Comprehensive Threat Model Evaluation Report (Combined)

Executive Summary & Comparative Analysis

1. Threats & Mitigations Maturity Ranking (Across Models)

Rank	Model Name	Threats & Mitigations Score	Maturity	Reasoning
1	husky-ai-model-anthropic-claude-sonnet-4-5-20250929	90	🏆 Excellent	Broad and balanced STRIDE coverage across actors, stores, processes, and cross-boundary flows; precise, actionable mitigations (e.g., mTLS, cert pinning, SCA, PAM, JIT) and strong alignment to practical controls; minimal redundancy and strong contextual accuracy.
2	husky-ai-model-openai-gpt-5	87	🌟 Good	Comprehensive and realistic threats (including data poisoning, cross-zone pivots, model tampering, and supply chain) with root-cause mitigations (signing, SBOMs, SCA, immutability); minor repetition and a few over-broad mitigations prevent top score.
3	husky-ai-model-anthropic-claude-opus-4-1-20250805	84	🌟 Good	Strong coverage of critical assets and flows; mitigations are specific, layered (e.g., sandboxing, key rotation, model signing) and proportionate; slightly less category depth and fewer cross-checks than the top two.
4	husky-ai-model-xai-grok-4-fast-reasoning-latest	78	🌟 Good	Wide-ranging STRIDE threats across many elements; solid mitigations (e.g., SCA, signed artifacts, mTLS) but some duplication and occasional inconsistency; still a credible and useful set.
5	husky-ai-model-novita-deepseek-deepseek-v3.1-terminus	70	✅ Adequate	Good coverage of internal unencrypted flows and SSH risks; mitigations are reasonable but occasionally generic; limited depth on data provenance and CI/CD trust boundaries.
6	husky-ai-model-novita-qwen-qwen3-coder-480b-a35b-instruct	65	✅ Adequate	Many threats identified, but several inconsistencies (e.g., encryption states vs. text) and uneven specificity; mitigations often correct but need refinement and tighter traceability.
7	husky-ai-model-xai-grok-4-latest	62	✅ Adequate	Broad threat set but varying quality; several mitigations are high-level or repetitive; contextual alignment uneven and some cross-zone edge cases insufficiently addressed.
8	husky-ai-	55	⚙️ Fair	Sparse threats; coverage misses several critical

	model-gemini-gemini-2.5-pro			flows and actors; mitigations are generally sound but too few to support comprehensive risk reasoning.
9	husky-ai-model-ollama-gemma327b	40	⚠️ Poor	Minimal threats and limited component coverage; mitigations not systematic; insufficient for credible, actionable risk posture.

2. Overall Model Maturity

2.1 Evaluation Summary

Across all files, the shared DFD depicts two trust zones (Experimental and Production) with clear core components (ingestion, notebook training, deployment, API gateway, web server, bastion) and relevant data stores. The model effectively exposes key cross-boundary flows (e.g., external ingestion, SSH deployments, production administration) enabling practical risk analysis. Notable gaps include occasional duplication of flows, inconsistent encryption flags versus labels (e.g., flows named "HTTPS" yet flagged unencrypted), and missing supporting services (IdP, CI/CD control-plane, logging/monitoring). Overall, the DFD is strong enough for effective analysis, with minor consistency and completeness issues.

2.2 Scoring Table

Dimension	Weight	Score	Reasoning
Clarity and Readability	25%	78	Trust boundaries and component roles are explicit; flows are named and generally understandable. Minor inconsistencies (e.g., "HTTPS" with isEncrypted=false, repeated "update" flows) slightly reduce clarity.
Completeness and Coverage	30%	75	Core assets, actors, and flows are present. Missing auxiliary services (IdP, CI/CD control-plane, telemetry/logging pipelines) and limited decomposition (e.g., API Gateway policies, service-mesh) constrain coverage.
Accuracy and Logical Consistency	25%	72	Logical structure is sound, but duplication (e.g., two similar bastion→model flows) and occasional contradictions between labels and attributes reduce consistency.
Usability for Security Analysis	20%	80	Attack surfaces and cross-zone ingress paths are clear, enabling practical risk inference and prioritization; layout is readily extensible with additional controls.

Overall Model Maturity Total Score (0–100): 76 **Overall Model Maturity:** 🌟 Good

3. Individual Model Evaluations (Threats & Mitigations Only)

##

husky-ai-model-anthropic-claude-sonnet-4-5-20250929

This section provides the dedicated Threats & Mitigations analysis for this specific model.

