

Extending the InGrid Metadata Catalogue with INSPIRE-Specific Elements: Technical and Organisational Challenges

Franz Schenk and Fred Kruse

*Coordination Center PortalU at the Lower Saxony Ministry of Environment and
Climate Protection*

*Archivstraße 2, D-30169 Hannover
kst@portal.u.de*

The INSPIRE regulations have by now become national law in most of the member states of the European Union. In order to be fully compliant with these regulations, a metadata editor for the public administration has to deal with additional requirements. The kind and amount of efforts that are necessary in order to fulfil these requirements vary depending on the complexity of the data model, the amount of data that has to be adjusted, and the number of users that have to be involved in the following process of data migration. In this paper we describe the challenges that PortalU faced when preparing the *InGrid* catalogue for full conformity with INSPIRE.

The *InGrid* metadata catalogue is the successor of the Umweltdaten-Katalog (abbreviated as UDK). It is planned, developed, and financed in a combined effort by the environmental administrations of both the federal states and the federal government in Germany. More than 15 years of expert knowledge in metadata catalogue development have flown into this project. Today, it is used by 14 of the 16 federal states, by two institutions of the federal government, furthermore in several research projects. More than ten thousand metadata objects can be found in this catalogue. It is part of an information brokering system for environmental data, called PortalU, the portal for environmental information in Germany.

Obviously, the data model of the *InGrid* catalogue was in existence long before INSPIRE came into play. It is also easy to see that there are different paradigms regarding the question which kind of information a metadata catalogue should provide: INSPIRE regards environmental information as just one kind of geographical information, whereas *InGrid* sees the geographical aspect being just one aspect among many others worth knowing about environmental information. The difference in attitudes can be found in a difference in data models as well. In *InGrid*, the classes describing datasets (or databases in general) and services have a broader scope than those that are targeted by INSPIRE. Moreover, there are classes for metadata objects covering projects, documents, organisational units, and tasks in the *InGrid* catalogue.

With the advent of INSPIRE, it became necessary to allow for the capturing of additional metadata elements for selected catalogue records. These elements, however, are not appropriate for all kinds of records in those classes covering geospatial services and datasets. Geospatial services, for example, are just one subset of information objects in the class *Service/Application/Information System* in *InGrid*. Similarly, datasets are part of the class *Geoinformation/Maps*, which is not limited to electronic information but subsumes all kinds of maps or geographical data, e.g. historical maps. Because of

the broader scope of classes in *InGrid*, the mandatory metadata elements as demanded by INSPIRE cannot be applied to the full set of elements in the classes *Service/Application Information System* and *Geoinformation/Maps*.

Besides the limitations due to differences in the data model, there are also different attitudes towards the question whether data quality and specification conformity information should be maintained in the metadata catalogue. Undoubtedly it is very useful to gather information of that kind where it is available. Yet, for many information objects in *InGrid* this information is not at hand. Rather to the contrary, data is often being described on a more abstract level and information like data quality or the encoding of the data is simply not available. In the same way, the INSPIRE metadata information model for datasets assumes that the person providing the metadata information is also familiar with the described data objects. This, however, is not always the case in administrative processes where metadata maintenance and data provision are often carried out in different departments or even different institutions. Keeping in mind that more than 50 people use the *InGrid* catalogue to maintain thousands of metadata objects, these considerations make it easy to see that each additional mandatory metadata element added to an existing data model means a lot of work for a lot of people.

In *InGrid*, the following upgrade strategy for the data model has been chosen: The class *Geoinformation/Maps* has been extended by additional elements, both mandatory and conditional. Furthermore, a new class for geospatial services has been defined. This new class is based on the *InGrid* data model and integrates the requirements of INSPIRE. In this way, geospatial services are now treated differently from arbitrary information systems. The migration of existing metadata records into the newly defined class can be realised in two ways. Either users identify suitable objects and assign them to the new class, or objects become reassigned automatically based on a number of keywords that have been given to metadata objects beforehand. Both variants demand for the filling in of the new mandatory metadata elements after the migration process. As regards datasets in the class *Geoinformation/Maps*, users will have to check every single object and add the missing information to the new metadata elements.

It has to be emphasised that it is not necessarily sufficient to simply add a set of metadata elements to an existing catalogue data model in order to obtain INSPIRE compliancy. Although *InGrid* is already ISO-compliant, the conceptual differences concerning the world of discourse, as it has been described here, lead to a partial incompatibility of the data models. A large number of catalogue objects are indeed relevant with regard to INSPIRE, yet the categorisation of these objects does not coincide with the classification of objects as provided by INSPIRE. Therefore, the evolution of an existing metadata catalogue can become a time and work consuming process.