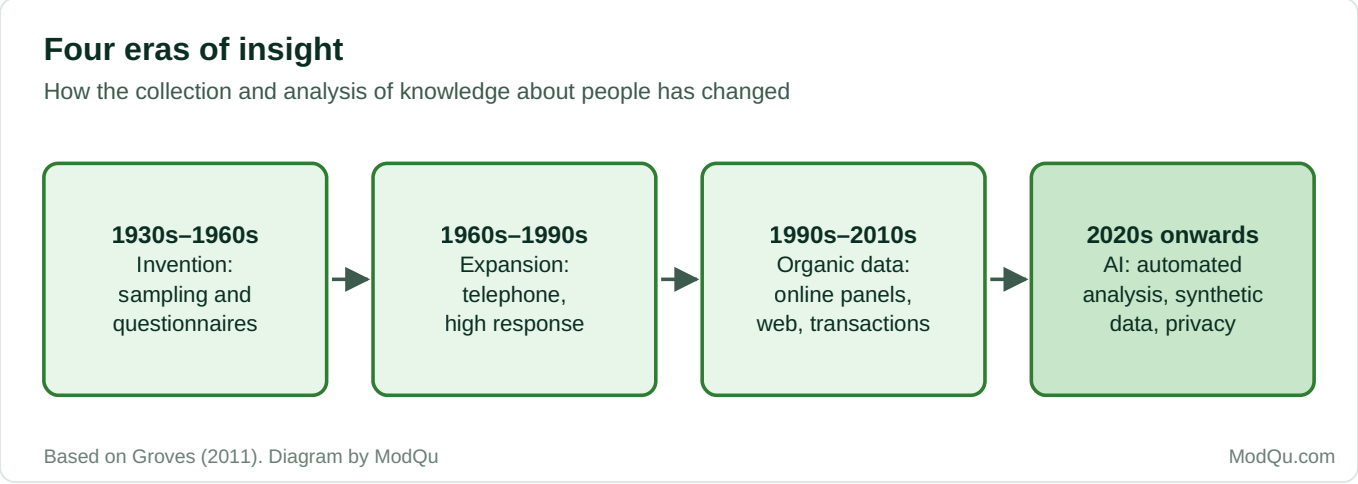


Figure 1. Four eras of insight. Diagram by ModQu, based on Groves (2011).

A fourth era may now be beginning, shaped by artificial intelligence. Its main features are automated analysis of large volumes of text and data, the generation of synthetic data and respondents, and tighter constraints on the collection and use of personal data.

From asking to observing

The balance between asking people and observing them has shifted. Couper (2017) reviewed new developments in survey data collection, including the move to online and mobile surveys and the growing use of sensors, apps and administrative data alongside or instead of questions. Varian (2010) described how computer-mediated transactions allow companies to observe behaviour at a scale and level of detail that was previously impossible. Wedel and Kannan (2016) reviewed how marketing analytics had developed to use these data-rich environments, and argued that the challenge had shifted from collecting data to analysing it well and using it to guide decisions.

Observation has clear advantages: it records what people do rather than what they say, at low marginal cost. But it rarely explains why people act as they do, and it is limited to the behaviour that digital systems capture. Surveys and qualitative research remain necessary to understand motivations, attitudes and the experiences of people who leave no digital trace.

The challenge of asking

Surveys themselves have become harder. Response rates to telephone surveys have fallen to very low levels in many countries. Keeter et al. (2017) reported that response rates to typical Pew Research Center telephone surveys had fallen to around 9%, although careful weighting still produced estimates that compared well with benchmarks on many measures.

Most commercial surveys now use online panels, groups of people who have agreed to take surveys, often for payment. Callegaro et al. (2014) reviewed research on the quality of such panels. Yeager et al. (2011) compared surveys using probability samples with surveys of non-probability online panels and found that the probability samples were consistently more accurate, and that weighting did not always correct the differences. Online samples also face problems of fraudulent and inattentive respondents. Kennedy et al. (2020) documented a "quality crisis" on one widely used online platform, when a rise in low-quality responses, some apparently from people outside the target country using tools to disguise their location, threatened the validity of studies, and showed that screening could identify many of them. Chandler et al. (2019) reviewed the growing range of online panels and the steps researchers can take to improve data quality.

The likely result is a growing premium on data quality: carefully recruited samples, verification of respondents and transparent reporting of how samples were built.

