Websites for the AI Era: How Discovery, Agents and Generative Interfaces Are Reshaping the Web

Oct 31, 2026 · @Logesh · Sysbot Technologies R&D

Research Study v1.1, 31 Oct 2026; v1.0 was frozen on 30 Oct 2026 and the changes are listed in the version history. Publisher: Sysbot Technologies, Research & Development. Author: Logesh Chandran. Classification: public research publication. This is an independent, evidence-informed industry research study and strategic technology assessment: not peer-reviewed academic research and not legal advice. Vendor-reported figures are identified as such throughout, every figure links inline to the page it was taken from, and every sourced claim is graded and re-checked in the companion evidence register.

Executive summary

A website in the AI era is a business interface served from one content store to several kinds of client: people who arrive already informed by an answer engine, search engines, AI assistants and software agents acting on a person's behalf. It is not becoming a backend for agents; it is becoming the place where a business's information and operations are exposed to all of them. Sysbot should build toward that dual-audience architecture now, because every one of the five shifts below is measurable today; the rules are acquiring binding deadlines, while the standards and protocols governing implementation are still evolving.

The five shifts.

1. **Discovery moves from results to answers.** AI Overviews reach over 2.5 billion people a month; by Similarweb's US clickstream measure 68% of Google searches end without a click to a site outside Google; and publishers in Chartbeat's network lost 40% of search traffic in a year. AI assistants send back a tiny fraction of that, but in the most recent measurements the visitors they send converted better. Being cited increasingly matters alongside traditional search rankings, and most citations go to third-party platforms such as Reddit and YouTube rather than brand sites.
2. **Agents become visitors.** By Cloudflare's and Imperva's counts, which include every automated request from search crawlers to malicious bots, more than half of web traffic is non-human, and user-directed agent requests on Cloudflare's network grew seventeenfold in a year. The standalone AI browser category consolidated in 2026, with agency moving into ChatGPT, Gemini, Claude and Copilot, and a set of complementary

protocols and mechanisms is emerging: MCP for tools, WebMCP in the page, signed requests for identity and commerce manifests for checkout. Many agents rely on semantic page structure and accessibility information, some also on screenshots, and they commonly hand off to a person for authentication, payment and security challenges.

3. **Interfaces become generative, inside a fixed catalog.** Leading production-oriented implementations, from Google's A2UI to Vercel's json-render to Apps in ChatGPT and MCP Apps, have the model compose from developer-defined components. Free-form page generation remains an experiment with minute-scale latency. Users want chat for comparison and research and a human for problems, and they do not find chat features unprompted.
4. **Building becomes a prompt.** Lovable reported a run rate above $600 million in September 2026 and Replit is estimated near $525 million. An independent assessment found database access-control misconfigurations in about 10% of 1,645 Lovable apps examined. Simple-site prices have collapsed; value has moved to the data model, integrations, security and operation.
5. **Trust, provenance and the rules arrive before the standards.** Crawler control is still voluntary and unevenly honoured. Bot identity is an IETF draft that Cloudflare already enforces. EU AI Act transparency duties apply since August 2026, EU accessibility requirements for covered products and services have applied since June 2025, and the next year brings further deadlines in the UK, California, the US public sector and India. Accessibility is the investment most likely to serve both the law and the agents, because accessible semantic interfaces are a common foundation for assistive technology and many browser agents, though WCAG compliance does not guarantee successful agent interaction.

What Sysbot should do. Build and test a reference architecture in which one structured content store feeds a fast, accessible, semantically marked human site; an agent surface of Schema.org data, an MCP server and WebMCP tools; a catalog-constrained generative layer for conversation and composition; and an edge policy layer that distinguishes search, training and agent traffic, verifies signed agents and logs violations. Position the result as AI-ready website infrastructure rather than another AI website builder; validate four candidate product opportunities separately, an AI-ready website, an agent gateway, AI website intelligence and agent access economics; and write a technical reference architecture before building the prototypes. Build the content store, the accessibility foundation and the agent gateway before investing heavily in generative interfaces.

What this study does not claim. Vendor figures for resolution rates, personalisation lift and agent task success are reported as claims, not findings. Whether AI referrals will ever replace lost search clicks, whether agentic checkout will reach volume, and whether any

AI company will honour licensing signals are open questions, listed with the bets Sysbot can place on them.

Why this study

Sysbot Technologies is starting an R&D initiative on websites for the AI era, and this study is its first deliverable: a written map of where websites are heading, so that the prototypes that follow test real shifts rather than guesses.

Scope. Public-facing websites of businesses and organisations: marketing sites, content sites, commerce sites and the service portals behind them. Internal tools and native apps are out of scope except where they set a pattern websites will follow.

Questions this study answers.

1. Who, or what, visits a website in 2026, and how is that mix changing?
2. How do people and AI systems find a website now that answers arrive before links?
3. What must a website expose so that AI agents can read it, act on it and transact through it?
4. Which parts of the visible interface are becoming generated on demand, and which parts stay fixed?
5. What rules, standards and trust mechanisms are forming around all of this, and which have deadlines?
6. What should Sysbot build first, and how will it know whether the bet paid off?

Method. Desk research over primary sources dated 2025 to October 2026: platform announcements, standards drafts, measured traffic studies and regulatory texts. Every figure in this document links to the page it was taken from. Where sources disagree, the disagreement is recorded as an open question rather than resolved by opinion. Every sourced claim carries an evidence grade, E1 to E5, in the evidence register, which also records when each source was last checked; the Methodology section explains the grades, the search strategy and the review process.

How to read it. The executive summary carries the argument. Each of the five shifts has its own section with evidence, implications and what it means for Sysbot. The reference architecture and prototype roadmap at the end turn the findings into a build plan for the next phase. Three kinds of statement appear and are kept apart: observed evidence, which is the linked and graded material in the five Shift sections; research interpretation, which is the position this study takes where the evidence conflicts; and strategic hypothesis, which is every bet, the strategic position and the prototype roadmap that exists to test them.

Shift 1: Discovery moves from search results to answer engines

The open web is losing clicks faster than AI assistants are sending them back, and the clicks that do arrive are worth more. For most websites the realistic plan is fewer, better visits, earned by being cited as well as ranked. Three things are changing at once and should not be conflated: organic clicks are falling, zero-click searches are growing, and consumption of AI-generated answers is growing. A monthly user of AI Overviews is not a displaced visit, and a shrinking share of informational clicks does not make websites irrelevant; they remain the destination for transactions, trust and first-party information.

Answer engines are now the default search experience. Google said at I/O in May 2026 that AI Overviews reach more than 2.5 billion monthly users and AI Mode passed 1 billion. Overviews appear on roughly a fifth of searches, 18% in Pew's March 2025 sample and over 20% per Similarweb in June 2026. AI Mode itself remains a small share of searches, about 0.34% in Similarweb's January to April 2026 clickstream, but it refers out on only 1.6 to 2.5% of queries against 17 to 19% for classic results.

The measured click loss.

Study	Period	Finding
Pew Research	Mar 2025, 68,879 searches	Users clicked a result on 8% of visits with an AI summary vs 15% without; clicked a link inside the summary on 1%
Ahrefs	Dec 2023 vs Dec 2025, 300k keywords	Position-1 desktop CTR on Overview keywords fell from 7.3% to 1.6%; about 58% of the drop attributable to Overviews, the control group itself halved
Seer Interactive	Jan 2025 to Feb 2026, 53 brands	Organic CTR on Overview queries fell from 3.19% to 1.31%, then recovered to 2.36%; being cited in the Overview yielded about 2.2x the clicks per impression of not being cited
Similarweb / SparkToro	Jun 2026	About two-thirds of Google searches (68% in this US panel clickstream) may end without an external website click, up from about 60% in 2024; the rate varies by dataset, market, device and the definition of an external click

Study	Period	Finding
Chartbeat	Jul 2025 to Jul 2026	Google search traffic to publishers in Chartbeat's network down 40.2%; search share of pageviews fell from about 9% to 5% in two years; small publishers down 60% over two years vs 22% for large. This applies to Chartbeat's publisher network and should not be extrapolated to e-commerce, corporate or small-business sites

AI assistants send little traffic, but it converts well. Chartbeat puts AI referrals at 0.01% of publisher traffic despite eightfold growth in two years. ChatGPT accounts for about 80% of chatbot referrals in Statcounter's September 2026 worldwide measure, with Gemini, Perplexity, Copilot and Claude sharing the rest. The quality signals point the other way: Similarweb measured ChatGPT referrals converting at 7% on transactional sites against 5% for Google, and Adobe Analytics found US retail AI-referred traffic in March 2026 converting 42% better than other traffic, a retail-specific, single-month figure on a small and fast-growing base, after converting 38% worse a year earlier; it says nothing about SaaS, services or small-business sites. Attribution is unreliable: ChatGPT links lost their tracking tag for a week in October 2026 and much AI traffic lands as direct.

