

## Automatic Collection System of Metadata based on RDF and Development of Solar-Terrestrial Science Data Semantic We

# Ken T. Murata[1]; Takuya Kubo[2]; Eizen Kimura[3]; Kazunori Yamamoto[4]; Mamoru Ishii[5]

[1] CITE, Ehime University; [2] Ehime Univ; [3] CITE, Ehime Univ.; [4] Ehime Univ.; [5] NICT

<http://www.infonet.cite.ehime-u.ac.jp>

**Background of current information web** Over the Internet, various information webs are independently managed. For example, when we come across an unknown word in an article on news web, we need to access to technical terms web. It means that relevance of information between webs must be judged by human. In the era of such flood of information, boundaries between required information and unnecessary information tend to be unclear. Taking into account of the limitation of man-power to process such large scale information flood, it is unreasonable that these judgments are done by human.

**Background of STP field** In the Solar-Terrestrial Physics fields, researchers analyze multi-points observation data in between the Sun and the Earth. For example, assume that there are three observation data; at an upstream point of the Earth, at the middle point in between the Sun and the Earth, and nearby the Earth. These time-sequential plots are created using observation data files managed by different data centers. Our data analysis environment is insufficient because each data center manages data files independently. In fact, when we create these data plots, we have to collect data files from these data centers by ourselves. Researcher has a difficulty in collecting data because they need to have a-priori knowledge for URL information, observed data and so on. It takes a lot of time.

**Our approach (meta-data database)** To overcome the problem of independent data file management, we have been developing a meta-data database system. It provides cross-over data file searches and download services. The meta-data database deals with the file name, URL, file access protocol, file permission, and so on. We have already stored more than 2 million data file records in the meta-data database. User is able to download data files without paying attention to data file location and file access permission. Meta-data collection system is designed based on RSS1.0. Using this system, collected meta-data in RSS1.0 (RDF) format are saved in an RSS database (Semame). In the present system we attempt to construct an integrated semantic web database for Solar-Terrestrial Science fields.