

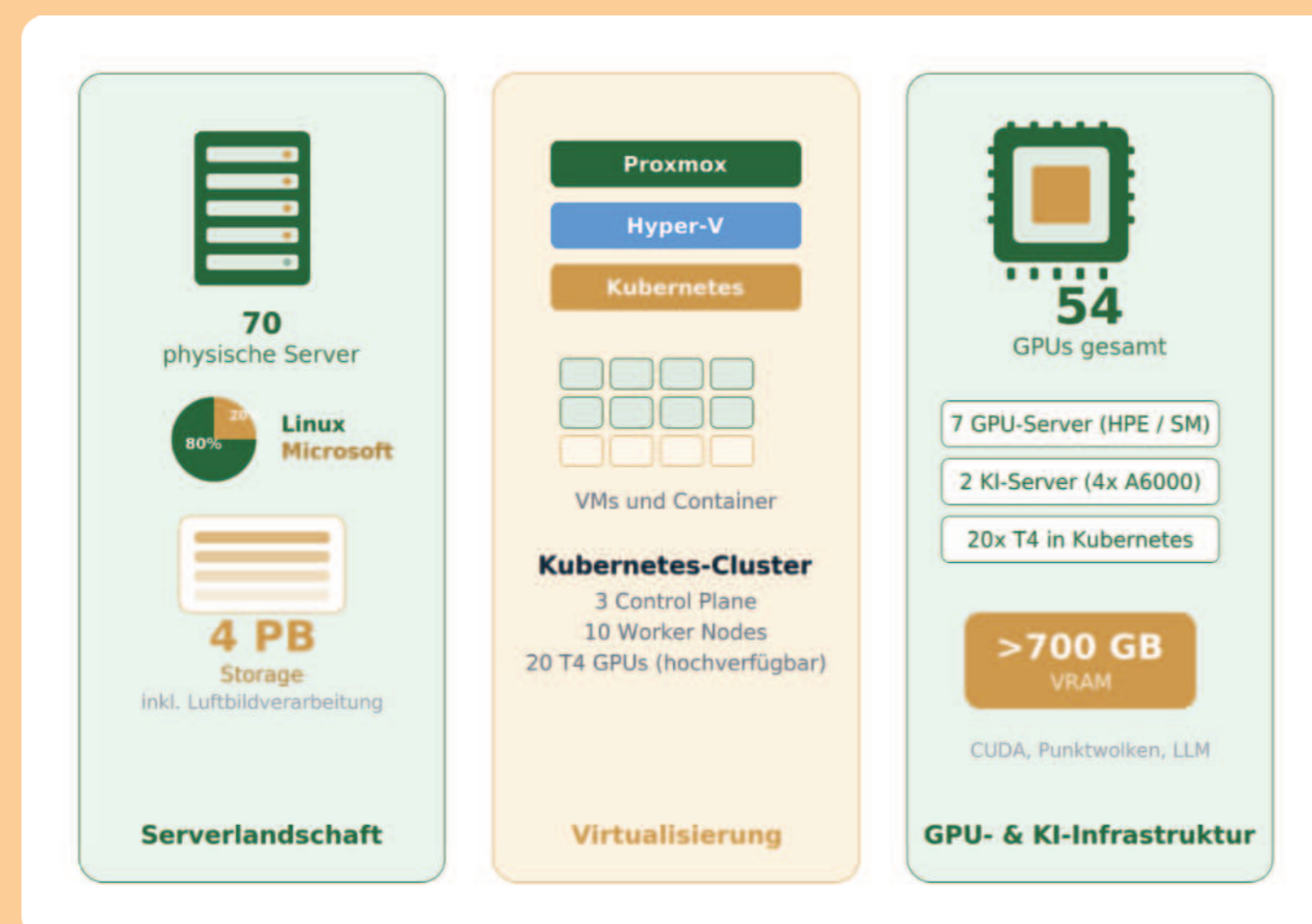
Bundesanstalt für Agrarwirtschaft und Bergbauernfragen

Einleitung

Die Bundesanstalt für Agrarwirtschaft und Bergbauernfragen (BAB) ist das sozioökonomische Forschungsinstitut des Bundesministeriums für Land- und Forstwirtschaft, Klima- und Umweltschutz, Regionen und Wasserwirtschaft (BMLUK).

Die Forschungsschwerpunkte umfassen Agrarpolitik und Ernährungswirtschaft, Berggebiete und nachhaltige Landwirtschaft, die Evaluierung der Gemeinsamen Agrarpolitik (GAP) sowie die Funktion als Datenkompetenzzentrum für INVEKOS, GIS und Buchführungsdaten.

Künstliche Intelligenz wird an der BAB systematisch in Forschung, Datenanalyse und Softwareentwicklung eingesetzt — auf einer eigenständigen IT-Infrastruktur mit 70 physischen Servern, 54 GPUs und über 700 GB VRAM.



