Five Blind Men and the Internet:
Towards an Understanding of Internet Traffic

Ege Cem Kirci
ETH Zürich

Ayush Mishra
ETH Zürich

Laurent Vanbever
ETH Zürich

Abstract

Our current view of traffic on the Internet—the world's largest and most pervasive network—comes from a variety of perspectives, each with its own blind spots and biases. In this paper, we make the case for using publicly available Internet exchange point (IXP) statistics to address some of these limitations. We show that IXP traffic data is fine-grained, accessible, and independently verifiable—and that it provides a global view of Internet usage patterns. We present results from a two-year study (2023–2024) from 472 IXPs worldwide, capturing approximately 300 Tbps of peak daily aggregate traffic by late 2024. Over this period, global IXP traffic increased by 49.2% (24.5% annualized), with regionally distinct diurnal patterns and event-driven anomalies. These results illuminate the interplay between network design and function and provide an accessible framework for researchers and operators to study the Internet's evolving ecosystem.

1 Introduction

Many (if not most) network decisions—be it about topology design, capacity planning, routing strategies, or monitoring—depend on the characteristics of carried traffic, such as total volume, expected growth, and inherent variability or *shape*.

This dependency holds at both the "micro" level, within a single Autonomous System (AS), and the "macro" level, the Internet at large. Internet-wide traffic characteristics directly impact CDN caching policies [6], peering and deployment decisions [28], network economics [10], and the scalability requirements of in-network algorithms [40]. For example, the surge of over 30% in global Internet traffic during COVID-19 lockdowns prompted Netflix and YouTube to temporarily reduce video quality in Europe to ease network strain [19].

Yet, while tracking traffic at the level of a single AS is (relatively) straightforward, doing so at Internet scale is challenging. The Internet is massive, diverse, and federated, which makes global, unbiased observability hard. As in the parable of *The Five Blind Men and the Elephant* (Figure 1), our understanding of traffic's size, shape, and form is shaped by vantage-point biases. At a high level, three primary sources inform our current picture of the Internet.

First, measurements by major platforms serving users worldwide capture a *global* view of traffic patterns: for instance, studies from Facebook [8, 12, 17] observed post-COVID-19 growth. While global, such measurements are necessarily biased toward specific applications [32, 18].

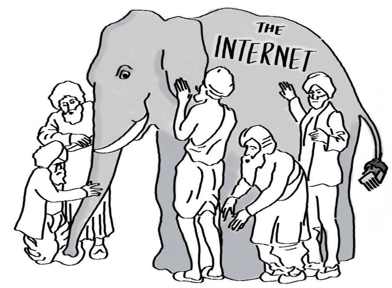

Figure 1: *As in the parable of the Five Blind Men and the Elephant, each of the vantages for studying the Internet touches different parts of it—not the whole—and carries its own blind spots and biases.*

Second, measurements by major transit networks capture an *application-agnostic* view but typically only at local or regional scales, introducing geographical bias [9, 22].

Finally, industry reports from companies such as Sandvine [24] and Cisco [16] aim to provide a global picture. However, these reports are coarse-grained, often forecast-based rather than measured, rely on undisclosed or unverifiable assumptions, and are infrequent (typically annual). They capture overall growth and volume but lack the granularity needed to study usage over shorter time horizons.

The missing *Blind Man*. A representative view of the Internet emerges only by combining *diverse*, *high-quality*, and—crucially—*complementary* sources. We argue that publicly available IXP traffic statistics provide a key missing vantage in that composite: IXP traffic is broadly application-agnostic, more globally distributed, and easily accessible and verifiable.

Prior work shows that IXPs capture a diverse range of Internet traffic, both in coverage [14] and application diversity [15, 6], and that IXPs play a significant role in how traffic is routed [29]. As shown in §4, these statistics are available from diverse locations worldwide, updated frequently (typically every 5 minutes), and publicly accessible.[1]

While IXP traffic is not a precise estimator of the *total* amount of Internet traffic, with sufficient coverage it is rich enough to capture usage patterns and even distinguish atypical behavior during significant global events. IXPs also carry substantial traffic in often under-studied regions such as Africa [39], improving visibility.

[1]Since IXPs have an economic incentive to publish live traffic statistics to attract new members, it is likely that this information will remain available.

We present a two-year measurement study of traffic exchanged by IXPs worldwide from January 2023 to December 2024. To collect these statistics, we built EYE-XP,[2] which crawls IXP websites that make traffic statistics available online. Overall, EYE-XP tracked traffic at 472 IXPs worldwide. These IXPs account for approximately 87% of the total port capacity across the 1065 IXPs listed in PeeringDB, a canonical registry of IXPs [27]. For these IXPs, we collected traffic volumes at 5-minute intervals throughout the entire study period. The result is a rich dataset that supports inferences about global traffic growth and region-specific usage patterns.

We make the following contributions and observations:

- We built EYE-XP, a reliable collection framework that gathers traffic volumes from 472 IXPs worldwide using API calls, HTML parsing, and optical character recognition (OCR).
- Average daily aggregate IXP traffic in our dataset rose by 49.2% over two years (from 138 Tbps to 200 Tbps). The growth trends mirror recent Internet-traffic approximations by Cloudflare [2, 3], suggesting that IXP traffic is a strong proxy for studying Internet traffic growth. In some respects, the "blind men" agree.
- While the blind men agree on some aspects (like growth), their different perspectives also reveal divergences: our IXP data captures an increase during the 2024 Olympics Opening Ceremony, whereas Cloudflare observed a dip at the same time [37].
- Thanks to EYE-XP's global coverage and fine-grained measurements, we observe distinct regional usage patterns, likely driven by differences in lifestyle and work schedules (§6).
- EYE-XP also observes shifts from typical usage in response to global events such as the Olympics, national elections, e-commerce sales, and releases of popular game updates (e.g., Fortnite).

To the best of our knowledge, this is the largest census of Internet traffic to date, surpassing prior work in scale, transparency, and duration. It marks the beginning of a broader initiative to map global Internet traffic dynamics. We plan to open-source our dataset for further research, and our measurement framework continues to collect IXP traffic statistics.

2 Background and Motivation

IXP data has been used to study the Internet in the past [23, 7]. In this section, we review prior studies, their insights, and some of their limitations, and use these observations to later motivate EYE-XP's design.

[2] Like the other blind men, EYE-XP also has its own blind spots; we discuss these in §7.

2.1 Prior IXP studies

Prior work on traffic volumes exchanged at IXPs has typically relied on close collaboration with large exchanges such as AMS-IX and DE-CIX [5, 13, 1]. These studies showed that peak volumes at individual exchanges can reach tens of Tbps [1, 13, 9] and, thanks to fine-grained measurements, revealed distinct diurnal usage patterns in their regions. Using anonymized sFlow records, they also quantified protocol and application-specific contributions to traffic volumes [30, 22, 38]; for example, Feldmann et al. observed increased streaming and collaborative-tool usage during COVID-19 [17]. Overall, this line of work demonstrates that IXP traffic is diverse, representative, and central to Internet connectivity [14, 6].

However, such studies have key limitations: because access often depends on operator agreements, vantage points are concentrated in a few regions, datasets are not easily reproduced or extended by others, and achieving global, longitudinal coverage without proprietary access is difficult. These constraints motivate a complementary approach.

2.2 Observing global usage patterns

Many IXPs publicly advertise aggregate traffic statistics to attract members. These feeds are mostly volumetric time series updated frequently (typically every 5 minutes) and are available across regions. Because they are public, they are accessible and independently verifiable. Aggregating them can yield a global, high-resolution, application-agnostic view of Internet usage patterns.

The challenge is practical rather than conceptual: formats vary (JSON APIs, HTML pages, and image-based plots), some feeds are federated, and care is required to avoid double counting and to normalize units, time zones, and sampling intervals. These realities shaped EYE-XP's design (§3).

It is important to note that we do not claim to see *all*, or even the *most significant* share of Internet traffic with this vantage. Like the other blind men, our data has blind spots as well (discussed in §7). Our goal is to capture the *shape* of Internet usage—its daily/weekly cycles, regional differences, and event-driven anomalies—rather than its absolute total. As we will show in §6, with sufficient coverage this vantage consistently captures characteristic regional habits and identifies atypical behavior in response to global events.

3 Methodology

Our objective is to build a comprehensive, longitudinal dataset of traffic volumes exchanged at IXPs worldwide. To this end, we continuously collect publicly available traffic profiles from all IXPs listed in PeeringDB as frequently as they are updated. Achieving this goal requires overcoming key challenges in *(i)* data discovery, *(ii)* extraction, and *(iii)* system reliability. In this section, we describe these challenges and how our system addresses them.

3.1 Identifying *traffic feeds*

PeeringDB is a community-maintained database widely recognized as the authoritative source for information on IXPs. It catalogs publicly registered IXPs, their geographic locations, region classifications, participating networks (ASNs), those networks' port capacities, and peering policies, serving as a valuable starting point for acquiring traffic volume data. PeeringDB also often provides URLs for IXPs that publicly share traffic statistics; however, we found this field to be sparsely populated and frequently inaccurate in practice. We therefore manually inspected the websites of all IXPs listed in PeeringDB. This multi-week effort yielded a curated list of URLs for publicly accessible traffic statistics pages, enabling continuous monitoring of IXP traffic volumes.

