

# Prototyping a metadata set based on the relationships description on Seismic Hazard

\*Hiroki Azuma<sup>1</sup>, Tadamichi Matsuyama<sup>2</sup>

1. National Research Institute for Earth Science and Disaster Resilience, 2. SOUM Corporation

## 1. Background

Following a series of efforts related to open data that began around the year 2013, the government sets forth the principle of public data and open data of data created by the government budget in the Basic Principles on Open Data (2017). Technically, LOD (Linked Open Data) is defined in 5 stages, and data set using RDF (Resource Description Framework) is cited as its highest position. Meanwhile, the Headquarters for Earthquake Research Promotion (HERP) has published the National Seismic Hazard Maps (SHM), and detailed data has been published through the J-SHIS (Japan Seismic Hazard Information Station) of the National Research Institute for Earth Science and Disaster Resilience (NIED) since 2005. Currently J-SHIS provides all the J-SHIS Web APIs that are also used for commercial use by third parties as well as portals and various documents showing Web maps and data download pages, their commentary and usage, as well as free of charge. In order to bridge these flows this time, we started prototyping J - SHIS RDF and introduce our efforts.

## 2. Purpose

The "Basic Principles on Open Data" approved by the IT Strategic Headquarters in May 2017, decided the definition of open data as follows: (1) Data published under a rule allowing secondary use, whether or not for commercial purposes (2) Data published in a machine-readable format (3) Data that can be used free of charge. Furthermore, the form of public data advocated by Tim Berners-Lee, the founder of the WWW (World Wide Web) In principle, the principle was set so as to make efforts to disclose in terms and formats that are more easily accessible, with reference to the index of the 5 stages (commonly known as "5 stars"). The J-SHIS RDF is an attempt to further promotes the J-SHIS Web API (currently, 3.5 stars), which has been published in accordance with the definition of open data, to 5-star data as a model of open data. At the same time, seismic hazard information which is also professional data required for advanced domain knowledge is used without misunderstanding, while partially referring to and understanding the user's own part by retroactivity of RDF If it is possible.

## 3. Method

J-SHIS RDF tried to refine the utilization method by reexamining all data related to J-SHIS in accordance with the definition of RDF (W3C compliant) and describing its relation. Since we went through a three-step process to create a trial version, it is shown below.

Step 1. Setting the basic granularity: Four layers with fine granularity are set as going downward, and the top two are prototyped. In the second layer, the Inter-Map Relations Organizing Layer, the basic granularity was taken as "map" in order to aim at what makes the overview of the whole project possible by carrying out the element arrangement of J-SHIS seen from the stakeholders.

Step 2. Establishment policy: Since the SHM needs to accompany scientific evidence regardless of where it is cut; it is based on the report. The basic policy for creating such RDF is summarized as the creation policy of this layer.

Step 3. Fact check by expert: A graph of just right size is outputted as a prototype and we received a fact check by experts. Furthermore, we corrected and clarified by receiving feedback from experts and tried better visualization based on pointed out issues.

#### 4.Result

The second layer named J-SHIS Map Relationship RDF consists of 66 triples (Figure). In comparison with the first layer named HERP Subcommittee Relationship RDF was able to describe in 5 triples, it became a fairly detailed graph. This two-layer RDF is a graph that overlooks the relationship between nodes having a huge data group inside, making the flow of how to make J-SHIS (SHM) easier to follow.

#### 5.Issues and considerations

Understanding of the data deepened in the course of work. Although it should reflect the facts based on the report, it was noticed that the cognitive level of the creator was naturally reflected, so it is considered that the fact check by the expert in Step 3 is indispensable in every layer. In relation to this, it is difficult to define predicates in particular, and it seems to be necessary to collate with the vocabulary base or to have our own defined schema. Also, when creating and making the current state as Layers 1 and 2 as described above, it was predicted that it would be possible to describe an individual point of the API if roughly Layer 4 was made. The extent to which these efforts of RDF are useful for promoting understanding of uncertainty in utilization of SHM and J-SHIS and reducing misunderstanding are many unknown parts. Therefore, that further improvement showing examples of utilization through the creation of Layers 3 and 4 is required in the future.

Keywords: Seismic Hazard, J-SHIS, RDF

