

The Greek e-shop technical audit

900 Greek domains measured on one day. What Greek online retail actually sends to a browser, and the two characters that inflated a market share nineteen-fold.

Nikolaos Broikos · broikos.gr August 2026 900 domains · 505 e-shops · 7 Aug 2026

36.8%

OF GREEK E-SHOPS SEND NONE OF SIX SECURITY HEADERS

11.1%

FAIL A VALIDATING TLS HANDSHAKE

5 of 505

PASS FIVE ELEMENTARY CHECKS

19×

HOW FAR ONE PATTERN INFLATED A MARKET SHARE

01 What this is

Discharges claims E001, E002, E003, E004.

Nobody publishes a technical measurement of Greek online retail. Greek e-commerce is written about constantly, turnover, growth, consumer surveys, and almost never *measured* at the level of what the sites actually send to a browser.

This paper measures it. **Nine hundred Greek domains, one polite request each, on 7 August 2026.** Seven hundred and thirty-eight responded. Five hundred and five carry the signatures of an online shop, and those five hundred and five are the subject.

Nothing here is a survey. No one was asked anything. Every figure is a property of an HTTP response, recorded and published.

No site is named

The dataset published with this paper identifies each site as GR0001 to GR0900 with a rank band, and carries **no domain name and no page title**. That is a deliberate choice and it costs the paper something: a reader cannot check an individual row against the site it came from.

The alternative was worse. Publishing "these named Greek shops have no security headers" is an accusation aimed at businesses that did not consent to being measured, several of which are competitors, customers or neither and deserve the same treatment. **The aggregate is the finding; the individual site is not.**

Everything needed to reproduce the aggregate is published: the sample frame, the crawler, the classifier and the anonymised per-site measurements. A reader who wants named results can run the crawler themselves in about twenty minutes.

What was measured

One HTTPS GET per domain, homepage only, no crawling beyond it, identifying user agent, twenty-second timeout, no retries. From each response:

- **transport**: TLS version, and whether the certificate validates against a default trust store
- **security headers**: HSTS, CSP, X-Frame-Options, X-Content-Type-Options, Referrer-Policy, Permissions-Policy
- **platform**: commerce platform and underlying CMS, from response signatures
- **markup basics**: mobile viewport, lang attribute, <h1> count, image alt coverage
- **structured data**: Product or Offer in JSON-LD
- **weight**: bytes on the wire, decompressed HTML size, <script> count, response time
- **consent machinery**: presence of a cookie-consent implementation

The headline results

Greek domains in the Tranco top 1M	3,418
measured	900
responded	738 (82%)
classified as e-shops	505
send none of six standard security headers	36.8%
send a Content-Security-Policy	18.4%
TLS certificate not valid to a default trust store	11.1%
have no <h1> at all	40.6%
have Product/Offer structured data on the homepage	1.8%
declare a mobile viewport	97.4%
fail two or more of five basic checks	48.5%
median decompressed homepage HTML	292 KB
heaviest homepage in the sample	2.46 MB

Two of those deserve reading twice. **More than a third of Greek online shops send no security headers at all:** not a weak policy, none of the six. And **one in nine presents a TLS certificate a default trust store rejects**, on a site that takes card payments.

The methodological finding

The first version of this audit reported that **64% of Greek e-shops run Magento**. That was wrong, and the error is instructive enough that section 06 is about it.

The platform classifier matched the pattern `mage-`, which also matches `image-`. It fired on 470 of 738 sites. The word "magento" appears on **30**. The true figure is **3.3%**, so a two-character pattern inflated a platform's apparent market share **nineteen-fold**.

Worse, the first crawler stored only the conclusion, not the evidence, so fixing it meant re-fetching all 900 sites rather than reclassifying offline. The published crawler now records every platform signal per site, including the diagnostic pair that caused the error.

What this paper does not claim

It is not a ranking of Greek retailers, and it names no one.

It is not a performance benchmark. A single unwarmed request from one machine in Athens measures the response, not the user experience. Where speed appears it is labelled as what it is.

It is not representative of *all* Greek e-commerce. The sample frame is the top 900 Greek domains by Tranco rank, which is a popularity list. **Small shops are under-represented by construction**, and section 07 says what that does and does not permit.

It is a first edition. The instrument is published so the second one can be compared with it.

02 The compliance that is visible, and the security that is not

Discharges claims E005, E006, E007, E008.

The single clearest pattern in the data is not that Greek e-shops are careless. It is that they are careful about exactly the things someone can see.

Two-thirds implement cookie consent. One in five of those sends a CSP.

SHARE OF E-SHOPS	
cookie-consent machinery present	67.7%
Content-Security-Policy header sent	18.4%
of the shops with consent machinery, also send a CSP	20.5%

Cookie consent is visible on the first paint of every page, is legally mandated, and is enforced by a regulator that issues fines. A Content-Security-Policy is invisible to every visitor, is mandated by nobody, and is the single most effective defence against the class of attack that steals card details from a checkout page.

The Greek market has implemented the first and not the second, and the gap is not small. Two-thirds against one-fifth.

This is not hypocrisy and it is not stupidity. It is a rational response to what gets checked. Consent banners are checked by regulators and by customers who complain. Response headers are checked by nobody until an incident.

