FRIDA 3.0
Multi-authored digital identification keys in the Web

Stefano Martellos

Abstract — FRIDA (FRiendly IDentificAtion) is a software for generating different types of digital identification keys starting from a database of characters. Its first release dates back to 2003. Since the first version, FRIDA was constantly developed to improve its core functions, and to add new features. At the beginning of 2010 it was decided to develop a new version, 3.0, written in PHP language for MySQL engines. The new version will have new and enhanced features, as well as the possibility to interact with other software developed in the framework of the European project *KeyToNature*. The development is currently in the beta-testing phase, and the final release is planned for the beginning of 2011.

Index Terms — biodiversity informatics, *Dryades*, *KeyToNature*.

◆

1 INTRODUCTION

The first approaches to digital identification are recent, dating back to the '70s of the last century, and especially to the beginning of the "explosion" of the World Wide Web, less than twenty years ago. Today there is a great and continuously increasing number of different digital identification keys, produced by several research centers: fixed- or free-pathway keys, with different querying systems, matrix keys, simple textual keys, etc. [1], [2]. They can be accessible on the Web, stored on CD- or DVD-ROMs, and some of them can run on mobile devices as PDAs and Smartphones [3], [4], [5], [6], [7], [8], [9], [10].

The production of identification keys, both "classic" and digital, to large groups of organisms (e.g. a national flora), normally requires the combined effort of several authors. In most of the classic, paper-printed keys, each author (or small groups of authors) develops one or few keys to families, species or genera, which are then connected in a hierarchical way by a "general key". This approach has been succesfully applied to digital keys as well, e.g. in the Flora of China project

S. Martellos is with the Department of Life Sciences, University of Trieste, I-34127, Italy. E-mail: martelst@units.it.

[10], in which a different dataset for each family or genus was developed by one or a few authors. A different approach is that of many authors working together on a common dataset, e.g. in the project LIAS for lichenized and lichenicolous fungi [11]. Another interesting experiment to develop a community approach in building multi-authored digital keys is the BioWikiFarm [12]. In this case, the keys are digital texts stored on a MediaWiki platform, and a potentially very large community of authors can edit them, while the system registers and keeps track of the changes.

FRIDA (FRiendly IdentificAtion) [13] was developed to allow several authors to work together, but rather independently from each other, while building a common database of morpho-anatomical data from which it is possible to produce a virtually unlimited number of different multi-authored digital keys. This software has been already used to generate hundreds of keys to plants, animals and fungi in the framework of the European project *KeyToNature*. FRIDA was available as a package running on Oracle databases only up to its last version (2.0). The new version, FRIDA 3.0, currently in beta stage of development, is written in PHP for MySQL databases, and has several new and improved features. Furthermore, it will be possible to use it both in stand-alone and on-line mode.

2 FRIDA

FRIDA (FRiendly IDentificAtion) has been developed since 2003 at the Department of Life Sciences of the University of Trieste (Italy), in the framework of project *Dryades* [7]. Up to its version 2.0 it was written in PL/SQL language, and developed on an Oracle Database engine.

The most interesting features of FRIDA [13] are:

1. *It does not require the learning of any code or programming language.* Input and management of data are in natural language, through simple Web interfaces written in HTML 4.0.
2. *Keys are immediately available on-line since their generation.* They are accessible from the Web by using any common web browser, through a single-access and a multi-entry query interface [8].
3. *Keys are independent from the original data.* When a key is produced, it does exist as a discrete entity, separated from the original database. In this way it is possible to modify the keys whithout affecting the original database, or vice-versa.
4. *The database of characters has a double-level architecture.* Characters are stored in two levels of information: a) a first level which is common to all taxa in the database, b) a second level, which is restricted to taxa belonging to a given "group" (see later). Organisms can be divided into more homogenous groups (e.g. genera and families, but also fully artifical groupings) by using several characters of the first level. These groups do exist as independent entities in the second level, and can be managed as independent databases by different authors (e.g. specialists of a given genus or family), which can thus work with a large degree of independence.
5. *The weight of characters in the generation of dichotomous keys is decided*

by the authors case by case. While it is possible to use an algorithm [14] to produce the "better" key (e.g. the key with the shortest branches), only an experienced taxonomist knows which is the weight of a given character in a particular group of organisms.

6. *Keys are portable in the field, both online and in stand-alone versions.* The latter, while less performing, are the better solution when internet connections are not available or poorly effective. Stand-alone keys can be also stored on CD- and DVD-ROMs.

3 NEW FEATURES IN VERSION 3.0

The development of FRIDA 3.0 started at the beginning of 2010. The new version is written in PHP language and runs on MySQL engines. It has the same core features of the previous version, plus some new and enhanced features:

1. *It can be also installed and used locally.* While the previous version required an Oracle Database and an Application Server to be installed and could be used on-line only, this version can run locally on any computer provided with MySql, PHP and Apache. It can be obviously installed on a server computer as well.

