INFRANEST RESEARCH · EDITION 1

Email security of the world's 100,000 most-visited websites

SPF and DMARC among the CrUX top 100,000

Jaspar Steenman https://orcid.org/0009-0000-7767-4036

InfraNest Research

Published	24 November 2026
Version	1.0
DOI	https://doi.org/10.5281/zenodo.24417472
Canonical page	https://infranest.io/research/email-security-2026/
Data collected	13–16 November 2026 (main scan 13–14 November 2026; second run 15–16 November 2026)
Pre-registration	https://doi.org/10.17605/OSF.IO/P6YTU (OSF Registries, protocol version 1.0)
Licence	Creative Commons Attribution 4.0 International (CC BY 4.0)
Contact	research@infranest.io

Suggested citation

Steenman, J. (2026). *Email security of the world's 100,000 most-visited websites* (Version 1.0) [Research report]. InfraNest Research. https://doi.org/10.5281/zenodo.24417472

Abstract

We measured SPF records and DMARC policies in public DNS for the 80,312 registrable domains behind the 100,000 origins that Chrome users worldwide load most (Chrome UX Report, October 2026), in a main scan on 13–14 November 2026 and a complete second scan on 15–16 November 2026. 57% of the mail domains studied lacked a DMARC policy requesting quarantine or rejection for all messages that fail DMARC, under the study's rules (n = 55,106 · 95% CI 56.8–57.6%): their effective DMARC policy was missing, invalid, partial or "none". Many organisations run several domains with one setup, so their domains are not independent. Counting one domain per domain group (the same name and the same mail or DNS provider) gives 58.1% (n = 51,348), outside this interval. Allowing for shared setups widens the interval to 56.7–57.7%. Applying RFC 9989's rules to the measured records gives 55.6% (n = 55,106), outside the headline's 95% interval (sensitivity check 5, added after the results were known). The newer standard ignores the legacy pct tag. Of the 55,060 mail domains whose SPF could also be read, 88% had a valid SPF record (n = 55,060 · 95% CI 87.5–88.0%), and 40% had both valid SPF and an enforced DMARC policy (n = 55,060 · 95% CI 40.1–40.9%). Counting one domain per domain group gives 39.6% (n = 51,309), also outside this interval; allowing for shared setups, the interval is 40.0–41.0%. DKIM was not measured. As registered beforehand, we expected enforced DMARC to be more common in band 1–1,000 than in all lower bands together; the data met this expectation (59% against 43%; n = 606 against 54,500; difference +16.8 points, 95% CI +12.8 to +20.7). A complete second scan agreed with the first for at least 99.8% of domains on every measure. We measured DNS configuration; message delivery and what receivers do were outside the study. The protocol was registered on OSF on 13 November 2026 at 23:16 UTC, after data collection began (19:22 UTC that day) and before any result was computed. No individual site is named.

1 Introduction

Anyone can send an email that claims to come from any domain. SPF lets a domain list the servers allowed to send its mail [13]. DMARC checks whether the visible From domain aligns with a domain authenticated by SPF or DKIM, and a domain can publish a preference for how receivers handle messages that fail DMARC: none, quarantine or reject [11, 21]. Only quarantine or reject asks receivers to act on those messages. This report asks how many of the domains behind the world's most-visited websites publish a DMARC policy that asks receivers to quarantine or reject messages that fail DMARC, and how that changes with popularity.

2 Prior work

A longitudinal study of DNS-based email authentication found DMARC on up to 11.5% of the top one million domains of three popular lists in May 2020, and that higher-ranked domains use more authentication mechanisms [24]. On the Alexa list, that study reported a more than tenfold rise between January 2015 and December 2019; its January 2015 figure, a valid DMARC record on 0.97% of the top one million domains, comes from Foster et al. [9, 24]. An analysis of SPF records from 12 million domains found errors, undefined content or ineffective rules in 2.9% of them, and many very lax configurations [4]. End-to-end tests with 35 popular email providers showed that most had the protocols to detect spoofing but still let forged mail reach the inbox [12].

Other work measured how receiving mail servers validate senders, and where that validation is vulnerable [2, 5]; the security of email delivery between providers [7, 9]; and the experience of running the mechanisms in a test system [21].

The closest recent comparison is Kondo et al. [14], who measured the Tranco top one million every ten days from February to May 2024, after Google's new requirements for senders to Gmail. Over that period the share of domains publishing DMARC rose, fewer domains met their definition of an effective DMARC policy than published DMARC at all, which they read as a preference for a policy of none, and adoption lagged among domains associated with China, South Korea and Japan. Their "effective DMARC policy" is defined differently: it also requires a reporting address, and its definition does not mention `pct` (`t=` did not exist yet). Tranco also ranks domains differently from CrUX, so their shares and ours are not directly comparable. The European Commission's Joint Research Centre reports the uptake of these standards in EU member states regularly; its 2025 report found a marked rise in DMARC [15]. Compared with this earlier work, this study measures the most-visited websites according to CrUX, compares popularity bands from the top 1,000 to the top 100,000, and counts the effective DMARC policy rather than the published one. Its plan was registered before any result was computed, and its totals are open data.

3 Population

The 100,000 origins (a scheme and host name, such as https://www.example.com) that Chrome users worldwide load most, from the Chrome UX Report (CrUX), October 2026 release [10], as published by the crux-top-lists project [8]. Of the popular top-site lists evaluated in 2022, CrUX matched Cloudflare's traffic data best [23]. Each origin becomes its registrable domain with the Public Suffix List [19], and a domain takes the most popular bucket of any of its origins. Several origins can share one domain, so the list gives 80,312 domains. CrUX places websites only in buckets [8, 10]; the bucket boundaries (1,000, 5,000, 10,000, 50,000, 100,000) are the rank values in the list file itself. We compare the disjoint bands between them, and within a band there is no ranking.

Domains per band (best bucket of any origin of the domain).

Band	Domains	Cumulative
1–1,000	816	816
1,001–5,000	3,370	4,186
5,001–10,000	4,183	8,369
10,001–50,000	32,310	40,679
50,001–100,000	39,633	80,312

4 Method

4.1 Measures

Definitions.

