

# Information Retrieval dan Perhitungan Kemiripan Dokumen pada Indonesian Heritage Library Menggunakan Vector Space Model

**Ginabila**

Program Studi Sistem Informasi, Universitas Bina Sarana Informatika

<sup>1</sup>Jalan Kramat Raya, Kota Jakarta Selatan

Email: [gina.gnb@bsi.ac.id](mailto:gina.gnb@bsi.ac.id)

## **Abstrak**

*Kebutuhan user untuk mencari suatu kumpulan atau pangkalan informasi secara otomatis saat ini sudah menjadi hal yang sering dilakukan, untuk memenuhi kebutuhan user menemukan kembali informasi-informasi yang dibutuhkan tersebut maka information retrieval system digunakan. Pencarian dokumen yang dilakukan oleh user pada sebuah database dengan cara menginputkan nama dokumen, maka semua dokumen dengan judul yang hampir mendekati dokumen yang user maksud akan ditampilkan. Hal ini dikarenakan dalam sistem pencarian tersebut, sistem belum dapat mengukur mana dokumen yang paling sesuai yang harus ditampilkan dan yang dimaksud oleh user. Maka dengan masalah seperti ini penulis menggunakan information retrieval. Dalam penelitian ini akan dilakukan perhitungan kemiripan dokumen menggunakan metode Vector Space Model. Dalam metode ini data akan melalui proses token dan indexing sehingga tingkat ketepatan dokumen yang dimaksud oleh user untuk temu kembali informasi akan lebih sesuai.*

**Kata Kunci:** *Information Retrieval, Vector Space Model, Perhitungan Kemiripan Dokumen, Perpustakaan Digital.*

## **Abstract**

*The requirements from users to search for a collection or base of information automatically has now become something that is often done, to meet the needs of the user to find the required information again, the information retrieval system is used. Document searches performed by the user in a database by inputting the name of the document, then all documents with a title that is almost close to the document that the user intended will be displayed. This is because in the search system has not been able to measure which documents are the most suitable that must be displayed and what is meant by the user. So with a problem like this the writer uses information retrieval. In this research, the document similarity calculation will be carried out using the Vector Space Model method. In this method, the data will go through a token and indexing process so the level of accuracy of the documents intended by the user for information retrieval will be more appropriate.*

**Keywords:** *Information Retrieval, Vector Space Model, Document Similarity Calculation, Digital Library.*

## **1. PENDAHULUAN**

Perkembangan ilmu pengetahuan dan teknologi membawa perubahan dalam berbagai bidang, salah satunya adalah maraknya pengembangan perpustakaan digital di Indonesia. Maraknya sistem perpustakaan digital di Indonesia ini tentunya sangat berpengaruh dalam dunia pendidikan. Dunia pendidikan akan sangat terbantu dengan modernisasi dari perpustakaan konvensional.

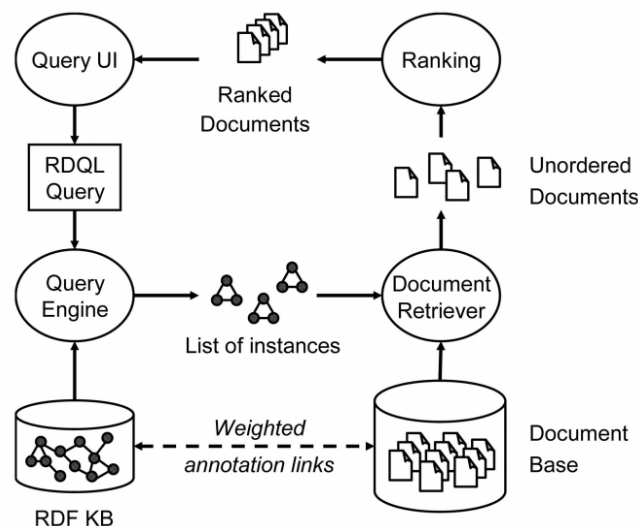
Perpustakaan adalah sarana yang sangat penting (vital) dalam upaya mencerdaskan kehidupan bangsa. Perpustakaan dari masa ke masa mengalami perkembangan yang signifikan sesuai kemajuan zaman dan kebutuhan penggunaannya. Paradigma perpustakaan yang kini berkembang yaitu dari fisik ke akses, memungkinkan perpustakaan untuk membantu mewujudkan visi perguruan tinggi mencapai taraf internasional. Saat ini perpustakaan digital semakin banyak dibicarakan. Hal tersebut terjadi karena arus globalisasi dan tingkat kebutuhan masyarakat yang semakin tinggi dalam mengakses informasi. Masyarakat semakin kritis dan ingin mengakses informasi secara cepat, tepat, akurat dan tentunya mudah. Solusinya dapat terpenuhi dengan mengakses informasi di Perpustakaan Digital.<sup>[1]</sup>

