

Visualisasi Tematik Al-Qur'an Berbasis Knowledge Graph

¹Lukman Nul Hakim, ²Winda Monika, ³Salhazan Nasution, ⁴Arbi Haza Nasution

^{1,4}Program Studi Teknik Informatika, Universitas Islam Riau

²Perpustakaan Universitas Islam Riau

³Program Studi Teknik Informatika, Universitas Riau

^{1,2,4}Jl. Kaharuddin Nst No.113, Simpang Tiga, Kec. Bukit Raya, Pekanbaru, Riau 2828428284

³Kampus Bina Widya Km 12,5 Simpang Baru, Pekanbaru, Riau 28293

¹lukmannulhakim@student.uir.ac.id, ²winda@tensaitech.id, ³salhazan@lecturer.unri.ac.id, ⁴arbi@eng.uir.ac.id

Abstrak— Al-Qur'an merupakan firman Allah SWT yang diturunkan kepada nabiullah baginda Nabi Muhammad SAW yang mana Al-Qur'an terdiri dari berbagai tema yang memiliki kaitan antar satu dengan yang lainnya. Penelitian ini menyajikan desain model manajemen publikasi tematik untuk mengintegrasikan metadata ilmiah berdasarkan *knowledge graph*. Berdasarkan model ini, platform media visualisasi tematik direalisasikan untuk pengambilan sumber daya ilmiah dan analisis, yang tujuannya adalah untuk meningkatkan efisiensi pencarian ilmiah dan mengurangi kesulitan dalam mempelajari tematik dalam Al-Qur'an. Pada penelitian ini, hubungan ayat-ayat Al-Qur'an dan tematiknya dirumuskan dalam bentuk *knowledge graph* dengan menggunakan basis data grafik Neo4j. Data tematik Al-Quran diambil dari Al-Qur'an Amazing (Cordoba) dan data Al-Qur'an diambil dari qurandatabase.org. Aplikasi web visualisasi hubungan ayat-ayat Al-Qur'an dengan tematiknya yang ditampilkan dalam bentuk grafik telah dibangun dengan nilai rata-rata precision 1 dan F-score 0.56 untuk pencarian berdasarkan surat dan ayat dan nilai rata-rata precision 1 dan F-score 1 untuk pencarian berdasarkan tema.

Kata kunci— Al-Qur'an, Tematik, Knowledge Graph, Visualisasi, Graph Database

I. PENDAHULUAN

Al-Qur'an 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. Pada saat ini 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].

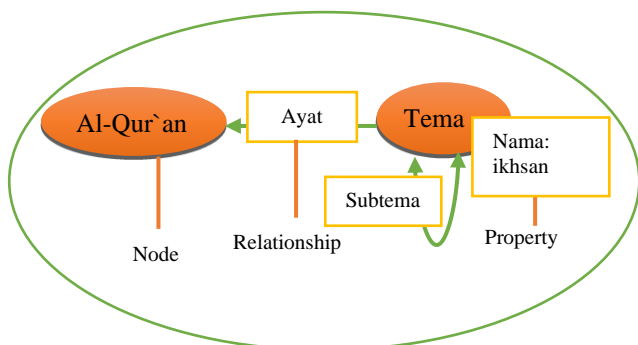
Penelitian-penelitian tersebut menjadi motivasi bagi penulis bahwa *knowledge graph* akan berpotensi untuk mempermudah pengguna dalam temu kembali informasi dan pembelajaran Al-Qur'an.

II. METODE PENELITIAN

Pada media visualisasi tematik ini, data diambil dari dua sumber yaitu <http://www.qurandatabase.org/> (Database Al-Qur'an) dan Al-Qur'an Amazing Cordoba (Data Al-Qur'an Tematik).

A. NEO4J

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.



Gambar. 1 Model Grafik Neo4j

Permodelan basis data grafik terdiri dari empat bagian seperti yang ditunjukkan pada Gambar 1 yaitu sebagai berikut:

- 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*

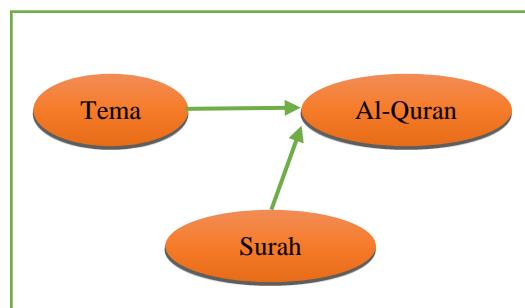
dapat memiliki satu atau lebih properti (yaitu, atribut yang disimpan sebagai pasangan kunci / nilai). Contoh *relationships* pada Gambar 1 ialah Ayat dan Subtema.

- 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.

B. 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 Model Node.

1) Tema

Tema merupakan *node* yang menampung semua data berbagai macam data tema dengan satu properti yaitu "nama".

