

Relevansi Ayat al-Quran Secara Tematik Menggunakan Pendekatan *Graph-Based Knowledge* dan *Lexical-Search*

M. Abd. Rouf¹, Abd. Charis Fauzan^{2,*}

¹Program Studi Ilmu Al-Qur'an dan Tafsir, Universitas Nahdlatul Ulama Blitar, Indonesia

²Program Studi Ilmu Komputer, Universitas Nahdlatul Ulama Blitar, Indonesia

¹abdrouf@unublitar.ac.id; ²abdcharis@unublitar.ac.id

* penulis korespondensi

INFO ARTIKEL

Sejarah Artikel

Diterima: 15 Desember 2022

Direvisi: 6 Maret 2022

Diterbitkan: 30 April 2023

Kata Kunci

Tematik

Ayat al-Quran

Kemiripan leksikal

Graph-Based Knowledge

ABSTRAK

Tematik Al-Qur'an merupakan berbagai kumpulan tema yang ada pada Al-Qur'an, yang mana bertujuan untuk menghubungkan satu tema dengan tema lainnya. Pada dasarnya 1 ayat dapat digolongkan dalam beberapa tema yang berbeda. Pengumpulan kelompok tema yang berbeda ini menunjang salah satu metode tafsir, yakni tafsir maudhu'i. Yaitu metode tafsir yang berusaha mencari jawaban al-Qur'an dengan cara mengumpulkan ayat-ayat al-Qur'an yang mempunyai tujuan yang satu, yang bersama-sama membahas topik/judul tertentu dan menertibkannya sesuai dengan masa turunnya dan selaras dengan sebab-sebab turunnya, kemudian memperhatikan ayat-ayat tersebut dengan penjelasan-penjelasan, keterangan-keterangan dan hubungan-hubungannya dengan ayat-ayat yang lain, kemudian mengistimbatkan hukum-hukum. Pada saat ini, sebagai penjunjang tema-tema dalam tafsir maudhu'i, dibutuhkan sebuah media yang bisa memberikan kemudahan pengguna untuk mencari berbagai macam tema yang ada pada Al-Qur'an, dan mengetahui hubungan antar tema dan hubungan ayat berdasarkan relasinya, yang mana relasinya berupa ayat dengan tema ataupun tema dengan tema. Metode knowledge graph secara garis besar merupakan implementasi data yang direalisasikan menggunakan visualisasi grafik dimana semua relasi antara tema akan dipaparkan. Penelitian ini mengusulkan bahwa knowledge graph akan berpotensi sebagai penunjang tafsir maudhu'i bagi para pengguna yang dalam hal ini adalah para ahli tafsir Al-Qur'an. Hasil dari uji coba dilakukan menunjukkan rata-rata *f-measure* yaitu 64.2% pada keseluruhan *query*, serta 63.1% dan 65.7% pada *short query* dan *long query*.

PENDAHULUAN

Al-Qur'an secara tematik merupakan berbagai kumpulan tema yang ada pada Al-Qur'an, yang mana bertujuan untuk menghubungkan satu tema dengan tema lainnya. Pada dasarnya 1 ayat dapat digolongkan dalam beberapa tema yang berbeda. Pengumpulan kelompok tema yang berbeda ini menunjang salah satu metode tafsir, yakni tafsir maudhu'i. Yaitu metode tafsir yang berusaha mencari jawaban al-Qur'an dengan cara mengumpulkan ayat-ayat al-Qur'an yang mempunyai tujuan yang satu, yang bersama-sama membahas topik/judul tertentu dan menertibkannya sesuai dengan masa turunnya dan selaras dengan sebab-sebab turunnya, kemudian memperhatikan ayat-ayat tersebut dengan penjelasan-penjelasan, keterangan-keterangan dan hubungan-hubungannya dengan ayat-ayat yang lain, kemudian mengistimbatkan hukum-hukum.

