Beneath the Divide: A Measurement Study of Minority-Serving Web Infrastructure

Bulut Ulukapi[1], Anna Sperotto[1], and Ralph Holz[1,2]

[1] University of Twente
[2] University of Münster

Abstract. Minority-serving websites, including those serving LGBT+, refugee, and religious minority communities, depend on Internet infrastructure to deliver services and maintain public visibility. These organizations frequently operate under political pressure and with limited resources, yet the infrastructure beneath them has not been measured. We present a multi-country, multi-dimensional measurement study of the hosting, DNS, and email dependencies of 438 minority-serving organizations across seven countries spanning distinct regulatory regimes and global regions, namely Australia, the Netherlands, Hungary, Turkey, India, the United States, and South Africa. We compare them against a baseline of national government web services in the same countries, which face the same domestic provider landscape and regulatory regime under materially fewer resource constraints. We use the term digital divide in a broader sense than its common framing as a client-side accessibility gap, extending it to the supply side.

We find that the divide begins before infrastructure does. 11.9% of these organizations operate no independent domain at all, rising to 24.3% of LGBT+ organizations and 61.1% of informal and grassroots initiatives, leaving them nothing of their own to configure and no portable identity to move. Among the organizations that do, the three dimensions concentrate in different shapes. Email is the most concentrated, with two US providers handling 68.3% of organizations and a majority in every country. One provider leads authoritative DNS in six of the seven countries. Hosting is the only dimension in which domestic providers retain a substantial share, and even there the same small group of US providers operates in every country. These dependencies compound, since one commercial entity operates at least two of the three for 49.4% of organizations. Against the government baseline, the two sets depend on different providers rather than on the same providers to different degrees, and the contrast is sharpest in email. The exception is DNS delegation, where 98.1% of minority-serving and 87.1% of government domains place every nameserver with a single operator, so being better resourced changes which provider an organization relies on but not that it relies on only one.

Keywords: Internet measurement · Infrastructure dependency · Digital divide · Provider concentration

1 Introduction

A website is the part of an organization's infrastructure that the public sees. What it rests on is not. Hosting, authoritative DNS, and mail are each procured from third parties, and those relationships determine whether the site can be reached, whether the name resolves at all, and whether mail arrives. For organizations serving LGBT+, refugee, disability, and religious and ethnic minority communities, these are the channels through which services are delivered and public visibility is maintained, and they are chosen under tighter constraints than most. What these organizations actually run on has not been measured.

Two literatures bear on the question. Internet measurement has documented the consolidation of the general web and of government digital services around a small number of hyperscalers, raising concerns about resilience, provider concentration and cross-border exposure [6,12,14,18]. That work studies well-resourced institutional actors whose procurement is formalized and publicly accountable. Research on digital inequality and on at-risk users studies minority communities closely, but takes the individual as its unit, characterizing who can get online and with what skills rather than what the organizations serving them depend on. We therefore use the term digital divide in a broader sense than its common framing as a client-side accessibility gap, extending it to the supply side as a disparity in the infrastructure conditions available to minority-serving websites relative to mainstream ones. Whether such a divide exists is an open empirical question, and three asymmetries suggest it might. Limited budgets may drive concentration toward affordable or free-tier providers by default. Differing jurisdictions and data protection frameworks may produce observable differences in infrastructure choice. And political pressure makes many of these organizations, particularly grassroots and civil society initiatives, more likely targets of censorship, takedown requests or adversarial interference, which raises the stakes of any fragility beyond operational inconvenience.

We study seven countries spanning distinct regulatory regimes and political contexts, and examine three dimensions of Internet infrastructure, namely hosting, DNS and email, which together underpin the core functions of any web service. We address three research questions. RQ1 asks what the hosting, DNS and email dependencies of minority-serving websites are, and how they vary across countries. RQ2 asks whether these websites exhibit different provider landscape and geographic exposure from national government web services in the same countries. RQ3 asks to what extent these services are delivered from foreign jurisdictions, and how that interacts with the applicable regulatory regime. We additionally examine provider stacking, the co-occurrence of two or more of these dependencies on a single commercial entity.

2 Background and Related Work

Digital inequality and at-risk users. Research on the digital divide has moved through three levels, from inequalitiesin physical access, to inequalities

in skills and use, to inequalities in the outcomes people derive from being online [9, 21, 25, 26], and a consistent finding is that ethnic minorities and other historically marginalized groups fall on the disadvantaged side of each [5, 27]. A complementary literature in security and human–computer interaction studies at-risk users, populations facing elevated or disproportionate digital-safety threats, including LGBTQ+ people, refugees, activists and survivors of violence. It identifies recurring contextual risk factors such as stigmatization and political targeting [3, 28], reviews the privacy needs of marginalized groups [20], and distills practical guidance from sustained engagement with these communities [17]. Both take the individual as the unit of analysis, characterizing who can get online, with what skills, and at what risk. Organizations enter only as attack targets, in traced campaigns against a human-rights NGO [16] and in practitioner programmes that treat civil society as targeted critical infrastructure and supply it with protective hosting [4], never as operators of infrastructure of their own. Rather than asking whether minority users can get online, we ask what the services built for them depend on.

