Email Authentication at Scale: A Systematic Review and 2026 Measurement of SPF and DMARC Deployment

Francis Davison

Founder, SpamCipher · https://spamcipher.com/insights/email-authentication-spf-dmarc-2026

Preprint · SpamCipher Insights · July 2026 · Systematic review + measurement

ABSTRACT

Email authentication through SPF, DKIM, and DMARC is a precondition for inbox delivery, but its deployment has been reported in scattered studies with incompatible denominators, and no measurement covered the period after the February 2024 Google and Yahoo bulk-sender requirements. This paper reports three linked measurements. We actively scanned a random sample of the Tranco top million in July 2026, extracted a single-source longitudinal series from the OpenINTEL DNS measurement platform for 2022 through 2026, and systematically reviewed the measurement literature from 2015 to 2026. In 2026, 62.8 percent of resolvable top-million domains published SPF (95 percent confidence interval 62.2 to 63.4) and 47.2 percent published DMARC (46.6 to 47.9); the OpenINTEL series agreed with the scan to within 1.5 percentage points from separate infrastructure. SPF deployment remained near 60 percent from 2022 onward. DMARC deployment held near 25 to 27 percent through January 2024 and reached 34.7 percent by July 2024, roughly four times the preceding rate of increase; the acceleration therefore falls within the six months containing the mandate. The share of DMARC domains with an enforcing policy was already about 47 percent by 2023 and remained near half through 2026, so the increase in the mandate window came from new adopters and the enforcing share did not rise. About a fifth of sampled domains returned no definitive DMARC answer, which places a lower bound of approximately 38 percent on the population DMARC rate. Inbox placement as a function of authentication has been measured by one study, which concerns forged mail. The scan code, extraction code, and data are released with the paper.

Keywords: email authentication, SPF, DKIM, DMARC, sender policy, bulk sender requirements, DNS measurement, inbox placement

1. Introduction

How widely email authentication is deployed, how it is configured, and how strictly it is enforced are basic quantities for mail operators, domain administrators, and security researchers. SPF, DKIM, and DMARC are prerequisites for inbox delivery at the major mailbox providers, and a companion analysis found that passing authentication is necessary but not sufficient for reaching the inbox [1]. Deployment of these protocols has been measured in snapshots taken in different years, over different domain populations, and from different vantage points, and the rates those studies report are not directly comparable. In addition, the February 2024 requirement that bulk senders to Gmail and Yahoo authenticate their mail and publish a DMARC record [8, 44, 45] postdates almost all of the published measurements.

This paper measures the current state of SPF and DMARC deployment and relates it to the prior literature. We addressed three questions: what fraction of top-million domains publish SPF and DMARC in 2026; how deployment changed between 2015 and 2026, in particular around the February

2024 mandate; and how policy strictness and configuration quality developed over the same period. We actively scanned a random sample of the Tranco top million in July 2026, recording SPF and DMARC publication, DMARC policy, and misconfiguration signals. We extracted a single-source longitudinal series from OpenINTEL, a long-running academic DNS measurement platform, covering 2022 to 2026 and bracketing the mandate. We also systematically reviewed the measurement literature from 2015 to 2026, holding population and vantage fixed within each comparison.

In the 2026 scan, 62.8 percent of resolvable top-million domains published SPF and 47.2 percent published DMARC, and the OpenINTEL series agreed with both figures to within about 1.5 percentage points. SPF deployment has remained near 60 percent since 2022. DMARC deployment rose roughly an order of magnitude over the decade, and the longitudinal series places its acceleration in the six months containing the mandate, when adoption moved from about 27 to 35 percent at roughly four times the prior rate. The share of DMARC domains with an enforcing policy reached about half by 2023 and has remained there. Because this is an observational measurement, we report the timing of these changes and do not attribute cause. The contributions are: (1) a measurement of top-million SPF and DMARC deployment after the 2024 mandate, with confidence intervals; (2) a single-source longitudinal series that times the DMARC acceleration and cross-validates the scan from separate infrastructure; and (3) a systematic review that reports each prior rate with its population, vantage, and year. Inbox placement as a function of authentication is covered by a single study of forged mail; Section 6 discusses this gap. We release the scan code, the OpenINTEL extraction, and the raw data.

2. Background

SPF, DKIM, and DMARC are DNS-published records that a receiving server can check to decide whether a message is authorized to use the domain in its headers.

SPF (Sender Policy Framework, RFC 7208) lets a domain publish which hosts may send mail in its name, in a TXT record at the domain apex that ends in an `all` mechanism whose qualifier sets the strictness: `-all` (hardfail), `~all` (softfail), or the permissive `+all` / `?all` [2]. DKIM (DomainKeys Identified Mail, RFC 6376) attaches a cryptographic signature to each message, validated against a public key published under a selector in DNS [3]. DMARC (RFC 7489, now superseded by RFC 9989) builds on the two: it ties authentication to the visible From domain and publishes, at `_dmarc.<domain>`, a policy of `none`, `quarantine`, or `reject`, together with optional aggregate (`rua`) and failure (`ruf`) reporting [4, 5]. Only an enforcing policy (quarantine or reject) instructs receivers to act on a failure; `p=none` requests monitoring only. The 2024 bulk-sender requirements pair DMARC publication with one-click unsubscribe, specified in RFC 8058 [36], and with low complaint rates. The Authenticated Received Chain [37] preserves authentication results across forwarding; it is related to these mechanisms but is not required by the mandate.

