

Semper-KI 3D-Druck-Plattform

Erleichterte Beschaffung von Bedarfsparametern,
Daten und Verfahrenswissen

KISS@DataWeek24, Johannes Frey, Dr. Hagen Jung, Akshay Shewatkar



Gliederung

KISS – Unser Projektauftrag

Semper-KI

Auftragserfassung

Modellierung

RDF – Datenakquise

Der Chatbot

Ausblick

KISS – Unser Projektauftrag

KI-basierte semantische Plattform zur rapiden Etablierung von Wertschöpfungsnetzwerken in einer Krise.

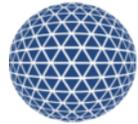
Gefördert durch:



aufgrund eines Beschlusses
des Deutschen Bundestages

Vor KISS	Nach KISS
Der Aufbau neuer Produktions- und Lieferketten, um den Bedarf an bestimmten Gütern in Krisenszenarien zu decken, dauert lange.	Die Markt- und Vernetzungsplattform ermöglicht dank KI-Methoden den zügigen Aufbau von neuen Wertschöpfungsketten, um den Bedarf bestimmter Güter in Krisensituationen zu decken.
Krankenhäusern und anderen Bedarfsstellen fällt es in Krisensituationen schwer zu spezifizieren, welche konkreten Produkte sie benötigen.	Durch einen KI-gestützten Chatbot wird die schnelle Spezifizierung der benötigten Güter erleichtert.
Die Bedarfsstellen haben Schwierigkeiten damit, optimale Produktions- und Lieferketten für ihre spezifischen Anliegen zu identifizieren.	Durch die KI-basierte Simulation und Bewertung potenzieller Wertschöpfungsketten, erhalten Bedarfsstellen einen schnellen Überblick über die Produktions- und Lieferketten, die ihren Anforderungen entsprechen.

Konsortium



InfAI[®]
Institut für Angewandte Informatik



TECHNISCHE UNIVERSITÄT
IN DER KULTURHAUPTSTADT EUROPAS
CHEMNITZ



WHZ Westsächsische
Hochschule Zwickau
Hochschule für Mobilität



Fraunhofer
IWU



magazin.semper-ki.org

USU

smart³ materials
solutions
growth



Semper-KI

Plattformmission

Besonderheiten 3D-Druck

Intelligent

Architektur

Geschäftsfall

Plattformmission

Ich, Nutzer möchte etwas in 3D drucken.

„unwissender“
Auftraggeber

Die Plattform begleitet mich schnell, intelligent und hilft mir einen **echten Auftrag** zu erstellen.

Ich, Anbieter habe ein Portfolio.

geprüfter
Auftragnehmer

Die Plattform erfasst vereinfacht, aktuell und multidimensional meine **Ausführungsmöglichkeiten**.

Matching

Preis, Qualität, Lieferung, Zuverlässigkeit

Vermittlung

Besonderheiten 3D-Druck

- Wünsche und Möglichkeiten für Kunden nicht ganz klar
- Daten-Vorbereitung notwendig, der Kunde muss liefern, kostspieliger Anpassungsbedarf, Kommunikationsaufwand, Analysesoftware
- Viele Materialien, viele Herstellungsverfahren, viele neuartige Möglichkeiten auf dem Markt vorhanden
- Starke Anpassungen der Verfahren und 3D-Drucker bei jedem Anbieter vor Ort
- Nachbearbeitung in der Herstellungskette notwendig (Schleifen, Fräsen, Lackieren, u.v.m.)
- Druckaufträge werden bei Anbietern optimiert (Bauraumbefüllung, Reihenfolge der Abarbeitung, Druck kann unter Umständen lange dauern)
- Hohe Fehlraten bei Fehldrucken, da stark individualisiert

Besonderheiten 3D Druck

multidimensional

hilft

begleiten

- Wünsche und Möglichkeiten für Kunden nicht ganz klar
- Daten-Vorbereitung notwendig, der Kunde muss liefern, kostspieliger An **echter Auftrag** Kommunikationsaufwand, Analysesoftware
- Viele Materialien, Viele Herstellungsverfahren, Viele neuartige Möglichkeiten auf dem Markt vorhanden **unwissend**
- Starke Anpassungen der Verfahren und 3D Drucker bei jedem Anbieter vor Ort **erfasst vereinfacht**
- Nachbearbeitung in der Herstellungskette notwendig (Schleifen, Frasen, Lackieren, u.v.m.)
- Druckaufträge werden bei Anbietern optimiert (Bauraumbefüllung, Folge der Abarbeitung, Druck kann unter Umständen lange dauern)
- Hohe Fehlraten bei Fehldrucken, da stark individualisiert

intelligent

Matching

Portfolio

3D Druck

geprüft

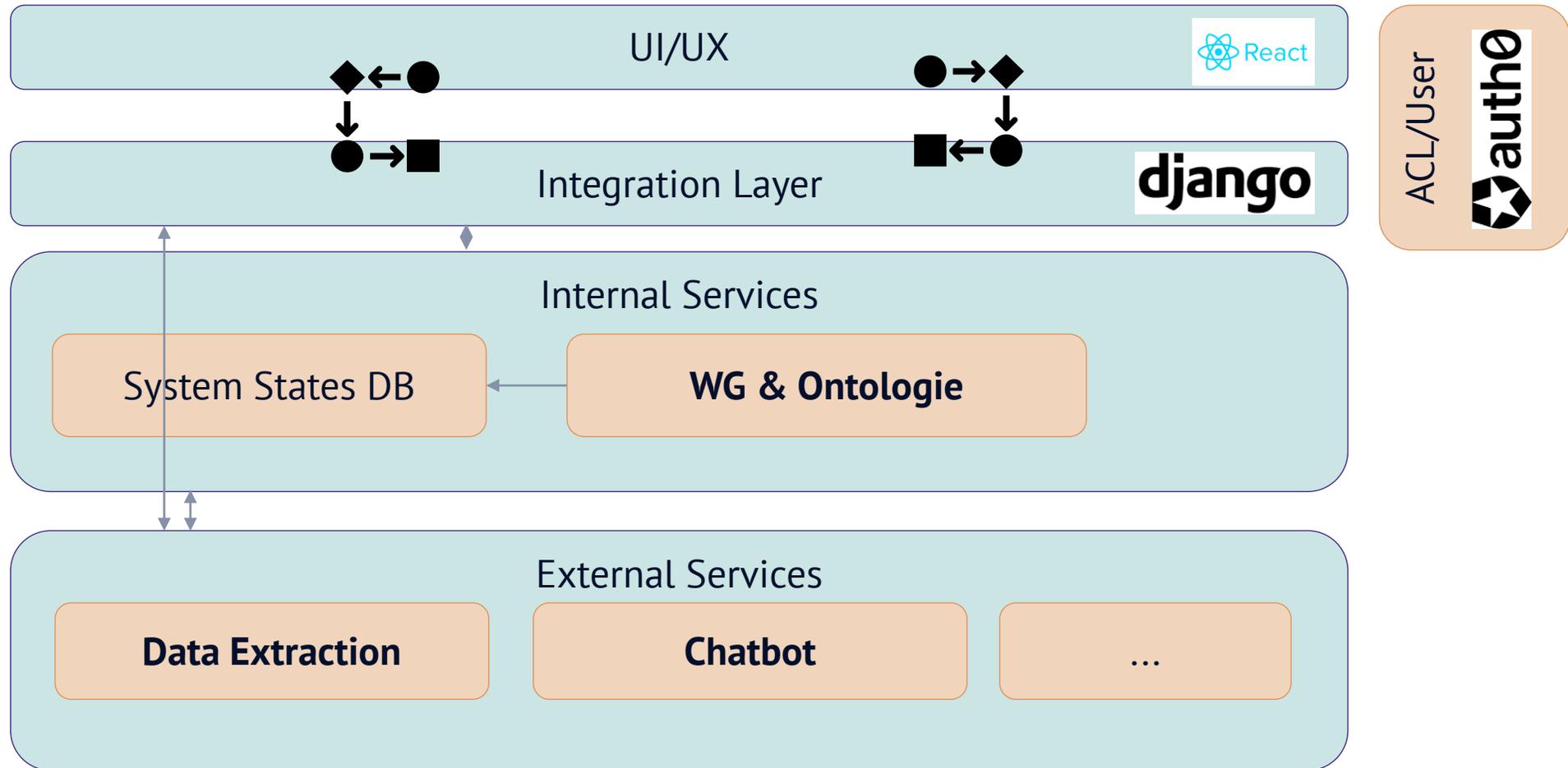
Intelligent

- **Chat-KI** Erfassung aller üblichen Bedarfsparameter, wie technische Uploads des 3D-Modells (mit Druckbarkeitsanalyse) oder Fertigungsskizze, Preis, Material, etc.
- Sofortige Erfragung aller wichtigen (**semantisches Wissen**) Zusatz-Bedarfsparameter (Rahmenbedingungen) für das spätere Matching, z.B. Bauteil-Einsatzzweck, Sichtorientierungen, o.ä.
- **Erfassung** dieses semantischen Wissens und der Daten **mit Hilfe von LLMs**
- Nicht nur bestes Preis-Matching (multidimensionale Bewertung, ML-Simulationen, Lieferkettenoptimierung)

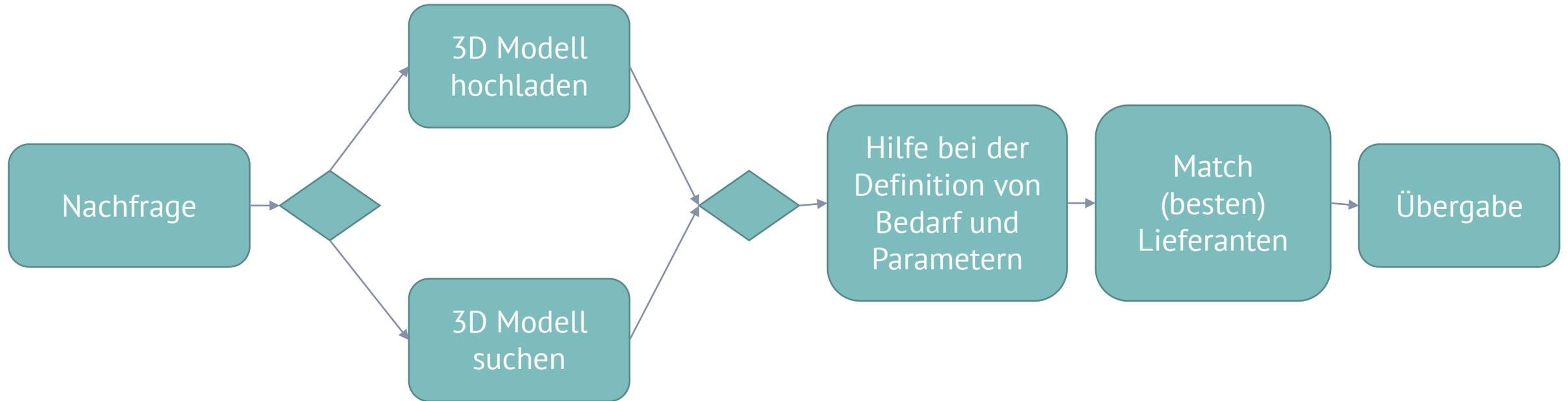
Aspekte von Intelligenz

- Semantischer Aspekt
 - Datenmodellierung- und integration via RDF Knowledge Graph / Linked Data
 - Daten/Objektschemata und (Prozess)wissen in Ontologie
- KI-Aspekt
 - logisches Schlussfolgern auf Basis von Ontologie und Daten im KG
 - NLP und Wissensextraktion und “Data lifting” via Large Language Models (LLM)

Architektur



Geschäftsfall aus Kundensicht



Wichtige Infos aus Herstellersicht

- Kundenanfrage (multimedial)
- Gewerblich vs. Privat
 - Privat: meist utopische Preisvorstellung oder wenig wirtschaftlichkeit für Hersteller
- Boxmaß/Volumen/Gewicht?
 - Fertigbar oder nicht
- 3D Modell vorhanden?
 - Ja/Nein
 - Falls nein – Service für 3D-Modellerstellung
 - Falls ja – NDA und Sharing Conditions
- Material?
- Einsatzzweck?
- Stückzahl?
- Gewünschter Liefertermin?

