



UNIVERSITÀ DEGLI STUDI DI MILANO

Dipartimento di Informatica
Via Celoria 18 - 20133 MILAN - ITALY

May 30, 2024

Prof. Ernesto Damiani
Editor-in-Chief
IEEE Transactions on Services Computing

Dear Prof. Borko, Dear Prof. Taghi,

We wish to submit an original research article for possible publication in Journal of Big Data.

Title: Balancing Data Quality and Protection in Distributed Big Data Pipelines

Authors: Marco Anisetti, Claudio A. Ardagna, Chiara Braghin, Ernesto Damiani, Antongiacomo Polimeno

I confirm that this work is original and has not been published elsewhere, nor is it currently under consideration for publication elsewhere.

In this paper, we report on a data governance framework designed to mitigate privacy and security risks in modern data-driven pipelines. This is significant because it addresses the dual challenge of maintaining high data quality while ensuring robust data protection in a multitenant cloud environment, which is increasingly important given the widespread adoption of cloud infrastructures.

We believe that this manuscript is appropriate for publication by the Journal of Big Data because it aligns with the journal's Aims and Scope, particularly in the areas of big data analytics, data privacy, and security.

The primary contributions of the paper can be summarized as follows:

- Enriching services with metadata that describes both data protection and functional requirements.
- Proposing a parametric heuristic tailored to address the computational complexity of the NP-hard service selection problem.
- Evaluating the performance and quality of the algorithm through experiments conducted using a real-world dataset.

We have no conflicts of interest to disclose. Please address all correspondence concerning this manuscript to me at claudio.ardagna@unimi.it. Thank you for your consideration of this manuscript.

Best regards,

Claudio Ardagna