



Impact Report

Q4-Q1 2022-2023

bright data

the **bright**
initiative

Table of Contents

03

Intro: Public Data for Public Good!

04

There Has Never Been a Better Time to Do a Lot of Good (with public web data)

05

The Growing Reliance on Publicly Available Data

06

Bright Data 2023 In Numbers

07

The Bright Initiative by Bright Data

09

Our Responsibility

10

Compliance & Ethics - Leading the Way

12

Leadership Appointments: Bright Data's First CRO

13

Employee Satisfaction: Promoting a Great Place to Work

17

Customer Satisfaction

19

Safeguarding the Environment: Meeting International Standards

22

Data for Public Well-Being: Fighting Hate Speech and Online Signs of Violence

26

Saving Lives with Data

29

Supporting Credible Journalism & Reporting

32

Looking Forward - Impacting Nationwide Data Skills

33

Developing Data-Driven Policies to Benefit Society

35

Driving Forward Future Data Professionals

37

Powering Critical Research with Web Data Tools & Expertise

39

Public Web Data At Your Service: Fueling Your Valuable Missions



Public Data for Public Good!

Dear friends,

Over the past year, The Bright Initiative has reached its all-time high with 700+ global partners, addressing some of the most pressing issues of our time. Whether it is fighting hate speech, signs of political manipulation or violence, all of which can easily translate into actual violence in reality, or fighting widespread cross-continent crimes, such as human or sex trafficking, or keeping our children safe, the Initiative partner organizations are tackling these issues and many more with the most creative public web data approaches I've ever seen.

Bright Data is a business, and as a business, we would like to deliver the best market-leading technological solutions to our customers. Seeing these multiple Initiative partners in action and advancing their missions, some of which are life-critical, is what truly reflects the power of data and, more specifically, publicly available web data.

Bright Data's team has been working consistently on keeping this public wealth of information public. Looking at this impact report will make it very evident as to why. Without public data, a lot of these massively important missions would not come to life. Without this kind of information, we would return to an age where there is no transparency in keeping brands, big and small, accountable. Just imagine what it would mean for your daily life if there were no open market competition. What would be its impact on the overall cost of living, on medical research or even on managing a pandemic on a global scale? The answer is unimaginable, and I hope we never get there.

Thank you for the wonderful partnership and for using web data to do a lot of good,

Or Lenchner

CEO, Bright Data



*The attempt to limit access to public web data:
Or Lenchner speaks at the Snowflake summit.*

There Has Never Been a Better Time to Do a Lot of Good

(with public web data)

Dear partners,

As someone who speaks to many of our partners on a daily basis, I can testify that all of you have one thing in common: the determination to make your mission a success and, with that, do a lot of good with publicly available web data.

Over the past year, that determination has initiated multiple creative data-driven programs that, simply put, look after us, the public, our environment, our safety, our well-being and more.

I can now say with the utmost certainty that there isn't a mission that publicly available web data cannot support; all our partners have demonstrated this.

These missions have spanned fighting climate change by building first-of-a-kind tools that can help predict droughts and floods to ensuring diversity within every industry sector, and from empowering those who are a little hidden at first glance to identifying possible youth in distress by accessing social media posts and publicly available stories.

This year, it has been made very clear – data's power of good is accelerating, and your determination has made it unstoppable.

This is a moment of great pride for all of us, and we are genuinely honored to serve and support these hugely valuable missions.

If you are one with a mission that can impact many, join us!

Thank you for this determination that drives us to be better,

Keren Pakes

General Manager, The Bright Initiative



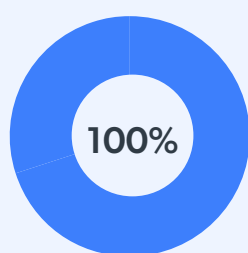
The Growing Reliance on Publicly Available Data

Both Commercial and Nonprofit Sectors Say Public Web Data Is Crucial to Their Mission

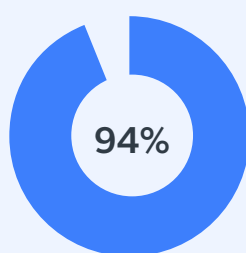
A recent [study](#) of data professionals in the U.S. who are working for both retail and technology businesses as well as nonprofit organizations found that data professionals and their employers rely significantly on publicly available web data to efficiently and effectively carry out their missions.

The study was commissioned by [Bright Data](#) and was conducted by the leading research firm [Vanson Bourne](#).

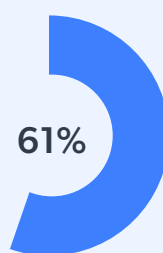
Key Highlights



Use publicly available web data in their operations to some extent



Use it in their day-to-day operations



Use public web data to compare pricing and inventory of competitors



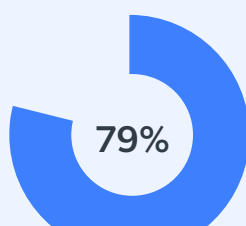
17%

Report using it for market research

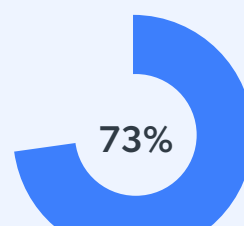
The survey also identified concerns among 68% of data professionals that the privatization of public web data would have a negative impact on both business and nonprofit operations.



Use public web data to identify harmful content



Agree that, without access to publicly available web data, their organization would not be able to operate effectively



agree that they could not accomplish their most important goals without public web data

Bright Data 2023

In Numbers



400+
employees

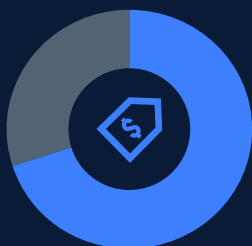


2 main offices,
USA (NYC) & Israel

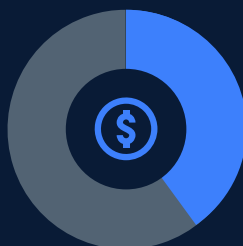


15,000+
customers

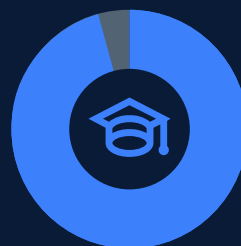
Who are Bright Data's customers?



