Traffic Centralization and Digital Sovereignty: An Analysis Under the Lens of DNS Servers

Demétrio Francisco Freitas Boeira, Eder John Scheid, Muriel Figueredo Franco,
Luciano Zembruzki, Lisandro Zambenedetti Granville

Institute of Informatics (INF)
Federal University of Rio Grande do Sul (UFRGS),
Porto Alegre, Brazil
{demetrio.boeira, ejscheid, mffranco, lzembruzki, granville}@inf.ufrgs.br

***Abstract*—The Domain Name System (DNS) service is one of the pillars of the Internet. This service allows users to access websites on the Internet through easy-to-remember domain names rather than complex numeric IP addresses. However, the concentration of DNS service providers on the Internet affects user security, privacy, and network accessibility as the reliance on a small number of large DNS providers can lead to (a) risks of data breaches and disruption of service in the event of failures and (b) concerns about the digital sovereignty of countries regarding DNS hosting. This work approaches the issue of DNS concentration on the Internet by presenting a solution to measure DNS hosting centralization and digital sovereignty in different countries, such as Brazil, India, China, Russia, and South Africa. With the data obtained through these measurements, relevant questions are answered, such as which are the top-10 DNS providers, if there is DNS centralization, and how dependent countries are on such providers to manage domains using their country code Top-Level Domains (ccTLD).**

***Index Terms*—DNS, Internet Access, Communication Protocols, Digital Sovereignty, Measurement**

I. Introduction

The Internet's Domain Name System (DNS) is a globally hierarchical naming mechanism that enables the association of networks, servers, and services to Internet Protocol (IP) addresses [1]. DNS enables, for example, accessing Websites through easy-to-remember domain names rather than IP addresses, meaning that *wikipedia.org* would be translated to `208.80.154.224`. The records that map domain names and IP addresses are maintained by authoritative DNS servers that provide authoritative and up-to-date records.

Because deploying a local DNS server requires technical expertise [2], companies not rarely have been delegating the task of maintaining their authoritative NameServers (NS) records to third-party DNS providers (*e.g.*, Cloudflare [3] and Akamai [4]). Such a delegation, which has been increasing over the years [5], led to the current scenario where DNS resolution is concentrated on a small number of large providers. And, for the sake of the business model, each large DNS provider multiplexes its Information Technology (IT) or data center infrastructure among its client companies [6]. As a result, DNS centralization inevitably leads to security and availability risks [7], such as user privacy and the inability to resolve domain names in case of an outage or service failure at one of the large providers. In addition, the overall digital dependency on a few IT service providers creates concerns regarding the *(a)* dependability and *(b)* digital sovereignty of countries [8], especially considering compliance regulations, such as Europe's General Data Protection Regulation (GDPR) and Brazil's Data Protection Law (LGPD).

DNS centralization has been widely investigated in the literature. There exist a number of research efforts on assessing the degree of centralization in authoritative DNS servers [7] [5] [9], showing, for example, that popular domains share the same authoritative DNS servers. Thus, disruptions (*e.g.*, due to cyberattacks or sabotage) on DNS infrastructure providers could lead to collateral damages to multiple DNS domains.

Although this centralization aspect has been previously addressed, further research and actions on digital sovereignty implications is necessary considering such a DNS dependency [10]. Analyzing digital sovereignty is crucial because it ensures a country's autonomy, control, and security over its digital infrastructure [11]. Efforts to quantify the dependency of different countries on DNS providers are, thus, required to uncover possible sovereignty risks for the nations and their critical infrastructures (*e.g.*, healthcare, banking, and education sectors), too.

In this paper, we investigate how country code Top-Level Domain (ccTLD) (*i.e.*, TLDs reserved for a country, sovereign state, or dependent territory) from two conglomerate of countries, *(i)* Brazil, Russia, India, China, and South Africa (BRICS) and *(ii)* the European Union (EU), are resolved and quantify their dependency on foreign DNS providers. For that, we define an approach to periodically collect measurements regarding `NS` records, `A` records, and `AAAA` records in order to find out and map the organizations responsible for managing such providers' infrastructure. These measurements use domains listed in the Tranco list [12]. Thus, we also analyze how domains are managed and discuss the implications on regulations, compliance, and digital sovereignty under the DNS scope. The results show that DNS centralization is a reality and a key concern for digital sovereignty, especially for countries that do not have relevant DNS providers and rely on infrastructure providers from countries or companies with different regulations and interests.

The rest of this paper is organized as follows. In Section II, we review background knowledge and discuss related work on DNS centralization and digital sovereignty. In Section III, we introduce our *DNS Measurement* approach and its components, including implementation details. In Section IV, we present the evaluation and results, followed by a discussion in Section V. Finally, in Section VI, we close this paper by presenting conclusions and discussions on future work.

II. Background and Related Work

Due to the damage DNS centralization may bring to the Internet infrastructure [5], academia has been addressing and discussing such a topic in recent years. Efforts and observations found an alarming concentration of DNS traffic, with more than 50% of the observed traffic being handled by only 10 Autonomous System (AS) operators [13]. Further, this leads to efforts toward emerging topics to build a responsible Internet [14], which proposes more transparency and trust within networks, independent of vendors and countries that run the underlying infrastructure. Thus, it is clear that companies from the technology and telecommunication sectors have a place to ensure secure communication and a key role in digital sovereignty.

In the context of DNS centralization, there are significant concerns about the impacts it may cause. One crucial concern is related to performance and how a centralized environment may negatively affect the response time of DNS in some regions of the globe [9]. Moreover, privacy is a main concern, as Internet Service Providers (ISP) typically operate DNS resolvers for their customers, meaning that they have access to users' DNS queries and can potentially monitor or manipulate the data, which is definitely a fact to watch. This centralization of DNS resolution can raise privacy and political concerns, mainly if ISPs engage in activities such as DNS filtering, censorship, or surveillance [15].

Furthermore, another concern that DNS centralization may bring is security since cyberattacks are evolving and becoming more sophisticated [16], including those that target or explore DNS (*e.g.*, tunneling, amplification, and flooding) to cause technical, economic, and societal impacts [17]. This security concern includes attacks on DNS authoritative servers [18], [19] and, also, availability of services worldwide, since the phenomenon of centralization is, in addition to being logical, also physical and geographical [20].

