

Intuitive AI Chat



Interact seamlessly with your local AI.

Limitless Open-Source AI

Choose from endless open-source models for any purpose.



razzfazz.ai
local matters



Local AI Powerhouse.
Uncompromising Security.

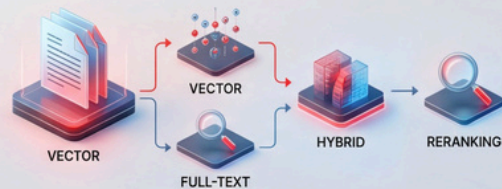
Our [razzfazz.ai Box](#) is powered by [razzfazz.io GmbH](#)

Agentic Workflow Automation



Bring local intelligence into your processes.

Advanced Knowledge Retrieval



Powerful local RAG databases with Hybrid Search & Reranking.

Total Data Sovereignty



Legal



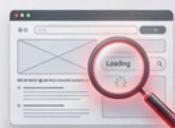
Development



Personal Data

Built for Legal, Development, and sensitive personal data.
Zero data leakage to the cloud.

Live Web Research



Deep, secure web research functionality.

Seamless Integration

SSO

REST API

MCP

Integrates perfectly into your existing IT environment.

Produktübersicht: razzfazz Box

Ihre private KI-Infrastruktur im eigenen Netzwerk.

Was ist die razzfazz Box?

Es handelt sich um einen vorkonfigurierten Rechner, der Ihnen den Einstieg in künstlichen Intelligenz – lokal, sicher und unabhängig - ermöglicht. Statt auf externe Cloud-Lösungen angewiesen zu sein, betreiben Sie mit razzfazz.ai lokal Open-Source-Modelle direkt in Ihrer eigenen Infrastruktur. Dies garantiert volle Datenkontrolle ohne Abfluss sensibler Informationen nach außen, bei gleichzeitig hoher Energieeffizienz durch optimierte Hardware-Architekturen.

Die **razzfazz.ai-Box** lässt sich als eine physische und technologische **KI-Werkzeugbox** charakterisieren, die nach dem Prinzip der Souveränität konzipiert ist. Anstatt ein starres, abgeschlossenes System zu konsumieren, wird hier ein modularer Baukasten bereitgestellt, der explizit auf die Anpassung und Weiterentwicklung durch den Anwender ausgelegt ist.

Aus der Perspektive einer veränderbaren Werkzeugbox lässt sich das Konzept wie folgt positionieren:

Die Basisausstattung: Ein einsatzbereites Fundament

Die Box wird als betriebsbereite Einheit geliefert, die Hardware und Software kombiniert. Man erhält ein System, das durch lokale Rechenpower (GPUs) und vorinstallierte Open-Source-Komponenten sofort funktionsfähig ist. Es handelt sich um ein Fundament, auf dem spezifische Lösungen aufgebaut werden können, ohne dass eine Abhängigkeit von externen Cloud-Infrastrukturen besteht.

Die Modularität: Werkzeuge wählen und austauschen

Ein wesentliches Merkmal dieser Werkzeugbox ist die Austauschbarkeit der Inhalte. Man ist nicht auf ein einziges KI-Modell festgelegt, sondern kann je nach Aufgabenstellung unterschiedliche „Werkzeuge“ (Open Source Large Language Models wie Llama, Mistral oder spezialisierte Modelle) integrieren. Diese Flexibilität erlaubt es, für jede Abteilung oder jeden Anwendungsfall das exakt passende Instrument bereitzustellen.

Die Individualisierung: Integration von Eigenwissen

Die Werkzeugbox ist darauf ausgelegt, mit unternehmenseigenen Daten „angereichert“ zu werden. Durch Verfahren wie **Retrieval Augmented Generation (RAG)** lassen sich interne Dokumente, Datenbanken und Prozessbeschreibungen in das System einspeisen. Dadurch verwandelt sich die Box von einem allgemeinen Assistenten in ein spezialisiertes Werkzeug, das den Kontext der jeweiligen Organisation präzise verstehen kann.

Die Skalierbarkeit: Erweiterung der Kapazitäten

Das Konzept sieht vor, dass die Werkzeugbox bei steigenden Anforderungen mitwachsen kann. Man kann die Box verstärken, indem man mehrere Einheiten (weitere Boxen oder Cloudinstanzen) zu einem Verbund zusammenschließt. Diese Form der Skalierbarkeit stellt sicher, dass die Infrastruktur nicht zu einem Engpass wird, wenn die Nutzung im Unternehmen zunimmt.

Die Souveränität: Kontrolle über den Werkzeugkasten

Im Gegensatz zu Software-as-a-Service-Modellen bleibt die vollständige Kontrolle über Konfiguration, Updates und Datenfluss beim Betreiber der Box. Man kann die Box modifizieren, Sicherheitseinstellungen individuell festlegen und sicherstellen, dass sensible Informationen den physischen Standort niemals verlassen. Die Box fungiert somit als privater Schutzraum für Innovationen.

Zusammenfassend: Die razzfazz.ai Box ist eine On-Premise-Plattform, die als flexibler Rahmen dient. Man nutzt sie nicht nur, sondern gestaltet, administriert und optimiert sie kontinuierlich, um sie perfekt in die bestehenden Arbeitsabläufe zu integrieren."

Technische Kernaspekte

- Hardware: Ausgestattet mit leistungsstarken Grafikkarten (z. B. 128 GB Unified Memory, wovon ca. 96 GB netto für KI nutzbar sind).
 - Software-Stack: Basiert auf Ubuntu Linux und nutzt die "razzfazz.ai Distribution", die Tools wie llama.cpp (zum Ausführen von Modellen wie Llama, Mistral oder DeepSeek) und PostgreSQL mit Vektorerweiterung enthält.
 - Benutzeroberfläche: Die Interaktion erfolgt über das Open WebUI, das eine ChatGPT-ähnliche Erfahrung im Browser bietet.
 - Skalierbarkeit: Mehrere Boxen können zu einem AI-Cluster zusammengeschaltet werden, um höhere Lasten zu bewältigen oder Redundanz zu schaffen.
-

Vorteile

- Datensouveränität: Da alle Daten lokal verarbeitet werden, bleibt die volle Kontrolle im Unternehmen. Dies ist besonders im Hinblick auf den EU AI Act und Compliance-Anforderungen relevant.
- Kosteneffizienz: Im Vergleich zu laufenden Cloud-Abonnements (Pay-per-Token) fallen nach der Anschaffung (Kauf) keine nutzungsabhängigen Kosten an.
- Automatisierung: Neben reinem Chat bietet die Box Werkzeuge zur Workflow-Automation (z. B. für Digitalisierungsprozesse, Speech-to-Text oder Text-to-Speech).
- Nachhaltigkeit: Die Box ist für zielgerichtete, lokale Anfragen energieeffizienter als gigantische Cloud-Rechenzentren.