A World Wide View of Browsing the World Wide Web

Kimberly Ruth
Stanford University
kcruth@cs.stanford.edu

Aurore Fass
Stanford University, CISPA Helmholtz Center for Information Security
aurore@cs.stanford.edu

Jonathan Azose*
OctoML
jazose@octoml.ai

Mark Pearson
Google, Inc.
mpearson@google.com

Emma Thomas
Google, Inc.
egthomas@google.com

Caitlin Sadowski
Google, Inc.
supertri@google.com

Zakir Durumeric
Stanford University
zakir@cs.stanford.edu

ABSTRACT

In this paper, we perform the first large-scale study of how people spend time on the web. Our study is based on anonymous, aggregate telemetry data from several hundred million Google Chrome users who have explicitly enabled sharing URLs with Google and who have usage statistic reporting enabled. We analyze the distribution of web traffic, the types of websites that people visit and spend the most time on, the differences between desktop and mobile browsing behavior, the geographical differences in web usage, and the most popular websites in regions worldwide. Our study sheds light on online user behavior and how the research community can more accurately analyze the web in the future.

ACM Reference Format:

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 (IMC '22), October 25–27, 2022, Nice, France.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3517745.3561418

1 INTRODUCTION

Despite the web's importance to modern society, we know relatively little about how users browse the web beyond what has been observed through local, small-scale studies [15, 18, 34], user surveys [3, 24, 32], and ranked lists of websites like the Alexa Top Million [25]. In this paper, we investigate how people use the web, the types of websites they frequent, and types of websites that they spend the most time on. Our study is based on traffic distribution data and ranked lists of popular websites anonymously aggregated from several hundred million global users of Google Chrome who have explicitly opted into browser history synchronization through Chrome and who have usage statistic reporting enabled. We analyze the types of sites and the most popular individual sites that users visit, broken down by platform (desktop vs. mobile), country, and popularity metric (page loads vs. time spent on page).

We start by analyzing the distribution of user traffic on the web. We find that browsing is concentrated toward a small handful of websites: six sites account for 25% of page loads on both desktop and mobile, and one site garners 17% of all desktop page loads globally. Time spent on page is even further concentrated, with ten sites accounting for about half of time spent on the web for desktop users. The top million sites capture over 95% of all page loads and time spent online, but they do so extremely unequally.

Top use cases of the web are strikingly similar globally—most countries have search engines, video sharing sites, social media, adult content, and entertainment sites in their top ten most visited sites. A handful of globally popular websites (e.g., google.com, netflix.com) drive large amounts of traffic in most countries. Across many countries, browsing behavior differs between desktop and mobile. Desktop users most frequently load search engines but spend the most time watching video streaming content. Beyond entertainment, desktop users spend more time browsing educational and work-related sites than mobile users do. On mobile, time on the web is also broadly spent on entertainment and leisure, but more time is spent browsing adult content than any other type of website by time on page. This result is likely exacerbated by common entertainment platforms like Netflix and YouTube having dedicated mobile apps that are not accessed through the web. Pornography, lifestyle, and weather websites are all accessed disproportionately on mobile devices compared to desktop.

However, we also observe significant differences in the specific sites visited based on where users are located. Beyond the small set of globally popular sites, most websites are country-specific. For example, of sites appearing in the top 1K for at least one country, over half do not rank in the top 10K for any other country, and even among the top 20 sites for a given country, at least half are only nationally popular. Complicating analysis, categories of websites are unevenly localized. For example, education, politics, and finance content tend to be country-specific, compared to technology, adult sites, and gaming content that tend to be more similar across countries. Some browsing trends are shared by countries that are either geographically close or share a common language, but these similarities are overshadowed by local differences.

Our study highlights several trends that may impact future research studies about the web. For instance, our findings on country differences suggest that global top lists underrepresent some categories of sites and some regions, and we emphasize the value of breaking down analyses of the web by country. Popularity lists

*Work performed while at Google, Inc.

IMC '22, October 25–27, 2022, Nice, France

ACM ISBN 978-1-4503-9259-4/22/10.
https://doi.org/10.1145/3517745.3561418

aggregated by page loads vs. time spent on page are only moderately correlated, underscoring the importance of choosing metrics appropriate to the analysis at hand. Due to concentration at the head of the web, studies calculated directly from a simple set of the top million sites place disproportionate focus on the long tail of the web. Furthermore, studies' choice of rank magnitude bucket affects the category composition of included sites: e.g., video streaming sites account for upwards of 40% of top-10 sites by time on page but less than 10% of the top 10K. Temporally, while top sites are typically stable between months, December shows noticeable bias towards e-commerce sites and away from educational sites, and may not be the best choice of analysis time window.

We hope that our results both shed light on how people use the web in 2022 as well as enable researchers to more accurately characterize the web in future studies.

2 RELATED WORK

Prior literature dedicated to web browsing behavior is surprisingly sparse. Most previous studies of browsing behavior are limited in geographic scope, rely on user-reported behavior, and/or predate the rise of mobile web usage.

The earliest studies of web usage are now more than a decade old. Montgomery et al. performed a coarse-grained analysis on the growth of the web between 1997 and 1999, highlighting increases in the number of web users and page views per user [22]. Other early measurement studies analyzed web traffic within the context of network-level traffic measurement from a local vantage point: Pang et al. [26] for enterprise network traffic and Maier et al. [20] for residential broadband traffic. The web has changed substantially since these early works, as have the insights we aim to gain.

There have been several small-scale qualitative studies of web usage. Early work by Sheehan et al. [32] (2002) described several use cases for the web based on user-completed journals. Chen et al. conducted a telephone survey in 2010 to study the Internet-usage patterns of Chinese immigrants in Singapore [3], finding that the longer immigrants stay in Singapore, the more likely they are to change their usage patterns and to browse websites from Singapore. In 2013, Mueller et al. compared smartphone vs. tablet web use in the U.S. based on user-completed journals [24]; they found that users primarily use their smartphone to communicate, while the tablet is popular for content consumption and entertainment.

Larger scale industry datasets yielded several studies of web browsing in the early 2010s. Kumar et al. [18] (2010) and Tikhonov et al. [34] (2015) studied how users navigate between pages based on 7-day search and toolbar logs from Yahoo! and 3-month toolbar logs from yandex.ru, respectively. Kumar et al. also categorized user pageviews based on a coarse taxonomy of content, communication, and search. In 2012, Goel et al. [15] leveraged the web histories of 250K anonymized users to analyze the frequency with which people browsed the web for different activities from 2009–2010, finding that the top 20% of users generated more than 60% of page views. Goel et al.'s study is closest to ours, but now more than ten years old. Torres et al. [11] (2014) investigated the search behavior of children using data collected through a browser toolbar. More recently (2018), Ng et al. studied browsing patterns between countries based on the Alexa Top 100 [25]. The study highlighted that countries with similar language and geographic proximity tend to have similar popular websites, in line with our findings.

In the commercial space, web analytics companies aim to characterize how people find web content. Researchers frequently treat publicly available top website lists such as the Cisco Umbrella 1 Million [6], the Majestic Million [21], and the now-defunct Alexa Top 1 Million [1] as indicative of web browsing behavior, but these lists have recently come under scrutiny due to brittleness [19] and inaccuracy [27]. Other commercial offerings (e.g., from SimilarWeb [33], Comscore [9], and Semrush [31]) are paywalled, rely on opaque and proprietary methodologies, and often cater to marketing and SEO customers rather than describing web behavior more broadly.

Overall, our work presents a holistic, large-scale view of how people browse the web, based on different metrics (page loads and time on page), for desktop and mobile, and across 45 countries.

3 METHODOLOGY AND ETHICS

We analyze user browsing behavior based on client telemetry data from Google Chrome, which we augment with website categorization data queried from Cloudflare. In this section, we describe these data sources and the ethics of our study.

3.1 Google Chrome Dataset

Our study is based on a dataset from the Google Chrome browser that consists of rank order lists of the top million most popular websites per month, broken down by country, platform, and popularity metric (e.g., top million websites by *Page Loads* for *Windows Users* in the *United States*) as seen by the browser. This data was collected only for users who chose to explicitly enable sharing browsing history with Chrome, and who have usage statistics reporting enabled. In total, our study is based on several hundred million users globally with both sharing URLs and usage reporting enabled.

Google additionally filtered the data used for this study to protect user privacy. First, the dataset excludes any websites with fewer visits from unique clients (i.e., browser installs) than a set threshold; this threshold was chosen both for privacy reasons and to ensure that we have enough samples to be confident in the statistical distributions for included pages. Second, when computing time spent on page, Chrome down-samples events such that each page foreground event has only approximately a 0.35% chance of being uploaded to Chrome, ensuring that we do not have a perfect view of browser history for any particular client. Last, Chrome excludes any visits to *non-public domains*—domains that are not hyperlinked from public websites or specify that they may not be crawled per `robots.txt`. We note that because of these safeguards, smaller countries and/or countries with low Chrome adoption often have fewer than 10K websites in the dataset.

Chrome specifically shared data for September 2021 to February 2022, aggregated monthly by domain and broken down along the following dimensions:

- **Platform/Operating System:** Chrome supports five platforms: Windows, Linux, Mac OS, Android, and Apple iOS. We limit our analysis to the two largest platforms, Windows and Android, due to a relatively small number of users on Mac OS, Linux, and iOS who share data within each region of our study. We use the traffic on Windows and Android to compare and contrast desktop versus mobile usage.

- **Popularity Metric:** Chrome tracks popularity across three metrics: initiated page loads, completed page loads, and time spent on page. We exclude analysis of initiated page loads since the metric is nearly identical to completed page loads (most page loads are successful). Page loads is defined as the number of times the content of a web page is loaded in the browser window (First Contentful Paint). Chrome filters out most non-user-initiated navigations (e.g., iFrames). Time on page is the total time with the window in the foreground, recorded in milliseconds each time the page is backgrounded.
- **Country:** We limit our analysis to the 80 countries with at least 10K websites throughout our study period because this helps to filter out countries without a large, representative install base. We further limit our analysis to 10 countries per continent, resulting in aggregate global metrics typically being computed as the median of 45 countries: 10 of which are in Asia, 10 in Europe, 9 in South America, 7 in North America, 7 in Africa, and 2 in Oceania (Appendix A). In most countries, 10K sites account for at least 70–85% of desktop traffic and 70–80% of mobile traffic by page load. This choice balances geographic diversity in the countries we study while also helping to ensure that we see sufficient, representative traffic from each country.

Unless stated otherwise, the analysis we present is based on Windows page loads from February 2022 based on the top 10K sites in the 45 countries we consider; in most cases we present median values calculated across the set of countries. As we discuss in Section 4.5, browsing behavior is relatively temporally stable and we expect February to be a representative month caveated by the study taking place during the COVID-19 pandemic.

There are several caveats and limitations to our dataset:

Dataset Representativeness. Prior work has shown that Google Chrome's public Chrome User Experience Report (CrUX) dataset [4], which is based on the same telemetry data as our study, provides the most accurate perspective on site popularity compared to other public datasets (e.g., Alexa Top Million) [27]. However, we note that our study is limited to users who *opt in* to data sharing and does not include data from private browsing sessions, which could bias our results (e.g., we may miss a larger share of adult content browsing [14]). In addition, Android telemetry misses traffic from most native apps, except those that use Custom Tabs or WebAPKs.

Aggregating Sites Across Domains. During our investigation, we find that many top sites are hosted under multiple ccTLDs, which creates noise when aggregating metrics globally. When comparing sites across countries, we merge websites when a secondary version exists under another eTLD (e.g., we aggregate `google.co.uk` with `google.com`), as defined by the Mozilla Public Suffix list [12] to address this problem. This process is imperfect (e.g., `top.com` is a social cryptocurrency exchange and `top.gg` is a ranking of Discord bots and servers). However, manual inspection finds that these errors are infrequent enough to not affect our results substantially.

Time-on-page Telemetry Error. Between the time of analysis and publication, a bug was discovered in the time-on-page telemetry collection [5], manifesting in a fraction of a single percent of domains. Given the small number of affected domains, we do not expect this to change any high-level results, but we conservatively use only the traffic distribution curves for page loads to model traffic volume in Section 4.2 and beyond.

