How the Fake Web Ruins Your Analytics Data

Uncover the disruptive nature of business data polluted with bots and bad actors

Table of Contents

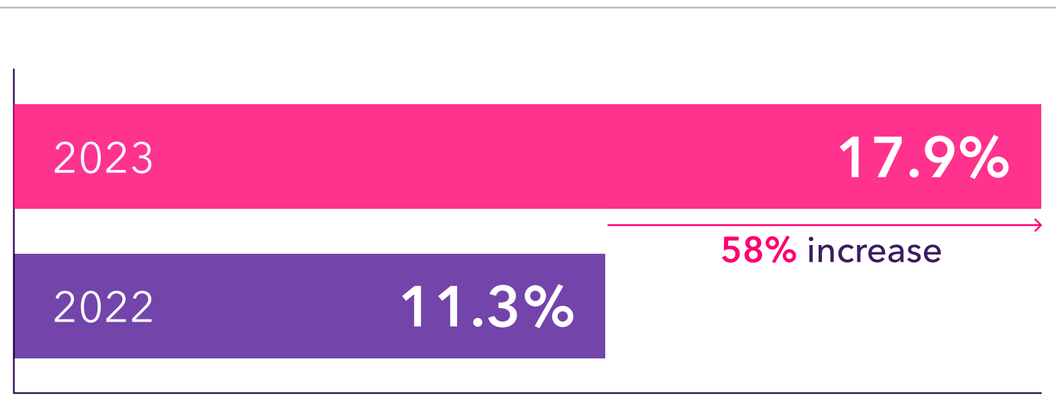
Introduction and Problem Space

Our annual report, The State of Fake Traffic 2024, analyzed 34 billion data points across hundreds of enterprise-level CHEQ clients to uncover how invalid web traffic is growing at an alarming rate. The presence of fake traffic in the digital ecosystem is typically the starting place for data integrity issues further down funnel.

Let's contextualize the pervasiveness of fake traffic for enterprises:

17.9%

of all traffic was fake



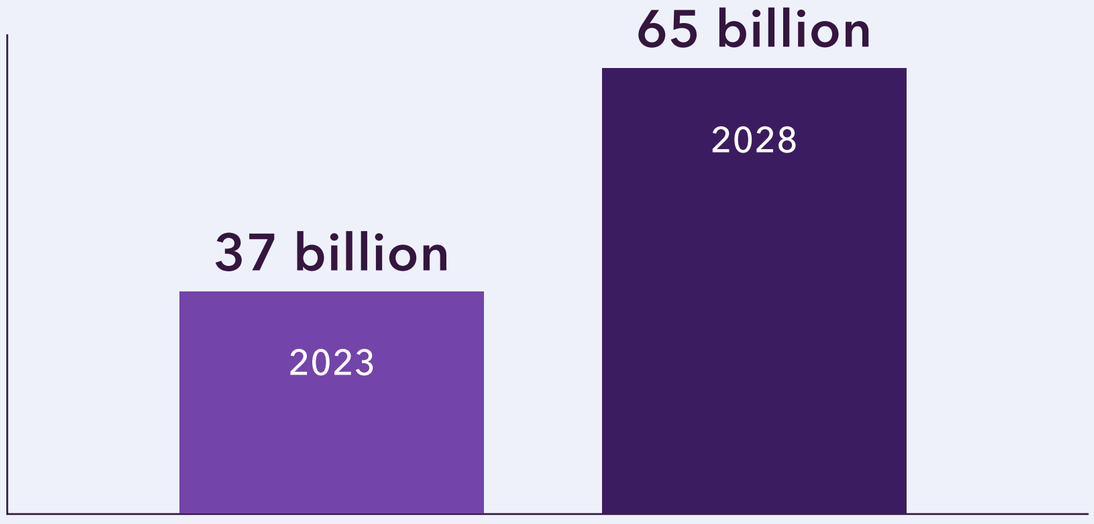
32% year-over-year increase in malicious bot traffic