What gets cited. Google's official guidance is that AI features need no new markup or files, only indexed, snippet-eligible pages. The empirical picture is more specific:

- Citations have decoupled from rankings. Ahrefs found the share of Overview citations drawn from the top 10 organic results fell from 76% to 38% between July 2025 and March 2026, with about 31% coming from outside the top 100.
- Third-party platforms dominate. Across 30 million citations, Reddit is the first or second most cited domain on all five major assistants; one analysis of 27 million citations reported brand-owned sites earning about 4%, a figure taken from a secondary roundup, not verified against the original dataset, and not relied on in this study's conclusions.
- Content shape matters. The Princeton generative engine optimisation work showed that adding statistics and quotations raised visibility by up to 40%, and a SIGIR 2026 study tests 18 such factors across six models. Cited pages are on average 26% fresher than organic results.
- llms.txt shows no measurable benefit yet, a provisional finding. Adoption is about 10% of 300k domains with no measurable link to citations; 97% of files received zero AI-crawler requests in a month; Google's John Mueller has said no AI system currently

uses it.

How publishers and businesses are responding. Cloudflare, which fronts over a fifth of web domains, blocks training and agent crawlers by default on ad-bearing pages for new domains since 15 Sep 2026, and is moving Pay Per Crawl toward pay per answer. Microsoft opened a Publisher Content Marketplace in February 2026 and Perplexity pays publishers an 80/20 share from Comet Plus. Antitrust suits by Penske Media and Chegg over Overviews were dismissed on 30 Sep 2026. The UK Competition and Markets Authority imposed a publisher conduct requirement on Google on 3 Jun 2026; the CMA's Notice states that it comes into force six months after that date, 3 Dec 2026, except the page-level controls, which come into force at nine months, 3 Mar 2027. Google is responding with a Search Console toggle to opt a site out of AI Overviews and AI Mode, UK first. Seven EU and Canadian publisher groups asked EU and Canadian leaders for legislation giving publishers tools to authorise, monitor and enforce the use of their content, in October 2026. Among small businesses, 33.9% noticed an AI-related traffic drop and 14.8% have seen an assistant state something incorrect about their business, yet only 13% are actively optimising for AI search.

Implications for Sysbot.

1. Design pages to close, not to accumulate pageviews. Informational traffic is shrinking; the visitors who arrive via assistants are closer to a decision.
2. Treat citation as the new ranking. Front-load quotable, dated, statistic-rich answers, keep key pages fresh and publish first-party data others will cite.
3. Manage the entity off-site. Reddit, YouTube, Wikipedia and review platforms carry most citations, and assistants already misdescribe one business in seven.
4. Do not bet on llms.txt. The real control surface is robots.txt categories, Cloudflare-style crawl policies, Search Console opt-outs and licensing rails.
5. Build measurement. AI referrals are undercounted and lumped into "Web" in Search Console; citation tracking and attribution for small businesses is an open product gap. Test AI-referral conversion separately by sector, since the only strong evidence is US retail.

Shift 2: Agents become visitors

More than half of web traffic is now automated, by counts that include search crawlers, monitoring and malicious bots alongside AI, and a growing slice of it is agents acting for a specific person rather than crawlers gathering training data. The standalone AI browser category consolidated within a year, with agency moving into the assistants themselves,

and the protocols for sites to serve agents are converging on semantic HTML, typed tools and signed requests. Protocol availability is not customer adoption: the infrastructure figures in this section measure developer uptake, not end users or transactions.

Browser agents: a consolidation year.

- OpenAI's Atlas browser stopped working on 9 Aug 2026; OpenAI's help centre now records Atlas as retired and says the older ChatGPT agent is no longer available, pointing users to the cloud browser inside ChatGPT. The earlier version of that article described a built-in and a cloud browser in the ChatGPT desktop app; the page has since been shortened.
- Google's Project Mariner ended on 4 May 2026, its technology folded into other Google products; the consumer browsing agent is Chrome auto browse, a US preview for paid tiers that pauses before purchases and posts and can sign in through Google Password Manager.
- Microsoft retired Copilot Mode on 13 May 2026 as Browse with Copilot arrived in Edge, US-only for Microsoft 365 Premium subscribers at launch; the business version, agentic browsing with Copilot, is a limited preview scoped by IT to approved sites.
- Anthropic's Claude in Chrome reached general availability on 26 Aug 2026 on paid plans, with a per-action safety classifier; Anthropic's own red-team figures put prompt-injection success at 23.6% in the 2025 pilot and 0 to 0.3% with current safeguards, a vendor evaluation rather than a general security guarantee.
- Perplexity's Comet has been free worldwide since Oct 2025. Trail of Bits showed four classes of injection, including fake CAPTCHAs, that exfiltrated Gmail when a user asked it to summarise an attacker's page, and Brave showed that near-invisible text in screenshots is executed as instructions.

The protocols a website can implement.

Protocol	Sponsor	Status, Oct 2026	What a site must do
MCP	Linux Foundation Agentic AI Foundation	Spec of 28 Jul 2026 made it stateless; SDKs near half a billion downloads a month, per the MCP maintainers; downloads measure developer uptake, not active servers, agents or transactions	Host an MCP server over HTTP exposing tools and resources, OAuth with client ID metadata documents

Protocol	Sponsor	Status, Oct 2026	What a site must do
WebMCP	Google and Microsoft, W3C community group	Chrome 149 and Edge 150 origin trials; Firefox and Safari have standards-position requests open and no stated position	Register typed tools in the page through an imperative API, or annotate forms declaratively; human-in-the-loop only, not headless
NLWeb	Microsoft, MIT licence	Reference implementation; launch partners Tripadvisor, Shopify, Eventbrite; adoption flat	Schema.org or RSS feeds, an ask endpoint that doubles as an MCP server
Web Bot Auth	IETF working group, Cloudflare and Google	IETF Internet-Draft 01 dated 29 Oct 2026, standards-track but not yet an RFC; OpenAI, Google and AWS sign agent requests	Verify signatures, or let the CDN do it; agents publish keys
Agentic Commerce Protocol	OpenAI and Stripe	Spec of 30 Jan 2026; in-chat Instant Checkout deprioritised in Mar 2026 for discovery in chat, purchase on site	A product feed; optionally five checkout REST endpoints
Universal Commerce Protocol	Google with Shopify, Etsy, Walmart, Target, Wayfair	Announced Jan 2026; merchants join a waitlist	A manifest at a well-known path, checkout over REST, MCP or A2A, a Merchant Center account
Agent Payments Protocol	Google, donated to the FIDO Alliance Apr 2026	v0.2 with human-not-present payments; no volumes	Accept mandate-backed payments through a payment provider
Markdown for Agents	Cloudflare	Live; one example cut 16k HTML tokens to 3k	Serve Markdown on an Accept header, preserving JSON-LD

Protocol	Sponsor	Status, Oct 2026	What a site must do
Search / Training / Agent controls	Cloudflare	Default since 15 Sep 2026; Pay Per Use in closed beta for US customers	Set a policy per category; optional pricing over HTTP 402. These classes are one vendor's classification, not an internet-wide one, and a bot can change behaviour

Agentic commerce is protocols ahead of adoption. Every network has a protocol: Visa's Intelligent Commerce Connect accepts four of them in one integration, and Mastercard's Agent Pay Trust Framework adds know-your-agent verification. The only merchant count found is Forrester's estimate, via secondary reporting, of about 30 live Shopify stores on Instant Checkout by February 2026, after which OpenAI itself retreated to discovery in chat and checkout on the merchant's site.

