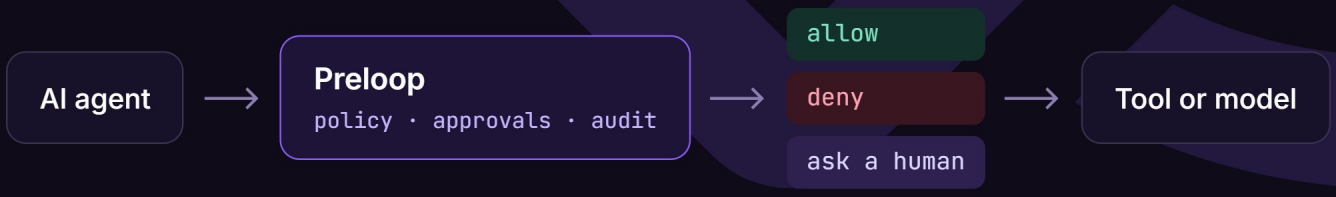
Preloop

COMPLIANCE BRIEF

PRELOOP AND THE EU AI ACT

Agents act. People decide. Records prove it.

How Preloop, the open-source control plane for AI agents, stands under the EU AI Act — and how it helps you meet your own obligations when agents act on your behalf.



Works with Claude Code, Codex CLI, Cursor, Gemini CLI, OpenClaw, OpenCode, Windsurf and any MCP-compatible agent.

EDITION

Version 1.0 · November 2026

PRODUCT ASSESSED

Preloop 1.0.0 — open source, Cloud and Enterprise

LAW

Regulation (EU) 2024/1689, as amended by Regulation (EU) 2026/1744

01 — AT A GLANCE

Your agents can act. You stay in control, and can prove it.

AI agents now open pull requests, query production databases and prepare refunds. The EU AI Act asks three things of that activity: who allowed it, who could stop it, and what actually happened. Preloop answers them at the point of action.

1 Preloop's own position

The core control plane applies rules written by people, so it is not an AI system. Its optional AI-assisted features are. We provide them as AI systems that are not high-risk, running on models you choose, with AI decisions always labelled and a human able to override.

2 Your obligations

When you use AI agents you are a deployer, and sometimes a provider. Preloop gives you the mechanisms the Act looks for: rules checked before an action runs, a named person on consequential steps, a stop control, and tamper-evident records kept for more than six months.

3 Clear lines

Preloop produces the records. You and your counsel decide classification and whether the evidence is sufficient. This brief says plainly what Preloop does not do.

Where the AI Act stands

as amended by Regulation (EU) 2026/1744, OJ L, 24.7.2026



The high-risk deadline moved. The design work did not.

Oversight, logging and retention take months to build into live workflows, and GDPR already applies to every agent log that names a person.

02 — WHERE PRELOOP STANDS

Our assessment, with the reasoning shown.

Classification under the Act is a legal determination. We publish ours for Preloop 1.0.0 so that your counsel can test it.

RULE-BASED

The control plane

MCP firewall, model gateway, policy engine (YAML + CEL), approvals and audit trail. It applies rules written by people; it does not infer.

Not an AI system

MODEL-BASED · OPTIONAL

AI-assisted features

AI review of approval requests, agent workflows (pull-request review, issue triage, security audits) and semantic session search. Used only if you enable them.

AI systems, not high-risk

QUESTION	OUR POSITION	BASIS
Our role	Provider of the AI-assisted features, which we place on the market under the Preloop name. Deployer of AI in our own operations. We do not rely on the open-source exclusion for paid editions.	Arts. 2(12), 3(3), 3(4)
General-purpose AI	Not a model provider. Preloop does not develop or train models; its features run on third-party models you select and can restrict.	Arts. 3(63), 53
Risk tier	No prohibited practice, and not high-risk. The intended purpose is governing software agents in engineering and business operations: not a safety component (AI used solely for automation or efficiency is not one), and not intended for any Annex III area. No EU database registration is therefore needed.	Arts. 5, 6, 6(1a), 49; Annexes I, III
Transparency	AI decisions are recorded and shown separately from human ones. AI output is attributed to a named agent, with the model used on record.	Art. 50
AI literacy	Every agent we run works under our own policies, approvals and audit trail. Our documentation supports your teams.	Art. 4, as amended

Changing the purpose changes the role.

