

Implementasi Algoritma BM25 Berbasis Auto-Tagging YAKE! untuk Pengembangan Sistem Pencarian Cerdas pada Website “Vokasi Product Hub” Fakultas Vokasi UNESA

Mohammad Avie Siena¹, Dodik Arwin Dermawan²

D4 Manajemen Informatika, Fakultas Vokasi, Universitas Negeri Surabaya, Surabaya, Indonesia

¹mohammadavie.22103@mhs.unesa.ac.id

²dodikdermawan@unesa.ac.id

Abstrak - Penelitian ini mengembangkan sistem pencarian cerdas pada platform Vokasi Product Hub sebagai repositori karya mahasiswa Fakultas Vokasi UNESA. Sistem mengintegrasikan algoritma YAKE! untuk auto-tagging metadata secara otomatis dan BM25 sebagai algoritma pemeringkatan dokumen berbasis relevansi. Pengembangan menggunakan model Waterfall dengan evaluasi fungsional melalui Black-Box Testing dan evaluasi relevansi menggunakan metrik Information Retrieval (IR) yaitu Precision@k, Mean Average Precision (MAP), dan Normalized Discounted Cumulative Gain (nDCG). Seluruh 20 skenario Black-Box Testing dinyatakan lulus. Evaluasi pada 120 karya simulasi dengan 30 kueri uji menghasilkan Precision@5 = 0,3533; Precision@10 = 0,2333; MAP = 0,5732; nDCG@5 = 0,6539; dan nDCG@10 = 0,7093. Nilai nDCG@10 yang relatif tinggi menunjukkan sistem mampu memeringkat dokumen relevan mendekati urutan ideal pada rentang sepuluh hasil teratas. Hasil tersebut menunjukkan bahwa integrasi auto-tagging YAKE! dan pemeringkatan BM25 layak digunakan sebagai fondasi awal sistem pencarian cerdas pada repositori karya akademik vokasi.

Kata kunci: BM25, YAKE!, Information Retrieval, Auto-Tagging, Vokasi Product Hub, Sistem Pencarian Cerdas

Abstract - This study develops an intelligent search system for the Vokasi Product Hub platform as a repository of student works at the Faculty of Vocational Studies, UNESA. The system integrates the YAKE! Algorithm for automatic metadata auto-tagging and BM25 as a relevance-based document ranking algorithm. The development process adopts the Waterfall model, with functional evaluation conducted through Black-Box Testing and relevance evaluation performed using Information Retrieval (IR) metrics, namely Precision@k, Mean Average Precision (MAP), and Normalized Discounted Cumulative Gain (nDCG). All 20 Black-Box Testing scenarios were successfully passed. Evaluation on 120 simulated documents with 30 test queries produced Precision@5 of 0.3533; Precision@10 of 0.2333; MAP of 0.5732; nDCG@5 of 0.6539; and nDCG@10 of 0.7093. The relatively high nDCG@10 value indicates the system is capable of ranking relevant documents close to the ideal ordering within the top-ten search results. These results indicate that integrating YAKE! auto-tagging and BM25 ranking is feasible

as an initial foundation for an intelligent search system in vocational academic-work repositories.

Keywords: BM25, YAKE!, Information Retrieval, Auto-Tagging, Vokasi Product Hub, Intelligent Search System

I. PENDAHULUAN

Fakultas Vokasi Universitas Negeri Surabaya menghasilkan beragam karya mahasiswa setiap tahunnya, mulai dari aplikasi, prototipe, desain produk, video dokumentasi, hingga laporan penelitian terapan. Karya-karya tersebut tersebar pada berbagai media penyimpanan pribadi seperti Google Drive, perangkat lokal, atau dokumentasi terpisah masing-masing program studi, sehingga sulit ditemukan kembali secara terpusat ketika dibutuhkan sebagai referensi oleh mahasiswa lain, dosen, atau mitra industri [1].

Repositori digital berfungsi sebagai sarana penyimpanan, publikasi, dan pelestarian karya ilmiah yang mendukung manajemen informasi akademik [2]. Dalam konteks Pendidikan vokasi, kebutuhan terhadap repositori semakin penting karena karya mahasiswa tidak hanya berupa laporan tertulis, tetapi juga mencakup artefak praktis seperti produk, prototipe, dan dokumentasi visual. Tinjauan literatur menunjukkan bahwa banyak repositori perguruan tinggi masih menghadapi keterbatasan pada aspek pencarian, khususnya dalam menyajikan hasil yang relevan dan terurut secara kontekstual [3].

Dalam bidang Information Retrieval (IR), algoritma BM25 (Best Matching 25) merupakan algoritma probabilistik yang mempertimbangkan frekuensi kemunculan istilah dan panjang dokumen [4]. BM25 masih digunakan sebagai *baseline* utama dalam berbagai sistem pencarian karena stabil, efisien, dan mampu meningkatkan kualitas relevansi hasil pencarian [5], [25], [30]. Selain pemeringkatan dokumen, kualitas metadata memiliki pengaruh signifikan terhadap performa pencarian. YAKE! (Yet Another Keyword Extractor) merupakan metode ekstraksi kata kunci *unsupervised*

yang memanfaatkan karakteristik lokal teks tanpa memerlukan data pelatihan [6]. Integrasi antara *auto-tagging* YAKE! dan algoritma pemeringkatan BM25 menawarkan pendekatan yang komprehensif dalam membangun sistem pencarian cerdas pada repositori karya mahasiswa.

Pada penelitian [6] mengkaji integrasi metode *keyword extraction* seperti YAKE! dan menunjukkan bahwa YAKE! secara konsisten mengungguli metode ekstraksi kata kunci *unsupervised* lainnya, seperti TextRank dan RAKE, tanpa memerlukan data pelatihan. Serta, penelitian [4] menegaskan bahwa algoritma BM25 masih menjadi *baseline* yang kuat dalam sistem pemeringkatan teks modern karena kemampuannya menilai relevansi berdasarkan frekuensi istilah dan panjang dokumen. Studi komparatif terhadap berbagai alat ekstraksi kata kunci tanpa pengawasan juga menempatkan YAKE! sebagai metode yang ringan dan kompetitif [13]. Selain sebagai pemeringkat tunggal, BM25 umum diposisikan sebagai retriever tahap pertama yang dapat dipadukan dengan pseudo relevance feedback maupun representasi semantik untuk meningkatkan kualitas pemeringkatan [14], [15].

