



Internet Society
Pulse

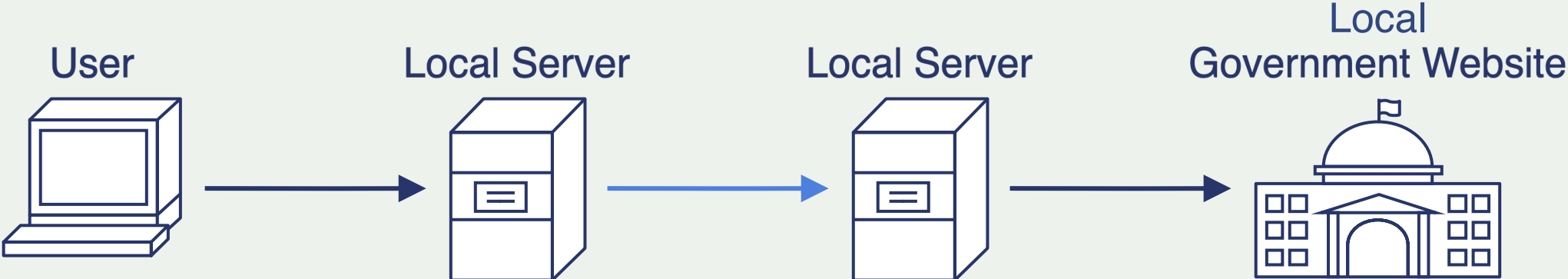
Towards Measuring Content Locality

James I. Madeley, Amreesh Phokeer,
Theophilus A. Benson, Aftab Siddiqui

ANRW'24
23rd July 2024

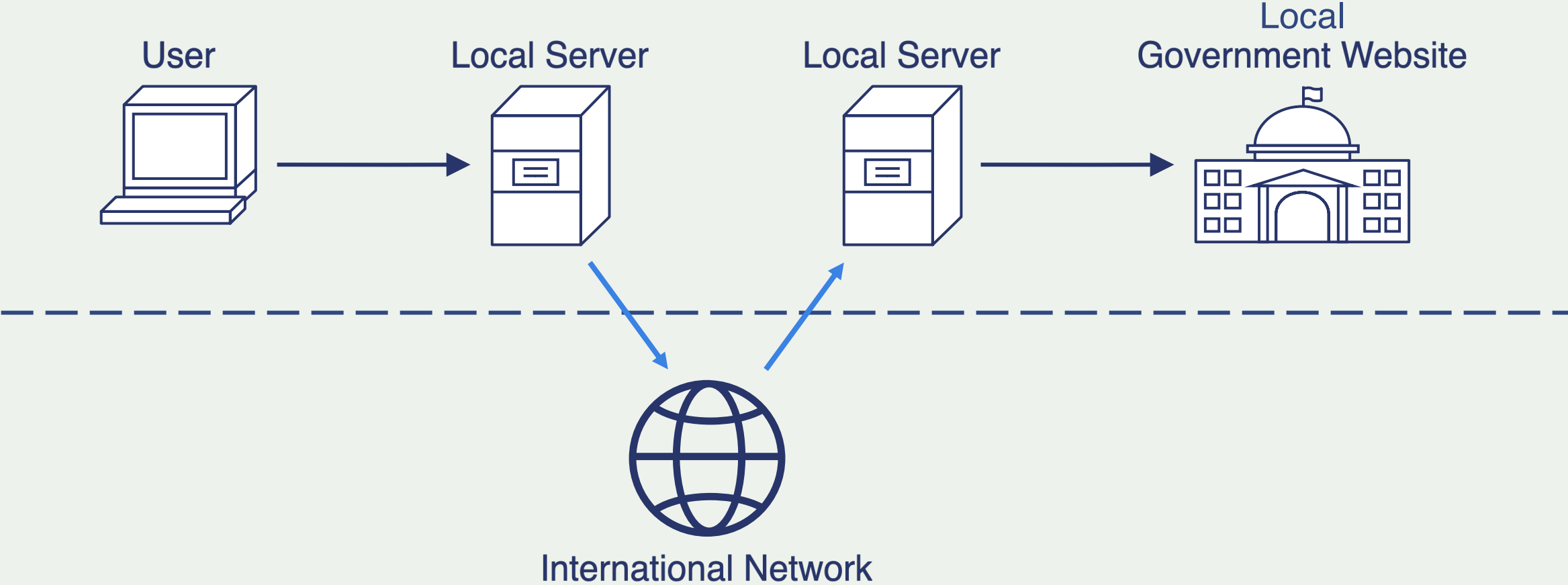


Local Traffic



International Network

External Traffic



Previous Works on CDN Deployment

Web Content Cartography

Bernhard Ager
T-Labs/TU Berlin
bernhard@net.t-labs.tu-berlin.de
Georgios Smaragdakis
T-Labs/TU Berlin
georgios@net.t-labs.tu-berlin.de

Wolfgang Mühlbauer
ETH Zurich

CacheLoc: Leveraging CDN Edge Servers for User Geolocation

Mingkui Wei¹[0000-0003-3606-3428], Khaled Rabieh²[0000-0003-2828-6971], and Faisal Kaleem²[0000-0001-6780-1759]

¹ Cyber Forensics Intelligent Center, Computer Science, Sam Houston State University, Huntsville, TX
² Computer Science and Cybersecurity, Metropolitan State University, Saint Paul, MN

The Central Problem with Distributed Content

Common CDN Deployments Centralize Traffic In A Risky Way

Kevin Vermeulen
TASCORP

Loqman Salamatian
Columbia University

Sang Hoon Kim
Columbia University

Ethan Katz-Bassett
Columbia University

Internet development in hosting and distribu

Enrico Calandro¹, Josiah Chavula², and Amreesh Phokeer³

¹ Research ICT Africa, Cape Town, South Africa
ecalandro@researchictafrica.net

² University of Cape Town, Cape Town, South Africa
jchavula@cs.uct.ac.za

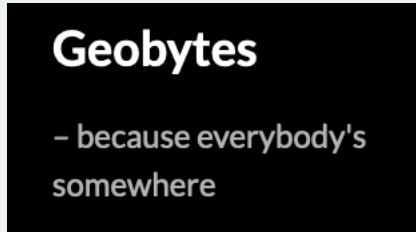
³ AFRINIC, Ebene, Mauritius
amreesh@afrrinic.net

Peering vs. Transit: Performance Comparison of Peering and Transit Interconnections

Adnan Ahmed and Zubair Shafiq
The University of Iowa

Harkeerat Bedi and Amir Khakpour
Verizon Digital Media Services

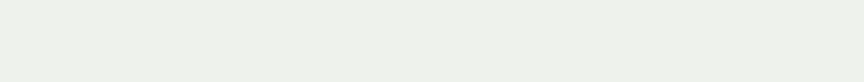
Previous Works on Geolocation



Locating CDN Edge Servers with HTTP Responses

Run Huang, Mengying Zhou, Tiar
Shanghai Key Lab of Intelligent Information Processing, Schc
{runhuang19,myzhou19,tcguo20,ch

Geolocation of IP Hosts in Large Computer Networks with Congestion



Towards Geolocation of Millions of IP Addresses*

Zi Hu John Heidemann Yuri Pradkin
USC/Information Sciences Institute {zihu, johnh, yuri}@isi.edu

Kishan B. Patel

Nadine Moukdad
Computer Science
University,
NY 10458
ukdad@gmail.com

S. Anand
Department of ECE
New York Institute of Technology
New York, NY 10023
Email: asanthan@nyit.edu

Motivation

Motivation

Policy

Performance

Persistence
(availability)

Motivation

Policy

South Africa's POPI
Act / EU's GDPR

Internet Society's
50/50 Vision

Performance

Persistence (availability)

Motivation

Policy

South Africa's POPI
Act / EU's GDPR

Internet Society's
50/50 Vision

Performance

Live streaming
favours local content

Persistence (availability)

Motivation

Policy

South Africa's POPI Act / EU's GDPR

Internet Society's 50/50 Vision

Performance

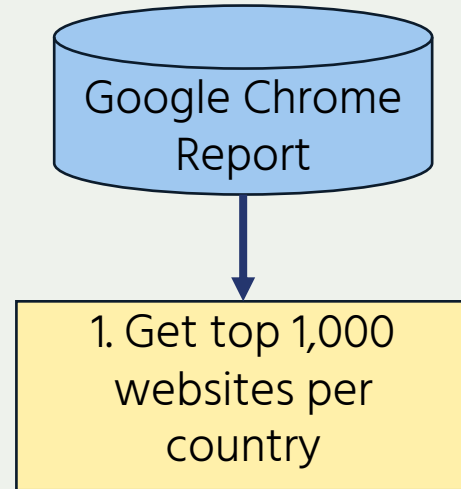
Live streaming favours local content

Persistence (availability)

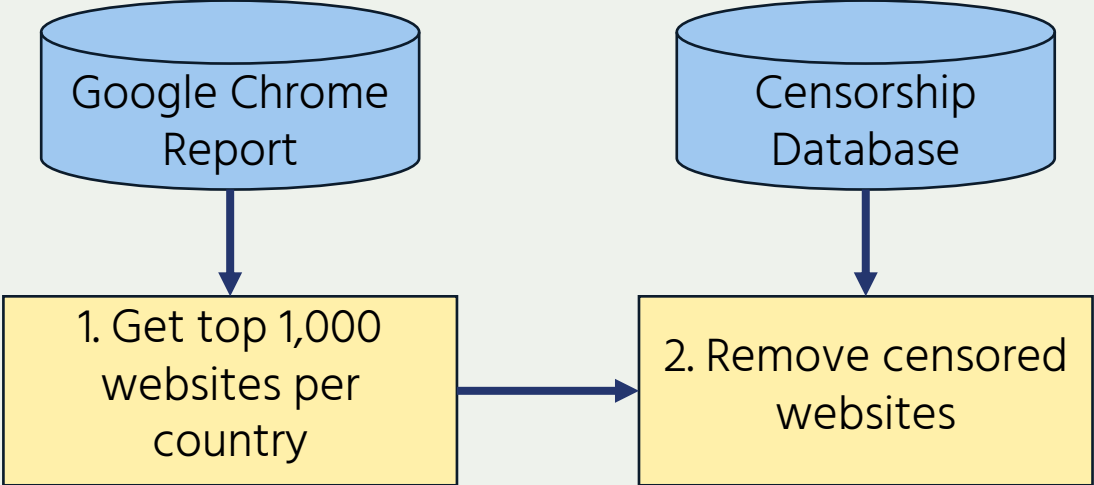
Resilience against faults e.g. cable cuts