Die Firewall als Grenze

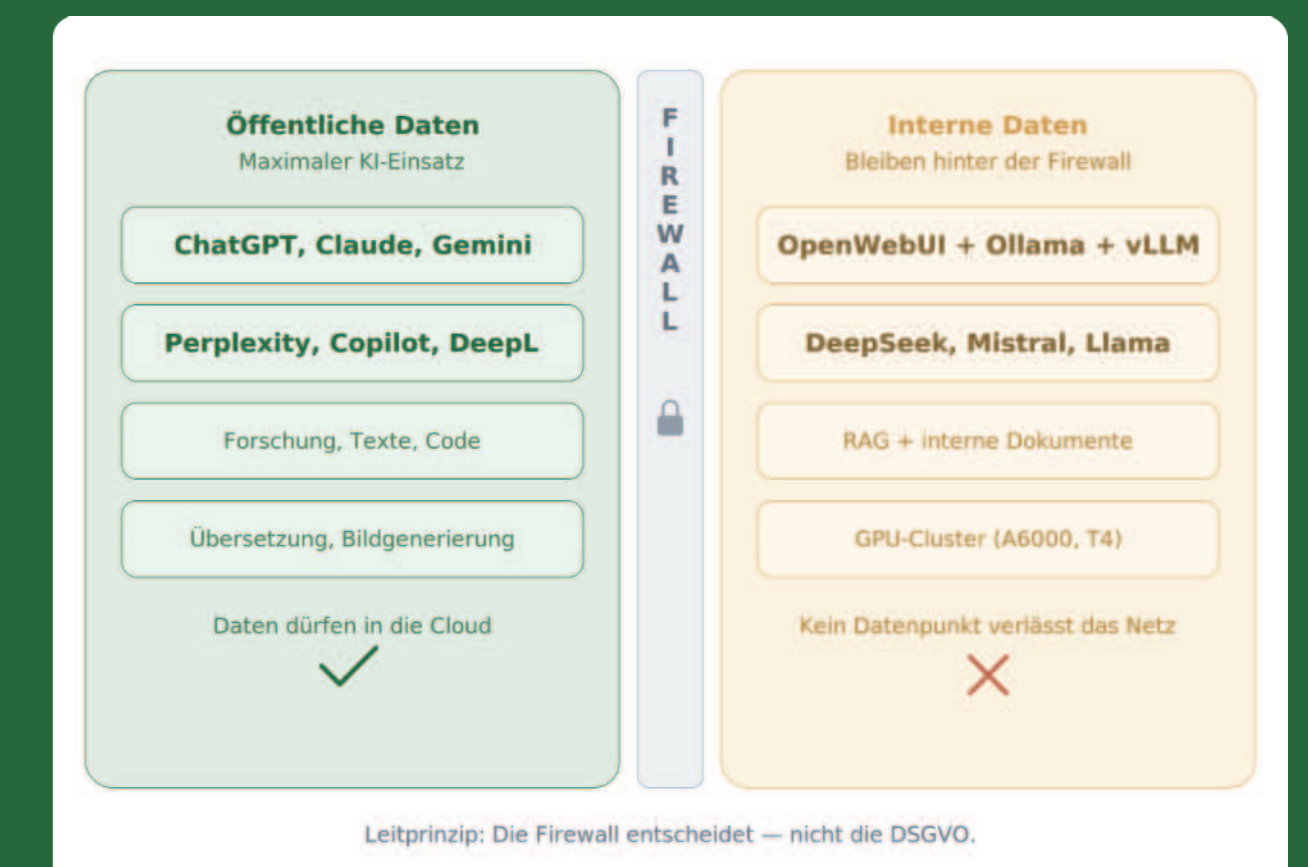
Nicht die DSGVO ist die Grenzlinie – sondern die BAB-Firewall.

Öffentliche Daten => Maximaler KI-Einsatz

Jede KI nutzen, die hilft, die Arbeit zu erledigen: ChatGPT, Claude, Gemini, Perplexity, Copilot, DeepL ... Forschung, Texte, Code, Übersetzung, Bildgenerierung

Interne Daten => Bleiben hinter der Firewall

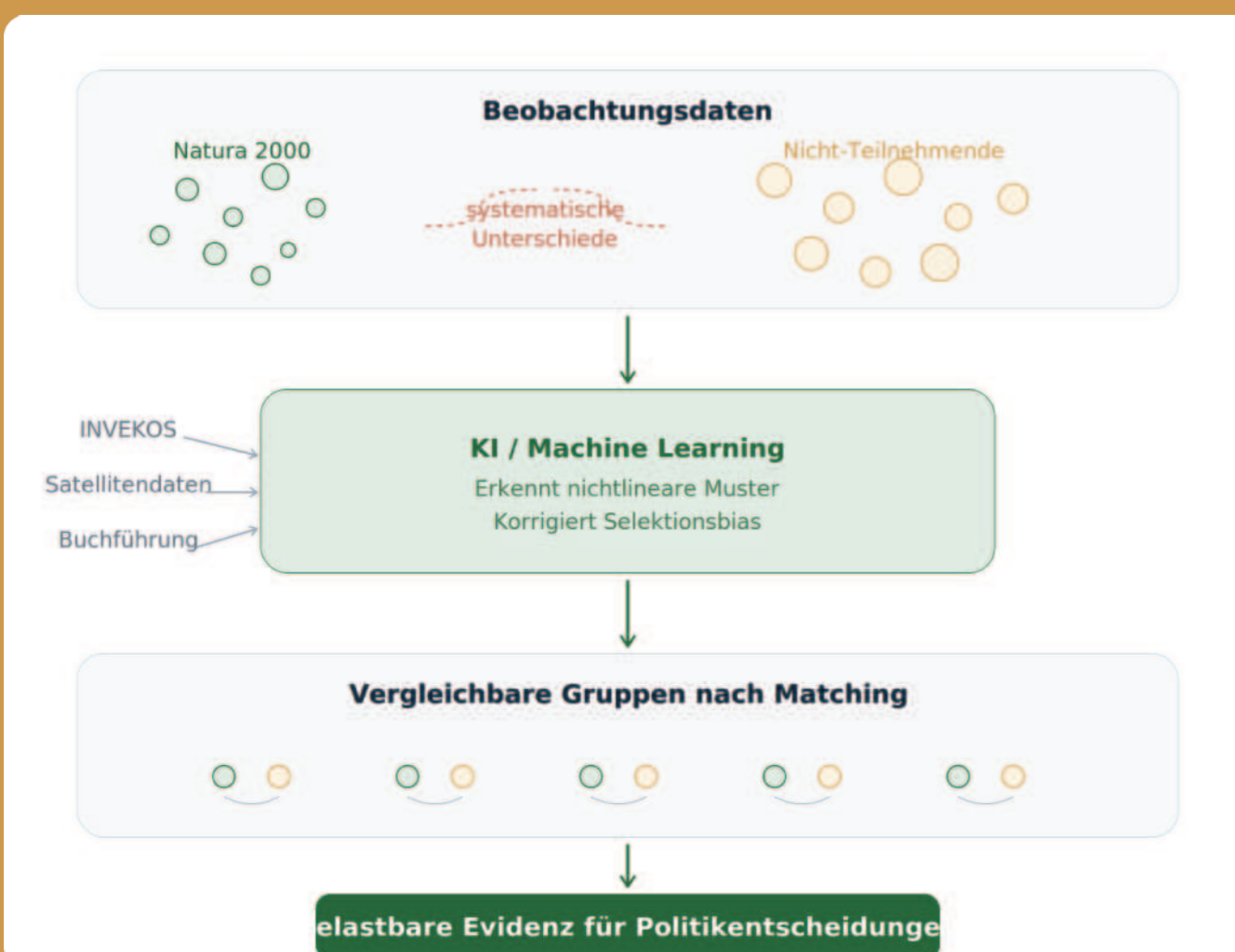
Eigene GPU-Server mit OpenWebUI + Ollama + vLLM, DeepSeek, Mistral, Llama – lokal, keine Cloud, RAG mit internen Dokumenten, Datenbank-Abfragen, Kein Datenpunkt quert die Firewall



