Version: November 2026

GEM CORTEX AI

POLICY-AS-CODE UND AI GATEWAY

Kaufmännische und technische Produktspezifikation

Policy-as-Code und AI Gateway bilden eine gemeinsame Kontrollebene für KI-Agenten und Anwendungen im Rahmen von GEM Cortex AI. Policy-as-Code überführt genehmigte Regeln in eine Form, die versioniert, getestet und zur Laufzeit ausgewertet werden kann. AI Gateway und angebundene Kontrollpunkte setzen diese Entscheidungen durch. Im vereinbarten Umfang steuert die Organisation damit, welche Daten KI verwenden darf, wohin sie diese übermitteln und welche Aktionen sie ausführen darf.

Gemeinsame Regeln haben einen eigenen, von einzelnen Agenten unabhängigen Lebenszyklus. Änderungen an Grenzwerten oder Berechtigungen werden an einer Stelle verwaltet und in die angebundenen Anwendungen übernommen. Jede Entscheidung dokumentiert die verwendete Version, die Begründung und den Bezug zum tatsächlichen Ergebnis.

Das Angebot verbindet Plattformkomponenten, Integrationen, die Formalisierung von Regeln und Leistungen von GEM für Einführung und Betrieb. Die konkreten Technologien, angebundenen Systeme, Modelle, der Autonomiegrad und die Serviceparameter werden im Lösungskonzept und im vertraglichen Leistungsumfang bestätigt. Die in dieser Spezifikation beschriebene technische Variante ist ein Referenzentwurf, der für die konkrete Bereitstellung validiert wird.

1. *Zusammenfassung für die Geschäftsleitung und Einordnung in das Angebot von GEM*

Die Lösung richtet sich an Organisationen, die KI in operativen Prozessen ausbauen und zugleich die Verantwortung für Zugriffe, Genehmigungen, Kosten und Ergebnisse wahren möchten. Die gemeinsame Kontrollebene verbindet die technische Steuerung mit den Verantwortlichen für die Geschäftsregeln.
Die Regel wird von der fachlich zuständigen Person für den Prozess, die Sicherheit oder die Compliance festgelegt. Das technische Team übernimmt die Überführung in eine ausführbare Form, die Tests und die Anbindung an Kontrollpunkte. Ein Agent kann Unterlagen vorbereiten und ein Vorgehen vorschlagen. Die Berechtigung zu einer geschützten Aktion ergibt sich jedoch aus genehmigten Regeln und dem aktuellen, verifizierten Kontext.

- ***Einheitliche Regeln für alle angebundenen Agenten und Anwendungen mit dokumentierten Regelverantwortlichen, Versionen und Gültigkeitsdaten.***
- ***Gesteuerte Modellaufrufe anhand von Identität, Zweck, Datenprofil und Budget.***
- ***Kontrolle betrieblicher Aktionen vor ihrer Ausführung und eine eindeutige Übergabe an einen Menschen, wenn eine Genehmigung erforderlich ist.***
- ***Ein nachvollziehbarer Ablauf von der Anfrage über die Entscheidung bis zum tatsächlichen Ergebnis.***
- ***Sichere Regeländerungen durch Review, Tests, schrittweise Einführung und Rückkehr zu einer vorherigen Version.***

GEM SYSTEM

GEM CORTEX AI
– POLICY-AS-CODE A AI GATEWAY

Einordnung in die Produktfamilie

GEM Cortex AI

Funktionsblock	Beschreibung der Funktionen
1 | Cortex Core	*Orchestrierung von Intent, Laufzeitpfad, Modell und Mandantenkontext.*
2 | Cortex Access	*SSO, Rollen, Workload-Identität, Anbieterkonten, Projekte und Lebenszyklusmanagement.*
3 | Cortex Knowledge	*Konnektoren, Ingestion, ACL-aware Retrieval, Quellenverweise und Lebenszyklus des Index.*
4 | Cortex Flow	*Workflows, Werkzeuge, Genehmigungen, menschliche Aufgaben und Prozessstatus.*
5 | Cortex Control	*Policies, KI-Register, Telemetrie, Evaluationen, Budgets und FinOps.*
6 | Cortex Workspace	*Benutzerkanal für genehmigte Chat-, Wissens- und Arbeitsaufgaben.*

Model Execution bildet eine eigenständige Architekturebene für Modelladapter, lokale Inferenz, Routing und Fallback.

Die Lösung kann an bestehende Agenten, Anwendungen und die Integrationsinfrastruktur des Kunden angebunden werden. Die Übersicht zeigt die Einordnung in die Produktfamilie; der konkrete Leistungsumfang umfasst ausschließlich die ausgewählten Funktionen. Die Bezeichnungen der Blöcke legen weder eigenständige Lizenzen noch eine Verpflichtung zur Einführung der gesamten Produktfamilie fest.

2. Kundenproblem und Qualifizierung der Geschäftschance

In einem ersten KI-Piloten können Regeln im Code, im Prompt oder in einer lokalen Konfiguration enthalten sein. Mit wachsendem Portfolio steigt der Änderungsaufwand: Teams ermitteln betroffene Anwendungen, setzen dieselben Anforderungen unterschiedlich um und testen jede Bereitstellung gesondert. Zugleich fällt es der Organisation schwer nachzuweisen, auf welcher Grundlage ein bestimmter Agent gehandelt hat.

Ein zweites Problem ist der fragmentierte Zugang zu Modellen und Werkzeugen. Es fehlt eine gemeinsame Verknüpfung von Identität, Zweck, Daten und Kosten. Ein Agent kann weitergehende Berechtigungen besitzen, als sein Prozess benötigt, oder es kann einen direkten Pfad geben, der die Kontrolle umgeht. Policy-as-Code bestimmt die Zulässigkeit einer Entscheidung; Integrationen und Zugriffskontrollen stellen ihre tatsächliche Durchsetzung sicher.

Geeignet sind Unternehmen mit wiederkehrenden Entscheidungen, gemeinsamen Regeln und verfügbaren Integrationspunkten. Typische Umgebungen sind Finanzdienstleistungen, Versicherungen, dokumentenintensive Organisationen, öffentliche Dienstleistungen und die Unternehmens-IT. Entscheidend sind die Auswirkungen des Prozesses und der Umfang der Steuerung, nicht die Branche an sich oder die Bezeichnung einer Anwendung als Agent.