2) Al-Qu'ran.

Al-Qur'an merupakan *node* yang menampung semua data Al-Quran dengan lima properti yang ditunjukkan pada Tabel I.

TABEL I
TABEL NODE AL-QUR'AN

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

3) Surah.

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

TABEL II
TABEL NODE SURAH

No	Property
1	Kode_surat
2	Nama_surat
3	Arti_Surat
4	Jumlah_ayat
5	Tempat_turun

C. 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 semua tema 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
CREATE (st)-[:Ayat]->(k1)
```

Gambar. 5 Load Tematik.

2) Cypher Hops.

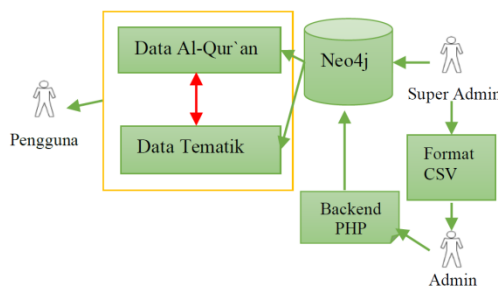
Cypher hops merupakan suatu kode *cypher* yang tersedia di platform Neo4j yang berguna untuk mencari semua relasi data yang terhubung dengan jarak hubungan yang diinginkan, seperti yang ditunjukkan pada Gambar 6.

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

Gambar. 6 Cypher Hops.

D. PENGEMBANGAN DAN PERANCANGAN MEDIA

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 Alur Kerja Media Visualisasi Tematik Al-Qur'an.

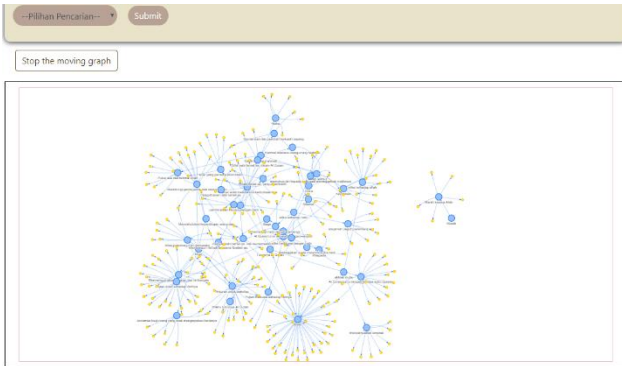
Proses aliran data yang terjadi pada media visualisasi diawali dengan konversi data file csv yang diinputkan oleh admin yang kemudian disimpan ke dalam database Neo4j menggunakan bantuan kueri *cypher load* yang diramu dalam aplikasi web berbasis php seperti yang ditunjukkan pada Gambar 7. Setelah data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

Qur'an dan tematik dengan menggunakan aplikasi berbasis web.

III. HASIL DAN PEMBAHASAN

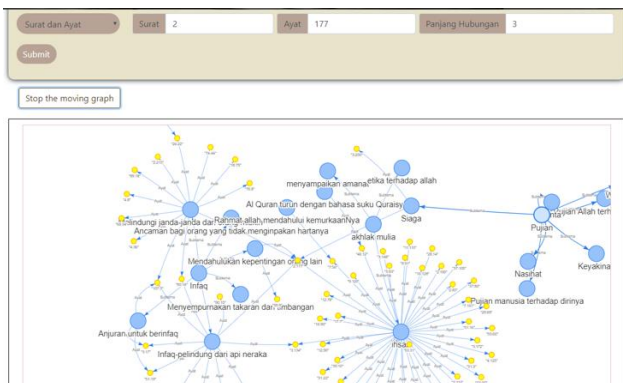
E. DESAIN ANTARMUKA

Berikut merupakan contoh tampilan hasil media visualisasi tematik Al-Qur'an:



Gambar. 9 Desain Antarmuka Pengguna Tematik.

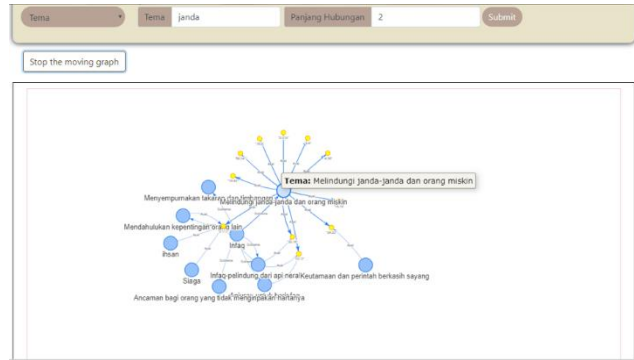
Gambar 9 menunjukkan tampilan media visualisasi dengan data tema sebanyak 46 tema dan 274 surat dan ayat. Untuk pencarian berdasarkan estimasi jarak bisa dilihat pada Gambar 10.



Gambar. 10 Pencarian Surat dan Ayat.

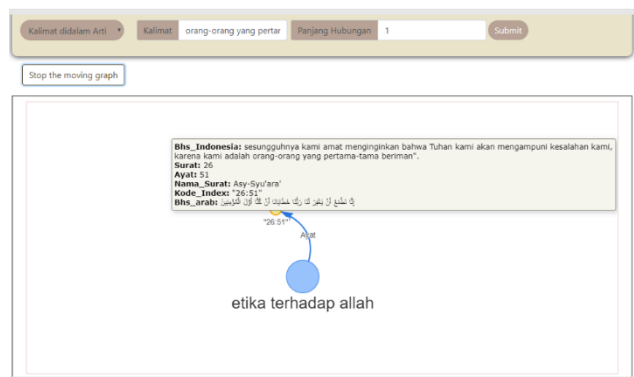
Data yang dicari pada Gambar 10 merupakan pencarian surat = 2 dan ayat = 177 dengan panjang hubungan sebanyak = 3. Untuk pencarian panjang hubungan menggunakan *cypher hops* yang berfungsi untuk mencari tetangga dari *node* yang dicari sebanyak 3 jarak hubungan.

Gambar 11 menunjukkan contoh pencarian tema dalam media visualisasi yang telah dibangun. Data yang dicari pada Gambar 11 merupakan pencarian tema = janda dengan panjang hubungan sebanyak = 2. Pada pencarian tema hanya perlu memaksukan key kata yang ingin dicari dan menampilkan node dengan panjang hubungan sebanyak 2.



Gambar. 11 Pencarian Tema.

Gambar 12 menunjukkan contoh pencarian kalimat lengkap atau kosakata dalam terjemahan dari surat dan ayat di terjemahan Al-Qur'an berbahasa Indonesia dalam media visualisasi yang dibangun.