Sistem perpustakaan digital adalah penerapan teknologi informasi sebagai sarana untuk menyimpan, mendapatkan dan menyebarkan informasi ilmu pengetahuan dalam format digital. Atau secara sederhana dapat dianalogikan sebagai tempat menyimpan koleksi perpustakaan yang sudah dalam bentuk digital.<sup>[2]</sup> Perpustakaan digital memiliki sejarah yang pendek namun bergejolak dan meledak. Sejumlah visioner awal, seperti Licklider berpendapat bahwa perpustakaan di masa depan sangat inovatif dan berbeda strukturnya.<sup>[3]</sup>

Dalam penggunaan sistem perpustakaan digital untuk mencari dokumen yang dimaksud, user yang menggunakan diharuskan untuk menginput nama dokumen yang dicari dalam kolom pencarian sistem perpustakaan digital tersebut dengan memasukkan keyword berupa judul buku, penulis, atau tema dari dokumen yang dibutuhkan.

Masalah utama dalam proses pencarian dokumen digital dibutuhkan waktu yang relatif lama karena pencariannya harus menyertakan isi judul dokumen secara lengkap dan benar, hal ini menjadi tidak relevan dalam sistem temu kembali informasi. Maka dari itu dibutuhkan sebuah *search engine* yang dapat mencari dokumen-dokumen tersebut secara lebih cepat dan mudah serta menghasilkan informasi yang relevan tanpa perlu menyertakan judul dokumen secara lebih terperinci.<sup>[4]</sup>

Dalam kasus ini jika pengguna perpustakaan digital tidak menyertakan judul dokumen secara terperinci maka hasil dari dokumen yang ditampilkan juga tidak akan maksimal dan dapat kurang sesuai dengan yang pengguna maksud. Untuk itu penulis menggunakan metode *vector space model* yang saat ini berperan penting untuk penelitian yang berhubungan dengan temu kembali informasi. Dalam penggunaannya untuk mendeteksi *similarity* yang diinput oleh pengguna dalam pencarian perpustakaan digital, metode ini digunakan sebagai representasi dari kumpulan dataset dokumen teks. *Vector space model* akan memberi bobot setiap kata pada seluruh dokumen yang nantinya akan digunakan untuk mengukur kepentingan kata tersebut dalam suatu dokumen.



Gambar 1. Proses Pencarian Dokumen Berbasis Teks<sup>[5]</sup>

Secara statistik dalam metode *Vector Space Model*, dokumen secara konseptual akan diwakili oleh vektor kata kunci yang diekstrak dari dokumen, dengan bobot terkait yang mewakili pentingnya kata kunci di dalam dokumen dan di dalam keseluruhan koleksi dokumen. Query juga berperan penting dalam metode ini, disini query dimodelkan sebagai daftar kata kunci dengan bobot terkait yang mewakili pentingnya kata kunci dalam dokumen yang dicari.<sup>[6]</sup>

## 2. METODE PENELITIAN

Metode penelitian merupakan rangka yang dirancang dan diimplementasikan oleh peneliti dengan tujuan untuk mendapatkan sebanyak mungkin informasi atau data dan melakukan penelitian pada data yang telah dikumpulkan tersebut. Metode penelitian dapat digunakan untuk memberikan penjelasan atau gambaran yang meliputi prosedur yang harus dilakukan dan langkah-langkah yang harus ditempuh dan berisi informasi sumber data yang didapatkan dalam penelitian serta bagaimana data tersebut diperoleh untuk selanjutnya digunakan dalam penelitian.

Pada penelitian ini penulis menggunakan metode *Vector Space Model* dengan melakukan perhitungan pada contoh dokumen yang ada. Metode ini akan memberikan bobot pada kata yang ada di setiap dokumen dan nantinya akan dibuat *ranking* untuk memberikan informasi dokumen yang paling relevan atau yang similiaritas nya paling tinggi dan paling dekat dengan *query* yang telah diinput oleh pengguna perpustakaan digital.

