

PENGARUH MESIN PENCARIAN DAN ARSITEKTUR DALAM MENGOPTIMUMKAN KINERJA DAN SKALABILITAS SISTEM KATALOG DIGITAL PERPUSTAKAAN

VINCENTIA DYAH KUSUMANINGTYAS



**PROGRAM MAGISTER ILMU KOMPUTER
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**

@Hak cipta milik IPB University

IPB University



- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



@Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



PERNYATAAN MENGENAI TESIS DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Saya menyatakan dengan sebenar-benarnya bahwa segala pernyataan dalam thesis saya yang berjudul Pengaruh Mesin Pencarian dan Arsitektur dalam Mengoptimumkan Kinerja dan Skalabilitas Sistem Katalog Digital Perpustakaan adalah karya saya dengan arahan komisi pembimbing dan belum pernah diajukan dalam berbagai bentuk kepada perguruan tinggi lainnya. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan penulis lain telah disebutkan dalam teks dan dicantumkan dalam daftar pustaka di bagian akhir tesis ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Juli 2026

Vincentia Dyah Kusumaningtyas
G6501221013



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

RINGKASAN

VINCENTIA DYAH KUSUMANINGTYAS. Pengaruh Mesin Pencarian dan Arsitektur dalam Mengoptimalkan Kinerja dan Skalabilitas Sistem Katalog Digital Perpustakaan. Dibimbing oleh IRMAN HERMADI dan TAUFIQ ABDUL GANI.

Sistem katalog digital perpustakaan dituntut untuk mengelola data bibliografis berskala besar sekaligus memberikan layanan pencarian yang cepat dan relevan. Penelitian ini mengevaluasi pengaruh konfigurasi mesin pencarian berbasis Apache Solr dan pilihan arsitektur sistem terhadap kinerja dan skalabilitas sistem katalog digital, serta mengembangkan strategi pencarian adaptif berdasarkan karakteristik kueri pengguna, menggunakan dataset 1.527.259 rekaman bibliografis dari sistem INLIS Perpustakaan Nasional Republik Indonesia (Perpusnas).

Lima pendekatan pencarian Apache Solr dievaluasi, terdiri dari empat konfigurasi tokenizer (*text_standard*, *text_edgengram*, *text_ngram*, *text_typo_tolerant*) dan satu teknik pencarian *fuzzy* berbasis operator kueri, menggunakan 30 kueri original dan 30 kueri *typo* yang diukur dengan metrik *precision*, *recall*, dan *F1-score*, kemudian dianalisis dengan uji *One-Way ANOVA*. Perbandingan arsitektur monolitik dan *Microservices* yang dibangun dengan Laravel 11 dilakukan menggunakan Apache JMeter pada tiga tingkat beban (200, 500, dan 1.000 *concurrent users*), mencakup skenario CRUD, pencarian katalog, dan beban campuran. Uji skalabilitas dilakukan melalui *Vertical Scaling* dan *horizontal scaling*.

Hasil evaluasi menunjukkan bahwa perbedaan konfigurasi tokenizer berpengaruh signifikan terhadap performa pencarian, terutama pada kueri *typo* ($p = 1,14 \times 10^{-5}$) dan kueri campuran ($p = 5,09 \times 10^{-5}$). Strategi pencarian adaptif yang diusulkan, yang mengombinasikan konfigurasi *text_standard*, *text_ngram*, dan *fuzzy search* secara dinamis berdasarkan panjang kueri, menghasilkan *precision* 0,735, *recall* 0,993, dan *F1-score* 0,795 — meningkat 21,6% dibandingkan konfigurasi terbaik tunggal (*text_ngram*, $F1 = 0,654$). Temuan penting lainnya adalah kueri dua kata merupakan titik paling kritis bagi sistem, bukan kueri terpanjang, karena ambiguitas pada level tersebut menjadi penyebab utama penurunan performa.

Dari sisi arsitektur, arsitektur monolitik menunjukkan keunggulan waktu respons pada beban rendah (200 pengguna, median 2.707 ms) dibandingkan *microservices* (3.220 ms), karena tidak adanya *overhead* komunikasi antar layanan. Sebaliknya, arsitektur *Microservices* secara konsisten menunjukkan *throughput* lebih tinggi dan waktu respons lebih rendah pada operasi pencarian katalog dan beban campuran di seluruh tingkat beban, dengan selisih mencapai 4.998 ms pada 1.000 pengguna. Pada uji *Vertical Scaling*, penambahan sumber daya menghasilkan *speedup* 1,29× pada *Microservices* dan 1,25× pada monolitik. Uji *horizontal scaling* pada *Microservices* menghasilkan *speedup* hingga 1,58× pada empat *instance*, meskipun pertumbuhan tidak linear akibat *bottleneck* pada komponen *shared data layer* dan *gateway*.

Analisis *cost-benefit* menunjukkan bahwa meskipun arsitektur *microservices* membutuhkan infrastruktur lebih besar, efisiensi *throughput* per vCPU-nya tetap lebih tinggi (15,9 req/s/vCPU dibandingkan 14,8 req/s/vCPU pada monolitik).

Simulasi proyeksi biaya sepuluh tahun menunjukkan *microservices* berpotensi lebih efisien secara kumulatif dalam jangka panjang karena mampu melakukan *horizontal scaling* tanpa lompatan biaya migrasi *instance* seperti pada monolitik, serta memberikan manfaat *tangible* dan *intangible* yang lebih besar, meski monolitik tetap unggul pada kesederhanaan pengelolaan operasional. Kombinasi strategi pencarian adaptif dengan migrasi bertahap menuju arsitektur *microservices*, khususnya pada layanan pencarian katalog, direkomendasikan sebagai pendekatan yang paling sesuai untuk sistem katalog digital perpustakaan berskala besar seperti Perpustakaan Nasional.

Kata kunci: Strategi Pencarian Adaptif, Apache Solr, Arsitektur *Microservices*, Skalabilitas Sistem, Kinerja Pencarian, Katalog Digital Perpustakaan

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

SUMMARY

VINCENTIA DYAH KUSUMANINGTYAS. The Role of Search Engine Technology and System Architecture in Optimizing Performance and Scalability of Digital Library Catalog Systems. Supervised by IRMAN HERMADI and TAUFIQ ABDUL GANI.

Digital library catalog systems are required to manage large-scale bibliographic data while delivering fast and relevant search services. This study evaluated the effect of Apache Solr search engine configuration and system architecture on the performance and scalability of a digital library catalog system, and developed an Adaptive search strategy based on query characteristics, using a dataset of 1,527,259 bibliographic records from the INLIS system of the National Library of Indonesia (Perpusnas).

Five Apache Solr search approaches were evaluated, consisting of four tokenizer configurations (`text_standart`, `text_edgengram`, `text_ngram`, `text_typo_tolerant`) and one fuzzy search technique based on a query-time operator, measured using 30 original and 30 typo queries evaluated with precision, recall, and F1-score metrics, then analyzed through One-Way ANOVA tests. A comparison of monolithic and microservices architectures, both built on Laravel 11, was conducted using Apache JMeter under loads of 200, 500, and 1,000 concurrent users, covering CRUD operations, catalog search, and mixed-load scenarios. Scalability was assessed through vertical and horizontal scaling experiments.

Results showed that differences in tokenizer configuration significantly affected search performance, particularly for typo queries ($p = 1.14 \times 10^{-5}$) and mixed queries ($p = 5.09 \times 10^{-5}$). The proposed Adaptive search strategy, which dynamically combines `text_standart`, `text_ngram`, and fuzzy search based on query length, achieved a precision of 0.735, recall of 0.993, and F1-score of 0.795 — a 21.6% improvement over the best single configuration (`text_ngram`, $F1 = 0.654$). A key finding was that two-word queries posed the greatest challenge to the system, rather than the longest queries, due to increased ambiguity at that complexity level.

Regarding architecture, the monolithic approach delivered lower response times under low load (200 users: median 2,707 ms) compared to microservices (3,220 ms), due to the absence of inter-service communication overhead. Conversely, the microservices architecture consistently showed higher throughput and lower response times for catalog search and mixed-load operations across all load levels, with a difference of up to 4,998 ms at 1,000 concurrent users. In the Vertical Scaling test, additional resources yielded a speedup of 1.29× for microservices and 1.25× for monolithic. Horizontal scaling of microservices achieved speedups of up to 1.58× with four instances, although growth was sub-linear due to bottlenecks at the shared data layer and gateway components.