Putting an AI-assisted feature to an Annex III use, such as AI review that screens job applicants, can make you the provider of a high-risk system. Talk to us first.

Art. 25(1)(c)

03 — SUPPORTING YOUR OBLIGATIONS

What the Act asks of you, and what Preloop provides.

Most agentic use is not high-risk. Where a use is, the deployer duties of Art. 26 apply from 2 December 2027. Either way, these are the controls auditors will ask to see.

THE ACT ASKS	PRELOOP PROVIDES	YOU STILL OWN
Human oversight Arts. 14, 26(2)	Rules marked `require_approval` pause the action until a named person decides, on mobile, Slack, Mattermost, email or CLI. Approvers see the tool, arguments and matching rule. AI and human decisions are recorded apart. **For every governed action this provides the capabilities Art. 14(4)(d) and (e) describe: a person can override the action, and can stop it.** Enterprise: teams, quorum, escalation	Which actions need a person; trained approvers with authority (AI literacy, Art. 4).
Stop and intervene Arts. 14(4)(e), 26(5)	A kill switch, for flows or for tools, blocks model and tool traffic, ends live agent sessions and freezes pending approvals, with audited recovery. Operator notes steer a running agent.	When to suspend, and whom to inform.
Records Arts. 12, 19, 26(6)	Every governed tool call, decision and approval is logged and linked in tamper-evident, hash-chained logs, kept for at least 183 days (beyond the six-month minimum), with legal hold. Signed exports become immutable on your object-lock (WORM) storage.	Your retention period; activity that bypasses Preloop.
Risk controls and robustness Arts. 9, 15, 26(1)	Versioned policy-as-code (YAML + CEL) checked before execution, outside the model; the rule in force is recorded with each decision. Allowed-model lists, budgets, redaction. Agents never hold provider keys.	Risk assessment and testing. Prompt-injection protection in Preloop is partial.
Monitoring and incidents Arts. 26(5), 72, 73	Session replay: step through every model request and response, and every tool call, that led to an action, with policy, approvals, spend and outcome. An export lists incident candidates.	Deciding what is a serious incident, and reporting it.
Data governance Arts. 4a, 10	Preloop trains no models. Allowed-model lists keep agents on approved models whose providers document their training data (Art. 53); every record shows which model handled which request.	Data governance for any high-risk system you train or fine-tune.
Impact assessments Art. 27; GDPR Art. 35	The facts: what each agent may do, who approves, what is recorded.	The FRIA and DPIA; the FRIA may cite the DPIA (Art. 27(4)).
Building high-risk AI Art. 25(4)	On request, we agree in writing the information, access and assistance you need from us as a supplier.	Your obligations as provider.

04 — SHARED RESPONSIBILITY

Preloop produces the record. Your counsel gives the verdict.

COUNSEL

Requirement

Which instrument and article? Which role and risk tier?

→

PRELOOP

Record

Produced automatically as the action passes the control. States its own limits.

→

COUNSEL

Verdict

Sufficient for this article? What is missing?

Engineering owns the configuration: rules, approvers and retention. When a rule, agent, model or law changes, the assessment is due again.

What Preloop does not do, by design

- ✕ **No verdicts.** No record says whether it is enough.
- ✕ **No classification** of your systems or of your role.
- ✕ **No conformity work** for your systems: no technical file, CE marking or Art. 50 notices.
- ✕ **No FRIA, DPIA or incident filing.** We supply the facts; you assess and report.
- ✕ **Only what passes through it.** Work that never reaches a governed tool, or credentials that bypass Preloop, stay out of view.
- ✕ **No WORM storage of its own.** Logs become immutable once exported to your object-lock storage.

Name the edition your assessment relies on

CAPABILITY	OPEN SOURCE	ENTERPRISE
Allow, deny or require approval before any governed call	Yes	Yes
Human approvals; runs wait up to 30 days for an answer	Yes	Yes
Retention, legal hold, hash chain, signed exports, records console	Yes	Yes
Audit collection, query API and console	Yes	Yes
Multiple users, invitations and teams	Yes	Yes
Member offboarding that revokes all of a person's access in one audited step	Yes	Yes
Roles and RBAC	No	Yes
Single sign-on (OIDC, SAML 2.0), SCIM provisioning, SSO enforcement with audited owner break-glass	No	Yes
Multi-approver routing, quorum, escalations	No	Yes
Persisted policy denials; budgets per user or team	No	Yes

Preloop Cloud, our hosted service, has the Enterprise capabilities. A capability in one edition is not evidence in another.