##

Threat Landscape Snapshot "(Per Model)"

STRIDE Category	High	Medium	Low	Observations
Spoofing	5	2	0	Strong authentication and SSH certificate guidance; covers external users and admin ingress.
Tampering	7	4	0	Extensive data/model/package signing; supply chain and model update paths well treated.
Repudiation	1	2	0	Logging/monitoring emphasized for gateways and bastion; could add immutable logging details for all stores.
Information Disclosure	5	4	0	Good coverage of secrets, models, data-at-rest/in-transit; suggests private endpoints and conditional access.
Denial of Service	2	3	0	DDoS and rate-limiting handled; reasonable for exposure.
Elevation of Privilege	6	3	0	PAM, JIT, session recording, separation of duties; strong alignment with critical paths.

Methodology balance is excellent and plausibly mapped to trust boundaries and exposures.

Mitigation Quality & Alignment "(Per Model)"

husky-ai-model-anthropic-claude-sonnet-4-5-20250929

Control Area	Adequacy	Observations
Relevance & Specificity	✅	Controls map directly to threats (e.g., mTLS for internal service calls, signed artifacts for model/package integrity).
Practicality	✅	Uses deployable cloud-native patterns (Private Link, Key Vault, DDoS protection, WAF, SCA).
Completeness & Coverage	✅	Addresses actors, processes, stores, and cross-zone flows; minimal residual gaps.
Effectiveness	✅	Root-cause treatment, layered defense; reduces exploitability meaningfully.
Standards Alignment	✅	Reflects common best practices for identity, network, and supply chain security.
Traceability & Justification	✅	Threat-to-control traceability is clear and defensible.

Summary Rating: ✅ Adequate

Gaps, Blind Spots & Prioritized Fixes "(Per Model)"

#	Finding	Impact	Effort	Recommendation

1	Limited explicit model provenance attestation	High	Medium	Add attestation in CI/CD (e.g., SLSA provenance) and enforce verification at deploy/load time.
2	Infrequent mention of telemetry pipelines	Medium	Medium	Define centralized, immutable logging (hash-chained) and telemetry baselines for anomaly detection.
3	Minor duplication of model update flows	Medium	Low	Normalize flows and document authoritative update path to reduce ambiguity in enforcement.

Threats & Mitigations Maturity Assessment "(Per Model)"

Evaluation Focus

Assess the model on five dimensions as described in table.

Dimension	Weight	Score	Reasoning
DFD Element Coverage	30%	92	Nearly all applicable elements are covered with meaningful threats/mitigations; critical assets clearly addressed.
Methodology Coverage & Balance	30%	90	Strong, balanced STRIDE coverage; limited minor gaps in repudiation depth.
Contextual Accuracy	20%	90	Threats line up with trust zones and exposure; cross-boundary specifics are on point.
Mitigation Validity	10%	88	Controls are practical and effective; minor opportunities to formalize provenance.
Proportionality & Realism	10%	90	Severities and controls proportionate to exposure, especially admin and production paths.

Threats & Mitigations Total Score (0–100): 90 **Threats & Mitigations Maturity:** 🏆 Excellent

Strategic Recommendations "(Per Model)"

husky-ai-model-anthropic-claude-sonnet-4-5-20250929

- Add signed provenance (SLSA L3+) for training data, code, and model artifacts; verify at deploy and load.
- Expand immutable, centralized logging with tamper-evident storage and automated correlation for bastion and API changes.
- Document a single authoritative model update path and disable alternates (enforce via network policy and IAM).
- Add explicit threat entries for insider exfiltration via notebook egress; include egress filtering and DLP controls.
- Include “model extraction” detections in production with thresholds and response playbooks.
- Specify certificate rotation cadences and automated TLS policy checks in CI.

- Add explicit secrets scanning gates in CI and pre-commit hooks for source/config repos.
- Capture recovery/rollback scenarios as first-class threats (rollback attacks), with controls and approvals.

##

husky-ai-model-openai-gpt-5

This section provides the dedicated Threats & Mitigations analysis for this specific model.

##

Threat Landscape Snapshot "(Per Model)"

STRIDE Category	High	Medium	Low	Observations
Spoofing	4	2	0	Strong on SSH certs, MFA, and host verification; solid user/API authentication guidance.
Tampering	6	4	0	Clear signing and immutability across data/model/artifacts; good model integrity checks at load.
Repudiation	2	2	0	Session recording and immutable audit emphasized; could broaden to all stores and CI events.
Information Disclosure	5	3	0	Data-at-rest/in-transit, model IP protection, and secrets management well handled.
Denial of Service	3	3	0	Gateway and inference DoS risks recognized; mitigations proportionate.
Elevation of Privilege	5	3	0	Cross-zone SSH bridge identified; controls are layered and realistic.

