DROPSOLID
ARTIFICIAL INTELLIGENCE

GUIDE 03

SOVEREIGN AI · TRANSPARENCY AND GOVERNANCE

Glass-box AI health check

Can you inspect, audit and steer the AI you already run?
Work through this checklist to identify where your system has strong governance controls, where you have blind spots, and which gaps need attention.

Traceability 1
Observability 2
Transparency 3
Feedback 4
Auditability 5
Steerability 6

Published by

Dropsolid AI · 2026

SEE

Can you see what the system is doing?

SYSTEM / USE CASE

1 Traceability

Can you reconstruct one interaction from request to response?

EVIDENCE TO ASK FOR

- [ ] Interaction ID or timestamp
- [ ] Original user request
- [ ] Retrieved passages or context
- [ ] Model used, version where available
- [ ] Response returned, citations or sources

ASSESSMENT

○ No ○ Partial ○ Yes

NOTES

e.g. evidence gathered, what's missing...

Red flag · Only the user request and final answer are stored. What happened between them cannot be reconstructed.

2 Observability

Can you see how the system is performing while it operates?

EVIDENCE TO ASK FOR

- [ ] Live monitoring or dashboard
- [ ] Usage, latency, cost
- [ ] Response-quality signals
- [ ] Decline or failure rate
- [ ] Drift or operational alerts

ASSESSMENT

○ No ○ Partial ○ Yes

NOTES

e.g. evidence gathered, what's missing...

Red flag · The first indication of a problem usually comes from a user complaint.

3 Transparency

Can an accountable reviewer inspect the evidence behind an answer?

EVIDENCE TO ASK FOR

- [ ] Retrieved source material, citations
- [ ] Retrieval information
- [ ] Confidence information, where available
- [ ] Model or configuration used
- [ ] Understandable data and model path

ASSESSMENT

○ No ○ Partial ○ Yes

NOTES

e.g. evidence gathered, what's missing...

Red flag · A vendor dashboard reports success, but your team cannot inspect the evidence behind the answer.

CONTROL

SYSTEM / USE CASE

Can you govern what you can see?

4 Feedback

Can a poor answer trigger a defined improvement process?

EVIDENCE TO ASK FOR

- [ ] User or expert feedback mechanism
- [ ] Named owner, triage process
- [ ] Record of source corrections
- [ ] Record of configuration adjustments
- [ ] Evidence recurring issues are acted on

ASSESSMENT

○ No ○ Partial ○ Yes

NOTES

e.g. evidence gathered, what's missing...

Red flag · Feedback is collected, but nobody owns what happens next.

5 Auditability

Can you produce the record of an interaction when governance or compliance asks for it?

EVIDENCE TO ASK FOR

- [ ] Sample interaction trace, export function
- [ ] Filtering by date, use case, audience
- [ ] Retention policy, access to records
- [ ] Agreed export format
- [ ] Named owner for audit requests

ASSESSMENT

○ No ○ Partial ○ Yes

NOTES

e.g. evidence gathered, what's missing...

Red flag · Producing an audit record requires engineers to reconstruct information manually from several systems.

6 Steerability

Can you change important system behaviour through configuration?

EVIDENCE TO ASK FOR

- [ ] Configurable prompts, retrieval thresholds
- [ ] Model selection, source scope
- [ ] Refusal and fallback settings
- [ ] Configuration history, approval process
- [ ] Behaviour adjusts without a rebuild

ASSESSMENT

○ No ○ Partial ○ Yes

NOTES

e.g. evidence gathered, what's missing...

Red flag · Meaningful behaviour changes require engineering work and a new deployment.

RESULT

Uncover your blind spots

SYSTEM / USE CASE

PROPERTY	NO Black box	PARTIAL Close the gaps	YES Glass box	MAIN EVIDENCE OR GAP	OWNER	NEXT ACTION
1 **Traceability**	○	○	○			
2 **Observability**	○	○	○			
3 **Transparency**	○	○	○			
4 **Feedback**	○	○	○			
5 **Auditability**	○	○	○			
6 **Steerability**	○	○	○			

Continue the conversation

Have a talk with Dropsolid about these questions, and see how we answer each one in practice.

Talk to Dropsolid AI

Whitepaper 03 — Glass box vs black box →

DROPSOLID
ARTIFICIAL INTELLIGENCE

Read Whitepaper 03, Glass box vs black box, for the architectural rationale behind traceability, observability, transparency, feedback, auditability and steerability.