Neo4j based apps with a dynamic model

# Introduction

Creating an application often leads to the following data workflow:

Serialized data 🡪 memory data 🡪 view data-> memory data -> serialized data

The application deserializes data then optionally process it in memory then display it in a view.

A user can modify a data through a view, which is optionally processed in memory and then stored

Because the serialized object, the memory objects, and view objects are different entities, the code will have some code like this:

myMemoryObject.memoryPropertyName = mySerializedObject.serializedPropertyName;

myViewObject.viewPropertyName = myMemoryObject.memoryPropertyName;

…

Which means that the syncronisation between the different entities are based on a static key value system.

Unfortunately, during the development the system will change a lot for different reasons among:

* The serialization system is changing
* The view components are changed or upgraded
* The structure of the objects has changed, to handle new features, or improvements

When a change occurs, the developer must update the bindings in the code using some find and replace operations and recompile the application.

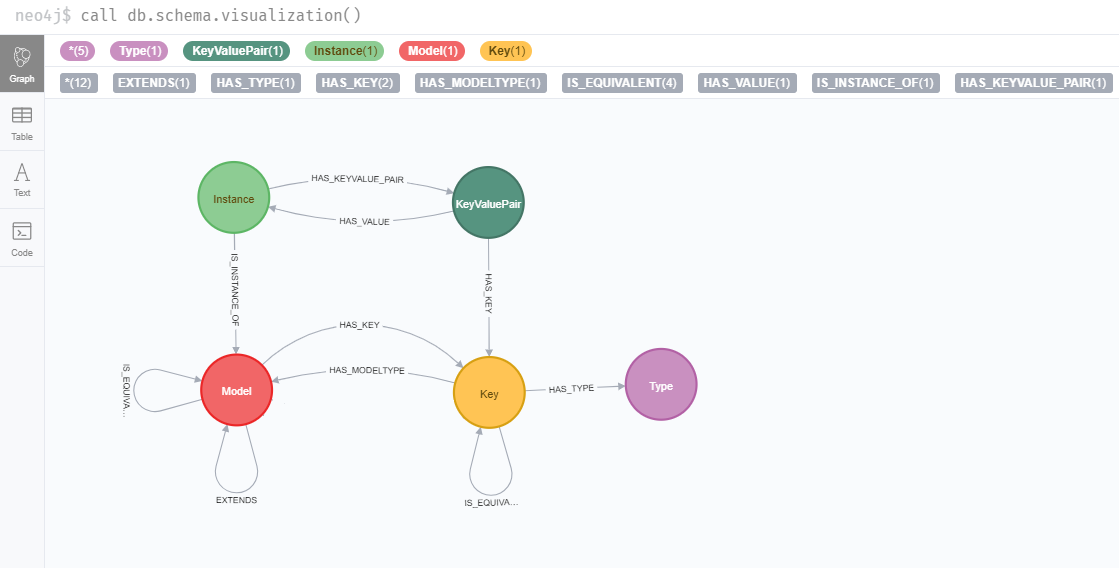
The current experimentation is trying to leverage on Graph database to have a dynamic binding system.

# Application Environment

The current sample will show a simple Angular web app, displaying data from a Neo4j Graph database.

# Database dynamic model

Here’s the schema of our graph database



It contains 5 labels:

* Model (properties: name)
* Key (properties: name, isArray)
* Instance (properties: id)
* KeyValuePair
* Type (properties: name)

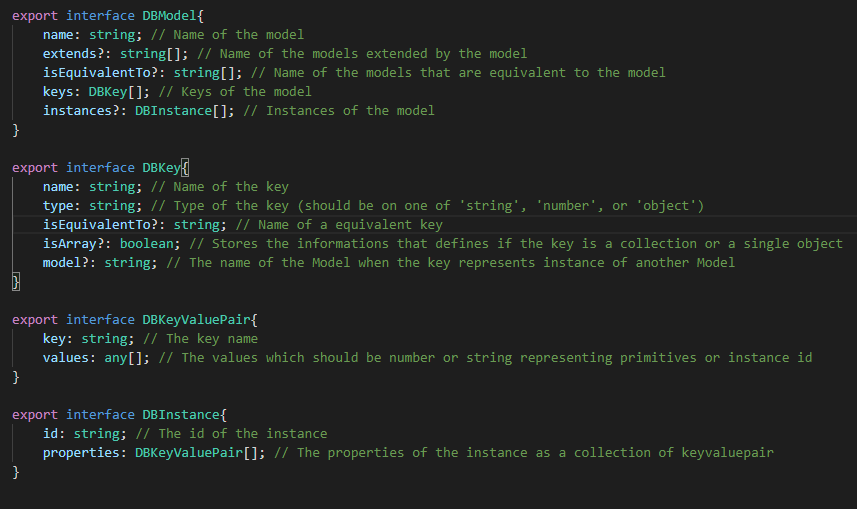
And 8 relationships:

* EXTENDS
* HAS\_TYPE
* HAS\_KEY
* HAS\_MODELTYPE
* IS\_EQUIVALENT
* HAS\_VALUE
* IS\_INSTANCE\_OF
* HAS\_KEYVALUE\_PAIR

This describe the following:

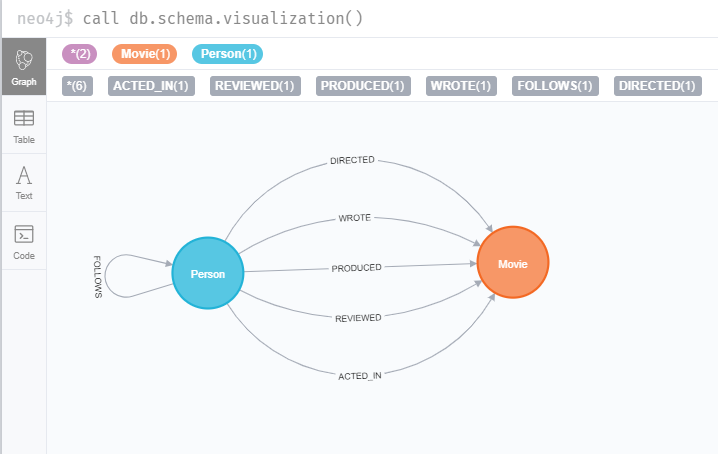
* A Model defines keys
* Instances are related to a Model
* Instances define a collection of KeyValuePair
* KeyValuePair is mapping a specific Key to a specific Instance
* A Key has a Type, and can be related to a Model (if the type is an object)
* A Model can extend other models
* A Model can be equivalent to another Model
* A Key of a model can be equivalent to the key of another Model
* Retrieving Instances of a model should give all the instances of the model but also the instances of the extending models and the instances of the equivalent models
* The Instances of a model should expose the keys of the model but also the key of extended models. It should use keys equivalency relationship to provide the appropriate key identifier

The database model should be expressed in typescript as the following:



It is considered in this document as a dynamic model, because the labels of the nodes are representing abstractions and the same model can be used whatever the objects it will store.

A static model is referring to a model with nodes and relationships based on the nature of stored objects. As an example, here is the schema of the Movie sample database in neo4j:

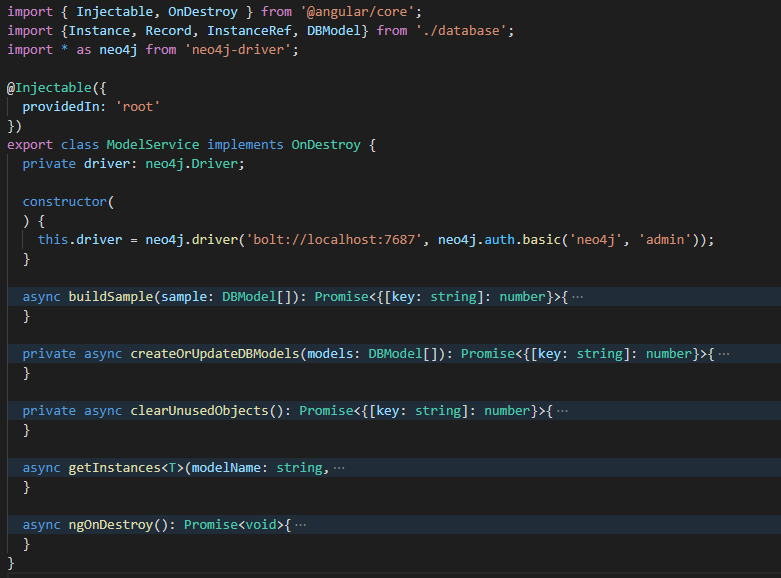


The database is storing ‘Person’ and ‘Movie’ objects which has relationships based on relationships between persons and movies. If we changed some concept in our structure, we probably have to add, remove or modify the nodes or the relationships to fit the new logic.