Privacy and regulation

Rules on the collection and use of personal data have tightened. The European Union's General Data Protection Regulation (Regulation (EU) 2016/679, 2016) set requirements for a lawful basis for processing, purpose limitation and data minimisation, and other jurisdictions have adopted similar rules. Goldfarb and Tucker (2011) found that earlier

European privacy rules reduced the effectiveness of online advertising that relied on tracking. Johnson et al. (2023) found that after the GDPR took effect, websites reduced their use of third-party technology vendors, and that the largest vendors gained market share, an unintended effect that increased concentration.

For insight, these changes reduce access to some behavioural data, especially data collected across websites by third parties, and increase the value of data that organisations collect directly from their own customers with consent. They also raise the importance of ethical practice: insight teams will increasingly need to show not only that their data are accurate but also that they were collected and used legitimately (ICC/ESOMAR, 2016).

Experimentation

Many questions that were once answered by asking customers what they would do are now answered by testing what they actually do. Kohavi et al. (2020) described how large online businesses run thousands of controlled experiments each year. Experimentation provides causal evidence that surveys and observational data cannot, and its spread is likely to continue as more of the customer experience becomes digital. The role of research shifts accordingly: from predicting what customers will do towards understanding why experiments produced the results they did, and generating better ideas to test.

Artificial intelligence

Artificial intelligence affects the insights industry in several ways at once.

Analysis. Language models can code open-ended responses, summarise interviews and draft reports far faster than people. Noy and Zhang (2023) found that access to such a model substantially reduced the time professionals took to complete writing tasks and raised the quality of their work.

Synthetic respondents. Argyle et al. (2023) found that a language model could reproduce the patterns of some survey responses when given the characteristics of real respondents, and Brand et al. (2023) found that a model reproduced some well-known patterns of consumer preference. Bisbee et al. (2024) found, however, that synthetic responses varied less than real ones and were sensitive to small changes in how questions were posed. Synthetic data may be useful for early exploration and testing questionnaires, but are unlikely to replace real respondents for questions where the answer is not already reflected in existing data.

Prediction. Agrawal et al. (2018) argued that AI makes prediction cheaper, which raises the value of complementary human skills, especially judgement about what to predict and what to do with the predictions.

Work. Eloundou et al. (2024) estimated that a large share of jobs in the United States contain tasks that language models could significantly speed up, with greater exposure in occupations involving writing and analysis. Autor (2015) argued that automation historically has replaced some tasks while increasing demand for complementary tasks that require judgement, adaptability and interpersonal skills. In the insights industry, routine tasks such as coding, tabulation and report drafting are most exposed; designing research, judging evidence, interpreting findings in context and persuading decision makers are less so.

