## Supervisor's statement of a final thesis

**Student:** Ondřej Hlaváč  
**Supervisor:** Ing. Jan Trávníček, Ph.D.  
**Thesis title:** SAP Hana database metadata extraction tool  
**Branch of the study:** Web and Software Engineering  
**Date:** 15. 6. 2020

<table>
<thead>
<tr>
<th>Evaluation criterion:</th>
<th>The evaluation scale: 1 to 4.</th>
</tr>
</thead>
</table>
| **1. Fulfilment of the assignment** | **1 = assignment fulfilled,**  
**2 = assignment fulfilled with minor objections,**  
**3 = assignment fulfilled with major objections,**  
**4 = assignment not fulfilled.** |

**Criteria description:**  
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.

**Comments:**  
The thesis text and related implementation fulfil the thesis topic. The topic is of moderate complexity.

<table>
<thead>
<tr>
<th>Evaluation criterion:</th>
<th>The evaluation scale: 0 to 100 points (grade A to F).</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2. Main written part</strong></td>
<td><strong>65 (D)</strong></td>
</tr>
</tbody>
</table>

**Criteria description:**  
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. 26/2017, 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.
Comments:
The thesis is written English of a good level. The text should avoid using non-technical terms like "reside" and text fragments like "To keep things simple...", "Beware that...", or colloquial words "one can come up with a pattern" -> "one can recognise a pattern", etc.

The introduction should also introduce Data lineage with a simple example so that the reader is familiar with the end goal.

The chapter Analysis of metadata would benefit from extensions in multiple ways:
- The section should start with relevant introduction of metadata description and access in Microsoft SQL Server and Oracle Database so that the reader is ready for the SAP Hana database details.
- Detailed reasoning about which metadata are needed and why should they be extracted for subsequent Dataflow analysis.
- I'm missing a section on metadata representation in the Manta system.

The chapter Metadata extraction can also be improved in many ways:
- The section should provide database queries to extract the selected metadata.
- Why it is important to filter custom types from tables? Is metadata extraction of virtual tables supposed to be handled anyhow specially?
- Do details about table storage approaches and history tables change the metadata extracted from these tables? If not, the details are not relevant here.
- View parameters are only briefly mentioned without providing details on how these parameters affect extracted metadata.
- Similarly to Sequences, other extracted entities should have their extracted metadata listed.
- Mostly it seems that this chapter only introduces the metadata to extract; this listing should rather be in the previous chapter.

The chapter Implementation mentions additional "setExtractInvalid" method in the SAPHanaExtractor interface with additional information about it provided later. This information should have been in an earlier section.

Similarly, the dependency manager should have been introduced sooner, in the chapter Design.

In overall, the sections in chapter Implementation are mostly standalone unconnected descriptions of individual functionalities of the extractor. I'm missing a bigger picture description of the extraction approach, like the order of steps of the extraction. The section "High-level idea" does that, unfortunately, very vaguely only.

In general, the text seems to be written quickly without time left for proofreading, which unfortunately degrades the overall impression from it.

An example of some unusual constructs: "The only this that is interesting here is the column SCHEMA_NAME." Page 16 line -2 "one have to" -> "one has to"

### Evaluation criterion:
The evaluation scale: 0 to 100 points (grade A to F).

#### 3. Non-written part, attachments

<table>
<thead>
<tr>
<th>Criteria description:</th>
<th>The evaluation scale: 0 to 100 points (grade A to F).</th>
</tr>
</thead>
<tbody>
<tr>
<td>The prototype implementation is good enough to be added to the Manta tool codebase and used to support the following stages of SAP Hana SQL scripts analysis.</td>
<td></td>
</tr>
</tbody>
</table>

#### 4. Evaluation of results, publication outputs and awards

<table>
<thead>
<tr>
<th>Criteria description:</th>
<th>The evaluation scale: 0 to 100 points (grade A to F).</th>
</tr>
</thead>
<tbody>
<tr>
<td>The resulting implementation will be used in the Manta project as the extractor of SAP Hana metadata which are later needed in the SAP Hana scripts parser, resolver and dataflow analyzer.</td>
<td></td>
</tr>
</tbody>
</table>

### Evaluation criterion:
The evaluation scale: 1 to 5.
### 5. Activity and self-reliance of the student

<table>
<thead>
<tr>
<th>5a:</th>
<th>1 = excellent activity, 2 = very good activity, 3 = average activity, 4 = weaker, but still sufficient activity, 5 = insufficient activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>5b:</td>
<td>1 = excellent self-reliance, 2 = very good self-reliance, 3 = average self-reliance, 4 = weaker, but still sufficient self-reliance, 5 = insufficient self-reliance</td>
</tr>
</tbody>
</table>

#### Criteria description:
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 (5a). Assess the student’s ability to develop independent creative work (5b).

#### Comments:
Throughout the student’s work on the implementation of the thesis, the student was exceptionally active and solved arising issues alone. On the other hand, a bit more effort while writing the text would have a huge positive impact.

#### Evaluation criterion:
The evaluation scale: 0 to 100 points (grade A to F).

### 6. The overall evaluation

| 85 (B) |

#### Criteria description:
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.

#### Comments:
The implementation quality is high, but unfortunately, with the lesser text quality, I recommend to evaluate the thesis with the overall grade very good (B).

Signature of the supervisor: