

Cross-Database Search for Interdisciplinary Use of Large-Scale, Multi-Domain and Heterogeneous Databases

Koji Zettsu, Eloy Gonzales, Bun Theang Ong and Yasuhiro Murayama
National Institute of Information and Communications Technology
3-5 Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-0289
e-mail: {zettstu, egonzales, ong_bt, murayama }@nict.go.jp

Abstract:

Recently, there are some efforts especially in the scientific community to share and access to scientific Big Data in order to discover scientific novelties not only within their own scientific domains but also within other related domains. That means the intrinsic interdisciplinary nature of the data-intensive science of “the fourth paradigm”. The main objective of our cross-DB search system is to facilitate search of interdisciplinary-correlated data sets from large-scale, multi-domain and heterogeneous data like WDS. Despite of conventional efforts such as system or portals for searching datasets, most of them are simple keyword-based search engines, thus the scientist must know the related dataset in advance or must find their relations manually. Our proposed system, instead, facilitates to search the datasets based on their latent associations by spatiotemporal, ontological, and citation correlations. Scientist do not need to know the relations beforehand, the system automatically provides enough set of datasets and their relations in order to appropriately respond to the scientist data requests. The basic concepts of our cross-DB search system are introduced as well as a demonstration of the prototype system.

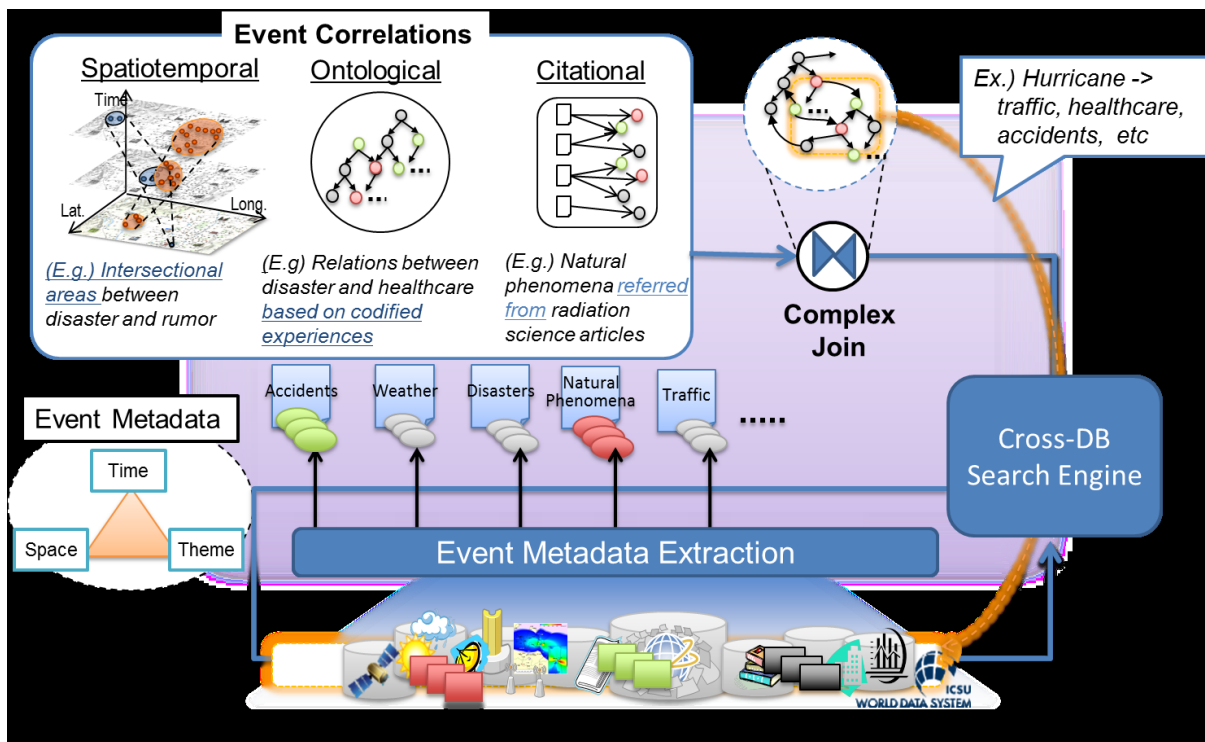


Figure 1: Cross-DB Search based on complex join of spatiotemporal, ontological and citational correlations among event metadata.