

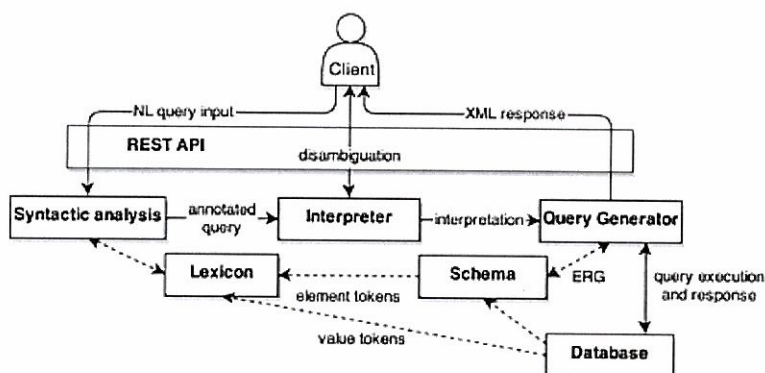
SPLASH Chatbot

Řešitel: Ing. Jakub Stejskal, FEL ČVUT

Vedoucí projektu: Ing. Jan Šedivý, CSc., FEL ČVUT

Databázová rozhraní v přirozeném jazyce (NLIDB) jsou systémy schopné překladu uživatelských dotazů napsaných v přirozeném jazyce do formálních databázových dotazů za účelem získání požadované informace z databáze.

Tento projekt řešil využití RESTových webových služeb jakožto báze znalostí pro systém NLIDB, a to místo běžně používaných relačních databází. Byl navržen framework umožňující snadné vytváření dotazovacích aplikací pro libovolné domény, který podporuje jak relační databáze, tak RESTová aplikační rozhraní. Důraz byl kladen na minimalizaci času a úsilí vynaloženého na konfiguraci pro novou doménu znalostí. Systém byl prakticky ověřen implementací aplikace NaLIDa, což je NLIDB poskytující rozhraní univerzitního informačního systému KOS prostřednictvím služby RESTful s názvem KOSapi.



Pro zodpovězení zadaného dotazu systém anotuje vstup v přirozeném jazyce za pomoci nástroje StanfordCoreNLP a výsledné anotace využívá ke generování a filtrování hypotéz sémantické interpretace vstupu. Po následné disambiguaci se interpretace mapuje na entity v grafové reprezentaci databáze, ve které je algoritmem pro hledání nejkratších cest nalezeno optimální spojení entit interpretace. Tato cesta je nakonec transformována do posloupnosti požadavků do RESTového API.



Welcome to NaLIDa a Natural Language Interface for Database

Ask a question from the domain of KOS, information system of CTU in Prague. NaLIDa knows stuff about students, teachers, divisions, courses, parallels and exams.

Check the [enumeration of supported entities, attributes and their natural language tokens](#) in XML format for inspiration. Click on the examples or submit your own natural language question in English. Use either a wh-question or a noun phrase.

KOSapi response

auto-disambiguate

Submit

Examples

What is phone of Jan Šedivý?

Which teachers are from division 13133?

SQL query

auto-disambiguate

Submit

Examples

What is phone of Jan Šedivý?

Which teachers are from division 13133?

Detailed log

auto-disambiguate

Submit

Examples

What is phone of Jan Šedivý?

Which teachers are from division 13133?

Experimenty s uživateli ukázaly, že pro platné vstupy systém vykazuje přesnost 96,8 %, úplnost 85,9 % a průměrnou dobu odezvy 1,6 vteřiny. NaLIDa tedy dokáže v reálném čase zodpovídat širokou škálu dotazů zadaných v přirozené angličtině, které se týkají studentů, učitelů, předmětů, rozvrhů, zkoušek apod. Může se tak studentům stát cennou pomůckou při rychlém získávání informací o studiu. Systém je podrobně popsán v diplomové práci řešitele s názvem „Data-oriented conversational system“ (FEL ČVUT, 2014).