Another concept that emerges from the discussions on DNS centralization is digital sovereignty. Digital sovereignty refers to a nation's ability to control its digital infrastructure, data, and digital technologies within its territorial borders [21]. It encompasses the idea that countries should be able to shape their digital policies, regulations, and frameworks to protect their national interests, security, and values in the digital realm. Digital sovereignty relies on certain aspects, such as data protection and privacy regulations, domestic digital infrastructure, digital trade and economic policies, and Internet governance [8]. Different works have focused on sovereignty from different perspectives, such as the usage o the decentralization provided by blockchain technology [22] as a potential ally for digital sovereignty [23]. However, it is unlikely that fundamental changes will become a reality in the short term since, besides enormous technological efforts and associated costs, it depends on convergence between technical and political spheres.

Digital sovereignty is intrinsically linked with Electronic-Government (e-gov) as services provided to citizens should be independent of foreign countries and highly resilient as they process and maintain personal and sensitive data of citizens. [24] assesses the DNS resilience of e-gov in four countries (*e.g.*, The Netherlands, Sweden, Switzerland, and the United States - US). The work measures the DNS structure of the list of domains used by the government of such countries and provides recommendations on how to improve the robustness of the DNS infrastructure of e-gov services.

Besides underlying communications infrastructure, many scopes and challenges must be explored and addressed towards digital sovereignty. Examples include *(i)* data protection, *(ii)* technological independence (*e.g.*, 5G and mobile communications), *(iii)* cybersecurity, and *(iv)* commercial software applications. Besides all technical and economic arguments possible [17], cybersecurity is a key concern for digital sovereignty because there is a considerable amount of money from a few groups and countries funding strategic cybersecurity companies worldwide. For example, in 2021, the US and Israel combined accounted for nearly 90% of all venture funding for cybersecurity companies [25]. Moreover, solutions towards data sovereignty [26] are also under discussion, which involves strong data regulations that ensure compliance with privacy standards and user control over personal data.

The analysis of [27] focuses on the scope of data ownership and control. It investigates the risks and vulnerabilities of using cloud services for hosting Australian government data (*e.g.*, financial services, intellectual property, and administrative decisions) in different locations and regulations. The authors concluded that the scenario is manageable for the Australian case but argue that it is important to continue fostering and developing its cloud technology while negotiating and designing trade rules and regulations for data flows.

The discussion on digital and data sovereignty is also on the agenda of European industry and academia. For example, [28] proposes Gaia-X, an ecosystem composed of a trusted infrastructure and data space that allow secure data exchanges through federated nodes and actors. Furthermore, a federated architecture is explored to ensure privacy and data protection among all data providers, thus, resulting in an ecosystem that favors the digital sovereignty of conferments and companies.

Further, there are efforts to understand and reduce the dependency on physical devices, especially considering novel technologies (*e.g.*, 5G and 6G). In [29], the authors discuss how the Slovak Republic is addressing such a concern in terms of allowing or prohibiting certain device suppliers from providing the 5G infrastructure in the country depending on security dimensions and taking the diversification of suppliers to avoid centralization in a single company.

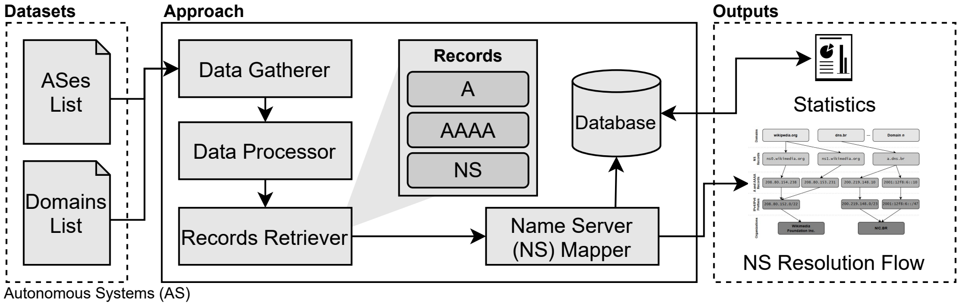


Fig. 1: Overview of the Approach to Analysis Domains

In conclusion, digital sovereignty is a complex and evolving topic, and there are debates around its implementation and potential trade-offs. However, to the best of our knowledge, this is the first effort to analyze and discuss digital sovereignty of several countries based on the infrastructure of DNS providers of their ccTLDs. Hence, this paper contributes to discussions about digital sovereignty under the technical lens of DNS.

III. Measurement Approach

The approach consists of the mapping of lists of popular Internet domains (*e.g.*, based on the publicly available rankings) to its authoritative NSes and organizations behind providing such a service. This allows us to identify *(i)* who provides the correct IPs, *(ii)* which organization operates the NS infrastructure, and *(iii)* to which country and regulations the operator is subjected. For that, the approach combines information from domains (*e.g.*, `A`, `AAAA`, and `NS` records) and AS records provided by Internet registries (*e.g.*, LACNIC, RIPE, and ARIN).

An AS is a network of interconnected computing devices that operate under the same policy. It is often managed by a single entity (*e.g.*, ISPs or technology organizations) and is identified by an AS Number (ASN). Each AS manages one or more unique IP ranges, for example, *Wikimedia Foundation Inc.* has an ASN 14907 and manages the IP range `208.80.152.0/22` in the US and `185.71.138.0/24` in the Netherlands. Thus, it is possible to associate the IP of any NS to an AS and, consequently, to its operator and region.

Therefore, the approach is able to determine the entire flow from the domain name to the organization handling the AS that manages the IP of the associated NS. This allows us to understand the different points where centralization and digital sovereignty risks might occur. For example, the owner of an NS can tamper with the DNS records, while the AS operator can outage the communication.

In both scenarios mentioned, a clear DNS-related dependence can be identified on a few players that maintain the underlying infrastructure (*e.g.*, those that operate ASes and NSes). This makes the need to analyze such players and centralization a key pillar for discussing digital sovereignty.

Figure 1 depicts the components that are part of the approach and the flow of information between them. They are organized into three main groups, namely **Datasets**, **Approach**, and **Outputs**. Datasets containing information regarding *ASes* and a list of *Domains* are used as inputs.

The ASes responsible for each NS are defined using the list provided by the Center for Applied Internet Data Analysis (CAIDA). For that, it was used the network prefix mapping to AS [30] and the mapping of AS to organizations [31]. This allows us to determine the AS, the organization managing the AS, and, thus, the country/region of DNS providers (based on the IP of the NS). For each measurement, an updated list of CAIDA is obtained by the *Data Gatherer* and processed so that the analysis rely on up-to-date information.

For the domains, the Tranco list [12] is used as a dataset since it provides an updated source of the top 1 million Websites on the Internet based on popularity and access traffic. The list is updated considering different sources, *e.g.*, Alexa, SimilarWeb, and Moz; with the latest list used in the experiments (*cf.* Section IV) generated on June 16, 2023. This offers a reliable and transparent list that can be used to conduct research using popular domains. The *Data Gatherer* obtains the updated Tranco list for each measurement (using a diff approach to identify changes) and the *Data Processor* organizes the information of both ASes and domains to be used in further steps.

Next, the *Records Retrieves* analyzes each one of the 1 million domains and retrieves information regarding the `A`, `AAAA`, and `NS` records. For example, for the domain *wikipedia.org*, the `A` is `208.80.154.224`, the `AAAA` is `2620:0:861:ED1A::1` and the `NS` is `ns0.wikimedia.org`. This information is sent to the mapper to understand the entire path to resolve the DNS in order to build the *NS Resolution Flow* and to collects statistics (*e.g.*, organizations concentration, measurement errors, and identified IPs) for further analysis.

The *NS Mapper* receives the records regarding the domain and obtains the IP of the NS. This information is then used to map the IP to the correspondent AS managing it. For that, the `A` record can be used in case of an IPv4 prefix or the `AAAA` for IPv6. Finally, the organization name is obtained using a lookup to the CAIDA AS organization rank mapping dataset [31]. Therefore, a complete analysis can be conducted to identify the nameserver region and relevant characteristics (*e.g.*, regulations and number of ASes being operated) so that insights from characteristics can be drawn.

The *NS Mapper* then stores information obtained in the *Database* and builds, as output, the *NS Resolution Flow*. This flow shows how the domain is resolved until discovering the organization or company that is managing the infrastructure, which is a point that may directly impact DNS resolutions in case of network disruption. Further, identifying NSes is crucial as they might tamper with DNS records, as they answer the requests in an authoritative manner.

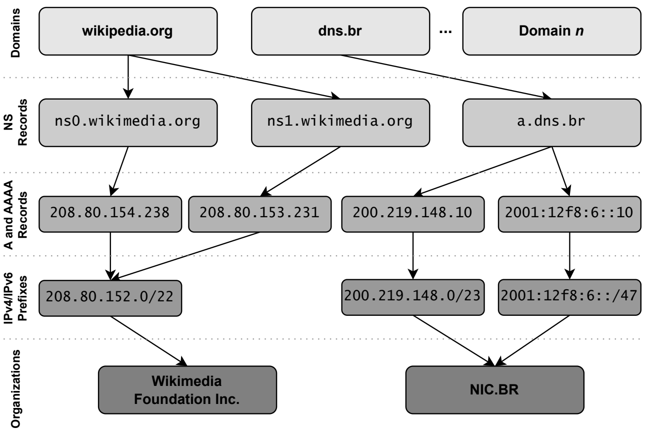


Fig. 2: NS Resolution Flow Example

Figure 2 illustrates a graph-like structure of the *NS Resolution Flow* for the domains *wikipedia.org* and *dns.br*. In the example, *wikipedia.org* has two `NS` records, *(i)* `ns0.wikimedia.org` and *(ii)* `ns1.wikimedia.org`, while *dns.br* has one, `a.dns.br`. This means these NSes are authoritative servers for these domains and are crucial to their operation.

This also applies to the organization that manages the IP addresses and advertises routing information of such servers (*i.e.*, their ASes). Such organizations, including the countries they are operating, are retrieved using the `A` and `AAAA` records of the NSes by identifying the resolved IPs using their prefixes and mapping them with the AS dataset list. Thus, in the example, *wikipedia.org* is managed by the Wikimedia Foundation Inc., located in the US, and *dns.br* is managed by NIC.BR, located in Brazil.

The implementation of the *DNS Measurement* and results of the evaluation of this paper are available at [32]. Python was used to implement the approach's components, with the *dnspython* [33], a Python library to request and manipulate DNS records, being used to implement the *Records Retriever*. The *NS Mapper* connects with a *SQLite3* database to store the data required to build the *NS Resolution Flow*.

Further, statistics can be retrieved and processed from such a database. Table I provides examples of information collected using the approach. Thousands of entries (*i.e.*, domains analysis) were stored, following these metrics and organization, as CSV files for further analysis (*cf.* Section IV).

IV. Evaluation and Analysis

The measurements considered all the 1 million domains from the Tranco list [12], using only the pay-level domains filter, with the latest Tranco list used in the experiments generated on June 16, 2023. To infer the AS names and countries, the CAIDA's AS-to-organization dataset [30] was used. A six-core AMD Ryzen 5-5500U @ 2.1 GHz with 8 GB of RAM connected to the Internet using an Ethernet cable to maintain a stable network connection was used to conduct the measurements. Its operation system was a Debian 11 "*bullseye*" stable distribution.

TABLE I: Example of Information Collected using the Measurement Approach

Information	Description	Example
Server Provider	Describe the nameserver allocated to resolve the domain	`ns3.google.com`
Organization Name	Identifies the organization owning the server	Google LLC
AS Number	Identifies the number of the AS managing the infrastructure of server provider	15 169
AS Country	Shows the country where the AS is operating	USA
AS Occurrences	Number of occurrences of the AS number controlling infrastructure of providers per country	30 812 (US) and 32 989 (RU) for BRICS
Centralization Percentage	The percentage that the AS occurrences represents in the data analyzed (*i.e.*, centralization)	35% (US) and 37% (RU) for BRICS

It is essential to mention that during the experiments, not all domains from the Tranco list were resolved correctly (*e.g.*, DNS records not found or incorrectly configured), and their NS or ASN was not identified; thus, hindering the possibility of identifying the country where their DNS was managed. However, such limitation does not invalidate the results and contributions provided herein as for the digital sovereignty analysis the percentage of unresolved domains was $< 5\%$.

A. Identifying Top-10 DNS Providers

Table II lists the ranking, using 10 positions, of the DNS providers identified during the analysis of the centralization aspect of the DNS traffic. The position in the ranking is based on the number of domains that rely on such DNS providers during the indicated period. Three periods were defined, **Period 1** from 16/12/2022 to 23/01/2023, **Period 2** from 23/01/2023 to 13/02/2023, and **Period 3** from 13/02/2023 to 15/03/2023. As can be seen in the table, the ranking remained stable during these periods, and there was only one change, rows highlighted in gray in the table indicate a change in the ranking, where TIGGEE was the 6th during the first two periods but replaced MICROSOFT-CORP-MSN-AS-BLOCK as 7th in the third period.

