Evidence for a digital divide? Measuring DNS dependencies in the context of the indigenous population of Australia

Ralph Holz[1,2][0000−0001−9614−2377], Niousha Nazemi[1][0000−0003−4085−7044], Omid Tavallaie[1][0000−0002−3367−1236], and Albert Y. Zomaya[1][0000−0002−3090−1059]

[1] School of Computer Science, The University of Sydney, Australia
[2] Department of Mathematics and Computer Science, University of Münster, Germany
{niousha.nazemi,omid.tavallaie,albert.zomaya,ralph.holz}@sydney.edu.au

Abstract. We recently presented a work-in-progress paper at the Workshop on Transparency, Accountability and User Control for a Responsible Internet (TAURIN 2023). Our paper investigates the relationship between the digital divide, Internet transparency, and DNS dependencies. This submission is a condensed version of the original paper, giving a summary of our results and our conclusions, which we would like to present for discussion. In a nutshell, we can indeed identify differences between the DNS dependencies, in particular with respect to the use of hyperscalers and dedicated government infrastructure. Our results show that Internet measurement can detect signals of a possible digital divide, and we believe that this is a worthwhile part in an agenda to make access to Internet services more equitable for different population groups.

1 Introduction

The concept of the digital divide has been the subject of much research and discussion over the past 20 years. The term was first introduced and defined in the mid-to-late 1990s in a series of reports titled "Falling through the net" [5–7]. The definition refers to the gap between individuals or groups who have access to and effectively use digital technologies and those who do not. This includes access to technologies such as the Internet [8]. Individuals who have access to and utilize these technologies are considered advantaged, while those who lack access or proficiency are at a disadvantage [4]. A digital divide often affects already economically disadvantaged groups. The indigenous people of Australia consist of two distinct cultural groups: the Aboriginal peoples of the Australian mainland and Tasmania and the Torres Strait Islander peoples from the seas between Queensland and Papua New Guinea.

The question we ask here is whether DNS dependencies impact how these vulnerable groups can access Internet-based services. We focus on analyzing the impact of the digital divide on indigenous communities in Australia regarding their DNS-mediated access to government websites. Given their geographically dispersed nature, service outages can significantly impact this vulnerable group. We examine the disparities in DNS dependencies of governmental services for the

indigenous and general populations. Our findings imply differences between the setups do exist. For example, sites for the indigenous population use different cloud providers, and while sites for the general population are sometimes run on what seems to be government-owned infrastructure, we find no such setups for sites for the indigenous population.

The following sections are largely from our original paper [3]. We refer to the original paper for a longer summary of related work as well.

2 Methodology

We create two lists: one with the domain names of Australian government websites that provide services to the general population and one with domain names of Australian government websites that provide services for the indigenous populations. We adopt a desk research approach to identify the domains of interest. Once the domain names are obtained, we also perform manual validation to ensure that the collected domain names align with the intended target audience. We finally obtain two lists with unique and relevant domains, each for the respective target audience (448 domains for the general population group and 54 domains for the indigenous group). We proceed to retrieve the authoritative name servers (NS) for the collected domain names by querying every authoritative NS to whom we observed a delegation. In addition, we also utilize the WHOIS command to gather information about the associated provider for each identified name server. We create the delegation graphs to analyze the dependencies. The relationship between domains and their name servers can be categorized as either direct or indirect dependencies. A direct dependency is a domain being directly associated with its designated name servers. These associations indicate an immediate connection between a governmental website and its corresponding DNS service provider. For the analysis presented here, we focus only on these.

3 Results

We analyze the dependency patterns for domains for the general and indigenous populations across various DNS providers. Table 1 provides key statistics on the dependencies we find for various provider types. The table also presents the percentage of domains with a dependency on a single provider versus a dependency on multiple providers. We distinguish between the following kinds of DNS providers:

Leading providers: We use the term “leading providers” to refer to prominent DNS service providers with a significant market presence and influence. These are widely known cloud providers often referred to as hyperscalers. Hyperscalers are a common choice when services must be reachable quickly across a wide geographic area. However, the fact that they are generally headquartered in another country also implies a certain amount of loss in digital sovereignty when they are chosen over a local, domestic provider.

