BlueFlag Security

DEVELOPER RISK & GOVERNANCE · SOLUTION BRIEF · AGENTIC AI CODING INSIGHTS · AUGUST 2026

# Agentic AI Economics: vom Token-Burn zum Business Value.

KI-Agenten committen, reviewen und mergen bereits Code — Copilot, Claude, Cursor, Code-Rabbit, Qodo, Lovable und weitere. Die meisten Organisationen können weder die Risikofrage beantworten (**welche Agenten können unsere Software verändern — mit welchen Privilegien?**) noch die Wirtschaftlichkeitsfrage (**was liefern Ausgaben und Token-Burn tatsächlich an akzeptiertem Code und ausgelieferten Features zurück?**). BlueFlags AI Insights beantworten beides an einem Ort: jeder Agent als Identität entdeckt und governed — und jeder Dollar KI-Ausgaben bis zum akzeptierten Code nachverfolgt.

Ausgaben → Token-Verbrauch → Akzeptierter Code → Ausgelieferte Features → **Business Value**

ENTDECKT
**7 Agenten-Plattformen**
Claude, Copilot, Cursor, Code-Rabbit, Qodo, Lovable u. a. — inkl. Agenten, die niemand registriert hat

ERFASST
**$354 / 90 Tage**
KI-Ausgaben mit Token-Verbrauch und akzeptierten Zeilen — pro Tool und pro Nutzer

IM EINSATZ
**15 Modelle**
unterschiedliche LLMs im Bestand — von claude-opus-4-8 bis zum selbst gehosteten phi-2

KARTIERT
**13 KI-Repos**
Repositories mit LLM-Nutzung, Frameworks, Providern und Hosting klassifiziert

*Alle Werte stammen aus einer Live-Beispielumgebung (90-Tage-Fenster) — zur Veranschaulichung des Produkts; kundenspezifische Werte liefert der Read-only-Scan.*

EXECUTIVE SUMMARY

## Agenten sind Identitäten. Ausgaben sind ein Steuerungsinstrument. Beides braucht Governance.

Das AI-Insights-Modul behandelt jeden Coding-Agenten als vollwertige Identität — entdeckt, attribuiert, access-reviewed und gescort — und KI-Ausgaben als Engineering-Economics-Datensatz: Seats vs. aktive Nutzer, Ausgaben vs. Tokens vs. akzeptierte Zeilen, Modell-Adoption über den gesamten Bestand. Das Ergebnis ist ein Gesamtbild, das CISO (Risiko), CFO (Kosten) und CTO (Velocity) gleichermaßen bedient.

### WAS DAS MODUL BEANTWORTET

- **Welche KI-Agenten existieren** in unserem SDLC — aktiv und im Verborgenen — und in welchen Repositories?
- **Wer hat was geschrieben:** agent-primäre vs. mensch-primäre vs. co-authored Commits und PR-Aktivität
- **Was kostet KI-Coding** — und was kommt als akzeptierter Code zurück?
- **Welche Modelle und Frameworks** sind tatsächlich im Einsatz — cloud-managed oder self-hosted?

### DIE FRAGEN, DIE DAS MANAGEMENT STELLT

- **Kann ein nicht registrierter KI-Agent in Produktion mergen — heute?**
- **Warum zahlen wir für Seats, die niemand nutzt, während unlizenzierte Tools Tokens verbrennen?**
- **Was hat dieses Feature tatsächlich an KI-Ausgaben gekostet?**
- **Wo gelangt KI-generierter Code in unsere regulierten Anwendungen?**

DIE FEATURE-MAP

# Jedes AI-Insights-Feature — Risiko eliminiert, Wert geschaffen

| FEATURE | WAS SIE IM PRODUKT SEHEN | ELIMINIERTES RISIKO · BUSINESS-NUTZEN |
|---|---|---|
| **SDLC Agent Activity** | Aktive Agenten, gerankt nach realer Aktivität — Commits, PR-Kommentare, Reviews (z. B. Code-Rabbit 4.358 Events, Qodo 238, Cursor 76) mit 90-Tage-Trendcharts. | **Risiko:** ungoverned Schreibaktivität von Agenten. **Nutzen:** wissen, welche Agenten in der Software-Fabrik wirklich arbeiten — und welche nur Rauschen sind. |
| **Shadow-AI / Agent Discovery** | Entdeckte Agenten, die die Organisation nie registriert hat — mit der Zahl der Identitäten und Repositories, die jeder berührt (z. B. Claude auf 4 Identitäten / 4 Repos, Lovable 1 / 0). | **Risiko:** Shadow AI — unverwaltete Agenten und Tokens mit Code-Zugriff. **Nutzen:** vollständiges Inventar vor der Governance; nichts verändert Code unbemerkt. |
| **Agent vs. Human Attribution** | Jeder Commit aufgeteilt in agent-primär, mensch-primär und co-authored; PRs eröffnet / approved / kommentiert / gemergt — pro Agent. | **Risiko:** unklare Provenienz KI-generierten Codes. **Nutzen:** audit-fähiger Autorschafts-Trail; Review-Aufwand dort fokussiert, wo Agenten schreiben. |
| **License Utilization** | Aktive Nutzer vs. lizenzierte Seats pro Tool im Zeitverlauf — inklusive Nutzung ohne Seats (z. B. Claude Code: 4 aktive Nutzer, 0 Seats). | **Risiko:** unlizenzierte / Schatten-Nutzung außerhalb des Procurements. **Nutzen:** Seats richtig dimensionieren, ungenutzte Lizenzen zurückholen, formalisieren, was Teams bereits nutzen. |
| **Usage Activity & AI Spend** | Gesamte KI-Ausgaben mit Tagestrend, umschaltbar zwischen Ausgaben, akzeptierten Zeilen und Tokens — pro Tool und pro Nutzer. | **Risiko:** unkontrollierter Token-Burn. **Nutzen:** $ pro akzeptierter Zeile — die Unit Economics des KI-Codings, messbar pro Team. |
| **Model Adoption** | Unterschiedliche Modelle im Einsatz (15 im Beispielbestand), gerankt nach Nutzern — claude-haiku-4-5, claude-opus-4-8, composer-2.5-fast, grok-Varianten, Defaults. | **Risiko:** nicht freigegebene oder veraltete Modelle schreiben Code. **Nutzen:** Standardisierung, Policy pro Modell-Tier, Verhandlungshebel gegenüber Anbietern. |
| **Agentic Application Discovery** | Repositories mit AI/LLM-Nutzung: externe LLM-APIs, agentische Frameworks (LangChain, LangGraph, baml, MCP, CrewAI), selbst gehostete offene Modelle; Provider gerankt (Anthropic, AWS Bedrock, Open Source …). | **Risiko:** unbekannte KI-Lieferkette in eigenen Produkten — EU-AI-Act- / CRA-Exposure. **Nutzen:** ein maschinenlesbares KI-Inventar (AI-BOM) pro Repository. |
| **Repository-Level Detection** | Badges pro Repo: LLM-Provider, Framework, Pakete, agentische Konfiguration, Hosting-Klasse (cloud-managed vs. self-hosted), exakte Modell-Strings, Dateien, letztes Erkennungsdatum. | **Risiko:** Self-hosted- vs. Cloud-Datenabflusspfade für AppSec unsichtbar. **Nutzen:** gezielte Architektur- und Datenfluss-Reviews statt pauschaler Audits. |
| **Agent Identity Governance** | Jeder Agent eine vollwertige Identität mit Insights- / Activity- / Access-Tabs — Zugriff reviewbar wie bei jedem menschlichen Konto. | **Risiko:** Agenten außerhalb jedes IAM-Prozesses. **Nutzen:** Least Privilege und Access Reviews auf KI ausgeweitet — ein Governance-Modell. |
| **OverPrivilege Score & Blast Radius** | OverPrivilege Score pro Agent (z. B. Cursor: 60) plus wichtigste menschliche Kollaborateure und aktivste Repositories. | **Risiko:** überprivilegierte Agenten mit großem Blast Radius. **Nutzen:** priorisierte Remediation — der riskanteste Agenten-Zugriff wird zuerst behoben. |

*Feature-Set wie in BlueFlag AI Insights dargestellt (Insights → AI, Identity Governance → AI Agents, Agentic Applications); Beispielwerte aus einem Live-Demo-Bestand.*

DIE PLATTFORM-SICHT

# Ein Modul. Fünf Perspektiven auf agentische KI.

| I | II | III | IV | V |
|---|---|---|---|---|
| **Agent Activity** | **Shadow-AI Discovery** | **AI Economics** | **AI-BOM** | **Agent Governance** |
| Commits, PRs, Reviews pro Agent | Nicht registrierte Agenten & Identitäten | Seats, Ausgaben, Tokens, akzeptierter Code | Modelle, Frameworks, Provider pro Repo | Zugriff, OverPrivilege, Blast Radius |

## I SDLC Agent Activity Graph

Aktive Agenten mit realer 90-Tage-Aktivität — Commits durch agentische KI (KI vs. co-authored), PR-Aktivität — und jeder Commit attribuiert: **agent-primär vs. mensch-primär**.

OUTCOME

*Provenienz für jede KI-berührte Änderung — der Audit-Trail, den Regulatoren und Reviewer verlangen.*

## II Shadow-AI & Agent Discovery

Entdeckte Agenten, die niemand registriert hat, mit Identitäts- und Repository-Zählern — die Differenz zwischen dem KI-Bestand, den Sie lizenziert haben, und dem, den Sie tatsächlich betreiben.

OUTCOME

*Erst vollständiges Inventar, dann Governance — kein Agent verändert Code unbemerkt.*

## III AI Coding Economics

License Utilization (aktiv vs. Seats), Gesamtausgaben mit Tagestrend, **umschaltbar Ausgaben / akzeptierte Zeilen / Tokens** — plus Modell-Adoption über den gesamten Bestand.

