

Analisis Dan Peningkatan Kualitas Framework Lex2kg Dalam Pembuatan Knowledge Graph Peraturan Perundang-Undangan = Analysis and Quality Improvement of Lex2kg Framework for Conversion of Legal Document to Knowledge Graph

Muhammad Ilham Al Ghifari, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920551510&lokasi=lokal>

Abstrak

Framework Lex2KG dibuat untuk mengekstraksi dokumen PDF peraturan perundang-undangan menjadi dokumen KG berbentuk RDF triple. Setelah pemeriksaan, framework masih memiliki permasalahan pada tahapan ekstraksinya mengakibatkan kurangnya kualitas KG pada aspek kualitas accuracy dan completeness. Sehingga dilakukan perbaikan framework Lex2KG untuk menghindari permasalahan yang muncul dan dibutuhkan sistem lain untuk menjaga dan meningkatkan kualitas. Pada penelitian dilakukan analisis serta perbaikan framework Lex2KG yaitu peningkatan jumlah dokumen yang dapat terekstraksi menjadi KG sehingga dapat mengekstraksi 1353 dokumen Undang-Undang (UU) dan 963 dokumen mempunyai Jumlah Pasal yang lengkap. Sementara itu, framework Lex2KG sebelum perbaikan hanya dapat mengekstraksi 784 dokumen dan 563 dokumen. Selain mengekstraksi dokumen UU, pada penelitian ini framework Lex2KG dapat mengekstraksi 3864 dari 4758 dokumen Peraturan Pemerintah (PP) menjadi data KG. Penelitian ini juga membuat SHACL shape untuk memvalidasi data KG sehingga ditemukan 60 dokumen UU yang tidak memiliki judul dikarenakan perbedaan format penulisan pada dokumen PDF nya. Untuk memahami dan menganalisis data Legal KG, dibuat kode visualisasi data KG. Visualisasi ini berbentuk statistik dan graph. Penulis juga membuat dataset yang berisikan pertanyaan beserta jawabannya untuk menjaga kualitas aplikasi Legal VA menggunakan sumber data hasil ekstraksi Lex2KG guna memastikan kualitas jawaban yang dikembalikan oleh aplikasi Legal VA akurat dan sesuai.

.....The Lex2KG framework was created to extract PDF documents of laws and regulations into KG documents in the form of triple RDF. After inspection, the framework still has problems at the extraction stage resulting in a lack of KG quality in terms of accuracy and completeness. So that the Lex2KG framework is improved to avoid problems that arise and other systems are needed to maintain and improve quality. In the research, an analysis and improvement of the Lex2KG framework was carried out, namely increasing the number of documents that could be extracted into KG so that 1353 Law documents were extracted and 963 documents had a complete number of articles. Meanwhile, the Lex2KG framework before the repair could only extract 784 documents and 563 documents. In addition to extracting law documents, in this study the Lex2KG framework was able to extract 3,864 out of 4,758 Government Regulation (PP) documents into KG data. This study also created a SHACL shape to validate KG data so that 60 UU documents were found that did not have titles due to differences in the writing format of the PDF documents. To understand and analyze Legal KG data, a KG data visualization code is generated. This visualization is in the form of statistics and graphs. The author also creates a dataset containing questions and answers to maintain the quality of the Legal VA application using data sources extracted from Lex2KG to ensure the quality of the answers returned by the Legal VA application are accurate and appropriate.