



Model as a Service – Zentrale AI-Modell-Verwaltung mit deutschem Proxy

Liebe Unternehmerinnen und Unternehmer,

Die Nutzung leistungsfähiger KI-Modelle wie GPT-4o, Gemini 2.5 oder Grok 4 stellt Unternehmen vor neue Herausforderungen: Datenschutz, Budgetkontrolle, technischer Zugang und Modellvielfalt.

Mit dem **MaaS Proxy** von **partimus** greifen Sie über eine zentrale API auf führende Modelle zu – abgesichert durch deutsche Infrastruktur, flexible Budgetverwaltung und optionalem Hosting für Open-Source-Lösungen.

Entwickeln, skalieren und integrieren Sie KI-Anwendungen mit maximaler Sicherheit und Kontrolle – ohne Wartelisten, versteckte Kosten oder Compliance-Risiken.

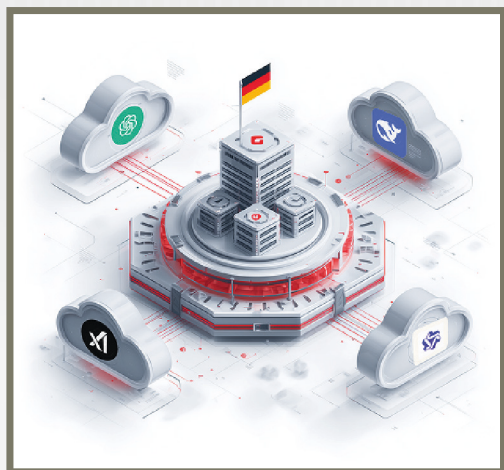
**DSGVO-konformer Zugang
zu Spitzenmodellen:**

-  DSGVO-konform gehostet in Deutschland
-  Zentrale Budgetkontrolle & API-Zugang
-  Sofort startklar – keine Wartezeit

Maximale Sicherheit & volle Kontrolle

Der **MaaS Proxy** wurde speziell für **höchste Datenschutzanforderungen** entwickelt und alle Anfragen werden über unsere Infrastruktur in deutschen Rechenzentren gehostet und verarbeitet. Der Proxy ist in Deutschland gehostet und verbindet lediglich zu den Partnern in den USA und der EU.

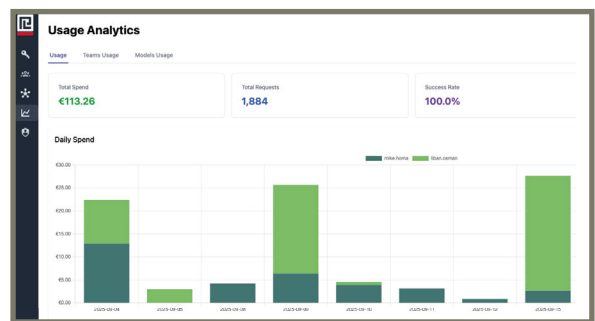
Durch umfangreiche Logging-, Audit- und Kontrollmechanismen behalten Sie jederzeit die Übersicht über Datenflüsse und Zugriffe. Auf Wunsch können sensible Modelle auch vollständig lokal gehostet werden – für volle Souveränität über Ihre KI-Anwendungen.



-  **Deutscher Proxy-Server**
mit deutschen Rechenzentren und Logging/Monitoring.
-  **Wahl zwischen EU und USA Modellen**
Für alle in der EU verfügbaren Premium Modelle.
-  **Optionales Deutschland-Hosting**
für maximale Datensouveränität.
-  **ISO 27001-Zertifizierung**
für unser Informations-Sicherheits-Managementsystem.
-  **Zentrale Zugriffssteuerung**
mit striktem Berechtigungs- und Rollenmanagement.

User Interface/Frontend Screenshots*

	Model**	Kosten pro 1M Input Tokens	Kosten pro 1M Output Tokens
	openai/gpt-5	1,40€	10,50€
	openai/gpt-5-mini	0,30€	2,10€
	openai/gpt-5-nano	0,10€	0,50€
	openai/gpt-oss-20b	0,20€	0,60€
	openai/gpt-oss-120b	0,20€	0,80€
	moonshotai/kimi-k2-instruct	1,10€	3,20€
	meta/llama-4-maverick-17b	0,30€	0,70€
	meta/llama3-70b	0,70€	0,90€
	meta/llama3-8b	0,10€	0,10€
	gemini/gemini-2.5-pro	1,40€	10,50€
	gemini/gemini-2.5-flash	0,40€	2,70€
	qwen/qwen3-coder-480b-a35b-instruct	0,50€	1,70€
	xai/grok-code-fast	0,30€	1,60€
	qwen/qwen3-coder-480b-a35b-instruct	0,50€	1,70€
	qwen/qwen3-235b-a22b-thinking	0,20€	0,70€
	zai/glm-4.5	0,60€	2,10€
	zai/glm-4.5-air	0,30€	1,20€
	deepseek/deepseek-v3.1	0,40€	1,10€
	deepseek/deepseek-r1-0528	0,60€	2,30€
	anthropic/claude-opus-4.1	15,80€	78,60€
	anthropic/claude-sonnet-4	6,30€	15,80€
	anthropic/claude-haiku-3.5	0,90€	4,20€



* Models and Pricing (links)
Usage Analytics (oben)
Virtual Keys (unten)