In a dynamic model, the structure shouldn’t change but able to represent a lot of different concepts

# Database Interaction

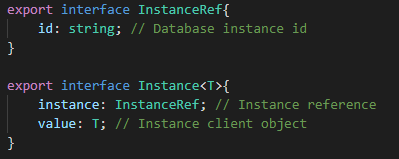
The application presented in the sample is an Angular application that will communicates directly with the graph database trough a service named ‘ModelService’:



That service is able to:

* Update the database according to a collection of DBModel objects
* Retrieve instances of a specific model with optional filtering by id option as a collection of Instance<T> objects where T will be a java object containing the key properties of the model

An Instance<T> is defined as following:



* instance property represents the instance database meta data (in the sample we will only use the id, but we can imagine adding other information like a creationDate, lastUpdate, etc)
* value property is a java object that will be used by the application

The model service has 2 public methods:

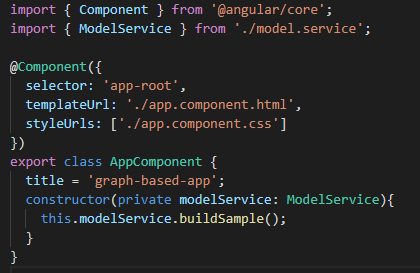
* buildSample: This will update the Database model using the private methods createOrUpdateDBModels and clearUnusedObjects providing a sample that declare the data structure and the instance
* getInstances will take a model name and will retrieve the instances for that model

It has also 2 private methods:

* createOrUpdateDBModels will be used to update the data structure of our project using a DBModel collection as an input
* clearUnusedObjects will removed object that have are no more used in the application

The queries were defined once and won’t be updated during the experimentation. They will be detailed at the end of the document.

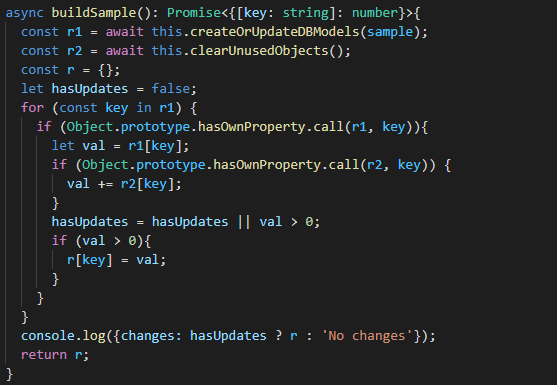
## Building the sample



The application will run the ModelService.buildSample() application once, when launching.

The purpose of the method is to adapt the database using a DBModel collection (referred as sample) representing the Models with their instance. The changes made to the database will be diffential: it means that the method will compare the existing database with the sample and create or remove the nodes and relationships to represent exactly the change. In a real project we wouldn’t use such a method, but it was created to check that a single generic query was able to modify dynamically the database.

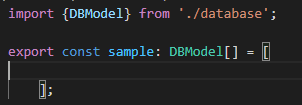
The buildSample code is detailed below:



It calls createOrUpdateDBModels method which will run a cypher query to the database providing the models then it calls clearUnusedObjects that will clean the database.

The changes of the 2 queries are aggregated and logged so we can track what are the changes made to the database.

At the beginning of the project the sample const was defined like this



Which means that de database contains nothing and should be empty

During the evolution of the project we will modify the sample to declare what we want in the database.

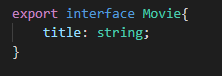
The buildsample method will apply that model to the Graph database and log the effective changes in the database.

# Let’s build a simple Movie application

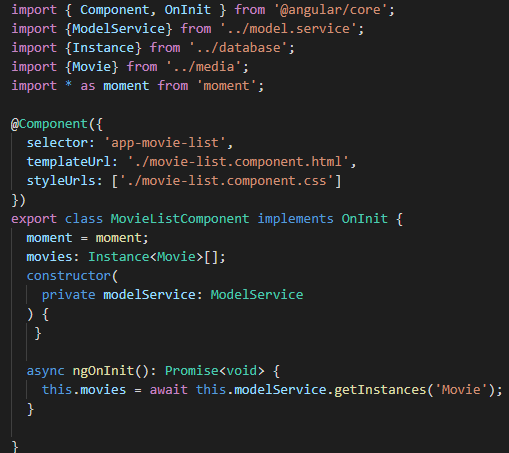
## Displaying the movies

We want to display a list of movies and we created a simple Movie interface that represents a movie which is defined by its title.

## Adding Movie model



We create then a MovieListComponent:



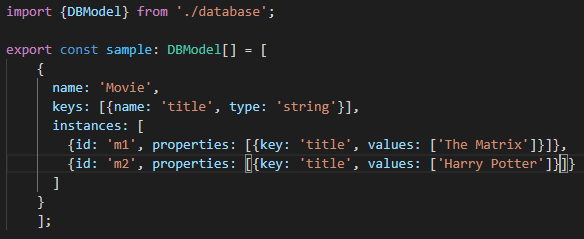
And its template

A close up of a screen

Description automatically generated

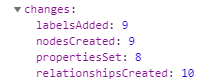
## Updating the sample

Let’s update the sample like this:



This tells the system, that we want a Model named ‘Movie’, containing the key named ‘title’ and having 2 instances ‘m1’ and ‘m2’ configured with ‘The Matrix’ and ‘Harry Potter’ as a title.

When looking into the logs, the first time the application is launched after having update the sample, we will see the following logs:



If we refresh the page or relaunch the application the logs will be:



This is because of the createOrUpdateDBModels is not incremential but differiental (the query will be detailed later).

NOTE: In the real life we probably never declare DBModel objects manually, but this is convenient for a simple experimentation.

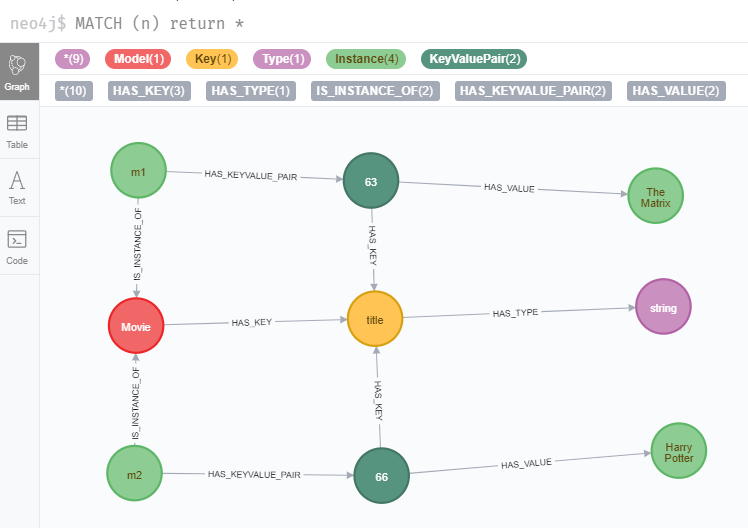
## Watching the changes

When running the application, it will display the movie list:

A picture containing drawing

Description automatically generated

Let’s now have a look at the database:



If you look to the labels we have:

* Model
* Key
* KeyValuePair
* Instance
* Type

If you look to the relationships, we have:

* IS\_INSTANCE\_OF
* HAS\_KEY
* HAS\_KEYVALUE\_PAIR
* HAS\_VALUE

And we can see exactly what we defined in the sample:

* A model, named ‘Movie’ which has a key named ‘title’
* 2 instances ‘m1’ and ‘m2’ having both a key/value pair binding the title to values which are instances having no model meaning that they are primitive
* The key as a type which is a string

Our model is considerred dynamic by the fact that ‘Movie’ is just a node labelled as ‘Model’ and ‘title’ a node labelled as ‘Key’. The abstraction model was defined statically but we can work with movies, fruits, or products, the labels and relationships will remain the same, which mean than we can have generic method based on model name that will dynamically create the desired object according to the specifications defined in the sample.

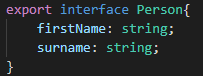
This way of doing creates more nodes than for a static model, and add complexity to the graph and some verbosity to the query, but due to the efficiency of Graph database to deal with relationships, it should also have some benefits that we will explore trying to improve the application.

# Improving the application

A movie should not only be defined by a title. It has also actors and we want to display the actors of a movie in our app.

## Creating the Person model

We create first an interface Person as following:



And a PersonComponent:

A screen shot of a smart phone

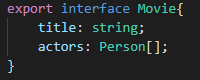
Description automatically generated

With its template:

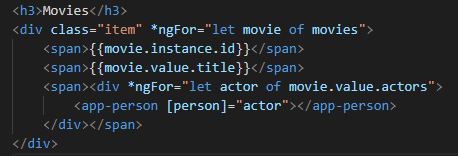


## Improving the Movie model

The movie model should know exposes an ‘actors’ property which is a Person collection:

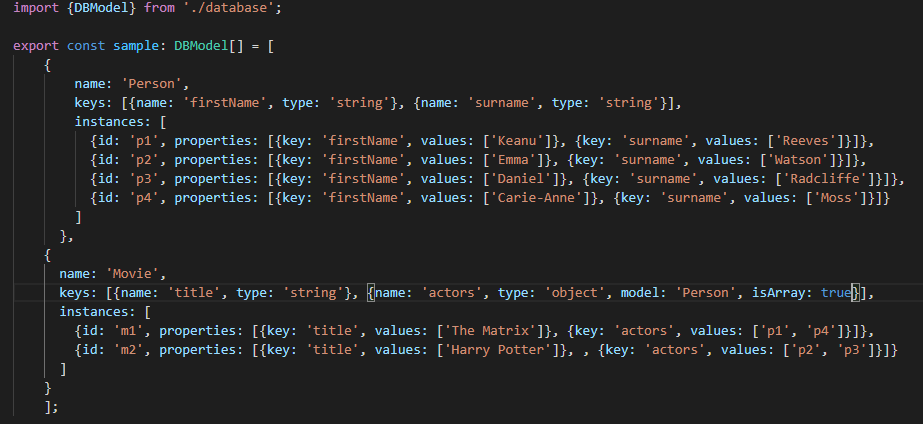


We can then improve the MovieListComponent template:



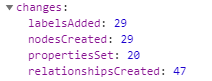
## Updating the sample

We will update the sample as following:



We have added the Model ‘Person’ and some instances of it, and added the ‘actors’ key

And the DB changes log will be:



Because of the separation between models, keys and instances it is quiet easy to update the model diffentialy, adding keys, and linking the node together.

