

# ***SRITEL-U: Pengembangan Platform Terintegrasi Google Scholar dan Pemodelan Topik Menggunakan Latent Dirichlet Allocation untuk Rekomendasi Peneliti dalam Perguruan Tinggi***

**Mochamad Nizar Palefi Ma'ady<sup>1</sup>, Desyka Widya Haryanti<sup>2</sup>, Rengganis Puspa Garini<sup>3</sup>,  
Lathifa Puteri Asy'ari<sup>4</sup>**

<sup>1,2</sup>*Program Studi Sistem Informasi, Telkom University, Kampus Surabaya*

<sup>3,4</sup>*Program Studi Teknik Logistik, Telkom University, Kampus Surabaya*

*e-mail: mnizarpm@telkomuniversity.ac.id<sup>1</sup>, desykawidya@student.telkomuniversity.ac.id<sup>2</sup>,  
rengganispuspa@student.telkomuniversity.ac.id<sup>3</sup>, lathifaasyari@student.telkomuniversity.ac.id<sup>4</sup>*

---

## **ABSTRAK**

### ***Kata Kunci:***

*Sistem Rekomendasi  
Text Mining  
Google Scholar  
Latent Dirichlet Allocation  
Perguruan Tinggi*

Kolaborasi penelitian multidisiplin di perguruan tinggi penting untuk inovasi ilmiah, namun menemukan partner peneliti yang tepat sering menjadi tantangan. Biaya sistem rekomendasi seperti *Elsevier* juga tinggi, sementara sistem berbasis *Google Scholar* bisa lebih terjangkau. Penelitian ini mengembangkan *SRITEL-U*, platform rekomendasi peneliti berbasis web yang terintegrasi dengan *Google Scholar*. Platform ini menggunakan text mining dan metode *Latent Dirichlet Allocation* (LDA) untuk mencocokkan bidang penelitian dengan keahlian dosen. *SRITEL-U* mendukung SDGs dalam kategori Pendidikan Berkualitas, Industri, Inovasi, dan Infrastruktur, serta Kemitraan untuk Mencapai Tujuan, menuju Visi Indonesia Emas 2045.

## **ABSTRACT**

### ***Keyword:***

*Recommendation System  
Text Mining  
Google Scholar  
Latent Dirichlet Allocation  
Higher Education*

*Multidisciplinary research collaborations in higher education are important for scientific innovation, but finding the right research partners is often a challenge. The cost of recommendation systems such as Elsevier is also high, while Google Scholar-based systems can be more affordable. This research develops SRITEL-U, a web-based researcher recommendation platform integrated with Google Scholar. The platform uses text mining and Latent Dirichlet Allocation (LDA) method to match research fields with lecturers' expertise. SRITEL-U supports the SDGs in the categories of Quality Education, Industry, Innovation, and Infrastructure, as well as Partnerships to Achieve Goals, towards the Golden Indonesia Vision 2045.*

## I. PENDAHULUAN

Inovasi dalam karya ilmiah sering kali muncul dari kombinasi berbagai disiplin ilmu dan pengembangan ide baru. Teknik *brainstorming* dan *literature review* menjadi kunci dalam memilih ide yang menarik, menentukan kesenjangan pengetahuan, dan mengidentifikasi penelitian potensial. Penelitian teoritis yang diperlihatkan terhadap Analisa konsep, teori dan ide – ide abstrak yang dikembangkan melalui penerapan *literature review* [1]. Pendekatan multidisiplin sangat penting dalam pendidikan di era Revolusi Industri 4.0, mengingat semakin kompleksnya konteks disiplin [2].

Perguruan tinggi di Indonesia menghadapi berbagai tantangan, termasuk kurangnya inovasi dalam lembaga penelitian. Kolaborasi penelitian mampu menjadi jembatan antar peneliti dalam hal inovasi dengan menambahkan ide – ide pengetahuan, keahlian dan peluang bagi peneliti [3]. Proses penentuan pembimbing untuk penelitian mahasiswa masih menggunakan metode manual yang tidak efektif, seperti *spreadsheet*, yang sering kali menyebabkan ketidaksesuaian antara penilaian dan bidang penelitian[4]. Selain itu, publikasi data yang tidak akurat dan perpindahan homebase juga menambah kompleksitas masalah yang dihadapi oleh perguruan tinggi [5].