Within this context, it was also investigated if the domains of such DNS providers (*e.g.*, *cloudflare.com*) were managed by them or if they relied on services from competitors. Table III presents the results of such investigation. The results indicate that not all DNS providers rely on their DNS services for their domains. For example, Amazon, the second largest DNS

TABLE II: Top-10 DNS Providers Identified

Position	Period 1	Period 2	Period 3
1st	CLOUDFLARENET	CLOUDFLARENET	CLOUDFLARENET
2nd	AMAZON-02	AMAZON-02	AMAZON-02
3rd	GODDADY-DNS	GODDADY-DNS	GODDADY-DNS
4th	ALIBABA-CN-NET	ALIBABA-CN-NET	ALIBABA-CN-NET
5th	GOOGLE	GOOGLE	GOOGLE
6th	TIGGEE	TIGGEE	MICROSOFT-CORP
7th	MICROSOFT-CORP	MICROSOFT-CORP	TIGGEE
8th	NSONE	NSONE	NSONE
9th	IONOS-AS	IONOS-AS	IONOS-AS
10th	OVH	OVH	OVH

Gray-highlighted rows indicate a change in the ranking.

provider according to Table II, uses Oracle's DNS services, and Godaddy, which employs its own DNS service but also relies on Akamai's DNS service. However, the majority of providers use their own DNS service.

TABLE III: DNS Providers Domains and their Providers

Domain	DNS Provider	Country
cloudflare.com	CLOUDFLARENET	US
amazon.com	ORACLE-BMC-31898	US
godaddy.com	GODADDY-DNS, AKAMAI-ANS2	DE, NL
alibaba.com	ALIBABA-CN-NET	US
google.com	GOOGLE	US
tiggee.com	TIGGEE	US
microsoft.com	MICROSOFT-CORP-MSN-AS-BLOCK	US
ns1.com	NSONE	US
ionos.com	IONOS-AS	DE
ovh.com	OVH	FR

B. Measuring DNS Centralization

Having identified the top-10 DNS providers that are responsible for hosting the highest amount of domains in the list, one question that arises is if there is an apparent centralization on those providers or if the DNS providing service is highly distributed to avoid Single Point of Failures (SPoF) or monopoly. To address this question, the concentration of domains resolved by the providers listed in Table II was measured from 16/12/2022 to 15/03/2023.

Figure 3 depicts the results from the performed concentration measurements. In the figure, the *x*-axis represents the date on which the concentration percentage was calculated, and the *y*-axis represents the concentration in the top-10 providers. Considering the period, the average concentration was 30% of the measured domains. This means that, on average, 30% of the one million domains of the Tranco list (*i.e.*, 300 000 domains) had their DNS records hosted by the top 10 DNS providers (*cf.* Table II). Further, considering that such a concentration peaked at 39% on 29/01/2023 and the fact that it was identified that around 3000 DNS providers were responsible for managing all of the one million domains, there is strong evidence that centralization in the DNS hosting industry is a reality.

C. Analyzing Digital Sovereignty

Narrowing down the discussion on DNS centralization to a country-based analysis, it is possible to analyze countries' dependency on these providers and quantify how sovereign its Internet infrastructure is in terms of DNS hosting. For that, domains from the Tranco list were selected based on their ccTLDs (*e.g.*, `.br` and `.cn`) and grouped into their political conglomerates. In total, 91 286 domains from 95 792 domains using the BRICS and EU ccTLDs were resolved, and their DNS hosting organization identified. This represents 9.1% and 9.5% of the Tranco list, respectively. Russia's ccTLD (`.ru`) represented 59% of the resolved domains, approximately 54 168 domains. Results from such analysis categorized by these groups are presented in the following sections.

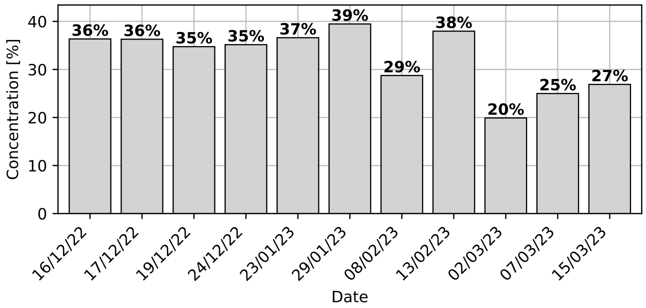


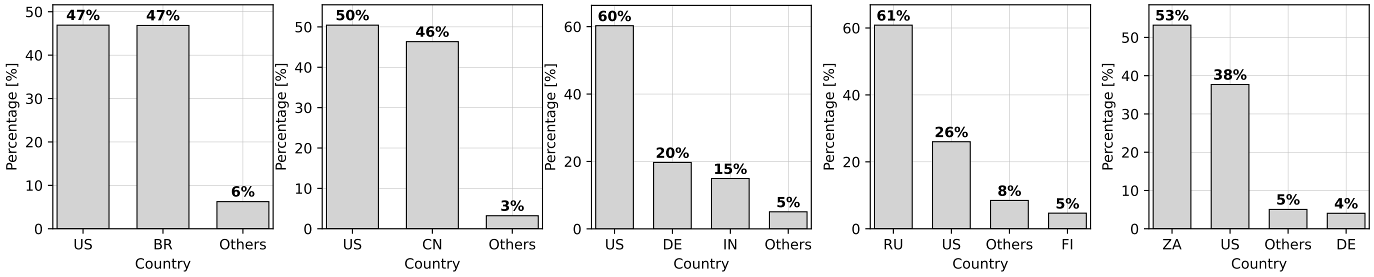
Fig. 3: Concentration on Top-10 DNS Providers over 3 Months

1) BRICS Domains: BRICS represents a conglomerate of five major emerging economies, namely *(a)* Brazil, *(b)* China, *(c)* India, *(d)* Russia, and *(e)* South Africa, formed to promote inter-economic cooperation and inter-political discussions. As BRICS does not have an official ccTLD as Europe, the ccTLD for the BRICS are, respectively, *(a)* `.br`, *(b)* `.cn`, *(c)* `.in`, *(d)* `.ru`, and *(e)* `.za`.

Figure 4 depicts the results of the BRICS analysis. For each chart, the *x*-axis represents where the AS operates, and the *y*-axis represents the percentage of domains having their authoritative servers relying on such an AS. The countries are represented as Alpha-2 ISO country codes [34], and countries with less than 4% of domains were aggregated in the "Others" category. For example, in Brazil (*cf.* Figure 4a), there was a tie between `.br` domains that relied on DNS providers from the US (*i.e.*, 47%) and domains that are provided by Brazilian-based companies (*i.e.*, 47%). The remaining share (*i.e.*, 6%) was located in other countries (*e.g.*, France and Germany).

It is possible to observe that US-based DNS providers, such as Cloudflare, Inc., Amazon.com, Inc., and Google LLC, represent a significant portion of the DNS hosting industry in the BRICS, with India presenting the highest dependence (*i.e.*, 60%) of the five nations. The exceptions are Russia (61%) and South Africa (53%), with most domains provided by national DNS companies (*e.g.*, Yandex.Cloud LLC for Russia and Xneelo (Pty) Ltd for South Africa). Thus, showing indications of concern regarding digital sovereignty.