These mechanisms give rise to three distinct measurement targets: whether a domain publishes a record, whether a receiving server validates one, and whether an authenticated message reaches the inbox. Rates for one target do not transfer to another, and this paper reports each separately.

3. Data and Methodology

We combined an active scan, a longitudinal extraction, and a systematic review. Figures measured by us are labeled as such throughout and kept distinct from cited figures. The scan resolves only public DNS

records, sends no email, and probes no host; the ethical exposure is limited to DNS query load, which we bounded with low query rates.

3.1 Active measurement of the Tranco top million

We drew a seeded random sample from the Tranco top-million list (pinned list 94VW2, generated 2026-07-05) [32]. The unit of measurement is the registrable domain, fixed by the Public Suffix List [38]. For each domain we resolved the apex TXT record for SPF and the `_dmarc` TXT record for DMARC over six public anycast resolvers. A domain was classified as resolvable if its apex query returned NOERROR, with or without a TXT record; a resolvable domain without an SPF record was counted as not publishing SPF, not as missing data. NXDOMAIN responses and timeouts were recorded separately and were never counted as absence of a record. We ran the scan twice with independent seeds to bound sampling variance; because the seeds draw different domains, the two runs bound sampling error, not vantage effects. All queries were issued from a single client origin, a vantage limitation discussed in Section 7. Records were parsed with the `checkdmarc` and `pyspf` libraries, and all proportions carry Wilson 95 percent confidence intervals [31]. Two constructs were excluded from the scan. DKIM cannot be measured without the non-public selector, so we report it from the literature only. Evaluating the SPF ten-lookup limit requires recursive expansion of include mechanisms, which we did not perform; we report the multiple-record and permissive-`all` misconfiguration signals instead. The independent variables of the analysis are calendar year, vantage, and population; the dependent variables are the deployment, policy, and misconfiguration rates. Top-list popularity bias is addressed by subgrouping results by population, denominator definition by Public Suffix List normalization, and resolver vantage by the multi-resolver set; DKIM selector blindness cannot be mitigated by scan design, so DKIM is reported from the literature. The pinned list identifier makes the sample reproducible.

3.2 OpenINTEL longitudinal series

To place the 2026 snapshot in a time series, we analyzed OpenINTEL, an academic active-DNS measurement platform that publishes daily forward-DNS records for the Tranco top million [33, 46]. For one snapshot in each year from 2022 through 2026, plus a pre-mandate and a post-mandate snapshot, we queried the public Parquet files with DuckDB [47] and computed the SPF and DMARC deployment shares and the DMARC policy distribution over the same-day SOA-apex domain set. The platform began querying the `_dmarc` name in mid-2023, so the DMARC series starts there. Each snapshot covers the full list of about one million domains, so sampling error is negligible, and the series provides a cross-check on our scan from separate infrastructure. Use of the data is non-commercial research under the CC BY-NC-SA license, with attribution given in the Declarations.

3.3 The prior-measurement corpus

We searched the ACM Digital Library, IEEE Xplore, USENIX open-access proceedings, dblp, and Crossref for measurement studies reporting SPF, DKIM, or DMARC deployment, configuration, enforcement, or placement between 2015 and 2026. Reporting follows PRISMA [26] and the systematic-review guidelines for computing research [27, 28]. We included peer-reviewed studies that report a numeric rate and excluded vendor and opinion material, which is used only as labeled context. Rates were extracted from the primary publications, and each reference was checked against Crossref, the publisher, USENIX, or the RFC Editor. Because reported rates depend on population and vantage, we did not pool a grand mean; each study is reported with its population, vantage, and year, and trends are read only within a fixed subgroup. A formal proportion meta-analysis was considered and rejected: the

standard random-effects estimator [39, 29] with a variance-stabilizing transformation [41] assumes denominator-comparable inputs that are not available across a decade of different populations, the transformation is contested under such heterogeneity [30], and the heterogeneity statistic [40] would be near its ceiling.

4. SPF and DMARC Deployment

4.1 Cross-sectional rates in 2026

Our 2026 scan finds SPF on 62.8 percent of resolvable top-million domains (95 percent CI 62.2 to 63.4; two-run pooled, n = 25,855) and DMARC on 47.2 percent (46.6 to 47.9; n = 23,961 with a definitive `_dmarc` answer). The OpenINTEL 2026 snapshot, measured independently over the full list, gives 61.3 and 45.9 percent, within about 1.5 and 1.3 points of the scan. About 13 percent of sampled domains returned no definitive apex answer and about 20 percent returned no definitive DMARC answer. Non-responding domains are weighted toward parked and defunct entries, which are less likely to publish records, so the definitive-answer DMARC rate is an upper estimate; treating every non-responder as not publishing DMARC gives a lower bound near 38 percent, and the population rate therefore lies between roughly 38 and 47 percent.

