Prisma Architecture (arc42)

Markus Glagla

Figure 1: prisma_logo.png

1. Einführung und Ziele

Vision

Datenintegrierende Analyseplattform mit starker Governance und Compliance-by-Design.

Ziele

- Sichere, nachvollziehbare, rechtskonforme Datenintegration & Analyse
- Case-first / Purpose-first Datenzugriff
- Transparenz und Nachvollziehbarkeit für alle Zugriffe

Nicht-Ziele

- Keine ungezielten Massenabfragen (Rasterfahndung)
- Kein Social Scoring
- Kein unregulierter KI-Einsatz

2. Randbedingungen

- **Rechtlich:** DSGVO, BDSG Teil 3 (LED 2016/680), EU-AI-Act, Rechtsprechung (BVerfG Rasterfahndung)
- **Technisch:** Einsatz von Open-Source-Komponenten, Cloud-native (Kubernetes), Mandantentrennung

- **Organisatorisch:** Rollen (Data Owner, Data Steward, DSB, Security Officer), Genehmigungs- und Kontrollprozesse

3. Kontextabgrenzung

Systemumfeld

- **Quellen:** operative Systeme, offene Daten, Datenbanken von Behörden/Unternehmen
- **Nutzer:** Analyst:innen, Ermittler:innen, Compliance, Management
- **Externe Systeme:** Identity Provider (Keycloak), Audit-/Logging-System, Justiz-Schnittstellen

Abgrenzung

Plattform ist Werkzeug, keine operative Datenquelle.

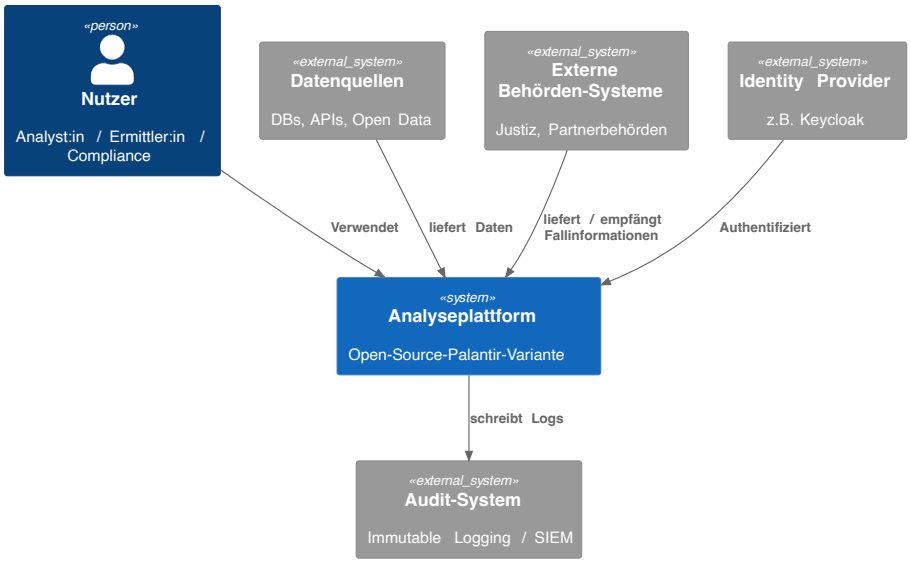

Figure 2: 03_context.svg

4. Lösungsstrategie

- Privacy- & Compliance-by-Design
- Case-first Zugriff, Policy Enforcement Layer, Immutable Audit, Datenlebenszyklus-Steuerung
- Polyglotte Persistenz (Warehouse + Graph)
- Policy Engine (OPA-artig), Low-Code-Frontend
- Sicherheitsstrategie: End-to-End-Verschlüsselung, ABAC/RBAC, Zero-Trust

5. Bausteinsicht

Ebenenmodell

- **Ingestion:** Airbyte/NiFi
- **Datenhaltung:** Postgres/Lake + Graph
- **Semantik/Entitäten:** Ontologien, Entity Resolution
- **Analyse/KI:** Jupyter, OR/ML-Pipelines
- **Visualisierung/Apps:** Superset, Low-Code
- **Governance & Security:** Policy Layer, Audit, Lineage

Querschnittskomponenten

- Logging
- Case-Management
- Rechtsgrundlagen-Registry
- Retention Engine

6. Laufzeitsicht

Szenario „Fallanalyse“

1. Nutzer legt Fall an → Zweck + Rechtsgrundlage werden registriert
2. Query Engine prüft Policy (Case, Zweck, Nutzerrolle)
3. Daten werden pseudonymisiert geladen
4. Ergebnisse nur im Case-Scope sichtbar
5. Alle Schritte werden im Audit-Log erfasst

Szenario „Rasterfahndung“

- Standardmäßig blockiert
- Nur via Anordnungs-Workflow + richterlichem Beschluss möglich

7. Verteilungssicht

- Kubernetes-Cluster mit isolierten Namespaces pro Mandant
- Trennung: Private Mandanten vs. Law-Enforcement-Mandanten (BDSG Teil 3 / LED)
- Externe Integrationen: Identity (Keycloak), SIEM, Backup/Archiv, externe Datenquellen

