

BASIS DATA GRAF NEO4J PADA SILSILAH KELUARGA

LAILA KUSUMA WARDANI¹, DHOMAS HATTA FUDHOLI²

^{1,2}Program Studi Informatika, Fakultas Teknologi Industri Universitas Islam
Indonesia, Indonesia

email: ¹18523010@students.uui.ac.id, ²hatta.fudholi@uui.ac.id

ABSTRAK

Perubahan sosial masyarakat Indonesia dipengaruhi oleh peran media sosial di era globalisasi saat ini. Fenomena yang terjadi, terutama di kalangan remaja, menunjukkan bahwa kebersamaan keluarga dan pengetahuan informasi semakin berkurang. Ini berarti bahwa seseorang tidak tahu sejarah keluarganya. Untuk menyelesaikan masalah tersebut, perancangan basis data graf diperlukan untuk menyimpan data keluarga. Dalam penelitian ini, teknologi basis data graf Neo4j dianggap sebagai alat yang efektif untuk menampilkan dan menyimpan banyak data kompleks. Untuk membuat properti dalam basis data, informasi perlu dikumpulkan. Informasi ini berasal dari buku, jurnal, dan wawancara dengan orang-orang di bidang yang relevan. Selanjutnya, data diekstrak dan digabungkan dengan data anggota keluarga. Pada graf, satu anggota keluarga merepresentasikan simpul dan hubungan antar anggota keluarganya merepresentasikan sisi. Sehingga, implementasi dan visualisasi graf yang menggambarkan simpul dari tiap-tiap anggota keluarga dan ragam relasi antar anggota keluarga akan didapatkan sebagai hasil akhir dari penelitian ini.

Kata kunci: Basis Data, Graf, Neo4j, Perancangan Basis Data, Silsilah Keluarga.

I. PENDAHULUAN

Keluarga merupakan gabungan dua orang atau lebih yang bersatu karena adanya kebersamaan, kedekatan emosional serta hal-hal lain yang mengidentifikasikan mereka sebagai sebuah keluarga (Friedman, 2010). Sebagai satu kesatuan unit terkecil dalam masyarakat, antar anggota keluarga pasti memiliki keterikatan hubungan darah baik dari ikatan perkawinan atau ikatan persaudaraan.

Pengaruh globalisasi yang berkembang pesat saat ini, salah satunya berdampak pada berkurangnya interaksi dan kebersamaan dalam keluarga. Didukung dengan kurangnya pengelolaan informasi mengenai silsilah keluarganya (Setialana, 2017) menyebabkan seseorang tidak mengenal anggota keluarga lainnya. Untuk mengatasi permasalahan tersebut, maka dibutuhkan sebuah basis data yang memudahkan dalam pengelompokan dan identifikasi data terkhusus data silsilah keluarga.

Pohon keluarga dapat berfungsi sebagai representasi bergambar dari data keluarga. Data ini mencakup informasi tiap-tiap anggota keluarga dan hubungan antar anggota keluarga dari berbagai generasi. Dari segi konseptual, pohon keluarga menggambarkan silsilah keluarga dari tiga generasi, yaitu generasi dirinya, orangtuanya, dan kakek neneknya (Fathonah, 2019).

Teori graf dalam basis data graf telah dibahas dalam berbagai bidang pada penelitian-penelitian sebelumnya. Diantaranya, penerapan basis data graf pada aplikasi mesin rekomendasi yang berperan penting dalam berbagai bidang industri (Azhar, 2018). Aplikasi ini digunakan untuk meningkatkan pengalaman berbelanja para pelanggan dengan menebak apa yang selanjutnya pelanggan tersebut inginkan. Pada aplikasi ini, basis data graf melakukan *relationship-based query* sehingga dapat menganalisis pola perilaku belanja pelanggan dengan mudah. Selain itu ada juga penelitian basis data graf di bidang kesehatan yakni untuk analisa struktur protein (Lokawati, 2016). Dimana penulis menjelaskan bahwa gambaran fungsi kerja sel dan penyakit-penyakit kompleks dapat dipelajari melalui jaringan interaksi antar protein. Jaringan tersebut kemudian direpresentasikan dalam graf. Dengan protein merujuk pada simpul dan PPI merujuk sebagai sisi atau hubungan. Representasi inilah yang nantinya digunakan ketika terdapat kasus analisa karakter struktural dari protein dengan skala besar.