Selanjutnya, studi oleh [7] meneliti peran konversi dan standarisasi metadata pada *findability* repositori universitas, dan hasil penelitian menunjukkan perbaikan skema metadata dan konsistensi kata kunci mampu meningkatkan kualitas temu balik informasi, terlihat dari peningkatan nilai nDCG. Penelitian oleh [8] menerapkan algoritma BM25 pada sistem pencarian lowongan pekerjaan dan melaporkan tingkat kepuasan pengguna hingga 91,2%. Lalu, penelitian oleh [5] memperluas kajian BM25 dengan pendekatan semantik berbasis *transformer*, namun menyimpulkan bahwa BM25 tetap unggul dari sisi kecepatan dan stabilitas, terutama pada *dataset* berukuran kecil hingga menengah.

Berdasarkan tinjauan tersebut, integrasi auto-tagging YAKE! dan pemeringkatan BM25 pada satu sistem repositori karya mahasiswa vokasi masih jarang dibahas dalam literatur yang tersedia. Kesenjangan inilah yang menjadi dasar penelitian ini. Penelitian ini bertujuan untuk (1) merancang dan membangun website repositori Vokasi Product Hub sebagai wadah publikasi karya mahasiswa Fakultas Vokasi UNESA; (2) mengimplementasikan algoritma YAKE! untuk auto-tagging dan BM25 untuk pemeringkatan hasil pencarian melalui integrasi berbasis REST API; (3) menguji fungsionalitas sistem menggunakan metode Black-Box Testing; serta (4) mengukur performa temu balik informasi menggunakan metrik Precision@k, Mean Average Precision (MAP), dan Normalized Discounted Cumulative Gain (nDCG).

No	Penelitian	Hasil Utama	Perbedaan dengan Penelitian Ini
1	Campos dkk. [6]	YAKE! mengekstrak kata kunci secara efisien tanpa data latih dan mengungguli TextRank serta	Menggabungkan YAKE! dengan BM25 pada repositori karya mahasiswa

		RAKE	
2	Lin dkk. [4]	BM25 tetap menjadi baseline yang kuat pada sistem pemeringkatan teks modern	Menerapkan BM25 pada konteks repositori karya vokasi
3	Nosé dkk. [7]	Konversi dan standarisasi metadata meningkatkan nilai nDCG temu balik	Menambahkan auto-tagging YAKE! untuk memperkaya metadata
4	Zibani dkk. [1]	Repositori fakultas efektif, tetapi aspek pencarian masih terbatas	Menawarkan sistem pencarian cerdas berbasis BM25 dan YAKE!
5	Kheng dkk. [8]	Penerapan BM25 meningkatkan kepuasan pengguna hingga 91,2%	Menerapkan BM25 pada domain repositori akademik vokasi
6	Hambarde & Proença [5]	BM25 unggul pada dataset kecil-menengah, cepat dan akurat	Memanfaatkan BM25 untuk sistem repositori karya vokasi

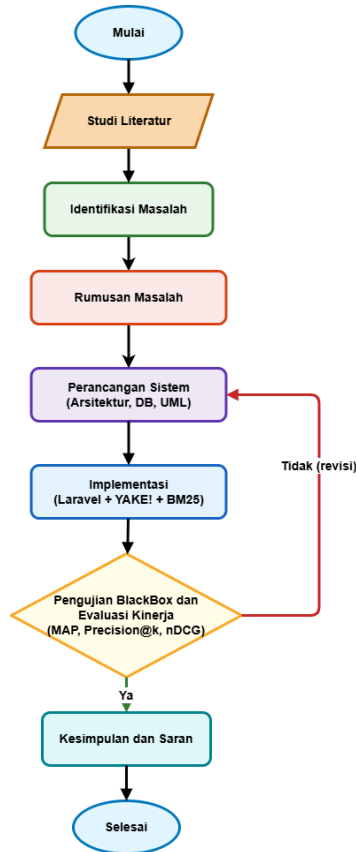
TABEL I
PERBANDINGAN DENGAN PENELITIAN TERDAHULU

Untuk memperjelas posisi penelitian, Tabel I merangkum sejumlah penelitian terdahulu yang relevan beserta perbedaannya dengan penelitian ini.

II. METODE PENELITIAN

A. Jenis dan Model Pengembangan

Penelitian ini merupakan penelitian pengembangan sistem (*development research*) dengan pendekatan kuantitatif evaluatif menggunakan model Waterfall [10], [27]. Lima tahap Waterfall yang diterapkan adalah analisis kebutuhan, perancangan sistem, implementasi sistem, pengujian sistem, dan evaluasi hasil pencarian.



Gambar 1. Alur Penelitian

B. Algoritma BM25

BM25 menghitung skor relevansi antara sebuah kueri q dan dokumen D berdasarkan kontribusi setiap istilah t yang muncul pada kueri tersebut. Secara matematis, skor BM25 dirumuskan sebagai [4]:

$$BM25(q, D) = \sum_{t \in q} \left[IDF(t) \cdot \frac{TF(t, D) \cdot (k_1 + 1)}{TF(t, D) + k_1 \cdot \left(1 - b + b \cdot \frac{|D|}{avgD}\right)} \right]$$

Di mana $TF(t, D)$ adalah frekuensi kemunculan kata t dalam dokumen D ; $|D|$ adalah panjang dokumen; $avgD$ adalah rata-rata panjang dokumen dalam koleksi; k_1 adalah parameter penghalus (biasanya 1,2-2,0); dan b adalah parameter normalisasi panjang dokumen (biasanya 0,75). Dalam penelitian ini, parameter yang digunakan adalah $k_1 = 1,2$ dan $b = 0,75$.