Consolidation of Internet infrastructure. A large measurement literature documents the consolidation of the general web. Popular websites depend critically on a handful of third-party DNS, CDN and certificate providers, such that a single provider failure can disrupt a large share of them [6, 13], and DNS traffic and hosting show the same pattern, with a share that has grown over time [2, 18, 29]. Consolidation is not uniform worldwide, since third-party dependency varies widely across 50 countries [15], and a formalized measure of centralization applied to 150 countries uncovers pronounced geographic variation together with sociopolitically driven regionalization [8]. Closest to our setting, work on government and public-interest services reveals reliance on few authoritative DNS providers and frequent single-provider dependencies [11, 23], and heavy dependence on transnational hosting with services routed through foreign jurisdictions [12, 14], alongside related exposure in privacy [19], HTTPS adoption [22] and resource cost [7]. Our earlier study extends this to the municipal tier, where email is the most concentrated dimension and provider stacking collapses otherwise independent services onto a single operator [1].

These two bodies of work leave a gap on both sides. The first studies minority communities but takes the individual as its unit. The second measures infrastructure but studies the general web and well-resourced institutional actors whose procurement is formalized and publicly accountable. Minority-serving organizations differ on exactly those dimensions, with limited budgets, informal or grassroots operation and, in several of our countries, political pressure, so their dependency structure cannot be assumed to mirror either. To our knowledge no prior work measures their hosting, DNS and email dependencies, nor compares them against a resourced reference population.

3 Methodology

3.1 Dataset Construction

Table 1. Country overview. [24].

Country	Regulatory regime	Population (M)	Region
Netherlands	GDPR	18.0	W. Europe
Hungary	GDPR	9.7	E. Europe
Turkey	Non-GDPR	86.0	SE Europe/W. Asia
India	DPDP Act	1428.6	South Asia
United States	Non-GDPR	335.9	North America
South Africa	POPIA	60.4	Sub-Saharan Africa
Australia	Privacy Act	27.4	Oceania

This section describes how the minority-serving domain dataset was constructed and reports its composition and coverage. Organizations that operate a resolvable, independent domain constitute the measurement-eligible set and admit DNS, hosting, and email measurement. Those that do not are reported separately in Section 4, since their absence from the measurement is itself a property of the population we study.

3.1.1 Community and Country Frame We organize the dataset along two dimensions. The first is the community category, of which we use six, namely LGBT+, disability, refugees and asylum seekers, ethnic minorities, linguistic minorities, and religious minorities. These span distinct bases of minority status rather than sampling a single one in depth, and they correspond to the remits of established international registries and equality bodies, which makes systematic enumeration feasible where ad hoc thematic groupings would not. They are unevenly distributed across the seven countries, with Australia holding 23 of the 68 disability organizations against 5 in the Netherlands, so comparing the infrastructure dimensions between categories would confound community with national provider landscape. We therefore use the categories to describe the composition of the dataset and the availability of an independent domain, and report the dimensions by country.

The second is the country. We study seven countries chosen to represent distinct regulatory regimes and political contexts, listed in Table 1, spanning GDPR and non-GDPR jurisdictions across five world regions and populations from 9.7 to 1428.6 million, and including cases where formal regulatory guarantees and political reality diverge. Each organization is assigned to exactly one country and one category. The few that plausibly span categories were assigned to their primary self-described community to preserve a partition.

3.1.2 Source registries and provenance We compiled organizations from three tiers of sources and recorded the originating source for every entry. The first tier is six supranational registries whose membership lists provide cross-national coverage across the LGBT+, disability, refugee, ethnic and linguistic minority categories, and which are most complete for the European countries in our frame. The second is national peak-body directories and equality or statutory-body registries, used where supranational coverage is thin, particularly outside the EU, and including bodies recognized under national law such as minorities notified

under India's National Commission for Minorities Act, and South African Chapter 9 institutions. The third is targeted manual curation via directory and web search for Turkey, India and South Africa, where centralized registries offer limited coverage, and to raise the population of under-represented country–category cells.

3.1.3 Domain identification and verification For each organization we sought its primary independent domain, meaning a registered domain operated on the organization's behalf rather than a presence maintained solely on a third-party platform, such as a social-media account or a shared subdomain like `*.wordpress.com`. The distinction is material, since an organization without one has no DNS, hosting or email configuration of its own to measure. We treat a domain as confirmed when observed during curation on the organization's own site or in an authoritative directory entry, and as provisional when identified from secondary evidence but not directly re-confirmed.

3.2 Dataset Composition

The dataset comprises 438 organizations across the seven countries and six categories. Of these, 386 (88.1%) operate a resolvable independent domain and are measurement-eligible, with the domain confirmed for 358, provisional for 27, and one flagged as resolving to a parked page. The remaining 52 organizations (11.9%) have no independent domain.

