05. AI-written code review standard

> **Template.** Part of the open-source AI App Policy Kit, version 1.0. Replace everything in [square brackets], delete what does not apply, and have the person who runs the business approve it. A starting point, not legal advice.

Business	[Business name]
Document owner	[Name and role]
Approved by, and date	[Name], [date]
Next review	[Date, at most 12 months on]

Why this matters

AI tools write code that works when everything goes right. They often leave out what keeps data safe when it does not: checking who is asking, refusing bad input, handling errors. A person who understands code must check it before customers depend on it.

Who it covers

All code that runs in [Business name]'s apps, whoever or whatever wrote it: staff, contractors, AI assistants and AI agents.

The rules

1. **Nothing goes live unreviewed.** Every change is read by a person who understands the code and did not write the prompt that produced it. If nobody in the business can, use an outside reviewer for anything that touches sign-in, payments or personal data.
2. **The person who merges owns it.** Responsibility for the change sits with a named person, never with the tool.
3. **Keep changes small.** A reviewer can check 200 lines properly, not 5,000. Ask the tool for one thing at a time.
4. **Tests come with the change.** New behaviour has an automated test, including at least one for what must be refused.
5. **Automated checks run on every change:** tests, a linter, a dependency vulnerability scan and secret scanning (document 04).
6. **Check every new dependency.** AI tools suggest packages that are abandoned, vulnerable or do not exist, and attackers register the invented names. Before adding one, confirm that it exists, is maintained and is the package you meant.
7. **AI agents work in a sandbox.** An agent that can run commands has no production access and no production secrets, and its changes arrive as a proposal a person approves.
8. **Know the licence.** Do not accept large blocks of code that look copied from a named project unless its licence allows it.

Review checklist

The reviewer checks each of these on any change that touches data, sign-in or money.

- ☐ **Who is asking?** Every route and function that returns or changes data checks the user is signed in.
- ☐ **Are they allowed?** A user can reach only their own records. Changing an id in the address does not show someone else's.
- ☐ **Is input checked on the server?** Type, length and allowed values are enforced on the server as well as in the browser.
- ☐ **Are database queries built safely?** Parameters are bound, not joined into strings.
- ☐ **Is output escaped?** User text and AI output are shown as text, never run as HTML or code.
- ☐ **Are errors handled?** A failure shows the user a plain message and logs the detail, with no stack trace and no personal data on screen.
- ☐ **Are there no secrets in the change?**
- ☐ **Is personal data kept out of logs?**
- ☐ **Are limits in place?** Rate limits on sign-in and on anything that costs money per call, such as an AI request.
- ☐ **Does it fail closed?** If the permission check itself fails, access is refused.

Record

Date	Change	Written with	Reviewed by	Outcome
[Date]	[Link or description]	[Tool]	[Name]	[Approved or sent back]

References

- OWASP, *Top 10* and *Application Security Verification Standard*.
- NCSC, *Guidelines for secure AI system development*.

AI App Policy Kit, version 1.0. © 2026 PugliseWeb LTD, licensed under CC BY 4.0: use it, change it and share it, with credit. The whole kit, and a free Software Health Check of your app against it: https://www.pugliseweb.com/ai-policy-kit