



BAB I

PENDAHULUAN

1.1 Latar Belakang

Pengelolaan dokumen di berbagai bidang organisasi telah mengalami transformasi secara signifikan. Perkembangan teknologi yang kian maju membuat proses pengelolaan dokumen ikut berubah dengan memanfaatkan teknologi-teknologi terbaru. Perubahan tersebut didorong oleh adanya kebutuhan akan efisiensi, kecepatan akses, dan keamanan data. Pengarsipan dokumen secara digital menjadi salah satu contoh dari penerapan teknologi terbaru untuk memaksimalkan proses pengelolaan dokumen. Sistem pengarsipan dokumen digital menawarkan solusi praktis untuk berbagai kendala yang dihadapi dalam proses pengarsipan dokumen secara konvensional, seperti risiko kerusakan fisik, kehilangan data, dan kesulitan dalam pencarian dokumen. Didorong dengan pesatnya kemajuan internet, pengolahan data pada arsip dokumen dalam sistem pengarsipan dokumen digital dapat bekerja lebih cepat dan menghasilkan lebih banyak informasi bermanfaat (Sunarni et al., 2020, p. 160).

RRI (Radio Republik Indonesia) adalah lembaga penyiaran publik milik pemerintah Indonesia yang menyelenggarakan kegiatan radio secara independen dan netral dengan memberikan pelayanan informasi, pendidikan, hiburan, kontrol sosial, dan perekat sosial dalam rangka melestarikan budaya bangsa untuk kepentingan seluruh lapisan masyarakat Indonesia (Maharani, 2021, p. 2). Dalam wilayah Kota Palembang, terdapat RRI Palembang yang menjadi lembaga penyiaran publik dengan menyelenggarakan kegiatan radio kepada masyarakat Kota Palembang secara khusus. Pada lingkungan internalnya, RRI Palembang memiliki banyak dokumen fisik dan digital yang tersimpan sebagai arsip, seperti surat tugas, surat kerja sama, laporan kegiatan, catatan absensi, memo internal, dan sebagainya. Namun, dokumen tersebut hanya disimpan tanpa adanya pengelolaan arsip yang terstruktur. Hal itu tentunya dapat menimbulkan banyak risiko pada masing-masing dokumen, terlebih RRI Palembang memiliki banyak dokumen



berharga. Oleh karena itu, diperlukan sistem pengelolaan arsip yang terorganisir agar setiap dokumen dapat tercatat dan tersimpan dengan aman. Sistem pengarsipan dokumen secara digital dapat diimplementasikan pada lingkungan internal RRI Palembang karena menawarkan keuntungan terkait keamanan, pengaksesan, dan integrasi antar-unit kerja. Dokumen fisik yang saat ini disimpan sebagai arsip dapat dijadikan dokumen digital melalui proses pemindaian (*scanning*).

Permasalahan utama yang dihadapi RRI Palembang terhadap arsip dokumen yang disimpannya adalah lamanya melakukan proses pencarian dokumen dan pengambilan informasi berdasarkan hasil pencarian tersebut karena belum memiliki sistem pengarsipan dokumen yang terorganisir. Melalui sistem pengarsipan dokumen digital, pencarian dokumen dapat dilakukan dengan mudah melalui identifikasi kode arsip atau nama dokumen. Proses pencarian tersebut dikenal sebagai pencarian berbasis kata kunci (*keywords*). Pencarian berbasis kata kunci adalah proses pencarian sederhana untuk mencari teks atau data terstruktur dengan memasukkan istilah relevan yang sesuai dengan target pencarian (Feddoul, 2020, p. 17). Kata kunci dalam pencarian berbasis kata kunci dapat berupa bagian-kata, kata, dan frasa atau klausa (gabungan kata-kata).

