Using Metadata for Interoperability of Species Distribution Models

Cleverton Ferreira Borba
University of Sao Paulo, Brazil
UNASP, Brazil
cleverton.borba@gmail.com

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

Keywords: Metadata, Species Distribution Modeling, Dublin Core, Darwin Core, architecture, interoperability.

1. Use of Metadata for Species Distribution Modeling (SDM)

This poster presents the use of metadata patterns for the field of Species Distribution Modeling and also presents a proposal for application of metadata to ensure interoperability between models generated by tools of Species Distribution Modeling (SDM).

According to Peterson et al. (2010) the area of Biodiversity Informatics is responsible, by the use of new technologies and computational tools, to meet the demand for support the biodiversity conservation. Portals of biodiversity, taxonomic databases, SDM tools help the scientists and researchers to decide the best for the biodiversity conservation. However, Berendsohn et al. (2011) says that one of the most serious problems in scientific biodiversity science is the need to integrate data from different sources, software applications and services for analysis, visualization and publication and thus offer an interoperability of data, information, application and tools.

In this context, the metadata patterns available, has been used to help the scientists and researchers to define vocabulary and data structure for analysis, visualization and publication of biodiversity data. Examples of metadata used in SDM are: Dublin Core (DC) (DCMI, 2012), Darwin Core (DwC) (Wieczorek *et al.*, 2012), Darwin Core Archive (DwC-A) (GBIF, 2010), EML – Ecological Metadata Language (Fegraus *et al.*, 2005), etc.

Biodiversity portals like GBIF (Global Biodiversity Information Facility - http://www.gbif.org/), ALA (Atlas of Living Australia – http://www.ala.org.au/), Specieslink (http://splink.cria.org.br/) also use metadata standards to support the data integrity, interoperability, and the data standardization. SDM tools, that use the data provided by this portals, to produce species distribution models, also support the metadata domain for their proposal.

Based on this information an application of metadata to ensure interoperability between models of SDM is presented below.

2. Application of Metadata for Ensure Interoperability between Models of Species Distribution Modeling

To support and ensure the interoperability between models generated by SDM tools, we propose the use of the DC and DwC metadata. The metadata information generate should have the minimum data for reuse in the same SDM tool or another one. The XML archive contain the biodiversity data used (occurrence points [presence/absence]), the algorithm and the parameter used for calculate the model, climatic package, and the model map provided for the SDM tool. The Figure 1 shows an example that how each metadata pattern contributes with the SDM domain in this study.

DCPAPERS

Proc. Int'l Conf. on Dublin Core and Metadata Applications 2015

Wieczorek *et al.* (2012) says that "an essential step towards understanding global patterns of biodiversity is to provide a standardized view of these heterogeneous data sources to improve interoperability", and that is the object to apply the metadata for Models of SDM.
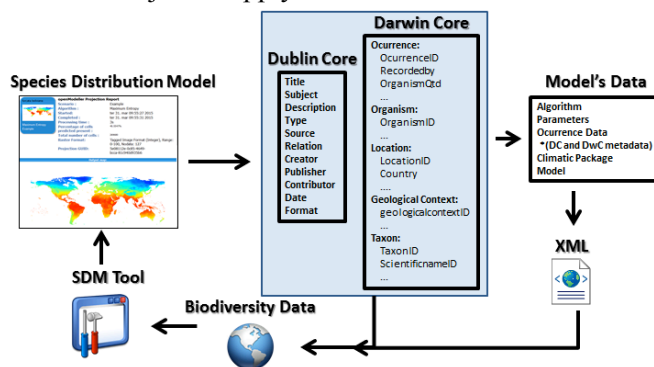

FIG. 1 - Use of Metadata for Species Distribution Modeling

The Figure 2 shows the proposed architecture for the SDM tools where the XML archive generated by a plug-in installed, has a metadata pattern to turn available the information used to make de model of species distribution (like algorithms, biodiversity data, parameters, climatic packages, etc.).
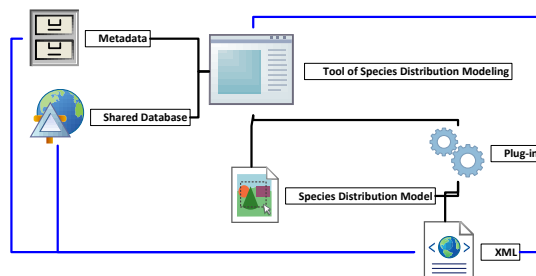

FIG. 2 - Proposed the use of metadata standards for species distribution modeling tools.

3. Conclusion

Through this research is possible to identify the importance of the metadata for the Biodiversity Informatics, specifically for Species Distribution Modeling. Just using a metadata is possible the interoperability between biodiversity data. In this study we proposed the use of metadata pattern for generate models of SDM tools through the development of a plug-in that presents a XML archive based in DC and DwC metadata to be reused, or offered in a portal of biodiversity.

As future work, we suggest the analysis with other metadata patterns and the use of the JSON archive for exportation of the model's data of species distribution.

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.

DCMI. (2012). Dublin Core Metadata Element Set, version 1.1: Reference description. Retrieved March 31, 2015, from http://dublincore.org/documents/dces/

Fegraus, E. H., Sandy, A., Matthew, J. B. & Mark, S. (2005) Maximizing the Value of Ecological Data with Structured Metadata: An Introduction to Ecological Metadata Language (EML) and Principles for Metadata Creation. Bulletin of the Ecological Society of America, p. 158-168.

GBIF. (2010) Darwin Core Archive – How-to Guide. Copenhagen: Global Biodiversity Information Facility.

DCPAPERS

Proc. Int'l Conf. on Dublin Core and Metadata Applications 2015

Peterson A. T., Knapp S., Guralnick R., Soberón J. & Holder M. (2010) Perspective: The big question for biodiversity informatics. Systematics and Biodiversity. The Natural History Museum. 8(2) 159-168.

Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., Robertson, T. & Vieglais, D. (2012). Darwin Core: An Envolving Community-Developed Biodiversity Data Standadrd. PLoS ONE 7.