

Penerapan Ontology Mining untuk Question Answering pada Domain Kesehatan Rahmad Fitrianta Eka Saputra¹, Shaufiah², Indra Lukmana Sardi³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹rahmadfitrianta@students.telkomuniversity.ac.id, ²shaufiah@telkomuniversity.ac.id,

³indraluk@telkomuniversity.ac.id

Abstrak

Kesehatan memerlukan eksplorasi yang efisien pada beragam referensi untuk menemukan solusi. Penggunaan ontologi dalam Question Answering System (QAS) merupakan solusi potensial. Sistem ini menafsirkan pertanyaan pengguna dalam bahasa alami dan mendapatkan respons yang tepat dari repositori pengetahuan. Membangun sistem ini melibatkan penggunaan pendekatan ontologi menggunakan RDF dan SPARQL untuk ekstraksi data melalui API PUBMED. Penilaian ini menggunakan pengujian Black Box, sebuah metode yang meneliti perangkat lunak tanpa menyelidiki cara kerja internalnya, dan memperlakukannya sebagai entitas yang tidak jelas. Meskipun ada pendekatan ini, kinerja sistem dievaluasi secara eksternal. Hasil menunjukkan bahwa sebagian besar jawaban akurat ditemukan pada posisi terdepan di antara pertanyaan yang diproses.

Kata kunci : question and answering system, ontology, SPARL, RDF

Abstract

Health necessitates an efficient exploration across diverse references to find solutions. Employing ontology in the Question Answering System (QAS) stands as a potential resolution. This system interprets user queries in natural language and derives precise responses from knowledge repositories. Constructing this system involves employing the ontology approach using RDF and SPARQL for data extraction via the PUBMED API. The assessment employs Black Box testing, a method that scrutinizes software without delving into its internal workings, treating it as an opaque entity. Despite this approach, the system's performance is evaluated externally. Results indicate that the majority of accurate answers are found at the foremost position among processed queries.

Keywords: question and answering system, ontology, SPARL, RDF

1. Pendahuluan

Evolusi aksesibilitas informasi yang pesat [1] telah membanjiri pengguna dengan banyak sekali data, namun seringkali sulit untuk diautentikasi. Untuk mengurangi masalah ini, Question Answering Systems (QAS) telah muncul, membantu pengguna dengan menafsirkan pertanyaan bahasa alami dan memberikan tanggapan tekstual [2]. Kesehatan, khususnya, merupakan domain yang memerlukan integrasi QAS yang kuat untuk memenuhi urgensinya [3]. Salah satu cara QAS adalah melalui Ontologi, memfasilitasi sintesis informasi dari permintaan pengguna [4].

Pengembangan QAS mencakup tiga pendekatan berbeda: pencocokan pola, statistik, dan berbasis ontologi [5]. Studi ini berpusat pada penggunaan pendekatan ontologi dalam QAS. Data kesehatan sebagian besar mengadopsi struktur Resource Description Framework (RDF), menggunakan skema ontologi yang rumit dan bahasa kueri SPARQL [6]. Model RDF unggul dalam menjelaskan hubungan data dan menawarkan jalur analitis langsung yang memanfaatkan ontologi [6]. Upaya penelitian seperti NITELIGHT [7] dan BioSPARQL [8] telah menyederhanakan sistem kueri dengan memanfaatkan ontologi berbasis RDT, memungkinkan eksplorasi biomedis melalui konsep ontologi dan subgraf relevan yang berkaitan dengan pertanyaan pengguna dalam biomedis. Query didekomposisi menjadi rangkap tiga RDF, kemudian diterjemahkan ke dalam SPARQL, mewakili subjek, hubungan, dan objek, menyelaraskan dengan konsep ontologi [6].

Studi ini memelopori QAS berbasis ontologi untuk kesehatan, dengan fokus pada kumpulan data kardiovaskular, memanfaatkan metode kueri RDF dan SPARQL. Penerapan pendekatan ontologi menjanjikan representasi pertanyaan dan jawaban yang lebih tepat, selaras dengan pengetahuan, semantik, dan konsep hierarki. Memproses pertanyaan dalam kerangka berbasis ontologi ini menawarkan peningkatan pemahaman terhadap hubungan dan konsep.