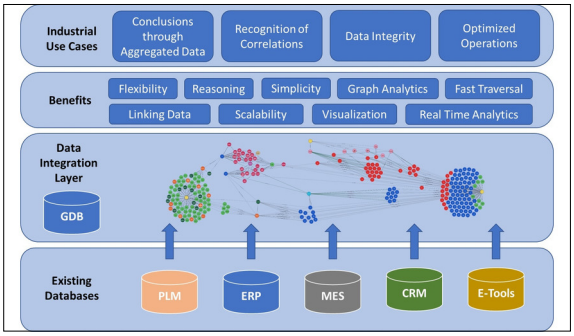
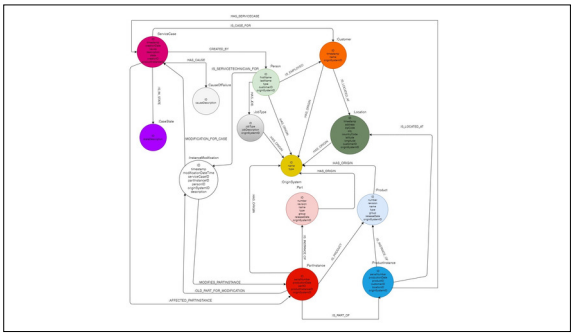
Daniel Nussbaumer

Graduate Candidate	Daniel Nussbaumer
Examiner	Prof. Dr. Daniel Patrick Politze
Co-Examiner	Ramon Schildknecht, SBB, Olten, Solothurn
Subject Area	Innovation in Products, Processes and Materials - Business Engineering and Productions

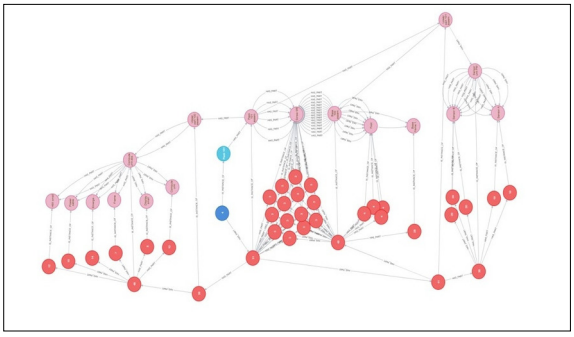
Industrial use cases for graph databases



Overview of the structure of the work, the benefits and the industrial use cases.
Own presentment



Latest version of the data model for the labelled property graph (LPG) database implemented on the neo4j software.
Own presentment



Example of the class of a product (light blue) with the instantiated product (dark blue) and all components (red).
Own presentment

Introduction: Our world is rapidly becoming more complex and connected. The amount of data piles up exponential, as more and more sensors collect raw information, anywhere and at any time. As the “internet of things” (IoT) - with its embedded sensors - finds new applications in nearly every sector of the industry, the data volumes got massively increased over the years.
Not only the amount of data, but also the number of databases is growing in companies, be due to new software tools, acquisitions or mergers. At the same time, the value of data is growing. All companies are trying to generate added value from the data and use it to optimize their offerings. All this leads to the fact that important data is increasingly scattered in individual data silos and a complete overview of the data is missing.

Definition of Task: The goal is to elaborate the basics for working with graph databases. Thereby use cases in the industrial environment shall be developed and tested. The procedure for creating the database should be documented so that it can be applied to further projects.

Conclusion: As a result of this thesis, nine benefits using a labeled property graph were elaborated. Out of these advantages then four use cases have been developed. These use cases are based on the idea of data integration via a graph database. All four use cases are theoretically realizable and can offer an effective added value through new insights to a company. However, the company must have a certain amount of data and databases to make the implementation in a graph database reasonable. Each database must be attached to the graph database in a time-consuming process.
In addition, an integrated graph database was created and tested using the example of the HSR Product Cockpit. The workflows, lessons learned and best practices could be used for further projects. For the development of a new data graph, the best practices of the manufacturer and further guidelines should be followed. The 101-Guideline for the development of ontologies from Stanford University is a very place to start.
Graph databases are gaining more and more ground in the world of computer science. Even on the latest Gartner hype-charts the technology is listed. Especially the labeled property graphs will become more and more widespread in the next years. A query language for graph databases is currently being certified as an ISO standard equivalent to SQL. As soon as the standards are set and a technology has become established, the databases will gain a foothold in the industry. Graphs according to the resource description framework (RDF), on the other hand, will continue to grow in the area of the World Wide Web, where huge amounts of data are stored according to these standards and are accessible to everyone.