

No more SQL: Anwendung von Machine Learning für Text2SQL

Damit Datenbanken durchsucht werden können, braucht es professionelle Anwender und Kenntnisse von Abfragesprachen wie SQL (Structured Query Language). Die natürliche Sprache kann als Ergänzung zu dieser Lücke ein mächtiges Tool sein. Da Computer natürliche Sprachen nicht verstehen und repräsentieren können, ist eine Schnittstelle für Benutzerabfragen in natürlicher Sprache nötig. Search Over Data Warehouses (SODA) ist eine Applikation, welche verschiedene Funktionalitäten bietet, um Informationen von Datenbanken zu durchsuchen oder zu visualisieren. SODA übersetzt Benutzerabfragen in SQL, damit Informationen aus Datenbanken extrahiert werden können.

In SODA können Schlagwörter, jedoch keine ganzen Sätze eingegeben werden. Durch die vorliegende Bachelorarbeit sollen Lösungswege gefunden werden, um Suchabfragen auch in natürlicher Sprache erstellen zu können. Dazu wurde SODA mithilfe von Natural Language Processing (NLP), Machine Learning (ML) und verschiedenen regelbasierten Methoden, welche in Java und Python implementiert wurden, erweitert. Durch diese Erweiterungen können einfache Suchabfragen in natürlicher Sprache erstellt und Datenbanken nach Informationen auf herkömmlichere Weise abgefragt werden. Die neu entwickelten Funktionalitäten in SODA vereinfachen für den Benutzer den Suchprozess nach Informationen in Datenbanken.



Diplomierende
Yusuf Koc
Erland Xhoxhaj

Dozierende
Mark Cieliebak
Kurt Stockinger

```
Where is London? Go
SELECT country.country
FROM sakila.city, sakila.country
WHERE
((city.city like "%london%")) AND
(city.country_id=country.country_id)
country
United Kingdom
Canada
2

What is the country of Dallas? Go
SELECT country.country
FROM sakila.city, sakila.country
WHERE
((city.city like "%dallas%")) AND
(city.country_id=country.country_id)
country
United States
1

In which country is Salzburg? Go
SELECT country.country
FROM sakila.city, sakila.country
WHERE
((city.city like "%salzburg%")) AND
(city.country_id=country.country_id)
country
Austria
1
```

Drei SODA-Suchabfragen in natürlicher Sprache.