

BAB 1 PENDAHULUAN

Bab ini menguraikan latar belakang penelitian, rumusan masalah, tujuan penelitian, ruang lingkup penelitian, tahapan penelitian, dan sistematika penulisan.

1.1. Latar Belakang

Indonesia sebagai negara hukum mempunyai berbagai jenis peraturan perundang-undangan. Peraturan perundang-undangan tersebut memiliki hirarki berdasarkan kekuatan hukum yang dimilikinya. Hirarki tersebut menjaga agar setiap peraturan perundang-undangan tidak saling bertentangan, baik yang setingkat maupun dengan yang lebih tinggi tingkatannya. Berdasarkan Ketetapan (TAP) MPR Nomor III/MPR/2000, hirarki peraturan perundang-undangan di Indonesia dimulai dari UUD 1945, diikuti dengan TAP MPR, Undang-Undang (UU), Peraturan Pemerintah Pengganti Undang-Undang (Perpu), Peraturan Pemerintah (PP), Keputusan Presiden (Kepres), dan Peraturan Daerah (Perda). Urutan hirarki tersebut mengalami perubahan sesuai dengan pasal 7 ayat (1) Undang-Undang Nomor 10 Tahun 2004 sehingga menjadi UUD 1945 sebagai urutan pertama, diikuti dengan UU atau Perpu, PP, Peraturan Presiden, dan Perda.

Peraturan perundang-undangan (selanjutnya disebut dokumen legal) yang dimiliki oleh negara Indonesia telah mencapai jumlah yang besar. Jumlahnya yang besar dan format penulisannya yang menggunakan bahasa alami dan tidak mengikuti aturan tertentu membuat proses pencarian informasi di dalamnya menjadi sulit dilakukan. Kesulitan tersebut ditambah dengan belum adanya suatu sistem pengelolaan dokumen yang mampu melakukan standardisasi aturan penulisan dokumen legal tersebut.

Untuk dapat melakukan standardisasi aturan penulisan dokumen legal, dibutuhkan suatu standar penulisan dokumen yang dapat mempermudah proses pencarian informasi. *Extended Markup Language* (XML) adalah suatu standar penulisan dokumen yang dapat memenuhi tujuan tersebut. Hal itu disebabkan oleh kemampuan bahasa XML yang dapat memisahkan antara isi dan struktur dari suatu dokumen.

Proses standardisasi dokumen legal dari format bahasa alami ke dalam format yang sudah terstandardisasi menggunakan XML dapat dilakukan secara manual. Namun mengingat jumlah dari dokumen legal yang besar dan keterbatasan sumber daya, maka proses manual tersebut akan memakan banyak waktu, tenaga, dan biaya. Oleh karena itu, dibutuhkan suatu sistem yang dapat melakukan standardisasi dokumen legal secara otomatis [MAW06].

Standardisasi dokumen legal secara otomatis dapat dilakukan dengan menggunakan Sistem Ekstraksi Informasi (SEI). SEI adalah suatu sistem yang dapat menganalisa tulisan dan menyediakan informasi yang lebih rinci sesuai dengan kebutuhan pengguna [CUN99]. SEI melakukan standardisasi dokumen legal dengan cara mengenali tiap bagian atau subbagian dari suatu dokumen legal dan melakukan pemberian *tag* secara otomatis. Contohnya SEI dapat mengenali bagian judul dari suatu dokumen legal Peraturan Pemerintah dan memberikan *tag* <JUDUL> dengan mengenali frase “PERATURAN PEMERINTAH TENTANG...” pada bagian judul dokumen legal tersebut.

Penelitian SEI dengan *domain* dokumen berbahasa Indonesia telah beberapa kali dilakukan. Salah satunya ialah penelitian yang dilakukan oleh Budi dan Wibowo yang melakukan pengenalan entitas bernama (*named entity recognition*) dengan berbagai metode [BUD05, WIB05]. Selain itu Budi juga melakukan penelitian SEI lain di bidang *conference resolution* dengan menggunakan pendekatan *machine learning* [BUD06]. Penelitian tersebut menggunakan dokumen berita sebagai *domain* dari penelitian.