4.2 Adoption from 2015 to 2026

Table 1 and Table 2 place the 2026 rates against the prior literature, and Figure 1 plots the top-list series. Across the top-list studies, SPF rose from about 37 percent to about 50 percent between 2015 and 2020, and the OpenINTEL series holds between 57 and 64 percent across 2022 to 2026. DMARC rose roughly an order of magnitude, from 1.1 percent in 2015 [6] to 47.2 percent in our 2026 scan. The cross-study points cannot resolve when the rise accelerated, because populations and denominators differ between studies; Section 4.3 addresses the timing with the single-source series.

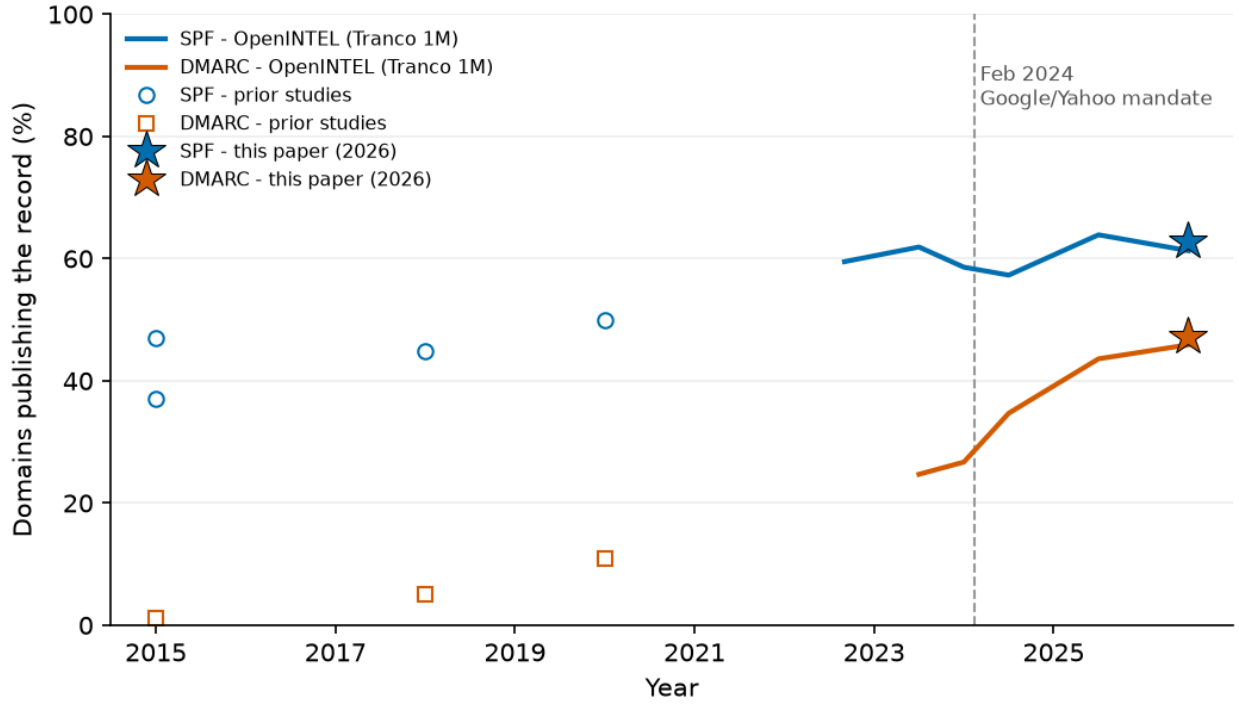

Figure 1. SPF and DMARC deployment among Tranco top-million domains, 2015 to 2026, as the percentage of domains publishing the record. Open markers show prior top-list studies (Table 1 and Table 2); solid lines show the single-source OpenINTEL series (Table 3); stars show this paper's 2026 active scan. The dashed vertical line marks February 2024, when the Google and Yahoo bulk-sender requirements took effect.

Table 1. SPF deployment, sender-side record publication, curated top-lists. The row for this paper is measured, with a 95 percent CI; the other rows are cited.

Study	Year	Population	SPF
Gojmerac et al. [11]	2015	Alexa 1M	~37%
Durumeric et al. [6]	2015	top-1M MX-filtered	~47%
Hu, Peng, Wang [12]	2018	Alexa 1M	44.9%
Tatang et al. [13]	2020	top-1M	~50%
Czybik et al. [14]	2023	Tranco 12M	56.5%
This paper (active scan)	2026	Tranco 1M sample	62.8% (62.2-63.4)

Table 2. DMARC deployment. Full-zone censuses are shown for contrast and are not pooled with top-lists; denominators are labeled because they are not matched across rows.