C. Algoritma YAKE! dan Penerapan Auto-Tagging

YAKE! mengevaluasi setiap kandidat kata menggunakan lima fitur utama; antara lain $TCase$ (sensitivitas huruf kapital), $TPos$ (posisi kemunculan kata), $TFNorm$ (frekuensi ternormalisasi), $TRel$ (ko-kemunculan kata), dan $TSent$ (penyebaran kata pada kalimat). Bobot kata dihitung menggunakan persamaan [6]:

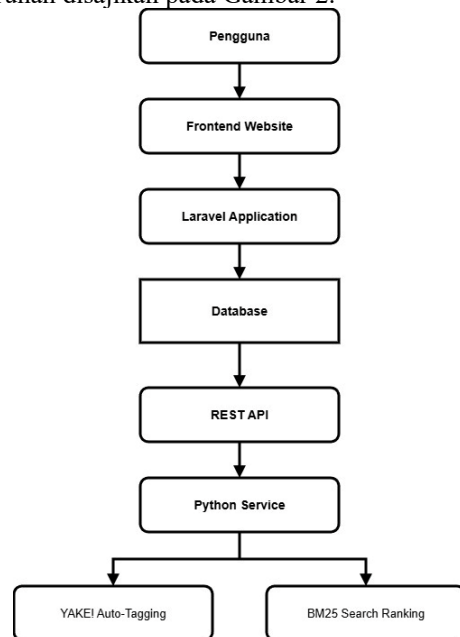
$$S(t) = \frac{TPos(t) \cdot TRel(t)}{TCase(t) + \frac{TFNorm(t) + TSent(t)}{TRel(t)}}$$

Nilai $S(t)$ yang lebih rendah menunjukkan bahwa kata tersebut memiliki tingkat kepentingan yang lebih tinggi. YAKE! efektif digunakan pada teks pendek hingga menengah, termasuk ringkasan, abstrak, maupun deskripsi singkat yang umum ditemukan pada sistem repositori akademik. Kualitas ekstraksi kata kunci unsupervised juga dipengaruhi oleh ketersediaan dataset berkualitas dan konteks kata [18], [23], dan YAKE! telah dimanfaatkan pada beragam tugas seperti ekspansi kueri [21] serta deteksi berita palsu [22].

YAKE! diterapkan khusus untuk mengekstraksi kata kunci dari teks deskripsi karya yang diinput pengguna saat proses unggah ke sistem repositori. Sistem tidak mengekstraksi dari isi *file* PDF, DOCX, video, atau arsip. Laravel mengirimkan teks deskripsi karya ke layanan Python melalui REST API, kemudian layanan Python menjalankan proses ekstraksi YAKE! dan mengembalikan hasil dalam format JSON. *Keyword* hasil ekstraksi disimpan sebagai metadata tambahan dan diikutsertakan dalam dokumen indeks pencarian BM25.

D. Arsitektur Sistem

Sistem dikembangkan menggunakan *framework* Laravel 12 (PHP) [26] sebagai backend dengan antarmuka *frontend* berbasis AJAX (JavaScript). Layanan Python digunakan sebagai pemroses algoritma YAKE! serta BM25, yang terintegrasi dengan Laravel melalui REST API [11], [29]. MySQL digunakan sebagai sistem basis data untuk menyimpan metadata karya, hasil *auto-tagging*, dan indeks pencarian [20]. Arsitektur sistem secara keseluruhan disajikan pada Gambar 2.



Gambar 2. Diagram Arsitektur Sistem Vokasi Product Hub

Metadata yang diproses dalam sistem pencarian meliputi judul karya, deskripsi karya, kategori, program studi, *tag* manual, dan *keyword* hasil YAKE!. Pendekatan ini dipilih agar BM25 dapat mempertimbangkan informasi yang lebih kaya dibanding hanya mengandalkan pencocokan kata pada deskripsi.

E. Metrik Evaluasi Information Retrieval (IR)

Evaluasi kinerja sistem IR dilakukan menggunakan tiga metrik standar [9], [16], [28], antara lain:

1) *Precision@k*: Mengukur persentase dokumen relevan pada posisi hasil pencarian 1 hingga *k*.

2) *Mean Average Precision (MAP)*: Menghitung rata-rata nilai *precision* pada setiap posisi relevan dalam daftar pencarian.

3) *Normalized Discounted Cumulative Gain (nDCG)*: Menilai kualitas urutan dokumen berdasarkan tingkat relevansi bertingkat, dinormalisasi terhadap urutan ideal.

F. Dataset dan Ground Truth

Dataset evaluasi terdiri atas 120 karya simulasi yang disusun berdasarkan karakteristik karya mahasiswa Fakultas Vokasi UNESA dari sepuluh program studi. Data simulasi digunakan karena belum tersedianya kumpulan karya mahasiswa yang terdokumentasi terpusat. *Ground truth* disusun menggunakan 30 kueri uji dengan 5 dokumen *judgment* per kueri (total 150 *relevance judgments*) menggunakan skala relevansi 0-3 (0 = tidak relevan, 1 = sedikit relevan, 2 = relevan, 3 = sangat relevan). Untuk evaluasi *Precision@k* dan MAP digunakan konversi ke relevansi biner (skor ≥ 2 dianggap relevan). Penyusunan *relevance judgment* dan koleksi uji mengacu pada praktik evaluasi temu balik informasi [17], [19].

Sebaran 120 karya simulasi tersebut terbagi secara merata pada sepuluh program studi di lingkungan Fakultas Vokasi UNESA, dengan setiap program studi diwakili oleh 12 karya. Komposisi tersebut disajikan pada Tabel II.

TABEL II
SEBARAN DATASET SIMULASI PER PROGRAM STUDI

No	Program Studi	Jumlah Dataset Simulasi
1	D4 Administrasi Negara	12
2	D4 Desain Grafis	12
3	D4 Pelatihan Olahraga	12
4	D4 Manajemen Informatika	12
5	D4 Tata Boga	12
6	D4 Tata Busana	12
7	D4 Teknik Listrik	12
8	D4 Teknik Mesin	12
9	D4 Teknik Sipil	12
10	D4 Transportasi	12
	Total	120

Setiap pasangan kueri–dokumen dinilai menggunakan skala relevansi bertingkat 0–3 sebagaimana dirinci pada Tabel III.

