

The 8 Years of Existence of Xper3: State of the art and future developments of the platform

Adeline Kerner[‡], Sylvain Bouquin[§], Rémy Portier^l, Régine Vignes Lebbe[§]

[‡] UMR 7207 – Centre de Recherche en Paléontologie - Paris CNRS – Sorbonne Université - MNHN, Paris, France

[§] Institut de Systématique, Evolution, Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles - France, Paris, France

^l UMS 3468 - Bases de données Biodiversité, Ecologie, Environnements Sociétés - MNHN-CNRS, Paris, France

Corresponding author: Adeline Kerner (kerner@mnhn.fr)

Abstract

The [Xper3](#) platform was launched in November 2013 (Saucède et al. 2020). Xper3 is a free web platform that manages descriptive data and provides interactive identification keys. It is a follow-up to Xper (Forget et al. 1986) and [Xper2](#) (Ung et al. 2010). Xper3 is used via web browsers. It offers a collaborative, multi-user interface without local installation. It is compatible with [TDWG's Structured Descriptive Data \(SDD\) format](#). Xper3 and its previous version, Xper2, have already been used for various taxonomic groups. In June 2021, 4743 users had created accounts and edited 5756 knowledge bases.

Each knowledge base is autonomous and can be published as a free access key link, as a data paper in publications or on websites. The risk of this autonomy and lack of visibility to already existing knowledge bases is possible duplicated content or overlapping effort. Increasingly, users have asked for a public overview of the existing content. A first version of a searching tool is now available online. [Explorer](#) lists the databases whose creators have filled in the extended metadata and have accepted the referencing. The user can search by language, taxonomic group, fossil or current, geography, habitat, and key words.

New developments of Xper3 are in progress. Some have a first version online, others are in production and the last ones are future projects. We will present an overview of the different projects in progress and for the future.

Calculated descriptors are a distinctive feature of Xper3 (Kerner and Vignes Lebbe 2019). These descriptors are automatically computed from other descriptors by using logical operators (Boolean operators). The use of calculated descriptors remains rare. It is necessary to put forward the calculated descriptors to encourage more feedback in order to improve them.

The link between Xper3 and [Annotate](#) continues to improve (Hays and Kerner 2020). Annotate offers the possibility of tagging images with controlled vocabularies structured in Xper3. Then, an export from Annotate to Xper3, allows automatic filling in of the

Xper3 knowledge base with the descriptions (annotations and numerical measures) of virtual specimens, and then comparing specimens to construct species descriptions, etc.

Future developments are in progress that will modify the Xper3 architecture in order to have the same functionalities in both local and online versions and to allow various user interfaces from the same knowledge bases. Xper2-specific features, such as merging states, adding notes, adding definitions and/or illustrations in the description tab, having different ways of sorting and filtering the descriptors during an identification (by groups, identification power, alphabetic order, specialist's choice) have to be added to Xper3.

A new tab in Xper3's interface is being implemented to give an access to various analysis tools, via API (Application Programming Interface), or R programming code:

- MINSET: minimum list of descriptors sufficient to discriminate all items
- MINDESCR: minimum set of descriptors to discriminate an item
- DESCRXP: generating a description in natural language
- MERGEMOD: proposing to merge states without loss of discriminating power
- DISTINXP, DISTVAXP: computing similarities between items or descriptors

One last project that we would like to implement is an interoperability between Xper3, platforms with biodiversity data (e.g., [Global Biodiversity Information Facility, GBIF](#)) and bio-ontologies. An ID field already exists to add Universally Unique IDentifiers (UUID) for taxa. ID fields have to be added for descriptors and states to link them with ontologies e.g., [Phenotypic Quality Ontology PATO](#), [Plant Ontology PO](#).

We are interested in discussing future developments to further improve the user interface and develop new tools for the analysis of knowledge bases.

Keywords

interactive key, computer-aided identification systems, knowledge base, collection, Annotate, Xper 2

Presenting author

Adeline Kerner

Presented at

TDWG 2021

Conflicts of interest

References

- Forget PM, Lebbe J, Puig H, Vignes R, Hideux M (1986) Microcomputer-aided identification: an application to trees from French Guiana. *Botanical Journal of the Linnean Society* 93 (2): 205-223. <https://doi.org/10.1111/j.1095-8339.1986.tb01021.x>
- Hays L, Kerner A (2020) Virtual Systematics with Annotate: The Mexican archaeocyaths example. *Biodiversity Information Science and Standards* 4 <https://doi.org/10.3897/biss.4.59106>
- Kerner A, Vignes Lebbe R (2019) Multi-context Knowledge Base using Calculated Descriptors from Xper3: the Archaeocyaths Knowledge Base example. *Biodiversity Information Science and Standards* 3 <https://doi.org/10.3897/biss.3.37083>
- Saucède T, Eléaume M, Jossart Q, Moreau C, Downey R, Bax N, Sands C, Mercado B, Gallut C, Vignes-Lebbe R (2020) Taxonomy 2.0: computer-aided identification tools to assist Antarctic biologists in the field and in the laboratory. *Antarctic Science* 33 (1): 39-51. <https://doi.org/10.1017/s0954102020000462>
- Ung V, Dubus G, Zaragueta-Bagils R, Vignes-Lebbe R (2010) Xper2: introducing e-taxonomy. *Bioinformatics* 26 (5): 703-704. <https://doi.org/10.1093/bioinformatics/btp715>