Formalizing Dependence of Web Infrastructure

Rumaisa Habib
Stanford University
Stanford, CA, USA

Kimberly Ruth
Stanford University
Stanford, CA, USA

Gautam Akiwate
Stanford University
Stanford, CA, USA

Zakir Durumeric
Stanford University
Stanford, CA, USA

ABSTRACT

Over the past decade, Internet centralization and its implications for privacy, resilience, and innovation have become a topic of active debate. While the networking community informally agrees on the definition of centralization, we lack a formal metric for quantifying it, which has limited in-depth analysis. In this work, we introduce a rigorous statistical metric for Internet centralization. In doing so, we also uncover how regionalization—geopolitical dependence on the Internet—fundamentally affects centralization. We argue that centralization and regionalization are intertwined forms of dependence that both affect the lived experiences of users and should be jointly studied. We develop a suite of statistical tools, which we use to better understand dependence across three layers of web infrastructure—hosting providers, DNS infrastructure, certificate authorities—in 150 countries. We hope that this statistical toolkit can serve as the foundation for future analysis of Internet behavior.

CCS CONCEPTS

• **Networks** → Network structure; • **Computing methodologies** → *Model development and analysis*; • **Information systems** → *Data mining*;

KEYWORDS

Centralization, Regionalization, Web, Hosting, DNS, PKI

ACM Reference Format:
Rumaisa Habib, Kimberly Ruth, Gautam Akiwate, and Zakir Durumeric. 2025. Formalizing Dependence of Web Infrastructure. In *ACM SIGCOMM 2025 Conference (SIGCOMM '25), September 8–11, 2025, Coimbra, Portugal.* ACM, New York, NY, USA, 22 pages. https://doi.org/10.1145/3718958.3750492

1 INTRODUCTION

Following the widespread adoption of cloud hosting and content distribution networks (CDNs), Internet centralization has become a hot topic of debate in the networking community. On the one hand, many large providers argue that their services improve performance, resilience, and security for their customers and, in turn, end users [14, 28]. On the other hand, researchers and members of the Internet standards community have noted that the Internet's original decentralization was one of the foremost reasons for its success and have warned that centralization can limit innovation, prevent competition, and pose privacy concerns [36, 43, 52].

Despite increased attention to the topic, our community lacks rigorous metrics to measure and reason about centralization. Prior work conceptually agrees that centralization is the *the concentration of an Internet function on a small number of providers* [18, 36, 43, 46, 51, 71], but, without a metric for directly measuring concentration, investigations have resorted to reporting measurements like the percentage of sites hosted by the top ten providers and analyzing the market share of well-known hyperscalers. While these studies have unquestionably shed light on Internet behavior, the lack of a rigorous metric makes it cumbersome to compare centralization across countries, infrastructure layers, and points in time. Existing descriptive measures also miss distributional nuances. For instance, measuring the total share of websites hosted by the top providers does not account for the distribution among these providers, which can dramatically affect resiliency and privacy.

We argue that we need a more rigorous approach for characterizing Internet behavior. Building on our community's informal consensus definition, we formulate requirements for a distilled, quantitative measure of Internet centralization. We formalize centralization as *the statistical distance of an observed distribution from a fully decentralized reference distribution*, which we quantify using Wasserstein distance, a popular integral probability metric. With this grounding, we measure and compare the centralization of three layers of web infrastructure—hosting/content delivery, authoritative DNS, and certificate authorities—across 150 countries.

Consistent with prior work [18, 33, 64, 68, 71], we observe that the largest providers like Cloudflare and Amazon play a crucial role in web dependence and drive centralization in a large part of the world. However, we also find that some countries are more centralized than their usage of top global providers suggests. For example, at times, large *regional* providers lead to greater consolidation than well-known hyperscalers like Google. Consolidation on regional providers makes users' lived experience of the web no less centralized, but implies a different locus of power that has gone undiscussed by focusing on global market share.

Investigating these local dependencies, we find that regionalization, the geopolitical dependence within and between countries, is deeply intertwined with centralization. Much like centralization, regionalization is a type of dependence that affects users' lived experiences on the web. However, understanding regionalization requires looking past the number of providers and analyzing those players in a broader geopolitical context. For instance, the Commonwealth of Independent States (CIS) countries (formed following the dissolution of Soviet Union) exhibit comparatively low centralization, but depend highly on Russian providers, which affects privacy, resilience, and innovation much like centralization. Understanding the interplay between centralization and regionalization requires a suite of statistical tools with which to ground analysis. Thus, in

SIGCOMM '25, September 8–11, 2025, Coimbra, Portugal

ACM ISBN 979-8-4007-1524-2/2025/09…$15.00
https://doi.org/10.1145/3718958.3750492

addition to our centralization metric, we present a set of measures for classifying the scale and global reach of providers in the broader context of geopolitical dependence.

In summary, our work contributes:

- The introduction of regionalization as a form of dependence deeply tied to centralization.
- A suite of metrics to enable researchers to capture Internet dependence systematically and comprehensively, including a statistical definition of centralization and new metrics supporting regionalization.
- An analysis of centralization and regionalization of hosting, DNS, and CA layers across 150 countries. We additionally describe TLD dependence in Appendix B. We surface geographical variations in centralization, the existence of regionally dominant providers, and cases of geopolitical dependence overlooked by prior work.
- The release of our data on web dependencies.

We conclude with observations across infrastructure layers and recommendations for the Internet community. Beyond shedding light on Internet dependency, we hope that our set of statistical tools can enable richer investigation of Internet behavior. Our data is available at https://doi.org/10.5281/zenodo.15733582.

2 RELATED WORK

There is a significant body of literature documenting centralization at different layers of the Internet, including cloud and hosting platforms [18, 33, 34, 64, 68, 71], DNS [3, 36, 51, 69], email [20, 49, 67, 70], certificate authorities (CAs) [43, 46], and third-party web resources [38, 65]. For example, Doan et al. [18] study web content (e.g., fonts and ads), finding that hosting services' increased dominance allows them to contribute to popularizing new Internet standards. Hoang et al. [33] find pronounced domain co-hosting at large providers. Several other studies highlight provider consolidation and single points of failure as a concern for web resiliency [3, 17, 43, 68]. Other work has considered why some customers gravitate to large providers and its implications [31, 35, 63, 66].

While most centralization research has focused on global concentration, several studies have noted country-level differences [32, 44, 46, 47]. Kumar et al. [46] analyze the Alexa Top 500 sites in 50 countries, finding that centralization has increased over time and is correlated with economic development. In a separate work, they also examine the hosting trends of government sites across 61 countries, uncovering bilateral relationships in serving content [47]. Zembruzki et al. [71] study hosting centralization for websites in 19 TLDs, finding evidence of geographic and language patterns in hosting provider dependence.

Another set of work has examined web infrastructure within specific regions. For instance, Jonker et al. [41] describe Russian domain infrastructure. Several prior studies describe web infrastructure in Africa, including single points of failure [42] as well as regional network infrastructure and trends of hosting abroad [26, 27]. Helles et al. [32] cluster top-150 EU websites with similar third-party dependencies to show regional variation. Li et al. study differences in China [48]. More globally, Fan et al. [25] measure how CDNs map users to front-end clusters. Our work is complementary, analyzing how centralization differs across regions and how sociopolitical patterns shape it.

Prior studies on web centralization have largely relied on summary statistics, ranging from collections of raw numbers [17, 18, 33, 34, 51, 64, 71] and top-k thresholds [18, 25, 36, 43, 46, 64, 68] to custom metrics [18]. While these works have unquestionably shed light on centralization, our work shows how a more quantitative, statistical approach can help us to uncover further nuance in centralization. Most notably, our work considers *all* providers in a region, unlike prior work that focuses on the top few above a specific cutoff. Most prior works [17, 43, 46, 51] consider only the top five providers when assessing centralization; one paper [47] considers 28 providers (the maximum we found).

Beyond web content, Moura et al. [51] analyze DNS queries seen at authoritative DNS servers to investigate centralization on cloud providers. Clark and Claffy [11] discuss topological, rather than geographical, regionalization in Internet routing. While there has been a significant body of work studying concentration trends in network topology (e.g., [5, 30]), here we address settings in which the data is not necessarily a graph (e.g., how many websites rely upon certain providers). Most closely related methodologically, two investigations [3, 36] use Herfindahl-Hirschman Index (HHI) to study centralization on DNS providers.

While this large body of prior work has effectively called attention to centralization, our work argues that to effectively move the discussion forward, we must formalize our definition of centralization and how we measure it.

3 QUANTIFYING DEPENDENCE

In this section, we introduce our suite of metrics for quantifying dependence on the web. We first formalize our community's intuitive notion of centralization (Sections 3.1 and 3.2) and introduce a statistical measure for quantifying it, which allows comparison across countries and layers of infrastructure. We then introduce additional measures (endemicity, usage, and insularity) to describe facets of regionalization (Section 3.3).

3.1 Defining Centralization

While the Internet community has not formalized a statistical definition of centralization, there is informal agreement that centralization is *the concentration of an Internet function on a small number of providers* [18, 36, 43, 46, 51, 71]. While intuitive, this definition is difficult to quantify: "concentration" and "small" are ill-defined, limiting researchers to self-evident cases. Without a direct measure, prior work has most often quantified centralization by describing the market share covered by the top N providers (e.g., top 10 providers). While a useful first-cut heuristic, this approach inherently captures just one point in the distribution and can be misleading.

For example, consider Azerbaijan (AZ) and Hong Kong (HK), which both have 59% of their top sites hosted by five providers. While seemingly very similar, looking at the distributions of providers (Figure 1), we see missing context: Azerbaijan's top two providers cover 42% and 5% of market share, while Hong Kong's cover 33% and 12%. Intuitively, Azerbaijan's steep drop-off should mean it is *more* centralized than Hong Kong, but the top-5 quantification

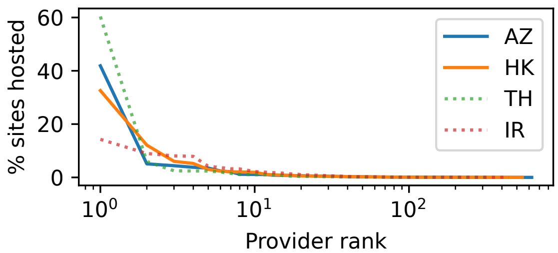


Figure 1: Example of Top-N Metric Shortcoming—Azerbaijan (AZ) and Hong Kong (HK) both have 59% of sites run by their top five hosting providers, but their distribution *within* the top five differs substantially. (Highly centralized Thailand (TH) and highly decentralized Iran (IR) shown for reference.)

says it is identically so. No single value of N is free of this problem, and juggling multiple values of N makes comparisons (e.g., across countries or infrastructure layers) unwieldy.

To design a more principled metric, we translate the community's informal definition into a set of requirements. Our metric should:

(1) account for both the number of providers and the distribution of the Internet function across those providers in one metric;
(2) accommodate comparing highly skewed distributions with few dominant providers and a long tail of others;
(3) facilitate fair comparisons across scenarios (e.g., across countries or over time), independent of the specific providers underlying the distribution;
(4) match human intuition of centralization in the context of Internet functionality.

We propose formalizing centralization as *the distance of an observed distribution of dependencies from a fully decentralized reference distribution.* Namely, we compare observed provider distributions against a hypothetical uniform distribution where every website has its own unique provider. Our decentralized distribution does not express an "ideal" or even a possible state of the world, but rather serves as a reference relative to which all other distributions can be compared. The observed distribution with the greatest distance from the fully-decentralized distribution is the most centralized.

This formulation allows us to consider a wide range of statistical measures for quantifying distance, including f-divergence functions (e.g., Kullback–Leibler (KL) divergence, Jensen–Shannon divergence, Hellinger distance, and total variation distance) and integral probability metrics (e.g., Wasserstein distance, Dudley metric, and maximum mean discrepancy). While f-divergence functions like KL-divergence are popular for measuring distribution dissimilarity, we find them unsuitable for our particular task. The f-divergence class works well only when comparing two distributions that are largely overlapping (the f-divergence between any two fully disjoint distributions is constant). In our case, we are comparing two fundamentally different distributions: a real-world observed distribution that is heavily skewed towards a few large providers and a hypothetical uniform distribution where every service has its own provider.

We note that we cannot simply compare observed provider distributions directly against each other since there would not be a clear directionality as to which distribution is more centralized; instead, we compare all observed distributions against a reference distribution of absolute zero centralization. Ultimately, we decide to use Wasserstein's distance, one of the—if not the—most commonly used integral probability metrics [7] that captures the distance between distributions even when they do not significantly overlap.

3.2 Quantifying Centralization

We use Wasserstein distance—also commonly known as *Earth Mover's Distance* (EMD) in the computer vision community [58, 59]—to measure centralization. The metric earns this name because it effectively considers two distributions as a mass of earth spread out in space and computes the minimum amount of "work" needed to transform the first distribution into the second, where a unit of work corresponds to transporting a unit of earth by a unit of ground distance according to a customizable ground distance function. The total work is the product of the mass of earth moved and the distance it was moved.

We apply EMD to our use case by measuring the work that would be required to transform an observed distribution of providers $A = (a_1, \ldots, a_n)$, where each a_i represents the number of websites using provider i, into the completely decentralized reference distribution $R = (r_1, \ldots, r_m)$ where every website uses a unique provider. In such a completely decentralized distribution, $r_j = 1\ \forall j$ (i.e., each provider j is only used by one website), and $m = C$ where $C = \sum_i a_i$ is the total number of websites considered. We define our ground distance metric d_{ij} between $a_i \in A$ and $r_j \in R$ as the difference between a_i and r_j, normalized by the total number of websites. Note, then, that websites using more popular providers must travel a greater distance toward a fully decentralized distribution compared with websites using less popular providers. Intuitively, the distribution that is the most centralized is the one that would require the greatest work to become fully decentralized and thus has the largest value for our centralization score (Figure 2).

Writing out this definition more formally yields the following expression for our *Centralization Score* ($\mathcal{S}$):

$$\mathcal{S} = \sum_{i=0}^{n} \left(\frac{a_i}{C}\right)^2 - \frac{1}{C}$$

Note that we can read $\frac{a_i}{C}$ as the fraction of websites that provider i hosts, and that the upper bound on $\mathcal{S}$ is $1 - \frac{1}{C}$, which approaches 1 as a larger C is chosen. Readers familiar with the Herfindahl–Hirschman Index (HHI), a metric used to measure competition between market entities in US antitrust law [53, 54], will recognize $\sum_{i=0}^{n} \left(\frac{a_i}{C}\right)^2$ as the formula for HHI applied to our setting. Thus, HHI is effectively a special case instantiation of EMD up to a constant. Further details are provided in Appendix A.

This instantiation of EMD fulfills our aforementioned goals for a centralization metric:

(1) The score accounts for both the size of providers (the larger the provider, the more it contributes to $\mathcal{S}$) and the number of providers (the longer the tail, the more providers contribute small d_{ij} terms to the sum).
(2) It makes no assumptions about the similarity or overlap between the distributions we are comparing, and provides

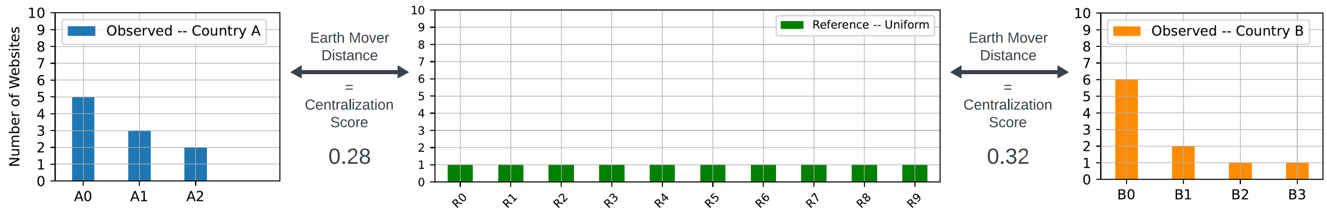


Figure 2: Centralization Comparison Example—To calculate our centralization score for the top websites in Countries A and B, we calculate the distance between the observed distribution in each to a reference uniform distribution using Earth Mover's Distance (EMD). In the example above, the EMD for Countries A and B are 0.28 and 0.32, respectively, indicating that Country A is *less centralized* than B.