Gambar. 12 Pencarian kalimat Dalam Terjemahan Surat dan Ayat.

F. PENGUJIAN MEDIA

Pengujian sistem dilakukan dengan cara membandingkan antara pencarian manual pada Al-Qur'an Amazing Cordoba dengan pencarian menggunakan media visualisasi tematik Al-Qur'an. Terdapat 2 cara pengujian yaitu (1) pencarian berdasarkan surat dan ayat untuk melihat ketepatan tema yang dihasilkan, dan (2) pencarian berdasarkan tema untuk melihat ketepatan surat dan ayat yang dihasilkan. Hasil pengujian pencarian berdasarkan surat dan ayat dapat dilihat pada Tabel III, dan pencarian berdasarkan tema dapat dilihat pada Tabel IV.

TABEL III
TABEL PENCARIAN BERDASARKAN SURAT DAN AYAT

No	Pertanyaan	Data Dari Al-Quran Amazing Cordoba	Data Dari Media Visualisasi Tematik	Precision	Recall	F1 Score
1	Coba cari surat = 2 dan ayat = 177, panjang hubungan= 1	Ada 20 tema yang dihasilkan { ihsan, hakikat kebaikan, jalan taqwa, ganjaran sabar, sabar dalam menghadapi sesuatu yang kita benci, keutamaan berhati-hati dalam berbuat jujur, sabar saat berperang, siaga, perintah memberi solat salah satu rukun islam, kewajiban zakat, menunaikan zakat salah satu rukun islam, tauhid uluhiyyah, kewajiban beriman kepada malaikat, kewajiban dan keutamaan beriman pada kitab-kitab, kewajiban beriman pada para rasul, kewajiban beriman pada hari akhir, iman adalah ucapan dan perbuatan, balasan dan pahala dari allah, memenuhi sumpah dan janji, Melindungi janda-janda dan orang miskin, infaq-pelindung dari api neraka }	Ada 5 tema yang dihasilkan {Mendahulukan kepentingan orang, ihsan, Melindungi janda-janda dan orang miskin, siaga, infaq-pelindung dari api neraka }	1	0.25	0.4
2	Coba cari surat = 3 dan ayat = 17, panjang hubungan = 1	Ada 6 tema yang dihasilkan { Anjuran untuk berinfaq, infaq pelindung dari api neraka, rendah diri dan khusuk, keutamaan berhati-hati dalam berbuat jujur, keutamaan istighfar, kapan disunnatkan istighfar }	Ada 2 tema yang dihasilkan {Anjuran untuk berinfaq, infaq pelindung dari api neraka }	1	0.33	0.5
3	Coba cari surat = 7 dan ayat = 199, panjang hubungan = 2	Ada 11 tema yang dihasilkan {Tawadhu, Rahmat diantara orang-orang beriman, memaafkan orang yang bersalah, berbaur dengan orang-orang bodoh, cara berdakwah kepada allah, amar ma'ruf dan nahi munkar, kesetiaan, cinta, lemah lembut, kelapangan dada orang alim, Menjauhi syubhat }	Ada 3 tema yang dihasilkan {Rahmat diantara orang-orang beriman, cinta, lemah lembut }	1	0.28	0.4
4	Coba cari surat = 37 dan ayat = 102, panjang hubungan = 1	Ada 9 tema yang dihasilkan { sabar dalam menghadapi sesuatu yang kita benci, keyakinan, Allah menguji nabi Ibrahim AS, mimpi menyembelih anaknya (Ismail), taat kepada orang tua, mematuhi perintah Allah & Rasulnya, Nabi Ismail AS. yang disembelih, Sifat nabi Ismail AS. dalam Al Quran, keyakinan }	Ada 3 tema yang dihasilkan {Nabi Ismail AS. yang disembelih,Sifat nabi Ismail AS. dalam Al Quran, keyakinan }	1	0.3	0.5
