peekbird

Peekbird on 1,009 real websites

Session replays, A/B tests, heatmaps, analytics and the statistics behind results, measured on the same 1,009 popular websites on a computer and on a phone.

Field report
October 2026

The sites
Drawn from the Tranco list of the most visited websites (list Y83KG, 3 October 2026)

Measured by
Peekbird, on its own product (version 423b009). Figures as measured, to one decimal.

SUMMARY

The numbers in brief

Each website was visited with Peekbird on, as a visitor would use it, and everything Peekbird made of the visit was checked against what happened: the replay against the live page, the click map against the clicks, the measurements against Google's own library.

REPLAYS

99.5%

of the live page drawn the same by a replay (median, Growth and Scale, at each page's worst moment)

HEATMAPS

97.9%

of 13,639 clicks and taps reached Peekbird; 80.2% drawn on the element clicked

ANALYTICS

100%

of Web Vitals (LCP, CLS, INP) agree with Google's web-vitals library, where both measured

A/B TESTS

125 ms

median time a page is held hidden while a variation applies; the original seen first on 3% of pages at most

SPEED

65 ms

of main-thread time over a 25-second visit with analytics, automatic events and recording on (median); the page's own scripts: 310 ms

STATISTICS

4.3%

of A/A tests called wrong on some day when checked daily (simulated); a fixed-horizon test checked daily: 22.7%

Where it falls short

- One page in ten replays below 74.2% the same at its worst moment. Worst are pages with streamed video (median 93.3%), script animations (98.4%), canvases (98.5%) and videos (98.5%); see section 2.
- Clicks on parts that stay on screen, like sticky headers, are drawn on the element clicked 70.8% of the time; menus and pop-ups opened after the page loaded mostly aren't on the page copy the heatmap is drawn on (section 4).
- 17.4% of sites send a Content-Security-Policy that blocks Peekbird's script until they list it, and 13% block its events: the install has to add it.
- 90.9% of searches in a site's own search box were detected with no setup; a site can send the rest from its own code.

SECTION 1

How it was measured

The websites

Drawn at random from the Tranco list (a research ranking of the most visited websites, combining several popularity lists), list Y83KG of 3 October 2026, by popularity: 600 from 1–1,000, 800 from 1,001–10,000, 800 from 10,001–100,000, 300 from 100,001–1,000,000, seed 1. One site per name (google.com, not also google.de), only domains that answer as websites (not content networks or APIs), and no adult sites: 2,426 sites.

Each was visited on a computer (Microsoft Edge, 1280 × 800) and on a phone (390 × 844, touch, 3× pixels), on its home page, as a first-time visitor. The run was stopped at about 87% of its visits, before most phone visits to the least popular band, so the report pairs the devices: **every figure comes from the 1,009 sites checked on both**, and the computer's and the phone's figures are from the same sites. In practice it covers websites in the top 100,000.

Visits to the 1,790 sites visited on both devices	Visits
Checked (the 1,009 sites checked on both)	**2,018**
Turned automated browsers away	655
Name lookup failed (mostly brief outages of the network the fleet ran on)	358
Other	106
Never finished loading	105
Unreachable	74
No page there (other HTTP 4xx)	15
Server error (HTTP 5xx)	14
Checked on one device only (left out)	235
All	3,580

Turned automated browsers away: a bot check, or an answer of 401, 403 or 429. Name lookups failed mostly in brief outages of the network the test ran on, at random, so they leave no bias.

What was checked on each page

- **Replays.** The page recorded as it loaded and scrolled, then replayed three ways; screenshots of the live page and of each replay at the same moments compared pixel by pixel.
- **A/B tests.** A simulated experiment on the page through Peekbird's real snippet, its variation forced: seven kinds of visual change on elements picked on the page, variation code, a split-URL test and a feature flag.
- **Heatmaps and analytics.** One visit with analytics, automatic events and recordings on, going down the page a screen at a time: on each of five screens, a click (a tap on phones) on a link or button and one on something else. Every click drawn on the dashboard's copy of the page (made from another visit's recording) by the dashboard's own code. Scroll depth, Web Vitals beside Google's web-vitals library in the same visit, automatic events, and a real search in the site's own search box.
- **Speed.** A CPU profile of that visit: how long Peekbird's code ran on the main thread, and the page's own scripts.