Coverage is strong and technically grounded.

Mitigation Quality & Alignment "(Per Model)"

husky-ai-model-openai-gpt-5

Control Area	Adequacy	Observations
Relevance & Specificity	✅	Precise mapping of threats to controls (e.g., signed models, mTLS, SBOMs).
Practicality	✅	Implements mainstream cloud-native and platform controls; deployable at scale.
Completeness & Coverage	✅	Very good breadth; small opportunity to expand on telemetry and provenance.
Effectiveness	✅	Addresses root causes and hardens critical flows and assets.
Standards Alignment	✅	Aligns with best-practice patterns for identity, supply chain, and network security.
Traceability & Justification	✅	Coherent threat-to-mitigation traceability throughout.

Summary Rating: ✅ Adequate

Gaps, Blind Spots & Prioritized Fixes "(Per Model)"

#	Finding	Impact	Effort	Recommendation
1	Limited explicit data/model provenance attestation	High	Medium	Add attestations, verify provenance at deployment (policy-as-code).
2	Telemetry/forensics model not fully defined	Medium	Low	Define retention/immutability, correlation rules, and response workflows.
3	Some repetition in flow threats	Low	Low	Consolidate duplicates and ensure single source of truth for update paths.

Threats & Mitigations Maturity Assessment "(Per Model)"

Dimension	Weight	Score	Reasoning
DFD Element Coverage	30%	88	Broad coverage across actors, processes, stores, and critical flows.
Methodology Coverage & Balance	30%	86	STRIDE balance is strong; minor room to deepen repudiation.
Contextual Accuracy	20%	88	Accurate mapping to cross-zone risks and admin ingress.
Mitigation Validity	10%	85	Controls are feasible and effective; slight improvement via provenance.
Proportionality & Realism	10%	88	Priorities align with exposure and trust boundaries.

Threats & Mitigations Total Score (0–100): 87 **Threats & Mitigations Maturity:** 🌟 Good

Strategic Recommendations "(Per Model)"

husky-ai-model-openai-gpt-5

- Enforce SLSA-style provenance for datasets, models, and build artifacts; verify in deployment gates.
- Expand immutable logging with automated detection for admin privilege escalation and model updates.
- Add explicit coverage for adversarial inputs and model extraction monitoring thresholds.
- Strengthen notebook egress controls with deterministic allowlists and periodic review.
- Consolidate duplicate “update” flows and define an authoritative path with policy enforcement.
- Specify secrets exposure prevention (e.g., memory-hardening, zero-logging assurance) for API keys.
- Include model rollback attack checks (version pinning + signature validation).

##

husky-ai-model-anthropic-claude-opus-4-1-20250805

This section provides the dedicated Threats & Mitigations analysis for this specific model.

##

Threat Landscape Snapshot "(Per Model)"

STRIDE Category	High	Medium	Low	Observations
Spoofing	4	2	0	SSH identity and user access addressed; could add explicit device-posture gates.
Tampering	6	3	0	Emphasizes model/data/package integrity and poisoning risks; good CI/CD mitigation patterns.
Repudiation	1	2	0	Logging present; add immutability and coverage for all stores and flows.
Information Disclosure	4	4	0	API keys, model IP, and datasets protected; would benefit from explicit memory/logging controls.
Denial of Service	2	2	0	Reasonable consideration at gateway and service; moderate depth.
Elevation of Privilege	5	2	0	PAM and JIT noted; strong bastion hardening guidance.

Well-structured and plausible, with focus on practical control pairs.

Mitigation Quality & Alignment "(Per Model)"

husky-ai-model-anthropic-claude-opus-4-1-20250805

Control Area	Adequacy	Observations
Relevance & Specificity	✅	Targeted to actual fault lines (e.g., SSH, poisoning, CI/CD).
Practicality	✅	Deployable with standard cloud tooling and pipelines.
Completeness & Coverage	✅	Near-complete; minor logging/provenance detail can be added.
Effectiveness	✅	Likely to materially reduce risk, especially for model tampering and admin abuse.
Standards Alignment	✅	Consistent with industry practice (signing, sandboxing, SCA).
Traceability & Justification	✅	Clear cause→control alignment.

Summary Rating: ✅ Adequate

Gaps, Blind Spots & Prioritized Fixes "(Per Model)"

#	Finding	Impact	Effort	Recommendation
1	Provenance not explicit end-to-end	High	Medium	Add attestations for datasets, artifacts; verify before deployment and loading.