28% year-over-year increase in all bots

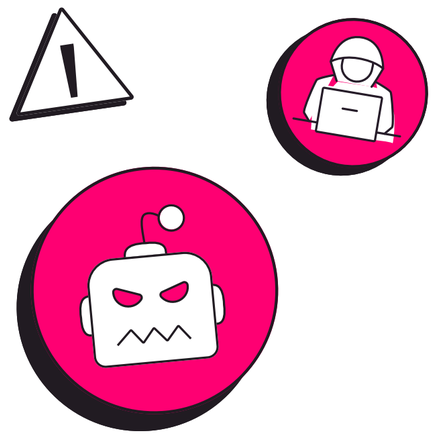
Understanding the scale of non-human interactions or engagement quickly ties data integrity issues back to the fake traffic problem. Clicks, or click-throughs, are foundational data points used to measure interest and engagement for go-to-market teams.

Clicks driven by nonhuman users such as click farms or automation tools, **wreck the denominator of every conversion metric** within paid media campaigns, website engagement reporting, and manipulates top of funnel demand. And, it's only getting worse: Juniper Research estimates fraudulent click sto grow from 37 billion to over 65 billion by 2028.

Projected fraudulent click growth from 2023 to 2028 per Juniper Research:



SAP estimates that the cost of bad data is a $3 trillion problem in the U.S. alone



How can you trust any of your data when it's polluted with faulty signals like fake clicks that ultimately misguide your decision-making and performance? In this report, we'll explore how the fake web has shaken the core of data-driven organizations and what you can do to restore trust and integrity back in your data.

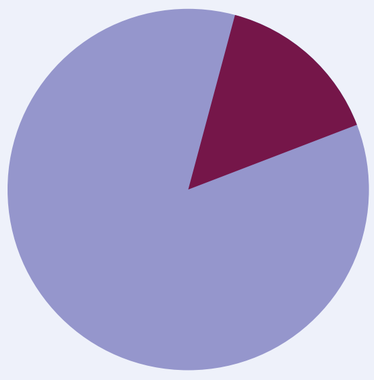
Sophisticated New Threats: Why You Should Care

> **"The top 10 threat types across industries in 2023 included scrapers, automation tools, and malicious bots ... the rise in generative AI makes it easier and more accessible for users to create bots."**
> **– State of Fake Traffic 2024**

Open source software and generative AI allows bad actors and new bots to adapt to various web conditions and mimic human behavior so closely that they're increasingly hard to spot in web analytics, CRMs, and other platforms.

The IAB Bots and Spiders List used by most analytics platforms is a good start to restoring data integrity, but it misses new or uncommon bots that mimic human-like behavior.

That means even with basic exclusions, **85% of fake users can go undetected within your data**.

