# Measuring OpenRTB Dialects in Client-Side Header Bidding

**Aleksander Sekowski**

Independent Researcher

rtblint.org . github.com/aleksUIX/rtblint

*Preprint, revised August 2026*

## Abstract

A browser can observe OpenRTB only where header-bidding libraries solicit demand client-side. We capture Prebid.js bid requests and responses from one US residential vantage across two samples: a probabilistic draw from the Tranco top 50,000 (n = 1,200, fixed seed) and 123 purposive publishers. The first-contact corpus is 10,457 payloads from 164 sites. It is not server-to-server OpenRTB, and it is not a census of programmatic traffic.

Validated with a pinned open-source linter under per-endpoint version attribution, 42.0 percent of bid requests on the random header-bidding sample are flagged on type, enumerated-value, or required-field checks (95 percent CI [36.0, 47.2], site-clustered). The same reading among major publishers is 40.4 percent [35.8, 45.0]; the intervals overlap. Counting unknown non-ext fields and legacy placements, which the specification tells receivers to tolerate, raises those rates to 69.4 percent [63.9, 75.0] and 55.4 percent [49.6, 61.4]. That 14-point gap is dialect and GDPR-path hangover, not core breakage. Both readings reproduce across three recrawls of each corpus over four days (strict Sample A 42.0, 42.6, 42.3). SSP responses are a different object. Thirteen endpoints emit no flagged response in any wave. A small set of endpoints omit nbr on every no-bid, so the aggregate response rate follows fill and is reported only at first contact (21.4 percent and 14.9 percent).

Requests are built by Prebid adapters; responses by SSP servers. On the random sample, 37.1 percent of request-side findings concentrate at one endpoint (adapter code), 2.0 percent at few publishers, and 60.9 percent is a leftover spread across both axes, led by shared-stack conventions such as deprecated imp.video.placement and site.ref sent as null. Publishers do not write those dialects. They enable adapters: 76 of 79 Sample A sites already have at least one endpoint at 0 percent strict-invalid on their own traffic. Inclusive 100 percent is often a moved GDPR path; strict 90 percent is type and value. Adapter-specific extra fields are dialects those endpoints parse, and they bid. Joining each request to the response it solicited, the naive pooled comparison inverts (Simpson's paradox). After stratifying by endpoint the Sample A association is +1.4 points (95 percent CI [-9.0, +13.4]), too wide to rule out a large bid-rate penalty or none. Cross-validating matched pairs flags 7.2 percent, including 169 bids whose deal identifiers were absent from the client-side request; that count does not distinguish SSP-side deal databases from unverifiable claims. Of 114 video bids joined to their impression, 86 percent are VAST wrappers with no inline duration or media file, a verifiability limit of wrapper architecture, reported as a case study. We release the capture and analysis code and a PII-free issue-level dataset covering 16,368 findings.

## 1. Introduction

Programmatic advertising clears through sealed-bid auctions conducted in the time a page takes to load, carried by the OpenRTB protocol maintained by the IAB Tech Lab [1]. The protocol has no reference validator and no public conformance test suite, and it does not use RFC 2119 requirement vocabulary. A companion study [13] classified every normative statement in OpenRTB 2.6 and 3.0 by what it would take to enforce it and found that roughly half of the specification is statically checkable, that between a quarter and a third has no machine-decidable criterion at all, and that some embedded examples diverge from the object tables.

That study analysed what the specification says. It could not say what traffic does. This paper measures the slice a browser can see.

The measurement is possible because some programmatic demand is solicited client-side. Header bidding libraries, overwhelmingly Prebid.js [11], run in the browser and issue OpenRTB bid requests directly to supply-side platforms, whose responses return to the same page. Each demand adapter translates the library's internal representation into the dialect its endpoint expects, a translation significant enough that the project maintains a dedicated conversion library [11]. That translation layer is the object this paper measures. Server-to-server hops downstream of an endpoint, including Google AdX, Amazon TAM, Open Bidding, and Prebid Server, are invisible to a browser. The population is client-side OpenRTB JSON, not OpenRTB.

We make seven contributions.

1. The first measurement of client-side OpenRTB JSON in live header-bidding traffic. Across 10,457 first-contact payloads from 164 sites and 101 distinct endpoints, captured at one US residential browser, 42.0 percent of bid requests are flagged on type, enumerated-value, or required-field checks. Including unknown and moved fields the specification tells receivers to tolerate raises the rate to 69.4 percent.
2. An attribution split no prior work separates, and a publisher menu. Requests are built by client-side adapters and responses by SSP servers. Publishers write almost none of the finding mass (about 2 percent). They do choose which adapters to enable. Inclusive versus strict rates by adapter separate moved GDPR paths from type and value divergence.
3. A decomposition of the sample gap. The 14-point difference between a random draw of header-bidding sites and major publishers lives entirely in unknown and legacy fields. Under the strict reading the two samples agree at about 41 percent.
4. Vantage as a measurement precondition. An identical harness through a datacenter address observes zero auctions on eight test sites. Invalid-traffic filtering of hosting space is documented; the implication for protocol capture is that a datacenter vantage silently records nothing. Repeated visiting from one residential profile also depresses fill across unrelated demand partners.
5. Stability of request construction, and a fill caveat. Three recrawls of each corpus over four days: the strict request rate moves 0.6 points on Sample A. That rules out a one-afternoon glitch in these adapters. It is not a claim about all OpenRTB or about other months.

Response-side fill statistics are only clean at first contact.

6. A bid-outcome comparison that inverts unless stratified. Joining each request to the response it solicited shows Simpson's paradox across endpoints. After stratification the Sample A interval is compatible with a large penalty, a large bonus, or zero. We do not claim the market is indifferent.
7. Public artifacts. Capture and analysis code, the sampling frame with its seed, and a PII-free issue-level dataset that reproduces every table below.

## 2. Background and Related Work

### 2.1 Client-side auction solicitation

In header bidding, a publisher page loads a library that requests bids from multiple demand partners before, or in parallel with, its primary ad server call [7]. Prebid.js dominates this layer. Each demand adapter translates the library's internal representation into the OpenRTB dialect its endpoint expects [11]. Scoring that dialect against the IAB object tables measures spec hygiene. It does not measure whether the endpoint's parser accepted the message. Both readings are reported.

### 2.2 What the specification can and cannot require

OpenRTB 2.6 is prose with field tables [1]. It never adopted the RFC 2119 requirement vocabulary [2, 3], and it mandates tolerant reading: implementers "must tolerate receiving new or unexpected fields and enumerated list values gracefully." The companion study [13] argues this removes the backpressure that would otherwise surface divergence, echoing the Internet Architecture Board's retrospective on the robustness principle [10], which is unrelated to the Interactive Advertising Bureau that maintains OpenRTB, and the LangSec position that an input language which is not decidable makes every parser its own de facto specification [9].