More than half
of the 10 leading
e-commerce websites



2 out of the 5
leading US banks



96 of 100
leading universities

The Bright Initiative

by Bright Data

A pro bono program leveraging web data and data expertise to drive positive change across the globe.

In Numbers

700+ partner organizations



The Bright Initiative goals

Public web data is a force for public good

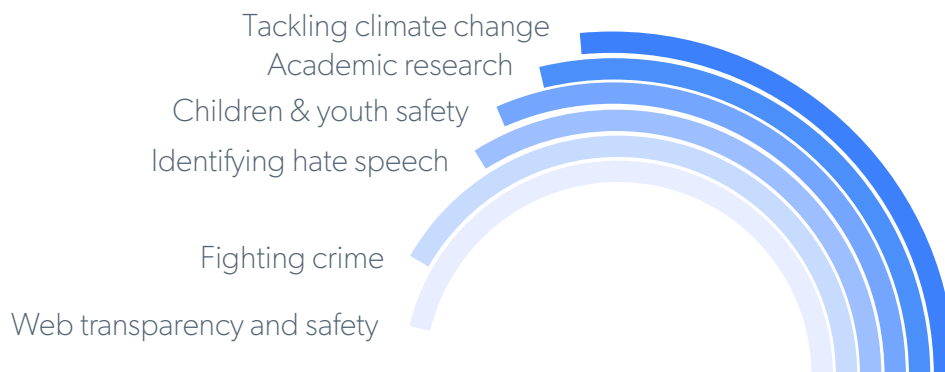
Shape the future of the public web
– keeping public web data public

Provide educational programs and support
research to impact our society and market

Use web data to promote public well-being and environmental protection

Develop data-driven policies and
regulation-driven initiatives to benefit all

Our Impact - Where the Data Is Going



The Bright Initiative Partners - Expanding the Impact

UK Government
Department of Digital,
Culture, Media & Sport

**DATA SKILLS
TASKFORCE**

Member of
WIPO Alert

MINISTÈRE
DE L'ENSEIGNEMENT
SUPÉRIEUR
ET DE LA RECHERCHE

BDV BIG DATA VALUE
ASSOCIATION

NCoc National Conference
on Citizenship

MRS Evidence
Matters™
Company Partner

UNIVERSITY OF
OXFORD

**PRINCETON
UNIVERSITY**

**Stanford
University**

Northwestern

BROWN

ETH zürich

N Northeastern
University

**KING'S
College
LONDON**

**Duke
UNIVERSITY**

UC San Diego

NYU

CORNELL UNIVERSITY
FOUNDED A.D. 1826

**THE UNIVERSITY OF
CHICAGO**

SciencesPo

UCSB

upReach
Transforming Social Mobility

**Manchester
Metropolitan
University**

**ROYAL
HOLLOWAY
UNIVERSITY
OF LONDON**

**Amsterdam University
of Applied Sciences**

**VIRGINIA
TECH**

RICE UNIVERSITY

SUBAK

**Climate
Policy
Radar**

**women
in
data**

Humans Against Trafficking

hti labs

ELEMIDOL

Our Responsibility



Preserve the right to access publicly available web data



Promote an openly competitive market using public web data that benefits businesses and society across all markets and industries



Enable information transparency and, with that, help multiple organizations effectively fight crime, social injustice, hate speech and more



Fuel academic research to tackle the most pressing issues of our time such as global warming, digital inclusion and more



Provide communities with the data tools, skills, and knowledge to drive our society forward

Compliance & Ethics - Leading the Way

Considering the uniqueness of the public web data industry and the lack of comprehensive guidelines, Bright Data has compiled first-of-a-kind compliance guidelines and practices from other sectors, such as Telecommunications, with regard to safeguarding our infrastructure and products such as the KYC process (Know Your Customer). We have also established an entire set of practices as well as an extensively global department to increase the overall standard of compliance and ethics for the public web data collection industry.

Some Key Highlights

ISO 27001

Bright Data has received its ISO 27001 certification, an international standard on how to manage information security in an organization.

CSA Star Level One

Striving to adhere to the most significant standards in our domain, Bright Data is proud to have achieved CSA Star Level One.

CSA (Cloud Security Alliance) is a global NPO that promotes security best practices for cloud service providers. Their STAR (Security, Trust, Assurance, and Risk) certification is part of their activity, serving as a security assessment for potential customers and partners.

We will proceed to achieve Level Two based on a 3rd-party audit. You can view it in the STAR [registry](#) as well as in the Bright Data [trust center](#).

CSA Star Level One

Striving to adhere to the most significant standards in our domain, Bright Data is proud to have achieved CSA Star Level One.

CSA (Cloud Security Alliance) is a global NPO that promotes security best practices for cloud service providers. Their STAR (Security, Trust, Assurance, and Risk) certification is part of their activity, serving as a security assessment for potential customers and partners.

We will proceed to achieve Level Two based on a 3rd-party audit. You can view it in the STAR [registry](#) as well as in the Bright Data [trust center](#).

Domain Category Block

Keeping our infrastructure and networks secure is an ongoing process. The Domain Category Blocker is taking this effort to the next level by exponentially expanding the ability to block any sites in the blacklisted categories while whitelisting and enabling immediate access to websites in pre-approved categories.

Up until now, Bright Data has practiced an approach of blocking first a long list of websites, focusing on well-known questionable verticals such as adult content, phishing and malware, unlawful streaming, crypto, gambling, etc. **Launching the Domain Category Blocker enables us to filter millions of websites in a simple, automated manner.**

Unknown new websites that have not been categorized by our expert compliance department require further manual inspection, while whitelisted categories are automatically approved.

Commitment to Promoting Ethical and Environmentally Aware Working Methods

Bright Data is committed to developing solutions that reduce the environmental impact of public web data collection tools while maintaining complete transparency and preventing their misuse for malicious intents.

