Metadata for Models Generated by openModeller

Agnei Silva
UNASP, Brazil
agnei.silva@hotmail.com

Cleverton Ferreira Borba
University of Sao Paulo -
USP, Brazil
UNASP - Brazil
cleverton.borba@usp.br

Pedro L. Pizzigatti Corrêa
University of Sao Paulo,
Brazil
pedro.correa@usp.br

Keywords: species distribution modeling; Dublin Core; interoperability; openModeller

1. Metadata for Models Tools

In the last years, the economic development has grown on a large scale and accelerating the destruction of the ecosystem process, increasing the demand for tools and methods to support decision making with regard to biodiversity conservation. According to Berendsohn et al. (2011), one of the most serious "bottlenecks" in the scientific workflows of biodiversity sciences is the need to integrate data from different sources, applications software, and services for analysis, visualization and publication.

The main reasons to use metadata patterns in modeling tools are: allow representation of clearer information, interoperate data between repositories, provide standardized structures, increase data accessibility (Dziekaniak, 2010), preserving information resources and documenting legal aspects of resources (Berendsohn et al., 2011). In this context we can explore the Dublin Core metadata, because this help us standardize the models, generated by species distribution modeling tools.

The Global Biodiversity Information Facility (GBIF) and Biological Collection Access Service (BioCASE) are examples of tools that make use of metadata can cite some of them: ABCD (Access to Biological Collection Data) and also the DwC (Darwin Core) metadata that is used to support information from the portals DNA Bank Network and the GeoCASE (Berendsohn et al., 2011). Now, EDIT Platform supports the export and import of data in the standards (ABCD, DwC and also in the SDD - structured Descriptive Data). Among other tools can also cite openModeller (Munoz et al., 2011), receiving information by GBIF and TAPIR/Darwin Core system, utilizing as metadata standards Darwin Core and ABCD.

2. Dublin Core Application for Models Generated by openModeller Tool

Among the existing tools, openModeller stands out with some advantage over other species distribution modeling tools because it allows different formats of data inputs for occurrence of species, environmental data and parameters for the algorithms, above all, different algorithms, simplifying thus to user/users group to reach your aim without needing to know different platforms and modeling tools. One of the problems of the other current tools of species distribution modeling is that they generate models with their standard independent and it cannot be used in other tools.

The need to use the metadata for models generated by OpenModeller tool, allows the data standardization to other platforms, producing data to be reused in the openModeller, and in future, in other tools.

This poster proposes the use of a Dublin Core metadata standard to present and make available the models generated by the species distribution modeling tool openModeller, in order to facilitate interoperability of the data generated by tool itself or other modelling tools.

3. Interoperability of Data Generated

Interoperability only happens, when a well-defined standard is implemented in the data that will be interoperable.

Using an ontology as a class and the Dublin Core metadata as standard, we can ensure that the export and import of the generated models also interoperate between any openModeller tool or any other tool that makes use of species distribution models. Models data will can be available on the Internet so any user will may have free access to this data to visualization or to any other task. Figure 1 clearly describes the idea described in this post:
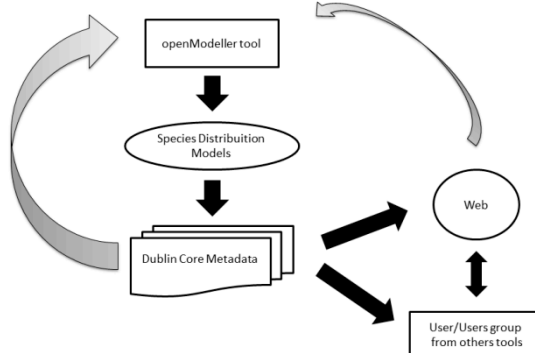

FIG. 1: Interoperability between openModeller tools, other tools and web, using standard Dublin Core Metadata

Therefore, when the openModeller tool generate a package with all information contained in a model, like, algorithm, parameters, occurrence data, etc., could be possible reuse this information to make a new modeling, entering other occurrence data, testing other algorithm and producing or formalizing results.

4. Conclusion

In conclusion, this poster presents a proposal for the reuse of models generated by the species distribution tool, openModeller. It is essential for the reuse of the model, use metadata pattern to ensure biodiversity data interoperability generated by this tools. In this poster we use the Dublin Core metadata for the initial stage of information that need to be reused.

Dublin Core metadata is an important domain to start the standardization of new tools particularly in data generated by species distribution tools that include: algorithm, parameters, climatic packages, biodiversity data, and the model.

Future Research: We suggest the use of new metadata patterns to make more studies case, and if possible, apply this model in other tools of species distribution modeling.

References

Berendsohn, W. G., Güntsch, A., Hoffmann, N., Kohlbecker, A., Luther, K. & Müller, A. (2011) Biodiversity information platforms: From standards to interoperability. ZooKeys, v. 150, p. 71-87.

Dziekaniak, Gisele Vasconcelos. (2010) Mapeamento do uso de padrões de metadados por comunidades científicas.

Muñoz, M.E.S., Giovanni, R., Siqueira, M.F., Sutton, T., Brewer, P., Scachetti, R.S., Canhos, D.A.L. & Canhos, V.P. (2011) "openModeller: a generic approach to species' potential distribution modelling". GeoInformatica.

Remsen, D., Ko, B., Chavan, B., Raymond, M. (2011) Getting Started, Overview of data publishing in the GBIF Network. GBIF.

Rodrigues, Fabrício Augusto. (2011) Modelagem da biodiversidade utilizando redes neurais artificiais. Doctoral Thesis presented at University of Sao Paulo.