Krzysztof Ostrzycki 193507, Łukasz Turek 193395

Graph Database project – raport

1. Business Domain

Descritpion: This project analyzes relationships between users, movies, actors, directors, and genres using a graph-based approach. The interconnected data structure reveals patterns and insights into user preferences and movie dynamics.

Goal: The goal is to uncover trends in ratings, genre preferences, and industry relationships.

2. Assumptions and constraints:

- All users have valid and complete data (name, surname, birth date, etc.).
- Every movie belongs to at least one genre.
- Ratings are integers between 1 and 5, inclusive.
- Users rate only movies they have watched.

3. CREATE statement

```
CREATE
// Users
(user1:User {name: "Alice", surname: "Johnson", birthDate: date("1999-05-15")}),
(user2:User {name: "Bob", surname: "Smith", birthDate: date("1994-07-20")}),
(user3:User {name: "Charlie", surname: "Brown", birthDate: date("2002-03-12")}),
(user4:User {name: "Diana", surname: "Green", birthDate: date("1996-11-30")}),
(user5:User {name: "Eve", surname: "White", birthDate: date("1989-02-28")}),
(user6:User {name: "Frank", surname: "Black", birthDate: date("1983-09-10")}),

// Movies
(movie1:Movie {title: "Inception", releaseYear: 2010}),
(movie2:Movie {title: "Titanic", releaseYear: 1997}),
(movie3:Movie {title: "The Matrix", releaseYear: 1999}),
(movie4:Movie {title: "Avatar", releaseYear: 2009}),

// Actors
(actor1:Actor {name: "Leonardo", surname: "DiCaprio"}),
(actor2:Actor {name: "Keanu", surname: "Reeves"}),
(actor3:Actor {name: "Joseph", surname: "Gordon-Levitt"}),
(actor4:Actor {name: "Kate", surname: "Winslet"}),
(actor5:Actor {name: "Carrie-Anne", surname: "Moss"}),
(actor6:Actor {name: "Laurence", surname: "Fishburne"}),
(actor7:Actor {name: "Matthew", surname: "McConaughey"}),

// Directors
(director1:Director {name: "Christopher", surname: "Nolan"}),
(director2:Director {name: "James", surname: "Cameron"}),
(director3:Director {name: "Lana", surname: "Wachowski"}),
(director4:Director {name: "Lilly", surname: "Wachowski"}),
```

```
// Genres
(genre1:Genre {name: "Sci-Fi"}),
(genre2:Genre {name: "Drama"}),
(genre3:Genre {name: "Action"}),

// Relationships
// Ratings
(user1)-[:RATED {rating: 1}]->(movie1),
(user2)-[:RATED {rating: 4}]->(movie1),
(user3)-[:RATED {rating: 5}]->(movie1),
(user1)-[:RATED {rating: 2}]->(movie2),
(user2)-[:RATED {rating: 4}]->(movie2),
(user4)-[:RATED {rating: 3}]->(movie2),
(user6)-[:RATED {rating: 4}]->(movie2),
(user5)-[:RATED {rating: 5}]->(movie2),
(user3)-[:RATED {rating: 5}]->(movie3),
(user4)-[:RATED {rating: 4}]->(movie3),
(user5)-[:RATED {rating: 4}]->(movie3),
(user6)-[:RATED {rating: 3}]->(movie3),
(user1)-[:RATED {rating: 4}]->(movie4),
(user5)-[:RATED {rating: 3}]->(movie4),
(user2)-[:RATED {rating: 2}]->(movie4),

//Actor to Movie Connections
(actor1)-[:ACTED_IN]->(movie1),
(actor3)-[:ACTED_IN]->(movie1),
(actor2)-[:ACTED_IN]->(movie2),
(actor4)-[:ACTED_IN]->(movie2),
(actor2)-[:ACTED_IN]->(movie3),
(actor5)-[:ACTED_IN]->(movie3),
(actor6)-[:ACTED_IN]->(movie3),
(actor4)-[:ACTED_IN]->(movie4),
(actor7)-[:ACTED_IN]->(movie4),
```

```
// Directors of Movies
(director1)-[:DIRECTED]->(movie1),
(director2)-[:DIRECTED]->(movie2),
(director3)-[:DIRECTED]->(movie3),
(director2)-[:DIRECTED]->(movie4),
(director4)-[:DIRECTED]->(movie3),

// Genres of Movies
(movie1)-[:IN_GENRE]->(genre1),
(movie1)-[:IN_GENRE]->(genre3),
(movie2)-[:IN_GENRE]->(genre2),
(movie2)-[:IN_GENRE]->(genre3),
(movie3)-[:IN_GENRE]->(genre1),
(movie3)-[:IN_GENRE]->(genre3),
(movie4)-[:IN_GENRE]->(genre2),
(movie4)-[:IN_GENRE]->(genre1),

// Users liking Genres
(user1)-[:LIKES]->(genre1),
(user1)-[:LIKES]->(genre3),
(user2)-[:LIKES]->(genre2),
(user3)-[:LIKES]->(genre1),
(user3)-[:LIKES]->(genre3),
(user4)-[:LIKES]->(genre3),
(user5)-[:LIKES]->(genre1),
(user6)-[:LIKES]->(genre2),
(user6)-[:LIKES]->(genre3);
```

