AI for Industry at the Computer Science Faculty of the Free University of Bozen-Bolzano

Diego Calvanese, Paolo Felli, Fabrizio Maria Maggi, Marco Montali
Faculty of Computer Science, Free University of Bozen-Bolzano, Italy
{calvanese,felli,maggi,montali}@inf.unibz.it

Abstract

We provide an overview of some of the main research lines and project activities focused on AI for industry and developed within the Faculty of Computer Science of the Free University of Bozen-Bolzano (unibz). Each section refers to a research line connected to a funded project. In particular, we highlight project details, participants, and a description of the executed and planned research activities connected to the project.

1 AI for Data-Aware Processes and Manufacturing

Project

- Title: *Automated process planning in cyber physical production systems of smart factories (SMART-APP)*;
- Funding scheme: UNIBZ Interdisciplinary Projects 2020;
- Period: 01/10/2020 - 30/09/2022;
- Budget: 130 K€ for unibz.

Participants

- From unibz - Faculty of Computer Science: Diego Calvanese (Faculty PI), Paolo Felli, Marco Montali, Sarah Winkler;
- From unibz - Faculty of Science and Technology: Erwin Rauch (project coordinator), Raphael A. Rojas, Benedikt G. Mark, Renato Vidoni.

Description. A central research line within the KRDB Research Centre for Knowledge and Data at unibz focuses on the formal specification, verification, reasoning, and mining of integrated models for processes and data. This is a long-standing problem at the intersection of AI, BPM, and database theory [Baral and De Giacomo, 2015; Dumas, 2011; Calvanese *et al.*, 2013].

One relevant application domain for these models and techniques is *manufacturing as a service* in Industry 4.0. In manufacturing as a service, products to be manufactured are not known in advance, batch sizes are small, and a facility may produce items belonging to heterogeneous product families for different customers. A key task in this setting is that of process planning, i.e., assessing whether a given product specification (including both data and control-flow aspects) can be realized in a given factory. The process plan controller, i.e., the control software that delegates the operations in the plan to the appropriate manufacturing resources, must then be synthesized.

Within SMART-APP we are approaching this task by connecting it to previous foundational results achieved within our group [Calvanese *et al.*, 2018; Ghilardi *et al.*, 2020]. In particular, we are studying verification and synthesis problems considering different models for data and processes [De Giacomo *et al.*, 2022; Felli *et al.*, 2022]. From the practical standpoint, the project aims at providing, testing and evaluating a proof-of-concept prototype for the synthesis of manufacturing controllers, ranging from the modelling of the equipment and the product specification to the generation of executable machine code.

2 AI for Process Mining

Project

- Title: *exPlaInable kNowledge-aware PrOcess INTelligence (PINPOINT)*;
- Funding scheme: PRIN 2020;
- Duration: starting in the first quarter of 2022;
- Budget: 976 K€ (194 K€ for unibz).

Participants

- From unibz: Marco Montali (project coordinator), Paolo Felli;
- Other members: Claudio Di Ciccio (Sapienza University of Rome), Rafael Peñaloza Nyssen (University of Milano-Bicocca), Luigi Pontieri (CNR), Francesco Ricca (University of Calabria);
- Additional supporters: Chiara Ghidini (Fondazione Bruno Kessler), three companies (providing use cases in logistics and customer care).

Description. Process mining is a collection of process intelligence techniques combining model-based and data-oriented analysis to obtain insights about the execution of business processes in reality [van der Aalst, 2016]. Several process mining tasks exist to discover processes from event data, check conformance of actual with expected behaviours, provide runtime operational support, and enhance models with insights from data. However, three crucial limitations hamper the process mining pipelines. First, data extraction and preparation often consist of handcrafted operations without documentation nor traceability, thereby making the provenance of process mining results obscure. Second, process mining pipelines do not include background, domain knowledge, and cannot properly contextualise the produced results in a specific organizational setting. This, process mining techniques

incorporate black-box components and handcrafted methods that are opaque and unable to explain why some results have been produced, and how they can be connected with process models.