More than a third send nothing at all

Six headers were measured: HSTS, Content-Security-Policy, X-Frame-Options, X-Content-Type-Options, Referrer-Policy and Permissions-Policy. Every one is free, every one is a single line of server configuration, and none requires changing a line of application code.

HEADERS SENT	E-SHOPS	SHARE
	0 of 6	186
	1 of 6	78
	2 of 6	87
	3 of 6	51
	4 of 6	52
	5 of 6	33
	6 of 6	18
		36.8%
		15.4%
		17.2%
		10.1%
		10.3%
		6.5%
		3.6%

Individually:

HEADER	SHARE SENDING IT
X-Content-Type-Options	45.9%
HSTS	44.6%
X-Frame-Options	36.8%
Referrer-Policy	19.6%
Content-Security-Policy	18.4%
Permissions-Policy	10.1%

Fewer than four in a hundred Greek online shops send a complete set. More than a third send none.

The headers that do get sent are the two that some hosting stacks and CDNs add by default. The ones that require a deliberate decision, CSP, Referrer-Policy, Permissions-Policy, are the three least common. **What the market has is what its infrastructure gave it, not what anyone chose.**

Almost half sit behind one company

SERVER HEADER	SHARE
Cloudflare	47.3%
nginx	18.6%
(none sent)	10.7%
Apache	9.5%
LiteSpeed	3.8%
Microsoft-IIS	3.6%

Nearly half of measurable Greek online retail terminates at a single American company. That is not a criticism: Cloudflare is why the HSTS and X-Content-Type-Options numbers are as high as they are, and it is the reason the median response time is 393 ms rather than something embarrassing.

It is a concentration finding. **A configuration change, an outage or a policy decision at one provider reaches about half of Greek e-commerce at once**, and no Greek institution has any say in it. For a sector that regulators and associations describe as strategically important, that is worth stating out loud.

It also means much of the security posture measured here is *inherited* rather than chosen, which is the same point as the previous section from a different direction.

The counter-argument, and why it only goes so far

A reasonable objection: security headers are defence in depth, and their absence does not mean a site is insecure. A shop on a managed platform, with a WAF in front and a payment page hosted by the acquirer, may be perfectly safe with none of the six.

That is true, and it is why this section reports what was sent rather than declaring anyone vulnerable. **No claim is made here that any Greek shop has been compromised or will be.**

But two things survive the objection. The headers are free and take minutes, so their absence measures attention rather than resources. And the sites that skip them are not the sites that did the other work: the shops sending zero headers are disproportionately the shops that also fail the other basic checks in section 05. **Absent headers are a marker, and what they mark is that nobody has looked.**

03 One in nine presents a certificate the browser should reject

Discharges claims E009, E010, E011.

TLS was measured separately from the page fetch, by opening a connection to port 443 and validating the certificate against the operating system's default trust store: the same check a browser performs before showing a padlock.

TLS 1.3	82.4%
TLS 1.2	6.5%
handshake failed or certificate rejected	11.1%

Fifty-six of five hundred and five Greek online shops did not complete a validating handshake. On sites that take payments.

What that figure does and does not mean

It does **not** mean 11% of Greek e-shops show a browser warning. The measurement was made from one machine, at one moment, over a plain default trust store, with no browser-specific fallbacks. Several distinct conditions land in this bucket:

- an expired certificate
- a certificate that does not cover the hostname as requested
- an incomplete chain: the server omits an intermediate that browsers often fetch anyway
- a host that simply refused the connection on that attempt
- geographic or bot filtering that rejects an unfamiliar client at the TLS layer

An incomplete chain is the most likely single explanation, and it is precisely the failure that many desktop browsers paper over and many mobile clients and API integrations do not. A shop in that state looks fine to its owner on a laptop and fails for a payment provider's server-to-server callback.

So the honest statement is narrower than the headline and still uncomfortable: **11.1% of measurable Greek online shops failed a validating TLS handshake from a standard client on 7 August 2026.** The rate at which real customers see a warning is lower than that and is not zero, and this instrument cannot separate the cases.

The good news is real

82.4% negotiated TLS 1.3. That is a genuinely strong number, better than the security-header result by a wide margin, and it did not happen because Greek merchants configured it.

It happened because TLS 1.3 is what modern servers and CDNs default to, and because Let's Encrypt made certificates free and automatic. The same reason the security-header numbers are poor: **the market runs on defaults**: is the reason the transport numbers are good. When the default is safe, Greek e-commerce is safe.

That is the practical lesson and it generalises past this sector: **for anything a merchant will not actively configure, the only intervention that works is changing what arrives switched on.**

It is not evenly distributed

Breaking the failures down by platform shows the problem concentrating:

PLATFORM	SHOPS	TLS HANDSHAKE FAILED	MEAN SECURITY HEADERS OF 6
OpenCart	19	26%	1.11
Magento	24	21%	2.88
PrestaShop	14	14%	1.43
no platform detected	356	9%	1.76
WooCommerce	81	7%	1.36

Two things stand out.

OpenCart shops are worst on both axes at once: the highest handshake-failure rate and the fewest security headers of any identified platform. The sample is nineteen shops, which is too few to be conclusive, but the direction is consistent across two independent measures.

Magento is the contradiction. Its shops send by far the most security headers (2.88 of six, more than double WooCommerce) and are second worst on certificate validity. That is the signature of a platform used by larger operations with real IT attention on the application, and neglected infrastructure underneath: the application team configured the headers, and nobody owns the certificate chain.

