DATA PROVIDER. COM

Structured web data for better equity research

Contents

Introduction

Managing multiple sources of information is crucial for effective investment decision-making and earning active returns. Structured web data allows you to systematically monitor technology use and the companies behind them. You can see which technologies are experiencing fast adoption and which ones are losing traction on the web. In this report, we'll share four concrete examples of how you can leverage web data to gain unique insights into private and publicly traded companies:

Track growth patterns of select technology

Monitor connection data to identify visitor and process traction

Identify how corporations use technology through SSL information

Track companies' assets with Reverse DNS

Web data is among the most sought-after data in the alternative data space, due to its ability to offer diverse and timely insights applicable across a wide range of industries. In this handout, we share how you can use web data to gain unique insights into how companies perform online. Using several case examples and a variety of indicators, we highlight the diversity of metrics and insights that can be obtained from web data.

At Dataprovider.com, we transform the web into an ever-growing structured database. Every month we collect data from more than 700 million hostnames worldwide and structure the data into over 200 data fields, from technical to marketing information, such as hosting, content management system, estimated number of web visitors and many more.

You can trust the integrity of our data as we can always trace it to its source: we don't use any third-party data (except for our Traffic Index). Additionally, all our processes, procedures and policies comply with the European General Data Protection Regulation and with the California Consumer Privacy Act.

Track growth and decline

In the world of website and content delivery, **Cloudflare, Inc. (NYSE: NET)** is a well-known name. The company offers a range of services, from content delivery networks to domain name system (DNS) management, to improved website performance and security without adding hardware, installing or modifying existing software. Cloudflare's DNS management includes features like load balancing, which can distribute traffic across multiple servers to prevent overloading, and DNSSEC, which adds a layer of security to the DNS system to prevent spoofing and other attacks.

With our data, we can track several of Cloudflare's core products and services, including how many domains are registered through their domain registrar, use their free SSL product, and rely on their DNS or web analytics tools.

In Figure 1, you can see the development of Cloudflare DNS over the past 24 months. As of March 2023, 12.1 million active domains were using Cloudflare DNS. However, this graph contains both the free and premium versions. As an investor, you may be more interested in how many domains make use of Cloudflare's paid offerings.
In Figure 2, we see the development of Cloudflare's interested in how many domains make use of Cloudflare's paid offerings. In Figure 2, we see the development of Cloudflare's Premium DNS, which shows an accelerating decline that started in the second half of 2022.

Figure 1: Domains using Cloudfare DNS

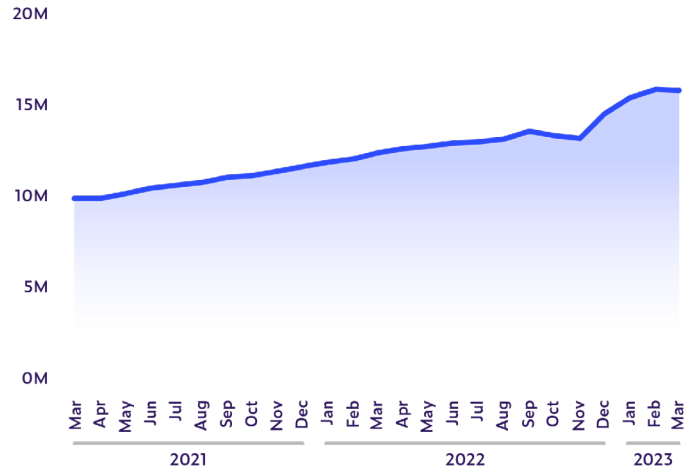
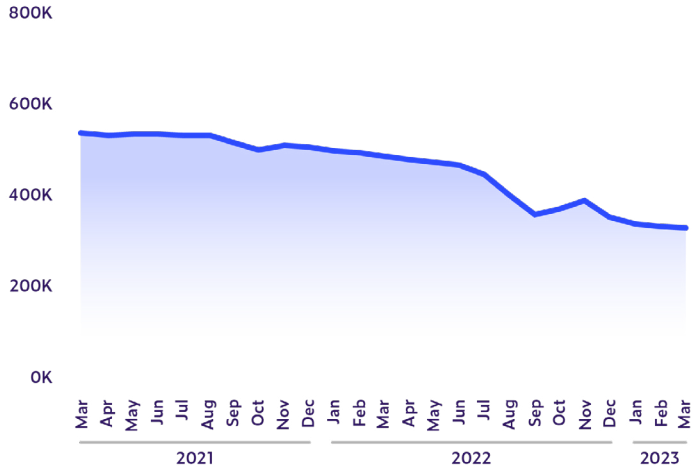


Figure 2: Domains using Cloudflare Premium DNS



To see what is going on, we compare the websites that no longer use Cloudflare Premium DNS with the websites that added it (see Figure 3). With a few exceptions, the percentage of websites that no longer have Cloudflare Premium DNS is consistently higher than the percentage of sites that add this product.