3.2.1 Operator type The dataset is predominantly non-governmental. Registered NGOs account for 403 organizations (92.0%), followed by informal or grassroots initiatives at 18 (4.1%), government-adjacent bodies at 13 (3.0%), and government or government-related bodies at 4 (0.9%). The informal category is small, but it behaves distinctly in our results, so we retain it as a separate stratum rather than folding it into the NGO group.

3.2.2 Independent domain availability The share of organizations operating an independent domain ranges from 77.6% in Turkey to 97.6% in the United States, and from 75.7% for LGBT+ organizations to complete in three of the six categories. Because organizations without one are by construction absent from the measurement-eligible set, and because availability co-varies with both country and category, we treat these as descriptive properties here and defer interpretation to Section 4.1.

3.2.3 Baselines. As a reference population we use national government web services in the same seven countries, 60 per country for 420 domains, drawn from national government registries and filtered to one domain per distinct national public body. They face the same domestic provider landscape, national registry and regulatory regime under materially fewer resource constraints, so differences between the two sets bound how much of the observed fragility belongs to the national environment and how much to the position of these organizations within it. Since our dataset is 92.0% non-governmental, the comparison spans sector as well as minority-serving status and establishes a resourced upper bound rather than isolating a single variable. It also supplies a floor for the finding of Section 4, since operating a domain is statutory for these bodies rather than a discretionary expense.

3.3 Anycast prevalence and geolocation validity

Geolocating an address presupposes that it has a single location. Anycast breaks that assumption, since one address is announced from many sites at once and an IP-to-country lookup returns a value defined only relative to the vantage point. We therefore classify every distinct address as anycast or unicast using LACeS [10], separately for the hosting (A), mail (MX) and nameserver (NS) addresses, before geolocating any of them.

The three dimensions differ by more than an order of magnitude (Table 2). Mail is almost entirely unicast at 3.0% anycast, with four of the seven countries showing none at all, hosting is just under half at 47.3%, and DNS is 76.3%, exceeding 80% in Australia, India and the United States.

Three consequences follow. We report no serving country for anycast addresses and compute geographic exposure over unicast addresses only, while retaining provider attribution for them, since the announcing organization remains well defined where the serving site does not. The coverage this leaves is effectively the whole population for mail, just over half for hosting and under a quarter for DNS, so a jurisdictional reading is firm for mail, weaker for hosting and not supportable for nameservers, where we rely on attribution alone and treat any gap between attribution and geolocation as unresolvable rather than as evidence of offshore operation. And because prevalence is itself uneven across countries, from 28.9% to 72.0% for hosting and 40.5% to 91.2% for DNS, a naive cross-country comparison of geolocated addresses would set populations of differing coverage against one another.

3.4 Measurements and Provider Attribution

We resolve MX, NS, and A/AAAA records per domain via Cloudflare and Google public resolvers, validated with distributed RIPE Atlas vantage points, and enrich results with ASN and geolocation data from IPinfo. Provider attribution follows a combined automated and manual procedure across hosting, email, and authoritative DNS, classifying each provider as private, government-related, or unknown, with domestic vs. foreign sub-classification. Attribution relies on

Table 2. Anycast prevalence among distinct addresses.

	Hosting		Mail		DNS	
Country	Addr.	Anycast	Addr.	Anycast	Addr.	Anycast
United States	100	72.0%	86	0.0%	283	91.2%
Australia	97	45.4%	73	1.4%	226	85.4%
India	68	48.5%	41	9.8%	181	83.4%
Turkey	42	40.5%	49	18.4%	109	68.8%
Netherlands	71	31.0%	89	0.0%	154	62.3%
Hungary	41	41.5%	41	0.0%	87	54.0%
South Africa	38	28.9%	93	0.0%	74	40.5%
Total	457	47.3%	472	3.0%	1114	76.3%

Table 3. Hosting infrastructure by country. Proxy = domains behind a reverse proxy or CDN; Hyper. = domains on Amazon, Google, Microsoft, Cloudflare, or Oracle.

Country	Dom.	Prov.	Single-IP	Domestic	Proxy	Hyper.
United States	81	19	43.2%	93.8%	53.1%	59.3%
Australia	80	26	65.0%	26.2%	27.5%	48.8%
Netherlands	56	31	78.6%	35.7%	14.3%	23.2%
India	53	22	50.9%	7.5%	24.5%	39.6%
Turkey	36	23	75.0%	36.1%	19.4%	19.4%
Hungary	35	18	77.1%	60.0%	22.9%	25.7%
South Africa	35	18	88.6%	65.7%	14.3%	20.0%

MX hostname patterns, SMTP banner collection, reverse DNS lookups, and CHAOS-class TXT queries, with ambiguous cases resolved manually.

4 Results

4.1 Hosting dependencies

We resolve A records for the 386 measurement-eligible domains, of which 376 return an address and are measured here, attribute each address to a provider, and geolocate the unicast subset. Most countries retain a domestic hosting sector, but the same small set of foreign providers sits above it in every one (Table 3).