Both readings are hypotheses at these sample sizes. What is not a hypothesis is the aggregate: **one in nine, on sites that take card payments.**

04 What Greek shops are actually built on

Discharges claims E012, E013, E014.

PLATFORM SIGNATURE	E - SHOPS	SHARE
none detected	356	70.5%
WooCommerce	81	16.0%
Magento	24	4.8%
OpenCart	19	3.8%
PrestaShop	14	2.8%
Shopify	4	0.8%
CS-Cart	4	0.8%
Wix	2	0.4%
Squarespace	1	0.2%

And the underlying content management system, where one is detectable:

CMS	SHARE OF E-SHOPS
WordPress	36.0%
Drupal	6.7%
Joomla	3.6%
Nuxt	2.8%
Next.js	2.6%
none detected	~48%

The number that matters is 70.5%

Seven in ten Greek online shops show no recognisable commerce platform signature at all. That is the finding, and it needs to be read carefully because it has two very different explanations and the data supports both.

Explanation one: the sample is popularity-ranked. These are the top Greek domains by traffic. Large retailers build custom, or buy from Greek software houses whose products carry no public fingerprint, or run platforms behind a CDN that strips identifying paths. A small shop is far more likely to be a stock WooCommerce install than a top-900 domain is.

Explanation two: detection from a homepage is weak. A commerce platform announces itself through asset paths, cookie names and script bundles that often appear on product and cart pages rather than the homepage. Only 1.8% of these shops carry Product structured data on the homepage, for the same reason: the homepage is a shop window, not a product page.

Both are true, and together they mean **70.5% is an upper bound on "custom or unidentified", not a measurement of custom builds.** Stated the other way round, which is the defensible direction: **at least 29.5% of top Greek e-shops run an identifiable off-the-shelf platform, and WooCommerce is more than half of those.**

WordPress is the substrate

36.0% of these shops run WordPress, against 16.0% carrying WooCommerce signatures. **More than twice as many Greek shops run WordPress as run the WordPress commerce plugin**, which means a large group is using WordPress for the site and something else, a custom cart, a hosted checkout, a marketplace listing, for the transaction.

That is a distinctively small-market pattern. It says the content problem was solved with the cheapest available tool and the commerce problem was solved separately, rather than a commerce platform being adopted wholesale.

For anyone selling to this market, it is the more useful of the two numbers: **the addressable base for a WordPress-adjacent commerce product is roughly twice the installed WooCommerce base.**

Shopify has not taken Greece

Four shops. 0.8%.

Shopify's global share of hosted commerce is an order of magnitude above what appears here. Whether that is a pricing story, a language and localisation story, a payment-integration story or simply a market that has not turned over yet, this instrument cannot say: it measures what is installed, not why.

What it does establish is that in the top 900 Greek domains, **the hosted-SaaS commerce model has close to no presence**, and Greek online retail runs on self-hosted software: WooCommerce, Magento, OpenCart, PrestaShop, and a large unidentified remainder.

Read the platform breakdowns as directional

Beyond WooCommerce at 81 shops, every platform group is small: Magento 24, OpenCart 19, PrestaShop 14, and single digits below that. **A difference of two or three shops moves a percentage point.**

The cross-platform comparisons in section 03 are reported because the direction is consistent across two independent measures, not because nineteen OpenCart shops settle anything. The aggregate figures in this paper rest on 505 shops. The per-platform figures do not, and are labelled wherever they appear.

05 Weight, markup and the compound failure

Discharges claims E015, E016, E017, E018.

Baseline hygiene: mostly solved, with two holes

CHECK	SHARE OF E-SHOPS
lang attribute on <html>	98.0%
mobile viewport declared	97.4%
response compressed	94.5%
exactly one <h1>	45.1%
no <h1> at all	40.6%
Product/Offer structured data	1.8%

The first three are effectively solved. Responsive design and compression are defaults now, and Greek e-commerce has them.

The <h1> result is not a default and it shows. **Two in five Greek e-shop homepages have no <h1> element**, which is both a search-engine signal and the primary landmark a screen-reader user navigates by. Another 14% have more than one. Fewer than half get the simplest structural element in HTML right.

Structured data at **1.8%** is partly an artefact of measuring homepages, as section 04 explains. It is still worth noting what it costs: Product and Offer markup is what produces price, availability and review stars in a search result, and it is what an answer engine reads when deciding whether a shop sells a thing. A market that does not emit it is invisible to that layer by construction.

Image alt text: a good median hiding a bad tail

median alt coverage	99%
shops at 100%	47.9%
shops below 50%	8.4%
shops at 0%	0.6%

Nearly half of Greek e-shops label every homepage image, and the median shop labels essentially all of them. That is a better accessibility result than this paper expected to find and it should be said plainly.

The tail is the problem. **Forty-two shops label fewer than half their images and three label none.** For a screen-reader user those sites are not degraded, they are unusable: an unlabelled product grid is a list of blank links.

The distribution matters more than the average here, and it is the reason this paper reports quantiles rather than means throughout. A market at 99% median and 8.4% below half is not "mostly accessible"; it is two markets.

Weight

Homepage only, single unwarmed request, one machine in Athens.