- **Statistics** can't be tested on real sites without real experiments: they were simulated, 2,000 experiments a scenario, judged by the results page's own code (section 7).

What it can't see

- What a page sends as it closes. Browsers send it, but the test harness can't catch it, so the 269 pages a click took elsewhere, or that loaded again while clicked, are left out of the click and analytics figures.
- Nothing reached Peekbird's servers: everything the snippet sent went to a stand-in inside the test browser. A site's Content-Security-Policy was set aside to measure the snippet, and counted on its own; a consent manager that holds scripts back until the site lists them did so on 0.4% of pages, also counted on its own.
- Logged-in pages, checkouts and sites behind a bot check (655 visits) weren't visited.

SECTION 2

Session replays

How much of the live page a replay draws the same: the share of pixels alike (a colour channel off by more than 40 of 255 counts as different), at each page's worst moment as it loaded and scrolled.

The same recording played three ways. **rrweb alone**: the open-source recorder Peekbird builds on, played by itself. **Free**: Peekbird's replay, loading pictures, fonts and stylesheets from the live site. **Growth and Scale**: from copies Peekbird keeps of them. Pages change by themselves too (slideshows, ads, feeds): two live visits of the same page, compared the same way, are a median 100% the same, and the last column takes those changes off.

	Pages	rrweb alone	Free	Growth and Scale	Growth and Scale: 99% or more	Within 1 point of the page's own changes
All pages	2,018	92.9%	98%	**99.5%**	58.9%	69.2%
On a computer	1,009	94.2%	98%	**99.5%**	59%	70.5%
On a phone	1,009	90.9%	97.7%	**99.5%**	58.8%	67.9%

Medians, but for the last two columns: the share of pages.

By popularity

	Pages	rrweb alone	Free	Growth and Scale	Growth and Scale: 99% or more	Within 1 point of the page's own changes
Tranco 1–1,000	636	92.2%	97.8%	**99.3%**	54.6%	65.3%
Tranco 1,001–10,000	914	92.6%	97.7%	**99.5%**	59.2%	69.3%
Tranco 10,001–100,000	468	95.4%	98.6%	**99.7%**	64.1%	74.4%

Pinch-zoomed on a phone

A moment zoomed in to three times, around a spot. Two live visits already differ here (a pinch never lands the same twice): they are a median 92.6% the same.

	Pages	rrweb alone	Free	Growth and Scale	Growth and Scale: 99% or more	Within 1 point of the page's own changes
Phone, zoomed in	1,009	63.9%	96.6%	**97.5%**	27.8%	72.4%

By what the site is built with

	Pages	rrweb alone	Free	Growth and Scale	Growth and Scale: 99% or more	Within 1 point of the page's own changes
jQuery	808	91.8%	97.7%	**99.4%**	57.1%	66.8%
React	515	89.9%	96.9%	**99.4%**	57.5%	68.7%

	Pages	rrweb alone	Free	Growth and Scale	Growth and Scale: 99% or more	Within 1 point of the page's own changes
Bootstrap	274	91.6%	97.6%	**99.3%**	53.6%	61.7%
WordPress	273	89.8%	98.4%	**99.6%**	65.9%	74%
Next.js	258	90.7%	96.8%	**99.5%**	61.6%	70.9%
Vue	121	85.5%	93.9%	**99.3%**	52.9%	64.5%
Drupal	68	92.7%	95.6%	**99.5%**	55.9%	69.1%
Webflow	52	96.1%	98.5%	**98.5%**	44.2%	63.5%
Nuxt	42	90.7%	99.1%	**99.7%**	57.1%	71.4%
Angular	32	83.1%	95.9%	**98.7%**	34.4%	68.8%
Astro	30	90.5%	93.6%	**98.9%**	46.7%	56.7%
Svelte	26	88%	96.5%	**98.8%**	46.2%	69.2%
Hugo	26	89.9%	94.6%	**97.1%**	46.2%	57.7%

Stacks found on 20 pages or more; a page can have several (React on Next.js, jQuery with WordPress).