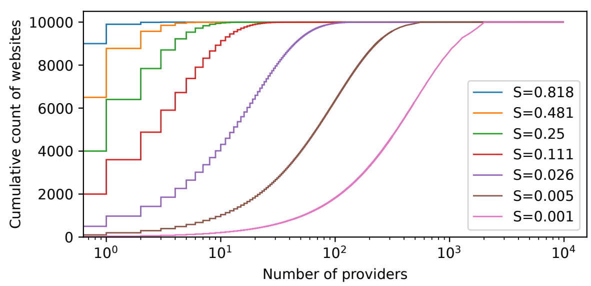


Figure 3: Example $\mathcal{S}$ Values—Centralization Score ($\mathcal{S}$) for multiple synthetic distributions. $\mathcal{S}$ values are most sensitive to differences between the highly centralized cases.

a meaningful measure of distance even for highly skewed distributions.

(3) As long as C is held constant throughout an analysis, the metric can fairly compare against multiple scenarios: it is based only upon the shape of the curve, not on the providers comprising it.

(4) The notion of quantifying the "work" required to flatten the data into a fully decentralized distribution is appealing in the context of centralization: the more concentrated the underlying distribution, the more "work" and the greater the $\mathcal{S}$ value. The largest providers are weighted most heavily in the metric: a provider's contribution to $\mathcal{S}$ grows quadratically with its market share among the set of websites considered, which empirically we find matches our intuition of which providers most contribute to centralization.

While we do not specify a cutoff for how large of $\mathcal{S}$ corresponds to "centralized," we provide a set of example curves in Figure 3 to aid intuitive interpretation of $\mathcal{S}$ values. Since $\mathcal{S}$ is closely related to HHI, the HHI interpretation guidelines used by the U.S. Department of Justice for defining market competition and monopolies can also provide context for how other fields think about concentration: 'competitive' (<0.10), 'moderately concentrated' (0.10–0.18), and 'highly concentrated' (>0.18) [53, 54]. We also note that the Internet community's interpretation norms for centralization scores may differ by the type of provider considered. For example, it might be expected, and even desired, that CAs are more centralized than hosting providers (and indeed, we observe this in Section 7).

Last, we note that EMD is customizable and provides a framework for future work to consider other aspects of Internet centralization. While we choose a specific reference distribution R and ground distance function d_{ij} that suit our goals, future work may want to explore other parameters. For example, a study looking at how countries rely on specific providers may wish to redefine d_{ij} and compare countries' distributions pairwise rather than using a reference distribution. Another natural extension would be to assign a weighted "mass" to each website (e.g., based on traffic), rather than weighting all sites equally. Other applications may involve customizing a_i: for instance, a study of provider redundancy may define it as the number of websites that require provider i to function, and a study of third-party trackers may define it as the number of websites that fetch tracker i.

3.3 Quantifying Regionalization

Centralization alone is not enough to describe Internet dependence, as it lacks geopolitical context. To describe regionalization—the geopolitical dependence within and between countries—we build metrics to quantify two concepts: the global reach of *providers* and the entanglement of *countries.* We use these metrics in Sections 5–7 to contextualize countries' dependence.

Providers. Providers vary both in their overall usage (i.e., sheer scale) and in their geographic usage distribution (i.e., concentration of usage in certain countries). To illustrate these differences, we consider each provider's *usage curve* similar to Ruth et al. [60]: we compute the percentage of popular websites in each country that use the provider, sort countries by decreasing usage, and plot the resulting list of usage values. We arrange these values as a nonincreasing sequence $(u_1, u_2, \ldots, u_n)$ where $n = 150$ is the total number of countries considered. Figure 4 shows two examples using our hosting provider data from Section 5. The first is of a large globally popular provider like Cloudflare, which is used by a significant number of websites in many countries; the second is of Beget LLC, a Russian provider mainly used in CIS countries. First, we define *usage* (U) to be the area under the usage curve: $U = \sum_{i=1}^{n} u_i$. This metric captures total usage across the countries in the dataset. In Figure 4, the total usage of the example global provider is larger than the regional one.

Second, we define *endemicity* ($\mathcal{E}$) to be the area between the usage curve and the flat line starting at the usage curve's maximum value: $\mathcal{E} = \sum_{i=1}^{n}(u_1 - u_i)$. This captures the deviation from absolute global consistency in usage (i.e., the "flatness" of the usage

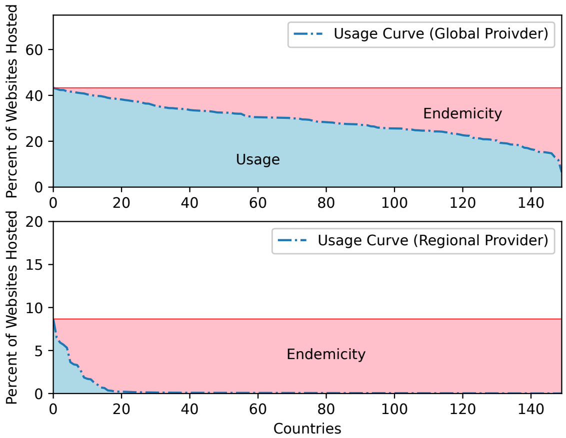


Figure 4: Usage and Endemicity—Usage (U) is the area under the usage curve and endemicity ($\mathcal{E}$) is the area between the usage curve and the horizontal line starting at the usage curve's maximum value. Usage captures popularity, while endemicity captures global consistency in usage. Regional providers have a higher endemicity than global providers.

curve), with a priority on capturing unusual popularity in a country rather than unusual unpopularity. We adopt this metric from Ruth et al. [60], with one modification: we normalize by provider size by computing the *endemicity ratio* as $\mathcal{E}_R = \frac{\mathcal{E}}{U+\mathcal{E}}$. Without this modification, the range of possible endemicity values is a function of the provider's maximum percent use in any country. Though endemicity ratio is not the only method of correcting for this [60], we find it to be the most natural for our use case. Values of $\mathcal{E}_R$ range from 0 to 1, with smaller values indicating more global reach and larger values indicating regional concentration. In Figure 4, the example regional provider is *more endemic* than the global one.

In summary, the area under the usage curve tells us how "large" the provider is globally, while the endemicity ratio captures how "regional" versus "global" the provider is.

Countries. Beyond providers, we also aim to describe countries' international dependence. To this end, we define the *insularity* of a layer for a country as the fraction of websites for which that layer is served by a provider based in the same country. For example, the hosting layer in the U.S. has an insularity of 92.1%, which is the fraction of its websites that are hosted in the U.S. Insularity captures how self-sufficient a country is in terms of infrastructure; it provides a foundation for investigating where countries have the most significant reliance on other countries.

3.4 Data Collection

Equipped with our suite of metrics for quantifying dependence, we analyze the centralization and regionalization of several layers of the web: hosting providers (Section 5), DNS providers (Section 6), Certificate Authorities (Section 7), and TLDs (Appendix B), across 150 countries. Here, we describe how we collect and enrich our data about web infrastructure.

Popular Websites. Our analysis is based on the public Chrome User Experience Report (CrUX) [10], which lists the websites most visited in each country, grouped in rank-magnitude buckets, as seen by Chrome browser telemetry. Prior work has shown CrUX most accurately captures popular websites compared to other top website lists [61]. Due to differing traffic volumes per country and Google's privacy safeguards, country popularity lists differ in length: smaller countries and countries with lower Chrome adoption have fewer websites in the dataset. To facilitate comparisons across countries, we analyze the top 10K websites in each of 150 countries whose top lists are at least that long (countries listed in Appendix E). This covers a substantial fraction of countries (63.3%) and traffic on the internet: the top 10K websites cover about 70–80% of web traffic globally [60].

Collecting Network Data. Building on the CrUX data, we perform active measurements for each website and annotate this dataset with additional third-party data. We resolve the 588K total domains using ZDNS [40] during May 2023, geolocate IPs with NetAcuity [24] and add origin AS using pfx2as [9], AS WHOIS organization and country using CAIDA's AS to Organization tool [8], and anycast configuration using `bgp.tools` anycast dataset [4]. Finally, we use ZGrab2 [19] to attempt a TLS handshake with the website on each resolved IPv4 address; following each handshake, we parse the leaf certificate and label CA ownership using the Common CA Database (CCADB) [6] per Ma et al. [50]. Using this enriched toplist data, we calculate the centralization on (1) hosting providers using the AS Organization of the IP address serving the content, (2) the DNS infrastructure using the AS Organization of the nameserver IP address, and (3) the CA using the CA Owner for the certificate served at the hosting IP.

Vantage Point. We perform measurements from Stanford University, but we validate our results are not significantly affected by vantage point selection in a secondary experiment. In December 2024, we conduct the same set of measurements through geographically distributed RIPE Atlas [57] probes. For each of the countries we consider, we conduct DNS A record lookups of sites in its CrUX list through randomly selected RIPE probes in that country. There were 14 countries with no RIPE probes; we select random probes for each measurement for these countries. We then recalculate centralization scores based on the AS Organization of the IP addresses returned by these probes. We find that the centralization scores for hosting through our university vantage point highly correlate with centralization scores based on RIPE measurements ($\rho = 0.96, p \ll 0.05$). As a result, we do not expect that our vantage point fundamentally affects the results in the study.

Interpreting Statistics. We apply our numerical definitions of centralization ($\mathcal{S}$), usage (U), and endemicity ratio ($\mathcal{E}_R$) as defined in Sections 3.2 and 3.3 to our data on hosting, DNS, CA, and TLD use. In addition, we use Pearson's correlation coefficient ρ as needed to compute correlation between rank-ordered sequences. We follow these guidelines for interpreting correlation coefficients: <0.30 is poor, 0.30–0.60 is fair, 0.60–0.80 is moderate, and >0.80 is strong [2].

Limitations. Our analysis relies on Chrome website popularity data, which is compiled from Chrome users who opt into data sharing; it may exhibit bias due to excluding incognito browsing, and it may be less representative of countries with lower Chrome adoption. Nevertheless, prior work has shown that it is the most accurate popularity data available [61]. Our NetAcuity geolocation data is imperfect, and we may mislabel IPs. However, prior work has shown that NetAcuity outperforms other geolocation providers, and that it is 89.4% accurate at the country level [29].

For our active measurements, we determine hosting providers by analyzing the autonomous system that serves the root page of each site. As such, we only have visibility into the last leg of content delivery. In the case of CDNs, we see the provider who *serves* the content, but acknowledge that this may not be the provider who authoritatively *hosts* the website's content. We label the provider that serves the website to end users as its "hosting provider" in this work. In the context of centralization, the last leg provider has significant ramifications for resilience, privacy, and innovation, even if the backend content storage differs.

4 ETHICS

Our work does not involve human subjects and according to our institution's policies, does not require IRB approval. On the server we use for performing active measurements, we follow the guidelines outlined by Durumeric et al. [21, 22] for identifying server purpose and ownership. We did not receive any opt-out requests. We seek to avoid making a value judgment on Internet dependency, but rather, to provide the research community with a suite of tools by which to measure centralization and regionalization, and make observations on how they are operationalized today. We also stress that we are a team of researchers located in the United States discussing the Internet architecture of many countries and world-regions that we are not a part of. Researchers in other regions and shaped by different cultural contexts may experience Internet dependencies in different and even personal ways. As such, our identities both shape and potentially limit our findings.

5 HOSTING PROVIDERS

In this section, we explore how centralized and regionalized the web is in terms of the providers that serve popular websites. We note that since many websites serve content using CDN or DDoS protection providers like Cloudflare, we are analyzing the provider that serves content to users, not the provider that authoritatively hosts each website. These CDNs are integral to the resiliency of these websites [15], and are hence important to discuss in the topic of dependence.

5.1 Centralization

Consistent with prior work, we observe the concentration of hosting on a small number of providers: 90% of websites are hosted by fewer than 206 providers in every country. However, as can be seen in Figure 5, there is significant variance in the degree to which countries have centralized on hosting providers. In the most extreme (i.e., most centralized) case, 60% of websites in Thailand ($\mathcal{S} = 0.3548$) are served by a single provider, Cloudflare. At the other extreme, Iran is the least centralized country with the top provider accounting for only 14% of websites and 90% of websites distributed across 80 providers ($\mathcal{S} = 0.0411$). The United States is the median country of the dataset ($\mathcal{S} = 0.1358$). We provide a full list in Appendix F.

While it is difficult to predict how centralized an individual country will be, there are subregional patterns, as can be seen in Figure 9. Southeast Asia is the most centralized ($\bar{\mathcal{S}} = 0.2403$) and includes the four most centralized countries: Thailand, Indonesia, Myanmar, and Laos. In contrast, Central Asia is the least centralized ($\bar{\mathcal{S}} = 0.0788$), with Turkmenistan as the second-least centralized country. European countries are, on the whole, consistently less centralized, with Slovakia, Russia, and Czechia in Eastern Europe being the least centralized countries in the region ($\bar{\mathcal{S}}_{EU} = 0.0994$, $\bar{\mathcal{S}}_{EastEU} = 0.0803$). While Africa, as a whole, does not tend toward any extreme ($\bar{\mathcal{S}} = 0.1553$), countries in Northern Africa (Libya) and Eastern Africa (Somalia) are more centralized than the global average ($\bar{\mathcal{S}} = 0.1429$, var = 0.003). As we will discuss in Section 5.3.2, this is largely due to using providers on other continents.

Class	Providers	Description	Example
XL-GP	2	Extra Large Global	Cloudflare
L-GP	6	Large Global	Akamai
L-GP (R)	2	Large Global (Regional)	OVH
M-GP	22	Medium Global	Incapsula
S-GP	73	Small Global	Wix
L-RP	174	Large Regional	Alibaba
S-RP	587	Small Regional	Loopia
XS-RP	11,548	Extra Small Regional	Forthnet

Table 1: Classes of Hosting Providers—We categorize providers based on how many top websites they serve (usage) and their usage consistency across countries (endemicity).

The total number of providers in a country has little effect on how centralized the country is: Thailand (most centralized) and Iran (least centralized) have the second (328) and sixth (444) fewest providers while the U.S. (median centralization) has the fourth most providers (834). Rather, centralization is most heavily affected by the breakdown of market share amongst the top providers in that country. When we break down the distribution amongst the top ten providers in each country, we see significant variation. The top provider in Thailand operates 60% of top websites, whereas the top provider in the U.S. operates only 29%, and, in Iran, 14% of websites. After the top 10–100 providers, countries have varying long tails of providers. For example, providers with fewer than 100 websites in our dataset host 17% of Iran's top sites but only 8% of Thai websites.

Anecdotally, we note that the most centralized countries tend to rely on large global providers like Amazon and Cloudflare, whereas the least centralized countries rely more on regional providers. We investigate how different types of providers (e.g., large global vs. small regional providers) contribute to country centralization in the next section.

5.2 Classes of Providers

Countries vary not only in their concentration on popular providers, but in the types of providers they use. For example, while seven of the top ten providers in Thailand are large global providers, seven of the top ten providers in Iran are local to the country. To understand the types of hosting providers that countries depend on and how this interacts with centralization, we build classes of hosting providers: we calculate the usage and endemicity ratio for each provider, then apply min-max scaling and cluster using affinity propagation. This results in 305 clusters (Figure 6), which we manually examine to identify 8 classes of providers (Table 1).

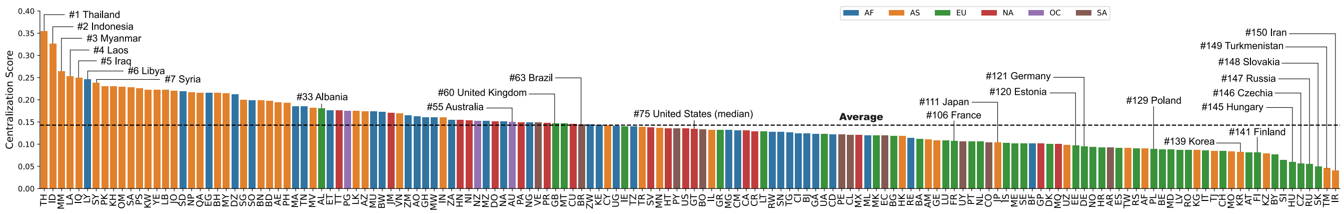


Figure 5: Hosting Provider Centralization by Country—Europe is consistently the least centralized, while Asia as a whole shows substantial variance. Other continents do not tend towards any extremes.