8. Querschnittliche Konzepte

- **Sicherheit:** ABAC mit Purpose/Case, Verschlüsselung, Zwei-Personen-Prinzip
- **Datenmanagement:** Klassifikation (normal/sensibel/besonders sensibel), Pseudonymisierung, Löschregeln

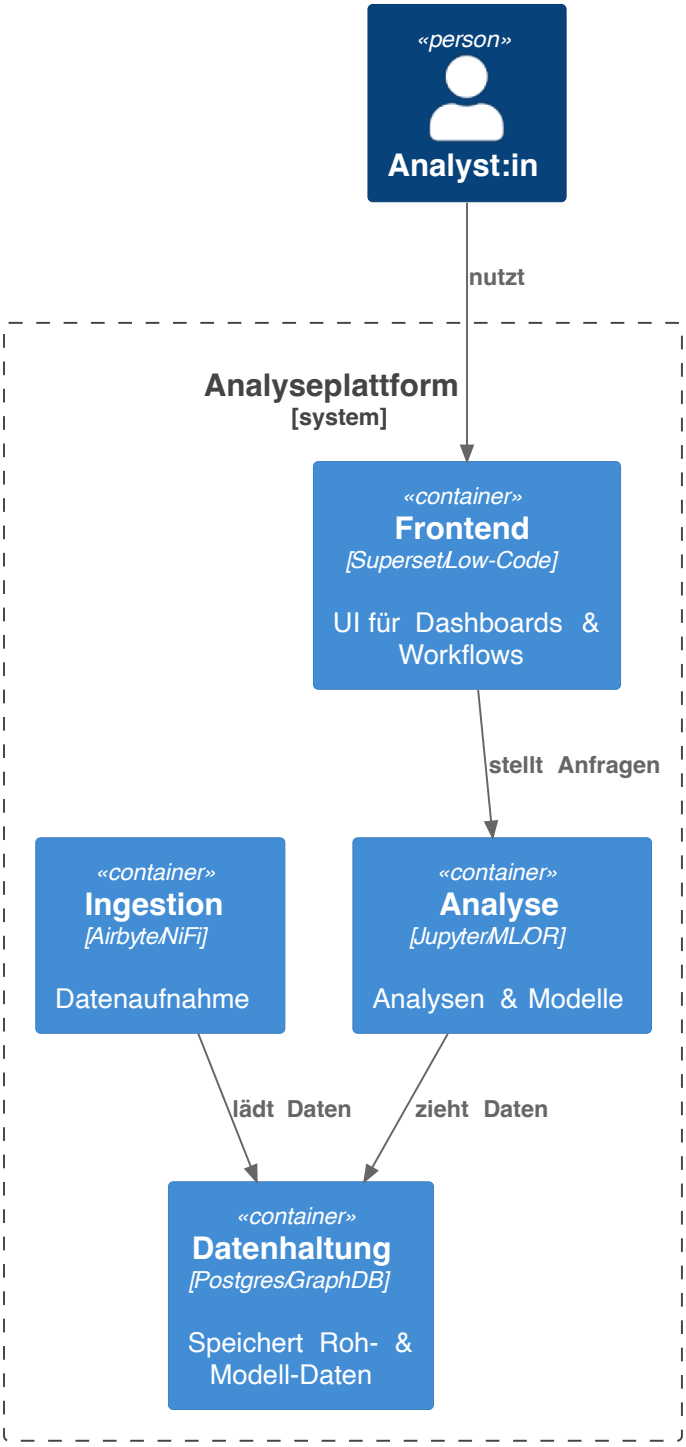

Figure 3: 02_container.svg

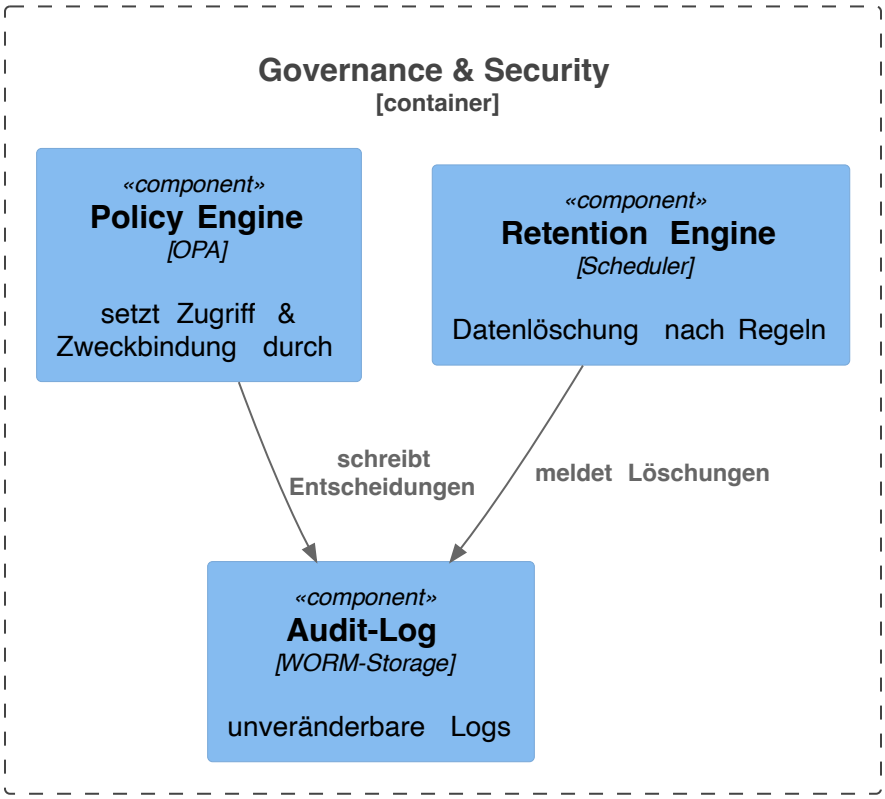

Figure 4: 03_component.svg

Laufzeitszenario: Fallanalyse

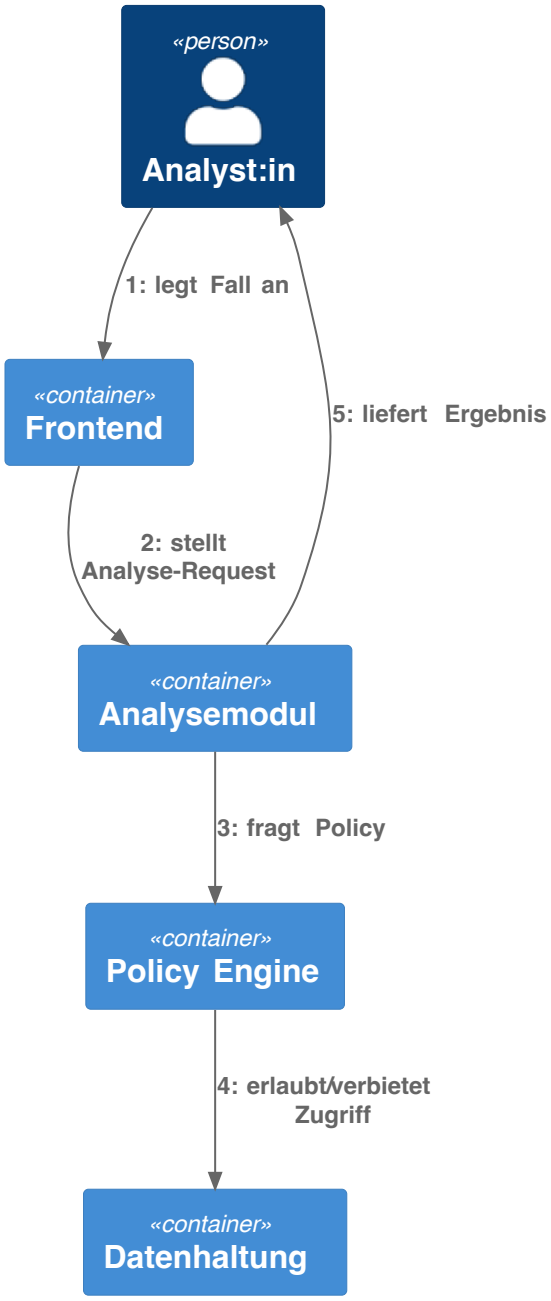

Figure 5: 04_dynamic.svg

- **Governance:** Rechtsgrundlagen-Registry, DPIA-Katalog, Modell-Katalog für KI (AI-Act)
- **Audit:** Immutable Logs, Transparenz-Reports, Auskunftspakete für Betroffene

9. Architekturentscheidungen

- Polyglotte Persistenz (Warehouse + Graph)
- OPA als Policy Engine
- Case-first Zugriff statt Data-first
- Rasterfahndung ohne Anordnung blockiert
- KI-Modelle nur nach AI-Act-Klassifizierung

10. Qualitätsanforderungen

- **Security:** Unautorisierter Zugriff $< 0{,}001$ % Wahrscheinlichkeit
- **Compliance:** 100 % Policy-Enforcement, keine unprotokollierten Abfragen
- **Usability:** Fallanlage < 1 min, Standardabfrage < 5 sec
- **Auditierbarkeit:** Jede Abfrage nachvollziehbar (Wer, Wann, Wozu, Rechtsgrundlage)

11. Risiken & technische Schulden

- Umgehung von Policies durch manuelle Workarounds
- Falsch klassifizierte Daten → fehlerhafte Schutzmaßnahmen
- KI-Bias → nicht rechtskonforme Entscheidungen
- Komplexität der Integration vieler Open-Source-Komponenten

12. Glossar

- **Case-Scope:** logische Einheit für Datenzugriffe mit Zweckbindung
- **Purpose Binding:** technische Durchsetzung der Zweckbindung (DSGVO)
- **LED:** Law Enforcement Directive (EU-Richtlinie für Strafverfolgung)
- **DPIA:** Datenschutz-Folgenabschätzung
- **AI-Act:** EU-Verordnung zur Regulierung von KI