*SRITEL-U* diusulkan sebagai solusi untuk meningkatkan efektivitas dalam publikasi penelitian mahasiswa dan penentuan pembimbing yang tepat. Menggunakan metode *Latent Dirichlet Allocation* (LDA) [6], platform ini mampu mengelompokkan kata dalam dokumen dan mengidentifikasi topik-topik utama. *SRITEL-U* mendukung tujuan *Sustainable Development Goals* (SDGs) [7], khususnya pada pilar ke-4 (*Quality Education*), pilar ke-9 (*Industry, Innovation, and Infrastructure*) [8], dan pilar ke-17 (*Partnerships for the Goals*), dengan tujuan meningkatkan akses dan kualitas pendidikan, membangun infrastruktur inovatif, dan meningkatkan kerja sama internasional [9].

## II. METODE

Kami mengusulkan *SRITEL-U*, platform rekomendasi untuk meringankan aktivitas pendidikan, terutama dalam publikasi penelitian mahasiswa dan penentuan pembimbing yang tepat berdasarkan bidang penelitian. Sistem ini menggunakan metode pemodelan topik *Latent Dirichlet Allocation* (LDA) dengan model probabilistik generatif untuk kumpulan data diskrit seperti topik *coherence*[10]. Metode pengumpulan dan pemrosesan data untuk informasi seperti penulis, judul publikasi, tanggal publikasi, dan abstrak telah berkembang pesat dalam beberapa tahun terakhir. LDA adalah metode efektif untuk menganalisis dokumen berukuran besar, mengelompokkan kata dalam dokumen berdasarkan frekuensi kemunculannya, sehingga mengidentifikasi topik-topik utama[11].

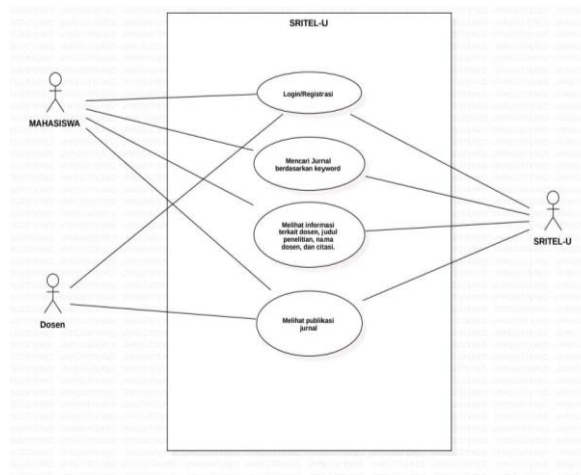
Metode ini diterapkan dalam platform sistem rekomendasi di *Telkom University* Surabaya untuk mendukung kebutuhan dan perkembangan perguruan tinggi. *SRITEL-U* merupakan platform

rekomendasi pembimbing untuk kolaborasi penelitian mahasiswa berdasarkan bidang penelitian mereka. Kolaborasi ini mendukung Tujuan Pembangunan Berkelanjutan (*SDGs*) khususnya pada pilar ke-4 (*Quality Education*), pilar ke-9 (*Industry, Innovation, and Infrastructure*), dan pilar ke-17 (*Partnerships for the Goals*). Tiga Tujuan Pembangunan Berkelanjutan (*SDGs*) merupakan serangkaian tujuan langkah besar secara universal untuk mengurangi kekurangan dari segi Pendidikan dan industry dengan melindungi seluruh makhluk hidup dan memastikan bahwa semua orang mendapatkan keadilan dan kedamaian [12].

