Enhancing Knowledge Graph Generation with Ontology Reshaping – Bosch Case

Dongzhuoran Zhou[1,2(✉)], Baifan Zhou[2], Zhuoxun Zheng[1,3], Egor V. Kostylev[2], Gong Cheng[4], Ernesto Jiménez-Ruiz[2,5], Ahmet Soylu[3], and Evgeny Kharlamov[1,2]

[1] Bosch Center for Artificial Intelligence, Renningen, Germany
dongzhuoran.zhou@outlook.com
[2] SIRIUS Centre, University of Oslo, Oslo, Norway
[3] Department of Computer Science, Oslo Metropolitan University, Oslo, Norway
[4] State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China
[5] City, University of London, London, UK

Motivation. In the context of Industry 4.0 [1] and Internet of Things (IoT) [2], modern manufacturing and production [3,4] lines are equipped with software systems and sensors that constantly collect and send a high volume of data. Knowledge Graphs (KGs) allow to represent these data in a semantically structured way and provide a convenient foundation for standardised AI and analytics solutions [5–13]. Thus, KGs opened new horizons [14] and were adapted for a wide range of industrial applications in Bosch [15–17], Siemens [18,19], Equinor [6] and other companies. An important industrial scenario for Bosch where we rely on KGs is *monitoring* of manufacturing processes, including e.g. analysing the quality of the manufactured products with heterogeneous data from various formats. In particular we rely on KGs for *welding quality monitoring* [5,9,15,20,21], where welding is performed with automated machines that connect pieces of metal together by pressing them and passing high current electricity through them [20]. The process is remarkably data intensive and requires efficient data infrastructure like KG databases [21–23].

Challenge. There is a number of approaches to enable (semi-) automatic construction of KGs over industrial data that is typically of high complexity and variety. These approaches typically rely on mappings from the raw data to a given KG schema, namely a domain ontology, and that can be used to construct (in the ETL fashion) the entities and properties of the KGs according to the ontology. However, the existing approaches to generate KGs are not always efficient enough and the resulting KGs are not sufficiently application and user-friendly. This challenge arises from a trade-off between the following two principles:

- It is in general considered a good practice to create domain ontologies in a *knowledge-oriented* way, namely to reflect the general domain knowledge rather than data particularities of arbitrary datasets [24,25].
- On the other hand, to properly reflect the raw data, a *data-oriented* KG schema is required. Raw data often come with diverse specificities, which

P. Groth et al. (Eds.): ESWC 2022 Satellite Events, LNCS 13384, pp. 299–302, 2022.
https://doi.org/10.1007/978-3-031-11609-4_45

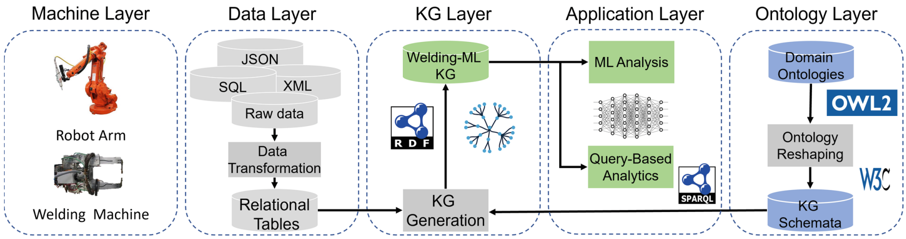

Fig. 1. An overview of our approach of ontology reshaping enhanced KG generation.

may differ significantly from the domain ontologies, e.g., not necessarily all classes in the latter can be mapped to attributes in the former. If a knowledge-oriented domain ontology is directly used as the KG schema, this can cause a series of issues, e.g., blank nodes created due to classes unmapped to the raw data.

Ontology Reshaping Enhanced KG Generation. To address this challenge, we propose our approach of ontology reshaping enhanced KG generation (Fig. 1). The core idea of our approach is to "reshape" a domain ontology to its (often) more compact data-oriented versions, and then to use the latter ones as the KG schemata to construct KGs from relational tables. Our system consists of five layers and several semantic modules.

