Advanced Database Systems: Part 3 – Exercise

Matthias Lanzinger

Institute of Logic and Computation
Databases & Artificial Intelligence Grroup

May 2019

Table of Contents

Exercise 1 – Distributed Joins

Try to compute the communication cost (in number of bytes transferred) of the distributed join strategies discussed in the lecture.

Exercise 1 – Submission

Document your computations, results and comparisions in your report.

Exercise 2 – Denormalization

Your task is to translate a relational data model to a data model for use with a document store.

- Denormalize in a way that optmizes the performance of the two queries + the updates described in the exercise sheet.
- You should produce only one data model that optimizes for all the the requirements at the same time (as much as possible).

Exercise 2 – Submission

In report:

- Brief description of your denormalizaiton decisions.
- In particular, describe new attributes and tradeoffs.

JSON files:

- Translate the given database instance into JSON files in the form of your new data model.
- You can have multiple (or even all) entities in the same file.
- Submit your JSON files with your report.
- Make sure they are valid JSON
  (e.g., using `https://jsonlint.com/`)

Exercise 3 – Graph Databases

- Learn about graph database concepts by experimenting with Neo4j and the Offshore Leaks Dataset.
- Download Neo4j includig the dataset: `https://offshoreleaks.icij.org/pages/database`.

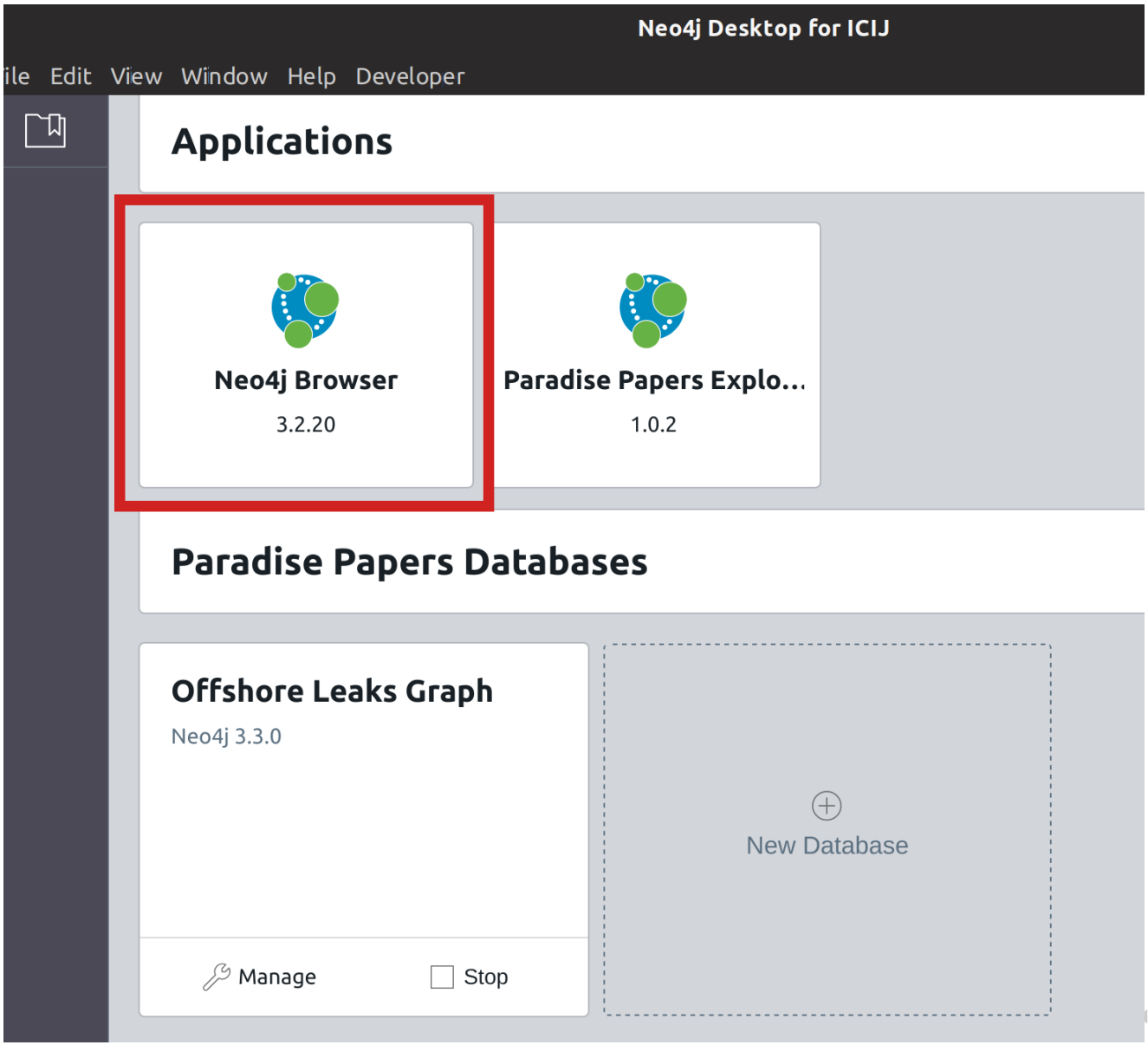
Exercise 3 – Installation

After downloading and launching Neo4j Desktop for ICIJ, all the required components and datasets are available. Once everything is installed the Neo4j Browser can be launched with the click of a button. An introductory guide that explains the structure of the data and how to use it for analysis and exploration will be available.

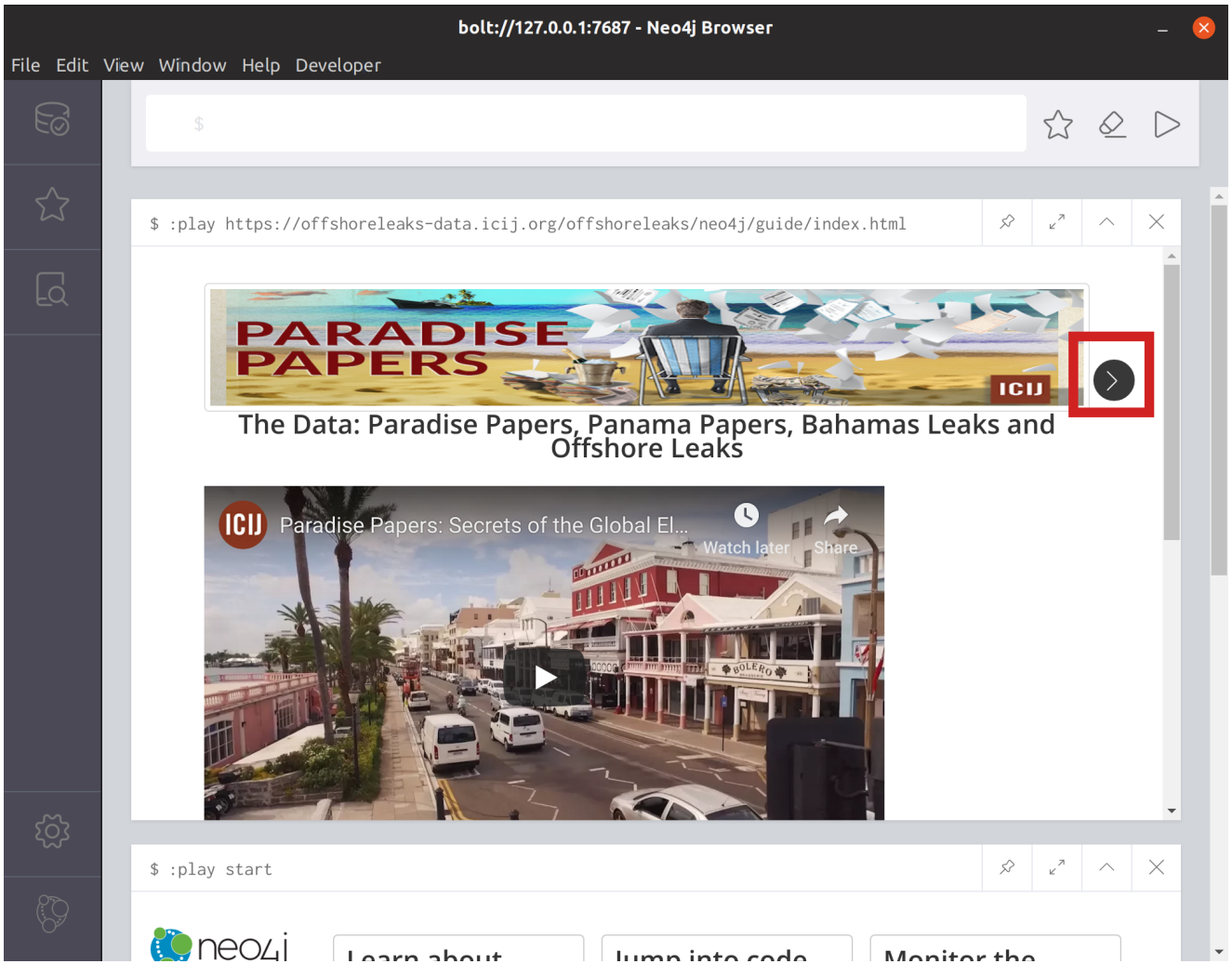
Download the right Neo4j Desktop executable from the following links:

- Windows
- Mac
- Unix

Exercise 3 – Neo4j Browser



Exercise 3 – Getting Started



Exercise 3 – Submission

- Document your queries and their return values in your report.
- Graph results are best read visually, use the Neo4j Browsers image export.

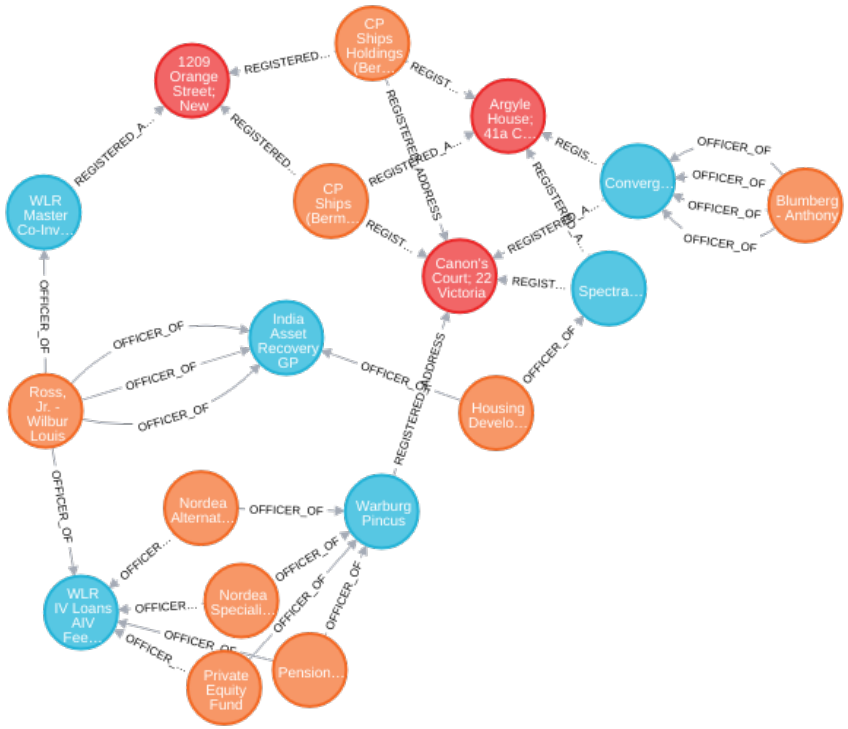


Table of Contents

Grading

- Same as last exercises.
- Grading of exercises is based on your understanding of the material.
- You get credits for serious attempts at solving the exercise.
- Every group member has to be able to demonstrate understanding of *all* of your submission.

Deadlines

- Upload your solutions by **17.06.2019 12:00 (lunchtime)**.
- Register for an exercise interview (as a group!) by **17.06.2019 23:55**.
- Interviews on June 18th and 19th.
- Be punctual for your exercise interview.

Question Sessions

There are two **voluntary** sessions for you to ask questions in person. These are intended for technical help, clarifications and to help you if you're stuck.
Not for checking your solutions.

- 04.06. 11:00-12:00: Informatiklabor Pong
- 05.06. 14:00-15:00: Informatiklabor Pong

Contact

- TUWEL forum
  We don't check any other forums / chats.
- Email: adbs@dbai.tuwien.ac.at
  (Please stop writing to individuals)