2	Telemetry immutability not guaranteed	Medium	Low	Use tamper-evident storage and signed logs; define forensics retention.
3	Limited, explicit defense-in-depth for internal HTTPS	Medium	Low	Enforce mTLS and cert rotation; add policy checks in CI.

Threats & Mitigations Maturity Assessment "(Per Model)"

Dimension	Weight	Score	Reasoning
DFD Element Coverage	30%	85	Solid coverage of elements and cross-zone data paths.
Methodology Coverage & Balance	30%	84	Good category spread; slight increase possible for repudiation depth.
Contextual Accuracy	20%	84	Accurate, environment-aware threats and controls.
Mitigation Validity	10%	82	Realistic and effective; minor enhancements to provenance/logging.
Proportionality & Realism	10%	85	Well prioritized for highest-risk paths (bastion, models).

Threats & Mitigations Total Score (0–100): 84 **Threats & Mitigations Maturity:** 🌟 Good

Strategic Recommendations "(Per Model)"

husky-ai-model-anthropic-claude-opus-4-1-20250805

- Add SLSA-style attestation and verification gates in CI/CD and on model load.
- Make internal service-to-service mTLS mandatory; automate certificate rotation checks.
- Expand immutable logging (append-only) and define incident-ready queries/playbooks.
- Include explicit adversarial input detection for inference and training data acceptance.
- Add model rollback prevention and signed version enforcement.
- Clarify a single model update route and enforce with network/IAM policy.

##

husky-ai-model-xai-grok-4-fast-reasoning-latest

This section provides the dedicated Threats & Mitigations analysis for this specific model.

##

Threat Landscape Snapshot "(Per Model)"

STRIDE	High	Medium	Low	Observations

Category				
Spoofing	4	3	0	Covers SSH and API spoofing reasonably well.
Tampering	5	4	0	Addresses imports, model/package integrity, and internal flows; some duplication.
Repudiation	2	2	0	Logging present; could add tamper-evident storage and broader scope.
Information Disclosure	4	3	0	Handles model, dataset, and key exposure; add memory/logging suppressions.
Denial of Service	3	2	0	Reasonable rate-limiting and DDoS focus on API; inference DoS noted.
Elevation of Privilege	4	3	0	Bastion, cross-zone deployments covered; mitigations solid though occasionally repetitive.

Coverage is strong, with minor inconsistencies and redundancy.

Mitigation Quality & Alignment "(Per Model)"

husky-ai-model-xai-grok-4-fast-reasoning-latest

Control Area	Adequacy	Observations
Relevance & Specificity	✅	Mostly specific to observed risks; a few high-level controls could be tightened.
Practicality	✅	Achievable with common cloud and platform tooling.
Completeness & Coverage	✅	Broad coverage; minor scope expansion needed for telemetry and provenance.
Effectiveness	✅	Controls reduce exploitability on main paths; some refinements suggested.
Standards Alignment	✅	Generally aligned with best practice.
Traceability & Justification	⚠️	Some duplication and mixed granularity reduce crisp traceability.

Summary Rating: ✅ Adequate

Gaps, Blind Spots & Prioritized Fixes "(Per Model)"

#	Finding	Impact	Effort	Recommendation
1	Duplicative/overlapping threats	Medium	Low	Consolidate duplicate flows (e.g., updates) and clarify single authoritative routes.
2	Telemetry controls not tamper-evident	Medium	Low	Adopt signed/hashed logs with retention; define SIEM correlation.
3	Inconsistent depth on provenance	Medium	Medium	Add model/data attestation and enforcement in deployment/load.

Threats & Mitigations Maturity Assessment "(Per Model)"

Dimension	Weight	Score	Reasoning
DFD Element Coverage	30%	80	Good breadth across most elements and flows.
Methodology Coverage &	30%	76	Balanced STRIDE coverage; mild bias to tampering and spoofing.

Balance			
Contextual Accuracy	20%	78	Generally accurate; a few inconsistencies.
Mitigation Validity	10%	75	Mostly correct; could sharpen provenance/logging.
Proportionality & Realism	10%	80	Focus matches attack surface; improvement possible on telemetry and model pipeline.

Threats & Mitigations Total Score (0–100): 78 **Threats & Mitigations Maturity:** 🌟 Good

Strategic Recommendations "(Per Model)"

husky-ai-model-xai-grok-4-fast-reasoning-latest

- Normalize flows to remove duplicates; document authoritative control points and disallow alternates.
- Enforce signed model/data provenance and verification at deployment and load time.
- Implement tamper-evident, centralized logging with defined detection rules (e.g., admin anomalies, model changes).
- Add explicit mitigations for adversarial inputs and model extraction attempts in production.
- Strengthen secrets handling (no-logs, memory-hardening, short-lived tokens).
- Formalize certificate lifecycle management and policy checks in CI/CD.