The economics of being crawled. Cloudflare reports that daily AI-agent requests grew over 1,700% year on year, that training crawlers' share of crawler requests rose from 22% in spring 2025 to 52% in June 2026, and that human traffic to heavily crawled categories fell by up to 40%. Thales and Imperva put bots at 53% of web traffic in 2025, with legitimate AI bots about 2% of sessions, split 85% crawlers to 15% user-directed fetches; the 53% is in the public press release, the breakdown is from the gated report and was not re-verified. Crawl-to-referral ratios differ by orders of magnitude between operators: Google near 5 to 1, OpenAI about 250 to 1, Anthropic about 1,900 to 1 in July 2026, by Cloudflare Radar's method, which excludes referrals from native apps and so overstates all three.

How agents actually read a site. Browser agents commonly work from semantic page structure and accessibility information, for example the accessibility tree with element references that Browserbase's agent browser exposes to the model, while some also act from screenshots. Agents commonly require human confirmation or hand-off for sensitive actions such as authentication, payment and certain security challenges, and the exact behaviour differs by product and task. A site is agent-hostile when it renders only through JavaScript without semantic HTML, leaves form controls unlabelled, walls itself behind CAPTCHAs or blocks all bots, which now also catches user-directed fetches. It is agent-friendly when it has semantic HTML and ARIA, Schema.org in JSON-LD, Markdown negotiation, stable URLs and explicit action surfaces.

Verified is not authorised. Web Bot Auth and signed agents establish who an agent is, not what it may do. A verified agent may still be unauthorised to read customer records, change an order or make a payment. Authorisation should be scoped by business operation, with consent captured separately, and the architecture needs authentication, authorisation and consent as three distinct layers. Recognising a signed agent is the least important of the three.

Agent operation	Permission
Read product catalogue	Allowed
Check stock	Allowed
Request quotation	Allowed with validation
Create purchase order	Requires approval
Modify customer account	Restricted
Transfer money	Explicit authorisation by the account holder, never by identity alone

Implications for Sysbot.

1. Treat user-directed agent traffic as a distinct audience from training crawlers: allow the fetches that serve a person, govern training separately, and never blanket-block.
2. Semantic HTML, labelled forms and Schema.org are the highest-leverage, lowest-risk investment. Most agent paths depend on them.
3. Expose structured information and reliable business operations first, through APIs and an MCP server; that is the immediate opportunity, not agentic checkout. Add browser-side WebMCP tools only when a customer needs them.
4. Design the human hand-off. Agents stop at login, payment and CAPTCHA, so make those steps short, resumable and clearly labelled, and plan for signed-agent verification instead of CAPTCHAs.
5. Measure agent sessions separately, because referral economics differ by operator by orders of magnitude.

Shift 3: Interfaces become generative, but inside a fixed catalog

Leading production-oriented generative UI implementations are constrained composition: the model picks and fills components from a developer-defined catalog, it does not write free-form HTML. Free-form generation exists only as a gated experiment with minute-scale latency.

What ships, as of October 2026.

Offering	Sponsor	Launched	Model of generation	Status
Gemini dynamic view / visual layout	Google	Nov 2025	Free-form page generation	Experiment, US, paid tiers, 18+, not in mobile app; "a minute or more" per page
A2UI v0.9	Google	Apr 2026	Declarative JSON intent rendered from a component catalog	Open spec; React, Flutter, Angular, Lit renderers; adopted by CopilotKit, AG2, Oracle Agent Spec
json-render	Vercel Labs	Jan 2026	JSON limited to a Zod-typed catalog	Open source, 13k+ stars by Mar 2026; the older AI SDK RSC streamUI path is marked experimental
Apps in ChatGPT	OpenAI	Oct 2025	MCP tools returning app UI inside ChatGPT	Directory open to submissions; production checkout enabled by OpenAI after sandbox testing and compliance attestation
MCP Apps	Anthropic / MCP	Jan 2026	Tools return UI rendered in a sandboxed iframe	Live in Claude on all plans; supported by ChatGPT, Goose, VS Code Insiders
Thesys C1, now OpenUI	Thesys	2025	Hosted endpoint streaming a declarative UI spec to a React SDK	Commercial; no usage disclosed

Google's own preference study ranked human-expert pages first and generative pages "close behind", both far above raw text, but the study ignored speed. Vercel built json-render specifically to stop a model emitting arbitrary or malicious code. The convergence on catalogs is the industry answer to brand, trust and security at once.

Chat-first sites: the evidence is thin and the critiques are strong. No independently verified case exists of a business replacing navigation with conversation. The closest, Agent Image's SiteAI for real-estate sites (Oct 2026), layers answers over a conventional crawlable site. Where a conversational layer is measured, the numbers come from the vendor:

- Shopify Sidekick, a merchant-facing assistant, handled about 34 million conversations in fiscal Q2 2026 (Apr to Jun 2026), with daily sessions up 4.8x year on year, per the earnings call.
- Intercom claims a 76% average resolution rate for its Fin agent, under a billing definition that counts a conversation as resolved when the customer confirms the answer or leaves without asking for more; no independent measurement of Fin's resolution rate was found. A cross-vendor benchmark puts median AI self-service deflection near 22%.
- Nielsen Norman Group found that none of its participants discovered Amazon's Rufus chat unprompted, and that most AI features tested lacked discoverability. Some participants in NN/g's Google AI Mode study did not realise follow-up questions were possible. NN/g's articulation barrier argument, made first for image generators, is that many users cannot phrase what they want, and it proposes hybrid GUI plus prompt as the remedy.

What users say stays fixed. The figures below are stated preferences from a survey, not observed behaviour, and should not decide an interface without a prototype. Baymard's October 2026 study of 2,273 US shoppers found a clear split by task: 63% prefer AI chat for feature comparison and 59% for stock checks, but 66 to 69% want a human for damaged, late or billing problems, and 60% say a product image they know is AI-generated lowers purchase confidence. NN/g's October 2026 survey of 521 practitioners scored AI time savings at 4.2 of 5 but sentiment at 3.57, with 69% of comments negative, mostly about unreliable output. Claimed personalisation lifts of 10 to 30% trace back to vendor and consultancy reports; no 2026 peer-reviewed trial of adaptive landing pages was found.

Cost and latency are design constraints, not afterthoughts. A modelled retail generative-UI session of three turns and three tool calls costs about $0.08 per session, with tool calls near 40% of the bill; prompt caching of a 2.5k-token prefix cuts it by roughly 38% but has to be switched on deliberately. An April 2026 paper on generative UI accessibility recommends specifying layout policies rather than layouts, because full-page regeneration is too slow without caching or smaller models.

Three layers, not one. The generative layer does not replace the conventional interface; it sits beside it, and the conventional interface remains the reliable fallback.

Layer	Responsibility
Traditional interface	Navigation, forms, dashboards, content; the fallback when generation is slow, costly or wrong
Conversational interface	Discovery, comparison, guidance; hybrid with suggested prompts and actions
Generative interface	Dynamic composition from approved components; the model selects and arranges, never writes production code

A request such as "the best manufacturing ERP package for a garment factory with 200 employees" is answered not by generating a new page but by composing a product comparison, a pricing component, a feature matrix, a customer case study and a request-demonstration action from the catalog.

Personalised composition is the larger opportunity. The same catalog can present one platform differently by industry or role without a new site. A garment manufacturer visiting an ERP site would see production planning, fabric inventory, cutting and sewing; a furniture maker would see timber inventory, cutting, assembly and finishing. The platform and the store are the same; only the composition changes. This is directly relevant to Sysbot's OnFleek ERP direction, and it is under-researched: no study in this review measured contextual composition against a fixed layout, so it belongs in the prototype roadmap rather than in the findings.

Implications for Sysbot.

1. Build generative UI as "model composes from our catalog" and make it a core R&D investment. The catalog pins brand components, trust signals and actions; the model only arranges them; the conventional interface remains the fallback.
2. Keep the scannable page. Add the conversational layer as a visible hybrid with suggested prompts and actions, because users do not find chat on their own.
3. Route by task: comparison and research to AI, exceptions and complaints to humans with a visible escalation path.
4. Treat per-session model cost and latency as product metrics from day one, with cached and static fallbacks.
5. Distrust vendor resolution and lift figures. Instrument holdout tests in every prototype.

Shift 4: Building a website becomes a prompt

Prompt-built sites have commoditised the template, copy and layout of a website in under two years; how large a share of a build that represents is an editorial judgement, not a measured figure, and it varies by project. The defensible work has moved to the data model, integrations, security review and maintenance, and the price of simple sites has collapsed while complex builds hold.

The builders, by the numbers. All figures are company claims unless stated.

Builder	Latest disclosed figure	Date	Note
Lovable	Over $600M annualised revenue; 1M new projects a week at $500M in June	Sep 2026	From $100M in Jul 2025; about 146 staff
Replit	About $525M annualised (Sacra estimate); Agent 4 launched with a $400M round at $9B	Mar to Apr 2026	Targets $1B by end of 2026
Vercel v0	Over 4M people have used v0; Teams and Enterprise over half of v0 revenue	Sep 2025	Vercel run-rate over $500M by Jul 2026; no standalone v0 figure
Wix	Q2 2026 revenue $563M, up 15%; Harmony builder on a proprietary model since Jan 2026	Aug 2026	Base44, bought for $80M in Jun 2025, reported at $150M ARR by May 2026
Bolt.new	About $40M ARR in its first five months	2025	No 2026 figure found
Shopify Sidekick, Squarespace Blueprint, Framer, Webflow AI, Hostinger Horizons, Durable	AI site building inside both Shopify (Sidekick) and Squarespace (Blueprint); Squarespace added an AI Visibility tool in Jul 2026	2026	No user or revenue disclosures found for the rest

Quality and security are where the cracks show. A researcher's 2025 assessment of 1,645 publicly discoverable Lovable applications identified 170 (10.3%) with database access-control misconfigurations involving Supabase row-level security, catalogued as CVE-2025-48757. The assessment did not establish which were exploited, active or later

fixed, so this is a misconfiguration rate, not a count of breaches. 2026 guides from security reviewers still list missing row-level security as the most common critical issue in Lovable apps. A December 2025 study, Vibe Coding in Practice, documents a trade-off between flow and debt: architectural inconsistency, security gaps and maintenance overhead. A 2026 paper on spec-driven development argues from the literature that prompt-built output is unreliable and hard to maintain. Widely repeated claims of 30 to 41% more technical debt have no verifiable primary source and are not relied on here.

What it does to agencies and freelancers. A GoodFirms survey found 98% of agencies using AI, with a third pricing AI-built sites at $500 to $1,500 and a quarter under $500, while complex builds held or raised prices. A Fast Company and AIGA survey of 1,273 designers and 150 hiring managers found only 14% of managers expecting freelance design needs to shrink, though two-thirds of designers are worried. Upwork's 2026 in-demand skills report shows demand for AI-referencing skills up 109% year on year with full-stack and graphic design holding. Agencies describe passing the efficiency on as faster and cheaper delivery of professional work.

The website as the agent's backend. A growing camp argues the site becomes a content store plus typed tools, with the interface generated on demand. The infrastructure for that view is real: Microsoft's NLWeb makes every instance an MCP server over Schema.org and RSS, with a one-click Cloudflare deploy; WebMCP lets a page register typed tools for a browser agent; A2UI and MCP Apps supply the on-demand UI half. The counter-arguments are also real: agentic browsers are still too small to move traffic, the agent protocols are described as very early even by their advocates, and every agent-first vendor still keeps a crawlable site for search. The realistic reading is dual-audience: a human site and an agent surface over one content store, not one replacing the other.

What customers now expect a website to do. The pressure on generic agencies is not only cheaper pages. Customers increasingly expect a site to perform business functions: generate quotations, capture and qualify leads, integrate with ERP, synchronise inventory, schedule appointments, process orders, provide support and connect to AI assistants. That is ground a website agency can move to and an AI builder cannot easily follow: business digital infrastructure rather than page production. The opportunity is not selling AI-generated websites; it is selling websites connected to actual business operations.

Implications for Sysbot.

1. Do not compete on producing pages. Move from website agency to business digital infrastructure provider: sites that generate quotations, qualify leads, integrate with ERP, synchronise inventory, schedule appointments, process orders, support customers and connect to AI assistants.
2. Security review of generated backends is a product opportunity in itself; about one in ten audited apps on a leading platform had database access-control misconfigurations.

3. Price on outcomes and speed, not pages.
4. Design every site dual-audience from the start: structured data and typed tools beside the human pages, with one content store behind both.
5. Keep export paths and own the content store. Credit-metered builders and proprietary models create lock-in.

Shift 5: Trust, provenance and the rules

The rules are arriving faster than the standards. Crawler control is still voluntary, bot identity is a draft, and content provenance is patchy, yet binding dates for accessibility, AI disclosure and data protection fall between now and April 2028. Accessibility is the one investment that serves the law and the agents at the same time, though evidence that it raises agent task success is still thin.

Crawler control: signals without enforcement. robots.txt is voluntary and, in the words of RFC 9309, not a form of access authorization. The IETF AI Preferences group's attach and vocabulary drafts remain working-group documents in Aug and Sep 2026, and its charter puts technical enforcement out of scope. Really Simple Licensing, launched in Sep 2025 with Reddit, Yahoo, Medium and O'Reilly, now has over 50 collective partners but no major AI company has adopted it. Cloudflare's Content Signals Policy ships three robots.txt signals on 3.8 million domains, framed as rights reservations under the EU Copyright Directive. Google's Search Console opt-out for AI Overviews and AI Mode, announced 2 Jun 2026 as a UK test, is a control that touches those features directly, because they ride on Googlebot rather than the Google-Extended token.

Compliance is uneven. Cloudflare documented Perplexity using undeclared user agents and rotating networks to evade no-crawl directives at 3 to 6 million requests a day, and Perplexity's own docs say its user-triggered fetcher generally ignores robots.txt. OpenAI states that for ChatGPT-User robots.txt rules may not apply. Anthropic honours robots.txt and Crawl-delay and publishes its IP ranges. Cloudflare now offers a robots.txt violations log.

Bot identity: a draft becoming a gate. The IETF Web Bot Auth group passed its April and August 2026 milestones without sending its drafts to the IESG, yet Cloudflare's signed agents class already verifies ChatGPT agent, Goose, Browserbase and Anchor by the draft, and Cloudflare has said Enterprise customers will be able to allow or block that class in security rules. Verified identity is not authorisation, and the vendors treat user-initiated fetchers as outside robots.txt.

Provenance and labelling have deadlines. EU AI Act Article 50 applies from 2 Aug 2026. The obligations fall on different parties: providers of AI systems must mark generated content in machine-readable form and make chatbots disclose themselves, while