These traffic feeds provide time-series data on outbound and inbound traffic exchanged over an IXP's public peering fabric, typically at five-minute intervals, consistent with the Open-IX OIX-1 standard [26, 25]. Discussions with 20 IXP operators confirm this practice represents total *public* peering volume and excludes private network interconnects (PNIs); accordingly, our analysis covers only the public fabric and does not capture PNIs. To avoid double-counting, we use inbound traffic, noting that at the fabric aggregate inbound and outbound volumes are equal (what enters must exit). Where both series were available, we observed they were virtually identical, with small differences attributable to packet loss within the fabric and measurement/reporting noise.

3.2 Extracting traffic data

The one-time, labor-intensive process of identifying reliable traffic feeds is essential for enabling continuous and accurate extraction of IXP traffic volumes. Upon establishing these feeds, our system methodically extracts traffic volumes from a diverse array of interfaces and formats. Depending on the traffic feed, we employ one of the following three extraction methods:

API-based. For some IXPs, we directly obtain their traffic volume data in JSON format at five-minute intervals via IXP Manager [21].

HTML-based. For IXPs not interfacing with IXP Manager, some embed traffic volume data within webpage HTML or accompanying JSON files. We develop specialized parsers using regular expressions specifically crafted for each individual IXP to extract this data.

OCR-based. For remaining IXPs, we extract traffic data from graphical plots within images. We use the Tesseract OCR engine [35] with two complementary techniques: digitizing plot axes and curve coordinates, and extracting explicitly rendered text values (e.g., "Current Traffic: X Gbps"), when available (see Figure 2).

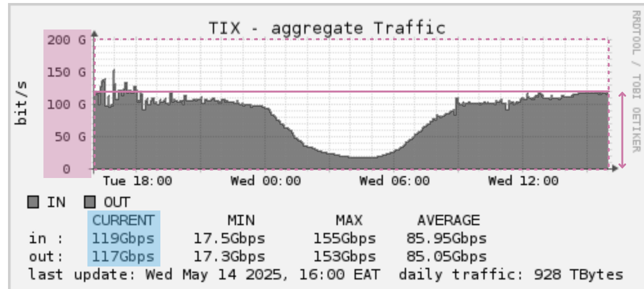

Figure 2: *Example of OCR-based data extraction from graphical plots (Tanzania IX [33]). The red box and line illustrate digitization of plot axes and curve coordinates, while the blue box shows extraction of explicit textual data, such as "Current: X Gbps".*

While API and HTML methods are generally stable once configured, OCR is inherently more fragile. Potential errors include misinterpreting digits or missing decimal points. To mitigate these, we implement a cross-check: when textual values are available on the plot, we compare them with the values derived from digitizing the plot curve. If only plot digitization is possible, accuracy can be affected by image resolution. We evaluated the accuracy of the OCR approach on a test set of images containing both plots and explicit text values. The results revealed a maximum average deviation of 2.4% between the OCR-extracted text values and the values derived from plot digitization, confirming the method's viability for our purposes.

The distribution of the 378 traffic feeds in our study across extraction methods is relatively balanced: 146 rely on OCR, 126 use IXP Manager APIs, and 106 employ HTML parsing. However, the share of traffic volume per method is skewed by operator size. Feeds accessed via the standardized IXP Manager API contribute only 7.2% of the total observed traffic volume, whereas OCR-based and HTML-parsed sources—typically bespoke portals used by larger, higher-traffic IXPs—account for roughly equal portions of the remainder. This suggests that larger IXPs more often expose traffic via custom web interfaces rather than default presentation methods.

3.3 Ensuring reliable collection

Continuous, long-term data collection requires a robust and fault-tolerant infrastructure. We designed and deployed a distributed scraping framework to automate data retrieval, validation, and storage.

Our system achieves fault tolerance primarily through redundancy. The system maintains two nodes: the Main Node and the Backup Node. Both nodes orchestrate scraping tasks

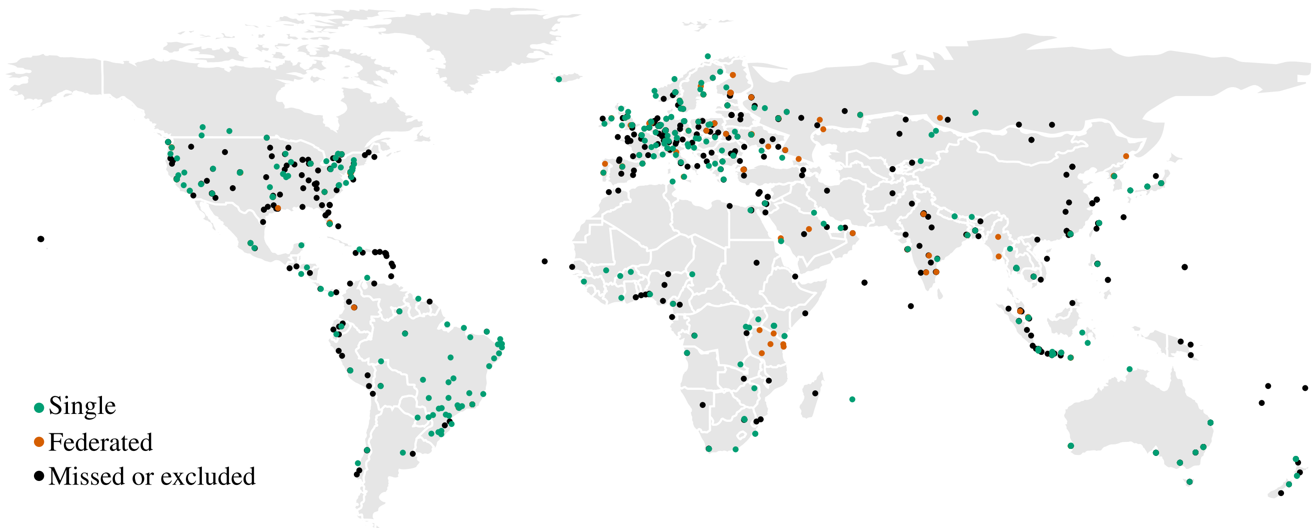

Figure 3: *Map of all IXPs listed in PeeringDB as of January 2023, classified by collection status.*

based on configured frequencies, execute appropriate scrapers (API, HTML, OCR), perform initial data sanity checks, and transmit validated data to a separate Data Node hosting the database. Careful job distribution prevents overload, ensuring smooth operation and consistent data quality.

While the Main Node actively performs these tasks, the Backup Node continuously monitors the Main Node's health by checking data freshness on the Data Node. If the Main Node becomes unresponsive for longer than one polling interval plus a small grace window, the Backup Node automatically assumes scraping responsibilities. This failover mechanism was triggered 57 times during the operational period, during scheduled maintenance and routine patching of the scraping hosts. The swift automated transition minimizes data gaps, ensuring the continuity of our longitudinal dataset.

The Data Node serves as the central repository for IXP traffic volume data, scraper configurations, historical PeeringDB snapshots [11], and operational logs. It handles weekly data archiving and provides a Grafana-based interface for internal monitoring and visualization [20]. Data storage employs a TimescaleDB database (a time-series optimized PostgreSQL extension) [36]. This combination of redundancy, proactive monitoring, and specialized database architecture ensures robust, reliable, and scalable long-term data collection.

4 Dataset

Our system enabled systematic collection of publicly available traffic statistics for IXPs worldwide over two years. Of the 1065 IXPs listed in PeeringDB as of January 2023 (the start of our longitudinal study), we successfully collected continuous traffic data for 472 IXPs. Although these represent slightly less than half of all IXPs listed in PeeringDB, they account for 87% of the total port capacity of all IXPs.

The 472 IXPs analyzed derive their data from two types of traffic feeds:

Single traffic feeds. For 352 of the analyzed IXPs, we identified individual traffic feeds providing a one-to-one mapping between an IXP and its traffic data.

Federated traffic feeds. For the remaining 120 IXPs, we identified traffic feeds that aggregate traffic for multiple IXPs managed by a single organization (e.g., Equinix). Consequently, these 120 IXPs were represented by 26 aggregate traffic feeds.

Thus, the 472 IXPs analyzed were represented by 378 traffic feeds, comprising single and federated feeds.

Missed IXPs. Of the 586 IXPs lacking traffic feeds, 166 had no identifiable websites, leaving their operational status uncertain. For the remaining 420 IXPs with websites, we found no publicly accessible traffic statistics. Our classification relied on best-effort assessment and might contain errors, as some IXPs could host data in non-standard locations or behind logins. Nevertheless, these missed IXPs, including those excluded due to attrition, constitute approximately 13% of global IXP port capacity, suggesting that our dataset captures the majority of global IXP traffic.

Excluded IXPs. Beyond the 472 IXPs with continuous data, 24 traffic feeds initially collected ceased operation during the two-year study due to unreachable websites or discontinued endpoints. Despite efforts to contact operators, the reasons remained unclear. Notably, 16 defunct feeds originated from

Table 1: *Regional distribution and coverage of IXPs.*

Region	IXPs (#)	Share (%)	Collected (#)	Coverage (%)
Africa	72	6.7	38	52.8
Asia-Pacific	225	21.1	65	28.9
Australia	54	5.1	33	61.1
Europe	372	34.9	185	49.7
Mid. East	23	2.2	10	43.5
N. America	227	21.3	88	38.8
S. America	92	8.7	53	57.6
Global total	1065	100	472	44.4

conflict-affected regions (Russia: 5; Ukraine: 11), suggesting geopolitical instability as a potential factor. Additionally, two IXPs, LL-IX and PIT-Arica, were later removed from PeeringDB, confirming their closure. Analysis of their final weeks showed stable traffic volumes (41.6 Gbps for LL-IX and 4.5 Gbps for PIT-Arica), indicating abrupt termination rather than gradual decline. Because PeeringDB is voluntary and lacks strict retirement protocols, some listed IXPs may be defunct, though confirmation is unavailable.