Funktionsblock	Beschreibung der Funktionen	Messbarer Nachweis
1 | CIO and CTO	*Wiederverwendung von Kontrollen und geringerer Änderungsaufwand über Anwendungen hinweg.*	*Dauer einer Regeländerung und Anzahl der Eingriffe in Anwendungen.*
2 | CISO and security	*Kontrolle der Datenpfade und Berechtigungen in der angebundenen Umgebung.*	*Kontrollabdeckung, Negativtests und Dokumentation von Ausnahmen.*
3 | Management and compliance	*Nachweisbare Verantwortung und Überblick über die KI-Nutzung.*	*Regelverantwortliche und Nachvollziehbarkeit einer konkreten Entscheidung.*
4 | Process owner	*Gewahrte Entscheidungsbefugnisse und vorhersehbare Genehmigungsverfahren.*	*Korrekte Eskalationen, Bearbeitungszeit und akzeptierte Ergebnisse.*
5 | CFO and procurement	*Transparente Kosten der Plattform und der einzelnen Prozesse.*	*Stückkosten und realisierter Nutzen gegenüber der Baseline.*

Die erste Geschäftschance muss eine verantwortliche Person, beschriebene Regeln, repräsentative Fälle und einen Punkt aufweisen, an dem sich eine Aktion stoppen lässt. Für kreative Assistenz oder ein einmaliges Experiment kann ein einfacheres Profil ausreichen. Wenn ein bestehender Workflow oder ein bereits lizenzierter Dienst denselben Bedarf sicher und wirtschaftlich abdeckt, schlägt GEM dessen Nutzung oder die Ergänzung gemeinsamer Kontrollen vor.

3. Kundenproblem und Qualifizierung der Geschäftschance

Policy-as-Code trennt die Frage, was unter bestimmten Bedingungen erlaubt ist, davon, wie ein Agent seine Arbeit ausführt. Eine ausführbare Regel kann beispielsweise Rollen, finanzielle Grenzwerte, Genehmigungsstatus, zeitliche Bedingungen, ein zulässiges Modell oder die Datenklassifizierung bewerten. Das generative Modell bereitet die Grundlagen vor; eine Regel prüft den genauen Grenzwert oder die Berechtigung anhand strukturierter Eingaben.

Dieselbe Regelversion sollte bei demselben vollständigen Kontext dieselbe Entscheidung liefern. Diese Eigenschaft erleichtert Tests und die Rekonstruktion eines Falls. Die Richtigkeit eines Eingabewerts und die korrekte Überführung einer Richtlinie bleiben Gegenstand gesonderter Kontrollen. Das Ergebnis der Extraktion aus einem Dokument ist daher möglicherweise kein ausreichender Nachweis, um einen Schreibzugriff zu erlauben.

Zwei Profile zur Steuerung von Agenten

Der Referenzentwurf unterscheidet nach Autonomie und Auswirkungen zwischen einem Assistenzprofil und einem Prozessprofil. Es handelt sich um Varianten der Einbindung der Kontrollebene, deren Eignung für den jeweiligen Prozess validiert wird.

Profil	Beschreibung der Funktionen	Messbarer Nachweis
1 | Assistenzprofil	*Zusammenfassungen, Textentwürfe oder Wissensarbeit ohne autonome Änderung eines Systems.*	*Jeder Modellaufruf erfolgt über AI Gateway; Identität, Daten und Verbrauch werden geprüft. Ein Mensch bewertet das Ergebnis.*
2 | Prozessprofil	*Der Agent bereitet mehrere Schritte vor und kann eine geschützte Aktion vorschlagen.*	*Begrenzung des Plans, Kontrolle von Modellaufrufen und Werkzeugen, erneute Entscheidung vor der Aktion und gegebenenfalls menschliche Genehmigung.*

Das vollständige Prozessprofil wird anhand der Auswirkungen und des Nachweisbedarfs gewählt. Zeitgeber, Warteschlangen, Wartezustände und der Workflow-Status verbleiben in der Prozessorchestrierung. Die Policy-Ebene bewertet die Bedingungen, die ihr der Orchestrator und vertrauenswürdige Quellen übermitteln.

Die Durchsetzung muss innerhalb des vereinbarten Umfangs auf jedem relevanten Pfad zu einem Modell, zu Daten oder zu einer geschützten Aktion erfolgen. Ein Gateway für Modellaufrufe allein verhindert weder einen Schreibzugriff auf ein ERP-System noch die Nutzung eines öffentlichen KI-Dienstes außerhalb der angebundenen Umgebung. Das Produktionskonzept erfasst deshalb auch Netzwerkpfade, Dienstkonten und Schlüssel, über die sich die Kontrolle umgehen lässt.

Eine Regeländerung erfordert keine Anpassung des Agenten, sofern die bestehende Integration alle benötigten Attribute liefert und die Entscheidung durchsetzen kann. Ein neuer Aktionstyp, ein neues Eingabefeld oder ein anderes Genehmigungsverfahren kann eine Änderung der Schnittstelle erfordern. Die Abgrenzung dieser Fälle gehört ebenfalls zum Leistungsumfang.

GEM Cortex AI

GEM Cortex AI

Abgrenzung des konkreten Leistungsumfangs

Der Angebotsumfang legt die angebundenen Agenten und Anwendungen, die ausgewählten Regeln, Modellendpunkte, betrieblichen Aktionen und Kontrollpunkte fest. Integrationen, Genehmigungsworkflows und Betriebsleistungen werden gesondert bestätigt. Weitere Prozesse werden auf Grundlage der Pilotergebnisse und einer genehmigten Erweiterung eingeführt.

PaC nutzt die Unternehmensidentität, den Kontext und die Ergebnisse weiterer Kontrollen. IAM, inhaltsbezogene Guardrails, Sicherheitsscanner und Orchestrierung haben jeweils eigene Verantwortungsbereiche. Ihr vollständiger Ersatz ist nicht Bestandteil des hier beschriebenen PaC-Prinzips.

4. Referenzarchitektur

Das technische Modell in den Kapiteln 4 bis 6 beschreibt eine Referenzvariante, die für die konkrete Bereitstellung zu validieren ist. Architektur, Auswahl der bereitgestellten Funktionen und Lösungsumfang werden in der Discovery-Phase und bei der Ausarbeitung des Leistungsumfangs bestätigt.

Die Referenzarchitektur trennt die Policy-Verwaltung von der Bewertung einer konkreten Anfrage. Die Steuerungsebene verwaltet die Regelquellen, deren Genehmigung, Tests und Veröffentlichung. Die operative Ebene bearbeitet Anfragen, beschafft den Kontext, trifft Entscheidungen und setzt das Ergebnis durch.