Wichtige Infos aus Herstellersicht

- Zielpreis?
- Machbarkeits Check
 - Qualitätscheck des 3D Modells
 - Materialcheck für Einsatz
 - Hohlräume
 - Spaltmaß
 - Wandstärke
 - Realisierbare Nacharbeiten
- Produktion, Nachbearbeitung und QM
- Lieferung

Heute möchte ich etwas in 3D drucken ...

Nachfrage

Das Völkerschlachtdenkmal



Auftragserfassung

Anforderungen

Komplexe Abhängigkeiten

Viele Datendimensionen

Video - Erneute Auswahl der Parameter

Video - Auswahl filtert Matching-Ergebnisse

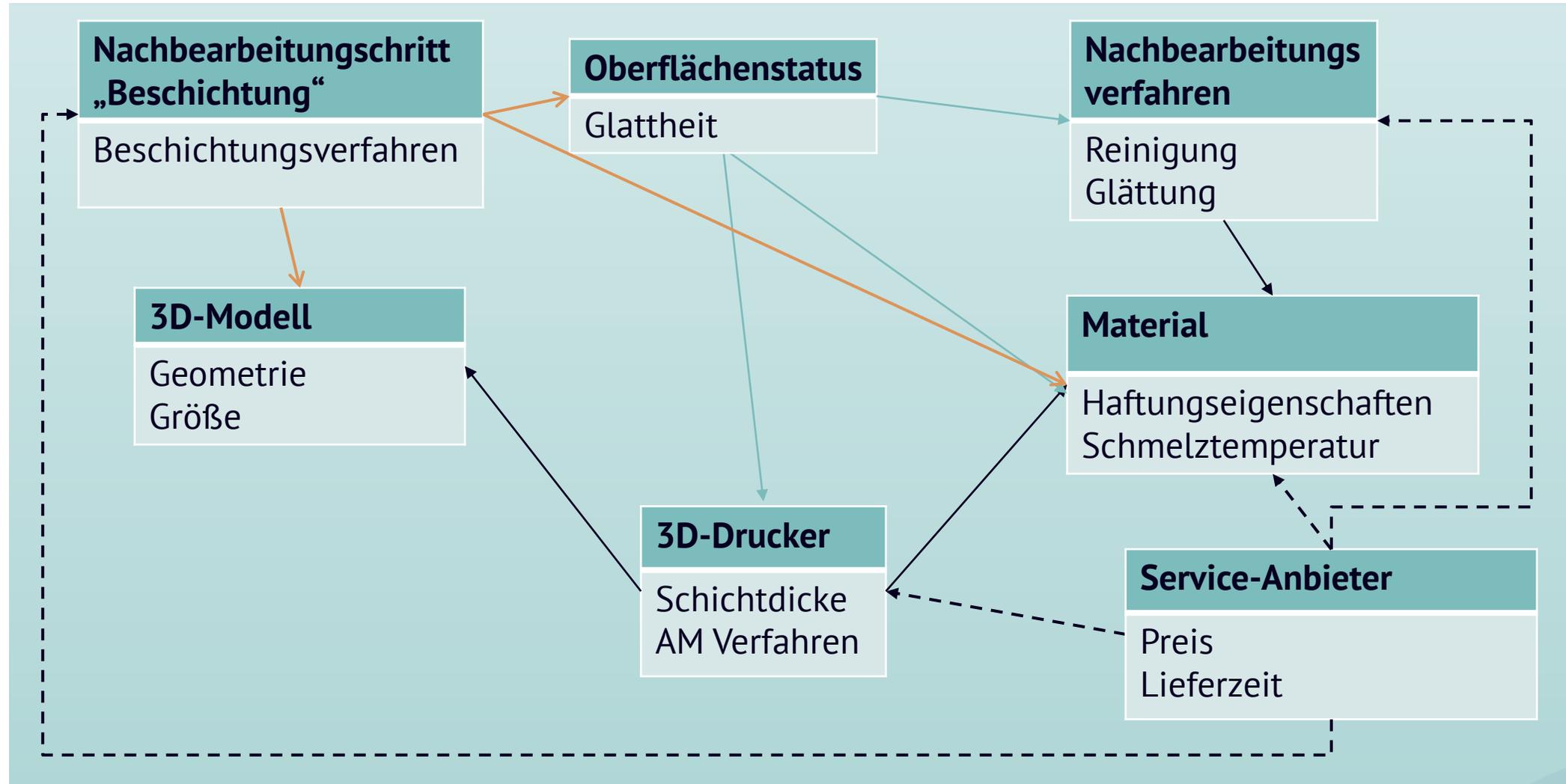
Anforderungen

Heute möchte ich das Völki in 3D drucken ...

- und ich habe ein 3D Modell als STL-File,
- das Modell soll schön glatt und lackiert sein
- es soll preiswert sein,
- Material ist egal,
- ich brauche zwei verschiedene Größen



Komplexe Abhängigkeiten „Beschichtung“



Viele Datendimensionen

- Additive Manufacturing (Additive Fertigung, AM) ist sehr breites Feld
- Hohe / tiefe Komplexität der einzelnen Prozesse (Vielzahl an Parametrisierungen)
- Komplexe Abhängigkeiten und Einflüsse
- Je nach Anforderungen und Voraussetzungen (Auftragnehmer, Auftraggeber) unterschiedliche Granularität benötigt / gegeben

→ Vielzahl an Daten und modelliertem Wissen wird für die Plattform benötigt

Viele Datendimensionen

- 3D Drucker + Spezifikationen (z.B. Materialien, max. Ausdehnung)
- Materialien
- Materialcharakteristik (z.B. UV- oder Hitzebeständigkeit)
- Einsatzzweck & Zertifizierungen (z.B. medizinische Implantate)
- AM-Prozesse (inkl. Parameter & Abhängigkeiten & Effekte)
- AM-Anbieter + Verfügbarkeiten (z.B. Herstellungskapazität)
- 3D Objekt-Modelle (Datenformate und Objektcharakteristik)

Video – Erneute Auswahl der Parameter

The screenshot shows the SEMPER-KI web interface. At the top left is the SEMPER-KI logo. At the top right, there is a navigation menu with 'Meine Projekte' and a hamburger menu icon. Below the logo, the breadcrumb '> Meine Projekte' is visible. The main content area is titled 'Projektübersicht' and contains a search bar with the placeholder text 'Suchen...'. Below the search bar is a checkbox labeled 'Eigene Projekte'. The text 'Keine Projekte vorhanden' is centered below the checkbox. A blue button labeled 'Projekt anlegen' is located in the top right corner of the project overview panel. The footer contains links for 'Impressum', 'Datenschutzerklärung', and 'Kontakt', along with social media icons for Instagram and Telegram, and the copyright notice '© 2023 Semper-KI'.

Video – Auswahl filtert Matching-Ergebnisse

The screenshot shows the SEMPER-KI web interface. At the top left is the SEMPER-KI logo. At the top right, there is a navigation menu with 'Meine Projekte' and a hamburger menu icon. Below the navigation, the page title is '> Meine Projekte'. The main content area is a white box titled 'Projektübersicht' with a 'Projekt anlegen' button in the top right corner. Inside this box, there is a search bar with the placeholder text 'Suchen...'. Below the search bar, there is a checkbox labeled 'Eigene Projekte' which is currently unchecked. In the center of the box, the text 'Keine Projekte vorhanden' is displayed. At the bottom of the page, there is a dark blue footer containing links for 'Impressum', 'Datenschutzerklärung', and 'Kontakt', along with social media icons for Instagram and LinkedIn, and the copyright notice '© 2023 Semper-KI'.

Modellierung und Daten

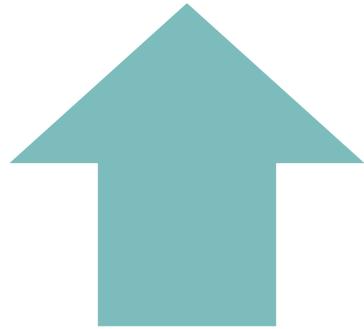
Modellierung

OWL-Ontologien

RDF-Wissengraphen

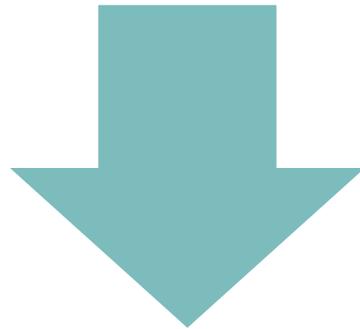
Beschleunigung der Modellierung

Semantische Daten- and Wissensmodellierung



Konzeptuelles Wissen

- OWL-Ontologien



Instanzdaten

- RDF Statements (KG)

Hintergrundwissen: Poll

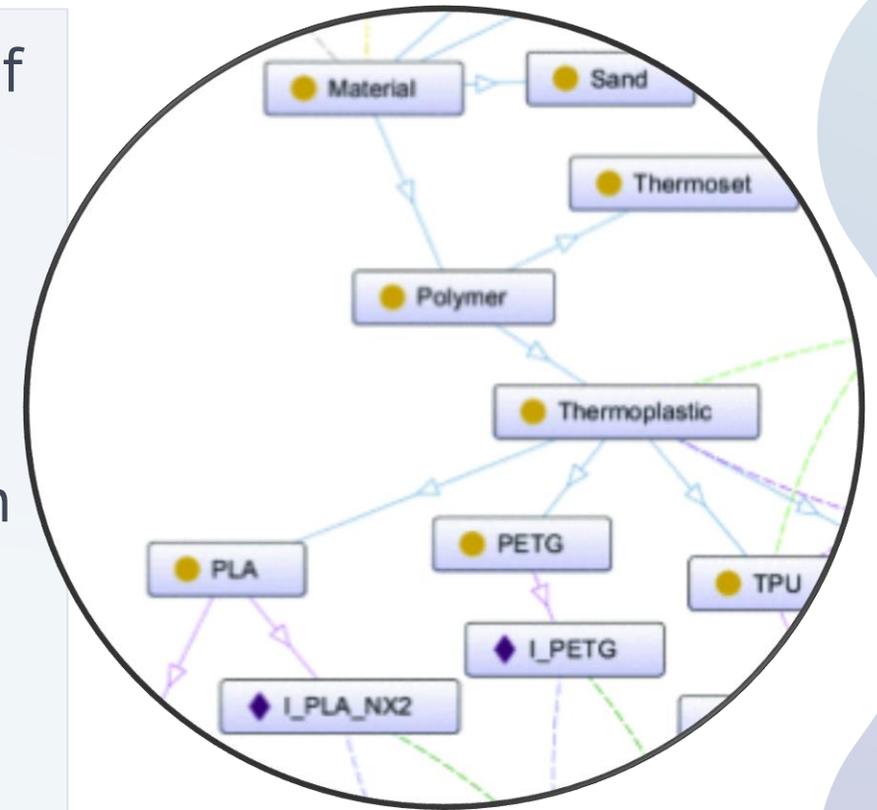
Bitte Hand heben:

Wer hat grobes Verständnis von?