### 2.3 Measurement of the advertising supply chain

Prior measurement has characterised header bidding traffic [7], fraud in exchanges [6], and the economics of real-time bidding [8]. Transparency studies have quantified unmatched spend across intermediaries [4, 5]. None of this work validates the protocol messages themselves. We are not aware of prior published measurement of OpenRTB against its object tables in live client-side traffic.

## 3. Method

### 3.1 Vantage

All measurements were taken from a residential consumer connection in one US metro. This is not incidental. An identical harness run through a datacenter network address observed zero auctions across eight test sites: Prebid loaded and reported its version, but registered no ad units and fired

no auction events, and one site additionally applied fingerprint-based bot detection. The same eight sites from a residential address produced OpenRTB traffic on six. Ad stacks treat datacenter address space as invalid traffic and suppress bidding, so a datacenter vantage silently measures nothing while appearing to work. We report this as a precondition for wire-level protocol capture, not as a discovery of invalid-traffic filtering.

### 3.2 Two samples

Sample A (probabilistic). From the Tranco top 50,000 [14], we removed 1,032 domains by documented mechanical exclusion, covering infrastructure, CDN, cloud, and ad-tech service hosts, and non-content top-level domains, leaving a frame of 48,968. From that frame we drew n = 1,200 uniformly at random with a published seed. Exclusions are substring and TLD rules applied programmatically, never by inspection, so the frame reproduces exactly. Residual non-content domains remain in the frame by design; hand-pruning them would reintroduce the selection bias the frame exists to remove.

Sample B (purposive). 123 publishers assembled by judgement as likely to run client-side header bidding. This sample is not representative and is used only where many payloads per endpoint are required. Every prevalence claim in this paper comes from Sample A.

### 3.3 Capture

A headless Chromium instance driven by Playwright, with network interception at the Chrome DevTools Protocol layer, records every POST body that parses as an OpenRTB bid request (a JSON object with id and a non-empty imp array) and every response body that parses as a bid response (seatbid or nbr present). Large bodies unavailable inline are retrieved through the protocol's post-data accessor. The harness accepts consent where a recognised consent platform button is present, scrolls to trigger lazily initialised ad slots, and then navigates to one article page, where ad unit density is higher than on a homepage. Prebid is detected even when the page renames its global object.

A separate lightweight pass, used for prevalence only, visits each homepage once, stores no payload bodies, and records only whether header bidding is present and whether OpenRTB crossed the wire.

The harness observes only client-side traffic. Server-to-server demand solicited downstream of an endpoint is invisible to it, as is any adapter using a proprietary or binary format. Capture already requires id and a non-empty imp array, so payloads missing that required core never enter the corpus. This bounds the population, and Section 6 treats it as a limitation.

### 3.4 Validation instrument

Payloads were validated with RTBlint 0.6.0, an open-source OpenRTB linter. It ships per-version field catalogs, documented value sets, single-message semantic rules, and request/response cross-validation. Attribution in Section 3.5 uses the 16 OpenRTB 2.x catalogs; 3.0 is a different object model and was not a candidate. The author maintains the instrument. Classifications can be recomputed from the released issue-level dataset.

Two of the instrument's severity choices matter to the results and are stated here rather than left implicit. Fields not defined in the version catalog, outside ext subtrees which the validator leaves open, are reported as errors, on the reading that the specification directs extensions into ext. Fields present at a location the specification later moved them from are likewise errors. Both are recommendations in the specification text rather than absolute requirements. The strict reading used as the headline drops both classes and keeps type mismatches, undocumented enumerated values, missing required fields, and the remaining error rules. Section 4.4 reports both readings.

The validator version must be pinned, and we report it for a concrete reason. An earlier analysis pass used a stale build whose 2.0 through 2.5 catalogs were known to be degraded. Because those catalogs under-reported unknown fields, the version attribution described below preferred them, producing an apparent cluster of endpoints speaking OpenRTB 2.2 that dissolved entirely when the validator was rebuilt. Any minimum-error version selection is only as trustworthy as the weakest catalog among its candidates. We report this because the failure is silent and would be easy for a replication to repeat.

### 3.5 Version attribution

96 percent of captured requests carry no x-openrtb-version header, so the version a payload should be judged against must be inferred. For each combination of endpoint and payload side, we validate all of its payloads against all 16 cataloged versions and assign the version that minimises total error count, breaking ties toward the newest. Attribution is per side because the client-side code that builds requests and the server that builds responses are different implementations and need not agree.

This procedure is deliberately charitable: every endpoint is judged against the dialect that flatters it most. It is not uniformly charitable. 172 Sample A requests (5.6 percent) and 364 Sample B requests are assigned 2.5. On those payloads, fields defined in 2.6, including device.sua (77 Sample A findings) and imp.video.plcmt (26), are scored as unknown. 67 Sample A requests are invalid only because device.sua or imp.video.plcmt scored as unknown, 2.2 points of the as-reported rate. The strict reading already drops unknown fields, so this contamination does not touch the headline. Issue-level counts of field.undefined should be read with the 2.5 assignments in mind.

### 3.6 Statistics

Measurements were taken in waves, where a wave is one complete crawl of one corpus. Each wave is analysed independently, including its own version attribution, so no wave inherits another's assumptions. Each sample has three waves, six captures in total; Section 4.4.3 reports what moved between them and what did not. The 10,457-payload corpus quoted above is first-contact Sample A plus Sample B, not the six captures pooled.

Payloads are not independent observations. One site fires many auctions and each auction fans out to many endpoints, so a pooled payload percentage overstates precision. All confidence intervals for invalid rates come from a cluster bootstrap that resamples whole sites with replacement (5,000 draws, seed 20260725, RNG reset per interval), with the site as the cluster

unit. Strict intervals drop unknown and moved-field errors before resampling. Prevalence proportions use Wilson score intervals. Per-site interquartile ranges are reported alongside pooled rates because site-level heterogeneity is large.

### 3.7 Ethics and data handling

The study observes traffic that the operator's own browser receives when visiting public pages. No accounts were used, no clicks were issued on advertisements, crawling was rate limited by a small concurrency bound, and each site was visited briefly. The raw payload corpus is not released: bid requests carry user identifiers, extended identifiers, address-derived geolocation, and commercial terms including floor prices and deal identifiers. The published dataset contains rule identifiers and structural JSON paths only, never payload values, with an additional filter that redacts any path resembling an identifier. Because values are withheld, claims about which undocumented skip values appear cannot be re-litigated from the public artifact.

## 4. Results

### 4.1 Prevalence

Table 1 gives the Sample A funnel.

