FACULTY
OF INFORMATION
TECHNOLOGY
CTU IN PRAGUE

Supervisor's statement of a final thesis

Supervisor:	Ing. Jan Trávníček, Ph.D.
Student:	Victor Petukhov
Thesis title:	Generic database metadata extractor
Branch / specialization:	Software Engineering
Created on:	6 June 2023

Evaluation criteria

1. Fulfillment of the assignment

▸ [1] **assignment fulfilled**
  [2] assignment fulfilled with minor objections
  [3] assignment fulfilled with major objections
  [4] assignment not fulfilled

The task was to analyze the JDBC interface, its possible limitations, and the complications of getting information about entities in the connected database. The analysis is very thorough, and the resulting implementation is able, to a high degree, to extract the desired metadata from the connected database.

I declare the assignment was fulfilled.

2. Main written part   88 /100 (B)

The text is written in English at an excellent level. Still, there are some factual issues, typographic issues, and general points for improvement:

The comparisons of the columns retrieved from the tested databases on individual JDBC calls described in section 2.6 is well structured into tables. However, I would also include the information about the null content of some of these columns that is common or even specific to the tested databases. Additionally, it would be nice if the information was summarized this way for all analyzed database objects, not just the more exhaustive ones.

"In order to know when the database does not have either catalogs or schemas, the output of the method getCatalog() should be considered." It is later discussed how the empty result means no catalogs or schemas in a given database system. I fail to see how one should know from practically one bit of information which of the three options it is: only catalogs, only schemas, or both.

Especially with corner cases of the implementation regarding the specifics of individually tested dialects, like routine overloading in PostgreSQL and no catalogs in Oracle, I would appreciate examples of the retrieved data to accompany the description of the issue.

Sentences with unbreakable words that cause some lines of text to be extensively spaced out should be reformulated. Similarly, lines of text that overflow to the right should too.

Minor issues:
A page break splits the Czech abstract.
Figure captions should be below the figures.
Code listing 2.1 is aligned inconsistently.
Figure 3.1 should be a vector, not a bitmap.
Section 4.2 should not be empty and on the very bottom of a page.

3. Non-written part, attachments   95 /100 (A)

The source of the prototype implementation of the JDBC extractor follows the coding standards of the Manta project and, as such, was reviewed with each change introduced to the main development branch,¨

Nevertheless, there are a few possible improvements to the code that can be made:
DictionaryWriterImpl.writeFunction   and DictionaryWriterImpl.writeFunctionWithoutParams can be unified to reduce code duplication; the same with procedures.
Similarly, the MetaDaoImpl.getFilteredCatalogsWithSchema could be simplified not to duplicate code.
MetaDao.hasCatalogs can be simplified using DeMorgan's law; at least to me, it would be clearer.

4. Evaluation of results, publication outputs and awards   95 /100 (A)

The thesis itself mentions some unfinished extra features, which will be finished in the upcoming weeks. The actual usefulness of the JDBC extractor will be determined in time, once the functionality is offered as a part of the Manta product.

5. Activity of the student

▸ [1] **excellent activity**
  [2] very good activity
  [3] average activity
  [4] weaker, but still sufficient activity
  [5] insufficient activity

The student worked on the thesis in the timeframe allocated for the creation of the Bachelor's thesis actively.

6. Self-reliance of the student

▸ [1] **excellent self-reliance**
  [2] very good self-reliance

[3] average self-reliance
[4] weaker, but still sufficient self-reliance
[5] insufficient self-reliance

The student actively proposed consultation topics and then solved the identified issues alone.

The overall evaluation $\qquad$ 92 /100 (A)

All in all, the thesis shows how JDBC can be used to extract metadata from a database system for future processing in a tool such as Manta implements. The identified limitations and their solutions are discussed. The text is exhaustive in the analytical chapter, which makes sense given the task at hand. The implementation of the prototype works the way it was designed to, and most of it is already present in the main development branch. Altogether the thesis text and prototype implementation are good enough for the overall evaluation to be as high as A (92 points),

Instructions

Fulfillment of the assignment

Assess whether the submitted FT defines the objectives sufficiently and in line with the assignment; whether the objectives are formulated correctly and fulfilled sufficiently. In the comment, specify the points of the assignment that have not been met, assess the severity, impact, and, if appropriate, also the cause of the deficiencies. If the assignment differs substantially from the standards for the FT or if the student has developed the FT beyond the assignment, describe the way it got reflected on the quality of the assignment's fulfilment and the way it affected your final evaluation.

Main written part

Evaluate whether the extent of the FT is adequate to its content and scope: are all the parts of the FT contentful and necessary? Next, consider whether the submitted FT is actually correct – are there factual errors or inaccuracies?

Evaluate the logical structure of the FT, the thematic flow between chapters and whether the text is comprehensible to the reader. Assess whether the formal notations in the FT are used correctly. Assess the typographic and language aspects of the FT, follow the Dean's Directive No. 52/2021, Art. 3.

Evaluate whether the relevant sources are properly used, quoted and cited. Verify that all quotes are properly distinguished from the results achieved in the FT, thus, that the citation ethics has not been violated and that the citations are complete and in accordance with citation practices and standards. Finally, evaluate whether the software and other copyrighted works have been used in accordance with their license terms.

Non-written part, attachments

Depending on the nature of the FT, comment on the non-written part of the thesis. For example: SW work – the overall quality of the program. Is the technology used (from the development to deployment) suitable and adequate? HW – functional sample. Evaluate the technology and tools used. Research and experimental work – repeatability of the experiment.

Evaluation of results, publication outputs and awards

Depending on the nature of the thesis, estimate whether the thesis results could be deployed in practice; alternatively, evaluate whether the results of the FT extend the already published/known results or whether they bring in completely new findings.

Activity of the student

From your experience with the course of the work on the thesis and its outcome, review the student's activity while working on the thesis, his/her punctuality when meeting the deadlines and whether he/she consulted you as he/she went along and also, whether he/she was well prepared for these consultations.

Self-reliance of the student

From your experience with the course of the work on the thesis and its outcome, assess the student's ability to develop independent creative work.

The overall evaluation

Summarize which of the aspects of the FT affected your grading process the most. The overall grade does not need to be an arithmetic mean (or other value) calculated from the evaluation in the previous criteria. Generally, a well-fulfilled assignment is assessed by grade A.