### III. HASIL DAN PEMBAHASAN

#### *Use Case Diagram (UCD)*

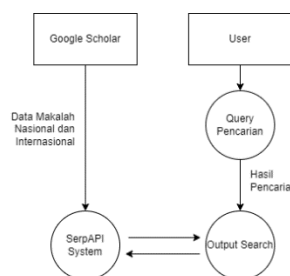
*Use Case Diagram (UCD)* merupakan representasi grafis dari interaksi antara pengguna sistem (entitas) dengan fungsi-fungsi utama yang disediakan oleh system [13]. Pada penelitian ini, sistem *SRITEL-U* memiliki beberapa entitas utama dan use case yang terkait.



**Gambar 1.** *Use Case Diagram*

#### *Data Flow Diagram (DFD)*

*Data Flow Diagram (DFD)* digunakan untuk menggambarkan aliran data dalam sistem dan proses-proses yang mengolah data tersebut. DFD membantu dalam memahami bagaimana data diproses dan disimpan di dalam sistem.



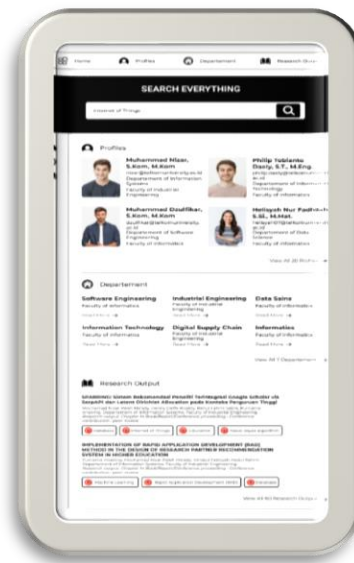
**Gambar 2.** *Data Flow Diagram*

## Pemaparan Hasil Mockup Program Sistem Rekomendasi Penelitian

Program sistem rekomendasi penelitian yang kami kembangkan direalisasikan dalam bentuk platform digital berbasis web. Pada halaman utama (*main page*), pengguna dapat menginputkan kata kunci (*keyword*) yang ingin dicari. Setelah melakukan pencarian, sistem akan menampilkan profil peneliti, departemen, dan output penelitian yang relevan dengan kata kunci atau topik yang dicari. Pengguna juga dapat melihat detail lebih lanjut dari hasil pencarian, seperti Riwayat publikasi, kolaborasi, dan informasi tambahan lainnya yang mendukung penelitian



Gambar 3. Main Page



Gambar 4. Mockup Search Result

## Validasi Pemodelan Topik

Pada bagian ini, akan dibahas proses validasi pemodelan topik menggunakan metode *Latent Dirichlet Allocation* (LDA). Proses ini dimulai dengan pengambilan data abstrak, yang kemudian diproses melalui LDA untuk mengidentifikasi topik yang terkandung dalam dokumen-dokumen tersebut. Sebagai contoh, digunakan data abstrak [14] dari beberapa penelitian yang akan di sitasi.



12	5	<i>Proposed</i>
13	6	<i>Furthermore</i>
14	6	<i>Time</i>
15	6	<i>Kmeans</i>
16	7	<i>Tablefilling</i>
17	7	<i>Procedure</i>
18	7	<i>Small</i>
19	8	<i>Retailer</i>
20	8	<i>Customer</i>

**Kamus Data:**

**Tabel 2.** Kamus Data

<b>Id</b>	<b>Word</b>
0	<i>Small</i>
1	<i>Procedure</i>
2	<i>Tablefilling</i>
3	<i>Time</i>
4	<i>Kmeans</i>
5	<i>Furthermore</i>
6	<i>Indicates</i>
7	<i>Efficient</i>
8	<i>Proposed</i>
9	<i>Retailer</i>
10	<i>Customer</i>
11	<i>Satisfaction</i>

Tabel 2 menunjukkan kosakata atau kamus yang digunakan. Kamus data adalah katalog fakta tentang kebutuhan data dan informasi suatu sistem informasi.

**Inisialisasi Topik:**

Langkah berikutnya adalah menginisialisasi topik untuk setiap token secara acak. Misalnya, jika terdapat empat topik, maka setiap token akan diinisialisasi ke topik satu hingga topik empat secara acak. Tabel 3 menunjukkan hasil inisialisasi topik untuk setiap token.