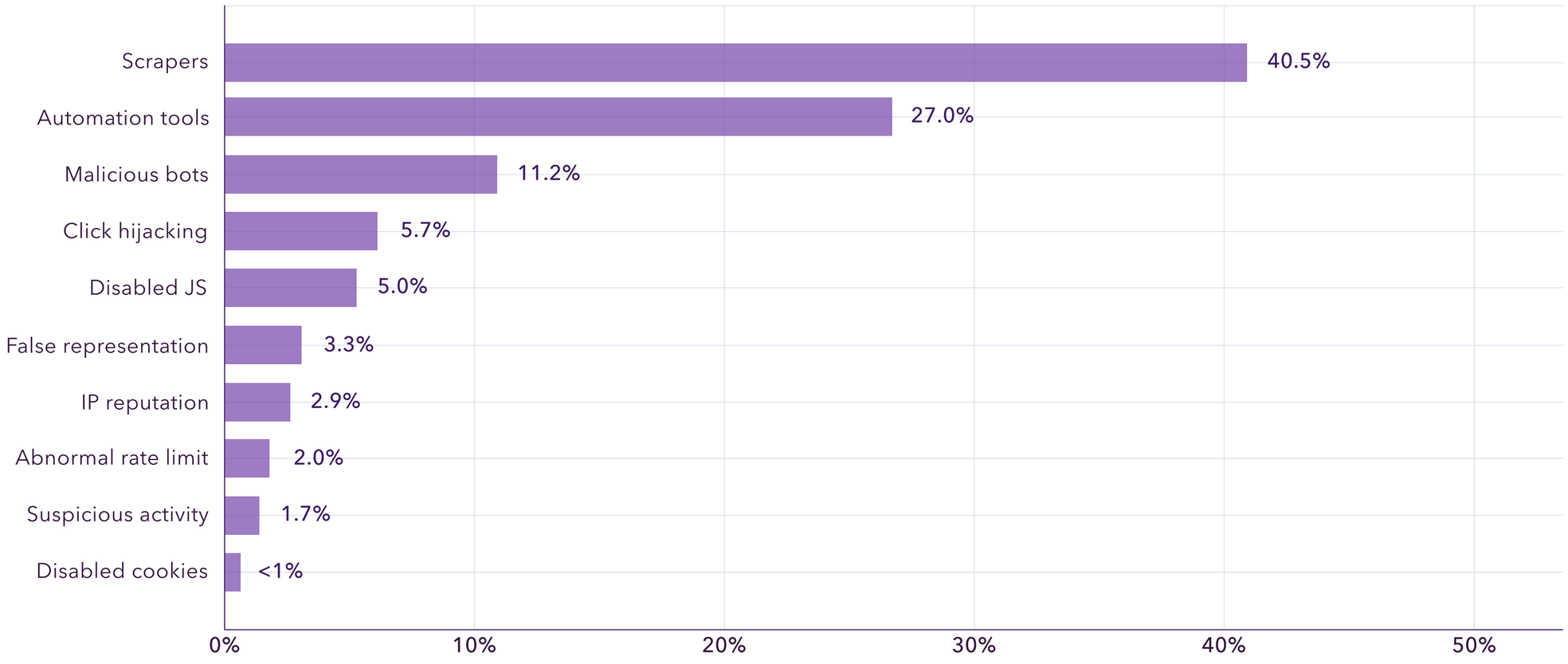


15% of all bots CHEQ detects aren't currently listed on the IAB Bots and Spiders list

Blurring the lines between human-like and human behavior

Today's bots pass security challenges and complete forms with a degree of realism that fool both systems and users, contributing to the **growing distrust** in data and **complicating efforts to remove fake traffic from testing, personalization, and audience profiling**. With nearly 1 in 4 web visits being fake, here are the most common threat types from the State of Fake Traffic 2024 report:

Top 10 threats types you should know that pollute your data



How Bad Data Breeds Mistrust and Ineffectiveness

> **From marketers to finance managers to CEOs, everyone is impacted by the fake traffic and data integrity problem when you consider how many essential business decisions and processes are based on data.**