This attrition resulted in 378 feeds providing continuous time-series data throughout 2023–2024. These active feeds consist of 352 single-IXP feeds and 26 federated feeds. Note that, except for the Equinix[3] feed (covering 42 IXPs distributed globally), all other federated feeds include IXPs located within the same region and time zone, as classified by PeeringDB. We summarize the global distribution of single-IXP, federated, and missed IXPs in Figure 3, with attrited IXPs counted as missed. IXP markers are positioned according to their PeeringDB coordinates; for multi-site IXPs within a city, the marker is placed at the median of those coordinates.

4.1 Data quality, coverage, and biases

Given the lack of formal documentation for most public IXP traffic statistics, we assessed data quality along two key dimensions:

Resolution. Most feeds (88%) report data at five-minute intervals, aligning with the MRTG/RRD tool default and the MRTG standard used by IXP Manager. Another 9% report at ten-minute intervals, with the remainder at coarser intervals, 30 minutes being the coarsest we observed.

Timeliness. Most feeds provide near real-time data, but four exhibit notable delays. The Equinix federated feed (covering 42 IXPs globally) and three IX.br locations (Campo Grande, Cascavel, Boa Vista) are five-minute time series, yet values are published as batch backfills and at uneven intervals rather than continuously. Our inquiries regarding these delays remain unanswered.

[3]Equinix later began sharing individual traffic feeds; however, to maintain continuity in our study, we still treat the data as federated.

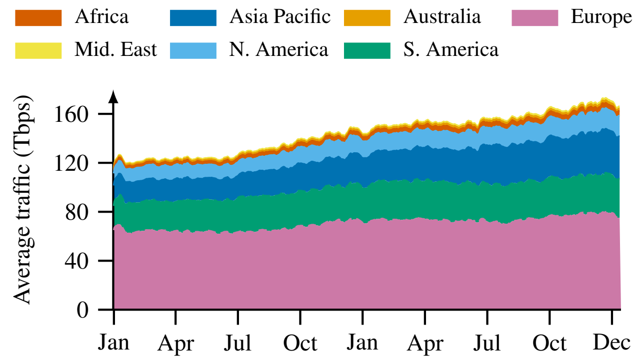

Figure 4: *Stacked plot of daily mean traffic volume by region, with Europe contributing 49% of observed traffic, based on PeeringDB's region classifications.*

The uneven landscape of global IXPs. IXP infrastructure is far from uniformly distributed. As shown in Table 1, Europe (34.9%), North America (21.3%), and Asia-Pacific (21.1%) host the majority (77.3%) of the world's IXPs. In contrast, Africa and the Middle East account for only 6.7% and 2.2%, respectively—an imbalance visible in Figure 3.

This imbalance intensifies at finer granularities. The United States leads all countries with 191 registered IXPs, while approximately half of all countries operate just one, and 50 nations have none. Indonesia, Brazil, Australia, Germany, India, and Russia each host over 40 IXPs. At the metropolitan scale, connectivity concentrates dramatically: a handful of hyper-connected cities—Jakarta (15 IXPs), Amsterdam (14), Frankfurt (13), London (12), and Singapore (11)—boast double-digit exchanges, whereas the median city with an IXP operates just one.

Measurement coverage and blind spots. Beyond regional counts, our measurement coverage also varies significantly (see Table 1). We achieved the highest coverage in Australia (over 60%)[4], South America (over 50%), and Africa (over 50%). All regions except North America and Asia-Pacific exceed 40% coverage. Coverage is notably lower in North America (38.8%) and especially in Asia-Pacific (28.9%). Overall, our dataset encompasses traffic data for 44.4% of all IXPs listed in PeeringDB.

The shortfall in Asia-Pacific is largely attributable to policy-driven opacity in two major economies: China contributes no public traffic feeds, and India provides data for only 23% of its IXPs. Attempts to access IXP websites in these countries often failed from our crawlers, suggesting potentially restrictive sharing practices or network filtering. However, opacity

[4]The Australia region refers to Australia and New Zealand.

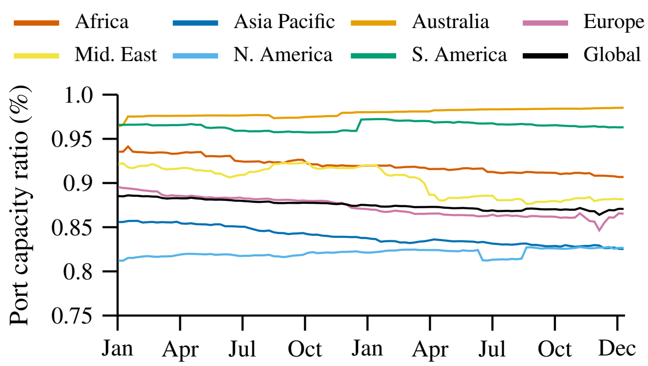

Figure 5: *Port capacity coverage of 472 monitored IXPs, representing 87% of global IXP port capacity.*

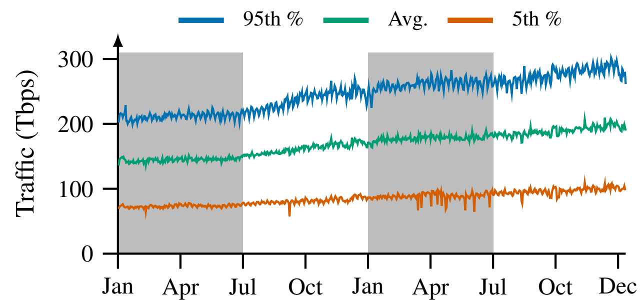

Figure 6: *Global IXP traffic increased significantly over two years, with a greater rise in 2023 than in 2024.*

is not solely an Asian phenomenon; for example, Fremont, California—a major U.S. technology hub—publishes data for only one of its ten registered IXPs. Addressing these coverage gaps remains a key direction for future work.

Traffic volume bias. Observed traffic shares vary across regions and reflect measurement coverage. In Figure 4, we present a stacked plot of average traffic volume by region, excluding Equinix due to its multi-continent footprint.[5]

European IXPs account for approximately 70.2 Tbps, or roughly 49% of the global average daily traffic observed (144 Tbps). This dominance stems from market realities—Europe hosts some of the world's largest IXPs—and high measurement accessibility from major exchanges like DE-CIX Frankfurt, AMS-IX, LINX, and Netnod Stockholm.

North America and Asia-Pacific rank second and third by volume, respectively, but with lower shares than one might expect given their IXP counts, due to limited public data from many large exchanges and the exclusion of Equinix from regional analyses. South America, with high coverage (57.6%), contributes 29.0 Tbps (20%). Africa (3.6 Tbps, 2.5%), the Middle East (1.3 Tbps, 1.0%), and Australia (1.7 Tbps, 1.2%) each contribute smaller shares, despite average or above-average coverage (Table 1). Nevertheless, each of these regions surpasses the 1 Tbps daily average threshold, which we consider sufficient for meaningful volume-based analysis within these regions (see §5.3).

4.2 Representativeness

Given the evident geographic and traffic volume disparities, how representative is our dataset of global IXP infrastructure?

As a measure of representativeness, we divided the total port capacities of all the IXPs we were successfully able to measure in a region by the total port capacities of *all* the IXPs in that region. We plot these ratios in Figure 5 for each of our seven regions. Since IXPs updated their port capacities during our measurement period, we plot these ratios over the entire measurement period.

As we can see, there is a strong evidence of representativeness. Globally, our monitored IXPs account 87% of all publicly announced IXP port capacity listed in PeeringDB. This high capacity coverage remained remarkably stable over 2023–2024, decreasing by only 1.5 percentage points despite considerable growth in global IXP traffic in this period. This high coverage persists regionally, with EYE-XP observing more than 80% available port capacity in all regions. Ranked by coverage, Australia leads (98%), followed by South America (96%) and Africa (92%).

We note that the analysis presented in Figure 5 exclude 76 IXPs (out of 1065) lacking port capacity data in PeeringDB during our study period; manual checks suggest these are typically small, geographically dispersed, or inactive IXPs, with likely negligible impact on aggregate results.

Also, because PeeringDB is self-reported, we assessed the freshness and internal consistency of the capacity data used for this analysis. For all IXPs included in our capacity analysis, the port-capacity-related fields were updated at least once per year during 2023–2024, including IXPs without public traffic feeds, indicating that these records are actively maintained rather than stale. As a plausibility check, where both traffic and capacity were available we verified that observed peaks did not exceed announced capacities; and we found no systematic contradictions.

5 Traffic Growth

In this section, we examine overall IXP-traffic growth during our two-year measurement period and analyze how trends vary seasonally and across regions.

[5]Equinix is likewise excluded from all further regional analyses in this paper.