Pada saat ini, sebagai penjunjang tema-tema dalam tafsir maudhu'i, dibutuhkan sebuah media yang bisa memberikan kemudahan pengguna untuk mencari berbagai macam tema yang ada pada Al-Qur'an, dan mengetahui hubungan antar tema dan hubungan ayat berdasarkan relasi nya, yang mana relasi nya berupa ayat dengan tema dan tema dengan tema. Metode knowledge graph secara garis besar merupakan implementasi data yang direalisasikan menggunakan visualisasi grafik dimana semua relasi antara tema akan dipaparkan. Dalam perancangan media visualisasi tematik Al-Qur'an ini, peneliti mengkaji beberapa penelitian sebelumnya yang berhubungan dengan implementasi teknologi knowledge graph dan visualisasi Al-Qur'an. Peneliti sebelumnya telah merancang aplikasi pembelajaran membaca Alqur'an dengan metode qiro'ah berbasis android [1]. Aplikasi pembelajaran membaca Al- Qur'an melalui smartphone tersebut dirasakan masih kurang

efektif karena hanya bisa diakses menggunakan android dengan versi minimal 2.3 (Gingerbread) sehingga tidak semua orang dapat mengaksesnya. Selain itu, aplikasi tersebut hanya memiliki fitur membaca Al-Qur'an saja.

Beberapa peneliti sebelumnya telah berupaya untuk membangun media representasi pengetahuan (knowledge representation) yang bertujuan mempelajari dan mengambil pengetahuan dari Al-Qur'an dengan membangun ontologi Al-Qur'an [2]–[5] dan aplikasi semantik web Al-Qur'an [6]. Peneliti lain juga menggunakan terjemahan bahasa Inggris tafsir Al-Qur'an untuk membangun ontologi tafsir Al-Qur'an [7]. Namun, penelitian-penelitian tersebut hanya terbatas pada pemodelan ontologi saja dan belum pada tahap implementasi menghasilkan aplikasi visualisasi yang dapat digunakan untuk mempelajari Al-Qur'an. Dalam penelitian ini, penulis tidak membuat ontologi baru, penulis hanya menggunakan data tematik Al-Qur'an yang diambil dari Al-Qur'an Amazing (Cordoba), namun, penelitian ini menghasilkan aplikasi visualisasi Al-Qur'an yang dapat digunakan dengan mudah.

Knowledge Graph juga digunakan dalam domain lain, seperti pendidikan [8] dan linguistik. Penelitian linguistic sebelumnya membuktikan bahwa pengelompokan bahasa daerah Indonesia secara tradisional dengan hierarchical clustering dan k-means clustering dianggap memiliki kelemahan dalam bidang fleksibilitas eksplorasi [9]. Untuk meningkatkan kemudahan dalam menelusuri kekerabatan antar bahasa daerah di Indonesia knowledge graph digunakan dalam visualisasi pengelompokan kekerabatan bahasa daerah Indonesia menggunakan pendekatan berbasis grafik untuk membuat dan memvisualisasikan pengelompokan kekerabatan leksikal bahasa daerah dengan memanfaatkan pangkalan data Automated Similarity Judgment Program (ASJP) untuk menghasilkan matrik kekerabatan bahasa daerah, dan kemudian merumuskan data dalam bentuk grafik [10]. Penelitian-penelitian tersebut menjadi motivasi bagi peneliti bahwa knowledge graph akan berpotensi sebagai penunjang tafsir maudhu'i bagi para pengguna yang dalam hal ini adalah para ahli tafsir Al-Qur'an.

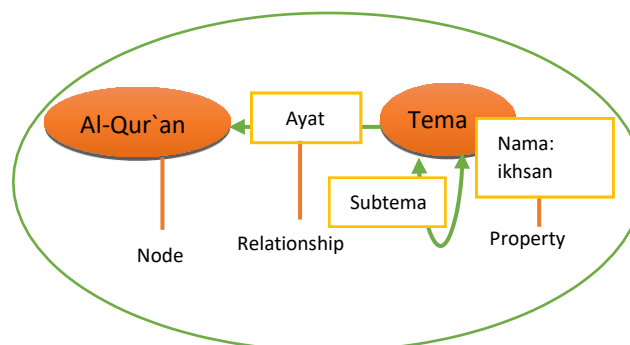
METODE

Pengumpulan Data

Pada media visualisasi tematik ini, data diambil dari Quran Kemenag yang dapat diakses dari situs quran.kemenag.go.id. Data yang diambil adalah Teks ayat al-Quran berbahasa arab dan terjemah berbahasa indonesia.

