13. Human oversight of AI decisions

> **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 systems are wrong some of the time, and they are wrong fluently. Where a wrong answer costs a customer money, an opportunity or their trust, a person must be able to check it, stop it and put it right. UK data protection law also gives people rights when a significant decision about them is made by automated means alone.

Who it covers

Every AI feature at [Business name] that decides, recommends or acts: chat assistants, scoring, screening, pricing, automated replies and AI agents.

The rules

1. **Sort every AI use into a level** using the table below, and record it in the register.
2. **Significant decisions about a person are made by a person.** Refusing a service, a job or credit, setting an individual's price, closing an account or resolving a complaint is never left to AI alone. The person deciding has the authority and the information to disagree with the AI, and the time to do so.
3. **People are told.** Customers know when they are dealing with an AI system, and how to reach a person.
4. **There is always a way to a person.** Every AI conversation offers a hand-over, and it is answered within [one working day].
5. **Results can be challenged.** A person affected by an AI-assisted decision can ask for an explanation and a review by someone who was not involved. Requests are logged.
6. **Actions have limits.** An AI agent that can send, book, pay, refund or delete has a ceiling on each (amount, number per day, who it can contact). Anything above the ceiling waits for a person's approval.
7. **There is an off switch.** [Name] can turn each AI feature off within minutes, and the business can carry on without it. The steps are written down.
8. **Check the output regularly.** Each [month], someone reviews a sample of AI outputs for accuracy, tone and fairness, and records what they find.
9. **Watch for unfair patterns.** Where AI affects how people are treated, compare outcomes across groups of customers at least every 12 months.

10. **Mistakes are corrected and learned from.** A wrong output that reached a customer is fixed with the customer, logged, and used to improve the prompt, the data or the limits.

Levels of oversight

Level	What the AI does	Oversight needed
1. Helps staff	Drafts, summarises, suggests. A member of staff decides what to use.	Staff check the output (document 01).
2. Talks to customers	Answers questions, with no effect on what the customer gets.	Customers are told. Hand-over to a person. Monthly sample.
3. Acts within limits	Books, sends, updates records, gives small refunds.	Written limits. Every action logged. Approval above the limit. Off switch tested.
4. Affects a person's rights or money	Screening, scoring, pricing, eligibility.	A person decides every case. Risk assessment first (document 09). Fairness check.

Register of AI uses

AI feature	Level	Limits	Who reviews, how often	Off switch held by	Last reviewed
[Website assistant]	[2]	[Cannot quote prices]	[Name, monthly]	[Name]	[Date]
[ ]	[ ]	[ ]	[ ]	[ ]	[ ]

Review record

Date	Feature	Sample size	Problems found	Action
[ ]	[ ]	[ ]	[ ]	[ ]

Checklist

- ☐ Every AI feature is in the register with a level.
- ☐ No level 4 decision is made without a person.
- ☐ The off switch for each feature was tested in the last 12 months.
- ☐ The monthly sample reviews are recorded.
- ☐ Challenges and their outcomes are logged.

References

- ICO, *Rights related to automated decision-making including profiling*.
- ICO and The Alan Turing Institute, *Explaining decisions made with AI*.

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