So, if Cloudflare Premium DNS is losing hostnames, perhaps other DNS services may be growing. In Figure 4, we compare the number of domains utilizing Cloudflare Premium DNS with those of **Microsoft's Azure DNS (NASDAQ: MSFT)**. While Cloudflare's Premium DNS is decreasing, we see a steady increase in the number of domains utilizing Azure DNS.

To identify domains that use Cloudflare DNS, Cloudflare Premium DNS or Microsoft Azure DNS, we scan our database of 700 million hostnames and filter out relevant records using information from the DNS TXT (text) record, the DNS cname (Canonical Name) record as well as the DNS NS (Name Server) domain.

Figure 3: % churn & newly added domains using Cloudfare Premium DNS

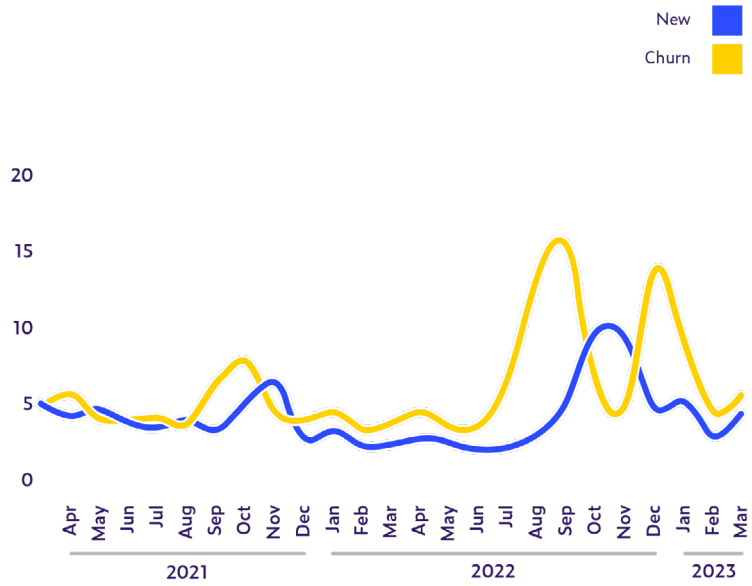
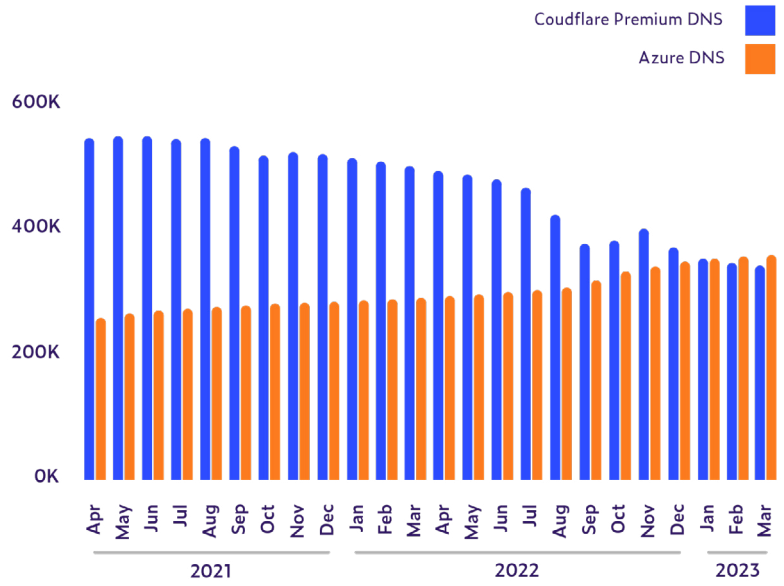