Methodology

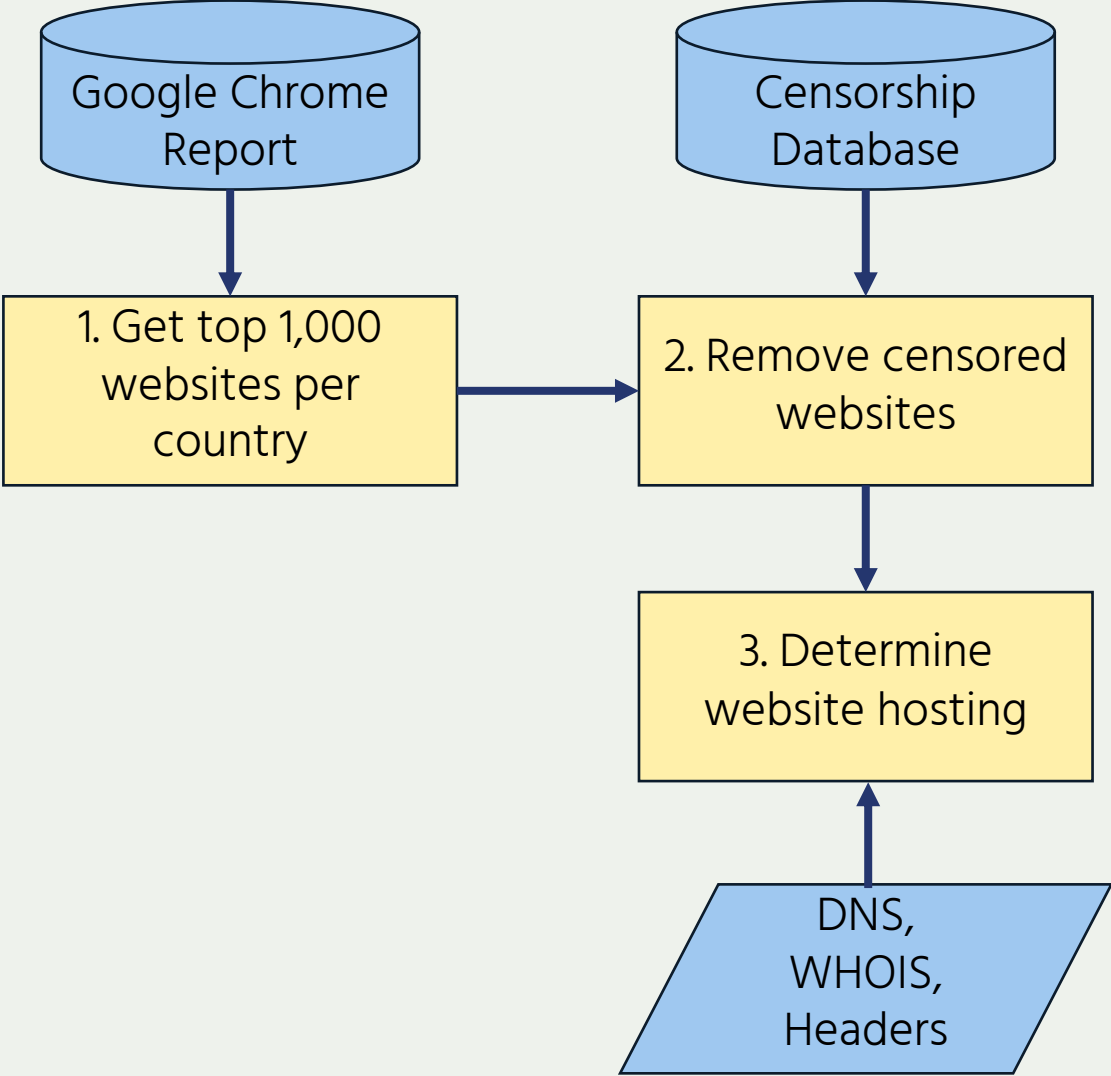
Methodology



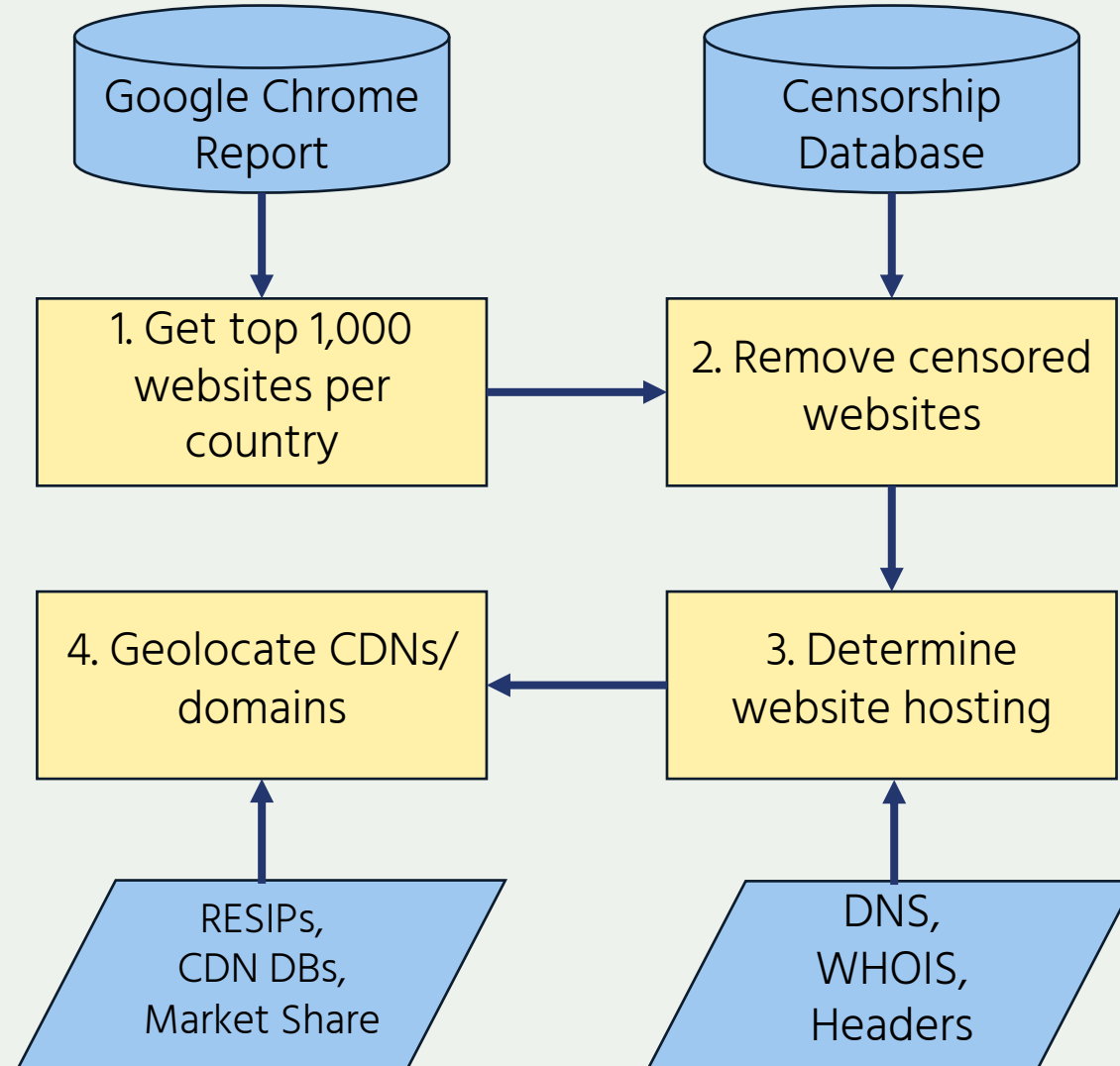
Methodology



Methodology



Methodology



Methodology – Determine Website Hosting

www.studysmarter.co.uk, NL (Netherlands)

Service	Returned	Parsed
WHOIS	CDN77 _, GB	CDN77
DNS CNAME	prod-web-uk.b-cdn.net.	BunnyCDN
	Result	Tie

Methodology – Determine Website Hosting

www.studysmarter.co.uk, NL (Netherlands)

Service	Returned	Parsed
WHOIS	CDN77 _, GB	CDN77
DNS CNAME	prod-web-uk.b-cdn.net.	BunnyCDN
	Result	Tie
Headers	BunnyCDN-DE1-1080	BunnyCDN
	Result	BunnyCDN

Methodology – Geolocate Domains without CDNs

wikipedia.org, NL (Netherlands)

Methodology – Geolocate Domains without CDNs

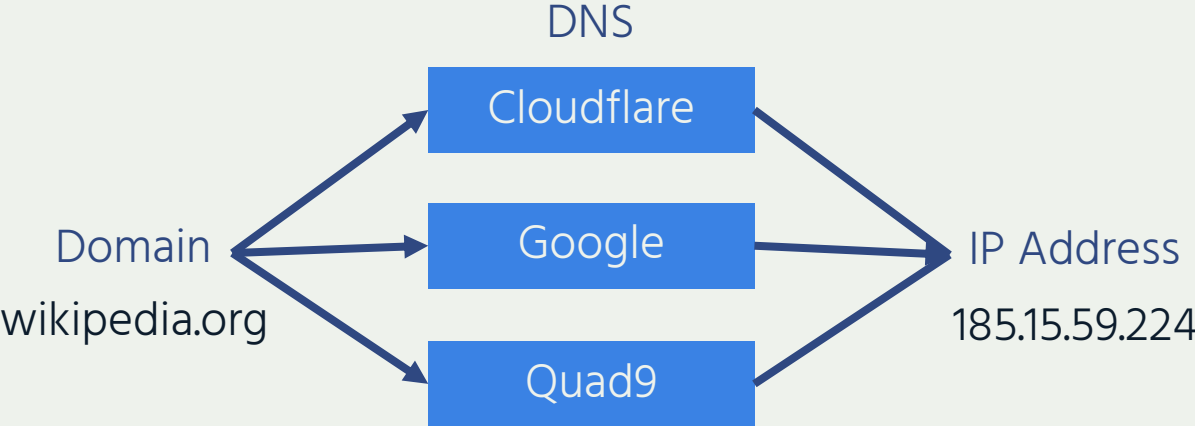
wikipedia.org, NL (Netherlands)

Domain

wikipedia.org

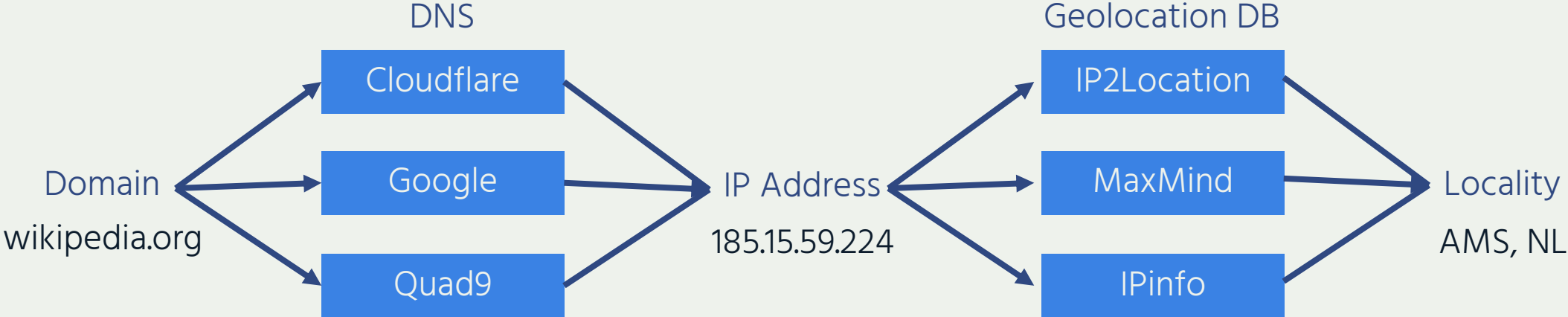
Methodology – Geolocate Domains without CDNs

wikipedia.org, NL (Netherlands)



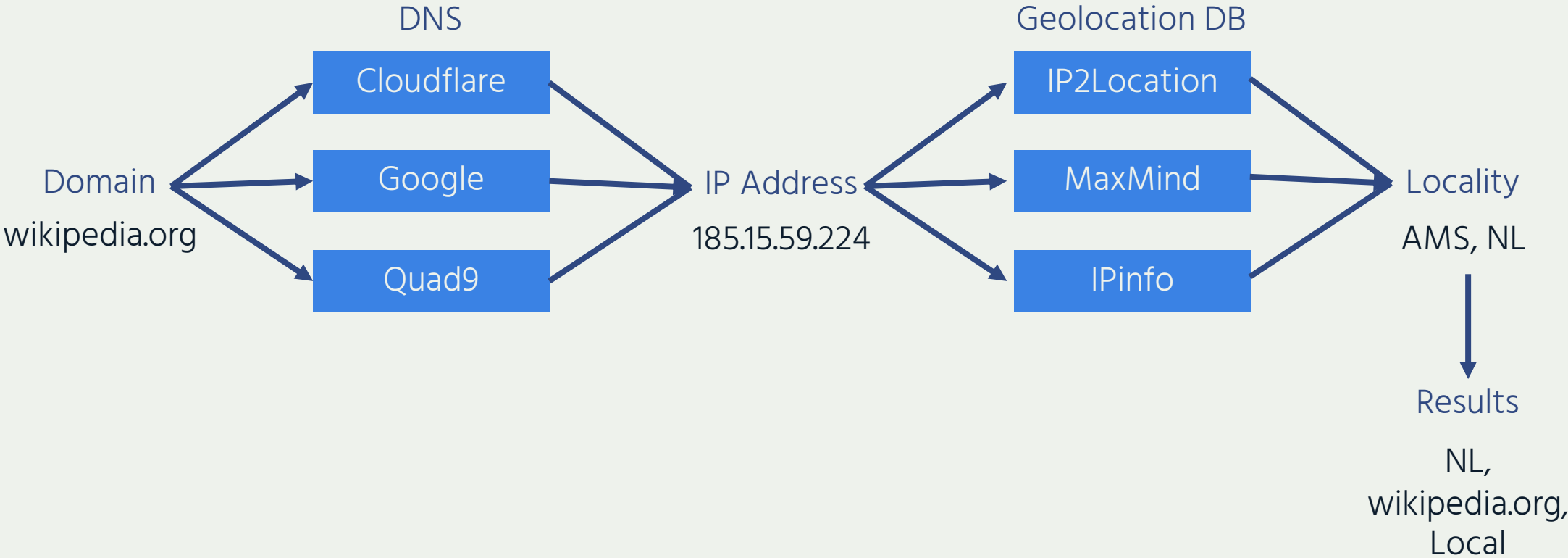
Methodology – Geolocate Domains without CDNs

wikipedia.org, NL (Netherlands)



