POLICY-READER TECHNICAL NOTE

For AI Office / Policy Reader

Article 50 Transparency as an Evidence Chain, Not a Label

Kotov Ivan
Bruxelles Belgique
2026

Document boundary

Status: Short policy-reader note. Technical implementation contribution.
Boundary: Not legal advice. Not certification. Not a conformity assessment. Not a formal Article 50 compliance claim. Formal compliance remains deployment-specific and subject to legal review.

Academic PDF edition. Prepared for reviewable technical reading.

For AI Office / Policy Reader

Status: Short policy-reader note. Technical implementation contribution.
Boundary: Not legal advice. Not certification. Not a conformity assessment. Not a formal Article 50 compliance claim. Formal compliance remains deployment-specific and subject to legal review.

Purpose

This note proposes a narrow implementation grammar for Article 50 transparency: **system → role → obligation → control → evidence → responsible actor**.

The practical problem is simple. Many organisations can state that “AI was used” or that their system is “transparent”. Such statements are rarely enough for implementation, supervision, procurement, or later review. A useful transparency claim should be traceable from a concrete AI system or AI-assisted workflow to a concrete actor role, a concrete obligation trigger, a concrete disclosure or marking control, a concrete evidence artifact, and a responsible human or organisation.

This is not a proposal to create a new legal category of AI. It is a proposal to make Article 50 implementation easier to inspect.

1. System boundary

Transparency should not be assessed against “AI” in general. The first question should be: **what is the system or workflow under review?**

A system may be a chatbot, a generative media tool, an AI-assisted publication workflow, an internal enterprise agent, a public-service interface, or a mixed workflow using third-party models, local tools, human editors, and automation. Without a system boundary, transparency becomes a slogan. With a boundary, the reviewer can ask what the system does, who uses it, who is affected by it, what outputs are produced, and where disclosure or marking is required.

Recommended minimum output: a short system description, scope limits, user groups, output types, deployment context, and excluded functions.

2. Actor role

The same organisation may act differently at different points in a workflow. It may be a provider for one system, a deployer for another, a publisher for an AI-assisted report, and an enterprise operator for an internal agent workflow. The role should therefore be classified per system or workflow stage, not by reputation, brand, or institutional identity.

Recommended minimum output: provider, deployer, publisher/editorial actor, enterprise operator, agent operator, researcher, protocol author, or mixed role; plus a short explanation of why that classification was selected.

3. Obligation trigger

After the system and role are defined, the implementation should identify the relevant transparency trigger. Examples include direct AI interaction, AI-generated or AI-manipulated content, public-interest AI-assisted publication, deepfake or synthetic media exposure, emotion recognition, biometric categorisation, or another specific Article 50-relevant event.

The trigger should be specific enough to determine what the natural person must see, what machines or auditors may need to inspect, and what evidence should be preserved. A generic statement that “AI is involved” should not be treated as equivalent to trigger analysis.

Recommended minimum output: trigger type, affected person or audience, point in the workflow where the trigger occurs, and uncertainty requiring legal review.

4. Control

A transparency control should match the trigger. Different controls serve different audiences and should not be collapsed into one checkbox.

A visible notice informs a natural person. A machine-readable mark supports detection. A provenance sidecar supports traceability. A witness or audit event supports later review. A human review record supports responsibility. A correction, takedown, dispute, or rollback route closes the feedback loop.

For AI-assisted workflows, especially tool-using or CLI/cloud agents, controls should also cover task contracts, permission scopes, sandbox boundaries, release gates, and human approval before publication. The question is not only which model generated an output, but who accepted it, under what authority, with what disclosure, and with what evidence.

Recommended minimum output: user-facing notice, marking or provenance mechanism where applicable, audit or witness event, human review gate, and correction route.

5. Evidence

Evidence of transparency is not the same as transparency itself. A log showing that a notice was displayed may help an auditor, but the person still needs a clear notice. Metadata may support detection, but it may be stripped downstream and may be invisible to the public. A human review statement may support accountability, but only if the review is substantive and attributable.

Useful evidence artifacts may include UI screenshots, versioned notice text, interaction-state logs, provenance sidecars, machine-readable metadata, release manifests, hashes, witness records, task contracts, sandbox records, human review records, and correction logs.

Evidence should also respect data minimisation. Auditability should not become surveillance. A good record should preserve boundary events, provenance state, review state, and responsible actor information without unnecessarily storing private prompts, secrets, conversations, or personal data.

Recommended minimum output: evidence list, retention class, privacy boundary, integrity mechanism, and review access route.

6. Responsible actor

Transparency without responsibility becomes decoration. The final question should be: **who is accountable for the control and the review?**

The answer should not be “the model”. It should be a named human, team, organisation, publisher, deployer, provider, or other accountable actor. Where human review is used, the record should identify the reviewed artifact, reviewer or responsible editor, review scope, decision, timestamp, and correction path.

Recommended minimum output: responsible human or organisation, role, review authority, escalation route, and legal review status.

Practical bridge

This pattern is intended to bridge policy language and engineering implementation. Article 50 transparency should be capable of moving from principle to inspection: from a legal obligation to an implemented control, and from an implemented control to evidence that a reviewer can examine.

In building terms, a label on an electrical panel is not the same as the wiring map, the inspection record, or the electrician’s signature. All four matter. AI transparency has the same structure. The notice is the label. Provenance is the wiring map. The audit or witness record is the inspection trail. The responsible actor is the signature. If any part is missing, the system may still be useful, but it should not be presented as fully evidenced transparency.

Recommended policy formulation

Article 50 guidance should encourage implementers to document transparency as an evidence chain:

system boundary → actor role → obligation trigger → disclosure or marking control → provenance / witness / audit evidence → human review where relevant → responsible actor → correction or review route.

This would help avoid decorative transparency, preserve the provider/deployer distinction, support mixed-role workflows, and make AI-assisted publication and agentic workflows easier to supervise in practice.