

Konzeption einer konversationellen Datenanalyse-Plattform mit agentenbasiertem KI-Workflow und Datenqualitätspipeline

A. Baskircev, A. Shanmugavel, B. Tamer, M. Karakilic, V. Becker, Studiengang Digital Business Technology, Hochschule Ravensburg- Weingarten University

Einführung

- Datenanalyse in klassischen BI-Systemen erfordert häufig fundierte technische Kenntnisse (SQL, Modellierung)
- Zugang zu datenbasierten Erkenntnissen bleibt dadurch für nicht-technische Nutzer eingeschränkt
- Eine durchgängige Verbindung von Dialog und mitwachsender Analysefläche fehlt bislang

Zielsetzung

- Konzeption, Entwicklung und Usability-Evaluation eines KI-gestützten Analysesystems
- Automatisierte Datenaufbereitung und dialogbasierte Analyse in einem Ansatz
- FF1: Wie gebrauchstauglich ist die durchgängige Verbindung aus Chat-Dialog und Dashboard
- FF2: Welchen Einfluss haben Alter und technische Vorerfahrung auf die wahrgenommene Usability?

Zusammenfassung

- Evaluation der Methodik im universitären Umfeld durch ausführende Tests
- Theoretischer Beitrag zur Verbindung von konversationeller Analyse und persistenter Visualisierung mit mitwachsenden Dashboards
- Praktischer Nutzen durch Konzeption einer zugänglichen Analyseplattform die nicht technischen User vereinfachte Datenanalyse ermöglicht

Methodik

Conversational-Analytics-Ansatz:

- Vollständige Interaktion über eine Chat-Oberfläche in natürlicher Sprache mit einem projektbasierten Aufbau: Prepare (Datenvorschau & Qualitätsbericht), Explore (Chat & Dashboard), Predict
- Zweischichtige Client-Server-Architektur: React-Frontend (ChatPanel, DashboardCanvas) kommuniziert per REST (/api/v1/datasets, /api/v1/chats) mit modularem FastAPI-Backend.

Datenqualitäts-Pipeline

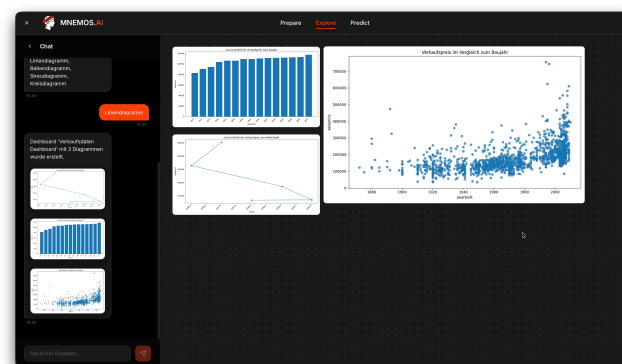
- Pipeline zur automatisierten Bereinigung von Datensätzen durch Ingestion-, Header-, Schema- und Type-Erkennung sowie Inconsistency Detection
- Verarbeitung erfolgt auf einer duplizierten cleaned-Kopie, der Originaldatensatz bleibt unverändert, ein strukturierter Quality Report fasst alle Ergebnisse zusammen.

LangGraph-Agent

- Interpretiert die Nutzeranfrage und steuert den Analyseprozess als gerichteten Zustandsgraph, mit einer regelbasierten Policy Engine für Rückfragen und Freigaben zur Erstellung der Graphen und Dashboards
- Text- oder FREE_CODE-Pfad je nach Anfrage — generierter Code läuft isoliert in einer Sandbox; GPT-4o-mini via LangChain mit deterministischen Fallback-Strategien.

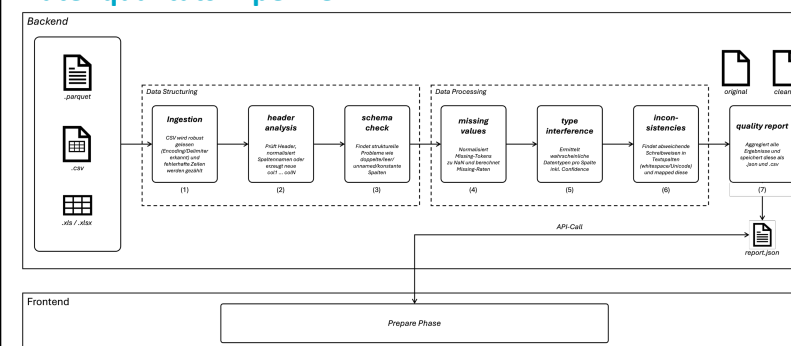
Ergebnisse

Conversational-Analytics-Ansatz:



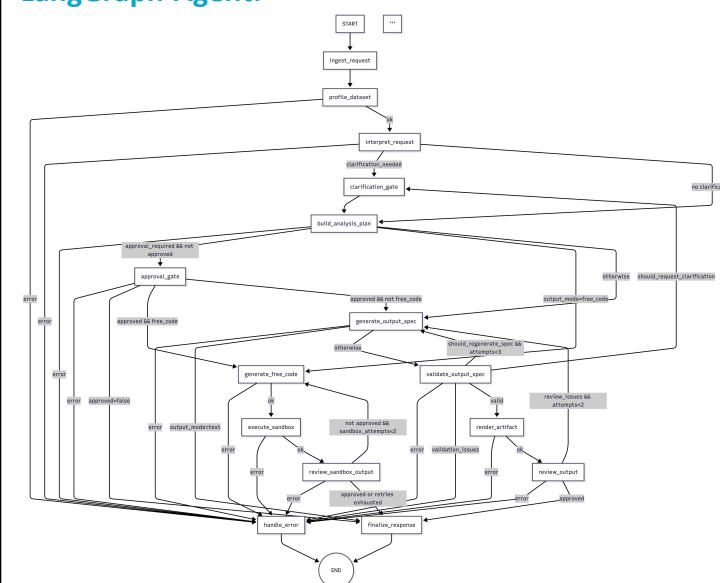
Die Explore-Ansicht von MNEMOS.AI: Chat und Dashboard-Fläche liegen nebeneinander im selben Arbeitsbereich.

Datenqualitäts-Pipeline:



Nach dem Upload durchläuft jeder Datensatz diese 7-Stufige Pipeline.

LangGraph-Agent:



Der agentische Workflow interpretiert die Nutzeranfrage, holt bei Bedarf über die Gates eine Rückfrage oder Freigabe ein und verzweigt je nach Ausgabemodus in einen Text- oder einen Code-Pfad mit isolierter Sandbox-Ausführung, bevor das Ergebnis final zusammengeführt wird.

Ressourcen

- Code Repository: <https://github.com/abi2310/mnemos>
- Webseite: <https://abi2310.github.io/mnemos/>

Kontakt

Prof. Dr. Wolfram Höpken
Leiter IDW / Forschungsgruppe Data Science
+49 751 501 9764
wolfram.hoepken@rwu.de