	MEDIAN	P90	MAX
response time	393 ms	1,200 ms	21,112 ms
bytes on the wire	51 KB	141 KB	400 KB
decompressed HTML	292 KB	868 KB	2.46 MB
<script> tags	34	77	466

The median Greek e-shop homepage is 292 KB of HTML before a single image, stylesheet or script is fetched, and carries 34 script tags. At p90 it is 868 KB and 77 scripts. One site sends 2.46 MB of HTML and another carries 466 script tags.

Compression is doing enormous work: 292 KB of HTML arrives as 51 KB on the wire. Without it the median page would be six times heavier.

The response times are respectable, 393 ms median, and that is largely the CDN concentration from section 02 rather than merchant tuning. **What the CDN cannot fix is the 292 KB of markup and 34 scripts it is caching,** which land on the customer's device and their battery regardless of how fast they arrive.

The compound failure

Five checks, all elementary: any security header at all, a mobile viewport, a valid TLS handshake, at least half of images labelled, and product structured data.

CHECKS FAILED	E-SHOPS	SHARE
0 of 5	5	1.0%
1 of 5	255	50.5%
2 of 5	203	40.2%
3 of 5	37	7.3%
4 of 5	4	0.8%
5 of 5	1	0.2%

Five Greek online shops out of 505 pass all five. One percent.

48.5% fail two or more.

The single most common failure is structured data, which is partly a homepage artefact and should not be over-read. But the shape is the point: the failures are not spread evenly across a competent market with a few stragglers. They stack. A shop with no security headers is disproportionately a shop with no <h1>, no structured data and a certificate that does not validate, because all four are symptoms of the same thing, which is that nobody has ever looked at the site with a technical eye.

That is the actionable finding in this paper. **The gap in Greek e-commerce is not sophistication, it is attention.** Every check in this section is free and takes an afternoon.

Does being bigger help?

The sample is ranked, so it can answer a question the aggregates cannot: does technical quality improve as a Greek shop gets more popular? The 505 e-shops split into quartiles by Tranco rank, Q1 being the most visited.

QUARTILE	N	SEND NO HEADERS	SEND A CSP	TLS INVALID	NO <H1>	MEDIAN MS
Q1 (most visited)	126	23.0%	23.8%	10.3%	47.6%	380
Q2	126	42.9%	13.5%	4.8%	37.3%	343
Q3	126	40.5%	19.8%	15.9%	40.5%	382
Q4 (least visited)	127	40.9%	16.5%	13.4%	37.0%	447

Security headers are the one thing that improves with size, and the step is large. The most-visited quartile is roughly half as likely to send nothing as any other, 23.0% against 40.9%, and the most likely to send a CSP. Below the top quartile the three remaining groups are indistinguishable from each other: Q2, Q3 and Q4 sit within two points on the no-headers measure.

So this is not a gradient. **It is a threshold.** Somewhere near the top of the Greek market, sites acquire someone whose job includes response headers, and below that line it makes no difference whether a shop is ranked two hundredth or nine hundredth.

Nothing else improves. Certificate validity does not track popularity at all: the second quartile is the best in the sample at 4.8% and the third is the worst at 15.9%, which is noise around a flat line rather than a trend. Median response time is flat to slightly better at the top, and the differences are inside what a single unwarmed request can resolve.

And one measure runs backwards. ****The most-visited Greek e-shops are the *most* likely to have no <h1> at all****: 47.6% against 37.0% at the bottom of the sample. That is worth stating plainly because it is the opposite of what a "bigger sites are better built" story predicts.

The likely explanation is visible elsewhere in the data: the busiest homepages are the heaviest and the most script-driven, built as component-assembled landing pages where the heading structure is an outcome of the component tree rather than a decision anyone made. A small shop running a stock theme gets an <h1> because the theme author put one there. A large retailer's bespoke homepage has whatever its components emit.

The reading that survives all four columns is narrow: popularity buys a security-header policy at the very top of the market and buys nothing else. Certificate hygiene, markup basics and speed are not functions of size, which means they are not functions of budget either, and that in turn is the strongest evidence in this paper for the claim in section 05 that **the gap is attention rather than resources**.

06 The two characters that inflated a market share nineteen-fold

Discharges claims E019, E020, E021.

The first version of this audit found that **64% of Greek e-shops run Magento**. It was published nowhere, because the number was checked before it was written up, and it is worth a section because of how ordinary the mistake was.

The bug

The platform classifier matched Magento with the pattern `mage-asset`. The second alternative was meant to catch Magento's `mage-` asset prefix.

It also matches `image-`.

`image-wrapper`, `image-slider`, `image-container`, `image-gallery`. Every site with a class name containing `image-`, which is nearly every site with pictures on it, which is every online shop, was labelled Magento.

The size of it

Both signals were re-measured on the corrected crawl and recorded per site, so the error can be quantified rather than described:

	SITES	SHARE OF REACHABLE
matched <code>mage-</code>	470	63.7%
contained the word "magento"	30	4.1%
matched <code>mage-</code> without "magento"	442	59.9%

Four hundred and forty-two sites out of 738 were mislabelled. The corrected Magento share among e-shops is **4.8%**, against a first-pass figure of 64%. **A two-character pattern inflated one platform's apparent market share by a factor of nineteen.**

Why it survived the first reading

Nothing about the result looked wrong. Magento is a real Greek e-commerce platform with real market presence. A dominant-incumbent finding is a plausible shape for a small market. And the number arrived inside a table where every *other* figure was correct: the security headers, the TLS rates and the weight percentiles from that same crawl all survived re-measurement unchanged.

