Stadt Zürich
Statistik

Mit den Daten sprechen: MeLODy, der zuverlässige KI-Bot der Stadt Zürich

TRANSFORM 2025
6. Mai 2025
Dr. Rolf Schenker

https://www.dbta.com/Editorial/Trends-and-Applications/How-to-Harvest-the-Worlds-Largest-Source-of-Data-Web-Data-132105.aspx

Agenda

- MeLODy kennenlernen
- Datengrundlage
- Funktionsweise
- Ausblick

MeLODy 1.2

Von community builder

Stelle irgendeine Frage

MeLODy 1.2

Von community builder

Stelle irgendeine Frage

Vorteile von MeLODy

Gegenüber dem herkömmlichen Web-Angebot bzw. anderen Bots

- Natürlichsprachliche Schnittstelle:
  - ist mehrsprachig
  - kennt Synonyme etc.
- Barrierefreier Zugang zu unseren Daten.
- Zuverlässige Antworten, insbesondere für einen Chatbot.
- Liefert Kontextwissen (wie beispielsweise Fertilitätsrate).
- Viel mehr Kontrolle darüber, welche Daten Benutzer*innen präsentiert bekommen.

Datengrundlage

Datengrundlage für MeLODy

Semantisch aufbereitete Daten

- Früher haben wir unsere Daten als Excel- und CSV-Dateien publiziert.
- Viele Metadaten (welches Jahr betreffen die Daten, was ist ihre Quelle, welche Definitionen verwenden sie, …) sind zwar dokumentiert, aber nicht direkt «bei den Daten».

T_1

Anteil der ausländischen Wohnbevölkerung und Bevölkerung nach Herkunft und Geschlecht

1901 – 2024

Quelle: Statistik Stadt Zürich, BVS

Jahr	Total	Anteil ausländische Bevölkerung (%)	Schweizer Bevölkerung	Schweizerinnen	Schweizer	Ausländische Bevölkerung
1901	150 301	28.0	108 291	...	...	
1902	153 741	28.0	110 707	...	...	
1903	158 719	28.9	112 915	...	...	
1904	164 060	29.5	115 612	...	...	
1905	168 877	30.6	117 236	...	...	
1906	173 283	31.6	118 600	...	...	
1907	177 899	32.1	120 861	...	...	
1908	180 722	32.1	122 662	...	...	
1909	185 041	32.4	125 001	...	...	
1910	190 092	33.5	126 470	...	...	
1911	195 525	33.8	129 396	...	...	
1912	200 462	34.2	131 976	...	...	