Virtual Keys

Create New Key

Filter: All teams

Key	Alias	Team	Created at	Expires	Spend	Budget
1a55a... ID		partimus	22.8.2025, 09:30:03	Never	€0.00	€20.00
485a4... ID		partimus	31.7.2025, 12:27:21	Never	€0.00	€50.00
65a4d... ID		partimus	17.7.2025, 15:36:53	Never	€0.00	€40.00
v18881... ID		partimus	11.7.2025, 16:10:51	Never	€0.00	€50.00
9b3f5a... ID		partimus	9.7.2025, 09:11:35	Never	€25.01	€25.00
827a63... ID		partimus	8.7.2025, 17:47:00	Never	€0.00	€20.00
29a08... ID		partimus	8.7.2025, 17:47:02	Never	€0.00	€20.00
80c91d... ID		partimus	8.7.2025, 17:46:53	Never	€10.71	€75.00
a77ba... ID		partimus	8.7.2025, 17:45:21	Never	€0.00	€50.00

Total Spend: €35.72
Total Budget: €410.00

** Alle Angaben und Preise Stand: 09/2025

Transparente Abrechnung – nutzungsbasiert oder pauschal Flexible Preismodelle für jedes Szenario

Ob Startup, Forschungsteam oder Konzern – unser Pricing passt sich Ihrem Bedarf an. Wählen Sie zwischen nutzungsbasierter Abrechnung pro Token, pauschalen Monatsplänen für selbstgehostete Modelle oder individuell vereinbarten Enterprise-Modellen. Sie zahlen nur, was Sie wirklich nutzen – bei voller Kostenkontrolle und klarer Übersicht.

	Starter	Professional	Enterprise
Entwickler-Accounts	5	25	Unbegrenzt
AI-Credit	50 €	200 €	Individuell
Proxy-Zugang	✓	✓	✓
Analytics	Basic	Advanced mit Reports	Advanced mit Reports
Team Managment	✗	✓	✓
Support	E-Mail Support	Priority Support	24/7 Premium Support
Custom SLAs	✗	✗	✓
Lokales Hosting	✗	✗	✓
Dedicated Infrastructure	✗	✗	✓
Preis	49 €/Monat*	199 €/Monat*	Individuell

* Alle Preise verstehen sich als monatliche/einmalige Gebühren zuzüglich Mehrwertsteuer.
Unser Angebot richtet sich ausschließlich an Gewerbetreibende.

Fundiert & praxisnah: Der partimus KI-Einführungsworkshop

Sie möchten moderne KI-Tools sicher, effizient und praxisnah in Ihrem Arbeitsalltag nutzen? In unserem Workshop erhalten Sie nicht nur einen fundierten Überblick über aktuelle Sprachmodelle und die Nutzung des zentralen Proxy-Zugangs, sondern lernen auch gezielt praxisbewährte Tools wie RooCode und Chatbox kennen. Gemeinsam setzen wir individuelle Use Cases um und zeigen Ihnen, wie Sie KI reibungslos und datenschutzkonform in Ihre Prozesse integrieren.

In interaktiven Übungen arbeiten Sie aktiv mit den Tools, erstellen eigene Chatbots und automatisierte Workflows und profitieren von konkreten Best Practices zur Kosten- und Budgetsteuerung. Nach nur einem Tag sind Sie bereit, KI verantwortungsvoll und produktiv in Ihrem Unternehmen einzusetzen.

Inhalte des Workshops:

-  **Einführung in moderne KI-Modelle**
Grundlagen zu aktuellen Sprachmodellen, Einsatzmöglichkeiten und Trends.
-  **Praxisnahe Tool-Einführung (RooCode & Chatbox)**
Vorstellung bewährter Tools für kollaboratives Coden und Workflow-Automatisierung.
-  **Interaktive Übungsphasen mit konkreten Use Cases**
Eigene Anwendungsfälle umsetzen und Tools praxisnah erleben.
-  **Schritt-für-Schritt-Anleitungen für Einsatzszenarien**
Klar verständliche Leitfäden für die Umsetzung im Arbeitsalltag.
-  **Fehlervermeidung & Best Practices im KI-Einsatz**
Praktische Tipps zur sicheren und effizienten Anwendung.
-  **Datenschutz & Compliance im Alltag**
Rechtssicherer Umgang mit KI in Ihrem Unternehmen.
-  **Kosten- und Budgetsteuerung in KI-Tools**
Strategien zur Kostenoptimierung und Ressourcenkontrolle.

Ihre Vorteile auf einen Blick

-  **Praxisnahes Online-Webinar (1 Tag) für Entwickler und Anwender**
-  **Einführung in moderne KI-Modelle und Unternehmens-Anwendungen**
-  **Schritt-für-Schritt-Anleitung zur Nutzung des AI-Proxys und der Budgetverwaltung**
-  **Schulung in Best Practices für Produktion, Sicherheit und Compliance**
-  **Aktive Tool-Demos: Umgang mit RooCode und Chatbox live erproben**
-  **Eigenständige Übungen & individuelle Fragen direkt im Workshop**
-  **Handout mit allen Workflows & Checklisten zur direkten Anwendung in Ihrem Team**
-  **Teilnahmezertifikat für alle Teilnehmenden**

Buchen Sie jetzt Ihren Workshop und machen Sie Ihr Team fit für den KI-Einsatz.



Ansprechpartner:

Jörg Müller

Telefonnummer:

+49 6431 9452-470

E-Mail:

vertrieb@partimus.com