**Tabel 3.** Random Topic Initialization

<b>Id_token</b>	<b>Doc</b>	<b>Word</b>	<b>Topic (Random)</b>
0	1	<i>Customer</i>	4
1	1	<i>Satisfaction</i>	4

2	1	<i>Retailer</i>	4
3	2	<i>Retailer</i>	4
4	2	<i>Customer</i>	4
5	2	<i>Satisfaction</i>	4
6	3	<i>Proposed</i>	3
7	4	<i>Time</i>	2
8	4	<i>Kmeans</i>	2
9	5	<i>Indicates</i>	3
10	5	<i>Efficient</i>	3
11	5	<i>Proposed</i>	3
12	6	<i>Furthermore</i>	2
13	6	<i>Time</i>	2
14	6	<i>Kmeans</i>	2
15	7	<i>Tablefilling</i>	1
16	7	<i>Procedure</i>	1
17	7	<i>Small</i>	1
18	8	<i>Retailer</i>	4
19	8	<i>Customer</i>	4

**Perhitungan Probabilitas Kata Topik (PWZ):**

Rumus Perhitungan Probabilitas Kata Topik :

$$\varphi_{j,t} = p(w = t | z = j) = \frac{n_j^{(t)} + \beta_t}{\sum_{t=1}^v n_j^{(t)} + w\beta_t} \quad (1)$$

Kata PWZ (*customer*) :

Word (*customer*), topic 1

$$= \frac{n_1^{(1)} + \beta_1}{\sum_{t=1}^v n_1^{(1)} + w\beta_1} = \frac{0+0.01}{3+(12 \times 0.01)} = 0.0032$$

Word (*customer*), topic 2

$$= \frac{n_2^{(1)} + \beta_1}{\sum_{t=1}^v n_2^{(1)} + w\beta_1} = \frac{0+0.01}{5+(12 \times 0.01)} = 0.0020$$

Word (*customer*), topic 3

$$= \frac{n_3^{(1)} + \beta_1}{\sum_{t=1}^v n_3^{(1)} + w\beta_1} = \frac{0+0.01}{4+(12 \times 0.01)} = 0.0024$$

Word (*customer*), topic 4

$$= \frac{n_4^{(1)} + \beta_1}{\sum_{t=1}^v n_4^{(1)} + w\beta_1} = \frac{3+0.01}{8+(12 \times 0.01)} = 0.3707$$

Perhitungan semua kata dan membuahkan hasil  $PWZ_{j,t}$

**Tabel 4.** *Probability* dari *Word Topics* (PWZ)

<b>Topic</b>	<b>Pelanggan</b>
1	0.0032
2	0.0020
3	0.0024
4	0.3707

### **Hasil Pemrograman:**

Berikut adalah hasil dari pemrograman LDA. Sistem ini menggunakan metode resmi LDA dan menghasilkan output sebagai berikut:

### **Topik yang Teridentifikasi:**

Topic 1 ['small' 'procedure' 'tablefilling']

Topic 2 ['time' 'kmeans' 'furthermore']

Topic 3 ['indicates' 'efficient' 'proposed']

Topic 4 ['retailer' 'customer' 'satisfaction']

### **Topik yang relevan pada Setiap Dokumen:**

Document 1 -- Topic: [4 3 2 1]

Document 2 -- Topic: [4 3 2 1]

Document 3 -- Topic: [2 3 4 1]

Document 4 -- Topic: [1 2 3 4]

Document 5 -- Topic: [3 2 4 1]

Document 6 -- Topic: [2 1 3 4]

Document 7 -- Topic: [1 3 2 4]

Document 8 -- Topic: [4 2 3 1]

Hasil diatas menunjukkan bahwa dokumen-dokumen yang dianalisis memiliki topik yang bervariasi, dengan topik tertentu menjadi topik utama dalam beberapa dokumen. Hasil ini membantu dalam memahami distribusi topik dalam Kumpulan dokumen yang dianalisis.