Figure 4: Installments Cloudflare Premium DNS and Azure DNS



Timely connection data

In this section, we will focus on online payment service providers **Adyen N.V. (AMS: ADYEN)** and **Stripe Inc**. Both companies specialize in mobile and web payment systems for e-commerce companies. Building on what we showed in the previous section, let's first have a look at the number of domains worldwide that use payment platforms provided by Adyen or Stripe. Figure 5 summarizes installments over the past 24 months. As you can see, Stripe vastly outnumbers Adyen. Yet, for payments, the number of transactions that occur is more important than the number of websites a payment service provider is serving.
In Figure 6, we compare the two payment providers based on the number of daily connections between January and March 2023. As you can see, despite being available on far fewer websites, the connections to Adyen's checkout URL are generally higher than to Stripe's checkout, suggesting this platform is likely being used on websites that process a higher volume of transactions.

Figure 5: Installments Adyen and Stripe

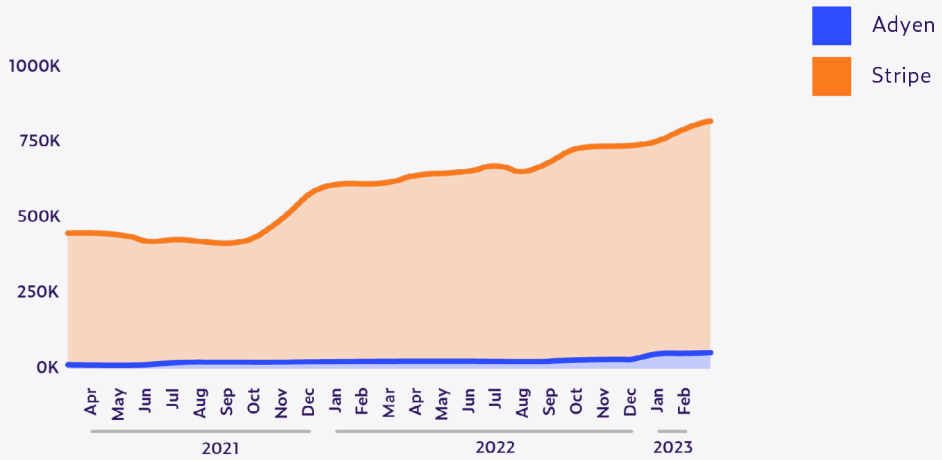
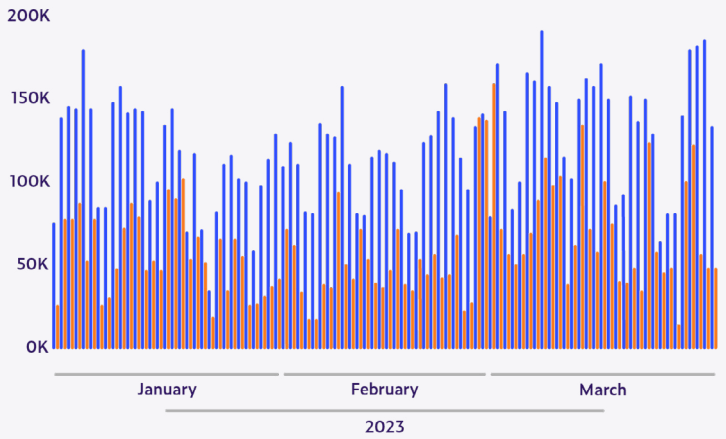


Figure 6: Daily traffic Adyen and Stripe



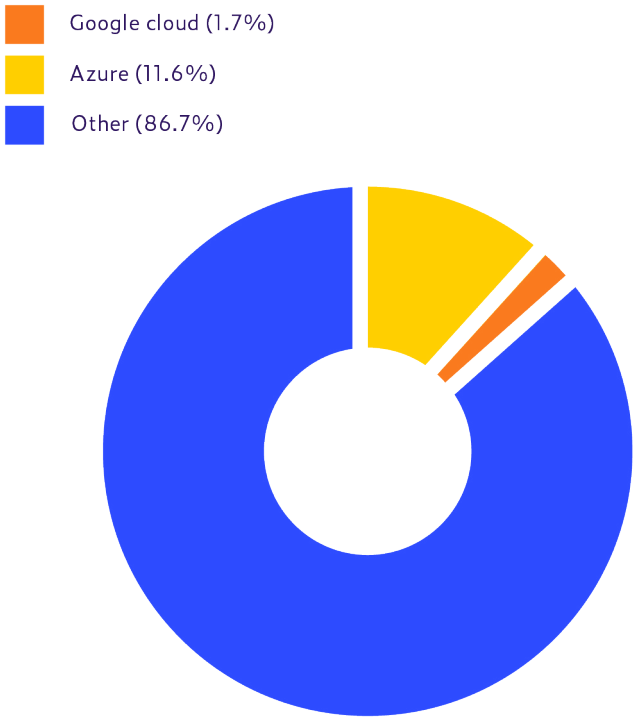
We use our proprietary Traffic Index developed from global connection data, which means it tracks not only human but also machine connections. The Traffic Index for any given domain is benchmarked against the traffic to www.google.com. For the current examples, we looked at connections to both the Stripe and Adyen checkout URL.