By what pages hold

	Pages	Growth and Scale	99% or more	Pages without it	Growth and Scale	99% or more
Canvases	121	98.5%	46.3%	1,897	99.5%	59.7%
Other sites' frames	934	99%	50.1%	1,084	99.8%	66.4%
Ads	345	99.1%	51%	1,673	99.6%	60.5%
A consent banner	477	99.2%	53.2%	1,541	99.6%	60.6%
Web components	616	99.4%	56.7%	1,402	99.6%	59.8%
Script animations	50	98.4%	42%	1,968	99.5%	59.3%
Videos	444	98.5%	43.5%	1,574	99.7%	63.2%
Streamed video	51	93.3%	33.3%	1,967	99.5%	59.5%
The device's own font	361	99.6%	62.3%	1,657	99.5%	58.1%
Right-to-left text	18	100%	72.2%	2,000	99.5%	58.8%

Other sites' frames (ads, embedded players, chat widgets) can't be recorded from the page, by browser design: replays show them as blank space. Streamed video (HLS, DASH) plays as its poster.

What differed

On Growth and Scale replays, the kinds of difference found, by the share of pages that showed each:

Difference	Pages
Content moved	40.8%
An element another height	33.9%

Difference	Pages
Laid out at another width	31.6%
Something hidden in the replay	10.8%
A font not loaded	7.3%

A recording of a page (about 7.2 seconds, loading and scrolling) is a median 88.5 KB as sent, compressed (90th percentile 260.3 KB).

SECTION 3

A/B tests

2,013 pages ran a simulated experiment through Peekbird's real snippet, its variation forced, and were watched as they loaded.

Change	Pages with something to change	Applied	Applied at (median, from the page's start)	Original seen first	For (median)
Text replaced	84.3%	**96.9%**	748 ms	0.2%	199 ms
Style changed	75.9%	**93.2%**	965 ms	1.1%	180 ms
Element hidden	67.3%	**95.3%**	810 ms	0%	–
Picture replaced	83.3%	**96.2%**	810 ms	0.7%	17 ms
Link changed	91.8%	**96.2%**	816 ms	0.2%	439 ms
Element moved	37.7%	**93.1%**	911 ms	3%	292 ms
Element added	84.3%	**96.9%**	748 ms	0.2%	199 ms
Variation code (JavaScript)	100%	**99%**	656 ms	0%	–

"Original seen first": the original was on screen before the change applied, the flicker A/B tools hide the page to avoid. Applied: the change was on the page when looked for. Text replaced and element added are on the same heading, in one variation: they apply together.

Page held hidden	A median 125 ms while the variation applied (90th percentile 839 ms).
Exposures	Exactly one sent for 98.2% of visual tests, 98.2% of code tests and 99.8% of tests run in the site's own code. A page that loads again can send a second one; results count each visitor once.
Feature flags	A server SDK, given the cookies the browser would send it, gave the browser's answers on 97.6% of 2,013 pages: every page that kept Peekbird's visitor cookie. The flag was counted exactly once on 99.8%.
Split URL	1,850 pages: redirected on 98.3%, after a median 806 ms; the original page shown first on 1.5%; one exposure on 98.9%.
Errors	The snippet threw an error on 0% of pages.
Security policies	40.5% of sites send a Content-Security-Policy; 17.4% would block Peekbird's script and 13% its events until they list it.

SECTION 4

Heatmaps and analytics

1,700 pages clicked and tapped as a visitor would, 13,639 clicks in all, each drawn on the heatmap by the dashboard's own code, on its copy of the page from another visit's recording.