- Wissensgraphen
 - RDF & Ontologien
- LLM
- JSON Daten von Hand angefasst

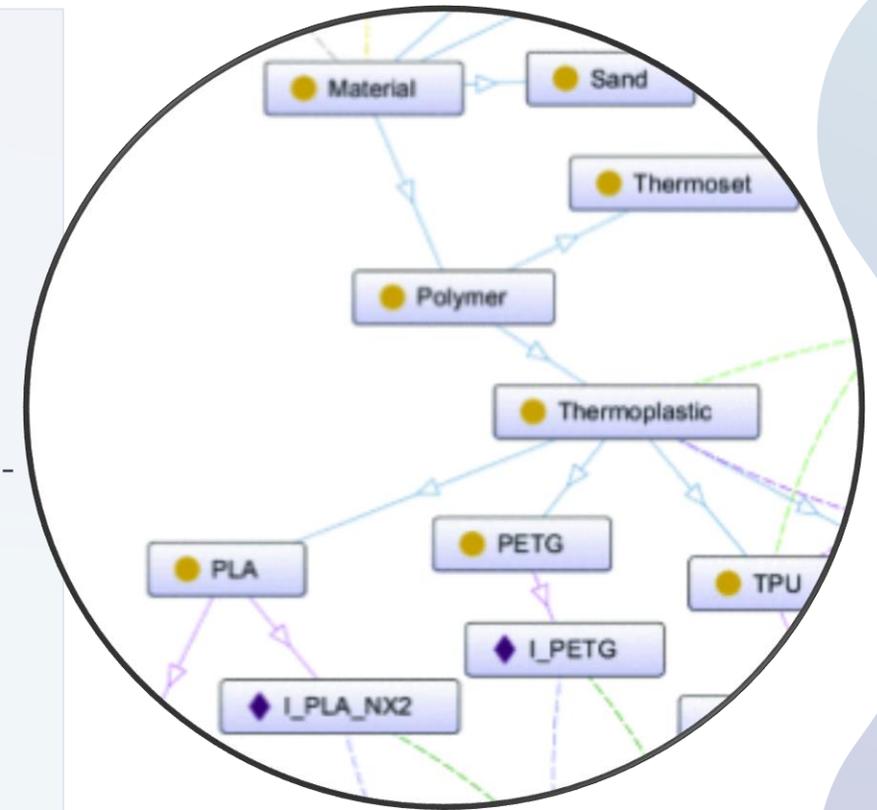
OWL-Ontologien

- Bereitstellung von Wissensbeschreibungen auf unterschiedlich tiefen Ebenen
 - z.B. “3D-Drucker” -> “plastik-3D-Drucker”
 - z.B. “**Polymer**” -> “**Thermoplastik**” -> “**PLA**”
- Erlaubt verschiedene Sichten auf ein Thema (z.B. Konstruktion vs. Fertigung) die verbunden sind und umgangssprachliche oder fachliche Terminologie erfassen z.B. steht der Term “Metall” umgangssprachlich auch für “Stahl”



OWL-Ontologien

- Definiert die Basis für das Schlussfolgern (Wissensabstraktionen/**Ableitungen**, welche nicht explizit gegeben sind)
 - Beispiel Schlussfolgern:
 - Axiome:
 - "Plastikdrucker" ist ein "3D-Drucker" welcher das Material "Polymer" unterstützt
 - "3D-Plastikdruckerhersteller" ist ein Hersteller von "3D-Plastikdrucker"
 - Instanz:
 - "Drucker XYZ"
 - des Herstellers "3dsystem"
 - druckt "PLA"
 - Abgeleitete Instanzen:
 - XYZ ist ein 3D- Plastikdrucker
 - 3dsystem ist ein "3D-Plastikdruckerhersteller"



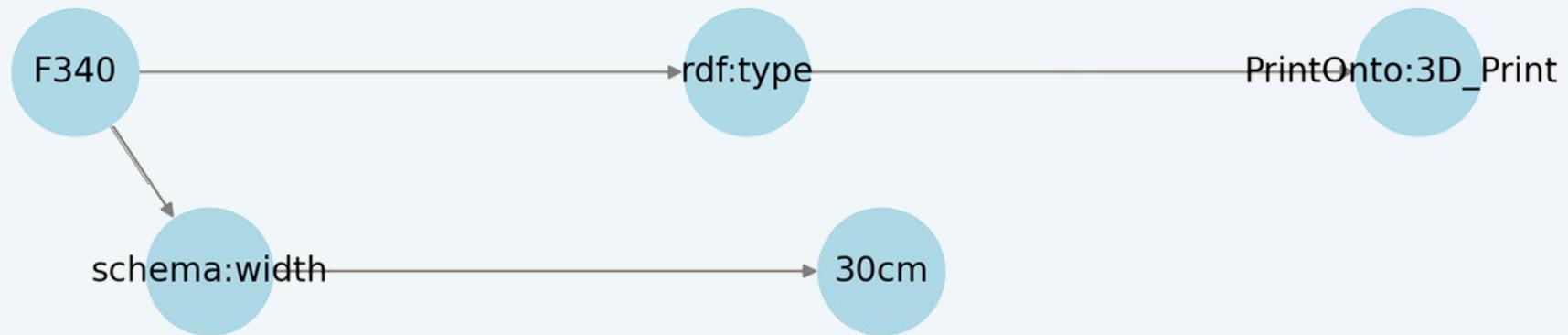
OWL-Ontologien

- Beispiel für eine Regel des Schlussfolgerns in Beschreibungslogik
- Erlaubt Klassifizierung von AM-Prozessen als SLS auch ohne explizite Zuweisung
- Erlaubt Inferenz partiell fehlenden Wissens (Open World Semantic)

SelectiveLaserSintering \equiv *AmProcess* \sqcap \exists *hasEnergySource.LaserBeam*
 \sqcap \exists *hasBuildMaterial.PowderMaterial*
 \sqcap \exists *hasBuildPlatform.PowderBed*
 \sqcap \exists *hasBuildMechanism.Sintering*

Was ist ein Wissensgraph (RDF)?

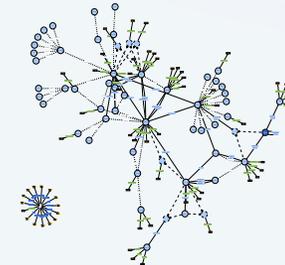
- Stellt einzelne Fakten als ein Tripel aus (Subjekt, Predikat, Objekt) dar



- Benutzt idealerweise Ontologie(n) um Typen von Knoten und Beziehungen (Prädikaten) einheitlich zu beschreiben und zu definieren
- Vielzahl an W3C standardisierten Vokabularen für Toolkompatibilität und Dateninteroperabilität

Beschleunigung bei der Modellierung

- Interview Assistenz System -> **Wissensaufnahme vor Ort**
- Expertenwissen benötigt zur Ontologie Erstellung
 - > Fachgebietsexperten müssen verstehen wozu Ontologie gebraucht wird und wie sie funktioniert
 - > Vermittlung des Wissens zu Kosten / Zeit intensiv
- Tool soll das Wissen der Onologieerstellung beinhalten
 - > KI gestützte Vorauswahl des Expertenwissens
 - > Interview Hilfe
 - > Zeit und Kosten Einsparung



LLM gestütztes Interview

- Eingabe des Kompetenzbereichs zur Erstellung der Ontologie:

Ontology Generator

Fachgebiet eingeben:

3D Druck

Generiere Ontologie

Ontology of 3D Druck

Wählen Sie die Hauptkategorie/n:

- 1. 3D Druckverfahren
- 2. 3D Drucker
- 3. Materialien
- 4. CAD-Software
- 5. Dateiformate
- 6. Druckparameter
- 7. Nachbearbeitung

- KI generierte Einteilung des Kompetenzbereichs:

RDF – Datenakquise mit LLMs

Aus PDF

Aus JSON

Hintergrund: Large Language Models

- AI that can process and generate sequences of tokens (~words)
 - trained on huge amounts of text documents on the web
 - Basic idea: predict the next token to follow given tokens before
 - “Der Himmel ist“
 - blau 60%
 - sternklar 5%
 - “Die Nacht ist“
 - dunkel 50%
 - sternklar 20%
- => ChatGPT3.5/4 is backed by
- OpenAI`s GPT3.5/4 Large Language Model

LLM “Chat”/”Instruction” Modelle

- “Nächste” Generation von LLMs
 - erstellt mit speziellen Trainingsdaten + Trainingsalgorithmen (z.B. RLHF)
 - erlauben mit dem LLM zu chatten (über die reine Vervollständigung von Texten hinaus)
 - Interaktion -> z.B. Beantwortung von Fragen
 - Beachtung oder gar Ausführungen von Anweisungen in Antworten (z.B. Übersetze in Englisch, Sätze maximal 10 Wörter, etc.)
 - Es gibt LLMs die gezielt fürs Programmieren trainiert worden sind
 - Anwendung auf Datenpro / LLMs für Daten
- “Naïve” Idee - Einsatz von LLMs auf der Plattform um Entscheidungen zu Treffen (z.B. Matching oder Machbarkeit)

Herausforderungen bei Verwendung von LLMs

Qualität der Antwort:

- Umgang mit Halluzinationen
- Knowledge cutoff in Trainingsdaten - aktuelle Informationen fehlen
 - Sep. 2022 for GPT3.5 Turbo
- Sehr spezifische Domäne / “Long Tail” Wissen wird nicht gut erfasst
- Geringe(re) Genauigkeit bei der Abfrage sehr spezifischer Details, die jedoch genau sein müssen (z. B. DOIs bei Papern, Dimensionen von Druckern etc.)
- Antwort ist Text der wieder interpretiert werden muss
 - Naïve Idee kann nicht in naher Zukunft mit guter und zuverlässiger Qualität realisiert werden
 - ABER Einbeziehung **externer Daten** verbessert Leistung signifikant

Einbeziehung ext. Information für bessere Qualität

Wie kann man Daten in ein LLM **einspeisen**
und wie kann man Wissen/Daten daraus **abrufen**?

1. Prompt Engineering (optimierte *Anweisungen für LLMs, um eine gute Antwort zu geben*)
2. Retrieval Augmented Generation (RAG) (*Versorgung von LLMs mit zusätzlichen relevanten Daten*)
3. Fine tuning / Training eines LLM
(*schwierig durchführbar, erfordert umfangreiche gelabelte Daten, sehr ressourcenintensiver Prozess*)

Ansatz zur Erlangung **RELEVANTER strukturierter Antworten**, die einem **VORGEBILDETEN SCHEMA** entsprechen, durch LLMs → **RAG + Prompt Engineering**

Einbeziehung ext. Information für bessere Qualität

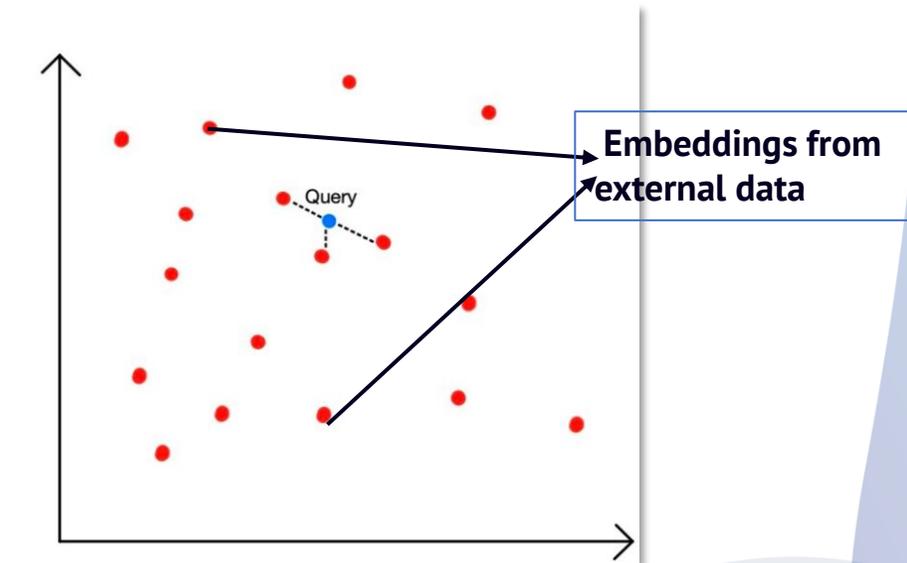
- **Prompt Engineering**

- Erarbeitung spezieller Aufforderungen, um das Verhalten der LLM zu steuern.
- Wie?
 - Mit *Pydantic models*, ein gewünschtes Datenmodell definieren, in das die Ausgabe des LLM gegossen werden kann.