Public Data Access. Although the data we use for this study is not public, a coarser-grained version is available publicly through the CrUX dataset [4, 10]. The public dataset consists of rank-order magnitude buckets of websites ranked by completed page loads and aggregated both per-country and globally.

3.2 Categorizing Websites

To identify broad trends in browsing behavior, we categorize websites using Cloudflare's Domain Intelligence API, which provides information about all domains (i.e., not limited to Cloudflare-hosted sites) [7]. The API places domains into 26 super-categories (e.g., Education) and 114 categories (e.g., Edu/Educational Institutions and Edu/Science).

To ensure the accuracy of our categorization, we manually validate ten random websites per category and keep only categories with at least 80% accuracy. As a result, we exclude 19 categories and merge their websites into our *Other/Unknown* category. In addition, we merge similar categories with a small number of sites or significantly overlapping definitions (e.g., Chat, Instant Messengers, and Messaging). This results in 22 super-categories and 61 categories. We detail the resulting taxonomy in Appendix B. Notable among the low-accuracy categories are *Search Engines* and *Social Networks*. Given that these categories are among the main use cases for the web, we manually verified sites in these categories within the top 100 sites for each country. We find that 56/60 search engine domains and 13/14 social network domains are correct, and we use only the sets of manually verified sites for these two categories in our study.

3.3 Ethics

To protect user privacy, we analyzed only aggregate web traffic distribution data and ranked websites for this study. For the distribution data, we analyzed only the number of websites accounting for varying percentiles of traffic (both globally and per country); no data on individual clients was shared. As for the ranked websites, data was aggregated at the country and platform level and no websites with fewer unique clients than Chrome's privacy threshold were included each month.

Our study considers only users who chose to explicitly enable sharing URLs with Chrome and have usage statistics reporting enabled. Chrome's privacy policy states in plain language (not only in EULA text) the situations in which URLs will be shared with Chrome, and users have the option to control whether this data is sent. Chrome's privacy policy also states that it uses collected information for research and development purposes. Chrome did not collect or store any additional data specifically for our study.

Stanford's IRB guidelines, which cover the Stanford researchers, indicate that analyzing aggregate data "without any individually identifiable information," as we used for this study, is not Human Subjects Research. Only aggregate data was analyzed for this study and the Stanford team governed never had access to any non-aggregate data. The research team never analyzed any individual clients or shared non-aggregate data outside of Chrome.

4 GLOBAL BROWSING BEHAVIOR

In this section, we investigate global trends in web browsing behavior, including how much browsing is concentrated on top websites, globally consistent uses of the web, and the differences between desktop and mobile browsing.

4.1 Distribution of Browsing Across Sites

We first examine the global distribution of page loads and time spent on websites (e.g., how many page loads do top sites see versus less popular sites?). We find that a significant portion of web traffic is concentrated on a small number of websites, but that there is no natural cutoff between top sites and the long tail of the web.

4.1.1 Traffic Distribution Data. The traffic volume data in this section is provided directly by Chrome. It is aggregated globally (rather than per country) and encompasses traffic to all websites, including sites excluded from subsequent analyses due to privacy thresholding, because the distribution data contains no identifying website or client data. As data is aggregated globally, rather than computed across countries, this section's results will be more heavily weighted towards countries with more web usage than other analyses in this paper.

4.1.2 Web Traffic Concentration. As can be seen in Figure 1, web browsing is heavily concentrated on popular websites. Indeed, a single website accounts for 17% of all Windows page loads globally. Beyond that, 25% of Windows page loads globally are served by only six sites, the top 100 websites capture just under 40%, top 10K around 70%, and the top million over 95% of page loads. Within each country, traffic is further concentrated towards top sites for that country, and the top ranked website in each country captures 12–33% of all page loads (median, 20%). On Windows, we see Google, YouTube, Facebook, WhatsApp, Roblox, and Amazon within the top six sites for at least ten countries. In nearly every country (44 / 45), Google is the top website by page loads, though Korean search engine Naver is top ranked in South Korea.

Time on Page is even more concentrated on popular sites than *Page Loads*, but captures a different set of sites. On Windows, the top site accounts for 24% of time spent on page, and half of user time is spent on just 7 sites; the top 100 sites capture over 60% of user time and the top 10K over 85%. The sites that appear in the top seven for at least 25 countries are YouTube, Google, Facebook, Netflix, WhatsApp, and Twitch. Users spend the most time on YouTube in 40 / 45 countries; Google is the top site for the remaining 5 countries, including the United States.

Despite this concentration, we find no natural cutoff between the "top sites" and "long tail" of the web by either time on page or page loads. Our results indicate that when studying the web, analyzing the top million websites as a uniform set captures the vast majority of websites that users visit on a regular basis, but it skews analysis towards a long tail of websites where users spend relatively little time.

Android exhibits less concentration towards top sites than Windows. Ten websites account for 25% of traffic versus six sites on desktop. Google ranks top for 41 / 45 countries, but other top websites begin to deviate. This is likely because popular sites like YouTube and Netflix have native mobile apps. Of the 114 sites ranking in the top 10 in at least one country by page loads on Windows but not Android, 93 (82%) have a dedicated Android app. Six websites rank in the top 10 for at least 20 countries: Google, Facebook, XNXX, XVideos, Pornhub, and AMP Project.[1] Android time spent is likewise more evenly distributed among the topmost sites, though 25% of traffic targets just 8 sites; the top 100 sites cover over 40% of user time and the top 10K just under 80%. We detail the differences between desktop and mobile behavior in Section 4.3.

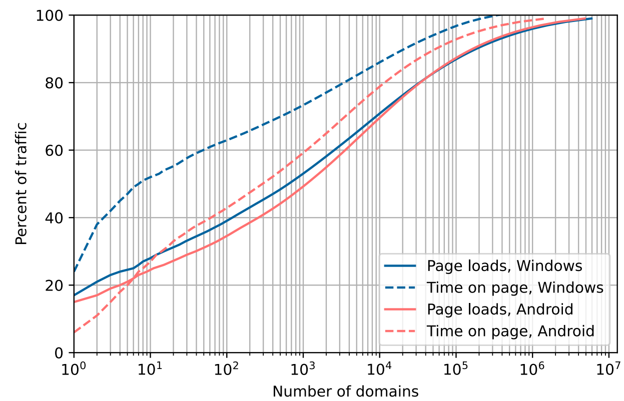

Figure 1: Distribution of Web Traffic By Website Rank—We show the distribution of web traffic towards popular sites as measured by both *Page Loads* and *Time on Page*. Browsing is heavily skewed towards top websites with a single website capturing 17% of all page loads, 10K sites 70%, and 1M sites over 95%. Time on Page is further concentrated: half of user time goes to just under 300 sites on Android and only 7 sites on Windows, and the top 10K capture about 80-85% of time spent.

In summary, users spend a significant portion of time on a small handful of websites, but even when computing simple metrics, we find that there are notable differences between mobile and desktop usage, page loads versus time on page, and between countries, which we explore further in the next several sections.

4.2 Use Cases for the Web

In this section, we aim to answer the question "What do people use the web for?" To answer this, we look at both the top ranked sites in each country as well as the categories of websites in the top 10K. We find that while there are many similarities in the types of top sites, the specific sites that users visit vary between countries, which we explore in Section 5.

4.2.1 Composition of the Top 10 Sites. We first investigate the top ten websites in each country, which typically account for a quarter to half of web traffic, depending on metric and platform. To ensure the accuracy of discussions about specific websites, we manually verified the identity of and categorized the top ten websites per (country, platform, metric) breakdown. Many of the top ten sites are

[1]Recall that we analyze site data only at domain granularity. AMP Project traffic is likely overwhelmingly composed of other sites' content served through AMP URLs, not from users intentionally navigating to the AMP Project site.

shared across the 45 countries in our study, and across the 1.8K domains found in the union of breakdowns, we identify and manually categorize 469 unique domains that belong to 402 websites.

Countries show striking similarity in the types of top sites. Even when exact sites do not match, we find that all 45 countries in our study have at least one search engine (e.g., Google, Baidu) and video sharing platform (e.g., YouTube, Bilibili) in the top ten. Most have social networking platforms (44 countries, e.g., Facebook, OK.ru), adult content (43, e.g., Pornhub, XVideo), e-commerce (32, e.g., Amazon, Taobao), chat/messaging (30, e.g., WhatsApp, Zalo), and classified ads (17, e.g., Craigslist). Many countries also have gaming-related sites like Twitch (31) and Roblox (26), news (20), banks (17), and multi-purpose portals (e.g., Naver) that combine email, search, news, weather, and beyond (21).

Except for Japan, Vietnam, and Russia, we see generic movie and TV streaming sites in all countries. Netflix shows by far the largest global adoption (41 of 42 countries). However, we also see a mix of 26 additional streaming sites, 18 of which provide free and/or pirated content in 15 countries. Anime (12 countries, incl. Russia) and Manga (15 countries, incl. Vietnam) are popular as well; the only video-related sites in Japan are Twitch and Nico, two video sharing sites. Beyond consumer uses, we see business platforms (22 / 45 countries, e.g., Sharepoint, Office 365), universities (10), government services (26), and then a long tail of other types of sites that include gig economy (3), EdTech (6), ISPs and telecoms (9, 4 of which provide TV service and 2 of which provide email), and job search sites (4), which we detail in Appendix Table 4.

4.2.2 Beyond the Top 10 Sites. Looking beyond the individual top sites, we consider the categories of websites in the top 10K websites per country, which typically capture 70–80% of user traffic (as noted in Section 3.1, we are precluded from analyzing beyond the top 10K based on privacy thresholds in most countries). Given the non-uniform distribution of traffic amongst sites, simply counting the number of sites per category skews our results towards the tail. As such, we additionally model the percent of page loads and time on page per category by computing a weighted count of sites per category with our traffic distribution data from Section 4.1. We present both perspectives, and we take a global view of category prevalence by taking the average of each category statistic across the 45 countries in our study. We show the breakdown for the top 100 and top 10K for both perspectives in Figure 2.

Perhaps unsurprisingly, given their number one rank in every country, search engines capture the plurality of page loads for both desktop and mobile users with 20–25% of top-10K page loads. However, while users load search engines most frequently, they do not spend the most time on them. Rather, we find that users spend the plurality of their time on video streaming sites on Windows (33% of time spent on top-10K websites).

Mobile time on page is similarly dominated by entertainment and leisure content, but the plurality of time is spent on adult content when browsing on mobile (18%). This is in part because users view more adult content on mobile devices than on desktop platforms [23], but also because users spend less time streaming non-adult video over the web using mobile browsers, because of native applications. We note that different per-platform usage rates of incognito mode (which is not captured in Chrome telemetry) may also cause us to underestimate adult content in particular. Other categories that attract large volumes of mobile traffic include News & Media (8.9–9.8% of weighted traffic volume, 6.5–14.3% of domains) and Video Streaming (11.0–15.5% of time on page). After adult content and video streaming, e-commerce and education are the next two most prominent use cases for the mobile web.

4.2.3 Category Prevalence by Rank. The notable differences in breakdown between top-100 and top-10K sites seen in Figure 2 leads us to further investigate the differences in composition between the head and tail of the web. In particular, some categories of sites may form a more concentrated ecosystem than others. That is, for two categories that receive a comparable amount of user traffic, one category may contain *fewer but more popular* sites than the other. To quantify this phenomenon, for a range of rank thresholds N, approximately logarithmically spaced, we compute the percentage of domains in the top N with each category label. We plot the median and 25–75% quartiles among 45 countries at each rank threshold for a selection of categories in Figure 3.

Category distributions vary by rank threshold. For instance, Video Streaming represents a higher proportion of top-10 sites than top-10K sites when ranking by time spent. By contrast, Business is disproportionately represented among long-tail sites, rising from just above 3% of top-30 sites to over 8% of top-10K desktop sites. Other categories are disproportionately represented in the middle of the range: News & Media peaks above 15% of top-50 sites and drops to less than 7% of top-10K sites. Some categories remain stable across rank (e.g., technology with 10–12% of desktop and 5–7% of mobile). Other disproportionate category representations occur for only one platform or one popularity metric. For example, adult content is disproportionately represented among top-50 sites on only mobile devices. Together, these variations begin to account for the different composition of top sites and long tail sites.