Hasil yang didapat melalui pencarian berbasis kata kunci hanya berupa dokumen yang diinginkan melalui kode arsip atau nama dokumen. Proses pencarian dapat dikembangkan lebih lanjut agar juga memungkinkan untuk mendapatkan dokumen yang diinginkan tanpa memasukkan kode arsip atau nama dokumen secara spesifik. Pencarian semantik dapat diimplementasikan pada sistem pengarsipan dokumen digital bersamaan dengan pencarian berbasis kata kunci untuk mewujudkan hal tersebut. Pencarian semantik adalah proses pengambilan informasi relevan dari teks berukuran panjang berdasarkan makna dan konteks pada kueri yang diberikan (Alqarni, 2024, p. 243). Sistem pengarsipan dokumen digital yang menerapkan pencarian semantik tidak mengharuskan pengguna untuk memasukkan kueri yang mengandung kata kunci sama dengan target pencarian karena proses pencarian dilakukan berdasarkan kesamaan makna dan konteks antara kueri dengan konten dokumen yang menjadi target pencarian.



Setelah mendapatkan solusi pencarian dokumen melalui pencarian semantik, RRI Palembang menyisakan satu permasalahan utama, yaitu pengambilan informasi dari dokumen secara cepat. Kumpulan dokumen yang menjadi hasil pencarian semantik selanjutnya dapat diolah kembali untuk dihasilkan ringkasan atau jawaban dari kueri yang diberikan sebagai informasi. *Retrieval-Augmented Generation* (RAG) adalah model bahasa yang mengombinasikan *prompt engineering* (teknik memberikan instruksi kepada model *machine learning*, umumnya berupa pertanyaan atau kueri) dengan proses pencarian dalam basis data untuk menghasilkan jawaban informatif, minim kesalahan, dan sesuai dengan data (Radeva et al., 2024, p. 1362). RAG merupakan sebuah metode pencarian lanjutan yang tidak hanya memberikan hasil pencarian berdasarkan kueri, tetapi juga memberikan informasi dalam bentuk kalimat atau paragraf sederhana berdasarkan pengetahuan yang didapat dari hasil pencarian. Sistem pengarsipan dokumen digital dapat menerapkan RAG dalam proses pencariannya untuk memberikan informasi secara cepat setelah didapatkan kumpulan dokumen dari hasil pencarian semantik. Informasi yang diberikan dapat digunakan untuk memahami isi seluruh dokumen yang ditemukan secara cepat sehingga tidak diperlukan membaca manual masing-masing dokumen untuk mendapatkan jawaban dari kueri yang diberikan.

Penulis mengembangkan sistem pengarsipan dokumen digital dalam lingkungan internal RRI Palembang dan mengimplementasikan metode RAG ke dalamnya untuk mengatasi permasalahan dalam pencarian arsip dokumen dan pengambilan informasi dari hasil pencarian tersebut. Sistem dirancang agar dapat digunakan di mana saja dengan perangkat apa pun sehingga disajikan dalam bentuk aplikasi web. Untuk mengorganisir proses pengembangan sistem, penulis menggunakan *V-Model* sebagai metode pengembangannya. *V-Model* adalah model pengembangan sistem yang berfokus pada prosedur pengujian dari awal hingga akhir tahap pengembangan dengan tahapan yang secara berurut dimulai dari analisis kebutuhan, desain sistem, desain arsitektur, desain modul, pengodean (*coding*), pengujian unit, pengujian integrasi, pengujian sistem, hingga pengujian



penerimaan (Abdel-Haq dan Asfoura, 2021, p. 1497). Tahap analisis kebutuhan, desain sistem, desain arsitektur, dan desain modul digolongkan sebagai tahap verifikasi. Sementara itu, tahap pengujian unit, pengujian integrasi, pengujian sistem, dan pengujian penerimaan digolongkan sebagai tahap validasi. Jadi, pengodean dilakukan setelah tahap verifikasi dan sebelum tahap validasi.

Berdasarkan uraian mengenai permasalahan yang dihadapi RRI Palembang beserta solusi untuk mengatasi permasalahannya, penulis melakukan penelitian tugas akhir dengan judul **“Implementasi Mesin Pencari Semantik dengan Metode *Retrieval-Augmented Generation* (RAG) dan *V-Model* pada Sistem Pengarsipan Dokumen Digital RRI Palembang”**.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, didapat rumusan permasalahan sebagai berikut.