A foreign provider group is common to every country. By provider registration, domestic hosting runs from 7.5% of Indian domains to 93.8% of U.S. ones, yet the same US-registered providers, chiefly Cloudflare, Amazon and Google, lead the foreign sector everywhere. Between 19.4% and 59.3% of domains per country sit on one of these hyperscalers, and the United States is the largest single foreign dependency for all six other countries, most pronounced in India at 70% and Australia at 68%. The pattern holds even where a domestic sector leads, as in South Africa, where domestic providers host 65.7% of domains yet 26% still depend on the United States. Registration, however, records the nationality of the operating company rather than where a site is served or whose law governs it, so we treat it as a first indication and turn to serving location.

What an address can and cannot show. Between 14.3% of domains in the Netherlands and South Africa and 53.1% in the United States resolve to a reverse proxy or CDN, most often Cloudflare and, in the United States, Fastly, for which the address identifies the front end rather than the origin. A further 0% to 13.6% sit on website builders such as Wix, Squarespace and GoDaddy, where no infrastructure choice exists below the platform, placing them between the domains analysed here and the platform-delegated organizations below.

Where the servers actually are. For the unicast A records, whose location is well defined, we additionally geolocate the serving address (Table 4). In every country a larger share of domains is served domestically than registration indicates, by 33 percentage points in Australia and 24 in India, and negligibly

in the United States, whose providers are domestic in both senses. The gap is largest where dependence on the foreign group is greatest.

Organizations with no independent domain The dependencies above presuppose infrastructure of one's own to measure, and 52 of the 438 organizations have none, their public presence delegated in full to a third-party platform, whether a social network, a page on a parent organization's site, or a hosted publishing service. The subset is overwhelmingly one category. LGBT+ organizations account for 46 of the 52 cases, so 24.3% of all LGBT+ organizations in the dataset have no independent domain, against 8.6% (5/58) for ethnic minorities and 2.4% (1/42) for linguistic minorities, while disability (0/68), refugees and asylum (0/41) and religious minorities (0/40) contain no cases at all. The pattern also tracks organizational form, at 61.1% (11/18) among informal or grassroots organizations against 9.7% (39/403) for registered NGOs, and it persists within the LGBT+ category taken alone, at 64.7% (11/17) against 19.9% (34/171), so form and community contribute separately rather than one standing in for the other.

Where an organization with its own domain concentrates risk in a provider it has chosen and can in principle leave, carrying its address and its correspondents with it, an organization without one has no portable identity to move. Deplatforming, policy change or account suspension removes its entire public presence, with no DNS record to repoint, and it falls hardest on exactly the communities most exposed to content moderation disputes and coordinated reporting campaigns.

4.2 Email dependencies

We resolve MX records for the measurement-eligible domains that operate mail, attribute each mail exchanger to a provider, and geolocate its serving address. Email is the most concentrated of the three dimensions, and the concentration takes the same form everywhere, two US providers serving the majority of organizations (Table 5).

Two providers carry most of the mail. Microsoft and Google together handle the mail of 237 of 347 organizations, or 68.3%, and their combined share exceeds half in every country, from 51% in India to 79% in Australia. Which of the two leads is the main national difference rather than whether they lead. No

Table 4. Domestic share of hosting by provider and by serving location. Reg. = provider in the organization's country; Served = address geolocated to it.

Country	Rec.	Reg.	Served	Largest foreign serving loc.
United States	28	93.8%	96.4%	Canada (4%)
South Africa	28	65.7%	82.1%	United States (11%)
Hungary	26	60.0%	80.8%	Germany (12%)
Australia	54	26.2%	59.3%	United States (31%)
Netherlands	49	35.7%	55.1%	Germany (18%)
Turkey	27	36.1%	44.4%	US / Germany (15% each)
India	35	7.5%	31.4%	United States (26%)

Table 5. Email infrastructure by country. Big two = domains whose mail is handled by Microsoft or Google; Self = domains self-hosting mail.

Country	Dom.	Leader	Big two	Self	Reg.	Served
Australia	77	MS 64%	79%	1.3%	8%	75%
United States	78	Goog 44%	77%	0.0%	97%	95%
South Africa	32	MS 59%	66%	15.6%	34%	56%
Netherlands	54	MS 46%	65%	0.0%	24%	87%
Hungary	31	Goog 45%	65%	3.2%	32%	32%
Turkey	34	Goog 47%	56%	2.9%	26%	26%
India	41	Goog 37%	51%	0.0%	12%	27%

third provider approaches them, the largest non-hyperscaler being Mimecast in the United States at 5%. Unlike hosting, where a domestic group persists beneath the foreign leaders, email offers little domestic alternative, and self-hosting has largely disappeared, remaining above 4% only in South Africa at 15.6%.