Summary. Countries often share major web use cases among their top sites, though specific individual sites and their ranks differ. While search engines see the largest fraction of page loads, users spend the most time online watching entertainment, though the type of content (e.g., streaming sites versus adult content) varies considerably across platforms. Beyond entertainment and search, the web consistently facilitates a spectrum of everyday activities, including work, education, and commerce.

4.3 Desktop vs. Mobile Browsing Behavior

In the previous section, we observed that browsing behavior differs between desktop and mobile platforms. In this section, we investigate which types of sites are *disproportionately* visited on desktop and mobile by examining the per-category variations between desktop and mobile platforms.

Estimating Traffic Volume. As in Section 4.2.2, we estimate the amount of traffic for websites in each category by computing a weighted sum per category for the top 10K sites, with weights drawn from the distributions in Section 4.1. We then compare traffic volumes per category across desktop and mobile by computing Fisher's binomial proportion test[2] ($p = 0.05$) with a Bonferroni

[2]Fisher's test requires each proportion to be out of a specified sample size. Because our dataset of ranked lists does not provide raw traffic volume to each site, we instead set a conservative sample size of $N = 10,000$ (the number of sites under consideration).

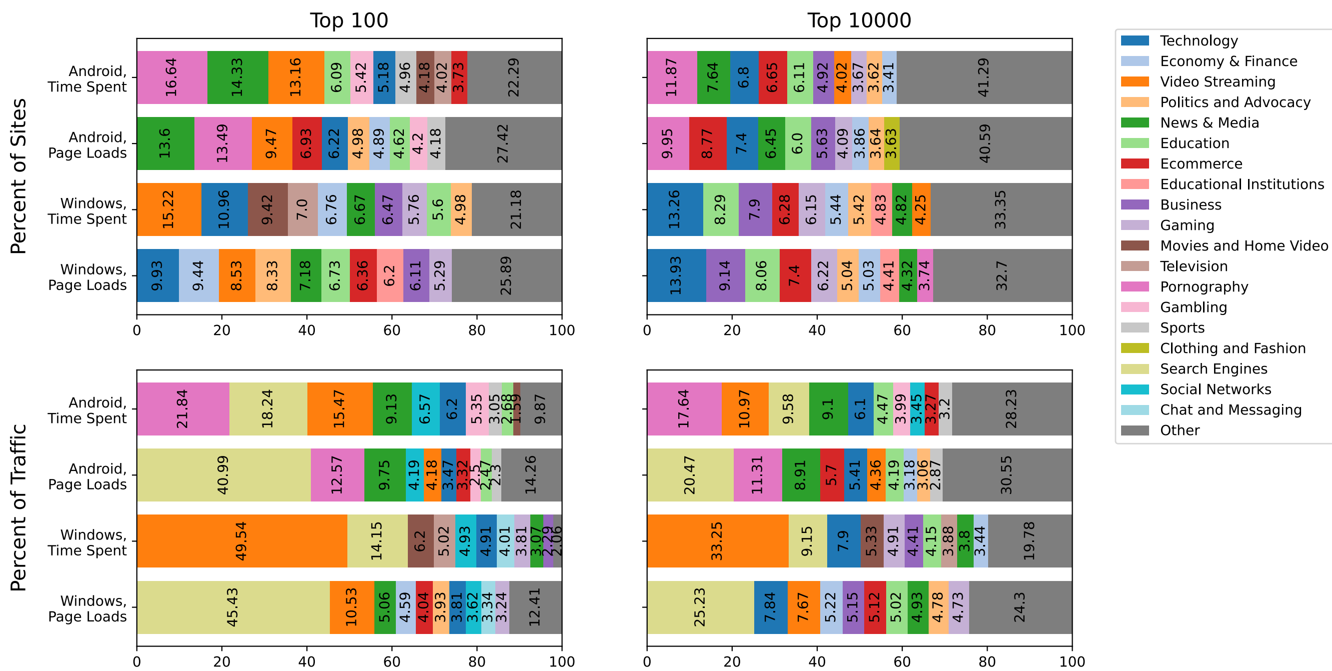

Figure 2: Types of Websites Receiving Most Traffic—We show the types of websites that users visit most based on both the breakdown of the top sites by rank and with user behavior modeled by our distribution of traffic towards top sites.

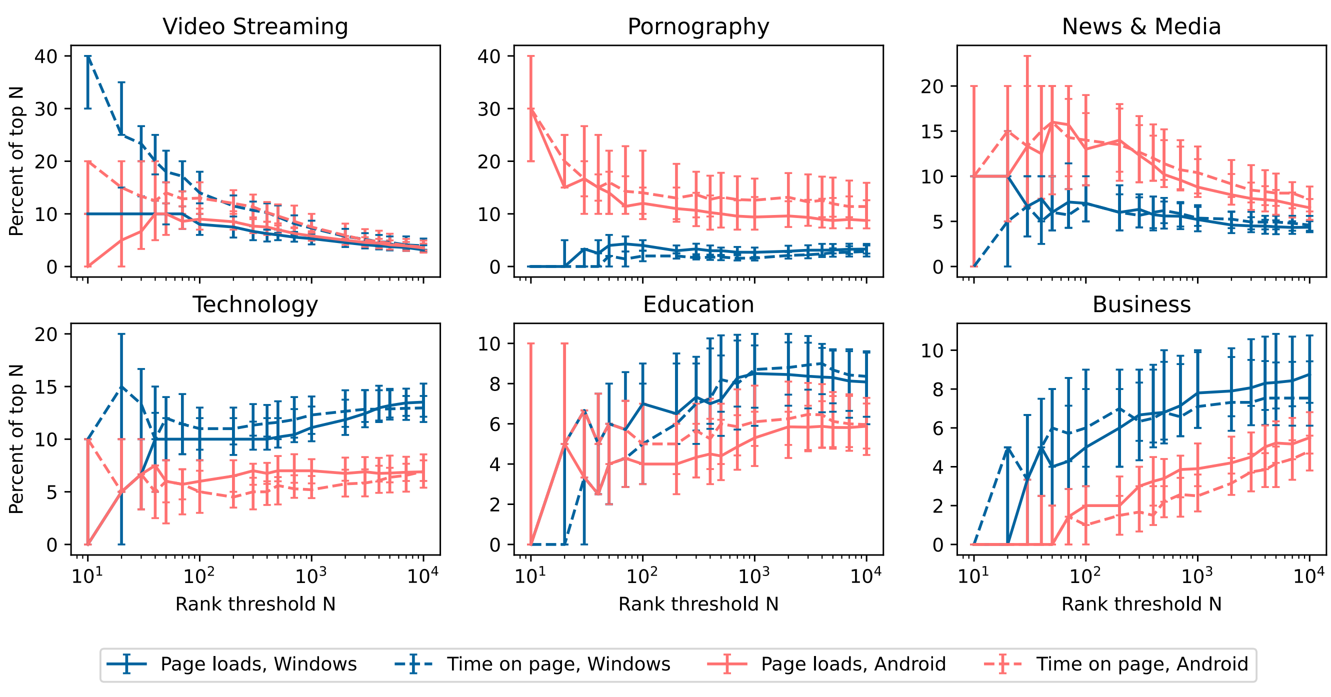

Figure 3: Category Prevalence By Rank— Category distribution varies by rank; some categories are disproportionately represented among top sites or long-tail sites. For most categories, though, there is also significant variation between countries that obscures global trends. Appendix Figure 14 shows these results split by popularity metric due to the overlapping lines.

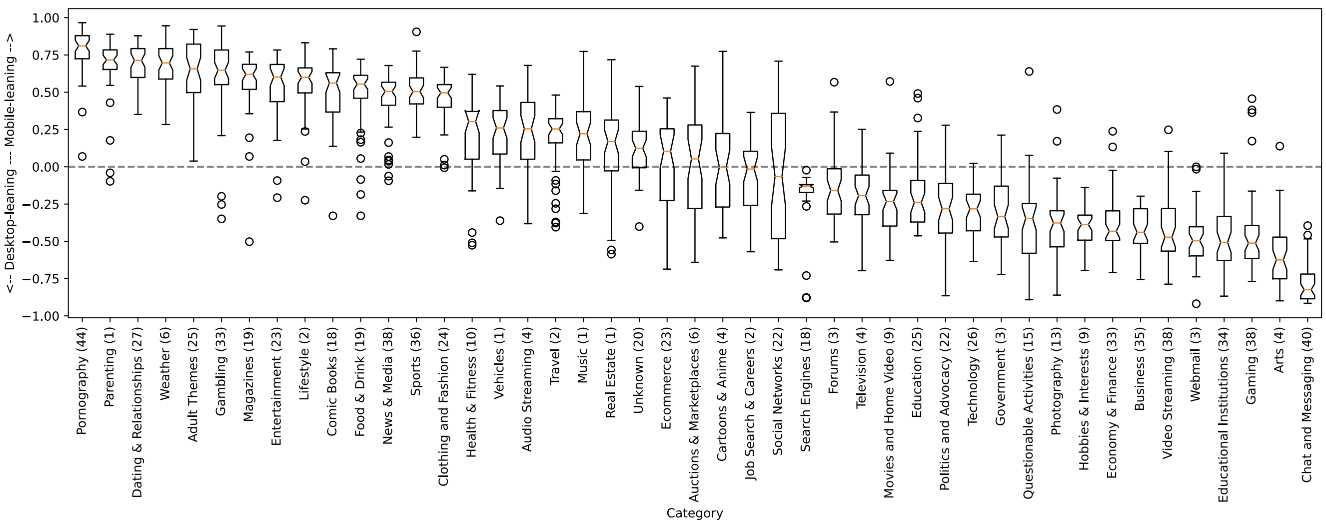

Figure 4: Types of Websites Disproportionately Visited on Desktop and Mobile—Comparing traffic volume for each category on different platforms as described in Section 4.3, we see that desktop and mobile users visit different types of websites. Disproportionately visited sites on mobile devices include Pornography, Dating & Relationships, and Gambling; desktop-focused categories include Gaming, Educational Institutions, and Economy & Finance. The number of statistically significant per-country results is listed with each category.

correction. Our final difference metric is then normalized to take a score from -1 to 1, based on the formula:

$$\frac{A - W}{\max(A, W)}$$

where A and W are the weighted traffic volume for Android and Windows, respectively. This formula expresses the difference in weighted traffic volume as a percentage of the larger value, with the sign representing which platform (Android or Windows) is more prevalent. Figure 4 shows the results for page loads for every category that demonstrated a statistically significant difference. Our results roughly hold for time on page as well (Appendix Figure 15).

As can be seen in Figure 4, the most disproportionately mobile-accessed categories include Pornography, Dating & Relationships, Gambling, Magazines, and other categories related to lifestyle and entertainment. By contrast, the most disproportionately desktop-accessed categories include Educational Institutions, Webmail, Gaming, Economy & Finance, and Business—primarily work- and school-related activities, along with some general knowledge content. These trends are consistent across the majority of countries. These results are consistent with our finding in Section 4.2.2 about entertainment being a primary use case on mobile devices, and additionally suggest that users disproportionately use mobile devices for lifestyle browsing. We again note that the browser's client telemetry does not capture mobile app traffic. As such, some categories like Gaming or Chat & Messaging may appear artificially desktop-oriented from a browser perspective vs. from a user perspective.

Summary. Browsing traffic to many categories occurs disproportionately on either mobile or desktop platforms. Mobile devices see the bulk of traffic in lifestyle, entertainment, and leisure categories, while desktop devices see more of work and school traffic. These different and complementary browsing habits highlight the benefit of researchers splitting their analyses by platform where these browsing category differences may confound results.

4.4 Time on Page vs. Page Loads

Page loads and time on page provide two differing perspectives on website popularity, and the lists of ranked sites by each vary non-negligibly. When comparing countries' top 10K lists across the two metrics, the median intersection is 65% of sites for desktop and 74% for mobile. Within the intersection, the median Spearman's correlation coefficient is 0.65 for desktop and 0.69 for mobile, representing only a modest rank order correlation. Correlation values remain in the same range within website categories, with 57–72% intersection and 0.5–0.8 Spearman correlation for desktop, and 67–82% intersection and 0.6–0.85 Spearman correlation for mobile.