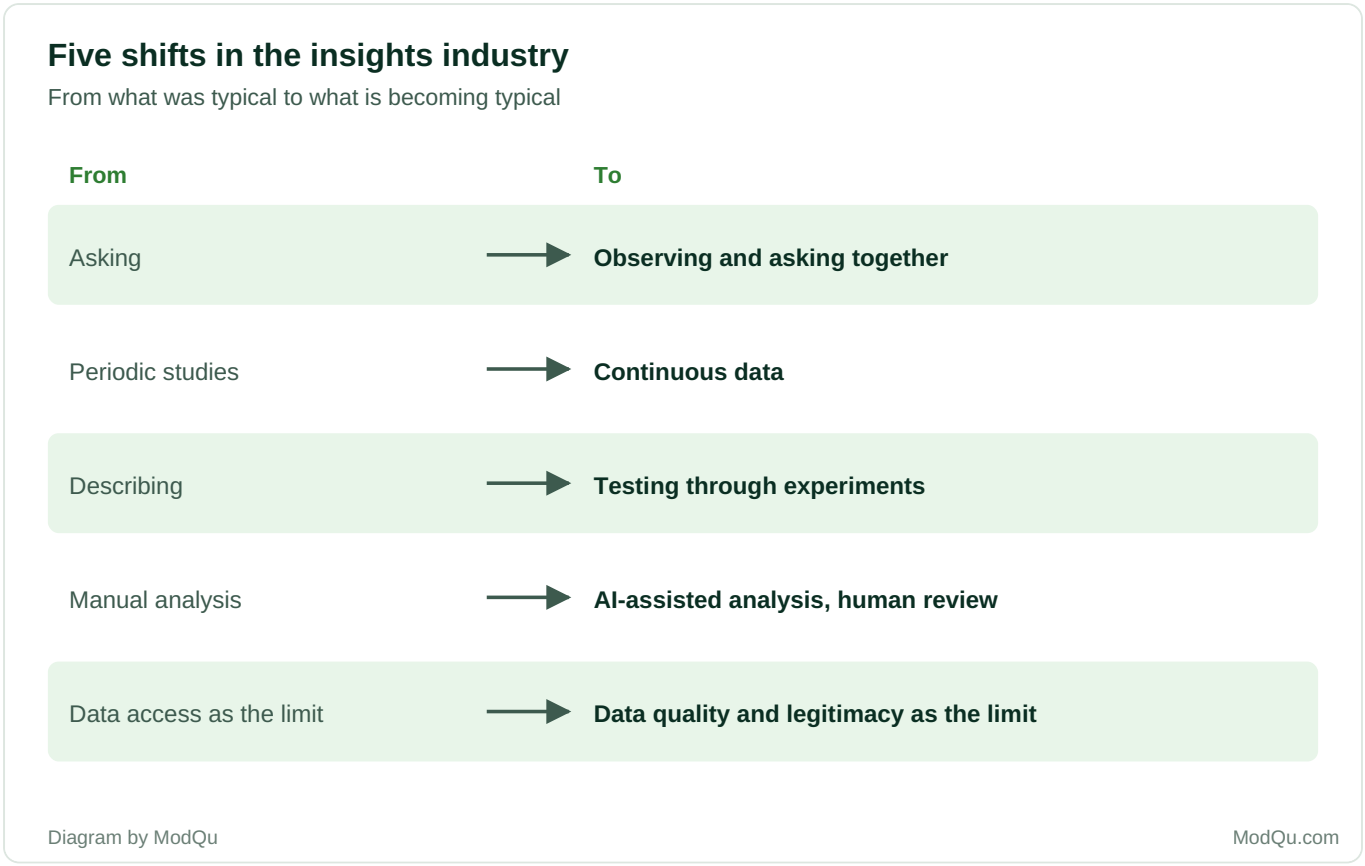


Figure 2. Five shifts in the insights industry. Diagram by ModQu.

Qualitative research in an automated world

Qualitative research, interviews, observation and group discussion, may become more rather than less important. As behavioural data and automated analysis answer more questions about what people do, the remaining questions are more often about why, about meaning and about situations that data do not capture. These are the questions qualitative methods address. AI tools are already changing how qualitative research is done: transcribing and summarising interviews, suggesting themes and searching large collections of transcripts. But the core of the work, building rapport with participants, asking follow-up questions that the participant did not expect and interpreting what is said in context, depends on human judgement. The likely future is qualitative researchers who use AI to handle volume while spending more of their time on the parts of the work that require people.

Quality and trust as differentiators

When data and analysis become cheap and plentiful, the scarce resources become quality and trust. Buyers of research will find it easier to obtain numbers and harder to know which numbers to believe. Suppliers and internal teams that can show how their samples were recruited, how respondents were verified, how AI tools were checked and how conclusions were reached are likely to gain an advantage over those that cannot. Transparency about methods, once a matter mainly for academic research, may become a commercial necessity.

Organisation and skills

As data and tools become more widely available, insight work is spreading beyond specialist teams. Davenport and Harris (2007) described organisations that compete on analytics, in which analytical capability is built into many roles. Moorman and Day (2016) reviewed how firms organise marketing capabilities and argued that the ability to sense and respond to markets depends on how knowledge is shared across the organisation, not only on specialist functions.

The skills likely to matter most combine technical and human abilities: designing studies that produce valid evidence; judging data quality; using AI tools critically; understanding causality and experiments; interpreting findings in the context of the business; communicating clearly; and handling data ethically. Brynjolfsson and McAfee (2014) argued that in an age of increasingly capable machines, people add most value where they combine with technology rather than compete with it.

Three possible futures

Because the direction of change is uncertain, it helps to consider several scenarios rather than a single forecast.

In the first, **automation dominates**. AI tools and continuous data handle most routine insight, and specialist teams shrink to a small core that designs studies, checks quality and handles complex questions. Insight becomes cheaper and faster, but the risk of confident errors rises if checks are weak.

In the second, **quality becomes the premium**. Concerns about synthetic data, fraud and misinformation lead buyers to pay more for verified, transparent research. The industry divides between cheap, automated products and expensive, carefully validated ones, much as other information industries have divided between free and premium content.

In the third, **insight is embedded everywhere**. Analytical skills spread through organisations, decision makers query data directly, and the insight function's main role becomes setting standards, training others and maintaining shared sources of truth.

Elements of all three are already visible. Insight teams that prepare for each, rather than betting on one, are more likely to adapt as the picture becomes clearer.

What insight teams can do now

Several steps help insight teams prepare for these changes.

Audit data sources. Identify which current sources depend on third-party tracking or low-quality samples, and build alternatives based on consented first-party data and well-recruited panels.

Experiment with AI, carefully. Use AI tools for tasks such as coding and summarising, with systematic checks against human work, and record where they help and where they fail.