Further, to have an overview of the digital sovereignty of the BRICS as a conglomerate, the five countries' results were aggregated and illustrated in Figure 5. Russia and the US appear to host the majority of the domains (*i.e.*, a total



(a) `.br` (b) `.cn` (c) `.in` (d) `.ru` (e) `.za`

Fig. 4: Results from the BRICS Domains Separated by ccTLD

of 73%), followed by Brazil, China, and Germany. This behavior is logical considering the division of Figure 4. Therefore, showing a dystopian view of digital sovereignty, where the BRICS is subject to and dependent on the US regarding DNS regulations and infrastructure.

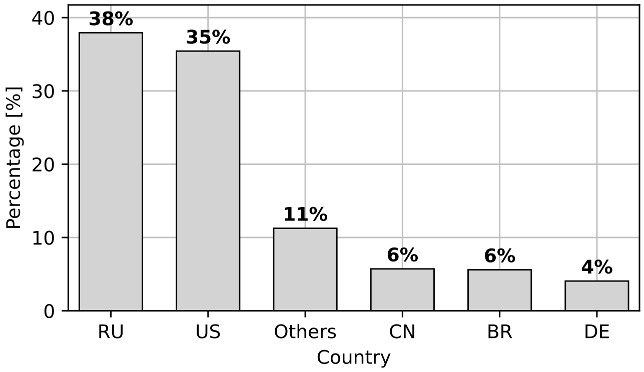


Fig. 5: Results from the Aggregated BRICS Domains

2) European Union: The EU is a political and economic union composed of 27 member states (*e.g.*, Portugal, Spain, France, Italy, Germany, and Hungary) located in Europe. For such countries, in the first moment, the `.eu` ccTLD was examined due to space constraints; however, the analysis of different European ccTLDs is planned for future work. Any person, company or organization within the EU may register domains with this ccTLD. Figure 6 illustrates a different scenario than the one from the BRICS (*cf.* Figure 5), where more countries share the DNS hosting infrastructure of the EU. Germany (*i.e.*, DE) represents a significant portion given its size and number of DNS hosting providers.

However, the US also concentrates a significant portion of the DNS hosting industry for `.eu` domains. After Germany, France, and the Netherlands appear as major countries hosting DNS domains for Europe, this supports the data presented in Table II, where OVH, a French cloud computing company, appears as the 10th DNS provider in the ranking. This concentration in a cloud provider might indicate that other services, besides DNS, are being hosted in France and the Netherlands, given the fact that such companies offer more services than DNS, such as virtual machines, Function-as-a-Service (FaaS), and web hosting that require a DNS provider.

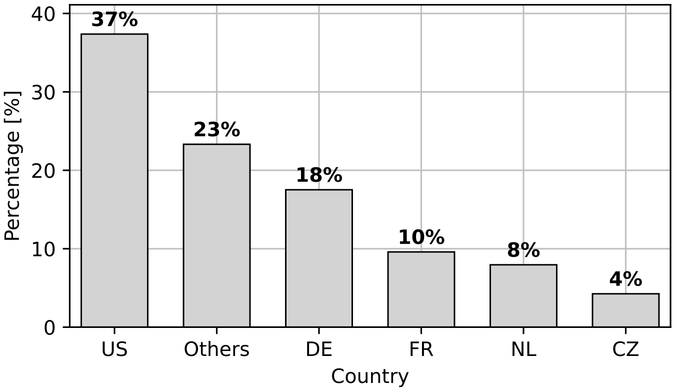


Fig. 6: Results from the `.eu` Domain

D. Hosting Governmental Domains

One analysis dimension that is highly relevant concerning digital sovereignty and centralization is to investigate where restricted TLDs, such as `.gov`, are hosted. These domains are intended to be used only by federal government institutions (*e.g.*, security agencies and institutes). Thus, their DNS should be hosted within federal organizations to maintain critical services for citizens and control over the infrastructure during critical periods (*e.g.*, global conflicts, pandemics, or sanctions).

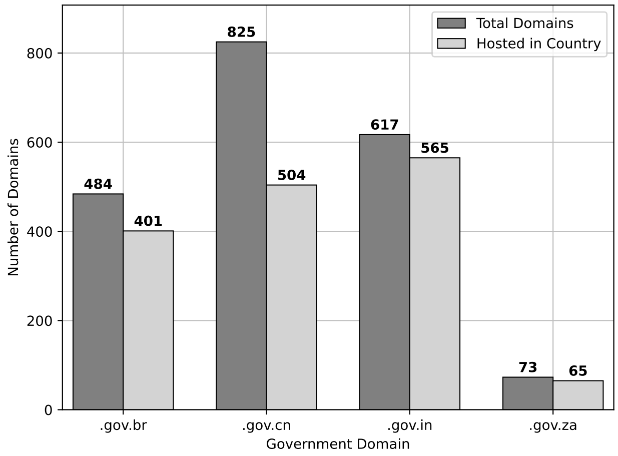


Fig. 7: Results from the Analysis of the `.gov.` Domains

Figure 7 depicts the results from the analysis of the BRICS domains: .gov.br, .gov.cn, .gov.in, and .gov.za. Russia did not present .gov domains in the Tranco list; hence, it is not presented in the results. It can be seen that Brazil's governmental domains are mostly resolved within Brazil, specifically in the Federal Data Processing Service (Serviço Federal de Processamento de Dados - SERPRO, in Portuguese), which is the biggest government-owned corporation of IT services in Brazil. Further, Indian and South African government domains are mostly hosted in their countries, with the National Informatics Centre (NIC) hosting most domains for India and the State Information Technology Agency (SITA) for South Africa. These results show a concern within BRICS about hosting governmental DNS domains for federal services within government organizations to avoid censorship, data leakage, and disruption of critical services.

V. Discussions and Key Observations

Different insights can be obtained from our experiments under different dimensions. From the technical dimension, we have shown that there is evidence of centralization on a few key players. Further, we showed that DNS centralization is economic in nature since big techs from developed countries lead the market. Moreover, several economic impacts (*e.g.*, business disruption and reputation harm) may happen in companies and governments in case of intentional or non-intentional disruption of the underlying DNS infrastructure. Our findings can also be explored from a legal dimension since digital sovereignty involves regulations and actions that can be done by policy-makers based on the technical analysis of the different protocols and dependence (*e.g.*, DNS and its centralization on a few companies and countries). The rest of this section provides a discussion on each dimension.