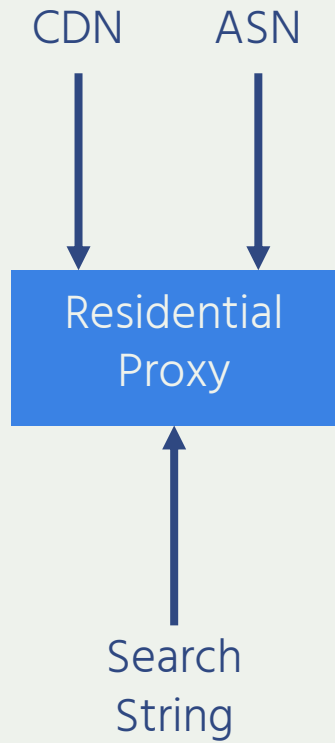
Methodology – Geolocate Domains without CDNs

wikipedia.org, NL (Netherlands)

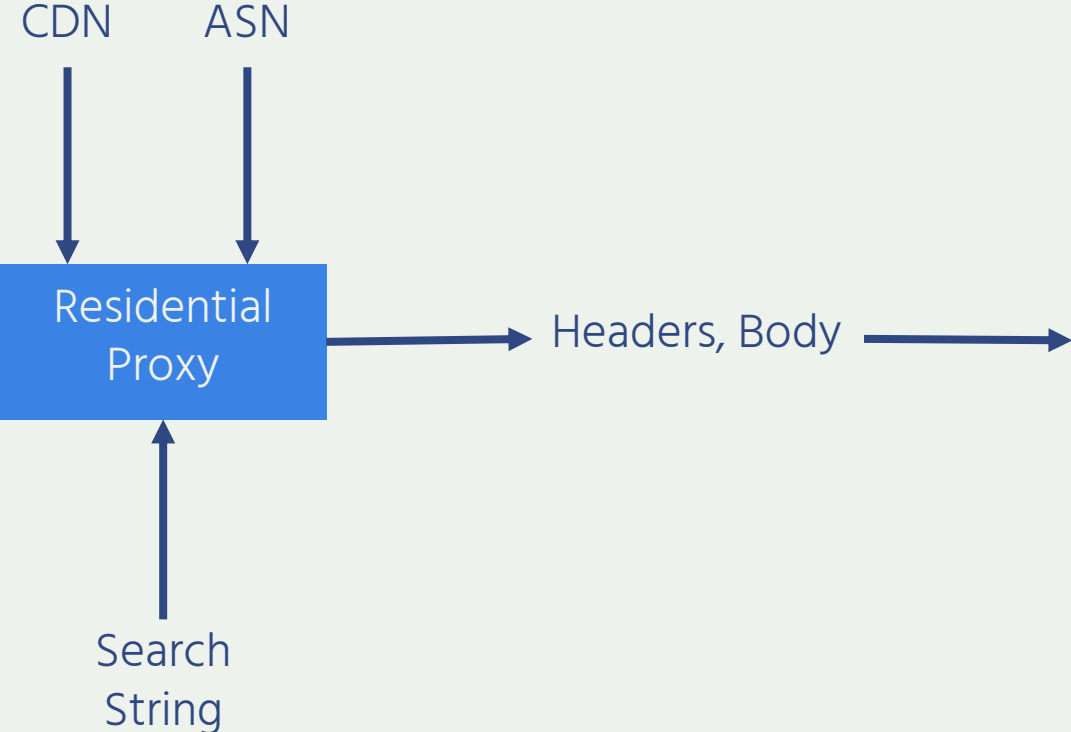


Methodology – Geolocate CDNs

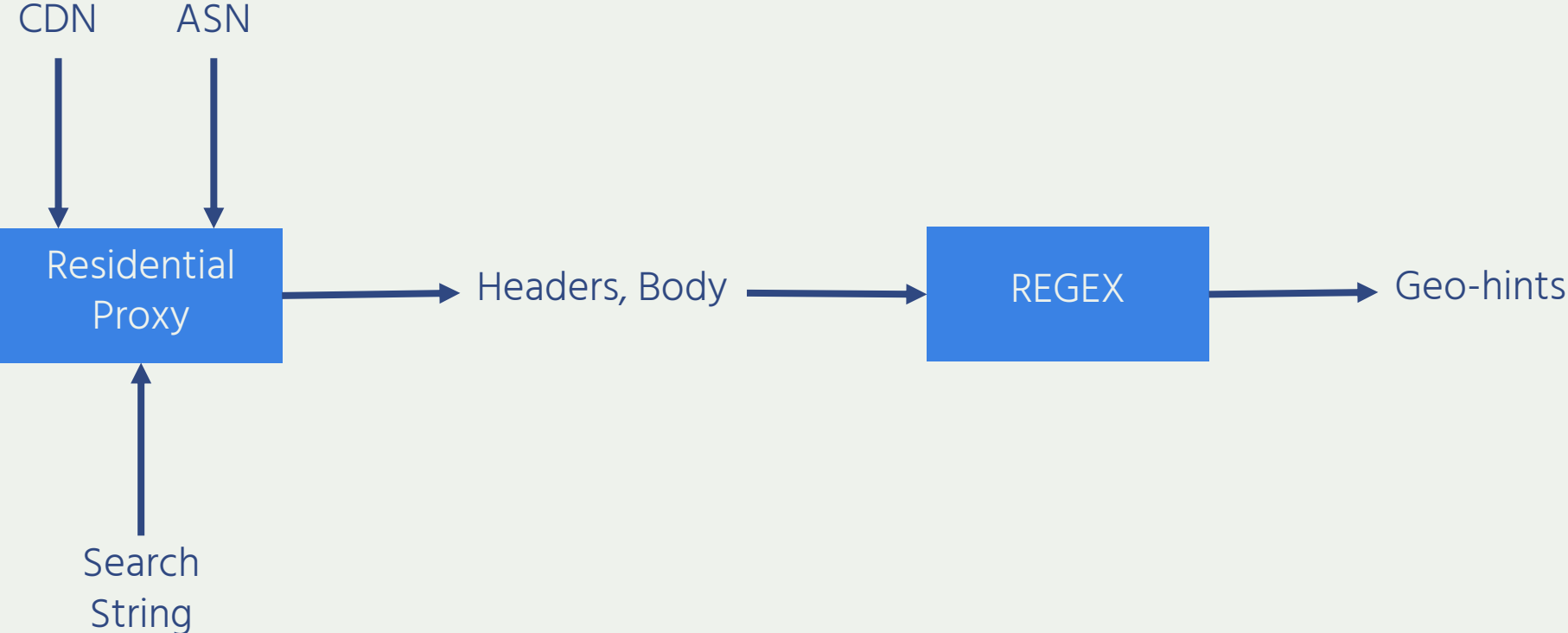
Methodology – Geolocate CDNs



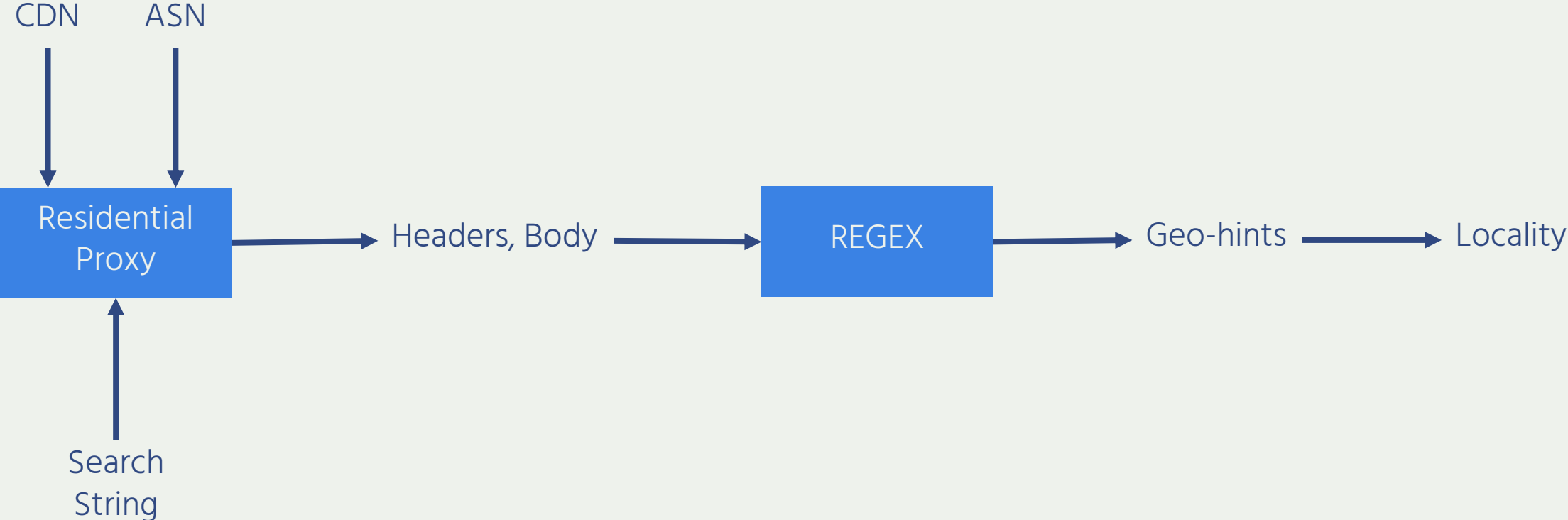
Methodology – Geolocate CDNs



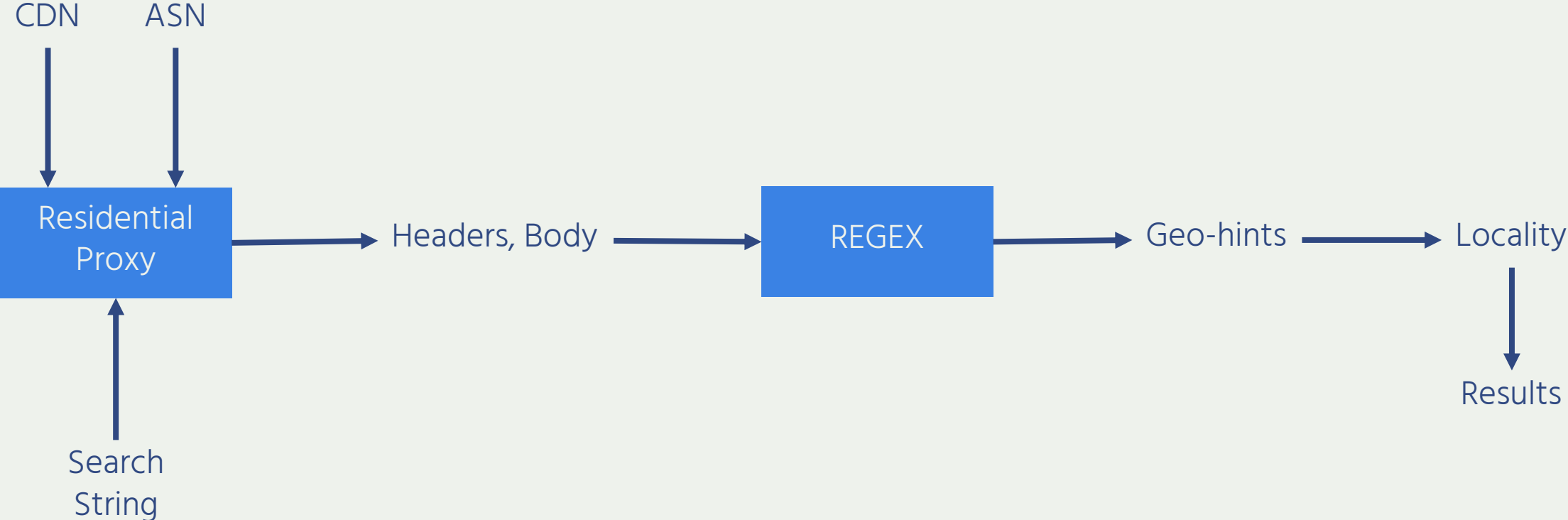
Methodology – Geolocate CDNs