Rolle	Bedeutung	Verantwortung
1 | PAP	*Policy-Verwaltung*	*Erstellung, Metadaten, Review, Tests, Genehmigung und Verteilung von Versionen.*
2 | PDP	*Policy-Entscheidungspunkt*	*Den genehmigten Regelsatz anhand eines strukturierten Kontexts bewerten.*
3 | PEP	*Policy-Durchsetzungspunkt*	*Die Anfrage abfangen und ihre Ausführung tatsächlich erlauben, stoppen oder an Bedingungen knüpfen.*
4 | PIP	*Kontextquelle*	*Vertrauenswürdige Rollen, Grenzwerte, Klassifizierungen und den aktuellen Fallstatus bereitstellen.*
5 | Audit-Ebene	*Nachweisführung und Aufsicht*	*Entscheidung, verwendete Version, Erfüllung der Bedingungen und tatsächliches Ergebnis verknüpfen.*

AI Gateway übernimmt die PEP-Rolle für Modellaufrufe. Für betriebliche Aktionen übernimmt diese Rolle ein API-Gateway, ein Integrationsdienst, ein Workflow, eine Pipeline oder unmittelbar das Zielsystem. Der Durchsetzungspunkt wird danach gewählt, wo sich die Operation zuverlässig stoppen lässt.

Die technische Referenzvariante verwendet Open Policy Agent als PDP und Rego für ausführbare Regeln. Regeln und Tests werden in Git erstellt, durchlaufen CI und werden als genehmigte, signierte Bundles in die Betriebsumgebung verteilt. Die Auswahl der konkreten Technologie und der unterstützten Versionen ist Bestandteil des Lösungskonzepts.

In der Referenzvariante sind Git und das Artefaktregister keine synchronen Abhängigkeiten jeder einzelnen Entscheidung. Der Entscheidungspunkt zur Laufzeit verwendet eine bereits geladene, verifizierte Version. Das Betriebskonzept legt deren Gültigkeit und das Verhalten bei Nichtverfügbarkeit weiterer Komponenten fest.

Die Integrationen binden das Unternehmensverzeichnis, Genehmigungsworkflows, Modellendpunkte, Datenkataloge und die Betriebsüberwachung an. Ein vorhandenes API-Gateway und die Infrastruktur des Kunden können genutzt werden. Eine REST-API oder ein Werkzeugprotokoll wie MCP stellt die Schnittstelle bereit; die Kontrolle von Identität, Berechtigungen und Auswirkungen der Aktion muss gesondert sichergestellt werden.

5. Regelverwaltung und sichere Änderungen

Die Regelverwaltung beruht auf fachlicher Verantwortung, Versionierung, Review, Tests und sicherer Einführung. Der Referenzentwurf konkretisiert diese Verwaltung im nachfolgenden Lebenszyklus. Die konkreten Metadaten, Werkzeuge und Veröffentlichungsverfahren werden im Bereitstellungskonzept festgelegt.

Jede Regel hat eine verantwortliche Person und einen Bezug zu einer genehmigten Quelle. Die Dokumentation verknüpft die Richtlinienversion, den Gültigkeitsbeginn, die Kritikalität, die Genehmigungen und die technische Implementierung, damit sich der Ursprung einer Entscheidung nachweisen lässt.

Maschinell auswertbare Bedingungen werden von Teilen getrennt, die fachliches Urteilsvermögen erfordern. KI kann bei der Textanalyse oder beim Entwurf einer Regel unterstützen; die Verantwortung für die Erstellung verbleibt jedoch bei einem Menschen. Eine unklare Richtlinie muss fachlich präzisiert werden, bevor sie in eine ausführbare Policy überführt wird.

Lebenszyklus einer Policy



- ***Die verantwortliche Person bestätigt die Quelle, die Auslegung, den Gültigkeitsbeginn und konkrete Beispiele für das gewünschte Verhalten.***
- ***Die erstellende Person bereitet die Regel, die Metadaten sowie Tests für zulässige, unzulässige, grenzwertige und unvollständige Eingaben vor.***
- ***Die Änderung durchläuft ein fachliches und technisches Review. Bei kritischen Regeln erfolgt die Genehmigung durch eine andere Person als diejenige, die die Regel erstellt hat.***
- ***CI prüft das Schema und den zusammengesetzten Policy-Satz. Bekannte Fälle werden mit der neuen Version erneut durchgespielt; jede Änderung des erwarteten Ergebnisses wird erläutert.***
- ***Das genehmigte Release wird als versioniertes Artefakt mit Signatur, Hashwert und nachvollziehbarer Herkunft erstellt.***
- ***Die neue Version wird im Shadow-Modus oder innerhalb einer begrenzten Gruppe validiert. Erst nach der Genehmigung setzt sie die neuen Bedingungen durch.***
- ***Der Betrieb überwacht die Auswirkungen der Änderung und die aktiven Versionen. Der Lebenszyklus umfasst Rezertifizierung, Ablauf der Gültigkeit, Rückkehr zum vorherigen Release und Außerkraftsetzung der Regel.***

Der Shadow-Modus vergleicht Entscheidungen und Auswirkungen der neuen Version, ohne deren blockierende Wirkung zu aktivieren. Bereits geltende Kontrollen bleiben dabei wirksam. Die Rückkehr zu einer früheren Regelversion stellt die vorherige Konfiguration wieder her; bereits ausgeführte betriebliche Aktionen werden dadurch nicht automatisch rückgängig gemacht.

GEM Cortex AI

Zusammensetzung und Auswahl von Regeln

GEM Cortex AI

Gemeinsame Regeln der Organisation werden mit domänen- und prozessbezogenen Regeln kombiniert. Die Hierarchie definiert Namensräume und die Aggregation der Ergebnisse ausdrücklich. Grundlegende Schutzmaßnahmen müssen Bestandteil jeder einschlägigen Entscheidung sein und dürfen durch eine Ausnahme nicht stillschweigend außer Kraft gesetzt werden.

In der Referenzvariante bestimmen das Register und der verifizierte Kontext innerhalb der Policy-Ebene den richtigen Regelsatz. Der Agent wählt die Regel nicht selbst aus. Fehlt eine benötigte Policy, wird die Anfrage abgelehnt; die Kontrolle wird nicht stillschweigend übersprungen.

Der Referenzentwurf sieht vor, die Signatur und den Geltungsbereich des Artefakts vor dessen Aktivierung zu prüfen. Er umfasst den Schutz der Signaturschlüssel, die Überwachung aktiver Versionen, die Rezertifizierung und zeitlich begrenzte Ausnahmen. Diese Mechanismen werden für die gewählte Bereitstellung validiert.