To accomplish this goal, Bright Data introduced three new features to its platform:

01

Site Health Monitor

An internal system that tracks the performance health of Bright Data's target sites to ensure they aren't overloaded when their performance decreases without Bright Data usage. Bright Data monitors website health by sending sample requests and monitoring the RTT (round-trip time). When sites are overloaded or their performance decreases, the RTT value soars. When performance is back to normal, the RTT will decline. On top of that, the Site Health Monitor tests websites from different locations to detect whether changes in website health occur in specific areas or globally, learning the usage boundaries per site and geolocation.

02

BrightBot and Webmaster Console

BrightBot, the crawler layer in Bright Data's platform tasked with accessing, retrieving, parsing, and qualifying public web information, is an industry-first ethical collector. BrightBot monitors the health of every domain it targets, prevents access to non-public information, and blocks interactive endpoints that could be abused, like ad clicks, reviews, likes, account management, etc. The Webmaster Console allows website owners to authenticate data properties on their platform and leverage Bright Data's Site Health Monitor free of charge. Websites can also whitelist BrightBot data center IPs for additional benefits such as reduced bot traffic, a dashboard showing data collection statistics, and data collection credit.

03

Cache Proxy

Bright Data customers can enable Cache Proxy and install the accompanying certificate so that users' usage is cached and served to other customers who make identical requests to the same target sites. Caching data results in target sites being under less load while users enjoy results 20 times faster than usual requests and pricing discounts for any proxy with cache proxy enabled. Cache Proxy ensures ethical usage and that no forbidden endpoints in the target site are accessed. It is important to note that only successful results obtained by any proxy network (Datacenter, ISP, Residential, or mobile) are cached, and the cached results shown are, at most, two hours old.

Leadership Appointments:

Bright Data's First CRO

(Chief Revenue Officer)



Bright Data is proud to share the appointment of Gunja Gargeshwari as its new Chief Revenue Officer. With over 25 years' experience in sales and marketing at leading tech companies, such as Oracle, AWS, and Zendesk, Gargeshwari will join the executive leadership team at Bright Data and oversee all global GTM (Go-To-Market) operations.

In this newly created role, Gargeshwari is now leading the company's extensive business development and sales teams, focusing on elevating Bright Data's valuable customer base. Gargeshwari is also focusing his and the team's efforts on expanding the company's market reach to new data territories — addressing the most large-scale emerging data requirements.

Gargeshwari holds an impressive track record of business leadership roles in IT services and SaaS (Software as a Service) at leading tech companies. Most recently, Gargeshwari served as Global VP of Platform Sales at Zendesk, where he led GTM for multiple product lines and drove Zendesk's Enterprise growth. Prior to Zendesk, Gargeshwari managed GTM operations at AWS, working with leading customers in the Digital Native and ISV space. He also spent 18 years at Oracle where he held various leadership positions.

Or Lenchener, CEO at Bright Data said: "I'm thrilled to have Gargeshwari join the Bright Data team as we expand our business. His appointment is a testament to Bright Data's commitment to delivering world-class solutions to our customers and growing the business in a competitive marketplace."

For further information, click [here](#).

Employee Satisfaction:

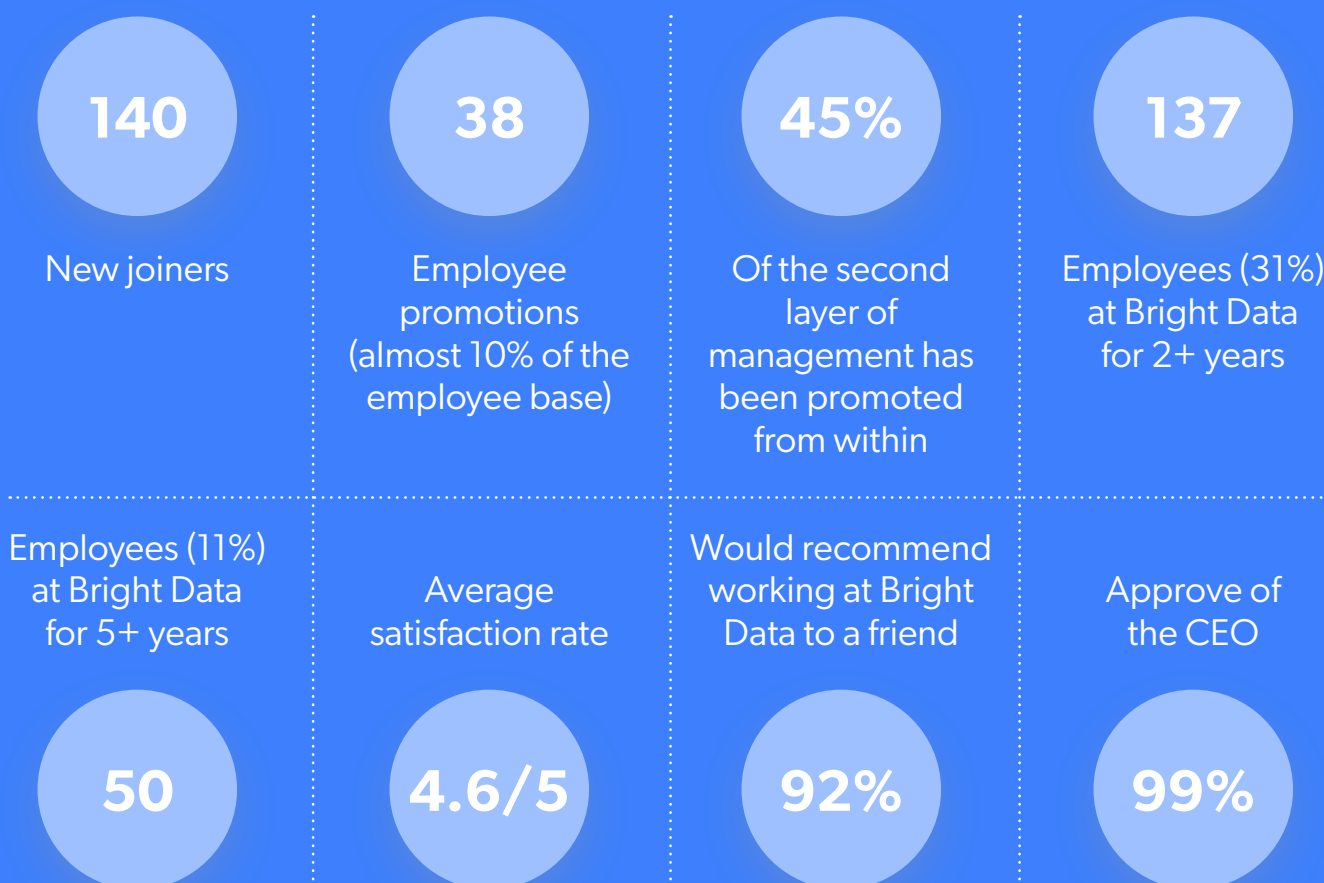
Promoting a Great Place to Work

We strongly believe that employee satisfaction is a key indicator of a company's overall market or industry leadership and success. Bright Data has consistently recognized that our most valuable asset is our people, and we take conscious steps to improve the employee experience.