Methodology – Geolocate CDNs

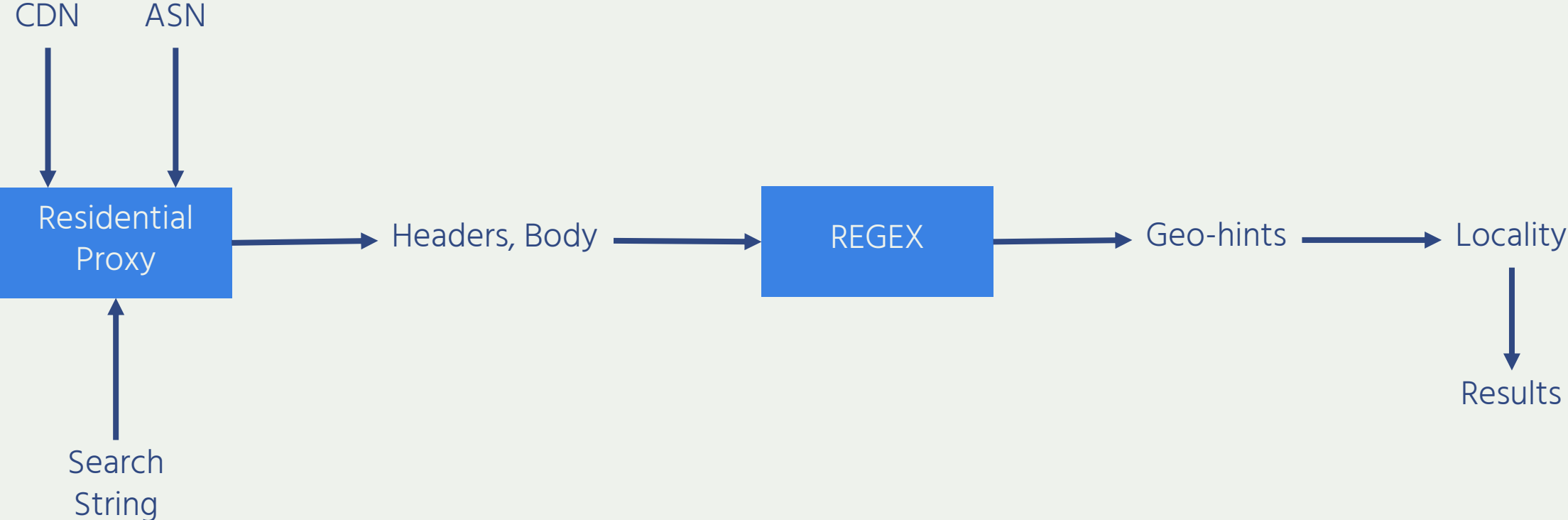


Methodology – Geolocate CDNs



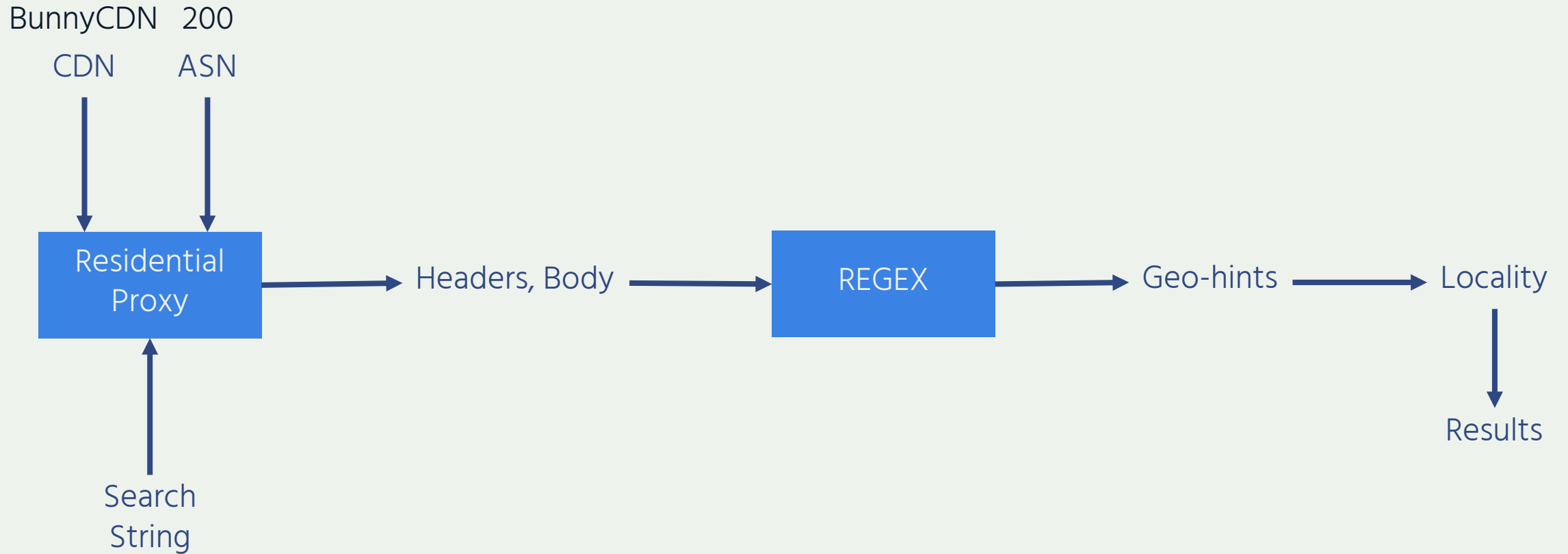
Methodology – Geolocate CDNs

BunnyCDN, AS 200, NL (Netherlands)



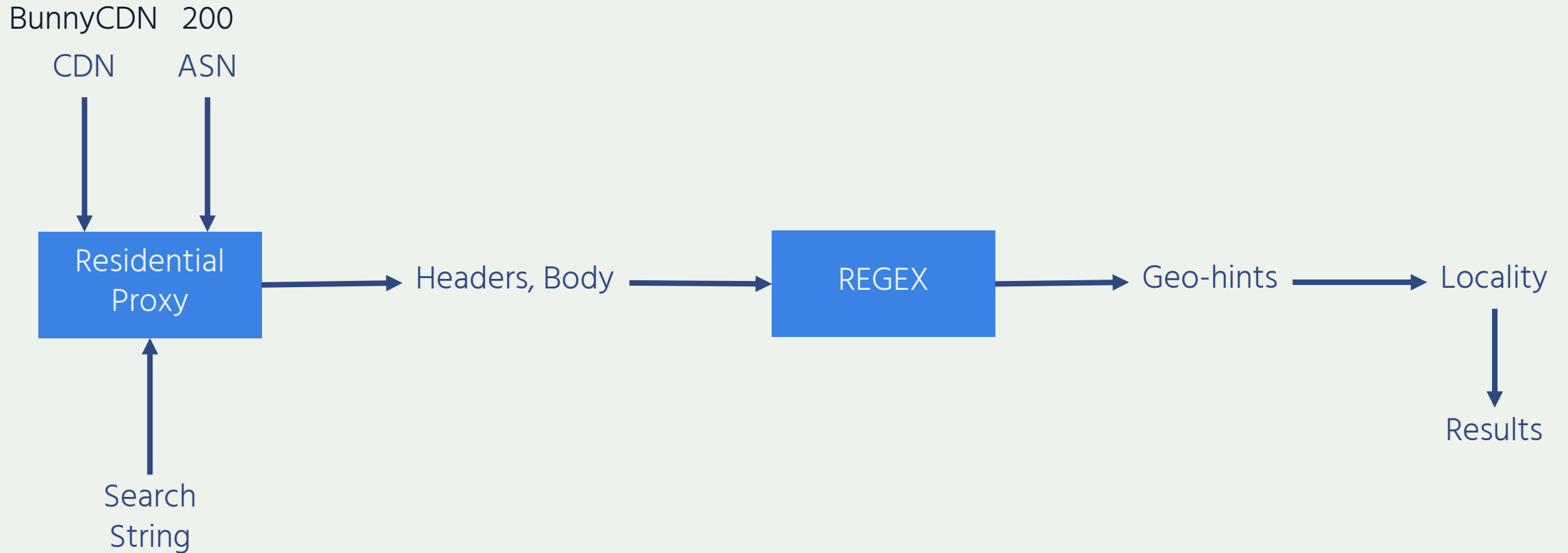
Methodology – Geolocate CDNs

BunnyCDN, AS 200, NL (Netherlands)



Methodology – Geolocate CDNs

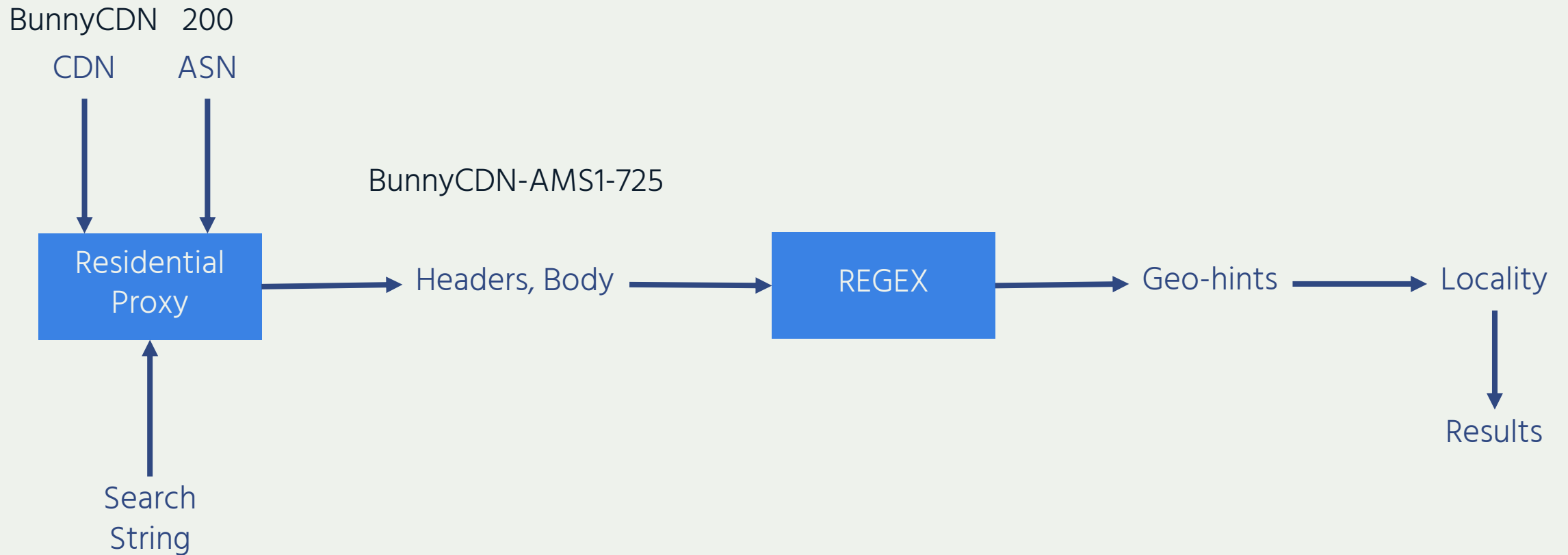
BunnyCDN, AS 200, NL (Netherlands)