Perancangan Neo4J sebagai Basis Visualisasi Tematik al-Quran

Neo4j merupakan suatu platform grafik yang mengambil pendekatan koneksi-pertama ke data. Pendekatan ini memperluas kemampuan pengguna mengenali pentingnya menjaga hubungan dan koneksi melalui setiap transisi eksistensi: dari ide, untuk merancang dalam model logis, untuk implementasi dalam model fisik, untuk operasi menggunakan bahasa kueri dan untuk persistensi dalam meningkatkan skala, suatu sistem basis data harus bisa diandalkan. Dasar representasi data yang terhubung dikenal dengan basis data berbentuk grafik. Permodelan basis data grafik terdiri dari empat bagian sebagaimana Gambar 1.

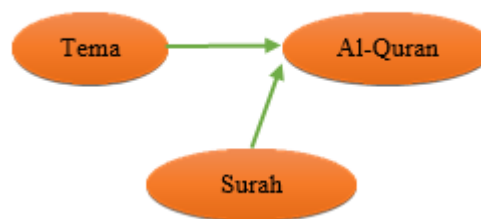


Gambar 1 Model Grafik Neo4j

- **Node.** Node merupakan elemen data utama yang saling terhubung ke node lain. Node dapat memiliki satu atau lebih properti (yaitu, atribut yang disimpan sebagai pasangan kunci / nilai). Node memiliki satu atau lebih label yang menggambarkan perannya dalam grafik, contoh: Node Al-Qur'an, Tema, dll.
- **Relationships.** Relationships adalah proses menghubungkan dua node bersifat dan terarah. Node dapat memiliki banyak hubungan, bahkan rekursif. Relationships dapat memiliki satu atau lebih properti (yaitu, atribut yang disimpan sebagai pasangan kunci / nilai). Contoh relationships pada Gambar 1 ialah Ayat dan Subtema.
- **Properties.** Properties merupakan tempat penyimpanan data (metadata) tentang Node ataupun Relationship berupa pasangan nama properti dan nilai properti. Nilai pada suatu Node berupa string, dapat diindeks dan dibatasi. Contoh properties dari node Tema ialah Nama: ikhsan.
- **Label.** Label digunakan untuk mengelompokkan node ke dalam himpunan/grup, yang mana label ini bagian dari nodes. Label diindeks untuk mempercepat pencarian node dalam grafik.

Perancangan Graph Database

Graph database merupakan suatu model database untuk penyimpanan menggunakan prinsip teori grafik, dimana semua data dilambangkan dengan node dan setiap node saling berkaitan yang sering disebut dengan edge. Adapun pemodelan node pada basis data grafik ini ditunjukkan pada Gambar 2



Gambar 2. Pemodelan Node

1. Tema merupakan *node* yang menampung semua data berbagai macam data tema dengan satu properti yaitu "nama".
2. Al-Qur'an merupakan *node* yang menampung semua data Al-Quran dengan lima properti yang ditunjukkan pada Tabel I.

No	Properti
1	Kode_Index
2	Kode_surat
3	Ayat
4	Bhs_arab
5	Bhs_Indonesia

3. Surah merupakan *node* yang menyimpan semua dataterjemahan surat dari Al-Qur'an dengan lima properti yang ditunjukkan pada Tabel II.

No	Property
1	Kode_surat
2	Nama_surat
3	Arti_Surat

4	Jumlah_ayat
5	Tempat_turun

Pembentukan Cypher

Cypher merupakan perintah yang digunakan pada Neo4j, yang mana teknik yang paling berguna dari bahasa kueri ini adalah menggunakan pencocokan pola antara kueri dan pola serupa yang ada dalam grafik. Fitur *cypher load* digunakan dalam penginputan data dengan bentuk inputan file dalam format csv ke basis data neo4j. Sedangkan dalam pencarian data dengan jarak tertentu digunakan fitur *cypher hops*.

1) Cypher Load CSV

Pada media ini *cypher load* digunakan untuk proses penginputan data sekaligus untuk mengelompokkan semuanya yang ada. Tahapan pengelompokan sebagai berikut:

- Load Tema.

Pada tahapan load tema ini, semua data tema akan dimasukkan pada label tema, seperti yang ditunjukkan pada Gambar 3.

```
[LOAD CSV WITH HEADERS FROM 'http://localhost/neo4j/tema.csv' AS row CREATE (:Tema {Name: row.tema})]
```

Gambar. 3 Load Tema.

- Load Subtema

Load subtema merupakan tahapan yang dilakukan untuk mencari hubungan antar berbagai tema yang ada, seperti yang ditunjukkan pada Gambar 4.

```
[LOAD CSV WITH HEADERS FROM "http://localhost/neo4j/tematik2.csv" AS row MATCH (t:Tema), (st:Tema) WHERE t.Name = row.tema AND st.Name = row.subtema CREATE (t)-[:Subtema]->(st)]
```

Gambar. 4 Load Subtema.