**Table 1**

| Stage | Rate | 95% CI |
|---|---|---|
| Reachable (page loaded) | 877 / 1,200 = 73.1% | [70.5, 75.5] |
| Header bidding present | 102 / 877 = 11.6% | [9.7, 13.9] |
| OpenRTB observed on the wire | 73 / 877 = 8.3% | [6.7, 10.3] |
| ... among header-bidding sites | 72 / 102 = 70.6% | [61.1, 78.6] |

Roughly one in nine reachable top-ranked sites runs client-side header bidding. Unreachability is dominated by DNS failure (190 of 323): Tranco ranks resolvable domains, many of which never served a homepage. This is a property of the frame and is reported rather than corrected.

The last row bounds the method. A single homepage visit observed OpenRTB on only 70.6 percent of sites that demonstrably run header bidding, so detection is a lower bound and the full harness recovers part of the remainder. One reachable site produced OpenRTB without a detected Prebid global, which is why the wire count is 73 against 72 among header-bidding sites, and why the deep-crawl list in Section 4.2 is 103 (102 Prebid-positive plus that one).

Prevalence by rank band, among reachable sites, runs 21.0 percent in ranks 1 to 5,000, 18.8 percent in 5,001 to 15,000, 5.6 percent in 15,001 to 30,000, and 10.1 percent in 30,001 to 50,000. The gradient is not monotonic and the bands overlap, so we claim only that header bidding concentrates toward the top of the ranking.

Across the homepage detection pass, 79 distinct endpoints appear on 73 sites. The most widely deployed are Index Exchange (52 sites), OpenX (40), The Trade Desk (32), Media.net (25), Magnite's Prebid Server (24), and Smart (21). Table 8 uses the deep crawl, which recovers additional adapters on the same sites.

## 4.2 Corpus

The 103 detection-positive sites from Sample A (102 with Prebid detected, plus one OpenRTB-only) were re-crawled with the full harness. The measurement corpus is therefore a random draw conditioned on running client-side header bidding: prevalence statements in Section 4.1 describe top-ranked sites generally, while dialect and validity statements describe top-ranked sites that run header bidding. Table 2 describes both corpora. Five sites appear in both samples (union 164).

**Table 2**

| | Sample A (random) | Sample B (purposive) |
|---|---|---|
| Sites crawled | 103 | 123 |
| Sites yielding traffic | 79 | 90 |
| Bid requests | 3,086 | 3,744 |
| Bid responses | 1,503 | 2,124 |
| Distinct endpoints | 88 | 66 |

Issue mass is concentrated. Media.net is 3.8 percent of Sample A requests and 25.7 percent of Sample A findings (6.0 percent and 33.8 percent on Sample B). Issue-level shares in later sections are media.net-weighted; payload-level rates are not.

## 4.3 Version signalling is unused

Of 3,744 requests in Sample B, 3,588 (96 percent) carry no x-openrtb-version header; the 156 that do declare 2.5. The specification's own version-signalling convention is effectively dead on the client side, which is why the attribution procedure of Section 3.5 is necessary at all.

Traffic is nevertheless firmly 2.6-shaped. Validating Sample B requests against 2.5 rather than 2.6-202606 raises the count of fields the version catalog does not recognise from 4,371 to 15,253, of which 10,316 are fields defined in 2.6 but absent from 2.5, and raises the invalid rate from 58.8 percent to 95.2 percent. Judging this traffic against the current specification is the right modern baseline, with the 2.5-assignment caveat in Section 3.5.

## 4.4 Two readings

Table 3 gives request invalid rates under the strict reading (type, enumerated value, required, and remaining non-disputable error rules) and under the inclusive reading (those plus unknown non-ext fields and moved fields). Intervals are site-clustered. Responses are shown separately because their aggregate follows fill (Section 4.5) and is not a conformance rate.

**Table 3**

| Reading | Sample A | Sample B |
|---|---|---|
| Strict (type, enum, required) | 42.0% [36.0, 47.2] | 40.4% [35.8, 45.0] |
| Including unknown and moved | 69.4% [63.9, 75.0] | 55.4% [49.6, 61.4] |
| Responses, first contact (follows fill) | 21.4% [16.4, 27.1] | 14.9% [12.3, 17.8] |
| Sites at 100% / 0% (inclusive) | 20 / 79, 5 / 79 | 12 / 90, 3 / 90 |
| Per-site IQR (inclusive, requests) | [51.7, 100] | [33.3, 80.0] |

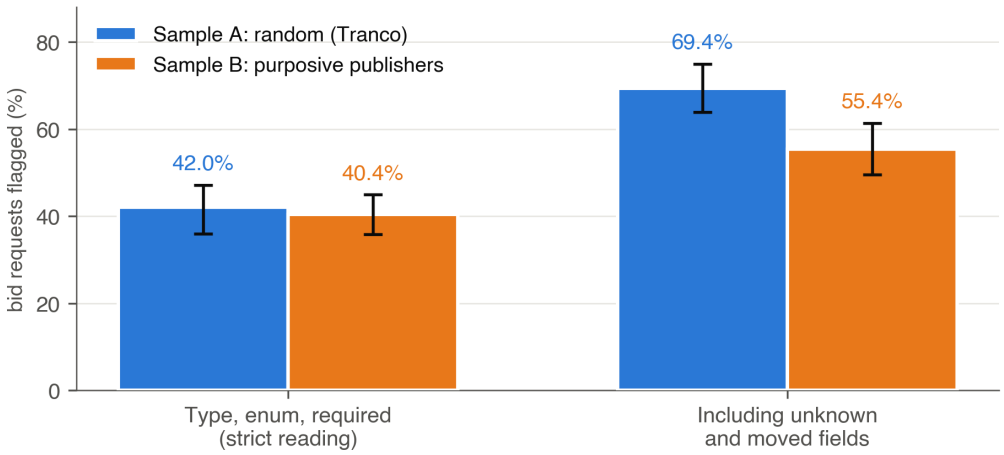


*Figure 1. Bid-request invalid rate under two readings, with site-clustered 95 percent confidence intervals. Strict drops unknown non-ext fields and moved fields. Sample A is a random draw of header-bidding sites; Sample B is purposive majors.*

The strict intervals overlap. Under type, enum, and required-field checks, a random draw of header-bidding sites matches major publishers at about 41 percent. The inclusive reading is the one whose intervals do not overlap, and Section 4.4.2 locates that gap. Per-site heterogeneity is large: on Sample A the inclusive per-site median is 70.0 percent with IQR [51.7, 100]. Twenty of 79 sites are at 100 percent. The pooled 69.4 percent is a mixture, not a typical site.

We report 42.0 percent as the headline. 69.4 percent is the inclusive sensitivity.