The PINPOINT project aims at attacking these limitations by developing a full-fledged set of techniques towards explainable, knowledge-aware process intelligence. It does so by relying on recent advancements in AI and declarative languages and techniques, which at unibz have been investigated for more than a decade, foundationally [Montali *et al.*, 2010] and in terms of tool support [Alman *et al.*, 2021]. In particular, within PINPOINT we intend to exploit declarative, multi-perspective formalisms combining time [Maggi and Westergaard, 2014], data [Artale *et al.*, 2019] and uncertainty [Maggi *et al.*, 2020a; Maggi *et al.*, 2020b] to express process knowledge, and targets end-to-end interpretable and explainable process mining pipelines through integrated learning and inference algorithms.

3 Virtual Knowledge Graphs for Data Access and Integration

Project

- Title: *Scalable End-user Access to Big Data (Optique)*;
- Funding scheme: EU FP7;
- Duration: 11/2012 - 10/2016 (48 months);
- Budget: 873 K€ for unibz.

Involved people

- From unibz: Diego Calvanese (local PI), Elena Botoeva, Bejamin Cogrel, Julien Corman, Elem Güzel Kalayci, Davide Lanti, Guohui Xiao.

Description. Most medium-sized and large organizations face the problem of having to deal with large and complex collections of data that need to be managed and queried in an integrated way. The Virtual Knowledge Graph (VKG) approach [Xiao *et al.*, 2019], also known as Ontology-based Data Access and Integration (OBDA/OBDI) [Xiao *et al.*, 2018a], provides a solution to the data integration challenge by relying on semantic technologies and the associated standards, specifically (i) RDF as a language to represent knowledge graphs, (ii) a lightweight ontology languages (OWL 2 QL) to encode domain knowledge, (iii) a language to define mappings between the ontology and the data sources (R2RML), and (iv) a language (SPARQL) to query through the ontology the underlying data sources exposed a virtual knowledge graph.

The Optique project [Kharlamov *et al.*, 2015] has developed an extensible platform providing a complete and generic solution to data access and integration challenges by relying on the VKG approach. The development of the platform has been evaluated within two large European companies in two industrial use cases: (i) At Equinor (formerly Statoil), a Norwegian multinational oil and gas company, VKGs have been used by geologists in Data Exploration, to integrate and uniformly query multiple complex and large data sources. (ii) Siemens Energy runs several service centers that remotely monitor several thousand appliances, such as gas and steam turbines, and the aim of Optique has been to access and integrate sensor data from these devices to perform reactive and predictive diagnostics.

Both use cases rely (among other technologies for query formulation and distributed query evaluation developed in Optique) on the VKG system Ontop [Calvanese *et al.*, 2017; Xiao *et al.*,], conceived and developed at unibz. Ontop is a state-of-the-art query reformulation engine that allows one to efficiently process queries posed over a (domain) ontology, by rewriting such a query with respect to the ontology axioms, and unfolding the rewritten query with respect to the data source mappings. A key challenge in Ontop has been to make such query reformulation process highly efficient, generating small queries that can also be executed efficiently by the underlying data source. For this purpose, Ontop relies on advanced query optimization techniques, which make extensive usage both of domain knowledge encoded in the ontology, and of constraints on the data sources to which the ontology is mapped [Xiao *et al.*, 2018b].

We observe that the VKG approach, and specifically Ontop, has also been used in other industrial settings. We mention here the project *OBDA Extensions for Bosch*, carried out in cooperation between Robert Bosch GmbH in Stuttgart (Germany) and unibz from 01/05/2019 to 30/11/2019, with a budget of 40,000 EUR. The aim of the project has been to study how to model within VKGs time-based manufacturing data so as to handle both concrete time (timestamps) as well as implicit time (based on sequences). The approach has been evaluated at Bosch on the Surface-mount Technology (SMT) process [Kalayci *et al.*, 2020].

Finally, notice that while Optique has mainly been concerned with relational data sources. However, in any data integration scenario, it is crucial to be able to deal with the heterogeneity coming from multiple types of data sources, such as csv tables, excel files, json documents, (RDF) graphs, or annotated text, and moreover to deal with different forms of data, such as geospatial data and notably temporal data. We are actively investigating these aspects in our current research.

4 AI for Semantic Interoperability

Project

- Title: *Semantic modelling and semantic interoperability / Smart Data Integration 4.0*;
- Funding scheme: industry funded applied research project;
- Period: 2018 - 2021;
- Budget: 180 K€ for unibz.

Participants

- From unibz: Enrico Franconi (project coordinator), Nony Ndefo, Francesco Sportelli, Nicola Pedot.

