The Beauty and the Beast: Patterns and Anti-Patterns in use of Data

About

This study investigates how companies in the software-intensive systems industry manage and utilize the vast amounts of data collected from their products. Through a multi-case study of eight companies, the research identifies beneficial practices, or 'patterns', that lead to successful data use, and detrimental practices, or 'anti-patterns', that companies should avoid.

Problem

Companies are collecting ever-increasing volumes of data from their products but often struggle to effectively transform this raw data into customer and business value. There is a lack of practical guidance on what constitutes best practices, leading to significant challenges in data management, quality assurance, access control, and cost management.

Study Outcome

- Successful companies follow beneficial patterns such as: exploring new data-driven services at low cost on internal infrastructure, operating monetized services with high-quality data on reliable public clouds, and collaborating with partners to generate new insights.
- Detrimental 'anti-patterns' that hinder success include: using raw, low-quality data for commercial services, allowing ungoverned data access across teams which leads to system failures, and incurring excessive costs from using public clouds for exploratory data analysis.
- Key operational challenges across companies are poor data quality, which consumes significant time for cleaning (up to 80% of effort in some cases), and the high costs associated with data storage and computation, especially in the cloud.

Keywords

data practices • software-intensive systems • patterns • anti-patterns • data management • case study • cloud computing