### **Analisis Komparatif**

Dalam kegiatan ini, kami menyajikan perbandingan antara sistem rekomendasi peneliti *Elsevier* dan *SRITEL-U*. Kami membandingkan jumlah profil peneliti dan output penelitian untuk beberapa kata kunci utama. Data dari kedua sistem ditampilkan dalam tabel berikut :



## Platform Lain

Tabel 5. Elsevier

No	Keyword	Profile	Research Output
1.	<i>Internet Of Things</i>	16	55
2.	<i>Basis Data</i>	8	146
3.	<i>Digital Bisnis</i>	4	5

## SRITEL-U

Tabel 6. SRITEL-U

N0	Keyword	Profile	Research Output
1.	<i>Internet Of Things</i>	20	60
2.	<i>Basis Data</i>	10	147
3.	<i>Digital Bisnis</i>	5	10

Berdasarkan hasil perbandingan ini, dapat disimpulkan bahwa *SRITEL-U* menawarkan beberapa keunggulan dibandingkan dengan sistem rekomendasi peneliti dan *Elsevier*. Keunggulan-keunggulan ini menunjukkan bahwa *SRITEL-U* dapat memberikan rekomendasi yang lebih akurat dan relevan, serta mendukung kolaborasi penelitian yang lebih efektif dan produktif di perguruan tinggi. Implementasi *SRITEL-U* diharapkan dapat meningkatkan efektivitas dan efisiensi proses penelitian, mempercepat penemuan ilmiah, dan mendukung perkembangan ilmu pengetahuan teknologi di Indonesia.

## IV. SIMPULAN

*SRITEL-U*, sebagai sebuah sistem rekomendasi peneliti yang terintegrasi dengan *Google Scholar* dan pemodelan topik menggunakan *Latent Dirichlet Allocation (LDA)*, telah berhasil menunjukkan potensinya dalam meningkatkan kolaborasi penelitian di perguruan tinggi. Penggunaan teknologi pemodelan topik LDA memungkinkan sistem untuk mengidentifikasi topik penelitian yang relevan secara otomatis, sehingga memudahkan peneliti dalam menemukan mitra kolaborasi yang memiliki minat penelitian serupa [15].

Dalam implementasinya, *SRITEL-U* menggunakan visualisasi grafik untuk memaparkan hasil rekomendasi, yang memberikan tampilan yang intuitif dan mudah dipahami oleh pengguna. Selain itu, pengembangan aplikasi berbasis *mobile* memungkinkan aksesibilitas yang lebih luas dan fleksibel bagi para peneliti, sehingga mereka dapat mengakses informasi dan rekomendasi kapan saja dan di mana saja.

Untuk meningkatkan akurasi rekomendasi, perlu dilakukan peningkatan pada pemodelan topik dan algoritma rekomendasi, sehingga hasilnya menjadi lebih relevan dan dapat diandalkan oleh peneliti. Selain itu, pengembangan fitur kolaboratif, seperti forum diskusi atau pesan pribadi, akan

mendukung komunikasi dan kolaborasi yang lebih efektif antar peneliti dalam platform *SRITEL-U*. Ekspansi integrasi data dengan lebih banyak sumber, seperti database jurnal internasional atau *repository* penelitian nasional, akan memperkaya informasi. Terakhir, optimasi antarmuka pengguna, terutama untuk aplikasi *mobile*, penting agar lebih responsif dan *user-friendly*.