##

husky-ai-model-novita-deepseek-deepseek-v3.1-terminus

This section provides the dedicated Threats & Mitigations analysis for this specific model.

##

Threat Landscape Snapshot "(Per Model)"

STRIDE Category	High	Medium	Low	Observations
Spoofing	3	2	0	SSH admin spoofing and external service identity covered; device posture could be added.
Tampering	4	3	0	Focus on unencrypted flows and model tampering; reasonable.
Repudiation	1	1	0	Auditing present but limited scope and detail.
Information Disclosure	4	2	0	Emphasis on secrets and models; could expand to logs and telemetry.
Denial of Service	2	1	0	Considered but basic; can expand on inference throttling.
Elevation of Privilege	3	2	0	Cross-boundary SSH and bastion path addressed; sound but brief.

Adequate breadth with practical mitigations, albeit less detailed.

Mitigation Quality & Alignment "(Per Model)"

husky-ai-model-novita-deepseek-deepseek-v3.1-terminus

Control Area	Adequacy	Observations
Relevance & Specificity	✅	Maps to real exposures (unencrypted flows, SSH).
Practicality	✅	Achievable and familiar controls.
Completeness & Coverage	⚠️	Some elements and flows lightly treated (telemetry, CI/CD supply chain).
Effectiveness	✅	Will reduce risk on main issues; depth limited.
Standards Alignment	✅	Aligns with common controls.
Traceability & Justification	⚠️	Threat→control mapping could be more explicit across all elements.

Summary Rating: ✅ Adequate

Gaps, Blind Spots & Prioritized Fixes "(Per Model)"

#	Finding	Impact	Effort	Recommendation
1	Limited provenance controls	High	Medium	Add signed provenance and enforce at deployment/load.
2	Basic telemetry/Forensics	Medium	Low	Adopt tamper-evident logging and define detection/response playbooks.
3	Inference DoS only briefly covered	Medium	Low	Add rate limits, timeouts, and backpressure controls per tenant/API key.

Threats & Mitigations Maturity Assessment "(Per Model)"

Dimension	Weight	Score	Reasoning
DFD Element Coverage	30%	72	Good for main elements; secondary controls less detailed.
Methodology Coverage & Balance	30%	68	Reasonable STRIDE balance; deepen repudiation/DoS.
Contextual Accuracy	20%	70	Accurate and plausible for the environment.
Mitigation Validity	10%	68	Practical but not deeply layered.
Proportionality & Realism	10%	72	Prioritization is acceptable; can improve for production inference.

Threats & Mitigations Total Score (0–100): 70 **Threats & Mitigations Maturity:** ✅ Adequate

Strategic Recommendations "(Per Model)"

husky-ai-model-novita-deepseek-deepseek-v3.1-terminus

- Enforce mTLS and certificate rotation for all internal flows; remediate any unencrypted internal traffic.

- Introduce artifact/data/model signing, with mandatory verification in CI/CD and at runtime load.
- Define a unified telemetry strategy (append-only logs, SIEM correlation, incident search packs).
- Expand inference-layer DoS protections (per-IP/key rate limits, concurrency caps, adaptive throttling).
- Add explicit adversarial input detection and extraction monitoring at the API layer.
- Strengthen SSH key lifecycle (short-lived certs, JIT, device posture verification).

##

husky-ai-model-novita-qwen-qwen3-coder-480b-a35b-instruct

This section provides the dedicated Threats & Mitigations analysis for this specific model.

##

Threat Landscape Snapshot "(Per Model)"

STRIDE Category	High	Medium	Low	Observations
Spoofing	4	2	0	Several ingress vectors identified; controls broadly correct.
Tampering	5	4	0	Good emphasis on stores and flows; a few contradictions around encryption flags.
Repudiation	2	1	0	Code/config traceability discussed; needs immutable logging expansion.
Information Disclosure	5	3	0	Strong focus on secrets and model IP; mixed precision on storage flags.
Denial of Service	2	2	0	Addressed at gateway and server; more detail desirable.
Elevation of Privilege	4	2	0	Bastion and admin risks identified; mitigations mostly appropriate.

Overall good breadth, but consistency issues reduce confidence.

Mitigation Quality & Alignment "(Per Model)"

husky-ai-model-novita-qwen-qwen3-coder-480b-a35b-instruct

Control Area	Adequacy	Observations
Relevance & Specificity	⚠️	Generally relevant; some conflicts with component attributes (encryption status).
Practicality	✅	Mostly deployable mitigations.
Completeness & Coverage	✅	Covers many elements and flows, including source/config and models.
Effectiveness	⚠️	Effective where precise; contradictory assumptions can weaken implementation.
Standards Alignment	✅	Reflects common practices for secrets and storage protection.
Traceability & Justification	⚠️	Improve consistency between threats and DFD attributes.