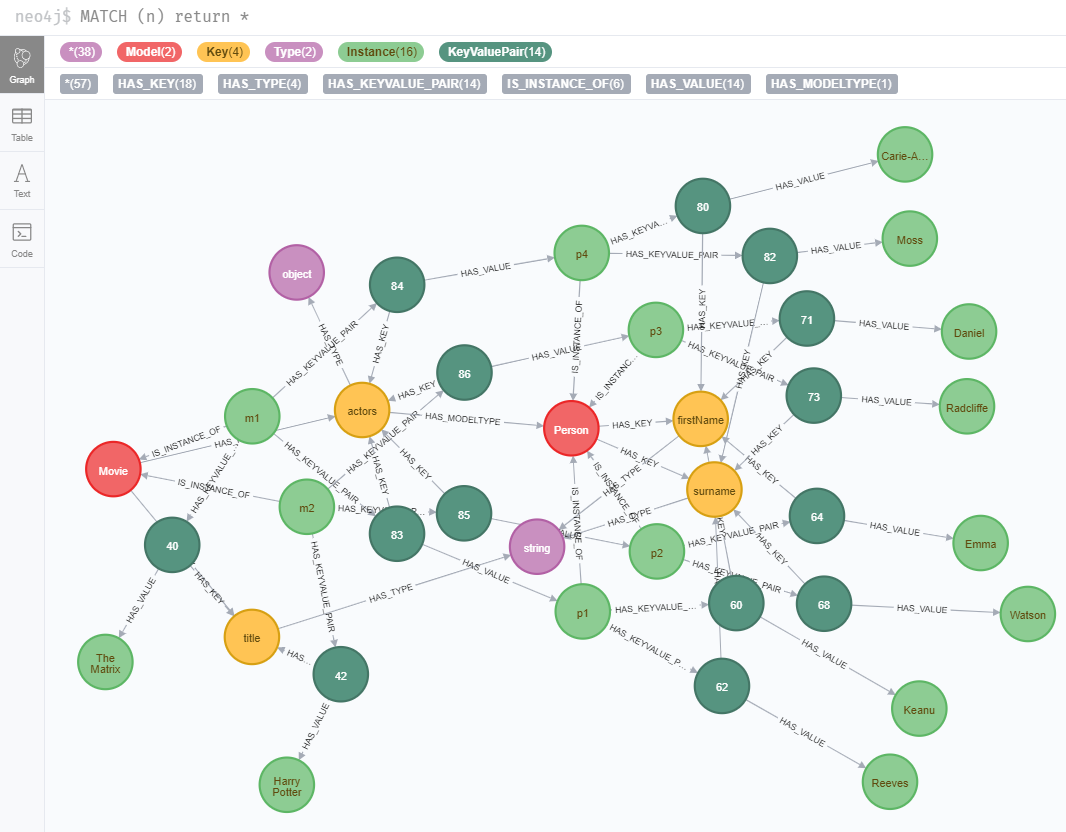
## Watching the changes

The application will now display the following:

A screenshot of a cell phone

Description automatically generated

In the database we have now



On the right there are all the primitives representing the ‘firstName’ and ‘surname’ which are related to the instances of Person. The movies have know new key/value pairs that stores the actors collection.

No new labels were added

One relationship was added: HAS\_MODELTYPE which tells to the system that actors are Person

Note that we didn’t have to change the modelService queries or the MovieListComponent. The changes in the sample has updated the database which has now all the information to return the updated Movie instances.

What will happen now, if we change the data structure of the sample?

# Changing the data structure

In our application, having the actors of a movie is nice but we should have also the character played in each movie. We need to update our model to reflect that need, creating a ‘Role’ model containing the ‘actor’ and the ‘character’ properties, and changing the ‘actors’ properties to ‘roles’ in the ‘Movie’ model.

## Adding Role model

This is the Role interface definition:

A picture containing sitting, holding, city, person

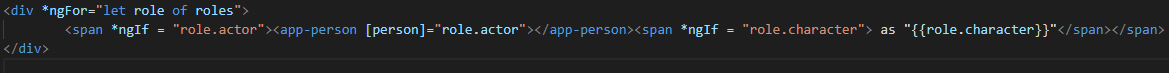
Description automatically generated

With a RoleListComponent:

A screenshot of a cell phone

Description automatically generated

And the template:



## Updating Movie model

The Movie interface is now like this:

A screenshot of a cell phone

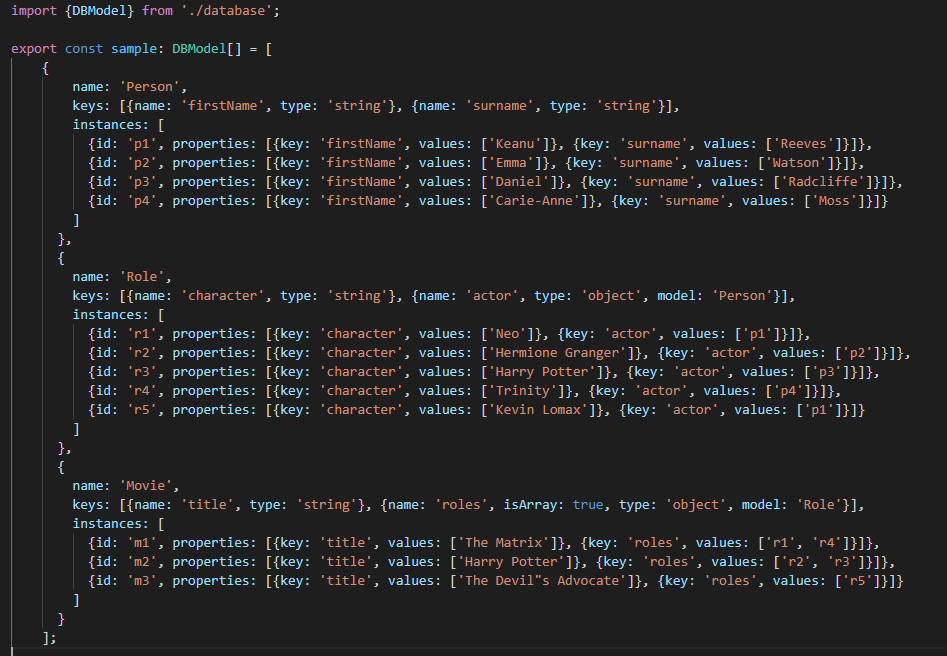
Description automatically generated

And the MovieListComponent template will be:

A screen shot of a person

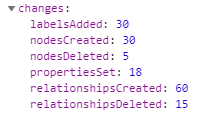
Description automatically generated

## Updating the sample



The sample reflects the changes in the model. We added the Role model, and the Movie update. We have also added a new movie.

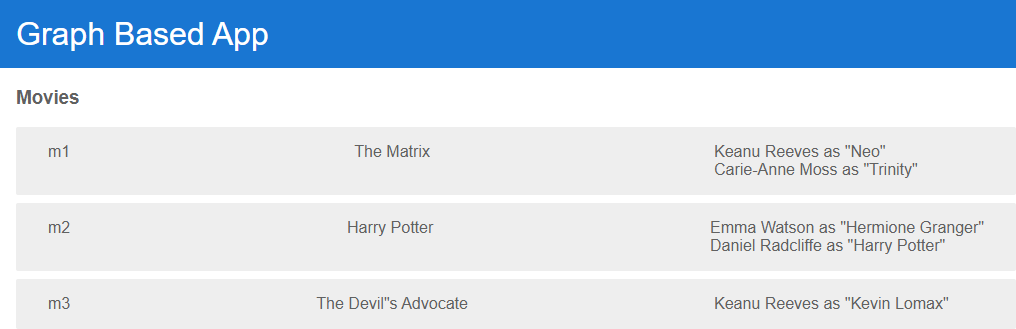
When running the application for the first time after updating the sample the changes in the databases logged by the application are:



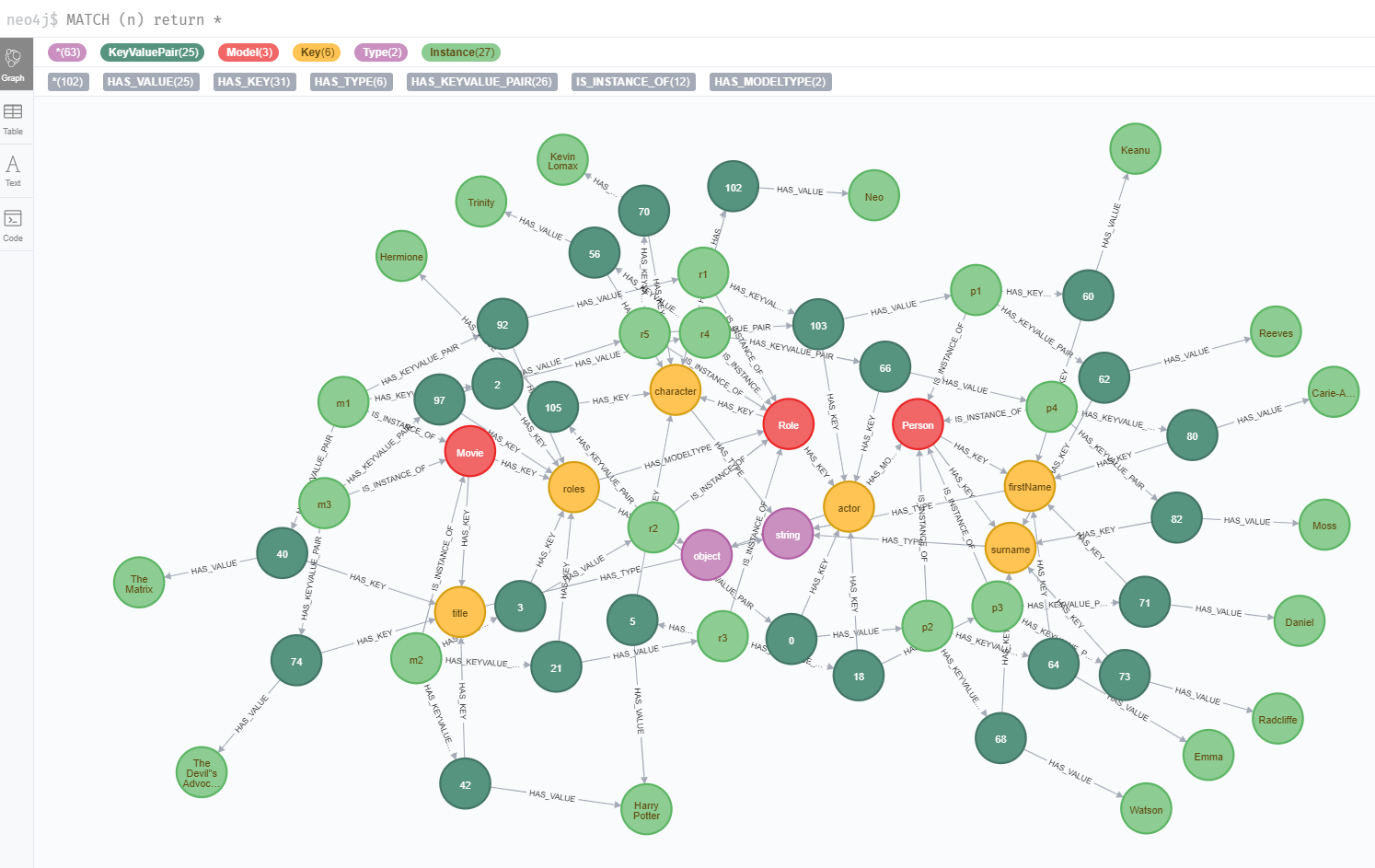
Some nodes and relationships were deleted: because the ‘actors’ properties is not used anymore, it was removed with all the key/value pairs related to it.

## Watching the changes

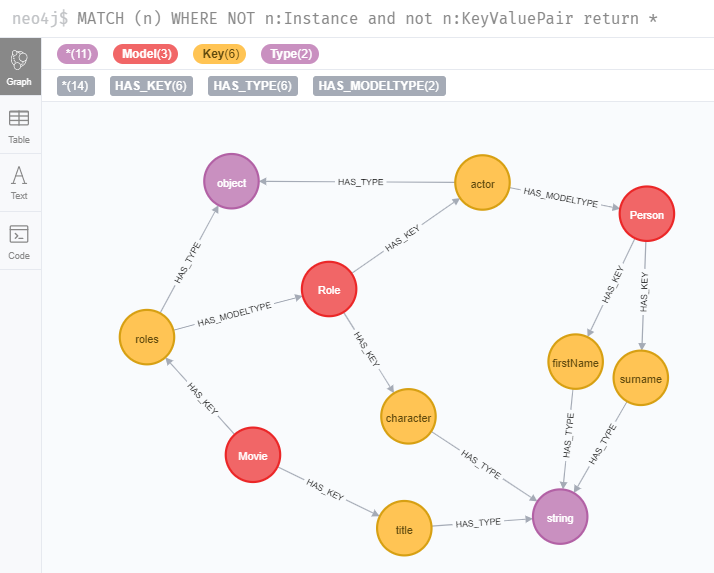
The application will now display:



The database looks now like:



Watching the entire graph beginning to become a little confusing, let’s focus on some nodes of interest

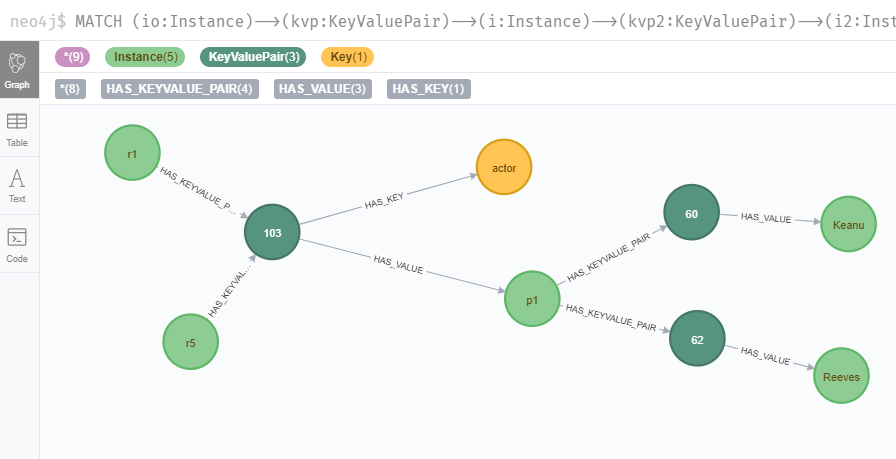


Excluding the instances and the key/value pairs we can focus on the data structure and we can see

That:

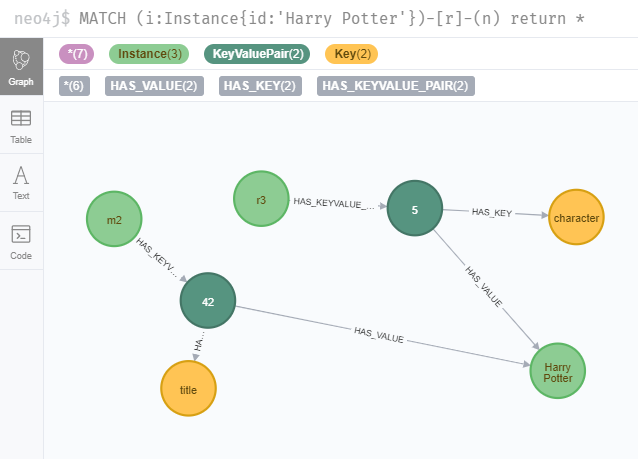
* the property ‘actors’ has disappeared from the model ‘Movie’ which has now the property ‘roles’ with ‘Role’ as model type
* The model ‘Role’ was added with 2 keys ‘character’ and ‘actor’ which is a ‘Person’
* The ‘roles’ and ‘actor’ keys have an ‘object’ type
* The other keys are defining primitives of type ‘string

Let’s have a look at Keanu Reeves now:



Because Roles ‘r1’ and ‘r5’ have both ‘p1’ as an actor (Keanu Reeves), they are sharing the same key/value pair

If we have a look at ‘Harry Potter’ primitive



Because ‘r3’ as ‘Harry Potter’ as character and ‘m2’ has ‘Harry Potter’ as title, they have both a key/value pair sharing the same primitive instance

The Model, Key, KeyValuePair, Instance system allows us to easily modify the entire structure of the data without having to destroy everything and recreate everything. We can destroy some relationships, add new ones, and then remove the nodes without specific relationships.

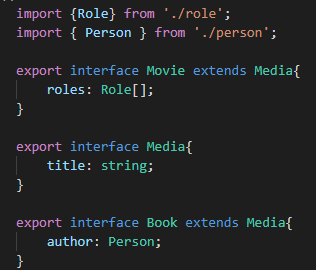
Let’s explore now how we can extend our model.

# Extending the model

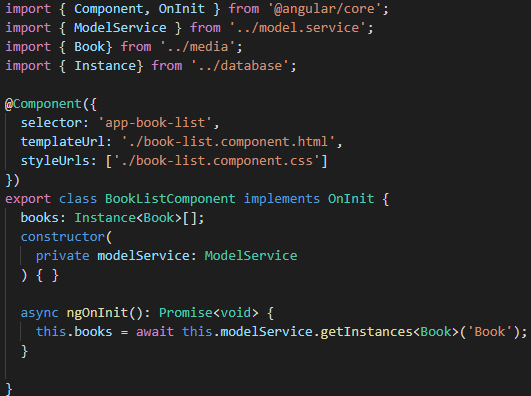
We have what we want in our movie list, but there is a new requirement for the application: display movies but also books.