Description. An activity which has been always at the core of the KRDB Research Centre for Knowledge and Data at unibz has been semantic interoperability and data integration. In this context, several applied research projects have been funded by industries in recent years. Here we mention the research activity funded by the European Space Agency (ESA) and the Datatellers startup company [Franconi, 2019].

The development and management of complex systems, such as those used by the European Space Agency, imply activities involving many parties among all the actors involved, both at global level (i.e. company/customer) and at local level (i.e. product/supplier), typically under-specified, in conflict with each other, distributed in space and time. Efficient and

effective interoperability is required throughout the system development (e.g., satellites) and operations (e.g., missions) lifecycle.

The European Cooperation for Space Standardisation committee addressed these difficulties mandating to address interoperability at the semantic level: it imposes a methodology and tools associated with the conceptual model at global and local levels. A workshop in 2019 aimed to coordinate the efforts to create an Overall System Modelling for System Engineering (OSMoSE) in the context of a Space System Ontology. The projects' activities focused on the support in modelling the "Space System Reference Database" of ESA, to capture, verify and validate all data required to operate the spacecraft in-flight, and on the design and development of a conceptual modelling tool based on Microsoft Visual Studio to support semantic integration in space systems engineering.

5 Predictive Maintenance

Project

- Title: *Predictive Maintenance for Industrial Equipment (PREMISE)*;
- Funding scheme: EFRE/FESR 2014-2020;
- Duration: 01/02/2021 - 30/09/2022;
- Budget: 466 K€ (130 K€ for unibz);
- Web: https://dbs.inf.unibz.it/projects/premise/

Participants

- From unibz: Johann Gamper (PI), Anton Dignös (Co-PI), Matteo Ceccarello;
- Other members: Michael Deflorian (company Durst), Markus Pfeifer (company TechnoAlpin).

Description. Today's availability of huge amounts of sensor data allows predictive maintenance, which is scheduled as-needed based on real-time conditions of industrial assets. Predictive maintenance tracks the performance of equipments during normal operation and detects possible defects before a failure occurs. Such a maintenance strategy brings significant cost savings due to a reduction of downtime, as maintenance steps can be better scheduled, and a reduction of resources, as parts are only changed if their performance degrades.

The PREMISE project aims at developing and implementing a framework for predictive maintenance for industrial devices and is composed of three main components: data preprocessing, data analysis, and prediction. The project is a collaboration between UNIBZ and the companies Durst and TechnoAlpin.

6 Wearable Computing

Project

- Title: *COMMunity-OrieNted WEARrable Computing Systems (COMMON-WEARS)*
- Funding scheme: PRIN 2020.
- Duration: starting in the first quarter of 2022.
- Budget: 944 K€ (186 K€ for unibz).

Participants

- From unibz: Antonio Liotta (local PI), Michele Segata, Giuseppe Di Fatta, Lucia Cavallaro;
- Other members: Giancarlo Fortino (project coordinator, University of Calabria), Mirko Viroli (University of Bologna), Emiliano Schena (Campus Bio-Medico University of Rome), Ferruccio Damiani (University of Turin).

Description. The project is about smart wearables and body sensor networks (BSN), with applications, e.g., in domestic, urban, manufacturing, emergency, and working environments. This will be achieved by combining recent research conducted by the proposers in aggregate computing, collective opportunistic IoT, machine learning, and WCS/BSN architectures. The project will demonstrate results on challenging use cases in the areas of healthcare, emergency response and pandemic management that are strongly impacted by smart wearables.

7 Hyperspectral Imaging

Project

- Title: *Hyperspectral Images for Inspection Applications (H2I)*
- Funding scheme: EFRE/FESR 2014-2020.
- Duration: 17/01/2019 - 16/03/2022.
- Budget: 503 K€ (202 K€ for unibz).
- Web: https://h2i.inf.unibz.it/.

Participants

- From unibz: Roberto Confalonieri (project coordinator), Buriro Attaullah, Markus Zanker, Youry Pii (Faculty of Science and Technology);
- Other members: Tammam Tillo (technical coordinator, IIIT-Delhi), Matteo Caffini, Simone Faccini, Marco Boschetti, Philipp Bock (Microtec), Salim Malek (Fondazione Bruno Kessler).