Study	Year	Population (denominator)	DMARC
Durumeric et al. [6]	2015	top-1M, MX-filtered	1.1%
Hu, Peng, Wang [12]	2018	Alexa 1M	5.1%
Tatang et al. [13]	2020	top-1M	~11%
Maroofi et al. [15]	2021	236M census	0.13%
Maroofi et al. [15]	2021	country top-500	34.3%
This paper (active scan)	2026	Tranco 1M, all resolvable	47.2% (46.6-47.9)

4.3 DMARC adoption around the February 2024 mandate

In the OpenINTEL series (Table 3), DMARC deployment moved from 24.7 percent in July 2023 to 26.7 percent in January 2024, reached 34.7 percent by July 2024, and continued to 45.9 percent by July 2026. The increase in the half year from January to July 2024, about eight points, is roughly four times the increase in the preceding half year, and that interval contains the February 2024 mandate. Our 2026 scan (Section 4.1) is consistent with the series endpoint.

Table 3. Single-source OpenINTEL Tranco series (full list, about 1M SOA-apex domains per snapshot). DMARC begins mid-2023, when the platform began querying `_dmarc`. Enforcing = quarantine plus reject as a share of DMARC domains.

OpenINTEL snapshot	SPF	DMARC	DMARC enforcing	rua
2022-09	59.5%	(not yet measured)	-	-
2023-07	61.9%	24.7%	47.4%	73.3%
2024-01 (pre-mandate)	58.6%	26.7%	49.2%	73.7%
2024-07 (post-mandate)	57.3%	34.7%	43.4%	67.6%
2025-07	63.9%	43.6%	45.8%	65.9%
2026-07	61.3%	45.9%	49.3%	65.0%

5. Policy Enforcement and Configuration

5.1 DMARC policy distribution over time

Early DMARC deployment was largely monitoring-only, and the enforcing share among publishing domains has risen since. Tatang et al. found about 70 percent of DMARC domains at `p=none` in 2020, implying roughly 30 percent enforcing, of which about 15 percent used reject [13]. Between 2021 and 2023, Maroofi et al. and Hureau et al. measured enforcing shares near 50 and 44 percent [15, 18]. Our 2026 scan finds 50.9 percent at none, 25.1 percent at quarantine, and 23.9 percent at reject (n = 11,318; 95 percent CIs about one point wide), an enforcing share of about 49 percent. The OpenINTEL series (Table 3, Figure 2) shows that the enforcing share was already about 47 percent by mid-2023 and stayed between 43 and 49 percent through 2026; growth in enforcement occurred mainly over 2020 to 2023 and then plateaued. Domains adopting DMARC in the mandate window skewed toward monitoring-only policies: the `p=none` share rose to 56.5 percent in July 2024 and returned to near 50 percent by 2026. Among the DMARC domains in our scan, 64.8 percent configure aggregate reporting (`rua`),

compared with the 49 percent Ashiq et al. measured on the .com, .net, and .org zones in 2023 [17]; the higher rate here is in line with a popularity-weighted sample.