With 140 new hires in 2022, we are constantly adjusting our organizational structure to accommodate growth while prioritizing promotions and regularly measuring employee satisfaction through various methods, including internal tools and external platforms. To further support this effort, we appointed Limor Kidron as Chief Human Resources.

Important to note: Over 30% of Bright Data's employees have been with the company for over 2 years, with 45% of the extended management team coming from within the company. Our diversity and inclusion rates are at an all-time high, making us a stronger, more creative team. As we look forward, we are confident that this combination of top talent and a positive and productive culture will allow us to evolve and innovate even more so than before.

2022 - Our People in Numbers



From Our Teams - The Voices of Bright Data

“Along with having the best technical product in the world, the company offers many promotion opportunities and a lot of flexibility for working remotely. I have the opportunity to work with the best professionals in the industry and with clients who are some of the most recognizable brands in the world. Overall, it is the best workplace I could have hoped for!”

Sales Director



“Bright Data is an excellent company to work for. They truly care for their employees, and it shows. The office is high-tech and provides healthy food and snacks all day. The work-life balance is also great. Once you are here, you truly feel part of the team; everyone is professional, friendly, and willing to help if needed.”

Compliance Officer

“Great to see your work affect people + intensive rhythm of development - fast learning + awesome community + good trend - company understand there to invest and that to improve + compatible compensation”

Software Engineer in Kazan



“There are few companies that understand work-life balance and actually respect that, and I can guarantee that Bright Data is one of them. You get to work with top thought leaders in several industries and smart and collaborative co-workers who are always there to help you. We are doing incredible things that impact the way business is done across the world. That's powerful!”

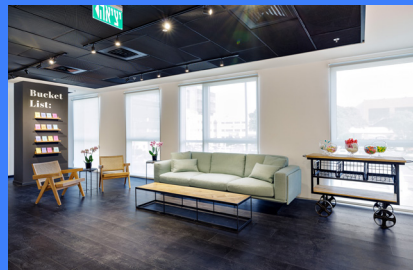
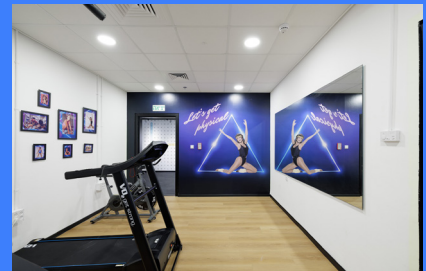
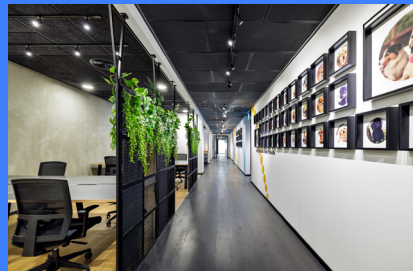
Sales Development Representative

For more employee voices please visit our [Glassdoor page](#).

A Bright New Office for The Bright Data Team

At the end of 2022, Bright Data moved to a brand new office to meet the growing needs of our expanding team. This beautiful new office features special corners and innovative design approaches best suited to a leading innovative team.

The pictures tell the story well.



Our Teams in Action - Giving Back to the Community

As in every year, we have been involved in various donation drives and initiatives over the past year, including:

Supplying various household items, such as furniture, clothing, and equipment, to boarding schools, rehabilitation centers, and hostels through “Guy Group” and also through “Home Share,” a cooperative community store in our HQ hometown of Netanya.

A significant portion of the treats and gifts for our employees were purchased from suppliers who employ people with disabilities or special needs, including Hameshakem, Yeladudes, and Kfar Idud.

When the war broke out in Ukraine, we collected clothing, toys, and equipment for children and donated to those in need through the Wizo Organization in Ra'anana.



Customer Satisfaction

Obtaining customer feedback is key to making sure we meet and exceed our own set goals. We constantly seek feedback from our customers to continuously innovate and improve our technology, products, and service and to give our customers the best solutions and experience possible.

Customers from a wide range of organizations worldwide have weighed in on their experience working with Bright Data – on multiple leading review sites such as G2, Capterra, and Trustpilot.

Based on their excellent reviews, we were pleased to receive multiple awards over the past six months, showcased below, including seven new winter badges from leading review platform G2. G2 has also rated us among the top software companies in the EMEA region (number 26) and awarded us the prestigious "Top 50" badge. Moreover, The Gartner Digital Markets Report has recognized us with three awards, including the Capterra Shortlist.

In addition, we have managed to maintain our score on the popular user feedback platform, Trustpilot, with 4.8 out of 5 stars and 540+ reviews from our customers.



“

“Public web data has become the most used source of data in businesses today across the board. As Bright Data’s survey with research firm Vanson Bourne indicates, the uses of web data range from competitive strategy building to pricing strategy, from measuring public sentiments to training operational systems. All this is done to ensure that customers or consumers get the most relevant services, offers, or products. After all, that is the goal of every business and organization.

“Bright Data is enabling the McDonalds Canada team to gain access to detailed competitor insights from pricing, menu offerings, and distributions, all available online.”

Amir Yazdi, Marketing Analyst at McDonald’s Canada



“Overall, we are genuinely pleased with the decision we made to go with Bright Data. The service has already added a lot of value to our business; it is easy to use and is stable and robust. The ease of use also really helps when onboarding new employees to use the system.

“In addition, the support service and communication are very agile and attentive, which is a very important factor because we live in an increasingly digital world, but we cannot forget the human touch.”