It was caught by a single question that had nothing to do with the code: **is 64% a believable share for Magento in Greece?** It is not. That is the whole check. No amount of re-reading the regular expression would have prompted it, because the regular expression looks fine.

The expensive part was not the bug

The first crawler stored the **conclusion**: platform: "Magento", and discarded the HTML it was derived from. So fixing a two-character pattern required re-fetching all 900 sites, because there was nothing left to reclassify.

The corrected crawler records every platform signal per site, plus the diagnostic pair `image-` and `image-` that caused the error. Any future change to the classification rule can now be applied to the stored data at no cost.

Store the evidence, not the verdict. A derived field is a claim; the signal it was derived from is the data. Keeping only the claim means every correction costs a full re-collection.

The pattern this makes four

This is the fourth measurement error caught in this programme in a week, and they are all the same shape: **a check that fires on something adjacent to what it means:**

PAPER	THE CHECK	WHAT IT ACTUALLY MATCHED
04	"the agent mass-deleted the project"	the agent clearing its own <code>__pycache__</code>
04	"the agent violated a read-only instruction"	the test harness's own session directory
05	"cost tripled at the last version"	a protocol change to three repetitions per task
06	"64% of Greek shops run Magento"	the string <code>image-</code>

Two of the four were published before being caught. All four were reproducible, all four were plausible, and every one made a better headline than the truth.

The defence that worked in all four cases was the same, and it is not code review: **recompute the number from the raw data and ask whether the answer is believable.** Three of the four were caught by the second half of that sentence.

07 How the anonymisation was done, and how it nearly failed

Discharges claims E029, E030, E031.

This paper measures businesses that did not agree to be measured. That imposes an obligation, and the obligation is not satisfied by deleting a column.

What is published

Each of the 900 measured sites appears as GR0001 to GR0900 with:

- a **rank band** of 200 (so a row says roughly where in the popularity ordering it sat, not where exactly)
- every technical measurement: platform, CMS, TLS, headers, markup, weight, timing

and **no domain name, no page title, no URL, and no server IP.**

What is not published, and why that changed

The first version of the dataset also published **the ranked list of 3,418 .gr domains** that formed the sample frame. That list is derived from a public source and adds nothing a reader could not regenerate, so publishing it seemed harmless.

It was not harmless. **The site_id values were assigned in rank order.** GR0001 was the highest-ranked domain in the frame, GR0002 the second, and so on. Publishing the ranked frame alongside a rank-ordered anonymised table means the two files **join**, and every "anonymous" row becomes a named site with a lookup.

The anonymisation was defeated by the ordering of the rows, which is not a field and does not appear in any column, and which no amount of inspecting the published columns would reveal.

Two changes were made:

- 01 **The rows were shuffled** with a fixed, recorded seed before site_id was assigned, so the identifier carries no ordering information. The seed is in the publication script, so the shuffle is reproducible without being invertible from the published file.
- 02 **The ranked frame was withdrawn from publication.** It is regenerable by anyone from Tranco list PYG5J in a few minutes, so nothing is lost except the join.

What anonymity remains, stated honestly

It is partial, and a determined reader with the crawler could re-identify many rows.

The rank band narrows a row to roughly 200 candidates. The technical fingerprint, platform, CMS, server, TLS version, header set, script count, byte counts, narrows it much further, and this paper should say by how much rather than leave it at "much".

Four columns are enough. Image count, script count, HTML bytes and wire bytes identify 726 of the 738 reachable rows uniquely, 98.4%. A reader who re-runs the published crawler against the frame can therefore match essentially every row, because the measurements are deterministic properties of public pages.

That is worth stating exactly, because it is the size of the thing being traded away and a reader is entitled to weigh it. It is also worth stating what it does *not* buy. **Re-identification requires the re-crawl, and the re-crawl already yields the security posture directly.** Nobody learns anything from this table that the frame and one afternoon would not give them without it.

So the published dataset is **not anonymous against a motivated analyst**. What it is:

- **anonymous against casual reading**, which is what prevents a headline listing named Greek businesses as insecure;
- **anonymous against republication**, since no one can lift a table of names out of this paper;
- **useless as an attack list**, since it contains no addresses and identifies no vulnerability, only the absence of hardening that any visitor could observe.

The alternative, publishing nothing per site, would have made every aggregate in this paper unverifiable. **The trade is deliberate: verifiability for the aggregate, obscurity rather than secrecy for the individual**, and the limits of that obscurity are stated here rather than implied.

Why the measurements are not sensitive in the first place

Nothing recorded here is private. Every field is what a web server sends to any visitor who opens the homepage: response headers, the certificate it presents, the markup it returns. A browser's developer tools show all of it, on any site, to anyone.

The audit made **one request per domain**, the load of a single page view, with an identifying user agent linking back to this research, no retries, no login attempt, no form submission and no crawling beyond the homepage.

No vulnerability was probed and none is reported. The absence of a Content-Security-Policy is not an exploit; it is a configuration choice observable from the outside. That distinction is why this audit could be published at all, and it is the line the second edition should stay behind when it extends to checkout pages.

The general lesson

The near-failure here is the same shape as the classifier error in section 06 and, for that matter, the errors recorded in the companion papers of this programme: **the defect was not in what the code computed but in a property nobody thought of as data.**

A regular expression matched something adjacent to its meaning. A row ordering carried an identity that no column contained. In both cases the published artefact looked correct on inspection, and in both cases the check that caught it was asking what the output would let someone do: not reading the code that produced it.

