SORS: Enhancing FAIRness in Harvard Dataverse with Variable-Level Metadata and Differential Privacy

Abstract: We highlight recent advances in the Dataverse Project aimed at extending support for variable-level metadata beyond tabular data, incorporating DDI, CDI-DDI, CroissantML, and other schemas to enhance discoverability and interoperability. Additionally, we explore the integration of differential privacy techniques to improve access to sensitive data while maintaining confidentiality, thereby promoting broader accessibility and compliance with FAIR principles.

Short Bio

Stefano M. Iacus is the Director of Data Science and Product Research at the Institute for Quantitative Social Science (IQSS), Harvard University. He also serves as the Managing Director of the Dataverse Project and is a member of the executive committee of the OpenDP Project. In addition, he oversees the Data Science Services and the Data Acquisition & Archiving teams at IQSS. Iacus is an affiliate faculty member of the Kempner Institute for the Study of Natural and Artificial Intelligence at Harvard.
Iacus began his academic career at the University of Milan (Italy), where he became a full professor of statistics in 2015. He founded and directed the Data Science Lab and created two master's programs in Finance and Economics, and Data Science for Economics. From 2019 to 2022, he served as an officer at the Joint Research Centre of the European Commission, leading a team that leveraged non-traditional data sources for evidence-based policy-making in areas like migration and demography, particularly during crises and to improve preparedness measures.
Since 2006, Iacus has held a recurring visiting position at the Graduate School of Mathematics at the University of Tokyo, where he co-leads the Yuima Project. He was a member of the R Core Team from 1999 to 2014 and remains a member of the R Foundation for Statistical Computing.
Beyond academia, Iacus played a critical role during the COVID-19 pandemic, managing a large-scale business-to-government project for the European Commission. This initiative used data from mobile network operators to generate insights for policy-making across European Union member states.
Iacus has published numerous books, scientific articles, and open-source software products, covering fields such as causal inference, sentiment analysis, stochastic processes, computational statistics, and quantitative finance. His work is widely cited, and he has founded two startup companies in social media analysis and quantitative finance.

Speakers

Speaker: Stefano M. Iacus is the Director of Data Science and Product Research at the Institute for Quantitative Social Science (IQSS), Harvard University.
Chair: Elena Rovenskaya, International Institute for Applied Systems Analysis
Host: Mercè Crosas, Head of Computational Social Sciences and Humanities, BSC
Barcelona Supercomputing Center - Centro Nacional de Supercomputación

---