- Load Tematik

Load tematik merupakan tahapan yang dilakukan menghubungkan antara tema dengan Al-Qur'an berdasarkan key yang ada pada Al-Qur'an berupa *kode_index*, seperti yang ditunjukkan pada Gambar 5.

```
[LOAD CSV WITH HEADERS FROM "http://localhost/neo4j/tematik.csv" AS row MATCH (t:Tema), (st:Tema), (k1:AlQuran) WHERE t.Tema = row.tema AND st.Tema = row.subtema AND k1.Kode_Index = row.tematik]
```

Gambar. 5 Load Tematik.

2) *Cypher Hops*.

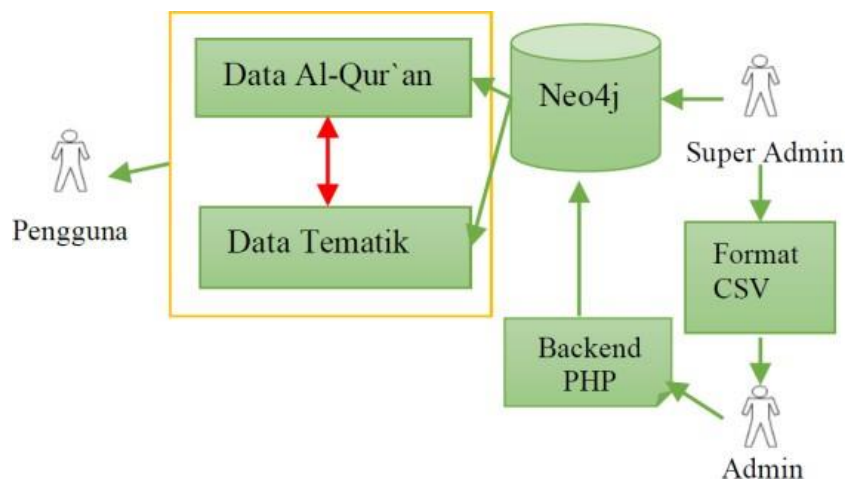
Cypher hops merupakan suatu kode *cypher* yang tersedia di platform Neo4j yang berguna untuk mencari semua relasi data yang terhubung dengan jarak hubungan yang diinginkan, seperti yang ditunjukkan pada Gambar 6.

```
[match (a:AlQuran {Surat:"2", Ayat:"177"})-
[r*1..3]-
(c) return c,a,]
```

Gambar. 6 Cypher Hops.

Pengembangan Visualisasi Tematik

Dalam membangun atau mengembangkan media visualisasi ini tentunya diperlukan perancangan agar media yang dibangun berjalan dengan baik. Alur kerja media visualisasi tematik Al-Quran berbasis knowledge graph dapat dilihat pada Gambar 7.



Gambar 7. Pengembangan Visualisasi Tematik

Proses aliran data yang terjadi pada media visualisasi diawali dengan konversi data file csv yang diinputkan oleh admin yang kemudian disimpan ke dalam database Neo4j menggunakan bantuan kueri cypher load yang diramu dalam aplikasi web berbasis php seperti yang ditunjukkan pada Gambar 7. Setelah data dikonversi dalam bentuk basis data grafik Neo4j, pengguna dapat mengakses data Al-Qur'an dan tematik dengan menggunakan aplikasi berbasis web.

Preprocessing Dokumen

Tahap pertama perangkaian dokumen adalah *preprocessing*. *Preprocessing* dilakukan pada dokumen dan *query* yang terdiri dari tiga tahap yaitu tokenisasi, *stopwords removal* dan *stemming*. Tokenisasi adalah proses pemecahan dokumen, di mana jika satu dokumen direpresentasikan sebagai satu halaman al-Quran maka dipecah menjadi kumpulan kalimat yang selanjutnya dipecah kembali menjadi kumpulan kata. Kemudian, kumpulan kata tersebut dilakukan penghapusan *diacritic* atau harokat yang terdapat pada bahasa Arab. Selanjutnya, tahap kedua adalah *stopwords removal* yaitu penghapusan kata-kata tidak penting. Penghapusan kata tidak penting dilakukan agar tidak mempengaruhi nilai bobot saat dilakukan pembobotan kata, di mana frekuensi kemunculan kata-kata tidak penting akan

muncul lebih banyak daripada kata pentingnya. *Library stopword* yang digunakan berdasarkan <https://pypi.org/project/Arabic-Stopwords/>. Tahap ketiga adalah *stemming* yaitu merubah suatu kata menjadi kata dasarnya. *Stemming* yang digunakan adalah *Arabic Light Stemmer*.