08 What this measures, and what it does not

Discharges claims E022, E023, E024, E025.

The sample is popularity-ranked, not representative

The frame is the Tranco top-1M list of 6 August 2026, filtered to .gr, taking the top 900 by rank. Tranco is built for research precisely because raw popularity lists are unstable, and it is citable and reproducible, but it is still a **popularity** list.

Small Greek shops are under-represented by construction. A shop doing thirty orders a month does not appear in a global top-1M ranking. Every figure in this paper therefore describes *the visible end* of Greek online retail, and the visible end is the end with the most resources.

That cuts one way only, and it is worth being explicit about the direction: **if the best-resourced 900 Greek domains produce a 36.8% no-security-headers rate and an 11.1% TLS failure rate, the rate across the whole population is very unlikely to be better.** This paper's numbers are a floor, not an estimate.

One request, one moment, one machine

Every measurement is a single unwarmed HTTPS GET from one machine in Athens on 7 August 2026, homepage only.

- **Response times are not user experience.** They are one server's response to one request from one network position. No rendering, no assets, no device.
- **A single fetch cannot distinguish a permanent condition from a bad moment.** Some of the 11.1% TLS failures and some of the 162 unreachable domains will have succeeded on a retry. No retries were made, deliberately, to keep the load at exactly one request per site, and that choice inflates the failure numbers somewhat.
- **The homepage is not the shop.** Product structured data, cart behaviour, checkout security and payment integration all live on pages this audit never requested. The 1.8% structured-data figure in particular should be read as "on the homepage", and section 04 says so.

Detection is a floor, not a census

Platform and CMS detection works by response signatures. A site behind a CDN that rewrites asset paths, or one built by a Greek software house whose product has no public fingerprint, will show as "none detected". **70.5% undetected is an upper bound on custom-or-unknown, not a measurement of custom builds.**

The defensible direction is the reverse: at least 29.5% run an identifiable off-the-shelf platform.

The classification rule is a choice, and both versions are published

A site counts as an e-shop if it shows a commerce platform signature, or Product/Offer structured data, or cart wording. That broad rule gives **505 shops from 738 reachable domains**.

A strict rule, platform signature or structured data only, no cart wording, gives **157**. The cart-wording signal is the loose one: a news site with a "καλάθι" in a sidebar advert will match it, so the broad rule certainly admits some sites that are not shops.

Every headline finding survives the stricter rule, and most get worse:

	BROAD RULE (N=505)	STRICT RULE (N=157)
send no security headers	36.8%	40.8%
send a CSP	18.4%	21.0%
TLS handshake invalid	11.1%	16.6%
no <h1>	40.6%	38.9%
cookie-consent machinery	67.7%	72.6%

The paper reports the **broad** rule throughout, which is the conservative choice: it produces the *lower* security-failure rates. Restricting to sites that are unambiguously shops raises the no-headers rate by four points and the TLS-failure rate by more than five. The consent-versus-CSP gap in section 02 widens rather than narrows.

Both counts are in the published data so a reader can recompute every figure under either rule.

Small groups are directional only

Beyond WooCommerce at 81 shops, every platform group is small: Magento 24, OpenCart 19, PrestaShop 14. Two or three shops move a percentage point. The cross-platform comparisons in section 03 are reported because two independent measures point the same way, not because these samples settle anything.

No named sites, and what that costs

The published dataset carries no domain names and no page titles, for the reasons in section 01. The cost is real: **a reader cannot verify an individual row**, and must either trust the aggregate or re-run the published crawler.

The crawler, the sample frame and the anonymised per-site table are all published, so re-running it takes about twenty minutes and produces named results the reader owns. That is the intended path for anyone who wants to check a specific claim.

Errors in this paper

One classifier error, corrected before publication and documented in section 06. The first pass reported Magento at 64%; the true figure is 4.8%. It required a full re-crawl to fix because the first crawler stored conclusions rather than signals.

One anonymisation defect, caught before publication. The per-site table originally assigned identifiers in Tranco rank order while the ranked frame was also published, so the two files joined and every row was re-identifiable. Rows are now shuffled under a recorded seed and the frame is withdrawn. Section 07 documents it in full.

One malformed record. The corrected crawl produced 901 lines of which one was truncated at a buffer boundary and could not be parsed. It was dropped, leaving 900 usable records, and this is recorded rather than silently ignored.

What the second edition should do

- 01 **Fetch a product page as well as the homepage.** It would settle the structured-data question and considerably improve platform detection.
- 02 **Retry once on failure**, to separate permanent TLS problems from transient ones.
- 03 **Sample below the top 900**, so the small-shop population is represented rather than excluded.
- 04 **Measure checkout**, which is where the security posture actually matters and where this audit does not look.
- 05 **Run it annually and compare.** One edition is a snapshot. The instrument is published so the second one can be a trend.

09 What follows

Discharges claims E026, E027, E028.

For a Greek merchant

Everything below is free, takes an afternoon, and is measured in this paper as absent from a large share of the market.