### 4.4.1 Sensitivity to the disputable classes

Two of the violation classes admit an objection. The specification says extensions "should" consistently be named ext, which is a recommendation rather than a requirement, and it separately requires receivers to tolerate unexpected fields. A reader may therefore argue that unknown non-ext fields, and legacy placements of fields that later moved, are ecosystem practice rather than violations. Our instrument treats both as errors, a choice we state rather than hide.

Table 4 recomputes the rate with those classes removed stepwise.

**Table 4**

| Reading | Sample A requests | Sample B requests |
|---|---|---|
| Inclusive (as scored) | 69.4% | 55.4% |
| Excluding unknown fields | 58.5% | 44.9% |
| Excluding unknown and moved (strict) | 42.0% | 40.4% |

The remaining 42.0 percent is type mismatches, undocumented enumerated values, and missing required fields on 1,286 of 1,296 strict-invalid requests, plus ten requests flagged only on mutually exclusive fields. Those are not extension practice under any reading. Of Sample A request findings, type mismatches are 1,246 (mostly site.ref as null, integer publisher.id, and imp.banner sent as an array), invalid values are 515 (509 of them imp.video.skip), and required-field errors are 33. Deprecated imp.video.placement is common (867 findings) and is scored as a warning, so it does not drive either rate.

### 4.4.2 What the sample gap actually is

The bottom row of Table 4 resolves the gap between the two samples. Under the strict reading the two are 42.0 and 40.4 percent, effectively identical. The entire difference between a random draw of header-bidding sites and a hand-picked set of major publishers lives in unknown and legacy field placement, not in core correctness: long-tail sites carry more dialect keys and outdated GDPR paths, while their type and value checks match those of major publishers.

Yahoo's Sample A request side is 100 percent invalid under the inclusive reading and is almost entirely regs.ext.gdpr and user.ext.consent, fields the 2.6 tables moved. Under the strict reading that endpoint is not a core-correctness failure.

Two mechanisms that would have predicted a broader inclusive gap were tested and both fail. It is not endpoint composition: restricting to the 20 endpoints present in both samples with at least 25 requests each, covering 75 percent of Sample A traffic and 87 percent of Sample B, leaves the inclusive gap essentially intact at 13.2 points, 70.3 against 57.1 percent, where the unrestricted gap is 14.0. Holding the endpoint mix fixed therefore accounts for less than a point of it. Fourteen of the 20 are worse on the random sample, by 19.1 points for OpenX, 22.1 for Magnite's Prebid Server, 26.9 for Xandr, 28.5 for Magnite's Fastlane, and 27.8 for Taboola; two are marginally better and four identical. Nor is it a clean age effect of the client library: across 128 sites with a detected Prebid.js version and at least five requests, median inclusive request-invalid rate by major release runs 62.0 percent on version 9, 64.7 percent on version 10, and 77.4 percent on version 11. Version 8 is four sites. The rank correlation between major version and legacy-field density is +0.128. This corpus does not support the claim that older Prebid releases are responsible. It also does not identify a causal effect of upgrading.

### 4.4.3 Stability across waves

Each corpus was recaptured on later days and analysed from scratch: six captures across four days, spanning weekend and weekday and hours from 14:00 to 00:15 local. These are recrawls of the same sites and adapters from the same address, not independent draws from a wider population. Request construction is deterministic, so stability is expected. It rules out a hung page or a one-afternoon artifact in this set of adapters. It does not buy seasonality, later Prebid releases, or other networks. Table 5 records the six captures.

**Table 5**

| Measure | Sample A | Sample B |
|---|---|---|
| Waves | 3 | 3 |
| Inclusive requests, by wave | 69.4%, 70.1%, 68.6% | 55.4%, 56.3%, 56.8% |
| Strict requests, by wave | 42.0%, 42.6%, 42.3% | 40.4%, 41.7%, 41.8% |
| Strict spread | 0.6 pts | 1.4 pts |
| Inclusive spread | 1.5 pts | 1.4 pts |
| Median per-endpoint drift (inclusive) | 2.5 pts (27 endpoints) | 3.8 pts (22 endpoints) |

The strict reading is the more stable of the two on Sample A. Large endpoints barely move on the inclusive reading. On Sample B, Sharethrough 74.8, 75.7, 76.6; Index 42.7, 43.0, 44.3; TTD 91.2, 91.8, 92.2; Lijit 100.0 in all three. Sample A is the same pattern at different levels: Sharethrough 90.2, 90.1, 88.8; Index 52.0, 49.5, 50.7.

The response side behaves differently, and Section 4.5 explains why: it moved from 14.9 to 31.6 percent on Sample B while request-side quantities held. That movement is not instability in the measurement but a consequence of no-bid handling at a few endpoints.

### 4.5 The construction-side split

Requests are constructed by Prebid adapters running in the page together with publisher configuration. Responses are constructed by SSP servers. Inclusive request rates and first-contact response rates do not overlap on either sample. Under the strict reading the request side remains far above the response side as well.

Response-side behaviour sorts into three classes, and the classes are stable even though the aggregate rate is not.

Conformant. Thirteen endpoints emit no flagged response at all, in any wave in which they appear. Six of them clear 400 responses: OpenX (1,772), Yahoo (861), Ozone (462), two Raptive Prebid Server hosts (411 and 404), and The Trade Desk (404). The remainder are cleanly conformant on smaller volumes: PubMatic (139), a second Raptive host (96), a second Trade Desk host (89), a.bids.ws (79), Amazon (72), Xandr (59, present in three waves), and Taboola (51). Several of those volumes are too small to treat 0 percent as a precise rate. Magnite's Prebid Server is close but not perfect, at 2 invalid responses in 1,367, both single instances in separate Sample B waves.

Empty seatbid, no nbr. A second group returns an empty seatbid array with no nbr no-bid reason, or omits required response fields, every time it declines to bid. For these endpoints the flagged rate is not a statistic at all; it equals the no-bid rate exactly. Across 17 endpoint-wave combinations covering Index, Magnite's Fastlane, nexx360, Sparteo, and 33Across, spanning six captures on four days, the invalid rate matched the no-bid rate to within 0.1 points in every single case (Table 6):

**Table 6**

| Endpoint | Wave | Responses | No-bid | Invalid |
|---|---|---|---|---|
| Index | A wave 1 | 256 | 52.3% | 52.3% |
| Index | A wave 2 | 238 | 50.8% | 50.8% |
| Index | B wave 1 | 350 | 51.7% | 51.7% |
| Index | B wave 2 | 317 | 99.7% | 99.7% |
| Index | A wave 3 | 219 | 99.1% | 99.1% |
| Magnite Fastlane | B wave 1 | 126 | 50.0% | 50.0% |
| Magnite Fastlane | B wave 2 | 133 | 96.2% | 96.2% |
| Magnite Fastlane | A wave 3 | 92 | 98.9% | 98.9% |
| Index | B wave 3 | 312 | 99.7% | 99.7% |
| Magnite Fastlane | B wave 3 | 128 | 99.2% | 99.2% |

Empty seatbid without nbr is a 2.x requirement with limited adoption. These endpoints omit it on every decline.