Machine Learning

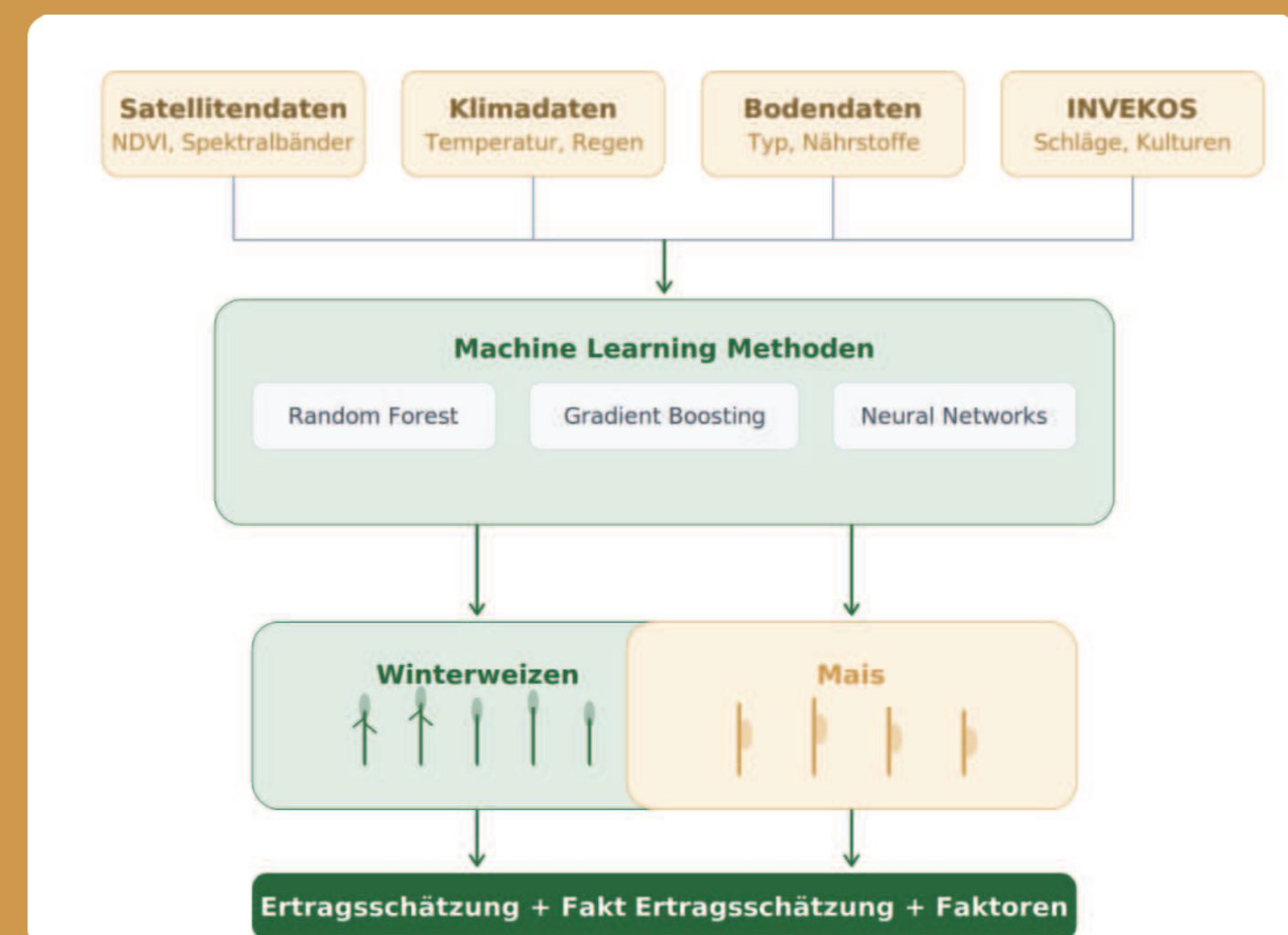
BAB 81/25: Der Einfluss von Natura 2000 auf die Produktivität und Effizienz landwirtschaftlicher Betriebe in Österreich (Kömle, Stickler, Tribl, Lakner).

