Data Quality Assessment Results as Decision-Relevant Enriched Metadata for Decision Support Systems

Trustworthy and impactful decision support systems (DSS) critically depend on data that are fit for their intended use. Here, data quality assessments should be understood not merely as a validation step, but as a systematic metadata enrichment process that produces structured, machine-actionable descriptors of data plausibility, and correctness.

Building on established data quality concepts in medical research, this talk generically conceptualises automated data quality assessment as a crucial pre-assessment layer that is analytically distinct from decision logic. Results from such assessments inform decision logic on appropriate data use.

To implement such assessment in an efficient and transparent manner, we have implemented a layered metadata model comprising: (i) metadata describing core data quality dimensions such as missing data, inadmissible data, contradictions, or unexpected associations; (ii) metadata capturing the formal rules and expectations used to detect data quality issues; (iii) metadata that classifies the severity and decision relevance of the detected issues within a specific application context - this classification is inherently normative and context dependent, it reflects explicit assumptions about how particular data limitations may affect decisions; and from the first three layers finally (iv) the derived fitness-for-decision metadata that explicitly characterises the suitability of a dataset for downstream decision processes.

Using the tools dataquieR (R) and dqrep (Stata) as concrete implementations, based on an example from the medical domain, we demonstrate how this separation of assessment and decision layers enables context-specific, automated generation of interoperable data quality metadata with minimal programming effort. We illustrate how such metadata can shape DSS behaviour—for example by suppressing unreliable alerts, or down-weighting model inputs.

We conclude by discussing implications for metadata standards, interoperability of data quality assessment results. Treating such results as decision-relevant metadata provides a structured mechanism to potentially increase the value of DSS.

ONLY WORKSHOPS - Proposed interaction format

Alternative Track

ONLY WORKSHOPS - Tentative audience

ONLY WORKSHOPS - Maximum number of participants

ONLY WORKSHOPS - Special technical requirements

Autoren: SCHMIDT, Carsten Oliver (Universitätsmedizin Greifswald); Dr. SALOGNI, Elena; Dr. STRUCK-MANN, Stephan

Vortragende(r): SCHMIDT, Carsten Oliver (Universitätsmedizin Greifswald)

Sitzung Einordnung: TALK SESSION

Track Klassifizierung: HMC Conference 2026 Track Topics: 1. Metadata in Action: Embedding Quality and Context into Research Infrastructures