Biblio-Tech (Books Database)

Fall 2024 - SOEN 363: Data Systems for Software Engineers
Phase 2 Submission

Team members
Vinisha Manek (40229456) – Team lead
Ziad Elsharkawi (40213438)
Adam Chami (40248165)
Mohammed Al Assad (40252007)

Database
https://drive.google.com/drive/folders/1dU_s_94cjaExObDDTMhKyx6vKED0cBph?usp=drive_link

Data Migration
- Source: PostgreSQL relational database from Phase 1
- Destination: Neo4j graph database
- Migration tool: Custom Python script (transfer.py)

Project Structure
- transfer.py: Script for migrating data from PostgreSQL to Neo4j
- cypher.txt: Contains all Cypher query implementations
- execute_cypher.py: Script for performance testing and query execution
- nosql.png: Neo4j database (graph) diagram
- bookdatabase_backup.sql: not submitted on moodle due to size (260MB)
https://drive.google.com/drive/folders/1dU_s_94cjaExObDDTMhKyx6vKED0cBph?usp=drive_link

Database Schema Changes
Modifications from relational model to graph model:
- Simplified book properties by merging PhysicalBook and EBook attributes (format, is_ebook, ebook_url) into the Book node
- Transformed weak entity (Ratings) as properties into the Book node

Performance Optimizations
- Created indexes for frequently queried properties in transfer.py
- Documented query execution times before/after indexing in execute_cypher.py
- Implemented full-text search capability
- Performance comparisons with relational database queries between execute_cypher.py and execute_relational.py (Phase 1)