Jul 2023

Fake Consent

The cost of bad data and why bots are often the cause.

Table of Contents

01 Introduction

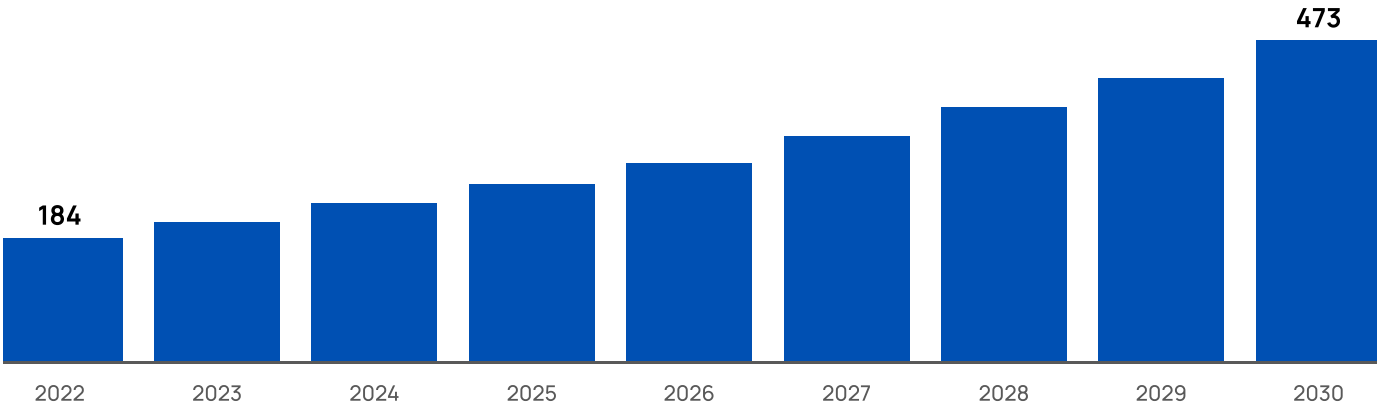
In the world of digitalization, data is often referred to as the "new gold". It serves as the fuel that drives modern businesses and provides the basis for informed strategic decisions.

Companies are continuously collecting more and more data, which is causing the global Big Data market to expand ceaselessly. It was valued at $184 billion in 2022 and is expected to reach $473 billion by 2030[1]. But more data does not automatically mean better decisions or greater success.

The quality of data is critical to business success, but **many suffer from the burden of "Bad Data".** This is **inaccurate, incorrect or incomplete information** that enters into decision-making processes and thus impacts business performance. In addition, **Bad Data creates a variety of hidden costs**, such as damage to corporate image and additional workload.

In this whitepaper, we would like to take a closer look at the impact of Bad Data and show you how bots are related to it and why they are often the cause of a distorted and unreliable data base.

Global Big Data market (USD billion)[1]



More than ever, online marketers need to be careful about which data points they optimize their campaigns on. While in the past it was possible to measure every user and every impression and session almost completely, there are now numerous restrictions and interferences in data collection. This makes it all the more important to have a bot-cleaned data basis in order to be able to determine high-quality and reliable KPIs. Today, clean data is not a luxury, but a necessity.

Daniel Distler
Managing Director
fraud0

(1): Acumen - Big Data Market Size

02 The cost of Bad Data

A study conducted by CrowdFlower shows that **Data Scientists spend around 60% of their time cleaning and organizing data**[2].

A similar study by Forrester found that nearly one-third of analysts spend over **40% of their time reviewing and validating their analytics data before it can be used to make strategic decisions**[3].

This time-consuming process imposes significant costs on companies. It is estimated that **companies in the U.S. alone lose three trillion dollars annually due to inaccurate data**[4].

3 trillion USD

is lost annually by companies in the U.S. alone due to inaccurate data[4].

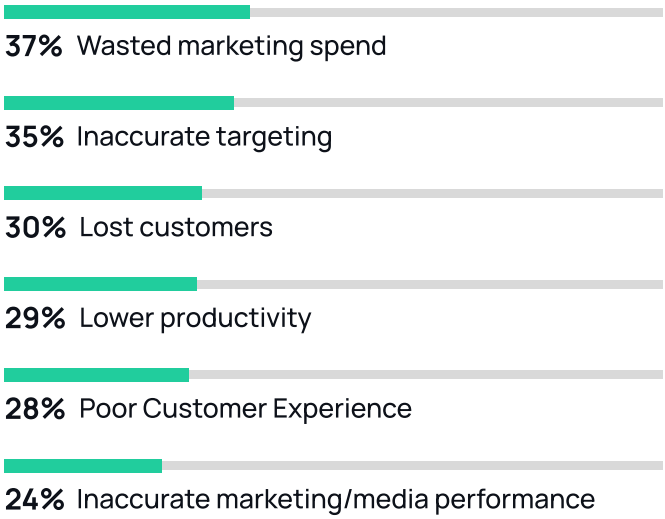
The consequences of poor data quality hit the marketing environment particularly hard. **For every dollar invested in marketing, 21 cents are lost due to poor data quality**[5].

In addition, 87% of all marketing managers surveyed stated that **high-quality data is indispensable for marketing success**[5].

Given the ever-increasing collection and use of data by companies, data quality is becoming a critical challenge that must be mastered.

What Data Scientists spend their time on

What is the experience of decision makers with poor quality marketing data?[5]



(2): Fortune Business Insights
(3): Forrester - Data Performance Management Is Essential To Prove Data's ROI
(4): IBM - The four V's of Big Data
(5): Forrester - Why Marketers Can't Ignore Data Quality

03 The importance of Clean Data

A clean, reliable database - also known as "Clean Data" - can unlock the true potential of collected data and offers a number of benefits in the process.

Marketing and Sales

Probably the most obvious advantage of Clean Data is in the area of marketing and sales. A direct marketing campaign based on high-quality data **reaches the intended target contact at the right time with relevant offers**. This increases the number of sales and the ROI of the campaign.

At the same time, **complete and accurate data enables sales to reliably contact existing customers.**

Improved decision-making

More accurate and reliable data leads to better business decisions. **Clean data provides sharper insights, enables more accurate predictions, and thus optimizes strategic decision making**.

Protection of brand reputation

Last but not least, **high data quality helps protect brand reputation**. For example, companies can be **prevented from contacting deceased persons** or otherwise making embarrassing mistakes that could damage the brand image.

Compliance and data privacy

Compliance with data protection regulations such as the GDPR requires companies to maintain clean data.

Contact forms pose a particularly high risk in this regard: In order to profit from cost-per-lead (CPL) campaigns, bots fill out forms with real user data now freely available on the Internet due to various data leaks in the past. As a result, companies find data of real people in their systems, but these people have often never heard of the company in question.

Companies thus run a major data protection risk if they do not protect themselves against such bot attacks.

By avoiding mistakes and excluding bots, companies can prevent fines and the associated damage to their reputation.

Overall, investing in Clean Data can bring significant benefits in many ways and help companies **realize the full potential of their data**.

But the challenge lies in collecting and maintaining clean data - a task that is more complex than ever in the current digital landscape.

04 The first step to Clean Data - Correct data collection

The General Data Protection Regulation (GDPR) has a profound impact on the way companies collect and process data. A central aspect of this regulation is the **consent to process personal data**. It forms a first essential touchpoint for a correct data basis in companies.

On websites and in apps, this consent is often obtained via so-called cookie banners. These are small frames that inform users that various technologies are being used to collect and process data. Users have the option of giving or rejecting their consent.

However, **if consent is already falsified through the use of bots**, this can have far-reaching consequences. **Incorrect data then runs like a thread through all systems** - from analytics to retargeting to customer relationship management (CRM).

This ultimately leads to distorted results and inefficient marketing activities.

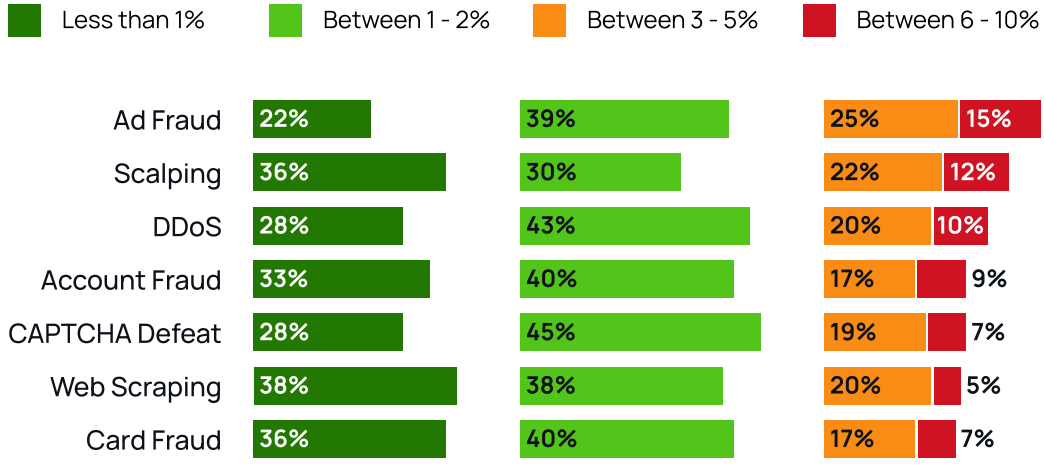
This challenge is compounded by the fact that the number of companies regularly affected by ad fraud in various forms is alarmingly high.

Currently, about **65% of companies are affected by ad fraud attacks at least weekly - 13% even daily**[6]. Ad fraud is fraudulent activity aimed at artificially inflating the numbers of advertising campaigns and thus providing fraudsters with valuable advertising budget.

Given this background, the importance of a clean, bot-cleaned database becomes clear - as does the essential role of correct acquisition of user consent to ensure this.

Financial impact of bot attacks[6]

What percentage of revenue was lost due to the corresponding type of attack?



(6): Forrester - State Of Online Fraud And Bot Management

05 Interaction of bots with cookie banners

The increasing sophistication and complexity of bots has ushered in a new era in the world of digital fraud. In this, bots are not only used to simulate website traffic or steal personal data, but **they are also programmed to interact with cookie banners**.

Currently, Internet traffic generated by bots is over 57%[7] worldwide - with a strong upward trend!

The reason for bots interacting with cookie banners is the monetary interests of the fraudsters who control these bots. In order to maximize the click price (CPC) or the cost per mille (CPM), bots try to collect as many cookies as possible. This **gets them into more expensive retargeting campaigns** and allows them to maximize their operators' profits.

This process is also known as retargeting fraud and poses a serious threat to businesses that rely on online marketing.

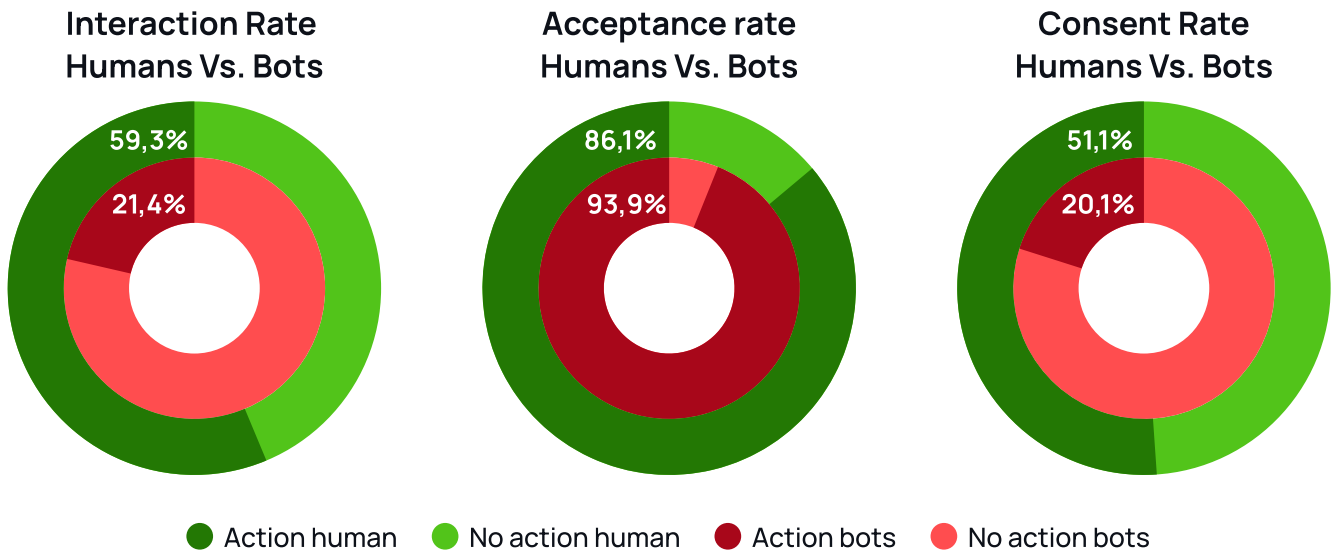
40% of companies lose between 3 - 10% of their revenue due to bot attacks. The impact on advertising budgets is equally frightening: **advertisers lose over $100 billion annually to ad fraud**[8].

Real-world experience shows that bots interacting with cookie banners is widespread. In our analyses, we found that **bots have an average interaction rate of over 20% with cookie banners**. Even more surprising is the acceptance rate: here, **bots accept cookie notices in 90% of cases on average**.

Overall, it can thus be said that **1 in 5 bots intentionally accepts all cookies via a cookie banner in order to be included in retargeting lists**.

Companies must face this challenge and take appropriate measures to counteract the negative consequences of bot traffic and ensure a clean database.

1 in 5 bots accepts all cookies intentionally via a cookie banner to be included in retargeting lists.



(7): Aggregierte Daten eMarketer, WSJ, Group M, Juniper Research, Dr. Augustine Fou & fraud0

(8): Juniper Research - Digital Advertising Spend Lost To Fraud

06 How retargeting fraud works

Retargeting has established itself as an effective strategy for reaching users who have already interacted with a website (for example, by viewing a product in an online store) again with personalized ads and thus increasing the conversion rate.

It has been proven that **retargeting campaigns are clicked on more often and convert better than traditional display campaigns**[9], which is why companies usually also offer a higher CPM for them.

But what happens **if the user is not actually a human, but a bot**? Since retargeting is predominantly based on cookies, bots only need to accept and save the corresponding cookies in order to be included in the respective retargeting campaign from then on.

In the following sections, we will shed light on how retargeting fraud works and explain why fraudsters are particularly interested in their bots imitating human behavior and thus being included in retargeting lists.

Step 1: Creation of fake websites (MFAs)

Fraudsters create **fake websites, so-called made-for-advertising sites or MFAs**, with copied content and register them with various advertising networks (e.g. Google Display Network). This allows them to **place ads on their websites and generate revenue**.

Nowadays, the process of creating such fake websites is **largely automated and takes only a few seconds**.

Step 2: Development of bots

Fraudsters program bots with instructions to behave like real users on a website. This includes **faking technical characteristics** such as the user agent string or IP address, but also many on-site factors such as navigating the website and, in particular, **interacting with and accepting an existing cookie banner**.

The bots change many of these characteristics randomly with each website visit. For example, the bot changes the IP address, the time spent on the website, and the selection of subpages it visits on each session.

Step 3: Collecting cookies

The programmed bots are released onto the Internet with the **aim of collecting as many cookies as possible from high-value industries** (such as insurance, legal services, etc.). By targeting the collection of cookies from high-value keywords and industries, the fraudsters try to **maximize their revenue**.

Step 4: Returning to the MFAs

After the bots have visited hundreds or even thousands of websites in this way, **they return to the fake MFAs**. There, the ad banners of the previously visited companies are already waiting for them, as the **bots have now landed in numerous retargeting lists**. The bots then click on the ad banners and with each click, the fraudsters earn money.

(9): Invesp - Ad Retargeting in Numbers

07 An intelligent approach to fraud detection and data cleaning

Given the increasing threat of bots and the resulting problems and costs, it is essential for companies to take appropriate measures to protect themselves and ensure a reliable database. An effective approach to this is to use specialized fraud detection and data cleaning software.

Software for detecting bots and ad fraud offers a comprehensive solution. Not only can it distinguish real human traffic from bot traffic, but it can also **exclude the detected bots from retargeting using a Negative Audience List** and block the IP address of this invalid traffic, e.g. in Google Ads, in real time using the corresponding API.

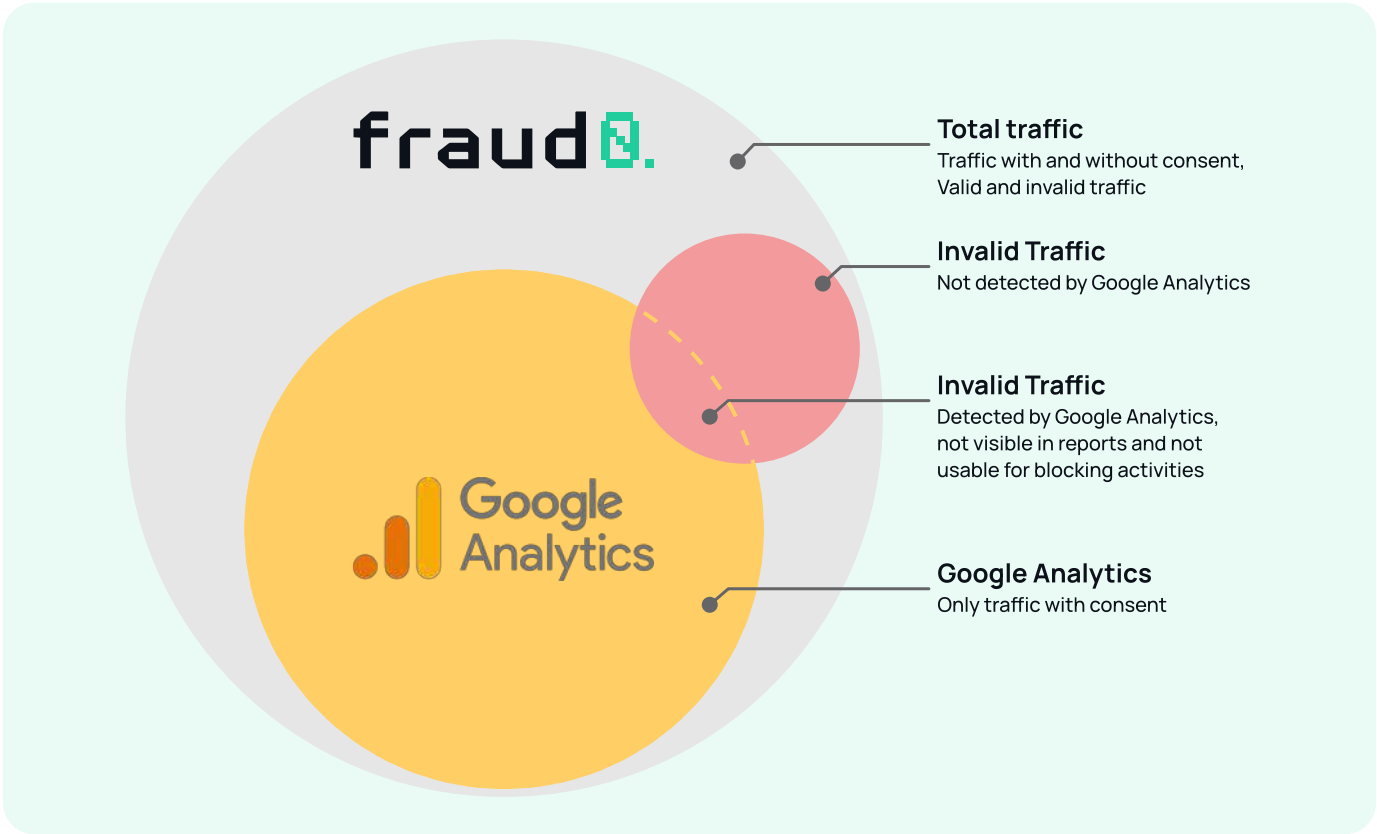
This leads to cleaned CMP statistics and thus eliminates one of the main sources of "bad data".

Another key advantage of such software is its compliance with the General Data Protection Regulation (GDPR). The software usually runs on the legal basis of legitimate interest and thus **enables GDPR-compliant tracking of 100% of traffic**.

In contrast, traditional analytics tools fall under the legal basis of consent. Since many visitors ignore the cookie banner, valuable data from these visitors is lost.

In summary, fraud detection software provides companies with a **trustworthy, real, and reliable data base** that is currently often lacking.

As a result, companies can reduce the cost of "bad data," increase their marketing efficiency, and minimize their privacy risk. All of these benefits ultimately lead to stronger growth and increased competitiveness in the digital marketplace.



08 About fraud0

With fraud0's AI-powered bot & ad fraud detection, you get full visibility into your bot traffic as well as fraudulent interactions of bots with your CMP and ads.

Prioritize real people in the future and make important decisions based only on real, valid data.

Your benefits from fraud0

Protection of your ad budget

Bots and fake users are automatically prevented from exhausting your advertising budget.

Clean retargeting lists

Excluding bots and fake users also saves you a lot of money in retargeting.

Clean overall data

Exclude bots and fake users from your CMP, analytics and CRM data and make business decisions based only on clean data.

Sign up for 7-day free trial