Measure	Counts as yes when
Mail domain	A usable explicit MX record. Not usable: a null MX, which declares that the domain accepts no mail [18], and an MX whose every host is `localhost` or a name without a dot. Only mail domains count in the email measures; a domain without an MX record can still receive mail at its address record (implicit MX), but is outside them (section 6).
DMARC published	Exactly one `v=DMARC1` record at `_dmarc.<domain>`; two records count as no record, because receivers discard both ([11], section 4.10).
Effective DMARC policy	One of six, decided in this order: (1) missing: no record, or two; (2) invalid: invalid characters, a missing `;`, or no or an invalid `p=`; (3) with `pct` 1–99 ([16], sections 6.3 and 6.6.4): `p=quarantine` is partial; `p=reject` counts as quarantine, or as none with `t=y`; (4) otherwise the `p=` value (none, quarantine or reject), one level lower with `t=y` ([11], section 4.7) and/or a legacy `pct=0`. An invalid `pct` is ignored, as receivers do. An invalid `sp=` is ignored by our choice, because subdomains are outside the study. Enforced = quarantine or reject. Combining RFC 7489's `pct` with RFC 9989's `t=` is this study's own classification, because receivers follow both today; sensitivity check 5 gives the result under RFC 9989 alone. A record without `p=` counts as invalid here and as none under RFC 9989; neither is enforced.
No enforcing DMARC policy (primary outcome)	Effective policy missing, invalid, partial or none: the domain does not ask receivers to quarantine or reject all mail that fails DMARC. Over mail domains whose MX and DMARC could be read after the retry pass. A domain-level classification: only the domain's own `_dmarc` record is read, not a policy higher up the DNS tree (section 6).
Valid SPF	Exactly one `v=spf1` record and no permanent error in anything it reaches: at most 10 DNS lookups, counted across every branch (section 5.3), at most 2 void lookups (the default RFC 7208 recommends, which our scanner applies), no include loop or missing include, no unknown term, bad address or invalid characters ([13], sections 4.5 and 4.6).
SPF near the limit	Valid with 8–10 lookups, as a share of valid records (our threshold).
SPF over the limit	A published record needing more than 10 lookups, as a share of mail domains publishing a readable SPF record.
Valid SPF and enforced DMARC	Valid SPF and an effective DMARC policy of quarantine or reject, as a share of mail domains whose SPF and DMARC could be read. Not the complement of the primary outcome; DKIM is not part of it (section 4.2).

The same scan also measured DNSSEC, HTTPS certificates and the providers behind mail, DNS and the web. Those totals are in the data files; the text about them follows in a companion study of Dutch, German and French websites, planned for December 2026.

4.2 Not measured

- DKIM, the signature a sending server adds to show that a message comes from the domain: its keys sit under selector names that cannot be listed from DNS.

- Whether real mail passes DMARC alignment, and what receivers do with mail that fails: we measure the policy a domain asks for in DNS, not what happens to its mail.
- Encryption between mail servers (STARTTLS): connecting to mail servers is outside our rules for public data.
- Subdomains, web redirects, and the registrar of each domain.

4.3 Collection

Public DNS only, through Cloudflare's unfiltered DNS over HTTPS service (`cloudflare-dns.com`, the 1.1.1.1 service), with Google Public DNS (`dns.google`) as fallback; both validate DNSSEC. Queries were sent from a scan server (DigitalOcean, Frankfurt). The scanner uses the same SPF and DMARC analysers as InfraNest's free SPF and DMARC checks; the effective policy and the totals are computed afterwards, from the scanner's results. Domains that could not be read were retried once, about a day later (the plan says 24 hours; see the deviations). Websites that gave the scan server no HTTPS handshake were tried again from a second location (Cologne, Germany). Main scan and its retry pass: 13–14 November 2026; second run (test-retest) and its retry pass: 15–16 November 2026; second-location re-check: 15 November 2026. Scanner commit e0c9063, tagged `scanner-e0c9063` in the public repository. List: CrUX global top 100,000 origins, October 2026 release (zakird/crux-top-lists, crux-202610).

4.4 Exclusions

A record counts as unreadable only when neither resolver could read it, even after the retry. A read error on the MX or `_dmarc` record removes the domain from all email measures; an unreadable SPF record removes it from the SPF measures and from "valid SPF and enforced DMARC" only. Domains with no MX, a null MX or an unusable MX are left out of the email measures. Owners who asked to be left out are skipped and only counted. The first six counts add up to the 80,312 domains of the list; the last is part of the mail domains.

Domains left in and out.

Mail domains	55,106
No MX, null MX or unusable MX	24,654
MX or DMARC unreadable	227
Email unreadable: the domain's own DNSSEC is broken	4
No longer registered	321
Opted out	0
SPF unreadable (of the mail domains)	46

4.5 Quality checks

Three checks test whether the numbers are right. The hand check tests the scanner's verdicts for single domains. The recount tests how the totals were added up from those verdicts. The test-retest measures how much the results change between two scans. Sensitivity checks then test whether the headline depends on choices we made.

- **Hand check.** We took 100 random readable domains, plus up to 20 domains in each rare state (DMARC invalid, SPF over the limit, DNSSEC broken, certificate expired). We re-checked them with

independent tools, through a resolver the scanner does not use. The target was at least 98% agreement per measure, with every disagreement explained. Result: mail (MX) 100.0%, SPF 100.0%, SPF lookups 100.0%, DMARC policy 100.0%, DNSSEC 100.0%, HTTPS 100.0% agreement (n = 100, checkdmarc 6.0.4 for SPF and DMARC and delv and dig (BIND 9.20) for DNSSEC, through Quad9; openssl for HTTPS). In the extra domains in rare states, every disagreement was traced to a difference of definition, a limit of the checking tools, or one of the scanner limitations under implementation notes.