Politikmaßnahmen wirken selten zufällig — Teilnehmende und Nichtteilnehmende unterscheiden sich häufig systematisch. Moderne Machine-Learning-Methoden können diese Selektionsverzerrungen erkennen und korrigieren, auch bei hochkomplexen, nichtlinearen Zusammenhängen und großen, satellitengestützten Datensätzen. Durch den Einsatz von ML-basiertem Matching wird aus Beobachtungsdaten belastbare Evidenz für politische Entscheidungen gewonnen — ohne aufwändige experimentelle Designs.



BAB 40/20: Biophysikalische Prozesse der landwirtschaftlichen Bodennutzung in Österreich (Stickler, Kömle).

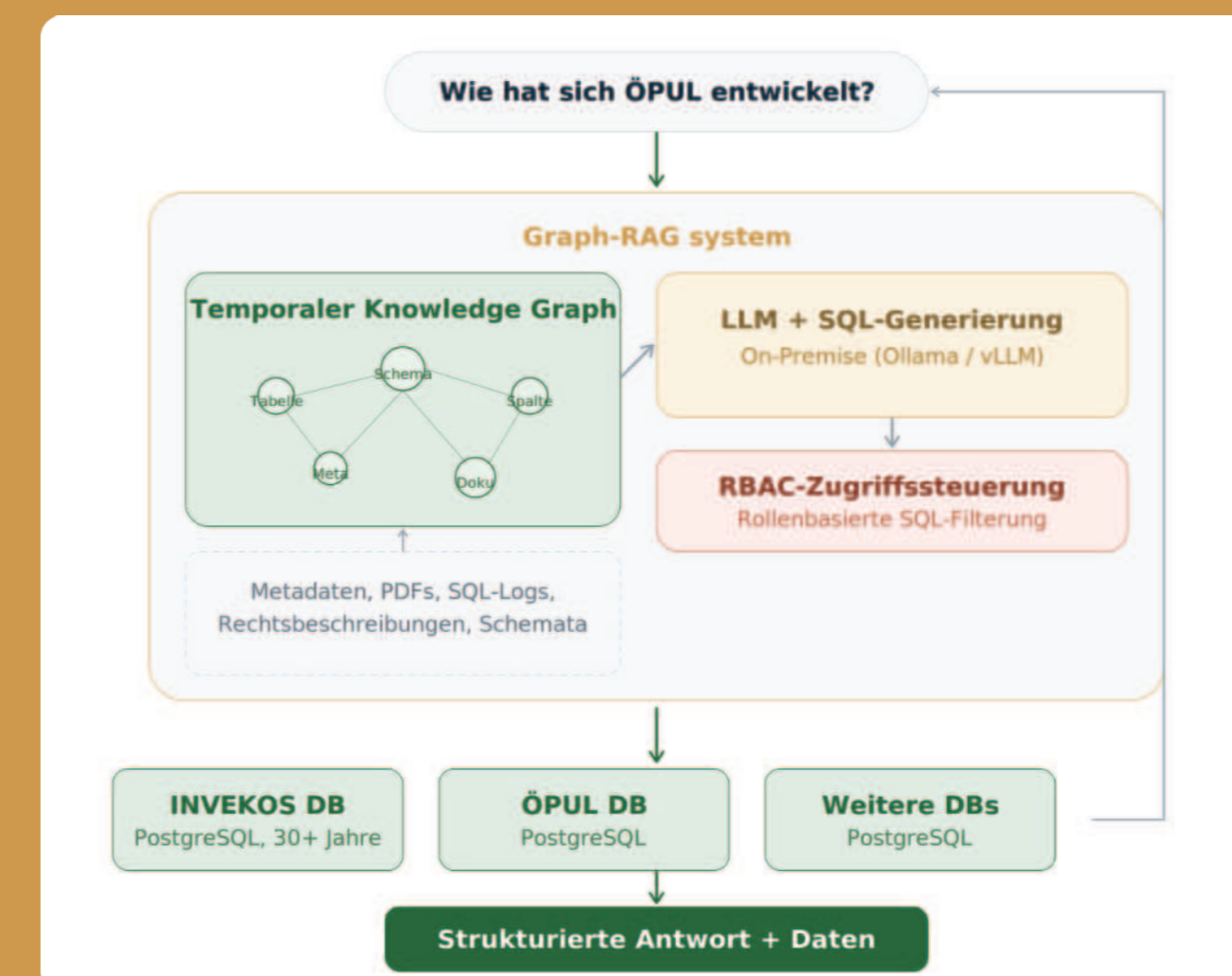
In diesem Projekt werden Ertragsschätzungen und Analysen der Bestimmungsfaktoren für Winterweizen und Mais durchgeführt. Zum Einsatz kommen verschiedene Machine-Learning-Methoden — auf Basis von Satellitendaten (NDVI, Spektralbänder), Klimadaten, Bodendaten und INVEKOS-Schlagdaten.