## Adding Media and Book model

We will create a Media interface that will store the title. The Book interface will have an ‘author’ property which is a Person. The Movie interface will be updated to remove the ‘title’ property and both Book and Movie will extend the Media interface



This will allow us to create a BookListComponent:



With its template:

A picture containing holding, sitting, person, orange

Description automatically generated

And a MediaListComponent:



With its template:

A screen shot of a person

Description automatically generated

## Adding navigation system

We added some buttons in the TopBarComponent template with different routes:

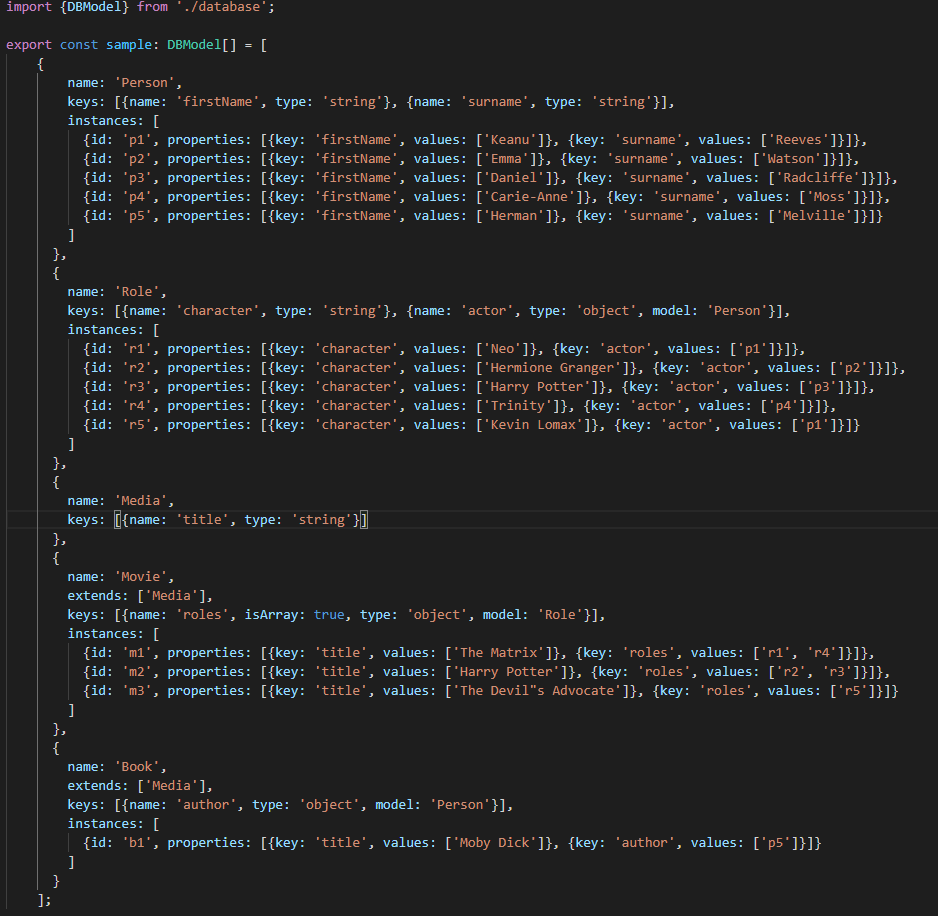
A close up of a sign

Description automatically generated

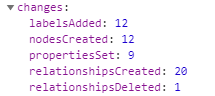
Our app having a routing module describing the following routes:



## Updating the sample

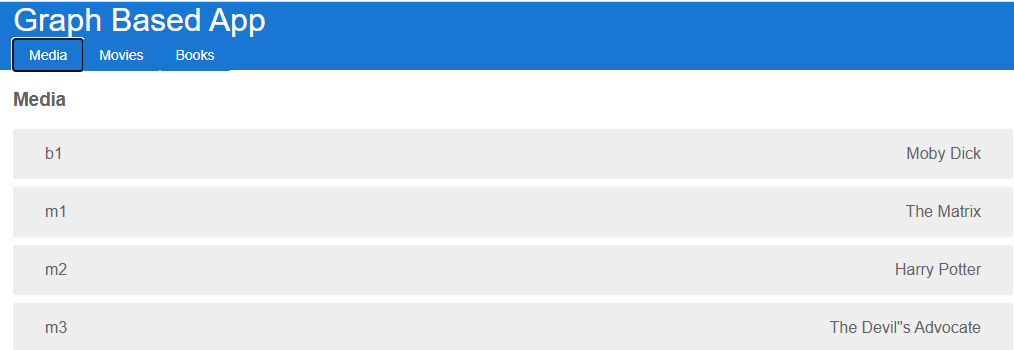


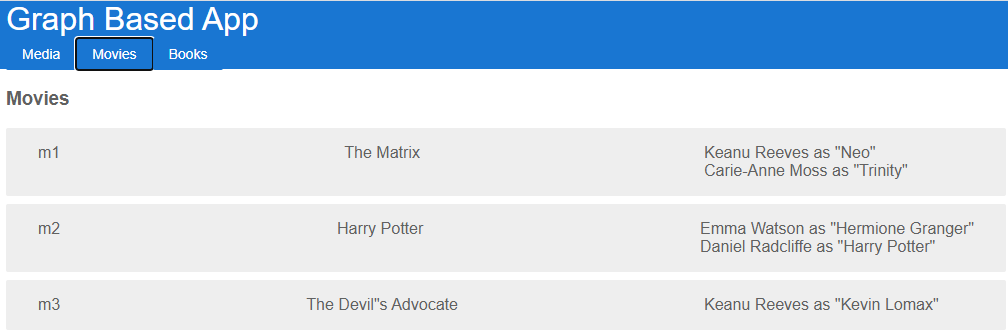
We added ‘Media’ and ‘Book’, added a ‘Person’ instance and a ‘Book’ instance and it gives us the following changes

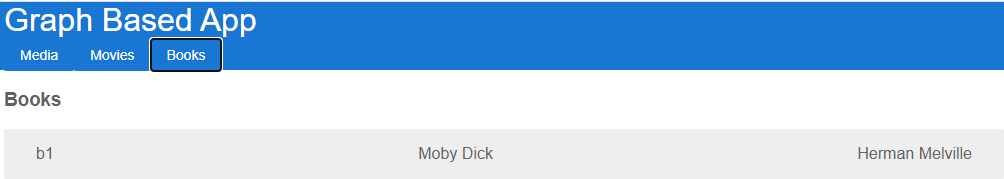


## Watching the changes

The app will display the following:

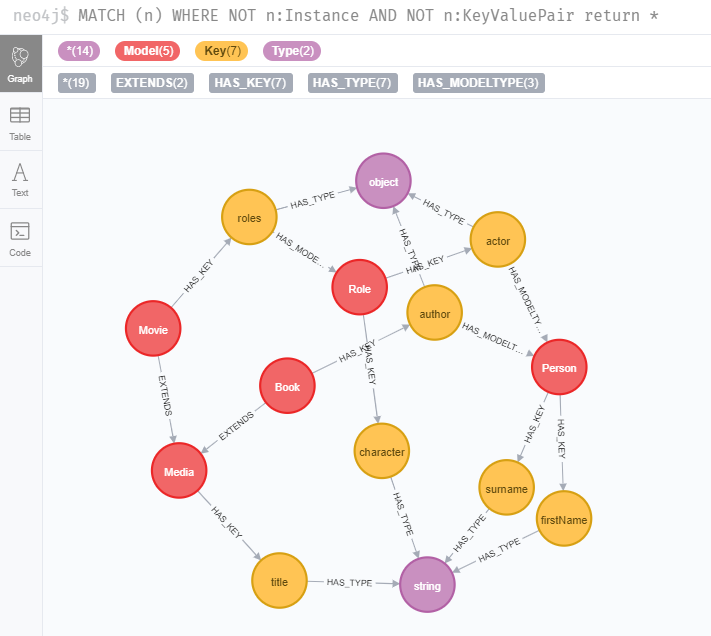






Note that the objects have been retrieved with the query getInstances, and the unique change in the query was the name of the model. Because the relation between models are described in the database the method can retrieve the appropriate objects for each model just knowing the model name.

The database model structure looks now like this:



A new relationship type was created:

* EXTENDS: tells the system how some models can extend other sharing their attributes.

The system is extensible, but what happens when introducing a new feature, that has a different model but is like existing ones.

# Merging new concepts

Let’s say that we have another app, displaying some ‘Items’ that are defined by their ‘name’.

We want to add these items to the application and be able to consider the ‘Media’ instances as ‘Items’.