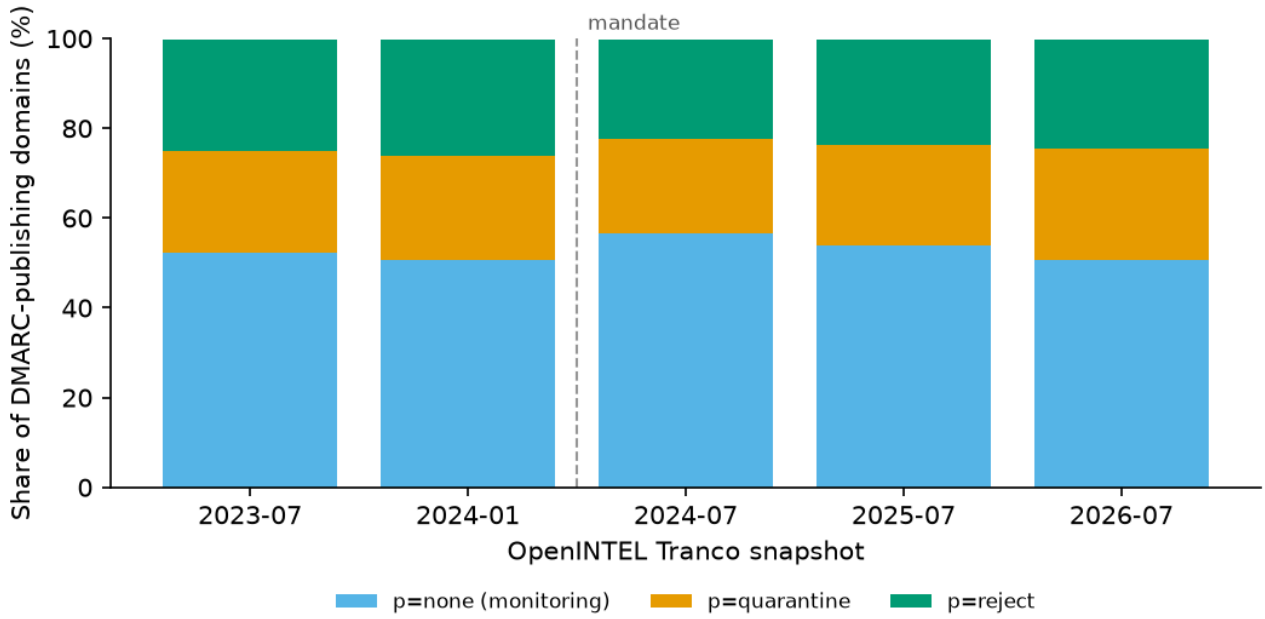

Figure 2. DMARC policy distribution among DMARC-publishing domains in the OpenINTEL Tranco snapshots (Table 3), as percentage shares that sum to 100. Bars stack p=none, p=quarantine, and p=reject; the enforcing share is quarantine plus reject. The dashed line marks the February 2024 mandate.

5.2 Configuration quality and validation

Czybik et al. found that 2.9 percent of SPF records contain errors and 34.7 percent authorize more than 100,000 IP addresses, with 5.9 percent lacking a restrictive `all` [14]; overly broad and permissive SPF also carries the amplification exposure documented by Scheffler et al. [43]. In our scan, of SPF-publishing domains (n = 16,229), 38.7 percent use `-all` (hardfail), 55.7 percent use `~all` (softfail), about 3.0 percent are permissive, and 1.6 percent carry the multiple-record error that invalidates SPF. DKIM detection lower bounds are about 28 percent in both 2015 (Gmail-observed [6]) and 2022 (selector-probed [16]); the 2022 figure is a floor because selectors are not public. Wang et al. found that 84 percent of DKIM-signing domains use keys of 1024 bits or shorter [16]. On the receiver side, Deccio et al. found that about 85 percent of mail servers validate SPF and 53 percent validate all three protocols, and that more than 25 percent violate the SPF ten-lookup limit [19]; Zhang et al. found that roughly 27 percent of servers cannot verify inbound SPF and DMARC at all [20].

6. Discussion

The results are conditioned by three variables that differ across the literature. The first is population: a rate measured on a curated top list is not comparable to a rate measured on a full zone. SPF has been reported at 31 percent on a large full-zone census [15] and at 56.5 percent on a twelve-million-domain top list [14]; both are internally valid, and the difference lies in the denominator. The second is vantage: publishing a record (about 63 percent of resolvable top-million domains for SPF in our scan) is distinct from validating one (about 85 percent of mail servers [19]), and both are distinct from inbox delivery. The third is enforcement: adoption counts records, and only an enforcing policy changes receiver behavior, so the 2026 state depends on the policy distribution as well as on the count.

The longitudinal series supports one time-resolved statement: DMARC adoption accelerated in the six months containing the February 2024 mandate, while the enforcing fraction did not change, so the mandate window coincides with an inflow of new, mostly monitoring-only adopters. We do not attribute cause. Isolating a causal effect would require a control population unaffected by the mandate and measured over the same window, and several co-occurring factors, including concerted provider action, improved tooling, and a pre-existing upward trend, cannot be separated in these data. The direction of the finding agrees with Kondo et al., who measured DMARC rising about five points over February to May 2024 in the opening months of the mandate [9].

Configuration and validation lag publication. Best-common-practice guidance has recommended strict alignment and an enforcing policy since at least 2020 [35], while the results in Section 5.2 show high rates of over-broad SPF records [14], short DKIM keys [16], and incomplete receiver-side validation [19, 20].

The literature measures deployment from several vantages but contains little on placement. The one end-to-end study of authentication and inbox placement, by Hu and Wang, tested forged mail and found that a strict policy on the spoofed domain limits impersonation, although 34 of 35 providers delivered at least one forged message to the inbox [10]. That study measures spoofing defense, not the placement of legitimate authenticated mail, and no comparable study of legitimate mail exists; vendor reports on this question do not disclose their methods or seed lists. This is the gap identified in the companion analysis [1], and the review found no study that closes it.

7. Limitations

Our scan and most of the cited literature sample popularity-ranked lists, which over-represent large active domains; the reported rates are therefore upper bounds on the full domain space. The scan resolved from one client origin at one point in time over public anycast resolvers, so it can miss geo-split or CDN-specific records; the two runs bound sampling variance only, not vantage. About a fifth of sampled domains returned no definitive DMARC answer; these are reported separately, and treating them all as non-publishers bounds the population DMARC rate below near 38 percent. DKIM and the SPF ten-lookup limit are outside the scan and are reported from the literature, so those figures carry the populations and years of the cited studies. The OpenINTEL series is a single homogeneous source covering 2022 to 2026, but its `_dmarc` measurement begins in mid-2023, so the DMARC path from about 11 percent in 2020 to 25 percent in 2023 rests on cross-study comparison rather than on a same-source series; OpenINTEL is also a fixed-vantage platform and shares the geo-vantage caveat at full-list scale. Placement is synthesized from a single forged-mail study; this paper contributes no placement measurement of its own.