Retrieval-Augmented Generation

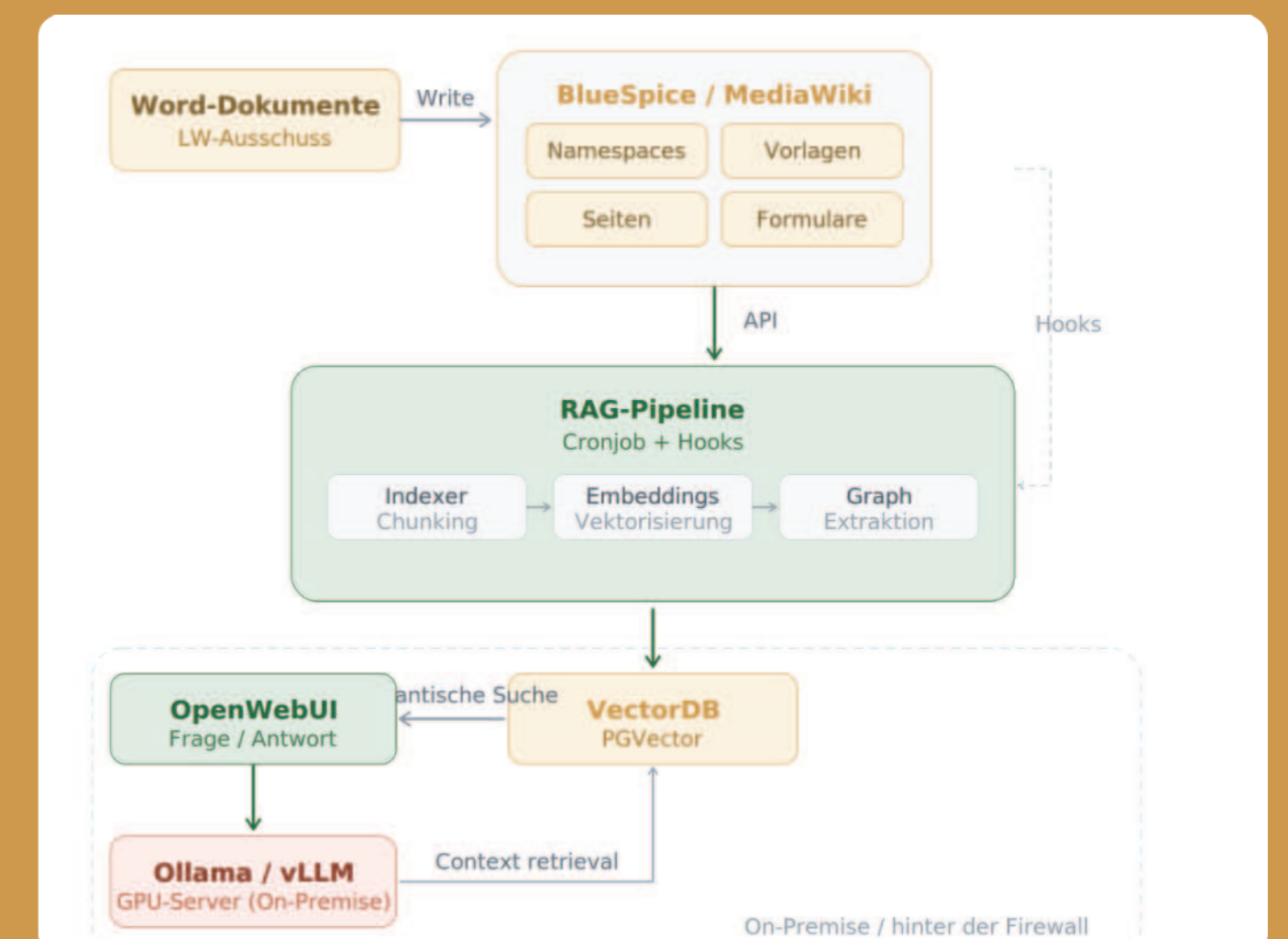
BAB 094/26: Conversations with Data Systems (Maria et al.).

Die landwirtschaftlichen Förderdaten der BAB — darunter INVEKOS und ÖPUL — umfassen über 30 Jahre Aufzeichnungen in mehreren PostgreSQL-Instanzen mit heterogenen Schemata und verteilter Dokumentation. Das Projekt entwickelt ein Graph-RAG-System, das einen temporalen Knowledge Graph über diese Datenbankstrukturen aufbaut, Schema-Evolutionen abbildet und natürlichsprachige Abfragen in RBAC-konformer SQL-Queries übersetzt. Die LLM-Inferenz erfolgt vollständig on-premise über Ollama und vLLM auf den institutseigenen GPU-Servern.



BAB 085/25: Ministerinformationssystem (Neissl, Maria et al.).