Datengrundlage für MeLODy

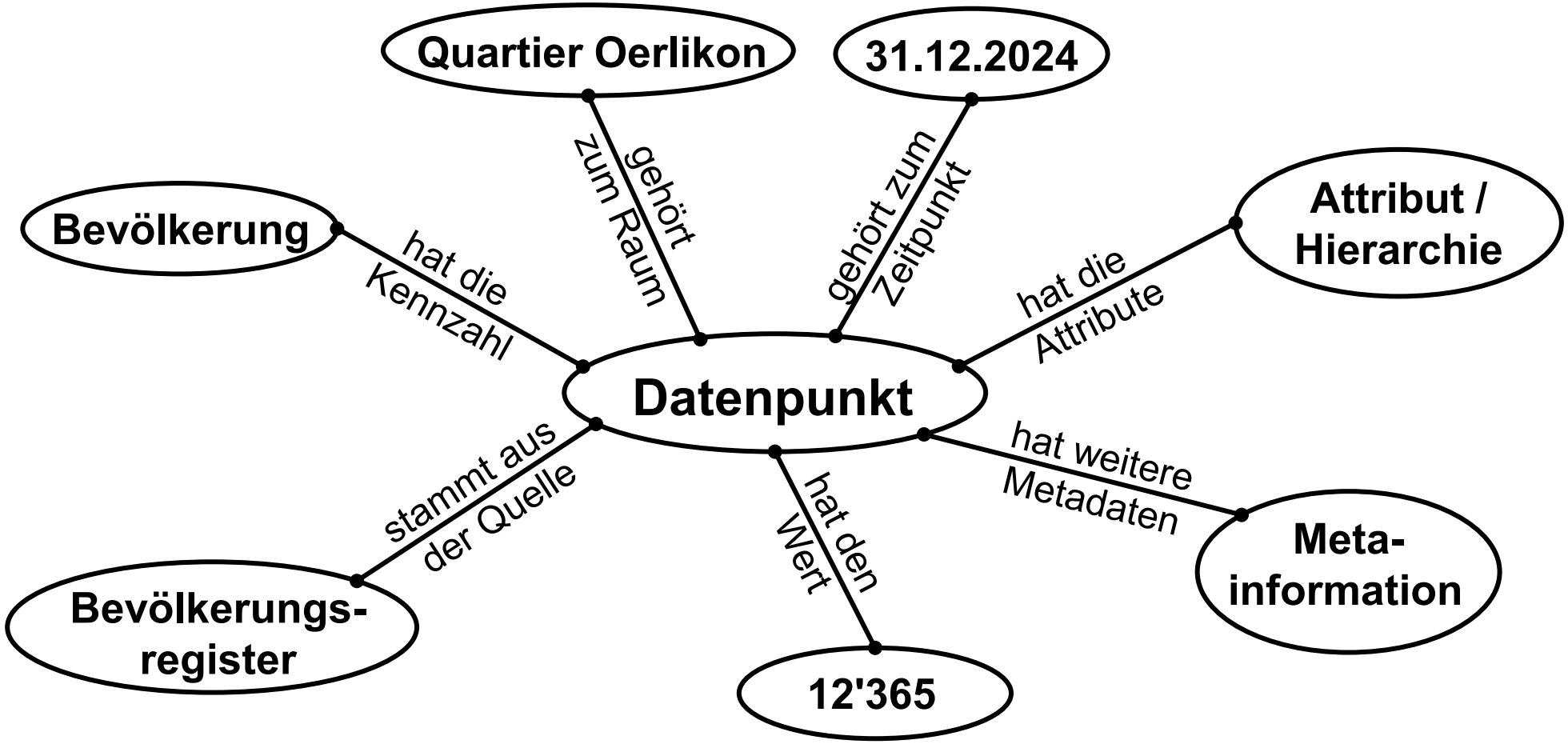
Semantisch aufbereitete Daten

- Früher haben wir unsere Daten als Excel- und CSV-Dateien publiziert.
- Viele Metadaten (welches Jahr betreffen die Daten, was ist ihre Quelle, welche Definitionen verwenden sie, …) sind zwar dokumentiert, aber nicht direkt «bei den Daten».

- Nun publizieren wir einen Teil unserer Daten als Linked Data.
- Damit sind sie vollständig semantisch beschrieben.
- KI-Systeme können Linked Data sehr viel besser interpretieren als Excel- oder CSV-Dateien.