6. Referenzansatz zur Steuerung von Plan und Aktionen

Das Prozessprofil sieht eine Kontrolle vor Beginn der Arbeit des Agenten und eine erneute Bewertung vor Eintritt einer geschützten Wirkung vor. Zwischen diesen Punkten können sich Fallstatus, Berechtigungen, Betrag oder Gültigkeit der Unterlagen ändern. Die Erlaubnis zu Beginn des Prozesses stellt deshalb keine Berechtigung für sämtliche nachfolgenden Aktionen dar.

PLAN legt den Arbeitsrahmen fest

Vor Arbeitsbeginn bestimmt die Policy die zulässigen Prüfungen, Werkzeuge, Vorlagen und Grenzwerte. Der Agent erhält einen definierten Rahmen, in dem er Unterlagen vorbereiten kann. Der Plan kann beispielsweise festlegen, welche Dokumente zu bewerten und welche Felder zu extrahieren sind und wann eine weitere Bearbeitung keinen Zweck erfüllt. So lassen sich unnötige Modelloperationen bereits vor ihrer Ausführung begrenzen.

Jeder nachfolgende Modellaufruf durchläuft eine eigene Kontrolle in AI Gateway. Die Nutzung eines Werkzeugs unterliegt dessen Berechtigungen und den aktuellen Regeln. Für den Prozessablauf, Wartezustände und die Aggregation der Ergebnisse ist die Orchestrierung verantwortlich.

DECIDE bewertet die konkrete Aktion

Unmittelbar vor einer geschützten Aktion bewertet der Entscheidungspunkt die Prüfergebnisse und den aktuellen Kontext. Ein technisches PASS-Ergebnis für ein Dokument reicht möglicherweise nicht aus, um einen Schreibzugriff zu erlauben: Berechtigungen, die zeitliche Gültigkeit oder eine Genehmigung können weiterhin fehlen. Eine unvollständige oder unbestimmte Angabe darf nicht stillschweigend als erfüllte Bedingung behandelt werden.

Das Ergebnis ist eine Erlaubnis, eine Ablehnung oder eine Erlaubnis unter verbindlichen Auflagen. Eine solche als Obligation bezeichnete Auflage kann eine menschliche Genehmigung, einen niedrigeren Grenzwert oder die Nutzung eines anderen zulässigen Modells verlangen. Der PEP führt die Operation erst aus, wenn die erforderliche Auflage erfüllt ist.

Inhalt der Entscheidung und Audit-Verknüpfungen

Der Referenzentwurf arbeitet mit Subjekt, Aktion, Objekt und Kontext. Die Entscheidung enthält zusätzlich eine Kennung, eine Begründung, die Policy-Version und gegebenenfalls Obligations. Das konkrete Schnittstellenschema wird im Integrationskonzept festgelegt.

Kontraktelement	Erforderlicher Inhalt
1 | Subject	*Verifizierter Benutzer oder Workload, Delegation und relevante Berechtigungen.*
2 | Action and object	*Konkrete Operation, Parameter, Zielsystem und Kennung des Falls oder der Ressource.*
3 | Context	*Zweck, Umgebung, Datenklassifizierung, Grenzwerte, aktueller Status und maßgeblicher Zeitpunkt.*
4 | Decision	*Ergebnis, Begründungscode, Erläuterung, Decision ID, Policy-Versionen und verbindliche Auflagen.*
5 | Execution	*Nachweis der erfüllten Auflagen, Status der Operation und Ergebniskennung im Zielsystem.*

GEM Cortex AI

Vertrauenswürdige Rollen, Grenzwerte und Genehmigungen stammen aus Unternehmensquellen. Bei der Entscheidung vor einer Aktion wird der aktuelle Kontext bewertet, der sich seit Prozessbeginn geändert haben kann.

Bei einer Obligation, die eine menschliche Genehmigung verlangt, sieht der Entwurf das Anhalten der Aktion und eine Audit-Verknüpfung mit der Entscheidung vor. Der Genehmigungsworkflow bleibt in der Verantwortung des Prozessteams. Für Werkzeugintegrationen gelten Anforderungen an Berechtigungen, Grenzwerte, Vorschau, Genehmigung, Idempotenz und Audit.

Der Pilot sieht standardmäßig Lesezugriffe oder die Erstellung von Vorschlägen vor. Irreversible und rechtlich bedeutsame Schritte bleiben unter menschlicher Kontrolle. Die Erweiterung der Autonomie erfordert eine gesonderte Entscheidung für den konkreten Prozess und validierte Kontrollpunkte.

7. AI Gateway Daten und Kostensteuerung

AI Gateway führt den Modellverkehr angebundener Anwendungen unter einer gemeinsamen Steuerung zusammen. Jeder Aufruf wird im Kontext von Identität, Zweck, Datenprofil, zulässigem Endpunkt und Kostenverantwortung bewertet. Das Modellregister dokumentiert die verantwortliche Person, die Region, die zulässige Datenklassifizierung, die vertraglichen Rahmenbedingungen und die Betriebsparameter.

Die Autorisierung eines Aufrufs und die Inhaltsprüfung erfüllen unterschiedliche Aufgaben. Die Policy legt fest, ob eine Anfrage ausgeführt werden darf. Inhaltsprüfungen, häufig als Guardrails bezeichnet, bewerten Eingaben und Ausgaben beispielsweise hinsichtlich sensibler Inhalte oder eines zulässigen Formats. Das Routing wählt ein genehmigtes Modell aus; die Verbrauchsmessung liefert die Grundlage für FinOps.

- ***Die Kontrolle vor der Übermittlung prüft die Zulässigkeit des Datenpfads. Je nach Anforderung werden Klassifizierung, Erkennung sensibler Daten oder Schwärzung ergänzt.***
- ***Routing und Fallback verwenden ausschließlich Endpunkte, die für die jeweilige Identität, den Zweck und das Datenprofil zulässig sind.***
- ***Grenzwerte beschränken Anfragegröße, Häufigkeit, Laufzeit, Token-Verbrauch und Budgetrahmen.***
- ***Die Messung ordnet den tatsächlichen Verbrauch einer Anwendung, einem Prozess, einem Projekt oder einer Kostenstelle zu und verknüpft ihn mit der abgeschlossenen Aufgabe.***
- ***Eine strukturierte Ablehnung erläutert den Grund und ermöglicht es der Anwendung, die Anfrage kontrolliert zu ergänzen, zurückzustellen oder an einen Menschen zu übergeben.***

Ein harter Budgetstopp gilt für den Verbrauch, der über das kontrollierte Gateway erfolgt. Die Schätzung vor einem Aufruf und die endgültige Abrechnung können voneinander abweichen; die Kostenaufzeichnungen werden deshalb mit den Anbieterdaten abgeglichen. Konkrete Grenzwerte und ihre Durchsetzung werden im Konzept und im Piloten validiert. Eine Optimierung des Modells oder des Kontexts wird nur unter Wahrung der geforderten Qualität übernommen.