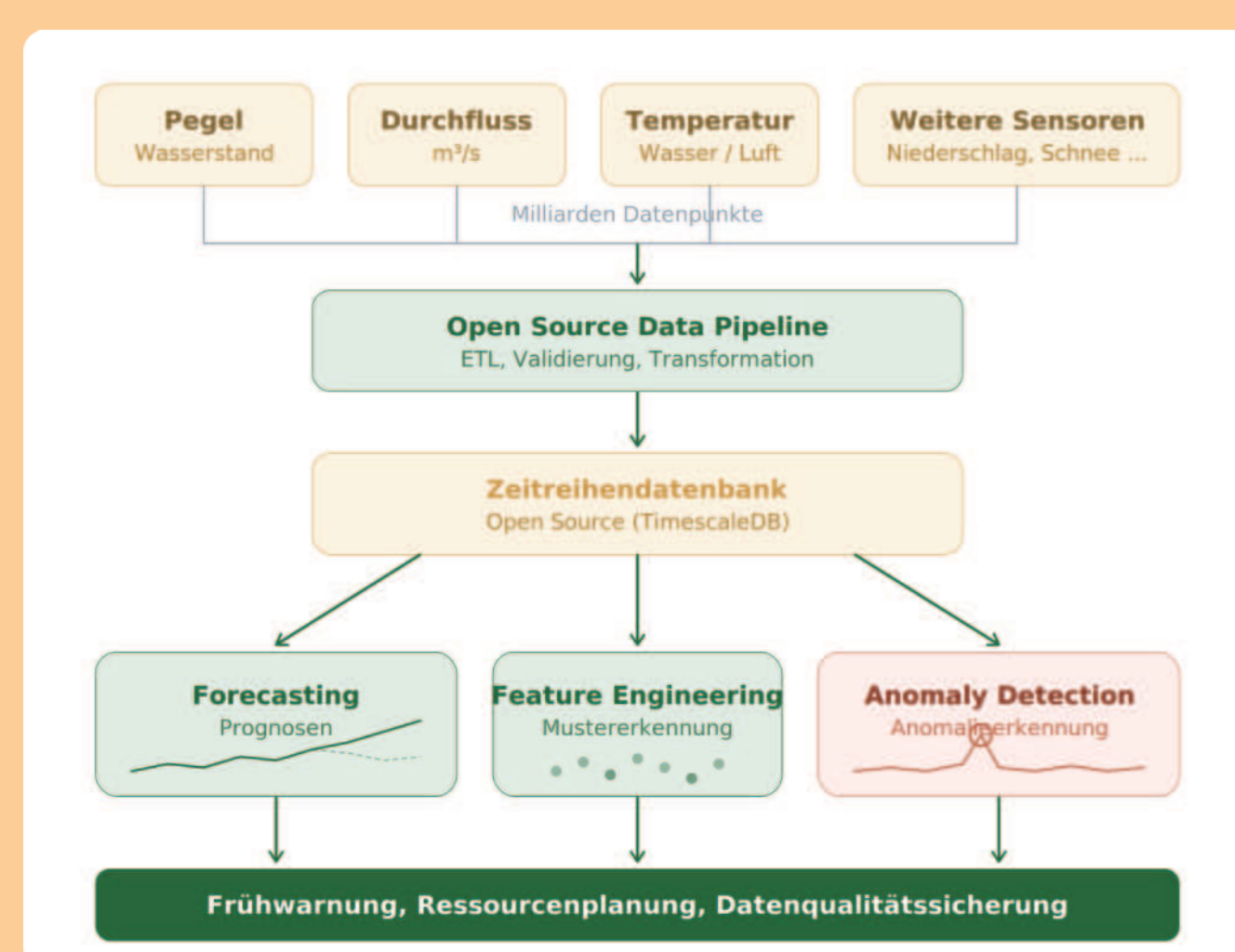
Aufbauend auf BlueSpice/MediaWiki als zentraler Knowledgebase wird ein RAG-System implementiert, das Wiki-Inhalte über eine automatisierte Pipeline (Cronjob und Hooks) indiziert, in Embeddings transformiert und in einer Vektordatenbank (PGVector) speichert. Die Abfrage erfolgt über OpenWebUI mit Ollama als LLM-Backend. Der gesamte Stack läuft containerisiert im Kubernetes-Cluster der BAB.



Forecasting & Modellierung

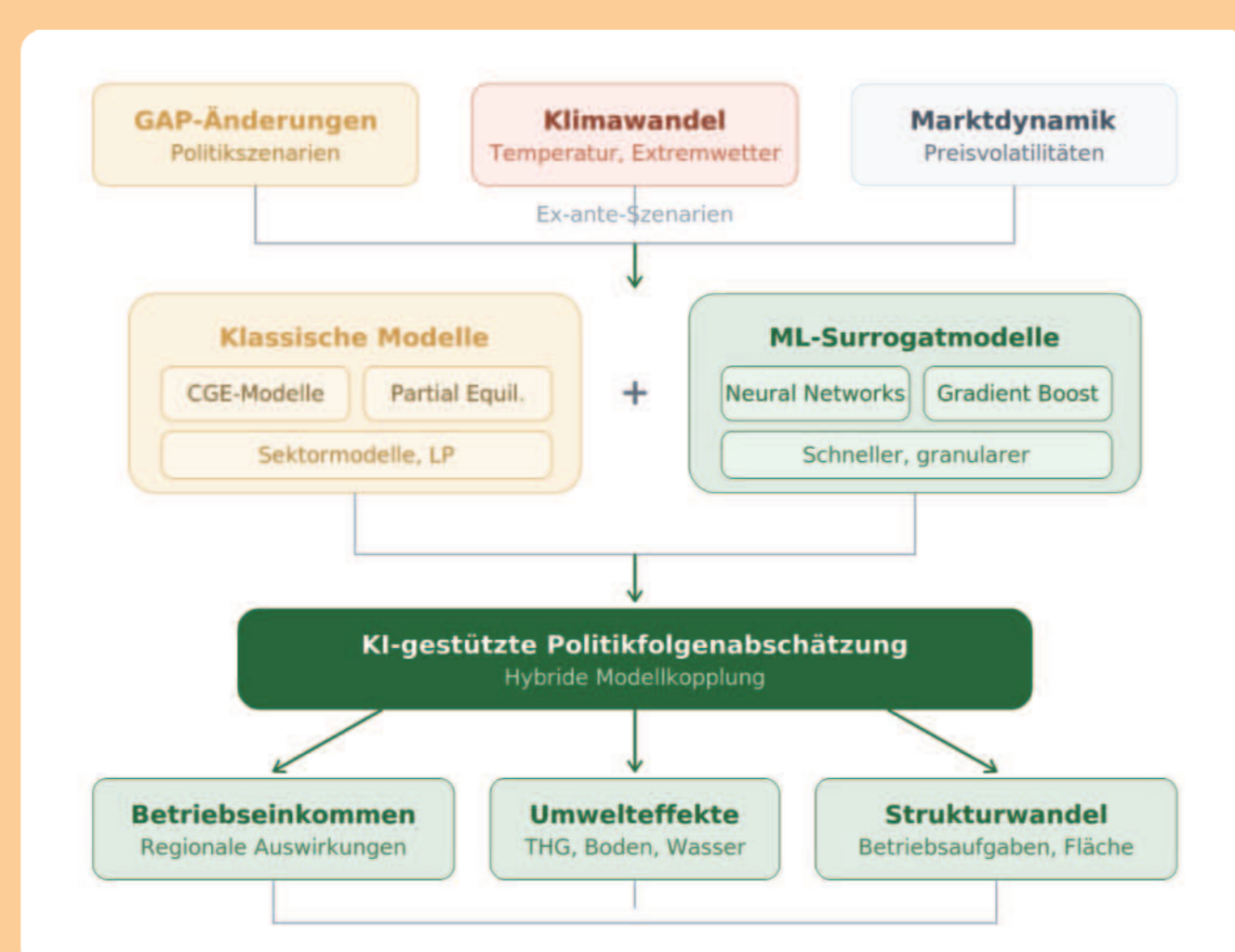
BAB 069/24: Open Source Data Pipeline & Database for HyDaMS (Zeglövits et al.).