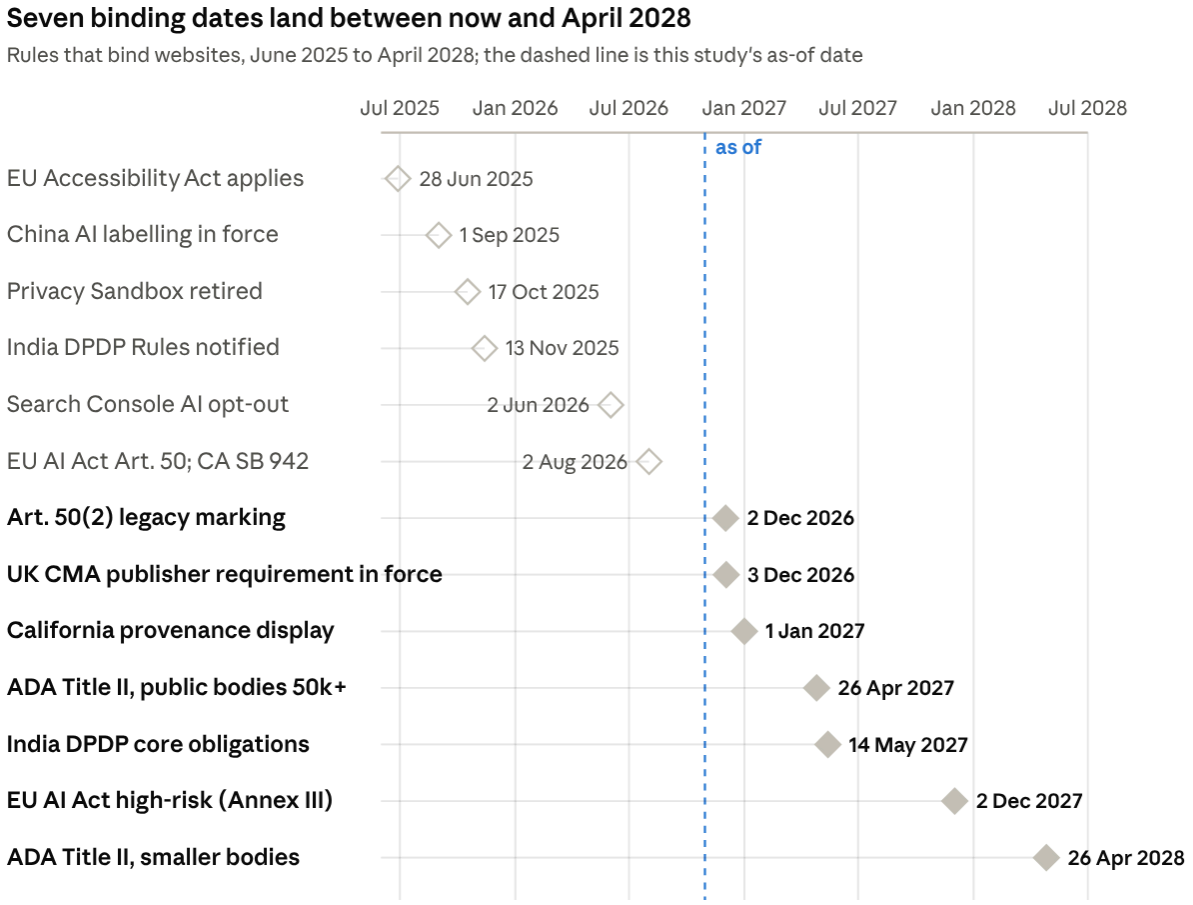
deployers, which includes a business running an AI assistant or publishing generated media on its site, must label deepfakes and AI-generated text on matters of public interest. One transitional provision applies: providers of systems that generate synthetic audio, image, video or text and were placed on the market before 2 Aug 2026 have until 2 Dec 2026 to meet the Article 50(2) marking obligation; no other Article 50 duty is extended, under the Digital Omnibus. China's labelling measures have been in force since 1 Sep 2025. California's AI Transparency Act became operative on 2 Aug 2026; it obliges covered AI providers to offer detection tools and latent disclosures, and a separate law requires large online platforms to display provenance data from 1 Jan 2027. C2PA Content Credentials are preserved by Cloudflare Images and applied across Adobe tools, and Google announced SynthID and C2PA verification in Search and Chrome in May 2026, but metadata is still stripped on upload by many platforms. Google's spam policy penalises scaled content abuse however it is created, and now names manipulation of generative AI responses as spam.

Accessibility is now law and interface at once. The European Accessibility Act has applied since 28 Jun 2025 to e-commerce, banking and transport, with German fines up to 100,000 euros; microenterprises are largely exempt. The US Department of Justice pushed ADA Title II web deadlines to 26 Apr 2027 and 2028; that rule binds state and local governments, not private businesses, whose exposure remains under Title III case law. WCAG 3.0 remains a working draft with years to go, so WCAG 2.2 AA is the benchmark. Research agents consume accessibility-tree snapshots that rely on WCAG-compliant markup; the widely quoted claim that agent success drops from 78% to 42% on inaccessible sites is a vendor figure without a locatable paper and is not relied on here.

Privacy: the cookie layer stays, consent fatigue grows. Google retired most Privacy Sandbox APIs in Oct 2025 and Chrome keeps third-party cookies. Only 5.4% of users configure cookie settings; the EU Digital Omnibus proposed a six-month no-re-ask rule, dropped from the Council's June 2026 compromise text and still in negotiation. India's Digital Personal Data Protection Rules were notified on 14 Nov 2025, per MeitY's rules and enforcement timeline; the accompanying commencement notification, a Gazette instrument dated 13 Nov 2025, brings the Act's provisions into force in phases from the publication date: specified Data Protection Board provisions took effect immediately; the consent-manager registration provisions take effect on 14 Nov 2026; and the core data-protection obligations, including notice, consent, breach and data-principal rights provisions, take effect on 14 May 2027. Which obligations apply to a given Sysbot site depends on the organisation's role, including whether it acts as a data fiduciary, data processor or consent manager.

Security: your page is an attack surface for someone else's agent. Brave showed hidden Reddit text making Comet read a user's email and leak a one-time code; OpenAI's Lockdown Mode concedes it cannot guarantee against exfiltration; EchoLeak was a zero-

click injection against Microsoft 365 Copilot. The OWASP Top 10 for LLM applications lists prompt injection, excessive agency and improper output handling. Site owners inherit two duties: sanitise user-generated hidden text they render, and treat their own chatbot as an injection target.



regulatory timeline · 12 dates, Jun 2025 to Dec 2027

The EU AI Act's transparency duties and California's transparency act are already in force; the next eighteen months bring the Article 50(2) marking deadline for older generative systems, the UK conduct requirement on Google, California's provenance display duty, the US public-sector web accessibility rule in two stages ending April 2028, and India's data-protection obligations.

Implications for Sysbot.

1. Ship a layered crawler policy as a product feature: per-bot robots.txt, Content Signals or RSL terms, the Search Console toggle, WAF rules and a violations log.
2. Verify agents cryptographically rather than blocking bots: allowlist signed agents, distinguish agent commerce from scraping.
3. Build to WCAG 2.2 AA as the engineering benchmark, while checking the specific accessibility standard and obligations applicable in each target jurisdiction. A clean accessibility tree supports assistive technology and many browser agents, though it

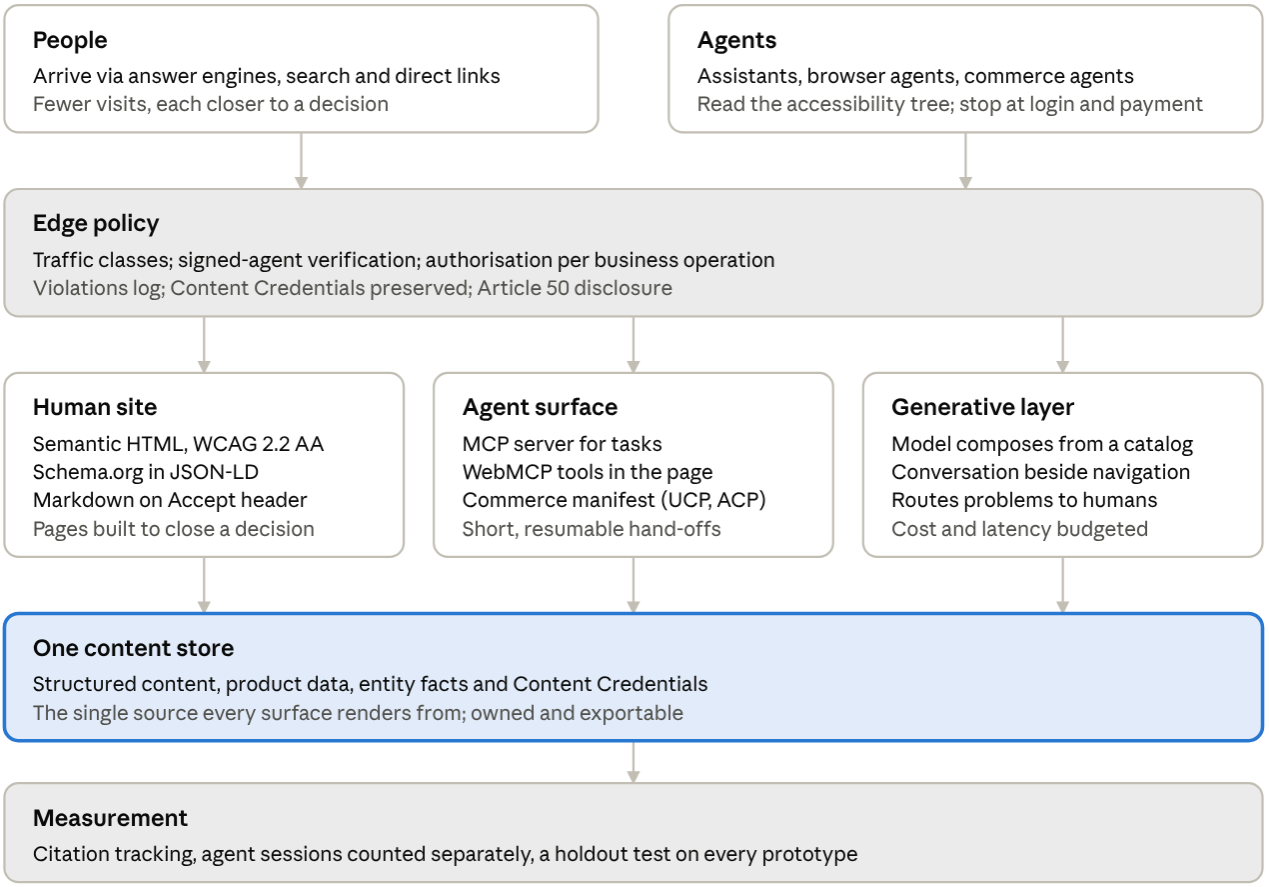
does not guarantee successful agent interaction.