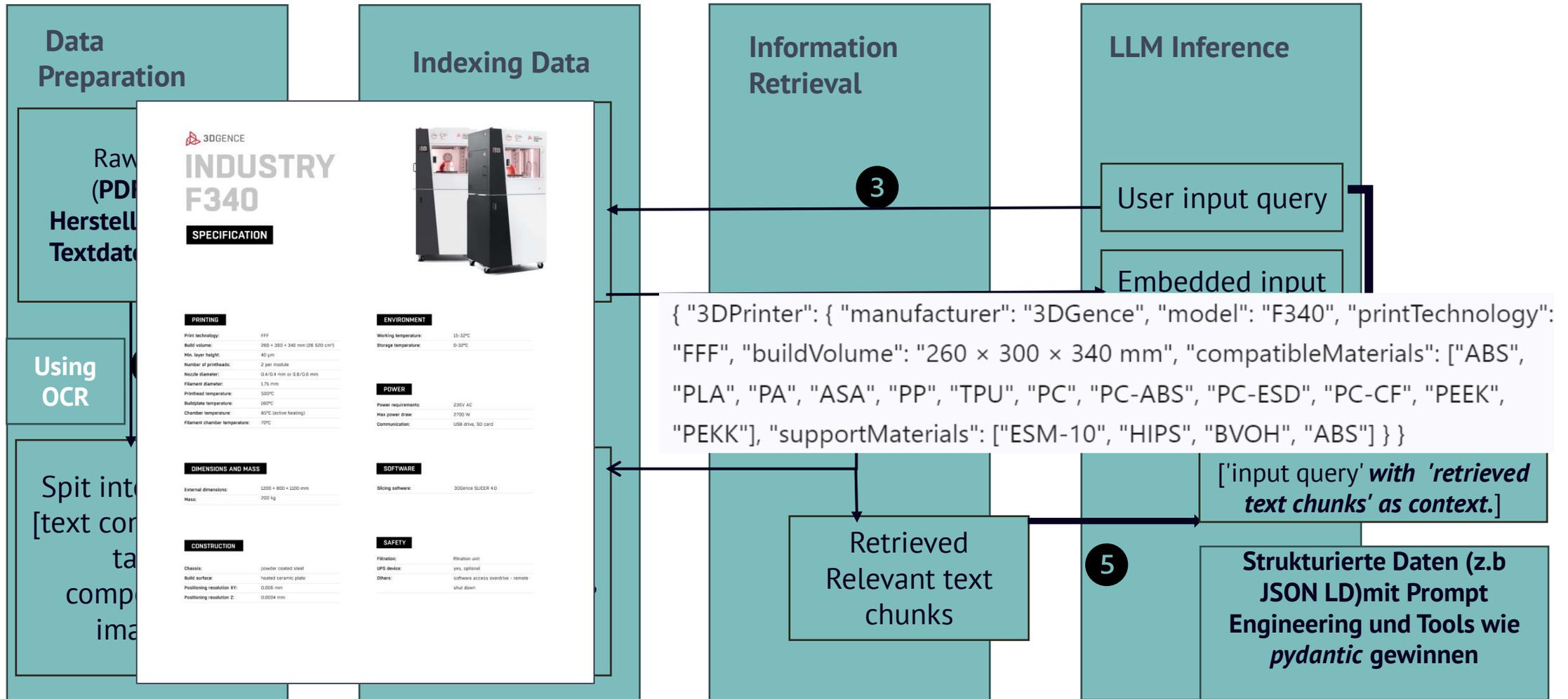
- **RAG**

- 'Retrieves' relevanten Daten aus Externen Daten ab und verwendet sie als 'Kontext', um die LLMs zu speisen.
- Wie?
 - Schritte: Vorverarbeitung der Daten, Embeddings, Ähnlichkeitssuche

```
class SpecificationsDetail(BaseModel):  
    manufacturer: str  
    printer_model: str  
    printing_technology: str  
    build_volume: str  
    compatible_material: str  
    support_material: str
```

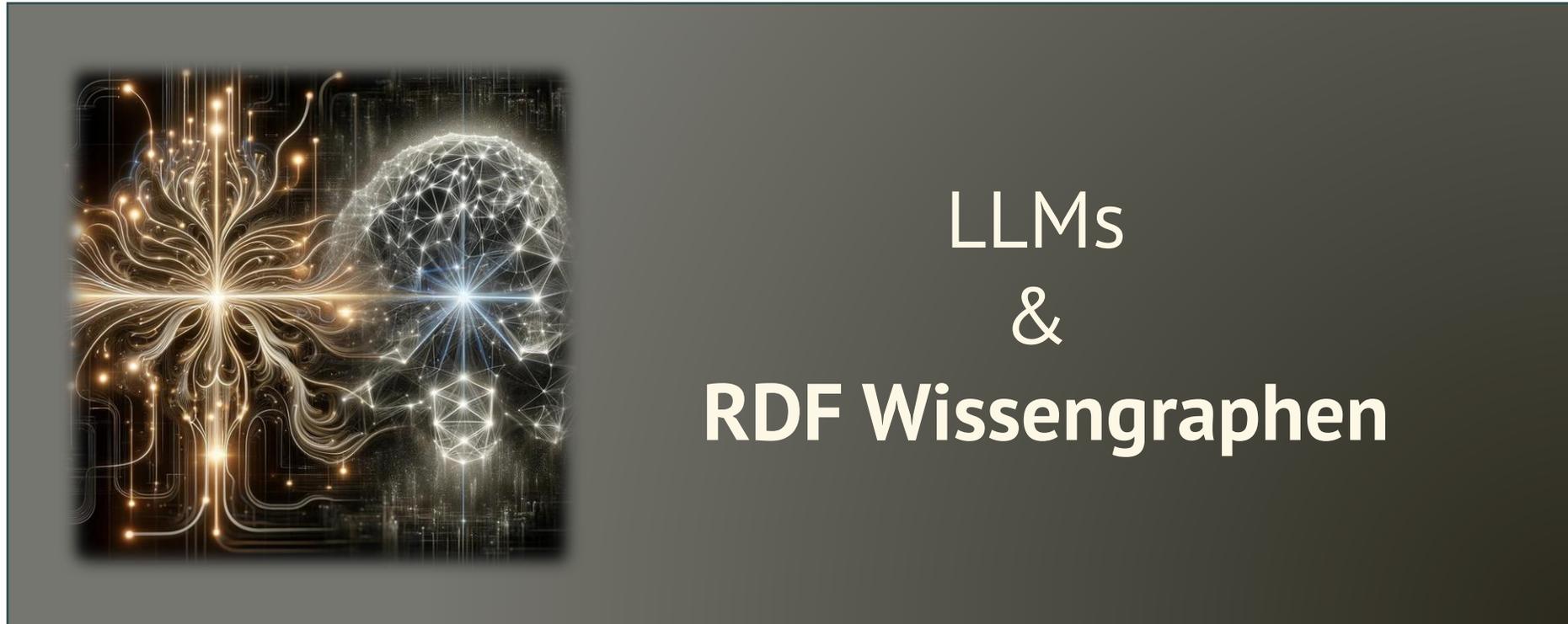


Typische RAG-Architektur via Vector-Embeddings



Ein weiterer “RAG” Lösungsansatz

“GraphRAG”: Kombiniere die Fähigkeiten von



Unsere LLM+KG Strategie

- Benutze LLM zur Datenaquise = Erstellung des KGs
 - Einbringung von externen Datenbeständen (Text, Bild) und Analyse/Auswertung dieser mit LLM Wissen
 - “Umwandlung” externer Informationen nach RDF direkt mit LLM oder LLM-gestützten Prozess
- LLM bezieht Informationen direkt aus KG um spezifische Frage zu beantworten

Vorteile gegenüber Vektor-RAG

- maschinenlesbares besser kontrollierbares Wissen (QM besser möglich)
- Inkrementelles, zyklisches Datenwachstum möglich
- hybride Erstellung/Verwertung mgl. (Nutzung von “traditionell daten-orientierter” Software und LLMs parallel)
- eingespielte Informationen prägnanter (potentiell günstiger bei tokenbasierter Abrechnung)
- “KG als Text” → Embedding/VektorRAG zusätzlich möglich

Bsp.: “Konstruiere KG-Entität aus Produktdatenblatt”

SPECIFICATION	
	
PRINTING	
Print technology:	FFF
Build volume:	260 × 300 × 340 mm (26 520 cm ³)
Min. layer height:	40 µm
Number of printheads:	2 per module
Nozzle diameter:	0.4/0.4 mm or 0.6/0.6 mm
Filament diameter:	1.75 mm
Printhead temperature:	500°C
Buildplate temperature:	160°C
Chamber temperature:	85°C (active heating)
Filament chamber temperature:	70°C
ENVIRONMENT	
Working temperature:	15-32°C
Storage temperature:	0-32°C
POWER	
Power requirements:	230V AC
Max power draw:	2700 W
Communication:	USB drive, SD card

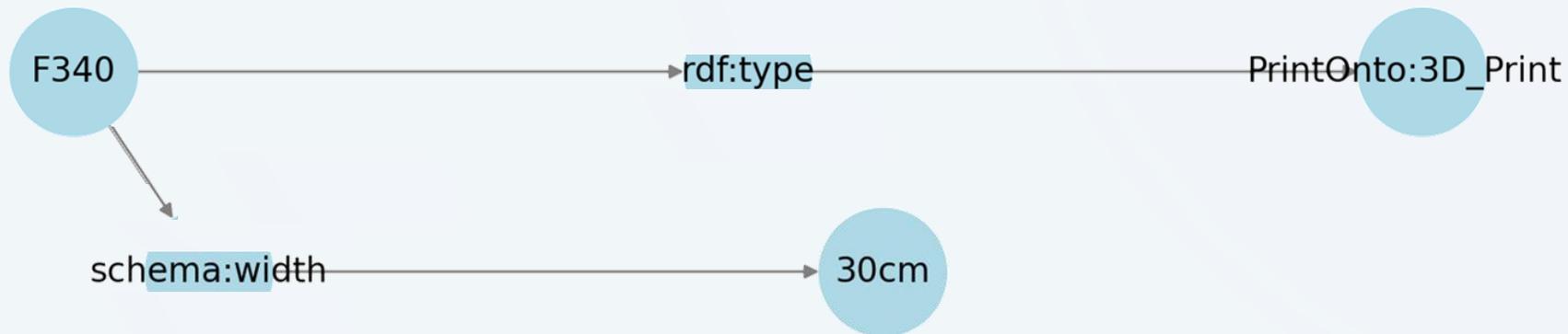
Präzisierte Forschungsfragen:

Wie kann man RDF KG Daten in ein LLM **einspeisen**
und wie kann das LLM Wissen in einen KG bringen?

