

 roundtable × **·OBSURVANT.**

Using AI to Fight AI

Balancing the Promise and the
Perils of Large Language Models

Alex Morrison Co-Founder, Obsurvant

Mayank Agrawal CEO, Roundtable

Mathew Hardy CTO, Roundtable

February 2025

Table of Contents

1. Introduction
2. Using AI to Detect AI Content
 - a. Analysing Responses Using Large Language Models
 - b. Characteristics of AI-Generated Survey Responses
3. Using AI to Detect AI Behaviour
 - a. Analysing Behaviour Using Biometrics
 - b. Understanding Typing Patterns
 - c. AI Agent Patterns
4. Conclusion
 - a. Securing the Future of Market Research

1. Introduction

Data quality has been a long-standing issue in the market research sector. The AI revolution has intensified this problem. While fraudulent survey responses have always been challenging, sophisticated AI tools enable bad actors to generate false data at unprecedented scale and sophistication.

Unlike traditional bots, this new generation of AI agents can produce nuanced, consistent responses, simulating authentic personal experiences and answers. These responses – often insightful and well-written – appear genuine at first glance. They present as a seemingly ideal data set. Traditional data quality measures and survey tools which typically flag responses for being short, repetitive, or nonsensical, are largely ineffective against these new forms of synthetic data.

The problem posed by AI-generated responses is particularly acute in survey environments with significant incentives. Online B2B surveys are especially vulnerable. On some projects up to 30% of responses are flagged as potentially AI-generated, either from automated agents or participants copying and pasting content. This phenomenon corrupts research results by introducing fabricated data which can completely distort a dataset if it is not removed.



Online B2B surveys are especially vulnerable, with up to 30% of responses flagged as potentially AI-generated, both from automated AI agents or participants copy-pasting AI content.

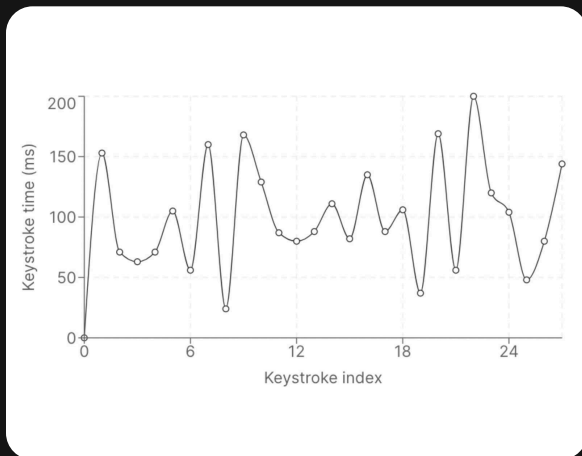
To address this growing issue, Observant and Roundtable have examined how these AI-generated responses corrupt research data, why traditional detection methods fall short, and what new approaches are needed to address this growing challenge. Building on our research, we've identified two key approaches that provide a solution for protecting data quality in modern market research: content analysis and behavioural biometric detection.

Both methods are different in approach but focus on analysing open-ended responses. Content pattern analysis leverages large language models to analyse response content, comparing it against patterns identified in AI-generated text. Behavioural signal analysis examines keystroke patterns to distinguish between human and machine-generated input.

These complementary approaches represent a new paradigm in data quality assurance: using AI to fight AI.

A Dual Approach to AI Detection

Real-Time Behavioural Tracking



Roundtable's real-time behavioural tracking detects unnatural typing and programmatic entries.

Instantly Identify AI Submissions

What's your ideal workspace?

As a language model, I do not have a preference. → 🚫 GPT Detected

Describe your ideal vacation.

one a look for times → 🚫 Gibberish

How do you choose jeans?

I only buy levis but I always wait for a sale → ✅ All checks pass

Roundtable and Obsurvant use language models to instantly classify responses and identify AI submissions.

2. Using AI to Detect AI Content

a. Analysing Responses Using Large Language Models

Despite the growing need, research has shown that identifying AI-generated text is unreliable. Language models are constantly improving, and users can edit or re-prompt to refine content until it appears more human. However, the constraints of online surveys provide a unique advantage for detection – online participants typically write responses quickly and make minimal edits. This creates distinctive patterns that differ from responses that leverage AI.

Building on these patterns, Roundtable collected a large dataset of human and AI-generated survey responses to create a specialised model that distinguishes between human and AI-generated content.

This detection system is part of a comprehensive response-quality framework that addresses both traditional and sophisticated data quality challenges. Beyond AI detection, the model also identifies 'off-topic' responses that don't address the question and 'gibberish' responses that are nonsensical or low-effort, ensuring all aspects of response quality are monitored and maintained.

b. Characteristics of AI-Generated Survey Responses

AI responses show clear stylistic patterns that distinguish them from human responses. These responses tend to be unusually long, with perfect grammar and paragraph structure. They often use the passive voice and other formal writing patterns that rarely appear in survey responses. A rushed survey participant typically writes more casually and makes minor errors.