Im Rahmen des Hydrographischen Daten-Management-Systems werden Milliarden von Sensordaten — Pegelstände, Durchfluss, Temperatur, Niederschlag — in eine Open-Source-Zeitreihendatenbank (TimescaleDB) exportiert. Darauf aufbauend werden KI-gestützte Methoden für Forecasting (Zeitreihenprognosen), Feature Engineering (automatische Mustererkennung) und Anomaly Detection (Erkennung ungewöhnlicher Messwerte) entwickelt. Die Ergebnisse dienen der Frühwarnung, Ressourcenplanung und Datenqualitätssicherung.



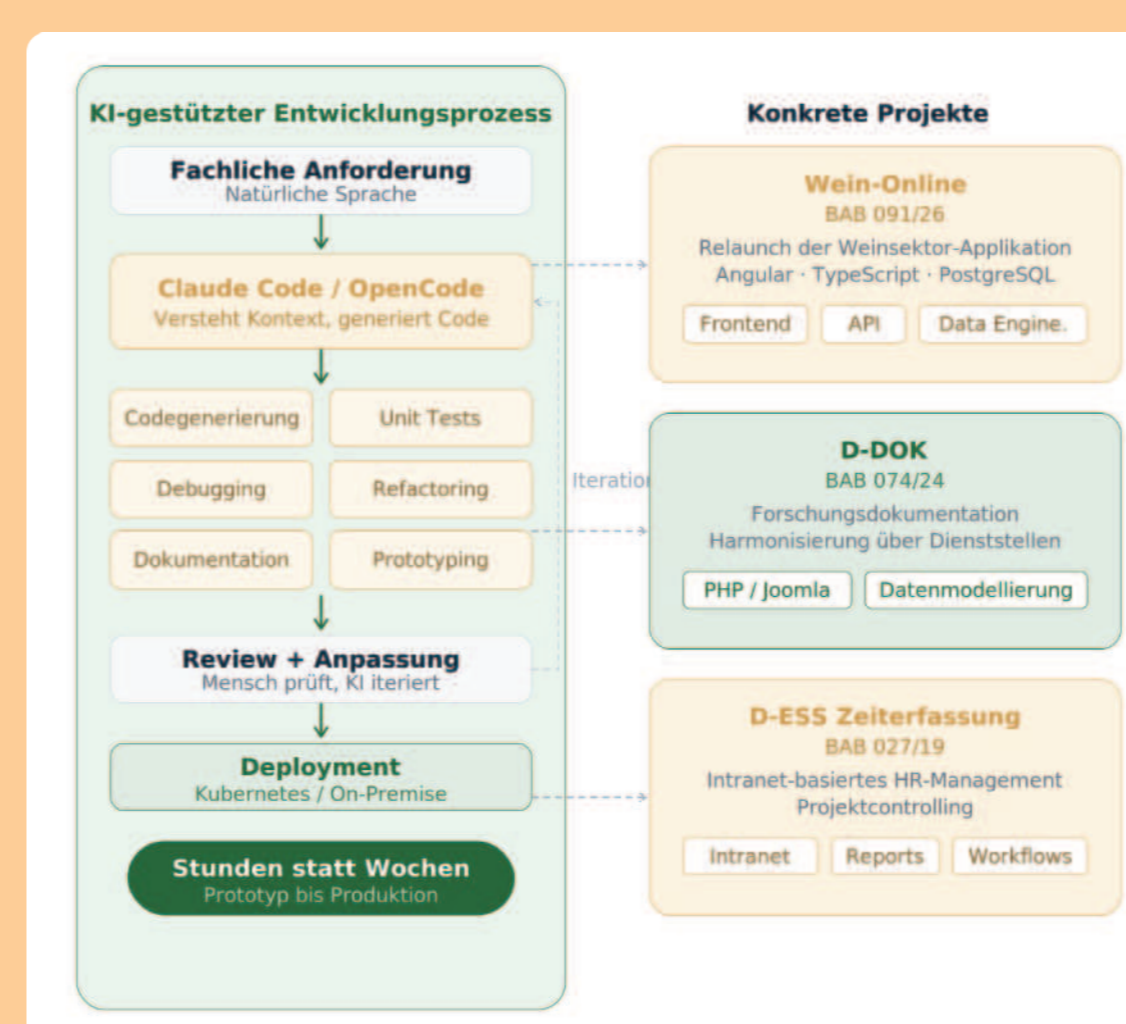
BAB 090/26: Grundlagen und Infrastruktur für modellgestützte Analysen — GRI-MA (Tribl et al.).

Zur Abschätzung der Folgen agrarpolitischer Maßnahmen oder externer Einflussfaktoren wie Klimawandel und Preisvolatilitäten wird ein hybrider Modellansatz verfolgt: Klassische Gleichgewichtsmodelle (CGE, Partial Equilibrium, nationale Sektormodelle) werden mit ML-basierten Surrogatmodellen (Neural Networks, Gradient Boosting) kombiniert. Diese Kopplung ermöglicht es, Ex-ante-Szenarien schneller und granularer durchzurechnen als mit traditionellen Methoden allein. Die Ergebnisse umfassen Auswirkungen auf Betriebseinkommen, Umwelteffekte (THG, Boden, Wasser) und Strukturwandel (Betriebsaufgaben, Fläche).