Perhitungan Kemiripan Secara Leksikal

Perhitungan kemiripan leksikal adalah perhitungan kemiripan menggunakan pembobotan kata. Pembobotan kata yang digunakan ialah *Term Frequency Inverse Document Frequency* (TF.IDF). Setelah pembobotan kata selanjutnya dilakukan perkalian vektor antara *query* dan dokumen. Nilai kemiripan leksikal antara tiap dokumen dan *query* dihitung menggunakan *Cosine Similarity*. Alur perhitungan kemiripan leksikal antara lain:

1) Pembobotan *Term Frequency Inverse Document Frequency* (TF.IDF)

Pembobotan $TF.IDF(w_i, d_j)$ ditunjukkan pada Persamaan (1). Bagian pertama menghitung frekuensi kemunculan suatu kata menggunakan $W_{tf}(w_i, d_j)$. $W_{tf}(w_i, d_j)$ adalah frekuensi kata dimana w_i adalah kata ke-i, sedangkan d_j adalah dokumen ke-j. Selanjutnya, menghitung *inverse document frequency* yang ditunjukkan pada Persamaan (2).

$$TF.IDF(w_i, d_j) = W_{tf}(w_i, d_j) * W_{idf}(w_i) \quad (1)$$

$$W_{idf}(w_i) = \log \frac{N}{df(w_i)} \quad (2)$$

dimana N adalah jumlah dokumen dalam koleksi korpus. Sedangkan $df(w_i)$ adalah jumlah dokumen yang berisi kata ke-i.

2) Perkalian Vektor antara *Query* dan Dokumen

Perkalian antara *query* dan tiap dokumen dilakukan untuk menghasilkan nilai bobot dari tiap kata pada suatu dokumen. Nilai bobot dari tiap kata dijumlahkan, di mana nilai ini dimasukkan ke dalam perhitungan kemiripan leksikal.

3) Perhitungan Kemiripan Leksikal

Perhitungan kemiripan leksikal didasarkan pada penggunaan *Cosine Similarity*. Perhitungan kemiripan tersebut memiliki nilai minimum yaitu 0 dan nilai maksimumnya sama dengan 1. Dalam pengukuran kemiripan ini, kemiripan antara dua objek dikatakan mirip jika nilainya mendekati 1. Sebaliknya, kemiripan antara dua objek dikatakan tidak mirip jika nilainya semakin mendekati 0. Persamaan (3) menunjukkan perhitungan kemiripan leksikal.

$$LexSim(Q, d_j) = \frac{\sum(TF.IDF(w_i, Q) \cdot TF.IDF(w_i, d_j))}{\sqrt{\sum(TF.IDF_Q)^2} \cdot \sqrt{\sum(TF.IDF_{d_j})^2}} \quad (3)$$

dimana Q adalah *query* yang berisi sejumlah kata dan d_j adalah suatu dokumen ke-j. $TFIDF_{(w_i, Q)}$ di mana w_i kata ke-i pada *query*, sedangkan $TF.IDF_{(w_i, d_j)}$ di mana w_i adalah kata ke-i pada dokumen ke-j. Sehingga $TF.IDF_{(w_i, Q)}$ dikali dengan $TF.IDF(w_i, d_j)$ adalah perhitungan TF.IDF antara *query* dan dokumen ke-j. Selain itu, $TF.IDF_Q$ adalah jumlah atau panjang kata pada *query* dan $TF.IDF_{d_j}$ adalah jumlah atau panjang kata pada dokumen ke-

j.

HASIL DAN PEMBAHASAN

Rancangan sistem dibangun menggunakan aplikasi berbasis web sebagaimana ditampilkan pada Gambar 8. Uji coba dilakukan terhadap 10 *query* berbahasa Indonesia. Uji coba dilakukan dengan menghitung kemiripan secara leksikal. Uji coba juga dilakukan terhadap hasil pencarian 5, 10, dan 15 teratas. Selain terhadap keseluruhan *query*, uji coba juga dilakukan terhadap frekuensi kata dalam *query* yang terbagi ke dalam dua bagian, yaitu: *query* dengan frekuensi kata kurang dari sama dengan 2 kata (*short query*) dan *query* dengan frekuensi kata lebih dari 2 kata (*long query*). Hal ini bertujuan untuk menunjukkan performa *query* sebagai bagian penting dalam pencarian dokumen.