Our approach consists of Machine Layer, Data Layer, KG Layer, Application Layer and Ontology Layer. In the Machine Layer, manufacturing data are constantly collected from running production lines of automated welding. In the Data Layer, the raw data of various formats e.g., json, SQL, XML are transformed into relational tables by the *Data Transformation* module. In the KG Layer, our Welding-ML KG is generated by populating KG schemata with the relational tables. We call our KG Welding-ML KGs because they contain information of the welding production and ML analysis [26]. The Welding-ML KG are used by a series of applications, e.g., ML Analysis, in the Application Layer. The KG schemata are data-oriented ontologies transformed by the Ontology Reshaping module from knowledge-oriented domain ontologies, in the Ontology Layer.

In particular, our Ontology Reshaping module [27] converts a given domain ontology to (often) compacter ontologies that serves as the KG schema [28]. The intuition behind it is to project the domain ontology to a labelled multi-graph, then select subsets of nodes and edges from it, and then connect the sub-graphs via some optimality criteria based on user heuristics, efficiency, simplicity, etc.

Evaluation at Bosch. We implement and evaluate our approach at Bosch for welding quality monitoring and data analysis [29,30]. The evaluation was done in the offline mode on several car manufacturing lines. These lines generated a great number of heterogeneous data such as historic data with sensor measurements, welding machine configurations, manufacturing specifications, and the quality estimates of finished welding operations. In this work, we select a section of the data for discussion. The selected section contains 4.315 million records. These data account for 1000 welding operations, estimated to be related to 100 cars. Thanks to ontology reshaping, our approach significantly increases time efficiency, space efficiency, and results in KGs with much better simplicity, compared to a baseline approach [27] without ontology reshaping: The KG generation becomes 7 to 8 times faster; the number of entities reduced to merely 1/2 to 1/6 of the baseline, and storage space reduced to 2/3; and the number of blank nodes are reduced to zero and queries over the KGs become shorter and simpler.

Outlook. Our approach is currently deployed in our Bosch evaluation environment, and we are considering to push it further into a more advanced and strict evaluation phase of production that runs in real-time. To show the benefits, we also plan to demonstrate our KG solution with more users and more use cases. In the Application Layer, we plan to develop Query-Based Industrial Analytics. In the future, we plan to develop formal theory of ontology reshaping, to enhance the KG generation modules to improve the compatibility of the KG schema to the domain ontologies; to extend the KG solution for more applications, e.g. question answering, visualisation, statistic analysis.

Acknowledgements. The work was partially supported by the H2020 projects Dome 4.0 (Grant Agreement No. 953163), OntoCommons (Grant Agreement No. 958371), and DataCloud (Grant Agreement No. 101016835) and the SIRIUS Centre, Norwegian Research Council project number 237898.

References

1. Kagermann, H.: Change through digitization—value creation in the age of Industry 4.0. In: Albach, H., Meffert, H., Pinkwart, A., Reichwald, R. (eds.) Management of Permanent Change, pp. 23–45. Springer, Wiesbaden (2015). https://doi.org/10.1007/978-3-658-05014-6_2
2. ITU, Recommendation ITU - T Y.2060: overview of the internet of things, Technical report, International Telecommunication Union
3. Chand, S., Davis, J.: What is smart manufacturing. Time Mag. Wrapper **7**, 28–33 (2010)
4. Wuest, T., Weimer, D., Irgens, C., Thoben, K.-D.: Machine learning in manufacturing: advantages, challenges, and applications. Prod. Manuf. Res. **4**, 23–45 (2016)
5. Zhou, B., Pychynski, T., Reischl, M., Kharlamov, E., Mikut, R.: Machine learning with domain knowledge for predictive quality monitoring in resistance spot welding. J. Intell. Manuf. **33**(4), 1139–1163 (2022)
6. Kharlamov, E., et al.: Ontology based data access in Statoil. J. Web Semant. **44**, 3–36 (2017)
7. Zhou, B.: Machine learning methods for product quality monitoring in electric resistance welding, Ph.D. thesis, Karlsruhe Institute of Technology, Germany (2021)
8. Zou, X.: A survey on application of knowledge graph. In: Journal of Physics: Conference Series, vol. 1487, p. 012016. IOP Publishing (2020)
9. Zhou, B., Svetashova, Y., Pychynski, T., Baimuratov, I., Soylu, A., Kharlamov, E.: SemFE: facilitating ML pipeline development with semantics. In: CIKM, pp. 3489–3492. ACM (2020)
10. Zhou, B., Pychynski, T., Reischl, M., Mikut, R.: Comparison of machine learning approaches for time-series-based quality monitoring of resistance spot welding (RSW). Arch. Data Sci. Ser. A (Online First) **5**(1), 13 (2018)