Cost-benefit analysis showed that the microservices architecture required higher development, start-up, and annual maintenance costs than the monolithic architecture due to the complexity of managing independent services, but was more cost-efficient in terms of annual infrastructure costs and throughput per vCPU (15.9 req/s/vCPU compared to 14.8 req/s/vCPU for monolithic). A ten-year cumulative cost projection simulation indicated that microservices could be more cost-efficient

in the long term, as it allows horizontal scaling without the instance-migration cost jumps required by monolithic architecture, while also delivering greater tangible and intangible benefits, although the monolithic architecture remained superior in terms of operational simplicity. The combination of the Adaptive search strategy with a phased migration toward microservices architecture, particularly for the catalog search service, is recommended as the most suitable approach for large-scale digital library catalog systems such as Perpustnas.

Keywords: Adaptive search strategy, Apache Solr, Microservices architecture, system scalability, search performance, digital library catalog



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2026
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

PENGARUH MESIN PENCARI DAN ARSITEKTUR DALAM MENGOPTIMUMKAN KINERJA DAN SKALABILITAS SISTEM KATALOG DIGITAL PERPUSTAKAAN

VINCENTIA DYAH KUSUMANINGTYAS

Tesis

Sebagai salah satu syarat untuk memperoleh gelar
Magister Komputer pada
Program Studi Ilmu Komputer

**PROGRAM MAGISTER ILMU KOMPUTER
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



Judul Tesis : Pengaruh Mesin Pencarian dan Arsitektur dalam Mengoptimalkan Kinerja dan Skalabilitas Sistem Katalog Digital Perpustakaan
Nama : Vincentia Dyah Kusumaningtyas
NIM : G6501221013

@Hak cipta milik IPB University

Disetujui oleh

Pembimbing 1:

Irman Hermadi, S.Kom., M.S., P.hD.

Pembimbing 2:

Dr. Taufiq A. Gani, S.Kom, M.Eng.Sc

Diketahui oleh

Ketua Program Studi:

Irman Hermadi, S.Kom., M.S., P.hD.
NIP 19750311 200604 1 009

Dekan Sekolah Sains Data, Matematika, dan Informatika:

Prof. Dr. Ir. Agus Buono, M.Si., M.Kom.
NIP 19660702 199302 1 001

Tanggal Ujian:
14 Juli 2026

Tanggal Lulus:

PRAKATA

Puji dan syukur penulis panjatkan kehadirat Tuhan Yang Maha Kuasa atas segala berkat dan penyertaan-Nya sehingga penulis dapat menyelesaikan tesis yang berjudul “Pengaruh Mesin Pencarian dan Arsitektur dalam Mengoptimalkan Kinerja dan Skalabilitas Sistem Katalog Digital Perpustakaan” ini sebagai salah satu syarat untuk memperoleh gelar Magister Komputer pada Program Studi Ilmu Komputer, Sekolah Pascasarjana, Institut Pertanian Bogor.

Penelitian ini membahas perbandingan performa konfigurasi Apache Solr serta arsitektur monolitik dan *microservices* dalam konteks sistem katalog perpustakaan digital, dan merupakan sebuah topik yang relevan dengan kebutuhan pengembangan sistem informasi perpustakaan di era digital saat ini, khususnya dalam mendukung pengelolaan koleksi yang mencapai jutaan data dan layanan yang efisien dan skalabel.

Penulis menyadari bahwa penyelesaian tesis ini tidak lepas dari bimbingan, dukungan, dan bantuan berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan terima kasih yang sebesar-besarnya kepada:

1. Bapak Dr. Irman Hermadi, S.Kom., M.S., Ph.D selaku ketua komisi pembimbing yang telah memberikan arahan, masukan, dan bimbingan dengan penuh kesabaran selama proses penyusunan tesis ini.
2. Bapak Dr. Taufiq A. Gani, M.Eng.Sc, selaku dosen pembimbing kedua atas segala bimbingan dan sarannya.
3. Pimpinan dan rekan-rekan di Perpustakaan Nasional Republik Indonesia, khususnya di Direktorat Deposit dan Pengembangan Koleksi Perpustakaan, atas dukungan dan pengertiannya selama penulis menjalani studi sekaligus bekerja.
4. Keluarga tercinta, Alm. Anskarius Anwar Sitanggang, Bapak Agustinus Agus Pamenang, Ibu Gabriella Susilowati, adik-adik Tiur, Petra, Mitha, Yosephine dan Ignatius atas doa, dukungan, dan pengertian yang tiada henti selama penulis menempuh pendidikan ini.
5. Sahabat-sahabat penulis, yang tidak dapat disebutkan satu per satu atas dukungan, semangat, dan kebersamaan yang selalu menjadi penguat bagi penulis dalam menyelesaikan tesis ini.
6. Seluruh dosen dan staf Program Studi Ilmu Komputer IPB yang telah memberikan ilmu dan bantuan administratif selama masa studi.
7. Rekan-rekan seperjuangan di Program Studi Magister Ilmu Komputer kelas khusus atas kebersamaan, diskusi, dan semangat yang saling menguatkan.

Penulis menyadari bahwa tesis ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan demi perbaikan di masa mendatang. Semoga tesis ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan, khususnya di bidang sistem informasi perpustakaan digital, serta bagi pihak-pihak yang membutuhkan.

Bogor, Juli 2026

Vincentia Dyah Kusumaningtyas



DAFTAR ISI

DAFTAR GAMBAR	x
DAFTAR TABEL	xi
DAFTAR LAMPIRAN	xii
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	3
1.4 Manfaat	3
1.5 Ruang Lingkup	3
1.6 Kebaruan	4
II TINJAUAN PUSTAKA	5
2.1 Sistem Katalog Digital Perpustakaan	5
2.2 Pengelolaan Metadata dalam Sistem Katalog Digital	5
2.3 Mesin Pencari dalam Sistem Katalog Digital	6
2.4 <i>Monolithic Architecture</i>	9
2.5 <i>Microservices</i>	10
2.6 <i>Framework Microservices</i>	11
2.7 CRUD (<i>Create, Read, Update, Delete</i>)	13
2.8 <i>Use Case Point (UCP)</i>	14
III METODE	16
3.1 Waktu dan Lokasi Penelitian	16
3.2 Data Penelitian	16
3.3 Tahapan Penelitian	16
3.4 Lingkungan Pengembangan	29
IV HASIL DAN PEMBAHASAN	31
4.1 Persebaran dan Karakteristik Data	31
4.2 Hasil Pengujian Mesin Pencari	32
4.3 Hasil Pengujian Arsitektur	45
4.4 Ringkasan Perbandingan Kinerja Arsitektur	62
4.5 Pengaruh Kondisi <i>Host</i> terhadap Stabilitas <i>Throughput</i>	63
4.6 Analisis <i>Cost-benefit</i> Arsitektur Sistem	63
V SIMPULAN DAN SARAN	77
5.1 Simpulan	77
5.2 Saran	77
5.3 Rekomendasi	78
DAFTAR PUSTAKA	79
LAMPIRAN	86
RIWAYAT HIDUP	98



