

## INTISARI

# IMPLEMENTASI *PARSING* KALIMAT BAHASA INDONESIA BERDASARKAN PENGENALAN PREDIKAT PADA SISTEM TANYA JAWAB DENGAN TEKNOLOGI WEB SEMANTIK

## STUDI KASUS: DOMAIN OTOMOTIF

Oleh

Mohammad Syarief

11/325406/PPA/03714

Ketersediaan dan teknik perolehan data serta bentuk penyajiannya merupakan hal yang penting dalam sistem informasi. Penyediaan dan teknik perolehan data dengan teknologi web semantik dapat mengantisipasi ambiguitas dan melakukan inferensi data. Ontologi sebagai tulang punggung web semantik diperlukan pada semua domain. Ontologi kendaraan bermotor berbahasa Indonesia sejauh ini belum ditemukan. Dari perspektif bisnis kendaraan, dibutuhkan ontologi yang mampu menjembatani kepentingan penyedia jasa dan data kendaraan. Selain ketersediaan dan perolehan data, cara penyajian data juga merupakan hal yang tak kalah pentingnya. Dalam rangka penyediaan data yang tepat sesuai dengan informasi yang hendak dicari oleh pengguna, maka penggunaan sistem tanya-jawab menggunakan kalimat tanya lebih tepat digunakan daripada sistem pencarian dengan kata kunci berbentuk frasa. Dalam sistem tanya-jawab, untuk mendapatkan jawaban yang benar, maka sistem harus bisa mengenali kalimat tanya dengan tepat, terutama pada kalimat kompleks.

Dalam penelitian ini, dikembangkan sistem tanya-jawab berbahasa Indonesia tentang informasi kendaraan bermotor menggunakan teknologi web semantik yang dikenal dengan SiTaSO. Ontologi yang dikembangkan mampu mengakomodasi kepentingan penyedia jasa penjualan dan persewaan kendaraan, disamping menangani hirarki *class* dan properti kendaraan. Pendekatan teknik *parsing* yang digunakan adalah *top-down parsing* yang bertumpu pada pengenalan predikat kalimat sederhana. Pengenalan kalimat kompleks dilakukan dengan cara memecah kalimat kompleks menjadi beberapa kalimat sederhana terlebih dahulu.

SiTaSO dirancang untuk mengenali lebih dari 250 pola kalimat tanya dengan 8 predikat utama. Dari hasil pengujian, SiTaSO berhasil menjawab dengan benar 77,5% dari pertanyaan.

Kata-kata kunci : sistem tanya-jawab, QA, konversi kalimat, pengenalan kalimat, web semantik, inferensi, ontologi, kendaraan bermotor, otomotif.

## ABSTRACT

# IMPLEMENTATION OF INDONESIAN SENTENCE PARSING BASED-ON PREDICATE-ARGUMENT RECOGNITION IN QUESTION-ANSWERING SYSTEM WITH SEMANTIC WEB TECHNOLOGY

## CASE STUDY: AUTOMOTIVE DOMAIN

By

Mohammad Syarief

11/325406/PPA/03714

Data availability, acquisition techniques and presentation is essential in information systems. Data provision and acquisition techniques with semantic web technologies can anticipate ambiguity and perform data inference. Ontology as semantic web backbone is required in all domains. Automotive ontology in Indonesian language has not been found so far. In the automotive business perspective, ontology which connect between service provider needs (dealers or rental agency) and automotive data is required. As data availability and acquisition, data presentation is important as well. In order to provide the right data based on information searched by user, then using question-answering system with interrogative sentence is more appropriate than search engine with phrases keyword. In order to get the right answer, a question-answering system must be able to recognize the interrogative sentence correctly.

In this research, question-answering system of automotive information using semantic web technology known as SiTaSO is developed in Indonesian Language. Main ontology (ontoKendaraan) is developed by building hierarchy of classes and properties in automotive domain and accommodating needs of dealer and rental agency as well. Parsing technique approach used in SiTaSO is top-down parsing. This approach is developed based on predicate recognition of a simple sentence. The recognition of complex sentence carried out by splitting it into simple sentences in advance.

SiTaSO is designed to identify more than 250 question patterns with 8 main predicates. Based on test results, SiTaSO managed to answer 77,5% of given questions correctly.

Keywords : question-answering system, QA, *parsing*, sentence recognition, semantic web, *inference*, ontology, automotive.