## Adding Item model

First, we create the ‘Item’ interface:

A picture containing object, clock, dark, sitting

Description automatically generated

With an ItemListComponent:

A screen shot of a social media post

Description automatically generated

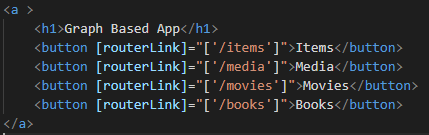
And its template:

A close up of a screen

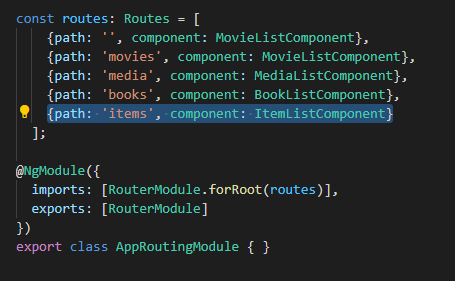
Description automatically generated

## Handling navigation

Then we add a navigation button to the TopBar:

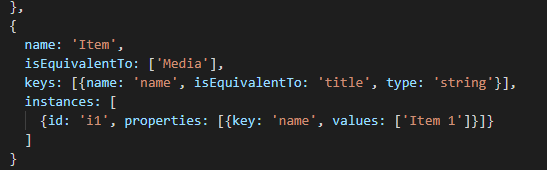


With the according route in the Routing module:

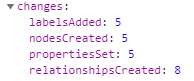


## Updating the sample

As there is no change in the other models, we just need to add the ‘Item’ concept in our DBModel collection

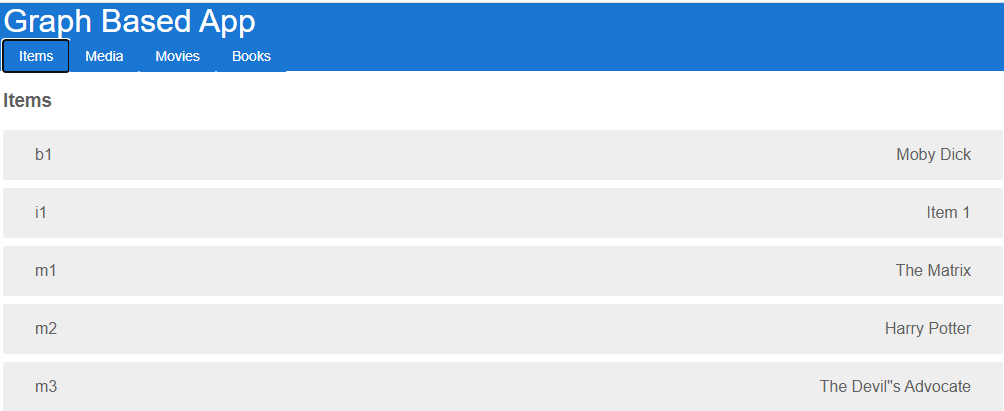


When saving the sample, the following changes appear in the logs:

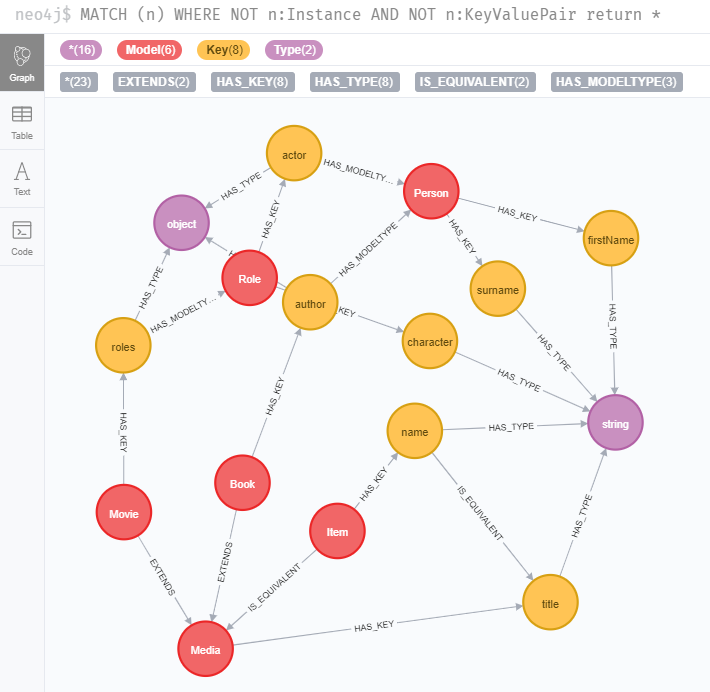


## Watching the changes

The application has now an Items panel displaying all the items and media objects



In the database, looking at the model structure:



We can see that a new relationship type was created:

* IS\_EQUIVALENT: tells the system that a model should consider the instance of other models as equivalent or that a key should be considered as equivalent to another

# Considerations about primitives

Primitives are instances without defined model. They will store number, or string. Having those primitives as nodes can be useful to retrieve transversal relationships by value between the different instances.

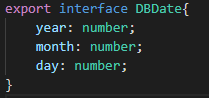
By exemple Harry Potter is a movie title or a character.

This can also interesting with dates.

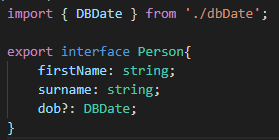
Let’s create a Date model containing the keys ‘year’, ‘month’, ‘day’

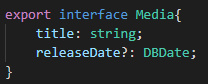
## Adding Date model

We will create first an interface DBDate model:



And add a ‘dob’ property to Person model and ‘releaseDate’ to Medial model:





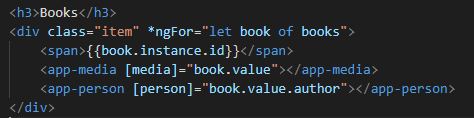
We will then create a MediaComponent:

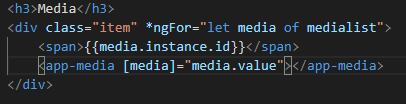


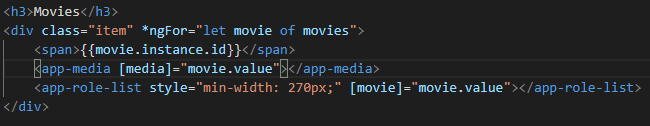
With its template:



And use the template in the MediaListComponent, BookListComponent and MovieListComponent templates:







We will also update the PersonComponent:

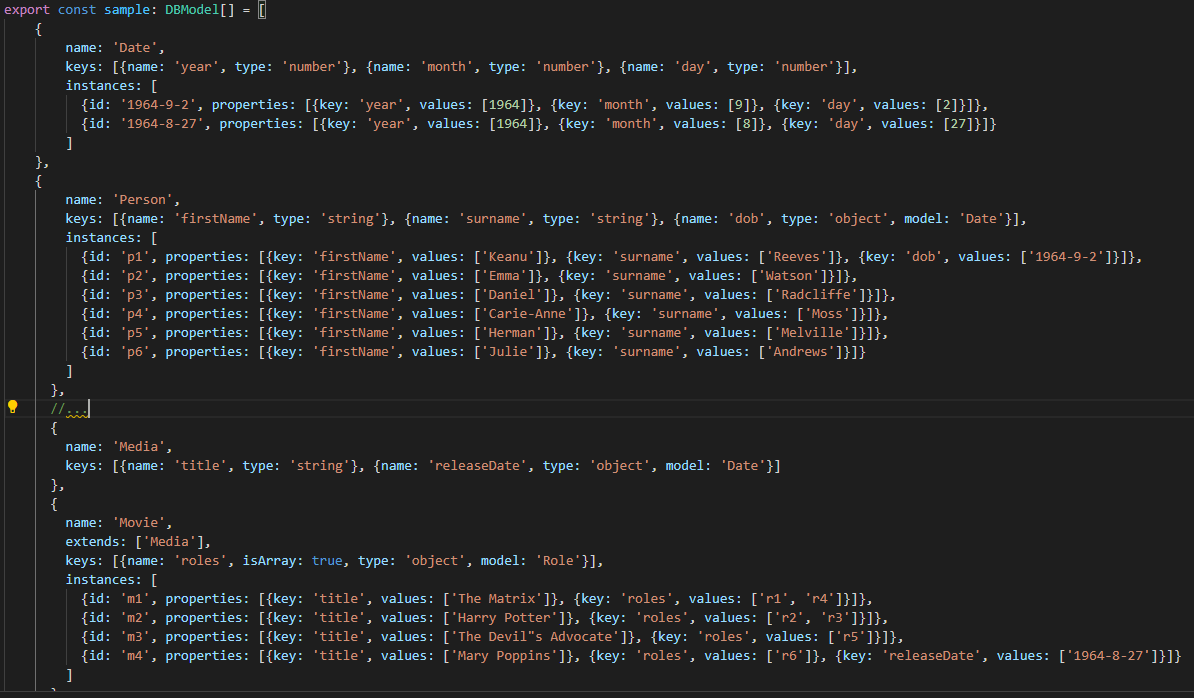


And its template:



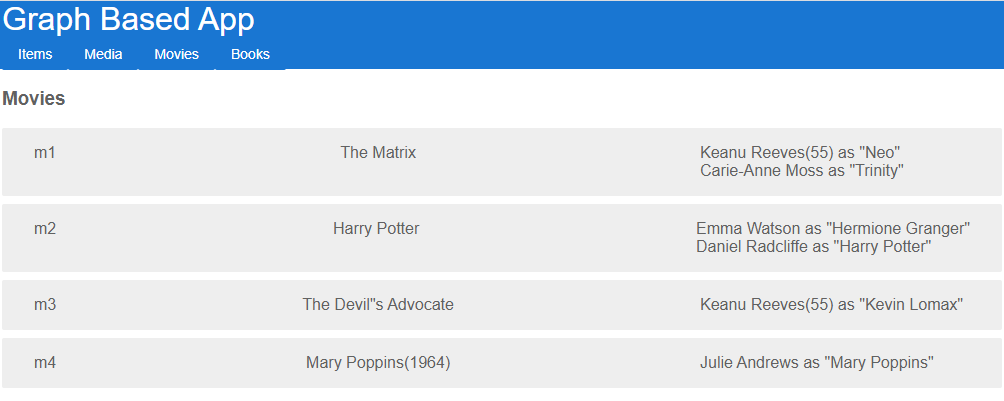
## Updating the sample

The sample is updated with some date instances:

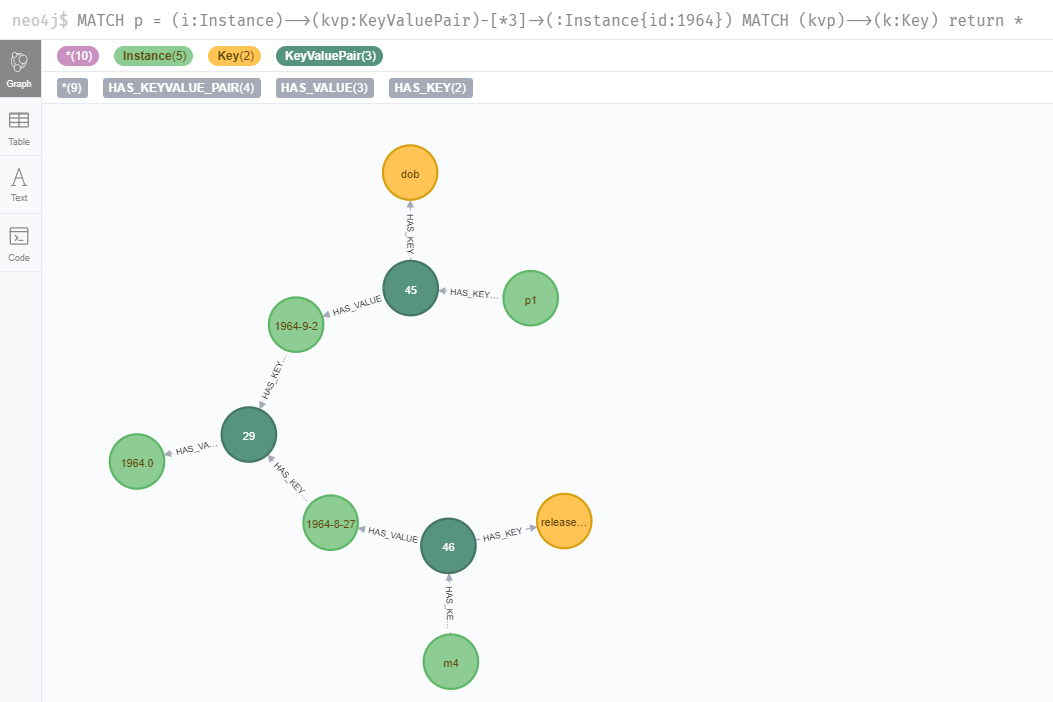


## Watching the changes

And the app is now displaying actor’s age or media release date:



In the database if we focus now on 1964 primitive:



We can find a correlation between a day of birth and the release of a movie.

The good thing having primitives as node instead as properties is that the objects of the database will define relationships by value that will be very useful when you’ll want to analyze or search your database. You can start from a Key, a Model or an Instance, you’ll always find all the object related to the node at a few hops.

Now that we have explored some typical application development use cases, let’s have a look at the queries of the ModelService.

# Queries in details

## Instances retrieval query

The ModelService has a method getInstances that will:

* Query all the instances of a specific model using a cypher query that will return json objects. The query can optionnaly filtered by instance id
* Transform json objects into Instances object that will contain the instance meta data and its object representation

### Cypher query

Here’s the code of the query:

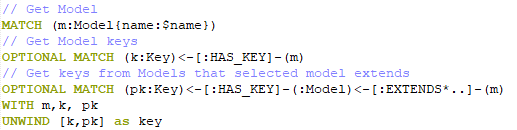


#### Getting the keys

The query first tries to reach the model by its name (the parameter of the query)

It looks then for the keys defined by the model and the keys defined by the parent models (if they exist) and uses those keys a unique collection named ‘key’.

Example: A ‘Book’ model will have ‘author’ key from book and ‘title’ key from the ‘Media’ model because ‘Book’ extends ‘Media’.

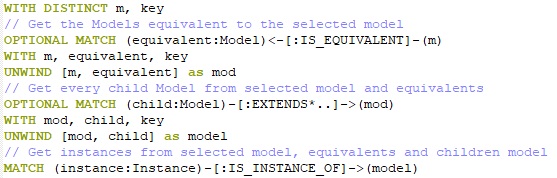


#### Getting the instances

With the model and all its keys, the query search for eventual equivalent model then for all the models extending them. The model, the equivalents and their children are gathered in a single collection a as model.

Example: an ‘Item’ model is equivalent to ‘Media’ and ‘Media’ is extended by ‘Book’ and ‘Movie’. These 4 models will be considered now as ‘model’.

The query can then get all the instances of the models

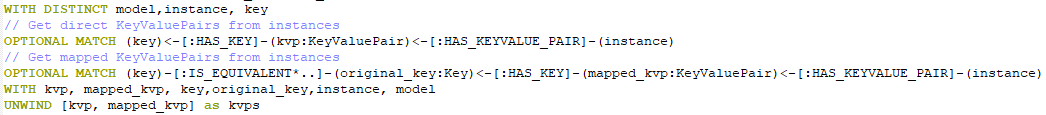


#### Getting the key/value pairs

The query will search for the KeyValuePair nodes of each instances that are related to a model key. It will also search for the KeyValuePair that might be equivalent to a model key.

Example: ‘Item’ model will have instances from ‘Book’, ‘Movie’, ‘Media’ and ‘Item’. ‘Item’ has the key ‘name’, but ‘name’ being equivalent to ‘title’, the query will gather all the KeyValuePair from the instances of ‘Book’, ‘Movie’, ‘Media’, or ‘Item’ that are related to ‘name’ or ‘title’

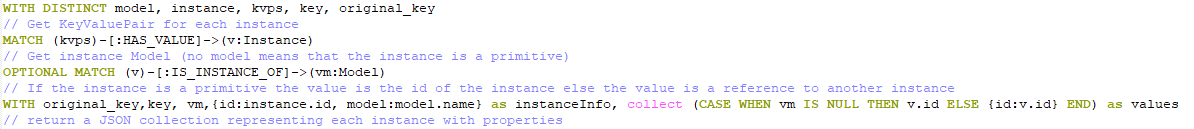
The KeyValuePair will be gathered as ‘kvps’



#### Getting the values

Now the query cand find the values associated to each KeyValuePair. It will also try to find the model associated with each value. If there is a model the instance should be considered as a reference otherwise as primitive. An instance reference will be an object defining the instance id. A primitive will return its id property as a value.

Example: A ‘Book’ has a ‘title’ and an ‘author’. The instance ‘b1’ has a value for ‘title’ which not related to any model so it will return the id of the instance which is ‘Moby Dick’. It has also a value for ‘author’ which is related to the model ‘Person’ so it will return {id: ‘p1’} which is a reference to the instance ‘p1’.

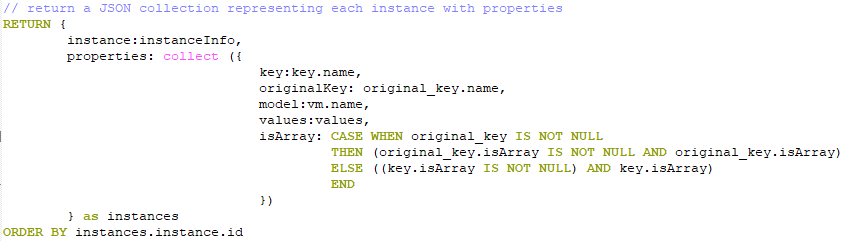


#### Formatting the result

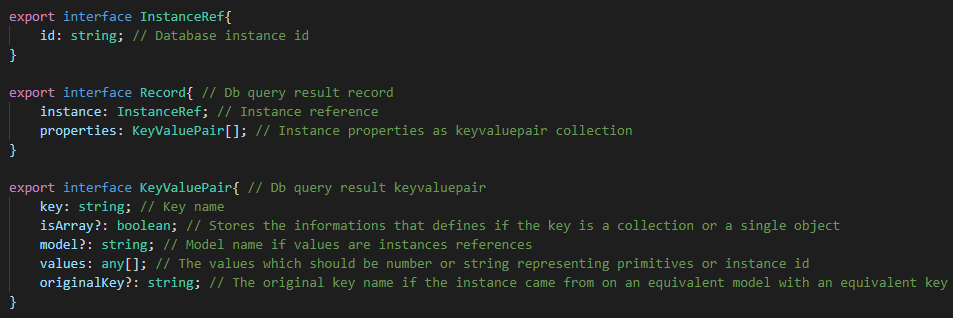
The instances will be return as an array of json object containing:

* The instance meta data (containing the id and the model name of the instance)
* The instance properties: an array of object containing:
  + The key identifier
  + The original key identifier in case of key mapping between equivalent models
  + The model in case the values are instance references
  + The values
  + The information if the key is an array or a single value

Here’s the code



In the application the result of the query is defined as collection of Record described as following:



### getInstances method

In the ModelService, the method is defined as following:



The method takes 4 arguments:

* The model’s name
* An optional InstanceReference collection that can be used as a filter
* An optional recursive index that defines how many levels of instances should be extracted (0 means all). Example: A Movie is composed by a ‘title’ and ‘roles’. The ‘roles’ property has instances of ‘Role’ which have instance of ‘Person’ in ‘actor’ property, which has an instance of ‘Date’ as ‘dob’ property. If the recursiveDepth is equal to 0, the method will recursively call itself until it reaches only primitives. If it is equal to 1, it will only return the Instances references for the ‘roles’ and not the Role objects. If it is equal to 2, it will return the Movie with the Roles objects, but the actor will be an InstanceReference.
* The current recursive level

It calls the instances Record objects using the query described above adding optionally a filter for the instances with that line:



Then for each Record, it creates an instance object setting the properties of the object dynamically using the key and values properties of the KeyValuePair collection of the Record.properties.

