25 security checks for your AI-built app

Mike Dufour · AI App Security · mikebillhaven@gmail.com · November 2026

Built with Lovable, Bolt, Supabase, Bubble or n8n? Go through this list before launch and after every big AI-generated change. Only check systems you own or have written permission to test.

Database and row-level security

- **1.** Row-level security is **on** for every table in the public schema.
- **2.** Every policy names who: `auth.uid() = user_id` or a role check, never `using (true)` for insert, update or delete.
- **3.** `select using (true)` exists only on tables where every column is meant to be public.
- **4.** No `grant insert/update/delete ... to anon`.
- **5.** `security definer` functions set `search_path` and check the caller.
- **6.** Storage buckets: private by default, with policies per user or folder.

Keys and secrets

- **7.** No secret keys in the front-end bundle (OpenAI, Anthropic, Stripe `sk_live_`, AWS, service-role keys).
- **8.** The Supabase service-role key is used only in server code or edge functions.
- **9.** No `.env` files in exports, builds or public folders.
- **10.** No secrets in prompt templates, logs or error messages.
- **11.** Any key that was ever shipped to a browser has been rotated.

Authentication and access

- **12.** Role and admin checks happen on the server or in RLS, not only in the UI.
- **13.** Edge functions and APIs verify the user's token (`verify_jwt` stays on unless a webhook checks its own signature).
- **14.** Users can't change their own role or plan by editing a request.
- **15.** Password reset and magic links expire and work once.

API and browser

- **16.** CORS allows only your own front-end origins, never `*` with credentials.
- **17.** No debug, admin or test routes reachable in production.
- **18.** User or model output is never rendered as raw HTML (`dangerouslySetInnerHTML`, `innerHTML`, `v-html`).

AI features and agents

- **19.** Every tool the agent has is listed, with its worst case written down.
- **20.** Tools that send, pay, delete or run code have an allow-list or a human confirmation.
- **21.** Text from users, web pages, e-mails and files is treated as untrusted and never placed in the system prompt.
- **22.** Model output is validated before it's used in a query, a URL or a command.
- **23.** Personal data sent to a model provider is minimised and covered by their data terms.

E-mail and domain

- **24.** DMARC, SPF and DKIM are set, so nobody can send convincing mail as your domain.
- **25.** Security headers are set on the site (HSTS, CSP, X-Content-Type-Options, frame-ancestors).

Free exposure check: I look only at what's already public (e-mail settings, headers, keys in your site's JavaScript) and send you a short report. No passwords, no testing of your systems without your written permission. Haven Digital Ltd · England & Wales · company no. 17063556