

# Data Quality Assessment: Challenges and Solutions

Lisa Ehrlinger<sup>1</sup>

<sup>1</sup>Hasso Plattner Institute, Potsdam, Germany

## Abstract

Artificial intelligence (AI) systems, such as chatbots or autonomous systems, are only as good as the data they are trained on, following the well-known “garbage-in, garbage-out” principle. Similarly, data quality directly impacts human decision-making across different domains, from healthcare to manufacturing. In the past, decision makers were still able to manually assess and interpret the quality of data at hand. However, with recent advances in digitalization and the deployment of AI systems in practice, the amount of data being collected, stored, and consequently used for automated decision-making, exceeds the capabilities of humans to process it. Hence, the need for automated data quality assessment and improvement methods has developed.

In this talk, we will examine the critical impact of data quality on both AI predictions and human decision-making through practical examples. Understanding the significance of data quality motivates the need for systematic assessment across different dimensions, such as completeness, accuracy, or understandability. The talk will highlight four current research challenges for automating data quality assessment and present possible solutions to address these challenges:

- (C1) Creating a unified data quality framework
- (C2) Developing metrics for data quality dimensions
- (C3) Managing metadata for data quality assessment
- (C4) Dealing with data evolution

## Acknowledgments

This invited talk is based on research together with Divya Bhadauria, Carolina Cortes, Lorena Etcheverry, Philipp Hildebrandt, Hazar Harmouch, Sedir Mohammed, Felix Naumann, and Divesh Srivastava in the context of the Metis project: [www.metisdq.org](http://www.metisdq.org).

---

36th GI-Workshop on Foundations of Databases (*Grundlagen von Datenbanken*), September 29 - October 01, 2025, Regensburg, Germany

✉ [lisa.ehrlinger@hpi.de](mailto:lisa.ehrlinger@hpi.de) (L. Ehrlinger)

🌐 <https://hpi.de/naumann/people/dr-lisa-ehrlinger.html> (L. Ehrlinger)

🆔 0000-0001-5313-0368 (L. Ehrlinger)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).