Datenplattform KI-ALLIANZ Baden-Württemberg

Die KI-Datenplattform Baden-Württemberg

Daten- und Modellinventare intelligent vernetzen

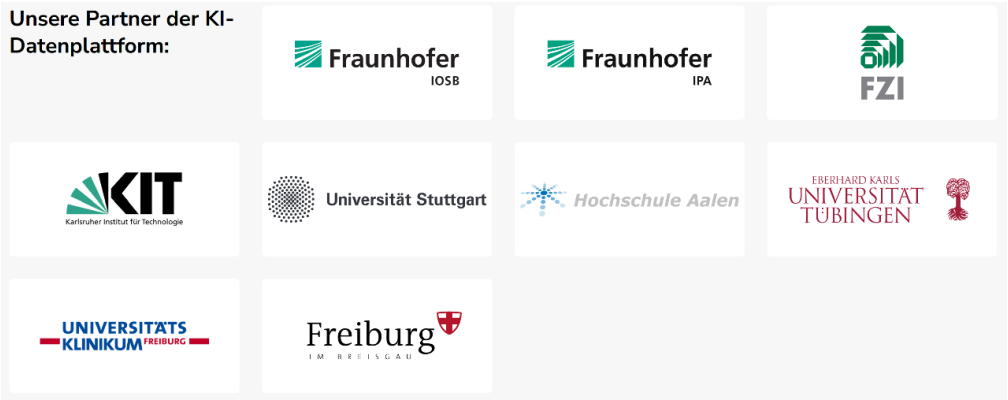
Ein Ergebnisbericht eines Teilvorhabens der KI-Allianz Baden-Württemberg eG gefördert durch das Ministerium für Wirtschaft, Handwerk und Tourismus Baden-Württemberg

---

Ausgangslage und Projektziele

Viele Unternehmen, insbesondere kleine und mittlere Unternehmen (KMU), aber auch Kommunen, Behörden und Forschungseinrichtungen, verfügen über wertvolle Datenbestände und KI-Modelle (KI-Assets), die jedoch oft verborgen, nicht auffindbar oder nicht nutzbar gemacht werden. Die Vision des Projekts *KI-Datenplattform (KI-DP)* war es, diesen verborgenen Schätzen der Datenwelt eine technische Infrastruktur zur Verfügung zu stellen, die den Zugang zu europäischen Datenräumen ermöglicht und die Wertschöpfung durch Künstliche Intelligenz (KI) fördert.

Das Projekt, gefördert mit einem Gesamtvolumen von 5,8 Mio. € und eingebettet in die **KI-Allianz Baden-Württemberg eG**, lief vom 20.07.2023 bis zum 30.06.2026. Der Fokus lag auf der Schaffung einer **Open-Source-Softwarelösung**, die Unternehmen, Wissenschaft und Verwaltung dabei unterstützt, ihre KI-Assets zu inventarisieren, zu teilen und zu monetarisieren – unter Wahrung der digitalen Souveränität, Beachtung ethischer und rechtlicher Anforderungen sowie Nutzung und Anbindung an internationale Standards der International Data Space Association (IDSA) und Initiativen (Gaia-X, Industrie 4.0, Manufacturing-X, Mobility Data Space, European Health Data Space, Smart Cities).



Erzielte Ergebnisse - Die Drei-Ebenen-Architektur

Die Lösung i unterstützt drei Stufen der Wertschöpfung in Datenökosystemen:

1. **Level 1 (KI-Asset Angebote):** Die Ebene der Datenerzeuger. Hier sorgen **Onboarding-Werkzeuge** (z.B. AAS2DCAT für die Industrie 4.0, FROST2DCAT für Sensordaten, Dienste zur Datenqualitätssicherung) dafür, dass Daten aus heterogenen Quellen (Maschinensteuerungen, Infrastrukturen, Excel) qualitativ aufbereitet und in einen standardisierten Metadatenstrom überführt werden, ohne das Eigentumsrecht an den Originaldaten zu verlieren.
2. **Level 2 (KI-Asset Kataloge):** Das Herzstück der Plattform. Ein metadatengestützter Katalog (basierend auf dem Metadatenmodell-Standard **DCAT-AP**), der Datensätze und KI-Modelle auffindbar macht. Hier arbeiten intelligente Suchmechanismen.
3. **Level 3 (Wissensplattform):** Die Erweiterung zum Wissensmanagement. Durch semantische Verknüpfung von Daten, Dokumenten und Expertenwissen über Retrieval Augmented Generation (RAG)-Systeme entstehen neue Beratungs- und Matchmaking-Dienste für die KI-Allianz und deren Mitglieder.

Technologiereifegrad - Kernkomponenten der Plattform

Die technischen Ergebnisse des Projekts wurden als Open-Source-Softwarepaket auf GitHub veröffentlicht unter https://github.com/KI-Allianz . Die Plattform basiert auf dem *Piveau*-Grundsystem und bietet folgende zusätzliche Schlüsselkomponenten:

- **Piveau.nextjs (Benutzeroberfläche):** Eine moderne, performante Web-Oberfläche auf Basis von Next.js 15. Sie ermöglicht intuitive Suche, Visualisierung von Daten und die Verwaltung von Favoriten direkt im Browser.
- **Semantische Suche:** Die Plattform versteht Konzepte und die Semantik von Begriffen. Eine Anfrage nach „Bahn" liefert auch Ergebnisse zu „Schienenverkehr". Die Suche kombiniert Vektorsuche mit klassischer Volltextsuche (Hybrid Search).
- **Datenqualitätsdienste:** Sie erkennen personenbezogene Daten (DSGVO-Compliance), ergänzt fehlende Werte, identifiziert Anomalien in tabellarischen Daten, kuratiert diese und bereitet sie damit für das Training von KI-Modellen vor.
- **Integrations-Hub (Branchenübergreifend):**
    - **AAS2DCAT:** Brückenwerkzeug für die industrielle Fertigung (Industrie 4.0), um Asset Administration Shells (AAS) in den KI-Asset-Katalog zu überführen.