Registration understates domestic service, but not everywhere. Because mail is almost entirely unicast, only 3.0% of MX addresses against 47.3% of hosting addresses (Section 3), we geolocate the serving exchanger directly rather than infer jurisdiction from the provider's nationality. Doing so does not uniformly soften the registration picture. Where the two providers answer from in-country infrastructure the correction is large, moving Australia from 8% domestic to 75%, the Netherlands from 24% to 87% and South Africa from 34% to 56%. Where they do not it is absent, and Turkey and Hungary stay at 26% and 32% on both measures, with India partial at 12% to 27%. The United States is domestic either way. The gap is thus not a constant offset but a marker of where a provider has placed regional infrastructure, and organizations on the wrong side of that choice remain served from outside their borders regardless of who their provider is.

Serving location follows the provider, not the organization. Where mail is served in Europe, very few operators serve it. Mail answered in Ireland is entirely Microsoft, in Germany 90% Google, and in the Netherlands 81% one of the two, and across our whole sample Microsoft and Google records are served from these three countries 48.4% of the time against 17.0% for all other providers combined. The Netherlands is the largest foreign serving location in four of the seven countries, including for U.S. organizations. Organizations are therefore not choosing a jurisdiction so much as inheriting whichever regional infrastructure their provider operates, which changes where the processing happens without changing which company performs it.

4.3 DNS dependencies

We resolve NS records for the 386 measurement-eligible domains, attribute each nameserver to a provider, and classify each of its addresses as anycast or unicast. Twelve domains return no nameserver, leaving 374 measured (Table 6).

One provider leads six of the seven countries. Cloudflare is the largest authoritative DNS provider in Australia, Hungary, India, the Netherlands, Turkey

and the United States, holding between 17.9% of Dutch domains and 35.8% of U.S. ones. South Africa is the exception, where the domestic operator xneelo leads at 32.4% against Cloudflare's 8.8%. Across the measured set Cloudflare serves 25.4% of domains and the next largest provider, GoDaddy, 8.8%, so the first and second operators stand further apart than the second does from the long tail behind it. The five providers we group as hyperscalers account for between 17.6% of South African domains and 39.5% of U.S. ones, and Cloudflare supplies most of that share in every country but South Africa, where it supplies half.

A domestic sector survives in some countries and not in others. Domestic attribution ranges from 9.3% of Indian domains to 92.6% of U.S. ones and reaches 53.2% across the measured set, remaining the majority position in four countries, the United States, Hungary, South Africa and the Netherlands. The United States is the largest foreign attribution for every country but itself, from 26.5% of Hungarian and South African domains to 74.1% of Indian ones, while no foreign operator reaches 4% of U.S. domains. Nameservers operated in house appear for two South African domains and nowhere else, and we count them as domestic. We attribute each nameserver to the company operating the service rather than to the network announcing it.

Serving location is not interpretable for nameservers. 238 of the 374 domains, 63.6%, are served entirely from anycast addresses, and none in this population mixes anycast and unicast nameservers, against 39 of the 418 government domains that do. Geolocation is therefore defined for 136 domains, covering between a tenth and three quarters of each country's population, and it is thinnest where concentration is highest, because the operators holding the largest shares announce from anycast prefixes and leave the located subset first. Comparing located addresses across countries would set one country's residual domestic operators against another's rather than against the populations of Table 6. We report no serving-location column for this dimension and rest it on attribution alone.

Delegation concentrates control over whether the name resolves at all. Hosting and mail can fail in part, since a proxied site still resolves and mail answered abroad still arrives. Authoritative DNS admits no such degradation, because a name that is not answered for does not exist for any purpose, and the hosting and mail dependencies reported above require it to resolve first. The concentration reported here therefore identifies one commercial actor per organization in the great majority of cases, whose withdrawal of service or compliance with an order directed at it removes the organization from the Internet in full rather than in part. For 53.2% of organizations that actor is incorporated in the organization's own country. For the remainder, and most sharply for the Indian and Australian organizations of which 74.1% and 55.1% respectively delegate to a U.S.-registered operator, it is incorporated elsewhere, and our measurement shows where an operator is registered and not which law reaches the delegation.

4.4 Cross-dimensional provider stacking

Table 6. Authoritative DNS by country. Prov. = distinct nameserver operators observed; Hyper. = Amazon, Google, Microsoft, Cloudflare or Oracle.

Country	Domains	Prov.	Domestic	Hyper.	Single	US
United States	81	24	92.6%	39.5%	96.3%	—
Australia	78	26	30.8%	38.5%	98.7%	55.1%
Netherlands	56	32	57.1%	17.9%	100.0%	26.8%
India	54	19	9.3%	37.0%	100.0%	74.1%
Turkey	37	20	37.8%	24.3%	97.3%	54.1%
Hungary	34	17	73.5%	20.6%	100.0%	26.5%
South Africa	34	16	70.6%	17.6%	94.1%	26.5%

Considered jointly, the three dimensions reveal provider stacking, meaning the same commercial entity operating more than one of an organization's three services. We report it over the 322 organizations whose hosting, mail and authoritative DNS are all attributed to an identifiable commercial provider (Table 7).

