

F13 in Sachsen

Der KI-Textassistent für die Verwaltung



Die Ausgangslage

Der Bedarf an KI-Unterstützung ist da

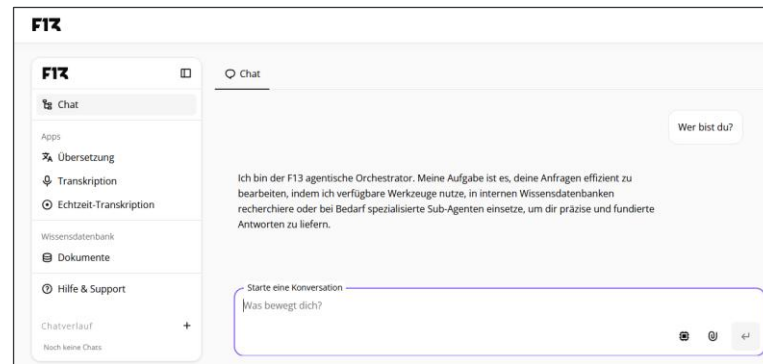


F13 – KI-Textassistenz für die Verwaltung

Was und Warum

- Entwickelt von Baden-Württemberg
- Modulare Microservice-Architektur, modellagnostisch
- Open Source seit Juli 2025,** Community-basierte Weiterentwicklung
- Unser Ziel für Sachsen:

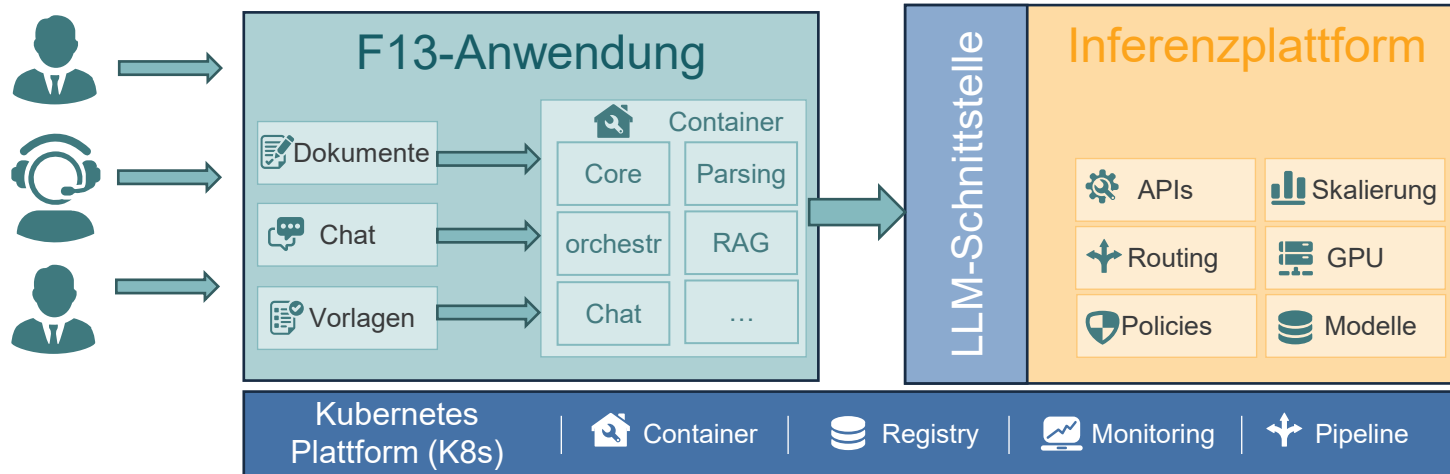
F13



- ➔ Unterstützung der Mitarbeiter bei typischen Büro-, Wissens- und Textaufgaben
- ➔ über einen sicheren, zentralen Zugang zu KI-Funktionen für die Landesverwaltung.