Statistische Daten als Linked Data

Datenstruktur



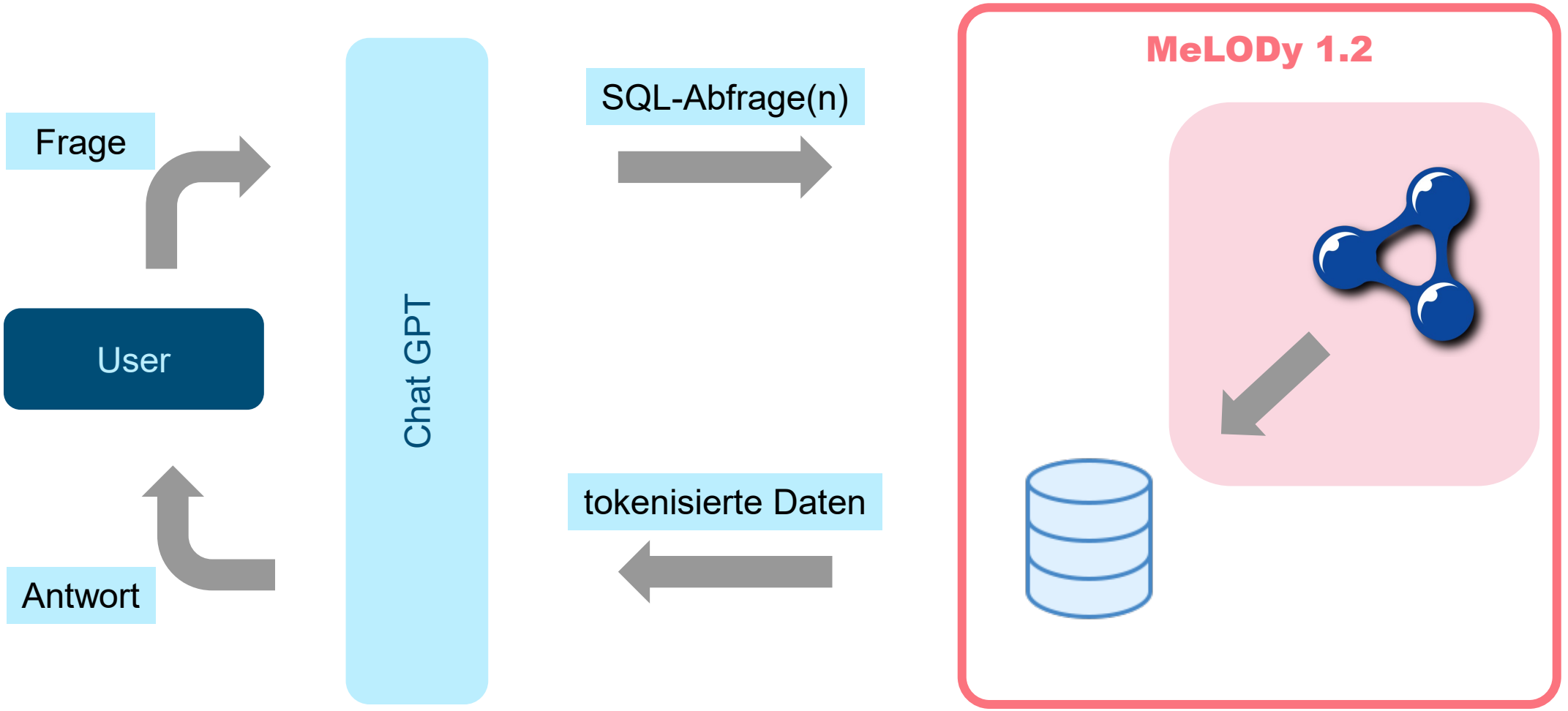
Funktionsweise

Was ist MeLODy

- Basiert auf ChatGPT, ein sogenannter «custom GPT».
- Dieser hat einige Anweisungen bekommen:
  - Verwende die Linked Data von Statistik Stadt Zürich als Grundlage.
  - Berücksichtige nur dann andere Quellen, wenn für die Frage die Angaben in den Linked Data nicht ausreichen.

- Im Hintergrund läuft ein RAG-Modell (Retrieval Augmented Generation).
- Wie andere Chatbots wird das jeweils nächste Wort mit einem stochastischen Modell berechnet.

Architektur MeLODy 1.2



Ausblick

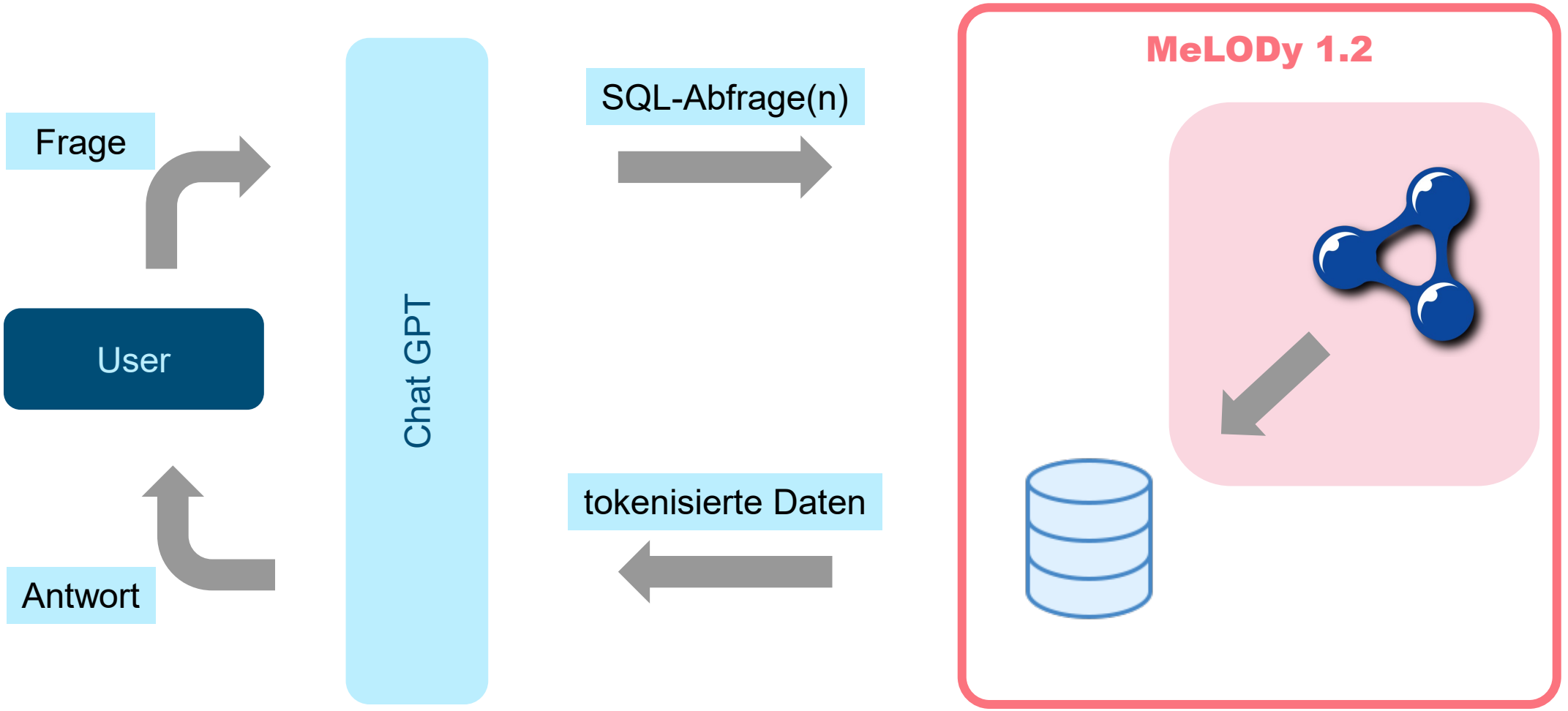
Ausblick

- MeLODy ist immer noch ein Proof of Concept und wird von der Firma Redlink entwickelt.