→ benötigen E/A-Schnittstellen(sprachen)
zwischen KG/Ontologie und LLM

Warum (RDF)Wissensgraphen?

- Natürlicher Weg um Fakten und Beziehungen zwischen Dingen zu enkodieren (Subject-Prädikat-Objekt)



- gut **standardisiert** (viele konsistente Trainingsdaten öffentlich verfügbar)
- maschinenlesbar mit Programmbibliotheken in vielen Programmiersprachen
- Sauberere **Semantik** durch eine Menge von standardisierten “Erweiterungen” (Ontologien, ...)
- Mächtige **Anfragesprache** SPARQL um Fakten abzufragen / Antworten zu finden

Eine Strategie: Verwendung vom RDF Turtle Format

- Turtle is an exchange format for RDF (“CSV” for knowledge graphs)
 - Good candidates as I/O Interface
 - **Paper “How well do LLMs speak Turtle?”**

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
```

```
@prefix schema: <http://schema.org/> .
```

```
@prefix PrintOnto: <http://example.org/PrintOnto#> .
```

```
<F340>
```

```
  rdf:type PrintOnto:3D_Printer ;
```

```
  schema:width "30cm" .
```

Schnittstellensprache: Erfahrungen erster Experimente

Experimente von Hand und Bewertung von RDF Experten

- JSON-LD
 - Gute Extraktion/Strukturierung der Daten
 - (zu) häufige Nutzung von schema.org als Vokabular/Ontologie
 - Modellierung von Einheiten z.B. hier problematisch
 - JSON Dokument mitunter nicht wohlgeformt – Parsing problematisch
- RDF OWL/XML
 - Ebenfalls oft nicht wohlgeformt (mitunter Abschnitte mit Turtle Syntax)
 - besonders teuer da lang
- Turtle
 - Prinzipien besser umgesetzt
 - Aber teilweise schlechtere Modellierung
 - Oft einzelne Fakten nicht sauber “aufgedröselt” sondern als Strings
 - Property identifier pro Drucker
 - Falsche Identifier für Millimeter (MilliM)
- Problem: Ontologien passen nicht in damalige Kontextgrößen der LLMs



Working conference on Artificial Intelligence Development for a Resilient and Sustainable Tomorrow
↳ AIDRST 2023: **First Working Conference on Artificial Intelligence Development for a Resilient and Sustainable Tomorrow**

Home > First Working Conference on Artificial Intelligence Development for a Resilient and Sustainable Tomorrow >

LLM-assisted Knowledge Graph Engineering: Experiments with ChatGPT

Lars-Peter Meyer , Claus Stadler, Johannes Frey, Norman Radtke, Kurt Junghanns, Roy Meissner, Gordian Dziwis, Kirill Bulert & Michael Martin

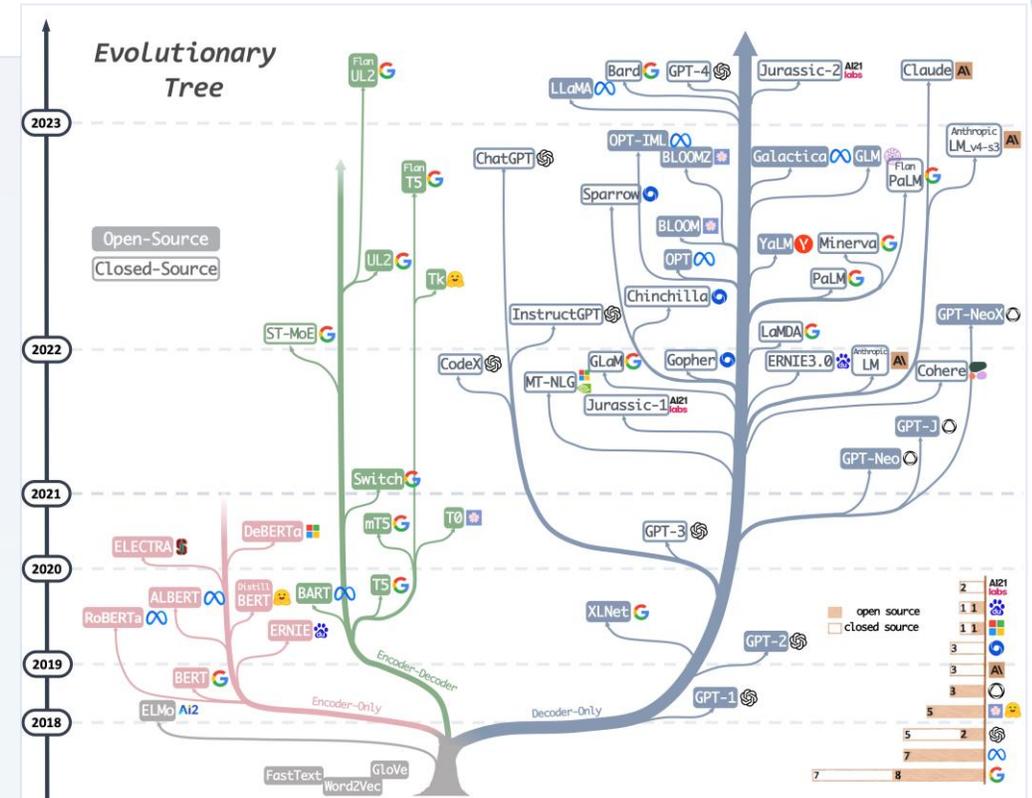
Conference paper | [Open Access](#) | First Online: 17 April 2024

Part of the book series: [Informatik aktuell](#) ((INFORMAT))

Fallstricke beim Experimentieren

Diverse Einflussfaktoren auf Ergebnis:

- Vielzahl von LLM Familien und trainierten Modellen
- Parametrisierung bei Anfrage (z.B. Temperature)
- Nicht deterministische Ergebnisse
- Neue Versionen von LLMs können Situation verbessern/verschlechtern
- Prompt Design
- Eingebbrachtes externes Wissen
- ➔ Qualität automatisiert messen

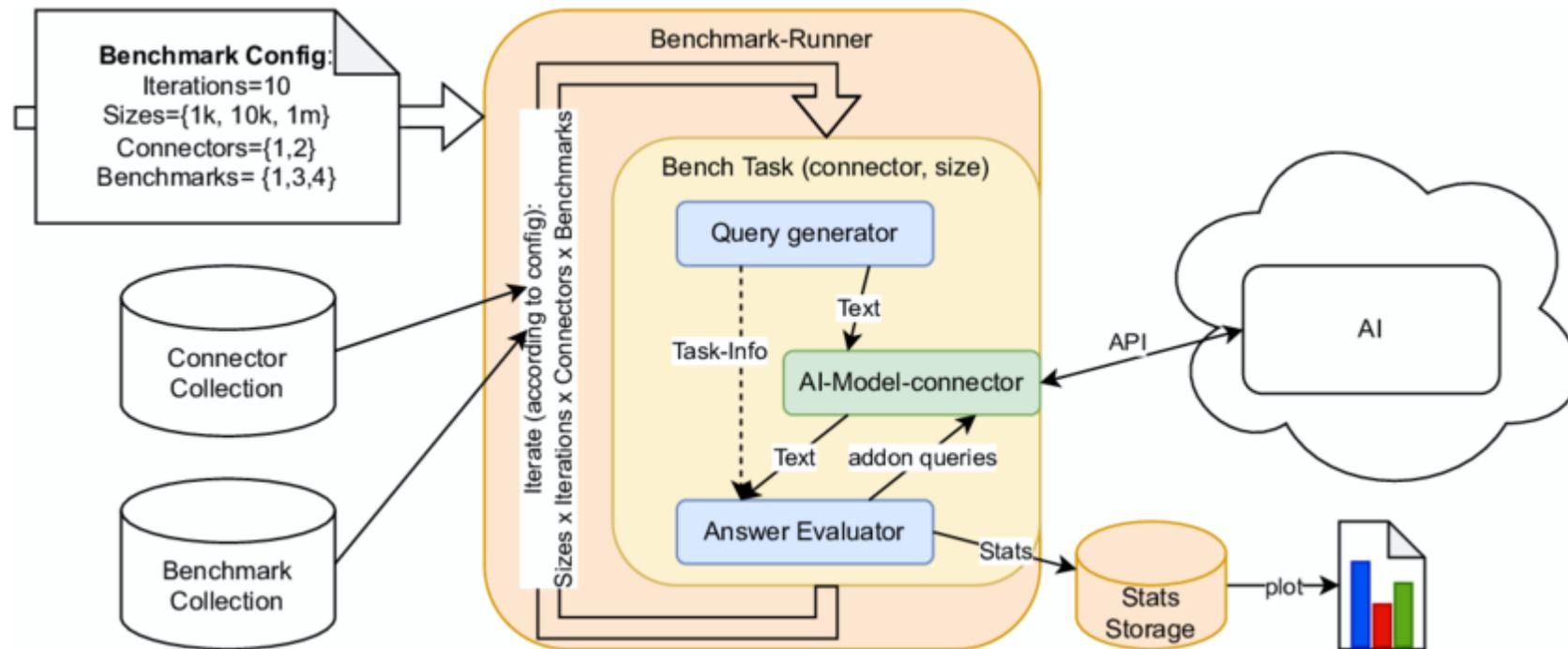


Quelle:

<https://github.com/Mooler0410/LLMsPracticalGuide>
J. Yang et al. 2023: „Harnessing the Power of LLMs in Practice: A Survey on ChatGPT and Beyond”

Benchmarking der LLM Performance (Qualität)

- Automated LLM assessment framework
 - Connectors for commercial models Claude and (Chat)GPT and GPT4all



Benchmark Task - RDF skill levels

- Task T1: Find a connection between Nodes in Turtle
- Task T2: Find syntax errors in Turtle
- Task T3: Generate Sample Person Graphs
- Task T4: Identify most known Person
- **Task T5: Extract Data from 3D Printer Factsheet**