1. Send the six headers. 36.8% of Greek e-shops send none of them; 3.6% send all six. This is server configuration, not development work, and if you are among the 47.3% behind Cloudflare, most of it is a toggle. **It is the cheapest competitive gap in this entire dataset.**

2. Check your certificate chain from something that is not your laptop. 11.1% failed a validating handshake from a standard client. The most likely cause is a missing intermediate, which desktop browsers often hide and mobile clients and payment callbacks do not. Test from a phone on mobile data and from an external checker.

3. Put an <h1> on the page. 40.6% have none. It is one tag, it is the primary landmark for screen-reader navigation, and it is a search signal.

4. Emit Product and Offer structured data. 1.8% of homepages carry it. This is what puts price and availability into a search result and what an answer engine reads when deciding whether you sell a thing. A market that does not emit it is invisible to that layer.

5. Look at the tail of your alt text, not the average. The median Greek shop labels 99% of homepage images. 8.4% label under half. If you are in the tail, your product grid is a list of blank links to a screen-reader user.

6. Weigh your homepage. The median is 292 KB of HTML and 34 scripts before any asset loads. A CDN makes that arrive fast; it does not make it smaller on the customer's device.

For anyone selling to this market

The addressable base is not what platform-share numbers suggest. 36.0% of these shops run WordPress and 16.0% carry WooCommerce signatures: **more than twice as many Greek shops run WordPress as run its commerce plugin.** A WordPress-adjacent commerce product has roughly double the market a WooCommerce-only view would show.

Hosted SaaS has not landed. Shopify is 0.8% of this sample. Greek online retail is self-hosted, and the reasons are not measured here.

Almost half of it terminates at one American company. 47.3% behind Cloudflare is a concentration worth understanding whether you are selling infrastructure, assessing sector risk, or writing policy.

For the sector

The pattern in section 02 is the one that should travel: **two-thirds of Greek e-shops implement cookie consent and one in five of those sends a Content-Security-Policy.**

That is not a market that does not care. It is a market that does exactly what is checked. Consent banners are visible, mandated and enforced; response headers are invisible, mandated by nobody, and checked by no one until an incident.

The intervention that follows is not an awareness campaign. It is that **the defaults have to change**, and the transport numbers prove it works. TLS 1.3 is at 82.4% not because Greek merchants configured it but because it arrived switched on, and free automated certificates removed the last excuse. When the safe thing is the default, this market has it.

For this paper

This is edition one of an annual index, and it is deliberately a floor rather than an estimate: the sample is the best-resourced end of Greek online retail, so the true population rates are unlikely to be better than what is reported here.

The instrument is published, sample frame, crawler, classifier and the anonymised per-site table, so that the second edition can be a comparison rather than another snapshot, and so that anyone who disputes a number can produce their own.

The most useful thing in it may not be any of the findings. **Five shops out of 505 pass five elementary checks.**
Not five per cent. Five.

10 **References and instruments**

One third-party source, used only as the sample frame. Everything else is an instrument built for this audit and published in full. Full URLs are printed as text so the PDF is standalone.

Sample frame

- 1 **Le Pochat, V., Van Goethem, T., Tajalizadehkhoob, S., Korczyński, M. and Joosen, W. Tranco: a research-oriented top sites ranking hardened against manipulation.** List PYG5J, generated 6 August 2026 over the 30 days to that date.
<https://tranco-list.eu/list/PYG5J/1000000>
*Filtered to .gr, giving **3,418 Greek domains**, of which the top 900 by rank were measured. Tranco is used because raw popularity lists are unstable and unciteable; it is designed to be reproducible.*

Instruments published with this paper

- 1 **Broikos, N. (2026). Greek e-shop technical audit: crawler.** One HTTPS GET per domain, homepage only, identifying user agent, 20-second timeout, no retries, no crawling beyond the homepage.
<https://broikos.gr/research/data/paper-06/crawl.py>
Records every platform signal per site rather than only the classification, including the mage- / image- pair that caused the error documented in section 06.
- 2 **Broikos, N. (2026). Anonymised per-site measurements.** 900 rows. **No domain name and no page title.**
<https://broikos.gr/research/data/paper-06/eshop-audit-anonymised.csv>
Sites are GR0001 to GR0900 with a rank band. Every figure in this paper is recomputable from this file under either classification rule.
- 3 **Broikos, N. (2026). Analysis script.**
<https://broikos.gr/research/data/paper-06/analyse.py>
Regenerates every table in sections 02 to 05.
- 4 **The ranked .gr frame is deliberately NOT published.** Publishing it alongside a rank-ordered per-site table would have made every anonymised row re-identifiable by a simple join. See section 07. It is regenerable from the Tranco list above in a few minutes.

Measurement date

All measurements were taken on **7 August 2026**. Response headers, certificates and page weight change continuously; a figure in this paper is a property of that day.

HOW TO CITE THIS PAPER

Broikos, N. (2026) *The Greek e-shop technical audit: what 900 Greek domains actually send to a browser*. Edition 1, 7 August 2026. Athens. Available at: <https://broikos.gr/research/greek-eshop-audit.pdf> (Accessed: DD Month YYYY).

VERSION

Edition 1, published 7 August 2026. This is intended as an annual index. The instrument is published so that edition 2 can be a comparison rather than another snapshot.

DATA AVAILABILITY

Published at <https://broikos.gr/research/data/paper-06/> under the same licence as the text: the crawler, the analysis script, and the anonymised 900-row measurement table. The ranked domain frame is deliberately withheld; section 07 explains why.

