Setting the Basis and Designing the Data Architecture for the "EN.I.R.I.S.S.T." Research Infrastructure

Amalia Polydoropoulou[a], Manolis Maragkoudakis[b], Ioanna Pagoni[a], Athena Tsirimpa[a], Annitsa Koumoutsidi, Ioannis Tsouros[a], Christos Tryfonopoulos[c]

[a] Department of Shipping, Trade and Transport,
University of the Aegean, Chios island, Greece

[b] Department of Information and Communication Systems Engineering,
University of the Aegean, Samos Island, Greece

[c] Software and Database Systems Lab, Department of Informatics & Telecommunications,
University of the Peloponnese, Greece

Within the past years, the transportation sector has been affected by the several advancements occurring in the era of big data. This has increasingly attracted the attention of both scientists and practitioners in the private and public sector and several studies and applications have been developed in this field. Despite the many advantages of big data, several challenges should be addressed to gain the full potential of big data. On the one hand, data analysis is very complex as it consists of multiple phases including data acquisition, processing, aggregation, and delivery. On the other hand, the appropriate big data tools and services should be applied. Therefore, designing an optimal data architecture is crucial.

This paper focuses on the Intelligent Research Infrastructure for Shipping, Supply chain, Transport and Logistics (EN.I.R.I.S.S.T.) which is a unique and innovative Research Infrastructure (R.I.) established in 2019 in Greece in the context of the National Strategy for Research, Technological Development and Innovation (2014-2020) in the National Roadmap of Large-Scale Research Infrastructures. The aim of this paper is two-fold. First, it will provide the current state-of-the-art on the established research infrastructures in Europe and worldwide. The review will identify the best practices adopted to formulate sustainable e-infrastructures that provide continuity addressing long-term changes but are flexible enough to respond to the modification of research priorities due to the rapid societal and technological developments. In addition, it will present the data architecture that will be developed in the context EN.I.R.I.S.S.T. Its goal is to become a widely acceptable, multi-purpose data analytics platform that will unify a wide variety of open data sources and enhance collaborations among different disciplines in the transportation sector. The Infrastructure relies on the NIST Big Data Reference Architecture supporting data cataloging, virtualization, analytics and visualization through the use of open source tools.

Keywords: research e-infrastructure, transport, logistics, supply chain, innovation, big data, data analytics.

Acknowledgments