- **Separate recount (AI-assisted).** A separate AI model run, which saw the plan, the scanner's source and the raw results but not our analysis code, recomputed the published numbers with its own code; any difference blocks publication until its cause is found and fixed. Result (20 November 2026): 3,699 of the 3,724 published values were recomputed on the final file, after the retry pass and the re-check from a second location; 3,621 are identical, and the other 78 differ only in labels or in how a definition was read (implementation note 16). On 16 November the recount added sensitivity check 4 and the test-retest: all 97 values are identical, after one correction to our code, which had counted unsigned domains in the test-retest of the DNSSEC state (implementation note 18). Not recomputed: September's own figures, which need September's per-domain results. On 20 November the recount also recomputed the two checks added after the outside review (sensitivity check 5 and the disjoint comparison in check 1): every value is identical, after one correction to our code, which had not applied RFC 9989's rule for an invalid sp= (implementation note 19)..
- **Test-retest.** We scanned the whole list a second time, 36–48 hours later, with the same code and a new order. On every measure, at least 99.8% of domains gave the same result in both scans. Cohen's kappa [3] measures agreement beyond chance; 1 is perfect agreement, and the plan calls agreement low below 0.80, or below 95% identical. The two runs are never combined: every headline comes from the first run and its retry pass. Share identical and kappa per measure, compared scan against scan:

Measure	Identical	Kappa	n
Mail domain (MX)	over 99.9%	0.999	79,735
Effective DMARC policy	99.9%	0.999	55,089
DMARC enforced	over 99.9%	0.999	55,089
SPF (none / valid / invalid)	99.9%	0.998	55,043
SPF near the limit	over 99.9%	0.999	48,302
Valid SPF and enforced DMARC	99.9%	0.999	55,043
DNSSEC signed	over 99.9%	1.000	79,976
DNSSEC state	99.9%	0.927	5,389
HTTPS (trusted / untrusted / no answer)	99.8%	0.974	77,767
HTTPS trusted	over 99.9%	0.984	75,739

  The published HTTPS result also includes the answers from the second location, which only the first run had; run 2 against that result gives 99.6% identical, kappa 0.935 (n = 77,767).
- **Sensitivity checks** (never headlines). (1) Lists: does the headline depend on the popularity list? Of the CrUX domains that are also in the top 100,000 of Tranco list 947N2, generated on 9 November 2026 [17, 25] (42,852 domains still registered), 51.3% have no enforcing policy (n = 33,475): −5.9

points, 95% CI −6.6 to −5.2 against the headline; valid SPF and enforced DMARC: 46.2%, +5.7 points, 95% CI +5.1 to +6.4. This subset lies inside the headline population; the interval treats the two as independent samples and does not account for the overlap; for a subset of the whole, this usually makes it wider. Both estimates are outside the headline's interval. So domains that are popular on both lists do better on both measures than the headline population. Compared as two separate groups (added after the results were known), the CrUX domains in Tranco against those not in it: no enforcing DMARC policy 51.3% against 66.4% (n = 33,475 against 21,631; −15.0 points, 95% CI −15.9 to −14.2; different by our rule); valid SPF and enforced DMARC 46.2% against 31.6% (n = 33,449 against 21,611; +14.6 points, 95% CI +13.8 to +15.4; different by our rule). (2) One domain per domain group: how do the headline measures look when brands with many domains count once? Domains are grouped when they share the name left of the public suffix and a mail or DNS provider, or, failing a known provider, the same set of mail or name servers; this approximates an organisation but cannot prove one. Each group is represented by one of its domains, the one in the most popular band, then the alphabetically first; a group whose representative is not a mail domain drops out of the email measures. These are estimates under a different weighting (each group counts once), not a test of whether grouping matters. One domain per group: 58.1% no enforcing DMARC policy (n = 51,348), outside the headline's 95% interval; headline interval allowing for shared setups 56.7–57.7%; 87.5% valid SPF (n = 51,309), inside the headline's 95% interval; headline interval allowing for shared setups 87.5–88.1%; 12.3% SPF near the limit (n = 44,896), inside the headline's 95% interval; headline interval allowing for shared setups 12.0–12.7%; 2.6% SPF over the limit (n = 47,659), inside the headline's 95% interval; headline interval allowing for shared setups 2.4–2.6%; 39.6% valid SPF and enforced DMARC (n = 51,309), outside the headline's 95% interval; headline interval allowing for shared setups 40.0–41.0%. The intervals allowing for shared setups come from 2,000 bootstrap resamples over the groups (seed 20261114); a measure outside its interval says so next to its headline number. (3) September 2026 bridge: section 5.5. (4) Second run: does a complete second scan give the same headline? The test-retest run, analysed on its own: 57.2% no enforcing DMARC policy (n = 55,111), 87.8% valid SPF (n = 55,063), 12.3% SPF near the limit (n = 48,340), 2.5% SPF over the limit (n = 51,223), 40.5% valid SPF and enforced DMARC (n = 55,063); each is inside its headline's 95% interval. (5) RFC 9989 handling of the measured records (added after the results were known, following an outside review): the primary outcome when the records we measured are read with the rules of the newer DMARC standard, which drops `pct` ([11], Appendix A.6) and reads a record without a valid `p=`, or with an invalid `sp=`, as none if it has a reporting address and as no DMARC policy otherwise (records we already count as invalid stay invalid) (section 4.10.1); neither is enforced. This is not full RFC 9989 processing: policies higher up the DNS tree are not looked up, and the scanner does not check `np=`, so an invalid `np=` is not caught. 55.6% have no enforcing policy (n = 55,106 · 95% CI 55.2–56.1%), 870 domains fewer than under the registered rule (−1.6 points), outside the headline's 95% interval. Classes changed: 842 partial → quarantine, 29 none → quarantine and 1 quarantine → none, which change the outcome; 156 quarantine → reject, 23 invalid → none and 1 partial → none, which do not.

4.6 Statistics

Each share has its denominator n. Headline measures have a 95% confidence interval (CI), computed with the Wilson method [27], which stays accurate for shares near 0% or 100%. We call two shares different only when the 95% interval of their difference (Newcombe's hybrid Wilson score method for the difference of two shares [20]) excludes zero and the difference is at least 3 points; otherwise we report

no difference by our rule, which is not a test of equivalence. The list is complete, not a random sample, so the intervals describe only binomial sampling uncertainty, as if the domains were drawn from a larger population of similar domains. They do not cover measurement error, which the hand check and the test-retest assess, or the choice of list, which sensitivity check 1 assesses. Only the primary outcome and the pre-registered expectation are confirmatory; the band comparisons are descriptive. A measure with more than 2% unreadable domains shows that share next to its number, and above 5% it is not used in the text. Cells with fewer than 100 domains are never shown. The text rounds shares to whole percentages; intervals, differences, the quality checks and the sensitivity checks keep one decimal, as do tables and data files.