05 — DATA PROTECTION AND SECURITY

Built to be checked, not just trusted.

Your data

Roles
On Preloop Cloud we process your records as your processor, with a data processing agreement signed at contract, and act as controller only for account and billing data. Self-hosted, we never receive your records.

Location
Preloop Cloud is hosted in Germany. Or run it yourself with Docker Compose or Helm, in your VPC or on-premises.

Subprocessors
Providers of hosting, flow execution, backup storage, email and push notifications, and the model providers you choose when you use hosted model credentials. The current list is in the hosting statement.

Zero retention by Preloop, self-hosted
Self-hosted, Preloop retains none of your prompts, responses or records. And Preloop trains no AI models, on your data or anyone else's.

Minimisation
Named fields are redacted before they are stored or sent in a notification, and the same rules can redact rows stored earlier. Retention is set per record class; deletion is off by default and is itself recorded.

Security and lifecycle

Credentials
Encrypted at rest or held in a Vault-compatible store. Agents receive gateway tokens, never provider keys.

Access
MCP clients connect over OAuth 2.1 with PKCE. Enterprise single sign-on with OIDC or SAML 2.0 uses your identity provider and its second factor; SCIM provisioning creates and removes members from the identity provider.

Supply chain
Releases carry three CycloneDX SBOMs derived from the build, a signed in-toto provenance attestation and an OpenVEX statement. Every change is scanned for vulnerabilities and secrets.

Support period
Security updates until 31 December 2031. Report issues privately to security@preloop.ai or via GitHub.

Open source
The code is public under Apache 2.0, so your engineers can read what governs your agents.

Certification
None yet. We do not claim ISO/IEC 27001, SOC 2 or ISO/IEC 42001.

Worth knowing
Agent logs name the people who approve, so they are personal data. Give them a retention period and a lawful basis.

PROOF POINT

We run Preloop on Preloop.

In September 2026 we pointed our own release audit at our own code. It failed: our SBOMs had been built from manifests, not from the build. A severity gate then paused on an unscored advisory and asked a person, who decided on the record, with a reason. We fixed the findings in public source.

The signed evidence pack can be checked by anyone, without asking us.

06 — HOW THE BRIEF DOCUMENTS ALIGNMENT

The functions behind the mapping.

All statements here are in the Preloop 1.0.0 code or changelog; none is planned work.

01

Rules before the action

Arts. 9, 15, 26(1)

Every tool call through the MCP firewall and every model call through the gateway is checked against policy before it runs. Policies are written by people in YAML with CEL, versioned and evaluated outside the model, so a prompt cannot change them. Each allow, deny or require-approval decision is recorded with the rule that produced it. Redaction rules strip named fields before storage, and can now be applied to older rows in audited batches.

Limit — Prompt-injection protection is partial: Preloop governs what an agent may do, not what it is told.

02

A person decides consequential steps

Arts. 4, 14, 26(2)

A `require_approval` rule pauses the action until a named person decides, with the tool, arguments and matching rule in view, on the console, mobile app, Slack, Mattermost, email or CLI. Long waits park the run without holding compute, for up to 30 days; an expired window resumes the agent with an "expired" answer on the record. Optional AI review is labelled as an AI decision and falls back to a person on low confidence or failure.

Enterprise — Multi-approver routing, quorum and escalation, so a team can hold the decision.

You own — Which actions need a person; approvers with the authority and training to decide.

03

Stop and intervene

Arts. 14(4)(e), 26(5)

A kill switch, for flows or for tools, blocks model and tool traffic, ends live agent sessions, refuses new ones and freezes pending approvals; activation and recovery are audited. Operator notes steer a running agent mid-task. On a developer's machine the approval hook fails closed when the control plane cannot be reached.

Limit — An operator may pause the hooks locally for up to 24 hours (one hour by default). While paused, Preloop neither governs nor records that machine.