Beberapa penelitian yang telah dilakukan menunjukkan bahwa belum banyak topik yang membahas mengenai basis data graf. Sedangkan Miller dalam penelitiannya berjudul “*Graph Database Applications and Concept with Neo4j*” menerangkan bahwa basis data graf telah membawa perubahan baru dalam pemodelan dan visualisasi data pada penyimpanan informasi skala besar yang saling berhubungan.

Oleh karena itu, data silsilah keluarga tepat dirancang menggunakan basis data graf sebagai penyimpanan data terhubung seperti informasi personal sebagai node yang terpisah (Yar & Tun, 2016). Penelitian ini nantinya dapat menjadi acuan penyimpanan data genealogi sesuai kebutuhan keluarga dan rancangan struktur data yang telah ditentukan. Pemilihan Neo4j sebagai basis data graf dalam

penelitian “Basis Data Graf pada Data Silsilah Keluarga” ini diharapkan dapat menjadi pembeda dari penelitian yang sudah ada, dimana sebagian besar basis datanya menggunakan relasional.

II. METODE PENELITIAN

Metode penelitian merupakan salah satu upaya memperoleh data untuk kemudian dapat dideskripsikan, dikembangkan, atau dibuktikan adanya hasil akhir yang sesuai dengan penelitian tersebut. Dalam penelitian ini metode kualitatif deksriptif dipilih sebagai metode yang tepat dalam menunjang perancangan basis data graf pada data silsilah keluarga. Tahapan yang dilakukan sebagai sarana pencapaian hasil akhir dalam penelitian ini adalah sebagai berikut:

1. Identifikasi dan analisa kebutuhan data

Tahap ini dilakukan dengan mencari informasi mengenai data apa saja yang diperlukan dalam sebuah silsilah keluarga. Informasi yang diperlukan adalah informasi umum baik dari identitas individunya maupun relasi dalam satu keluarganya.

2. Pemodelan basis data graf secara konseptual

Setelah data silsilah keluarga berhasil diidentifikasi, maka langkah selanjutnya adalah memodelkan basis data graf. Hal ini perlu dilakukan agar dapat dilakukan implementasi penerapan *syntax* untuk menunjukkan hasil akhir visualisasi dari silsilah keluarga.

3. Implementasi pemodelan basis data graf

Pengimplementasian model basis data graf dalam penelitian ini adalah dengan menggunakan Neo4j. Dimana Neo4j merupakan basis basis data transaksional bersifat *open source* yang tidak melakukan penyimpanan struktur data dalam bentuk tabel melainkan dalam bentuk graf (Fernandes & Bernardino, 2018).

III. HASIL DAN PEMBAHASAN

Hasil proses pembangunan basis data graf keluarga menggunakan Neo4j. Secara sistematis hasil penelitian ini disajikan dalam susunan sebagai berikut.

A. Identifikasi dan Analisa Kebutuhan Data

Graf terdiri dari 2 himpunan berhingga (Daniel & Taneo, 2019), yaitu:

1. Verteks atau simpul (*vertex*), adalah himpunan tak kosong dari simpul-simpul.
2. Sisi (*edge*), adalah himpunan (mungkin kosong) dari sisi yang menghubungkan antar simpul.

Visualisasi graf ialah sekumpulan titik atau *nodes* yang saling terhubung melalui garis atau *relationships*. Dalam penelitian ini, simpul merepresentasikan seseorang dalam sebuah keluarga, sedangkan sisi merepresentasikan hubungan antar anggota keluarganya. Contoh graf yang memperlihatkan hubungan antar anggota keluarga seperti terlihat pada Gambar 1.