TABEL III
SKALA PENILAIAN RELEVANSI PADA GROUND TRUTH

Skor	Kategori	Keterangan
0	Tidak relevan	Dokumen tidak berkaitan dengan kueri
1	Cukup relevan / marginal	Dokumen memiliki keterkaitan lemah terhadap kueri
2	Relevan	Dokumen secara substansial sesuai dengan maksud kueri
3	Sangat relevan	Dokumen sangat sesuai dan menjawab maksud kueri secara langsung

Contoh kueri uji beserta tipe dan dokumen relevannya disajikan pada Tabel IV.

TABEL IV
CONTOH KUERI UJI DAN GROUND TRUTH

No	Query Uji	Tipe Query	Contoh Dokumen Relevan
1	Administrasi digital	Administrasi	Digitalisasi Layanan Administrasi; Sistem Administrasi Terpadu
2	Alat pendeteksi arus otomatis	Kompleks	Prototype Sensor Arus; Alat Ukur Tegangan Digital
3	Analisis kekuatan beton	Kompleks	Analisis Kekuatan Beton; Laporan Uji Material Konstruksi
4	Aplikasi absensi mahasiswa	Aplikasi	Aplikasi Absensi Mahasiswa dan karya pendukungnya
5	Pengelolaan dokumen digital	Dokumen	Aplikasi Manajemen Dokumen; Tata Kelola Arsip Digital

G. Pengujian Black-Box Testing

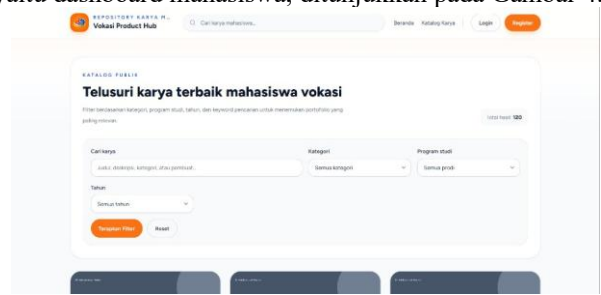
Black-Box Testing dilakukan sesuai standar ISO/IEC/IEEE 29119-1 untuk memverifikasi fungsi sistem [12]. Fitur yang diuji meliputi akses katalog, unggah karya, *auto-tagging* YAKE!, validasi karya, pengelolaan *role*, *reindexing*, *health check* layanan Python, dan *search analytics*.

III. HASIL DAN PEMBAHASAN

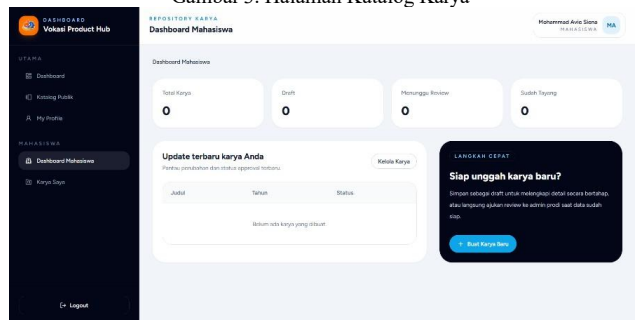
A. Implementasi Website Vokasi Product Hub

Website Vokasi Product Hub telah berhasil diimplementasikan menggunakan *framework* Laravel 12 sebagai repositori karya mahasiswa Fakultas Vokasi UNESA. Sistem mendukung alur pengelolaan karya yang lengkap, mulai dari unggah karya, pengisian metadata, validasi oleh pengelola, hingga publikasi pada katalog. Sistem menerapkan lima *role* pengguna, yaitu Guest, Mahasiswa, Operator Prodi, Admin Fakultas, dan Super Admin, beserta pembatasan hak akses sesuai fungsi masing-masing. Tampilan halaman katalog karya yang

dapat diakses oleh seluruh pengguna ditunjukkan pada Gambar 3. Salah satu contoh antarmuka berbasis peran, yaitu dashboard mahasiswa, ditunjukkan pada Gambar 4.



Gambar 3. Halaman Katalog Karya



Gambar 4. Dashboard Mahasiswa

B. Dukungan Format File Berdasarkan Analisis Kebutuhan

Hasil kuesioner analisis kebutuhan pengguna menjadi dasar penentuan dukungan format berkas pada fitur unggah karya. Setiap program studi menghasilkan jenis karya dan format file yang berbeda, sehingga sistem dirancang untuk mengakomodasi beragam format mulai dari dokumen, gambar, video, arsip, file desain dan teknis, hingga tautan eksternal. Ringkasan jenis karya dan format file yang dihasilkan tiap program studi disajikan pada Tabel V.

TABEL V
JENIS KARYA DAN FORMAT FILE PER PROGRAM STUDI

Program Studi	Jenis Karya	File yang Dihasilkan
D4 Administrasi Negara	Laporan penelitian, website/web app, video dokumentasi, desain grafis, modul pembelajaran	PDF/DOC, PPT, JPG/PNG, MP4
D4 Desain Grafis	Desain grafis (poster/branding), animasi/motion graphic, video	JPG/PNG, PDF/DOC, MP4, PPT, file desain (PSD/AI/CDR), audio
D4 Kepelatihan Olahraga	Program latihan olahraga, video dokumentasi	MP4, PPT
D4 Manajemen Informatika	Aplikasi mobile, website/web app, sistem informasi,	Source code (ZIP/GitHub), link eksternal

	rancangan sistem (UML)	
D4 Tata Boga	Produk fisik (makanan), video, desain grafis, laporan, analisis data	JPG/PNG, MP4, PDF/DOC, PPT, Excel
D4 Tata Busana	Produk fisik/desain busana, hasil jadi busana, video, desain grafis	JPG/PNG, PDF/DOC, file desain (PSD/AI/CDR), MP4
D4 Teknik Listrik	Prototype alat, IoT/embedded, gambar teknik, simulasi teknik, laporan, video	JPG/PNG, MP4, PDF/DOC, source code, CAD, file simulasi, CSV/JSON
D4 Teknik Mesin	Prototype alat, model 3D, gambar teknik/blueprint	JPG/PNG, PDF/DOC, PPT, CAD (AutoCAD/Sketch Up)
D4 Teknik Sipil	Prototype alat, gambar teknik/blueprint, laporan, video	JPG/PNG, MP4, PDF/DOC, CAD
D4 Transportasi	Gambar teknik/blueprint, prototype alat, model 3D, desain grafis, laporan	PDF/DOC, CAD, Excel, JPG/PNG, PPT

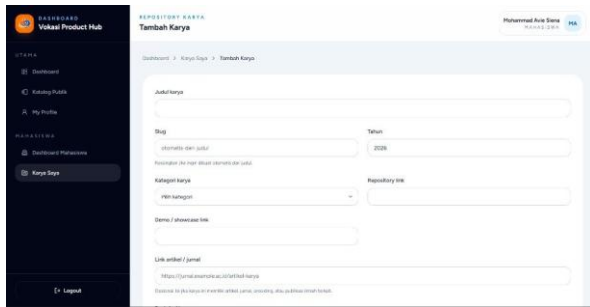
Berdasarkan keberagaman tersebut, fitur unggah karya dirancang mendukung banyak format file sehingga setiap program studi dapat menyimpan karyanya sesuai jenis dan format yang dihasilkan, sekaligus menjadi dasar penyusunan dataset simulasi pada penelitian ini.