Der Datenschutz umfasst die Prüfung der Zugriffsberechtigungen für Quellen, bevor der Modellkontext zusammengestellt wird. Eine Prüfung der Antwort kann die Übermittlung einer unzulässigen Eingabe nicht nachträglich verhindern. Zero Retention, Region und der Ausschluss einer Datennutzung für das Training werden für den konkreten Dienst und die vertragliche Konfiguration bestätigt.

Öffentliche Chatdienste, Desktop-Werkzeuge und SaaS-Assistenten erfordern eine gesonderte Bewertung der Integrationsmöglichkeiten und Unternehmenskontrollen. Bei Entwickler-Copiloten kann das Ergebnis im Pull Request und in CI geprüft werden; dies weist jedoch keine Kontrolle über die gesamte Kommunikation des Werkzeugs mit dem Anbieter nach.

8. Audit Sicherheit und betriebliche Verantwortung

Der Audit-Trail verknüpft den Plan, Modellaufrufe, Entscheidungen vor einer Aktion, die Erfüllung von Auflagen und das tatsächliche Ergebnis. Ein Erlaubnisvermerk bestätigt keine erfolgreiche Ausführung. Diese muss das Zielwerkzeug oder das Unternehmenssystem nachweisen. Eine eindeutige Korrelation ermöglicht es festzustellen, an welcher Stelle der Prozess gestoppt wurde und warum.

Ein Audit-Eintrag verknüpft Zeitpunkt, Identität, Prozess, Aktion, maßgeblichen Kontext, Decision ID, Regelversion, Begründung und Durchsetzungsstatus. Der Entwurf bestimmt anhand der Datenklassifizierung und des Kontrollzwecks den Umfang der gespeicherten Eingaben, die Möglichkeit einer erneuten Auswertung, die Aufbewahrungsdauer und die Zugriffsberechtigungen für die Nachweise.

Die übliche Telemetrie enthält nicht automatisch vollständige Prompts oder Dokumente. Zugriffe für Betrieb, Audit und Forensik werden nach Rollen getrennt. Aufzeichnungen müssen vor unbefugter Änderung geschützt sein; erforderlich sind außerdem ein definierter Export und eine Vollständigkeitsprüfung einschließlich Ablehnungen und Fehlern. Die Nachweise unterstützen das interne Audit und die regulatorische Bewertung. Der Kunde bestätigt die Anwendbarkeit rechtlicher Anforderungen und die Verantwortung für deren Auslegung.

GEM Cortex AI

Sicherheitsanforderungen

Der Entwurf bewertet die Umgehung von PEP, gefälschte Identitäten oder Kontexte, kompromittierte Regelpakete, die Manipulation von Audit-Aufzeichnungen und die Nichtverfügbarkeit des Entscheidungsdienstes. Risiken durch Prompt Injection und Werkzeugmissbrauch werden gesondert behandelt. Der Zugriff auf Schlüssel, Repositories und Policy-Änderungen folgt dem Prinzip der geringsten Berechtigungen und der Funktionstrennung.

Bei geschützten Aktionen sieht der Referenzentwurf ein sicheres Anhalten vor, wenn keine Entscheidung verfügbar ist, sowie die Nutzung der zuletzt verifizierten Policy in einem definierten Betriebsmodus. Das Verhalten bei Ausfällen, die Gültigkeit von Policies, etwaige Caches und die Wiederherstellung werden für den konkreten Betrieb bewertet und getestet.

Betriebsmodell und Verantwortlichkeiten

Die folgende Aufteilung stellt ein Referenzmodell der Verantwortlichkeiten dar. Verantwortlichkeiten und Serviceverpflichtungen werden für den gewählten Leistungsumfang bestätigt.

Bereich	Verantwortung des Kunden	Verantwortung von GEM im vereinbarten Umfang
1 | Regeln und Prozess	*Fachliche Auslegung, Grenzwerte, Zuständigkeit und Genehmigung der Autonomie.*	*Formalisierung, Implementierung und technische Tests.*
2 | Daten und Identitäten	*Klassifizierung, Zweck, Berechtigungen und Zugang zu Quellen.*	*Integration und korrekte Übermittlung des Kontexts.*
3| Plattform	*Auswahl des Betriebsmodells und Bereitstellung kundenseitiger Abhängigkeiten.*	*Anwendungskomponenten, Konfiguration und vereinbarte Administration.*
4| Abnahme	*Bestätigung korrekter Ergebnisse und Übernahme des Prozesses.*	*Testnachweise, Mängelbeseitigung und Übergabe.*
5| Vorfälle und Änderungen	*Prozesseskalationen und Genehmigung fachlicher Änderungen.*	*Technische Diagnose, Wiederherstellung und Versionsverwaltung.*

GEM CORTEX AI
– *POLICY-AS-CODE A AI GATEWAY*

Infrastruktur und Modellendpunkte werden von der vertraglich bestimmten Partei betrieben. Die Lösung kann in der Cloud des Kunden, als Dienst von GEM, in einer hybriden Konfiguration oder in einer eingeschränkten lokalen Umgebung betrieben werden. Jede Variante hat eigene Datenflüsse, Kapazitätsanforderungen und Verantwortlichkeiten.

Das Servicekonzept legt Verfügbarkeit, Latenz, Kapazität sowie die Wiederherstellungsparameter RTO und RPO entsprechend der Kritikalität des Prozesses fest. Im vereinbarten Betriebsumfang umfasst es Monitoring, Backups, Wiederherstellung, Versionsrücknahme und Verfahren für Vorfälle. Die Leistung wird anhand eines realen Regelsatzes und der erwarteten Betriebslast validiert.

Haben Sie Ihre KI-Agenten unter Kontrolle?

Auf Grundlage welcher Regel hat der Agent entschieden?

Wohin dürfen Unternehmensdaten übermittelt werden?

Können Sie die Entscheidung gegenüber einem Auditor belegen?

Policy-as-Code und AI Gateway helfen, Regeln durchzusetzen und Entscheidungen nachzuweisen.