SISTEM PENCARIAN AYAT ALQURAN

Masukkan konteks yang anda cari Search

Konteks: Kesucian Jiwa (Tazkiyah)

Al-Tawba : 103
Madaniyah
خُذْ مِنْ أَمْوَالِهِمْ صَدَقَةً تُطَهِّرُهُمْ وَتُزَكِّيهِمْ بِهَا وَصَلِّ عَلَيْهِمْ إِنَّ صَلَاتَكَ سَكَنٌ لَهُمْ وَاللَّهُ سَمِيعٌ عَلِيمٌ
Artinya:
Ambillah zakat dari sebagian harta mereka, dengan zakat itu kamu membersihkan dan mensucikan mereka dan mendoalah untuk mereka. Sesungguhnya doa kamu itu (menjadi) ketenteraman jiwa bagi mereka. Dan Allah Maha Mendengar lagi Maha Mengetahui.

Ash-Shams : 9
Makiyah
فَذُوقْ مِنْ زَجَائِمِهَا
Artinya:
sesungguhnya beruntunglah orang yang mensucikan jiwa itu.

Al-Baqara : 155
Madaniyah
وَلَنَبْلُوَنَّكُمْ بِشَيْءٍ مِنَ الْخَوْفِ وَالْجُوعِ وَنَقْصٍ مِنَ الْأَمْوَالِ وَالْأَنْفُسِ وَالثَّرَاتِ وَبَشِّرِ الصَّابِرِينَ
Artinya:
Dan sungguh akan Kami berikan cobaan kepadamu, dengan sedikit ketakutan, kelaparan, kekurangan harta, jiwa dan buah-buahan. Dan berikanlah berita gembira kepada orang-orang yang sabar.

Aal-Imran : 164
Madaniyah
لَقَدْ مَنَّ اللَّهُ عَلَى الْمُؤْمِنِينَ إِذْ بَعَثَ فِيهِمْ رَسُولًا مِنْ أَنْفُسِهِمْ يَتْلُو عَلَيْهِمْ آيَاتِهِ وَيُزَكِّيهِمْ وَيُعَلِّمُهُمُ الْكِتَابَ وَالْحِكْمَةَ وَإِنْ كَانُوا مِنْ قَبْلُ لَفِي ضَلَالٍ مُبِينٍ
Artinya:
Sungguh Allah telah memberi karunia kepada orang-orang yang beriman ketika Allah mengutus diantara mereka seorang rasul dari golongan mereka sendiri, yang membacakan kepada mereka ayat-ayat Allah, membersihkan (jiwa) mereka, dan mengajarkan kepada mereka Al Kitab dan Al Hikmah. Dan sesungguhnya sebelum (kedatangan Nabi) itu, mereka adalah benar-benar dalam kesesatan yang nyata.

Gambar 8. Sistem pencarian tematik ayat al-Quran

Tabel 1 menunjukkan query uji coba yang dilakukan. Hasil penelitian dievaluasi menggunakan *recall*, *precision* dan *f-measure*. *Recall*, *precision* dan *f-measure* biasa digunakan untuk mengukur performa hasil *retrieved* dokumen dalam sistem temu kembali informasi. *Recall* adalah tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi. Dalam konteks ini, *recall* adalah rasio dokumen yang di-*retrieve* dan relevan berdasarkan *query* terhadap dokumen yang relevan secara keseluruhan atau dengan kata lain dokumen yang berhasil di-*retrieve* di antara dokumen yang relevan. *Precision* adalah rasio dari dokumen yang di-*retrieved* dan relevan berdasarkan *query* terhadap dokumen yang di-

retrieved oleh sistem dengan kata lain dokumen yang relevan di antara semua dokumen yang berhasil di-retrieved, dengan kata lain *precision* adalah tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem. Serta, *f-measure* merupakan perhitungan kombinasi dari *precision* dan *recall* untuk mencari nilai tengah antara *precision* dan *recall*.