4.7 Pre-registration and deviations

The research questions, definitions and comparisons were fixed in a plan, registered on OSF Registries as version 1.0 on 13 November 2026 at 23:16 UTC, timestamped and read-only, before any number in this study was computed: https://doi.org/10.17605/OSF.IO/P6YTU. As the plan itself states: data collection began on 13 November 2026 at 19:22 UTC, with scanner code that implements draft 0.9 of the plan, which was being finalised at the time. Version 1.0 was registered after that, before any result was computed or viewed. The measure definitions and the headline analysis did not change between the two.

Deviations from the plan:

- Retry passes (timed during data collection, 13–14 November 2026, before any result was computed): the retry passes of the main scan and of the September re-scan were timed 24 hours after each scan started, so they ran about 22 hours after each scan ended, not 24 hours later as protocol section 6 asks. A retry pass only re-reads domains that could not be read or whose DNSSEC looked broken; we expect no effect on any result.
- Choices made after the first results were seen (15–16 November 2026), documented in implementation notes 17 and 18: the test-retest compares scan against scan, to compare like with like, and run 2 against the published HTTPS result is reported beside it; and the reporting code was corrected in three places to match the plan's definitions (prose rounded from counts, one domain added to the email unreadable flag, and the DNSSEC state test-retest limited to signed domains). None changes a headline number or a conclusion.
- Changes after an outside review of the draft (20 November 2026), all after the results were known: sensitivity check 5 added (the primary outcome with RFC 9989's rules applied to the measured records), and a second comparison in check 1 (CrUX domains in Tranco against those not in it, two disjoint groups). Neither is in the plan; both are reported beside the registered checks and never as headlines. The names of two measures were made more exact without changing their definitions: "not protected against spoofing" is now "no enforcing DMARC policy", and "fully configured" is now "valid SPF and enforced DMARC". Definitions, rules and every registered number are unchanged (implementation note 19).

4.8 Implementation notes

How the code applies the plan, including the choices the plan left open, is documented with the code, in numbered implementation notes (https://github.com/InfraNest-Infrastructure-Organized/infranest-research/blob/main/docs/study1-implementation-notes.md); the notes this report cites by number are there. The checks found these limitations of the scanner. We publish the values as scanned, and fix the scanner before the next study:

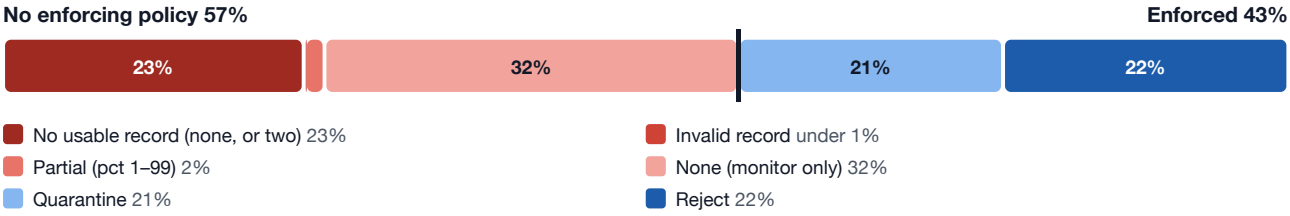
- Note 14: DNS-over-HTTPS answers carry text records in escaped form (a tab arrives as \009), and the scanner did not decode the escapes. One DMARC record that separates its tags with tabs, which the standard allows, is counted as invalid; it should count as “none”. Both count as having no enforcing policy, so the headline does not change, and no published share changes at one decimal, also in that record’s band. Only the records with a syntax finding were queried again; a record with escapes but no syntax finding, or one that starts with an escape, could have been missed.
- Note 15: A signed zone counts as insecure when the first resolver that answers, normally Cloudflare, does not validate it. When we checked them again the next day, most of these zones validated through Google and Quad9. Insecure is part of the DNSSEC totals in the data files; the text does not use it.

5 Results

5.1 Headline

In November 2026, 57% of the mail domains studied lacked a DMARC policy requesting quarantine or rejection for all messages that fail DMARC, under the study’s rules (n = 55,106 · 95% CI 56.8–57.6%). Their effective DMARC policy was missing, invalid, applied to only part of the mail, or set to “none”. This is what the domains ask for in DNS. Receivers decide what happens to a message: RFC 9989 asks them to weigh other information even when a domain asks for reject ([11], section 5.4), and they may filter forged mail by other means. We did not measure delivery. Many organisations run several domains with one setup, so their domains are not independent. Counting one domain per domain group (the same name and the same mail or DNS provider) gives 58.1% (n = 51,348), outside this interval. Allowing for shared setups widens the interval to 56.7–57.7%. Applying RFC 9989’s rules to the measured records gives 55.6% (n = 55,106), outside the headline’s 95% interval (sensitivity check 5, added after the results were known). The newer standard ignores the legacy pct tag.

Figure 1. Effective DMARC policy of all mail domains. Left of the line: no enforcing DMARC policy; right: enforced (quarantine or reject). Data: https://infranest.io/research/email-security-2026/figure-1.csv



5.2 Effective DMARC policy

We count the effective policy: the policy the standard asks receivers to apply. A record in test mode (`t=y`, or the older `pct=0`) counts one level lower [11, 16]. With `pct` 1–99 [16], `p=quarantine` counts as partial, which is not enforced; `p=reject` counts as quarantine, or as none in test mode. Receivers that still follow the first specification ignore `t=` [16].

Of the 55,106 mail domains, 22% ask receivers to reject a forged message and 21% to quarantine it: 43% in all. The largest group is the 32% that have an effective policy of “none”, which stops nothing (it can only ask for reports). The other groups, as shares of all mail domains: 23% have no usable DMARC record (none at all, or two, which receivers discard), 2% apply quarantine to only part of their mail and

under 1% publish a record receivers may not be able to parse. Together, the groups other than reject and quarantine make up the 57% with no enforcing DMARC policy.