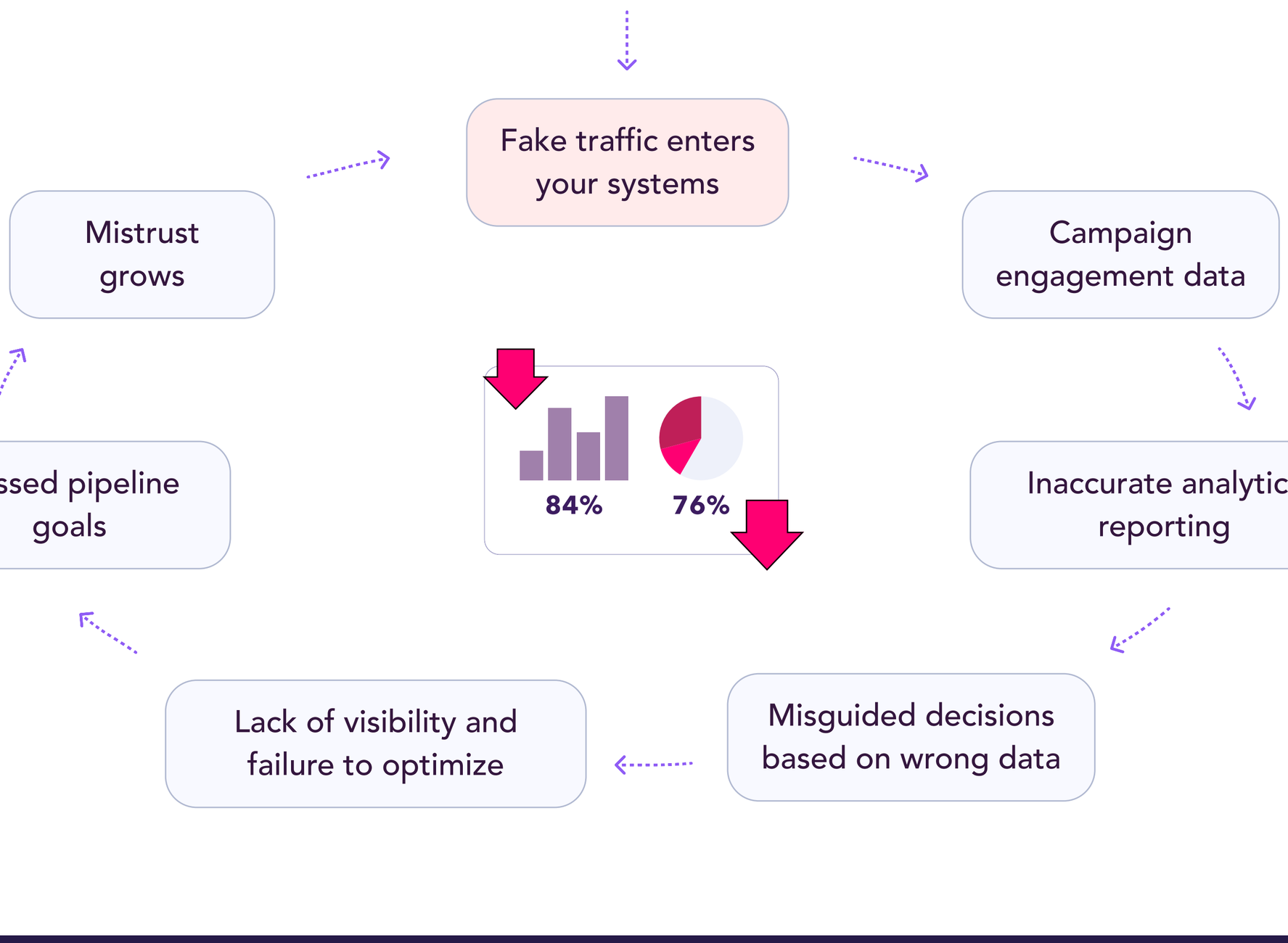
When your organization no longer trusts its data, teams begin to second-guess their decisions. Flawed data also diminishes your ability to clearly see what is happening in your funnel. Eventually, this snowballs into an avalanche of inefficiency and ineffectiveness which undermines the ability to grow the business. Let's review a few real-life examples:

Bad data means wasted resources on ineffective demand gen campaigns

Your marketing team launches a new paid social campaign. After a few weeks, reporting shows a record high of new site visits and ad clicks. At first glance, this looks like a successful ad campaign, so you share the initial performance with the wider org:

- The marketing analytics team recommends double-downing on the campaign
- Your peers optimize new and existing content to match the campaign insights
- Your CMO reallocates resources from other teams to accelerate the campaign
- Your web strategy team adjusts the user experience based on the feedback

However, a deeper dive into the data shows **that nearly 30% of the traffic is from bots**. Your marketing team wasted hundreds of thousands of dollars on an ineffective campaign and missed pipeline goals because polluted data directed you to **place the wrong bet**.



How Bad Data Breeds Mistrust and Ineffectiveness (cont.)