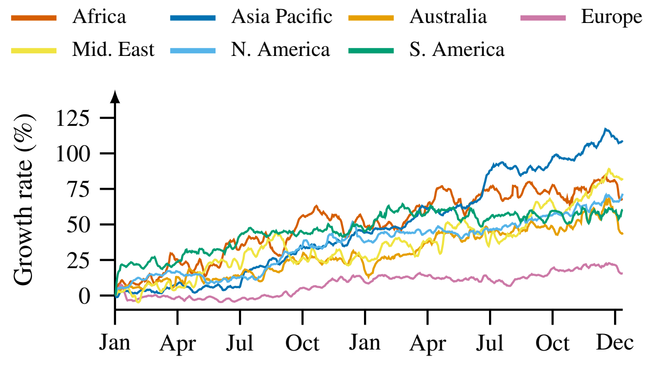

Figure 7: *Traffic surged substantially across all regions, with the Asia-Pacific region experiencing the greatest increase and Europe the most modest rise.*

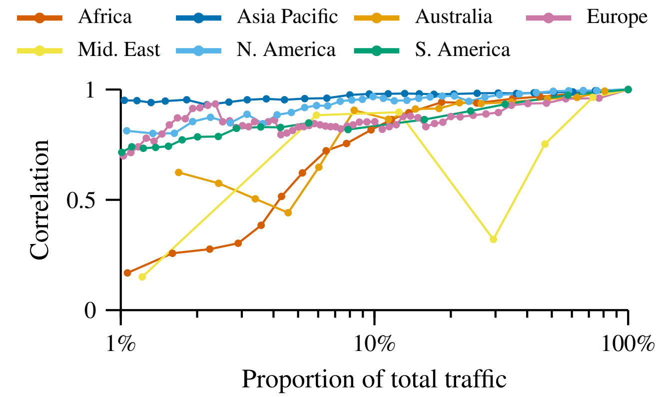

Figure 8: *Convergence of regional growth signals as more observed traffic is included. The Middle East shows irregular convergence due to a large step change at a single IXP; other regions converge smoothly.*

5.1 Global traffic expansion

Between January 2023 and December 2024, aggregate IXP traffic exhibited substantial and consistent growth. We illustrate this trend in Figure 6, which sums the daily mean, peak (95^{th} percentile), and trough (5^{th} percentile) traffic volumes across all traffic profiles to provide a comprehensive view of global IXP traffic volume. The daily mean traffic rose from approximately 138 Tbps to 200 Tbps, a 49.2% increase over the two-year period. Peak and trough traffic experienced similar increases of 47.6% and 54.3%, respectively.

While substantial, growth was not uniform: average traffic increased by 23.4% in 2023 but by a more moderate 16.9% in 2024, based on a linear fit applied to each year. These figures align closely with independent observations (Cloudflare reporting 25.0% and 17.2% global increases, respectively) [2, 3], suggesting that IXP traffic is a strong proxy for overall Internet-traffic growth. Peak traffic followed a similar deceleration pattern (20.0% in 2023 vs. 13.6% in 2024), while trough traffic slowed less dramatically (27.7% vs. 22.2%).

Beyond growth rates, the day-to-day volatility of these metrics reveals distinct characteristics. Peak traffic exhibits the largest absolute daily fluctuations, with typical swings of $\pm$6.0 Tbps (standard deviation). Conversely, trough traffic is proportionally the most volatile, with daily changes averaging 5.1% relative to its mean—more than double the relative volatility of peak (2.4%) or average (2.0%) traffic. Average traffic is the most stable metric, with absolute fluctuations of $\pm$3.4 Tbps and the lowest relative volatility, making it ideal for tracking overall trends reliably.

5.2 Seasonal cycles and holiday impacts

Traffic growth varies significantly within each year of the study period. As evident in Figure 6, global traffic growth remains relatively stagnant in the first half of the year (shaded in gray) but accelerates in the second half. For example, in 2023, average traffic grew by only 3.4% in the first half of the year, compared to 14.3% in the second half.

Regional analysis clarifies this trend. As illustrated in Figure 7, northern hemisphere regions closely follow the global pattern, whereas South America exhibits the opposite trend. For instance, in 2023, South America experienced accelerated traffic growth of 42.5% in the first half of the year, compared to only 17.1% in the second half. Australia, another southern hemisphere region, showed balanced growth throughout the year. A common pattern emerges: traffic growth tends to accelerate during cooler months of the year, possibly due to increased indoor online activity during autumn and winter.

Short-term slowdowns appear around specific holidays and events, including the Christmas period in Europe and North America, New Year's Eve/Day in North America, and Eid al-Fitr (April 21-23, 2023, and April 10-12, 2024), marking the end of Ramadan, in the Middle East. These slowdowns likely reflect reduced business activity and shifts in user behavior during holidays. Note that Figure 6 uses a seven-day rolling (weekly) window average for visual clarity, which smooths short-lived, extreme fluctuations; the unsmoothed series shows these holiday dips more prominently.

These examples demonstrate that, while broad seasonal cycles provide a baseline, specific cultural and regional events introduce predictable, short-term variations in aggregated IXP traffic data. Understanding both large-scale seasonal waves and short-term, event-driven ripples is crucial for accurately interpreting traffic dynamics, a theme further explored in §6.

5.3 Regional divergence

Global averages mask substantial regional heterogeneity (Figure 7). Europe, the largest region by observed volume, shows the most modest growth (15.5%) over two years, while Asia-

Table 2: *Regional growth trends stabilize rapidly, requiring only a fraction of observed traffic to achieve strong correlation with the final regional trend.*

Region	% Obs. traffic for R=0.90	% Obs. traffic for R=0.95
Africa	18.4%	35.0%
Asia-Pacific	0.2%	0.3%
Australia	8.3%	48.1%
Europe	1.9%	56.8%
Mid. East	72.6%	72.6%
N. America	5.3%	8.1%
S. America	23.9%	57.8%

Pacific (108%), Middle East (82%), North America (72%), Africa (68%), and South America (60%) grow markedly faster. Australia shows moderate growth (43%).

Given the disparities in observed traffic volumes across regions (§4.1), we investigated whether these different growth rates were genuine or artifacts of varying coverage and observed traffic volumes across regions. To do so, we analyzed how quickly each region's aggregate growth pattern converges as more IXP data is included. Starting with the IXP contributing the least traffic in a region and progressively adding larger ones by volume, we calculated the Pearson correlation (denoted R) between the two-year growth trend of this cumulative partial aggregate and the growth trend of all observed IXPs in that region. The results are plotted in Figure 8.

For all regions except the Middle East and Australia, correlations increase steadily as more traffic is included, indicating high self-similarity within subsets of IXPs. In the Middle East, Manama-IX (Bahrain), a major contributor to regional traffic, experienced an abrupt 50% drop in average traffic (from 300 Gbps to 150 Gbps) around September 2023, likely due to a network disconnection or contract termination. This sudden shift in a single IXP, diverging from smoother regional trends (yellow line in Figure 8), delays cumulative correlation. A similar, less pronounced event occurred in Australia with Megaport Brisbane in early January 2024, possibly due to a contract termination. These disruptions, while skewing convergence metrics for the Middle East and Australia, reinforce confidence in the data's fidelity, suggesting that, absent such events, convergence would likely occur earlier.

Table 2 summarizes the observed-traffic share needed to achieve $R>0.90$ and $R>0.95$ with the final regional trend. In most regions, growth trends stabilize remarkably quickly. For instance, Asia-Pacific and North America reach $R>0.95$ with less than 10% of observed traffic volume. South America and Europe require 57.8% and 56.8%, respectively, while Africa and Australia achieve $R=0.95$ at 35.0% and 48.1%. The Middle East, due to the Manama-IX disruption, requires 72.6% for both thresholds.

Takeaway: the growth-rate signal converges quickly, so we do not need complete coverage to recover the regional trend; once a modest share of traffic is included (often $< 10\%$), adding more IXPs primarily narrows uncertainty rather than changing the trajectory.

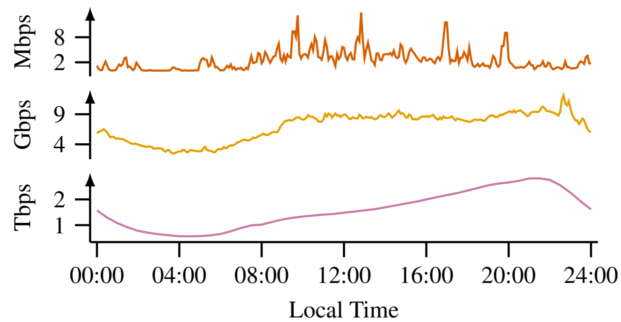

Figure 9: *Same-day, local-time traffic for three IXPs, illustrating how diurnal patterns emerge as average volume increases.*

Overall, the dominant growth signal emerges well before all data is aggregated, varying by region. This internal consistency, combined with our high capacity coverage (§4.2), strongly supports the representativeness of our findings. The convergence suggest that the observed growth rates and regional trajectories reported here accurately reflect the dominant dynamics of the broader regional ecosystems, rather than being artifacts of incomplete measurement. This robustness is further explored from an infrastructure perspective in §A.

6 Usage Patterns

Beyond volumetric trends, our dataset captures daily and weekly traffic rhythms that reflect synchronized Internet user behaviors. This section examines the "shape" of traffic profiles in our dataset to understand underlying drivers of network demand and usage patterns.

We first investigate how discernible traffic patterns emerge. We demonstrate that, as traffic volume passing through an IXP increases, observed traffic patterns shift from unstructured fluctuations to structured diurnal cycles. We then characterize and compare these typical rhythms across global regions. Finally, we show how significant deviations from these traffic patterns indicate real-world events affecting network usage.