GEM Cortex AI

GEM System a. s. | Štětkova 1638/18, 140 00 Praha 4 – Nusle, Tschechien | sales@gemsystem.cz | www.gemsystem.cz

9. Anwendungsfälle und erste Vertriebsszenarien

Die folgenden Szenarien zeigen Einsatzbereiche der gemeinsamen Kontrollebene. Der erste Prozess, seine Integration und die Abnahmemetriken werden bei der Qualifizierung und im Pilot Charter festgelegt.

Dokumenten- und Finanzprozesse

KI bereitet strukturierte Daten aus Dokumenten auf und prüft, ob die erforderlichen Unterlagen vorhanden sind. Die Policy bewertet präzise Bedingungen, etwa die Gültigkeit eines Dokuments, eine Unterschrift, einen zeitlichen Schwellenwert oder einen finanziellen Grenzwert. Ein unvollständiger Fall wird an eine Fachkraft übergeben. Ein Schreibzugriff oder eine finanzielle Aktion erfolgt nach der erforderlichen Genehmigung im kontrollierten System.

Für den ersten Piloten wird ein Dokumentenprozess mit begrenztem Umfang ausgewählt. Validiert werden die Vorbereitungszeit, die Übergabe an einen Menschen, regelbasierte Entscheidungen und der Audit-Trail. Der Nutzen entsteht durch schnellere Vorbereitung und einheitliche Kontrolle bei gewahrter Verantwortung für die Entscheidung.

KI-gestützte Entwicklung und Bereitstellung

Ein Entwickler oder Agent bereitet eine Änderung vor. Pull Request und CI-Pipeline prüfen Tests, Ergebnisse von Sicherheitsscannern, Secrets im Code und genehmigte Abhängigkeiten. Die Policy bewertet diese Ergebnisse und die Bedingungen für den nächsten Schritt. Ein Mensch genehmigt die Änderung; die Bereitstellung in der Produktion erfolgt im zulässigen Verfahren.

Der Pilot kann auf ein ausgewähltes Repository und eine Pipeline begrenzt werden. Er validiert die regelkonforme Blockierung von Änderungen und die Nachvollziehbarkeit der Genehmigungen. Der Agent erhält nur Zugang zu erlaubten Repositories und Umgebungen. Policy-as-Code verwendet die Ergebnisse der Scanner als Eingaben.

Schwachstellen und Steuerung der Behebung

Der Agent klassifiziert Befunde und erstellt Aufgabenvorschläge. Die Policy berücksichtigt die Kritikalität des Assets, den Schweregrad des Befunds, interne Fristen und Eskalationsregeln. Eine Änderung in der Produktion erfordert ein genehmigtes Vorgehen. Eine erfolgreiche Empfehlung zur Behebung wird von der Berechtigung zu ihrer Umsetzung getrennt.

Der Pilot umfasst einen begrenzten Bereich des Schwachstellenmanagements und der Behebung. Er validiert Entscheidungen anhand von Fristen und Eskalationsregeln sowie die Audit-Verknüpfung. Gemeinsame Regeln erleichtern es, eine neue interne Methodik in den angebundenen Teams anzuwenden.

Datenschutz und Zugang zu Modellen

Eine Anwendung fordert eine Zusammenfassung oder Analyse interner Inhalte an. Das Gateway bewertet Zweck, Klassifizierung und zulässigen Endpunkt. Die Wissensebene prüft vor der Weitergabe von Inhalten die Berechtigungen des Benutzers. Gibt es keinen geeigneten Pfad, wird die Anfrage gestoppt oder einer genehmigten Prüfung zugeführt.

Der Pilot wählt eine angebundene Anwendung und einen Datenumfang aus. Er validiert die Abdeckung der Modellaufrufe, die Zulässigkeit des Routings, den Fallback und die Kostenmessung. Sofern der Umfang Inhaltsprüfungen einschließt, wird deren Wirksamkeit gesondert validiert.

Kundenprozesse und kaufmännische Befugnisse

KI erstellt einen Vorschlag für eine Antwort, einen Rabatt oder die Bearbeitung einer Reklamation. Die Regeln prüfen Zuständigkeiten, Grenzwerte und Fallstatus. Eine verbindliche Zusage oder Vertragsänderung wird an eine berechtigte Person übergeben und als kontrollierter Schritt im zuständigen System ausgeführt.

Der Pilot wählt eine konkrete Entscheidungssituation und das zugehörige Genehmigungsverfahren aus. Er validiert die Einhaltung von Zuständigkeiten, Grenzwerten und Eskalationen. Das Ergebnis der Vorbereitung und die endgültige kaufmännische Entscheidung bleiben getrennt nachvollziehbar.

GEM Cortex AI

Öffentliche Dienstleistungen und sensible Verwaltungsverfahren

KI ordnet Eingaben zu und bereitet Registerdaten oder eine Prüfliste vor. Die Regeln steuern Zugriffe, Übergaben zwischen Rollen und Workflow-Bedingungen. Die fachliche Bewertung und rechtlich bedeutsame Entscheidungen verbleiben bei der verantwortlichen Fachkraft.

Der Pilot arbeitet mit einem einzelnen Verwaltungsverfahren und einer repräsentativen Fallmenge. Er validiert Berechtigungen, Fehlklassifizierungen, die Übergabe an eine Fachkraft und Audit-Verknüpfungen. KI trifft keine Sachentscheidungen.

10. Leistungserbringung Implementierung und Abnahme

GEM beginnt mit einem konkreten Prozess und einer überprüfbaren Nutzenhypothese. Die Discovery-Phase erfasst Regeln, bestehende Zugriffsregelungen, Datenpfade, Konten und sämtliche Stellen, an denen eine geschützte Wirkung entstehen kann. Sie ermittelt, welche Kontrollen aus bestehenden Systemen genutzt werden können und wo eine Integration erforderlich ist.

Phase	Arbeitsinhalte	Ergebnis für den Kunden
1 | Qualifizierung und Discovery	*Prozess, verantwortliche Person, Baseline, Regeln, Risiken und Integrationsmöglichkeiten.*	*Abgrenzung der Geschäftschance und empfohlener Umfang.*
2 | Konzeption und Pilot	*Entscheidungskontrakt, priorisierte Regeln, historische Fallmenge und Shadow-Modus.*	*Pilot Charter und Nachweis von Funktion und Auswirkungen.*
3| Begrenzter Produktionsbetrieb	*Identitäten, Durchsetzung, Genehmigungen, Monitoring, Sicherheit und Wiederherstellung.*	*Abnahme und Betriebsübernahme.*
4| Weiterentwicklung und Betrieb	*Weitere Prozesse, Rezertifizierung, FinOps und gesteuerte Änderungen.*	*Portfolio mit gemessenen Ergebnissen und Wiederverwendung von Komponenten.*

Der Pilot Charter legt Eingaben und Ausgaben, die einbezogenen und ausgeschlossenen Bereiche (in scope und out of scope), Verantwortliche, Kontrollpunkte, Testfallmenge, Kosten und die Kriterien für die Entscheidung Go, Adjust oder Stop fest. Repräsentative Fälle schließen auch Grenzfälle und unzulässige Situationen ein. Die zur Abstimmung verwendete Fallmenge und die unabhängige Validierungsmenge werden getrennt gehalten.