Asset discovery

In this section, we provide an example of how to use data obtained from publicly available SSL certificates to track web technology in the cloud. To ensure security, all servers need a verified SSL certificate. Tracking SSL certificates and structuring the data provides a unique view behind the firewall and allows the discovery of all machines connected to a certain SSL certificate, including but not limited to servers on a network. Certain web technologies have to be mentioned on an SSL certificate. In this example (Figure 7), we show that among the 24 million SSL certificates registered by **MongoDB Inc. (NASDAQ: MDB)**, 12% are related to **Azure cloud servers (NASDAQ: MSFT)** and about 2% to **Google's cloud platform (NASDAQ: GOOGL)**.

We scan over six billion SSL certificates and structure the public certificate information such as domains, subdomains, top-level domains and hostnames of the server name protected by the certificate, type and version of the SSL certificate, the date on which the certificate was validated and when it's going to expire. This data can then be used to search for specific companies and technologies. For this example, we looked for all hostnames associated with MongoDB and then further dissected the data to identify those certificates associated with Google Cloud and Microsoft Azure (NASDAQ: MSFT).

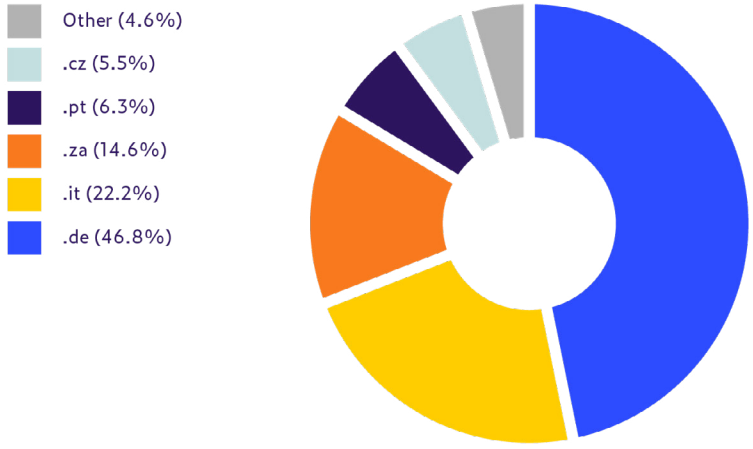
Figure 7: Share of SSL certificates owned by MongoDB associated with select cloud providers



Reversed search

Figure 8: Number of IPv4 addresses for major US telecommunications companies

Figure 9: Share of Vodafone associated IPv4 addresses



Typically when accessing a website, you type the name in the browser, e.g., www.nyse.com, but what you are actually connecting to is an IP address, e.g., 104.16.104.50. In reverse search, you start with the IP address and then identify the domain. As IPv4 addresses range from 1.1.1.1 to 255.255.255.255, it is possible to systematically iterate through all numbers to discover all hardware (servers, routers, mobile phones, etc.) associated with a certain domain. In Figure 8, we show the number of identified IPv4 associated hostnames for three top US telecommunications companies. Even though **Verizon Communications Inc. (NYSE: VZ)** is the second largest provider based on annual revenue, in our data we detect fewer IPv4 addresses compared to both **AT&T Inc. (NYSE: T)** and **Comcast Corporation (NASDAQ: CMCSA)**. In Figure 9, we show a break-down of of **Vodafone Group Plc (LON: VOD)** associated hostnames by top-level domain.

Biweekly we perform a reverse DNS lookup to map a domain name to the IP address of the server hosting that domain and to discover what networks surround it. For this example, we specifically searched for domains associated with AT&T, Verizon, Comcast and Vodafone.

Elevate your equity analysis with our data

The examples above only provide a snapshot of how web data can strengthen your equity research. To discover the full scope of what you can do with the data access our platform or explore the data via our partners at Snowflake and MScience.

Search the entire web in one big database

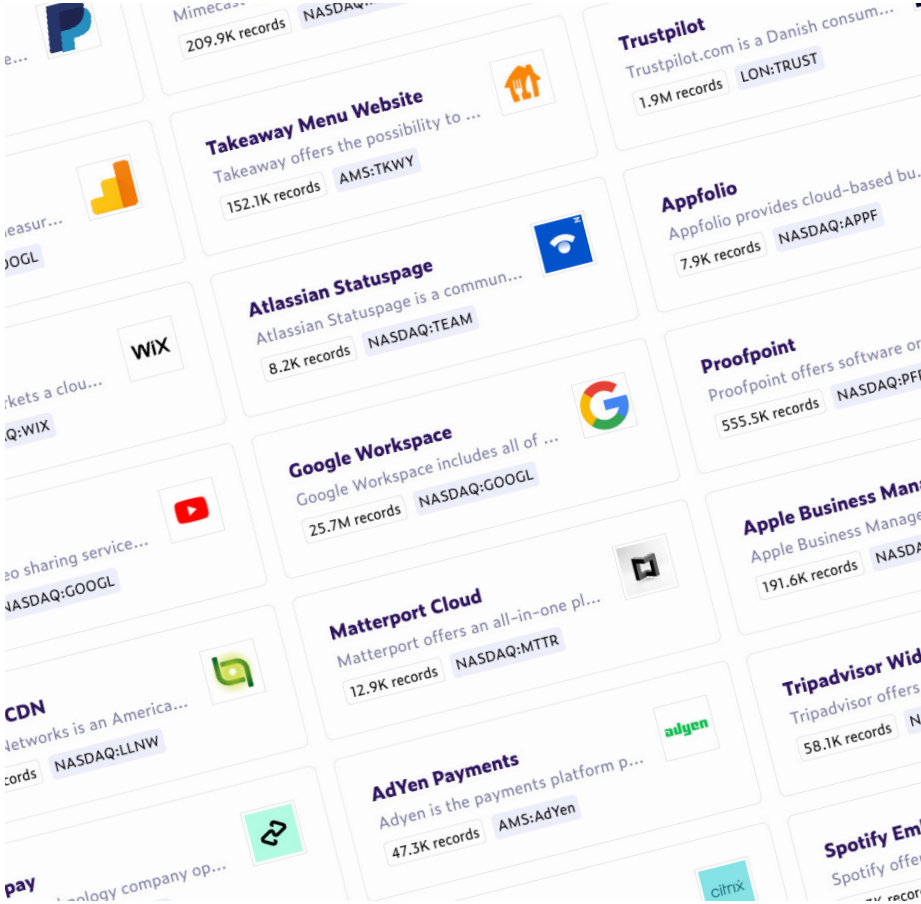
Our Search Engine grants you access to data from over 700 million hostnames, more than 200 data fields per hostname and up to four years of historical data. Perform custom searches with a range of B2B data fields and specific keywords to refine your results, helping you audit the performance of tech companies and gain insights, such as whether their customer base is growing or shrinking or where their customers are located.

Track technologies and predict future trends

Save time by using our pre-filtered custom datasets, also known as Recipes: with one click, discover the market share of tech companies such as Fastly Inc. (NYSE: FSLY), Stripe Inc., DocuSign Inc. (NASDAQ: DOCU) or Shopify Inc. (NASDAQ: SHOP) and specific web technologies such as Amazon AWS (NASDAQ: AMZN), Google Cloud Platform (NASDAQ: GOOGL), Facebook Advertisement (NASDAQ: FB) and many more.

Gain perspective with our proprietary scores

Put data into perspective for smarter insights: measure the digital world and analyze the data in a comprehensive way using our unique Proprietary Scores. Track website traffic with our Traffic Index, determine the economic activity of a website with the Economic Footprint and monitor the growth of a website with the Heartbeat. Find fraudulent websites with our Trust Grade.




DATA
PROVIDER.
COM