F13 – Mehr als eine einfache Webanwendung

Ein Ökosystem



Was hinter initial „bereitstellen“ steckt

Recht & Datenschutz



- | Datenschutzfolgeabschätzung
- | Nutzungsregeln
- | Haftung und Verantwortung
- | Rechtliche Rahmenbedingungen

Sicherheit & Betrieb



- | Betriebs- und Rollenkonzept
- | BSI-Grundschutz / Sicherheitskonzept
- | Protokollierung & Nachvollziehbarkeit
- | Absicherung gegen KI-spezifische Angriffe

Organisation & Governance



- | Beteiligung der Personalvertretung
- | Rollen und Verantwortlichkeiten
- | Modellauswahl und Freigabe
- | Wissensbereitstellung behördenübergreifend/spezifisch

Technik



- | K8s-Plattform produktiv für F13
- | GPU-Anbindung eingerichtet
- | Skalierbare Inferenzplattform
- | F13-Anwendung: Installation + Konfiguration + Monitoring

Was hinter initial „bereitstellen“ steckt

Recht &

- | Datenschutz
- | Nutzungs
- | Haftung
- | Rechtlich

Organis

- | Beteiligu
- | Rollen u
- | Modellauswahl und Freigabe
- | Wissensbereitstellung behördenübergreifend/spezifisch

„Bereitstellung“ ist der Anfang: Danach Operativer Betrieb

Service – Komponenten warten, Weiterentwicklung

Modelle – Bewerten – aktualisieren – testen

Prompts – Templates – optimieren – standardisieren

Wissen – RAG – Wissensbasen – Berechtigungen

Daten & Qualität – Aktualisierungen – Monitoring – Nutzerfeedback



| Skalierbare Inferenzplattform

| F13-Anwendung: Installation + Konfiguration + Monitoring

Unser Weg zur produktiven F13-Plattform

Produktiv aufbauen, parallel erproben

Ziel

Produktive K8s-Plattform

Cluster – GPU – Harbor – Git-Pipeline - Betriebsbasis



F13 Release 2

**Neuaufbau auf
der Zielplattform**



(statt Migration von
Release 1)

Erprobung

K3s-Umgebung

Aufbau F13 Release 1 – Inferenzplattform - Schnittstellen

Aktueller Roll-Out-Pfad

Basis bereitstellen – parallel weiterentwickeln



Was F13 nicht ist

Grundsätzliche Grenzen und bewusster Umfang der ersten Rollout-Stufe

Grundsätzlich nicht

- Keine automatisierte Verwaltungsentscheidung
- Kein Ersatz für Fachverfahren
- Ergebnisse sind keine finalen Dokumente, sondern Arbeitsentwürfe

In Rollout-Stufe 1 noch nicht

- Kein Zugriff auf interne Wissensquellen
- Keine Internetrecherche
- Keine Transkription oder Bildverarbeitung
- Keine Historisierung und Wissensdatenbanken

Systembedingt

- Modelle können plausibel klingende Fehler erzeugen
- Veraltetes Modellwissen
- Ergebnisse müssen fachlich geprüft werden

Erkenntnisse aus dem bisherigen Aufbau

Was sich in der praktischen Umsetzung bestätigt und konkretisiert hat

- 1. F13 bereitstellen heißt Plattform aufbauen**
(Die Anwendung ist nur ein Teil – Inferenz, Infrastruktur und Betrieb gehören dazu.)
- 2. KI-Betrieb ist kein Standard-IT-Betrieb**
(GPU, Modelle und Inferenz bringen eigene Betriebsparadigmen und Fehlerbilder mit.)
- 3. Im Kleinen erproben zahlt sich aus**
(Deployment, Schnittstellen und Fehlerbilder lassen sich früh und risikoärmer kennenlernen.)
- 4. Technik allein bestimmt nicht das Tempo**
(Datenschutz, Sicherheit, Mitbestimmung und organisatorische Klärungen müssen parallel mitwachsen.)
- 5. Das geht nur gemeinsam**
(Plattform, Anwendung, KI, Sicherheit und Organisation müssen als ein Team arbeiten)



F13 in Sachsen

Der KI-Textassistent für die Verwaltung

Vielen Dank! Fragen?