Clicks	Made	Reached Peekbird	Drawn on the element clicked	Within 10 px	Within 50 px	Element not on the copy	Not drawn
All clicks	13,639	**97.9%**	**80.2%**	66.4%	76.4%	5%	5.7%
Links and buttons	7,541	**98%**	**81.5%**	67.3%	76.7%	4.4%	6%
Other parts (pictures, headings, text)	6,098	**97.9%**	**78.7%**	65.2%	76%	5.7%	5.4%
Parts that stay on screen (sticky headers, chat buttons)	4,456	**97%**	**70.8%**	60.2%	67.5%	7.4%	14.6%
Parts not there as the page loaded (menus, pop-ups)	264	**97%**	**7%**	85.5%	90.2%	77.3%	2.7%
On a computer	6,920	**98.6%**	**81.7%**	68.8%	78.9%	5.3%	3.9%
On a phone (taps)	6,719	**97.2%**	**78.7%**	63.8%	73.8%	4.6%	7.6%

Element not on the copy: placed by its position instead. Not drawn: its element is on the copy but wasn't showing there (a menu that was open).

Counts and scrolling

- Clicks counted right on 97.5% of pages; link and button texts sent for 96.9% of those clicked.
- Rage clicks reported where none were made: 0% of pages.
- Scroll depth within 5 points of how far the visit went on 97.5% of 1,663 pages (median gap 0 points).
- Clicks a page's own script made (a slideshow pressing its own button) weren't counted as the visitor's: they happened on 3.4% of pages.

Web Vitals

	Pages	Peekbird sent one	Agree
LCP	1,684	98%	**100%**
CLS	1,694	97.6%	**100%**
INP	1,627	97.7%	**100%**

Against Google's web-vitals library in the same visit; agree: within 10 ms or 2%, whichever is more (LCP), 0.005 (CLS), 8 ms (INP).

Automatic events

Detected with no setup: purchases, add to cart, checkouts, discount codes, error messages, sign-ups and searches.

- On visits that bought, signed up and searched for nothing, events detected anyway (share of pages): an error message 0.5%, a search 0.3%, a checkout started 0.1%.
- 21.8% of pages show a search box; of 339 searches made in one, 90.9% were detected; a site can send the rest from its own code.
- Recordings sent on 98.7% of pages, a median 101.3 KB over the visit.

SECTION 5

Speed

How long Peekbird kept the page's main thread busy over a visit of about 25 seconds, with analytics, automatic events and a recording of the visit on.

Main thread over the visit	Median	90th percentile
Peekbird (snippet and modules)	**65 ms**	211 ms
The page's own scripts	310 ms	–
Peekbird's share of all script time	19.7%	–

A CPU profile of 1,969 visits, sampled every 0.25 ms.

SECTION 6

Safari

The same checks in WebKit, Safari's engine, on 186 of these websites, beside Chromium's figures on the very same sites.

WebKit as Playwright builds it (version 26.6), telling sites it is Safari on a Mac or on an iPhone, on 200 of the report's websites drawn at random; the 186 checked on both devices in both browsers give every figure here.

On the same sites, both devices	Safari (WebKit)	Chromium
Replays: Growth and Scale, median	**99.6%**	**99.6%**
Replays: Free, median	99.5%	98.3%
Replays: rrweb alone, median	97.2%	93%
Clicks and taps that reached Peekbird	**99.1%**	**98.5%**
Drawn on the element clicked	82.4%	79.4%
LCP sent, and agreeing with web-vitals	99%, 100%	98.4%, 100%
INP sent, and agreeing with web-vitals	99.3%, 100%	98.3%, 100%
CLS	not measured by the browser	98.4%, 100%
Scroll depth within 5 points	97.4%	97.4%
A/B changes applied, by kind	91.3% to 98.4%	93.4% to 98.9%
Original seen first, at most	3%	2.8%
Feature flags as a server SDK answers	97.8%	97.3%
Split URL redirected	99.7%	98.2%
Recordings sent	99.3%	99.4%
Searches detected	89.7%	88.4%

CLS: Safari doesn’t measure layout shifts, so no tool can report it for Safari’s visitors. The time a page is held hidden isn’t given for Safari: WebKit doesn’t draw a hidden page, nor run the frames the check measures with until it draws; whether the original was seen first is measured, and is given.