To understand the impact of these differences, we investigate the sites that show the greatest difference in the two metrics. To do so, we estimate the percent of page loads vs. time that users spend on each site (again using the traffic distribution data from Section 4.1) and take the ratio of these two values. A high value for this ratio means the site captures more of users' page loads than of their time. We consider the highest and lowest 20% of these ratios to be the most page-loads-leaning or time-on-page-leaning sites, respectively. Finally, we count how many sites of each category appear among page-loads-leaning, time-on-page-leaning, and all other sites.

Categories that disproportionately appear among page-loads-leaning sites on desktop are E-commerce, Educational Institutions, and Economy & Finance; for time-on-page-leaning sites, overrepresented categories are Video Streaming, Movies & Home Video, and News & Media (Figure 5). This makes intuitive sense: users streaming a video are likely to load the page once and stay there to view the video. These results are almost all consistent on mobile

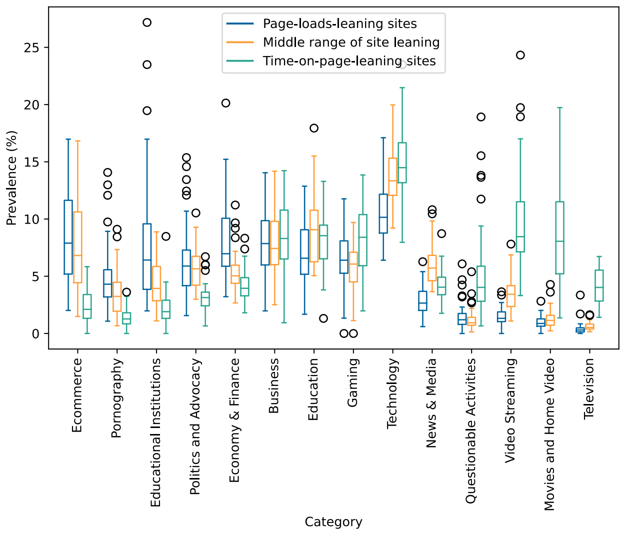

Figure 5: Category Distribution of Page-Load-Leaning, Time-on-Page-Leaning, and all Other Sites—Each boxplot aggregates results across 45 countries. Only categories with a median of at least 3% prevalence in any of the three sets of sites are shown.

(Appendix Figure 16). The exception is Pornography, which is page-loads-leaning on desktop and time-on-page-leaning on mobile.

Summary. Types of sites vary considerably in how users interact with them, ranging from brisk page navigation on e-commerce sites to long page interaction times for video streaming. Neither page loads nor time on page is a one-size-fits-all metric for measuring the web, as they yield very different results.

4.5 Temporal Stability

We next investigate the temporal stability of website popularity from September 2021 to February 2022. While prior work [28] has shown that weekday and weekend traffic differs, our dataset is aggregated monthly. This prevents us from analyzing day-to-day changes but enables us to analyze longer-term trends.

Similarity Between Adjacent Months. First, we examine the consistency of ranked lists across months of our study. We measure consistency in two complementary ways: (1) percent intersection between the set of websites in each list, and (2) Spearman's rho, a rank-order correlation coefficient that operates on the intersection between lists. For Spearman's rho, we follow Cohen's guidelines for interpreting correlation coefficients: <0.10 is negligible, 0.10–0.30 is small, 0.30–0.50 is moderate, and >0.50 is strong [8]. For each rank bucket (e.g., top 10K), for each popularity list (combination of platform, popularity metric, and country), we computed percent intersection and Spearman's rho between pairs of adjacent months as well as between September 2021 and each of the five subsequent months. All numbers presented here are based on the median and 25–75% quantiles among the 45 countries we consider.

List correlations are strong between adjacent months, especially among the highest ranks. Except for December 2021, adjacent months exhibit about 85–95% intersection for the top 20 websites, 82–90% for the top 100, and 80–90% for the top 10K. Spearman coefficients are also very strong, ranging from about 0.90–0.99 for the top 20, to 0.89–0.97 for the top 100, to 0.85–0.95 for the top 10K. December is the most different from adjacent months, with 35–85% intersection and 0.82–0.93 Spearman coefficient for the top 10K; the lowest correlations occur for time spent on Windows. December's uniqueness tends to be somewhat more pronounced for the time on page metric compared to page loads. By contrast, January and February are the adjacent months in our dataset that are the most similar to each other, consistently occupying the high end of the above similarity score ranges.

Stability of Category Distributions. Next, we examine how stable the distribution of site *categories* is over time. For each rank bucket N, we count the occurrences of each category and express them as a percent of N; we then plot the median category percent over all countries. We find that the distribution of categories is mostly stable. The most noticeable changes occur in December, when, e.g., Education drops from 8.4% to 6.8% of sites and Ecommerce rises from 5.0% to 6.1% for desktop top 10K time on page.

Website popularity is relatively stable, and there is more variation between countries than between months in our dataset. Based on these results, we rely on February data for the remainder of this paper. However, based on December's uniqueness in browsing patterns (likely due to holidays, increased e-commerce, and decreased work/school traffic), we caution researchers against making generalized claims about analyses conducted on only December data.

4.6 Summary: Global Browsing Behavior

Users spend a significant portion of time on a handful of globally popular websites (e.g., six websites account for 25% of Windows page loads). The individual top sites between countries vary, but main use cases are relatively consistent globally. All 45 countries in our study rank a search engine and video sharing platform in their top 10 sites. Most also include social networks and adult content. While web use cases are consistent, we observe differences between desktop and mobile browsing behavior and pages frequently loaded vs. pages users spend time on. For example, people tend to use desktops disproportionately for gaming, work/education, and general knowledge-related activities, whereas mobile is dominated by entertainment and adult content. These results inform various facets of what it means to be representative of web browsing globally.

5 GEOGRAPHICAL DIFFERENCES

In this section, we show that despite broad global browsing patterns and consistent popular sites at the highest ranks, there are significant geographic differences in web usage. We first showcase that while some sites are globally popular, the popularity of many others is localized to specific countries. Then, we further explore web usage similarities and differences across countries.

5.1 Determining Global vs. National Popularity

While some of the top sites are consistently popular globally, many have dramatically different rankings across countries, which we aim to quantify. The main challenge is that many sites are endemic (i.e., local) to only one country or region. For example, of the 24K sites that rank in the top 1K for any one country, 13K (53.9%) do not

Shape	Description	Examples
Shallow slope	Similar rank presence across almost all countries	google.com, wikihow.com
Concave	Consistently high popularity in many countries, lower popularity afterwards	whatsapp.com, crunchyroll.com
Moderate slope	Steady decrease in popularity	amazon.com, tupperware.com
Convex	High popularity in a small number of countries, lower popularity otherwise	singaporeair.com
Steep slope	Rapid, steady decrease in popularity	mercadolibre.com, whitepages.com
Inflection points	Irregular distribution: highly popular in some countries, moderately in other countries, low popularity afterwards	hbomax.com

Table 1: Types of Website Popularity Curves—Website popularity curves fall into one of six distinct shapes, which drive their global vs. national popularity.

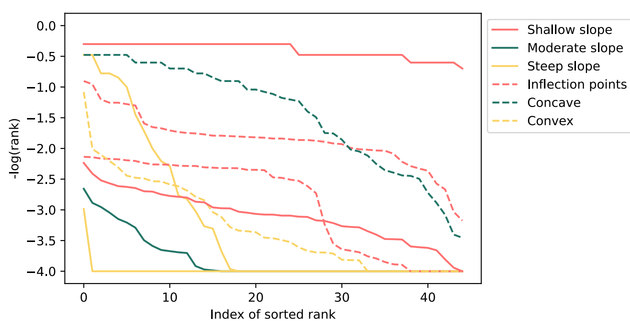

Figure 6: Shapes of Website Popularity Curves—Each curve visualizes the set of ranks that a site achieves in 45 countries. All curves belong to anonymized sites in the dataset.

appear in the top 10K for any other country; 17K sites (73.2%) appear in the top 10K for at most 3 countries. Given this geographic diffusion, we cannot simply count number of countries in which a site is ranked within the top 10K since this does not account for the rank it achieves in those countries; summary statistics (e.g., mean and standard deviation) are also unreliable due to the typically small number of countries in which a website is popular.

To address this challenge, we define an *endemicity score* for each website that satisfies the following properties:

(1) We wish to measure *evenness* of rank, not rank itself. A site with rank 10 in all countries and a site with rank 1,000 in all countries should have the same score.
(2) We aim to quantify unusual popularity in a country, not unusual unpopularity. A site with rank 100 in one country and 1K in all others should be considered much more endemic than one with rank 1K in all countries except 10K in one.
(3) The score should be more sensitive at the top of rankings. Given how unevenly site traffic is distributed (Section 4), a difference between rank 1 and rank 5 should be scored more significantly than a difference between rank 1001 and 1005.
(4) The score should accommodate sites present in few rankings. Even a site that only appears in one country's rank list should still be assigned a usable score.

We derive a score with these properties in two steps:

Step 1: Building Website Popularity Curves. For each site (recall from Section 3.1 that we drop the eTLD for cross-country domain comparisons), we generate a sorted list of its per-country ranks using the top 10K rank list for every country. This creates a vector $[r_1, r_2, \ldots, r_{45}]$ per site, sorted by the smallest rank (i.e., most popular) to the largest rank (i.e., least popular). As noted earlier, not every site appears in every country's top 10K list. To address this, we denote the rank for that country as 10,001 (i.e., the lowest possible rank value + 1); this will ensure Property 4. Then, for each value i from 1–45, we plot the inverse log of the rank of the site (i.e., $-\log_{10}(r_i)$) across the rank vector to generate what we call a *website popularity curve*. We choose the inverse log as it presents a normalized scale (from -4 to 0) to compare sites, it amplifies differences among the highest ranks according to Property 3, and it supports clearly visualizing a site's popularity across countries.

Figure 6 shows examples of the 6 shapes of resultant website popularity curves, and Table 1 describes each shape in detail. At a high level, each distinct shape describes a unique popularity pattern per site. For example, a website popularity curve with a shallow slope describes a site with similar rank presence across all countries (e.g., google.com, twitter.com, facebook.com), while a curve with multiple inflection points denotes a site consistently popular in a few countries but not on others (e.g., hbomax.com).

Step 2: Computing Endemicity Scores. From each popularity curve, we compute a site's *endemicity score*, which is a single metric that captures a website's nationalization across the 45 countries we study. Our objective is to distill each popularity curve to a single, comparable metric, which we can then use to compare domains. We define the endemicity score $\mathcal{E}_w$ of a website w as the area between the theoretically flattest possible curve starting at the highest rank it achieves in any country (i.e., if all countries listed w with the same rank r_1) and the actual curve. This formulation intuitively captures the "difference" from maximum global consistency (Property 1), and it is particularly sensitive to a website popularity curve that drops rapidly away from its baseline (Property 2).

Our *endemicity score* ranges from 0–180 based on the bounds described earlier. A smaller score corresponds to a website that is more globally popular, whereas a larger score corresponds to a website that is more endemic to one area. We compute endemicity scores for sites that appear in the top 1K of at least one country (23,785 websites) and consider the remaining sites to be in the long tail of all countries.