WHY NO SITE IS NAMED

The published dataset carries no domain names and no page titles. Publishing a list of named Greek businesses with weak security posture would be an accusation aimed at companies that did not consent to being measured. The aggregate is the finding; the individual site is not.

The cost is stated in section 07: a reader cannot verify an individual row, and must either accept the aggregate or re-run the published crawler, which takes about twenty minutes and produces named results the reader owns.

METHOD AND ETHICS

One HTTP request per domain. Homepage only, no crawling beyond it, an identifying user agent carrying a link to this research, a 20-second timeout and no retries. The load placed on any measured site was a single page view.

No login was attempted, no form submitted, no personal data collected, and nothing was accessed that a browser would not fetch when a visitor opens the homepage.

DECLARATION OF INTERESTS

The author operates e-commerce businesses in Greece and therefore competes with some of the sites measured here. Neither the author's own shops nor any other site is identified in the published data, and the classification and measurement code was written before the results were seen.

The author also sells web development services, which makes a paper reporting widespread technical neglect in the Greek market a commercially convenient finding. That is exactly why every figure is recomputable from the published table and why the crawler is published: **the appropriate response to a self-interested finding is to check it, and the means to check it is provided.**

LICENCE

This paper and the datasets published with it are released under a **Creative Commons Attribution 4.0 International licence (CC BY 4.0)**. You may copy, redistribute, quote, chart and build on this material, including commercially, provided you credit the source. Licence text:

<https://creativecommons.org/licenses/by/4.0/>, and served beside this paper at

<https://broikos.gr/research/LICENSE-CC-BY-4.0.txt>. What the grant covers and what it withholds is itemised in <https://broikos.gr/research/LICENSING.md>. Analysis code is MIT:

<https://broikos.gr/research/LICENSE-MIT.txt>

FUNDING AND INDEPENDENCE

Unfunded. No sponsor, client or commissioning party paid for, commissioned, reviewed or approved this research. No platform vendor, hosting provider or CDN was contacted.

CORRECTIONS

If a figure here is wrong, write and it will be corrected with the date and the reason recorded in the public change log at <https://broikos.gr/research/corrections.md> rather than silently edited. **Section 06 records an error caught during production** that would have overstated one platform's market share nineteen-fold.

Contact: <https://broikos.gr/contact> · <https://broikos.gr>

ABOUT THE AUTHOR

Nikolaos Broikos operates e-commerce businesses in Greece, works in web development and digital strategy, and builds agent harnesses. He writes from Athens.

No academic affiliation, no institutional backing, no funding, and no client commissioned any of this.

The papers therefore ask to be judged on their published instruments, data and corrections rather than on credentials: every dataset, every claim ledger including the claims that failed, and every retraction is published alongside the text, so a reader who distrusts the author can check the work instead.

THE OTHER PAPERS IN THIS PROGRAMME

Independent, unfunded and published free under CC BY 4.0, each with its underlying data. Read separately; they share a method, not an argument.

1 Who publishes the numbers: a census of corporate research in Greece

<https://broikos.gr/research/who-publishes-the-numbers.pdf>

68 Greek publication programmes coded on a published schema: who produces the numbers business decisions rest on, what motivates them, how far the press carries them, and which of them answer engines actually cite.

2 What a Greek online order actually costs

<https://broikos.gr/research/what-a-greek-online-order-costs.pdf>

What the €36.1bn e-commerce figure counts and what it does not, the convergence of Greek and European online buying, and a landed-cost model built from published tariffs rather than quoted rates.

3 The adoption gap: what businesses say about AI and what the statistics measure

<https://broikos.gr/research/the-adoption-gap.pdf>

Where Greek firms actually sit on AI adoption once the size classes are separated, why the skills-gap explanation does not survive the spending data, and what Greek firms bought instead.

4 What actually makes an agent harness work

<https://broikos.gr/research/agent-harness.pdf>

Techniques, prompt engineering and measured results from three agent harnesses, two built by the author and one not, across 45 graded sessions on a single model, including three faults found in the measuring apparatus itself.

5 The Greek SME digital bill

<https://broikos.gr/research/greek-sme-digital-bill.pdf>

What it costs per year to sell online in Greece, priced entirely from published pages: EUR 157.50 at the floor. Of 25 cost lines, 7 have no published price at all, and they are disproportionately the mandatory ones.

6 What an agent actually costs

<https://broikos.gr/research/what-an-agent-costs.pdf>

Seven versions of one production agent on the same benchmark and the same model, including the two versions that got worse, the failure taxonomy, and the chart that would have shown a cost explosion that never happened.

ABOUT THE MEASUREMENTS

900 Greek domains, one HTTPS request each, homepage only, on 7 August 2026. Sample frame: the Tranco top-1M list of 6 August 2026 filtered to .gr, which yields 3,418 domains. 738 responded; 505 classify as e-shops.

Load placed on any measured site: a single page view. No login, no form submission, no crawling beyond the homepage.

NO SITE IS NAMED

The published table identifies sites as GR0001 to GR0900 with no domain name and no page title, and the rows are shuffled so the identifier carries no ranking information.

Section 07 states exactly how far that anonymity goes and where it stops.

CORRECTIONS

Section 06 records an error caught before publication that would have overstated one platform's market share nineteen-fold, and section 07 records an anonymisation defect that would have made every row re-identifiable.

broikos.gr