- Task T6: Übersetzung von Text Abfragen an KG via SPARQL

	T1	T2	T3	T4	T5	T6
Turtle Read	+	++	/	+	/	/
Turtle Write	/	++	+	+	++	+
Graph Understanding	++	/	+	++	++	++
Vocabulary Knowledge	/	/	/	/	+	+

T5: “Konstruiere WG-Entität aus Produktdatenblatt”

Auf Grundlage textueller Beschreibung im Stile von Key-Value-Paaren einer RDF-Entität!

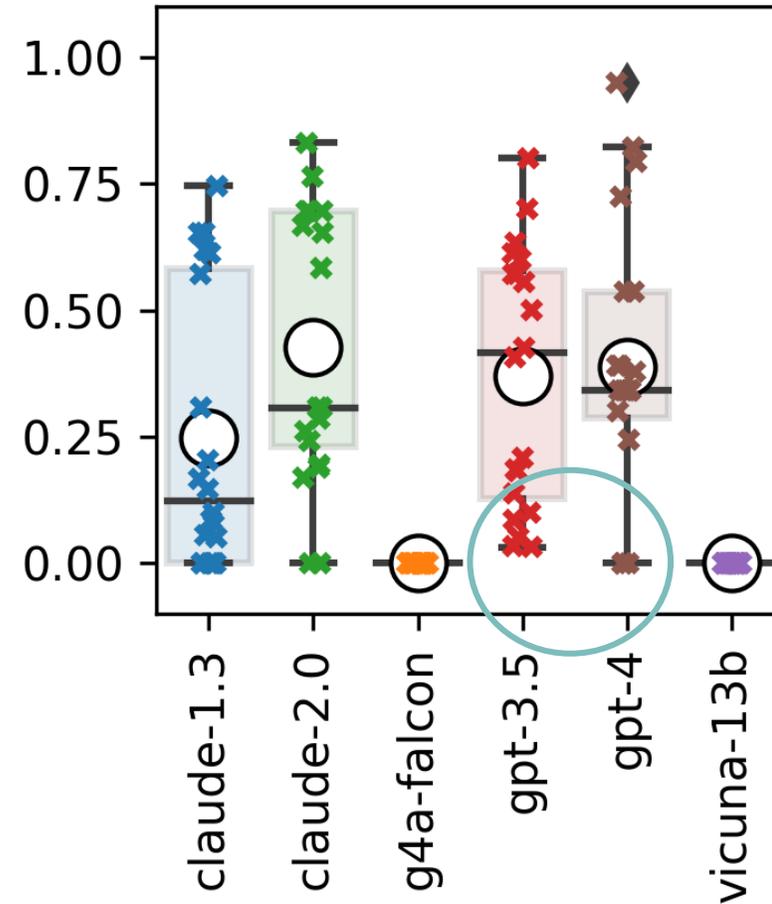
- Eingabe ist ein Klartextauszug aus einem 3D-Drucker-PDF-Datenblatt
- Prompt/Actionen sind sehr komplex
 - Detailliert bezüglich dessen, wie IRIs aussehen sollten (für einen klaren Vergleich)
 - Erfordert Vokabularwissen
 - Starke Verwendung der wichtigsten Prompt-Engineering-Techniken



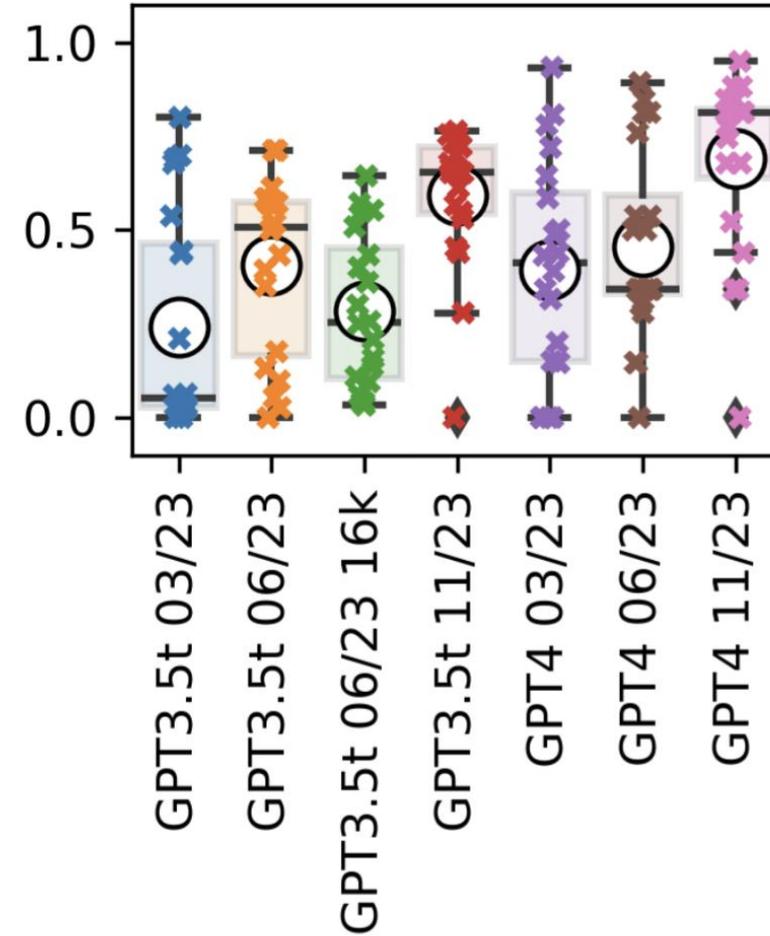
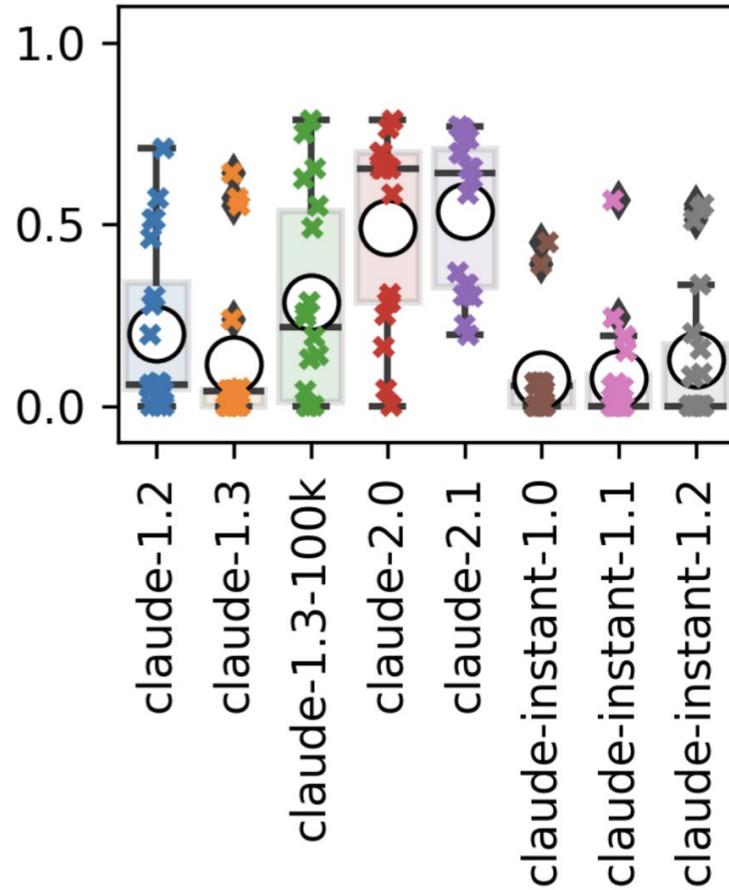
SPECIFICATION	
PRINTING	
Print technology:	FFF
Build volume:	260 × 300 × 340 mm (26 520 cm ³)
Min. layer height:	40 µm
Number of printheads:	2 per module
Nozzle diameter:	0.4/0.4 mm or 0.6/0.6 mm
Filament diameter:	1.75 mm
Printhead temperature:	500°C
Buildplate temperature:	160°C
Chamber temperature:	85°C (active heating)
Filament chamber temperature:	70°C
ENVIRONMENT	
Working temperature:	15-32°C
Storage temperature:	0-32°C
POWER	
Power requirements:	230V AC
Max power draw:	2700 W
Communication:	USB drive, SD card

Evaluation T5 mit LLM-KG-Benchmark (Aug 23)

- GPT4 has best almost perfect solution (outlier)
- GPT4 and Claude 1.3 frequently **unparseable** content, claude 2.0 sometimes too
- GPT3.5 no syntactic errors !!! !!!
- Falcon stuck in repetitive prefix gibberish patterns
- Vicuna – nothing looks like turtle



Evaluation T5 mit LLM-KG-Benchmark Evolution - OpenAI



Improving 3D Printer Data Extraction Quality

- **Initial Phase: Data Extraction from PDF Files in JSON-LD Format**
 - Data gathered from manufacturer websites.
 - Processed into text files using LLM for content extraction.
- **Challenges in the First Phase**
 - Issues with non-uniform data.
 - Varying JSONLD structures.
 - Navigating through different manufacturer terminologies.
- **Second Phase: Data Structuring**
 - Creating an Expected JSONLD Structure.
 - Consistent naming conventions for features.



The image shows a technical specification sheet for the 3DGENCE INDUSTRY F340 3D printer. The sheet is titled 'INDUSTRY F340 SPECIFICATION' and includes a photograph of the printer. The specifications are organized into several categories:

- PRINTING**
 - Print technology: FFF
 - Build volume: 260 x 300 x 340 mm (26 5/10 cm³)
 - Min. layer height: 40 µm
 - Number of printheads: 2 per module
 - Nozzle diameter: 0.4/0.4 mm or 0.8/0.6 mm
 - Filament diameter: 1.75 mm
 - Printhead temperature: 520°C
 - Buildplate temperature: 120°C
 - Chamber temperature: 85°C (active heating)
 - Filament chamber temperature: 70°C
- ENVIRONMENT**
 - Working temperature: 15-32°C
 - Storage temperature: 0-32°C
- POWER**
 - Power requirements: 230V AC
 - Max power draw: 2700 W
 - Communication: USB drive, SD card
- DIMENSIONS AND MASS**
 - External dimensions: 1200 x 800 x 1100 mm
 - Mass: 200 kg
- SOFTWARE**
 - Slicing software: 3DGENCE SLICER 4.0
- CONSTRUCTION**
 - Chassis: powder coated steel
 - Build surface: heated ceramic plate
 - Positioning resolution XY: 0.006 mm
 - Positioning resolution Z: 0.0004 mm
- SAFETY**
 - Filtration: filtration unit
 - LPS device: yes, optional
 - Others: software access override - remote shut down

Datenstrukturierung durch LLM-Ansatz

1. Erstellung eines Templates im erwarteten JSON-LD-Format mit vordefinierten Merkmalen