4. Treat Article 50 as a product requirement: AI disclosure on chatbots, Content Credentials preserved through the image pipeline, visible labels on generated media.
5. Harden in both directions: sanitise rendered user content, and apply OWASP controls to any chatbot or agent the site runs.

The anatomy of an AI-era website

The architecture Sysbot should build toward has one owned content store underneath three surfaces, with a single edge policy deciding who and what gets through. The thesis is not that websites become backends for agents; it is that a website becomes a business interface consumed by several kinds of client: people, search engines, AI assistants and autonomous agents. Each component is grounded in an existing standard, platform capability or measured behaviour from the five shifts; the integrated architecture is Sysbot's proposed implementation of those capabilities, not an existing industry standard.

One content store, three surfaces, one edge policy



reference architecture · 2 audiences, 1 policy layer, 3 surfaces, 1 store

People and agents both pass the edge policy, which classifies traffic, verifies identity and then applies authorisation per business operation and consent before any surface sees a request. Authentication, authorisation and consent are three separate layers; a verified agent is not an authorised one. The human site, the agent surface and the generative layer all render from the same store, so a fact changes in one place and every surface, and every assistant that cites it, picks it up. Measurement sits under the store so that citations, agent sessions and human conversions are counted against the same content.

Why each part is there.

Part	Earned by	Standard or evidence it rests on
Edge policy	Shift 2, Shift 5	Cloudflare search / training / agent classes, Web Bot Auth signed agents, robots.txt violations log, EU AI Act Article 50
Human site	Shift 1, Shift 5	Citation studies favouring front-loaded, dated, structured content; WCAG 2.2 AA as the legal benchmark; accessible structure as a common foundation for many browser agents
Agent surface	Shift 2	MCP stateless spec of Jul 2026, WebMCP origin trials, UCP manifest and ACP product feeds; vendors stop at login and payment
Generative layer	Shift 3	Catalog-constrained composition in A2UI, json-render, Apps in ChatGPT and MCP Apps; Baymard's task split between AI and humans
One content store	Shift 4	Value moving to the data model; lock-in in credit-metered builders; NLWeb and commerce protocols reading Schema.org
Measurement	Shift 1, Shift 3	AI referrals undercounted and lumped into Web in Search Console; vendor resolution rates unreliable without holdouts

Sysbot must own four things outright: the canonical content and business data model, the contracts through which it is exposed, the permission model governing who may read and change customer information, and the export and portability mechanism. Infrastructure such as the CDN, hosting, search indexing, model providers, analytics storage and the MCP transport can begin as thin layers over existing services and be replaced as the standards settle; the business rules governing access to customer data are never outsourced.

Open questions and bets

The evidence is thin or contradictory in nine places. Each is written as a question, the position this study takes, and the bet Sysbot can place without waiting for the answer. The position is research interpretation and the bet is strategic hypothesis; neither is a finding, and the evidence register lists what would overturn each.

Question	Where the evidence conflicts	Position taken	Bet for Sysbot
How much of the click decline is AI Overviews, and how much secular Google change?	Ahrefs' control group halved without Overviews; Seer saw CTR fall on no-Overview queries too; Google rejects all three methods	Overviews explain roughly half; the rest is Google keeping users on Google (a hypothesis, not a measured split)	Plan for informational traffic to keep falling either way; build pages that close
Does AI-referred traffic convert better, or just differently?	Similarweb 7% vs 5%, Adobe +42% in 2026 after −38% in 2025; different baselines and windows	Better for now, partly a novelty and intent-mix effect	Instrument AI referrals separately before investing in them
Are citations decoupling from rankings or just noisy?	Ahrefs 76% to 38%, BrightEdge about 17%, earlier claims of 93% overlap	Decoupling is real but the series is unstable across model changes	Optimise content for citation; do not abandon ranking
Will free-form generative UI reach commerce-grade latency and accuracy?	Google ships it as an experiment at a minute per page; everyone else ships catalogs	Catalog composition is the production ceiling for at least two years	Build the catalog; keep free-form generation to internal experiments
Will agentic checkout produce merchant volume?	Every network has a protocol; the only hard number is about 30 live Shopify stores; OpenAI retreated to buy-on-site	Discover in AI, buy on site is the equilibrium through 2027	Publish product feeds and a manifest; defer in-chat checkout

Question	Where the evidence conflicts	Position taken	Bet for Sysbot
Will WebMCP reach Firefox, Safari and headless agents?	Chrome and Edge origin trials; Firefox and Safari have no stated position; headless is a stated non-goal	A Chrome and Edge in-page feature for the foreseeable future	Expose the same tools over MCP first, WebMCP second
Will any AI company honour licensing signals such as RSL or AI Preferences?	Over 50 publishers in the RSL Collective; zero AI companies; IETF drafts without enforcement	Licensing settles through litigation, bilateral deals and CDN tolls	Implement the signals cheaply; expect enforcement from the edge, not the protocol
Does Web Bot Auth become a gate before the IETF finishes it?	Cloudflare, Google and AWS already sign and verify; the working group has passed its milestones without delivering	Yes, enforced by CDNs	Allowlist signed agents; design for small unsigned agents to be treated as users
Will provenance survive metadata stripping, and will enforcement reach websites?	C2PA preserved by Cloudflare and Adobe, stripped by most social platforms; Article 50 enforcement untested	Patchy survival; first enforcement targets model providers, then large platforms	Preserve credentials through the pipeline; disclose AI on chatbots now

Research gaps

Seven areas deserve investigation that this study did not reach. Several are potentially more valuable to Sysbot than the statistics already reported.

Missing research area	Why it matters to Sysbot
Agent identity and permissions	Determines how agents can safely interact with business systems; authentication, authorisation and consent as separate layers

Missing research area	Why it matters to Sysbot
API-first website architecture	Makes business functionality accessible to agents without scraping
AI search attribution	Tells a business whether AI discovery produces leads and revenue, not just visits
Personalised UI composition	Turns catalog-constrained generative UI into a product rather than a feature
AI-generated application security	A potential recurring service, given the misconfiguration rates in prompt-built apps
Business data portability	Prevents customers being locked into AI website builders and credit-metered models
The economics of agent-readable websites	A person visits once; an agent may fetch product data, pricing, stock, documents and policies repeatedly, and the owner pays for compute, API calls, database queries, bandwidth, model usage and security while the agent may never send a customer back

The last gap is the one with no settled answer anywhere in the sources reviewed. Pay Per Crawl, pay per use and signed-agent allowlists are the first attempts to price machine access to a business's data and operations. How a business should monetise, ration or subsidise that access deserves its own research section before any agent surface is offered commercially. It is also a fourth candidate product, agent access economics: who accesses a business's information, how often, what resources it consumes, which agents generate business and which only consume. The first step is to measure actual agent consumption on real customer sites, not to build a monetisation platform.

Prototype roadmap

Six prototypes, each sized at two to four weeks of elapsed time for a team of two to three engineers, each testing one hypothesis from the study against a measurable outcome, all built against the technical reference architecture proposed below. They are ordered so that each one builds on the store and instrumentation the previous one leaves behind, and so that the cheapest, highest-confidence bets come first.