6.1 The emergence of diurnal patterns

Analysis of IXP traffic at varying scales reveals a fundamental trend: the emergence of clear, repeating structures in usage patterns correlates strongly with the aggregate traffic volume. At lower volumes, typically below 1 Gbps, daily traffic profiles often appear as relatively unstructured fluctuations without a consistently repeating shape (top plot, Figure 9). As traffic volume increases and more individual user activities are aggregated, these random variations tend to average out, and distinct diurnal patterns reflecting collective human

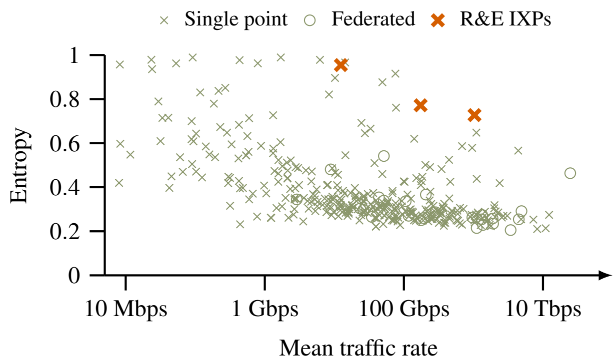

Figure 10: *As average traffic increases, daily patterns become more regular (lower spectral entropy), with notable outliers.*

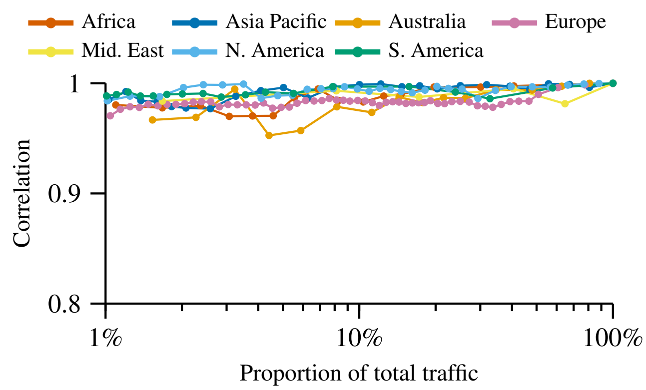

Figure 11: *Regional weekly traffic patterns stabilize rapidly, often requiring less than 2% of observed traffic to strongly correlate with the overall regional pattern.*

activity typically become apparent (middle plot). Once average traffic levels exceed approximately 10–15 Gbps, these daily and weekly patterns generally become well-defined and predictable (bottom plot).

To quantitatively assess when these fluctuations give way to repeating structures, we turn to information theory. We use Shannon entropy (a measure of randomness) to characterize daily traffic patterns. In our context, lower entropy indicates structured traffic patterns and the existence of diurnal cycles, whereas higher entropy reflects erratic and irregular usage throughout the day.

Plotting Shannon entropy against average traffic rate for each traffic profile confirms this trend (Figure 10). Entropy generally decreases as traffic volume increases. This demonstrates that, as IXPs aggregate more diverse user activities, the resulting collective behavior smooths into lower-entropy diurnal patterns. Figure 10 also reveals notable outliers: a cluster of vantage points characterized by both high traffic volumes and persistently high entropy. Manual investigation identifies several of these as major research and education (R&E) IXPs (e.g., CERN IX, PacificWave, and VIX) as depicted with red markers. To our knowledge, these are the only R&E IXPs represented in our dataset. This suggests that these profiles may capture significant volumes of machine-driven traffic, such as large-scale scientific data transfers or inter-campus backups. Unlike typical human-driven traffic, such automated flows can result in continuous or irregular bursts of data at any time, leading to a flatter traffic distribution and thereby maintaining higher entropy despite substantial volumes. This suggests that metrics like Shannon entropy could differentiate human-driven and machine-driven traffic in networks, potentially enabling passive detection of botnets or other large-scale malicious activity.

6.2 Characterizing regional patterns

Across our dataset, the majority of our 378 traffic feeds from 472 tracked IXPs consistently exhibit diurnal patterns at sufficiently high traffic volumes, reflecting the rhythmic nature of global Internet usage. This subsection examines how diurnal patterns vary across regions in our dataset.

To analyze traffic pattern shapes, we adjust each profile, including single and federated profiles that are clustered by timezone, to their local time. These adjusted profiles are then aggregated to derive regional traffic patterns.

To determine how representative an IXP's traffic pattern is of its region, we conduct a detailed analysis—analogous to the growth trend stability analysis in §5.3—assessing the consistency of temporal patterns across IXPs within each region. Specifically, starting with the IXP contributing the least traffic in a region and progressively adding larger ones by volume, we calculate the Pearson correlation between the weekly pattern of the growing partial aggregate and the final pattern derived from all observed IXPs in that region. Figure 11 shows the percentage of observed regional traffic required to achieve near-perfect correlation with the final, fully aggregated regional pattern.

The results reveal a striking contrast to the stabilization of growth trends. Characteristic weekly patterns emerge and converge almost immediately: in most regions, a correlation coefficient of $R > 0.95$ is achieved with less than 2% of observed traffic—often with under 0.1%. Even $R > 0.99$ typically requires sampling under 10%. This rapid stabilization demonstrates that the weekly "shape" of Internet usage is a strong, consistent signal within a region, captured by observing only a small fraction (approximately 10–15 Gbps) of its traffic in our dataset.

Figure 12 presents the average traffic profile for each day of the week, computed per region and normalized by the regional average traffic. These profiles reveal distinct regional rhythms and highlight consistent contrasts between weekdays and weekends.

South America—the boom-and-bust continent. On weekdays, the traffic surges with a peak-to-trough ratio of 4.7×—a full 40% higher than any other region. Yet by Saturday, the peak diminishes by 8.8%, marking the steepest weekend drop observed. The trough is equally pronounced, bottoming out at 0.34× the regional mean at 05:10. Marked by explosive evenings and silent nights, South American traffic is emblematic of a residential-heavy region.

Middle East—where night is the prime time. Cultural and religious customs shift the weekend to Friday-Saturday—and the Internet follows. The weekday peak hits at 18:15, but on the regional weekend, it shifts a remarkable 4 hours and 35 minutes later, to 22:50. Despite this shift, the amplitude remains below 1.9×—the flattest profile observed in our dataset. More striking still, the *late-night share* (23:00-02:00) rises 17% above the daily mean, the highest worldwide. When we add the latest trough (07:35), a pattern emerges: users stay up—and wake up—later than anywhere else.

Asia-Pacific—always on, scarcely off. In Asia-Pacific, the weekend barely registers: peak traffic drops just 3.7%, and the peak time shifts by just 5 minutes. Amplitudes stay high (weekday 3.47×, weekend 3.34×), but the trough never dips below 0.42×. Traffic barely dips after midnight, with troughs arriving at 04:45.

Australia—early birds with sturdy weekends. Down Under, the evening peak arrives earliest at 20:25, and—uniquely—rises slightly on weekends (0.28%). Weekday amplitude reaches 3.47×, dipping only slightly to 3.34× on Saturday and Sunday. During the 08:00-12:00 window, weekend traffic averages approximately 10% below weekday levels—twice the mid-morning dip in Africa (5%) and exceeding other regions (two of which show increases). In short, while much of the world eases into its Saturday scroll, users in Australia are emphatically offline.

Europe vs. North America—same clock, different motors. Both regions hit their trough at 04:05, but that's where the symmetry ends. Europe exhibits a sharper weekday amplitude (3.15×) and a modest weekend dip (1.6%), while North America flattens out with a lower amplitude (2.19×) and virtually no weekend change (+0.2%).

Africa—weekdays wired, weekends rising. Africa's peaks barely move (-0.05% weekend change), yet weekday amplitude hits a solid 3.26×. The trough (04:14, 0.41×) is deeper than Europe's but less extreme than South America's.