<https://bunny.net/images/bunnynet-the-best-content-delivery-network-cdn.jpg>

Methodology – Geolocate CDNs

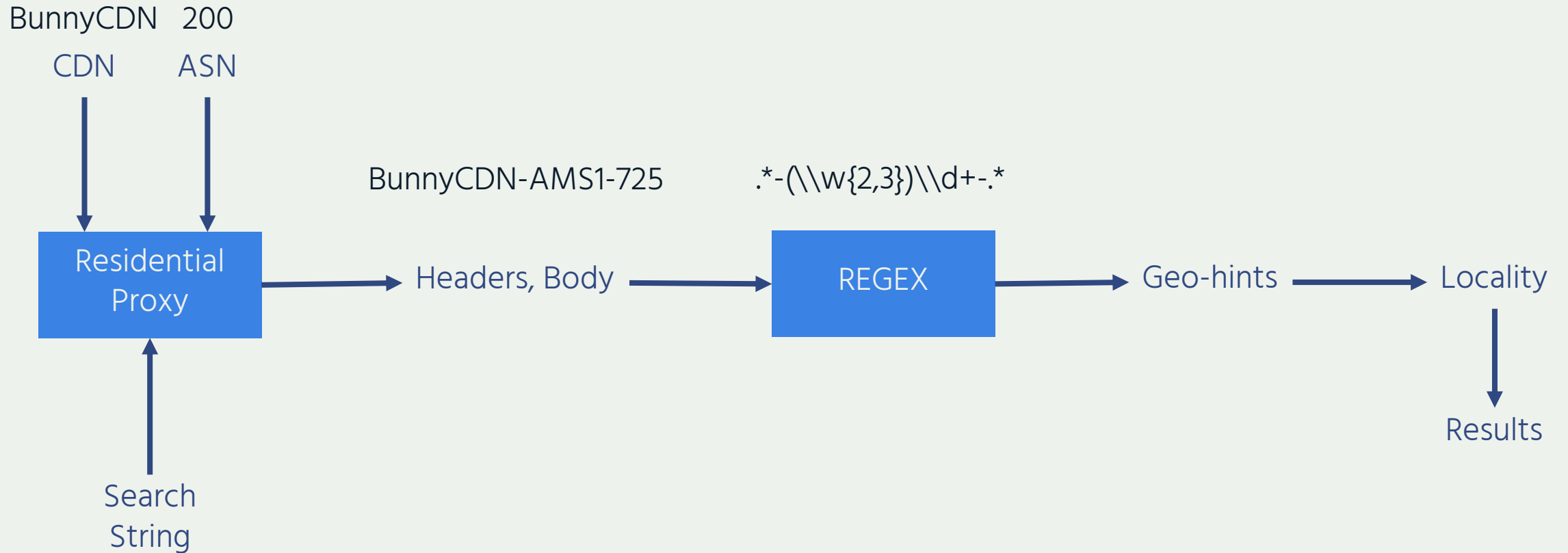
BunnyCDN, AS 200, NL (Netherlands)



<https://bunny.net/images/bunnynet-the-best-content-delivery-network-cdn.jpg>

Methodology – Geolocate CDNs

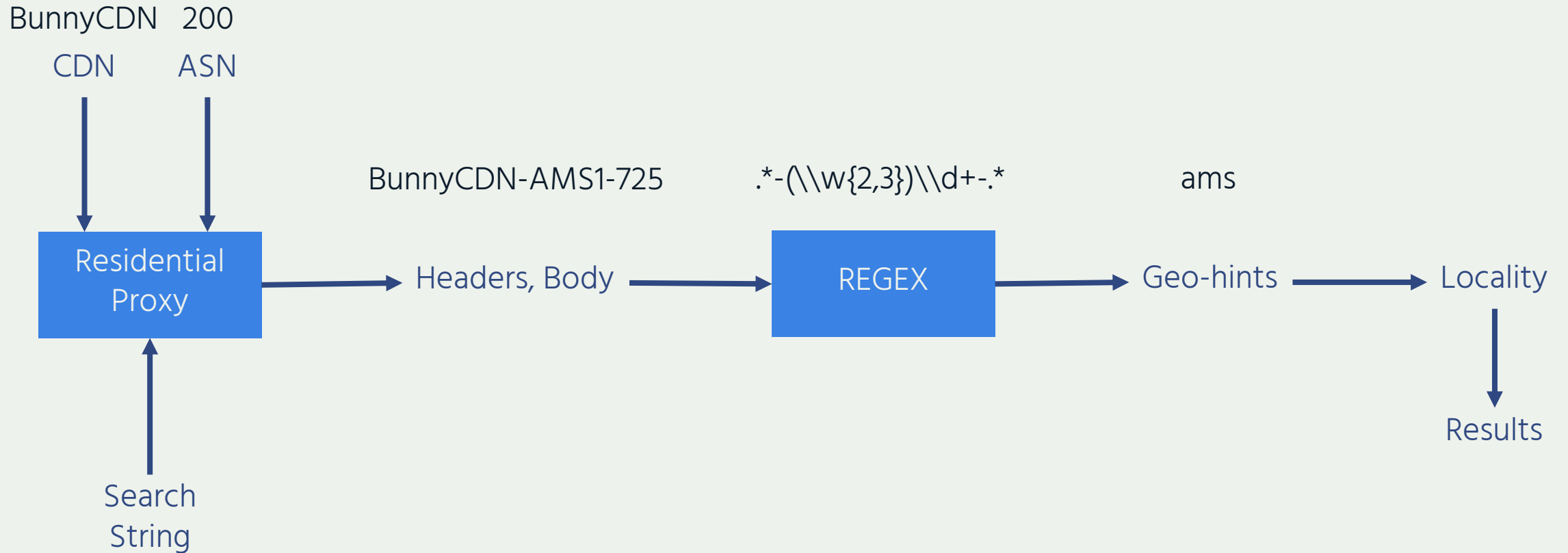
BunnyCDN, AS 200, NL (Netherlands)



<https://bunny.net/images/bunnynet-the-best-content-delivery-network-cdn.jpg>

Methodology – Geolocate CDNs

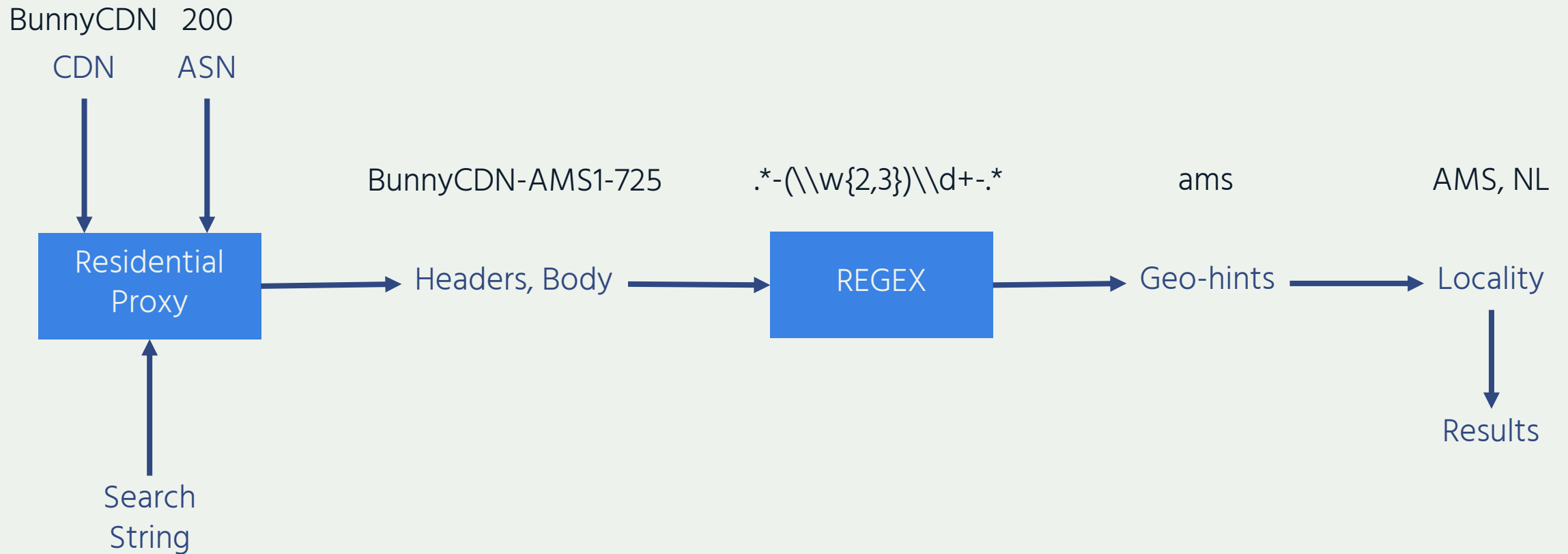
BunnyCDN, AS 200, NL (Netherlands)



<https://bunny.net/images/bunnynet-the-best-content-delivery-network-cdn.jpg>

Methodology – Geolocate CDNs

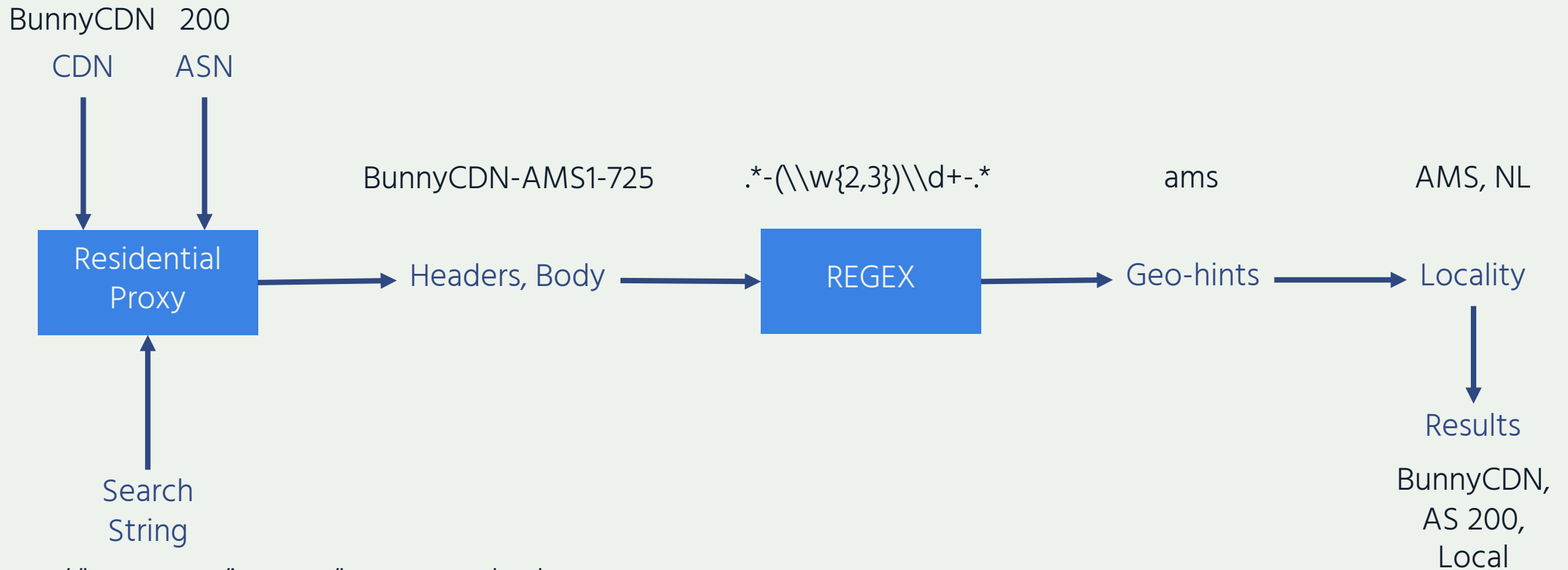
BunnyCDN, AS 200, NL (Netherlands)



<https://bunny.net/images/bunnynet-the-best-content-delivery-network-cdn.jpg>

Methodology – Geolocate CDNs

BunnyCDN, AS 200, NL (Netherlands)



<https://bunny.net/images/bunnynet-the-best-content-delivery-network-cdn.jpg>

Methodology – Combining Results

Methodology – Combining Results

CDN	ASN	Market Share	Local (Y/N)
BunnyCDN	AS 200	50%	Y
	AS 250	50%	N

Methodology – Combining Results

CDN	ASN	Market Share	Local (Y/N)
BunnyCDN	AS 200	50%	Y
	AS 250	50%	N

Hosting Method	Domain Count	Locality	Local Domains
BunnyCDN	10	50%	5
Native	5	20%	1