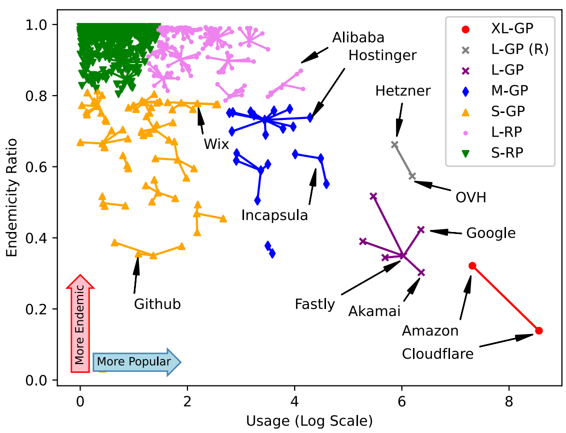


Figure 6: Classification of Providers—We cluster and classify providers based on their size and their endemicity. We visualize 7 classes, with XS-RP class not visualized for clarity. L-GP indicates a large (based on use in the top websites) global provider while L-RP indicates a large regional provider (based on endemicity).

One of the two largest global providers (XL-GP) is the largest provider in every country. With the exception of Japan, which relies most on Amazon, Cloudflare is the top provider in every country. As can be seen in Figure 7, the centralization of each country is strongly correlated with the dominance of these two XL-GPs (Pearson Correlation Coefficient ρ=0.90, p $\ll$ 0.05). Perhaps surprisingly, the use of large global providers (L-GP) beyond Amazon and Cloudflare (e.g., Google, Akamai, and Microsoft) has only a poor correlation with centralization (ρ=0.19, $p < 0.05$). This class serves 11–41% of websites between countries, but never sees outsized adoption in any country; rather, websites are split across a number of L-GPs.

Beyond the global providers, countries' use of regional providers varies substantially, from 12% (Trinidad and Tobago, Caribbean) to 68% (Iran). Among the least centralized countries, a cluster of regional providers overshadows the global providers, as shown in Figure 7. Generally, regional providers diffuse the provider ecosystem: indeed, the use of large regional providers is moderately correlated with lower centralization in countries (ρ=−0.72, p $\ll$ 0.05). However, in a few cases, there is a single dominant large regional provider, such as SuperHosting.BG in Bulgaria (22%, $\mathcal{S}$= 0.1188) and UAB in Lithuania (22%, $\mathcal{S}$= 0.1286), which never outrank Cloudflare but come a close second, contributing meaningfully to centralization in one country despite not being significant players globally.

Taking a regionalization perspective allows us to surface these large regional providers which have been missed by prior work.

5.3 Regionalization

Most regional providers achieve substantial market share in only their home country. However, at times, provider dependence follows other geopolitical patterns, which may be of concern even absent centralization. In this section, we look at how insular this dependence is, as well as external, regional patterns of dependence.

5.3.1 Insularity. The U.S. is most insular country (92.1%) because the largest global providers are American. The reliance on large global providers also means that most countries are not very insular. Indeed, U.S. providers host the largest share of sites in all but five countries: Iran, Czechia, Russia, Hungary, and Belarus. Further, the three most insular countries in our dataset after the U.S. are Iran ($\mathcal{S}$= 0.0411) with 64.8%, Czechia ($\mathcal{S}$= 0.0561) with 54.5%, and Russia ($\mathcal{S}$= 0.0554) with 51.1% of websites hosted by local providers. Lack of reliance on the U.S. does not necessarily mean countries are insular, though. Turkmenistan and Slovakia, which are among the countries with the least U.S. presence, are not particularly insular. Only 4% of websites in Turkmenistan use in-country providers, instead relying heavily on Russian providers (33%). In Slovakia, 25.7% of sites use Czech providers.

Looking at subregions broadly, East Asia (South Korea and Japan) and Europe are more insular. Countries in Africa have low insularity (average of 3%), indicating dependence on external providers, which we explore in Section 5.3.2. Insularity and centralization are orthogonal measures—a highly insular nation need not be highly centralized. However, given that countries that do not rely on a single large global provider tend to have websites split across a handful of providers, higher insularity is moderately correlated with lower centralization (ρ=-0.61, p $\ll$ 0.05).

5.3.2 Regional Patterns. Grouping providers by their home continent (Figure 8a), we find that European countries tend to rely on European providers, with the exception of Northern Europe (e.g., U.K., Sweden), which primarily relies on large global providers. While Asia broadly uses global providers, Eastern Asia (e.g., Japan, South Korea) uses regional Asian providers and Central Asia (e.g., Turkmenistan and other former USSR countries) use Russian providers (Section 5.3.3). Top websites in Africa are primarily hosted by American and European providers.

We stress that the use of providers headquartered in one country does not necessarily mean that the hosted content is not geographically proximate to users. Breaking down top websites by

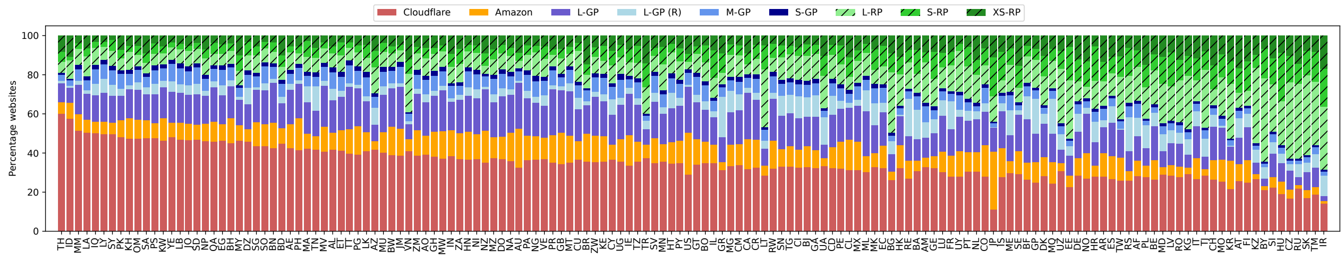
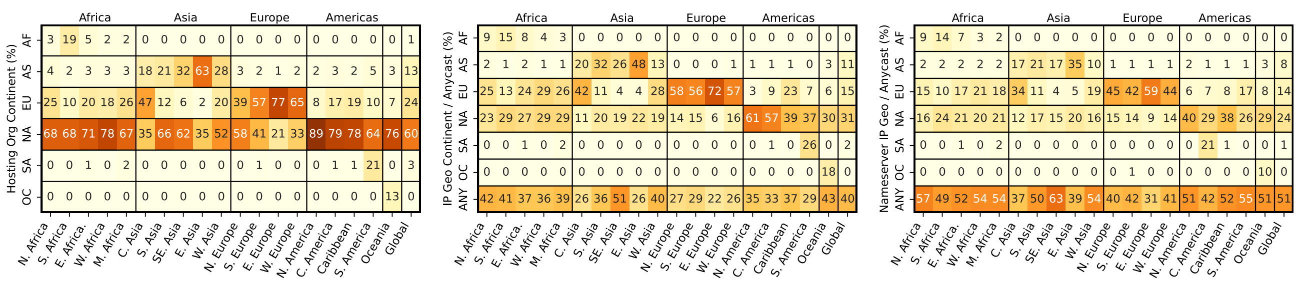


Figure 7: Breakdown of Hosting Provider Types—Cloudflare is the most popular provider in every country except Japan. While the most centralized countries overtly rely on Cloudflare, the least centralized countries tend to rely on a range of regional providers (hatched bars).



(a) Hosting Provider H.Q. **(b) IP Geolocation** **(c) DNS Nameserver Geolocation**

Figure 8: Regional Dependencies on Other Continents—Since most global providers are headquartered in the U.S., we see a strong reliance on N. America. While Europe and Eastern Asia (Japan, Korea) are mostly self-reliant, Central Asia uses providers and services in Europe (Russia). Africa primarily relies on providers and services in North America and Europe.

the continent in which the IP address is geolocated, we see that many websites using North American providers (Figure 8b) are served from the same region where the website is popular. While less commonly used for hosting, anycast plays an important role in the DNS ecosystem. In fact, we find that some European countries use large global DNS providers while relying on large regional providers for hosting (Section 6). For most regions, if not anycast, the IP addresses geolocate primarily in the same continent (or else in North America). The exception is Africa, where most websites geolocate to North America and Europe. It is unclear whether this is due to heavy use of American and European sites or if African sites use American or European hosting services.

5.3.3 Regional Case Studies. Looking beyond the dominance of large providers in the U.S., we see regional dependencies emerge. We manually analyze cases where countries rely disproportionately on select countries, and identify the following patterns:

Russia. Russian providers are heavily used by the Commonwealth of Independent States (CIS): Turkmenistan (33%), Tajikistan (23%), Kyrgyzstan (22%), Kazakhstan (21%) and Belarus (18%). However, not all post-Soviet states heavily use Russia: Ukraine (2%), Lithuania (3%), and Estonia (5%).

France. France is the second most relied on country after the U.S., largely due to reliance on OVH. Beyond OVH, eight French large regional providers (e.g., Online S.A.S) are extensively used in French administrative regions: Réunion Island (36%), Guadeloupe (34%), and Martinique (35%). Several African countries that were formerly French colonies, such as Burkina Faso (21%), Côte d'Ivoire (18%), and Mali (18%), also rely heavily on French hosting providers.

Czechia. 26% of the Slovak top sites are hosted by Czech providers. Czechia is quite insular and does not rely heavily on Slovakia.

Germany. The large German provider Hetzner is used globally (2% of all sites). There are also 7 large regional providers that see use in Austria (3%) as well, which is consistent with prior work [71] that attributes this to the fact that German is the predominant language in both countries.

Iran. More than 20% of Afghan top sites are hosted by Iranian providers. This may be due to the shared Persian language (we detect the language of the website using LangDetect [16]): we note that 31.4% of the websites in Afghanistan's top list are in Persian, of which 60.8% are hosted in Iran. It is unclear whether the Iranian-hosted sites popular in Afghanistan are owned by entities in Iran or in Afghanistan.

Interestingly, we see similar cross-country patterns in TLDs, despite countries being more insular on the TLD layer (Appendix B). This suggests that when countries look beyond their borders, they often look at the same set of external countries across multiple layers of infrastructure, even when the technical barriers are lower (as is the case for acquiring a ccTLD-based domain).

5.4 Longitudinal Change

We conducted our initial measurements in May 2023. We conducted a secondary measurement in May 2025 on the same vantage point to measure change over time. We find minimal overall changes in concentration since our initial measurements ($\rho = 0.98, p \ll 0.05$). We see the largest increase in centralization in Brazil ($\mathcal{S}_{old}$=0.1446, $\mathcal{S}_{new}$=0.2354), which is largely attributed to significant Cloudflare adoption (36% in May 2023 to 46% in May 2025). Cloudflare usage has generally increased (average +3.8% pts.) for all countries save for four (Russia, Belarus, Uzbekistan, Myanmar), with the largest increase seen in Turkmenistan (+11.3% pts), followed by Brazil. The largest *decrease* is seen in Russia (-2.0% pts). While Cloudflare usage has generally increased, the usage of U.S.-based providers in general has not increased significantly, with 56 out of 150 countries decreasing their reliance on the U.S since our initial measurements.

The largest decrease in centralization occurred in Russia, which dropped from 0.0554 to 0.0499. Here, we see slight movement away from U.S. providers (from 30% to 29%) and an increased use of Russian providers (from 50% to 56%). This is not inherently surprising given Russia's increased focus on reducing external dependencies after sanctions were imposed due to its invasion of Ukraine in 2022 [41]. It is important to note that the websites in the Russian CrUX top list have changed since the first measurement. Indeed, we note that the Jaccard index (which measures the similarity between two sets) of the domains lists is 0.4, indicating a significant change. For context, the average Jaccard index across all countries is 0.37, suggesting that many countries experienced similar levels of churn. However, the specific decrease in Russia's reliance on U.S. providers may reflect a deliberate reduction in visits to U.S.-hosted websites, rather than simply variation in their top sites.

5.5 Summary

We see much regional variation in hosting provider usage. European countries are consistently less centralized and Southeast Asian countries more centralized. Although Cloudflare usage frequently drives the centralization in a country, and use of regional providers is correlated with lower centralization, we also see cases of large regional providers rivaling market share of large global players in individual countries. We also observe regionalization patterns in cross-border dependence, such as CIS countries relying on Russia.

6 DNS INFRASTRUCTURE

In this section, we explore the centralization and regionalization of DNS infrastructure—a direct comparison made simple by our suite of metrics. While DNS centralization is close to that of hosting, we also see that the providers underlying that centralization tend to be larger than for web hosting.

6.1 Centralization

DNS centralization is similar to hosting (Appendix C.2, Figure 17). In large part, this is because most websites use the same provider for hosting and DNS. Indeed, Cloudflare content serving is predicated on using their DNS service in most cases [13]. In the most centralized case, 65% of the websites (up from 57% in hosting) in Indonesia ($\mathcal{S}$= 0.3757) are served by Cloudflare (Appendix C.1, Figure 14). Thailand, which was the most centralized country for hosting providers, is a close second with 62% of the websites (up from 60%) having DNS served by Cloudflare. In Czechia ($\mathcal{S}$= 0.0391), the least centralized country, Cloudflare DNS usage remains the same as for

Class	Count	Description	Example
XL-GP	2	Extra Large Global	Cloudflare
L-GP	10	Large Global	NSONE
L-GP (R)	2	Large Global (Regional)	OVH
M-GP	17	Medium Global	DNSimple
S-GP	78	Small Global	Sucuri
L-RP	273	Large Regional	Alibaba
S-RP	578	Small Regional	Scalaxy
XS-RP	9,049	Extra Small Regional	Forthnet

Table 2: Classes of DNS Infrastructure Providers—We categorize DNS providers by usage and endemicity.

hosting (17%). In contrast, Iran drops to the tenth least centralized country given increased use of Cloudflare.

Beyond Cloudflare, some countries that depend on global hosting providers shift to *large* global providers that provide managed DNS such as NSONE and Neustar UltraDNS (Section 6.2), and countries that depend on small regional providers for hosting often shift to larger regional DNS providers. Even with these underlying shifts, the broader regional trends are similar to the hosting layer.

6.2 Classes of Providers

Similar to hosting, we classify providers based on their usage and endemicity ratio metrics (Table 2). While we see roughly the same number of global providers across hosting and DNS, the use of managed DNS providers (e.g., Neustar UltraDNS, NSONE) shifts numbers towards large global providers. Both UltraDNS and NSONE are seen in the top ten providers for more than a hundred countries. We also observe a shift away from small regional providers towards large regional ones. In Czechia (the least centralized country) the share of websites using large regional providers goes up from 39% (hosting) to 47% (DNS). This shift towards large providers does not markedly change centralization ($\bar{\mathcal{S}}_{DNS}$ = 0.1379, $\bar{\mathcal{S}}_{Host}$ = 0.1429) given the reliance on multiple providers.

The shift towards larger providers may explain an increased anycast use for nameservice infrastructure compared to hosting (Figure 8c). While we see global adoption, Europe lags behind in anycast adoption, likely due to heavy reliance on regional providers in Europe who do not employ it.

7 CERTIFICATE AUTHORITIES

Websites' dependencies on certificate authorities (CAs) follow dramatically different patterns than on other layers, with near-universally high centralization and low insularity (Figure 11).

Class	Count	Description	Example
L-GP	7	Large Global	DigiCert
M-GP	2	Medium Global	Entrust
L-RP	11	Large Regional	Asseco
S-RP	10	Small Regional	SSL.com
XS-RP	15	Extra Small Regional	TrustCor

Table 3: Classes of Certificate Authorities—We cluster CAs based on their use and endemicity to identify classes of providers.

