#Valerie Restat#Meike Klettke#Uta Störl#

$A Review on Data Cleansing Methods for Big Data$Social-minded Measures of Data Quality: Fairness, Diversity, and Lack of Bias$Taming technical bias in machine learning pipelines$Kurz erklärt: Measuring data changes in data engineering and their impact on explainability and algorithm fairness$Data Wrangling: Making data useful again$Foundations of Data Quality Management$Four Generations in Data Engineering for Data Science: The Past, Presence and Future of a Field of Science$Towards Automated Data Cleaning Workflows$CleanML: A Study for Evaluating the Impact of Data Cleaning on ML Classification Tasks$Effective Data Cleaning with Continuous Evaluation$The Challenges of Data Quality and Data Quality Assessment in the Big Data Era$Data quality: A survey of data quality dimensions$The FAIR Guiding Principles for scientific data management and stewardship$Towards reliable interactive data cleaning: a user survey and recommendations$Detecting Data Errors: Where are we and what needs to be done? $Data Preparation: A Survey of Commercial Tools$FAIR Computational Workflows$Nullius in Verba: Reproducibility for Database Systems Research$Artifact review and badging – version 2.0$Improving Reproducibility of Data Science Pipelines through Transparent Provenance Capture$A Link is not Enough - Reproducibility of Data$Data Management in Machine Learning Systems$BoostClean: Automated Error Detection and Repair for Machine Learning$HoloClean: Holistic Data Repairs with Probabilistic Inference$HoloDetect: Few-Shot Learning for Error Detection$Raha: A Configuration-Free Error Detection System$Baran: Effective Error Correction via a Unified Context Representation and Transfer Learning$TFX: A TensorFlow-Based Production-Scale Machine Learning Platform$Data Cleaning: Overview and Emerging Challenges$Towards Explaining the Effects of Data Preprocessing on Machine Learning$Towards Generating Consumer Labels for Machine Learning Models$A Survey on Bias and Fairness in Machine Learning$Beyond Fairness Metrics: Roadblocks and Challenges for Ethical AI in Practice$EvoBench: Benchmarking Schema Evolution in NoSQL$An Inquiry into Machine Learning-based Automatic Configuration Tuning Services on Real-World Database Management Systems$