Penelitian dengan *domain* dokumen legal sendiri telah banyak dilakukan. Diantaranya yaitu:

- Sebuah sistem yang mampu memberikan ringkasan dari suatu keputusan hukum telah dihasilkan di Kanada. Sistem tersebut diberi nama *LetSum (Legal Text Summarizer)* [FAR04].
- Sebuah sistem yang mampu mengekstraksi referensi yang dimiliki oleh suatu dokumen legal secara otomatis telah dihasilkan di Spanyol [MAR05].
- Penelitian yang bertujuan untuk melakukan standardisasi dokumen legal telah dilakukan di Italia. Penelitian tersebut dilakukan dengan menandai tiap bagian dan referensi dalam suatu dokumen legal [PAL05].

- Sebuah sistem yang mampu melakukan fungsi *semantic annotation*, *effective indexing*, dan *legal document retrieval* telah dihasilkan pula di Italia. Sistem tersebut diberi nama SALEM (*Semantic Annotation for Legal Management*) dan bertujuan untuk mempermudah manajemen dokumen legal [SOR07].
- Sebuah *web portal* yang merupakan kumpulan penelitian dokumen legal menggunakan XML telah dihasilkan di Spanyol. *Web portal* tersebut dapat mengorganisasikan dan memanipulasi isi suatu dokumen legal. *Web portal* ini dapat digunakan baik oleh ahli hukum (*law professional*) maupun penduduk biasa yang tidak mempunyai pengetahuan tentang hukum (*non expert user*) [BLA07].

Penelitian SEI yang menggunakan *domain* dokumen legal Indonesia masih belum banyak dilakukan. Saat ini penulis hanya menemukan tiga penelitian yang menggunakan dokumen legal Indonesia sebagai *domain*-nya, yaitu penelitian yang dilakukan Mawadah [MAW06], Nuraminah [NUR07], dan Violina [VIO08]. Mawadah berusaha untuk melakukan standardisasi dokumen legal Indonesia dengan pendekatan *knowledge engineering*. Nuraminah melakukan penelitian dengan tujuan yang sama, yaitu melakukan standardisasi dokumen legal, tetapi dengan pendekatan *machine learning*. Violina berusaha menyempurnakan penelitian Mawadah sehingga sistem yang dihasilkan mampu menerima masukan UU Perubahan dan menampilkan keterkaitan antar dokumen legal dengan graf.

Mawadah memilih menggunakan pendekatan *knowledge engineering* karena belum tersedianya dokumen pelatihan untuk *domain* dokumen legal Indonesia [MAW06], walaupun dalam pendekatan pendekatan *knowledge engineering* terdapat permasalahan *adaptability*, yaitu pakar baru yang mengalami kesulitan membuat penyesuaian karena pakar yang membuat aturan sudah tidak ada [APP99]. Penelitian yang dilakukannya ialah penelitian yang pertama kali menggunakan *domain* dokumen legal Indonesia. Pendekatan *knowledge engineering* yang digunakannya melakukan ekstraksi informasi berdasarkan aturan yang dibuat oleh pakar (*expert*) yang mengerti *domain* yang diteliti. Mawadah menggunakan bantuan informasi kontekstual untuk membuat aturan yang dibutuhkan oleh SEI-nya. Informasi kontekstual adalah informasi yang sudah tersedia atau melekat pada dokumen. Informasi tersebut terdiri dari

informasi kata kunci, informasi kata hubung, informasi tanda baca, dan informasi kata khusus. SEI yang dikembangkannya berpusat pada dokumen legal jenis UU, tetapi belum mampu untuk menerima masukan UU Perubahan.

Permasalahan *adaptability* dalam pendekatan *knowledge engineering* dan sudah tersedianya dokumen pelatihan yang dibutuhkan membuat Nuraminah lebih memilih menggunakan pendekatan *machine learning* [NUR07]. Dalam pendekatan *machine learning*, sistem belajar dari suatu korpus dokumen pelatihan (*training*) yang relevan dengan dokumen asli untuk selanjutnya dibuat aturan-aturan (*rules*) secara otomatis. SEI yang dihasilkan Nuraminah mampu untuk melakukan standardisasi dokumen legal dengan baik, walaupun secara umum tingkat akurasi masih di bawah SEI buatan Mawadah.