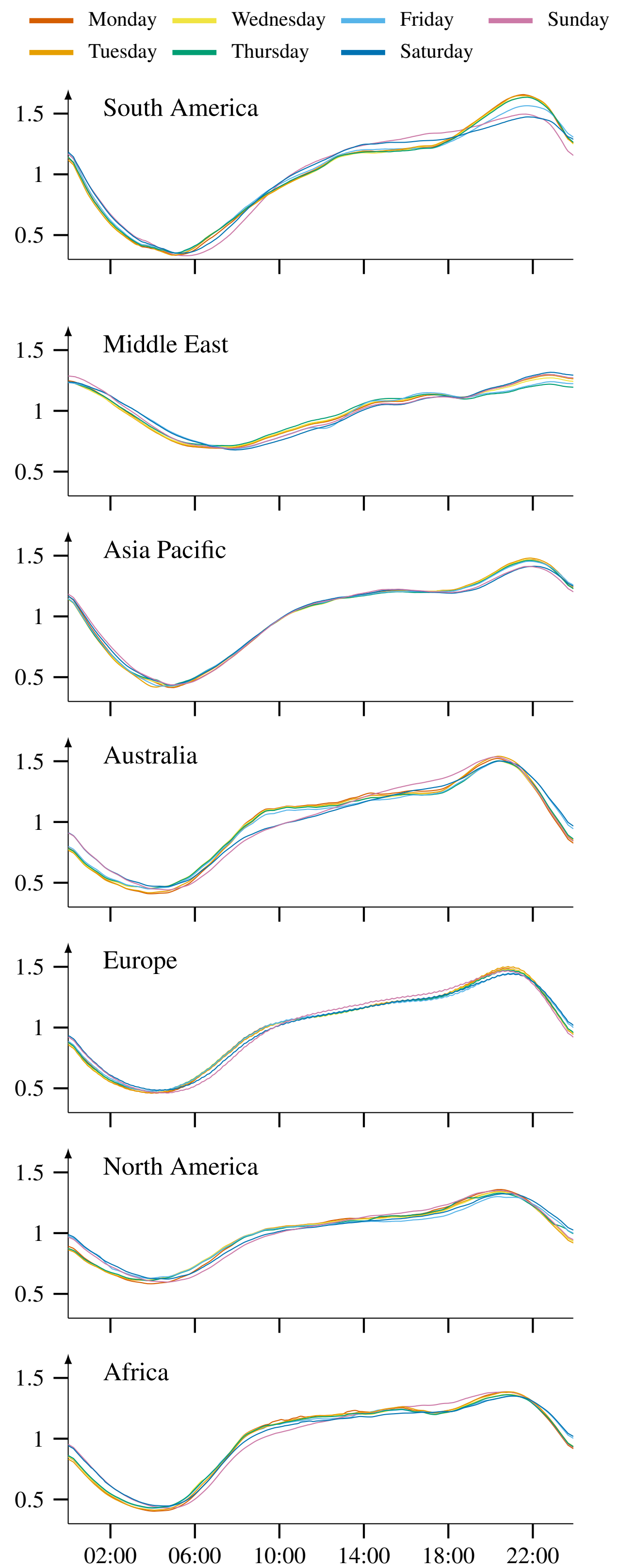

Figure 12: *Distinct weekly traffic patterns across global regions (normalized by each region's mean), reflecting cultural schedules and usage behaviors.*

6.3 Detecting outliers in diurnal patterns

The inherent stability of daily and weekly usage patterns, established in the preceding analyses, provides a robust baseline against which significant deviations can be identified. Such deviations can signal anomalous network events driven by impactful real-world occurrences. Building on this observed pattern regularity, we identify anomalous days by detecting substantial variations in their daily traffic shape compared to a region's typical profile for that day across other weeks. Once an anomalous day is flagged, its intra-day traffic curve is examined against the regional average to pinpoint specific event timings and quantify their impact. For this paper, we validate and select outliers manually; nonetheless, we observe hundreds of such candidates across regions. Work is underway on an automated, scalable pipeline to discover, validate, and characterize these events systematically.

Our dataset reveals numerous instances where aggregated IXP traffic patterns clearly reflect the impact of major societal events. We present a manually curated subset below, contrasting the traffic profile on the event day with typical patterns. For instance, in Africa, traffic on July 26, 2024, deviated significantly from the regional norm, coinciding with the Olympic Games opening ceremony, with observed traffic peaking at twice the typical Friday evening level during ceremony hours. A similar impact was observed in South America on August 7, 2024, during the Olympics Women's football semi-final (Spain vs. Brazil), a popular derby, where traffic during the match hours rose to nearly twice the usual volume for that period. The A-League Men's Grand Final in Australia (May 24, 2024) also produced a noticeable, albeit more moderate, traffic increase of 20-25% during match time (dashed line).

Gaming events also leave a clear imprint. In Australia, a major Fortnite update on November 30, 2024, led to sustained traffic elevation on December 1, 2024, with peak download and gaming hours showing traffic levels over twice the typical Sunday afternoon level. This event also had a significant impact in Europe on December 1, 2024, with observed traffic during peak gaming periods approximately 10% above typical Sunday levels. The re-release of Fortnite's "OG" mode on December 6, 2024, had a further global impact, visibly elevating traffic in the Middle East (approximately 25% above the typical baseline during evening gaming) and similarly affecting North America (around 20% increase during peak engagement hours). E-commerce events, such as Singles' Day on November 13, 2024, in Asia-Pacific, resulted in sustained traffic elevation throughout the day, approximately 30-40% above the regional average for a comparable weekday. Reality television, such as the Big Brother Brasil finale on April 4, 2024, also drove a significant traffic increase in our dataset.

These diverse examples demonstrate that aggregated IXP traffic patterns are highly sensitive and responsive to a wide range of real-world events that drive collective online activity.

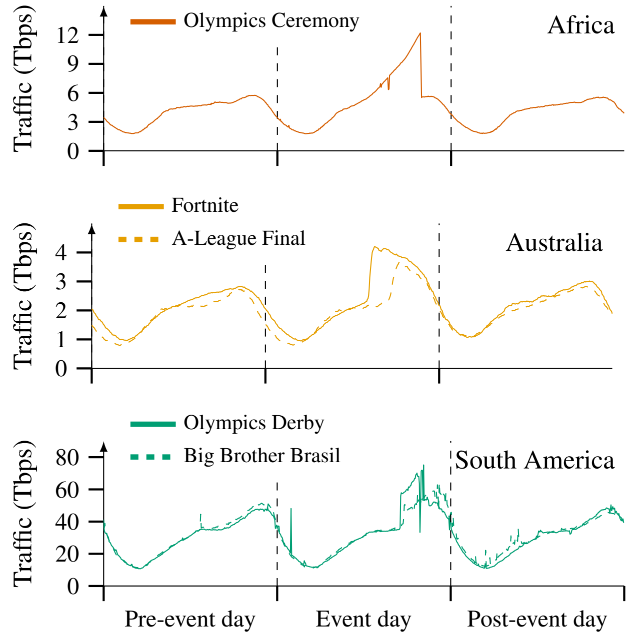

Figure 13: *Sensitivity of regional IXP traffic patterns to major real-world events, showcasing distinct deviations from typical daily profiles.*

The ability to detect and quantify these deviations validates the richness of IXP traffic data beyond simple volume metrics. Such pattern deviation analysis not only provides insights into the network-level impact of different societal phenomena but also highlights the potential for using this approach in near-real-time monitoring to identify widespread service-demand shifts or other significant network anomalies. Understanding the shape of traffic, therefore, offers a complementary and often more nuanced view of Internet dynamics compared to volumetric analysis alone.

7 Discussion

Having introduced our dataset and examined its characteristics across regions, we now discuss its implications: what applications and future research it enables, and its inherent limitations—what it *cannot* do.

Scope and limits of what we measure. At its peak, we observe more than 300 Tbps of IXP traffic globally. However, we do not know what fraction of all Internet traffic this represents, nor how that fraction varies by region. The reported values capture only traffic on the *public* IXP fabric; private network interconnections (PNIs) are outside our measurement scope and may constitute a substantial share in some exchanges.

As argued in §4.2, however, capacity metadata indicates our visibility covers the vast majority of *public* IXP capacity for each region—providing strong support for representativeness of public-fabric dynamics. Prior IXP studies reported tens of Tbps at major exchanges [1, 13] and found IXP traffic to be diverse and broadly representative of Internet exchange activity [14, 6]; our contribution is to provide this view at global, longitudinal scale using public data instead of one-off collaborations.

What our dataset enables. First, it offers a transparent, application-agnostic baseline for monitoring Internet growth. Our measured annual growth (23.4% in 2023; 16.9% in 2024) closely aligns with independent reports such as Cloudflare Radar (25%, 17.2%) [2, 3] and with longer-horizon industry forecasts (e.g., Cisco's 20–30% CAGR) [16]; Sandvine's reports on video's growth (24% in 2023) are consistent with our aggregate increase [24, 31]. While those sources rely on proprietary telemetry, our public-IXP vantage provides a complementary, verifiable baseline. This complementarity is evident during the Paris 2024 Olympics opening ceremony: we observe traffic spikes at African IXPs, whereas Cloudflare reports a pronounced drop on its network [37]—consistent with live-streaming load appearing at IXPs while other Cloudflare-served content temporarily declines—highlighting vantage-specific biases and the value of combining views in underrepresented regions such as Africa.

Second, it enables *regional and seasonal analyses* that are often absent from global reports: a high-level view of regional and country-level Internet traffic growth and patterns can inform policy-makers and researchers about socio-economic and technological trends in a given geography. Furthermore, these reports usually lack the traffic shape and patterns that can inform network management, more sustainable infrastructure solutions, and investment and provisioning—enabling robust regional "shape" baselines useful for capacity planning, energy-aware operations, and benchmarking. Even where overall Internet visibility is limited, the IXP vantage can illuminate underrepresented regions; for Africa in particular, the prevalence of IXPs makes them a useful proxy for overall traffic [39].

Third, understanding usage patterns, or what normal traffic looks like at a vantage point, is especially important to detect *outliers*. Our future work will cover detecting, documenting, and characterizing the outliers we see worldwide. Such a dataset enables research on the reliability of the Internet, the prevalence of security or failure challenges, and informs policy-makers in this age where the Internet has become a human right and a non-negotiable piece of modern life.

What the dataset cannot do. Our vantage cannot *(i)* quantify PNI traffic or infer total interconnection volume; *(ii)* reveal application mix or content/provider shares; *(iii)* attribute causes from traffic alone (corroboration is needed); *(iv)* provide per-AS, per-flow, or path-level insights; or *(v)* guarantee uniform timeliness (a small number of feeds exhibit batching/lag §4.1). Federated feeds can also blur regional assignment in a few cases (e.g., Equinix aggregates), and regional coverage remains lower in some markets (notably Asia-Pacific and North America; §4.1). Accordingly, we interpret capacity coverage as an *upper bound* on provisioned public-fabric capacity, and avoid extrapolating to the entirety of Internet traffic.