Flagged on every response. Smart is 100 percent flagged while never declining a bid, a pattern unrelated to no-bid handling.

This resolves the one unstable quantity in the study, though not in the direction we first supposed. The aggregate response-side rate ranges from 14.9 to 32.5 percent across waves entirely because

the no-bid rates of these endpoints range from about 50 percent to about 99 percent. Our initial reading was that this reflected demand, and we tested it by scheduling one capture at midday. It failed that test: Index's no-bid rate was 99.7 percent at 14:00 and 99.7 percent at 20:54, while three captures at three different evening hours on the first crawl day all sat near 51 percent. The variation is a step, not a cycle, and the step falls at the end of the first day of crawling.

The most likely cause is the measuring profile. Bid rates fell simultaneously across unrelated demand partners, who do not coordinate: comparing the first crawl day with later days, Index fell from 48.4 to 0.5 percent, PubMatic from 88.4 to 64.0, Yahoo from 31.5 to 17.8, OpenX from 21.0 to 10.1, and Magnite's Prebid Server from 19.0 to 8.6. The reading is that after roughly a day of repeated visiting the address and browser profile stop attracting bids, through frequency capping or invalid-traffic scoring.

Three confounds were tested and excluded. It is not a changed network: captured requests carry device.ip, and the same residential address appears in every capture, including the later ones. It is not a shifting site mix: restricting to the 81 Sample B sites that yielded responses in all three waves, with comparable response counts per wave, reproduces the decline unchanged (Index 47.9 to 0.3 to 0.3 percent, Magnite's Prebid Server 27.6 to 13.4 to 7.5). It is not the hour, as the midday capture established.

One alternative cannot be settled from a single vantage: a genuine market-wide change coincident with our first crawl day. We regard it as less likely, because the decline is graded across endpoints in proportion to how heavily we exercised them rather than uniform, but a capture from a fresh residential or carrier address would discriminate cleanly. A datacenter or VPN exit would not: Section 3.1 shows those observe no auctions. We report the effect as observed and the mechanism as probable rather than established. A second vantage is untested.

This has an asymmetric consequence that matters for anyone repeating this work. Requests are constructed client-side and do not depend on whether anyone bids, so request-side rates are untouched, which the wave stability in Section 4.4.3 confirms directly. Response-side statistics that depend on fill are only clean at first contact, so we report the first-contact figures (14.9 percent for Sample B, 21.4 percent for Sample A) as descriptive of that first day and treat later waves as confirming the per-endpoint behaviour rather than the rate. They should not be quoted as a fixed property of SSP responses.

The behavioural law is unaffected by any of this, because it is defined per response rather than per rate: across 17 endpoint-wave combinations spanning six captures on four days, these endpoints omitted nbr on every no-bid.

This is the construction-side result. Exchanges largely emit IAB-shaped OpenRTB on the response path, and where they diverge, the pattern is specific, repeatable, and attributable to a named endpoint. The layer that diverges from the object tables is the client-side code soliciting bids on their behalf, which is also the layer whose job is to speak each endpoint's dialect.

### 4.6 Where client-side findings originate

Client-side has two possible origins: an SSP's own adapter, which is shared code appearing wherever that adapter runs, and a publisher's configuration, which would appear across all of that

publisher's endpoints. The two are distinguishable by how a defect signature spreads. Over the signatures with at least 20 occurrences (35 covering 5,854 findings on Sample A, 46 covering 7,278 on Sample B) we split three ways (Table 7). The third bin is a leftover: signatures that meet neither concentration rule. Leading examples in that leftover are shared-stack conventions. The 60.9 percent figure is not a positively identified owner.

**Table 7**

| Origin | Sample A (random) | Sample B (purposive) | Signature |
|---|---|---|---|
| SSP adapter code | 37.1% | 49.8% | one endpoint, many publishers |
| Leftover (both axes) | 60.9% | 48.6% | not concentrated on either axis |
| Publisher configuration | 2.0% | 1.6% | many endpoints, few publishers |

Both columns replicate on their later waves, to within roughly 2 points: Sample A gives 34.7 / 63.1 / 2.2 and Sample B gives 47.1 / 51.4 / 1.5. The publisher share is the most stable quantity in the study, sitting between 1.5 and 2.2 percent in all four measurements. Including signatures below the 20-occurrence cutoff does not change it (1.8 percent publisher-like on Sample A).

The adapter and leftover shares differ by sample, and the difference is informative rather than noise. On the random draw of header-bidding sites, leftovers are the majority at 61 percent; among major publishers, adapter-specific findings rise to roughly half. Long-tail sites carry proportionally more of the shared stack's extra keys, which is consistent with Section 4.4.2. Where a single figure is wanted, the random sample is the one to quote.

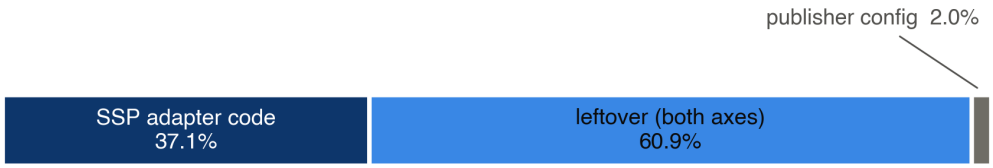


*Figure 2. Origin of request-side findings on the random sample, over the 35 signatures with at least 20 occurrences covering 5,854 findings. The middle bin is a leftover, not a named mechanism. Sample B shifts toward adapter code; both are reported in the text.*

Publishers almost never write the dialect. They select it. Examples of pure adapter origin, each at complete endpoint concentration: Media.net emits imp.all, imp.ortb2Imp, imp.transactionId, imp.bidfloors, ortb2, site.canonical_url, site.isTop, and site.topMostLocation across 28 Sample A publishers, and also sends imp.banner as an array of banner objects where the specification defines a single object. That array is the dialect that endpoint parses. We name it because it is a type mismatch against the IAB tables, not because the auction failed.

The leftover tier is the more consequential half for spec authors, because no single adapter owns it. Counting on Sample B, deprecated imp.video.placement appears 1,108 times across 38 endpoints and 57 sites; undocumented values for imp.video.skip appear 576 times across 25 endpoints; site.ref is sent as JSON null, where the specification expects the field to be omitted, 501 times across 69 sites; and the legacy regs.ext.gdpr placement appears 300 times across 19 endpoints. The same signatures lead Sample A at proportionate volumes, 867 and 509 for the two video fields. These are conventions of the shared client stack rather than independent vendor mistakes, which is

exactly why they persist.