Der Pilot validiert eine ausgewählte Regel von der Genehmigung bis zur Anwendung an den angebundenen Kontrollpunkten, die Richtigkeit der Entscheidungen, den Audit-Ablauf und die betrieblichen Auswirkungen. Das Prozessprofil ergänzt die Kontrolle der konkreten Aktion und der erforderlichen menschlichen Genehmigung.

Abnahme im vereinbarten Umfang

Die Validierungsbereiche werden anhand des Profils und der Architektur des Piloten ausgewählt. Konkrete Szenarien, erwartete Ergebnisse und numerische Schwellenwerte werden im Pilot Charter genehmigt.

GEM Cortex AI

Validierungsbereich	Nachweis für die Abnahme
1 | Regelbasierte Entscheidungen	*Zulässige, unzulässige und grenzwertige Fälle liefern das erwartete Ergebnis.*
2 | Policy-Änderung und Rücknahme	*Review, Tests, Veröffentlichung und Nachvollziehbarkeit der verwendeten Version; validierte Rücknahme.*
3 | Modellpfad	*Aufrufe der angebundenen Anwendung erfolgen über das festgelegte Gateway und zulässige Modelle.*
4 | Prozessaktionen und Genehmigungen	*Der PEP setzt die Entscheidung durch; bei erforderlicher Genehmigung wartet die Aktion auf einen Menschen.*
5 | Ausfälle und Wiederherstellung	*Das Verhalten entspricht dem vereinbarten sicheren Betriebsmodus und den Wiederherstellungsregelungen.*
6 | Audit-Trail	*Identität, Regelversion, Begründung und Ergebnis eines ausgewählten Falls sind nachvollziehbar.*
7 | Betrieb und Wirtschaftlichkeit	*Gemessene Latenz, Verbrauch, Aufwand und Auswirkungen gegenüber der Baseline.*

Negativtests prüfen, dass innerhalb der vereinbarten Testfallmenge keine unzulässige Aktion auftritt. Das Bestehen der Tests ist weder eine allgemeine Garantie für einen fehlerfreien künftigen Betrieb noch eine automatische Compliance-Zertifizierung.

Das Übergabepaket wird entsprechend der Phase und dem Leistungsumfang festgelegt. Es umfasst die Abgrenzung der Lösung, Architektur und Datenflüsse, Sicherheits- und Testnachweise, betriebliche Verantwortlichkeiten, Wirtschaftlichkeit und vereinbarte Exporte bei Beendigung des Dienstes.

11. Wirtschaftlichkeit Angebot und Nutzenmessung

Die Wirtschaftlichkeitsbetrachtung unterscheidet zwischen der gemeinsamen Plattform und der Einführung einzelner Prozesse. Für die Plattform fallen anfängliche und wiederkehrende Kosten an. Weitere Prozesse können bereits eingeführte Identitäten, die Regelverteilung, das Gateway und Audit-Funktionen nutzen. Prozessspezifische Regeln, ihre Genehmigung und Validierung erfordern dennoch einen eigenen Aufwand.

Das Angebot trennt Discovery, Plattformimplementierung, Integrationen und Prozess-Onboarding von Modellverbrauch, Infrastruktur, Support und laufender Regelverwaltung. Der Preis hängt insbesondere von der Anzahl der zu kontrollierenden Pfade, der Dokumentationsqualität, der Verfügbarkeit von Identitäten, dem Genehmigungsumfang und dem geforderten Servicelevel ab. Eine verbindliche Kalkulation wird nach der Qualifizierung des Umfangs erstellt.

Phase	Arbeitsinhalte	Ergebnis für den Kunden
1 | Regeländerungen	*Änderungshäufigkeit, Aufwand und Anzahl betroffener Anwendungen.*	*Eine gemeinsame Regel ersetzt tatsächlich wiederholte Arbeit.*
2 | Tests und Aufsicht	*Zeitaufwand für Regressionstests, Korrekturen und betriebliche Kontrolle.*	*Die erforderliche Qualität und der Validierungsumfang bleiben erhalten.*
3| Audit-Vorbereitung	*Zeit für das Auffinden und Zusammenstellen von Nachweisen.*	*Die Nachweise sind vollständig und für den jeweiligen Zweck verwendbar.*
4| Modellverbrauch	*Aufrufe, Tokens und Kosten je akzeptiertem Ergebnis.*	*Die Optimierung wahrt Qualität und Adoptionsgrad.*
5| Prozessdurchsatz	*Bearbeitungsdauer je Fall, Nacharbeit und abgeschlossenes Volumen.*	*Die Organisation nutzt die freigesetzte Kapazität.*

Zeitersparnisse werden als Kapazität erfasst, bis deren Nutzung oder eine Senkung tatsächlicher Kosten nachgewiesen ist. Eine Risikominderung wird in einem gesonderten Szenario dargestellt. Ein geeigneterer Plan, Kontext oder ein geeigneteres Routing kann den Modellverbrauch senken; die Einführung von PaC allein garantiert keine allgemeine prozentuale Einsparung.

Der Business Case stellt den mehrjährigen Nutzen den Kosten für Implementierung, Onboarding und wiederkehrenden TCO gegenüber. Er berücksichtigt das erwartete Portfolio, den Adoptionsgrad und realistische Einführungstermine. Der Pilot soll die einzelnen Annahmen validieren und die Kalkulation präzisieren. Die Wirtschaftlichkeit eines konkreten Kundenangebots lässt sich nicht automatisch auf eine andere Bereitstellung übertragen.

12. Vertriebskommunikation und nächster Schritt

Kernbotschaft

GEM Cortex AI gibt der Organisation gemeinsame Regeln für die Nutzung von Modellen und die Steuerung von Aktionen der KI-Agenten. Die Regeln haben Verantwortliche, Versionen und Tests. Angebundene Gateways und betriebliche Kontrollpunkte wenden sie vor einer konkreten Operation an; das Ergebnis lässt sich nachweisen. Der Kunde kann weitere KI-Prozesse auf der bereits etablierten Kontrollebene aufbauen.

