

The geochemical database “DODAI” with high-precision location coordinates

*Satoru Haraguchi¹, Kenta Ueki¹, Kenta Yoshida¹, Tatsu Kuwatani^{1,3}, Mika Mohamed¹, Hikaru Iwamori^{1,2,4}

1. Japan Agency for Marine-Earth Science and Technology, 2. Earthquake Research Institute, the University of Tokyo, 3. PRESTO, 4. Tokyo Institute of Technology

We constructed the geochemical database “DODAI” which compiled data published in the domestic journals and bulletins, and published its details recently (Haraguchi et al., 2018, J. Geol. Soc. Japan, doi: 10.5575/geosoc.2018.0027). This database has converted the sampling position of individual geochemical data into the GPS-familiar location coordinates, i.e. latitude and longitude, by image-based geocoding and improved the positional accuracy drastically compared to the previous databases.

There are some international databases of geochemical data, for example: "GEOROC" mainly involving land samples, "PetDB" mainly involving samples from seafloor. Recent progress on the data-driven science shed a light on a large number of sample data with a multitude of variables. As for heterogeneity of the mantle compositions (e.g. Iwamori et al, 2018), the world-wide scale of the geochemical structure are discussed on the basis of such databases. However, considerable number of data in these international databases provided the location coordinates only as the representative position of the geologic map of the original articles, and thus the resolution of these location data often can represent only the order of degree unit. This precision may have no problem for the discussion of world-wide scale, while interpretation in the detailed spatial scale may not be possible. The precision of location data in the “DODAI” is mostly less than 1 km, some of which have < 10m of errors, based on the spatial resolution of the map provided in the original articles. This precision allows us to correlate sample positions to the correct geological units described in the digital geologic map such as the Seamless Geological Map (AIST). We also defined a standard data format based on the metadata available in the compiled papers covering such as geographical information, sampling, and analytical methods. In this presentation, we will show the concept and details of DODAI dataset, and discuss the potential of data-driven analysis in the geochemical study.

Keywords: geochemical database, geocoding, normalization of format