04

Records that can be checked

Arts. 12, 19, 26(6)

Every governed tool call, decision, approval and model request is logged with a correlation id, the agent, the model and the person involved. Rows are sealed in a per-account hash chain, so an edit, deletion or reordering breaks every later hash. Exports and evidence packs are signed with a per-account Ed25519 key whose public half is published, so anyone can verify them without Preloop. Retention per record class has a 183-day floor and a 365-day default, with legal hold. Redaction for GDPR erasure keeps the chain verifiable and shows which rows were redacted.

Limit — The chain proves integrity from sealing, not that a record was true when written. No WORM storage of its own. Bypass is detected for managed runs, not on a developer's laptop.

06 — HOW THE BRIEF DOCUMENTS ALIGNMENT, CONTINUED

05

Attribution and transparency

Art. 50

AI decisions and human decisions are recorded apart and shown apart. Every action is attributed to a named agent, with the model that handled the request on record. Upstream MCP servers receive a Preloop User-Agent header and a correlation id matching the audit rows, so their operators can trace a call back to its record.

06

Identity, access and offboarding

Art. 26; NIS2 Art. 21(2)

In every edition, removing a member revokes that person's sessions, tokens, CLI logins, MCP OAuth tokens, personal API keys and registered runners in one audited transaction. Model credentials rotate in place, audited without the key. Hosted agents run as an unprivileged user and the approval hook is owned by root, so an agent cannot remove the control that governs it.

Enterprise — Single sign-on with OIDC or SAML 2.0, verified email domains, identity-provider groups mapped to roles and teams, SCIM provisioning, and enforced SSO with audited owner break-glass. Preloop Cloud has the same.

07

Monitoring and incidents

Arts. 26(5), 72, 73

Session replay steps through every model request and response and every tool call that led to an action, with the policy decision, approvals, spend and outcome. An incident-candidates export lists failed executions, kill-switch activations, budget denials and gateway upstream failures for a period, with timestamps and correlation ids. A flow alerts its owners after a set number of consecutive failures.

Enterprise — Persisted policy denials appear in the export.

You own — Deciding what is a serious incident, and reporting it. The export lists candidates only.

08

Preloop's own product

Art. 15; Cyber Resilience Act

Releases carry three CycloneDX SBOMs derived from the build (backend, CLI, frontend), a signed in-toto provenance attestation, an OpenVEX statement and checksums; release v0.17.1 of 21 October 2026 is a public example. Every change is scanned for vulnerabilities and secrets. Vulnerabilities are reported privately, security updates are committed until 31 December 2031, and the code is public under Apache 2.0.

Preloop produces records. Your counsel decides what they prove.

Preloop does not decide whether a record satisfies an article. It gives no classification, technical file, CE marking, Art. 50 notice, FRIA, DPIA or incident filing for your systems. Features differ by edition: the table on page 5 names which.

07 — NEXT STEPS

Start with one consequential action.

01
Pick one real workflow and name its owner.

02
Write the rule: what the agent may do, and what waits for a person.

03
Test the allowed, denied, expired and interrupted paths.

04
Ask someone who did not watch to rebuild the outcome from the records.

Beyond the AI Act

GDPR
Redaction, retention per record class, legal hold and a data processing agreement signed at contract. EU-based hosting.

DORA
Exports feed your Art. 8 inventory and Art. 28 register, and list Art. 17 incident candidates. Classification stays yours.

NIS2
Approvals on agent paths into production support your Art. 21(2) measures. Preloop's own development applies Art. 21(2)(d)–(e): supply-chain security, secure development, vulnerability disclosure.

Cyber Resilience Act
Presets verify the SBOM your build produced. For Preloop itself: a dated support period and Art. 14 reporting runbooks.

Available to your counsel and procurement

- Edition and feature matrix
- A sample signed evidence pack
- Hosting statement
- Security policy and support period

Preloop

See them, govern them, cut their cost.

preloop.ai
docs.preloop.ai
github.com/preloop/preloop

Preloop engaged **Law & Tech, Attorneys at Law** (www.lawandtech.eu) to review the legal content of this brief.

This brief sets out Preloop's own assessment of its product under Regulation (EU) 2024/1689, as amended by Regulation (EU) 2026/1744, as at November 2026. It is general information, not legal advice, a certification or a conformity assessment, and it does not state that any record satisfies a legal requirement. Your obligations depend on your role, use case and sector: consult your own counsel. Features differ by edition and version. This brief describes Preloop 1.0.0. © 2026 Preloop.