Faulty signals ruin optimization efforts and customer experience

Let's look at another example: At the beginning of the quarter, your web optimization team deploys a 50/50 split homepage design test, where 80% of all traffic enters your inbound funnel. Six weeks later, the team declares the variation the winner based on a 10% lift in form fills compared to the control, and ships this variant as the new default experience.

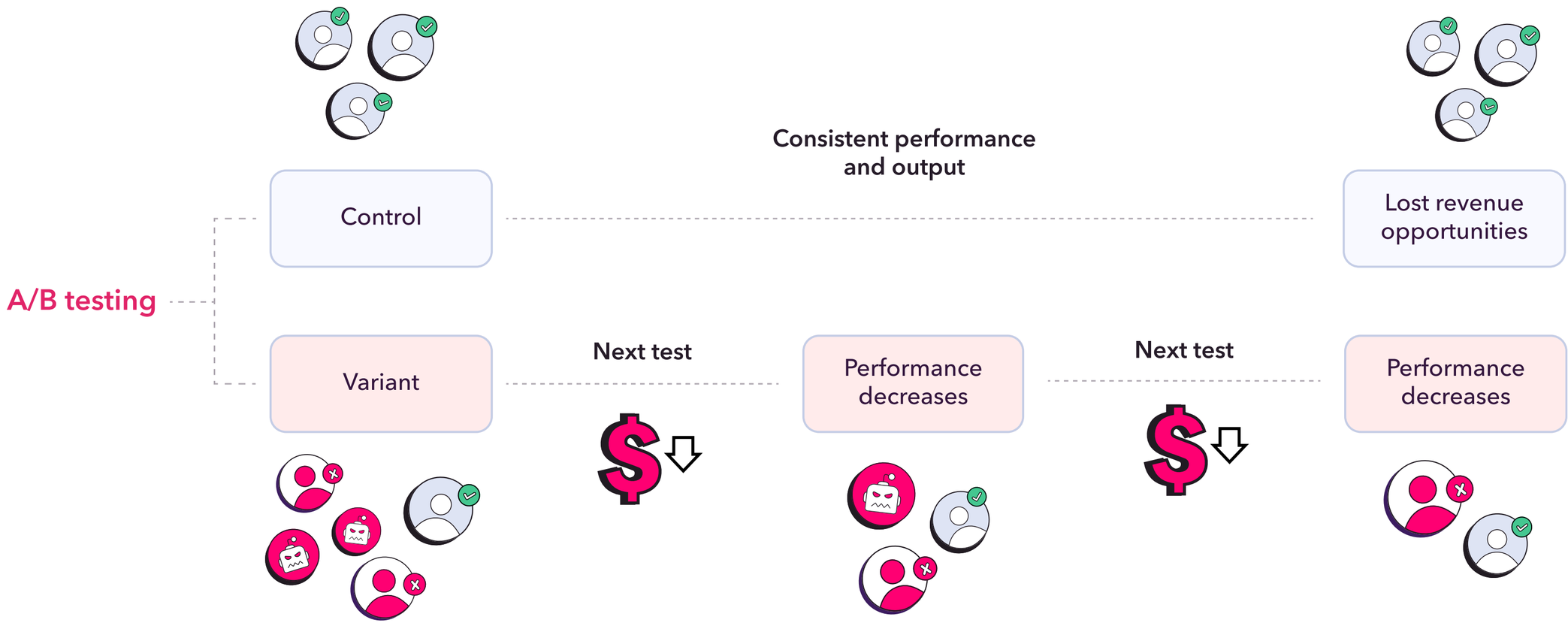
At the end of the quarter, form fill-to-paid users has **actually dropped 20% QoQ**, leaving the team confused and blind to the fact that the rise in form fills were caused by automation tools, not users. **If you had filtered out nonhuman users from the test, the control (or existing) experience should have been declared the winner.**

Without the proper visibility into the health of your funnel, the team scrambles to spin up a flurry of 50/50 split tests based on the ineffective variant, but all this does is introduce more friction for human users:

- Your customer experience is degraded
- The volume of paid users continues to spiral downward after each test
- Fake users are STILL skewing your results along the way, sowing confusion and distrust
- You've wasted months of time, resources and revenue opportunities

This very scenario happens because your testing roadmap and strategy was anchored to the **wrong experience due to faulty signals.**

Removing the noise from your data is critical for growth and providing the best experience for your customers. In the next section, we'll cover how you can protect your data integrity and downstream workflows.