```
{
  "@context": "http://schema.org/",
  "@type": "Product",
  "name": "",
  "manufacturer": {
    "@type": "Organization",
    "name": ""
  },
  "additionalProperty": [
    {
      "@type": "PropertyValue",
      "name": "Build Volume",
      "value": {
        "@type": "QuantitativeValue",
        "width": {
          "@type": "QuantitativeValue",
          "value": "",
          "unitText": ""
        },
        "height": {
          "@type": "QuantitativeValue",
          "value": "",
          "unitText": ""
        },
        "depth": {
```

2. Anwenden des LLM zum Ausfüllen des Templates

SPECIFICATION	
PRINTING	
Print technology:	FFF
Build volume:	260 x 300 x 340 mm (26 520 cm ³)
Min. layer height:	40 µm
Number of printheads:	2 per module
Nozzle diameter:	0.4/0.4 mm or 0.6/0.6 mm
Filament diameter:	1.75 mm
Printhead temperature:	500°C
Buildplate temperature:	160°C
Chamber temperature:	85°C (active heating)
Filament chamber temperature:	70°C

```
{
  "@context": "http://schema.org/",
  "@type": "Product",
  "name": "INDUSTRY F340",
  "manufacturer": {
    "@type": "Organization",
    "name": "3DGence"
  },
  "additionalProperty": [
    {
      "@type": "PropertyValue",
      "name": "Build Volume",
      "value": {
        "@type": "QuantitativeValue",
        "width": {
          "@type": "QuantitativeValue",
          "value": "260",
          "unitText": "mm"
        },
        "height": {
          "@type": "QuantitativeValue",
          "value": "300",
          "unitText": "mm"
        },
        "depth": {
```

Aktuelle Herausforderungen



Halluzinationen in der Ausgabe:

Herausforderung: Sicherstellung der Genauigkeit bei der Datenextraktion, insbesondere wenn die angegebenen Schlüsselwörter oder ihre semantischen Synonyme im Produktdatenblatt völlig fehlen.

Zukünftige Arbeit: Erforschung von Techniken zur Verbesserung der Genauigkeit durch Unterscheidung gültiger Informationen von fehlerhaften Daten.



Extraktion von Bildinhalten:

Herausforderung: Extrahieren von Text aus Bildern in PDF-Dateien zur Verbesserung der Datengenauigkeit und -vollständigkeit.

Zukünftige Arbeit: Erkunden der OCR- und Bildverarbeitungsmethoden von LLMs zur Verbesserung der Datenqualität durch Extraktion von Text aus bildbasierten Inhalten.

Aus dem Labor: Transformation CSV/JSON zu KG via LLM

- RML ist Mappingsprache in Turtle
 - erlaubt Abbildung von JSON Keys oder CSV Spalten auf Ontologie
- Wie gut kann LLM diese Mapping generieren?

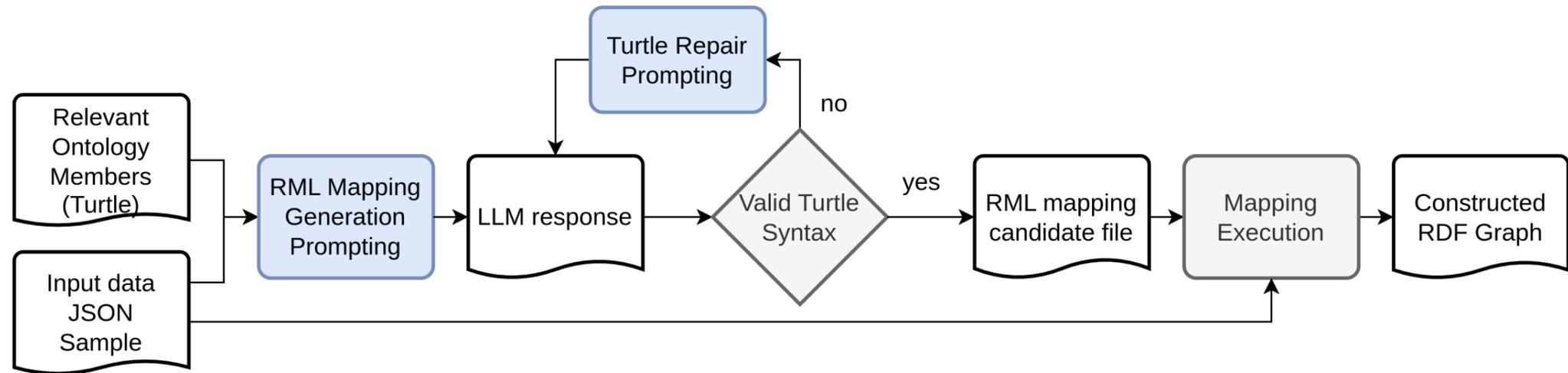
```
{
  "id": "tt0167423",
  "originalTitle": "Diamonds",
  "runtimeMinutes": 91,
  "startYear": 1999,
  "genre": ["Comedy", "Mystery"],
  "titleTyp": "movie",
  "isAdult": 0,
  "involvedPeople": [{
    "id": "tt0167423",
    "ordering": 1,
    "name": "Kirk Douglas"
    "birthYear": 1916,
    "deathYear": 2020,
    "category": "actor" }, ...]
}
```

```
# @prefix ...
@base <http://mykg.org/resource/>

<tt0167423> a dbo:Film ;
  dbo:title "Diamonds" ;
  dbo:genre "Comedy", "Mystery" ;
  dbo:startYear "1999"^^xsd:gYear ;
  dbo:Work/runtime "91"^^dtd:minute ;
  dbo:starring <nm0000018> , ... ;
  dbo:director <nm0038875> ;
  ...

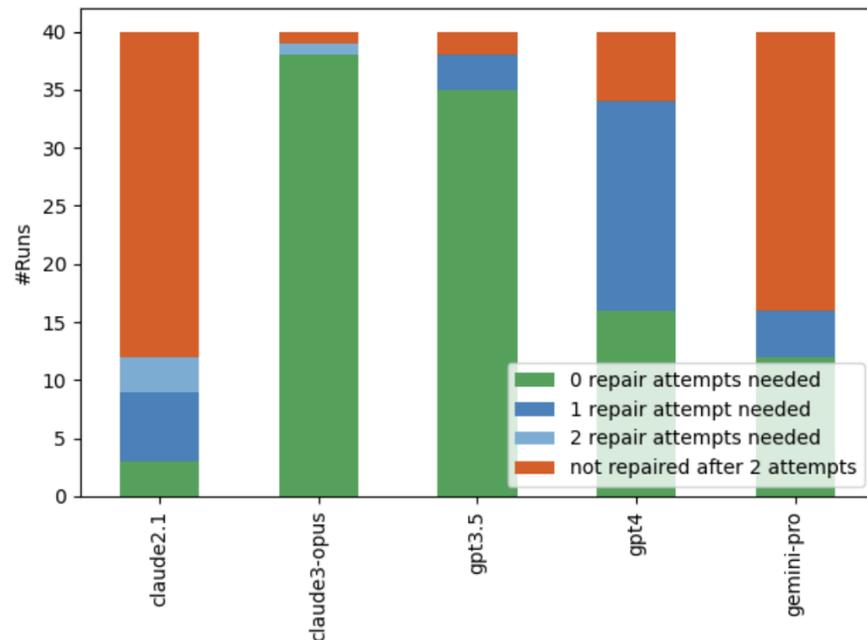
<nm0000018> a dbo:Person ;
  dbo:name "Kirk Douglas" ;
  dbo:birthYear "1916"^^xsd:gYear ;
  dbo:deathYear "2020"^^xsd:gYear .
```

Workflow CSV/JSON zu KG via LLM-generiertes RML Mapping

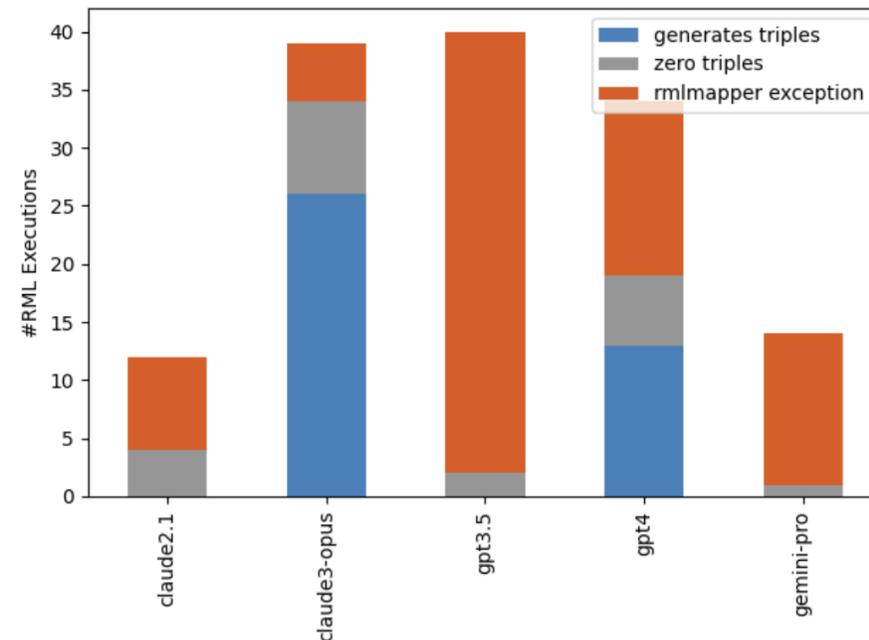


Aus dem Labor: Mapping Sinnhaftigkeit

- Wohlgeformt?
- Produziert Ergebnisse?



(a) RML Turtle Syntax Validity

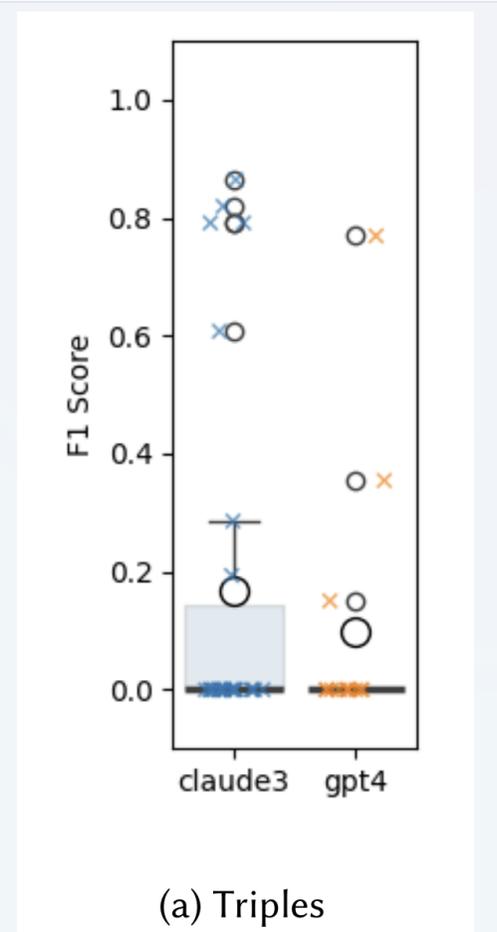


(b) Mapping Soundness

Aus dem Labor: Mapping Korrektheit

- Messungen der Korrektheit der RDF Tripel

	claude3	gpt4
mappings with triples	26	13
full predicate coverage	4	0
only ontology mapped	20	2
isAdult ordering not mapped	26	13
usage of any / custom funct.	0/0	3/3



Der Chatbot

Vision

Beispiel-Dialog

Auswertung

Datenerfassung

Die Chatbot Vision

Konfigurationsassistent der Auftragserstellung:

- Wenig erfahrene Kundschaft startet Anfrage via Chatbot
- Anforderungen werden interaktiv und an Verwendungszweck angepasst erfasst
- Ggf. Ableitung von Folgefragen auf Basis der Antworten bzw. damit einhergehendem Wissensbedarf
- Einbindung von Wissen aus dem Matching der Plattform
- Hilfestellung bei Antworten (inhaltliche Rückfragen mit Erklärungen möglich)

➔ Realisierung mit Hilfe von Large Language Models (LLMs)

Beispiel-Dialog

Hallo! Hast du eine 3D-Modelldatei für dein Bauteil? (Ja oder Nein)

Ja

Verfügst du auch über eine Fertigungszeichnung? (Ja oder Nein)

Nein

Okay, du hast eine 3D-Modelldatei, aber keine Fertigungszeichnung. Lass uns daher mit der Erfassung der Bauteilinformationen fortfahren. Wie lauten die Maße deines Bauteils in Millimetern? (Bitte gib mindestens ein Maß an: Höhe, Breite oder Länge)

300x15x10

Verstanden, Maße: 300x15x10 mm. Welche Mindest- und Maximaltemperatur muss dein Bauteil aushalten können? (Bitte in Grad Celsius angeben)

-50°C bis 90°C

Beispiel-Dialog

Maße: 300x15x10 mm, Temperaturanforderungen: -50°C bis 90°C. Nimmt das Bauteil relevante Last auf? (z.B. Biegelast, Stützlast, Zuglast, Druck) oder nimmt es wenig Last auf?

Es sollen 25kg schwere Gegenstände dran gehangen werden.

Verstanden, das Bauteil nimmt eine relevante Last von 25kg auf. Soll das Bauteil aus Kunststoff oder Metall gefertigt werden? (Falls nicht sicher, ich schlage aufgrund der Last und Temperatur Metall vor.)

Dann Metall, bitte.

Gut, dann Metall. Welche Genauigkeit benötigt deine Fertigung? (Bitte in mm Abweichung angeben)

0,1mm

Verstanden, Genauigkeit: 0,1 mm. Wie steht es um die Rauigkeit der Oberfläche? (Bitte in sa angeben oder, falls nicht relevant, sag einfach "nicht relevant")

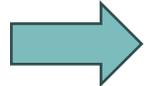
Beispiel-Dialog

Alles klar, die Rauigkeit der Oberfläche ist nicht relevant. Nun zur Produktion:
Wie viele Stück benötigst du und bis wann?

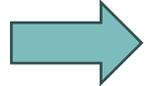
20 Stück bis 10.10.2024