Methodology – Combining Results

CDN	ASN	Market Share	Local (Y/N)
BunnyCDN	AS 200	50%	Y
	AS 250	50%	N

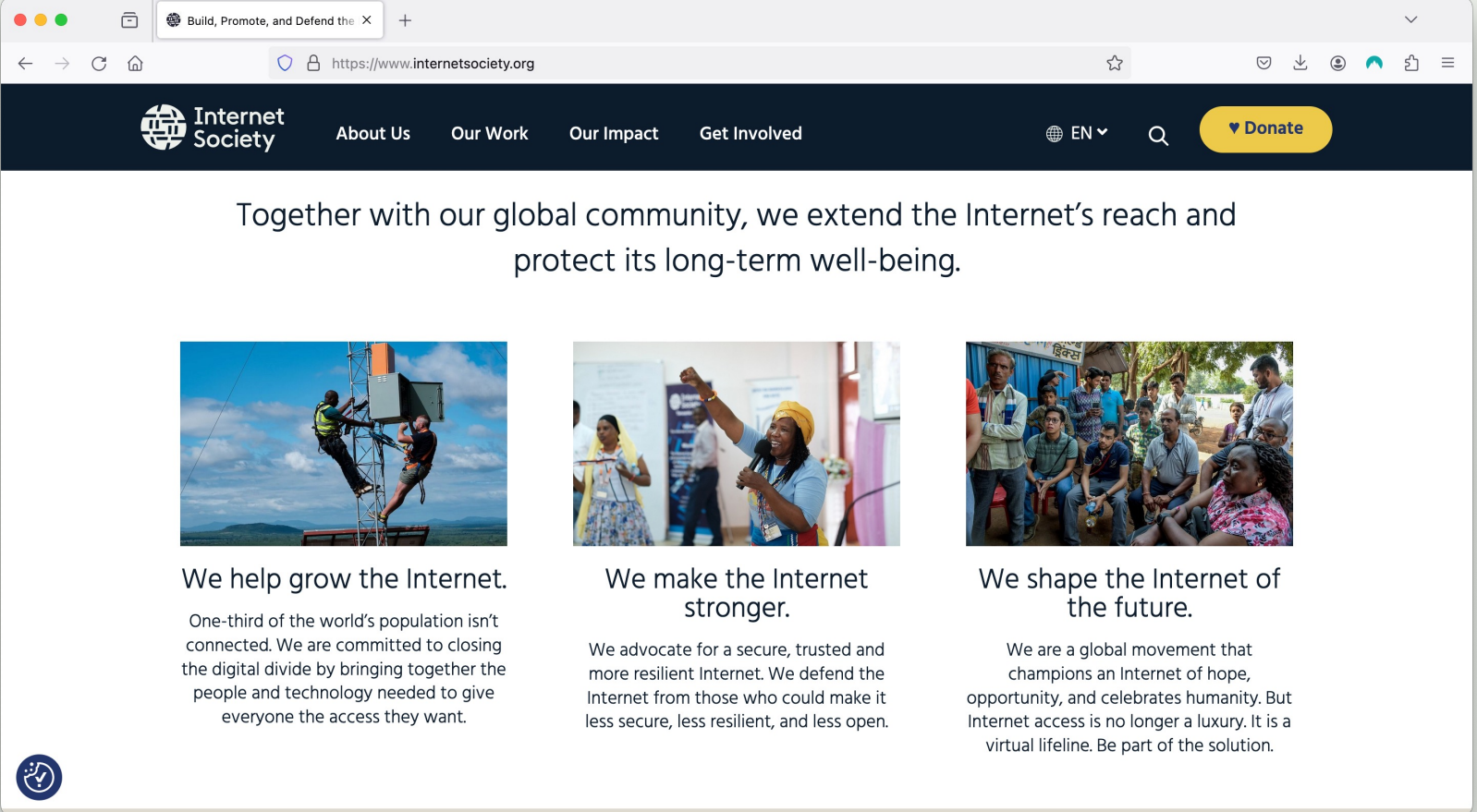
Hosting Method	Domain Count	Locality	Local Domains
BunnyCDN	10	50%	5
Native	5	20%	1

Country Locality	40%
------------------	-----

Limitations

Limitations


Website Complexity





The screenshot shows the Internet Society website homepage. The browser address bar displays "https://www.internetsociety.org". The navigation menu includes "About Us", "Our Work", "Our Impact", and "Get Involved". A "Donate" button is visible in the top right corner. The main content area features a central headline: "Together with our global community, we extend the Internet's reach and protect its long-term well-being." Below this headline are three columns, each with an image, a heading, and a paragraph of text.

Internet Society About Us Our Work Our Impact Get Involved EN [Donate](#)

Together with our global community, we extend the Internet's reach and protect its long-term well-being.

 **We help grow the Internet.**
One-third of the world's population isn't connected. We are committed to closing the digital divide by bringing together the people and technology needed to give everyone the access they want.

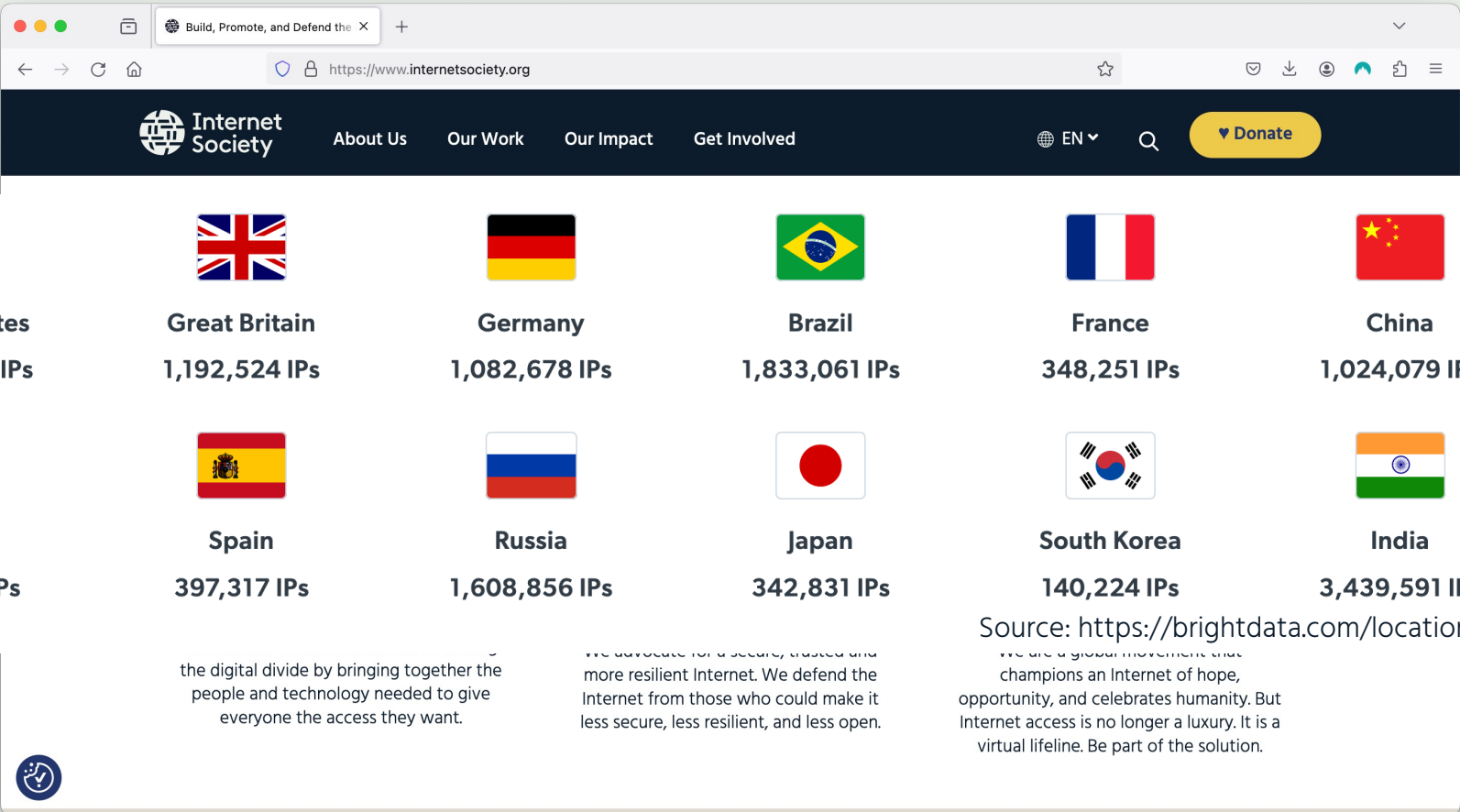
 **We make the Internet stronger.**
We advocate for a secure, trusted and more resilient Internet. We defend the Internet from those who could make it less secure, less resilient, and less open.

 **We shape the Internet of the future.**
We are a global movement that champions an Internet of hope, opportunity, and celebrates humanity. But Internet access is no longer a luxury. It is a virtual lifeline. Be part of the solution.

Limitations

Website Complexity

Proxy Probes



The screenshot shows the Internet Society website with a grid of 12 countries and their IP counts. The website header includes the Internet Society logo, navigation links (About Us, Our Work, Our Impact, Get Involved), a language selector (EN), a search icon, and a yellow 'Donate' button. The grid lists the following countries and IP counts:

Country	IPs
United States	3,291,464
Great Britain	1,192,524
Germany	1,082,678
Brazil	1,833,061
France	348,251
China	1,024,079
Canada	260,871
Spain	397,317
Russia	1,608,856
Japan	342,831
South Korea	140,224
India	3,439,591

Source: <https://brightdata.com/locations>

the digital divide by bringing together the people and technology needed to give everyone the access they want.

we advocate for a secure, trusted and more resilient Internet. We defend the Internet from those who could make it less secure, less resilient, and less open.

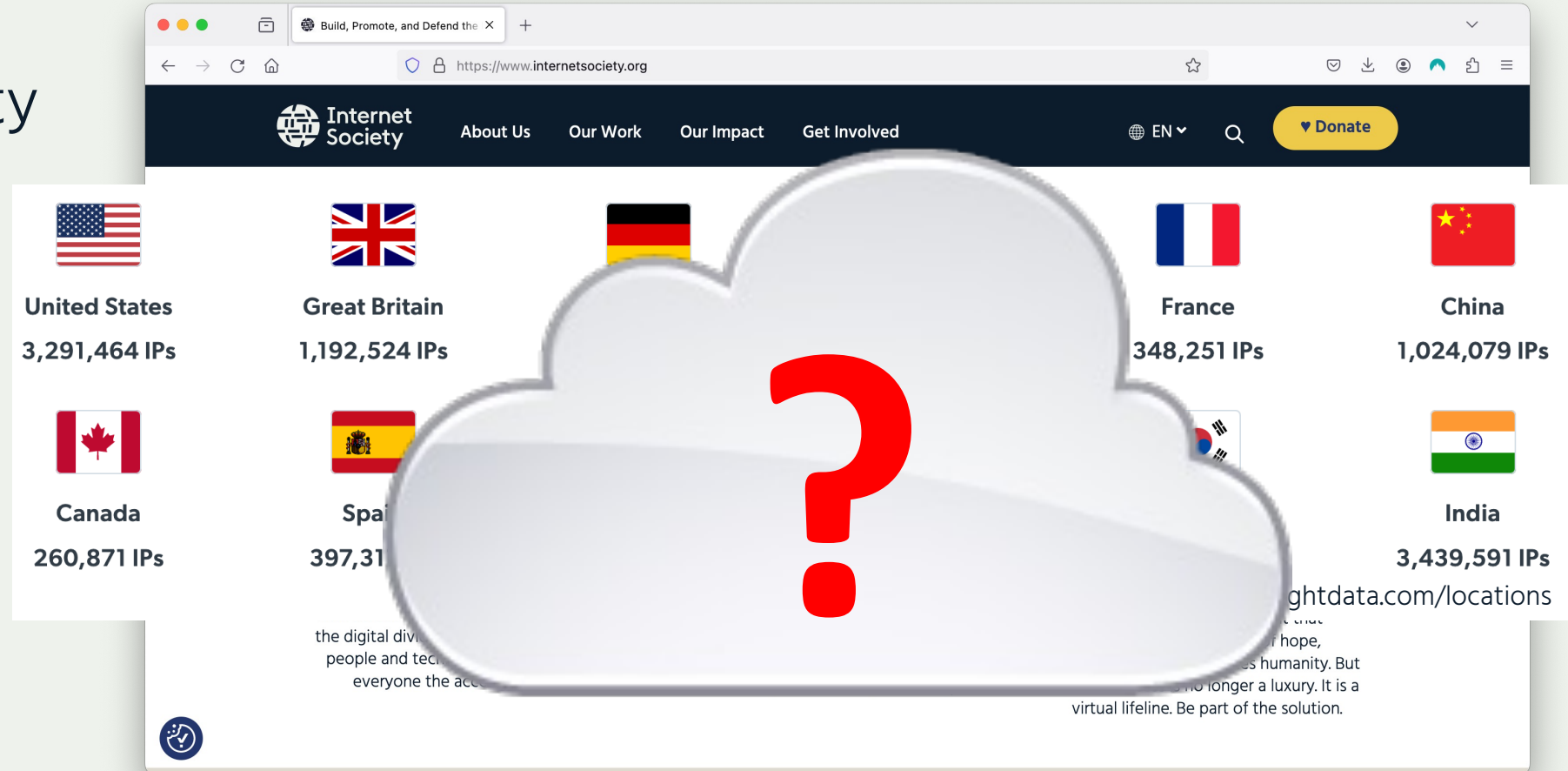
we are a global movement that champions an Internet of hope, opportunity, and celebrates humanity. But Internet access is no longer a luxury. It is a virtual lifeline. Be part of the solution.