Future directions. In addition to the automated event-detection pipeline, we see opportunities to: *(i)* improve coverage in opaque markets; *(ii)* model PNI dynamics indirectly (e.g., combining PeeringDB port churn, BGP announcements, and capacity utilization to bound PNI growth); *(iii)* integrate energy/CO_2 models with diurnal profiles for sustainability planning; and *(iv)* study resilience by correlating traffic anomalies with routing incidents and outages [34, 4]. These directions build on the strengths of a public, longitudinal, global IXP vantage while respecting its scope limits.

The network evolves in lockstep with demand. One of the remarkable things we observed (but do not present in detail due to space constraints) is how closely the port capacity of these IXPs grows with the growth in overall traffic. Our investigations suggest that if this trend is consistent, port capacity could be used as a rough proxy for traffic growth even in IXPs that do not make their traffic statistics publicly available. We provide a deeper analysis of this trend in the supplementary materials (§A).

8 Conclusions

We present a longitudinal study of traffic volumes exchanged at IXPs around the world between January 2023 - December 2024 using EYE-XP. We demonstrate that IXP data, while having its own blind spots and biases, can form a strong complement to our current understanding of Internet traffic. Not only is IXP data global, available, and verifiable, but it is also fine-grained enough to capture regional usage patterns and atypical network usage in response to global events. Using the *daily mean* as our canonical metric, we also observed global aggregate traffic rise from 138 Tbps to 200 Tbps over 2023–2024 (+49.2%; CAGR $\approx$ 24.5%), which is in line with many other estimates of traffic growth on the Internet. We will release this rich dataset and the collection framework (EYE-XP) to support reproducible measurement, operational benchmarking, and future work on event detection, sustainability, and resilience.

Availability. The dataset will be available at `https://github.com/anonymous/ixp-traffic-dataset`.[6]

[6]Obfuscated for double-blind review.

References

[1] Bernhard Ager, Nikolaos Chatzis, Anja Feldmann, Nadi Sarrar, Steve Uhlig, and Walter Willinger. Anatomy of a large european IXP. In *Proceedings of the ACM SIGCOMM 2012 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, Sigcomm '12, pages 163–174, New York, NY, USA, 2012. Association for Computing Machinery.

[2] David Belson. Cloudflare 2023 Year in Review, December 2023.

[3] David Belson. Cloudflare 2024 Year in Review, December 2024.

[4] {Leandro Marcio} Bertholdo, {Joao Marcelo} Ceron, {Lisandro Zambenedetti} Granville, and {Roland Martijn} {van Rijswijk - Deij}. Forecasting the impact of IXP outages using anycast. In *Proceedings of the 5th Network Traffic Measurement and Analysis Conference (TMA)*. IFIP, September 2021.

[5] Timm Böttger, Gianni Antichi, Eder L Fernandes, Roberto di Lallo, Marc Bruyere, Steve Uhlig, Gareth Tyson, and Ignacio Castro. Shaping the internet: 10 years of IXP growth. *arXiv preprint arXiv:1810.10963*, 2018.

[6] Timm Böttger, Felix Cuadrado, Gareth Tyson, Ignacio Castro, and Steve Uhlig. Open Connect Everywhere: A Glimpse at the Internet Ecosystem through the Lens of the Netflix CDN. *SIGCOMM Comput. Commun. Rev.*, 48(1):28–34, April 2018.

[7] Timm Böttger, Felix Cuadrado, and Steve Uhlig. Looking for hypergiants in peeringDB. *SIGCOMM Comput. Commun. Rev.*, 48(3):13–19, September 2018.

[8] Timm Böttger, Ghida Ibrahim, and Ben Vallis. How the Internet reacted to Covid-19: A perspective from Facebook's Edge Network. In *Proceedings of the ACM Internet Measurement Conference*, IMC '20, pages 34–41, New York, NY, USA, October 2020. Association for Computing Machinery.

[9] Samuel Henrique Bucke Brito, Mateus A. S. Santos, Ramon dos Reis Fontes, Danny A. Lachos Perez, and Christian Esteve Rothenberg. Dissecting the largest national ecosystem of public internet eXchange points in brazil. In Thomas Karagiannis and Xenofontas Dimitropoulos, editors, *Passive and Active Measurement*, pages 333–345, Cham, 2016. Springer International Publishing.

[10] Lloyd Brown, Albert Gran Alcoz, Frank Cangialosi, Akshay Narayan, Mohammad Alizadeh, Hari Balakrishnan, Eric Friedman, Ethan Katz-Bassett, Arvind Krishnamurthy, Michael Schapira, and Scott Shenker. Principles for Internet Congestion Management. In *Proceedings of the ACM SIGCOMM 2024 Conference*, ACM SIGCOMM '24, pages 166–180, New York, NY, USA, August 2024. Association for Computing Machinery.

[11] CAIDA UCSD. CAIDA UCSD PeeringDB dataset, 2023-01-01 – 2024-12-31, 2017.

[12] Massimo Candela, Valerio Luconi, and Alessio Vecchio. Impact of the COVID-19 pandemic on the Internet latency: A large-scale study. *Computer Networks*, 182:107495, 2020.

[13] Juan Camilo Cardona Restrepo and Rade Stanojevic. IXP traffic: A macroscopic view. In *Proceedings of the 7th Latin American Networking Conference*, Lanc '12, pages 1–8, New York, NY, USA, 2012. Association for Computing Machinery.

[14] Nikolaos Chatzis, Georgios Smaragdakis, Jan Böttger, Thomas Krenc, and Anja Feldmann. On the benefits of using a large IXP as an internet vantage point. In *Proceedings of the 2013 Conference on Internet Measurement Conference*, IMC '13, pages 333–346, New York, NY, USA, October 2013. Association for Computing Machinery.

[15] Nikolaos Chatzis, Georgios Smaragdakis, Anja Feldmann, and Walter Willinger. There is more to IXPs than meets the eye. *SIGCOMM Comput. Commun. Rev.*, 43(5):19–28, November 2013.

[16] Cisco Systems. Cisco annual internet report (2018–2023) white paper. White Paper C11-741490, Cisco Systems, Inc., March 2020.

[17] Anja Feldmann, Oliver Gasser, Franziska Lichtblau, Enric Pujol, Ingmar Poese, Christoph Dietzel, Daniel Wagner, Matthias Wichtlhuber, Juan Tapiador, Narseo Vallina-Rodriguez, Oliver Hohlfeld, and Georgios Smaragdakis. The lockdown effect: Implications of the COVID-19 pandemic on internet traffic. In *Proceedings of the ACM Internet Measurement Conference*, Imc '20, pages 1–18, New York, NY, USA, 2020. Association for Computing Machinery.

[18] Petros Gigis, Matt Calder, Lefteris Manassakis, George Nomikos, Vasileios Kotronis, Xenofontas Dimitropoulos, Ethan Katz-Bassett, and Georgios Smaragdakis. Seven years in the life of Hypergiants' off-nets. In *Proceedings of the 2021 ACM SIGCOMM 2021 Conference*, Sigcomm '21, pages 516–533, New York, NY, USA, 2021. Association for Computing Machinery.

[19] Hadas Gold. Netflix and YouTube are slowing down in Europe to keep the internet from breaking | CNN Business, March 2020.

[20] Grafana Labs Contributors. Grafana/grafana: The open-source platform for monitoring and observability, May 2025.

[21] INEX – Internet Neutral Exchange Association. IXP manager — world's most trusted IXP platform, 2025.

[22] Craig Labovitz, Scott Iekel-Johnson, Danny McPherson, Jon Oberheide, and Farnam Jahanian. Internet inter-domain traffic. In *Proceedings of the ACM SIGCOMM 2010 Conference*, Sigcomm '10, pages 75–86, New York, NY, USA, 2010. Association for Computing Machinery.

[23] Aemen Lodhi, Natalie Larson, Amogh Dhamdhere, Constantine Dovrolis, and kc claffy. Using peeringDB to understand the peering ecosystem. *SIGCOMM Comput. Commun. Rev.*, 44(2):20–27, April 2014.

[24] Samir Marwaha. Sandvine's 2024 global internet phenomena report: Global internet usage continues to grow, April 2024.

[25] Open-IX Association. IXP technical requirements — OIX-1 (version 2.0). Standard OIX-1, Open-IX Association, October 2016.

[26] Open-IX Association. About us – open-IX association, 2025.

[27] PeeringDB Community. PeeringDB — the interconnection database, 2025.

[28] Maxime Piraux, Louis Navarre, Nicolas Rybowski, Olivier Bonaventure, and Benoit Donnet. Revealing the evolution of a cloud provider through its network weather map. In *Proceedings of the 22nd ACM Internet Measurement Conference*, Imc '22, pages 298–304, New York, NY, USA, 2022. Association for Computing Machinery.

[29] Philipp Richter, Georgios Smaragdakis, Anja Feldmann, Nikolaos Chatzis, Jan Boettger, and Walter Willinger. Peering at Peerings: On the Role of IXP Route Servers. In *Proceedings of the 2014 Conference on Internet Measurement Conference*, IMC '14, pages 31–44, New York, NY, USA, November 2014. Association for Computing Machinery.

[30] Matthew Roughan, Albert Greenberg, Charles Kalmanek, Michael Rumsewicz, Jennifer Yates, and Yin Zhang. Experience in measuring backbone traffic variability: Models, metrics, measurements and meaning. In *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet Measurment*, Imw '02, pages 91–92, New York, NY, USA, 2002. Association for Computing Machinery.

[31] Sandvine. Sandvine's 2023 global internet phenomena report shows 24% jump in video traffic, with netflix volume overtaking YouTube, January 2023.