On the **technical** dimension, based on the results, it can be assumed that there is a clear indication of a DNS centralization, which can lead to a scenario where the Internet's infrastructure and management are directly dependent on a few key players (*e.g.*, governments and companies with different technical and political characteristics). This is not the best scenario since it can lead to the issues discussed in Section II, such as security, assurance, and operational risks. Moreover, allowing such centralization in a given country, region, or company increases the risk of Internet censorship, as such a control can be achieved by injecting fake DNS replies to block access to certain content [35]. Thus, the DNS infrastructure and its distribution concentrated on a few authoritative servers may lead to Internet outages (due to misconfigurations) and Internet censorship, as the technical enablers for implementing this control are in place.

When discussing the economic dimension of DNS centralization, one point that relates is the possibility of DNS providers profiting from DNS lookup data. [36] advocates that DNS providers do not commercialize such information because of the potential consumer and regulatory backlash of such a monetization. However, suppose the DNS provider's centralization occurs in a country with not-so-well-defined regulations concerning commercializing user-sensitive data. In that case, further monopoly is risky as DNS lookup can be valuable for advertisement. Thus, monitoring and addressing DNS centralization and digital sovereignty is critical to tackling such an economic perspective. Further, most DNS providers (*e.g.*, Amazon, Google, and Microsoft) are also major cloud provider companies [37], where their business is strongly tied to providing a reliable DNS infrastructure to access such cloud instances. However, such a combined service offering leads to a vendor lock-in issue [38] and even further dependence on their infrastructure, in which clients are subject to such companies' pricing policies.

In addition to these possible economic impacts, DNS centralization has an economic motivation since big techs (often based in the US) offer DNS infrastructure, resolvers and associated services as part of their business core. In 2020, the DNS market was worth USD 372 million, and it is expected to be worth USD 862 million by 2025 [39]. This growth expectation is attributed to the increasing number of domain name registrations and Web traffic. Concerns about security, centralization, and digital sovereignty may be part of the marketing and product development strategies for DNS providers and big techs operating the underlying infrastructure.

Lastly, in the **legal and political** dimension, there are different efforts from the EU to strengthen its digital sovereignty, such as the GDPR for the idea of data sovereignty and the action plan for more digital sovereignty called by governments of Germany, Estonia, Denmark, and Finland [40]. Cybersecurity experts, entrepreneurs, and decision-makers also moved to the discussion to highlight the need to develop and promote digital infrastructures under European technological sovereignty [41]. Even though digital sovereignty is receiving much political attention around the world, discussions still need to evolve to find a common understanding to succeed in such dimensions.

In Brazil, the topic is being discussed among debates on different regulations that are required to increase national cybersecurity and digital sovereignty [42]. Thus, as seen with these examples and discussions, digital sovereignty is a matter that many stakeholders (*e.g.*, governments, companies, and society) have to address from technical, economic, and legal perspectives. Otherwise, digital colonialism may become more prominent and dangerous in the following years, providing mechanisms to increase censorship and digital warfare.

Thus, as shown in this work and experiments, we advocate that the analysis and discussions on digital sovereignty under different lenses are needed. In parallel to the discussion on the centralization of protocols, such as DNS, different aspects, such as cybersecurity, regulations and investments for technology, and mobile communications and its vendors, must be investigated to lead the discussions of digital sovereignty.

In summary, digital sovereignty relates to different layers (*cf.* Figure 8); layers depicted in white color are not covered in the discussions of this paper, while gray-highlighted layers were covered. Thus, the research and discussions presented herein address the **data**, **technical**, **operational**, and **assurance** sovereignty layers within the *self-determination* aspect.

As pointed out in [43], data sovereignty is the enabler for organizations to achieve digital sovereignty fully. In this sense, companies and countries should ensure that DNS records and related data are processed and treated within country legislation and rules. Further, the technical aspect relates to striking a balance between companies and governments becoming dependent on in-house solutions that might become legacy systems and becoming dependent on single DNS provider services. The operational aspect relates to DNS services providing transparent information about DNS records and monitoring their infrastructure. Lastly, as discussed in Section II, the DNS infrastructure is crucial for the availability of services and, in the case of DNS centralization, this infrastructure becomes a SPoF, which might affect not only a single service but its entire supply chain (*i.e.*, all services that rely on such DNS provider). Thus, critical services (*e.g.*, governmental and financial) must be resilient regarding DNS availability, allowing users and interested stakeholders to reach a specific service using human-readable names within the country's infrastructure.

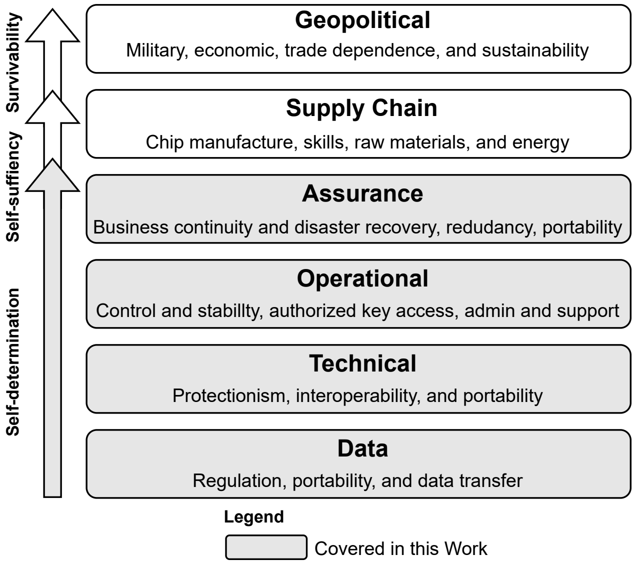

Survivability
Self-suffiency
Self-determination
Geopolitical
Military, economic, trade dependence, and sustainability
Supply Chain
Chip manufacture, skills, raw materials, and energy
Assurance
Business continuity and disaster recovery, redudancy, portability
Operational
Control and stabillty, authorized key access, admin and support
Technical
Protectionism, interoperability, and portability
Data
Regulation, portability, and data transfer
Legend
Covered in this Work

Fig. 8: Digital Sovereignty Stack. Adapted from [44]

VI. Conclusion and Future Work

The Domain Name System (DNS) infrastructure plays an essential role in the Internet access infrastructure by allowing content and services to be reached using easy-to-remember names (*i.e.*, domains). However, during its development, it was never imagined that such a system would become a market of global proportions. Thus, aspects such as its centralization and governmental regulations were disregarded. In this sense, given its central role in society and concerns regarding the level of control that DNS providers could enforce if the system becomes centralized, understanding and identifying DNS centralization is a key concern.