8. Related Work

Large-scale measurement of email security began with Durumeric et al., who combined Gmail SMTP logs with Internet-wide scans to report transport and authentication adoption at top-million scale [6], and Foster et al., who measured provider-side support and enforcement [7]. These studies predate the growth of DMARC and the 2024 mandate; our scan updates their adoption estimates by a decade on a comparable population.

A second group measured individual protocols and constructs: SPF configuration at twelve-million-domain scale [14] and vulnerabilities in SPF validators [21]; DKIM deployment and key strength [16]; and DMARC adoption over time [13], its reporting ecosystem [17], obstacles to its wider deployment [18], and

analysis of its reports [25]. Each covers one protocol or one construct in one period; our measurement reports SPF and DMARC jointly on a current sample with a single denominator definition.

A third group measured at census scale: anti-spoofing adoption across hundreds of millions of domains with country and sectoral breakdowns [15], and protection of high-profile domains and their subdomains [24]. These use full-zone denominators, which we report alongside top-list rates but never pool with them.

A fourth group measured the receiving side: validation behavior at mail servers [19, 20], and authentication bypasses that arise when components are composed inconsistently [22, 23]. These studies measure validation; our scan measures publication, and Section 5.2 relates the two sides.

A parallel line measured the transport security of email delivery [34, 42]. Placement has one end-to-end measurement, of forged mail [10], and Kondo et al. measured implementation in the months immediately after the sender guidelines took effect [9]. Our study differs from all of these in measuring top-million SPF and DMARC after the 2024 mandate, in adding a single-source longitudinal series that times the DMARC acceleration and cross-validates an independent scan, and in reporting every rate within a fixed population and vantage subgroup.

9. Conclusion

This paper set out to measure SPF and DMARC deployment on the top million domains after the February 2024 bulk-sender mandate and to relate the result to a decade of prior measurement. On the first question, 62.8 percent of resolvable top-million domains publish SPF and 47.2 percent publish DMARC in 2026, with agreement to within 1.5 points between two independent infrastructures. On the second, SPF deployment has been near 60 percent since 2022, while DMARC deployment accelerated in the six months containing the mandate and reached 45.9 percent in the July 2026 OpenINTEL snapshot. On the third, about half of DMARC domains enforce, a share reached by 2023 and stable since, while configuration errors, over-broad SPF records, and short DKIM keys remain common. A controlled measurement of authentication and reputation against live inbox placement for legitimate senders is not available in the literature; it is planned as the next study in this series.

Declarations

Competing interests. The author is the founder of SpamCipher, a commercial cold email and deliverability platform. This is a material competing interest. The analysis rests on primary sources, our own reproducible measurement, and public data; no product is evaluated, recommended, or required by any result.

Author contributions. Francis Davison is the sole author and is responsible for the conception, the measurement and its code, the systematic review and its verification, the analysis, and the accuracy of every figure and value.

Data and code availability. The pinned Tranco list identifier, resolver set, scan code, sampled domains, raw per-domain results, the OpenINTEL extraction and derived series, and the figure-generation code are released with this paper. The cited literature and the OpenINTEL data are public.

Acknowledgment. The longitudinal analysis was made possible by OpenINTEL (https://www.openintel.nl/), a joint project of the University of Twente, SIDN, NLnet Labs and SURF; OpenINTEL open-access data is licensed CC BY-NC-SA 4.0 and used here for non-commercial research.

Funding. None external; produced by SpamCipher.

References

[1] F. Davison, "Gmail's 2026 Inbox Decision for High-Volume Cold Email: A Three-Layer Authentication, Reputation, and Engagement Model," SpamCipher Insights, 2026. https://spamcipher.com/insights/how-gmail-decides-inbox-vs-spam

[2] S. Kitterman, "Sender Policy Framework (SPF) for Authorizing Use of Domains in Email, Version 1," RFC 7208, IETF, 2014. https://www.rfc-editor.org/info/rfc7208

[3] D. Crocker, T. Hansen, M. Kucherawy (Eds.), "DomainKeys Identified Mail (DKIM) Signatures," RFC 6376 (STD 76), IETF, 2011. https://www.rfc-editor.org/info/rfc6376

[4] T. Herr, J. Levine (Eds.), "Domain-based Message Authentication, Reporting, and Conformance (DMARC)," RFC 9989, IETF, 2026 (obsoletes RFC 7489 and RFC 9091). https://www.rfc-editor.org/info/rfc9989

[5] M. Kucherawy, E. Zwicky (Eds.), "Domain-based Message Authentication, Reporting, and Conformance (DMARC)," RFC 7489 (Informational, obsoleted), IETF, 2015. https://www.rfc-editor.org/info/rfc7489

[6] Z. Durumeric, D. Adrian, A. Mirian, J. Kasten, E. Bursztein, N. Lidzborski, K. Thomas, V. Eranti, M. Bailey, J. A. Halderman, "Neither Snow Nor Rain Nor MITM...: An Empirical Analysis of Email Delivery Security," ACM IMC 2015. https://doi.org/10.1145/2815675.2815695