Figure 7 shows the distribution of $\mathcal{E}_w$ for all websites w we consider, based on the highest rank that w achieves. To determine whether a website is globally or nationally popular, we measure the distance between each point in Figure 7 and the upper bound on the endemicity score, and then perform outlier detection on this set

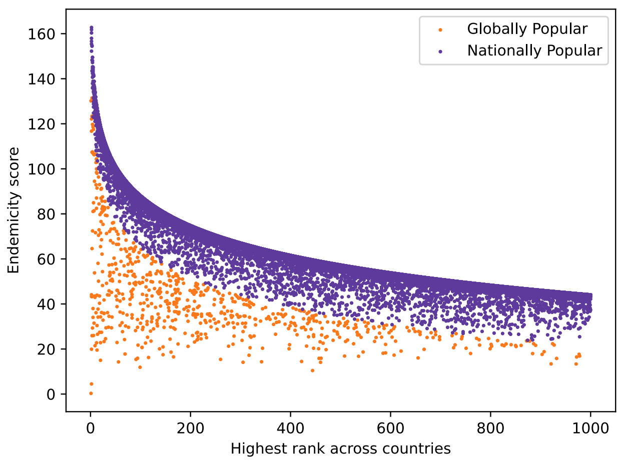

Figure 7: Globally Popular vs. Nationally Popular Websites— We show visually the globally popular websites (in orange, bottom region) vs. nationally popular websites (in purple, top region). Globally popular websites have much lower endemicity scores due to their distance from the theoretical maximum endemicity.

Platform	Popularity Metric	Top 1K in Any Country		Long Tail
		Globally Pop.	Nationally Pop.	
Windows	page loads	669 (2.8%)	23,116 (97.2%)	337,436
Windows	time on page	594 (2.7%)	21,074 (97.3%)	127,549
Android	page loads	457 (1.8%)	24,283 (98.2%)	433,778
Android	time on page	323 (1.4%)	22,739 (98.6%)	267,162

Table 2: Rarity of Globally Popular Websites—The vast majority of websites we label are nationally popular (98%) compared to just 2% of websites that are globally popular. Most websites fall within the long-tail, which we do not include in our analysis.

of distances (based on Smirnov-Grubbs [16]). The core intuition is that outlier detection will identify deviation from the upper bound on how endemic a site can be, which should correspond to more globally popular websites. In Figure 7, we color websites in orange if they are globally popular and purple if they are nationally popular.

5.2 Global vs. National Popularity

The vast majority of websites are nationally popular (average of 98% across countries and metrics) compared to an average of 2% that are globally popular (Table 2). Interestingly, the categories of websites that are globally versus nationally popular differ significantly (Figure 8). For Windows, globally popular websites relate to technology, pornography, gaming, hobbies, messaging, and photography. By contrast, content related to educational institutions, politics, and economy & finance tends to be more regional. Our qualitative observations also hold for Android devices, albeit with different proportions. For example, adult content represents 20–25% of globally popular websites on Android compared to just 3–6% on Windows. Ultimately, our results point to notable differences in global and national popularity based on website categories.

Finally, we turn to examining *how popular* the globally vs. nationally popular sites are. For each of several rank buckets, we compute

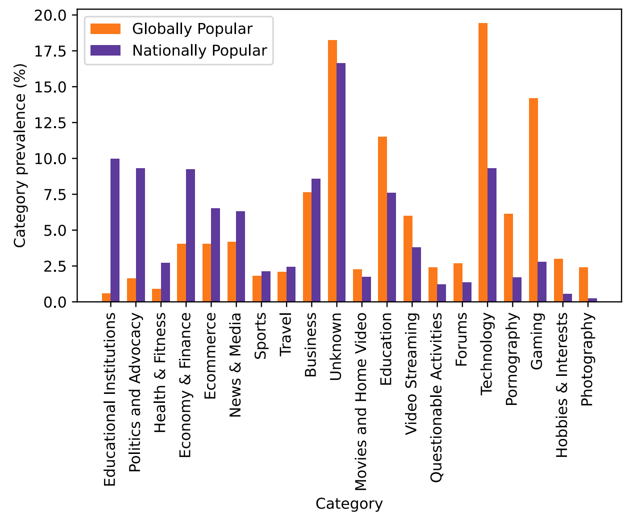

Figure 8: Types of Globally vs. Nationally Popular Websites— Some categories skew towards global popularity—such as technology, adult, and gaming websites. In contrast, educational websites, political websites, and economy & finance websites are much more likely to be locally popular, highlighting a divide between browsing behaviors based on category and country.

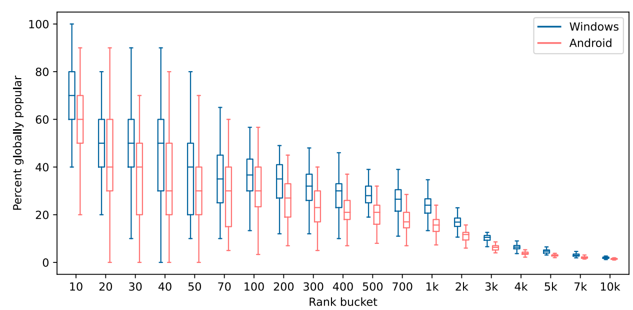

Figure 9: Globally Popular Websites in Each Rank Bucket— Globally popular sites are most prevalent at high ranks, but also many high-ranking sites are highly localized. Each boxplot covers 45 countries and expresses the proportion of sites that are globally popular per rank bucket.

the percentage of sites in that rank bucket that are globally popular. As shown in Figure 9, globally popular sites predominate among the top 10 (median of 6–7 / 10 across 45 countries). However, nationally popular sites also have a strong presence among extremely popular sites. Already in the top 10, a median of 3–4 sites are nationally popular, and starting at top 20, there are at least as many (if not more) nationally popular sites compared to globally popular sites; among sites ranked 101–200, a median of 65–73% of sites are nationally popular. This observation holds for most countries. We observe similar findings when comparing globally and nationally popular sites ranked by time spent (Appendix Figure 17).

5.3 Website-Based Country Comparisons

While some websites are globally popular, the majority of each country's most popular sites are regional. In this section, we quantify the similarities and differences across countries' top 10K sites. In addition, we identify clusters of similar countries based on geographic proximity and language.

5.3.1 Clusters of Countries. We analyze pairs of per-country top 10K lists by using a variation on Rank-Biased Overlap (RBO) [2, 35]. RBO is a list comparison method where agreement at the top of the list is weighted more heavily than agreement lower down the list. Instead of using a geometric distribution for weighting, we leverage our web traffic distribution from Section 4.1. Figure 10 presents a pairwise comparison of countries for Windows based on page loads; we include heatmaps for Android and time-spent in Appendix G.

There are several clusters of similar countries that are visually apparent in Figure 10: for example, Algeria, Egypt, Morocco, and Tunisia (first columns). To more rigorously identify sets of more similar browsing profiles, we cluster countries with similar web usage using the affinity propagation algorithm [13, 29] on the pairwise weighted RBO values. Affinity propagation is a clustering algorithm that does not require specifying the expected number of clusters and accommodates an arbitrary similarity score matrix with clusters of potentially varying density (DBSCAN struggles with varying-density clusters). Affinity propagation defines clusters by sending messages between pairs of data points until convergence. We observe 11 clusters of countries (Figure 11). To measure the strength of clusters, we use Silhouette Coefficient [17, 30] (*SC*), which, given cluster labels and pairwise distances between data points, quantifies how dense and well separated clusters are on a $[-1, 1]$ scale. Although present, clusters are only weakly bound together, with an average *SC* of only 0.11.

Where present, clusters of web browsing behavior follow patterns of shared geography and shared language, in line with prior work [25]. For example, countries from Central and South America—covering a large geographical region but almost all primarily Spanish-speaking—are clustered together (average *SC* of 0.14), and Brazil (primarily Portuguese-speaking) is easily the least similar to the other countries in the cluster ($SC = -0.03$).

Geographically proximate countries in Europe form two clusters: France, Belgium, and the Netherlands group together (average $SC = 0.18$), while the remaining countries cluster extremely weakly (-0.02). Also, the English-speaking countries of Australia, Canada, New Zealand, the UK, and the US are together in one cluster (0.08) despite their vast geographic differences. The tightest clusters are North Africa (0.31) and Taiwan/Hong Kong (0.21), while the sub-Saharan Africa/India (-0.01) and southeast Asia (0.04) clusters are among the loosest. We also observe two outliers, Japan and South Korea, which have distinct browsing patterns separating them from all other country clusters. Further details are in Appendix Figure 21.

5.3.2 Countries and Categories. While popular *categories* of websites are consistent across countries, individual sites can be vastly different. To measure this, we manually review the 10 most visited websites for each country. We find that, while Google is globally popular, ranked in the top 10 for every country, 21 countries have a second top-10 search or portal site. 30 countries have chat or messaging domains in their top 10 (total of 4 sites, e.g., Facebook Messenger), and 22 countries have sites related to business and professional tools (5 sites, e.g., Sharepoint). While 3 pornography sites appear in the top 10 for at least 10 countries (`pornhub`, `xnxx`, and `xvideos`), 4 countries (South Korea, Turkey, Vietnam, and Russia) do not have any of those three in their top 10. In fact, these countries make a policy of censoring adult content, though with varying efficacy—e.g., Vietnam has a pornography site (`sex333`) in its top 10. Social networks and forums also vary in country similarity. Of the 8 social networking sites we observe, 3 sites (Facebook, Instagram, and Twitter) make the top 10 in at least 10 countries, even considering that native mobile app traffic is not included in Chrome telemetry. Forums are more country-specific, with 7 forums spanning 5 countries: e.g., South Korea has 4 forums in its top 10 (`arca.live`, `dcinside.com`, `fmkorea.com`, and `inven.co.kr`).

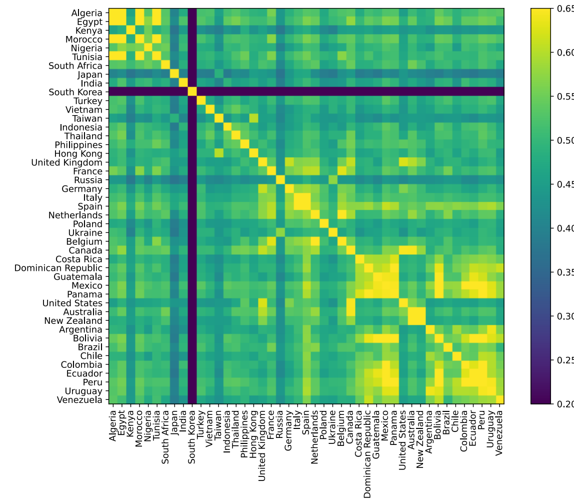

Figure 10: Traffic-Weighted RBO Values Between Countries— Countries show varying amounts of rank-biased overlap, with South Korea exhibiting the most distinct traffic patterns. The color scale corresponds to RBO value.

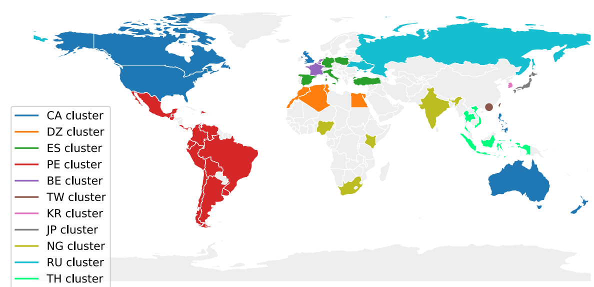

Figure 11: Countries with Similar Top Sites—Countries with similar browsing patterns tend to have geographic or language similarities. Clusters are based on traffic weighted RBO (Figure 10).

Some website categories contribute both to country similarities and differences. For instance, many e-commerce domains are popular in multiple countries: 15 unique companies span 32 countries (e.g., Amazon and AliExpress). However, these companies almost exclusively have one distinct eTLD for every country in which they operate (e.g., `shoppee` in southeast Asia has `.vn`, `.tw`, `.co.id`, and `.co.th` domains). In contrast, classified ads sites are national: 15 of 17 domains are top-10 for only one country (e.g., `2dehands` in Belgium, `ouedkniss` in Algeria, or `yapo` in Chile). Another example is video streaming. Of the 27 streaming sites, a few are household names like Netflix and Prime Video, but 19 are only top-10 in one country and typically offer content in a specific language. By comparison, all 11 television sites appear in only one country; these are primarily broadcasters who reach an area within their country of business, e.g. TVNZ in New Zealand, TV Globo in Brazil.

Other website categories are highly specific to a particular country. For example, government sites (43 sites in 26 countries) are only ever top-10 in one country; the same holds for news (33 sites in 20 countries, e.g., BBC in the UK, VnExpress in Vietnam) and banks (25 sites in 17 countries).