Tabel 1. Contoh Query untuk Uji Coba

	Query Bahasa Arab	Query Bahasa Indonesia
Short Query	“دَبَّحُ الْأَضَاجِي”، “الْمَحَارِمُ لِلْمَرْأَةِ”، “صَوْمُ النَّطْوُعِ”، “الْمَحَارِمُ لِلرِّجَالِ”، “تَشْرِبُ الْخَمْرِ”، “مُطَبَّلَاتُ الصَّلَاةِ”، “صَلَاةٌ”، “الْعَقِيقَةُ لِلنَّبَاتِ”، “أَزْكَانُ الْإِيمَانِ”، “غَسْلُ الْمَيِّتِ”، “صَلَاةُ التَّهَجُّدِ”، “الْجَنَازَةُ”، “أركان الوضوء”	“menyembelih hewan qurban”, “mahram anak perempuan”, “mahram anak laki-laki”, “puasa sunnah”, “meminum minuman keras”, “rukun iman”, “aqiqah anak perempuan”, “sholat jenazah”, “sholat tahajud”, “memandikan jenazah”, “rukun wudhu”
Long Query	“الطَّوَافُ فِي الْحَجِّ”، “صِيَامُ 6 أَيَّامٍ فِي شَوَّالٍ”، “أَكَلَ لَحْمٍ”، “رَمَى الْجُمَرَاتِ فِي الْحَجِّ”، “الصلوات الخمس المفروضة”، “الخنزير”، “قضاء صيام رمضان”، “إخراج زكاة الفطر”، “الإفطار في الطائرة”	“puasa 6 hari di bulan syawal”, “tawaf ibadah haji”, “lempar jumroh ibadah haji”, “memakan daging babi”, “shalat fardhu lima waktu”, “membayar zakat fitrah”, “qada puasa ramadhan”, “berbuka puasa di atas pesawat”

Seperti yang telah dijelaskan sebelumnya bahwa *query* yang diinputkan mewakili dokumen pencarian yang akan di-retrieve. Salah satu contoh *short query* misalnya “غَسْلُ الْمَيِّتِ” yang berarti “memandikan jenazah”. Contoh *long query* misalnya “صِيَامُ 6 أَيَّامٍ فِي شَوَّالٍ” yang artinya “puasa 6 hari di bulan Syawal”. Dari 20 *query* uji coba, terdapat 12 *query* yang termasuk ke dalam golongan *short query* dan 8 *query* lainnya termasuk ke dalam golongan *long query*. Seluruh *short query* dilakukan uji coba yang hasilnya dapat dilihat pada Tabel 1. Sedangkan, hasil uji coba *long query* terdapat pada Tabel 2. Hasil retrieve dokumen pencarian yang dihasilkan oleh sistem divalidasi oleh pakar yang mampu berbahasa Arab dan ahli tafsir al-Quran serta berkualifikasi sebagai dosen Ilmu Al-Quran dan Tafsir di Universitas Nahdlatul Ulama Blitar.

Tabel 2. Recall dan Precision untuk short query

Query bahasa indonesia	Recall(%) All Query			Precision(%) All Query		
	Top 5	Top 10	Top 15	Top 5	Top 10	Top 15
Kesucian Jiwa	62.16%	74.49%	83.82%	78.00%	82.50%	66.33%
Kebersihan	70.50%	82.43%	90.10%	78.00%	86.50%	66.33%
Hewan Qurban	40.10%	62.99%	77.28%	65.00%	51.00%	42.33%
Shalat	72.56%	84.49%	93.82%	78.00%	88.50%	66.33%
Minuman Keras	24.06%	34.80%	48.40%	37.00%	41.50%	49.00%

Berdasarkan Tabel 2 pengujian terhadap seluruh *query*, untuk *query* shalat memiliki hasil tertinggi untuk *recall* dan *precision*. Tercatat pada hasil pencarian 5,10, dan 15 teratas yaitu 72.56%, 84.49% dan 93.82% pada *recall* secara berturut-turut, serta *precision* pada hasil pencarian 5, 10 dan 15 teratas yaitu 78.00, 88.50% dan 66.33%. Sedangkan, untuk *query* minuman keras memiliki hasil terendah untuk *recall* dan *precision*. Tercatat pada hasil

pencarian 5,10, dan 15 yaitu 24.06%, 34.80% dan 48.40% pada *recall* secara berturut-turut, serta *precision* pada hasil pencarian 5, 10 dan 15 teratas yaitu 37.00, 41.50% dan 49.00%. Hasil *query* minuman keras memiliki *recall* dan *precision* yang rendah dikarenakan algoritma pencarian belum menerapkan pencarian berbasis semantik, yaitu makna kata. Contoh dalam kata dalam al-Quran yang memiliki makna sepadan dengan minuman keras adalah kata “khamer” dan “memabukkan”.