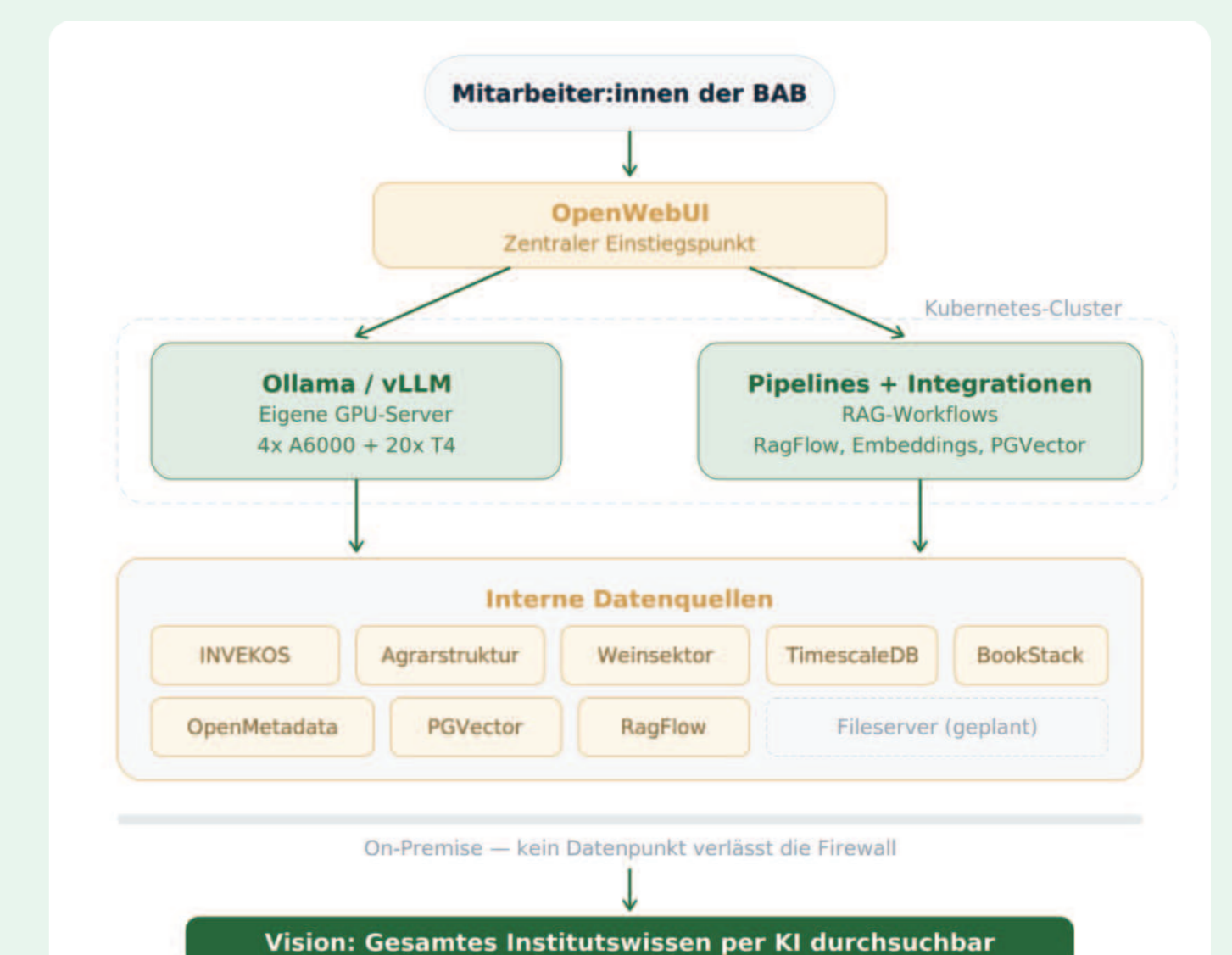


Softwareentwicklung

Die BAB setzt KI-Coding-Assistenten — insbesondere Claude Code und OpenCode — systematisch in der internen Softwareentwicklung ein. Der Arbeitsablauf folgt einem iterativen Dialog: Fachliche Anforderungen werden in natürlicher Sprache formuliert, die KI generiert Code, Tests und Dokumentation, der Mensch prüft und steuert den Prozess. Die sechs Kernfunktionen umfassen Codegenerierung, automatisches Debugging, Refactoring, Dokumentation, automatisiertes Unit-Tests und Rapid Prototyping. Drei aktive Projekte profitieren von diesem Ansatz: Wein-Online (BAB 091/26) — der komplette Relaunch der Weinsektor-Applikation mit Angular, TypeScript und PostgreSQL; D-DOK (BAB 074/24) — die Digitalisierung und Harmonisierung der Forschungsdokumentation über mehrere Dienststellen; sowie D-ESS (BAB 027/19) — ein intranetbasiertes HR-Management- und Projektcontrolling-System.



BABsi



Die BAB baut eine integrierte KI-Plattform auf, die über OpenWebUI als zentralen Einstiegspunkt alle internen Datenquellen zugänglich macht. Im hochverfügbaren Kubernetes-Cluster laufen Ollama und vLLM als LLM-Backend, verbunden über Pipelines und Integrationen mit PostgreSQL (INVEKOS, Agrarstruktur, Weissektor), TimescaleDB, PGVector, BookStack, OpenMetadata und RagFlow. Langfristiges Ziel ist die schrittweise Anbindung aller Fachdatenbanken, der Forschungsdokumentation und perspektivisch des gesamten Fileservers — indiziert und per RAG durchsuchbar. Das gesamte Institutswissen soll den Forschenden als intelligenter Assistent zur Verfügung stehen: auf eigener Infrastruktur, mit eigenen Modellen, unter voller Kontrolle.