Interestingly, some categories show regional patterns rather than national or global. For instance, there are 15 university sites in the dataset, and 9 of the 10 countries they appear in are located in the global south (including 8 in South and Central America, e.g., Universidad de la República is Uruguay's largest university), while only one university is from the global north (`kuleuven` in Belgium). Similarly, for the 25 sites dedicated to gambling, sports betting, and lottery, 11 of 14 countries are in the global south; and for the 13 sports-related sites (e.g., `cricbuzz` in India), 7 of 9 countries are also in the global south. E-books and free fiction writing (6 countries) total 9 sites, 5 of which are in Taiwan (`ixdzs.com`, `uukanshu.com`, `czbooks.net`, `sto.cx`, and `twfanti.com`).

South Korea. South Korea is notably different from the other countries we consider, in part because South Korea has country-localized alternatives to popular services. For instance, four South Korean forums appear in the top 10 sites for only South Korea (two resemble Reddit, one is centered on sports and gaming, and one is a cross between 4chan and Pinterest). While Roblox does not reach top 10, Nexon (a Korean game publisher) does and is unique to South Korea's top 10. For video streaming, in addition to Netflix, South Korea also has `wavve` and `noonoo.tv` (the latter providing free content), and `afreecatv` in addition to `youtube`. South Korea also has an equivalent of Wikipedia (`namu.wiki`) and two unique search/portal sites (Naver and Daum). We also note that South Korea is the only country where Google does not rank number 1 by page loads on Windows (Section 4.1). Together, these sites differentiate South Korean from other countries. Note that, by design, our RBO metric is heavily weighted toward top sites (based on the traffic distribution from Figure 1), so Naver likely has a significant impact on the difference we observe in Figure 10.

5.3.3 Impact of Top List Size, Platform, and Popularity Metric. Section 5.3.1 provides a traffic-weighted view of country similarities over the top 10K. We now wish to examine how the strength of cross-country similarity differs at the head vs. the long tail of the web. To do so, we take an unweighted percent intersection between rank lists for each pair of countries. This results in a set of $\binom{45}{2} = 990$ values. Rather than expressing them in a heatmap, we succinctly plot by sorting the values in descending order and plotting their cumulative sum. We repeat over several rank buckets (Figure 12).

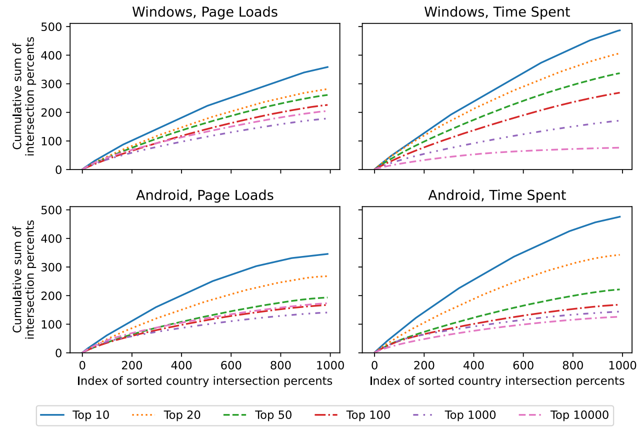

Figure 12: Cross-Country Similarity per Rank, Platform, and Popularity Metric—Cumulative sum of percent intersection for each pair of countries, for different rank thresholds; perfect intersection would correspond to the line $y = x$. Countries' popular sites show greater overlap among the topmost ranks than in the long tail. Lists by time spent show greater overlap than for page loads.

Countries' popular sites are more similar among the topmost ranks than among the long tail. This is likely due to globally popular websites (e.g., google.com, netflix.com), and dovetails with Figure 9. Although typically there is smaller intersection as rank-bucket size increases, this effect seems to bottom out or even reverse as the rank bucket approaches 10K, suggesting a saturation point in country differentiation from locally popular websites and a moderating effect by noise in the global long tail.

5.4 Summary: Geographical Differences

Despite global similarities in browsing behavior, there are also significant web browsing differences depending on user locale. For example, a large portion of the most popular websites differ across countries: of the 24K websites that are in the top 1K of at least one country, 54% do not appear in the top 10K of any other country. While some websites are globally popular (related to, e.g., technology, adult content, or gaming), nationally popular websites also rank in the top 10 of their country (3–4 sites); these concern, e.g., education, politics, or economy & finance content. We also uncover clusters of countries with similar browsing patterns, driven often by country geographic proximity and/or shared language. Finally, we manually review the top 10 websites for all countries and explain web browsing similarities and differences across countries: while website categories tend to stay similar across countries, websites related to, e.g., government, news, or banks are locally popular.

6 DISCUSSION

Our results illustrate that browsing behavior is more complex than often assumed. Here, we summarize our results and discuss both future research avenues and lessons for the research community.

Browsing is heavily concentrated on top sites. Top websites receive several orders of magnitude more traffic than lower ranked sites. Indeed, the single top ranked website accounts for 17% of all Windows page loads and 6 websites account for 25% of all page loads globally. Top million lists capture well the vast majority of user traffic (≈95%), but studies that focus on the top million sites evenly are skewed heavily towards the long tail of the web. We encourage future studies to consider how the distribution of page loads and time on page can be used to develop more representative metrics for characterizing web behavior.

Globally ranked lists underrepresent regional sites. While use cases for the web are fairly consistent between countries, not all are fairly captured by a global rank order list. Global lists do not capture sites that are important to only a small number of regions, which biases them against categories of websites that tend to be localized. Educational institutions, politics, health, and finance sites all tend to have more distinctly national user bases, compared to adult content, technology, and gaming sites, which have global user bases. This may bias studies away from characterizing the typical browsing experience of a user. For example, a study on vulnerabilities in top sites may benefit from an analysis of per-country prevalence, as this would better capture the security posture of hospitals and banks in addition to describing which regions are most vulnerable. Researchers may want to consider collecting a representative set of websites from different categories, or comparing results from a global set and a handful of countries.

Most sites' traffic is regional. Sites that are popular in only one country or region outnumber sites with a broad global reach. For instance, among sites ranking in the top 1K for at least one country, over half do not appear in the top 10K for any other country. At the extreme, South Korea has nationally endemic platforms driving a large share of the traffic among its top sites. Geographic proximity and shared language only partially explain country differences. This serves as a reminder that a global list of sites is not representative of any individual user and likely biases results toward populous, industrialized countries. As one concrete consequence, system builders optimizing based on global data (e.g., fitting statistical models) should not assume that their systems are also optimized for those on the margins, and should design and validate their systems to be more geographically equitable.

Popularity metric matters. There are multiple ways to calculate website popularity, and results vary based on the metric. For example, the top 10K sites as defined by page views and time on page overlap by about 65–75%. While other work analyzes levels of agreement between more popularity metrics [27], we see some of the reasons that metrics differ, particularly that search engines and video platforms both skew metrics. Measuring engagement with the web—either broadly or for a particular site—is complex, and a thoughtful choice of metrics is required for meaningful insight.

Web usage differs by platform. Desktop devices are used disproportionately for work and school, while mobile sees a disproportionate share of entertainment and lifestyle traffic. We thus advise researchers to split analyses by platform where feasible if these category differences may introduce a confounding factor. There may also be an opportunity for web designers to tailor features by platform in response to users' per-platform motivations. We caution, though, against oversimplifying: these global trends may break down in regions where mobile devices are users' primary modes of web access.

Not every month is representative. Web browsing is similar across months but not uniformly so. December is the most unusual month in our September–February dataset, exhibiting noticeable changes in education and e-commerce traffic. Our measurement period does not cover summer months in the northern hemisphere, which may also exhibit different patterns from the months we study. We advise researchers against using only December data for general-purpose analyses, as well as to be aware of other potentially unusual periods if studying a specific geography or user demographic. Longitudinal studies provide the best assurance of temporal representativeness, but for many studies, a snapshot of a "typical" month likely generalizes to adjacent months.

Lessons for geo-aware methodology. Drawing from the above observations, we offer several recommendations for those studying the web. First, we encourage breaking down analyses by country where possible. In particular, an observation made using globally aggregated data should be validated against geographic splits to show that it is not only a property of a populous, industrialized country or of sites with globally distributed traffic. We encourage researchers to consider data sources with per-country breakdowns (e.g., the public CrUX data [4]). Second, for studies with a more inherently localized measurement vantage point (e.g., a network tap), we remind researchers to be cautious when generalizing results, and we encourage the community to welcome replication studies from other geographical regions. Third, we call out the need for further exploration of how to fairly represent both multiple platforms and regions in research studies. For instance, one could hypothesize that taking the global top 1K together with the top 1K from each country may lead to more geographically generalizable conclusions than taking simply the global top 10K.

7 CONCLUSION

In this paper, we conduct an in-depth study of user web browsing behavior based on Chrome browser telemetry data. Among our results, we find that global browsing behavior is highly concentrated toward top sites, identify entertainment as the most dominant use case for the web, and highlight differences in browsing patterns for desktop and mobile platforms. In addition to these global trends, we also show that there are tremendous differences in browsing patterns depending on where users are located: the majority of the most popular websites in a given country are endemic to that country. We hope that our work can serve as a foundation for researchers and practitioners looking to refine hypotheses, build better systems, and conduct more thoughtful analyses of the web.

ACKNOWLEDGMENTS

We thank Deepak Kumar for his help framing the paper, Hans Hanley, Liz Izhikevich, and Parker Ruth for helpful discussions about statistics, and members of the Stanford Empirical Security Research Group for valuable conversations. We also thank Adriana Porter Felt, our shepherd Philipp Richter, and the anonymous reviewers for their feedback. This work was supported in part by an

NSF Graduate Research Fellowship DGE-1656518 and a gift from Google, Inc.