#### DAFTAR RUJUKAN

- [1] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén, “Systematic Literature Reviews Bt - Experimentation In Software Engineering,” C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, And A. Wesslén, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, Pp. 45–54. Doi: 10.1007/978-3-642-29044-2\_4.
- [2] M. Ridwan dan Sitti Rizki Mulyani, “Determination Of Price And Service Quality On Purchasing Decisions In The Jurnal Ilmu Multidisiplin,” *Dinasti Int. J. Digit. Bus. Manag.*, Vol. 4, No. 2, Pp. 402–410, 2023, Doi: 10.31933/Dijdbm.V4i2.1776.
- [3] J. F. Rusdi, T. G. Baktina, R. G. Hadiningrat, B. Sunaryo, P. Fannya, and F. Laurenty, “Kolaborasi Peneliti Di Era Teknologi Informasi Dan Komunikasi,” *Baca J. Dokumentasi Dan Inf.*, Vol. 41, No. 1, P. 23, May 2020, Doi: 10.14203/J.Baca.V41i1.585.
- [4] S. Keputusan Dirjen Penguatan Riset dan Pengembangan Ristek Dikti, Y. Sahria, And D. Hatta Fudholi, “Terakreditasi Sinta Peringkat 2 Analisis Topik Penelitian Kesehatan Di Indonesia Menggunakan Metode Topic Modeling Lda (Latent Dirichlet Allocation),” *Masa Berlaku Mulai*, Vol. 1, No. 3, Pp. 336–344, 2017.
- [5] N. A. Sultan dan D. B. Abdullah, “Investigating Scientific Collaboration Networks In Iraq Using Cloud Computing And Data Mining,” *Int. J. Appl. Sci. Eng.*, Vol. 20, No. 4, 2023, Doi: 10.6703/Ijase.202312\_20(4).003.
- [6] S. J. Blair, Y. Bi, dan M. D. Mulvenna, “Aggregated Topic Models For Increasing Social Media Topic Coherence,” *Appl. Intell.*, Vol. 50, No. 1, Pp. 138–156, Jan. 2020, Doi: 10.1007/S10489-019-01438-Z.
- [7] “Sdgs-Series-Pilar-Sosial”.
- [8] N. Humaida, M. Aula Sa’adah, H. Huriyah, dan N. Hasanatun Nida, “Pembangunan Berkelanjutan Berwawasan Lingkungan (Sustainable Development Goals) Dalam Perspektif Islam,” *Khazanah J. Stud. Islam Dan Hum.*, Vol. 18, No. 1, P. 131, Jun. 2020, Doi: 10.18592/Khazanah.V18i1.3483.
- [9] L. Salsabila, K. Teovani Lodan, dan E. Khairina, “Public Engagement Impact On Sustainable Waste Management In Indonesia: Examining Public Behavior,” *J. Adm. Publik Public Adm. J.*, Vol. 13, No. 2, Pp. 158–178, Dec. 2023, Doi: 10.31289/Jap.V13i2.10391.
- [10] J. Shen, W. Huang, dan Q. Hu, “Picf-Lda: A Topic Enhanced Lda With Probability Incremental Correction Factor For Web Api Service Clustering,” *J. Cloud Comput.*, Vol. 11, No. 1, Dec. 2022, Doi: 10.1186/S13677-022-00291-9.

- [11] H. Jelodar *Et Al.*, “Latent Dirichlet Allocation (Lda) And Topic Modeling: Models, Applications, A Survey,” *Multimed. Tools Appl.*, Vol. 78, No. 11, Pp. 15169–15211, 2019, Doi: 10.1007/S11042-018-6894-4.
- [12] S. Prabu Aji dan D. T. Kartono, “Kebermanfaat Adanya Sustainable Development Goals (Sdgs),” *J. Soc. Res.*, Vol. 1, No. 6, Pp. 507–512, 2022, Doi: 10.55324/Josr.V1i6.110.
- [13] Y. Andriyani, I. D. Id, E. Mahdiyah, dan A. Aminuddin, “Use Case Realization In Software Reverse Engineering,” *Ing. Des Syst. D’information*, Vol. 27, No. 2, Pp. 335–341, 2022, Doi: 10.18280/Isi.270218.
- [14] M. R. Fahlevvi dan A. Sn, “Topic Modeling On Online News.Portal Using Latent Dirichlet Allocation (Lda),” *Ijccs (Indonesian J. Comput. Cybern. Syst.*, Vol. 16, No. 4, P. 335, Oct. 2022, Doi: 10.22146/Ijccs.74383.
- [15] L. W. Narendra, “Topic Modeling In Conversational Dialogs For Naming Intent Labels Using Lda Pemodelan Topik Pada Dialog Percakapan Untuk Penamaan Label Intent,” 2022.