Summary Rating: ⚠️ Partially adequate

Gaps, Blind Spots & Prioritized Fixes "(Per Model)"

#	Finding	Impact	Effort	Recommendation
1	Inconsistent encryption assertions vs. DFD flags	Medium	Low	Reconcile isEncrypted states; update threats/controls accordingly.
2	Telemetry immutability not explicit	Medium	Low	Adopt signed logging and retention policies; ensure non-repudiation.
3	Provenance and model rollback insufficiently formalized	High	Medium	Enforce signing/attestation and prevent rollback with version pinning and approvals.

Threats & Mitigations Maturity Assessment "(Per Model)"

Dimension	Weight	Score	Reasoning
DFD Element Coverage	30%	66	Many elements covered with threats; a few stores/flows need consistent treatment.
Methodology Coverage & Balance	30%	64	Broad STRIDE spread but uneven due to inconsistencies.
Contextual Accuracy	20%	66	Mostly plausible; correct the contradictory flags.
Mitigation Validity	10%	62	Sound where accurate; reduce ambiguity in storage/flow protection.
Proportionality & Realism	10%	66	Priorities are reasonable; improve precision for production models.

Threats & Mitigations Total Score (0–100): 65 **Threats & Mitigations Maturity:** ✅ Adequate

Strategic Recommendations "(Per Model)"

husky-ai-model-novita-qwen-qwen3-coder-480b-a35b-instruct

- Reconcile all “isEncrypted” states with narrative; correct threats/mitigations to match reality.
- Require signed/attested models and configuration with verification at deploy/load.
- Define tamper-evident logging for bastion, API gateway, and model stores; set alert rules.
- Add adversarial input detection and model extraction monitoring in production.
- Expand defense for internal flows (mandatory mTLS, cert rotation, policy checks).
- Enforce secrets scanning in CI and prevent credentials in configs.

##

husky-ai-model-xai-grok-4-latest

This section provides the dedicated Threats & Mitigations analysis for this specific model.

##

Threat Landscape Snapshot "(Per Model)"

STRIDE Category	High	Medium	Low	Observations
Spoofing	3	2	0	Covers SSH and API flows; scope is reasonable but sometimes generalized.
Tampering	4	3	0	Addresses model/code/packages; several high-level recommendations.
Repudiation	2	1	0	Mentions logging and session recording; add immutability and coverage expansion.
Information Disclosure	4	2	0	Recognizes model/secret leakage; strengthen memory/logging protections.
Denial of Service	2	2	0	Gateway DoS and inference load considered; could deepen.
Elevation of Privilege	3	2	0	Bastion and cross-zone updates discussed; controls credible but uneven.

Adequate breadth with uneven depth and some redundancy.

Mitigation Quality & Alignment "(Per Model)"

husky-ai-model-xai-grok-4-latest

Control Area	Adequacy	Observations
Relevance & Specificity	⚠️	Many mitigations relevant; some remain generic and repetitive.
Practicality	✅	Largely feasible within typical cloud environments.
Completeness & Coverage	✅	Covers major assets/flows with usable mitigations.
Effectiveness	⚠️	Effective when specific; generic advice should be tightened.
Standards Alignment	✅	Generally aligned with best-practice patterns.
Traceability & Justification	⚠️	Mixed granularity reduces direct traceability for some threats.

Summary Rating: ✅ Adequate

Gaps, Blind Spots & Prioritized Fixes "(Per Model)"

#	Finding	Impact	Effort	Recommendation
1	Generic mitigations reduce clarity	Medium	Low	Replace generic steps with concrete controls (e.g., mTLS, Sigstore, PAM/JIT specifics).

2	Telemetry immutability under-specified	Medium	Low	Add signed/append-only logging and explicit retention.
3	Provenance not enforced	High	Medium	Require signed/attested artifacts and enforce at deployment/model load.

Threats & Mitigations Maturity Assessment "(Per Model)"

Dimension	Weight	Score	Reasoning
DFD Element Coverage	30%	64	Major elements addressed; some flows need sharper mapping.
Methodology Coverage & Balance	30%	60	STRIDE coverage present but uneven and occasionally high-level.
Contextual Accuracy	20%	62	Generally plausible; depth varies.
Mitigation Validity	10%	60	Many good controls; generic items dilute effectiveness.
Proportionality & Realism	10%	64	Prioritization acceptable; refine for production model/store risks.