Description. Hyperspectral images allow one to inspect the composition of objects in a scene in a non-destructive way. Hyperspectral images are used in remote sensing applications, astronomy, and recently also in the wood and food processing industry, and in agriculture. The aim of the H2I project is to develop an end-to-end framework for the classification of hyperspectral images. The proposed hyperspectral framework can be used in online decision support systems in a variety of industrial applications, for instance to detect wood defects and fruit diseases, and to predict the status of agriculture fields for precision agriculture. Results on wood hyperspectral images classification are provided in [Htun *et al.*, 2021].

8 Temporal Databases

Project

- Title: *Enabling Industrial-Strength, Open-Source Temporal Query Processing (ISTEP)*
- Funding scheme: Research Südtirol/Alto Adige 2019, Province of Bozen-Bolzano
- Duration: 01/10/2020 - 30/09/2023
- Budget: 260 K€ for unibz
- Web: https://dbs.inf.unibz.it/projects/istep/

Participants

- From unibz: Johann Gamper (PI), Anton Dignös (Co-PI);
- External supporters: Michael Böhlen (University of Zurich), Christian Jensen (Aalborg University).

Description. Temporal data is ubiquitous, and its importance has been witnessed by the recent introduction of temporal features in the SQL standard and commercial DBMSs. Despite such efforts, current DBMSs offer little support for query formulation and processing. This makes application

code and SQL queries complex and inefficient. Temporal alignment is the first and only framework that provides comprehensive temporal query support over interval-timestamped data and allows a tight integration into existing DBMSs (see http://tpg.inf.unibz.it).

The ISTEP project aims at boosting the efficiency of the temporal alignment framework as follows: (1) Instead of using two primitives to transform temporal queries into non-temporal queries, which might produce redundancy in the intermediate relations, we are exploring new alignment primitives that are customized for specific temporal operators. Results on built-in computation for temporal joins have been obtained [Dignös *et al.*, 2021]. (2) Since current cost estimates for intermediate relations are very conservative, we are developing more realistic cost estimates. (3) We are studying new algebraic equivalence rules for query optimization. If successful, the project will provide the foundation for the first industrial-strength, open-source RDBMS with comprehensive temporal query support.

References

[Alman *et al.*, 2021] Anti Alman, Claudio Di Ciccio, Fabrizio Maria Maggi, Marco Montali, and Han van der Aa. , *Proc. of the 19th Int. Conf. on Business Process Management (BPM 2021)*, volume 12875. Springer, 2021.

[Artale *et al.*, 2019] Alessandro Artale, Alisa Kovtunova, Marco Montali, and Wil M. P. van der Aalst. , *Proc. of the 17th Int. Conf. in Business Process Management (BPM 2019)*, volume 11675. Springer, 2019.

[Baral and De Giacomo, 2015] Chitta Baral and Giuseppe De Giacomo. , *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI 2015)*. AAAI Press, 2015.

[Calvanese *et al.*, 2013] Diego Calvanese, Giuseppe De Giacomo, and Marco Montali. , *Proceedings of the 32nd ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS 2013)*. ACM, 2013.

[Calvanese *et al.*, 2017] Diego Calvanese, Benjamin Cogrel, Sarah Komla-Ebri, Roman Kontchakov, Davide Lanti, Martin Rezk, Mariano Rodriguez-Muro, and Guohui Xiao. Ontop: Answering SPARQL queries over relational databases. *Semantic Web J.*, 8(3), 2017.

[Calvanese *et al.*, 2018] Diego Calvanese, Giuseppe De Giacomo, Marco Montali, and Fabio Patrizi. First-order μ-calculus over generic transition systems and applications to the situation calculus. *Inf. Comput.*, 259(3), 2018.

[De Giacomo *et al.*, 2022] Giuseppe De Giacomo, Paolo Felli, Brian Logan, Fabio Patrizi, and Sebastian Sardiña. Situation calculus for controller synthesis in manufacturing systems with first-order state representation. *Artificial Intelligence*, 302, 2022.

[Dignös *et al.*, 2021] Anton Dignös, Michael H. Böhlen, Johann Gamper, Christian S. Jensen, and Peter Moser. *VLDB Journal*, 0, 2021.

[Dumas, 2011] Marlon Dumas. , *Proc. of the 15th Int. Conf. on Advances in Databases and Information Systems (ADBIS 2011)*, volume 6909. Springer, 2011.

