

Lex2KG: Automasi Pembuatan Knowledge Graph dari Dokumen Peraturan Perundang-undangan Indonesia = Lex2KG: Automated Legal Knowledge Graph Population from Indonesian Legal Documents

Muhamad Abdurahman, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920534258&lokasi=lokal>

Abstrak

Dokumen peraturan perundang-undangan pada umumnya tersedia dalam bentuk PDF yang bersifat tidak machine-readable, sehingga data tidak dapat diproses secara otomatis dan dalam skala besar oleh komputer untuk dimanfaatkan dalam berbagai teknologi digital. Oleh karena itu diperlukan struktur data yang dapat memuat informasi peraturan perundang-undangan, beserta sistem yang melakukan konversi dari PDF menjadi struktur data tersebut. Dengan alasan tersebut, pada penelitian ini penulis mengembangkan Lex2KG, framework untuk mengonversi dokumen PDF peraturan perundang-undangan di Indonesia (Lex berasal dari Bahasa Latin yang berarti hukum) menjadi knowledge graph. Knowledge graph (KG) adalah graph yang menggambarkan entitas dunia nyata beserta keterkaitannya dan memberikan informasi terstruktur yang machine-readable. Pada penelitian ini KG dipilih dari berbagai struktur data yang tersedia karena KG terkategori sebagai 5-star data menurut 5-star deployment scheme for Open Data, yaitu data dengan jenis informasi paling bermanfaat, memberikan data dalam bentuk open license, terstruktur, tersedia dalam open format, menggunakan URI sebagai notasi data, dan dapat dihubungkan (linked) dengan data lain. KG peraturan perundang-undangan mengandung berbagai data terstruktur konten tekstual, struktur dokumen, seperti metadata, serta relasi antara peraturan seperti amendemen dan rujukan. Lex2KG memungkinkan pemanfaatan data peraturan perundang-undangan secara advanced, otomatis, dan dalam skala besar pada berbagai lingkup digital terutama pada industri hukum dan pengacara. Contoh pemanfaatan data dapat berupa search engine, sistem question answering, dan analisis statistik peraturan perundang-undangan. Menggunakan Lex2KG, penulis berhasil mengonversi 784 undang-undang menjadi KG dengan ukuran total lebih dari 1,1 juta triple.

Salah satu peraturan yang berhasil dikonversi adalah UU 11/2020 tentang Cipta Kerja yang kontennya bersifat relatif kompleks dan berukuran besar. Penulis juga menunjukkan use case dari KG peraturan perundang-undangan yaitu chat bot sederhana, SPARQL query, dan visualisasi peraturan perundang-undangan

.....Most of the legal documents are available as PDF which is not machine-readable, which means the data could not be processed automatically and in large scale by a computer to be utilized in various digital technology. Therefore, we need a data structure that can contain a legal information, and also a system which converts PDF into that structure. For that reason, in this research, author developed Lex2KG, a framework wh converting legal PDF documents in Indonesia (Lex comes from Latin which means law) into a Knowledge Graph. A knowledge graph (KG) is a graph that describes real-world entities and their relationships as machine-readable and structured information, and linkable to another KG on different domain. In this research KG is chosen from various data structure available because KG it categorized as 5-star data according to 5-star deployment scheme for Open Data, which data comes with most beneficial information, available under an open licen-

se, structured, open format, uses URI to denote things, and linkable to other data. The legal KG contains various kinds of structured data such as textual content, document structures, metadata, and relations between law such as amendments and citations. Lex2KG enables the advanced and automatic utilization of legal data on a large scale on a various digital scope especially on legal industry and lawyer. The utilization could be in form of search engine, question answering system, and statistics analytics for legals. Through Lex2KG, author have successfully converted 784 Indonesian laws into a KG with a total size of over 1.1 million triples. One of the regulation that was successfully converted was Law 11/2020 on Job Creation, which the content is relatively complex and large. Author also shows use cases of the legal KG for simple chatbots, SPARQL querying, and legal visualizations.