Half of the organizations we can measure depend on one provider for at least two dimensions. 159 of the 322 organizations, 49.4%, have at least two of their three services operated by the same entity, from 37.5% in South Africa to 67.7% in Turkey. All three rest on one provider for 6.2% overall, reaching 18.9% in India and falling to zero in South Africa. These country shares rest on populations of 24 to 78 organizations, so one organization moves a South African figure by more than four percentage points and the ordering should be read with that in mind.

Stacking runs through hosting and DNS rather than through email. The hosting and DNS pair accounts for 131 of the 159 stacked organizations, is the most common pair in every country, and alone covers more than a third of organizations in five of them. The two pairs involving mail are far rarer, at 12.1% for mail with DNS and 9.3% for hosting with mail. This follows from the structure reported above, since Microsoft and Google carry the mail of more than two thirds of these organizations but lead neither other dimension, so the provider handling an organization's mail is usually not the provider handling anything else. Concentration in email is therefore high but largely self-contained, whereas concentration in hosting and DNS compounds.

One provider accounts for most of the stacking we observe. Cloudflare appears in the stacked provider set of 64 of the 159 organizations, more than seven times the next entity, Microsoft, at nine, and excluding it the share with any stacking falls from 49.4% to 29.5%. This reflects how the service is delivered rather than a repeated procurement decision. Where a domain is proxied, the address observable from DNS belongs to Cloudflare whatever the origin, and proxying is sold alongside authoritative DNS as one product, so these organizations have not chosen the same company twice but one service spanning two of our dimensions. We report the figure as observed, noting that for them the two dependencies were never independent.

Table 7. Provider stacking. H = hosting, E = email, D = authoritative DNS, Any counts organizations with at least one matching pair rather than pairs.

Country	Org.	Any	H+E	H+D	E+D	All three
United States	78	51.3%	6.4%	42.3%	5.1%	1.3%
Australia	75	48.0%	6.7%	40.0%	9.3%	4.0%
Netherlands	48	41.7%	8.3%	37.5%	12.5%	8.3%
India	37	56.8%	24.3%	48.6%	21.6%	18.9%
Turkey	31	67.7%	16.1%	51.6%	25.8%	12.9%
Hungary	29	41.4%	6.9%	31.0%	10.3%	3.4%
South Africa	24	37.5%	0.0%	29.2%	12.5%	0.0%
Total	322	49.4%	9.3%	40.7%	12.1%	6.2%

A stacked organization has one relationship to lose rather than three. Where services are spread across operators, each dependency can be renegotiated or terminated on its own terms and the loss of any one is partial. Suspension or withdrawal is generally decided at the level of the account rather than the service, so one entity holding two dimensions reaches both, and holding all three reaches an organization's public presence, its correspondence and its name resolution together. Read alongside the near-universal single delegation of Section 4.3, this leaves a substantial part of the population with one commercial relationship carrying the whole of their online presence.

5 Discussion

5.1 Comparison with government services

The two sets differ in kind rather than in degree. Hosting is domestic for 47.3% of minority-serving domains against 79.7% of government ones, and hyperscaler dependency in DNS runs the other way at 30.5% against 16.4%. Mail separates them furthest. Microsoft and Google handle 68.3% of minority-serving organizations and a majority in every country, from 51% in India to 79% in Australia, against 25.3% of the government domains that publish an exchanger, from 0.0% in India and Turkey to 67.3% in the United States. The minority-serving side also carries a common dependency the baseline does not, since Cloudflare leads hosting in six of the seven countries, from 14.3% of domains in the Netherlands and South Africa to 43.2% in the United States. The two sets are not at different points on one distribution of providers. Only in the United States, where both hosting and mail converge, do they substantially overlap.

The gap is widest where the minority-serving population is most foreign dependent. Domestic hosting reaches 96.7% of Indian government domains against 7.5% of Indian minority-serving ones, a difference of 89 percentage points, with Turkey at 100.0% against 36.1% and the Netherlands at 85.0% against 35.7%. It narrows in Hungary and South Africa and closes almost entirely in the United States, where both sets are domestic because the providers

they use are American. Australia inverts it, government hosting sitting at 18.3% against 26.2%, and the inversion is confined to hosting, since Australian government mail is 65.0% domestic against 8%. The ordering follows neither national wealth nor the resources of either set, so we read the baseline as a resourced reference population rather than as isolating a single variable.

Delegating to one operator is the measure on which the two sets agree. 87.1% of government domains place every nameserver with one provider against 98.1% of minority-serving domains, and in six of the seven countries the baseline sits between 91.7% and 100.0%. The Netherlands is the sole exception at 35.0%, and the arrangement there is uniform rather than incidental, since 34 of the 60 domains delegate to the same three operators and one of the three is registered outside the Netherlands. Resources change which operator the delegation lands on. In six of the seven countries they do not change that it lands on one.