Table 1. Dependency on third-party DNS providers for general and indigenous domains.

Population group	General		Indigenous	
	Absolute	Relative	Absolute	Relative
Number of domains	448	100%	54	100%
Depends on...				
... leading providers	219	48.9%	29	53.7%
... non-leading providers	140	31.3%	25	46.3%
... intra-government providers	113	25.2%	0	
... single provider	412	92%	54	100%
... multiple providers	36	8%	0	0
... intra-government + 3rd party providers	19	4.2%	0	0
Undisclosed	8	1.8%	0	0

Non-leading providers is our term for DNS providers outside the group of the leading (hyperscaler) providers. They generally have a smaller market share and fewer cloud resources and represent a wide and diverse range of DNS service providers. Many domestic (Australian) providers fall into this category.
Intra-government providers are those where the respective governmental sections are responsible for hosting and managing their DNS infrastructure, including offering DNS provisioning for other government sections.
Undisclosed providers: For about two percent of general domains, we could not further identify the DNS providers from either the WHOIS or the domain names of the NS records.

3.1 Analysis by Provider Type

Fig. 3 illustrates the relationships between domains and DNS providers that we group as "leading", "non-leading", and "intra-government" dependencies. While some general domains have implemented a multi-provider strategy, possibly to mitigate risks associated with a single, critical dependency, the practice is not widespread. It is particularly noteworthy that it is absent for domains for the indigenous population.

Single-provider setups We first investigate how many domains rely on a single DNS provider, which is a critical metric: outage of this provider will make the relying services unavailable. We find that 92% of all domains for the general population rely on a single provider. *All* of the domains for the indigenous populations do so. This implies a generally unsatisfactory state across all government domains, but it is also a first hint that there is a difference between the services for the two population groups.
Multi-provider setups Having multiple DNS providers offers benefits in terms of redundancy and resilience. In the event of a service outage or disruption from one provider, the availability of DNS services can be maintained through the alternative provider. Inequalities in the use of multi-provider strategies hence reflect differences in access to information and online services. Fig. 1 shows the distribution of domains with a multi-provider dependency for the general

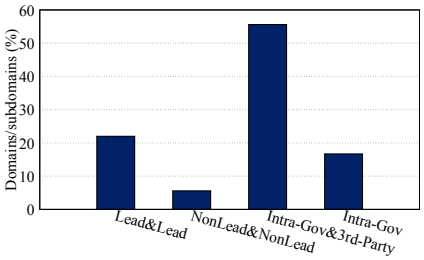

Fig. 1. Multi-DNS-provider setups. Note that no domains for the indigenous population use such a setup.

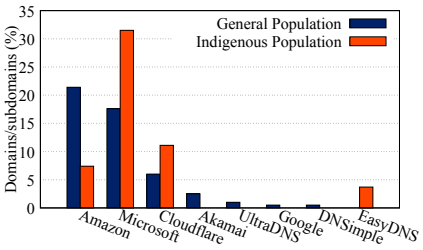

Fig. 2. Leading DNS providers.

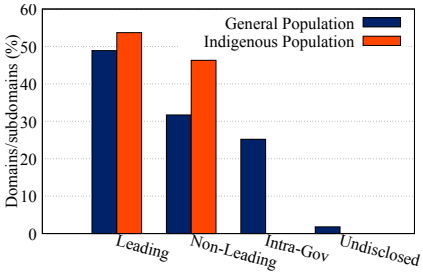

Fig. 3. DNS providers by category.

population (none of the indigenous websites have multiple DNS providers). We find that 20% of setups have a dependency on two distinct leading DNS providers (Amazon and Microsoft); this was observed for eight domains of the Victorian government. More than 50% of setups use a governmental provider along with a third-party DNS as an alternative server.