It is using recursivity to optionnaly retrieve the complete objects.

It returns the Instance collection that can then be used directly by the application components.

## Data Base update query

### Cypher query

The cypher query takes a DBModel array as parameter and update the database accordingly.

It removes the Model nodes and the relationships existing in the database and not in the DBModel objects and create the nodes and relationships that doesn’t exist in the database.

The query is quite long because it does a lot of thing. It was created more to be able how the structure can be updated using queries. I discovered cypher creating this project and I have to admit that I was amazed how the language is intuitive and allows to perform complex tasks in a single query.

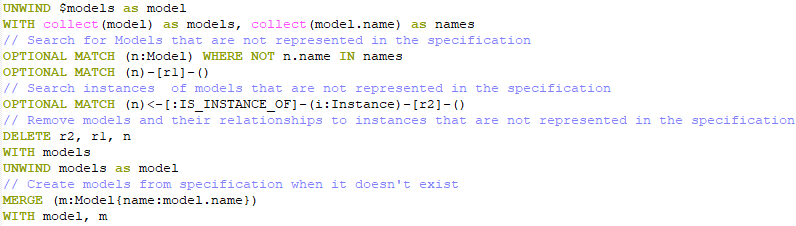
NOTE: In the real life, we won’t probably never use such a query, but it’s convenient for a POC.

IMPORTANT: To be able to perform a single query, I had to change the following setting in the database and set it to true:

cypher.lenient\_create\_relationship

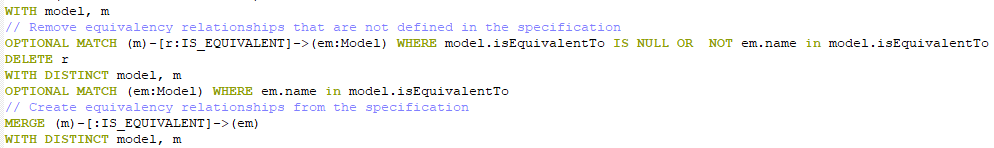
#### Removing or Creating the Models

The query first looks for the Model objects existing in the database and not in the DBModel collection and remove them. Then it merges the Model objects using the name property of the node.



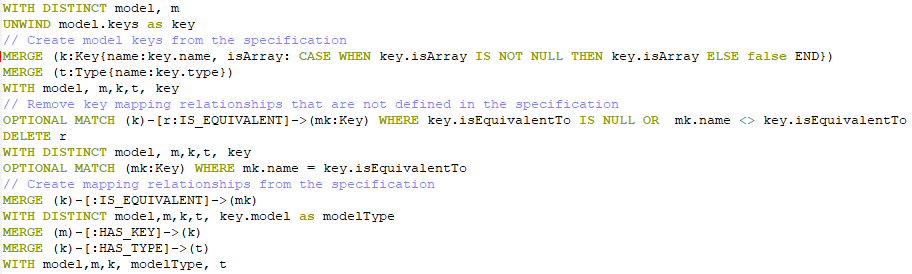
#### Removing or Creating Equivalency relationships

The query first looks for the relationships that are not defined in the specification (the DBModel collection) and remove the relationships. It will then create the [:IS\_EQUIVALENT] relationship between the Model objects according to the specification.



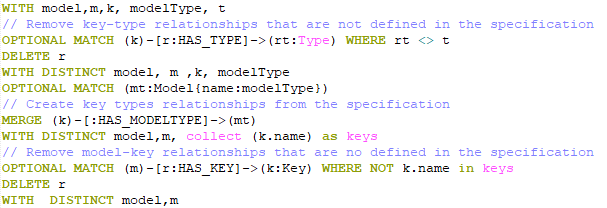
#### Creating Keys, Types and Removing or Creating Equivalency relationships between Keys

The query creates then the Key nodes related to a model. It merges the Type object for each key and remove or merge the [:IS\_EQUIVALENT] relationships according to the specification. It creates then [:HAS\_KEY] and [:HAS\_TYPE] relationships that links models to keys and keys to types



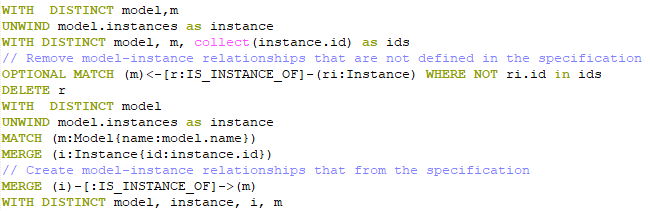
#### Updating Key and Model relationships

The query delete then the relationships that are not defined in the specification and create optionally the [:HAS\_MODELTYPE] relationships between a key and its model when a key is related to another object. It removes then all the [:HAS\_KEY] relationships between model and keys that are not defined in the specification



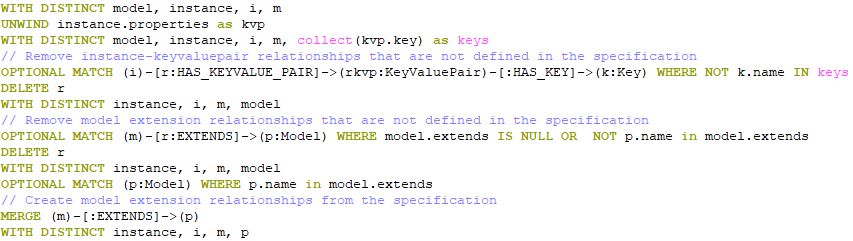
#### Merge Instance and Update Instance and Model relationships

The query then remove the relationships [:IS\_INSTANCE\_OF] that are not defined in the specification. Then merge Instance defined for each model and create the relationship between them.



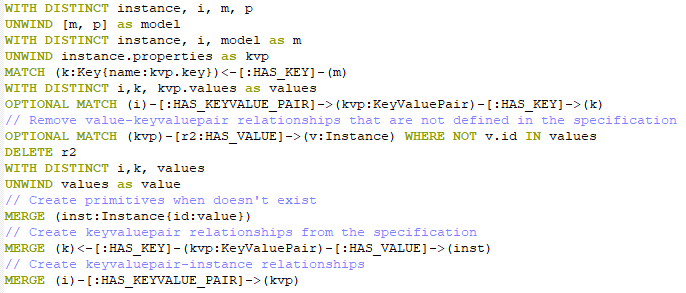
#### Updating Instance/KeyValuePair relationships and extends relationships between Model objects

The query continues removing unspecified relationships between the Instance and KeyValuePair nodes then update the [:EXTENDS] relationships between Model objects according to the specification



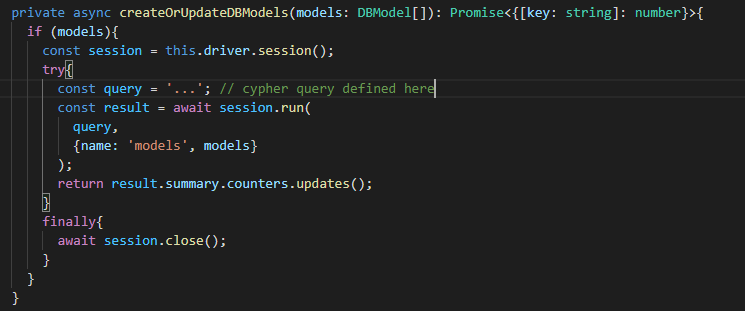
#### Merging Value Instances, KeyValuePair and updating relationships

Finally, the KeyValuePair / Value relationships are updated, merging the Instance by id.



### createOrUpdateDBModels method

The method simply the cypher query described above and return the update information from the server (number of node,labels,relationships created or deleted).



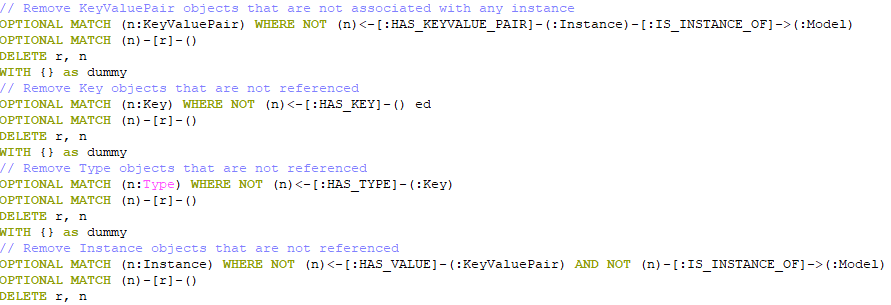
## Remove unused objects query

### Cypher query

The previous query might have removed some relationships and the database will have nodes that have no more meaning. That query ensures that every unwanted object is removed from the database.

The query removes:

* The KeyValuePair nodes that are not associated with any instance
* The Key nodes that are not referenced
* The Type nodes that are not referenced
* The Instance nodes that are not referenced



### clearUnusedObjects method

In the ModelService the method is simply calling the cypher query described above returning the counters.