[Felli *et al.*, 2022] Paolo Felli, Marco Montali, and Sarah Winkler. , *Proceedings of the 36-th AAAI Conference on Artificial Intelligence (AAAI 2022)*. AAAI Press, 2022. To appear.

[Franconi, 2019] Enrico Franconi. , *Ital-IA Workshop on AI for Space*, 2019.

[Ghilardi *et al.*, 2020] Silvio Ghilardi, Alessandro Gianola, Marco Montali, and Andrey Rivkin. , *Proc. of the 18th Int. Conference on Business Process Management (BPM 2020)*, volume 12168. Springer, 2020.

[Htun *et al.*, 2021] Phyu Phyu Htun, Marco Boschetti, Attaullah Buriro, Roberto Confalonieri, Boyuan Sun, A. Htwe, and Tammam Tillo. , *Proc. of the 2021 IEEE Int. Conf. on Multimedia & Expo Workshops (ICME)*, Los Alamitos, CA, USA, 2021. IEEE Computer Society.

[Kalayci *et al.*, 2020] Elem Güzel Kalayci, Irlan Grangel González, Felix Lösch, Guohui Xiao, Anees ul Mehdi, Evgeny Kharlamov, and Diego Calvanese. , *Proc. of the 19th Int. Semantic Web Conf. (ISWC)*, volume 12507. Springer, 2020.

[Kharlamov *et al.*, 2015] Evgeny Kharlamov, Ernesto Jiménez-Ruiz, Christoph Pinkel, Martin Rezk, Martin G. Skjæveland, Ahmet Soylu, Guohui Xiao, Dmitriy Zheleznyakov, Martin Giese, Ian Horrocks, and Arild Waaler. , *Proc. of the 14th Int. Semantic Web Conf. (ISWC) Posters & Demonstrations Track*, volume 1486 of *CEUR Workshop Proceedings,* http://ceur-ws.org/. CEUR-WS.org, 2015.

[Maggi and Westergaard, 2014] Fabrizio Maria Maggi and Michael Westergaard. Using timed automata for *a Priori* warnings and planning for timed declarative process models. *Int. J. Cooperative Inf. Syst.*, 23(1), 2014.

[Maggi *et al.*, 2020a] Fabrizio Maria Maggi, Marco Montali, and Rafael Peñaloza. , *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI 2020)*. AAAI Press, 2020.

[Maggi *et al.*, 2020b] Fabrizio Maria Maggi, Marco Montali, Rafael Peñaloza, and Anti Alman. , *Proc. of the 18th Int. Conf. on Business Process Management*, volume 12168. Springer, 2020.

[Montali *et al.*, 2010] Marco Montali, Maja Pesic, Wil M. P. van der Aalst, Federico Chesani, Paola Mello, and Sergio Storari. Declarative specification and verification of service choreographies. *ACM Trans. Web*, 4(1), 2010.

[van der Aalst, 2016] *Process Mining - Data Science in Action, Second Edition*. Springer, 2016.

[Xiao *et al.*,] Guohui Xiao, Davide Lanti, Roman Kontchakov, Sarah Komla-Ebri, Elem Güzel-Kalayci, Linfang Ding, Julien Corman, Benjamin Cogrel, Diego Calvanese, and Elena Botoeva. , *Proc. of the 19th Int. Semantic Web Conf. (ISWC)*.

[Xiao *et al.*, 2018a] Guohui Xiao, Diego Calvanese, Roman Kontchakov, Domenico Lembo, Antonella Poggi, Riccardo Rosati, and Michael Zakharyaschev. , *Proc. of the 27th Int. Joint Conf. on Artificial Intelligence (IJCAI)*. IJCAI Org., 2018.

[Xiao *et al.*, 2018b] Guohui Xiao, Roman Kontchakov, Benjamin Cogrel, Diego Calvanese, and Elena Botoeva. , *Proc. of the 17th Int. Semantic Web Conf. (ISWC)*, volume 11136. Springer, 2018.

[Xiao *et al.*, 2019] Guohui Xiao, Linfang Ding, Benjamin Cogrel, and Diego Calvanese. Virtual Knowledge Graphs: An overview of systems and use cases. *Data Intelligence*, 1(3), 2019.