4. Examples of nodes

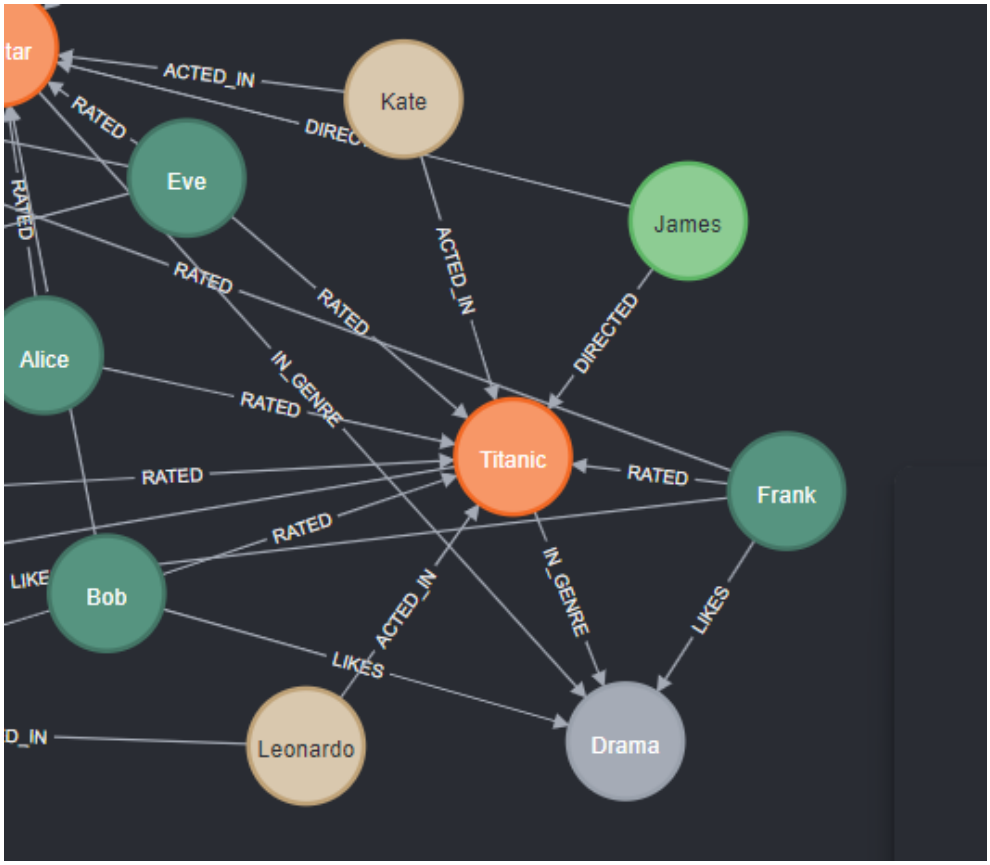

```
{
  "identity": 18,
  "labels": [
    "User"
  ],
  "properties": {
    "surname": "Black",
    "name": "Frank",
    "birthDate": "1983-09-10"
  },
  "elementId": "4:9c901ccd-cc9d-4d15-ad4c-be65d5676129:18"
}
```

```
{
  "identity": 20,
  "labels": [
    "Movie"
  ],
  "properties": {
    "title": "Titanic",
    "releaseYear": 1997
  },
  "elementId": "4:9c901ccd-cc9d-4d15-ad4c-be65d5676129:20"
}
```

```
{
  "identity": 23,
  "labels": [
    "Actor"
  ],
  "properties": {
    "surname": "DiCaprio",
    "name": "Leonardo"
  },
  "elementId": "4:9c901ccd-cc9d-4d15-ad4c-be65d5676129:23"
}
```

```
{
  "identity": 5,
  "labels": [
    "Genre"
  ],
  "properties": {
    "name": "Drama"
  },
  "elementId": "4:9c901ccd-cc9d-4d15-ad4c-be65d5676129:5"
}
```

```
{
  "identity": 1,
  "labels": [
    "Director"
  ],
  "properties": {
    "surname": "Cameron",
    "name": "James"
  },
  "elementId": "4:9c901ccd-cc9d-4d15-ad4c-be65d5676129:1"
}
```

5. Relations

RATED: Connects a user to a movie with a rating attribute.