Figure 2. Effective DMARC policy of mail domains, by popularity band. Left of the line: no enforcing DMARC policy; right: enforced (quarantine or reject). The left side is shortest in band 1–1,000 (41%) and longest in band 50,001–100,000 (61%). Data: https://infranest.io/research/email-security-2026/figure-2.csv

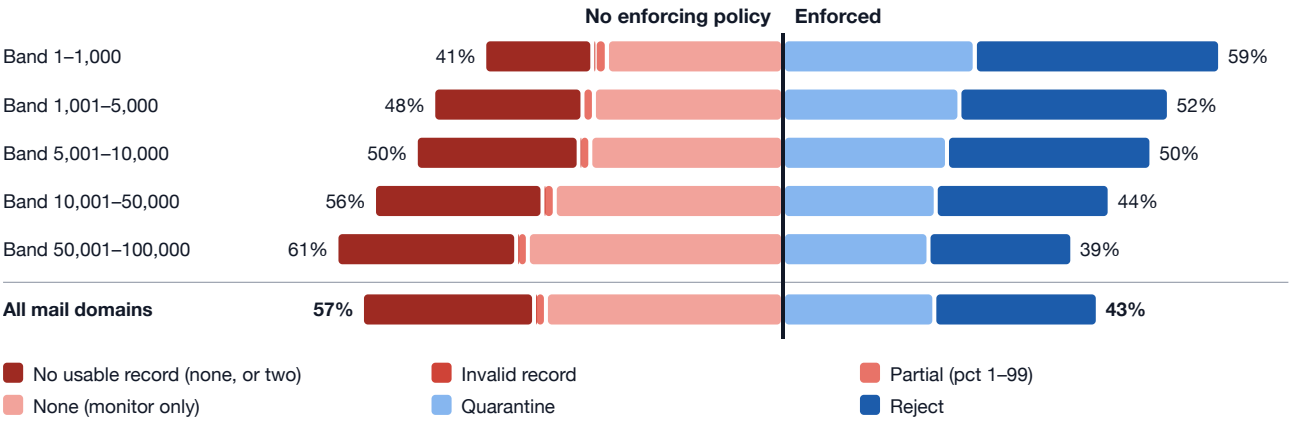


Figure 2 in numbers: effective DMARC policy by band (% of mail domains).

Policy	1–1,000	1,001–5,000	5,001–10,000	10,001–50,000	50,001–100,000	All (top 100,000)
Reject	33.3%	28.5%	27.8%	23.6%	19.5%	22.2%
Quarantine	26.1%	24.0%	22.3%	20.8%	19.8%	20.6%
None (monitor only)	23.9%	25.7%	26.2%	31.0%	34.7%	32.2%
Partial (pct 1–99, p=quarantine)	1.7%	1.6%	1.5%	1.6%	1.5%	1.5%
Invalid record	0.3%	0.0%	0.1%	0.1%	0.1%	0.1%
No usable record (none, or two)	14.7%	20.3%	22.1%	22.9%	24.4%	23.4%
Mail domains (n)	606	2,505	2,997	22,335	26,663	55,106

Counted as published, before test mode and `pct` are applied, and as a share of the 42,220 mail domains with exactly one DMARC record: 29% publish `p=reject`, 29% `p=quarantine` and 42% `p=none`. Of these domains, under 1% are in test mode (`t=y` or `pct=0`) and 3% set `pct` below 100. Records with an invalid `sp=`: 2. These shares differ from the ones above because they count published records, not effective policies.

5.3 SPF

Checking an SPF record takes DNS lookups, one for each `include`, `a`, `mx`, `ptr`, `exists` or `redirect` in it and in the records it includes. More than 10 lookups ([13], section 4.6.4) or two SPF records (section 4.5) give a permanent error, so SPF cannot pass. Our scanner counts DNS-querying terms across all branches of the record and the records it includes. A receiver applies the RFC limit to the terms it evaluates for a particular message: it checks them in order and stops at the first match. The two can differ: the scanner also counts terms written after an `all`, which a receiver never reaches, and counts an include built from a macro as one lookup without following it. “Over the limit” is therefore a finding under our scanner’s counting rule. A record near the limit (8 to 10 lookups) can fail when one or two more sending services are added. 88% of the mail domains have a valid SPF record, and 7% publish no SPF record. Of

the published records, 6% give a permanent error; for 2% of records the cause is being over the limit. Of the valid records, 12% are near the limit.

Figure 3. SPF on the mail domains. Left of the line, SPF cannot pass (no record, or an error). Right of the line, SPF works; the lighter part is near the limit of 10 lookups. Shares in the bar are of the mail domains whose SPF could be read; the 12% near the limit in the text is a share of valid records. Data: https://infranest.io/research/email-security-2026/figure-3.csv

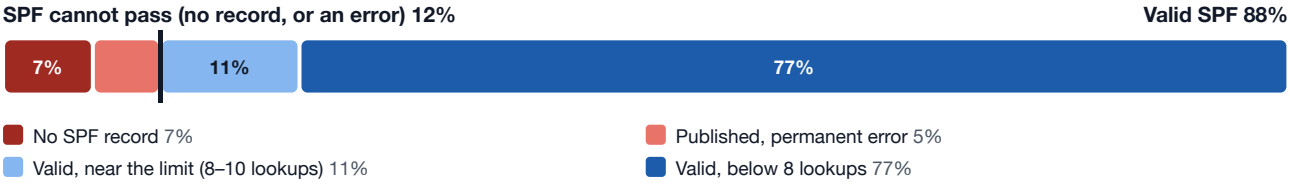


Figure 3 in numbers: SPF on the mail domains, all bands.

Measure	Share	Of
SPF record published	93.0%	mail domains
Valid SPF 95% CI 87.5–88.0%	87.8%	mail domains
Over the limit: more than 10 lookups 95% CI 2.4–2.6%	2.5%	SPF records
Near the limit: 8–10 lookups 95% CI 12.0–12.6%	12.3%	valid records
Permanent error (any)	5.6%	SPF records
… two SPF records	1.4%	SPF records
… more than 2 void lookups	0.1%	SPF records
Ends in `+all` (anyone may send)	0.0%	SPF records
Mail domains whose SPF could be read: 55,060. SPF records: 51,217. Valid records: 48,325.		