5	Coba cari surat = 37 dan ayat = 107, panjang hubungan = 1	Tema {Nabi Ismail AS. disembelih, Pengorbanan nabi Ismail }	Tema {Nabi Ismail AS. disembelih, Pengorbanan nabi Ismail }	1	1	1

TABEL IV
TABEL PENCARIAN BERDASARKAN TEMA

No	Pertanyaan	Data Dari Al-Quran Amazing Cordoba	Data Dari Media Visualisasi Tematik	Precision	Recall	F1 Score
1	Coba cari tema = etika terhadap allah, panjang hubungan = 1	Surat dan ayat {11:47, 11:45, 20:84, 26:51, 20:12, 27:19, 2:23, 18:24, 7:143}	Surat dan ayat {11:47, 11:45, 20:84, 26:51, 20:12, 27:19, 2:23, 18:24, 7:143}	1	1	1
2	Coba cari tema = pembicaraan khusus dengan nabi, panjang hubungan = 1	Surat dan ayat {58:12, 58:13}	Surat dan ayat {58:12, 8:13}	1	1	1
3	Coba cari tema = khusus, panjang hubungan = 1	Surat dan ayat {58:13, 58:12}	Pembicaraan khusus dengan nabi {58:13, 58:12}	1	1	1
4	Coba cari tema = Nasihat, panjang hubungan = 1	Surat dan ayat {103:3, 24:17, 11:34, 90:17, 7:159, 36:20, 4:63, 28:76, 28:20, 5:79}	Surat dan ayat {103:3, 24:17, 11:34, 90:17, 7:159, 36:20, 4:63, 28:76, 28:20, 5:79}	1	1	1
5	Coba cari tema = Siaga, panjang hubungan = 1	Surat dan ayat {3:200, 2:177}	Surat dan ayat {3:200, 2:177}	1	1	1

Berdasarkan Tabel III, hasil pencarian berdasarkan surat dan ayat untuk menentukan ketepatan tema, rata-rata F1 Score yang dihasilkan adalah:

$$\text{Rata-rata} = \frac{0.4+0.5+0.4+0.5+1}{5} = 0.56$$

Walaupun semua pengujian menghasilkan *precision* bernilai 1, namun nilai F1 Score media visualisasi tematik Al-Qur'an berbasis knowledge graph dalam pencarian berdasarkan surat dan ayat hanya bernilai 0.56, hal ini dikarenakan tidak semua data pada Al-Qur'an Cordoba dimasukkan ke dalam media visualisasi. Namun, berdasarkan Tabel IV, hasil pencarian berdasarkan tema untuk menentukan ketepatan surat dan ayat, seluruh pengujian memiliki F1 Score bernilai 1. Hal ini dikarenakan surat dan ayat pada beberapa tema yang sudah ditetapkan sebagai *sample* diinputkan semua ke dalam media visualisasi.

IV. KESIMPULAN

Dalam penelitian ini, media visualisasi dibuat untuk memudahkan pengguna untuk melakukan pencarian berbasis tematik pada Al-Quran dan semua hubungan antara tema dengan tema yang lain, tema dengan ayat, dan ayat dengan ayat lain berdasarkan tema. Aplikasi web visualisasi hubungan ayat-ayat Al-Qur'an dengan tematiknya yang ditampilkan dalam bentuk grafik telah dibangun dengan nilai rata-rata precision 1 dan F-score 0.56 untuk pencarian berdasarkan surat dan ayat dan nilai rata-rata precision 1 dan F-score 1 untuk pencarian berdasarkan tema.

UCAPAN TERIMA KASIH / ACKNOWLEDGMENT

Penelitian ini didanai oleh Universitas Islam Riau melalui Lembaga Penelitian dan Pengabdian Kepada

Masyarakat dengan Nomor: 687/KONTRAK/LPPM-UIR/5-2019.

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.