OntoGuard Layer Facts

Positioning: OntoGuard is AI infrastructure: a Semantic Control Plane / Decision Authorization layer for evaluating AI outputs before they become business consequence.

Movement Evaluated

OntoGuard evaluates an AI output or proposed release moving toward business action.

Examples include customer-facing responses, claims rationales, policy interpretations, reviewer recommendations, workflow decisions, compliance-sensitive summaries, eligibility language, denial rationales, financial explanations, and other high-stakes AI-generated outputs.

Input Accepted

OntoGuard can evaluate prompts, responses, AI logs, proposed outputs, workflow context, risk signals, evidence references, policy/citation context where available, and runtime metadata.

Initial value can begin from AI outputs/logs alone, without requiring internal document upload.

Output Produced

OntoGuard returns a governed decision state:

ALLOW / BLOCK / ESCALATE

The output includes release authorization state, release status, routed-to value, reason codes, semantic scope, evidence posture, uncertainty state, risk posture, audit identifiers, and proof-packet artifacts.

What It Allows

OntoGuard allows outputs that are sufficiently supported, scoped, low-risk, and appropriate for the intended release path.

Allowed outputs can move forward with recorded authorization, audit identifiers, and packetized proof.

What It Blocks

OntoGuard blocks outputs that are unsafe, unsupported, out of scope, misleading, hallucination-prone, policy-inconsistent, or not appropriate for release.

Blocked outputs do not proceed autonomously into business consequence.

What It Escalates

OntoGuard escalates plausible outputs that may be useful but are not strong enough for autonomous release.

Escalation is triggered by evidence gaps, uncertainty, weak policy or citation support, high downstream reliance, sensitive workflow context, benchmark failure, ambiguity, or need for expert review.

Receipt Produced

OntoGuard produces portable proof artifacts, including:

- Decision Authorization Packet
- Buyer governance JSON
- Governance PDF
- Decision receipt JSON
- Artifact manifest JSON
- Packaged governance ZIP
- Refusal / escalation receipt where release is withheld

Replay / Proof Surface

The public proof surface shows a synthetic high-stakes AI output evaluated before release, escalated to human review, and packaged into portable proof artifacts.

The proof surface demonstrates:

plausible output → OntoGuard evaluation → ESCALATE / release withheld → receipt created → packet exported

Deeper replay and runtime walkthroughs can be shown in controlled diligence settings without exposing protected implementation logic.

Changed-Condition Behavior

If conditions change, OntoGuard can produce a new governed decision state.

Changed conditions may include stronger evidence, updated policy scope, improved citation support, reduced risk, human-review resolution, new workflow context, or changed release destination.

The new decision can be re-evaluated and packetized as a separate governed state.

Protected Internals Not Disclosed

The public layer facts and proof kit do not disclose protected OntoGuard runtime internals.

Not disclosed:

source code, semantic routing internals, scoring formulas, BM25 or retrieval logic, ontology mappings, ranking methods, citation infrastructure, benchmark internals, private configuration, proprietary thresholds, customer data, or implementation file structure.

Boundary statement: The proof surface is public. The protected runtime remains protected.