Gambar 1. Graf Keluarga yang Dibentuk Manual

Bahri dalam bukunya berjudul “*SILSILAH KELUARGA KIAI ENDING ZAHIDI, Suatu Jejak yang Hampir Terlupakan*” menggambarkan silsilah keluarganya cukup dengan informasi berupa nama. Namun melalui hasil wawancara dengan Kepala Bidang Pengelolaan Informasi Administrasi Kependudukan dan Pemanfaatan Data Dinas Kependudukan dan Pencatatan Sipil Kabupaten Wonogiri, Totok Sugiyarto, terdapat beberapa informasi penunjang lainnya yang dapat melengkapi data seseorang seperti tahun lahir, tempat lahir, alamat, jenis kelamin, dan posisi dalam keluarga. Beliau juga menjabarkan status hubungan antar anggota keluarga inti paling tidak terdiri dari, suami-istri, ibu-anak, kakak-adik, dan sepupu. Sehingga dirincikanlah informasi-informasi tersebut dalam Tabel 1.

Tabel 1. Lingkup Informasi pada Basis Data Graf

Data	Tipe dalam graf	Properti	Jenis Properti
Individu	Simpul	Nama Tahun Lahir Tempat Lahir Alamat Jenis Kelamin Posisi	
Hubungan	Sisi	Status Hubungan	Suami-Istri Ibu-Anak Kakak-Adik Sepupu

B.Pemodelan Basis Data Graf Secara Konseptual

Silsilah keluarga merupakan susunan individu dalam sebuah keluarga mulai dari keluarga inti (ayah, ibu, dan anak) sampai keluarga besar. Silsilah keluarga biasanya mencakup sampai dengan 3 generasi. Dalam sebuah keluarga, terdapat kemungkinan seseorang memiliki lebih dari satu hubungan dengan anggota keluarga lain. Misal seorang anak yang sudah menikah memiliki hubungan suami-istri dengan istrinya, memiliki hubungan ibu-dan anak dengan ibunya, dan memiliki hubungan kakak-adik dengan adiknya. Sehingga pada pemodelan ERD, antar individu dalam keluarganya memiliki hubungan *one to many* seperti Gambar 2.



Gambar 2. Model ERD Basis Data Graf pada Data Silsilah Keluarga

Setelah ERD dimodelkan, maka kemudian dapat dirancang model *propertygraph* seperti pada Gambar 3. Pada gambar tersebut terdapat simpul yang saling berhubungan dengan sifat yang sama. Dimana simpul tersebut mendeklarasikan seorang individu. Hubungan yang mengikat antar individu memiliki arah yang menunjukkan bahwa individu satu memiliki status hubungan tertentu dengan individu lainnya.



Gambar 3. *Propertygraph* Basis Data Graf pada Data Silsilah Keluarga

C.Implementasi Pemodelan Basis Data Graf

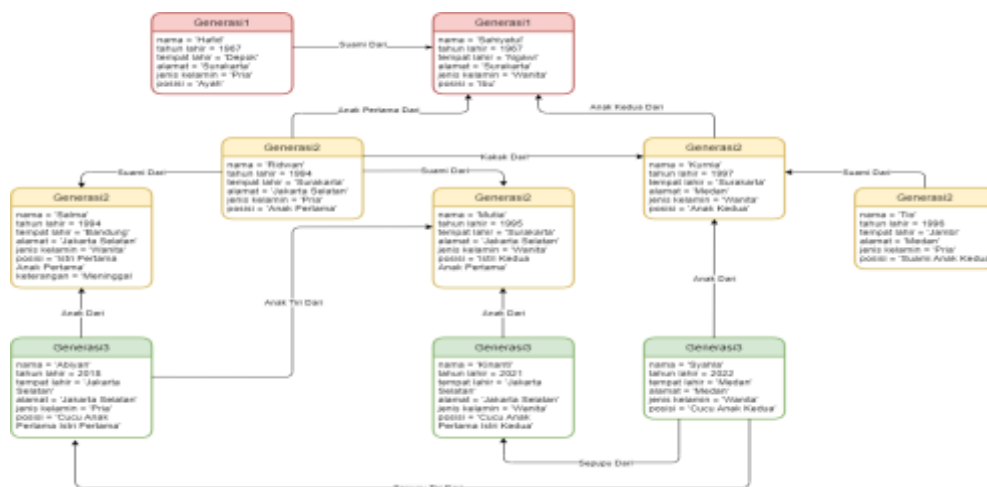
Pada tahap ini dilakukan dengan mengimplementasikan model *propertygraph* pada Gambar 3 ke dalam basis data graf Neo4j. Contoh data yang akan diterapkan ke dalam basis data graf adalah Gambar 1. Pada gambar tersebut terdapat empat individu dengan nama masing-masing, inilah yang nantinya akan dikuerikan menjadi simpul.