Treat synthetic data as a hypothesis generator. Use synthetic respondents to explore and refine questions, and confirm anything that matters with real people.

Build experimentation capability. Work with product and marketing teams to test ideas rather than only research them in advance.

Invest in judgement and communication. Train analysts in study design, causal reasoning, critical use of AI and storytelling, the skills least likely to be automated.

Strengthen ethics and governance. Document how data are collected and used, and make ethical review a routine part of projects.

What will not change

Some things are likely to remain constant. Decisions will still depend on understanding why people behave as they do, which requires asking and listening as well as observing. Evidence will still need to be judged for quality, and patterns tested before they are trusted. Insight will still have value only when it changes decisions, which depends on trust between those who produce it and those who use it. And the ethical obligations to the people whose data and words form the raw material of insight will grow, not diminish.

Key takeaways

The insights industry is moving from asking towards combining asking with observing, from periodic studies towards continuous data and experiments, and from manual towards AI-assisted analysis. Surveys face falling response rates and data quality problems, privacy rules restrict some data while raising the value of consented first-party data, and synthetic respondents offer speed but not a substitute for real people. Routine tasks are most exposed to automation; judgement, study design, interpretation, communication and ethics are likely to grow in importance. The core purpose of insight, helping organisations understand people well enough to decide better, remains the same.

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

Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*, *29*(3), 3–30. https://doi.org/10.1257/jep.29.3.3

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.

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton.

Callegaro, M., Baker, R., Bethlehem, J., Göritz, A. S., Krosnick, J. A., & Lavrakas, P. J. (Eds.). (2014). *Online panel research: A data quality perspective*. Wiley. https://doi.org/10.1002/9781118763520

Chandler, J., Rosenzweig, C., Moss, A. J., Robinson, J., & Litman, L. (2019). Online panels in social science research: Expanding sampling methods beyond Mechanical Turk. *Behavior Research Methods, 51*(5), 2022–2038. https://doi.org/10.3758/s13428-019-01273-7

Couper, M. P. (2017). New developments in survey data collection. *Annual Review of Sociology, 43*, 121–145. https://doi.org/10.1146/annurev-soc-060116-053613

Davenport, T. H., & Harris, J. G. (2007). *Competing on analytics: The new science of winning*. Harvard Business School Press.

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science, 384*(6702), 1306–1308. https://doi.org/10.1126/science.adj0998

Goldfarb, A., & Tucker, C. E. (2011). Privacy regulation and online advertising. *Management Science, 57*(1), 57–71. https://doi.org/10.1287/mnsc.1100.1246

Groves, R. M. (2011). Three eras of survey research. *Public Opinion Quarterly, 75*(5), 861–871. https://doi.org/10.1093/poq/nfr057

ICC/ESOMAR. (2016). *ICC/ESOMAR international code on market, opinion and social research and data analytics*. ESOMAR.

Johnson, G. A., Shriver, S. K., & Goldberg, S. G. (2023). Privacy and market concentration: Intended and unintended consequences of the GDPR. *Management Science, 69*(10), 5695–5721. https://doi.org/10.1287/mnsc.2023.4709

Keeter, S., Hatley, N., Kennedy, C., & Lau, A. (2017). *What low response rates mean for telephone surveys*. Pew Research Center.

Kennedy, R., Clifford, S., Burleigh, T., Waggoner, P. D., Jewell, R., & Winter, N. J. G. (2020). The shape of and solutions to the MTurk quality crisis. *Political Science Research and Methods, 8*(4), 614–629. https://doi.org/10.1017/psrm.2020.6

Kohavi, R., Tang, D., & Xu, Y. (2020). *Trustworthy online controlled experiments: A practical guide to A/B testing*. Cambridge University Press. https://doi.org/10.1017/9781108653985

Moorman, C., & Day, G. S. (2016). Organizing for marketing excellence. *Journal of Marketing, 80*(6), 6–35. https://doi.org/10.1509/jm.15.0423

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

Regulation (EU) 2016/679. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). *Official Journal of the European Union, L 119*, 1–88.

Varian, H. R. (2010). Computer mediated transactions. *American Economic Review, 100*(2), 1–10. https://doi.org/10.1257/aer.100.2.1

Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. *Journal of Marketing, 80*(6), 97–121. https://doi.org/10.1509/jm.15.0413

Yeager, D. S., Krosnick, J. A., Chang, L., Javitz, H. S., Levendusky, M. S., Simpser, A., & Wang, R. (2011). Comparing the accuracy of RDD telephone surveys and internet surveys conducted with probability and non-probability samples. *Public Opinion Quarterly, 75*(4), 709–747. https://doi.org/10.1093/poq/nfr020