A *Mathematical Programming (MP) model* is a common representation for virtually all types of real-world entities, e.g., production lines, delivery schedules, personnel assignment. The MP model consists of variables that correspond to, e.g., production volume, constraints that represent the relationships between the variables, e.g., production conditions, and an objective function that corresponds to the outcome, e.g., production cost. A feasible solution is a vector of values of variables that satisfies all constraints. The feasible solution that minimizes the objective function is the optimal solution. The use of MP models is fully automated thanks to the solvers — the tools that produce solutions, either feasible or optimal. However, building the correct MP model requires intense training and expertise, and often turns out to be time-consuming and error-prone. The errors in MP models are difficult to spot and often remain unidentified until the optimal solution turns out to be inapplicable in practice, resulting in many iterations of modeling, conformance checking, and model enhancement. MP models have many advantages over many other representations, e.g., interpretable structure offering explainable decisions, executable semantics, and mature solver tools.

The project aim is to help human experts by developing new *MP Model Mining* algorithms. Our contribution is divided into four areas: discovery, conformance checking, enhancement, and verification in real-world settings. We use easy-to-provide data, such as exemplary solutions recorded by computers managing the modeled entity, available quantities, e.g., parameters, sets, variables, and/or other background knowledge information. *Novel discovery algorithms* will build from the data the MP models that maximize three criteria: fitness, precision, and generalization. Fitness evaluates how well the MP model embraces the data, precision assesses how tight the MP model is, and generalization measures how well the MP model describes different instances of the modeled entity. *Novel conformance checking algorithms* will calculate these measures given the existing MP model and the data. *Novel enhancement algorithms* will identify deficiencies in the existing MP model and propose fixes. We will develop algorithms that employ high-level modeling languages, e.g., AMPL and ZIMPL. We will also propose new measures for fitness, precision, and generalization, as the contemporary Artificial Intelligence measures are largely inappropriate to MP models, e.g., award the models equally for all examples and ignore syntax while for most MP models concise representation and tight constraints are more important. We will also seek new algorithms that propose fixes in MP models driven by, e.g., irreducible inconsistent subsystem and counterexample data. To date, most of the work in the field has been done in the context of the discovery problem. However, the existing algorithms are immature and are not ready to be used in large-scale real-world problems. They suffer from the curse of dimensionality for even 6-8 variables, do not handle well one-class data, and/or have unacceptable computation complexity.

Our ultimate goal is to overcome these challenges. The only way to reliably validate the developed algorithms is their use to build MP models for complex real-world *Operations Research (OR) problems*. To this aim, we will investigate current and past OR challenges, e.g., ROADEF and PACE challenges. They consist of problems of high-importance in industry and data donated by international companies. Exemplary past problems in these challenges are truck loading by Renault, maintenance planning by RTE, cutting by Saint-Gobain, and inventory routing by Air Liquide. We will use the proposed algorithms to develop proof-of-concept MP models ready to solve these OR problems. Our contributions will make foundations for better modeling and optimization tools for almost every sector of economy. Note that solving the problems originating in industry is not the project goal per se but only a side effect of basic research on algorithms for MP model mining.