Bahasa pemrograman yang digunakan Neo4j untuk melakukan interaksi perintah dalam merancang basis data adalah *Chyper Query Language*. Tabel 2 menjabarkan perintah apa saja yang digunakan dalam merancang simpul dan hubungan pada Neo4j.

Tabel 2. Penjelasan Perintah Dasar pada Neo4j

Perintah	Penjelasan
CREATE	Membuat simpul, hubungan, dan properti
MATCH	Mengambil kembali atau <i>retrieve</i> data simpul, hubungan, atau properti yang telah dibuat
WHERE	Memberikan kondisi untuk menyaring data yang diambil
RETURN	Mengembalikan kondisi ke hasil kueri

Untuk keperluan pembuatan basis data graf dan pengaksesannya, graf yang digunakan ditunjukkan secara lebih rinci beserta propertinya pada Gambar 4. Seperti pada Gambar 1, Gambar 4 menunjukkan data individu satu dengan individu lainnya beserta status hubungannya.



Gambar 4. ERD Silsilah Keluarga

Untuk memvisualisasikan graf tersebut melalui Neo4j, maka langkah pertama yang dilakukan adalah menyusun *chyper query* dalam rangka membentuk simpul. Perintah yang perlu diberikan adalah sebagai berikut:

```
CREATE (:Generasi1 {nama:'Hafid', tahun_lahir:1967, tempat_lahir:'Depok', alamat:'Surakarta', jenis_kelamin:'Pria', posisi:'Ayah'})
```

Perintah di atas dilakukan secara berulang menyesuaikan data masing-masing anggota keluarga yang tersedia. Sesuai data pada Gambar 4, maka perintah pembuatan simpul tersebut dilakukan sebanyak 10 kali.

Langkah selanjutnya adalah membuat hubungan antar simpul. Perintah yang perlu diberikan adalah seperti berikut:

```
MATCH (g1:Generasi1), (g:Generasi1)
WHERE g1.nama = 'Hafid' AND
      g.nama = 'Sahiyatul'
CREATE (g1)-[:STATUS_HUBUNGAN {peran: ['Suami Dari']}]-(g)
```

Sama seperti saat membuat simpul, perintah pembuatan hubungan di atas dilakukan secara berulang menyesuaikan status hubungan apa saja yang ada dalam data silsilah keluarga. Sesuai data pada Gambar 4, maka perintah pembuatan hubungan tersebut dilakukan sebanyak 13 kali.

Simpul dan hubungan yang telah dibentuk kemudian divisualisasikan ke dalam graf yang menampilkan data beserta hubungan ibu dan dua orang anak seperti pada Gambar 6. Untuk keperluan tersebut, perintah di bawah ini diberikan:

```
MATCH p=()-[r:STATUS_HUBUNGAN]->() RETURN p LIMIT 25
```

Dengan perintah di atas, terlihatlah hasil akhir visualisasi silsilah keluarga sesuai data pada Gambar 4 yang menampilkan 10 anggota keluarga dengan 13 status hubungan yang menghubungkan masing-masing antar anggota keluarga. Perbedaan warna yang terdapat pada Gambar 6 merepresentasikan perbedaan generasi. Warna merah menunjukkan generasi pertama yang berisi ayah dan ibu. Warna krem menunjukkan generasi kedua yang berisi 2 anak dan 3 menantu. Warna hijau menunjukkan generasi ketiga yang berisi 3 orang cucu.