REFERENCES

[1] Alexa top million. https://www.alexa.com/. Accessed on 09-12-2022.
[2] C. Chen. RBO implementation – similarity measures between two different ranked lists. https://github.com/changyaochen/rbo/blob/master/rbo/rbo.py. Accessed on 05-10-2022.
[3] W. Chen. Internet-Usage Patterns of Immigrants in the Process of Intercultural Adaptation. *Cyberpsychology, Behavior, and Social Networking*, 2010.
[4] Chrome User Experience Report. https://developers.google.com/web/tools/chrome-user-experience-report. Accessed on 09-12-2022.
[5] Chromium Bugs. Issue 1321253: PageForegroundSession.ForegroundDuration has a high percentage of negative values. https://bugs.chromium.org/p/chromium/issues/detail?id=1321253.
[6] Cisco Umbrella 1 million. https://umbrella.cisco.com/blog/cisco-umbrella-1-million. Accessed on 09-12-2022.
[7] Cloudflare. Cloudflare API v4 documentation – get domain details. https://api.cloudflare.com/#domain-intelligence-get-domain-details. Accessed on 03-28-2022.
[8] J. Cohen. *Statistical power analysis for the behavioral sciences.* 1998.
[9] Comscore latest rankings. https://www.comscore.com/Insights/Rankings.
[10] CrUX methodology. https://developer.chrome.com/docs/crux/methodology/.
[11] S. Duarte Torres, I. Weber, and D. Hiemstra. Analysis of search and browsing behavior of young users on the web. *ACM Transactions on the Web (TWEB)*, 2014.
[12] M. Foundation. Public suffix list. https://publicsuffix.org/list/. Accessed on 04-03-2022.
[13] B. J. Frey and D. Dueck. Clustering by Passing Messages Between Data Points. *Science*, 2007.
[14] X. Gao, Y. Yang, H. Fu, J. Lindqvist, and Y. Wang. Private browsing: An inquiry on usability and privacy protection. In *13th Workshop on Privacy in the Electronic Society*, 2014.
[15] S. Goel, J. M. Hofman, and M. I. Sirer. Who Does What on the Web: A Large-scale Study of Browsing Behavior. In *AAAI Conference on Weblogs and Social Media*, 2012.
[16] F. E. Grubbs. Sample Criteria for Testing Outlying Observations. *Annals of Mathematical Statistics*, 1950.
[17] P. J.Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53—65, 1987.
[18] R. Kumar and A. Tomkins. A Characterization of Online Browsing Behavior. *WWW*, 2010.
[19] V. Le Pochat, T. Van Goethem, S. Tajalizadehkhoob, M. Korczyński, and W. Joosen. Tranco: A research-oriented top sites ranking hardened against manipulation. In *Network and Distributed Security Symposium (NDSS)*, 2019.
[20] G. Maier, A. Feldmann, V. Paxson, and M. Allman. On dominant characteristics of residential broadband internet traffic. In *ACM Internet Measurement Conference (IMC'09)*, 2009.
[21] The Majestic million. https://majestic.com/reports/majestic-million. Accessed on 09-12-2022.
[22] A. L. Montgomery and C. Faloutsos. Identifying web browsing trends and patterns. *Computer*, 34, 2001.
[23] A. Morichetta, M. Trevisan, and L. Vassio. Characterizing web pornography consumption from passive measurements. In *Passive and Active Measurement (PAM)*, 2019.
[24] H. Müller, J. L. Gove, J. S. Webb, and A. Cheang. Understanding and Comparing Smartphone and Tablet Use: Insights from a Large-Scale Diary Study. In *Australian Computer-Human Interaction Conference*, 2015.
[25] Y. M. M. Ng and H. Taneja. Mapping User-Centric Internet Geographies: How Similar are Countries in Their Web Use Patterns. *Journal of Communication*, 2019.
[26] R. Pang, M. Allman, M. Bennett, J. Lee, V. Paxson, and B. Tierney. A first look at modern enterprise traffic. In *Proceedings of the 5th ACM SIGCOMM Conference on Internet Measurement (IMC'05)*, 2005.
[27] K. Ruth, D. Kumar, B. Wang, L. Valenta, and Z. Durumeric. Toppling top lists: Evaluating the accuracy of popular website lists. In *ACM Internet Measurement Conference (IMC)*, 2022.
[28] W. Rweyemamu, T. Lauinger, C. Wilson, W. Robertson, and E. Kirda. Clustering and the weekend effect: Recommendations for the use of top domain lists in security research. In *Passive and Active Measurement (PAM)*, 2019.
[29] scikit-learn. sklearn.cluster.affinitypropagation. https://scikit-learn.org/stable/modules/generated/sklearn.cluster.AffinityPropagation.html. Accessed on 04-09-2022.
[30] scikit-learn. sklearn.metrics.silhouette_samples. https://scikit-learn.org/stable/modules/generated/sklearn.metrics.silhouette_samples.html.
[31] Semrush top websites ranking. https://www.semrush.com/website/top/.
[32] K. Sheehan. Of Surfing, Searching, And Newshounds: A Typology Of Internet Users' Online Sessions. *Journal of Advertising Research*, 2002.
[33] SimilarWeb top websites ranking. https://www.similarweb.com/top-websites/.
[34] A. Tikhonov, L. O. Prokhorenkova, A. Chelnokov, I. Bogatyy, and G. Gusev. What can be Found on the Web and How: A Characterization of Web Browsing Patterns. In *ACM Web Science Conference*, 2015.
[35] W. Webber, A. Moffat, and J. Zobel. A Similarity Measure for Indefinite Rankings. *ACM Trans. Inf. Syst.*, 2010.

A LIST OF COUNTRIES

Below is the list of countries we analyze by continent.

Africa. Algeria (DZ), Egypt (EG), Kenya (KE), Morocco (MA), Nigeria (NG), Tunisia (TN), South Africa (ZA)

Asia. Japan (JP), India (IN), South Korea (KR), Turkey (TR), Vietnam (VN), Taiwan (TW), Indonesia (ID), Thailand (TH), Philippines (PH), Hong Kong (HK)

Europe. United Kingdom (GB), France (FR), Russia (RU), Germany (DE), Italy (IT), Spain (ES), Netherlands (NL), Poland (PL), Ukraine (UA), Belgium (BE)

North America. Canada (CA), Costa Rica (CR), Dominican Republic (DO), Guatemala (GT), Mexico (MX), Panama (PA), United States (US)

Oceania. Australia (AU), New Zealand (NZ)

South America. Argentina (AR), Bolivia (BO), Brazil (BR), Chile (CL), Colombia (CO), Ecuador (EC), Peru (PE), Uruguay (UY), Venezuela (VE)

B CATEGORY DATA ACCURACY ANALYSIS

Below we include further details on our accuracy analysis and final category taxonomy.

Figure 13 shows the results of our manual accuracy analysis. We randomly selected 10 websites of each category and labeled them as being definitely correct ("Yes"), somewhat correct ("Maybe"), or definitely incorrect ("No"). We dropped categories that did not have more than 8 / 10 plausibly or definitely correct labels, as well as those with not a single definitely correct label.

Next, we manually merged categories that were small and had similar semantics. This resulted in 22 supercategories and 61 categories, shown in Table 3.

C LONG TAIL OF CATEGORIES AMONG TOP-10 SITES

Table 4 shows the long tail of site categories we encountered during our manual analysis of top-10 sites.

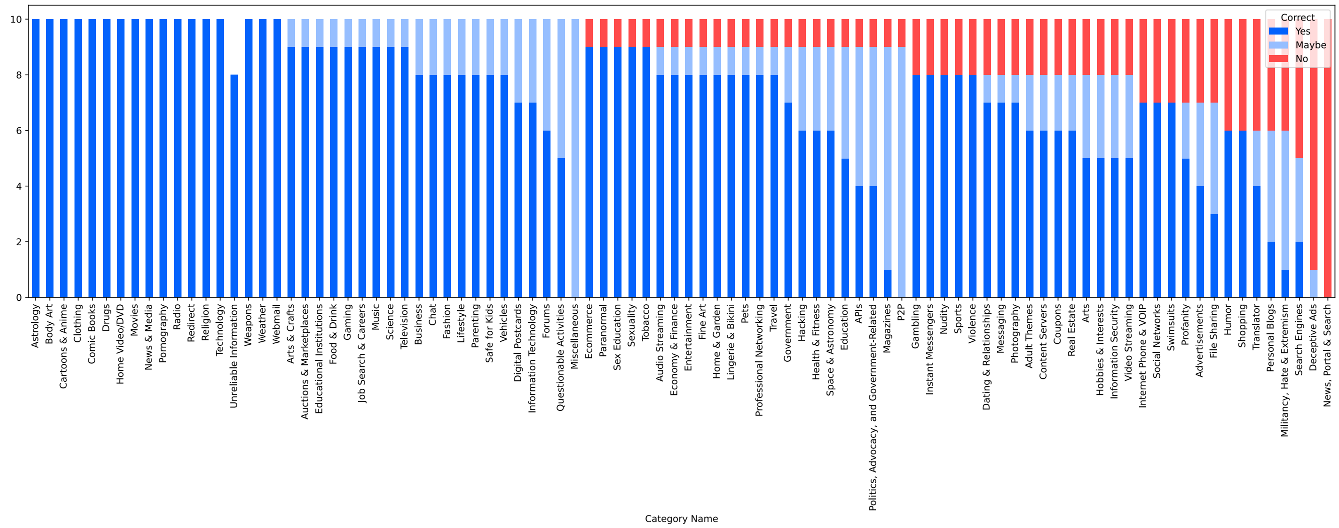

Figure 13: Category API Accuracy Analysis—Manually labeled correctness results for 10 randomly chosen websites in each category.

Supercategory	Categories
Adult Themes	Pornography; Adult Themes
Business & Economy	Business; Economy & Finance
Education	Educational Institutions; Education; Science
Entertainment	News & Media; Audio Streaming; Music; Magazines; Cartoons & Anime; Movies & Home Video; Arts; Entertainment; Gaming; Video Streaming; Television; Comic Books; Paranormal
Gambling	Gambling
Government & Politics	Government & Politics; Politics, Advocacy, and Government-Related
Health	Health & Fitness; Sex Education
Internet Communication	Forums; Webmail; Chat & Messaging
Job Search & Careers	Job Search & Careers
Miscellaneous	Redirect
Questionable Content	Drugs; Questionable Content; Hacking
Real Estate	Real Estate
Religion	Religion
Shopping & Auctions	Ecommerce; Auctions & Marketplaces; Coupons
Society & Lifestyle	Lifestyle; Clothing and Fashion; Food & Drink; Hobbies & Interests; Home & Garden; Pets; Parenting; Photography; Astrology; Dating & Relationships; Arts & Crafts; Sexuality; Tobacco; Body Art; Digital Postcards
Sports	Sports
Technology	Technology
Travel	Travel
Vehicles	Vehicles
Violence	Weapons; Violence
Weather	Weather
Unknown	Unknown

Table 3: Final Category Taxonomy

Site description	Count
Accounting service for small business	2
Artist community	1
Graphic design platform	1
Marketplace for new and used cars	2
Informational site and forum about cars	1
Cryptocurrency site	2
Financial investment site	2
Dating platform	2
Educational technology platform	6
Exam preparation	2
Homework Q&A (academic misconduct)	3
Job search platform	4
Event/movie ticket purchasing	2
Gig economy (microtasking, freelance)	3
Insurance (including health insurance)	3
Digital identity provider	1
ISP/telecom	9
Mobile payment platform	2
Postal service	2
Proxy/anonymizer	1
Real estate	3
Travel booking	1
Videoconferencing	1

Table 4: Other Categories Among Top-10 Sites—Top 10 sites include a long tail of site categories.

D SUPPLEMENTARY FIGURES FOR CATEGORY PREVALENCE PER RANK

Figure 14 shows the data from Figure 3 split out by popularity metric for clarity.

E POPULARITY METRIC DIFFERENCES ON MOBILE

Figure 16 shows the category distribution of sites with different popularity metric leaning on mobile.

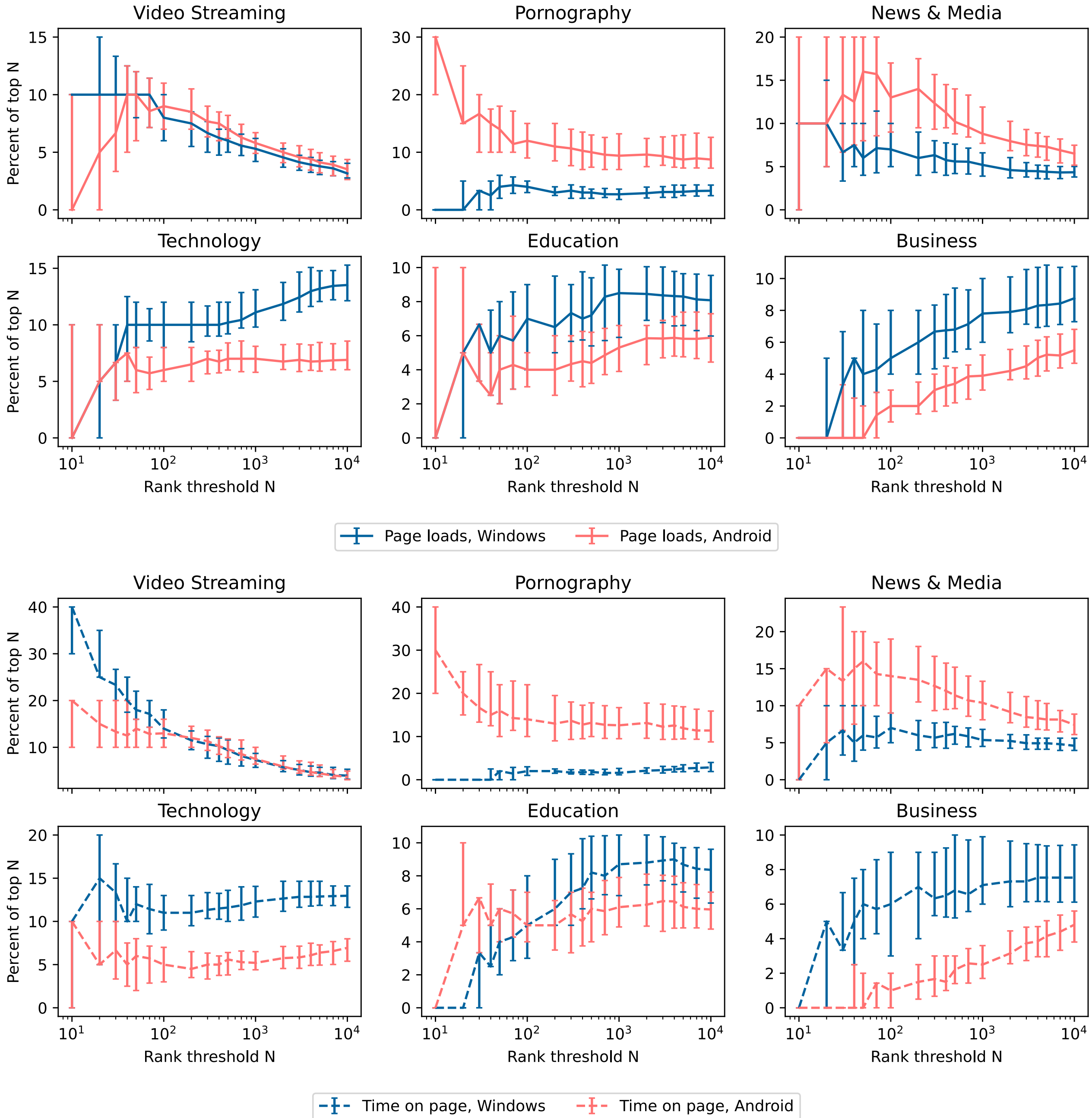

Figure 14: Category Prevalence per Rank, Split by Metric— Data from Figure 3 split by popularity metric to make overlapping lines more visible.