1. Bagaimana membangun model *Retrieval-Augmented Generation* (RAG) agar dapat melakukan pencarian semantik terhadap arsip dokumen digital RRI Palembang dan memberikan informasi dokumen berdasarkan pengetahuan yang didapat dari hasil pencarian semantik secara cepat?
2. Bagaimana mengimplementasikan metode pengembangan *V-Model* pada sistem pengarsipan dokumen digital RRI Palembang berbasis aplikasi web agar sistem dapat dikembangkan secara terstruktur, terdokumentasi, dan teruji?
3. Bagaimana menghubungkan model RAG yang telah dibangun dengan sistem pengarsipan dokumen digital RRI Palembang berbasis aplikasi web agar sistem dapat melakukan pencarian semantik dengan RAG berdasarkan dokumen yang tersedia dalam basis data sistem tersebut?

1.3 Batasan Masalah

Penelitian tugas akhir ini memiliki beberapa batasan masalah agar pembahasan tidak menyimpang dan melebar dari permasalahan yang ada. Berikut adalah batasan masalah pada penelitian ini.



1. Sistem dibangun sebagai aplikasi web.
2. Sistem hanya diperuntukkan lingkungan internal RRI Palembang.
3. Sistem akan dikelola oleh LPP RRI Palembang Bagian Tata Usaha.
4. Data yang digunakan untuk mengembangkan model RAG adalah dokumen fisik dan digital RRI Palembang.
5. Dokumen yang diarsipkan berupa nota dinas, laporan kegiatan, laporan kehadiran, surat tugas, surat keputusan, undangan, pengumuman, dan lainnya dari seluruh departemen di LPP RRI Palembang.
6. Arsip dibagi menjadi empat jenis berdasarkan kriteria retensinya: arsip vital yang akan disimpan abadi, arsip penting yang akan disimpan selama 25 tahun, arsip berguna yang akan disimpan selama 10 tahun, dan arsip sementara yang akan disimpan selama 2 tahun.
7. Teks yang diolah pada dokumen hanya konten utama dari dokumen, mengecualikan bagian *header* dan *footer*.
8. Model RAG yang dibangun hanya bekerja untuk dokumen yang kontennya sebagian besar berupa teks berbahasa Indonesia.
9. Pengujian dalam *V-Model* hanya berfokus pada alur pencarian dengan RAG.

1.4 Tujuan dan Manfaat

1.4.1 Tujuan

Tujuan akhir yang hendak dicapai dalam penelitian tugas akhir ini adalah sebagai berikut.

1. Menghasilkan model *Retrieval-Augmented Generation* (RAG) yang dapat melakukan pencarian semantik terhadap arsip dokumen digital RRI Palembang dan memberikan informasi dokumen berdasarkan pengetahuan yang didapat dari hasil pencarian semantik.
 2. Membangun sistem pengarsipan dokumen digital RRI Palembang berbasis aplikasi web dengan metode pengembangan *V-Model*.
 3. Mengintegrasikan model RAG yang telah dibangun dengan sistem pengarsipan dokumen digital RRI Palembang berbasis aplikasi web.
-



1.4.2 Manfaat

Manfaat yang akan didapat dari tercapainya tujuan akhir penelitian tugas akhir ini adalah sebagai berikut.

1. Memudahkan pihak LPP RRI Palembang Bagian Tata Usaha dalam mengelola arsip dokumen, termasuk proses penyimpanan, pendistribusian, pemanfaatan, pemindahan, dan pemusnahan.
2. Menjadikan arsip dokumen RRI Palembang yang disimpan dapat bertahan dalam waktu lama tanpa kerusakan karena berbentuk digital dan tersimpan pada server yang terjaga.
3. Memudahkan seluruh pihak RRI Palembang yang menjadi pengguna sistem dalam menyimpan arsip dokumennya secara terpusat sehingga dapat dipantau oleh pihak internal lain yang berwenang.
4. Mempercepat proses pencarian arsip dokumen beserta informasi tambahan berdasarkan kueri yang diberikan pengguna.

1.5 Metodologi Penelitian

1.5.1 Lokasi dan Waktu Penelitian

Penulis melaksanakan penelitian tugas akhir di LPP RRI Palembang yang beralamatkan Jl. Radio No. 2 RT 27 RW 11 KM 4 Kelurahan 20 Ilir D-IV Kecamatan Ilir Timur I Kota Palembang 30128. Penelitian dilakukan dari tanggal 29 April 2024 hingga 30 Juni 2024.

1.5.2 Metode Pengumpulan Data

Penulis mengumpulkan data yang diperlukan dalam penelitian melalui metode pengumpulan data berikut ini.