C. Implementasi Auto-Tagging YAKE!

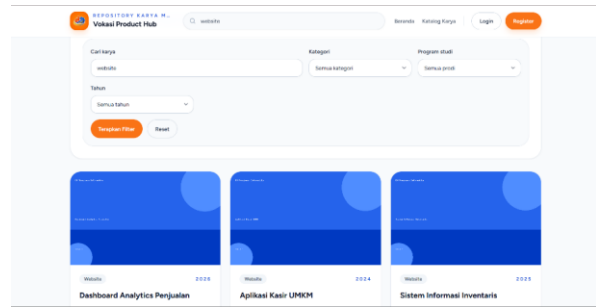
Algoritma YAKE! berhasil diimplementasikan sebagai fitur *auto-tagging* melalui layanan Python yang terintegrasi dengan Laravel menggunakan REST API. Proses ekstraksi kata kunci dipicu secara otomatis ketika mahasiswa mengisi kolom deskripsi pada halaman unggah karya, seperti ditunjukkan pada Gambar 5. Contoh hasil *auto-tagging* ditunjukkan pada Tabel VI.

TABEL VI
CONTOH HASIL AUTO-TAGGING YAKE!

No	Judul Karya	Keyword YAKE!
1	E-Commerce Produk Mahasiswa	e-commerce, e-commerce produk mahasiswa, produk mahasiswa, manajemen informatika, kategori website
2	REST API Layanan Publik	REST API, REST API layanan publik, layanan publik, API layanan, manajemen informatika

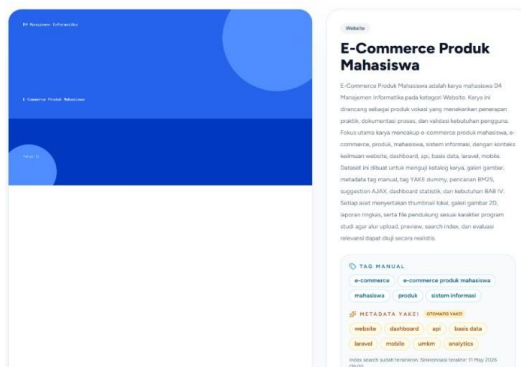


Gambar 5. Halaman Unggah Karya



Gambar 7. Pencarian Karya oleh Pengunjung

Hasil implementasi menunjukkan bahwa YAKE! dapat menghasilkan kata kunci yang representatif terhadap topik karya ketika deskripsi diisi dengan baik, sebagaimana terlihat pada metadata keyword yang ditampilkan di halaman detail karya pada Gambar 6. Keterbatasan yang teridentifikasi adalah kualitas keyword bergantung pada kelengkapan dan kejelasan deskripsi yang diinputkan mahasiswa. Deskripsi yang terlalu singkat atau kurang spesifik cenderung menghasilkan *keyword* yang kurang relevan terhadap isi karya. Oleh karena itu, kualitas metadata yang dihasilkan sangat dipengaruhi oleh kualitas informasi yang diberikan pengguna saat proses unggah karya.



Gambar 6. Hasil Metadata YAKE! pada Detail Karya

D. Implementasi BM25 pada Fitur Pencarian

Algoritma BM25 diterapkan untuk menghitung relevansi antara kueri pengguna dan metadata karya. Proses pencarian berjalan secara asinkron menggunakan AJAX [24], sehingga hasil pencarian diperbarui tanpa memuat ulang halaman. Sistem juga dilengkapi *filter* berdasarkan program studi, kategori, tahun, dan tipe karya untuk mempersempit hasil pencarian. Tampilan halaman pencarian karya oleh pengunjung ditampilkan pada Gambar 7.

E. Hasil Black-Box Testing

Hasil pengujian Black-Box Testing menunjukkan bahwa seluruh 20 skenario pengujian fungsional memperoleh status lulus (100%). Skenario yang diuji mencakup akses katalog oleh Guest, registrasi dan unggah karya oleh mahasiswa, *Auto-tagging* YAKE!, Validasi karya oleh operator prodi, pembatasan akses admin fakultas, dan pengelolaan *search tools* oleh super admin. Rekapitulasi hasil pengujian ditunjukkan pada Tabel VII.

TABEL VII
REKAPITULASI HASIL BLACK-BOX TESTING

Role	Jumlah Skenario	Status
Guest	3	Lulus
Mahasiswa	6	Lulus
Operator Prodi	4	Lulus
Admin Fakultas	3	Lulus
Super Admin	4	Lulus
Total	20	Lulus (100%)

F. Hasil Evaluasi Relevansi Pencarian

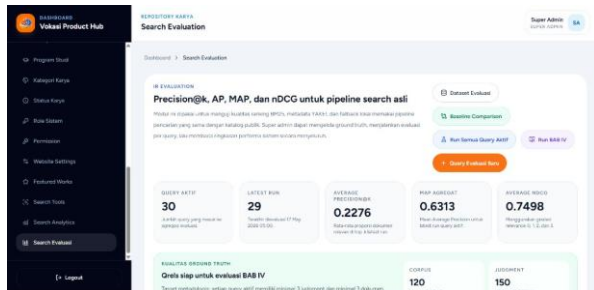
Evaluasi relevansi pencarian dilakukan menggunakan dataset simulasi berisi 120 karya, 30 kueri aktif, dan 150 *relevance judgment* pada skala 0-3. Rekapitulasi hasil evaluasi metrik IR ditunjukkan pada Tabel VIII.