Valid SPF and an enforced DMARC policy (quarantine or reject): 40% (n = 55,060 · 95% CI 40.1–40.9%). Many organisations run several domains with one setup, so their domains are not independent. Counting one domain per domain group (the same name and the same mail or DNS provider) gives 39.6% (n = 51,309), outside this interval. Allowing for shared setups widens the interval to 40.0–41.0%.

5.4 By popularity

These comparisons are descriptive. In band 1–1,000, 41% of mail domains have no enforcing DMARC policy; in band 50,001–100,000, 61% (band 1–1,000 minus band 50,001–100,000: −20.1 points, 95% CI −24.0 to −16.1): different by our rule. From the most-visited band to the least, the share is 41%, 48%, 50%, 56% and 61%. Band 1–1,000 has a higher share than band 50,001–100,000 for two other measures: valid SPF and enforced DMARC (56% against 37%; +19.0 points, 95% CI +15.0 to +22.9) and valid SPF (92% against 87%; +4.5 points, 95% CI +2.0 to +6.5). For SPF near the limit and SPF over the limit, the two bands show no difference by our rule.

Expectation E2, registered before any result was computed: enforced DMARC (quarantine or reject) more common in band 1–1,000 than in all other bands together (1,001–100,000). The data met this expectation: 59% against 43% (n = 606 against 54,500 mail domains; difference +16.8 points, 95% CI +12.8 to +20.7). E1 and E3 belong to the companion study.

Figure 4. Two measures by popularity band. Dot: the band's share; line through it: its 95% interval; dashed line: all mail domains together. Valid SPF and enforced DMARC is a share of mail domains whose SPF and DMARC could be read. Data: https://infranest.io/research/email-security-2026/figure-4.csv

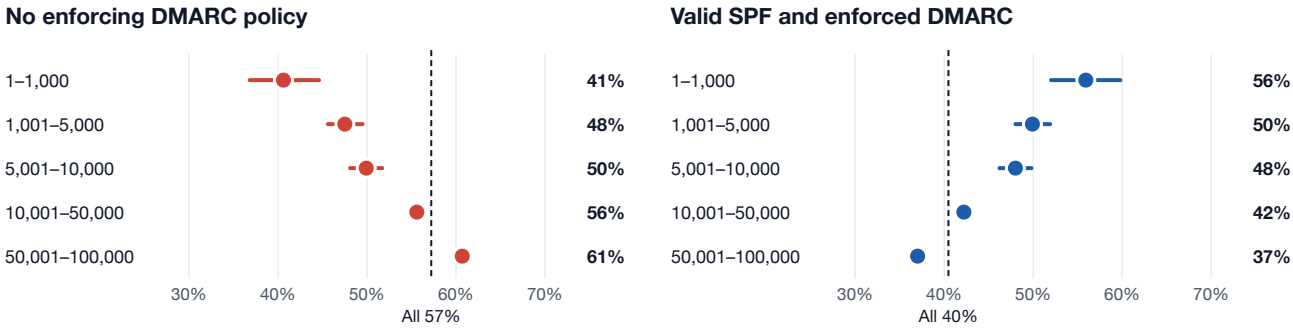


Table 1. All measures by popularity band (% of the measure's population; 95% interval below the headline measures). Bands are disjoint. Data: https://infranet.io/research/email-security-2026/table-1.csv

Share of mail domains, by band	1–1,000	1,001–5,000	5,001–10,000	10,001–50,000	50,001–100,000	All (top 100,000)
Domains in the list	816	3,370	4,183	32,310	39,633	80,312
Mail domains (n)	606	2,505	2,997	22,335	26,663	55,106
… whose SPF could be read (n)	606	2,502	2,996	22,317	26,639	55,060
No enforcing DMARC policy	40.6% 36.8–44.6%	47.5% 45.6–49.5%	49.9% 48.1–51.7%	55.6% 54.9–56.2%	60.7% 60.1–61.3%	57.2% 56.8–57.6%
No DMARC record (or two)	14.7%	20.3%	22.1%	22.9%	24.4%	23.4%
DMARC policy: reject	33.3%	28.5%	27.8%	23.6%	19.5%	22.2%
Valid SPF	91.7% 89.3–93.7%	88.8% 87.6–90.0%	89.4% 88.3–90.5%	87.9% 87.5–88.3%	87.3% 86.9–87.7%	87.8% 87.5–88.0%
SPF at 8–10 lookups (of valid)	10.4%	12.0%	13.5%	12.5%	12.1%	12.3%
SPF over 10 lookups (of SPF records)	2.4%	2.0%	2.5%	2.5%	2.5%	2.5%
Valid SPF and enforced DMARC	55.9% 52.0–59.8%	49.9% 48.0–51.9%	48.0% 46.2–49.8%	42.2% 41.5–42.8%	37.0% 36.4–37.5%	40.5% 40.1–40.9%

Table 2. No enforcing DMARC policy, cumulative tiers (each includes the ones before). Data: https://infranet.io/research/email-security-2026/table-2.csv

Tier	Top 1,000	Top 5,000	Top 10,000	Top 50,000	Top 100,000
Domains in the list	816	4,186	8,369	40,679	80,312
No enforcing DMARC policy	40.6% 36.8–44.6%	46.2% 44.4–47.9%	48.0% 46.8–49.3%	54.0% 53.4–54.5%	57.2% 56.8–57.6%

5.5 September 2026

In September we published SPF and DMARC figures for the top 10,000 domains of a different list (Tranco list K9PXW, generated on 24 September 2026 [17, 26]), using an older analyser. The list and definitions differ, so we do not compare this study's results with September's. Instead, we re-scanned September's list with this study's code (sensitivity check 3) and compare that re-scan with September's figures; the differences mix two months of change with the change of analyser. Two definitions also differ: September counted a DMARC record as published even when a domain had two, and September's

SPF base was all mail domains. Both columns describe largely the same domains, but the interval treats them as independent samples, as registered, and does not account for that pairing; a paired interval would usually be narrower.