Untuk melakukan penelitian ini, penulis melakukan pengumpulan data-data yang dibutuhkan. Pengumpulan data yang ditempuh untuk mendapatkan sumber data yang akurat dalam penelitian ini antara lain:

1) Observasi

Dalam penelitian ini penulis melakukan pengamatan pada beberapa sistem perpustakaan digital yang telah ada dan dapat diakses oleh pengguna umum untuk memilih mana data yang paling baik untuk digunakan dalam penelitian.

2) Studi Pustaka

Penulis juga mencari informasi tentang penelitian-penelitian sebelumnya dengan lingkup yang hampir sama untuk menambah dan membandingkan referensi penelitian baik dalam bentuk jurnal maupun buku yang berkaitan dengan penelitian dan dengan metode *Vector Space Model*.

Berdasarkan teknik yang dilakukan dalam penelitian untuk mengumpulkan sumber data atau informasi yang telah dijabarkan oleh penulis sebelumnya, maka penulis memutuskan untuk mengambil 10 *sample* judul dokumen pada tema *Indonesian Heritage Digital Library* untuk selanjutnya akan dilakukan perhitungan dan pemberian bobot pada dokumen-dokumen tersebut. Sebelum dilakukan perhitungan, penulis mengolah data tersebut melalui tahapan *tokenizing*, *filtering* dan *stemming* untuk memaksimalkan perhitungan pada data.

Setelah dilakukan *preprocess* pada data dengan melakukan *tokenizing*, *filtering* dan *stemming*, selanjutnya penulis melakukan *indexing* yang berfungsi untuk memisahkan masing-masing kata dalam setiap nama dokumen untuk nantinya dilakukan proses perhitungan. Setelah proses *indexing* selesai, maka proses perhitungan akan dimulai. Langkah-langkah yang dilakukan dalam proses perhitungan antara lain:

1) Mengitung *Term Frequency* (TF)

Nilai TF didapatkan dengan cara mengitung banyaknya kemunculan term *i* pada dokumen *j*.

$$tf = tf_{(i,j)} \dots \dots \dots (1)$$

Keterangan:

- tf = *Term Frequency* (Frekuensi banyaknya muncul term)
- tf<sub>(i,j)</sub> = Banyaknya kemunculan *term i* (t<sub>i</sub>) dalam dokumen *j* (d<sub>j</sub>)

2) Menghitung TF-IDF (*Term Frequency – Inverse Document Frequency*)

Perhitungan TF-IDF digunakan untuk mendapatkan bobot hasil perkalian atau kombinasi antara *tf* dan *idf*.

$$W_{ij} = tf_i \cdot \log \left( \frac{N}{df_i} \right) \dots \dots \dots (2)$$

Keterangan:

- W<sub>ij</sub> = Bobot Dokumen
- tf<sub>i</sub> = Banyaknya kemunculan *term i* pada dokumen *j*
- N = Jumlah dokumen yang terambil
- df<sub>i</sub> = Banyaknya dokumen dalam koleksi dimana *term i* muncul di dalamnya.

3) Perhitungan Jarak dokumen dengan *query*

Jarak *query* dihitung untuk mendapatkan jarak *query* dari bobot dokumen.

$$|q| = \sqrt{\sum_{j=1}^t (w_{i,q})^2} \dots \dots \dots (3)$$

Keterangan:

- |q| = Jarak *Query*
- (w<sub>i,q</sub>) = Bobot *Query* Dokumen ke – *i*

Jarak dokumen adalah bobot dokumen ke- i yang dihitung untuk mendapatkan jarak dokumen dari bobot dokumen yang terambil oleh sistem.

$$|d_j| = \sqrt{\sum_{i=1}^t (w_{ij})^2} \dots\dots\dots(4)$$

Keterangan:

- |d<sub>j</sub>| = Jarak Dokumen
- (w<sub>ij</sub>) = Bobot Dokumen ke – i

4) Perhitungan Similiaritas dari Query 1 dan 2

$$sim(q, d_j) = \frac{q \cdot d_j}{|q| \cdot |d_j|} = \frac{\sum_{i=1}^t w_{iq} \cdot w_{ij}}{\sqrt{\sum_{j=1}^t (w_{iq})^2 * \sum_{i=1}^t (w_{ij})^2}} \dots\dots\dots(4)$$

Keterangan:

- sim(q,dj) = Similiaritas antara query dan dokumen
- q = Query
- dj = Bobot dokumen
- (|q|) = Akar jumlah kuadrat q
- (|dj|) = Akar jumlah kuadrat dokumen