TABEL VIII
REKAPITULASI EVALUASI RELEVANSI PENCARIAN

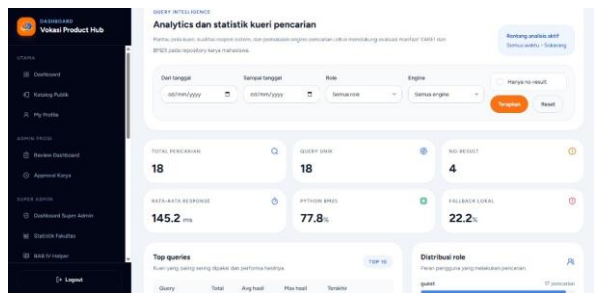
Metrik	Nilai	Interpretasi
Precision@5	0,3533	Relevansi top-5
Precision@10	0,2333	Relevansi top-10
MAP	0,5732	Rata-rata <i>overall</i>
nDCG@5	0,6539	Kualitas urutan top-5
nDCG@10	0,7093	Kualitas urutan top-10

Nilai Precision@5 lebih tinggi daripada Precision@10 karena hasil pada lima urutan awal lebih banyak memuat dokumen yang relevan. Penurunan nilai Precision@10 dipengaruhi oleh karakteristik *ground truth* yang menggunakan 5 dokumen *judgment* per kueri, sehingga membatasi nilai Precision pada k besar. Nilai MAP sebesar 0,5732 menunjukkan bahwa dokumen relevan sudah cukup sering muncul pada posisi awal. Nilai nDCG@10 yang lebih tinggi dari nDCG@5 menunjukkan bahwa dokumen dengan relevansi bertingkat masih ditemukan pada rentang sepuluh hasil teratas, dan urutan

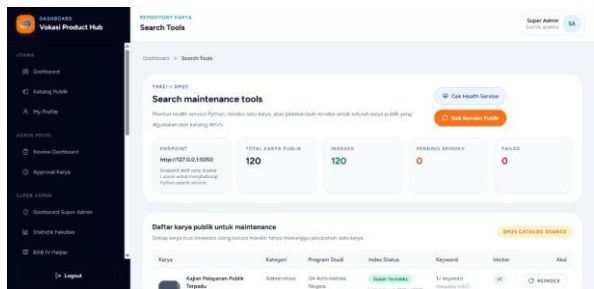
hasil mendekati urutan ideal. Tampilan antarmuka evaluasi relevansi pencarian pada sistem ditunjukkan pada Gambar 8. Statistik performa pencarian secara keseluruhan dapat dilihat pada Gambar 9. Serta, infrastruktur pengelolaan indeks pencarian ditunjukkan pada Gambar 10.



Gambar 8. Halaman Search Evaluation



Gambar 9. Halaman Search Analytics



Gambar 10. Halaman Search Tool Indexing Reindexing

G. Contoh Perhitungan Metrik Evaluasi

Sebagai ilustrasi proses perhitungan, berikut diuraikan perhitungan metrik pada salah satu kueri uji, yaitu kueri “dashboard inventaris gudang”. Pada ground truth, kueri ini memiliki lima dokumen dengan skor relevansi 3, 2, 2, 2, dan 1, dan empat di antaranya tergolong relevan (skor ≥ 2). Sepuluh dokumen teratas yang dikembalikan sistem beserta kontribusi DCG-nya disajikan pada Tabel IX.

TABEL IX
PERHITUNGAN DCG KUERI “DASHBOARD INVENTARIS GUDANG”

Peringkat	Skor Relevansi	Kontribusi DCG
1	3	7,000

2	2	1,893
3	0	0,000
4	0	0,000
5	0	0,000
6	0	0,000
7	2	1,000
8	0	0,000
9	0	0,000
10	2	0,867

Dokumen relevan (skor ≥ 2) berada pada peringkat 1, 2, 7, dan 10. Dengan demikian, $Precision@5 = 2/5 = 0,400$ dan $Precision@10 = 4/10 = 0,400$.

Average Precision (AP) merata-ratakan nilai precision pada keempat peringkat dokumen relevan, yaitu $(1,000 + 1,000 + 0,429 + 0,400) / 4 = 0,707$.

Berdasarkan kontribusi pada Tabel IX, $DCG@5 = 7,000 + 1,893 = 8,893$ dan $DCG@10 = 8,893 + 1,000 + 0,867 = 10,760$. Urutan ideal (skor 3, 2, 2, 2, 1) menghasilkan $IDCG = 12,072$, sehingga $nDCG@5 = 8,893/12,072 = 0,737$ dan $nDCG@10 = 10,760/12,072 = 0,891$. Nilai-nilai ini menjadi salah satu baris pada rekapitulasi seluruh kueri uji yang menghasilkan nilai rata-rata akhir pada Tabel VIII.

H. Pembahasan

Hasil evaluasi menunjukkan bahwa integrasi YAKE! dan BM25 dapat diterapkan pada Website Vokasi Product Hub untuk mendukung pencarian cerdas berbasis metadata karya. YAKE! berperan menghasilkan *keyword* otomatis dari deskripsi karya, sedangkan BM25 berperan menghitung relevansi dan mengurutkan hasil pencarian. Istilah “pencarian cerdas” dalam penelitian ini merujuk pada proses pencarian informasi yang memanfaatkan *auto-tagging* metadata dan pemeringkatan relevansi, bukan pada pencarian semantik berbasis model bahasa skala besar.

Nilai Precision perlu ditinjau dengan mempertimbangkan karakteristik ground truth yang digunakan. Untuk setiap kueri hanya tersedia 5 dokumen judgment dari total 120 dokumen, sehingga 115 dokumen lain tidak memiliki penilaian eksplisit dan secara konvensi dianggap tidak relevan. Apabila sebuah kueri hanya memiliki 2 sampai 3 dokumen relevan, batas atas teoretis $Precision@10$ berada pada rentang 0,20 sampai 0,30 meskipun seluruh dokumen relevan telah diperingkat pada posisi teratas. Dengan demikian, nilai $Precision@10$ sebesar 0,2333 lebih mencerminkan keterbatasan jumlah relevance judgment per kueri pada skema single-rater, bukan kelemahan algoritma pemeringkatan. Nilai nDCG yang dinormalisasi terhadap urutan ideal (IDCG) lebih tahan terhadap keterbatasan ini, sehingga $nDCG@10$ sebesar 0,7093 memberikan gambaran kualitas pemeringkatan yang lebih akurat.