Thus, in this paper, we presented an approach to measure DNS centralization and digital sovereignty based on DNS domain resolution. The approach relies on a list of 1 million popular domains (*i.e.*, the Tranco list) and, for each one, identifies the name server responsible for hosting the domain (*i.e.*, its authoritative server) and, based on its IP address, maps it to the Autonomous System (AS) managing the IP address.

Further, with the AS information, the approach identifies the country in which the AS is located to analyze which regulations the AS is subject to. Consequently, with that information, the approach infers the top-10 DNS providers, the percentage of centralization of the Tranco list in these providers, and the portion of domains that are managed within their country based on its country-code Top-Level Domain (ccTLD).

Results from the analysis show that most of the top-10 DNS providers identified in the Tranco list are in the US, with Cloudflare being the 1st DNS provider. Further, the analysis of how centralized the DNS hosting industry is revealed that the concentration of domains resolved in the identified top-10 providers peaked at almost 40%, which shows signals of centralization.

Lastly, the results of measuring digital sovereignty in Brazil, Russia, India, China, and South Africa (BRICS) and the European Union (EU) unveiled a scenario where a significant percentage of domains within these countries are not hosted by national companies but hosted on US-based organizations; exceptions being Russia and South Africa. Based on the results, it can be said that not only is DNS centralization occurring on the Internet as previous literature showed (*cf.* Section II), but also that countries are becoming less sovereign in terms of control over the national DNS infrastructure.

Considering future work, it is planned to *(a)* analyze such DNS providers distribution with additional countries (*e.g.*, different European countries) that are discussing digital sovereignty, *(b)* address the limitations of the work discussed in Section IV, including the research on the DNS decentralization aspect, and *(c)* create a tool (similar to [13]) to analyze DNS providers distribution periodically. Furthermore, our measurement approach can be extended to analyze additional protocols and technologies to provide a more granular technical view of the digital sovereignty landscape. This also includes exploring the measurement opportunities enabled by programmable networks, such as In-band Network Telemetry (INT) and P4-based programs.

Acknowledgments

This work was supported by the São Paulo Research Foundation (FAPESP) under the grant number 2020/05152-7, the PROFISSA project.

References

[1] P. Mockapetris and K. J. Dunlap, "Development of the Domain Name System," in *Symposium Proceedings on Communications Architectures and Protocols (SIGCOMM 1988)*, August 1988, pp. 123–133.

[2] S. Hao, H. Wang, A. Stavrou, and E. Smirni, "On the DNS Deployment of Modern Web Services," in *IEEE International Conference on Network Protocols (ICNP 2015)*, San Francisco, United States of America, November 2015, pp. 100–110.

[3] Cloudflare, Inc., "Cloudflare DNS - Authoritative and Secondary DNS," 2023, https://www.cloudflare.com/dns/.

[4] K. Schomp, O. Bhardwaj, E. Kurdoglu, M. Muhaimen, and R. K. Sitaraman, “Akamai DNS: Providing Authoritative Answers to the World’s Queries,” in *Annual Conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication (SIGCOMM 2020)*, Virtual Event, USA, July 2020, pp. 465–478.

[5] G. C. M. Moura, S. Castro, W. Hardaker, M. Wullink, and C. Hesselman, “Clouding up the Internet: How Centralized is DNS Traffic Becoming?” in *ACM Internet Measurement Conference (IMC 2020)*, Virtual Event, USA, October 2020, pp. 42–49.

[6] M. Allman, “Comments on DNS Robustness,” in *Proceedings of the Internet Measurement Conference (IMC 2018)*, Boston, USA, October 2018, pp. 84–90.

[7] L. Zembruzki, A. S. Jacobs, G. S. Landtreter, L. Z. Granville, and G. C. M. Moura, “dnstracker: Measuring Centralization of DNS Infrastructure in the Wild,” in *Advanced Information Networking and Applications (AINA 2020)*, L. Barolli, F. Amato, F. Moscato, T. Enokido, and M. Takizawa, Eds., Caserta, Italy, April 2020, pp. 871–882.

[8] P. Roguski, “Layered Sovereignty: Adjusting Traditional Notions of Sovereignty to a Digital Environment,” in *11th International Conference on Cyber Conflict (CyCon)*, vol. 900, Tallinn, Estonia, 2019, pp. 1–13.

[9] T. V. Doan, J. Fries, and V. Bajpai, “Evaluating Public DNS Services in the Wake of Increasing Centralization of DNS,” in *IFIP Networking Conference (IFIP Networking 2021)*, Espoo, Finland, 2021, pp. 1–9.

[10] E. Kantas and M. Dekker, “Security and Privacy for Public DNS Resolvers,” February 2022, ENISA Report. https://www.enisa.europa.eu/publications/security-and-privacy-for-public-dns-resolvers.

[11] J. Pohle and T. Thiel, “Digital Sovereignty,” *Journal of Internet Regulation*, vol. 9, no. 4, pp. 76–88, 2020.

[12] V. Le Pochat, T. Van Goethem, S. Tajalizadehkhoob, M. Korczyński, and W. Joosen, “Tranco: A Research-Oriented Top Sites Ranking Hardened Against Manipulation,” in *26th Network and Distributed System Security Symposium (NDSS 2019)*, San Diego, United States of America, Feb. 2019.

[13] P. Foremski, O. Gasser, and G. C. Moura, “DNS Observatory: The Big Picture of the DNS,” in *ACM Internet Measurement Conference (IMC 2019)*, Amsterdam, Netherlands, October 2019, pp. 87–100.

[14] C. Hesselman, P. Grosso, R. Holz, F. Kuipers, H. Xue, M. Jonker, J. Ruiter, A. Sperotto, R. Rijswijk-Deij, G. Moura, A. Pras, and C. Laat, “A Responsible Internet to Increase Trust in the Digital World,” *Journal of Network and Systems Management*, vol. 28, pp. 882–922, October 2020.

[15] R. Li, X. Jia, Z. Zhang, J. Shao, R. Lu, J. Lin, X. Jia, and G. Wei, “A Longitudinal and Comprehensive Measurement of DNS Strict Privacy,” *IEEE/ACM Transactions on Networking*, pp. 1–16, 2023.

[16] M. Franco, J. von der Assen, L. Boillat, C. Killer, B. Rodrigues, E. J. Scheid, L. Granville, and B. Stiller, “SecGrid: A Visual System for the Analysis and ML-Based Classification of Cyberattack Traffic,” in *IEEE 46th Conference on Local Computer Networks (LCN 2021)*, Edmonton, Canada, October 2021, pp. 1–8.

[17] M. F. Franco, L. Z. Granville, and B. Stiller, “CyberTEA: a Technical and Economic Approach for Cybersecurity Planning and Investment,” in *36th IEEE/IFIP Network Operations and Management Symposium (NOMS 2023)*, Miami, USA, 2023, pp. 1–6.

