

Het PHLIQA Question-Answering Project

J. Landsbergen 1989

Omschrijving

PHLIQA (PHiLips Question Answering) was de naam van een project dat tot doel had gegevensbestanden eenvoudiger toegankelijk te maken door in natuurlijke taal (Engels) geformuleerde vragen toe te laten. Hiertoe werd een systeem ontwikkeld dat de ingevoerde vraag via een aantal tussenstappen omzette in een formele query die met behulp van de gegevens in de data base beantwoord werd.

Kenmerkend voor PHLIQA was het gebruik van drie opeenvolgende logische talen:

- EFL (English-oriented Formal Language), waarvan de termen corresponderen met betekenissen van woorden,
- WML (World Model Language), waarvan de termen staan voor begrippen uit het wereldmodel,
- DBL (Data Base Language), waarvan de termen de objecten in de data base aanduiden.

Het systeem bestond voor een groot deel uit vertaalslagen: van Engels naar EFL, van EFL naar WML en van WML naar DBL, gevolgd door de evaluatie van de DBL-expressie, die het antwoord opleverde.

Het project heeft twee prototype-systemen opgeleverd, in 1975 PHLIQA1 en in 1979 een tweede versie PHLIQA1+. De systemen werden ontwikkeld op een P1400 computer in de programmeertaal SPL. PHLIQA1+ was voor die tijd een groot programma, dat 480 K geheugen nodig had, wat met veel moeite d.m.v. overlay technieken teruggebracht kon worden tot 220 K.

Aanleiding

Het onderzoek was gestart bij Philips-Electrologica, Apeldoorn, in de afdeling GSO (Geavanceerde Systeemontwikkeling). Men was geïnteresseerd in het nieuwe vak Artificiële Intelligentie en het gebruik ervan voor het beter structureren en ontsluiten van data bases. In 1972 werd de afdeling GSO ontmanteld, maar in overleg met het Nat. Lab. werd besloten het PHLIQA-werk daar voort te zetten, d.m.v. detachering van vier Electrologica. medewerkers en toevoeging van twee Nat.Lab. medewerkers. Na 1975 werd het een volledig Nat.Lab. project.

Bijdrage aan onderzoeksgebied

PHLIQA was destijds uniek door de systematische behandeling van de verschillende aspecten die bij de interpretatie van een vraag een rol spelen, zoals de syntaxis, de logische semantiek, het model van het onderwerpgebied en de structuur van de data base. Het is een opzet waarmee men nog steeds voor de dag kan komen. Nog onlangs is bij het Amerikaanse onderzoeksinstituut Bolt, Beranek and Newman een systeem ontwikkeld met dezelfde drie interpretatie-niveaus (EFL, WML en DBL) als in PHLIQA. Het gebruik van logische semantiek in question-answering systemen, dat in de eerste helft van de jaren zeventig niet populair was, is later gemeengoed geworden.

Omvang van het onderzoek

Het project liep van 1972 t/m 1979 en telde gemiddeld 5 à 6 medewerkers.

Gevolgen

Het project heeft niet tot een Philips-product geleid. De interesse van Data Systems in dit soort geavanceerde toepassingen nam in de loop van het project snel af. PHLIQA is wel een goed uitgangspunt geweest voor twee volgende onderzoeksprojecten: het Rosetta-project aan automatisch vertalen en het SPICOS-project, waarin een question-answering systeem met spraak-invoer en -uitvoer wordt ontwikkeld. Het taalinterpretatie-stuk van SPICOS (grotendeels door Siemens ontwikkeld) is rechtstreeks op PHLIQA gebaseerd.