[32] Brandon Schlinker, Hyojeong Kim, Timothy Cui, Ethan Katz-Bassett, Harsha V. Madhyastha, Italo Cunha, James Quinn, Saif Hasan, Petr Lapukhov, and Hongyi Zeng. Engineering egress with edge fabric: Steering oceans of content to the world. In *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, Sigcomm '17, pages 418–431, New York, NY, USA, 2017. Association for Computing Machinery.

[33] Tanzania Internet eXchange (TIX). TIX traffic statistics portal, 2025.

[34] Malte Tashiro, Romain Fontugne, and Kensuke Fukuda. Following the data trail: An analysis of IXP dependencies. In Philipp Richter, Vaibhav Bajpai, and Esteban Carisimo, editors, *Passive and Active Measurement*, pages 199–227, Cham, 2024. Springer Nature Switzerland.

[35] Tesseract OCR Contributors. Tesseract-ocr/tesseract: Tesseract open source OCR engine, November 2024.

[36] TimescaleDB Contributors. Timescale/timescaledb: A time-series database for high-performance real-time analytics, April 2025.

[37] João Tomé. Paris 2024 Olympics recap: Internet trends, cyber threats, and popular moments, August 2024.

[38] Matthias Wichtlhuber, Eric Strehle, Daniel Kopp, Lars Prepens, Stefan Stegmueller, Alina Rubina, Christoph Dietzel, and Oliver Hohlfeld. IXP scrubber: Learning from blackholing traffic for ML-driven DDoS detection at scale. In *Proceedings of the ACM SIGCOMM 2022 Conference*, Sigcomm '22, pages 707–722, New York, NY, USA, 2022. Association for Computing Machinery.

[39] Semebia Y. Wurah, Theophilus A. Benson, and Edwin Mugume. Poster: Analysis of the internet ecosystem in east africa: DNS resolution, ASN peering, and utilization of ixps. In *Proceedings of the 2024 ACM on Internet Measurement Conference*, Imc '24, pages 785–786, New York, NY, USA, 2024. Association for Computing Machinery.

[40] Tong Yang, Jie Jiang, Peng Liu, Qun Huang, Junzhi Gong, Yang Zhou, Rui Miao, Xiaoming Li, and Steve Uhlig. Elastic sketch: Adaptive and fast network-wide measurements. In *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*, Sigcomm '18, pages 561–575, New York, NY, USA, 2018. Association for Computing Machinery.

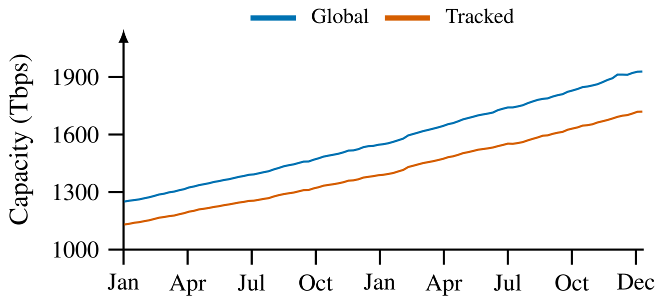

Figure 14: *Tracked and global provisioned port capacity trajectories align, reflecting ecosystem-wide scaling.*

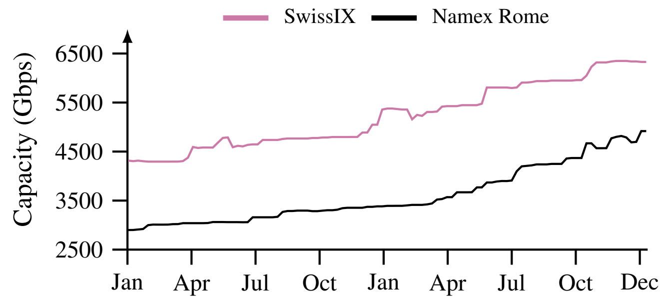

Figure 15: *Port capacity of IXPs increases significantly in a stepwise manner.*

A Port Capacity as a Proxy for Traffic

In §4.2, we used port capacity from PeeringDB to demonstrate the representativeness of our dataset, covering approximately 87% of global IXP port capacity. This supplementary material provides further evidence that port capacity serves as a reliable proxy for inferring traffic volumes and growth trends, especially for unmonitored IXPs.

We first examine the evolution of provisioned IXP port capacities globally (§A.1) and show that capacity expansion is synchronized between the measured and missed/excluded IXPs discussed in this study. We then compare these capacity growth trends with traffic growth (§A.2) and reveal how increase in IXP port capacity is highly correlated with overall traffic growth, albeit with different (but steady) levels of utilization regionally. This overall tight coupling between infrastructure scaling and observed traffic demands gives us confidence that the aggregate traffic growth observed by us can potentially be applied to the entire IXP ecosystem.

A.1 Capacity expansion

How is the global capacity of IXPs increasing worldwide? To answer this, we looked at the growth trajectories of the public peering port capacities of all the IXPs listed on PeeringDB as of January 2023. As a comparison, we also plot the port capacities of the 472 tracked IXPs in Figure 14. Over our two-year study, global provisioned capacity in PeeringDB increased by 54%, from 1125 Tbps to 1910 Tbps. In parallel, our 472 tracked IXPs showed nearly identical growth of 52%, from 1140 Tbps to 1700 Tbps.

This synchronization implies that infrastructure scaling across the IXP ecosystem is driven uniformly by common factors such as anticipated demand, technological advancement, and competitive market pressures, further validating port capacity as a proxy for traffic growth.

While aggregate port capacity increases smoothly and linearly, individual IXP port capacity grows in steps, as expected. Figure 15 plots weekly port capacity updates for SwissIX and Namex Rome. For these IXPs, weekly port capacity increases range from 10 Gbps to 300 Gbps. These updates are not always monotonic, as IXPs occasionally report slight reductions in port capacity. However, overall, IXPs typically consistently increase their port capacities.

A.2 Regional utilization

To assess whether user traffic drives IXP capacity provisioning, this subsection examines regional utilization of our 472 monitored IXPs. Regional network utilization is defined as the ratio of total average daily traffic observed from our 472 monitored IXPs in a region to their total provisioned public peering port capacity, based on contemporaneous PeeringDB data. As depicted in Figure 16, we find striking stability in regional utilization rates over our two-year study. For example, Africa maintains exceptionally consistent utilization, averaging 8.75% with minimal overall fluctuation ($\pm$0.18%). Similarly, Australia (mean 3.57%, -0.41% change) and North America (mean 4.96%, +0.86% change) also demonstrate strong consistency.

Some regions show slightly more pronounced trends, but changes remain modest. Europe's utilization declined from 9.27% to 7.05% (-2.22%), while South America and the Middle East showed decreases of -1.19% and -2.60%, respectively. The Asia Pacific region uniquely experienced an increase, rising from 7.99% to 10.71% (+2.72%).

Despite significant traffic growth in our dataset (§5.1), stable utilization patterns highlight consistent provisioning strategies by IXPs and their peering networks. This stability suggests a deliberate operational approach to maintaining predictable capacity headroom relative to observed demand. Although average utilization setpoints vary significantly by region, reflecting local economic conditions, peering density, and engineering practices, their consistent intra-regional stability points to a high degree of operational predictability within the global IXP landscape, reinforcing port capacity as a proxy for traffic dynamics.

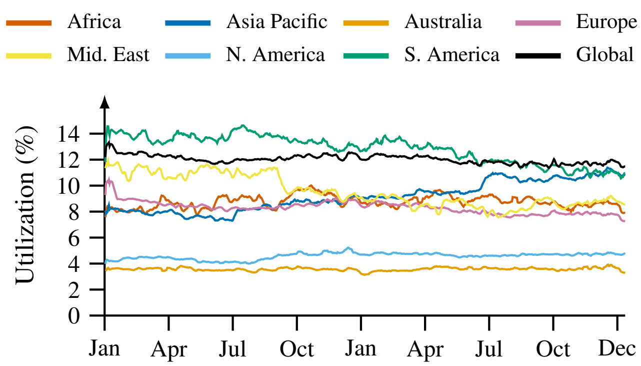

Figure 16: *Regional network utilization varies but remains stable or exhibits modest trends over two years.*

Global utilization of our 472 monitored IXPs remained stable during the two-year measurement period. Overall, our 472 tracked IXPs experienced a 52% capacity increase, closely paralleled by a 49% rise in aggregate observed daily traffic. This proportionality underscores a direct, robust relationship between physical infrastructure investment and realized traffic growth within our extensive sample. Such tightly coupled scaling confirms a foundational aspect of the public IXP ecosystem: infrastructure provisioning closely anticipates and matches traffic demand, reflecting a well-coordinated approach to capacity management across the IXP community. Thus, although we cannot measure their traffic, the port capacity of missed IXPs may provide insights into their overall traffic volumes.

For example, if our dataset, covering approximately 87% of global IXP port capacity, observes approximately 300 Tbps of peak traffic, a naive extrapolation suggests an additional 45 Tbps potentially traversing the unmonitored 13% of capacity. It is likely that the complete picture of the IXP ecosystem is slightly more nuanced. Future work should refine this capacity-based estimation by developing and validating robust regional utilization benchmarks. Such an approach can enable more comprehensive assessments of global IXP traffic. Ultimately, the stable nexus between infrastructure and usage offers a powerful tool for illuminating the broader Internet ecosystem. We believe the rich dataset uncovered in this paper can pave the way towards such an analysis.