Limitations

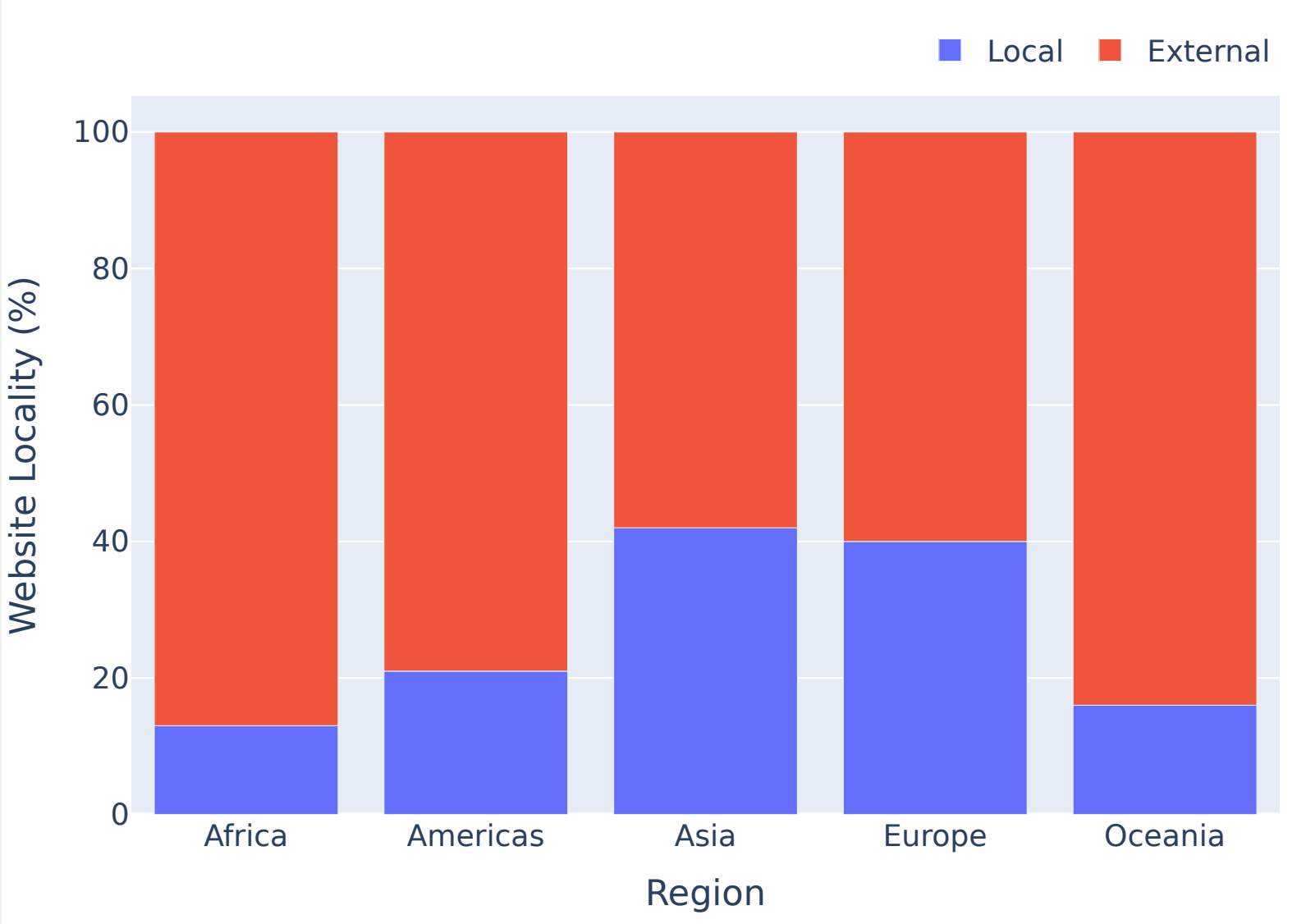
Website Complexity

Proxy Probes

“Fog of Cloud”