1. Wawancara

Wawancara adalah teknik pengumpulan data secara tatap muka melalui proses tanya jawab antara pengumpul data dengan narasumber (Trivaika dan Senubekti, 2022, p. 34). Wawancara dilakukan untuk mendapatkan informasi dengan cara bertanya secara langsung dengan responden. Penulis melakukan



wawancara dengan Bapak Oktav Zulkarnain selaku Staf Subbag SDM Bagian Tata Usaha LPP RRI Palembang. Wawancara dilaksanakan secara langsung di kantor LPP RRI Palembang Bagian Tata Usaha dalam dua kali pertemuan langsung. Penulis mempersiapkan pertanyaan wawancara yang terstruktur untuk kemudian ditanyakan agar mendapatkan informasi mengenai permasalahan yang ada dan kebutuhan sistem.

2. Observasi

Observasi adalah teknik pengumpulan data melalui pengamatan secara langsung terhadap partisipan dan lingkungannya dengan tujuan tertentu untuk mengungkap dan memprediksi landasan munculnya perilaku tertentu (Adhandayani, 2020, p. 3). Observasi dilakukan secara langsung di LPP RRI Palembang Bagian Tata Usaha dengan mengamati lingkungan kerja. Penulis melakukan pengamatan bagaimana organisasi dalam LPP RRI Palembang bekerja dan jenis dokumen yang sering dikelola oleh staf bagian tata usaha untuk kemudian dijadikan arsip. Siklus hidup arsip juga menjadi fokus observasi penulis untuk memahami bagaimana arsip bekerja, mulai dari proses penciptaan dokumen, pendistribusian dokumen, pengarsipan dokumen, pemanfaatan arsip, perpindahan arsip antar-unit kerja, dan pemusnahan arsip.

3. Studi Literatur

Studi literatur adalah teknik pengumpulan data yang menelaah sumber-sumber tertulis terpercaya yang berkaitan dengan penelitian, seperti jurnal ilmiah, buku referensi, ensiklopedia, dan lainnya (Sabrina, 2021, p. 19). Studi literatur dilakukan dengan mempelajari sumber-sumber pustaka dan penelitian-penelitian sebelumnya mengenai pencarian semantik, implementasi *Retrieval-Augmented Generation* (RAG), dan implementasi *V-Model*. Penulis mendapatkan data, informasi, dan konsep yang bersifat teoritis dari buku, artikel dalam jurnal, dan referensi internet resmi berdasarkan permasalahan yang ada. Penulis juga melakukan studi pustaka pada dokumentasi *framework* dan *library* yang digunakan untuk mengembangkan sistem.



1.6 Sistematika Penulisan

Agar pembahasan tugas akhir ini dapat memberikan gambaran secara jelas dan terstruktur sesuai dengan tujuan akhir yang telah diuraikan, penulisan tugas akhir ini disusun dengan sistematika penulisan seperti berikut.

BAB I PENDAHULUAN

Pada bab ini, penulis menguraikan bagian awal tugas akhir yang terdiri dari latar belakang, rumusan masalah, batasan masalah, tujuan dan manfaat, metodologi penelitian, serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Pada bab ini, penulis memberikan penjelasan mengenai teori-teori yang digunakan dalam penelitian, di antaranya adalah teori umum, teori judul, teori khusus, dan teori program. Penulis juga mencantumkan referensi penelitian sebelumnya yang berkaitan dengan penelitian tugas akhir ini.

BAB III METODOLOGI PENELITIAN

Pada bab ini, penulis memberikan gambaran dan informasi umum mengenai instansi tempat penelitian, tahapan penelitian, metode pengolahan data, dan metode pengembangan sistem.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini, penulis menjelaskan proses pengembangan sistem mulai dari tahap analisis kebutuhan, desain sistem, desain arsitektur, desain modul, pengodean, pengujian unit, pengujian integrasi, pengujian sistem, hingga pengujian penerimaan.

BAB V KESIMPULAN DAN SARAN

Pada bab ini, penulis memaparkan kesimpulan yang didapat dari hasil penelitian untuk menyelesaikan permasalahan yang ada dan saran yang ditujukan kepada pihak-pihak terkait serta penelitian selanjutnya.