- **Fixed from it.** The recorder read the document of other sites' frames and of sandboxed ones (ads, players, captchas), which have none to give; Safari logged each try as an error naming Peekbird's recorder, on pages with an ad, a player or a captcha. It no longer reads them.
- **Security policies.** On 1.6% of pages the test couldn't set the site's own policy aside in WebKit as it does in Chromium (a few sites require Trusted Types), so Peekbird's recorder and detector weren't let in: as in any browser, until the site lists Peekbird.
- **What WebKit on a computer can't show.** Safari's limit on cookies set by a page's script: they last seven days there, so a visitor who comes back after a week counts as new, and can see another variation of a test. The iPhone's own address bar and touch, and Apple Pay.

SECTION 7

Statistics, simulated

Results are sequential: they stay right however often they're looked at, and a test ends at its planned sample. Checked by simulating 2,000 experiments a scenario, judged each day for 28 days by the results page's own code.

Traffic in each scenario is set so a fixed-horizon test's planned sample (to find a 10% change at 95% confidence) comes in after 14 days; Peekbird's sequential plan, a little larger, in a median 17. A wrong call is a winner or a loser where there's no difference, or the wrong one.

Scenario	Right call at the plan	Wrong call at the plan	Wrong call on some day, checked daily	Split alarm on some day
A/A, 3% conversion	97.4%	2.6%	4.3%	0.5%
A/A, 1% conversion	97.4%	2.6%	4.1%	0.5%
A/A, revenue per visitor	98.8%	1.2%	3.4%	0.6%
A/A, five secondary goals	97%	3%	4.6%	0.6%
+10% lift, 3% conversion	81.3%	0%	0%	1%
+20% lift, 3% conversion	99.8%	0%	0%	1%
−10%, 3% conversion	86%	0%	0%	1.1%
+10% order value, revenue per visitor	99.8%	0%	0%	1%
+10% lift, one of five secondary goals down 10%	24.5%	0%	0%	0.8%
+10% lift, a shop with many repeat buyers	79.2%	0%	0%	0.8%
A/A, the split itself 1% off	99.3%	0.7%	1.8%	70.7%

At 95% confidence. For A/A tests (no real difference) the right call is "no difference". The split alarm warns that visitors weren't divided as set: with the split itself 1% off, it's right to raise it.

Confidence levels, and checking a fixed-horizon test daily

	A/A tests called wrong on some day, checked daily	A real +10% found at the plan
Sequential, 90% confidence	7.8%	81.5%
Sequential, 95% (the default)	**4.3%**	81.3%
Sequential, 99%	0.7%	82.1%
Fixed-horizon at 95%, called whenever significant	22.7%	81.5%

The last row is what checking an ordinary test every day and stopping at the first significant result does.

SECTION 8

What this run changed in Peekbird

Measuring the product on real sites found things no test of its own had. Each was fixed for every site, with a test that failed first, before the figures here were taken.

- Visual changes now win over a site's own `!important` styles.
- The snippet starts on pages that can't keep cookies (sandboxed pages).
- INP is reported for visits whose interactions were all quick, as Google's library does.
- The recorder starts whole on pages that replace the browser's `Event` or `Response`, and never throws an error into the page.
- Scroll depth is read as visitors scroll, not as the page loads, and on app-like pages from the box their content scrolls in.
- Search detection finds searches known by what their box says, in the main languages, searches with no form, and terms in the address.
- Clicks a page's own script makes aren't counted as the visitor's; one standing in for a tap still is.
- Automatic events run on pages that make `fetch` read-only.
- Error messages are detected only where they are ones: not a page title read out to screen readers, a player's status or a price shown in red.
- The visual editor never saves a change by an id the page makes up on each load.
- In Safari, the recorder no longer reads other sites' frames or sandboxed ones, which Safari logged as an error naming it.

Reproducing it

Sites	Tranco list Y83KG (3 October 2026), drawn by band as above, seed 1; adult sites left out by word list and two public adult-domain lists.
Devices	Microsoft Edge (Chromium) through Playwright: 1280 × 800; phone 390 × 844, touch, 3× pixels. Home page, first visit, English.
Dates	4 and 5 October 2026.
Peekbird	Version 423b009: analytics, automatic events and recordings on (every session recorded); A/B tests in visits of their own.
Questions	peekbird.com