David Barros, Data Scientist at Worten

”



“

“The reason we chose Bright Data is the combination of the breadth of the web data, the different datasets available, the freshness of the web data, and the ability to bundle all those relevant datasets into a reasonable price for us.

“If we didn’t have access to this web data, it would certainly make our lives a lot more difficult.”

Danny Issroff, Cred Investments



Safeguarding the Environment

Meeting International Standards

Stepping Up Our Net Zero Journey

Together with Normative, a science-based carbon accounting software granting companies tools to reduce their carbon footprints, we are committed to reaching Net Zero by 2026. For that reason, we have set our science-based targets for 2023, which include reducing GHG emissions by 38% by 2030 from a 2021 base year and measuring and reducing our scope 3 emissions.

Normative

Identifying the Causes for Droughts and Floods

The Bright Initiative by Bright Data is an official partner in the [UN's Integrated Management Drought Program](#). For the past following months, joint expert teams from across the globe are looking for ways to utilize publicly available web data to build a first-of-a-kind tool that will be able to closely study droughts and floods so we can all better understand how to best respond to the growing challenge and even predict it.

The teams have been looking into versatile web sources such as public social media posts, media reports as well as subject matter reviews.



INTEGRATED DROUGHT
MANAGEMENT PROGRAMME

Navigating Environmental Risk - The Global Centre for Risk and Innovation

The Global Centre for Risk and Innovation (GCRI) is an international non-profit network committed to accelerating public participation in research, policymaking, and development programs. Its goal is to create a unique nexus where civil society meets and engages the academy, industry, and government to make smarter decisions toward more resilient, sustainable, and prosperous communities.

The network turned to The Bright Initiative by Bright Data to help collect public social media data to analyze the discussion and deliberation process around environmental risk. It looks to see how people gravitate toward different sentiments and the reasoning behind them through a detailed cognitive process.



"Because we operate in a zero-trust environment, web data is vital for us, stakeholders, and the public. Consequently, we must provide evidence beforehand and make as much public information available as possible. And that's where The Bright Initiative is our best friend - it gives us the technology to delve into, for example, social media discussions, bringing epistemic knowledge to our work. There's so much more to it than just statistics."

Saeed Valadbaygi, The Global Centre for Risk and Innovation

BRight app - Combating Climate Change Effortlessly

At Bright Data, we're committed to using our expertise and technology to make a positive impact on the world around us. That's why we've developed the [BRight app](#), which enables anyone to plant trees and combat climate change absolutely effortlessly. By simply downloading the app, you help us plant trees across the globe. Our partnership with dots.eco ensures that the impact is felt globally, and the trees are planted where they're needed the most.

Our app works by utilizing your device's unused bandwidth to retrieve publicly available web data for vetted companies and non-profits, while we plant trees in your name. Additionally, when charging your phone, you can offer your extra resources for environmental research and other valuable initiatives that we committed to empower through The Bright Initiative. With the BRight app, we're empowering everyone to make a difference in the fight against climate change, one tree at a time.



Screen-Save the Planet



Data for Public Well-Being

Fighting Hate Speech and Online Signs of Violence

The Network Contagion Research Institute (NCRI) aims to identify and forecast cyber-social threats targeting individuals, organizations, and communities.

The Institute is working towards facilitating honest conversations around the spread of political deception, hate, and manipulation, especially on social media. NCRI uses The Bright Initiative by Bright Data's platform to monitor publicly available web content on TikTok, Instagram, and other social media platforms.

It's often been said that the internet is the new public square. People meet, shop and form bonds there – but bad actors lurk in the shadows. In this digital town square, accessing publicly available data is how NCRI can monitor for threats to help protect vulnerable communities, public health, democratic institutions, and civil society as a whole.



The National Contagion Research Institute (NCRI) develops contagion models and utilizes machine learning fueled by public data to identify who is plotting harm so they can identify cyber-social threats and forecast where those threats may become more prolonged and cause real world harm. Then, they put that information into the hands of those at risk so that bad actors can be stopped in their tracks. In simpler terms: Public web data enables NCRI to identify, forecast and mitigate online threats before they escalate into real-world violence.

Unfortunately, in the digital age, online manipulation is rampant. For example, children face exploitation through the transmission of unwanted and inappropriate images across social media platforms.



“NCRI has seen firsthand the positive impact — and significant value — of partners like Bright Data. Bright Data provides public data that has statistical significance across social media platforms and other publicly accessible digital forums. Using NCRI’s specific skill set and pioneering methodologies in evaluating scaled social media data to identify trends that could lead to real-world violence, we have successfully forecasted potential threats. For instance, NCRI was able to leverage Bright Data’s datasets to track emerging antisemitic threats, which are unfortunately ever-increasing. The work Bright Data and other companies do can help inform law enforcement officials of potential threats to places of worship.

“Given the rise of criminal networks and misinformation online, local communities and law enforcement must be fully prepared. NCRI’s research provides one step, but it needs the foundational data to develop these analyses. Bright Data provides the open-sourced data to companies, organizations and researchers to develop solutions to thwart the potential kinetic online activity from escalating into real-world violence.

“We know this firsthand. Using our machine learning and scaled data analysis technology, NCRI quantitatively forecasted a wide range of troubling new social media related challenges including the rapid growth of social media ‘self harm groups’ among teenagers, rampant manipulation that helped to crash the crypto market, even where potential reluctance to COVID-19 vaccinations can be most pronounced down to a county level. This vital work doesn’t happen without the partnership of companies like Bright Data.”

Adam Sohn, CEO, NCRI

Combating Antisemitism Head-On

The Anti-Defamation League

The Anti-Defamation League (ADL) is an international Jewish non-governmental organization based in the United States and specializing in civil rights law. The ADL fights all forms of antisemitism and bias, using innovation and partnerships to drive impact.

The ADL uses The Bright Initiative by Bright Data's platform to track and follow individuals with radical views on social media platforms and their activity. By identifying and tracking these individuals and their network, they can enhance their investigations and identify more individuals with extreme antisemitic and racist views.