Gambar 6. Visualisasi Graf pada Neo4j

IV.KESIMPULAN

Teori graf dapat dimanfaatkan dalam memodelkan berbagai permasalahan termasuk untuk memodelkan data silsilah keluarga. Salah satu keunggulan yang ditunjukkan dengan menggunakan Neo4j adalah pengguna hanya perlu menuliskan kueri berdasarkan data yang dimiliki dan Neo4j yang akan mengolah kemudian menampilkannya ke dalam bentuk graf. Hal ini sangat berguna bagi pengelolaan data berskala besar.

Pada penelitian ini telah dijabarkan dasar perancangan basis data graf yaitu pembuatan simpul dan hubungan secara sederhana menggunakan CQL (*Chyper Query Laguage*). Dimana simpul dan hubungan tersebut memiliki ragam atribut yang dapat disesuaikan. Syarat utama dalam pembuatan simpul dan hubungan adalah masing-masing harus memiliki minimal satu properti. Selain itu, simpul harus memiliki label. Label tersebut yang berperan dalam pengelompokkan simpul. Dimana pada penelitian ini terdapat 3 label yaitu generasi 1, generasi 2, dan generasi 3.

DAFTAR PUSTAKA

- Azhar, I. I. (2018). *Teknologi Basis Data Berorientasi Graf* (Makalah IF2120 Matematika Diskrit - Sem.I Tahun 2018/2019).
- Bahri, I. S. (2020). *SILSILAH KELUARGA KIAI ENDING ZAHIDI: Suatu Jejak yang Hampir Terlupakan*.
https://books.google.com/books?hl=en&lr=&id=I2rXDwAAQBAJ&oi=fnd&pg=PA19&dq=alat+musik+sangkakala&ots=ncesfXMZ1h&sig=CDKEEoIJHKgw_MeFrKE4cinvce8

- Daniel, F., & Taneo, P. N. L. (2019). *Teori Graf* (1st ed.). DEEPUBLISH.
- Fathonah, N. (2019). Penggunaan Metode Genogram untuk Meningkatkan Kemampuan Perencanaan Karir. *Prophetic : Professional, Empathy and Islamic Counseling Journal*, 2(2), 279. <https://doi.org/10.24235/prophetic.v2i2.5814>
- Fernandes, D., & Bernardino, J. (2018). Graph Databases Comparison: AllegroGraph, ArangoDB, InfiniteGraph, Neo4J, and OrientDB. *DATA 2018 - Proceedings of the 7th International Conference on Data Science, Technology and Applications*, 373–380. <https://doi.org/10.5220/0006910203730380>
- Friedman, M. M. (2010). *Buku ajar Keperawatan Keluarga : Riset, Teori, & Praktik* (5th ed.). Jakarta EGC.
- Lokawati, H. (2016). Aplikasi Teori Graf dalam Analisa Struktur Protein. In *Makalah IF2120 Matematika Diskrit – Sem. I Tahun 2016/2017*.
- Miller, J. J. (2013). Graph database applications and concepts with Neo4j. *Proceedings of the Southern Association for Information Systems Conference, Atlanta, GA, USA, 2324*, 36.
- Setialana, P. (2017). Pencarian hubungan kekerabatan pada struktur data genealogy dalam graph database. Universitas Gadjah Mada.
- Yar, K. T., & Tun, K. M. L. (2016). Searching Personnel Relationship from Myanmar census data using Graph database and Deductive Reasoning prolog rules. *2016 International Conference on Computer Communication and Informatics (ICCCI)*, 1–7. <https://doi.org/10.1109/ICCCI.2016.7479945>