- **mds2hdcat:** Konvertierungstool für Gesundheitsdaten (NFDI4Health → Health-DCAT-AP).
- **FROST-Plugin:** Export von Beschreibungen von IoT-Sensordaten (z.B. für Smart City-Anwendungen) in standardisierter KI-Asset-Kataloge.

Ethische und rechtliche Anforderungen

Ein wesentlicher Begleitfaktor war die frühzeitige Integration von ethischen und rechtlichen Aspekten („compliance by design“). Die *Task Force Ethik & Recht* erarbeitete ein Whitepaper, das die Plattformarchitektur, Governance-Strukturen und Metadatenanforderungen an Datensouveränität, Transparenz und Fairness ausrichtet. Die Plattform ist somit nicht nur technisch funktional, sondern auch regulatorisch robust (u.a. AI Act, DSGVO). Das Whitepaper ist öffentlich verfügbar unter DOI: 10.6094/UNIFR/276339

Testbetrieb und Verfügbarkeit

Die technischen Ergebnisse werden seit Anfang 2026 im Echtbetrieb unter **ki-daten.hlrs.de** getestet. Hier werden Metadaten von über 7.000 Datensätzen vorgehalten. KI-Modelle aus HuggingFace wurden importiert. Der Testbetrieb wird noch bis Ende 2026 aufrechterhalten. Gespräche mit Rechenzentren und IT-Dienstleistern zur Überführung in einen nachhaltigen, operativen Betrieb laufen.

Anwendungsfelder und Nutzen für die Wirtschaft

Um den Praxistransfer für die open source-Software zu sichern, wurde parallel zur Entwicklung ein **Anwenderforum** mit zehn Partnern etabliert. Unternehmen wie Averbis, EDI AG, PLEXPERT, Matchory GmbH, NuCOS GmbH, raumobil GmbH, SSC-Services und Studio Fluffy testeten die Nutzungsszenarien in den Bereichen Medizintechnik, industrielle KI, metallischer 3D-Druck, Mobilitätsdienste, Wissensmanagement und synthetische CAD-Daten.

Strategische Positionierung

Die Ergebnisse der KI-DP sind strategisch wichtig für den Anschluss an weitere Initiativen:

- **Manufacturing-X:** Das Fraunhofer IPA entwickelte das AAS2DCAT-Tool, das gemeinsam mit Standard-AAS-Infrastrukturen (FA^{3}ST-Server des Fraunhofer IOSB) den einfachen KMU-Einstieg in industrielle Datenräume auf der Basis ermöglicht.
- Die Stadt Freiburg ist als strategischer Partner für offene und skalierbare **urbane Datenräume** unmittelbar eingebunden und technisch über das Level 1 angeschlossen.

- Das Universitätsklinikum Freiburg untersucht und ermöglicht als innovativer Akteur in der Gesundheits- und Medizinbranche die **Anschlussfähigkeit an MEDI:CUS**, der multi-cloud-basierten Medizininfrastruktur des Landes Baden-Württemberg.
- **Wissensplattform ask-bw:** Die auf den Hannover Messen 2025/26 präsentierte Wissensplattform dient als leistungsfähiges Instrument für ein regionales Matchmaking von KI-Experten und Anwendern, u.a. für die KI-Allianz Baden-Württemberg.
- Das **Whitepaper zu ethischen und rechtlichen Anforderungen** dient als vorbereitendes Fachdokument, um für den operativen Betrieb verbindliche Gutachten gemäß den nationalen Gesetzen und europäischen Verordnungen einholen zu können.

Die geplante **strategische Erweiterung** der KI-DP führt zu einem offenen **Multi-Agenten-System**, das sowohl die automatisierte Recherche als auch das Management von KI-Assets in Wirtschaft und Verwaltung ermöglicht.

Fazit und Ausblick

Das Projekt KI-Datenplattform hat eine skalierbare, technologische Software-Basis geschaffen, die es ermöglicht, verteilte KI-Assets sicher, standardisiert und auffindbar zu machen. Durch den Open-Source-Ansatz und die Anbindung an nationale und internationale Datenräume stärkt die Plattform die digitale Souveränität des Wirtschaftsstandorts Baden-Württemberg und senkt die Einstiegshürden für KMU in die datengetriebene Wertschöpfung in eine KI-Ökosystem.

Die Integration mit bestehenden Initiativen wie Manufacturing-X und die Ausweitung des urbanen Datenraums Freiburg auf weitere Kommunen (u.a. Karlsruhe) demonstrieren die Skalierbarkeit. Zusammen mit der KI-Allianz Baden-Württemberg eG laufen Gespräche mit interessierten Akteuren zum Übergang in einen operativen Betrieb, um die Plattform als zentrale Anlaufstelle und erste standard-basierte interdisziplinäre technische Infrastruktur für ein Datenökosystem Baden-Württembergs zu etablieren.

Weiterführende Links:

- Projektseite: ki-allianz.de/ki-datenplattform
- GitHub: github.com/KI-Allianz
- Testbetrieb: ki-daten.hlrs.de
- Magazinbeitrag: https://ki-allianz.de/opensource-veroeffentlichung-2025/
- Whitepaper Ethik und Recht: DOI: 10.6094/UNIFR/276339