[7] I. Foster, J. Larson, M. Masich, A. Snoeren, S. Savage, K. Levchenko, "Security by Any Other Name: On the Effectiveness of Provider Based Email Security," ACM CCS 2015. https://doi.org/10.1145/2810103.2813607

[8] Google, "Email sender guidelines," Gmail Help. Accessed 2026-07-06. https://support.google.com/a/answer/81126

[9] D. Kondo, Y. Shibuya, R. S. Yamaguchi, T. Ishihara, Y. Sekiya, T. Nakata, T. Asami, "Who Did Not Implement Email Security Measures After Google's New Email Sender Guidelines? A Large-Scale Measurement Study," IFIP TMA 2025. https://doi.org/10.23919/TMA66427.2025.11097004

[10] H. Hu, G. Wang, "End-to-End Measurements of Email Spoofing Attacks," USENIX Security 2018. https://www.usenix.org/conference/usenixsecurity18/presentation/hu

[11] I. Gojmerac, P. Zwickl, G. Kovacs, C. Steindl, "Large-Scale Active Measurements of DNS Entries Related to E-Mail System Security," IEEE ICC 2015. https://doi.org/10.1109/ICC.2015.7249513

[12] H. Hu, P. Peng, G. Wang, "Towards Understanding the Adoption of Anti-Spoofing Protocols in Email Systems," IEEE SecDev 2018. https://doi.org/10.1109/SecDev.2018.00020

[13] D. Tatang, F. Zettl, T. Holz, "The Evolution of DNS-based Email Authentication: Measuring Adoption and Finding Flaws," RAID 2021. https://doi.org/10.1145/3471621.3471842

[14] S. Czybik, M. Horlboge, K. Rieck, "Lazy Gatekeepers: A Large-Scale Study on SPF Configuration in the Wild," ACM IMC 2023. https://doi.org/10.1145/3618257.3624827

[15] S. Maroofi, M. Korczynski, A. Holzel, A. Duda, "Adoption of Email Anti-Spoofing Schemes: A Large Scale Analysis," IEEE Transactions on Network and Service Management 18(3), 2021. https://doi.org/10.1109/TNSM.2021.3065422

[16] C. Wang, K. Shen, M. Guo, Y. Zhao, M. Zhang, J. Chen, B. Liu, X. Zheng, H. Duan, Y. Lin, Q. Pan, "A Large-scale and Longitudinal Measurement Study of DKIM Deployment," USENIX Security 2022. https://www.usenix.org/conference/usenixsecurity22/presentation/wang-chuhan

[17] M. I. Ashiq, W. Li, T. Fiebig, T. Chung, "You've Got Report: Measurement and Security Implications of DMARC Reporting," USENIX Security 2023. https://www.usenix.org/conference/usenixsecurity23/presentation/ashiq

[18] O. Hureau, J. Bayer, A. Duda, M. Korczynski, "Spoofed Emails: An Analysis of the Issues Hindering a Larger Deployment of DMARC," PAM 2024, LNCS 14538. https://doi.org/10.1007/978-3-031-56249-5_10

[19] C. Deccio, T. Yadav, N. Bennett, A. Hilton, M. Howe, T. Norton, J. Rohde, E. Tan, B. Taylor, "Measuring Email Sender Validation in the Wild," ACM CoNEXT 2021. https://doi.org/10.1145/3485983.3494868

[20] H. Zhang, L. Chen, M. Liu, Y. Shi, S. Wu, Z. Xue, "Both Sides Needed: A Two-Dimensional Measurement Study of Email Security Based on SPF and DMARC," IEEE MSN 2023. https://doi.org/10.1109/MSN60784.2023.00126

[21] N. Bennett, R. Sowards, C. Deccio, "SPFail: Discovering, Measuring, and Remediating Vulnerabilities in Email Sender Validation," ACM IMC 2022. https://doi.org/10.1145/3517745.3561468

[22] J. Chen, V. Paxson, J. Jiang, "Composition Kills: A Case Study of Email Sender Authentication," USENIX Security 2020. https://www.usenix.org/conference/usenixsecurity20/presentation/chen-jianjun

[23] K. Shen, C. Wang, M. Guo, X. Zheng, C. Lu, B. Liu, Y. Zhao, S. Hao, H. Duan, Q. Pan, M. Yang, "Weak Links in Authentication Chains: A Large-scale Analysis of Email Sender Spoofing Attacks," USENIX Security 2021. https://www.usenix.org/conference/usenixsecurity21/presentation/shen-kaiwen

[24] S. Maroofi, M. Korczynski, A. Duda, "From Defensive Registration to Subdomain Protection: Evaluation of Email Anti-Spoofing Schemes for High-Profile Domains," IFIP TMA 2020.

[25] K. Konno, N. Kitagawa, N. Yamai, "False Positive Detection in Sender Domain Authentication by DMARC Report Analysis," ACM ICISS 2020. https://doi.org/10.1145/3388176.3388217

[26] M. J. Page, J. E. McKenzie, P. M. Bossuyt, et al., "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," BMJ 372:n71, 2021. https://doi.org/10.1136/bmj.n71

[27] B. Kitchenham, S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering," EBSE Technical Report EBSE-2007-01, 2007.

[28] W. Fei, H. Ohno, S. Sampalli, "A Systematic Review of IoT Security: Research Potential, Challenges, and Future Directions," ACM Computing Surveys 56(5), 2023. https://doi.org/10.1145/3625094

[29] J. J. Barendregt, S. A. Doi, Y. Y. Lee, R. E. Norman, T. Vos, "Meta-analysis of prevalence," Journal of Epidemiology and Community Health 67(11), 2013. https://doi.org/10.1136/jech-2013-203104

[30] G. Schwarzer, H. Chemaitelly, L. J. Abu-Raddad, G. Rucker, "Seriously misleading results using inverse of Freeman-Tukey double arcsine transformation in meta-analysis of single proportions," Research Synthesis Methods 10(3), 2019. https://doi.org/10.1002/jrsm.1348

[31] E. B. Wilson, "Probable Inference, the Law of Succession, and Statistical Inference," Journal of the American Statistical Association 22(158), 1927. https://doi.org/10.1080/01621459.1927.10502953

[32] V. Le Pochat, T. Van Goethem, S. Tajalizadehkhoob, M. Korczynski, W. Joosen, "Tranco: A Research-Oriented Top Sites Ranking Hardened Against Manipulation," NDSS 2019. https://doi.org/10.14722/ndss.2019.23386

[33] M. Jonker, A. Sperotto, R. van Rijswijk-Deij, R. Sadre, A. Pras, "Measuring the Adoption of DDoS Protection Services," ACM IMC 2016 (OpenINTEL platform). https://doi.org/10.1145/2987443.2987487

[34] O. van der Toorn, R. van Rijswijk-Deij, T. Fiebig, M. Lindorfer, A. Sperotto, "TXTing 101: Finding Security Issues in the Long Tail of DNS TXT Records," IEEE EuroS&P Workshops 2020. https://doi.org/10.1109/EuroSPW51379.2020.00080

[35] M3AAWG, "Email Authentication Recommended Best Practices," Messaging, Malware and Mobile Anti-Abuse Working Group, 2020. https://www.m3aawg.org/sites/default/files/m3aawg-email-authentication-recommended-best-practices-09-2020.pdf

[36] J. Levine, T. Herkula, "Signaling One-Click Functionality for List Email Headers," RFC 8058, IETF, 2017. https://www.rfc-editor.org/info/rfc8058

[37] K. Andersen, B. Long, S. Blank, M. Kucherawy (Eds.), "The Authenticated Received Chain (ARC) Protocol," RFC 8617, IETF, 2019. https://www.rfc-editor.org/info/rfc8617

[38] Mozilla Foundation, "Public Suffix List." Accessed 2026-07-06. https://publicsuffix.org/

[39] R. DerSimonian, N. Laird, "Meta-analysis in clinical trials," Controlled Clinical Trials 7(3), 1986. https://doi.org/10.1016/0197-2456(86)90046-2

[40] J. P. T. Higgins, S. G. Thompson, "Quantifying heterogeneity in a meta-analysis," Statistics in Medicine 21(11), 2002. https://doi.org/10.1002/sim.1186

[41] M. F. Freeman, J. W. Tukey, "Transformations Related to the Angular and the Square Root," Annals of Mathematical Statistics 21(4), 1950. https://doi.org/10.1214/aoms/1177729756

[42] F. Holzbauer, J. Ullrich, M. Lindorfer, T. Fiebig, "Not that Simple: Email Delivery in the 21st Century," USENIX ATC 2022. https://www.usenix.org/conference/atc22/presentation/holzbauer

[43] S. Scheffler, S. Smith, Y. Gilad, S. Goldberg, "The Unintended Consequences of Email Spam Prevention," PAM 2018, LNCS 10771. https://doi.org/10.1007/978-3-319-76481-8_12

[44] Yahoo, "Sender Best Practices," Yahoo Sender Hub. Accessed 2026-07-06. https://senders.yahooinc.com/best-practices/

[45] Google, "New Gmail protections for a safer, less spammy inbox," The Keyword, 2023. https://blog.google/products-and-platforms/products/gmail/gmail-security-authentication-spam-protection/

[46] R. van Rijswijk-Deij, M. Jonker, A. Sperotto, A. Pras, "A High-Performance, Scalable Infrastructure for Large-Scale Active DNS Measurements," IEEE Journal on Selected Areas in Communications 34(6), 2016. https://doi.org/10.1109/JSAC.2016.2558918

[47] M. Raasveldt, H. Muhleisen, "DuckDB: an Embeddable Analytical Database," ACM SIGMOD 2019. https://doi.org/10.1145/3299869.3320212

Competing interest: the author is the founder of SpamCipher, a commercial cold email and deliverability platform. This analysis is grounded only in primary and peer-reviewed sources. Canonical version: https://spamcipher.com/insights/email-authentication-spf-dmarc-2026