

WARMUP INBOX NETWORK · ORIGINAL BENCHMARK

The Deliverability Benchmark

What 20 million emails a month reveal about inbox placement, measured from real mail landing in real inboxes, not seed-list tests.

88.7%

true Gmail primary rate

2 wk

to stable placement

84.7%

of drops hit Gmail first

Day 15

the recovery cliff

Five numbers that define deliverability in 2026

Every figure in this report comes from first-party placement data across the Warmup Inbox network: real mail flowing to real inboxes at a volume no seed list can reach. Here is the headline picture.

98.9%

MISLEADING

Reported Gmail inbox rate by standard IMAP measurement, which reads as near-perfect and overstates the truth.

88.7%

THE REAL NUMBER

True Gmail primary-tab rate once you measure by label through OAuth. The gap is tab routing, not spam.

2 wks

Median time to stable placement. Inbox rate hits 95% in Week 1 and 98% by Week 2, not the folklore eight weeks.

84.7%

Share of degradation events that surface at Gmail first. Gmail is the earliest warning signal in email.

Day 15

THE ONE THAT CHANGES HOW YOU OPERATE

Day 15 is the recovery cliff. Catch a placement drop before it and recovery takes 7 to 34 days. Cross it and recovery time multiplies about **8x**, while the odds of recovering at all fall to a coin flip.

First-party placement data, not a seed list

Most deliverability statistics you have read come from **seed-list tests**: a few hundred planted addresses a vendor sends to and measures. Seed lists are small, synthetic, and easy to game, and they are the basis for most published deliverability reports.

This report is built differently. Every number here comes from real mail flowing to real inboxes inside the Warmup Inbox network, placement no seed list can see.

For current-state benchmarks we use roughly **18,600 active, customer-facing inboxes** on paid plans, excluding quarantined and internal test accounts. Across a 30-day window those inboxes exchanged about **20 million sends**. For historical trends we use the full network history: about 117,000 inboxes over 3.7 years.

One principle runs through everything: we report what we can actually see, and we flag what we cannot.

18.6k

Active customer inboxes in the current-state sample

20M

Monthly sends measured across a 30-day window

3.7yr

Of continuous network placement history

Where a number is easy to misread, we say so, and give the honest version.

That is what makes this a benchmark you can plan against, rather than a marketing figure.

The Gmail measurement gap

Read quickly, standard measurement says Gmail almost never filters warmup mail, a 98.9% inbox rate. It doesn't. Gmail applies labels (Promotions, Updates) to messages that stay in the inbox folder, and an IMAP reader counts that tab-routed mail as inbox-landed. Measured honestly through OAuth, the real primary-inbox rate is **88.7%**. The missing 9 points are tab routing, not spam.

Gmail placement: what IMAP reports vs. what actually lands in the primary tab

Share of Gmail volume · OAuth subset: 379 inboxes, 224K sends

Standard IMAP: reads everything as
inbox

98.9%

counted as inbox

Measured by label: OAuth,
honest

88.7%

primary inbox

Primary inbox 88.7% Category tabs 9.2% Spam 2.0%

GMAIL AND OUTLOOK BEHAVE THE SAME WAY

Gmail → category tabs 9.2%

Outlook → Other tab 12.4%

Both major inboxes route roughly a tenth of bulk mail to secondary tabs. This is normal industry behavior, **not a Warmup Inbox problem.**

WHAT TO DO WITH IT

Hold your own reporting to the same standard. If your dashboard reads 99% at Gmail, it is measuring folders, not tabs. Plan around **~89% primary placement**, with another ~9% reachable in secondary tabs.

Source: Warmup Inbox network, 30-day window. Gmail honest-measurement subset via Google API label access.