11. Soylu, A., et al.: TheyBuyForYou platform and knowledge graph: expanding horizons in public procurement with open linked data. Semant. Web **13**(2), 265–291 (2022)
12. Zhou, B., Chioua, M., Schlake, J.-C.: Practical methods for detecting and removing transient changes in univariate oscillatory time series. IFAC-PapersOnLine **50**(1), 7987–7992 (2017)
13. Zhou, B., Chioua, M., Bauer, M., Schlake, J.-C., Thornhill, N.F.: Improving root cause analysis by detecting and removing transient changes in oscillatory time series with application to a 1,3-butadiene process. Ind. Eng. Chem. Res. **58**, 11234–11250 (2019)
14. Horrocks, I., Giese, M., Kharlamov, E., Waaler, A.: Using semantic technology to tame the data variety challenge. IEEE Internet Comput. **20**(6), 62–66 (2016)
15. Zhou, B., Svetashova, Y., Byeon, S., Pychynski, T., Mikut, R., Kharlamov, E.: Predicting quality of automated welding with machine learning and semantics: a Bosch case study. In: CIKM (2020)
16. Zhou, B., et al.: Method for resistance welding, US Patent App. 17/199,904 (2021)
17. Kalaycı, E.G., et al.: Semantic integration of Bosch manufacturing data using virtual knowledge graphs. In: Pan, J.Z., et al. (eds.) ISWC 2020. LNCS, vol. 12507, pp. 464–481. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-62466-8_29
18. Kharlamov, E., et al.: Semantic access to streaming and static data at Siemens. J. Web Semant. **44**, 54–74 (2017)
19. Hubauer, T., Lamparter, S., Haase, P., Herzig, D.M.: Use cases of the industrial knowledge graph at siemens. In: ISWC (P&D/Industry/BlueSky) (2018)
20. Zhou, B., et al.: SemML: facilitating development of ML models for condition monitoring with semantics. J. Web Semant. **71**, 100664 (2021)
21. Svetashova, Y., et al.: Ontology-enhanced machine learning: a Bosch use case of welding quality monitoring. In: Pan, J.Z., et al. (eds.) ISWC 2020. LNCS, vol. 12507, pp. 531–550. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-62466-8_33
22. Svetashova, Y., Zhou, B., Schmid, S., Pychynski, T., Kharlamov, E.: SemML: reusable ML for condition monitoring in discrete manufacturing. In: ISWC (Demos/Industry), vol. 2721, pp. 213–218 (2020)
23. Zhou, B., Svetashova, Y., Pychynski, T., Kharlamov, E.: Semantic ML for manufacturing monitoring at Bosch. In: ISWC (Demos/Ind), vol. 2721, p. 398 (2020)
24. Smith, B.: Ontology. In: The Furniture of the World, pp. 47–68. Brill (2012)
25. Guarino, N., Oberle, D., Staab, S.: What is an ontology? In: Staab, S., Studer, R. (eds.) Handbook on Ontologies. IHIS, pp. 1–17. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-540-92673-3_0
26. Zhou, B., Zhou, D., Chen, J., Svetashova, Y., Cheng, G., Kharlamov, E.: Scaling usability of ML analytics with knowledge graphs: exemplified with a Bosch welding case. In: IJCKG (2021)
27. Zhou, D., Zhou, B., Chen, J., Cheng, G., Kostylev, E.V., Kharlamov, E.: Towards ontology reshaping for KG generation with user-in-the-loop: applied to Bosch welding. In: IJCKG (2021)
28. Zheng, Z., et al.: Query-based industrial analytics over knowledge graphs with ontology reshaping. In: ESWC (Posters & Demos) (2022)
29. Zhou, B., et al.: The data value quest: a holistic semantic approach at Bosch. In: ESWC (Demos/Industry) (2022)
30. Yahya, M., et al.: Towards generalized welding ontology in line with ISO and knowledge graph construction. In: ESWC (Posters & Demos) (2022)