Manual inspection confirms the large type-mismatch classes are genuine divergences rather than detector artifacts: null where a string is required, integers where identifiers are declared as strings at four separate SSPs, a currency array sent as a bare string, and one adapter serialising banner.format as an object with numeric keys.

### 4.6.1 Adapter menu

The origin split says publishers are not the authors of request-side findings. The Prebid config still decides which authors run. Table 8 is the Sample A first-contact menu for every request endpoint present on at least 10 sites. Inclusive includes unknown and moved fields. Strict is type, enumerated value, and required. This is not a revenue ranking and it is not a recommendation to drop demand.

**Table 8**

| Adapter | Sites | Requests | Inclusive | Strict |
|---|---|---|---|---|
| Index | 55 | 304 | 52% | 18% |
| OpenX | 43 | 260 | 52% | 27% |
| The Trade Desk | 34 | 194 | 97% | 93% |
| Magnite PBS | 32 | 167 | 78% | 65% |
| Media.net | 28 | 116 | 100% | 86% |
| Smart | 25 | 202 | 100% | 88% |
| PubMatic | 25 | 133 | 50% | 42% |
| Sharethrough | 17 | 153 | 90% | 90% |
| Criteo | 17 | 88 | 57% | 47% |
| Lijit | 17 | 70 | 100% | 0% |
| Xandr | 15 | 71 | 56% | 17% |
| Magnite Fastlane | 13 | 89 | 40% | 1% |
| Sparteo | 12 | 38 | 8% | 0% |
| Yahoo | 11 | 151 | 100% | 0% |
| Adform | 10 | 83 | 100% | 70% |
| Marphezis | 10 | 33 | 79% | 61% |
| Ogury | 10 | 24 | 54% | 0% |

The two columns are two kinds of divergence from the tables. Yahoo and Lijit are 100 percent inclusive and 0 percent strict: requests carry a moved GDPR path or an unknown field, and none is flagged on type, enum, or required checks. The Trade Desk, Sharethrough, Media.net, and Smart are flagged on the strict reading on most requests. Magnite Fastlane is 40 percent inclusive and 1 percent strict. Inclusive 100 percent on a site is often the Yahoo row: a path the 2.6 tables moved, not a type error.

76 of 79 Sample A sites already load at least one endpoint that is 0 percent strict-invalid on that site's own traffic. Adapters with high strict rates sit alongside those. 38 of 79 sites still send regs.ext.gdpr. That is a consentManagement module path, not an SSP choice, and a US vantage understates how much that path matters in the EEA. Section 4.7 shows we cannot price either choice as fill. The menu is still the config a publisher already controls.

## 4.7 Bid outcome, stratified

If type and value mismatches suppressed demand, publishers would see it as lost fill. We join each request to the response it solicited and compare bid rates. Because bid outcome is fill, and Section 4.5 shows fill decays as the measuring profile ages, only first-contact waves are used: 1,503 matched pairs in Sample A and 2,124 in Sample B. Inclusive request validity is the split, matching the instrument's default scoring.

The naive comparison points the wrong way. Pooled across all endpoints, requests the linter called valid received a bid 41.6 percent of the time against 47.4 percent for flagged ones in Sample A, and 47.6 against 53.8 percent in Sample B. Taken at face value this says extra-field requests do slightly better, which is not a credible mechanism.

It is a composition artifact. Endpoints differ enormously in both their baseline bid rate and in the dialects of the adapters that address them, so pooling across them inverts the comparison. The Trade Desk illustrates it: on Sample A it bid on 98.7 percent of the flagged requests it received against 50.0 percent of the unflagged ones, not because extra fields help, but because the adapters that address it are both prolific and extra-field-heavy, and it bids on nearly everything. Stratifying by endpoint removes the effect and reverses the sign (Table 9):

**Table 9**

| | **Sample A (random)** | **Sample B (purposive)** |
|---|---|---|
| Matched pairs at first contact | 1,503 | 2,124 |
| Naive pooled difference | -5.8 points | -6.2 points |
| Endpoint-stratified difference | +1.4 points | +11.3 points |
| 95% CI, site-clustered | [-9.0, +13.4] | [+3.2, +20.2] |

On the random sample the interval is compatible with a 9-point penalty, a 13-point bonus, or zero. On the purposive sample the interval excludes zero, but that is not a result we treat as exchange policy.

The per-endpoint detail is what settles the interpretation. Individual endpoints do show large gaps on Sample B, Index at +24.7 points and Magnite's Prebid Server at +22.0, but neither reproduces on Sample A, where the same endpoints give +3.6 and -7.4. Of the twelve endpoints with both strata populated in both samples, the four highest-volume ones all fail to hold their sign or magnitude across samples. Only endpoints bidding identically in both strata replicate cleanly: The Trade Desk, Xandr, and Amazon sit at 100 percent regardless of validity, which is the behaviour a tolerant reader produces. An exchange applying a table-conformance bidding policy would produce a stable per-endpoint effect; what we observe instead moves with which adapters happen to populate each corpus.

We do not claim causation. Inclusive validity is overwhelmingly a property of the adapter rather than of an individual request, as Section 4.6 establishes. Within one endpoint, the two strata therefore arrive largely from different adapters, not from one adapter varying. Any difference is confounded with everything else that differs between adapters, including floor prices, targeting signal richness, and inventory quality. Isolating a real effect would require adapters that emit both IAB-shaped and dialect-shaped requests to the same endpoint, which our corpus does not contain in useful numbers.

What the result does establish is the Simpson's paradox and the width of the Sample A interval. It

does not establish that dialects are free, and it does not establish that exchanges penalise them.

### 4.8 Cross-message findings

Single-message validation cannot see whether a response is coherent with the request that solicited it. Cross-validating all 2,124 matched pairs in Sample B at first contact, 153 pairs (7.2 percent) carry findings that genuinely require the request: 169 bids reference deal identifiers the client-side request did not list (some pairs carry more than one such bid), 8 responses carry an identifier that does not match the request they answer, 5 bids are for impressions that were never offered, and the remainder are markup and currency incoherences.

We restrict this count to rules that cannot fire without the request. A bid declaring no markup type is a common finding, 768 instances here, but it surfaces during ordinary single-message response validation and so is not evidence for what pair mode adds; including it would inflate the figure to 19.7 percent.

These counts also depend on fill, since a bid must exist before it can cite a deal. In the later Sample B wave, where the measuring profile had aged and bid volume collapsed (Section 4.5), the same analysis gives 63 flagged pairs (3.7 percent) and 64 deal-identifier findings. We therefore report first-contact figures here as well.