Table 3. The same Tranco top 10,000, two analysers. Data: https://infranest.io/research/email-security-2026/table-3.csv

Same domains	September 2026 old analyser	November 2026 this study's code	Difference points, 95% CI
SPF record published (of mail domains)	92.9%	93.2%	+0.2 points, 95% CI −0.6 to +1.1
DMARC record published (of mail domains)	82.3%	82.8%	+0.6 points, 95% CI −0.7 to +1.8
More than 10 SPF lookups (of domains publishing SPF)	2.7%	2.8%	+0.1 points, 95% CI −0.5 to +0.6
Published p=reject (of domains with a DMARC record)	42.5%	43.1%	+0.6 points, 95% CI −1.2 to +2.3

September 2026: 7,204 mail domains, scanned 25 September 2026; counted from September's own results with September's definitions. Re-scan: Tranco K9PXW top 10,000.

6 Limitations

- CrUX counts only Chrome users on desktop and Android who sync their history and share usage statistics; Chrome on iOS and other Chromium-based browsers are not included. Websites must also be publicly indexable and pass a popularity threshold Google does not disclose [10].
- Most websites are popular in only one country, so a global list leans towards websites that are popular in many countries and away from local ones such as government, education and finance [22]. The companion study uses each country's own list for that reason.
- The domains within one band cannot be ranked against each other.
- One organisation can have many domains (one brand under many country domains). They usually share one setup, so they are not independent; sensitivity check 2 counts one domain per domain group and recomputes the interval over the groups.
- A domain without its own DMARC record counts as having none, even when a policy higher up the DNS tree would apply to it (for example under a shared hosting suffix). This can overstate the share with no enforcing policy slightly.
- Receivers that still follow the first DMARC specification ignore test mode, so for them a few records count one level higher than here.
- One moment in time, from one scan server through two public resolvers: a domain can change its records the next day.
- Domains without a usable explicit MX record are outside the email measures, although mail can still reach some of them through their address record (implicit MX), and mail can be forged in the name of any of them. This study does not count how many of them publish the protection a non-mail domain needs (a null MX, `v=spf1 -all` and a DMARC policy of reject).

7 Ethics

Public DNS answers through public resolvers, and for HTTPS one TLS handshake per name; no logins, no probing, no email sent. The scanner identifies itself and limits its rate per network. Results are published only as totals by popularity; no individual site is named, and no list is published from which one could be read off. Per-domain results are used for nothing else, stay private, and are deleted no later than 13 months after the scan. Site owners can ask what was measured or be left out of future editions at research@infranest.io; the scanner is described at https://infranest.io/research/scanner/. No ethics board reviewed this study; it follows the principles of the Menlo Report [1] and established guidance for internet-wide scanning [6]: public data, a low rate, clear identification and an opt-out.

8 Funding, interests and roles

The research is funded and run by InfraNest, a software company whose platform lets teams manage their domains, DNS, servers, certificates and monitoring in one place. The platform includes SPF and DMARC checks, and published totals may be used in its marketing. To limit this interest, the definitions, primary outcome and decision rules were fixed in the registered plan, every pre-specified total is published whatever its direction, and the code and input lists are public. Scan, analysis and text are by Jaspar Steenman (InfraNest Research), who founded InfraNest. The separate recount (section 4.5) and two reviews of this text, one for facts and sources and one for clarity, were AI-assisted: each was a separate AI model run that had not seen the analysis code or the other reports. They are not external peer review. An outside review of the draft, also AI-assisted, led to sensitivity check 5, the second comparison in check 1 and clearer definitions; the deviations list these changes.

9 Data and code availability

Totals as CSV and JSON, the numbers behind each figure and table (Figure 1, Figure 2, Figure 3, Figure 4, Table 1, Table 2, Table 3), the input list (domain and CrUX bucket) and this report are released under CC BY 4.0 at https://doi.org/10.5281/zenodo.24417472 and at https://infranest.io/research/email-security-2026/; input list: https://github.com/InfraNest-Infrastructure-Organized/infranest-research/blob/scanner-e0c9063/lists/crux-global100k-202610.csv. The scanner, the summariser and the plan are at https://github.com/InfraNest-Infrastructure-Organized/infranest-research. In the data files, the primary outcome is named `not_protected` and valid SPF with enforced DMARC `fully_configured`. No per-domain results are published.

Version history

- 1.0 · 24 November 2026 · First publication.

References

[1] Michael Bailey, David Dittrich, Erin Kenneally, and Doug Maughan. 2012. The Menlo Report. *IEEE Security & Privacy* 10, 2 (2012), 71–75. https://doi.org/10.1109/MSP.2012.52

[2] Nathaniel Bennett, Rebekah Sowards, and Casey Deccio. 2022. SPFail: Discovering, Measuring, and Remediating Vulnerabilities in Email Sender Validation. In *Proceedings of the 22nd ACM Internet Measurement Conference (IMC '22)*. ACM, 633–646. https://doi.org/10.1145/3517745.3561468

[3] Jacob Cohen. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* 20, 1 (1960), 37–46. https://doi.org/10.1177/001316446002000104

[4] Stefan Czybik, Micha Horlboge, and Konrad Rieck. 2023. Lazy Gatekeepers: A Large-Scale Study on SPF Configuration in the Wild. In *Proceedings of the 2023 ACM Internet Measurement Conference (IMC '23)*. ACM, 344–355. https://doi.org/10.1145/3618257.3624827

[5] Casey Deccio, Tarun Yadav, Nathaniel Bennett, Alden Hilton, Michael Howe, Tanner Norton, Jacob Rohde, Eunice Tan, and Bradley Taylor. 2021. Measuring Email Sender Validation in the Wild. In *Proceedings of the 17th International Conference on emerging Networking EXperiments and Technologies (CoNEXT '21)*. ACM, 230–242. https://doi.org/10.1145/3485983.3494868

[6] Zakir Durumeric, Eric Wustrow, and J. Alex Halderman. 2013. ZMap: Fast Internet-wide Scanning and Its Security Applications. In *Proceedings of the 22nd USENIX Security Symposium (USENIX Security 13)*. USENIX Association, 605–620. https://www.usenix.org/conference/usenixsecurity13/technical-sessions/paper/durumeric

[7] Zakir Durumeric, David Adrian, Ariana Mirian, James Kasten, Elie Bursztein, Nicolas Lidzborski, Kurt Thomas, Vijay Eranti, Michael Bailey, and J. Alex Halderman. 2015. Neither Snow Nor Rain Nor MITM...: An Empirical Analysis of Email Delivery Security. In *Proceedings of the 2015 Internet Measurement Conference (IMC '15)*. ACM, 27–39. https://doi.org/10.1145/2815675.2815695