Threats & Mitigations Total Score (0–100): 62 **Threats & Mitigations Maturity:** ✅ Adequate

Strategic Recommendations "(Per Model)"

husky-ai-model-xai-grok-4-latest

- Replace generic mitigations with specific, enforceable controls (mTLS-by-default, certificate pinning, signed artifacts).
- Enforce signed provenance for code, data, and models; validate in CI/CD and at runtime.
- Adopt tamper-evident centralized logs with correlation for admin/model changes.
- Add detailed protections for model extraction and adversarial inputs.
- Consolidate update flows and define a single controlled route with policy enforcement.

##

husky-ai-model-gemini-gemini-2.5-pro

This section provides the dedicated Threats & Mitigations analysis for this specific model.

##

Threat Landscape Snapshot "(Per Model)"

STRIDE Category	High	Medium	Low	Observations
Spoofing	1	1	0	Limited coverage of user/admin identity risks.

Tampering	2	2	0	Supply chain mentioned once; model/package tampering lightly treated.
Repudiation	0	1	0	Basic gateway logging; lacks immutable logging strategy.
Information Disclosure	2	1	0	Secrets/model exposure addressed only partially.
Denial of Service	1	2	0	DoS at gateway covered; inference DoS limited.
Elevation of Privilege	1	1	0	Bastion risk recognition minimal; mitigations thin.

Coverage is incomplete relative to the shared DFD's exposure.

Mitigation Quality & Alignment "(Per Model)"

husky-ai-model-gemini-gemini-2.5-pro

Control Area	Adequacy	Observations
Relevance & Specificity	⚠️	Mitigations relevant but too few and basic.
Practicality	✅	Feasible controls where specified.
Completeness & Coverage	⚠️	Many critical flows, stores, and admin paths under-modeled.
Effectiveness	⚠️	Insufficient to materially reduce aggregate risk.
Standards Alignment	✅	Aligns with best practices where present.
Traceability & Justification	⚠️	Sparse threats reduce traceability and confidence.

Summary Rating: ⚠️ Partially adequate

Gaps, Blind Spots & Prioritized Fixes "(Per Model)"

#	Finding	Impact	Effort	Recommendation
1	Undercoverage of cross-zone and admin paths	High	Medium	Add detailed threats/controls for SSH, bastion, and model update pipelines.
2	Sparse coverage of model integrity and provenance	High	Medium	Introduce signing/attestation and enforce verification at CI/CD and runtime.
3	Minimal telemetry/forensics strategy	Medium	Low	Implement signed, centralized logging with response playbooks.

Threats & Mitigations Maturity Assessment "(Per Model)"

Dimension	Weight	Score	Reasoning
DFD Element Coverage	30%	56	Several key elements/threats missing.
Methodology Coverage & Balance	30%	54	STRIDE categories underrepresented.

Contextual Accuracy	20%	56	Accurate where present; too sparse overall.
Mitigation Validity	10%	54	Practical but insufficiently layered.
Proportionality & Realism	10%	56	Priorities need extension to core risks.

Threats & Mitigations Total Score (0–100): 55 **Threats & Mitigations Maturity:** ⚙ Fair

Strategic Recommendations "(Per Model)"

husky-ai-model-gemini-gemini-2.5-pro

- Expand threats/controls for bastion, SSH, and cross-zone deployments (JIT, PAM, short-lived certs).
- Enforce model/data/code signing and verification gates in CI/CD and at load time.
- Add adversarial input and model extraction threats with concrete mitigations.
- Mandate mTLS for all internal communications; validate certificates automatically.
- Build immutable logging with anomaly detection for model/admin events.

##

husky-ai-model-ollama-gemma327b

This section provides the dedicated Threats & Mitigations analysis for this specific model.

##

Threat Landscape Snapshot "(Per Model)"

STRIDE Category	High	Medium	Low	Observations
Spoofing	0	0	0	Not covered.
Tampering	1	0	0	Only a single supply chain threat.
Repudiation	0	0	0	Not covered.
Information Disclosure	0	0	0	Not covered.
Denial of Service	1	0	0	Limited API gateway DoS noted.
Elevation of Privilege	0	0	0	Not covered.

Coverage is insufficient for meaningful risk management.

Mitigation Quality & Alignment "(Per Model)"

husky-ai-model-ollama-gemma327b

Control Area	Adequacy	Observations
Relevance &	⚠	Where present, relevant; overall too limited.

