Web Analytics in the Age of the Fake Web

How invalid traffic ruins your web analytics, audiences, optimization, and personalization efforts

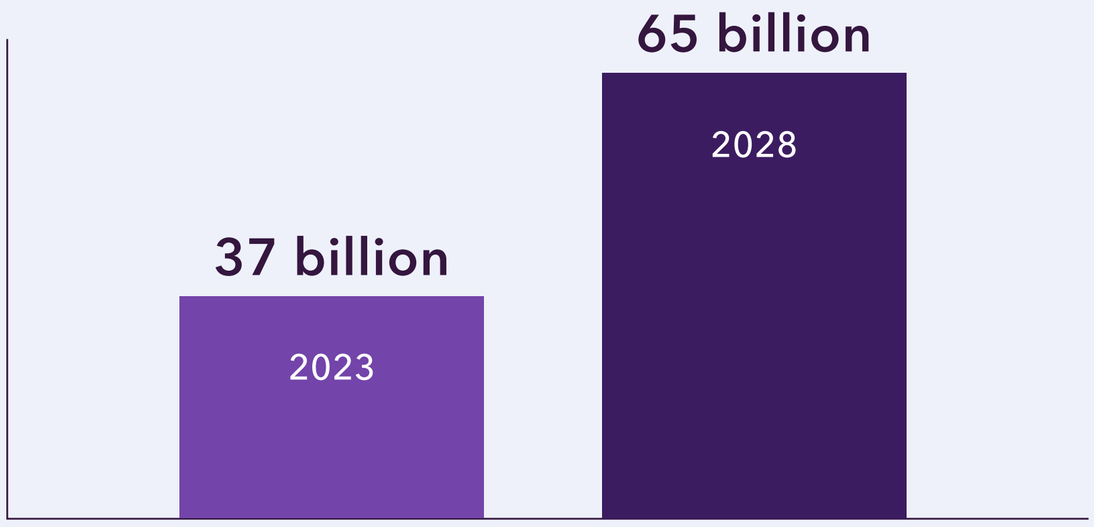
Table of Contents

Introduction and Problem Space

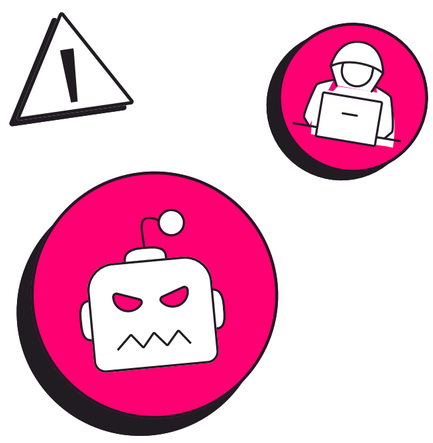
The presence of fake traffic in your web analytics data is often the starting place for data integrity issues further down funnel. Faulty signals from fake traffic infiltrate your campaigns, conversion actions, martech platforms, and ultimately, your decision making.

Understanding the scale of faulty signals from invalid interactions quickly ties data integrity issues back to fake traffic. Clicks driven by click farms or automation tools, **inflate the denominator of every conversion metric** within your campaigns and website engagement reporting, and manipulate top-of-funnel demand. The problem is only getting worse: Juniper Research estimates fraudulent clicks will grow from 37 billion today 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

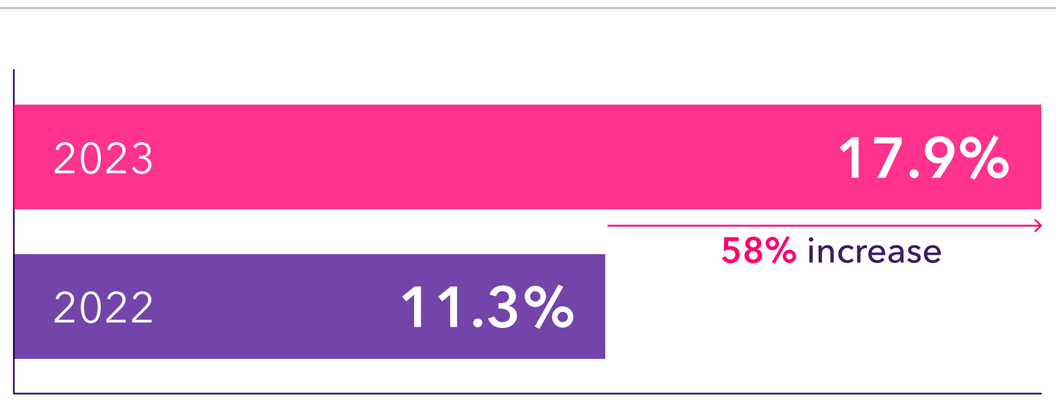


Our annual report, *The State of Fake Traffic 2024*, analyzed 34 billion data points across hundreds of enterprise-level CHEQ clients to uncover how invalid web traffic is growing at an alarming rate.

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

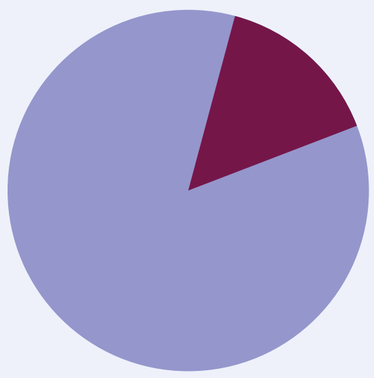
Emerging Threats: Why You Should Care

> **"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."**
> **– *The State of Fake Traffic 2024***

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

The IAB Bots and Spiders List used by most analytics platforms is a helpful resource in restoring data integrity, but it misses emerging and sophisticated bots, especially those that mimic human-like behavior.

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

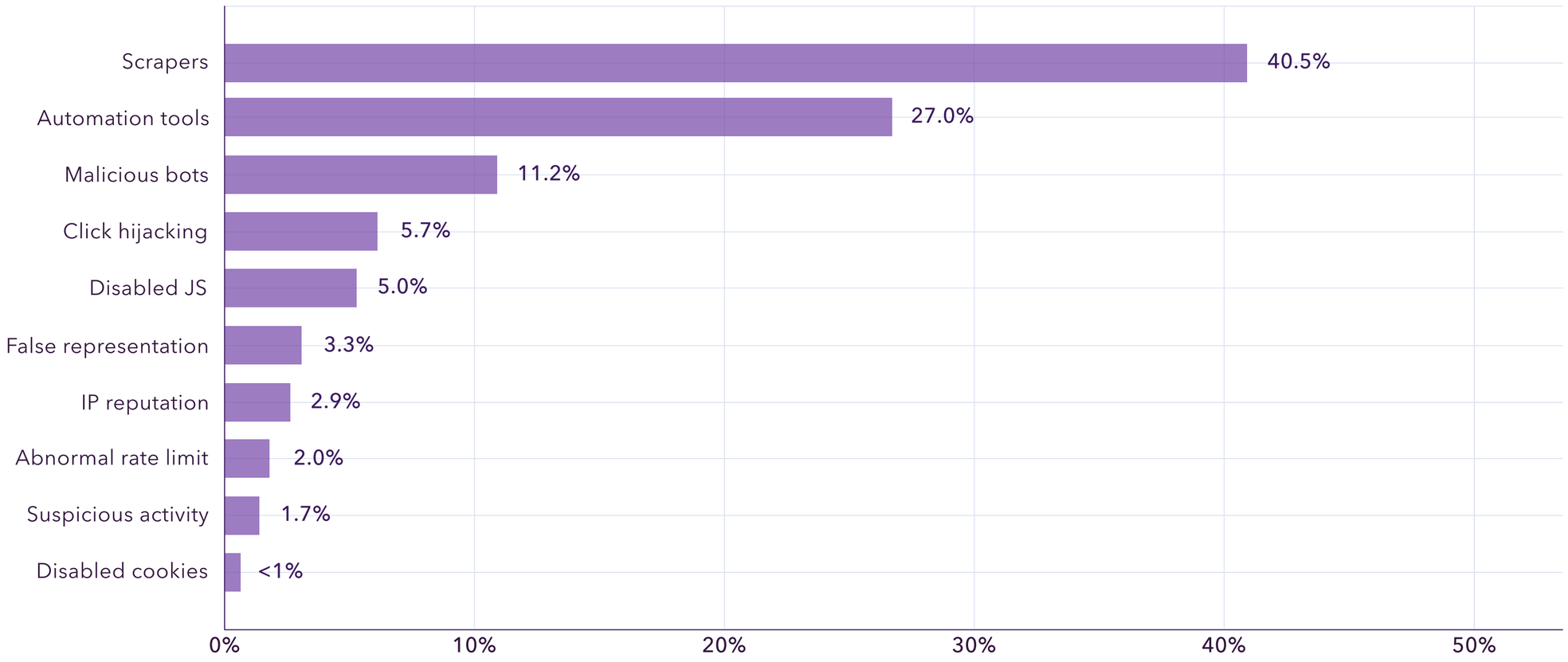


15% of all bots CHEQ detects aren't currently listed on the IAB Bots and Spiders list

Blurring the lines between human-like and human behavior

Today's bots can easily solve basic security challenges and complete unprotected forms with a degree of realism that can fool marketing and analytics leaders, contributing to the 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*:

Top 10 threats polluting web 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 web analytics data.

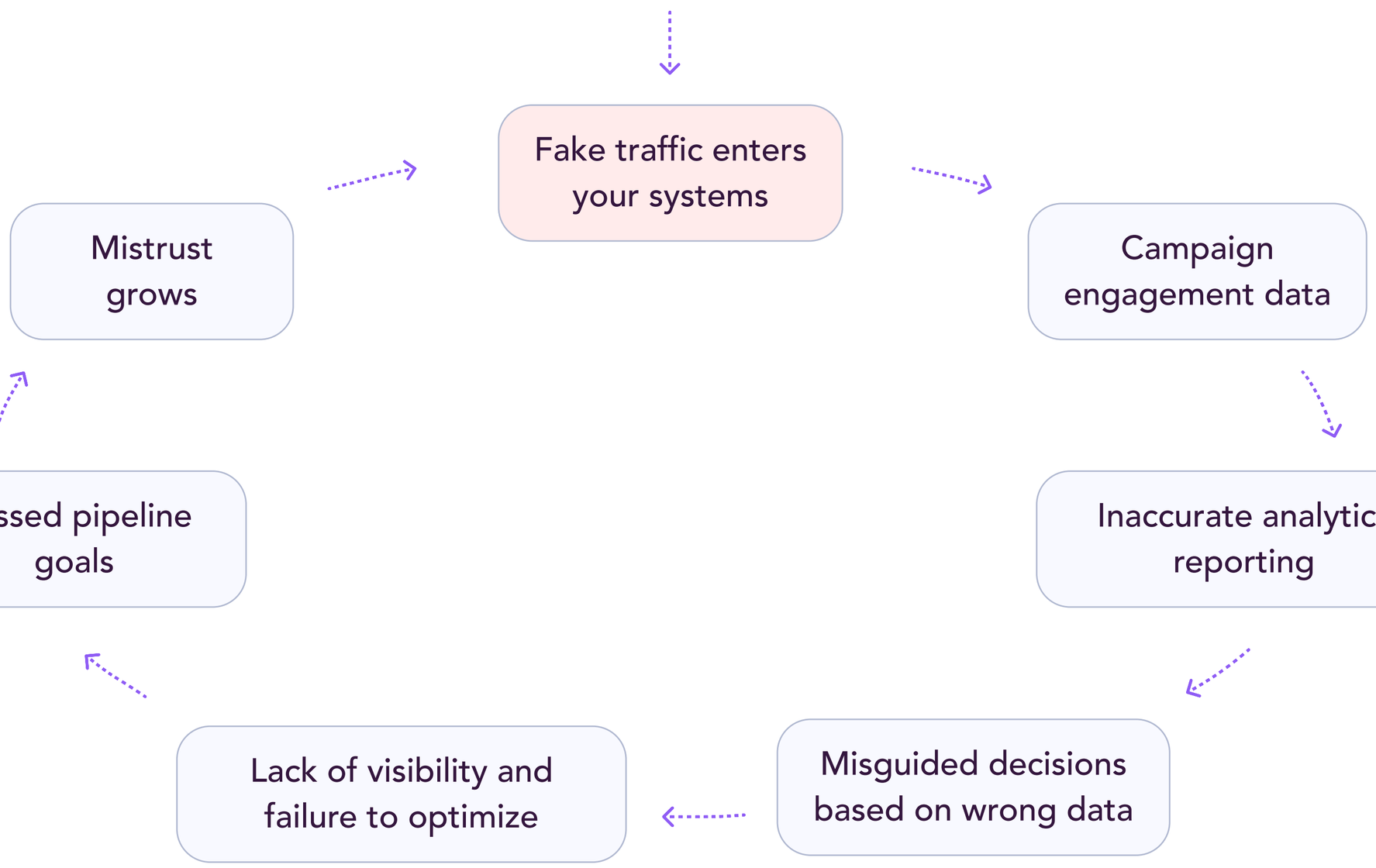
When your organization loses trust in its data, decision-making becomes convoluted. Polluted data also obscures your view of the customer journey, making it difficult to identify what drives meaningful engagement. Over time, this creates a snowball effect of inefficiency and ineffectiveness, ultimately hindering business growth. Let's review a few real-life examples:

Bad data means wasted resources on ineffective acquisition campaigns

Your marketing team launches a new paid social campaign. After a few weeks, reporting shows a record high of new site visits and sign-ups. 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**. You've wasted hundreds of thousands of dollars on an ineffective campaign and missed performance 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 60% of all traffic enters your website. Six weeks later, the team declares the variation as the winner based on a 10% lift in form fills compared to the control, and ships this variant as the new default experience.

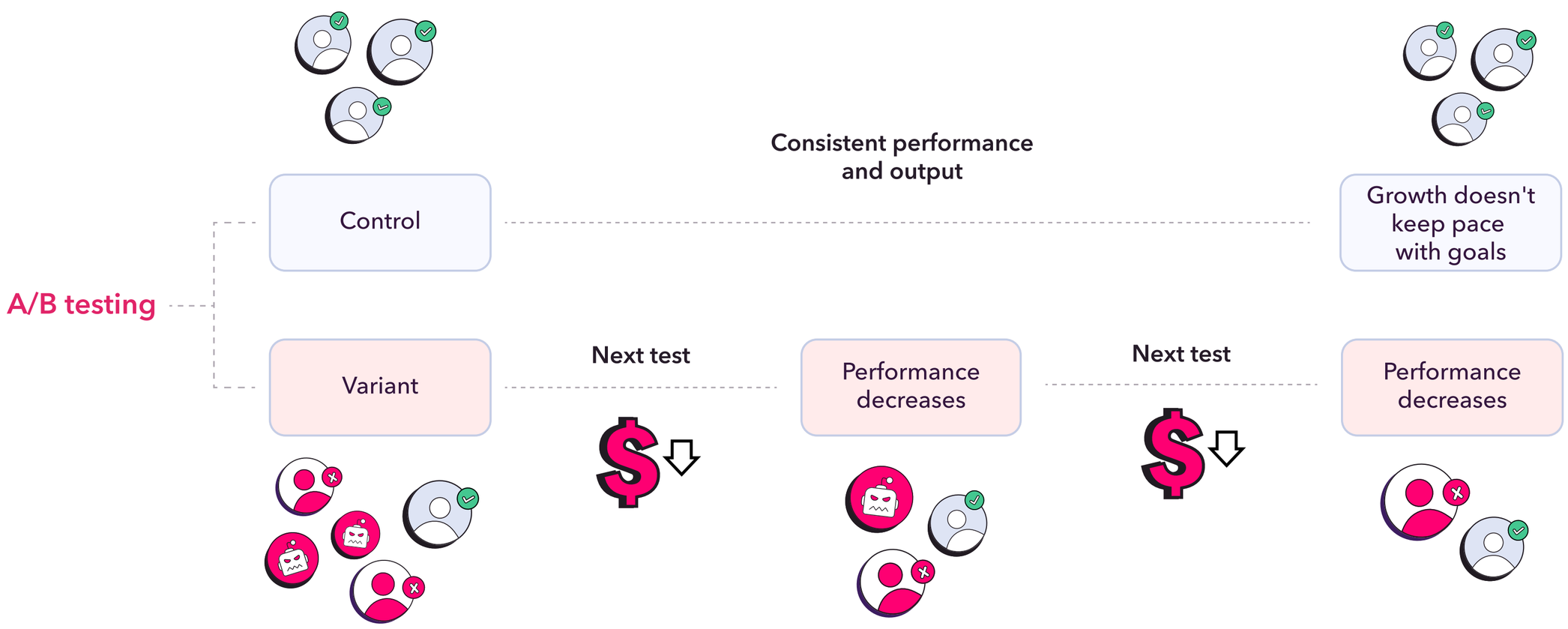
By the end of the quarter, sign-ups-to-customers have dropped **20% QoQ**, and after an arduous analytics audit, the team has no real answers. Little do they know that the rise in sign-ups was caused by **automation tools, not real users**, and had they been able to remove nonhuman users from the test, the control (or existing) experience would have been declared the winner.

Without proper visibility into the efficacy of their engagement metrics, the team then tries to correct last quarter's miss by spinning up a flurry of A/B 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 customers 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

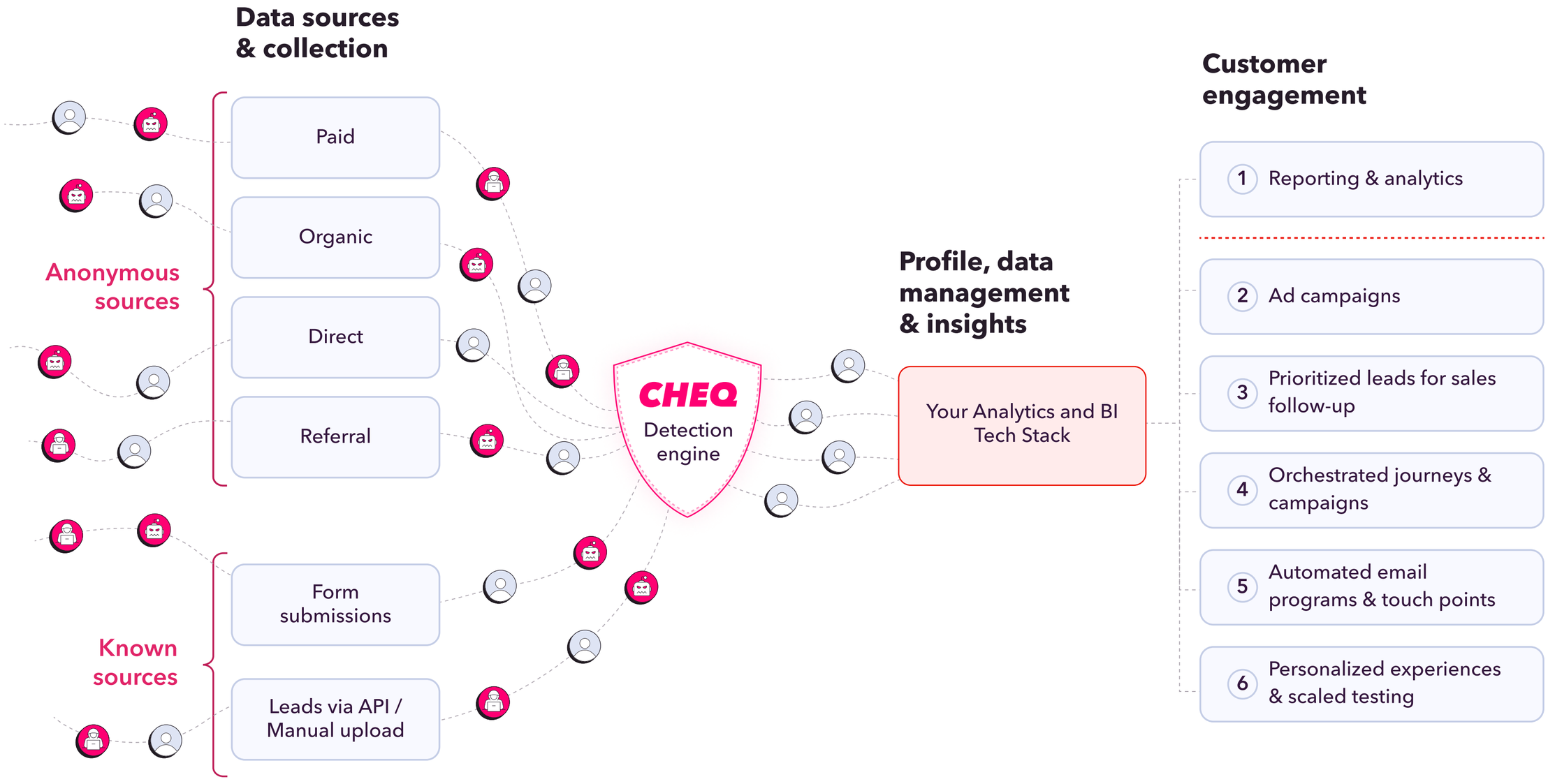
This very scenario happens because your testing roadmap and strategy was anchored to the **wrong experience due to faulty signals.**

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

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 with faulty signals. The graph below visualizes how these widespread issues are neutralized by CHEQ:



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

Backed by 6T signals, CHEQ's detection engine conducts over 2,000 real-time cybersecurity challenges on each site visit to verify its authenticity, distinguishing good from 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.

The CHEQ detection engine can also be calibrated to your unique business requirements and integrate with your web analytics and BI systems of record so that you can protect your data where you work most:

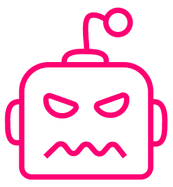
Adobe Analytics · Analytics 360 · amazon S3 · snowflake · API

Case Study: Making Better Decisions with CHEQ

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

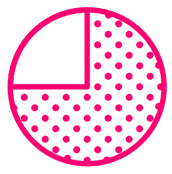
Key Metrics

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

Leader 2024

Easiest To Do Business With 2024

Best Support 2024

Momentum Leader 2024