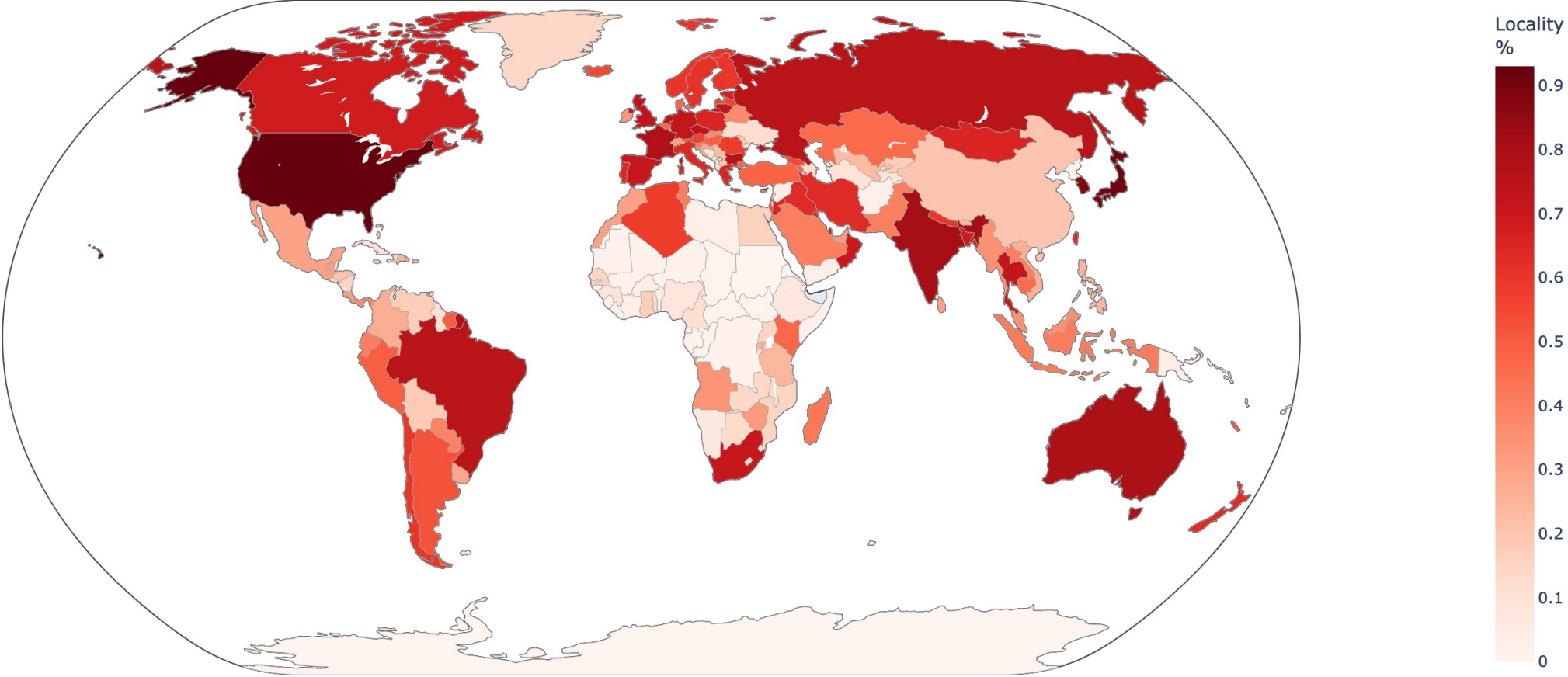


Preliminary Results – Regional Locality

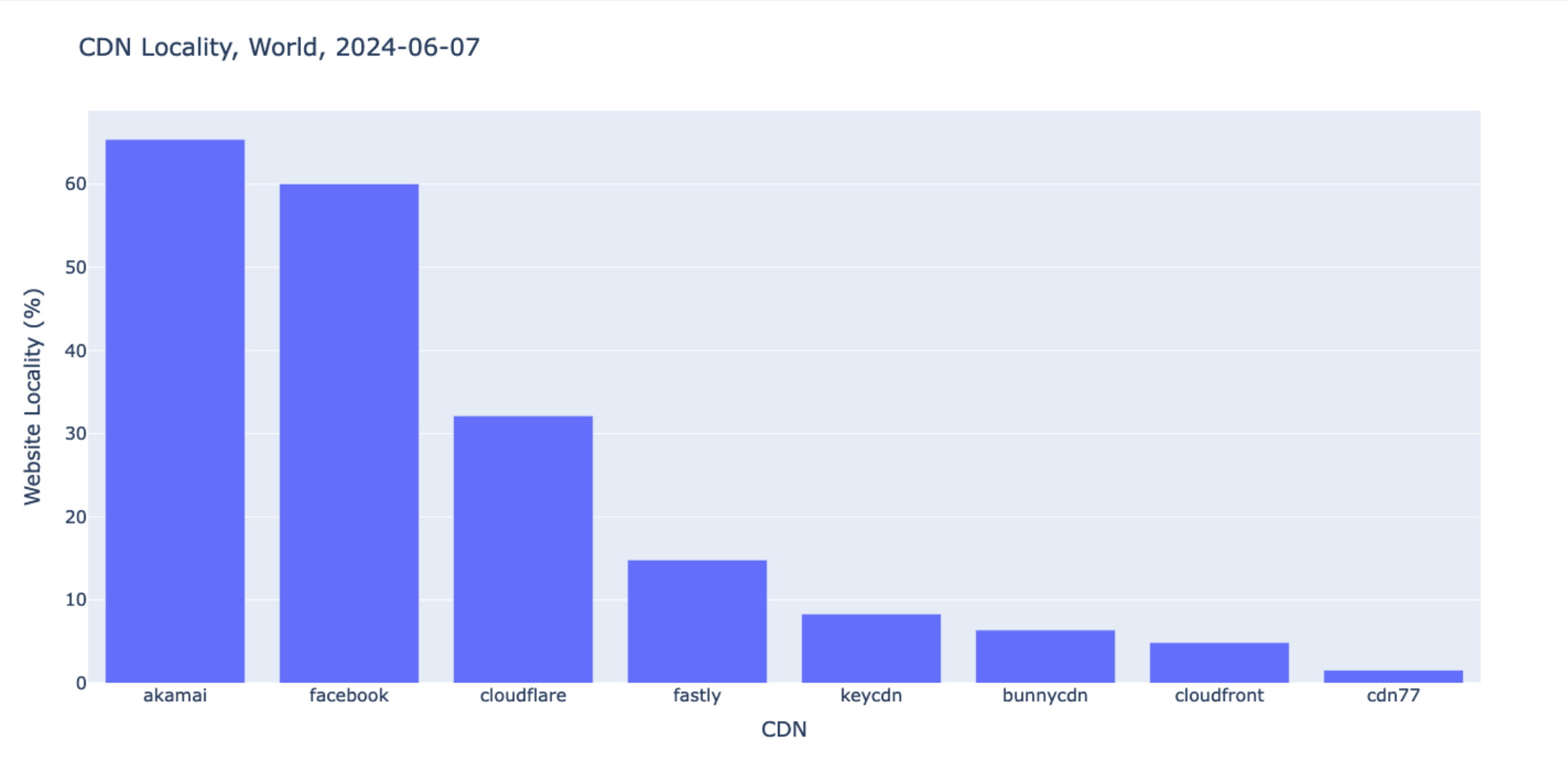


Preliminary Results – Global Locality

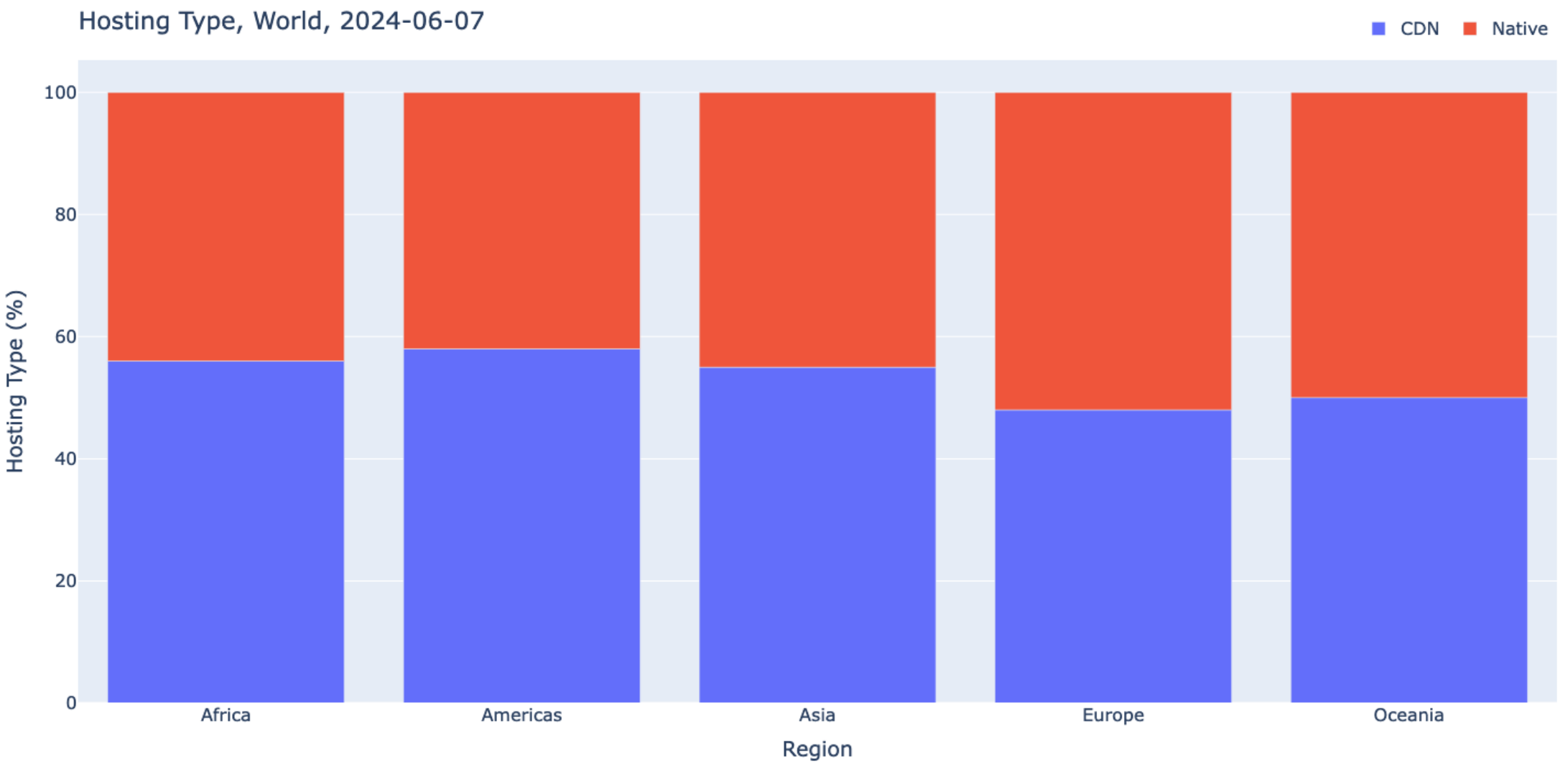
Locality Results: 2024-06-07



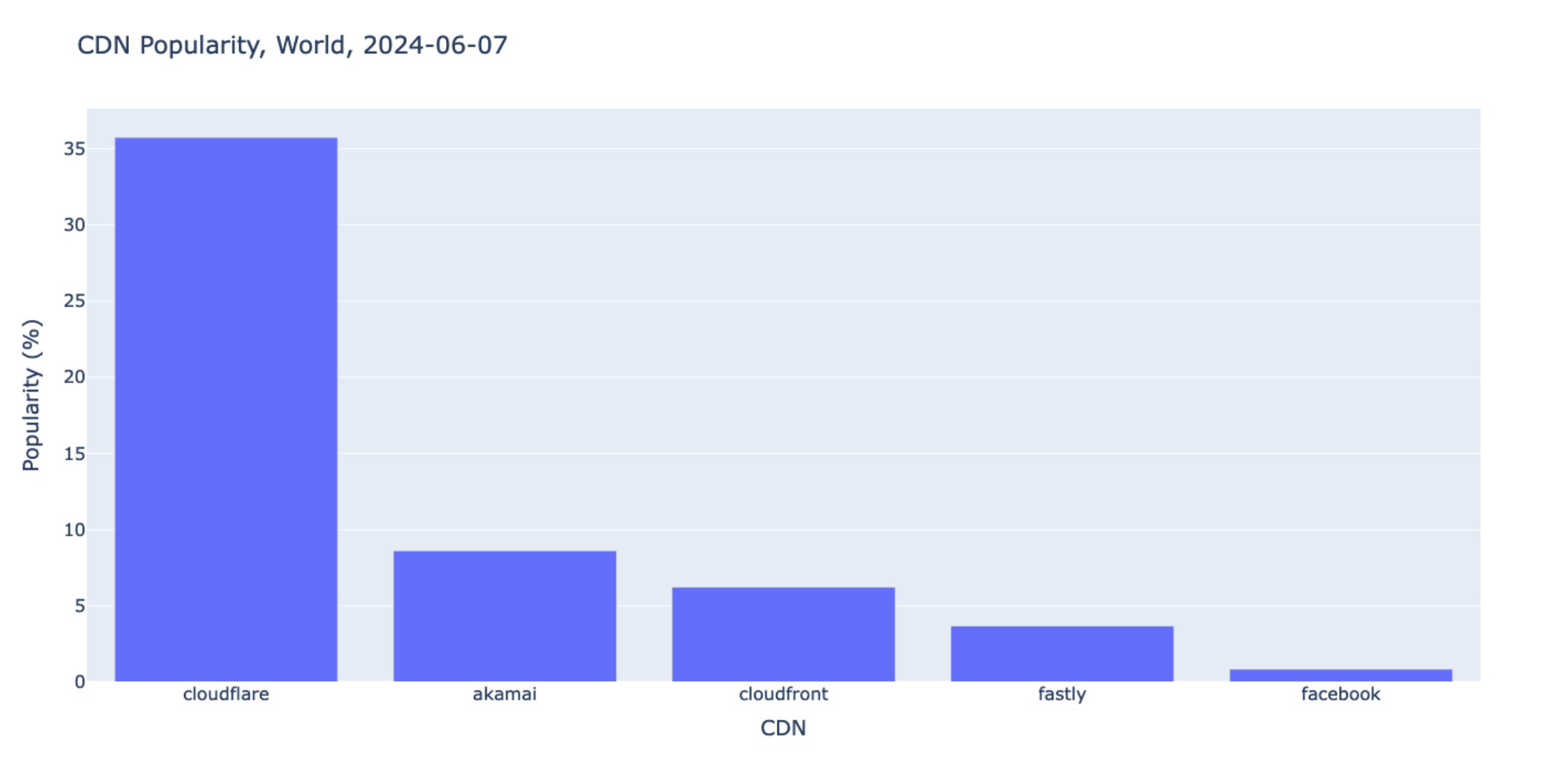
Preliminary Results – CDN Locality (World)



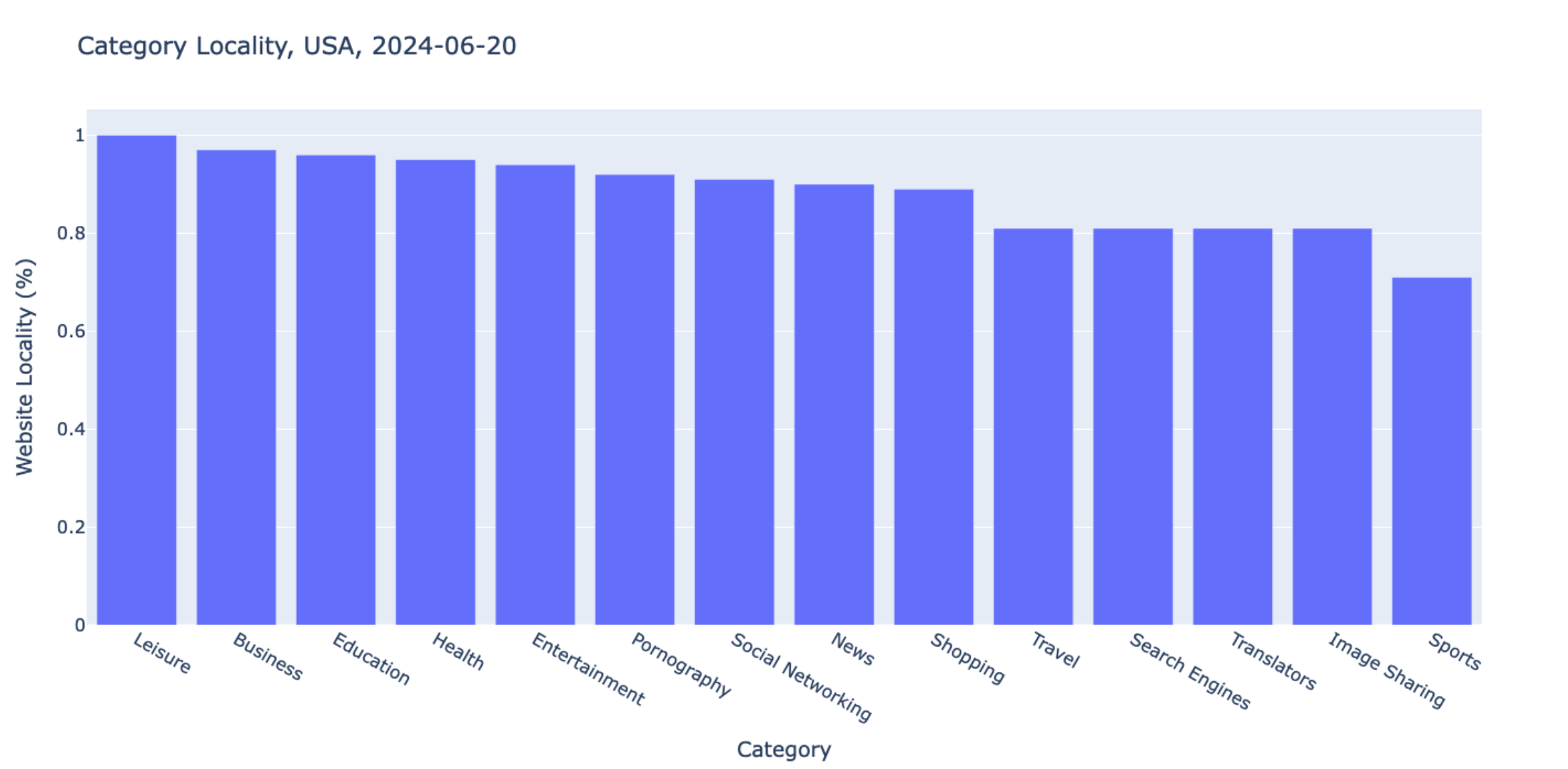
Preliminary Results – Hosting Type (World)



Preliminary Results – 5 Most Popular CDNs (World)



Preliminary Results – Categories Locality (USA, 100 Sites)



Future Work

Future Work

Recursive Search

Traceroute/Latency

Future Work

Recursive Search

Traceroute/Latency

Reasons for Locality

Visualisation

Future Work

Recursive Search

Traceroute/Latency

Reasons for Locality

Visualisation

Categorization

Summary

Local traffic is (often) beneficial

Conducting global Internet traffic locality measurements

Expanding methodology

Any questions?



Check out the [paper](#)



Learn about the
[50/50 Vision](#)