DAFTAR GAMBAR

1	Tahapan-tahapan yang dilakukan pada penelitian	17
2	Arsitektur sistem katalog digital Perpustakaan saat ini	21
3	Arsitektur sistem katalog digital Perpustakaan yang diusulkan	22
4	Topologi lingkungan pengujian	24
5	Perbandingan <i>F1-Score</i> berdasarkan konfigurasi dan variasi kueri	33
6	Perbandingan <i>F1-Score</i> berdasarkan kompleksitas kueri dan konfigurasi	37
7	Algoritma pencarian adaptif	42
8	Perbandingan <i>F1-Score</i> semua konfigurasi dengan pencarian adaptif	45
9	Grafik perbandingan median <i>response time</i> pada operasi CRUD	46
10	Grafik perbandingan p95 <i>response time</i> pada operasi CRUD	47
11	Grafik perbandingan <i>throughput</i> pada operasi CRUD	48
12	Grafik perbandingan median <i>response time</i> pada operasi pencarian katalog	49
13	Grafik perbandingan p95 <i>response time</i> pada operasi pencarian katalog	50
14	Grafik perbandingan <i>throughput</i> pada operasi pencarian katalog	51
15	Grafik perbandingan median <i>response time</i> pada operasi campuran	52
16	Grafik perbandingan p95 <i>response time</i> pada operasi campuran	53
17	Grafik perbandingan <i>throughput</i> pada operasi campuran	53
18	Grafik perbandingan <i>throughput</i> vs vCPU pada pengujian <i>Vertical Scaling</i>	58
19	Grafik perbandingan <i>speedup</i> vs vCPU pada pengujian <i>Vertical Scaling</i>	59
20	Grafik perbandingan <i>throughput</i> dan <i>instance</i> pada uji <i>horizontal scaling</i>	61
21	Grafik perbandingan <i>speedup</i> (S_{hor}) dan <i>instance</i> pada uji <i>horizontal scaling</i>	61
22	Diagram <i>use case</i> inti sistem katalog digital	64
23	Simulasi biaya kumulatif 10 tahun	71

@Hak cipta milik IPB University



DAFTAR TABEL

1	Kolom tabel <i>catalogs</i> yang digunakan dalam penelitian	17
2	Distribusi katalog berdasarkan kelas DDC	31
3	Hasil evaluasi rata-rata kueri <i>original</i>	33
4	Hasil evaluasi rata-rata kueri <i>typo</i>	33
5	Hasil evaluasi gabungan seluruh kueri	34
6	Hasil evaluasi kueri gabungan per jenis kueri	34
7	Hasil evaluasi kueri <i>original</i> per jenis kueri	35
8	Hasil evaluasi kueri <i>typo</i> per jenis kueri	36
9	ANOVA berdasarkan varian kueri	39
10	ANOVA berdasarkan kompleksitas kueri	39
11	Kompleksitas komputasi tiap pendekatan pencarian	43
12	Hasil Evaluasi rata-rata performa konfigurasi pencarian	44
13	Hasil uji kinerja operasi CRUD pada arsitektur <i>monolitik</i> dan <i>Microservices</i>	46
14	Hasil uji kinerja pencarian katalog	49
15	Hasil uji kinerja campuran operasi CRUD dan pencarian katalog	52
16	Efisiensi <i>throughput</i> per vCPU pada skenario campuran	55
17	Hasil uji <i>vertical scaling</i> pada arsitektur <i>microservices</i>	56
18	Skalabilitas vertikal pada arsitektur <i>microservices</i>	56
19	Hasil uji <i>vertical scaling</i> pada arsitektur monolitik	57
20	Skalabilitas vertikal pada arsitektur monolitik	57
21	Hasil uji kinerja <i>horizontal scaling</i>	60
22	Skalabilitas horizontal	60
23	Efisiensi <i>throughput</i> terhadap CPU	62
24	Daftar deskripsi <i>use case</i> inti sistem katalog digital	64
25	Klasifikasi aktor dan UAW	65
26	Klasifikasi <i>use case</i> dan UUCW	65
27	Perbandingan <i>Technical Complexity Factor</i> (TCF)	66
28	Perhitungan <i>Environmental Factor</i> (EF)	66
29	Estimasi UCP, effort dan biaya pengembangan	67
30	Estimasi biaya <i>procurement</i>	67
31	Estimasi biaya <i>start-up</i>	68
32	Estimasi biaya <i>ongoing</i>	69
33	Ringkasan total biaya implementasi	69
34	Proyeksi biaya kumulatif 10 tahun	70
35	Manfaat <i>intangible</i> implementasi sistem	72

@Hak cipta milik IPB University

DAFTAR LAMPIRAN

1	<i>Ground Truth</i> Dokumen	87
2	Daftar Skrip Program	89
3	Konfigurasi <i>Schema</i> Apache Solr	90
4	Antarmuka Prototipe Sistem	91
5	<i>Oracle Virtual Box Manager</i>	92
6	Deskripsi Rinci <i>Use Case</i> Inti Sistem Katalog Digital	93
7	Rincian perhitungan <i>Technical Complexity Factor</i> (TCF)	96
8	Kueri SQL untuk Ekstraksi Data Tabel <i>Catalog</i>	97