The deal-identifier count is a pair-mode observation, not an economic conclusion. Header-bidding requests often omit pmp.deals because deals are configured in the SSP and attached on the server. A bid whose dealid is absent from the client-side request may still name a deal that exists in that SSP's database. Distinguishing that case from a fabricated identifier requires the SSP's deal table, which this vantage does not have.

### 4.9 The creative layer, as a case study

Video bids carry VAST markup inside the bid, but no OpenRTB requirement makes that markup valid or consistent with what the request asked for. Of 131 VAST creatives captured, 73 (55.7 percent) contain at least one VAST error, reproducing on this inventory the premise of an earlier study of the creative layer [12]. Inventory is display-heavy; this is a case study, not a video census. It was not re-measured on later waves.

Of 114 creatives that could be joined to their originating impression, every one of the corresponding requests declared its maxduration, mimes, protocols, and api terms, so the request stated a contract. But 98 of 114 (86 percent) are VAST wrappers carrying no inline duration and no media file: the delivered duration, codec, and API framework only become knowable after resolving the wrapper chain at render time, long after the auction has cleared. The pattern holds across all 21 endpoints that served video.

This is how VAST wrappers work. It is a verifiability limit at bid time, not an OpenRTB violation. Of the 16 creatives that did expose their facts inline, all complied, and one still served a Flash media file.

### 4.10 Mapping onto the companion taxonomy

Mapping every rule to the enforceability class assigned by the companion study [13] describes where this instrument's findings sit. It is not an independent test of that study's predictions. The companion classification, the linter rules, the rule-to-class map, and this measurement share an author. Capture already requires id and a non-empty imp array, so findings on the required core are selected against. Issue-level shares are media.net-weighted (Section 4.2). Unknown-field findings are the instrument counting what the specification tells receivers to ignore. Table 10 places those shares next to the companion taxonomy.

**Table 10**

| Mapping | Sample A (random) | Sample B (purposive) |
|---|---|---|
| Required-core share of findings | 1.3% of findings, 1.9% of payloads | 1.6% of findings, 2.3% of payloads |
| Class A issue share | 64.6% of findings | 74.3% of findings |
| Undefined fields as a share of findings | 37.9% | 51.0% |

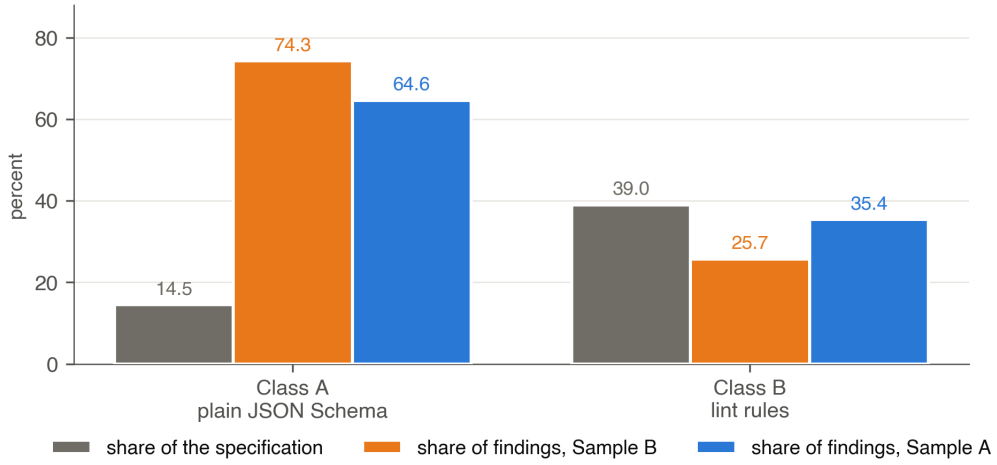


*Figure 3. Share of the specification each enforceability class covers, against the share of findings it carries in each sample. Issue-level class A is inflated by per-field unknown-field hits and by media.net. On Sample A, payload-level class A does not exceed class B.*

The required core is 22 of 417 fields, 5.3 percent of the surface, and absorbs 1.3 percent of findings on Sample A and 1.6 percent on Sample B. That is expected given the capture filter and given that adapters and servers need id and imp to transact.

Class A, the 14.5 percent of the specification a plain JSON Schema can enforce, carries 64.6 percent of findings on the random sample and 74.3 percent on the purposive one at issue level. Issue counts inflate this class because unknown-field findings fire once per field. Payload-level incidence: class A appears in 38.2 percent of Sample A payloads against class B's 39.9 percent, and in 34.9 percent of Sample B payloads against 29.8 percent. On the random sample the conservative measure does not separate the two classes.

Undefined fields form the largest single finding class in both samples, at 37.9 percent of Sample A findings and 51.0 percent of Sample B. That is the class the tolerant-reader mandate instructs receivers to accept, and the class Section 3.5 notes is contaminated by 2.5-assigned 2.6 fields. It dominates because we scored it as an error.

## 5. Discussion

The results describe two wire protocols sharing a name. SSP response sides largely match the IAB tables, except for a documented no-bid pattern at a few endpoints (seatbid empty, no nbr). Client-side request sides speak Prebid dialects: extra keys, legacy GDPR paths, and a 42 percent floor of type, enum, and required-field mismatches that is the same on a random draw of header-bidding sites and on major publishers. Receivers parse those dialects and they bid. The IAB tables and the endpoint parser are different contracts. This paper measures the first; production uses the second.

The origin leftover is the piece spec authors can use. A deprecated field that every adapter still sends is shared vocabulary, not one vendor's bug. Putting extras in ext, or absorbing documented shared keys into the tables, matches how this stack actually speaks. A JSON Schema generated from the specification's field tables and applied where Prebid constructs a request would reject adapters whose endpoints parse those dialects on purpose. The cost is operational, not CPU. Per-endpoint lint, or an official dialect appendix, fits the data better than a global schema.

The piece ad ops can use is Table 8. Enabling The Trade Desk, Sharethrough, Media.net, or Smart is a type-and-value dialect. Enabling Yahoo or Lijit is mostly a GDPR path. Most sites in the random sample already have a 0 percent strict endpoint on the page. Dropping or keeping an adapter is a demand decision this paper does not price (Section 4.7). It is still a config decision, which writing the dialect is not.

Section 4.7 does not put a market penalty on an empirical footing. The Sample A interval is too wide, the comparison is adapter-confounded, and the pooled figure inverts. The Simpson's paradox is the reusable methods result. A publisher who wants dollars needs an A/B on their own inventory, not another crawl.

