

## PhD Topic “Data and Knowledge Structuring in the Enterprise”

While enterprises and holdings acquire and manage more and more data, the data assets often become less manageable through as data grows faster than its accumulation can be documented and structured. Many companies, like ÖBB, therefore currently recognize Data Cataloguing and Data Governance Processes as a key requirement, in order to establish a function data ecosystem across departments and divisions.

### *Status quo:*

ÖBB is already looking proactively into structuring and cataloging their data landscape and making it accessible via a common “Data Lake”, current projects involve (real-time) data ingestion and tying building sustainable Data Science pipelines, enable Digital Twinning, estimated to be at TRL2 of 5 at the moment. Challenges to be addressed include automated Data Quality checks, (automating) ticketing, or integrated Data Management, but also (organization) cultural obstacles to data governance, as well as (personal) data protection. Analytics solutions could enable condition-based, dynamic maintenance where both internal as well as external data sources could be used (e.g. weather, climate data, simulations).

### *Research potential:*

We see a lot of potential in improving Data Catalogs and Data Governance frameworks by organizing the organizational knowledge in Knowledge Graphs [1]. Our experiences in the institute for Data, Process and Knowledge Management (**Prof. Axel Polleres**) in the area of Open Data Catalogues and quality monitoring frameworks, as well as using Semantic Technologies for improving Data Quality, building Semantic Search Engines could be a potential enabler for the current focus challenges outlined above [2,3,4]. Secondly, cataloguing not only data sources, but also provenance/lineage data as well as usage policies attached to the data, as well as investigating how such policies can be automatically checked/enforced, based on concrete use cases, would be a potential research topic, where we could draw from related research in the area of personal data management and protection [5,6]. As for organizational culture aspects, the team of **Prof. Alexander Kaiser**, also in the institute for Data, Process and Knowledge Management potentially could support the above-mentioned topics [7]

Joint activities and showcases could potentially be exploited and disseminated within the Data Intelligence Offensive,<sup>1</sup> where both ÖBB and WU are already members, in order to potentially steer data markets and platforms where the exchange of data with external players becomes relevant.

### *Selected relevant publications and projects:*

1. Aidan Hogan, Eva Blomqvist, Michael Cochez, Claudia d'Amato, Gerard de Melo, Claudio Gutierrez, José Emilio Labra Gayo, Sabrina Kirrane, Sebastian Neumaier, Axel Polleres, Roberto Navigli, Axel-Cyrille Ngonga Ngomo, Sabbir M. Rashid, Anisa Rula, Lukas Schmelzeisen, Juan Sequeda, Steffen Staab, and Antoine Zimmermann.

---

<sup>1</sup> <https://dataintelligence.at/>

Knowledge graphs. *ACM Computing Surveys (CSUR)*, 2021. Extended pre-print available at <https://arxiv.org/abs/2003.02320>.

2. Sebastian Neumaier, Axel Polleres, Simon Steyskal, and Jürgen Umbrich. Data integration for open data on the web. In Giovambattista Ianni, Domenico Lembo, Leopoldo E. Bertossi, Wolfgang Faber, Birte Glimm, Georg Gottlob, and Steffen Staab, editors, *Reasoning Web. Semantic Interoperability on the Web (Reasoning Web 2017)*, volume 10370 of *Lecture Notes in Computer Science (LNCS)*, pages 1--28. Springer, London, United Kingdom, July 2017.
3. Sebastian Neumaier, Jürgen Umbrich, and Axel Polleres. Automated quality assessment of metadata across open data portals. *ACM Journal of Data and Information Quality (JDIQ)*, 8(1):2, November 2016.
4. Sebastian Neumaier and Axel Polleres. Enabling spatio-temporal search in open data. *Journal of Web Semantics (JWS)*, 55:21--36, March 2019.
5. Simon Steyskal and Axel Polleres. Towards formal semantics for ODRL policies. In 9th International Web Rule Symposium (RuleML2015), number 9202 in *Lecture Notes in Computer Science (LNCS)*, pages 360--375, Berlin, Germany, August 2015. Springer.
6. Giray Havur, Simon Steyskal, Oleksandra Panasiuk, Anna Fensel, Víctor Mireles, Tassilo Pellegrini, Thomas Thurner, Axel Polleres, and Sabrina Kirrane. Automatic license compatibility checking. In Mehwish Alam, Ricardo Usbeck, Tassilo Pellegrini, Harald Sack, and York Sure-Vetter, editors, *Proceedings of the Posters and Demo Track of the 15th International Conference on Semantic Systems (SEMANTiCS 2019)*, volume 2451 of *CEUR Workshop Proceedings*, Karlsruhe, Germany, September 2019. CEUR-WS.org
7. Rositsa V Ivanova, Clemens Kerschbaum, Sebastian Neumaier, Alexander Kaiser, Axel Polleres. How much Knowledge is in Knowledge Graphs? - A Knowledge Management Perspective. Under submission for the *Semantic Web Journal*, cf. <http://www.semantic-web-journal.net/content/how-much-knowledge-knowledge-graphs-knowledge-management-perspective>

#### *Potential Research Questions for Student Projects:*

- Data Governance and Data Quality – Automated Data Understanding/Machine Learning: piloting which data producing and collecting processes can be automated and how.
- Mapping Data Sources, Roles and Functions and exploring routes for efficient data exchange between different divisions
- Definitions of Requirements (e.g. real-time requirements for sources to be included in the Data Lake), interfaces and data access policies
- Support the identification and verification of data availability and quality demands by defining concrete use cases and data products
- How can we prepare and publish data internally to make integration possible “with a few clicks”?
- How to (support) build(ing) an effective Enterprise Knowledge Graph?