```
{  
  "model_given": true,  
  "model_data_file_given": false,  
  "manual_model_data":  
  {  
    "temp_min": -50,  
    "temp_max": 90,  
    "physical_stress": 25,  
    "resolution": 0.1,  
    "surface_quality": null,  
    "material": "metal"  
  },  
  "amount": 20,  
  "date_delivery": "2024-10-10"  
}
```

Dialog-Auswertung



Systematische Erfassung des Auftrags als JSON



Zu erhebender Basisfragenkatalog aus Herstellersicht wird an LLM als Text übermittelt

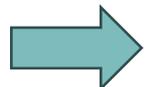
3D Modell ja/nein

Preisvorstellung ja/nein

Stückzahl

Lieferzeitraum (zeitkritisch/zeitnah/hat Zeit)

Ungefähre Abmaße vorhanden? ja/nein



LLM kann Hilfestellung bei Antworten geben aufgrund des eigenen Wissens

Überleitung des JSONs zu normaler UI, aber fertig ausgefüllt

Erinn.: Herausforderungen & wiss. Fragestellungen

- Kommerzielle LLMs besitzen bereits großes Sprach- und Prozesswissen, trotzdem wird Modellierung der Parameter und Abhängigkeiten bzgl. der Realität und ggf. individueller Hersteller-spezifischer Bedingungen benötigt
 - (Textuelle) Requirements des Kunden können deutlich variabler sein und können nicht alle a-priori mit vorgegebenen Zielfragen erfasst werden
- Benötigen Verknüpfung mit Wissensgraph und Ontologie

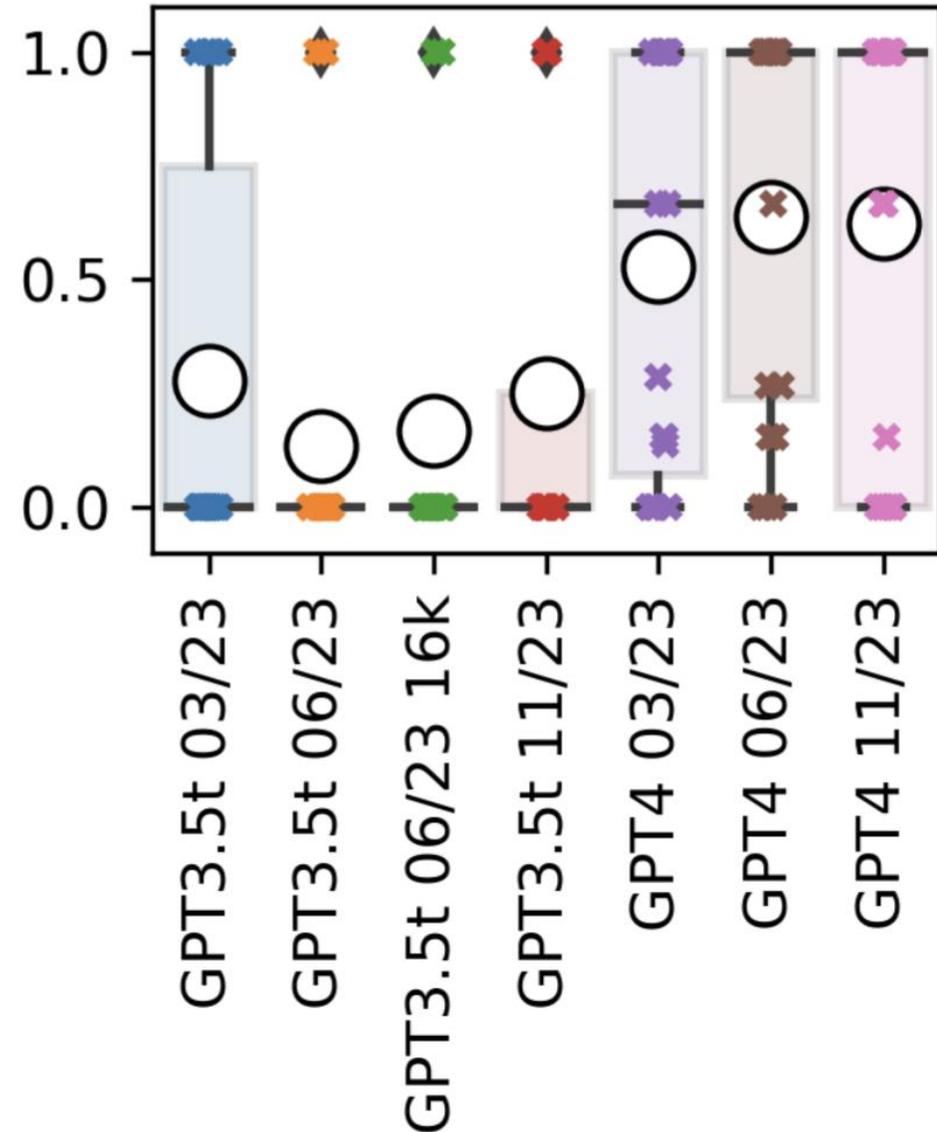
Ausblick: Komplexe Anforderungen des Kunden

- Ansatz LLM übersetzt darstellbare Aspekte (aufgrund der Ontologie / Datenbestands) in (SPARQL) Abfragen bzgl. des KGs
- "looking for *German 3d plastic print service*" translates into
 - `<AMService provider> <isLocatedIn> <Germany>`
 - `AND <AMService provider> <hasPrinterModel> <x> AND <x> <isA> <3dPLasticPrinter>`

→ Benötigt Identifizierung von Entitäten in Instanzdaten im KG („Germany“) oder Konzepten in der Ontologie

Native SPARQL-Fähigkeiten von LLMs

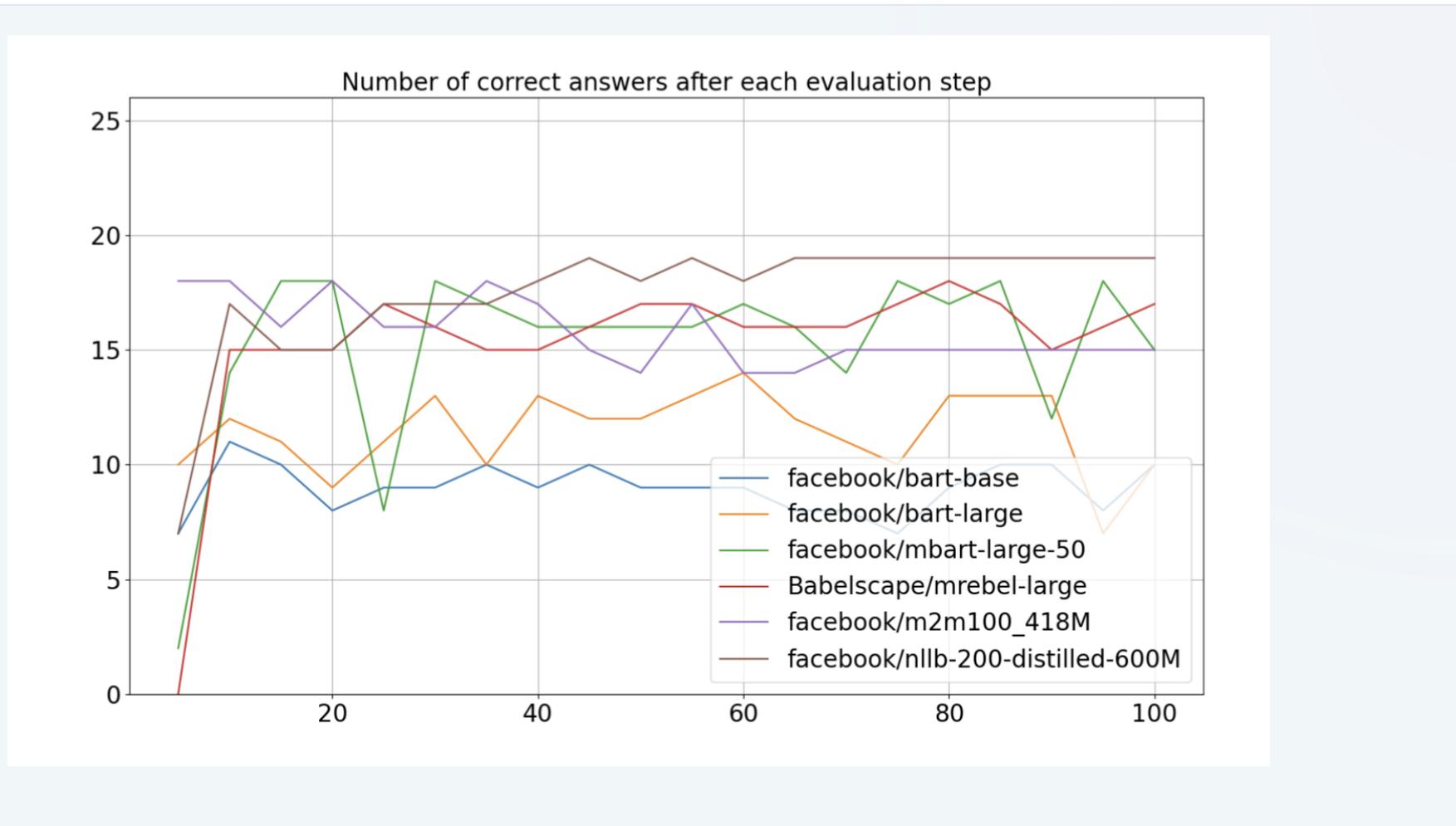
- SPARQL ist standardisierte Abfragesprache (ähnlich SQL) für KGs
- Untersuchung von Fähigkeit der LLMs Informationen aus Wissensgraphen (via Kenntnis der Ontologie) mit SPARQL abzufragen (Text2Sparql) auf Basis von Wikidata



Native LLM-SPARQL-Fähigkeiten?

- Probleme
 - Ontologie (noch) nicht bekannt für LLM
 - Ontologiewissen zu detailliert um alle daten exakt zu Wissen (Wikidata)
 - Ontologie zu groß/zu teuer um in Prompt einzubetten (Wikidata)
- Idee 1: trainieren LMs / (sehr) kleiner LLMs gezielt auf einen Wissensgraphen
- Idee 2: nötiges Ontologiewissen über RAG einbinden (Future Work)

Aus dem Labor: Finetuning von LLMs für SPARQL bzgl. spez. KGs



Semper-KI Vision ...

- Vision 2 – Verbessertes Matching
 - 4 wesentliche Matching-Dimensionen (Zeit der Auftragsbearbeitung, Resilienz-Score der Anbieter, Qualität und Preis ... erweiterbar)
 - Simulationsverfahren für aggregierte Bewertung der Dimensionen
 - Assistenz für Preiskalkulation für Anbieter
- Vision 3 - Aufzeigen von Alternativen Lösungen
 - Kunde weiß z. B. nicht, dass ähnliches Material, für seinen Anwendungsfall hinreichend gute Eigenschaften hat, dafür aber vielleicht verfügbarer oder preiswerter ist, usw.
 - Mögliche Nacherfassung mit Chat-Dialog notwendig
- Vision 4 – Anbindung weitere Module und Produktionsverfahren
 - z.B. nicht nur 3D-Druck, neue Nachbearbeitungsverfahren, Live-3D-Experten-Dialog, u.v.m.