Data Citation Wiki for Harnessing Collective Intelligence on Document-to-Data Associations to Interdisciplinary Data Access

Eloy Gonzales, Xinpeng Zhang, Yuhei Akahoshi, Yasuhiro Murayama and Koji Zettsu.

National Institute of Information and Communications Technology
3-5 Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-0289, Japan
e-mail: {egonzales ,xp.zhang, akahoshi, murayama, zettsu}@nict.go.jp

Poster Presentation (Semantic Web and Linked Data)

Abstract:
Data citation (DC) allows users to edit and create references between web documents and scientific datasets. It finds the citation correlations. DC has been increasingly considered by the scientific community as a necessity for data sharing, data re-use and attribution. DC is considered as the appropriate tool to encourage data publication for scholarly acknowledgment. DC provides not only data access and retrieval, but also validation, verification and replication of data underlying published studies. Moreover, DC may also be used for increasing the comprehensibility of the contents of web documents such as articles, news or scientific papers and also for extending the information necessity to support the reliability and accuracy of web documents. The proposed **DC-Wiki System** allows users to edit and create references between web documents and scientific datasets via an interactive interface. DC-Wiki System aims to extend the use and practice of data citations especially among scientists and researchers. The tool provides a robust meta-data to a search engine in order to find the most closely scientific datasets which can be also be accessed and visualized. Most of the other related projects do not focus so much in the consistency of citations. In the proposed system, three levels of consistency are defined to ensure the validity of citations. Moreover, a citation update mechanism is also provided within a user-friendly environment. We will demonstrate the prototype of our DC-Wiki system and its application to cross-DB search for interdisciplinary use of large-scale, multi-domain and heterogeneous databases.

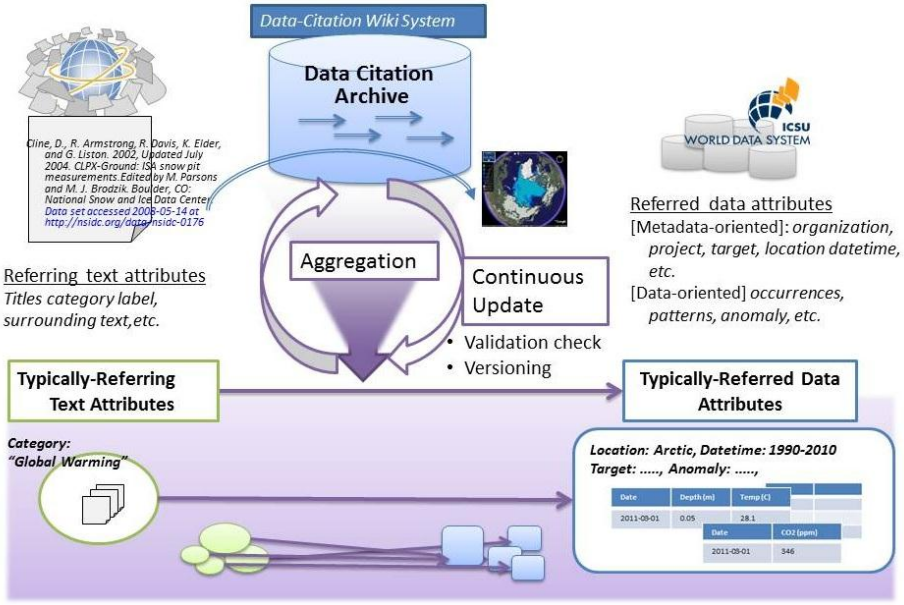

Fig. 1. Document-to-Data correlation analysis mechanism.