The comparison is asymmetric in observability, in the same direction as the difference. Minority-serving domains sit behind a content delivery network far more often than government domains, from 14.3% against 3.3% in the Netherlands to 53.1% against 38.3% in the United States, and are more often anycast in every dimension, 47.3% against 14.5% for hosting and 76.3% against 52.5% for DNS. A proxied or anycast address identifies the front end or the announcing operator rather than the origin, so our domestic share for the minority-serving set is a lower bound in a way the government figure largely is not, and part of every gap in Table 8 is visibility rather than infrastructure. Coverage is asymmetric the same way, since an own nameserver delegation was a condition of entry to the control set, so the baseline cannot exhibit the delegation failure that leaves twelve minority-serving domains unmeasured. We therefore read the comparison as bounding how concentrated a resourced population is and not how reliably it can be reached. Minority-serving websites do exhibit different provider concentration and geographic exposure from national government services in the same countries, and the difference is widest in mail and narrowest in the number of operators a domain delegates to.

5.2 Jurisdiction and the regulatory regime

The regime does not order the outcome, and the widest divergence falls inside a single regime. Of the seven regimes, only GDPR operates a destination list within its data protection framework. Turkey adopted an adequacy requirement in 2024 (Law No. 7499, amending Art. 9 of Law No. 6698) and has recognised no country under it. Australia may, since December 2024, prescribe countries whose law is substantially similar (Privacy Act 1988 (Cth), APP 8, as amended 2024), and has prescribed none. South Africa conditions transfer on the recipient being bound to substantially similar protection, assessed case-by-case (POPIA s. 72), with no adequacy findings made. India permits transfer to any country not notified as restricted (DPDP Act 2023, s. 16), and none had been notified during our measurement. The United States has no general data protection transfer rule. Were the regime governing placement, the countries whose destination rules operate in practice would show the most domestic service. They

Table 8. Minority-serving (M) vs government services (G) in the same country. Big two = mail handled by Microsoft or Google.

	Hosting		Mail		DNS			
	Domestic		Big two		Domestic		Single	
Country	M	G	M	G	M	G	M	G
Domains	376	418	347	383	374	420	374	420
India	7.5%	96.7%	51%	0%	9.3%	98.3%	100.0%	98.3%
Turkey	36.1%	100.0%	56%	0%	37.8%	91.7%	97.3%	93.3%
Netherlands	35.7%	85.0%	65%	25%	57.1%	95.0%	100.0%	35.0%
Hungary	60.0%	84.5%	65%	14%	73.5%	90.0%	100.0%	95.0%
South Africa	65.7%	78.3%	66%	34%	70.6%	88.3%	94.1%	91.7%
United States	93.8%	95.0%	77%	67%	92.6%	66.7%	96.3%	96.7%
Australia	26.2%	18.3%	79%	30%	30.8%	50.0%	98.7%	100.0%
Total	47.3%	79.7%	68.3%	25.3%	53.2%	82.9%	98.1%	87.1%

do not. Australia and Turkey, with opposite regime designs and equally empty destination lists, sit 49 percentage points apart on mail served domestically, at 75% against 26%. The Netherlands and Hungary, under the same instrument and the same adequacy decisions, sit 55 points apart on the same measure and invert on DNS. Within-regime variation exceeds between-regime variation, on the pairs our frame admits.

What tracks the outcome is the provider's regional footprint. The correction from provider registration to serving location reported in Section 4.2 is large for Australia, the Netherlands and South Africa, partial for India, and absent for Turkey and Hungary, where neither dominant provider answers mail in-country. The government baseline corrects in the same places and not in others, and Hungary is the clearest case, where mail stays at 32% for minority-serving domains and 89.7% for government ones on both measures. Two populations that share almost no mail providers diverge from registration in the same countries. We rest this on mail alone, since it is the dimension for which anycast leaves the serving location defined for effectively the whole population, and it is the strongest evidence we have that the determinant of jurisdictional exposure is the provider's infrastructure map.

Limitations Our dataset is enumerated from registries and directories of uneven completeness rather than sampled, most thinly outside Europe, so it describes the organizations we could identify rather than a defined population, and because the six categories are distributed unevenly across the seven countries we report the dimensions by country only, leaving any variation between communities inside a country unobserved. The domestic shares we report are lower bounds, since between 14.3% and 53.1% of domains resolve to a reverse proxy or CDN and anycast removes serving location for 47.3% of hosting and 76.3% of nameserver addresses, and the shortfall is largest where dependency on

large providers is greatest. Provider attribution identifies the entity operating each service and rests in part on manual resolution of ambiguous cases.

6 Conclusion

The divide begins before infrastructure does. A substantial minority of these organizations operate no independent domain at all, concentrated among LGBT+ and grassroots initiatives, so the most exposed have no configuration to measure and no identity to move. Those that do hold few relationships rather than many, since one commercial entity commonly carries more than one of hosting, mail and authoritative DNS, and nearly all place their whole delegation with a single operator. Against a resourced baseline the difference is one of kind rather than degree, except in delegation, where the two populations converge.