## 6. Limitations

Each sample was captured three times from the same residential address. Request-side rates in this adapter set reproduce to within 1.5 points inclusive and 0.6 points strict on Sample A, so those numbers are not artifacts of a single capture. Two coverage limits remain. All measurement comes from one vantage on one metropolitan residential network, and a second vantage is untested. Captures span 14:00 to 00:15 local across weekend and weekday in July 2026. The remaining response-side caveat is not temporal but procedural: Section 4.5 shows the measuring profile stops attracting bids after about a day of repeated visiting, so fill-dependent statistics are reported from first contact and a study wanting many clean fill measurements would need a fresh residential or carrier address per wave, not a datacenter or VPN. The response-side aggregate follows fill and should not be quoted as a fixed rate.

The harness sees client-side traffic only. Server-to-server demand solicited downstream of an endpoint is invisible, as is any adapter using a proprietary or binary encoding. The population is client-side OpenRTB JSON. That is the point of a browser vantage, and it is not all of OpenRTB.

Inventory is display-heavy. The 131 video creatives support the cross-layer analysis as a case

study, not as a general claim about video.

Best-fit version attribution depends on catalog quality, as Section 3.4 documents concretely, and 2.5 assignments score some 2.6 fields as unknown, as Section 3.5 documents. The strict headline already drops those findings.

The gap between the two samples is confined to dialect and legacy field placement rather than type and required-field checks, and two mechanisms that would have predicted a broader inclusive gap were tested and rejected. Why long-tail sites carry more such fields is not established here.

The bid-outcome analysis in Section 4.7 is observational, underpowered on Sample A, and confounded with adapter identity. It is reported as a stratified comparison and a Simpson's paradox, not as an estimate of the cost of a mismatched request. Table 8 is a config menu on the same terms: first-contact Sample A, not a forecast of lost revenue if an adapter is dropped.

Finally, the author maintains the validator used as the instrument. The dataset is published at the level of individual findings with structural paths so that any classification can be recomputed or contested, and the specification-side classification it refers to comes from a separately published study with its own independent recoding check [13].

## 7. Conclusion

At a US residential browser, on sites that run client-side header bidding, 42.0 percent of OpenRTB bid requests are flagged on type, enumerated-value, or required-field checks against the IAB tables, 95 percent CI [36.0, 47.2], and that strict rate is 42.0, 42.6, 42.3 across three recrawls of the same adapters over four days. Including unknown and moved fields the specification tells receivers to tolerate raises the random-sample rate to 69.4 percent; the extra is dialect, and it is also the entire gap versus major publishers. Bid responses at first contact are flagged at 21.4 percent, a mixture that follows fill at a few no-bid endpoints and should be read that way. The divergence from the object tables is overwhelmingly client-side. Publishers write about 2 percent of it and select the rest by which adapters they enable: 76 of 79 sites already have a 0 percent strict option on the page. Inclusive 100 percent is often a moved GDPR path. Strict rates near 90 percent are type and value dialects, listed in Table 8. SSP servers mostly match the tables, except where named endpoints omit nbr on every decline.

Adapters exist to speak SSP dialects. The written spec and the Prebid wire have diverged, in a slice a browser can see, and the type-and-value floor is about 41 percent on both a random draw and a purposive one. The next measurement that would change this picture is a fresh residential or carrier vantage for fill, deals, and geo, or a publisher A/B for revenue, not a datacenter crawl and not another recrawl of the same adapters.

## Data and Artifact Availability

Capture and analysis code, the sampling frame with its seed, and PII-free issue-level datasets for both samples (16,368 findings across 164 sites, with rule identifiers and structural JSON paths but no payload values) accompany this preprint. RTBlint is open source at github.com/aleksUIX/rtblint.

The raw payload corpus is withheld because it contains user identifiers and commercial terms, and is available to researchers on request.

## Suggested Citation

Sekowski, A. (2026). Measuring OpenRTB Dialects in Client-Side Header Bidding. Preprint. doi:10.13140/RG.2.2.26572.78720.

## References


[1] IAB Tech Lab. OpenRTB 2.6 Specification, releases 2.6-202204 through 2.6-202606. github.com/InteractiveAdvertisingBureau/openrtb2.x.

[2] Bradner, S. Key words for use in RFCs to Indicate Requirement Levels. RFC 2119, BCP 14, IETF, March 1997. doi:10.17487/RFC2119.

[3] Leiba, B. Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words. RFC 8174, BCP 14, IETF, May 2017. doi:10.17487/RFC8174.

[4] ISBA and the Association of Online Publishers, carried out by PwC. ISBA Programmatic Supply Chain Transparency Study. ISBA, May 2020.

[5] Association of National Advertisers. Programmatic Media Supply Chain Transparency Study: Complete Report. ANA, December 2023.

[6] Stone-Gross, B., Stevens, R., Zarras, A., Kemmerer, R., Kruegel, C., and Vigna, G. Understanding Fraudulent Activities in Online Ad Exchanges. In Proceedings of ACM IMC 2011, pp. 279-294. doi:10.1145/2068816.2068843.

[7] Pachilakis, M., Papadopoulos, P., Markatos, E. P., and Kourtellis, N. No More Chasing Waterfalls: A Measurement Study of the Header Bidding Ad-Ecosystem. In Proceedings of ACM IMC 2019, pp. 280-293. doi:10.1145/3355369.3355582.

[8] Wang, J., Zhang, W., and Yuan, S. Display Advertising with Real-Time Bidding (RTB) and Behavioural Targeting. Foundations and Trends in Information Retrieval, 11(4-5):297-435, 2017. doi:10.1561/1500000049.

[9] Sassaman, L., Patterson, M. L., Bratus, S., and Shubina, A. The Halting Problems of Network Stack Insecurity. ;login: (USENIX), 36(6):22-32, December 2011.

[10] Thomson, M., and Schinazi, D. Maintaining Robust Protocols. RFC 9413 (Informational), IAB Stream, RFC Editor, June 2023. doi:10.17487/RFC9413.

[11] Prebid.org. Prebid.js and the ortbConverter library. github.com/prebid/Prebid.js.

[12] Sekowski, A. VAST XML Validation at Bid-Time Scale: Latency Analysis and Integration Patterns for Programmatic Video Pipelines. Preprint, April 2026. doi:10.13140/RG.2.2.11404.27520.

[13] Sekowski, A. How Machine-Checkable Is OpenRTB? Classifying the Normative Content of the

Protocol That Clears Real-Time Advertising. Preprint, July 2026. doi:10.13140/RG.2.2.27937.57448.

[14] Le Pochat, V., Van Goethem, T., Tajalizadehkhoob, S., Korczynski, M., and Joosen, W. Tranco: A Research-Oriented Top Sites Ranking Hardened Against Manipulation. In Proceedings of NDSS 2019. doi:10.14722/ndss.2019.23386.

Capture performed July 2026 from a residential connection in the United States. Validator: RTBlint 0.6.0, pinned. Analysis scripts and datasets accompany this preprint.