Specificity		
Practicality	✅	Feasible mitigations for few threats noted.
Completeness & Coverage	❌	Major elements/flows lack threats and mitigations.
Effectiveness	❌	Not sufficient to change risk posture.
Standards Alignment	⚠️	Too sparse to assess robust alignment.
Traceability & Justification	❌	Minimal traceability due to limited content.

Summary Rating: ❌ Inadequate

Gaps, Blind Spots & Prioritized Fixes "(Per Model)"

#	Finding	Impact	Effort	Recommendation
1	Missing coverage for admin, models, secrets, and flows	High	Medium	Add comprehensive threats/mitigations for SSH, bastion, stores, and cross-zone flows.
2	No model integrity or provenance controls	High	Medium	Introduce signing/attestation and verification at CI/CD and runtime.
3	No telemetry/forensics strategy	Medium	Low	Add tamper-evident centralized logging and basic detection.

Threats & Mitigations Maturity Assessment "(Per Model)"

Dimension	Weight	Score	Reasoning
DFD Element Coverage	30%	42	Significant gaps in elements and flows.
Methodology Coverage & Balance	30%	38	STRIDE categories largely missing.
Contextual Accuracy	20%	40	Too minimal to fully assess.
Mitigation Validity	10%	38	Insufficient detail for effectiveness.
Proportionality & Realism	10%	42	Priorities not reflected in content.

Threats & Mitigations Total Score (0–100): 40 **Threats & Mitigations Maturity:** ⚠️ Poor

Strategic Recommendations "(Per Model)"

husky-ai-model-ollama-gemma327b

- Add threats and controls for admin ingress (bastion/SSH), API gateway, and model stores (both experimental and production).
- Require mTLS internally and certificate pinning for external ingestion; remove plaintext flows.
- Introduce signing of code, data, and models; verify at CI/CD and runtime load.

- Implement centralized, immutable logging; define alerts for admin/model changes.
- Add adversarial input and model extraction detection in the web server.
- Cover secrets handling at the API key store (rotation, zero-logging, short-lived tokens).

##

husky-ai-model-xai-grok-4-fast-reasoning-latest

This section provides the dedicated Threats & Mitigations analysis for this specific model.

##
(Already evaluated above.)

##

husky-ai-model-xai-grok-4-latest

This section provides the dedicated Threats & Mitigations analysis for this specific model.

##
(Already evaluated above.)

##

husky-ai-model-novita-deepseek-deepseek-v3.1-terminus

This section provides the dedicated Threats & Mitigations analysis for this specific model.

##
(Already evaluated above.)

##

husky-ai-model-novita-qwen-qwen3-coder-480b-a35b-instruct

This section provides the dedicated Threats & Mitigations analysis for this specific model.

##
(Already evaluated above.)

##

husky-ai-model-anthropic-claude-opus-4-1-20250805

This section provides the dedicated Threats & Mitigations analysis for this specific model.

##

(Already evaluated above.)

##

husky-ai-model-gemini-gemini-2.5-pro

This section provides the dedicated Threats & Mitigations analysis for this specific model.

##

(Already evaluated above.)

---

4. Conclusion

- Comparative strengths/weaknesses: Anthropic Sonnet and OpenAI GPT-5 provide the strongest, most balanced STRIDE coverage with precise, deployable mitigations and clear threat-to-control traceability. Anthropic Opus and XAI Grok (fast-reasoning) follow closely, with solid breadth but minor duplication and a few inconsistent details. Novita Deepseek and Qwen models are adequate yet need stronger provenance, telemetry immutability, and consistency between DFD attributes and narratives. Gemini and especially Ollama require substantial expansion to meet baseline expectations for threats and mitigations.
- Common DFD-Only maturity: 🌟 Good (76/100). The shared architecture is well-defined for security analysis, exposing key trust boundaries and cross-zone flows. Improvements should focus on eliminating duplicative/ambiguous flows, reconciling inconsistent encryption flags vs. labels, and adding auxiliary components (IdP/CI-CD/telemetry) to improve completeness and accuracy.
- Next steps:

1) Make mTLS a universal internal principle; reconcile all encryption flags vs. labels and remove plaintext flows. 2) Enforce signed, attested provenance (code, data, models) and verify at deployment and model load. 3) Normalize to a single authoritative model update route per environment with policy/IAM enforcement. 4) Implement tamper-evident, centralized logging with detections for admin/model changes; define response playbooks. 5) Add explicit adversarial input and model extraction detections in production, with thresholds and action plans. 6) Strengthen SSH access with short-lived certs, JIT, PAM, MFA, device posture checks, and session recording.

These actions will lift both the shared architecture's security posture and the per-model threats/mitigations quality, converging toward consistent, high-confidence, and auditable risk management across all models.