Posisi hasil penelitian terhadap penelitian terdahulu dianalisis secara kualitatif karena perbedaan domain korpus, skema penyusunan ground truth, dan metrik evaluasi yang digunakan. Penelitian [8] yang menerapkan BM25 pada sistem pencarian lowongan pekerjaan

diposisikan sebagai penguat konteks pemilihan metode, bukan pembanding kuantitatif. Pemilihan kombinasi metrik berbasis posisi (Precision@k dan nDCG) serta metrik berbasis rata-rata (MAP) sejalan dengan praktik evaluasi Information Retrieval [9], sedangkan karakteristik unsupervised YAKE! yang menghasilkan kata kunci representatif tanpa data pelatihan konsisten dengan temuan [6].

Kualitas pencarian BM25 sangat dipengaruhi oleh kelengkapan metadata dan pemilihan istilah pada kueri. Keterbatasan sistem antara lain:

- YAKE! hanya mengekstrak *keyword* dari deskripsi teks, belum dari isi file,
- BM25 belum memiliki kemampuan pemahaman semantik penuh,
- Dataset masih berupa simulasi dan perlu divalidasi dengan data aktual, serta
- Eksplorasi *parameter tuning* BM25 belum dilakukan.

IV. KESIMPULAN

Penelitian ini berhasil mengembangkan Website Vokasi Product Hub sebagai repositori karya mahasiswa Fakultas Vokasi UNESA dengan sistem pencarian cerdas berbasis integrasi BM25 dan YAKE!. Kesimpulan utama penelitian ini adalah:

- Website Vokasi Product Hub berhasil dibangun menggunakan Laravel 12 dengan lima *role* pengguna (Guest, Mahasiswa, Operator Prodi, Admin Fakultas, Super Admin) beserta alur pengelolaan karya yang lengkap.
- Algoritma YAKE! berhasil diimplementasikan sebagai fitur *auto-tagging* melalui layanan Python menggunakan REST API. Kualitas *keyword* bergantung pada kelengkapan deskripsi yang diinputkan mahasiswa.
- Algoritma BM25 berhasil diterapkan untuk pemeringkatan hasil pencarian berdasarkan relevansi metadata karya, berjalan secara asinkron menggunakan AJAX.
- Seluruh 20 skenario Black-Box Testing memperoleh status lulus (100%), membuktikan bahwa fungsi-fungsi inti sistem berjalan sesuai rancangan.
- Evaluasi relevansi pada dataset simulasi menghasilkan Precision@5 = 0,3533; Precision@10 = 0,2333; MAP = 0,5732; nDCG@5 = 0,6539; dan nDCG@10 = 0,7093. Nilai nDCG@10 yang relatif tinggi menunjukkan sistem mampu memeringkat dokumen relevan mendekati urutan ideal pada sepuluh hasil teratas.

Penelitian lanjutan dapat difokuskan pada perbandingan YAKE! dengan metode lain, seperti KeyBERT atau TextRank. Selanjutnya, eksplorasi terkait *parameter tuning* BM25, dan implementasi BM25F dengan bobot *field* yang berbeda. Lalu, evaluasi menggunakan dataset yang aktual karya mahasiswa dengan *ground truth multi-rater*.

REFERENSI

- [1] P. Zibani, M. Rajkoomar, and N. Naicker, "A systematic review of faculty research repositories at higher education institutions," *Digital Library Perspectives*, vol. 38, no. 2, pp. 237–248, 2022, doi: 10.1108/DLP-04-2021-0035.
- [2] A. A. Bamgbose, M. Mohd, T. S. M. Tengku Wook, and H. Mohamed, "Are digital repositories for university libraries trustworthy?: A systematic review," *Business Information Review*, vol. 42, no. 2, pp. 125–132, 2025, doi: 10.1177/02663821251328830.
- [3] E. Orduña-Malea, C. I. Font-Julián, and J. Serrano-Cobos, "Open access publications drive few visits from Google Search results to institutional repositories," *Scientometrics*, vol. 129, no. 11, pp. 7131–7152, 2024, doi: 10.1007/s11192-024-05175-0.
- [4] J. Lin, R. Nogueira, and A. Yates, "Pretrained Transformers for Text Ranking: BERT and beyond," *Synthesis Lectures on Human Language Technologies*, vol. 14, no. 4, pp. 1–325, 2021, doi: 10.2200/S01123ED1V01Y202108HLT053.
- [5] K. A. Hambarde and H. Proenca, "Information Retrieval: Recent Advances and beyond," *IEEE Access*, vol. 11, pp. 76581–76604, 2023, doi: 10.1109/ACCESS.2023.3295776.
- [6] R. Campos, V. Mangaravite, A. Pasquali, A. Jorge, C. Nunes, and A. Jatowt, "YAKE! Keyword extraction from single documents using multiple local features," *Information Sciences*, vol. 509, pp. 257–289, 2020, doi: 10.1016/j.ins.2019.09.013.
- [7] M. Nosé et al., "Enhancing Findability and Searchability of Research Data: Metadata Conversion and Registration in Institutional Repositories," *Data Science Journal*, vol. 23, pp. 1–9, 2024, doi: 10.5334/dsj-2024-040.
- [8] T. Kheng, J. S. Asri, S. Wahyu, and Y. Yulhendri, "Penerapan Algoritma BM25 dalam Pencarian Lowongan Pekerjaan pada Website Job Portal," *Bulletin of Computer Science Research*, vol. 5, no. 5, pp. 1029–1038, 2025, doi: 10.47065/bulletinscr.v5i5.760.
- [9] N. Rastogi, P. Verma, and P. Kumar, "Performance evaluation of reformulated query for information retrieval using real estate ontology," *Int. J. Intelligent Information and Database Systems*, vol. 14, no. 4, pp. 365–378, 2021, doi: 10.1504/IJIDS.2021.118547.
- [10] A. L. Santonanda et al., "Online Craftsman Ordering Application Development using Waterfall Methodology," in *Proc. IC2IE 2022*, pp. 150–155, 2022, doi: 10.1109/IC2IE56416.2022.9970021.
- [11] O. D. Arianto and Y. A. Susetyo, "Penerapan Restful Web Service Dengan Framework Laravel Untuk Pembangunan Sistem Informasi Manajemen Sumber Daya Manusia," *JUPI*, vol. 7, no. 2, pp. 522–532, 2022, doi: 10.29100/jupi.v7i2.2870.
- [12] M. A. Puspita, M. Mardiana, R. Ariesta, and M. A. Muhammad, "User Acceptance Testing Dan Performance Testing Pada Pengembangan Website Kanal Pengetahuan Dikti," *Journal of Information Engineering and Educational Technology*, vol. 8, no. 2, pp. 96–101, 2025, doi: 10.26740/jieet.v8n2.p96-101.
- [13] M. Nadim, D. Akopian, and A. Matamoros, "A Comparative Assessment of Unsupervised Keyword Extraction Tools," *IEEE Access*, vol. 11, pp. 144778–144798, 2023, doi: 10.1109/ACCESS.2023.3344032.
- [14] M. Kim and Y. Ko, "Multitask Fine-Tuning for Passage Re-Ranking Using BM25 and Pseudo Relevance Feedback," *IEEE Access*, vol. 10, pp. 54254–54262, 2022, doi: 10.1109/ACCESS.2022.3176894.
- [15] Q. Chen, W. Zhou, J. Cheng, and J. Yang, "An Enhanced Retrieval Scheme for a Large Language Model with a Joint Strategy of Probabilistic Relevance and Semantic Association in the Vertical Domain," *Applied Sciences*, vol. 14, no. 24, art. 11529, 2024, doi: 10.3390/app142411529.
- [16] A. Moffat, "Batch Evaluation Metrics in Information Retrieval: Measures, Scales, and Meaning," *IEEE Access*, vol. 10, pp. 105564–105577, 2022, doi: 10.1109/ACCESS.2022.3211668.
- [17] S. Althammer, S. Hofstätter, S. Verberne, and A. Hanbury, "TripJudge: A relevance judgement test collection for TripClick health retrieval," in *Proc. 31st ACM Int. Conf. on Information and Knowledge Management (CIKM)*, 2022, doi: 10.1145/3511808.3557714.

