Editorial
Volume 15:4, 2022

Journal of Computer Science & Systems Biology

ISSN: 0974-7230 Open Access

Editorial on Data Sharing and Environmental Monitoring

Jose Rodriguez*

Department of Information Science, University of Berlin, Berlin, Germany

Editorial

Climate change and human activities are having an increasing impact on ecosystems, and documenting global change patterns has been identified as a first step to dealing with the changing environment and understanding global biodiversity loss, necessitating the need for quick, dependable access to high-quality data. It emphasises the importance of standards in data and metadata, for example, in ecology, without being fully defined. Even though a number of repositories for storing shared data have been established, and ongoing developments are attempting to define common data and metadata standards, existing data are not exposed in a standard, machine-readable format using a common vocabulary.

As Deligiannis explain in, many stakeholders lack the resources for infrastructure and/or computing expertise and continue to rely on antiquated methods such as I storing data in spreadsheets or raw files, (ii) sharing data with colleagues via email, cloud uploads of zip files, or even snail mailing electronic copies in removable media, and (iii) analyzing data using sub-standard tools and trial software. They also stated that the proposed data storage solutions typically necessitate significant computing infrastructure and necessitate the constant support and active participation of Information Technology (IT) experts, even for minor tasks such as incorporating new data. Data heterogeneity [1-3] and interoperability remain common challenges.

However, this necessitates the creation of infrastructures such as data repositories that allow data to be shared and valued. Ontologies and Findable, Accessible, Interoperable, and Reusable (FAIR) systems can currently handle these challenges effectively. According to Wilkinson et al., there are numerous and diverse stakeholders who have much to gain from researchers wanting to share, obtain credit for, and reuse each other's data and interpretations; funding agencies (private and public) are increasingly concerned with long-term data stewardship; and the data science community is interested in mining, integrating, and analysing new and existing data to advance discovery.

This type of system is well suited for environmental studies and smart agriculture because it allows for long-term monitoring by deploying a monitoring system in a remote area where a high data rate is not required and where minimal power consumption is sufficient (see for example for environment applications and for smart agriculture application). The frequency of transmission is important in near real-time processing, but the total processing time (from acquisition to visualization) can be measured in minutes (up to ten), especially if the environmental process under study is stable. A spatial analysis and representation are commonly implemented in the environmental domain, at least by georeferencing with GPS WSG84 latitude and longitude combined with a UTC timestamp [4,5].

As a result, much research is being conducted on cloud data storage, but few works take into account the complete integration of data from the sensor to the cloud for long term with various types of sensors and on multiple sites. Nonetheless, long-term data collection is being worked on in large national data centres or data hubs. Along with new technologies that enable data analysis, storage and processing architecture play an important role in the big data ecosystem. Whatever architectures are proposed, their goal is to make storing, manipulating, analyzing, and accessing structured and unstructured data easier. Among these modern architectures, data lakes are of particular interest because their design aims to store a large volume of data in any format and structure, as well as to provide services for data access and analysis.

They are especially useful in environmental monitoring and agriculture because they can manage raw data, whether structured or unstructured. A data lake is a low-cost data storage system that can improve the data analysis process from data ingestion to storage, exploration, and exploitation in its native format. In this paper, we present the "Cloud Environmental au Bénéfice de l'Agriculture" (CEBA), one deliverable of the Challenge "Sustainable agroecosystems in the Context of Global Change" of the Initiatives Science–Innovation–Territories–Economy (I-Site) Clermont CAP 20-25 Project.

The CEBA proposes an on-premise data lake to fill the gap in Auvergne (France) for a regional tool to store, visualize, analyze, and share heterogeneous data (structured or unstructured) collected from research programmes and studies related to the environment and agriculture, whether these data come from files, databases, or sensor networks, with an authentication service and data ownership. The CEBA's contributions include an innovative on-premise data lake with the following features: a data catalogue managing spatially referenced resources, an ingestion platform to collect IoT data regardless of format and structure, a storage platform to upload and store any type of data file with query solutions, and a visualisation platform to access hot data in near real-time.

Conflict of Interest

None.

References

1. Audic, Stéphane and Jean-Michel Claverie. "The significance of digital gene expression profiles." *Genome Res* 7 (1997): 986-995.
2. Henderson, Ian R, Xiaoyu Zhang and Cheng Lu, et al. "Dissecting *Arabidopsis thaliana* DICER function in small RNA processing, gene silencing and DNA methylation patterning." *Nat Genet* 38 (2006): 721-725.
3. Glynn, Shawn M., Gita Taasoobshirazi and Peggy Brickman. "Science motivation questionnaire: Construct validation with nonscience majors." J Res Sci Teach 46 (2009): 127-146.
4. Moorthy, N., P.C. Jobe Prabakar, S. Ramalingam and G.V. Pandian, et al. "Vibrational, NMR and UV–visible spectroscopic investigation and NLO studies on benzaldehyde thiosemicarbazone using computational calculations." *J Phys Chem Solids* 91 (2016): 55-68.
5. Ponhong, Kraingkrai, Sam-ang Supharoek and Watsaka Siriangkhawut, et al. "A rapid and sensitive spectrophotometric method for the determination of benzoyl peroxide in wheat flour samples." *J Food Drug Anal* 23 (2015): 652-659.

***Address for Correspondence:** Jose Rodriguez, Department of Information Science, University of Berlin, Berlin, Germany, E-mail: JoseRodriguez125@gmail.com*

***Copyright:** © 2022 Rodriguez J. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.*

Received: 03-Apr-2022, Manuscript No. jcsb-22-66581; **Editor assigned:** 04-Apr-2022, Pre QC No. P-66581; **Reviewed:** 09-Mar-2022, QCNo.Q-66581; **Revised:** 14-Apr-2022, Manuscript No.R-66581; **Published:** 19-Apr-2022, DOI: 10.37421/0974-7230.2022.15.409.

How to cite this article: Rodriguez, Jose. "Editorial on Data Sharing and Environmental Monitoring." J Comput Sci Syst Biol 15 (2022): 409.