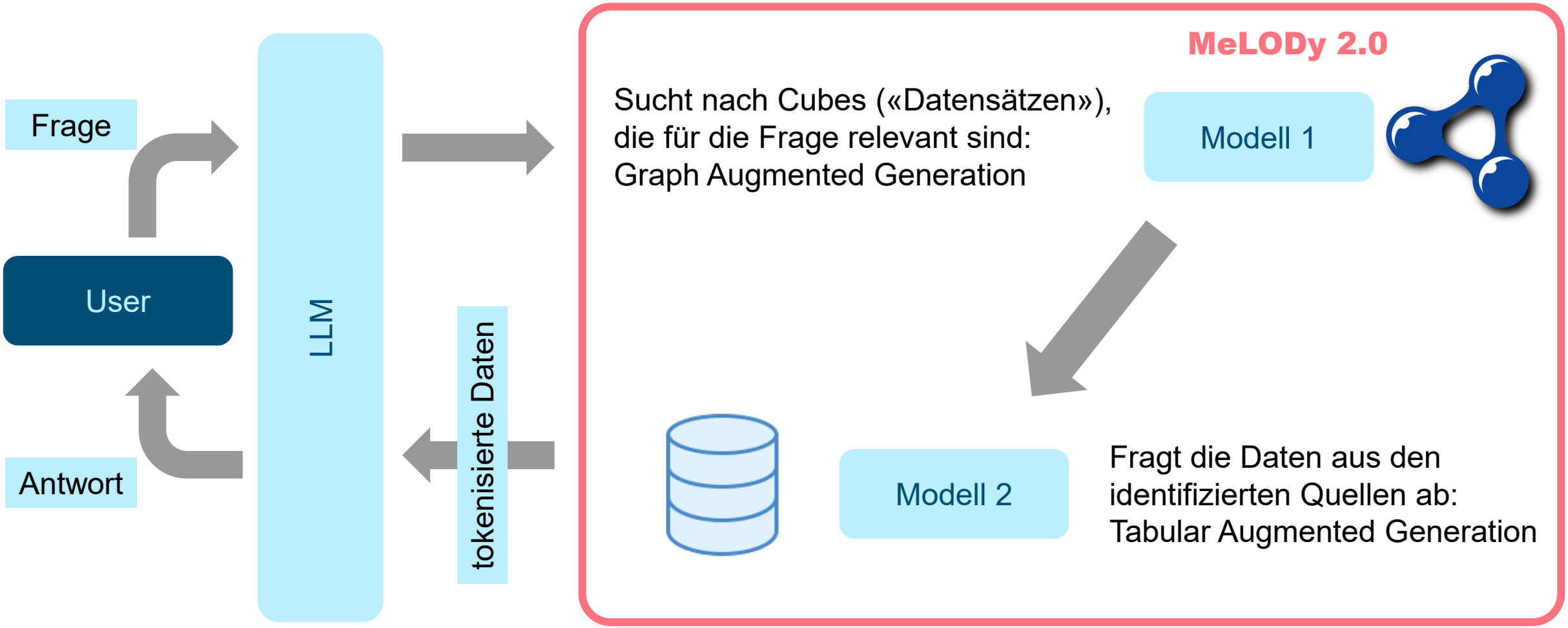
- Wir werden immer mehr Daten als Linked Data publizieren, damit stehen MeLODy immer mehr Daten zur Verfügung.

- MeLODy soll nicht mehr auf Chat GPT basieren:
  - Aktuell ist MeLODy nur mit einem Chat GPT Account zugänglich.
  - Wir möchten MeLODy möglichst komplett öffentlich zur Verfügung stellen.

Architektur MeLODy 1.2



Architektur MeLODy 2.0



Fazit

Fazit

MeLODy…

– spricht fliessend zahlreiche Sprachen.
– liefert faktenbasierte Antworten.

– greift auf unsere Linked Data zu, deshalb müssen wir die Daten nicht «extra dafür bereitstellen».

– ist OpenSource: https://github.com/StatistikStadtZuerich/MeLODy.
– ist frei zugänglich: https://tinyurl.com/4ye98twn.

Vielen Dank.

Bern, 06.05.2025
Dr. Rolf Schenker

rolf.schenker2@zuerich.ch