OUTCOME

*Kosten pro akzeptierter Zeile und pro Team — Seats richtig dimensioniert, Schatten-Nutzung formalisiert, ROI messbar.*

## IV Agentic Application Discovery (AI-BOM)

Repositories mit LLM-Nutzung, klassifiziert nach API-Typ, agentischem Framework (LangChain, LangGraph, baml, MCP, CrewAI), Provider und Hosting — bis hin zu exakten Modell-Strings pro Repo.

OUTCOME

*Ein maschinenlesbares KI-Inventar — das Artefakt, das EU AI Act, CRA und Supply-Chain-Reviews voraussetzen.*

## V Agent Identity Governance & OverPrivilege

Jeder Agent ist eine vollwertige Identität mit Insights-, Activity- und **Access**-Tabs, einem **OverPrivilege Score**, den wichtigsten menschlichen Kollaborateuren und den aktivsten Repositories — reviewbar im selben Access-Review-Zyklus wie jeder Mitarbeiter.

OUTCOME

*Least Privilege auf KI ausgeweitet — der riskanteste Agenten-Zugriff zuerst remediert, der Blast Radius vor dem Incident bekannt.*

AGENTIC AI ECONOMICS — DIE WERTSCHÖPFUNGSKETTE, DIE DAS MODUL MESSBAR MACHT

**Ausgaben** → **Token-Verbrauch** → **Akzeptierter Code** → **Ausgelieferte Features** → **Business Value**

*Jedes Glied ist instrumentiert: Ausgaben und Tokens pro Tool/Nutzer · akzeptierte Zeilen pro Tool · agent-attribuierte Commits & PRs pro Repository · Repository-zu-Applikation-Mapping.*

## DER RETURN ON AI VISIBILITY

### 4 aktiv · 0 Seats

Claude Code produktiv im Einsatz — ohne einen einzigen lizenzierten Seat. **Schatten-Nutzung sichtbar gemacht** und in eine Procurement-Entscheidung überführt.

### $ / akzeptierte Zeile

Ausgaben, Tokens und akzeptierte Zeilen in einer Ansicht — **die Unit Economics**, die entscheiden, welche Tools und Modelle ihre Kosten verdienen.

### 15 Modelle

Modell-Adoption sichtbar gemacht — **Konsolidierung und Policy** pro Modell-Tier statt unkontrollierten Wildwuchses.

### Score 60

OverPrivilege Score eines Agenten plus seine aktivsten Repos — **Remediation priorisiert** nach Blast Radius, nicht alphabetisch.

## WARUM JEDER STAKEHOLDER PROFITIERT

| | |
|---|---|
| **CISO** | Jeder Agent eine Identität mit Access-Trail und OverPrivilege Score — **Shadow AI als Kategorie eliminiert**, agenten-geschriebener Code in jedem Incident oder Audit attribuierbar. |
| **CFO / FinOps** | Seats vs. aktive Nutzer, Ausgaben vs. Tokens vs. akzeptierter Code — **License-Right-Sizing und $-pro-akzeptierter-Zeile** ersetzen das Bauchgefühl in KI-Budgetentscheidungen. |
| **CTO / Platform** | Welche Agenten, Modelle und Frameworks tatsächlich liefern — **Velocity-Daten pro Tool**, um den Stack zu standardisieren, ohne Teams auszubremsen. |
| **IAM-Leitung** | Agenten treten in denselben Access-Review-Zyklus ein wie Menschen — **ein Least-Privilege-Modell** über Mitarbeiter, Contractors, NHIs und KI hinweg. |
| **GRC / Compliance** | AI-BOM pro Repository und Provenienz pro Commit — **die Evidenz, die EU AI Act, CRA, NIS2 und interne KI-Richtlinien voraussetzen**, kontinuierlich erzeugt. |

EXECUTIVE MESSAGE

*„Agentische KI ist die am schnellsten wachsende Belegschaft in Ihrer Software-Fabrik — und die einzige ohne Ausweis, ohne Vorgesetzten und ohne Kostenstelle. BlueFlag gibt ihr alle drei: eine Identität, ein Governance-Modell und eine P&L.“*

**BlueFlag Security** · Prevent first. Prove always.

**Sehen Sie Ihren realen agentischen KI-Bestand — Agenten, Ausgaben und Schatten-Nutzung.**

Read-only AI-Insights-Scan · Deployment in <30 Minuten · erste Ergebnisse in 48 Stunden

**blueflagsecurity.com**

*Informatives Solution Brief. Alle referenzierten Werte und Screenshots stammen aus einer Live-Beispielumgebung (90-Tage-Fenster) und dienen der Veranschaulichung; kundenspezifische Werte liefert der Read-only-Scan. Feature-Verfügbarkeit gemäß BlueFlags aktuellem AI-Insights-Modul (Beta-Features im Produkt gekennzeichnet).*