Salah satu kelemahan SEI yang dikembangkan oleh Mawadah ialah tidak mampu untuk menerima masukan UU Perubahan. Hal tersebut dikarenakan struktur batang tubuh dari dokumen UU Perubahan yang sangat berbeda dengan dokumen UU pada umumnya. Akibatnya aturan-aturan yang diterapkan pada SEI tidak mampu untuk mengenali bagian batang tubuh UU Perubahan. Melihat kelemahan tersebut, Violina berusaha untuk memperbaiki SEI buatan Mawadah dengan penyempurnaan di beberapa bagian [VIO08]. Selain itu Violina juga menambahkan fitur yang dapat menghasilkan graf keterkaitan atau referensi antar dokumen legal. Graf tersebut berfungsi untuk mempermudah proses pengecekan kemungkinan terjadinya kontradiksi atau *overlapping* antar dokumen legal, yang mana hal tersebut tidak boleh terjadi [BUD07]. Fitur penghasil graf tersebut masih memiliki keterbatasan, yakni graf yang dihasilkan masih bersifat statis. Setiap graf yang dihasilkan hanya berupa gambar dan tidak dapat memberikan informasi lebih dari itu.

Walaupun sudah mampu untuk menangani dokumen UU Perubahan, SEI yang dikembangkan oleh Mawadah dan disempurnakan oleh Violina masih bermasalah dalam hal adaptifitasnya. Hal tersebut karena aturan yang digunakan oleh sistem masih bersatu dengan sistem itu sendiri, sehingga menyulitkan proses pengembangan dan adaptasi sistem. Selain itu, penelitian dokumen legal Indonesia yang dilakukan oleh Mawadah, Nuraminah, dan Violina hanya

menggunakan jenis dokumen legal UU. Belum ada penelitian yang menggunakan jenis dokumen legal Indonesia yang lainnya seperti Perpu, PP, atau Perda.

Berdasarkan latar belakang tersebut, penelitian ini bermaksud untuk mengembangkan SEI berdasarkan SEI yang dikembangkan oleh Mawadah dan disempurnakan oleh Violina. Pengembangan dilakukan dengan memodularisasi aturan sehingga diharapkan SEI akan lebih adaptif. Adaptivitas SEI akan diuji dengan jenis dokumen legal Indonesia yang lain, yaitu dokumen legal Peraturan Pemerintah (PP). Alasan pemilihan dokumen PP sebagai studi kasus baru dikarenakan kemudahan pencarian dan kelengkapan koleksi dari dokumen PP itu sendiri, selain itu juga karena PP memiliki keterkaitan referensi langsung dengan UU, sehingga dapat dengan mudah dilihat keterkaitannya.

Penelitian ini juga bermaksud untuk menambah informasi yang ditampilkan oleh graf referensi dan membuatnya lebih interaktif dengan pengguna. Hal tersebut dikarenakan graf referensi yang dihasilkan oleh sistem sebelumnya hanya berupa gambar statis dan hanya dapat menampilkan informasi mengenai keterkaitan referensi antar dokumen legal. Pengguna nantinya diharapkan dapat berinteraksi dengan graf tersebut, misalnya melakukan klik pada suatu node dalam graf, dan mendapatkan informasi yang lebih banyak melalui interaksi tersebut.

1.2. Rumusan Masalah dan Tujuan Penelitian

Aturan yang masih bersatu dengan sistem mengurangi adaptivitas dari sistem itu sendiri, maka dari itu masalah pertama yang timbul adalah bagaimana cara melakukan modularisasi aturan, sehingga adaptivitas sistem dapat meningkat. Setelah dilakukan modularisasi aturan, adaptivitas sistem akan diuji dengan melakukan adaptasi sistem ke jenis dokumen legal PP. Masalah kedua yang timbul adalah bagaimana membuat aturan-aturan yang dapat mengenali setiap objek dalam dokumen PP, sehingga sistem dapat diadaptasikan ke jenis dokumen legal PP.