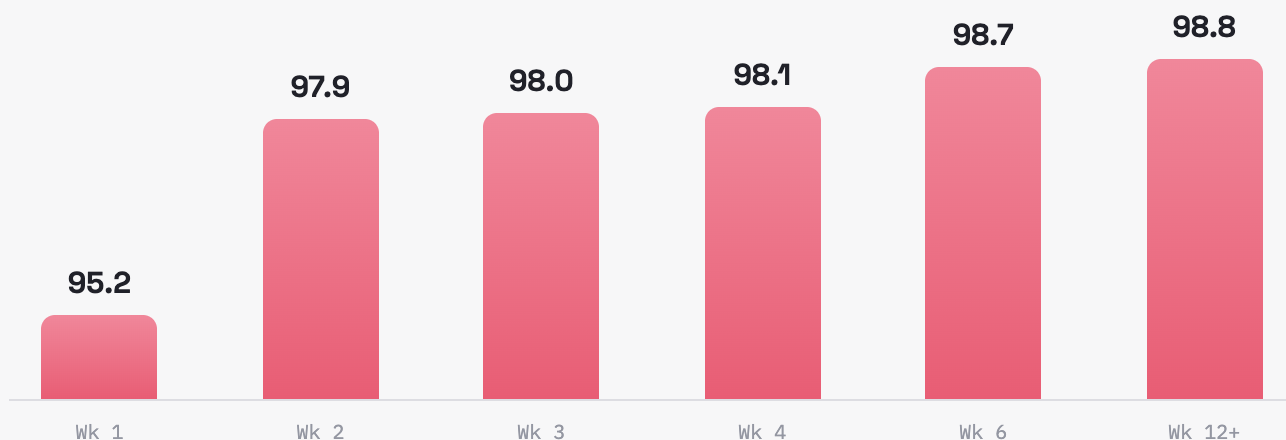
Warmup is done in two weeks, not eight

The industry rule of thumb is that warmup takes eight weeks or more. The data does not support it.

Bucketing 17,247 inboxes by weeks since warming started, the median inbox rate is already **95% in Week 1** and **98% by Week 2**. The gains come fast, but they are not permanent. Placement holds only as long as warming continues.

Median inbox rate by weeks since warming started
17,247 in-scope inboxes • snapshot methodology

y-axis 94–99.5% • IMAP measurement



Predictability, not placement, is what improves

In Week 1 inboxes scatter across a 6.5-point band; by Week 12 that tightens to about **1 point**. Warmup gets you to the right neighborhood in two weeks and makes you tightly predictable over the following months.

Spam collapses fast and stays low

Median spam rate falls from **1.77% in Week 1** to ~0.4% by Week 5 and just **0.24%** at Week 12+. Warmed inboxes rarely see the spam folder at all.

MAINTENANCE

Warmup is a habit, not a one-time fix

The two-week result is real, and like the gym, it only lasts while you keep it up. Stop warming and reputation drifts back toward where it started; providers re-assess sending behavior continuously, not once. Ongoing warming is what holds an inbox in that tight 98–99% band month after month.

Source: Warmup Inbox network. Intermediate weeks omitted for space; the curve is unchanged.

Gmail is the canary, and the harshest filter

When an inbox starts to degrade, Gmail reacts first. Across 1,292 degradation events with a detectable per-provider spike, Gmail's spam rate spiked first **84.7%** of the time, about a day ahead of Outlook and three days ahead of long-tail providers. And it hits hardest.

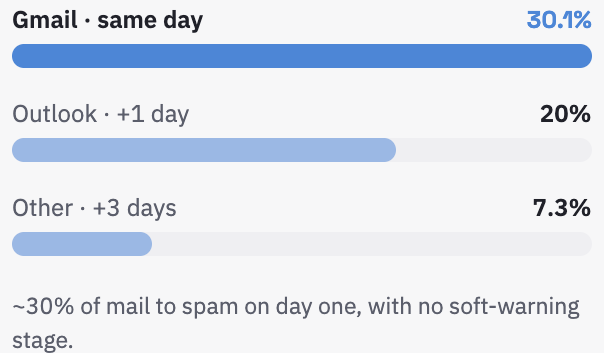
Which provider flips first

Share of degradation events



...and how hard it hits

Median spam rate at the moment of the flip



WHAT TO DO WITH IT

Make Gmail placement your primary deliverability KPI. By the time Outlook or a smaller provider reflects a problem, Gmail has been showing it for one to three days.

