

Conference Abstract

National Scientific Validation for Data in the Information System on Nature and Landscapes (Système d'Information sur la Nature et les Paysages - SINP)

Rémy Jomier[‡], Solène Robert[‡], Pascal Dupont[‡]

[‡] Natural History Museum, Paris, France

Corresponding author: Rémy Jomier (remy.jomier@mnhn.fr), Solène Robert (solene.robert@mnhn.fr)

Received: 16 Apr 2018 | Published: 22 May 2018

Citation: Jomier R, Robert S, Dupont P (2018) National Scientific Validation for Data in the Information System on Nature and Landscapes (Système d'Information sur la Nature et les Paysages - SINP) . Biodiversity Information Science and Standards 2: e25881. <https://doi.org/10.3897/biss.2.25881>

Abstract

As part of the Biodiversity Information System on Nature and Landscapes (Système d'Informations Nature et Paysages or SINP), the French National Natural History Museum has been appointed by the French ministry in charge of ecology to develop mechanisms for biodiversity data exchange, especially taxon occurrences (there are also elements on habitat occurrences, geo-heritage, etc.). Given that there are thousands of different sources for datasets, containing over 42 million records, such a development brings into question the underlying quality of data. To add complexity, there can be several layers of quality assurance: one by the producer of the data, one by a regional node, and another one by the national node.

The approach to quality issues was addressed by a dedicated working group, representative of biodiversity stakeholders in France. The resulting documents focus on core methodology elements that characterize a data quality process for, in the first instance, taxon occurrences only. It may be extended to habitats, geology, etc. in the near future.

For scientific validation, two processes are used:

- One automated process that uses expertise upstream (automated validation based on previous databases created through the use of said expertise), with several criteria such as comparison with a national taxonomic reference database (TAXREF), and with species reference distributions. The outcomes of this process will indicate error potential and can be used to automatically flag data above a certain threshold for the following process.
- A second, manual process, that allows for further scrutiny in order to reach a conclusive evaluation.

The combination of both processes allows experts to focus on data that has a higher likelihood of being erroneous, thus saving time and resources.

One objective of the INPN (Inventaire National du Patrimoine Naturel, or National Inventory of Natural Heritage), after one or both approaches, is to have each record assigned a confidence level.

The poster will be about National scientific validation of data in the SINP. It will show for whom and why it is done, whether the expertise lies upstream or downstream (manual validation through expert networks), what documents may exist, and what attributes have been considered to be added to the national standards so as to convey the information derived from these processes.

Keywords

Data quality, biodiversity data, validation

Presenting author

Rémy Jomier