Order	Prototype	Hypothesis tested	What gets built	Success measure	Shift
1	Dual-audience reference site	A site served from one structured store, with semantic HTML, Schema.org JSON-LD and Markdown negotiation, is read correctly by both people and agents	A real reference implementation, not a toy: product catalogue, business information, pricing, FAQs, contact forms, structured data, semantic HTML, WCAG 2.2 AA, Markdown on Accept header, on Next.js over a typed store	At least three browser agents run five predefined business tasks ten times each, 150 executions, reporting task completion rate, incorrect-action rate, human-intervention rate and execution time; completion above 80% with no incorrect action that changes data; accessibility scan clean; Markdown response under a quarter of the HTML token count	2, 5
2	Agent and citation analytics	AI referrals, agent sessions and citations can be measured separately and tied to business outcomes	Edge logging by bot class, signed-agent verification, a dashboard of which assistants discover the site, which pages they retrieve, which agents act, which sessions produce leads and which leads become customers; a weekly citation probe across five assistants	Dashboard distinguishes search, training and agent traffic and attributes leads; probe runs unattended for four weeks	1, 2
3	Agent surface	Exposing a small set of typed business operations lets agents complete tasks deterministically instead of scraping	An MCP server over the store with exactly three tools, search_products, check_availability and request_quote, with per-operation authorisation; no payments and no unrestricted operations; a product feed and a UCP manifest	Task completion rate and steps per task for agents using tools versus the HTML path, on the same 150-execution protocol as prototype 1; human hand-off completes in one screen	2

Order	Prototype	Hypothesis tested	What gets built	Success measure	Shift
4	Catalog-constrained generative UI	A conversational layer composing from a small approved catalog improves task completion over conventional navigation without hurting trust	json-render or A2UI over a limited catalog from prototype 1, suggested prompts, visible escalation to a human, prompt caching, a per-session cost budget; one industry-specific composition variant	Holdout test on comparison tasks against navigation; cost per session under ceiling; median turn under five seconds; fallback suite green	3
5	Citation-shaped content	Front-loaded, dated, statistic-rich pages earn more assistant citations than the same facts written conventionally	Matched content pairs: ten pages, each in two versions with equivalent facts and controlled differences in presentation only (front-loading, dating, cited statistics), versions assigned at random so topic does not contaminate the result; run as a controlled experiment in parallel with prototype 2	Measurable difference in citation frequency between matched versions across five assistants after four weeks	1
6	Policy and compliance kit	Crawler policy, agent verification, AI disclosure and Content Credentials can be packaged as a reusable layer that customers will pay for	robots.txt generator by bot and category, Content Signals, Search Console toggle guidance, violations log, Article 50 chatbot disclosure, C2PA preserved through the image pipeline; built only after a customer has asked for it	Kit applied to the reference site and one paying client site with no manual steps beyond configuration	5

What the roadmap deliberately leaves out. Free-form page generation, in-chat checkout and llms.txt are not prototyped; the evidence says they are not where value lies in the next two years. Security review of prompt-built apps is a strong product candidate but a separate initiative, since it does not depend on the reference architecture.

Decision points. After prototype 2 the team will know whether assistants and agents reach real Sysbot client sites in numbers that justify prototypes 3 and 4. After prototype 4 the team will know the real cost per session of a generative layer and can decide whether

to productise it. Prototype 6 waits for a customer willing to pay for it. The biggest mistake would be building all six without validating that customers need them.

Three validations, not one. A prototype that passes its success measure has achieved technical validation only. Before anything is productised it must also pass user validation, evidence that it solves a real customer problem, and commercial validation, a customer who pays for it. A successful prototype is never treated as evidence of product-market fit; each of the four candidate products carries its own validation question in the strategic position section.

Strategic position: AI-ready website infrastructure

Sysbot should not position itself as an AI website builder; that market is crowded and the price of its output is collapsing. The stronger position is a platform that turns conventional business websites into AI-readable and agent-operable business systems.

Product layer	What Sysbot provides
Content infrastructure	Structured business content and product information, owned and exportable
Human experience	Fast, accessible, search-friendly websites
AI discovery	Schema.org, structured content and citation measurement
Agent infrastructure	MCP server and approved business tools
Generative experience	Catalog-constrained dynamic UI and personalised composition
Trust infrastructure	Permissions, crawler controls and security
Analytics	AI discovery, agent actions and business conversion

Four candidate products, validated separately.

Product	For whom	Revenue model	Validation question
AI-ready website	Businesses that need a conventional site with structured content, AI discoverability and agent compatibility	Setup fee plus monthly maintenance	Will customers pay a premium over a template site for discoverability they cannot yet measure?
Agent gateway	Businesses whose information and approved operations should be reachable by assistants and agents	Subscription plus usage-based pricing	Do assistants reach these businesses in numbers that justify the gateway, and who pays for agent traffic?
AI website intelligence	Any business wanting to know how assistants discover, retrieve and act on its site	SaaS subscription	Is attribution accurate enough to show revenue, not just visits?
Agent access economics	Businesses whose data and operations agents consume repeatedly	Measurement first; pricing, rate-limiting or monetisation only once consumption is understood	Which agents generate business and which only consume resources, and does any customer want to charge for access?

Each product should be validated independently with paying customers before any of them are combined. The commercial conclusions in this study are partially validated at best: the evidence shows the shifts are real, not yet that businesses will pay to respond to them.

Next step. Before the prototypes, write a second document, a Sysbot technical reference architecture, defining component boundaries, data contracts, the security model with its authentication, authorisation and consent layers, APIs, MCP integration, the generative UI catalog and the measurement framework. That document becomes the engineering blueprint all six prototypes are built against.

Methodology and limitations

This is desk research over primary sources dated 2025 to October 2026: platform announcements, standards drafts, measured traffic studies, regulatory texts and company filings. It is an independent, evidence-informed industry research study and strategic technology assessment. It is not peer-reviewed academic research and does not report primary user research or production measurement; the prototype programme that follows it exists to produce that evidence.

How the research was done. Four research threads ran in parallel: discovery (Shift 1), agents (Shift 2), interfaces and builders together (Shifts 3 and 4), and trust and rules (Shift 5). Each thread started from the primary publishers for its topic and used the trade press only to find them: standards bodies (W3C, IETF, OWASP), regulators and legal texts (EUR-Lex and the European Commission, the US Federal Register and state legislatures, the UK CMA, MeitY), the engineering blogs and documentation of the platform companies (Google, OpenAI, Anthropic, Microsoft, Cloudflare, Shopify, Vercel), the measurement companies (Similarweb, Ahrefs, Chartbeat, Seer Interactive, Cloudflare Radar), independent research institutes (Pew Research Center, Nielsen Norman Group, Baymard Institute) and arXiv. Searches were run in English, with one French-language survey, for the period January 2025 to 30 October 2026.

Inclusion and exclusion. A figure was included when it could be read on a page that states it, with its date and the scope it was measured over. Where the original dataset was public, it was linked in preference to reporting about it. A press or blog page was linked only when it named its original source and the original was paywalled or unavailable, or when the page was itself the subject of the claim. Figures were excluded when they were unattributed or circulated without a traceable origin, however widely repeated; the claimed 30 to 41% increase in technical debt from AI-assisted coding is one example named in the text. Vendor figures were included only when identified as vendor claims.

Evidence grading and verification. Every sourced claim in the five Shift sections, 146 claims from 139 distinct sources, is listed in the companion evidence register with the grade of its source: E1 for standards, regulations and primary research; E2 for independent measurement; E3 for vendor self-reports; E4 for secondary reporting and estimates; E5 for this study's own interpretations and hypotheses, which are listed separately with the evidence they rest on and what would overturn them. On 31 October 2026 every source cited in v1.0 was requested again and each page that could be read was checked for the figure the study attributes to it, 143 claims in all: 104 claims were confirmed as stated, 24 confirmed with a caveat now recorded in the register, 4 were not supported by the page cited and have been corrected in this version, and 11 could not be re-read because the page refused automated access. By grade, 33 claims rest on E1 sources, 17 on E2, 48 on E3 and 48 on E4.

Review. Two independent critical audits were run against the draft on 30 October 2026 by reviewers who took no part in writing it, and their corrections were applied; one of their objections, to the UK conduct-requirement dates, was resolved by reading the regulator's Notice directly. A third audit, of research credibility, was received on 31 October 2026; its recommendations produced the evidence register, this methodology section and the changes listed in the version history. None of the reviewers has a commercial interest in Sysbot's products.

Three kinds of statement. The study keeps apart observed evidence (the Shift sections, each figure linked and graded), research interpretation (the position taken on each open question, and readings marked as such in the text) and strategic hypothesis (the bets, the strategic position and the prototype roadmap, which exists to test them). A prediction in this document is a hypothesis to be tested, not a finding.