The content itself also shows distinct patterns. Most notably, AI responses are markedly optimistic, which can bias results and lead to bad business decisions. The responses may also include unnecessary background information or tangential facts, a rare occurrence in human responses.

3. Using AI to Detect AI Behaviour

a. Analysing Behaviour Using Biometrics

While content analysis provides valuable signals for detecting AI-generated responses, sophisticated users can edit or refine AI content to evade detection. To address this possibility, Roundtable and Obsurvant developed a complementary approach by focusing on behavioural biometrics.



AI responses tend to be unusually long, use perfect grammar and paragraph structure, and are written using the passive voice and other formal writing patterns.

This method analyses how text is entered, rather than the text itself, to flag AI responses. Together, the two approaches outlined below create a robust detection system that's difficult to circumvent – even if an AI response can mimic a human writing style, it cannot easily mimic human typing patterns.

b. Understanding Typing Patterns

Human typing patterns have distinct characteristics that emerge from our cognitive and motor processes. Humans typically write in spurts, entering text one character at a time with organic pauses interspersed. Pauses align with the natural flow of the language and are most frequent at the ends of words, sentences, and phrases.

The granularity of this behavioural data is remarkable. The system captures each keystroke and text modification, creating a detailed history of how the response evolved.



AI agents often show unnaturally consistent intervals between keystrokes with no backspaces or other corrections, patterns that emerge from their programmatic functioning rather than the cognitive and physical process of human typing.

Another identifiable pattern emerges from survey participants who use AI to generate and modify responses. These cases typically follow two patterns. Either the participant writes some text and then pastes AI-generated content, or they paste the AI content before editing it. Both patterns create distinctive signatures in the text entry timeline that differ from human composition.

The captured data includes when letters are typed and the timing of backspaces, corrections, pauses, and paste events. These patterns create an informative behavioural fingerprint that complements other biometric markers like mouse movements and device characteristics.

c. AI Agent Patterns

AI agents interact with text fields in markedly different ways from real users. While AI can edit and re-edit content to appear more natural, the underlying mechanics of how it enters text fundamentally differ from human typing. AI agents often show unnaturally consistent intervals between keystrokes with no backspaces or other corrections, patterns that emerge from their programmatic functioning rather than the cognitive and physical process of human typing. Even when users try to adjust the timing or add variations manually, they cannot fully replicate the complex patterns of natural human composition.

4. Conclusion

a. Securing the Future of Market Research

Artificial intelligence's impact on market research is a double-edged sword. While AI has enabled unprecedented advances and new methodologies, it has also created new challenges for data quality. The proliferation of advanced AI tools has made it easier for bad actors to generate survey responses that appear genuine but systematically bias and corrupt research results.

The partnership between Obsurvant and Roundtable demonstrates how AI can be used to address this threat. Combining content analysis with behavioural biometric detection allows us to identify AI-generated responses with unprecedented accuracy.

The content analysis system identifies suspiciously polished or structured responses, while the behavioural biometrics system detects unnatural patterns in text entry. Together, these complementary approaches provide a robust defence against AI-generated responses.

However, this is a cat-and-mouse game. As our models improve, so do the methods bots and other bad actors use. Success requires constant vigilance and adaptation, and Roundtable and Obsurvant continuously update their models and processes to address emerging threats.

The future of market research depends on our ability to harness AI's benefits while mitigating its potential for misuse. As AI tools become increasingly sophisticated, the best defence may be AI itself. By using AI to fight AI, Obsurvant and Roundtable are working to ensure that market research preserves its human voice and remains a strong foundation for business decision-making.

·OBSURVANT

Obsurvant is a dynamic insights agency dedicated to making the latest research technologies and sampling methods accessible to everyone, regardless of budget or experience. Specialising in tailored quantitative polling solutions, Obsurvant places innovation, data quality, and representative audiences at the core of every study. Through a rigorous research methodology and a consultative approach, Obsurvant delivers targeted B2B and consumer respondents globally within tighter timeframes, empowering clients to make confident, data-driven decisions.

roundtable

Roundtable is a behavioural intelligence and fraud detection platform built to protect businesses from deceptive and fraudulent activity online. By combining artificial intelligence, behavioural biometrics, and device fingerprinting, Roundtable provides companies with the confidence to trust their data and make smarter decisions.

Curious to learn more? Get in touch today.



Alex Morrison
Co-Founder, Obsurvant
alex@obsurvant.com



Mayank Agrawal
CEO, Roundtable
mayank@roundtable.ai



Mathew Hardy
CTO, Roundtable
matt@roundtable.ai