Use of leading providers Approximately half of the domains for both the general and indigenous populations rely on a single leading DNS provider. Only around 2% of the domains for the general population employed *two leading* providers, with the remainder using either a second non-leading or intra-government provider. Fig. 2 shows a breakdown of the leading DNS providers for our domains. For the general population, 48.9% of domains rely on leading providers, with Amazon being the most utilized provider at 21.4%. Microsoft is the second most commonly used provider at around 17%, followed by Cloudflare at 6%. Other leading providers, such as Akamai, UltraDNS, Google, DNSimple, and EasyDNS, are used in less than 5% of DNS services for the general population. Regarding domains for the indigenous population, 53.7% of them rely on leading providers. Microsoft is the most utilized provider at 31.5%, followed by Cloudflare (11.1%) and Amazon (7.4%). No other leading providers are in use for these domains. Comparing the two groups of domains, we identify a common preference for leading providers, although the preferred providers differ starkly. Cloudflare offers a free tier, which may explain this common choice in the second group of domains. There is slightly less variety in the chosen providers in the case of the domains for the indigenous population.

Use of non-leading providers and intra-government providers: As we see in Fig. 3, slightly more than half of domains for the general population rely on at

least one non-leading provider or an intra-government provider, with an almost equal split between the latter two. We do not observe this for the domains of the indigenous population: here, 46.3% of the domains rely on non-leading providers, and none use intra-government providers. Government-hosted providers would be required to comply with Australian standards and government regulations, and using these providers implies a certain level of coordination and collaboration. While we observe only about 15 government agencies operating name servers, we see that they serve well over 100 different domains. It seems curious that no single service for the indigenous population is among these.

4 Limitations

Our study has several limitations owing to the early stage of our work. We did not investigate indirect dependencies yet, for example, and our observations are not yet longitudinal. There are significantly fewer domains for the indigenous population than for the general population, so one needs to pay attention when comparing small percentages between the groups.

5 Implications

We set out to identify possible disparities in the DNS dependencies for sites for different population groups. We find evidence that dependencies for the indigenous population are indeed differently configured, and we view our evidence as indicative of different provisioning concepts being employed. However, the exact implications of this are much less clear. In particular, does this result in a tangible digital divide? It seems clear to us that follow-up measurements will be needed to decide this question. The lack of intra-government provisioning for indigenous population domains is noteworthy, but there may be practical or legal reasons why we do not find such setups. A qualitative study could shed light on this. As single-provider setups are so common, it is too early to speak of a digital divide in terms of availability. In particular, it is unclear whether intra-government provisioning or the use of smaller domestic providers will improve availability, which can be decided with Internet measurements. Cloudflare was a more common choice among the leading providers in the case of domains for the indigenous populations (possibly because of their free tier). Together with the fact that over 40% of indigenous domains use domestic DNS providers, this may indicate a desire to improve DNS resolution but an inability or unwillingness to move to the cloud. Again, a qualitative study could help illuminate this.

We argue that it is a worthwhile undertaking to add measurements of digital divides to the agenda, using both quantitative and qualitative methods. In addition to investigating DNS dependencies, we recognize the significance of considering other measurements that might contribute to a comprehensive assessment of the digital divide. These include availability measurements by using datasets such as Common Crawl [1] or OONI (Open Observatory of Network Interference) [2], routing measurements, and measuring the use of web content

management systems. In the future, we need to qualitatively assess the criticality of services for different population groups and explore the correlation between popularity and criticality.

References

1. Common Crawl. https://commoncrawl.org/ (2023), accessed on August 17, 2023
2. OONI Data. https://ooni.org/data/ (2023), accessed on August 17, 2023
3. Nazemi, N., Tavallaie, O., Zomaya, A.Y., Holz, R.: DNS dependencies as an expression of the digital divide: the example of Australia. In: Workshop on Transparency, Accountability and User Control for a Responsible Internet (TAURIN 2023). Den Haag, Netherlands (10 2023)
4. Rogers, E.M.: The digital divide. Convergence **7**(4), 96–111 (2001)
5. U.S. Department of Commerce: Falling through the net: A survey of the" have nots" in rural and urban america (July 1995), https://ntia.gov/page/falling-through-net-survey-have-nots-rural-and-urban-america
6. U.S. Department of Commerce: Falling through the net ii: New data on the digital divide (July 1998), https://ntia.gov/page/falling-through-net-ii-new-data-digital-divide
7. U.S. Department of Commerce: Falling through the net: Defining the digital divide (July 1999), https://ntia.gov/report/1999/falling-through-net-defining-digital-divide
8. Van Dijk, J.: The digital divide. John Wiley & Sons (2020)