Example: (User1)-[:RATED {rating: 4}]->(Movie1)

LIKES: Connects a user to genres they prefer.

Example: (User1)-[:LIKES]->(Genre1)

ACTED_IN: Connects actors to movies they acted in.

Example: (Actor1)-[:ACTED_IN]->(Movie1)

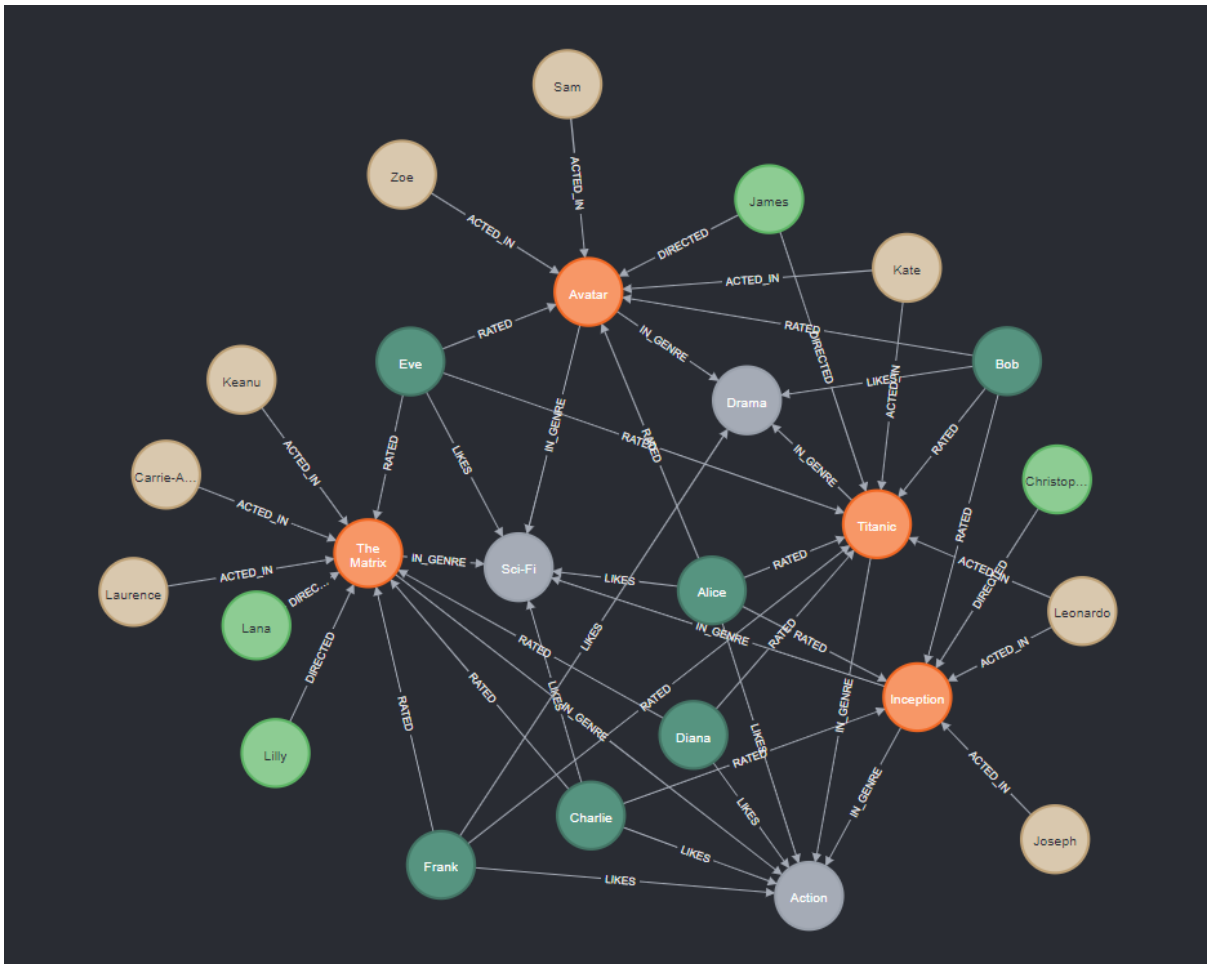
DIRECTED: Connects directors to the movies they directed.