The logo for the Anti-Defamation League (ADL) is displayed in a white, rounded rectangular box. The letters "ADL" are in a bold, sans-serif font, with the "A" and "D" connected.

CyberWell

CyberWell is a non-profit organization dedicated to democratizing the fight against online antisemitism. It serves the public as the premier research, monitoring, and alert center for tracking online antisemitism across social media platforms in multiple languages by launching the first-ever open database of online antisemitism.

CyberWell is now working with Bright Data's Web Scraper IDE – through the Bright Initiative – to help with the automatic collection of publicly available web data from social media platforms at scale.

The organization then creates real-time alerts that are tailored for each social media platform and shared for them to "address and act upon."

CyberWell has already identified over **100,000 pieces of content** based on key phrases and narratives guided by the working definition from the International Holocaust Remembrance Alliance (IHRA).

After identifying cases of antisemitism, the team then reports infringing content and monitors how long it takes for that content to be removed.

In 2022, CyberWell vetted and reported more than 2,000 instances of antisemitic content, identifying an average removal rate of around 24% across the five platforms.

When Ye – formerly known as Kanye West – made several antisemitic comments last year, CyberWell found that this led to an almost 128% increase in antisemitic content vetted by the business compared with the previous month.

The CyberWell team also identified five antisemitic narratives and a series of hashtags (e.g., #KanyelsRight, #YelsRight, #SynagogueOfSatan) that appeared in content supporting Ye or his statements.

Recently, Cyberwell and The Bright Initiative by Bright Data held a special Holocaust remembrance data event, led and inspired by Holocaust survivor and TikTok influencer, Gidon Lev (who has over 400K followers).



"Since the social media channels themselves are failing to remove this dangerous content effectively, our initiative relies on public web data to identify, track and report hate content that is increasingly amplified by millions when left online. Access to open-source public social media data is key to stopping these escalating waves of hate."

Tal-Or Cohen Montemayor, founder and executive director, CyberWell

Saving Lives with Data

HTI Labs Combats Human Trafficking

Since 2021 The Bright Initiative has been supporting an incredibly impressive organization: [HTI Labs](#). This organization focuses on developing an intelligent infrastructure for fighting violence and exploitation. The recognition from them as part of our partnership through The Bright Initiative by Bright Data humbles us, and we will remain committed to delivering top-notch solutions and publicly accessible web data to aid them.



“We have been partnering with Bright Data, an industry-leading web data platform, through The Bright Initiative by Bright Data since April 2021. Their state-of-the-art proxies are one of the cornerstones of solving the engineering challenges involved with scraping commercial sex ad data. Their partnership is critical to the work we do.”

[HTI Labs LinkedIn page](#)



hti labs

Identifying Youth in Distress

ELEM Israel - Youth in Distress is a 35-year-old nonprofit organization that has taken up the critical responsibility of helping at-risk youth in Israel. As part of its cutting-edge program, ELEM has teamed up with The Bright Initiative by Bright Data, aiming to use public social media data to possibly spot young girls who may be at risk of or have already endured domestic abuse.

A team of social workers and data engineers worked together to develop an AI model that analyzes social media posts for keywords, phrases, hidden messages, tags, and other potential distress signals. Training and improving this AI model requires a vast amount of data. Therefore, ELEM utilized Bright Data's public datasets to build a data collector that can extract new information and create data on a large scale.

Once all targeted profiles are accessible through a platform, trained social workers will be able to create an intervention plan and help make it accessible to girls in need. It's essential to remember that, sometimes, saving lives begins with reaching the right real-time information to provide the much-needed help!



Tackling Gender Equality in the Workplace - Senior IT

Shay Yahal, a team leader of data engineers, scientists, and analysts at Grip Security, wanted to hire a data scientist for her team and noticed that out of the 30 applications she received, only one was a woman. Along with Shira Asa-El, a Business and Computer Science M.A. student at Harvard University, they embarked on a mission to find how to bring in more women to fill senior positions in hi-tech – Senior IT.



The Israeli Central Bureau of Statistics shows that 33% of university computer science graduates are women, but the average tech team doesn't reflect the same statistics. To find a solution, Yahal and Asa-El first had to conduct thorough research. They turned to The Bright Initiative for help, and thanks to Bright Data's preexisting LinkedIn dataset, they could analyze the information and identify market trends.

Through the data collected from Bright Data's platform and interviews of women in various stages of their careers in tech, Yahal and Asa-El found that larger companies employ a higher percentage of women. They concluded that there's a tendency for women to look for stability in their workplace and to want more mentorship, and women tend to find that in larger companies.

Another critical analysis of the data showed the difference between how men and women brand themselves on LinkedIn, such as how many years of experience, on average, it takes men to claim to be seniors compared to women. The difference they found later translates to better work opportunities and higher salary demands.

senior.it

“

“There are a lot of projects aimed at bringing women to tech, but few projects focus on the part of the funnel of getting women to advance from junior to senior positions. The value we got from Bright Data is immense; it powered our research, and without Bright Data's help, I don't think we would have had any information.”

Asa-El

“I was surprised by how comfortable and easy the process had been with The Bright Initiative. In several hours, we had all the data we needed from LinkedIn. The existence of this initiative is extraordinary, especially since data is something so valuable - the fact that people can simply use it to do good should not be taken for granted.”

Shay Yahal

”

Supporting Credible Journalism & Reporting

Empowering Journalists with Reliable Real-Time Data

The Bright Initiative powered by Bright Data recently supported a [study by The New York Times \(NYT\)](#), one of the world's leading newspapers. Ever since Tesla CEO Elon Musk took over Twitter and lifted the curtain on accounts barred from sharing their extremist views on the social media platform, NYT technology reporter [Stuart A. Thompson](#) wanted to review the effect of those once-banned users, so he turned to The Bright Initiative for help.

Thompson provided Bright Data with a list of 1500+ accounts that underwent reversals of permanent suspensions, collected by Travis Brown, a Berlin-based software developer who has tracked extremism on Twitter. After analyzing the list and identifying Thompson's needs, Bright Data deployed its Web Scraper IDE to compile a list that contained the date of when the users were restricted and unbanned, their following, the content of their tweets, links to the original posts, and the engagement those tweets generated.

