A Methodology for Constructing Patterns for the Management of Data Science Projects

Christian Haertel, Sarah Schramm, Matthias Pohl, Sascha Bosse, Daniel Staegemann, Christian Daase, and Klaus Turowski
Institute of Technical and Business Information Systems, Otto-von-Guericke University Magdeburg, Magdeburg, Germany
christian.haertel, sarah.schramm, matthias.pohl, sascha.bosse, daniel.staegemann, christian.daase, klaus.turowski}@ovgu.de

Introduction

As the amount of generated data is steadily increasing, businesses of various domains aim to derive potential advantages and enhance their competitive positioning (de Medeiros et al., 2020). Data Science (DS) aims to extract knowledge and insights from data using various methods and techniques (Chang and Grady, 2019) and thus, has gained increasing significance (Cao, 2017). Organizations often encounter challenges in implementing these projects (Martinez et al., 2021) and there is no widely accepted and applied approach. This is reflected in the low DS project success rate (VentureBeat, 2019), demanding improvements for DS project management (Saltz and Krasteva, 2022). As the literature identifies common problems in the execution of DS projects (Martinez et al., 2021a), the adaptation of the pattern concept to DS appears promising.

Pattern Construction Process for Data Science Project Management

In the literature, there are different approaches to the development of patterns. The process according to (Fehling et al., 2014) was selected and adapted to the field of DS project management since its applicability to various domains has already been demonstrated. The method consists of three phases: pattern identification, pattern authoring, and pattern application, as illustrated. Each stage involves several iteratively traversed activities to continuously improve and adapt the developed results (Fehling et al., 2014). In particular, the first two phases are repeated multiple times to discover and form patterns. Finally, the third phase involves refining the patterns for specific use cases or application environments.

Pattern Search and Recommendation: accompany DS patterns with a summary of the problem and solution of each of the patterns to facilitate navigation and identify suitable patterns.

Pattern-based Solution Design: expand patterns by the section ‘Examples’ to provide a reference solution of the DS problem within the given pattern.

Refinement of the Solution Design: patterns are constrained and adapted to a specific environment where they should be applied.

Instantiation of the Solution Design: determine the means to manage, configure, and deploy the patterns.

References


Contact: Christian Haertel, Magdeburg Research and Competence Cluster, christian.haertel@ovgu.de, www.mrcc.eu