Source: 92-day per-send placement records; 1,292 events with a detectable per-provider spike.

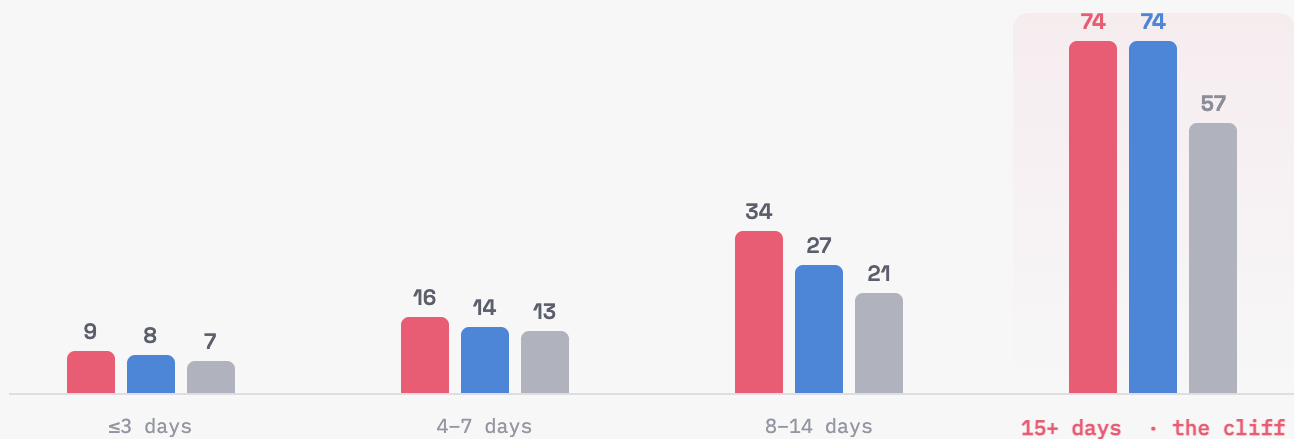
Reputation recovery and the 15-day cliff

Walking 3.7 years of history across 71,581 degradation events, one finding is counterintuitive: **duration matters far more than depth**. A severe 40-point drop caught in three days recovers in about a week, the same speed as a mild drop of the same duration. But once an inbox stays degraded for 15+ days, recovery time jumps to 57–74 days regardless of depth.

Median days to recover, by how long the inbox stayed degraded

Grouped by drop depth. Note how little depth matters next to duration

10–20pp 20–40pp 40+pp



THE OPERATIONAL THRESHOLD

The goal is not to avoid deep drops. It is to catch drops early. A deep drop caught in the first few days recovers faster than a mild drop left to sit for two weeks. Everything in recovery is a race against day 15.

67%

of all detected drops recover, median **20 days**. Most degradations are short and fully recoverable when caught in time.

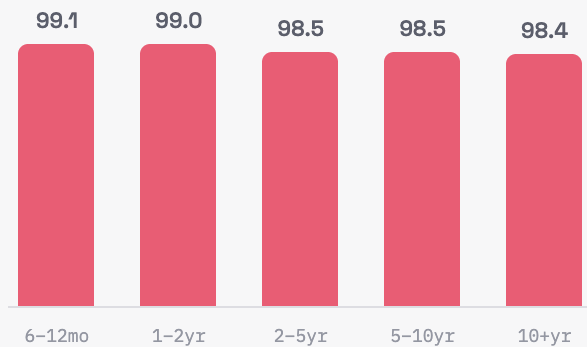
Source: 3.7 years of network history, 71,581 detected events. 33% censored (recovery not observed) and reported honestly.

Two things that don't move placement

Two persistent beliefs in cold email, that you need an aged domain and that send-day matters, do not hold up in the data. Both are essentially flat lines. Spend that attention on monitoring instead.