With the data supplied, Thompson could see how most reinstated accounts went back to spreading the same messaging that got them banned in the first place, knowing that under Musk's permissive code of conduct, they were no longer in danger of censorship.

Bright Data's public web data tools played a key role for Thompson in conducting his research, producing insights, and supporting his argument with credible and conclusive public information, the building blocks of investigative journalism.

Click [here](#) for the full article.

Uncovering Inequality with Data-Driven Investigative Journalism

The Bright Initiative powered by Bright Data supported a [study conducted by The Markup, a nonprofit newsroom](#). This organization makes it its mission to run data-driven investigations. This particular study was based on recent research also supported by The Bright Initiative and published by Princeton University. Investigative reporters Leon Yin and Aaron Sankin wanted to uncover the digital inequality in internet package deals offered by leading ISPs in different parts of major American cities, so they turned to The Bright Initiative to use Bright Data's platform to conduct their research.

"As became painfully obvious when pandemic lockdowns began, fast, reliable internet service is a necessity today. A [Pew survey](#) released last year found that 90 percent of Americans said the internet has been essential or important to their lives over the course of the pandemic," says the study's introduction.

Yin and Sankin reviewed the broadband availability tools, taken directly from ISPs' websites, to show that the racial and wealth makeup of each area's residents, among other factors, led to a digital divide for high-speed internet infrastructure. The reporters' analysis also examined pricing in conjunction with speed and the demographics of areas that are offered the same price for worse service, a practice called "tier flattening."

Using Bright Data's infrastructure, the reporters gathered and examined 1.1 million service offers made by different internet carriers. They found that not only is tier flattening rampant but also that people who can least afford internet services are being offered substantially lower quality internet than other people living in the same city for the same price.

Bright Data's platform and infrastructure played a key role in assisting The Markup's reporters in their research by granting them access to public web data, using it to uncover the disparities between different socio-economic areas in the same city in internet package deals, which is an important step towards combating inequality.

The Markup

Combating Fake News

[Marcin Sawinski](#) and [Milena Stróżyna](#) are researchers at the Department of Information Systems at the [Poznan University of Technology](#). Together, they are leading the [OpenFact project](#) in collaboration with Goggle, Meta, Harvard University as well as leading fact-checking organizations in Poland.

With funding from the Polish National Research Center, the end goal of their research is to increase the capacity of fact-checkers when scanning the network for potential fakes.

Through our partnership, Sawinski and Strozyna have access to [Bright Data](#)'s platform which allows them to orchestrate and manage all the processes locally while fetching the public web data they need from different zones.

The results of the first phase of the project were delivered as a dataset containing fake news in Polish for training artificial intelligence algorithms as well as possible detection methods. The researchers aim to develop a solution that will inform the public when the credibility of messages is in question while using search engines and social networks.



"Bright Data was key for our research because the data we need to conduct our research is not easy to find. We're partnering with quite a few organizations, but the support we've been getting from The Bright Initiative is a game-changer for us. We're amazed that The Bright Initiative exists to help researchers because when it comes to academic research, it's not that easy to get funding."

Marcin Sawinski



POZNAŃ UNIVERSITY
OF ECONOMICS
AND BUSINESS

Looking Forward - **Impacting Nationwide Data Skills**

An important mission for The Bright Initiative is to work closely with policy makers, using our data expertise and innovative tools to truly contribute to the development of the required skill set of tomorrow.

Here are some highlights:

France's Education Ministry - Being Market Ready

The Bright Initiative has been working closely with the French government's Ministry of Higher Education, Research and Innovation. Relying on public web data, the ministry aims to be ready for the future and better assess the required skills needed in a market greatly driven by technological innovation.



**MINISTÈRE
DE L'ENSEIGNEMENT
SUPÉRIEUR
ET DE LA RECHERCHE**

Developing Data-Driven Policies to Benefit Society

Supporting the UK's National Data Strategy

Throughout 2022, The Bright Initiative by Bright Data continued to support the UK's National Data Strategy (NDS), helping to ensure that data benefits the country's economy and society. Most notably, the Initiative was an active member of the NDS Forum - a body established by the Department of Digital, Culture, Media, and Sport (DCMS) to drive the NDS forward. Through the Forum, Bright Data regularly provided UK Government officials with expertise and advice on complex industry issues like skills and the role of data intermediaries.

Complementing membership of the NDS Forum in 2022, The Bright Initiative also accepted an invitation to join the Data Skills Taskforce. Charged with ensuring that the UK builds the skills base needed to develop as a global data leader, the Taskforce brings together leaders from industry, the public sector, and educational institutions to develop and deliver high-impact initiatives. The Bright Initiative took up an invitation to give a presentation to the Taskforce about the work it does with universities around the world.

In 2022, The Bright Initiative also became a sponsor and steering group member of the All Party Group for Data Analytics Inquiry into data and AI ethics. Over the coming year, this will see The Bright Initiative work alongside members of the UK Parliament to collect and consider evidence on how ethical considerations can inform regulation policy around the use of data and AI.

In addition to these formal positions, The Bright Initiative has helped to support the NDS through a range of partnerships and events. These include:

- Working with the Women in Data network to convene a summit looking at how data can be used to tackle violence against women and girls;
- Ensuring students from less advantaged backgrounds are supported to embark on data careers by working with the social mobility charity upReach;
- Building cross-party support for the power of data with events at the conferences of both main UK political parties.

We also published expert opinion articles in specialist media outlets, including [LocalGovernment.com](https://www.localgovernment.com), Times Higher Education, Business Leader, and The Sunday Times "Future of Data" report, and produced and published a white paper on how data can help fight climate change.

2022 NDS Support in Numbers:

- 7 major events
- 8 expert consultation meetings
- 9 expert opinion articles

**DATA SKILLS
TASKFORCE**

The Skill Set of Tomorrow Depends on Data Expertise

As part of The Bright Initiative's ongoing support of the UK's NDS (National Data Strategy) Initiative, we recently convened a meeting of UK business leaders with the country's [Education Secretary, Rt. Hon. Gillian Keegan MP](#). Posted in partnership with the Enterprise Forum – a respected business network – the event was held at the heart of Westminster in the prestigious Broadway House building.