1.3. Ruang Lingkup Penelitian

Penelitian ini memakai SEI yang sudah dikembangkan sebelumnya oleh Mawadah [MAW06] dan disempurnakan oleh Violina [VIO08]. SEI tersebut

dikembangkan dengan pendekatan *knowledge engineering* dan terdiri dari dua fitur utama, yaitu fitur standardisasi dokumen dan fitur pembuatan graf referensi. Penelitian ini berpusat pada dua kegiatan utama, yaitu modularisasi aturan yang digunakan oleh sistem dan membuat aturan-aturan baru sehingga sistem dapat diadaptasikan pada dokumen legal PP. Selain itu, sebagai kegiatan tambahan penelitian ini juga akan menambah informasi yang ditampilkan oleh graf referensi dan membuatnya lebih interaktif dengan pengguna.

Modularisasi aturan dilakukan pada tiap aturan yang digunakan oleh fitur standardisasi dokumen, yaitu aturan untuk mengenali tiap bagian dan sub bagian dari judul, pembukaan, batang tubuh, dan penutup. Pembuatan aturan-aturan baru juga dilakukan pada tiap aturan yang digunakan oleh fitur standardisasi dokumen. Aturan-aturan baru tersebut diharapkan dapat mengenali tiap bagian dan subbagian dari dokumen PP, yaitu judul, pembukaan, batang tubuh, dan penutup. Selain itu akan dilakukan juga penambahan aturan untuk mengenali keterkaitan antar-bagian dokumen PP serta keterkaitannya dengan dokumen lain.

Penyempurnaan fitur pembuatan graf referensi dilakukan dengan menambah informasi yang ditampilkan dan membuatnya lebih interaktif dengan pengguna. Graf yang dihasilkan akan disimpan ke dalam format SVG (*Scalable Vector Graphic*) yang berbasis kode. Penyempurnaan akan dilakukan pada kode SVG tersebut, sehingga perubahan tampilan yang diinginkan pada graf referensi dapat terjadi.

Penelitian ini akan diujicobakan pada sejumlah dokumen PP yang didapat dari berbagai sumber.

1.4. Tahapan Penelitian

Penelitian ini dilakukan dengan tahapan sebagai berikut:

- Melakukan studi pustaka untuk mempelajari sistem dokumen legal berbasis XML yang pernah dikembangkan sebelumnya.
- Merumuskan masalah penelitian.
- Melakukan analisa dan perancangan sistem.
- Mengimplementasikan hasil analisa dan perancangan sistem dengan bahasa pemrograman Java.

- Melakukan uji coba dan evaluasi sistem.
- Melakukan analisis terhadap hasil uji coba sistem.
- Merumuskan kesimpulan penelitian.

1.5. Sistematika Penulisan

Laporan penelitian tugas akhir ini mempunyai sistematika sebagai berikut:

BAB 1 PENDAHULUAN

Bab ini menguraikan latar belakang penelitian, rumusan masalah, tujuan penelitian, ruang lingkup penelitian, tahapan penelitian, dan sistematika penulisan.

BAB 2 LANDASAN TEORI

Bab ini menjelaskan teori-teori dan konsep yang dipakai dalam penelitian, yaitu konsep ekstraksi informasi, XML, DOT, gambaran umum Peraturan Pemerintah Republik Indonesia, pendekatan yang digunakan, dan cara evaluasi sistem.

BAB 3 ANALISIS DAN PERANCANGAN SISTEM

Bab ini menjelaskan tentang analisis dan perancangan sistem pengenalan struktur PP dan pembuatan graf referensi PP. Bab ini juga menjelaskan arsitektur sistem, setiap proses dalam sistem, dan rancangan evaluasi sistem.

BAB 4 IMPLEMENTASI SISTEM

Bab ini menjelaskan struktur implementasi sistem dan setiap kelas yang diimplementasikan.

BAB 5 UJI COBA DAN ANALISIS SISTEM

Bab ini menjelaskan lingkungan uji coba, karakteristik dokumen yang digunakan, hasil uji coba, dan analisisnya.

BAB 6 PENUTUP

Bab ini menjelaskan kesimpulan yang diperoleh dari penelitian dan saran yang dapat diberikan untuk pengembangan sistem lebih lanjut.