Domain age vs. inbox rate

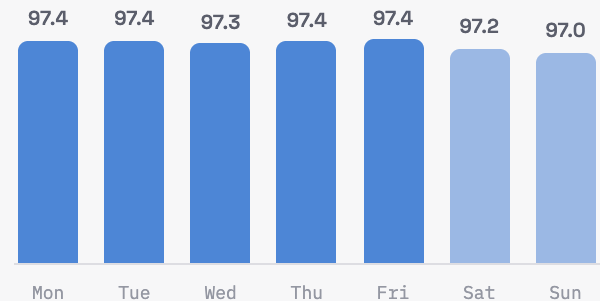
Pearson $r = -0.096$ · no relationship · y-axis 90-100%



A six-month-old domain lands the same as a ten-year-old one. **A new domain is not a handicap.**

Day of week vs. inbox rate

0.41pp spread, best to worst · y-axis 90-100%



Placement quality per send holds steady all week. **Send-day is not a lever.**

Both non-findings are good news for senders: there is no waiting period to warm a fresh domain, and no magic send-day to chase. The levers that *do* move placement, early detection and fast recovery, are the ones Warmup Inbox is built to pull.

Source: 4,406-inbox subsample with validated domain-registration data; 90-day placement window.

3.7 years of Gmail policy events, and the one that mattered

Every few months a Gmail policy announcement sends the deliverability world into a scramble. We overlaid the four best-known events on 3.7 years of network placement. Of the four, only the November 2025 enforcement ramp produced a clearly visible signal, and even that recovered in about 16 weeks.

POLICY EVENT	DATE	VISIBLE EFFECT IN THE NETWORK	SIGNAL
Google / Yahoo joint announcement	Oct 2023	A drop ~5 weeks later; timing suggestive, not proof	MARGINAL
Gmail bulk-sender requirements	Feb 2024	No visible degradation	NONE
One-click unsubscribe deadline	Jun 2024	No visible effect	NONE
Gmail full enforcement ramp	Nov 2025	97.7% → 96.4% inbox, spam ~2×, 16-week recovery	CLEAR

The big moves weren't policy

The two largest inflections in the entire series, a climb in January 2023 and a 5.3-point jump in late September 2024, do not line up with any known external policy event. Policy dates explain far less of real-world placement variance than the headlines imply.

THE NETWORK AS AN INDUSTRY SENSOR

In July 2025 the network caught a spam-rate spike (~1.3% → 4.5%) three weeks before it recovered, with no announced policy behind it. At 20M sends a month, Warmup Inbox often sees provider shifts before the industry names them.

A NOTE ON HONESTY

We can see what happened in our network and when. We cannot see individual receivers' filter decisions, so throughout, timing correlations are exactly that: correlations, not proof of causation.

One operating model for deliverability

1 Use Gmail as your primary

KPI

It catches problems first, in 85% of events, and reacts hardest. If Gmail is healthy, you usually are.

2 Measure Gmail

honestly

Plan around ~89% primary inbox plus ~9% tab routing, not a phantom 99% from IMAP.

3 Warm to two weeks, then keep it up

Placement converges fast, but holds only while warming continues. Like the gym, stop and you drift back. Treat it as ongoing maintenance.

4 Catch drops before day

15

The cliff is the single most important operational threshold in the data. Duration beats depth.

5 Stop over-indexing on domain age & send day

Neither predicts placement. Spend that attention on monitoring instead.

6 Keep policy events in

perspective

In 3.7 years, exactly one produced a clear network-wide signal.

CONCLUSION

If you take one thing from 2026, take the cliff.

It is where good and bad deliverability outcomes diverge. Warmup Inbox is built to help you act before day 15: catch the drop early, keep warming, and recover in a week instead of two months. [See it on your own inboxes free for a week.](#)

START FREE · NO CARD REQUIRED

Try Warmup Inbox free for 7 days

Connect an inbox to the network of 20,000+ real inboxes, watch your Gmail placement climb in the first two weeks, and get alerted the moment reputation starts to slip, while recovery is still fast. Or book a demo to see exactly how Warmup Inbox can help you.

Start free trial →

Book a demo