KESIMPULAN

Penelitian ini dapat disimpulkan bahwa metode knowledge graph secara garis besar merupakan implementasi data yang direalisasikan menggunakan visualisasi grafik dimana semua relasi antara tema akan dipaparkan. Penelitian ini mengusulkan bahwa knowledge graph akan berpotensi sebagai penunjang tafsir maudhu’i bagi para pengguna yang dalam hal ini adalah para ahli tafsir Al-Qur’an. Hasil uji coba menunjukkan bahwa untuk *query* shalat memiliki hasil tertinggi untuk *recall* dan *precision*. Tercatat pada hasil pencarian 5,10, dan 15 teratas yaitu 72.56%, 84.49% dan 93.82% pada *recall* secara berturut-turut, serta *precision* pada hasil pencarian 5, 10 dan 15 teratas yaitu 78.00, 88.50% dan 66.33%. Sedangkan, untuk *query* minuman keras memiliki hasil terendah untuk *recall* dan *precision*. Tercatat pada hasil pencarian 5,10, dan 15 yaitu 24.06%, 34.80% dan 48.40% pada *recall* secara berturut-turut, serta *precision* pada hasil pencarian 5, 10 dan 15 teratas yaitu 37.00, 41.50% dan 49.00%. Hasil *query* minuman keras memiliki *recall* dan *precision* yang rendah dikarenakan algoritma pencarian belum menerapkan pencarian berbasis semantik, yaitu makna kata. Contoh dalam kata dalam al-Quran yang memiliki makna sepadan dengan minuman keras adalah kata “khamer” dan “memabukkan”.

REFERENSI

- [1] T. T. Nur, “Perancangan Aplikasi Pembelajaran Membaca al- Qur’an dengan Metode Qiro’ah Berbasis Android,” Universitas Islam Negeri Alauddin Makassar, 2013.
- [2] B. Ali and M. Ahmad, “Al-Quran themes classification using ontology,” *Icoci. C. Net. My*, vol. 74, pp. 383–389, 2013.
- [3] M. F. Noordin, T. M. T. Sembok, R. Othman, and R. H. Gusmita, “Constructing an ontology-based and graph-based knowledge representation of English Quran,” *J. Teknol.*, vol. 78, no. 8–2, 2016.
- [4] Z. Farihan Elyana, “Ontology application for the Al-Quran,” Universiti Utara Malaysia, 2009.
- [5] A. R. Yauri, R. A. Kadir, A. Azman, and M. A. A. Murad, “Quranic verse extraction base on concepts using OWL-DL ontology,” *Res. J. Appl. Sci. Eng. Technol.*, vol. 6, no. 23, pp. 4492–4498, 2013.
- [6] A. R. Yauri, R. A. Kadir, A. Azman, M. Azrifah, and A. Murad, “Semantic Web Application for Historical Concepts Search in Al- Quran,” *Int. J. Islam. Appl. Comput. Sci. Technol.*, vol. 2, no. 2, pp. 1–7, 2014.
- [7] N. K. Farooqui, M. F. Noordin, and R. Othman, “Ontology Matching: A Case of English Translation of Al-Quran Tafsir,” in 2018 International Conference on Information and Communication Technology for the Muslim World (ICT4M), 2018, pp. 1–6.
- [8] Y. Chi, Y. Qin, R. Song, and H. Xu, “Knowledge Graph in Smart Education: A Case Study of Entrepreneurship Scientific Publication Management,” *Sustainability*, vol. 10, no. 4, p. 995, Mar. 2018.
- [9] A. H. Nasution, Y. Murakami, and T. Ishida, “Generating Similarity Cluster of Indonesian Languages with Semi-Supervised Clustering,” *Int. J. Electr. Comput. Eng.*, vol. 9, no. 1, pp. 1–8, 2019.
- [10] A. H. Nasution and Y. Murakami, “Visualizing Language Lexical Similarity Clusters: A Case Study of Indonesian Ethnic Languages,” *J. Data Sci. Its Appl.*, vol. 2, no. 2, pp. 45–59, Nov. 2019.