KU LEUVEN

FACULTY OF ENGINEERING TECHNOLOGY

DE NAYER (SINT-KATELIJNE-WAVER) CAMPUS

Inverting knowledge graphs back to raw data

How can we leverage the rules we use to construct knowledge graphs to do the inverse?

Tijs VAN KAMPEN

Promotor: Anastasia Dimou

Master thesis submitted to obtain the degree of Master of Science in the engineering technology: E-ICT Software Engineer - option ICT

academic year 2023 - 2024

©Copyright KU Leuven

This master's thesis is an examination document that has not been corrected for any errors.

Reproduction, copying, use or realisation of this publication or parts thereof is prohibited without prior written consent of both the supervisor(s) and the author(s). For requests concerning the copying and/or use and/or realisation of parts of this publication, please contact KU Leuven De Nayer Campus, Jan De Nayerlaan 5, B-2860 Sint-Katelijne-Waver, +32 15 31 69 44 or via e-mail iiw.denayer@kuleuven.be.

Prior written consent of the supervisor(s) is also required for the use of the (original) methods, products, circuits and programmes described in this Master's thesis for industrial or commercial purposes and for the submission of this publication for participation in scientific prizes or competitions.

Contents

Contents iii

1 Thesis details 1

Chapter 1

Thesis details

Title: Inversion: from knowledge graphs to raw data

Subject:

- **RQ1:** How can we leverage RML to construct raw data from heterogeneous data.
- **RQ2:** How can we extend an existing system like RML or create a new system to construct raw data from knowledge graphs.

Description: Knowledge graphs are gaining traction nowadays and more and more companies use them, such as Amazon, Bosch, IKEA, Facebook, Google, LinkedIn, SIEMENS, Zalando, etc. Most knowledge graphs are nowadays constructed from other heterogeneous data sources, such as tables in relational databases, data in XML files or in JSON format derived from a Web API. While the construction of knowledge graphs from heterogeneous data was theoroughly investigated so far, the inverse, namely constructing raw data from knowledge graphs is not explored so far.
Initially we will investigate reconstructing tables back to their initial form from an RDF graph, later extending this to other formats like JSON, XML,...

Company: internal research group EAVISE, CAMPUS DE NAYER

University promotor: Anastasia Dimou

Planning: See gantt chart in figure 1.1 for a visual overview.

Goals/Deadlines:

- **22nd October:** Gathering information
  - Defining the problem
  - Refine research questions
- **12th November:** Write related work
- **3th December:** Prototype
- **22nd December(deadline intermediate report):** Expand and improve related work
- **3th March:** Implementation
- **5th May:** Evaluation and refining implementation
- **4th June(deadline final report):** Finalize report

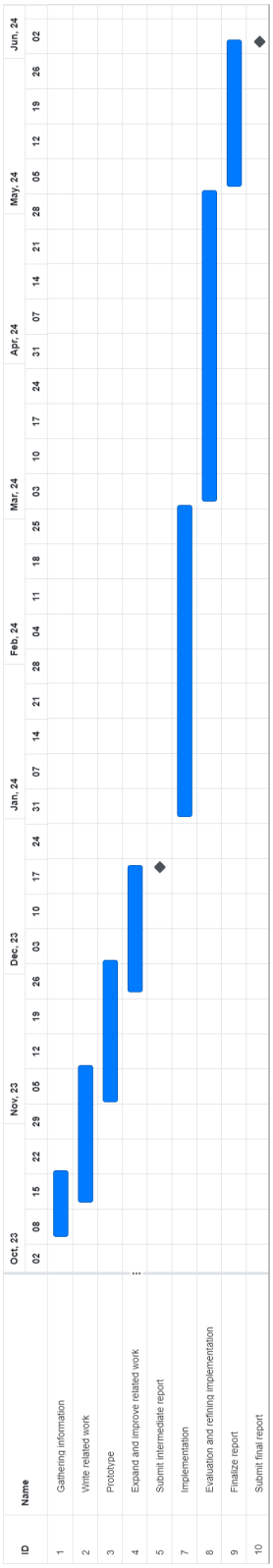

Figure 1.1: Gantt chart