Protecting Your Data Integrity and Downstream Workflows

You need to trust the integrity of your data to move fast and win in today's market. CHEQ Analytics detects fake traffic, surfaces trends, and pinpoints specific issues so that you can trust your insights, and make faster, more accurate optimization decisions.

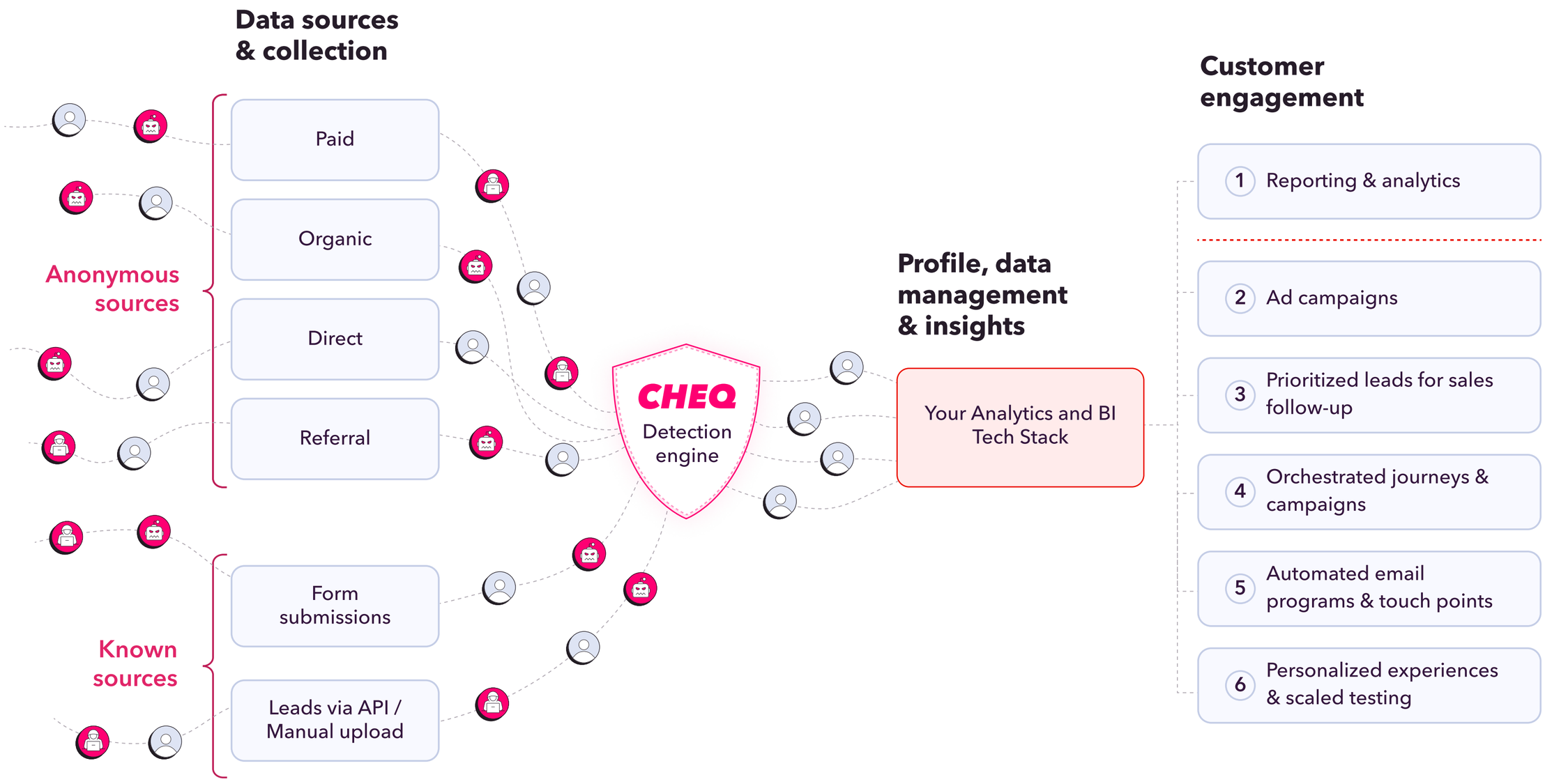
We conduct over 2,000 real-time cybersecurity challenges on each site visitor to verify authenticity and legitimacy. We also perform behavioral analysis of traffic patterns backed by 6T signals to help distinguish between good and bad traffic.

This level of granularity enables you to track invalid traffic patterns across KPIs, regions, and channels over time and dive deep into the impact of various threats on specific campaigns, pages, user journeys, and more.

You can also calibrate detection to your team's specific requirements and integrate with your web analytics and BI systems of record so that you can protect your data where it's used most:

Adobe Analytics · Analytics 360 · amazon S3 · snowflake · API

As we've covered in this report, the implications of fake traffic goes far beyond analytics data. It neutralizes effective lookalike modeling and remarketing audiences for ad campaigns, infiltrates lead automation processes, and compromises testing optimization with faulty signals. The graph below visualizes how these widespread issues are ultimately protected by CHEQ:

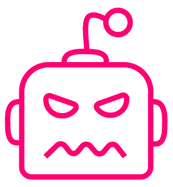


Case Study: Unlocking Better Decisions With CHEQ

A Leading B2B eCommerce Brand Uses CHEQ to Empower their Decision-Making

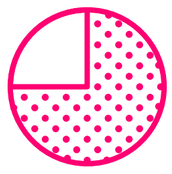
Key Metrics

CHEQ Analytics revealed how the company's decisions were significantly misinformed:



40%

Total web traffic identified as bots by CHEQ



0-50%

Variance in each web page's level of bot traffic

Challenge

A leading B2B eCommerce brand depends on an optimized website experience to drive its prospective customers through their journey to purchase. To facilitate constant optimization, they built out an expensive and sophisticated data infrastructure consisting of technology, people, and processes. This allowed the company to continually monitor, test, and execute changes to its website and marketing strategy aimed at improving its conversion rates... or so they thought.

When inexplicable drops in conversion rates on several of its product pages occurred, the company struggled to explain what caused the fluctuations and, more importantly, how to handle the affected products. Should they discontinue them? Lower their price? No test or adjustment would reveal the answer.

Solution

Upon implementing **CHEQ Analytics**, it was revealed that roughly 40% of the site's overall traffic was made up of bots. To make matters worse, this invalid traffic was asymmetric across each product page, with some experiencing over 50%. This made it impossible for the company to know which of its products, campaigns, or user journeys performed best.

By integrating **CHEQ Analytics** directly into their marketing intelligence systems, they immediately gained full clarity into how invalid traffic was affecting the performance of every element of their web experience.

This enabled the company to finally make well-informed decisions about how best to maximize conversions and increase revenue.

Ready to protect your data? Request a CHEQ Analytics demo today.

About CHEQ

CHEQ is the global leader in Go-to-Market Security. Trusted by more than 15,000 companies, ranging from emerging brands to the Fortune 50, CHEQ protects business-critical digital interactions from malicious, automated, and human-driven threats.

Powered by its unrivaled, context-specific detection engine, CHEQ offers the most comprehensive set of solutions for securing go-to-market operations from threats to business continuity, brand reputation, privacy compliance, and marketing effectiveness. It's why CISOs trust CHEQ, marketers love CHEQ, and more businesses choose CHEQ.

See how bots and bad intent are harming your go-to-market efforts. Request your CHEQ Analytics demo today.

Customers Love Us on G2