### 3. HASIL DAN PEMBAHASAN

Penulis mengumpulkan data dari perpustakaan digital *Indonesian Heritage Library* dengan mengambil sampel 10 dokumen untuk kemudian dilakukan preprocess dan dilakukan perhitungan dengan metode *Vector Space Model*. Dokumen-dokumen yang diambil dari perpustakaan digital adalah dengan judul-judul berikut;

- 1) *Early Views of Indonesia* : Pemandangan Indonesia di Masa Lampau
- 2) *Golden Letters : Writing Traditions of Indonesia*. Surat Emas Budaya Tulis di Indonesia.
- 3) Ragam Hias Pada Busana dan Peralatan Kesenian Tradisional
- 4) Permainan Tradisional Indonesia
- 5) Album Kerajinan Tradisional
- 6) Pustaka Wisata Budaya Perahu Tradisional Nusantara
- 7) Pameran Keliling Seni Topeng Nusantara
- 8) Adat dan Upacara Perkawinan Daerah Sumatera Utara
- 9) *Early Views of Indonesia* : Pemandangan Indonesia di Masa Lampau
- 10) Arti Lambang dan Fungsi Tata Rias Pengantin Dalam Menanamkan Nilai-Nilai Budaya Provinsi Daerah Istimewa Yogyakarta

Setelah melalui tahapan *preprocessing* yaitu *tokenizing*, *filtering*, dan *stemming* pada dokumen, selanjutnya penulis melakukan *indexing* pada tiap nama dokumen dan diperoleh hasil *indexing* sebagai berikut :

Tabel 1. Indexing Dokumen

Doc 1	Doc 2	Doc 3	Doc 4	Doc 5	Doc 6	Doc 7	Doc 8	Doc 9	Doc 10
Early	Budaya	Alat	Indonesia	Album	Budaya	Keliling	Adat	Early	Arti
Indonesia	Emas	Busana	Main	Rajin	Nusantara	Nusantara	Daerah	Indonesia	Budaya
Indonesia	Golden	Hias	Tradisional	Tradisional	Perahu	Pameran	Kawin	Indonesia	Daerah
Lampau	Indonesia	Ragam			Pustaka	Seni	Sumatra	Lampau	Fungsi
Masa	Letter	Seni			Tradisional	Topeng	Upacara	Masa	Istimewa
Pandang	Surat	Tradisional			Wisata		Utara	Pandang	Lambang
View	Traditions							Views	Nilai
	Tulis								Pengantin
	Writing								Provinsi
									Rias
									Tanam
									Tata
									Yogyakarta

Tabel 1 diatas menunjukkan bahwa, hasil *indexing* setelah dokumen melalui proses *preprocessing* dan *indexing* maka nama dokumen akan dipisahkan per-kata dan nama dokumen yang

terdapat diatas telah dihilangkan imbuhan nya serta kata-kata yang tidak penting dalam dokumen juga dibuang dalam proses *stoplist* di *preprocessing*. Setelah tahap *indexing* selesai, kemudian data tersebut melalui proses perhitungan, berikut adalah tahap tahap perhitungan yang dilakukan:

### 3.1. Perhitungan TF (*Term Frequency*)

Setelah proses *indexing* selesai, maka data tersebut akan melalui proses penghitungan TF. Untuk melakukan perbandingan perhitungan yang nanti nya akan didapatkan, maka dalam penelitian ini penulis mengambil 2 sampel *query* yaitu “Budaya Indonesia” dan “Ragam Seni Nusantara Daerah Yogyakarta”. Berikut adalah hasil perhitungan dari TF:

Tabel 2. Hasil perhitungan dari *Term Frequency*

Token	Q1	Q2	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Df
Adat	0	0	0	0	0	0	0	0	0	1	0	0	1
Album	0	0	0	0	0	0	1	0	0	0	0	0	1
Arti	0	0	0	0	0	0	0	0	0	0	0	1	1
Budaya	1	0	0	1	0	0	0	1	0	0	0	1	3
Busana	0	0	0	0	1	0	0	0	0	0	0	0	1
Daerah	0	1	0	0	0	0	0	0	0	1	0	1	2
Early	0	0	1	0	0	0	0	0	0	0	1	0	2
Emas	0	0	0	1	0	0	0	0	0	0	0	0	1
Fungsi	0	0	0	0	0	0	0	0	0	0	0	1	1
Golden	0	0	0	1	0	0	0	0	0	0	0	0	1
Hias	0	0	0	0	1	0	0	0	0	0	0	0	1
Indonesia	1	0	2	2	0	1	0	0	0	0	2	0	7
Istimewa	0	0	0	0	0	0	0	0	0	0	0	1	1
Kawin	0	0	0	0	0	0	0	0	0	1	0	0	1
Keliling	0	0	0	0	0	0	0	0	1	0	0	0	1
Lambang	0	0	0	0	0	0	0	0	0	0	0	1	1
Lampau	0	0	1	0	0	0	0	0	0	0	1	0	2
Letter	0	0	0	1	0	0	0	0	0	0	0	0	1
Main	0	0	0	0	0	1	0	0	0	0	0	0	1
Masa	0	0	1	0	0	0	0	0	0	0	1	0	2
Nilai	0	0	0	0	0	0	0	0	0	0	0	1	1
Nusantara	0	1	0	0	0	0	0	1	1	0	0	0	2
Pameran	0	0	0	0	0	0	0	0	1	0	0	0	1
Pandang	0	0	0	0	0	0	0	0	0	0	1	0	1
Pengantin	0	0	0	0	0	0	0	0	0	0	0	1	1
Perahu	0	0	0	0	0	0	0	1	0	0	0	0	1
Provinsi	0	0	0	0	0	0	0	0	0	0	0	1	1
Pustaka	0	0	0	0	0	0	0	1	0	0	0	0	1
Ragam	0	1	0	0	1	0	0	0	0	0	0	0	1
Rajin	0	0	0	0	0	0	1	0	0	0	0	0	1
Rias	0	0	0	0	0	0	0	0	0	0	0	1	1
Seni	0	1	0	0	1	0	0	0	1	0	0	0	2
Sumatra	0	0	0	0	0	0	0	0	0	1	0	0	1
Surat	0	0	0	1	0	0	0	0	0	0	0	0	1
Tanam	0	0	0	0	0	0	0	0	0	0	0	1	1
Tata	0	0	0	0	0	0	0	0	0	0	0	1	1
Topeng	0	0	0	0	0	0	0	0	1	0	0	0	1
Tradisional	0	0	0	0	1	1	1	1	0	0	0	0	4
Traditions	0	0	0	1	0	0	0	0	0	0	0	0	1
Tulis	0	0	0	1	0	0	0	0	0	0	0	0	1
Upacara	0	0	0	0	0	0	0	0	0	1	0	0	1
Utara	0	0	0	0	0	0	0	0	0	1	0	0	1
View	0	0	1	0	0	0	0	0	0	0	1	0	2
Wisata	0	0	0	0	0	0	0	1	0	0	0	0	1
Writing	0	0	0	1	0	0	0	0	0	0	0	0	1
Yogyakarta	0	1	0	0	0	0	0	0	0	0	0	1	1

Tabel 2 menunjukkan hasil dari frekuensi kemunculan dari semua *term* yang ada diatas. Semakin besar jumlah *term* yang ada dalam dokumen, maka nanti nilai bobot yang didapat akan semakin besar dan nantinya akan memberikan nilai kesesuaian yang semakin besar. Pada tabel 2 *term* yang paling banyak muncul adalah *term* “Tradisional” dengan 4 frekuensi kemunculan.

3.2. Perhitungan TF-IDF (*Term Frequency-Inverse Document Frequency*)

Perhitungan yang dilakukan setelah mendapatkan nilai dokumen frekuensi pada perhitungan sebelumnya, maka selanjutnya akan dilakukan perhitungan TF-IDF. Berikut adalah hasil yang didapat dari perhitungan TF-IDF:

Tabel 3. Hasil Perhitungan TF-IDF

Idf	tf*idf												
	log(D/df)	Q1	Q2	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
1	0	0	0	0	0	0	0	0	0	0	1	0	0
1	0	0	0	0	0	0	0	1	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	1
0.522879	1	0	0	1	0	0	0	0	1	0	0	0	1
1	0	0	0	0	1	0	0	0	0	0	0	0	0
0.69897	0	1	0	0	0	0	0	0	0	0	1	0	1
0.69897	0	0	1	0	0	0	0	0	0	0	0	1	0
1	0	0	0	1	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	0	0	1	0	0	0	0	0	0	0	0	0
1	0	0	0	0	1	0	0	0	0	0	0	0	0
0.154902	1	0	00.07	00.07	0	1	0	0	0	0	0	0,485	0
1	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	0	0	0	0	0	0	0	0	0	1	0	0
1	0	0	0	0	0	0	0	0	0	1	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	1
0.69897	0	0	1	0	0	0	0	0	0	0	0	1	0
1	0	0	0	1	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	1	0	0	0	0	0	0	0
0.69897	0	0	1	0	0	0	0	0	0	0	0	1	0
1	0	0	0	0	0	0	0	0	0	0	0	0	1
0.69897	0	1	0	0	0	0	0	0	1	1	0	0	0
1	0	0	0	0	0	0	0	0	0	1	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	1	0
1	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	0	0	0	0	0	0	0	1	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	0	0	0	0	0	0	0	0	0	0	0	1
0.39794	0	0	0	0	1	1	1	1	1	0	0	0	0
1	0	0	0	1	0	0	0	0	0	0	0	0	0
1	0	0	0	1	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	1	0	0	0
1	0	0	0	0	0	0	0	0	0	0	1	0	0
0.69897	0	0	1	0	0	0	0	0	0	0	0	1	0
1	0	0	0	0	0	0	0	1	0	0	0	0	0
1	0	0	0	1	0	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	0	0	0	0	1

Tabel 3 menunjukkan bobot yang didapatkan dari setiap term yang dihitung berdasarkan metode TF-IDF.

3.3. Perhitungan Jarak Dokumen dengan *Query*

Setelah bobot dari *term* didapatkan pada perhitungan TF-IDF, selanjutnya penulis melakukan penghitungan jarak antara dokumen dan *query*. Berikut adalah hasil dari perhitungan jarak antara dokumen dengan *query*:









Dari tabel 7 diatas dapat diketahui nilai *similarity* dari yang paling rendah sampai yang paling rendah yang didapatkan dari masing-masing dokumen berdasarkan pengujian dari sampel *query* 1 dan *query* 2. Tabel diatas juga mengurutkan ranking *query* 1 dan *query* 2 pada masing-masing dokumen.

#### 4. KESIMPULAN

Berdasarkan hasil penelitian yang telah penulis lakukan dengan menggunakan metode *Vector Space Model*, maka kesimpulan yang diperoleh adalah sebagai berikut:

- 1) Dari serangkaian langkah-langkah perhitungan yang dilakukan dengan menerapkan sampel *query* 1 dan *query* 2 pada data dokumen yang dikumpulkan dari *Indonesian Heritage Library* maka penelitian ini mendapatkan hasil bahwa, hasil *similarity* tertinggi dari uji coba *query* 1 yang paling cocok adalah dari dokumen 4 dengan judul “Permainan Tradisional Indonesia”. Sedangkan untuk uji coba *query* 2, hasil *similarity* tertinggi yang paling cocok adalah dokumen 3 dan dokumen 7 yaitu “Ragam Hias Pada Busana dan Peralatan Kesenian Tradisional” dan “Pameran Keliling Seni Topeng Nusantara”.
- 2) Dalam menyelesaikan *Information Retrieval* pada *Indonesian heritage Digital Library*, metode *Vector Space Model* terbukti dapat menyelesaikan masalah tersebut. Metode ini dapat dengan akurat membantu menampilkan dokumen-dokumen sesuai yang diinginkan oleh pengguna *digital library* tanpa harus memasukkan *query* yang sangat terperinci sehingga tidak menyulitkan pengguna.

#### DAFTAR PUSTAKA

- [1] Prabowo, T. T. (2013). Mengenal Perpustakaan Digital. *Ilmu Perpustakaan Dan Informasi*, 8(1), 108–120.
- [2] Subroto, G. (2010). Perpustakaan Digital. *Jurnal Pustakawan Indonesia*, 10(2), 1–11.
- [3] Saracevic, T. (2000). Digital library evaluation: Toward an evolution of concepts. *Library Trends*, 49(2), 350–369.
- [4] Zain, M. Y., & Suswati. (2016). *INFORMATION RETRIEVAL SYSTEM PADA PENCARIAN FILE DOKUMEN BERBASIS TEKS DENGAN METODE VECTOR SPACE MODEL DAN ALGORITMA ECS STEMMER*. Vol. 1, No.
- [5] Castells, P., Fernández, M., & Vallet, D. (2007). An adaptation of the vector-space model for ontology-based information retrieval. *IEEE Transactions on Knowledge and Data Engineering*, 19(2), 261–272. <https://doi.org/10.1109/TKDE.2007.22>
- [6] Lee, D. L., Chuang, H., & Seamons, K. (1997). Document ranking and the vector-space model. *IEEE Software*, 14(2), 67–75. <https://doi.org/10.1109/52.582976>