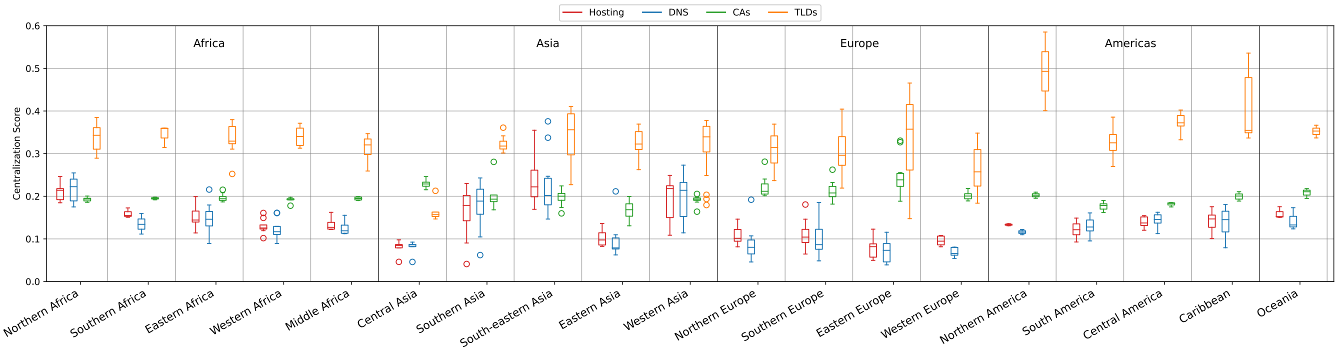


Figure 9: Centralization Across Layers and Subregions—Distribution of $\mathcal{S}$ across the hosting and DNS layers look roughly similar, while CA usage shows minimal variance in centralization given the limited number of CAs, and TLD use shows a higher degree of centralization and variance compared to other layers.

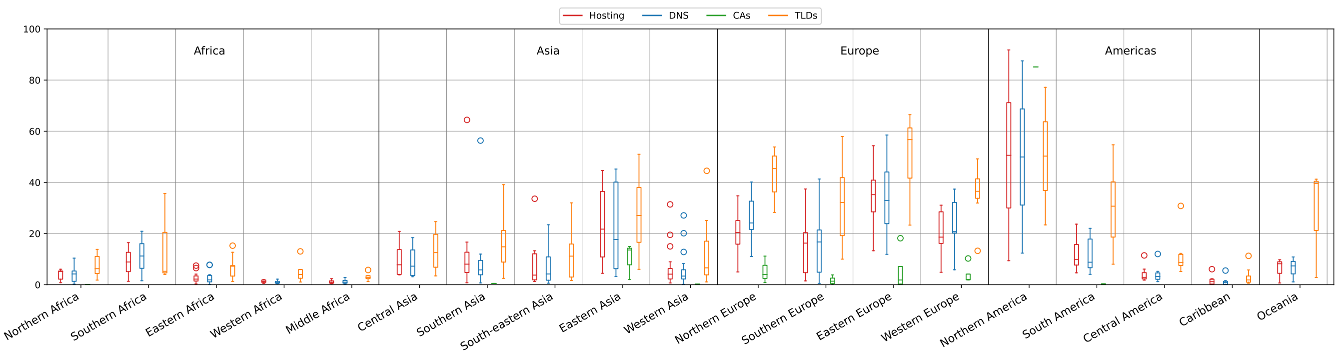


Figure 10: Insularity Across Layers and Subregions—Most global providers are based in the U.S., which drives North America to be the most insular region. Countries in Europe and Eastern Asia are consistently the most insular countries across all layers. Countries in the Global South show insularity at the TLD layer, but have low insularity in other layers since equivalent providers do not exist locally.

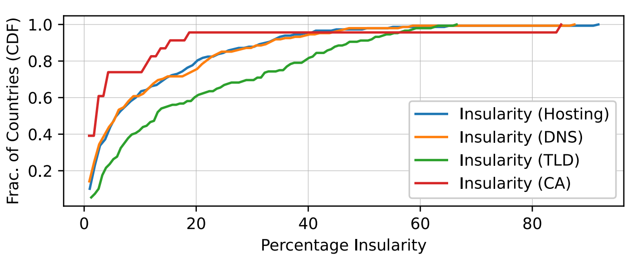


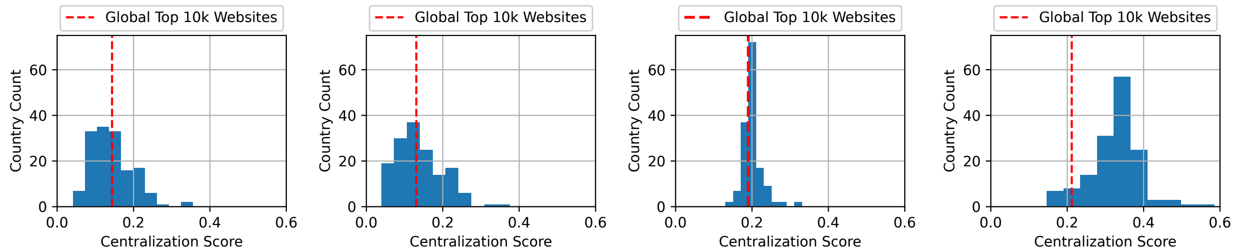
Figure 11: CDF of Insularity Across Layers—Countries tend to be more insular in their usage of TLDs compared other layers. Insularity in hosting and DNS track closely to each other. The small number of CAs and the relative domination of the large global CAs lead to a skewed distribution of insularity.

7.1 Centralization

Websites are far more centralized in their dependencies on certificate authorities than they are at any other layer (Figures 9 and 10). Unlike the tens of thousands of hosting and DNS providers that an operator can choose from, websites in our dataset use only 45 CAs. Note that while a top-N centralization metric would struggle to make a fair comparison across these two domain sizes—the top 10 hosting providers and the top 10 CAs are not an apples-to-apples comparison—our centralization metric makes the comparison straightforward. As a result of fewer available CAs, centralization on a small number of CAs is consistent across all countries: $\bar{\mathcal{S}}$ = 0.2007, var = 0.0007 (Figure 12 and 18). However, even within the small set of CAs, use is heavily skewed to a handful of large authorities: seven CAs account for 98% of websites.

At a high level, consistent with prior work [43, 46], we see that a handful of large global CAs (e.g., Let's Encrypt [1]) dominate the ecosystem. Additionally, a handful of countries heavily rely on regional CAs such as Asseco (a Polish CA). In the most centralized case, Slovakia ($\mathcal{S}$ = 0.3304), Let's Encrypt accounts for 55% of websites, three CAs account for 97% of websites, and seven CAs account for 98% of websites (Appendix C.1, Figure 15). The next most centralized countries, Czechia ($\mathcal{S}$ = 0.3268) and Estonia ($\mathcal{S}$ = 0.2811), show similar distributions. Potentially exacerbating centralization, hosting providers often partner with CAs to issue certificates for hosted websites. For example, Cloudflare uses Let's Encrypt, DigiCert, Google, and Sectigo while Incapsula (Imperva) uses GlobalSign [12, 39].

Similar to hosting and DNS, we classify CAs based on usage and endemicity, identifying five classes (Table 3). Based on this



(a) Hosting Providers **(b) DNS Providers** **(c) Certificate Authorities** **(d) TLDs**

Figure 12: Centralization Distributions By Country—The distributions for hosting and DNS providers are roughly similar. In contrast, the distribution for CAs has very little variance given the smaller number of providers and dominant use of large global CAs. The histogram for TLDs shows higher centralization scores all across, indicating that countries depend on few TLDs. (For full analysis of TLDs, see Appendix B.) Also marked is $\mathcal{S}$ for the Global Top 10k websites and shows that while it is representative of the average centralization in hosting, DNS, and CA layers, it is not representative for TLDs.

classification, the most important class of CAs is L-GP, which includes the seven large global CAs that together dominate the web: Let's Encrypt, DigiCert, Sectigo, Google, Amazon, GlobalSign, GoDaddy. These seven providers account for 80–99.7% of websites (Iran and Russia, respectively, at the extremes). DigiCert and Let's Encrypt are the two most popular CAs accounting for most of the use, ranging from 40-75% usage in individual countries, and 57% usage overall across all websites in the 150 countries we consider.

In the least centralized cases, Taiwan ($\mathcal{S}$= 0.1308) and Japan ($\mathcal{S}$= 0.1499), the seven large global CAs account for 82% and 85% of websites respectively; the remaining sites are secured by regional CAs. We note that Let's Encrypt is heavily used in European countries, especially Eastern European countries that use regional hosting providers. Interestingly, Russia shows the most use of global providers even though it is one of the most insular countries in other layers. This is consistent with prior work that has observed that Russia has heavily depended on Let's Encrypt and GlobalSign once DigiCert largely moved out of Russia following the full-scale invasion of Ukraine in 2022 [41].

7.2 Regionalization

Because most countries do not have their own CAs, insularity is near zero almost everywhere. This leads to the countries that had the highest insularity in other layers to be most different at the CA layer. In absolute terms, Europe (like other continents) is less insular at the CA layer than for other layers. Still, as with the hosting (Figure 20), DNS (Figure 21), and TLD layers (Figure 22), Europe is more insular at the CA layer than other continents. Whereas higher insularity correlated with lower centralization at the hosting layer, here Europe goes against that trend, exhibiting both higher insularity (Figure 13) and higher centralization ($\bar{\mathcal{S}}_{EU}$ = 0.2220, Figure 18) than other continents at the CA layer. For example, while Europe had 7 out of the 10 *least* centralized countries for hosting, it has 8 of the 10 *most* centralized countries in the CA layer. Czechia and Slovakia, which place among the least centralized at the hosting ($\mathcal{S}_{CZ}$ = 0.0561, $\mathcal{S}_{SK}$ = 0.0497) and DNS layers ($\mathcal{S}_{CZ}$ = 0.0391, $\mathcal{S}_{SK}$ = 0.0429), are the most centralized at the CA layer ($\mathcal{S}_{CZ}$ = 0.3268, $\mathcal{S}_{SK}$ = 0.3304).

The most popular regional CA is a Polish CA, Asseco [1], which is used not only in Poland (19%) but also in Iran (19%) and Afghanistan (5%). This may be due to these websites avoiding U.S.-based CAs. With the exception of Iran and Afghanistan, the use of regional CAs is concentrated in their home country. Poland, Taiwan, and Japan are the most insular countries after the U.S., with 19%, 17%, and 14% of the websites using local CAs, respectively.

Centralization on a small number of CAs is not inherently bad for the security of the Internet since many small CAs struggle with operational requirements and a moderate degree of centralization reduces the attack surface of the WebPKI [45]. However, prior work has also hinted that concentrated power was a factor for noncompliance [62]. Regardless, one clear consequence is that the vast majority of countries are dependent on infrastructure in the U.S. and have few alternatives to choose from. Recently, several countries and regions have tried to change this. When major Western CAs left Russia following the invasion of Ukraine [23], Russia attempted to stand up a state-sponsored root CA in 2022, but the root certificate was never accepted by major web browsers. The EU also recently attempted to regulate trust decisions made by web browsers, which was met with significant backlash from both web browsers and the broader security community [55, 56].

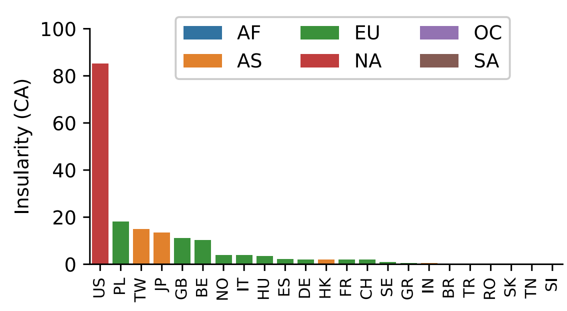


Figure 13: CA Insularity by Country—Countries sorted by % of websites using CAs based in the same country. Only 24 countries in our dataset use a CA in their own country. Large global CAs in the U.S. dominate the ecosystem.

[1]Previously Unizeto Centrum.

8 DISCUSSION AND CONCLUSION

In this work, we introduced a suite of dependency metrics, including centralization, usage, endemicity, and insularity, which we used to investigate the centralization and regionalization of web infrastructure. We hope that a reliable toolkit for quantifying Internet dependency enables the networking community to better understand and discuss how the Internet is structured. Here, we highlight implications and open questions from our work.

The Role of Regional Providers. Cloudflare (and, to a lesser degree, a handful of other large providers) significantly affect the centralization of the web. However, when we focus our discussion on the overall size of providers, especially global providers, it is easy to lose the context of how centralization impacts people's lived experiences. Although a country's use of regional providers typically contributes to a more diffuse ecosystem, this is not universal: in some cases, a country can be centralized on a single regional provider that appears small on the global stage. For example, SuperHosting.BG in Bulgaria and UAB in Lithuania have not seen substantial attention in prior discussion of centralization, despite driving centralization in their respective countries more so than large global providers like Google, Akamai, and Fastly. This dynamic implies a different locus, but the same level of power.

Geopolitically Concentrated Dependence. Locus of power extends beyond reliance on any one provider. Even where we observe a constellation of only moderately sized regional providers, from the perspective of these countries, dependence on a set of providers in a single foreign country also carries similar risks. For instance, the CIS countries' dependence on Russian providers creates a power dynamic that could be affected by contemporary geopolitics. It is in this context that we urge the research community to look beyond centralization and to consider users' experience of *concentrated dependence* more holistically.

Dependence Across Layers. Formalizing a suite of metrics for quantifying dependence enables us identify geographic patterns across multiple infrastructure layers and to hypothesize about their causes. For instance, we see patterns of cross-country dependence in hosting providers—such as former French colonies using French providers—that recur at the TLD layer, suggesting that historical and linguistic ties contribute to countries coalescing around a particular set of providers. In addition, countries' centralization on hosting providers is often tightly coupled with DNS: hosting and DNS services are often bundled (indeed, Cloudflare's service is typically predicated on using their DNS service). Although CA centralization is clearly influenced by browsers' selective trust in CAs, we also see hosting providers partnering with the largest global CAs (Section 7). Thus, we can hypothesize that part of the centralization we see on the web is a result of provider, not operator, choice.

Root Causes. While our findings suggest possible explanations, they do not establish causality. There is much still to be done to understand reasons underlying the dependency patterns that we observe. This requires handling challenging confounds: for example, do Slovak websites choose to use Czech providers, or do Slovak people tend to frequent Czech websites? Do the choices by a provider, such as a hosting provider's default CA configuration or DNS offering, cause a synergy in centralization patterns across infrastructure layers? Exploring these complexities is key to understanding and affecting the future of Internet dependency. As such, future work should analyze such covariates to uncover causal relationships.

Consequences for the Networking Community. The geopolitical dynamics that we observe shed light on how we can better design for and evaluate networks globally. For instance, while major providers like Google may roll out new protocols and optimizations to improve user experience, these benefits may not be felt uniformly. In many countries (such as Iran or Bulgaria), these deployments affect only a modest fraction of the sites users frequent. This underscores the need for worldwide adoption of new standards to improve the web as a whole. Moreover, researchers studying web resilience would benefit from understanding how availability and performance could be impacted not only by a provider outage, but also by a geopolitical schism between two countries. Moving forward, we hope that the quantitative metrics we introduce for measuring dependence on the web enable future studies to more deeply understand the Internet and shape how it is evolving.

ACKNOWLEDGMENTS

We thank Thea Rossman, Aleksander Durumeric, and members of the Stanford Empirical Security Research Group. This work was supported in part by a Sloan Research Fellowship, a Stanford Grduate Fellowship (SGF), and the National Science Foundation under Grant Number #2319080. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF or other funding organizations.

REFERENCES

[1] Josh Aas, Richard Barnes, Benton Case, Zakir Durumeric, Peter Eckersley, Alan Flores-López, J. Alex Halderman, Jacob Hoffman-Andrews, James Kasten, Eric Rescorla, Seth Schoen, and Brad Warren. 2019. Let's Encrypt: An Automated Certificate Authority to Encrypt the Entire Web. In *ACM SIGSAC Conference on Computer and Communications Security*.