F GLOBALLY POPULAR SITES BY RANK, TIME ON PAGE

Figure 17 shows the distribution of globally popular websites for time-on-page ranked lists.

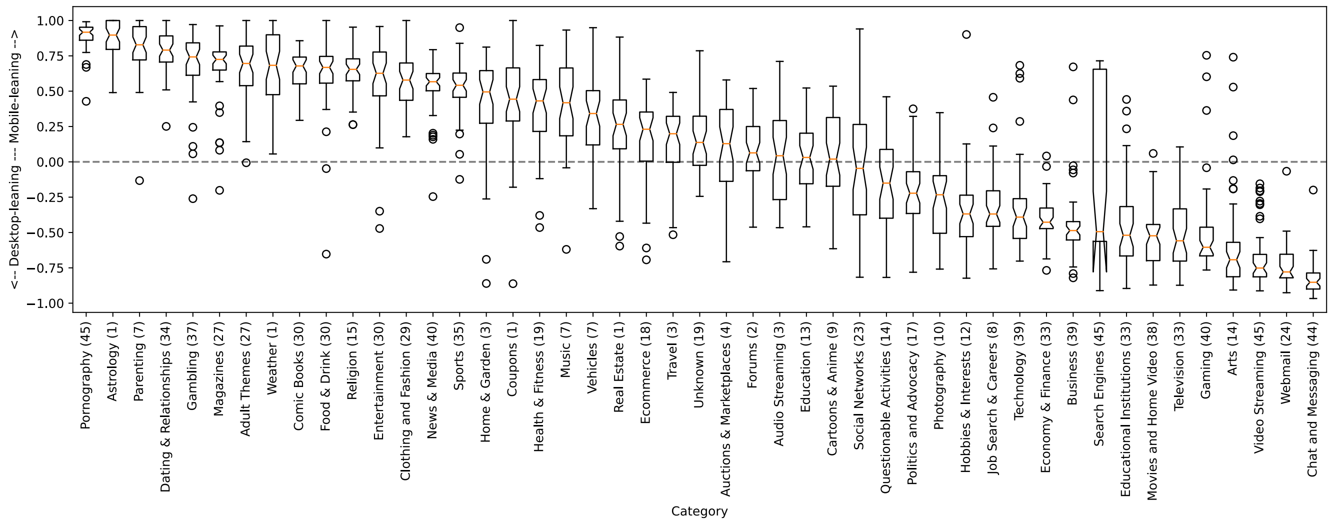

Figure 15: Types of Websites Disproportionately Visited on Desktop and Mobile—Time on page metric. The number of statistically significant per-country results is listed with each category. See Section 4.3 for the corresponding figure for page loads.

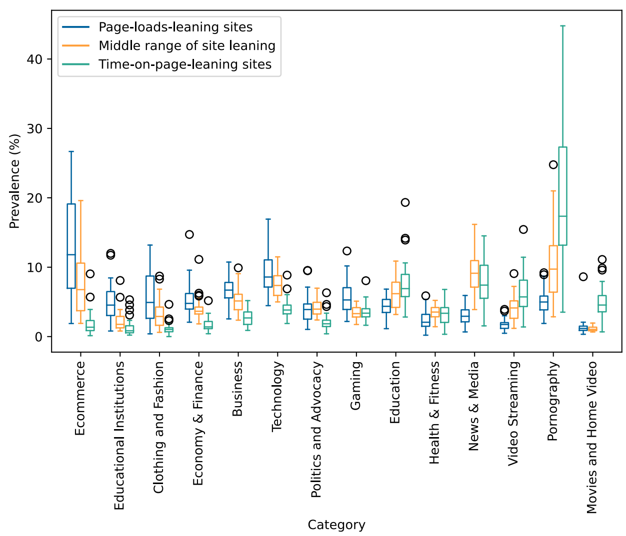

Figure 16: Category Distribution of Page-Load-Leaning, Time-on-Page-Leaning, and all Other Sites—Android top 10K. Each boxplot aggregates results across 45 countries. Only categories with a median of at least 3% prevalence in any of the three sets of sites are shown. See Section 4.4 for the corresponding figure for desktop.

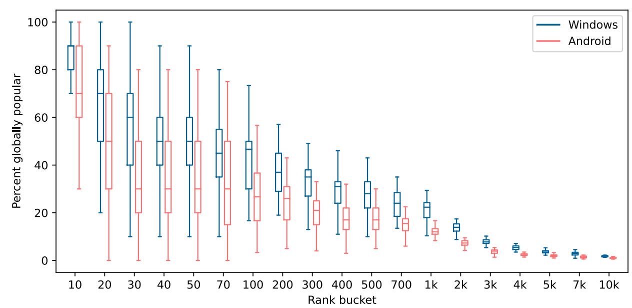

Figure 17: Distribution of Globally Popular Websites by Rank—Time on page metric. See Section 5.2 for the corresponding completed page loads graph.

G WEIGHTED COUNTRY SIMILARITIES, ANDROID AND TIME SPENT

We include for completeness the weighted country similarity heatmaps for Windows time on page (Figure 18), Android page loads (Figure 19), and Android time on page (Figure 20), using the methodology described in Section 5.3.1. We advise cautious interpretation: time on page traffic is likely sensitive to particular video streaming platforms, and Android telemetry does not capture mobile app traffic and so misses user interaction with major sites' native apps.

H SILHOUETTE COEFFICIENT DETAILS

Figure 21 shows the Silhouette Coefficient plot for the country clusters described in Section 5.3.1.

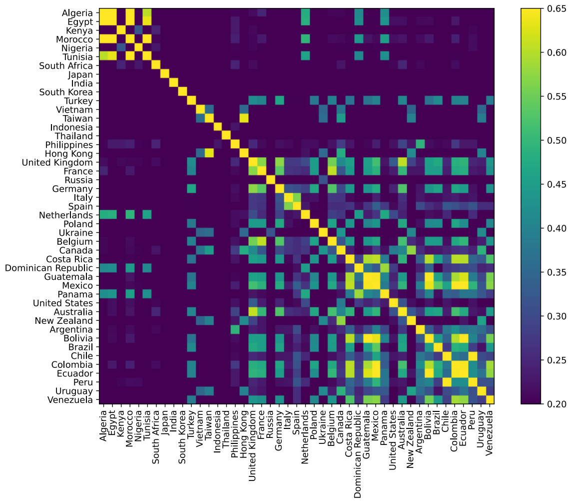

Figure 20: Traffic-Weighted Country Similarities, Android time on page—Country similarities are much lower than for other (platform, metric) pairs, likely due in part to popular streaming services having native apps. See Section 5.3.1 for the corresponding figure for Windows page loads.

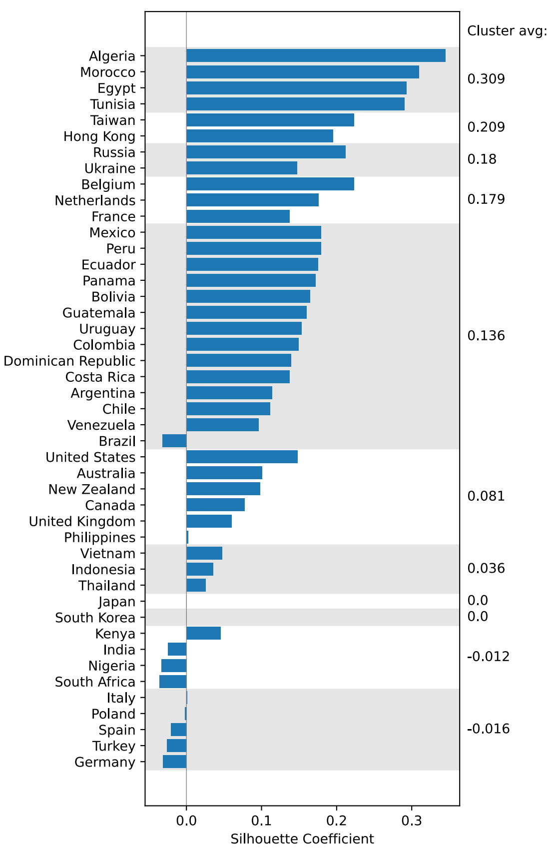

Figure 21: Silhouette Coefficient Plot—Cluster validation for affinity propagation results (Windows completed page loads). Average Silhouette Coefficient for each cluster is shown on the right. Country clusters are loose but some patterns emerge.

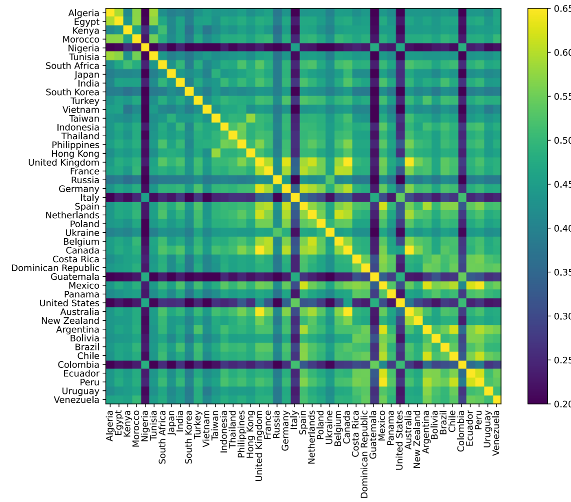

Figure 18: Traffic-Weighted Country Similarities, Windows time on page—Country similarities vary, with the starkest differences likely driven by a mix of sites vying for the number-1 rank by time on page. See Section 5.3.1 for the corresponding figure for Windows page loads.

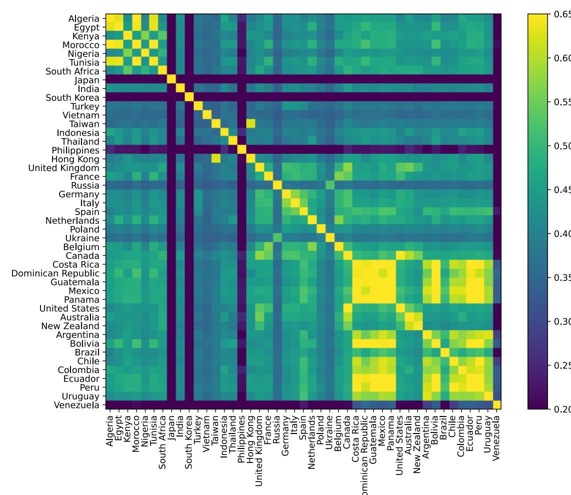

Figure 19: Traffic-Weighted Country Similarities, Android page loads—Comparing against Windows page loads, South Korea is joined by Japan, the Philippines, and Venezuela as outliers. Despite missing traffic from native apps, these similarities are not as low as for Android time on page, possibly due to different category composition of top sites by mobile page loads. See Section 5.3.1 for the corresponding figure for Windows page loads.