Two things follow for a regulatory reading. Jurisdictional exposure does not track the applicable regime, and the widest divergence we observe falls between two countries under the same instrument; what tracks it is where a provider has placed regional infrastructure, which no instrument in our frame addresses. And obligations that presuppose an organization choosing a provider do not reach one whose entire presence sits inside a platform's terms. The exposure we describe is decided at a layer that regulation, on this evidence, does not reach.

References

1. Anonymized authors: Who runs the town? a multi-dimensional analysis of municipal internet infrastructure (2026), under review
2. Bates, S., et al.: Evidence of decreasing internet entropy: the lack of redundancy in dns resolution by major websites and services. Tech. rep., National Bureau of Economic Research (2018)
3. Bellini, R., et al.: Sok: Safer digital-safety research involving at-risk users. In: IEEE Symposium on Security and Privacy (SP) (2024)
4. Cloudflare: Cyberattacks against civil society: Project galileo anniversary report. Tech. rep., Cloudflare (2026)
5. DiMaggio, P., Hargittai, E., et al.: From the 'digital divide'to 'digital inequality': Studying internet use as penetration increases. Princeton: Center for Arts and Cultural Policy Studies (2001)
6. Doan, T., et al.: An empirical view on consolidation of the web. ACM Transactions on Internet Technology **22**(3), 1–30 (2022)
7. Habib, R., et al.: A first look at public service websites from the affordability lens. In: ACM Web Conference (2023)
8. Habib, R., et al.: Formalizing dependence of web infrastructure. In: Proceedings of the ACM SIGCOMM Conference. pp. 1132–1153 (2025)
9. Hargittai, E.: Second-level digital divide: Differences in people's online skills. First Monday (2002)
10. Hendriks, R., et al.: Laces: An open, fast, responsible and efficient longitudinal anycast census system. IMC '25 (2025)
11. Houser, R., Hao, S., Cotton, C., Wang, H.: A comprehensive, longitudinal study of government dns deployment at global scale. In: IEEE/IFIP DSN (2022)

12. Jansen, B., et al.: Pushing boundaries: An empirical view on digital sovereignty of six governments. Government Information Quarterly **40**(4), 101862 (2023)
13. Kashaf, A., et al.: Analyzing third party service dependencies in modern web services: Have we learned from the mirai-dyn incident? In: ACM IMC (2020)
14. Kumar, R., et al.: Of choices and control: A comparative analysis of government hosting. In: ACM IMC (2024)
15. Kumar, R., Asif, S., Lee, E., Bustamante, F.E.: Each at its own pace: Third-party dependency and centralization around the world. ACM on Measurement and Analysis of Computing Systems (2023)
16. Le Blond, S., Uritesc, A., Gilbert, C., Chua, Z.L., Saxena, P., Kirda, E.: A look at targeted attacks through the lense of an {NGO}. In: USENIX Security 14 (2014)
17. Matthews, T., et al.: Supporting the digital safety of at-risk users: Lessons learned from 9+ years of research and training. ACM Transactions on Computer-Human Interaction (2025)
18. Moura, G., et al.: Clouding up the internet: How centralized is DNS traffic becoming? In: ACM IMC (2020)
19. Samarasinghe, N., Adhikari, A., Mannan, M., Youssef, A.: Et tu, brute? privacy analysis of government websites and mobile apps. In: ACM web conference (2022)
20. Sannon, S., Forte, A.: Privacy research with marginalized groups: what we know, what's needed, and what's next. Proceedings of the ACM on Human-Computer Interaction (2022)
21. Scheerder, A., Van Deursen, A., Van Dijk, J.: Determinants of internet skills, uses and outcomes. a systematic review of the second-and third-level digital divide. Telematics and informatics (2017)
22. Singanamalla, S., Jang, E.H.B., Anderson, R., Kohno, T., Heimerl, K.: Accept the risk and continue: Measuring the long tail of government https adoption. In: ACM IMC (2020)
23. Sommese, R., Jonker, M., van der Ham, J., Moura, G.C.: Assessing e-government dns resilience. In: CNSM (2022)
24. United Nations: UN E-Government Knowledgebase (UNEgovKB). https://publicadministration.un.org/egovkb (2024)
25. Van Deursen, A.J., Helsper, E.J.: The third-level digital divide: Who benefits most from being online? (2015)
26. Van Deursen, A.J., Van Dijk, J.A.: The first-level digital divide shifts from inequalities in physical access to inequalities in material access. New media & society (2019)
27. Van Dijk, J.: The digital divide. John Wiley & Sons (2020)
28. Warford, N., et al.: Sok: A framework for unifying at-risk user research. In: 2022 IEEE Symposium on Security and Privacy (SP)
29. Zembruzki, L., Sommese, R., Granville, L.Z., Selle J., A., Jonker, M., Moura, G.C.M.: Hosting industry centralization and consolidation. In: NOMS (2022)