[2] Haldun Akoglu. 2018. User's Guide to Correlation Coefficients. *Turkish Journal of Emergency Medicine* (2018).

[3] Samantha Bates, John Bowers, Shane Greenstein, Jordi Weinstock, Yunhan Xu, and Jonathan Zittrain. 2018. Evidence of Decreasing Internet Entropy: The Lack of Redundancy in DNS Resolution by Major Websites. https://www.nber.org/system/files/working_papers/w24317/w24317.pdf.

[4] bgp.tools. 2023. Anycast Prefixes. https://github.com/bgptools/anycast-prefixes. Accessed: 2023-05-10.

[5] Timm Böttger, Gianni Antichi, Eder L Fernandes, Roberto di Lallo, Marc Bruyere, Steve Uhlig, Gareth Tyson, and Ignacio Castro. 2019. Shaping the Internet: 10 Years of IXP Growth. https://arxiv.org/abs/1810.10963.

[6] CA Certificates in Firefox 2023. CA Certificates in Firefox. https://ccadb.my.salesforce-sites.com/mozilla/CACertificatesInFirefoxReport.

[7] Yuhang Cai and Lek-Heng Lim. 2022. Distances between probability distributions of different dimensions. *IEEE Transactions on Information Theory* (2022).

[8] CAIDA. 2023. AS to Organizations Mappings. https://catalog.caida.org/dataset/as_organizations. Accessed: 2023-05-18.

[9] CAIDA. 2023. Routeviews Prefix to AS mappings Dataset for IPv4 and IPv6. http://www.caida.org/data/routing/routeviews-prefix2as.xml.

[10] Chrome User Experience Report 2023. Chrome User Experience Report. https://developer.chrome.com/docs/crux/.

[11] David Clark and KC Claffy. 2021. Trust zones: A path to a more secure internet infrastructure. *Journal of Information Policy* 11 (2021), 26–62.

[12] Cloudflare. 2023. Certificate Authorities. https://web.archive.org/web/20230520233658/https://developers.cloudflare.com/ssl/reference/certificate-authorities/.

[13] Cloudflare. 2024. Manage Domains. https://developers.cloudflare.com/fundamentals/setup/manage-domains/.

[14] Cloudflare, Inc. [n. d.]. The Cloudflare global network. https://www.cloudflare.com/network/.

[15] Cloudflare, Inc. 2025. Cloudflare Service Outage on June 12, 2025. https://blog.cloudflare.com/cloudflare-service-outage-june-12-2025/

[16] Michal Mimino Danilak. 2021. LangDetect. https://pypi.org/project/langdetect/.

[17] Matteo Dell'Amico, Leyla Bilge, Ashwin Kayyoor, Petros Efstathopoulos, and Pierre-Antoine Vervier. 2017. Lean On Me: Mining Internet Service Dependencies From Large-Scale DNS Data. In *Annual Computer Security Applications Conf.*

[18] Trinh Viet Doan, Roland vanRijswijk Deij, Oliver Hohlfeld, and Vaibhav Bajpai. 2022. An Empirical View on Consolidation of the Web. *ACM Transactions on Internet Technology* (2022).

[19] Zakir Durumeric, David Adrian, Ariana Mirian, Michael Bailey, and J Alex Halderman. 2015. A search engine backed by Internet-wide scanning. In *22nd ACM SIGSAC Conference on Computer and Communications Security.*

[20] Zakir Durumeric, David Adrian, Ariana Mirian, James Kasten, Elie Bursztein, Nicolas Lidzborski, Kurt Thomas, Vijay Eranti, Michael Bailey, and J Alex Halderman. 2015. Neither snow nor rain nor MITM... an empirical analysis of email delivery security. In *ACM Internet Measurement Conference.*

[21] Zakir Durumeric, David Adrian, Phillip Stephens, Eric Wustrow, and J. Alex Halderman. 2024. Ten Years of ZMap. In *Proceedings of the 2024 ACM on Internet Measurement Conference* (Madrid, Spain) *(IMC '24)*. Association for Computing Machinery, New York, NY, USA, 139–148. https://doi.org/10.1145/3646547.3689012

[22] Zakir Durumeric, Eric Wustrow, and J Alex Halderman. 2013. ZMap: Fast Internet-wide scanning and its security applications. In *22nd USENIX Security Symposium.*

[23] EFF. 2022. You Should Not Trust Russia's New Trusted Root CA. https://www.eff.org/deeplinks/2022/03/you-should-not-trust-russias-new-trusted-root-ca.

[24] Digital Element. 2023. NetAcuity IP Geolocation Data. https://www.digitalelement.com/geolocation/. Accessed: 2023-05-18.

[25] Xun Fan, Ethan Katz-Bassett, and John Heidemann. 2015. Assessing affinity between users and CDN sites. In *Traffic Monitoring and Analysis (TMA).*

[26] Rodérick Fanou, Gareth Tyson, Eder Leao Fernandes, Pierre Francois, Francisco Valera, and Arjuna Sathiaseelan. 2018. Exploring and Analysing the African Web Ecosystem. *ACM Tranactions on the Web* 12, 4 (2018).

[27] Rodérick Fanou, Gareth Tyson, Pierre Francois, and Arjuna Sathiaseelan. 2016. Pushing the Frontier: Exploring the African Web Ecosystem. In *Proceedings of the 25th International Conference on World Wide Web.*

[28] Fastly, Inc. 2025. Content Delivery Network. https://www.fastly.com.

[29] Manaf Gharaibeh, Anant Shah, Bradley Huffaker, Han Zhang, Roya Ensafi, and Christos Papadopoulos. 2017. A look at router geolocation in public and commercial databases. In *Proceedings of the 2017 Internet Measurement Conference.*

[30] Phillipa Gill, Martin Arlitt, Zongpeng Li, and Anirban Mahanti. 2008. The flattening internet topology: Natural evolution, unsightly barnacles or contrived collapse?. In *Passive and Active Network Measurement (PAM).*

[31] Catherine Han, Deepak Kumar, and Zakir Durumeric. 2022. On the infrastructure providers that support misinformation websites. In *International AAAI conference on web and social media.*

[32] Rasmus Helles, Stine Lomborg, and Signe Sophus Lai. 2020. Infrastructures of tracking: Mapping the ecology of third-party services across top sites in the EU. *New Media & Society* 22, 11 (2020), 1957–1975.

[33] Nguyen Phong Hoang, Arian Akhavan Niaki, Michalis Polychronakis, and Phillipa Gill. 2020. The web is still small after more than a decade. *SIGCOMM Comput. Commun. Rev.* 50, 2 (2020), 24–31.

[34] Ralph Holz, Jens Hiller, Johanna Amann, Abbas Razaghpanah, Thomas Jost, Narseo Vallina-Rodriguez, and Oliver Hohlfeld. 2020. Tracking the deployment of TLS 1.3 on the Web: A story of experimentation and centralization. *ACM SIGCOMM Computer Communication Review* 50, 3 (2020), 3–15.

[35] Florian Holzbauer, Johanna Ullrich, Martina Lindorfer, and Tobias Fiebig. 2022. Not that Simple: Email Delivery in the 21st Century. In *2022 USENIX Annual Technical Conference (USENIX ATC 22)*. USENIX Association, Carlsbad, CA, 295–308. https://www.usenix.org/conference/atc22/presentation/holzbauer

[36] Geoff Huston. 2022. Looking at Centrality in the DNS. https://circleid.com/posts/20221121-looking-at-centrality-in-the-dns.

[37] ICANN. 2020. ICANN and Verisign Announce Proposed Amendment to .COM Registry Agreement. icann.org/en/announcements/details/icann-and-verisign-announce-proposed-amendment-to-com-registry-agreement-3-1-2020-en.

[38] Muhammad Ikram, Rahat Masood, Gareth Tyson, Mohamed Ali Kaafar, and Roya Ensafi. 2022. A Study of Third-party Resources Loading on Web. https://arxiv.org/pdf/2203.03077.pdf.

[39] Imperva (Incapsula). 2023. Web Protection - SSL/TLS. https://docs.imperva.com/bundle/cloud-application-security/page/more/ssl-certificate.htm.

[40] Liz Izhikevich, Gautam Akiwate, Briana Berger, Spencer Drakontaidis, Anna Ascheman, Paul Pearce, David Adrian, and Zakir Durumeric. 2022. ZDNS: a fast DNS toolkit for internet measurement. In *ACM Internet Measurement Conference.*

[41] Mattijs Jonker, Gautam Akiwate, Antonia Affinito, kc Claffy, Alessio Botta, Geoffrey M. Voelker, Roland van Rijswijk-Deij, and Stefan Savage. 2022. Where .ru? Assessing the Impact of Conflict on Russian Domain Infrastructure. In *ACM Internet Measurement Conference.*

[42] Aqsa Kashaf, Jiachen Dou, Margarita Belova, Maria Apostolaki, Yuvraj Agarwal, and Vyas Sekar. 2023. A first look at third-party service dependencies of web services in Africa. In *Passive and Active Network Measurement.*

[43] Aqsa Kashaf, Vyas Sekar, and Yuvraj Agarwal. 2020. Analyzing Third Party Service Dependencies in Modern Web Services: Have We Learned from the Mirai-Dyn Incident?. In *ACM Internet Measurement Conference.*

[44] Michael Kende and Karen Rose. 2015. Promoting local content hosting to develop the internet ecosystem.

[45] Deepak Kumar, Zhengping Wang, Matthew Hyder, Joseph Dickinson, Gabrielle Beck, David Adrian, Joshua Mason, Zakir Durumeric, J Alex Halderman, and Michael Bailey. 2018. Tracking certificate misissuance in the wild. In *IEEE Symposium on Security and Privacy (S&P).*

[46] Rashna Kumar, Sana Asif, Elise Lee, and Fabian E. Bustamante. 2023. Each at its Own Pace: Third-Party Dependency and Centralization Around the World. *the on Measurement and Analysis of Computing Systems* (2023).

[47] Rashna Kumar, Esteban Carisimo, Lukas De Angelis Riva, Mauricio Buzzone, Fabián E. Bustamante, Ihsan Ayyub Qazi, and Mariano G. Beiró. 2024. Of Choices and Control - A Comparative Analysis of Government Hosting. In *ACM on Internet Measurement Conference.*

[48] Zhenyu Li and Yiwen Huang. 2023. A First Look into the Third-Party Web Dependencies in China. In *Asian Internet Engineering Conference.*

[49] Enze Liu, Gautam Akiwate, Mattijs Jonker, Ariana Mirian, Stefan Savage, and Geoffrey M Voelker. 2021. Who's got your mail? characterizing mail service provider usage. In *ACM Internet Measurement Conference.*

[50] Zane Ma, Joshua Mason, Manos Antonakakis, Zakir Durumeric, and Michael Bailey. 2021. What's in a Name? Exploring CA Certificate Control. In *30th USENIX Security Symposium (USENIX Security 21)*. USENIX Association, 4383–4400. https://www.usenix.org/conference/usenixsecurity21/presentation/ma

[51] Giovane C. M. Moura, Sebastian Castro, Wes Hardaker, Maarten Wullink, and Cristian Hesselman. 2020. Clouding up the Internet: how centralized is DNS traffic becoming?. In *Proceedings of the ACM Internet Measurement Conference.*

[52] Mark Nottingham. 2023. Centralization, Decentralization, and Internet Standards. https://datatracker.ietf.org/doc/rfc9518/.

[53] US Department of Justice. 2010. Antitrust Division – Horizontal Merger Guidelines. https://www.justice.gov/atr/horizontal-merger-guidelines-08192010.

[54] US Department of Justice. 2024. Antitrust Division – Herfindahl-Hirschman Index. https://www.justice.gov/atr/herfindahl-hirschman-index.

[55] European Parliament. 2023. Qualified certificates for website authentication. https://www.europarl.europa.eu/RegData/etudes/ATAG/2023/739285/EPRS_ATA(2023)739285_EN.pdf.

[56] The Record. 2023. EU urged to drop new law that could allow member states to intercept and decrypt global web traffic. https://therecord.media/eu-urged-to-drop-law-website-authentication-certificates.

[57] RIPE NCC. 2014. RIPE Atlas. https://atlas.ripe.net. Accessed: 2025-06-13.

[58] Y. Rubner, C. Tomasi, and L.J. Guibas. 1998. A Metric for Distributions with Applications to Image Databases. In *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*. 59–66. https://doi.org/10.1109/ICCV.1998.710701

[59] Yossi Rubner, Carlo Tomasi, and Leonidas Guibas. 2000. The Earth Mover's Distance as a Metric for Image Retrieval. *Intl. Journal of Computer Vision* (2000).

[60] Kimberly Ruth, Aurore Fass, Jonathan Azose, Mark Pearson, Emma Thomas, Caitlin Sadowski, and Zakir Durumeric. 2022. A World Wide View of Browsing the World Wide Web. In *ACM Internet Measurement Conference.*

[61] Kimberly Ruth, Deepak Kumar, Brandon Wang, Luke Valenta, and Zakir Durumeric. [n. d.]. Toppling Top Lists: Evaluating the Accuracy of Popular Website Lists. In *22nd ACM Internet Measurement Conference.*

[62] Nicolas Serrano, Hilda Hadan, and L Jean Camp. 2019. A complete study of PKI (PKI's Known Incidents). *SSRN Electronic Journal* (2019).

[63] Raffaele Sommese, Gautam Akiwate, Mattijs Jonker, Giovane C Moura, Marco Davids, Roland van Rijswijk-Deij, Geoffrey M Voelker, Stefan Savage, Anna Sperotto, et al. 2021. Characterization of anycast adoption in the DNS authoritative infrastructure. In *Network Traffic Measurement and Analysis Conference.*

[64] Samaneh Tajalizadehkhoob, Maciej Korczyński, Arman Noroozian, Carlos Ganán, and Michel Van Eeten. 2016. Apples, oranges and hosting providers: Heterogeneity and security in the hosting market. In *IEEE/IFIP Network Operations and Management Symposium (NOMS).*

[65] Tobias Urban, Martin Degeling, Thorsten Holz, and Norbert Pohlmann. 2020. Beyond the Front Page: Measuring Third Party Dynamics in the Field. In *Proceedings of The Web Conference.*

[66] Anh V Vu, Alice Hutchings, and Ross Anderson. 2024. No easy way out: The effectiveness of deplatforming an extremist forum to suppress hate and harassment. In *IEEE Symposium on Security and Privacy.*

[67] Chuhan Wang, Yasuhiro Kuranaga, Yihang Wang, Mingming Zhang, Linkai Zheng, Xiang Li, Jianjun Chen, Haixin Duan, Yanzhong Lin, and Qingfeng Pan. 2024. BREAKSPF: How Shared Infrastructures Magnify SPF Vulnerabilities Across the Internet. *NDSS'24* (2024).

[68] Synthia Wang, Kyle MacMillan, Brennan Schaffner, Nick Feamster, and Marshini Chetty. 2024. Measuring the Consolidation of DNS and Web Hosting Providers. https://arxiv.org/abs/2110.15345.

[69] Luciano Zembruzki, Arthur Selle Jacobs, and Lisandro Zambenedetti Granville. 2022. On the Consolidation of the Internet Domain Name System. In *IEEE GLOBECOM Global Communications Conference.*

[70] Luciano Zembruzki, Arthur Selle Jacobs, Lisandro Zambenedetti Granville, and Ricardo Jose Pfitscher. 2023. Examining the Centralization of Email Industry: A Landscape Analysis for IPv4 and IPv6. In *IEEE Symposium on Computers and Communications (ISCC)*.

[71] Luciano Zembruzki, Raffaele Sommese, Lisandro Zambenedetti Granville, Arthur Selle Jacobs, Mattijs Jonker, and Giovane C. M. Moura. 2022. Hosting Industry Centralization and Consolidation. In *NOMS IEEE/IFIP Network Operations and Management Symposium*.

Appendices are supporting material that has not been peer-reviewed.

A EMD REFERENCE

We details how EMD is formally defined and how we arrive at a simple expression for our instantiation of it. We begin with a general definition of EMD in its discretized formalization.