## PhD Topic “Digital products and interfaces in business ecosystems”

Many business processes including interactions with clients and partners are carried out predominantly online, or at least leave a rather comprehensive digital trail. This poses new challenges for data management and storage – and new opportunities for innovative digital products and services.

### *Status quo:*

ÖBB is looking proactively into structuring and cataloging their data landscape and making it accessible via a common “Data Lake”. External data sources as well as client and partner business platforms already are or will be connected with the ÖBB digital ecosystem. ÖBB owns or maintains a variety of data (e.g. operations, maintenance) that could be bundled and provided to clients, partners or third parties as digital data products. Aside from challenges arising from the data management side, there are also questions regarding the interaction design of the ÖBB digital ecosystem for both customers and employees. ÖBB aims to increase workplace digitalization and explores digital options to deal with the attendant cultural and organizational obstacles (e.g., chatbots, frontend optimization). Designing the digital experience for employees, but also for customers, also requires taking into account how expectations and mindsets are changing with the rapid spread of the so-dubbed low-touch economy.

### *Research potential:*

Designing data products and services requires answering questions from a number of different perspectives, e.g. product definition, pricing, bundling, marketing, or product configuration. These questions are closely related to interface design decisions. Customer interfaces can be used to generate more insight into customer behavior, in turn feeding into improvements of the data products regarding pricing, configuration etc. Employee interfaces, especially in complex environments, need to provide guidance, structure and easy access to expert help, whether provided by (semi-)automated or human agents. At the institute of Digital Ecosystems (**Prof. Verena Dorner**), we can draw on our experience in algorithm and system design [8-11], (strategic) interaction design on digital platforms [8, 12] and pricing [13, 14], and user behavior [15, 16] to generate insights into these questions, and view them in a holistic fashion. Following a use case-based approach, we could define and investigate a range of data products with ÖBB and aggregate the insights thus created in a more generic process model and design templates. For the optimization of the digital experience for customers and employees, we could design and implement an experimental approach towards well-structured and effective improvement processes.

### *Selected relevant publications and projects:*

8. Lux E., Adam M.T.P., Dorner V., Helming S., Knierim M., Weinhardt C. (2018): Live biofeedback as a user interface design element: A review of the literature. Communications of the Association for Information Systems, 43
9. Scholz, M., Dorner, V., Schryen, G., Benlian, A. (2017): A configuration-based recommender system for supporting e-commerce decisions. European Journal of Operational Research 259, 205-215

10. Jung, D., Dorner, V., Weinhardt, C., Puzmaz, H. (2017): Designing a robo-advisor for risk-averse, low-budget consumers. *Electronic Markets*, 1-14
11. Dorner, V., Morana, S., Mädche, A., Weinhardt, C. (2016): Towards designing individual value-oriented decision support for selecting internet-based services. *Workshop on value-sensitive design of Internet-based services*, 46. Jahrestagung der Gesellschaft für Informatik 2016, Klagenfurt (A)
12. Dorner, V., Giamattei, M., Greiff, M. (2020). The Market for Reviews: Strategic Behavior of Online Product Reviewers with Monetary Incentives. *zfbf Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung - Schmalenbach Business Review (sbr)*, 72 (3), 397-435
13. Scholz, M., Dorner, V., Franz, M., Hinz, O. (2015): Measuring Consumers' Willingness-to-Pay with Utility-based Recommendation Systems. *Decision Support Systems* 72: 62-71
14. Dorner, V., Scholz, M. (2013): Predicting and Economically Exploiting Utility Thresholds with Utility-based Recommendation Systems. In: *Proceedings of the 21st European Conference on Information Systems*, Utrecht (NL)
15. Schuhbeck, V., Siegfried, N., Dorner, V., Benlian, A., Scholz, M., Schryen, G. (2019): Walking the Middle Path: How Medium Trade-off Exposure leads to higher Consumer Satisfaction in Recommender Agents. In: *Proceedings of 14th International Conference on Wirtschaftsinformatik*, Siegen (D)
16. Jung, D., Erdfelder, E., Bröder, A., Dorner, V. (2019): Differentiating Motivational and Cognitive Explanations for Decision Inertia. *Journal of Economic Psychology*, 72, 30-44

*Potential Research Questions for Student Projects:*

- Pricing approaches for data products in digital ecosystems
- Configuration of data products and utility-based pricing
- Roles, procedures and access requirements for data product creation
- Integration of external data sources in data products – feasibility, value and cost analyses for specific use cases
- Test data product design process and templates in use cases
- Develop data product design interface for employees
- Improve on interface design for (semi-)automated support processes (e.g. chatbots)
- Designing seamless “multichannel” experiences
- Experimental investigation of changes to ticket shop frontend