In her remarks, the Education Secretary welcomed Bright Data's commitment to collaborating with educational institutions and emphasized the need for others to follow suit. She highlighted measures that the government has introduced, including [T Levels](#) and [Skills Bootcamps](#), as ideal opportunities for businesses to support the development of cutting-edge skills for future generations.

A full discussion with the business leaders in attendance followed, addressing the range of nationwide initiatives. All agreed with Bright Data CRO Gunja Gargeshwari's premise of the importance of collaborating with educators, and with the Secretary calling for all in attendance to leverage the government's initiatives.

The hosting of this event further demonstrated Bright Data's commitment to bringing social and economic benefits from data – and specifically publicly available web data – to governments worldwide.



Driving Forward Future Data Professionals

As mentioned, The Bright Initiative runs over 10 educational sessions a month, helping the future data professional gain real insights into the market of today.

We recently worked with several universities, including Brown, Queen Mary, Kings College and more.

Here are some highlights:

Data4Good Turned Educational

The entire essence of The Bright Initiative is to use data4good. We recently turned it into a practical course and ran our first such session at Brown University. See [here](#).



Educating Students on How Public Web Data Can Bring Positive Change

As part of our extensive partnership with Queen Mary University, we hosted several virtual workshops titled "The Web Data Revolution – At Our Service" for the university's students.

During one session, Bright Data's VP of Product, Erez Naveh, offered rare insights into the tech industry and how it is now part of our everyday lives, and Chief Brand Officer, Keren Pakes, discussed how data and brands are intrinsically linked nowadays.

Organized in collaboration with [Queen Mary Careers](#), we have also offered to provide expertise and use of our platform across the University, including the possibility of managing a Q-tech project.

Helping students understand the power of data and its part in their careers and lives is essential to Bright Data's mission to show how data offers insights that can bring about positive action and change.



Coders at Data Action - Nod Coding Going Global

Nod is an intensive 10-week-long data analytics coding bootcamp based in Stockholm, Sweden. The program focuses on learning the most sought-after and up-to-date technologies, where students eventually build a substantial coding portfolio to showcase to future employers.

The Bright Initiative coordinated a face-to-face session with Nod Coding to teach its students how to use the Bright Data platform pro bono to gain access to real-time, quality public web data to augment their GitHub portfolio and add it to their list of projects produced during the bootcamp.

The Bright Initiative session at the Nod Coding Bootcamp illustrates Bright Data's commitment to educating future data analysts on using public web data to power many projects and research.



Powering Critical Research with Web Data Tools & Expertise

Rethinking GDPR Compliance Measurements to Impact the Future

[Amit Zac, Ph.D.](#), is a Postdoctoral Researcher at the Swiss Federal Institute of Technology, ETH Zürich. Zac's research revolves around creating a general method for measuring compliance with the General Data Protection Regulation (GDPR).

Zac turned to The Bright Initiative for help in his attempt to change the existing automated compliance measuring practices, which are limited to specific types of websites. Zac intends to use the Bright Data platform to scrape over a million public websites to get a representative sample of the internet.

Leveraging Bright Data's platform through The Bright Initiative is of utmost importance to Zac's research as the data provided is unbiased and unaffected by where Zac is based. Then, using Natural Language Processing and Machine Learning methods, he can analyze the data extracted, layer several compliance practices together, and compare them with GDPR requirements.

“

“If private companies truly wanted to make society better, they would grant access to their data for researchers to conduct their studies. The Bright Initiative is the best we can ask for in today's world to get closer to the unbiased data we need. There are many ways to realize the vision of combining business with making a social impact as The Bright Initiative does.”

Amit Zac

A Deep Dive into Diversity, Equity, and Inclusion

[Kaustav Sen, Ph.D.](#), is a professor of accounting at the Newman School of Business at [Pace University](#) in New York City, who analyzes historical data to study and use econometric techniques. As part of Prof. Sen's study on corporate finance and looking at incentives and corporate executives and governors in a broader way, he is researching how global corporation employers treat their employees concerning diversity, equity, and inclusion.

To power his research, The Bright Initiative provided Prof. Sen with public datasets from two public domains crucial for his study: the [EEOC](#) (Equal Employment Opportunity Commission), a part of the US government disclosure, and [Glassdoor](#).

The two datasets allowed Prof. Sen to study how employers treat their employees and the priorities of major global corporations in Corporate America regarding their workforce. Specifically, Prof. Sen analyzed the data and identified how companies emphasize the issues of diversity, equity, and inclusion at work.



"The Bright Initiative helped me collect and create two major datasets crucial for my research and presented them in a clean format to help save a lot of time. I am keen to take this partnership forward because there are more interesting questions to investigate and more data I'm interested in looking at, including climate-related data released by climate agencies."

Kaustav Sen, Ph.D., Pace University

PACE
UNIVERSITY

Making the Internet More Transparent

As part of his Ph.D. studies at the University of Toronto, [Shawn Shuang](#) is working on creating software that automatically detects and identifies scripts that track cross-website user activity. Much like the Chrome adblocker extension, Shuang hopes to create a similar tool that would alert users when their activity is being tracked.

To create such a tool, Shuang needed to access websites without being blocked and identify the scripts' activity. Leveraging Bright Data's platform through his partnership with The Bright Initiative, Shuang's vision of making the internet transparent is possible.



"I'm trying to get a general sample of the internet, but I focus on websites that are known to track users, like news or shopping sites. I want to make the internet more transparent and let people know when they're tracked and what private information is collected. Bright Data's platform is easy to use, and the results have been reliable."

Shawn Shuang

 UNIVERSITY OF
TORONTO

Public Web Data At Your Service

Fueling Your Valuable Missions

As we enter our third year, we are very much looking forward to further expanding our partnerships and making an even greater impact on the world as a whole.

Public web data really is the foundation of everything today, and we are committed to continuing to enable its accessibility where it matters most – by helping to drive market competition, promoting public well-being, advancing research, assisting life-saving missions, and imparting the internet for good – all for the greater good of society. As market leaders, we see this as our duty.

If you're interested in partnering with The Bright Initiative, please reach out to:

info@brightinitiative.com

bright data

the **bright**
initiative