Limitations.

- Vendor-reported figures (revenue run rates, resolution rates, red-team results, personalisation lifts) are identified as such and were not independently verified.
- Traffic and conversion studies differ in market, device, period and definitions; the study reports each with its scope and does not combine them into a single estimate.
- Several figures were available only through secondary reporting; where the original dataset could not be read, the figure is marked as unverified and not relied on in conclusions. The register names the original wherever the cited page does.
- Regulatory statements summarise primary texts but are not legal advice; applicability depends on the entity, its role and its jurisdiction, and a legal review is recommended before relying on any deadline.
- The field is moving quickly; statements about protocol status, browser-agent products and crawler policies are accurate as of the dates given and may change within months.
- No primary user research or production measurement was performed; the prototype roadmap exists to produce that evidence, and no commercial opportunity named here is validated until it does.

Version history.

Version	Date	Change
1.0	30 October 2026	First public release, frozen after two independent critical audits.

Version	Date	Change
1.1	31 October 2026	Evidence register published and every sourced claim re-checked against its source. Methodology expanded with search strategy, inclusion and exclusion rules, evidence grades and review process. Positioning stated as an evidence-informed industry study. The "first 80%" estimate removed from Shift 4. The AI Overviews attribution marked as a hypothesis. Twelve claims corrected where the cited page did not support the wording: the Intercom resolution-rate comparison, the Squarespace AI Visibility citation, agency pricing under AI, the Lovable scanner statement, the ChatGPT citation-position figure, the AI SDK version note, the Cloudflare Pay Per Crawl cell, the Web Bot Auth milestone and signed-agents wording, the UCP waitlist scope, the publisher coalition's request, the Digital Omnibus status and the DPDP commencement instrument; the Atlas, Mariner, Edge, WebMCP, Thales, Browserbase, OpenAI commerce, NN/g and Search Console sentences were qualified to match their sources. Four citations moved from secondary reporting to the original source and three primary sources added.

Disclaimer. This document is an independent technology research publication by Sysbot Technologies' Research & Development function. It is not legal, financial or investment advice. Product and company names are the property of their owners.

Sources

Pages opened for this study, grouped by shift. Every figure in the text links to its page inline; this list is the index. Vendor and company figures are self-reported unless a study is named. The evidence register grades every linked claim and records when each page was last checked.

Shift 1: discovery

- Google I/O 2026 keynote, May 2026
- Pew Research, Google users are less likely to click when an AI summary appears, Jul 2025
- Ahrefs, AI Overviews reduce clicks, update, Feb 2026

- Seer Interactive, AIO impact on Google CTR, 2026 update, Apr 2026
- Similarweb and SparkToro, zero-click marketing, Jun 2026
- Semrush, Google AI Mode SEO impact, Jul 2025
- Chartbeat 2026 Publisher Playbook via Search Engine Journal and Axios
- Google Search Central, AI features and your website, updated Dec 2025
- SIGIR 2026, What gets cited
- Ahrefs, 38% of AI Overview citations pull from the top 10
- Peec AI, AI search and GEO statistics, Mar 2026
- Ahrefs, do AI assistants prefer fresh content
- SE Ranking, does llms.txt impact AI visibility and citations and llms.txt adoption 2026
- Search Engine Roundtable, Mueller on llms.txt, Jun 2025
- Cloudflare, Content Independence Day and PPC Land on pay per answer
- Search Engine Land, AI company and publisher partnerships; Engadget on Perplexity Comet Plus
- Search Engine Land, Penske Media lawsuit dismissed, Sep 2026
- PPC Land, publisher groups ask EU and Canada to act, Oct 2026
- BCC Interactive, local AI search statistics; Statcounter, AI chatbot market share worldwide
- CMA press release, publisher conduct requirement imposed on Google, 3 Jun 2026; CMA Notice under section 19(1), paragraph 6 on commencement; CMA case page

Shift 2: agents

- Anthropic, piloting Claude for Chrome, Aug 2025
- OpenAI forum, Atlas scheduled to stop working; OpenAI help, evolving Atlas; ChatGPT agent
- Gigazine, Google shuts down Project Mariner; Google, Gemini 3 auto browse in Chrome
- Microsoft Edge, May 2026 updates and Edge for Business agentic browsing
- Anthropic, Claude in Chrome generally available, Aug 2026
- PPC Land, Comet free worldwide; Trail of Bits, auditing Comet; Brave, unseeable prompt injections
- MCP blog, July 2026 specification
- Chrome, WebMCP origin trial; WebMCP docs; implementation status
- Presenc, NLWeb adoption 2026; Microsoft, optimise your site for agents; Cloudflare, NLWeb and AutoRAG
- IETF, Web Bot Auth HTTP signature protocol draft; Cloudflare, signed agents; Cloudflare, the agentic web

- Agentic Commerce Protocol changelog and seller guide; OpenAI, powering product discovery in ChatGPT; Digital Silk on ChatGPT commerce
- Google, Universal Commerce Protocol under the hood; Merchant Center UCP; Agent Payments Protocol to FIDO
- Visa Intelligent Commerce Connect; Mastercard Agent Pay trust framework
- Cloudflare, accountable mixed-use AI crawlers; Cloudflare, the Internet has a second audience; Markdown for Agents
- Thales and Imperva 2026 Bad Bot Report; Ecorp IT, crawl-to-referral ratios 2026
- Browserbase, agent browser introduction

Shift 3: interfaces

- Google Research, generative UI; Gemini help, dynamic view; A2UI v0.9
- Vercel AI SDK RSC overview; InfoQ on json-render
- OpenAI, apps in ChatGPT; app submissions; commerce production guide
- MCP Apps announcement; Anthropic, interactive tools in Claude; OpenUI (formerly Thesys C1) docs
- Inman, Agent Image SiteAI; Shopify Q2 2026 earnings call
- Intercom, from resolutions to outcomes; Simpli, the Fin resolution number; HappySupport deflection benchmarks
- Nielsen Norman Group: discoverability of AI at Amazon, Google AI Mode, the articulation barrier, AI in UX survey 2026
- Baymard, AI engagement quantitative UX insights 2026; Shno, AI personalisation statistics
- Very Good Ventures, run-rate economics of generative UI; arXiv 2604.25455, generative UI and accessibility

Shift 4: building

- TechCrunch on Lovable: $500M, $600M
- Bleek, CVE-2025-48757; Launch Ready Code, is Lovable safe
- Sacra on Replit and Vercel; AI Magazine on Replit's valuation
- Wix Q2 2026 results, SEC filing; Latka on Base44
- Altar, builder comparison; Squarespace, July 2026 product updates
- arXiv 2512.11922, Vibe coding in practice; arXiv 2607.16680, spec-driven development
- GoodFirms website cost survey; Fast Company and AIGA freelance report; Upwork in-demand skills; 925 Studios on agencies
- VentureBeat on WebMCP in Chrome; Digiday on agentic browsers and publishers; The New Stack, the agentic web

Shift 5: trust and rules

- RFC 9309, Robots Exclusion Protocol, Sep 2022; IETF AI Preferences charter
- Cloudflare, control content use for AI training; Content Signals Policy; robots.txt violations tracking; Perplexity stealth crawlers
- IETF AI Preferences documents; IETF Web Bot Auth charter
- RSL standard press release; Digiday on the RSL Collective
- Google common crawlers; 9to5Google on the AI features opt-out; Google spam policies
- OpenAI bots; Anthropic crawler support page; Perplexity crawlers
- European Commission, code of practice on transparency of AI-generated content; ActuIA on Omnibus 2026/1744
- Loeb, China AI labelling measures; Orrick, California AI Transparency Act delay; Beebom on SynthID and C2PA in Search and Chrome
- Greenberg Traurig, European Accessibility Act; Federal Register, ADA Title II extension; W3C WCAG 3.0 working draft; arXiv 2511.19477, agents and the accessibility tree
- Privacy Sandbox update, Oct 2025; The Media Leader on cookie settings; Usercentrics on the Digital Omnibus; S.S. Rana summary of India's DPDP Rules (secondary)
- Brave, Comet prompt injection; IT Brief on OpenAI Lockdown Mode; Vectra on EchoLeak; OWASP Top 10 for LLM applications
- MeitY, Digital Personal Data Protection Rules 2025 and enforcement timeline; Gazette commencement notification, dated 13 Nov 2025