[18] Y. Jin, M. Tomoishi, and S. Matsuura, “Detection of hijacked authoritative dns servers by name resolution traffic classification,” in *IEEE International Conference on Big Data (Big Data)*. Los Angeles, USA: IEEE, 2019, pp. 6084–6085.

[19] C. Aishwarya, M. Sannidhan, and B. Rajendran, “DNS Security: Need and Role in the Context of Cloud Computing,” in *International Conference on Eco-friendly Computing and Communication Systems*, 2014, pp. 229–232.

[20] B.-S. Lee, Y. S. Tan, Y. Sekiya, A. Narishige, and S. Date, “Availability and Effectiveness of Root DNS servers: A Long Term Study,” in *IEEE Network Operations and Management Symposium (NOMS)*, Osaka, Japan, 2010, pp. 862–865.

[21] A. Aydın and T. K. Bensghir, “Digital Data Sovereignty: Towards a Conceptual Framework,” in *1st International Informatics and Software Engineering Conference (UBMYK)*, Ankara, Turkey, 2019, pp. 1–6.

[22] E. J. Scheid, B. Rodrigues, C. Killer, M. Franco, S. R. Niya, and B. Stiller, *Blockchains and Distributed Ledgers Uncovered: Clarifications, Achievements, and Open Issues*, ser. IFIP AICT Festschrifts. Cham, Switzerland: Springer, Aug 2021, no. 1, pp. 1–29.

[23] S. Manski and B. Manski, “No Gods, No Masters, No Coders? The Future of Sovereignty in a Blockchain World,” *Law and Critique*, vol. 29, pp. 151–162, 2018.

[24] R. Sommese, M. Jonker, J. van der Ham, and G. C. M. Moura, “Assessing e-Government DNS Resilience,” in *18th International Conference on Network and Service Management (CNSM 2022)*, Thessaloniki, Greece, October 2022, pp. 118–126.

[25] Crunchbase, “The Rise Of Global Cybersecurity Venture Funding,” 2021, https://about.crunchbase.com/cybersecurity-research-report-2021/.

[26] P. Hummel, M. Braun, M. Tretter, and P. Dabrock, “Data Sovereignty: A Review,” *Big Data & Society*, vol. 8, no. 1, pp. 1–17, 2021.

[27] A. D. Mitchell and T. Samlidis, “Cloud services and government digital sovereignty in Australia and beyond,” *International Journal of Law and Information Technology*, vol. 29, no. 4, pp. 364–394, 01 2022. [Online]. Available: https://doi.org/10.1093/ijlit/eaac003

[28] A. Braud, G. Fromentoux, B. Radier, and O. Le Grand, “The Road to European Digital Sovereignty with Gaia-X and IDSA,” *IEEE Network*, vol. 35, no. 2, pp. 4–5, 2021.

[29] T. Gábriš and O. Hamuľák, “5G and Digital Sovereignty of the EU: The Slovak Way,” *TalTech Journal of European Studies*, vol. 11, no. 2, pp. 25–47, September 2021.

[30] CAIDA, “Routeviews Prefix-to-AS Mappings (pfx2as) for IPv4 and IPv6,” 2013, https://publicdata.caida.org/datasets/routing/routeviews-prefix2as/.

[31] ——, “Inferred AS to Organization Mapping Dataset,” 2014, https://www.caida.org/catalog/datasets/as-organizations/.

[32] D. F. Boeira, L. Zembruzki, E. J. Scheid, M. F. Franco, “DNS Sovereignty Repository,” 2023, https://github.com/ComputerNetworks-UFRGS/DNS-Sovereignty.

[33] Dnspython Contributors, “dnspython Library,” 2020, https://www.dnspython.org/.

[34] International Organization for Standardization, “ISO 3166 - Country Codes,” 2023, https://www.iso.org/iso-3166-country-codes.html.

[35] P. Pearce, B. Jones, F. Li, R. Ensafi, N. Feamster, N. Weaver, and V. Paxson, “Global Measurement of DNS Manipulation,” in *26th USENIX Conference on Security Symposium (SEC 2017)*, Vancouver, BC, Canada, August 2017, pp. 307–323.

[36] K. Borgolte, T. Chattopadhyay, N. Feamster, M. Kshirsagar, J. Holland, A. Hounsel, and P. Schmitt, “How DNS over HTTPs is Reshaping Privacy, Performance, and Policy in the Internet Ecosystem,” in *47th Research Conference on Communication, Information and Internet Policy (TPRC)*, Washington, DC, July 2019, pp. 1–9.

[37] L. Zembruzki, R. Sommese, L. Z. Granville, A. Selle Jacobs, M. Jonker, and G. C. M. Moura, “Hosting Industry Centralization and Consolidation,” in *IEEE/IFIP Network Operations and Management Symposium (NOMS 2022*, Budapest, Hungary, April 2022, pp. 1–9.

[38] J. Opara-Martins, R. Sahandi, and F. Tian, “Critical Analysis of Vendor Lock-in and its Impact on Cloud Computing Migration: a Business Perspective,” *Journal of Cloud Computing*, vol. 5, no. 4, pp. 1–18, April 2016.

[39] MarketsAndMarkets, “Managed Domain Name System (DNS) Services Market,” 2023, https://www.marketsandmarkets.com/Market-Reports/dns-service-market-240632025.html.

[40] Handelsblatt, “Appell von vier Regierungschefinnen an die EU: “Europa muss seine digitale Souveränität stärken”,” 2021, https://goo.by/xIVUn.

[41] G. D. Rodosek, M. Broy, U. Helmbrecht, “Quo Vadis European Digital Sovereignty?” 2021, https://www.concordia-h2020.eu/blog-post/quo-vadis-european-digital-sovereignty/.

[42] L. Belli, B. Franqueira, E. Bakonyi, L, Chen, N. Couto, S. Chang. N. da Hora, W. Gaspar, “Cibersegurança: Uma Visão Sistêmica Rumo A Uma Proposta De Marco Regulatório Para Um Brasil Digitalmente Soberano,” 2023, https://goo.by/32fNL.

[43] R. Nasir, “The Evolution of Digital Sovereignty: Moving Beyond Data and Cloud,” January 2023, https://blog-idceurope.com/the-evolution-of-digital-sovereignty-moving-beyond-data-and-cloud/.

[44] R. Nasir, R. Duncan, R. Helkenberg, and M. Claps, “IDC’s Worldwide Digital Sovereignty Taxonomy, 2023: Cloud Sovereignty,” May 2023, https://www.idc.com/getdoc.jsp?containerId=EUR150601123.

All links visited on September 2023.