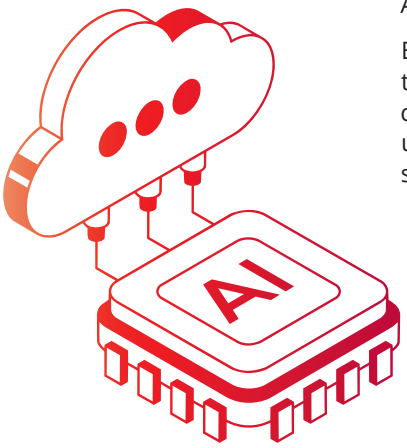
Kurze Vorstellung des Angebots

Wenn sich eine Unternehmensregel ändert, verursachen ihre unterschiedlichen Umsetzungen in KI-Anwendungen Aufwand und Risiken. GEM verbindet Ihre Agenten mit einer gemeinsamen Regelverwaltung und einem kontrollierten Zugang zu Modellen über ein Gateway. Präzise Grenzwerte und Berechtigungen werden vor einer Aktion bewertet, Änderungen durchlaufen Tests und risikobehaftete Schritte werden an einen Menschen übergeben. Wir beginnen mit einem Prozess und validieren Kontrolle, betriebliche Auswirkungen und Wirtschaftlichkeit.

Umgang mit typischen Einwänden

Ein bestehender Agent oder Workflow kann erhalten bleiben, sofern geeignete Integrationspunkte vorhanden sind. In der Discovery-Phase werden die Regeln ausgewählt, die einen gemeinsamen Lebenszyklus benötigen, und der erforderliche Kontext erfasst. Eine schrittweise Integration begrenzt den Änderungsumfang und ermöglicht den Vergleich der neuen Entscheidungslogik mit dem aktuellen Zustand.

Eine lokale Konfiguration ist für sich genommen kein Problem. Entscheidend sind Verantwortung, Tests, Veröffentlichung und Durchsetzung der Regel. Ein B2B-Nutzen entsteht dort, wo eine einheitliche Verwaltung wiederholte Arbeit beseitigt und nachweisbare Kontrolle über mehrere Anwendungen hinweg schafft.



Eine Policy Engine übernimmt die Auswertung von Regeln. Eine produktive Bereitstellung ergänzt Kontext, Kontrollpunkte, Genehmigungen, sichere Veröffentlichung und betriebliche Verantwortung. Dieser Umfang unterscheidet einen bereitgestellten Plattformdienst von einer eigenständigen Bibliothek oder einem Demonstrationsprototyp.

GEM Cortex AI

GEM CORTEX AI
– POLICY-AS-CODE A AI GATEWAY

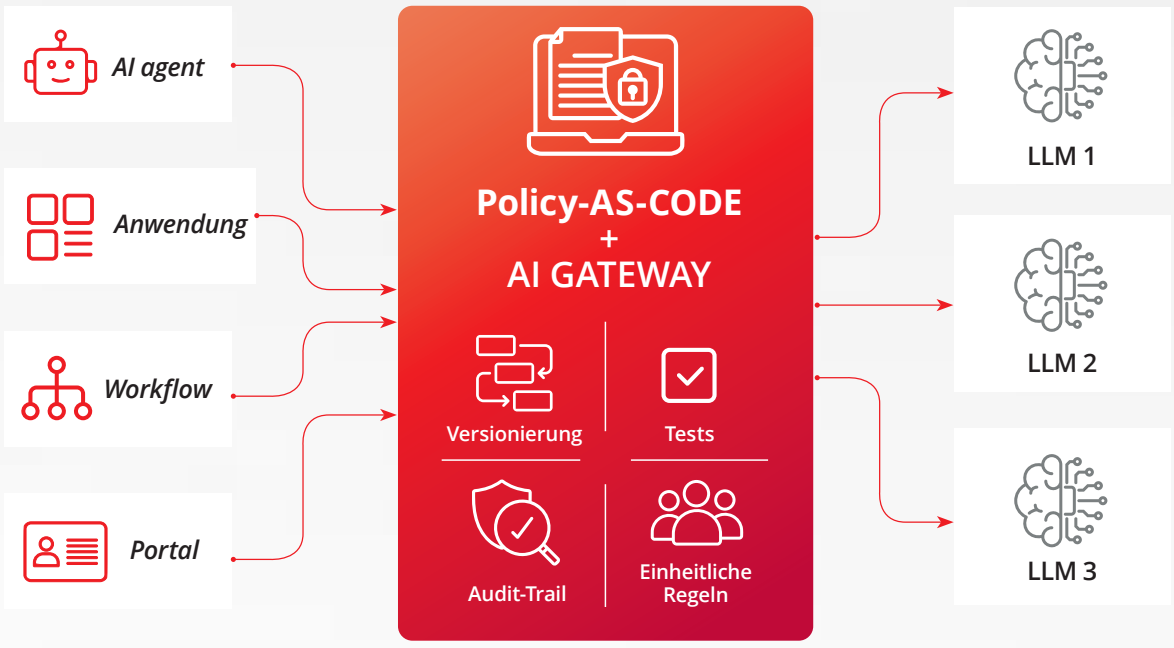
Fragen für das erste Gespräch

- *Welchen Prozess und welche Entscheidung soll die Lösung verbessern, und wer trägt dafür die Verantwortung?*
- *Welche Regeln wiederholen sich in Anwendungen, und welchen Aufwand verursachen ihre Änderungen derzeit?*
- *Welche Daten, Modelle und Unternehmenswerkzeuge nutzt der Prozess, und wer verwaltet die Zugriffe?*
- *An welcher Stelle lässt sich die tatsächliche Aktion stoppen, und welche Schritte erfordern eine menschliche Genehmigung?*
- *Lassen sich für einen konkreten Fall die verwendete Version, die Begründung und das tatsächliche Ergebnis nachvollziehen?*
- *Welchen Nachweis muss der Pilot liefern, damit der Kunde über die Einführung in den Produktionsbetrieb entscheiden kann?*

Ein KI-Architekt von GEM und die prozessverantwortliche Person bereiten einen begrenzten Piloten mit konkreten Regeln, Kontrollpunkten und einer Messung gegenüber der Baseline vor. Das Ergebnis ist eine durch Nachweise gestützte Entscheidung über die Einführung und Erweiterung der Lösung.

Eine Regeländerung gilt für alle

Regeln getrennt vom Code der KI-Agenten



GEM Cortex AI

GEM System a. s. | Štětkova 1638/18, 140 00 Praha 4 – Nusle, Tschechien | sales@gemsystem.cz | www.gemsystem.cz