Example: (Director1)-[:DIRECTED]->(Movie1)

IN_GENRE: Connects movies to their genres.

Example: (Movie1)-[:IN_GENRE]->(Genre1)

6. Graph



7. Competency questions:

- **What is the average rating for each movie?**

```
MATCH (movie:Movie)←[r:RATED]-(user:User)
RETURN movie.title AS Title, round(AVG(r.rating), 2) AS AverageRating;
```

	Title	AverageRating
1	"Inception"	3.33
2	"Titanic"	3.6
3	"Avatar"	3.0
4	"The Matrix"	4.0

- **How many movies did each user rate, and what is their average rating?**

```
MATCH (u:User)-[r:RATED]→(m:Movie)
RETURN u.name AS User, COUNT(r) AS NumberOfRatings, round(AVG(r.rating), 2) AS AverageRating
ORDER BY AverageRating DESC;
```

	User	NumberOfRatings	AverageRating
1	"Charlie"	2	5.0
2	"Eve"	3	4.0
3	"Diana"	2	3.5
4	"Frank"	2	3.5
5	"Bob"	3	3.33

- **What movies belong to each genre?**

```
MATCH (genre:Genre)←[:IN_GENRE]-(movie:Movie)
RETURN genre.name AS Genre, collect(movie.title) AS Movies;
```

	Genre	Movies
1	"Sci-Fi"	["Inception", "The Matrix", "Avatar"]
2	"Action"	["Inception", "Titanic", "The Matrix"]
3	"Drama"	["Titanic", "Avatar"]

- **Which movie has the most ratings?**

```
MATCH (movie:Movie)←[:RATED]-(user:User)
RETURN movie.title AS Title, COUNT(user) AS NumberOfRatings
ORDER BY NumberOfRatings DESC
LIMIT 1;
```

	Title	NumberOfRatings
1	"Titanic"	5

- **Which actors have acted in Sci-Fi or Action movies?**

```
MATCH (actor:Actor)-[:ACTED_IN]→(movie:Movie)-[:IN_GENRE]→(genre:Genre)
WHERE genre.name IN ["Sci-Fi", "Action"]
RETURN DISTINCT actor.name AS FirstName, actor.surname AS LastName;
```

	Name	Surname
1	"Leonardo"	"DiCaprio"
2	"Joseph"	"Gordon-Levitt"
3	"Keanu"	"Reeves"
4	"Carrie-Anne"	"Moss"
5	"Laurence"	"Fishburne"
6	"Kate"	"Winslet"

- **What is the best-rated movie in each genre?**

```
MATCH (user:User)-[r:RATED]→(movie:Movie)-[:IN_GENRE]→(genre:Genre)
WITH genre, movie, ROUND(AVG(r.rating),2) AS AverageRating
ORDER BY genre, AverageRating DESC
WITH genre, COLLECT(movie { .title, AverageRating })[0] AS BestMovie
RETURN genre.name AS Genre, BestMovie.title AS Title, BestMovie.AverageRating AS AverageRating
```

	Genre	Title	AverageRating
1	"Sci-Fi"	"The Matrix"	4.0
2	"Drama"	"Titanic"	3.6
3	"Action"	"The Matrix"	4.0

- **What is the average rating for movies rated by users older than 30?**

```
MATCH (user:User)-[r:RATED]→(movie:Movie)
WHERE duration.between(user.birthDate, date()).years ⩾ 30
RETURN movie.title AS Title, ROUND(AVG(r.rating),2) AS AverageRating, COUNT(r.rating) AS NumberOfRatings
```

	Title	AverageRating	NumberOfRatings
1	"Inception"	4.0	1
2	"Titanic"	4.33	3
3	"Avatar"	2.5	2
4	"The Matrix"	3.5	2

- **Movie Recommendations Based on Shared Users.**

```
MATCH (m:Movie {title:'The Matrix'})←[:RATED]-(u:User)-[:RATED]→(rec:Movie)
RETURN
    rec.title AS recommended_movie,
    COUNT(u) AS shared_users,
    COLLECT(u.name) AS shared_user_names
ORDER BY shared_users DESC
```

	recommended_movie	shared_users	shared_user_names
1	"Titanic"	3	["Diana", "Eve", "Frank"]
2	"Inception"	1	["Charlie"]
3	"Avatar"	1	["Eve"]