Let $A = (a_1, a_2, \ldots, a_n)$ be the observed distribution of data, and let $R = (r_1, r_2, \ldots, r_m)$ be the reference distribution. For simplicity, assume $\sum_{i=1}^{n} a_i = \sum_{j=1}^{m} r_j$. We define a ground distance function d_{ij} for all $1 \leq i \leq n$, $1 \leq j \leq m$. Transforming A into R requires assigning nonnegative *flow* values f_{ij} for all i, j, which define the amount of earth moved from bucket i to bucket j. Naturally, the total flow out of a pile equals amount originally in the pile ($\sum_{j=1}^{m} f_{ij} = a_i\ \forall i$), and the total flow into a pile equals the ending size of the pile ($\sum_{i=1}^{n} f_{ij} = r_j\ \forall j$). There may be many solutions satisfying these constraints, but we aim to find one that minimizes the total *work* involved in transporting the earth, which is defined as the product of flow and distance:

$$w = \min_{[f_{ij}]} \sum_{i=1}^{n} \sum_{j=1}^{m} f_{ij} d_{ij}$$

Taking the f_{ij} that solve this minimization problem, the EMD is defined as this minimum work, optionally normalized to $[0, 1]$ by means appropriate to the chosen distance metric and total flow. If $0 \leq d_{ij} \leq 1\ \forall i, j$ above, then the expression becomes:

$$EMD(A, R) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} f_{ij} d_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{m} f_{ij}} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} f_{ij} d_{ij}}{\sum_{i=1}^{n} a_i}$$

For our instantiation of EMD in this paper, we can use our uniform reference distribution and simple vertical-difference distance function to enable further simplification. Specifically:

- Let $A = (a_1, a_2, \ldots, a_n)$ be a nonincreasing sequence representing the counts of websites using each provider, with $\sum_i a_i = C$ the total number of sites.
- Let R be a uniform distribution across C buckets, each of size 1, representing a fully decentralized reference distribution.
- Let $d_{ij} = \frac{1}{C}(a_i - r_j) = \frac{a_i - 1}{C}\ \forall i, j$. This represents the vertical height difference between a_i and r_j, normalized by the total number of sites.

Next, we determine the optimum flow. For each i, we need to move each of a_i units of size 1 (each website) into a separate bucket in R. (It does not matter which, as d_{ij} is only dependent on i, not j.) So for each i we have $f_{ij} = 1$ for a_i values of j and $f_{ij} = 0$ for all other j. This means the total work is

$$w = \sum_{i=0}^{n} a_i (1 \cdot \frac{a_i - 1}{C}) = \sum_{i=0}^{n} \frac{a_i^2 - a_i}{C} = \sum_{i=0}^{n} \frac{a_i^2}{C} - \frac{\sum_{i=0}^{n} a_i}{C} = \sum_{i=0}^{n} \frac{a_i^2}{C} - 1$$

To normalize to $[0, 1]$, we divide by the total flow, which also equals C. This yields our final expression for our centralization score $\mathcal{S}$:

$$\mathcal{S} = \frac{1}{C} \left(\sum_{i=0}^{n} \frac{a_i^2}{C} - 1 \right) = \sum_{i=0}^{n} \left(\frac{a_i}{C} \right)^2 - \frac{1}{C}$$

B TOP LEVEL DOMAINS (TLDS)

Overall, there is greater centralization in countries' use of TLDs ($\bar{\mathcal{S}}$ = 0.3262) compared to other layers (Figure 9). Despite a large number of top level domains, TLD centralization is primarily driven by global usage of `.com` and countries' usage of ccTLDs. In the most centralized case, 77% of the top websites in the U.S ($\mathcal{S}$= 0.5853) use the `.com` TLD, while in the the least centralized case, Kyrgyzstan ($\mathcal{S}$= 0.1468), `.com` accounts for 29% of the top websites while `.ru` and `.kg` account for 22% and 12% respectively. The top four most centralized countries, all in North America ($\bar{\mathcal{S}}$= 0.4930) and Caribbean ($\bar{\mathcal{S}}$= 0.4042), all rise to the top because of their `.com` usage.

After the handful of top countries that rely heavily on `.com`, we begin to see countries rise to be most centralized based on their usage of a ccTLD. While there are a few exceptions (e.g., use of `.ru` by CIS countries), most countries centralize on their own ccTLDs. Eastern Europe ($\bar{\mathcal{S}}$= 0.3361) stands out with Czechia ($\mathcal{S}$= 0.4656), Hungary ($\mathcal{S}$= 0.4450), and Poland ($\mathcal{S}$= 0.4265) being the 5th–7th most centralized countries due to their reliance on their local ccTLDs. We also see that countries tend to be most insular at the TLD layer (Figure 11). This may be because the choice of TLD has relatively few technical or financial implications compared to other network layers, providing operators more flexibility to choose an in-country option.

When we look at insularity across layers and regions, we see two classes of behavior (Figure 10). First, countries in Europe, Eastern Asia (Japan and Korea), and North America are consistently the most insular countries across all layers. In these countries, the use of local ccTLD is also coupled with a strong network of regional providers (e.g., Japan, Czechia). As such, if a country overcomes the infrastructure barrier and is insular in its hosting infrastructure, we expect it to strongly indicate insularity in the TLD layer. In fact, insularity in the hosting layer is moderately correlated with insularity in the TLD layer (ρ=0.70, p $\ll$ 0.05). In the second case, countries in the Global South show insularity most directly at the TLD layer but have low insularity in other layers since equivalent providers do not exist in-country (or, in some cases, even in-region).

Finally, when we look at countries with a large external dependency on their ccTLD, we see evidence of countries depending on a few countries across different layers:

France. Similar to regional clusters in hosting, we find the `.fr` ccTLD is commonly used across 14 countries: Burkina Faso, Benin, Congo, Côte d'Ivoire, Cameroon, Algeria, Guadeloupe, Haiti, Madagascar, Mali, Martinique, Réunion, Senegal, and Togo, where `.fr` is also more popular than their own local ccTLDs.

Russia. Similar to hosting providers, we see the dependence of some old Soviet states — countries that are now part of the Commonwealth of Independent States (CIS) on `.ru`. Not only do these countries use `.ru` but also their own ccTLD in addition to `.com` making them the least centralized countries given their lack of overt use of a single TLD.

Germany. Several countries where German is a dominant language use the `.de` ccTLD. These include Germany (44%), Austria (14%), Luxembourg (8%), and Switzerland (7%). This observation is consistent with prior work [71].

C DNS, CA, AND TLD LAYERS

C.1 Cluster Usage Distribution

Figures 14, 15, and 16 show the category breakdowns of DNS, CA, and TLD use by country, sorted by countries' centralization scores. See Figure 7 in Section 5.2 for corresponding results for hosting.

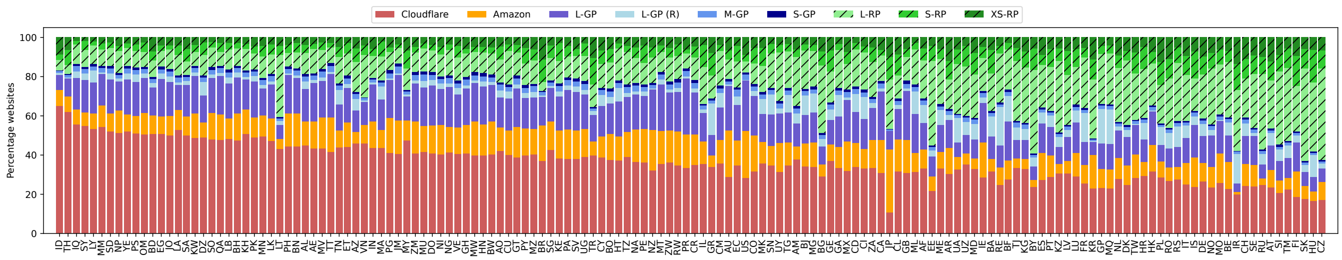


Figure 14: Percentage of websites in all the countries (sorted by $\mathcal{S}$) broken down by DNS provider type—Similar to the hosting provider case, Cloudflare dominates in every country save for Japan.

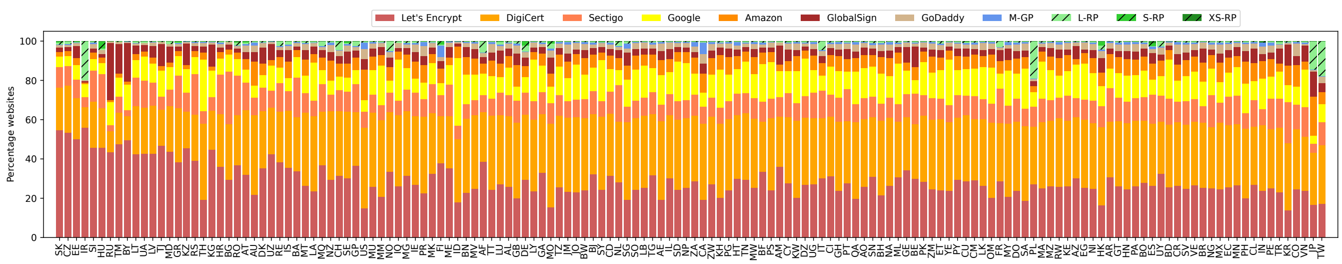


Figure 15: Percentage of websites in all the countries (sorted by $\mathcal{S}$) broken down by CA type—Notably, the 7 large CAs account for an average of 98% of websites across countries.

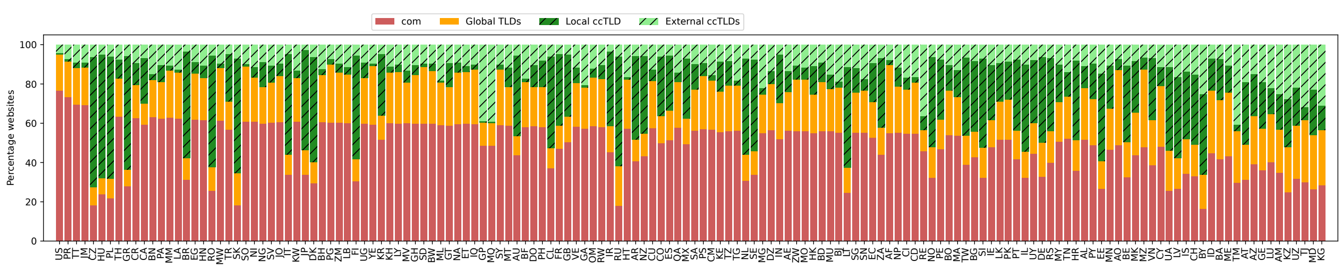


Figure 16: Percentage of websites in all the countries (sorted by $\mathcal{S}$) broken down by TLD type—The usage of external ccTLDs is strongly correlated with lower centralization.

C.2 Centralization Continental Trends

The distributions of countries' centralization scores color-coded by continent for the DNS, CA, and TLD layers, respectively. See Figure 5 in Section 5.1 for corresponding results for hosting.

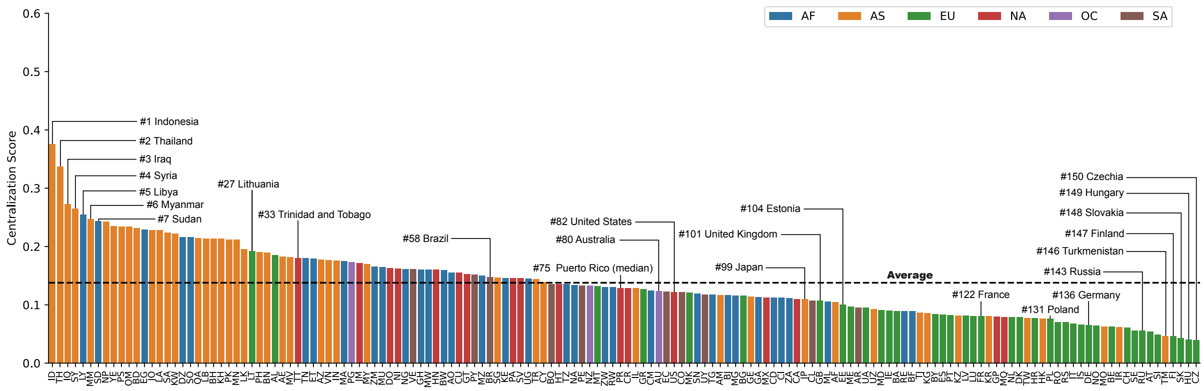


Figure 17: DNS Centralization (sorted by $\mathcal{S}$, color coded by continent)—It shows similar trends to that of hosting providers i.e. European countries tend to be less centralized, while Southeast Asian countries tend to be more centralized.

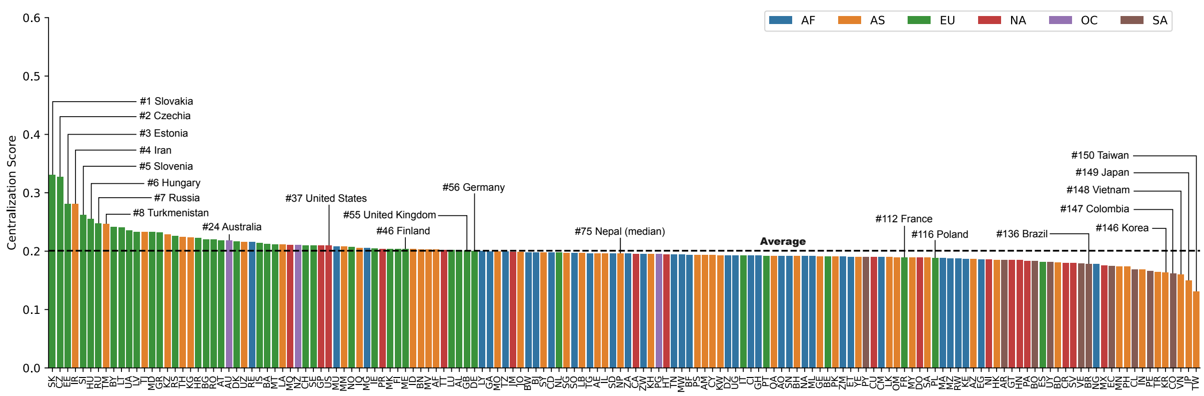


Figure 18: CA Centralization (sorted by $\mathcal{S}$, color coded by continent)—We see a shift from the hosting and DNS patterns. European countries tend to be more centralized in this layer.

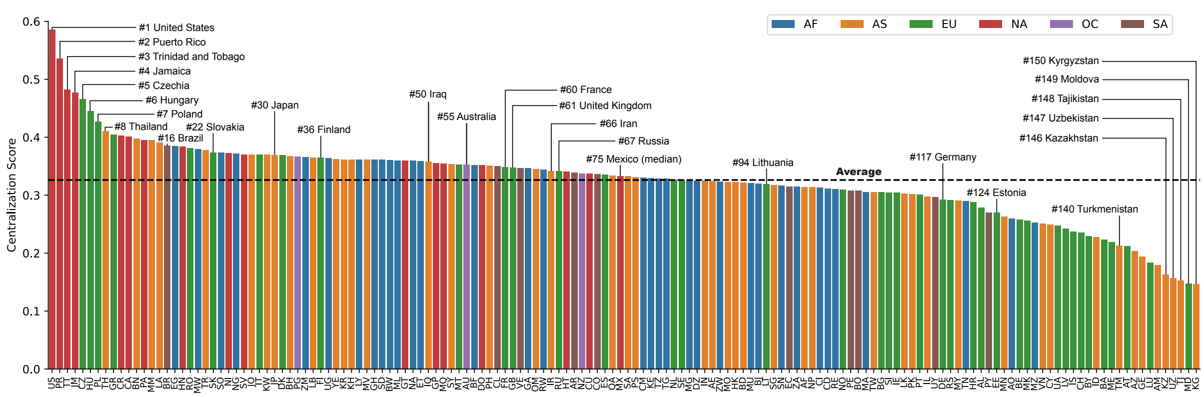


Figure 19: TLD Centralization (sorted by $\mathcal{S}$, color coded by continent)—North American countries show a tendency to be centralized. The CIS countries are on the other extreme, due to their dependence on a TLD apart from their own (`.ru`) beyond globally popular TLDs.