- [18] Z. H. Amur, Y. K. Hooi, G. M. Soomro, H. Bhanbhro, S. Karyem, and N. Sohu, "Unlocking the potential of keyword extraction: The need for access to high-quality datasets," *Applied Sciences*, vol. 13, no. 12, art. 7228, 2023, doi: 10.3390/app13127228.
- [19] D. Ganguly and E. Yilmaz, "Query-specific variable depth pooling via query performance prediction," in *Proc. 46th Int. ACM SIGIR Conf. on Research and Development in Information Retrieval*, 2023, doi: 10.1145/3539618.3592046.
- [20] C. A. Györödi, D. V. Dumșe-Burescu, R. Györödi, D. R. Zmaranda, L. Bandici, and D. E. Popescu, "Performance impact of optimization methods on MySQL document-based and relational databases," *Applied Sciences*, vol. 11, no. 15, art. 6794, 2021, doi: 10.3390/app11156794.
- [21] M. Z. Hamidi, D. Purwitasari, and R. N. E. Anggraini, "Kombinasi ekstraksi kata kunci dan ekspansi kueri untuk deteksi isu etik pada ringkasan penelitian kesehatan," *Techno.Com*, vol. 22, no. 1, pp. 58–67, 2023, doi: 10.33633/tc.v22i1.7268.
- [22] A. R. Kamila, G. H. Derhass, Surianto, and V. Budiyo, "Evaluation of keyword extraction using YAKE and KeyBERT in text preprocessing for hoax news detection based on Bi-LSTM," *Riwayat: Educational Journal of History and Humanities*, vol. 8, no. 3, pp. 4490–4500, 2025, doi: 10.24815/jr.v8i3.48626.
- M. Q. Khan, A. Shahid, M. I. Uddin, M. Roman, A. Alharbi, W. Alosaimi, J. Almalki, and S. M. Alshahrani, "Impact analysis of keyword extraction using contextual word embedding," *PeerJ Computer Science*, vol. 8, art. e967, 2022, doi: 10.7717/peerj-cs.967.
- [23] P. D. Nusantara, "Peningkatan kinerja fitur pencarian dengan teknologi AJAX pada aplikasi web," *Jurnal Informatika & Komputasi*, vol. 14, pp. 66–75, 2020.
- [24] G. M. Rosa, R. C. Rodrigues, R. Lotufo, and R. Nogueira, "Yes, BM25 is a strong baseline for legal case retrieval," in *Proc. COLIEE 2021 Workshop*, 2021. [Online]. Available: [arXiv:2105.05686](https://arxiv.org/abs/2105.05686).
- [25] Z. Subecz, "Web-development with Laravel framework," *Gradus*, vol. 8, no. 1, pp. 211–218, 2021, doi: 10.47833/2021.1.csc.006.
- [26] Sugiyono, *Metode Penelitian Kuantitatif, Kualitatif dan R&D*. Bandung, Indonesia: Alfabeta, 2016.
- [27] N. Thakur, N. Reimers, A. Rücklé, A. Srivastava, and I. Gurevych, "BEIR: A heterogeneous benchmark for zero-shot evaluation of information retrieval models," in *Proc. 35th Conf. on Neural Information Processing Systems (NeurIPS) Datasets and Benchmarks Track*, 2021.
- [28] Yuda Syahidin and R. Ramadhan, "REST API architecture design on multi-platform device development," *Jurnal E-Komtek*, vol. 5, no. 2, pp. 178–189, 2021, doi: 10.37339/e-komtek.v5i2.762.
- [29] Z. Zhang, "An improved BM25 algorithm for clinical decision support in precision medicine based on co-word analysis and Cuckoo Search," *BMC Medical Informatics and Decision Making*, vol. 21, no. 1, pp. 1–15, 2021, doi: 10.1186/s12911-021-01454-5.