[8] Zakir Durumeric and David Adrian. 2022. Chrome (CrUX) Top Million Websites. Dataset, data/global/202610.csv.gz at commit a67d0c8f09326df09739560bc43205b4f949f8e6 (2026-11-10). GitHub. Retrieved November 11, 2026 from https://github.com/zakird/crux-top-lists

[9] Ian D. Foster, Jon Larson, Max Masich, Alex C. Snoeren, Stefan Savage, and Kirill Levchenko. 2015. Security by Any Other Name: On the Effectiveness of Provider Based Email Security. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security (CCS '15)*. ACM, 450–464. https://doi.org/10.1145/2810103.2813607

[10] Google. 2026. Chrome UX Report (CrUX). Dataset. BigQuery dataset chrome-ux-report; release 202610. https://developer.chrome.com/docs/crux/methodology

[11] Herr, T., Ed. and J. Levine, Ed., "Domain-Based Message Authentication, Reporting, and Conformance (DMARC)", RFC 9989, DOI 10.17487/RFC9989, May 2026, <https://www.rfc-editor.org/info/rfc9989>.

[12] Hang Hu and Gang Wang. 2018. End-to-End Measurements of Email Spoofing Attacks. In *Proceedings of the 27th USENIX Security Symposium (USENIX Security 18)*. USENIX Association, 1095–1112. https://www.usenix.org/conference/usenixsecurity18/presentation/hu

[13] Kitterman, S., "Sender Policy Framework (SPF) for Authorizing Use of Domains in Email, Version 1", RFC 7208, DOI 10.17487/RFC7208, April 2014, <https://www.rfc-editor.org/info/rfc7208>.

[14] Daishi Kondo, Yuya Shibuya, Rie S. Yamaguchi, Tomohiro Ishihara, Yuji Sekiya, Toshiyuki Nakata, and Tohru Asami. 2025. Who Did Not Implement Email Security Measures After Google's New Email Sender Guidelines?: A Large-Scale Measurement Study. In *Proceedings of the Network Traffic Measurement and Analysis Conference (TMA 2025)*. IFIP. https://tma.ifip.org/2025/wp-content/uploads/sites/14/2025/06/tma2025_paper36.pdf

[15] Vasileios Kouliaridis and Ioannis Kounelis. 2025. Email Communication Security Standards: An Analysis of Uptake in the EU: 2025. JRC143105. Publications Office of the European Union. https://doi.org/10.2760/5576432

[16] Kucherawy, M., Ed. and E. Zwicky, Ed., "Domain-based Message Authentication, Reporting, and Conformance (DMARC)", RFC 7489, DOI 10.17487/RFC7489, March 2015, <https://www.rfc-editor.org/info/rfc7489>. Obsoleted by RFC 9989, RFC 9990 and RFC 9991.

[17] Victor Le Pochat, Tom Van Goethem, Samaneh Tajalizadehkhoob, Maciej Korczyński, and Wouter Joosen. 2019. Tranco: A Research-Oriented Top Sites Ranking Hardened Against Manipulation. In *Proceedings of the 26th Annual Network and Distributed System Security Symposium (NDSS 2019)*. Internet Society. https://doi.org/10.14722/ndss.2019.23386

[18] Levine, J. and M. Delany, "A "Null MX" No Service Resource Record for Domains That Accept No Mail", RFC 7505, DOI 10.17487/RFC7505, June 2015, <https://www.rfc-editor.org/info/rfc7505>.

[19] Mozilla Foundation. n.d. Public Suffix List. Dataset, version 2026-11-11_13-52-29_UTC (commit 7786aac8931b7069d6fe2174a7471a54d273d3e6), ICANN and private sections, SHA-256 55d282c86a00f85cd5c686c91e8ba26a3b156154640d93853d3d389a5e949b07. publicsuffix.org. Retrieved November 11, 2026 from https://publicsuffix.org/

[20] Robert G. Newcombe. 1998. Interval estimation for the difference between independent proportions: comparison of eleven methods. *Statistics in Medicine* 17, 8 (1998), 873–890. https://doi.org/10.1002/(SICI)1097-0258(19980430)17:8<873::AID-SIM779>3.0.CO;2-I

[21] Stephen J. Nightingale. 2017. Email Authentication Mechanisms: DMARC, SPF and DKIM. NIST Technical Note 1945. National Institute of Standards and Technology. https://doi.org/10.6028/NIST.TN.1945

[22] Kimberly Ruth, Aurore Fass, Jonathan Azose, Mark Pearson, Emma Thomas, Caitlin Sadowski, and Zakir Durumeric. 2022. A World Wide View of Browsing the World Wide Web. In *Proceedings of the 22nd ACM Internet Measurement Conference (IMC '22)*. ACM, 317–336. https://doi.org/10.1145/3517745.3561418

[23] Kimberly Ruth, Deepak Kumar, Brandon Wang, Luke Valenta, and Zakir Durumeric. 2022. Toppling Top Lists: Evaluating the Accuracy of Popular Website Lists. In *Proceedings of the 22nd ACM Internet Measurement Conference (IMC '22)*. ACM, 374–387. https://doi.org/10.1145/3517745.3561444

[24] Dennis Tatang, Florian Zettl, and Thorsten Holz. 2021. The Evolution of DNS-based Email Authentication: Measuring Adoption and Finding Flaws. In *Proceedings of the 24th International Symposium on Research in Attacks, Intrusions and Defenses (RAID '21)*. ACM, 354–369. https://doi.org/10.1145/3471621.3471842

[25] Tranco. 2026. Tranco list 947N2. Dataset. Retrieved November 11, 2026 from https://tranco-list.eu/list/947N2

[26] Tranco. 2026. Tranco list K9PXW. Dataset. Retrieved November 13, 2026 from https://tranco-list.eu/list/K9PXW

[27] Edwin B. Wilson. 1927. Probable Inference, the Law of Succession, and Statistical Inference. *Journal of the American Statistical Association* 22, 158 (1927), 209–212. https://doi.org/10.1080/01621459.1927.10502953