D INSULARITY CONTINENTAL TRENDS

Figures 20, 21 and 22 shows the distribution of countries' insularities color-coded by continent for hosting, DNS and TLD layers, respectively. See Figure 13 for corresponding results for CAs.

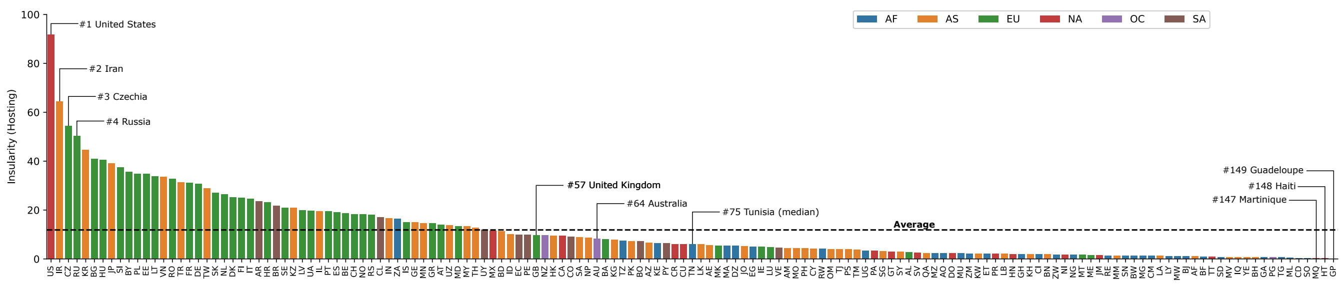


Figure 20: Hosting Provider Insularity—Countries sorted by % of websites served by hosting providers based in the same country color coded by continent. The U.S. is the most insular. While Iran is the second most insular it is also the least centralized because the top websites use a long tail of local providers. Countries in Europe tend to be more insular while countries in Africa do not show much insularity primarily due to a lack of equivalent providers in the country.

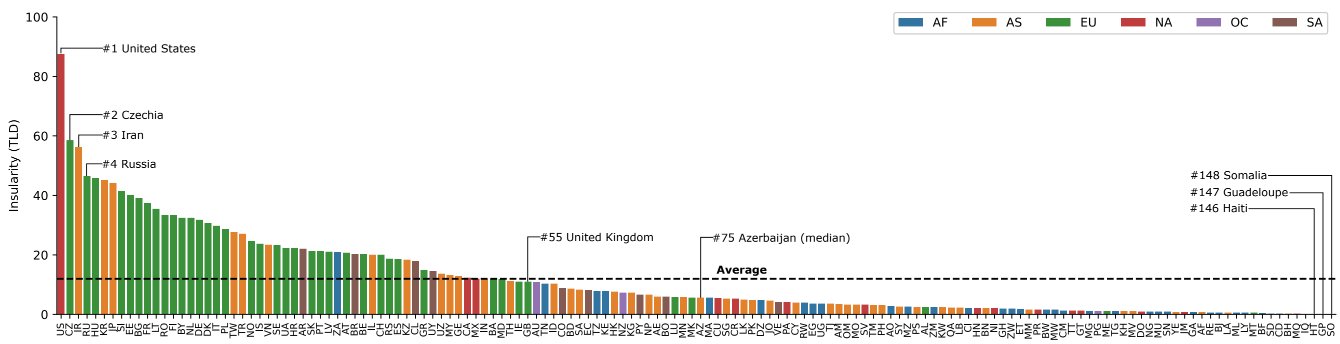


Figure 21: DNS Insularity by Country—Countries sorted by % of websites served by DNS infrastructure providers in the same country color coded by continent. The DNS shows similar trends to that of hosting providers.

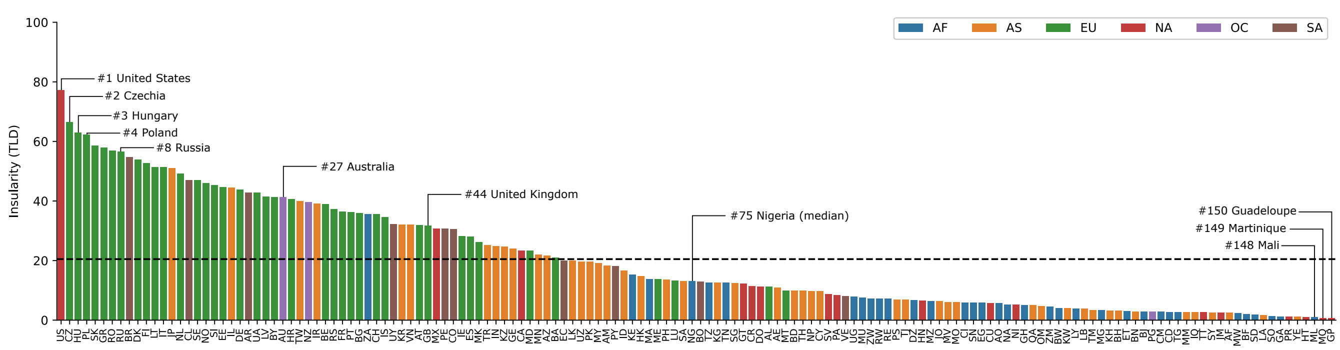


Figure 22: TLD Insularity by Country—Countries sorted by their % use of their own local ccTLD. We consider the use of `.com` to be insular to the U.S. given the historical role of the U.S. Government until recently in its operation [37]. Countries in Europe (especially Eastern Europe), Eastern Asia, and South America (especially Brazil) make heavy use of their local ccTLDs.

E COUNTRY REFERENCE

Table 4 lists the 150 countries in our dataset and their abbreviations.

CC	Country	Region	Continent	CC	Country	Region	Continent
AE	United Arab Emirates	Western Asia	AS	LT	Lithuania	Northern Europe	EU
AF	Afghanistan	Southern Asia	AS	LU	Luxembourg	Western Europe	EU
AL	Albania	Southern Europe	EU	LV	Latvia	Northern Europe	EU
AM	Armenia	Western Asia	AS	LY	Libya	Northern Africa	AF
AO	Angola	Middle Africa	AF	MA	Morocco	Northern Africa	AF
AR	Argentina	South America	SA	MD	Moldova	Eastern Europe	EU
AT	Austria	Western Europe	EU	ME	Montenegro	Southern Europe	EU
AU	Australia	Oceania	OC	MG	Madagascar	Eastern Africa	AF
AZ	Azerbaijan	Western Asia	AS	MK	North Macedonia	Southern Europe	EU
BA	Bosnia and Herzegovina	Southern Europe	EU	ML	Mali	Western Africa	AF
BD	Bangladesh	Southern Asia	AS	MM	Myanmar	South-eastern Asia	AS
BE	Belgium	Western Europe	EU	MN	Mongolia	Eastern Asia	AS
BF	Burkina Faso	Western Africa	AF	MO	Macao	Eastern Asia	AS
BG	Bulgaria	Eastern Europe	EU	MQ	Martinique	Caribbean	NA
BH	Bahrain	Western Asia	AS	MT	Malta	Southern Europe	EU
BJ	Benin	Western Africa	AF	MU	Mauritius	Eastern Africa	AF
BN	Brunei Darussalam	South-eastern Asia	AS	MV	Maldives	Southern Asia	AS
BO	Bolivia	South America	SA	MW	Malawi	Eastern Africa	AF
BR	Brazil	South America	SA	MX	Mexico	Central America	NA
BW	Botswana	Southern Africa	AF	MY	Malaysia	South-eastern Asia	AS
BY	Belarus	Eastern Europe	EU	MZ	Mozambique	Eastern Africa	AF
CA	Canada	Northern America	NA	NA	Namibia	Southern Africa	AF
CD	Congo	Middle Africa	AF	NG	Nigeria	Western Africa	AF
CH	Switzerland	Western Europe	EU	NI	Nicaragua	Central America	NA
CI	Côte d'Ivoire	Western Africa	AF	NL	Netherlands	Western Europe	EU
CL	Chile	South America	SA	NO	Norway	Northern Europe	EU
CM	Cameroon	Middle Africa	AF	NP	Nepal	Southern Asia	AS
CO	Colombia	South America	SA	NZ	New Zealand	Oceania	OC
CR	Costa Rica	Central America	NA	OM	Oman	Western Asia	AS
CU	Cuba	Caribbean	NA	PA	Panama	Central America	NA
CY	Cyprus	Western Asia	AS	PE	Peru	South America	SA
CZ	Czechia	Eastern Europe	EU	PG	Papua New Guinea	Oceania	OC
DE	Germany	Western Europe	EU	PH	Philippines	South-eastern Asia	AS
DK	Denmark	Northern Europe	EU	PK	Pakistan	Southern Asia	AS
DO	Dominican Republic	Caribbean	NA	PL	Poland	Eastern Europe	EU
DZ	Algeria	Northern Africa	AF	PR	Puerto Rico	Caribbean	NA
EC	Ecuador	South America	SA	PS	Palestine	Western Asia	AS
EE	Estonia	Northern Europe	EU	PT	Portugal	Southern Europe	EU
EG	Egypt	Northern Africa	AF	PY	Paraguay	South America	SA
ES	Spain	Southern Europe	EU	QA	Qatar	Western Asia	AS
ET	Ethiopia	Eastern Africa	AF	RE	Réunion	Eastern Africa	AF
FI	Finland	Northern Europe	EU	RO	Romania	Eastern Europe	EU
FR	France	Western Europe	EU	RS	Serbia	Southern Europe	EU
GA	Gabon	Middle Africa	AF	RU	Russia	Eastern Europe	EU
GB	United Kingdom	Northern Europe	EU	RW	Rwanda	Eastern Africa	AF
GE	Georgia	Western Asia	AS	SA	Saudi Arabia	Western Asia	AS
GH	Ghana	Western Africa	AF	SD	Sudan	Northern Africa	AF
GP	Guadeloupe	Caribbean	NA	SE	Sweden	Northern Europe	EU
GR	Greece	Southern Europe	EU	SG	Singapore	South-eastern Asia	AS
GT	Guatemala	Central America	NA	SI	Slovenia	Southern Europe	EU
HK	Hong Kong	Eastern Asia	AS	SK	Slovakia	Eastern Europe	EU
HN	Honduras	Central America	NA	SN	Senegal	Western Africa	AF
HR	Croatia	Southern Europe	EU	SO	Somalia	Eastern Africa	AF
HT	Haiti	Caribbean	NA	SV	El Salvador	Central America	NA
HU	Hungary	Eastern Europe	EU	SY	Syria	Western Asia	AS
ID	Indonesia	South-eastern Asia	AS	TG	Togo	Western Africa	AF
IE	Ireland	Northern Europe	EU	TH	Thailand	South-eastern Asia	AS
IL	Israel	Western Asia	AS	TJ	Tajikistan	Central Asia	AS
IN	India	Southern Asia	AS	TM	Turkmenistan	Central Asia	AS
IQ	Iraq	Western Asia	AS	TN	Tunisia	Northern Africa	AF
IR	Iran	Southern Asia	AS	TR	Turkey	Western Asia	AS
IS	Iceland	Northern Europe	EU	TT	Trinidad and Tobago	Caribbean	NA
IT	Italy	Southern Europe	EU	TW	Taiwan	Eastern Asia	AS
JM	Jamaica	Caribbean	NA	TZ	Tanzania	Eastern Africa	AF
JO	Jordan	Western Asia	AS	UA	Ukraine	Eastern Europe	EU
JP	Japan	Eastern Asia	AS	UG	Uganda	Eastern Africa	AF
KE	Kenya	Eastern Africa	AF	US	United States	Northern America	NA
KG	Kyrgyzstan	Central Asia	AS	UY	Uruguay	South America	SA
KH	Cambodia	South-eastern Asia	AS	UZ	Uzbekistan	Central Asia	AS
KR	Korea	Eastern Asia	AS	VE	Venezuela	South America	SA
KW	Kuwait	Western Asia	AS	VN	Viet Nam	South-eastern Asia	AS
KZ	Kazakhstan	Central Asia	AS	YE	Yemen	Western Asia	AS
LA	Laos	South-eastern Asia	AS	ZA	South Africa	Southern Africa	AF
LB	Lebanon	Western Asia	AS	ZM	Zambia	Eastern Africa	AF
LK	Sri Lanka	Southern Asia	AS	ZW	Zimbabwe	Eastern Africa	AF

Table 4: Reference for all the country codes and their regions.

F CENTRALIZATION SCORES FOR 150 COUNTRIES

Tables 5, 6, 7, and 8 show each country's centralization score for the hosting, DNS, CA, and TLD layers, respectively.

Rank	Country		CS	Rank	Country		CS	Rank	Country		CS	Rank	Country		CS
1	TH	AS	0.3548	41	JM	NA	0.1702	81	CM	AF	0.131	121	DE	EU	0.0947
2	ID	AS	0.3258	42	VN	AS	0.1694	82	CA	NA	0.1308	122	NO	EU	0.0937
3	MM	AS	0.2641	43	ZM	AF	0.1653	83	CR	NA	0.1287	123	HR	EU	0.0931
4	LA	AS	0.2526	44	AO	AF	0.1623	84	LT	EU	0.1286	124	AR	SA	0.0928
5	IQ	AS	0.249	45	GH	AF	0.1608	85	RW	AF	0.1275	125	ES	EU	0.0918
6	LY	AF	0.2462	46	MW	AF	0.1603	86	SN	AF	0.1273	126	TW	AS	0.0914
7	SY	AS	0.2379	47	IN	AS	0.16	87	TG	AF	0.1266	127	RS	EU	0.0905
8	PK	AS	0.23	48	ZA	AF	0.1549	88	CI	AF	0.1247	128	AF	AS	0.0904
9	KH	AS	0.2299	49	HN	NA	0.1545	89	BJ	AF	0.1244	129	PL	EU	0.0887
10	OM	AS	0.2287	50	NI	NA	0.1537	90	GA	AF	0.1232	130	BE	EU	0.088
11	SA	AS	0.2282	51	NZ	OC	0.1524	91	UA	EU	0.1228	131	MD	EU	0.0876
12	PS	AS	0.2254	52	MZ	AF	0.1519	92	CD	AF	0.1219	132	LV	EU	0.0873
13	KW	AS	0.2228	53	DO	NA	0.1511	93	PE	SA	0.1218	133	RO	EU	0.0869
14	YE	AS	0.2219	54	NA	AF	0.1508	94	CL	SA	0.1213	134	KG	AS	0.0868
15	LB	AS	0.2219	55	AU	OC	0.1504	95	MX	NA	0.1203	135	IT	EU	0.0859
16	JO	AS	0.2198	56	PA	NA	0.1495	96	ML	AF	0.1193	136	TJ	AS	0.0844
17	SD	AF	0.2188	57	NG	AF	0.1493	97	MK	EU	0.1192	137	CH	EU	0.0842
18	NP	AS	0.2167	58	VE	SA	0.1488	98	EC	SA	0.1192	138	MO	AS	0.0839
19	QA	AS	0.2161	59	PR	NA	0.1478	99	BG	EU	0.1188	139	KR	AS	0.0825
20	EG	AF	0.2155	60	GB	EU	0.1463	100	HK	AS	0.118	140	AT	EU	0.0816
21	BH	AS	0.2151	61	MT	EU	0.1462	101	RE	AF	0.114	141	FI	EU	0.0815
22	MY	AS	0.2143	62	CU	NA	0.1459	102	BA	EU	0.1121	142	KZ	AS	0.079
23	DZ	AF	0.2126	63	BR	SA	0.1446	103	AM	AS	0.1103	143	BY	EU	0.0766
24	SG	AS	0.2003	64	ZW	AF	0.1443	104	GE	AS	0.1086	144	SI	EU	0.0645
25	SO	AF	0.1991	65	KE	AF	0.1431	105	LU	EU	0.108	145	HU	EU	0.0604
26	BN	AS	0.1983	66	CY	AS	0.1418	106	FR	EU	0.1069	146	CZ	EU	0.0561
27	BD	AS	0.1971	67	UG	AF	0.1406	107	UY	SA	0.1066	147	RU	EU	0.0554
28	AE	AS	0.1937	68	IE	EU	0.1398	108	PT	EU	0.1065	148	SK	EU	0.0497
29	PH	AS	0.1934	69	TZ	AF	0.1395	109	NL	EU	0.1062	149	TM	AS	0.0461
30	MA	AF	0.1852	70	TR	AS	0.1394	110	CO	SA	0.1044	150	IR	AS	0.0411
31	TN	AF	0.1848	71	SV	NA	0.1374	111	JP	AS	0.1036				
32	MV	AS	0.1823	72	MN	AS	0.136	112	IS	EU	0.1025				
33	AL	EU	0.1806	73	HT	NA	0.1359	113	ME	EU	0.102				
34	ET	AF	0.1764	74	PY	SA	0.1359	114	SE	EU	0.1018				
35	TT	NA	0.1755	75	US	NA	0.1358	115	BF	AF	0.1018				
36	PG	OC	0.1755	76	GT	NA	0.134	116	GP	NA	0.1011				
37	LK	AS	0.1749	77	BO	SA	0.1335	117	DK	EU	0.101				
38	AZ	AS	0.1743	78	IL	AS	0.132	118	MQ	NA	0.1007				
39	MU	AF	0.1737	79	GR	EU	0.1319	119	UZ	AS	0.0978				
40	BW	AF	0.1727	80	MG	AF	0.1318	120	EE	EU	0.097				

Table 5: Country x Provider Centralization Scores

Rank	Country		CS	Rank	Country		CS	Rank	Country		CS	Rank	Country		CS
1	ID	AS	0.3757	41	JM	NA	0.1712	81	EC	SA	0.1227	121	LU	EU	0.0808
2	TH	AS	0.3374	42	MY	AS	0.1700	82	US	NA	0.1221	122	FR	EU	0.0805
3	IQ	AS	0.2730	43	ZM	AF	0.1651	83	CO	SA	0.1214	123	KR	AS	0.0804
4	SY	AS	0.2653	44	MU	AF	0.1643	84	MK	EU	0.1212	124	GP	NA	0.0797
5	LY	AF	0.2548	45	DO	NA	0.1628	85	SN	AF	0.1189	125	MQ	NA	0.0793
6	MM	AS	0.2469	46	NI	NA	0.1624	86	UY	SA	0.1179	126	NL	EU	0.0793
7	SD	AF	0.2439	47	NG	AF	0.1611	87	TG	AF	0.1173	127	DK	EU	0.0792
8	NP	AS	0.2430	48	VE	SA	0.1610	88	AM	AS	0.1168	128	TW	AS	0.0775
9	YE	AS	0.2346	49	GH	AF	0.1607	89	BJ	AF	0.1164	129	HR	EU	0.0774
10	PS	AS	0.2340	50	MW	AF	0.1601	90	MG	AF	0.1157	130	HK	AS	0.0760
11	OM	AS	0.2340	51	HN	NA	0.1600	91	BG	EU	0.1155	131	PL	EU	0.0760
12	BD	AS	0.2317	52	BW	AF	0.1594	92	GE	AS	0.1142	132	RO	EU	0.0704
13	EG	AF	0.2291	53	AO	AF	0.1553	93	GA	AF	0.1135	133	RS	EU	0.0703
14	JO	AS	0.2281	54	CU	NA	0.1549	94	MX	NA	0.1124	134	IT	EU	0.0676
15	LA	AS	0.2281	55	GT	NA	0.1531	95	CD	AF	0.1123	135	IS	EU	0.0660
16	SA	AS	0.2241	56	PY	SA	0.1517	96	CI	AF	0.1119	136	DE	EU	0.0656
17	KW	AS	0.2217	57	MZ	AF	0.1499	97	ZA	AF	0.1113	137	NO	EU	0.0644
18	DZ	AF	0.2159	58	BR	SA	0.1472	98	CA	NA	0.1099	138	MO	AS	0.0625
19	SO	AF	0.2157	59	SG	AS	0.1466	99	JP	AS	0.1097	139	BE	EU	0.0624
20	QA	AS	0.2140	60	KE	AF	0.1461	100	CL	SA	0.1072	140	IR	AS	0.0620
21	LB	AS	0.2139	61	PA	NA	0.1457	101	GB	EU	0.1072	141	CH	EU	0.0611
22	BH	AS	0.2136	62	SV	NA	0.1456	102	ML	AF	0.1052	142	SE	EU	0.0556
23	KH	AS	0.2136	63	UG	AF	0.1451	103	AF	AS	0.1047	143	RU	EU	0.0556
24	PK	AS	0.2115	64	TR	AS	0.1444	104	EE	EU	0.1001	144	AT	EU	0.0543
25	MN	AS	0.2115	65	CY	AS	0.1393	105	ME	EU	0.0966	145	SI	EU	0.0485
26	LK	AS	0.1956	66	BO	SA	0.1359	106	AR	SA	0.0953	146	TM	AS	0.0460
27	LT	EU	0.1919	67	HT	NA	0.1354	107	UA	EU	0.0953	147	FI	EU	0.0459
28	PH	AS	0.1900	68	TZ	AF	0.1352	108	UZ	AS	0.0924	148	SK	EU	0.0429
29	BN	AS	0.1892	69	NA	AF	0.1342	109	MD	EU	0.0907	149	HU	EU	0.0404
30	AL	EU	0.1855	70	PE	SA	0.1332	110	IE	EU	0.0897	150	CZ	EU	0.0391
31	AE	AS	0.1827	71	NZ	OC	0.1327	111	BA	EU	0.0894				
32	MV	AS	0.1817	72	MT	EU	0.1321	112	RE	AF	0.0894				
33	TT	NA	0.1805	73	ZW	AF	0.1305	113	BF	AF	0.0893				
34	TN	AF	0.1803	74	RW	AF	0.1300	114	TJ	AS	0.0868				
35	ET	AF	0.1796	75	PR	NA	0.1287	115	KG	AS	0.0862				
36	AZ	AS	0.1772	76	CR	NA	0.1286	116	BY	EU	0.0841				
37	VN	AS	0.1769	77	IL	AS	0.1284	117	ES	EU	0.0836				
38	IN	AS	0.1755	78	GR	EU	0.1266	118	PT	EU	0.0819				
39	MA	AF	0.1750	79	CM	AF	0.1246	119	KZ	AS	0.0818				
40	PG	OC	0.1732	80	AU	OC	0.1235	120	LV	EU	0.0813				

Table 6: Country x DNS Infrastructure Centralization Scores

Rank	Country		CS	Rank	Country		CS	Rank	Country		CS	Rank	Country		CS
1	SK	EU	0.3304	41	IQ	AS	0.2054	81	HT	NA	0.1945	121	AZ	AS	0.1863
2	CZ	EU	0.3268	42	MG	AF	0.2051	82	TN	AF	0.1943	122	EG	AF	0.1859
3	EE	EU	0.2811	43	IE	EU	0.2043	83	MW	AF	0.1943	123	NI	NA	0.1853
4	IR	AS	0.2807	44	PR	NA	0.2041	84	BF	AF	0.1937	124	HK	AS	0.1852
5	SI	EU	0.2623	45	MK	EU	0.2039	85	PS	AS	0.1937	125	AR	SA	0.1850
6	HU	EU	0.2555	46	FI	EU	0.2038	86	AM	AS	0.1936	126	GT	NA	0.1848
7	RU	EU	0.2474	47	ME	EU	0.2035	87	CY	AS	0.1932	127	HN	NA	0.1845
8	TM	AS	0.2462	48	ID	AS	0.2035	88	KW	AS	0.1930	128	PA	NA	0.1833
9	BY	EU	0.2418	49	BN	AS	0.2032	89	DZ	AF	0.1928	129	BO	SA	0.1828
10	LT	EU	0.2404	50	MV	AS	0.2030	90	UG	AF	0.1926	130	ES	EU	0.1816
11	UA	EU	0.2354	51	AF	AS	0.2030	91	IT	EU	0.1924	131	UY	SA	0.1810
12	LV	EU	0.2332	52	TT	NA	0.2022	92	CI	AF	0.1923	132	BD	AS	0.1804
13	TJ	AS	0.2331	53	LU	EU	0.2020	93	GH	AF	0.1922	133	CR	NA	0.1798
14	MD	EU	0.2329	54	AL	EU	0.2012	94	PT	EU	0.1920	134	SV	NA	0.1795
15	GR	EU	0.2323	55	GB	EU	0.2012	95	QA	AS	0.1920	135	VE	SA	0.1786
16	KZ	AS	0.2289	56	DE	EU	0.2005	96	AO	AF	0.1920	136	BR	SA	0.1779
17	RS	EU	0.2259	57	LY	AF	0.2004	97	SN	AF	0.1918	137	NG	AF	0.1779
18	TH	AS	0.2243	58	GA	AF	0.1996	98	BH	AS	0.1917	138	MX	NA	0.1750
19	KG	AS	0.2235	59	MO	AS	0.1995	99	NA	AF	0.1917	139	EC	SA	0.1745
20	HR	EU	0.2222	60	TZ	AF	0.1992	100	ML	AF	0.1913	140	MN	AS	0.1738
21	BG	EU	0.2200	61	JM	NA	0.1988	101	GE	AS	0.1910	141	PH	AS	0.1738
22	RO	EU	0.2198	62	JO	AS	0.1984	102	BE	EU	0.1910	142	CL	SA	0.1683
23	AT	EU	0.2183	63	BW	AF	0.1978	103	PK	AS	0.1908	143	IN	AS	0.1683
24	AU	OC	0.2179	64	BJ	AF	0.1976	104	ZM	AF	0.1907	144	PE	SA	0.1657
25	DK	EU	0.2165	65	SY	AS	0.1975	105	ET	AF	0.1903	145	TR	AS	0.1639
26	UZ	AS	0.2154	66	CD	AF	0.1974	106	YE	AS	0.1902	146	KR	AS	0.1631
27	RE	AF	0.2153	67	NL	EU	0.1973	107	PY	SA	0.1901	147	CO	SA	0.1618
28	IS	EU	0.2137	68	SG	AS	0.1971	108	CU	NA	0.1900	148	VN	AS	0.1599
29	BA	EU	0.2123	69	SO	AF	0.1967	109	CM	AF	0.1899	149	JP	AS	0.1499
30	MT	EU	0.2116	70	LB	AS	0.1966	110	LK	AS	0.1897	150	TW	AS	0.1308
31	LA	AS	0.2113	71	TG	AF	0.1963	111	OM	AS	0.1895				
32	MQ	NA	0.2107	72	AE	AS	0.1962	112	FR	EU	0.1891				
33	NZ	OC	0.2106	73	IL	AS	0.1958	113	MY	AS	0.1889				
34	CH	EU	0.2101	74	SD	AF	0.1956	114	DO	NA	0.1887				
35	SE	EU	0.2097	75	NP	AS	0.1956	115	SA	AS	0.1887				
36	GP	NA	0.2096	76	ZA	AF	0.1956	116	PL	EU	0.1884				
37	US	NA	0.2096	77	CA	NA	0.1953	117	MA	AF	0.1879				
38	MU	AF	0.2084	78	ZW	AF	0.1953	118	MZ	AF	0.1874				
39	MM	AS	0.2077	79	KH	AS	0.1952	119	RW	AF	0.1870				
40	NO	EU	0.2074	80	PG	OC	0.1949	120	KE	AF	0.1868				

Table 7: Country x CA Centralization Scores

Rank	Country		CS	Rank	Country		CS	Rank	Country		CS	Rank	Country		CS
1	US	NA	0.5853	41	LY	AF	0.3610	81	TG	AF	0.3284	121	HR	EU	0.2878
2	PR	NA	0.5358	42	MV	AS	0.3609	82	NL	EU	0.3270	122	AL	EU	0.2781
3	TT	NA	0.4821	43	GH	AF	0.3609	83	SE	EU	0.3258	123	PY	SA	0.2700
4	JM	NA	0.4771	44	SD	AF	0.3608	84	MG	AF	0.3254	124	EE	EU	0.2694
5	CZ	EU	0.4656	45	BW	AF	0.3600	85	DZ	AF	0.3252	125	MN	AS	0.2624
6	HU	EU	0.4450	46	ML	AF	0.3595	86	IN	AS	0.3250	126	AO	AF	0.2592
7	PL	EU	0.4265	47	GT	NA	0.3595	87	AE	AS	0.3245	127	BE	EU	0.2573
8	TH	AS	0.4108	48	NA	AF	0.3591	88	ZW	AF	0.3233	128	MK	EU	0.2560
9	GR	EU	0.4044	49	ET	AF	0.3586	89	MO	AS	0.3227	129	MZ	AF	0.2524
10	CR	NA	0.4022	50	IQ	AS	0.3579	90	HK	AS	0.3223	130	VN	AS	0.2506
11	CA	NA	0.4008	51	GP	NA	0.3552	91	BD	AS	0.3214	131	CY	AS	0.2486
12	BN	AS	0.3979	52	MQ	NA	0.3539	92	MU	AF	0.3203	132	UA	EU	0.2470
13	PA	NA	0.3951	53	SY	AS	0.3535	93	BJ	AF	0.3200	133	LV	EU	0.2421
14	MM	AS	0.3945	54	MT	EU	0.3530	94	LT	EU	0.3186	134	IS	EU	0.2367
15	LA	AS	0.3903	55	AU	OC	0.3530	95	SG	AS	0.3174	135	CH	EU	0.2356
16	BR	SA	0.3856	56	BF	AF	0.3521	96	SN	AF	0.3166	136	BY	EU	0.2289
17	EG	AF	0.3846	57	DO	NA	0.3517	97	EC	SA	0.3144	137	ID	AS	0.2272
18	HN	NA	0.3837	58	PH	AS	0.3510	98	ZA	AF	0.3143	138	BA	EU	0.2228
19	RO	EU	0.3811	59	CL	SA	0.3496	99	AF	AS	0.3142	139	ME	EU	0.2192
20	MW	AF	0.3797	60	FR	EU	0.3481	100	NP	AS	0.3138	140	TM	AS	0.2128
21	TR	AS	0.3776	61	GB	EU	0.3470	101	CI	AF	0.3128	141	AT	EU	0.2123
22	SK	EU	0.3731	62	VE	SA	0.3469	102	CD	AF	0.3108	142	AZ	AS	0.2035
23	SO	AF	0.3729	63	GA	AF	0.3468	103	RE	AF	0.3106	143	GE	AS	0.1936
24	NI	NA	0.3723	64	OM	AS	0.3450	104	NO	EU	0.3098	144	LU	EU	0.1838
25	NG	AF	0.3713	65	RW	AF	0.3439	105	PE	SA	0.3077	145	AM	AS	0.1794
26	SV	NA	0.3701	66	IR	AS	0.3418	106	BO	SA	0.3076	146	KZ	AS	0.1629
27	JO	AS	0.3701	67	RU	EU	0.3416	107	MA	AF	0.3055	147	UZ	AS	0.1569
28	IT	EU	0.3700	68	HT	NA	0.3407	108	TW	AS	0.3054	148	TJ	AS	0.1526
29	KW	AS	0.3699	69	AR	SA	0.3391	109	BG	EU	0.3051	149	MD	EU	0.1475
30	JP	AS	0.3693	70	NZ	OC	0.3369	110	SI	EU	0.3043	150	KG	AS	0.1468
31	DK	EU	0.3692	71	CU	NA	0.3367	111	IE	EU	0.3040				
32	BH	AS	0.3668	72	CO	SA	0.3364	112	LK	AS	0.3024				
33	PG	OC	0.3666	73	ES	EU	0.3355	113	PK	AS	0.3015				
34	ZM	AF	0.3658	74	QA	AS	0.3339	114	PT	EU	0.3009				
35	LB	AS	0.3647	75	MX	NA	0.3326	115	IL	AS	0.2971				
36	FI	EU	0.3646	76	SA	AS	0.3325	116	UY	SA	0.2966				
37	UG	AF	0.3635	77	PS	AS	0.3311	117	DE	EU	0.2920				
38	YE	AS	0.3620	78	CM	AF	0.3302	118	RS	EU	0.2914				
39	KR	AS	0.3613	79	KE	AF	0.3293	119	MY	AS	0.2905				
40	KH	AS	0.3610	80	TZ	AF	0.3284	120	TN	AF	0.2893				

Table 8: Country x TLD Centralization Scores