2. *Enhanced multilingual support.* While in the previous version it was possible to input the data into two languages only, in the new version the number of languages can be decided by the authors, and is virtually unlimited. The languages are selected in the settings interface, and new languages can be added at any time.

3. *New templates.* The development of the previous version focused almost exclusively on software performance, the interfaces for data management were rather poor and sometimes tricky for the users. In the new version, thanks to a new template, the functions are accessible in a far more rational way. For example, the management of characters can be done by using a few interfaces, which group all the character and states functions (Fig. 1), while the old version had a different interface for each function.

4. *Improved management of character images.* In the previous version the images of characters could be stored and reused, but the process was tricky, and sometimes forced the authors to create "dummy" characters with several fake states. In the new version character images can be stored in a image archive, and searched by file name or keyword.

5. *Easier management of records.* In FRIDA each taxon can be described by several records. Each record, while having the same name, and referring to the same taxon, differs for at least a character state. In the new version all the records referring to a taxon can be viewed together, and managed in a more efficient way (Fig. 2).

6. *Commented functions.* All the interfaces are provided with short textual information, to give the authors reminders on the use of the different functions. A full reference manual to FRIDA 3.0 is in preparation.

Fig. 1 – The new interface for the management of characters and values. Several functions are grouped together.
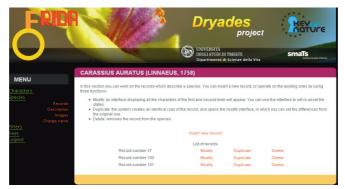

Fig. 2 – Each taxon can be described by several records, differing for at least a character state. The records are managed in a simple interface, which permits to edit, duplicate and delete them, as well as to add new records.

4 Conclusion

FRIDA 3.0 will be accessible to several research centers, including those without an Oracle system. It will permit to export the keys both in the Open Key Editor [15], [16], [17], in the Open Key Player [18] and in the BioWikiFarm [12] formats, which were developed in the framework of *KeyToNature* [19], to contribute to the development of integrated, open networks for digital identification.

The estimated roadmap for the future development of FRIDA 3.0 is:
- November, 2010 – FRIDA 3.0 Beta 2,
- December, 2010 – FRIDA 3.0 Beta 3,
- January, 2011 – FRIDA 3.0 Release Candidate (RC) 1,
- February, 2011 – FRIDA 3.0 RC 2,
- March, 2011 – FRIDA 3.0 - official release.

While the beta testing phase is closed, the RC versions will be available upon request to the author.

REFERENCES

[1] M. J. Dallwitz, T. A. Paine, and E. J. Zurcher, *Principles of interactive keys.* (http://delta-intkey.com), 2000 (onwards).

[2] M. J. Dallwitz, T. A. Paine and E. J. Zurcher, *Interactive identification using the Internet*. (http://delta-intkey.com), 2002 (onwards).

[3] G. Agarwal, H. Ling, D. Jacobs, S. Shirdhonkar, W. J. Kress, R. Russell, P. Belhumeur, A. Dixit, S. Feiner, D. Mahajan, K. Sunkavalli, R. Ramamoorthi and S. White, "First steps towards an electronic field guide for plants." *Taxon,* vol. 53 (3), pp. 597-610, 2006.

[4] A. R. Brach and H. Song, "ActKey: a Web-based interactive identification key program". *Taxon,* vol. 54 (4), pp. 1041-1046, 2005.

[5] D. F. Farr, "On line keys: more than just paper in the web". *Taxon,* vol. 53 (3), pp. 589-596, 2006.

[6] K. Chang-Sheng and H. Song, "Interactive key to Taiwan grasses using characters of leaf anatomy – the ActKey approach". *Taiwania,* vol. 50, pp. 261-71, 2005.

[7] S. Martellos and P. L. Nimis, "*KeyToNature*: Teaching and Learning Biodiversity. Dryades, the Italian Experience". In: M. Munõz, I. Jelínek, and F. Ferreira (eds.), *Proceedings of the IASK International Conference Teaching and Learning 2008*, pp. 863-868, 2008.

[8] R. D. Stevenson, W. A. Haber and R.A. Morris, "Electronic field guides and user community in the ecoinformatics revolution". *Conservation Ecology,* vol. 7(1), pp. 3, 2003

[9] M. J. Dallwitz, "*A comparison of interactive identification programs*". (http.//delta-intkey.com), 2000 (onwards).

[10] A. R. Brach and H. Song, "eFlora: New directions for online floras exemplified by the Flora of China Project". *Taxon,* vol. 55(1), pp. 188-92, 2006.

[11] G. Rambold, "LIAS – The concept of an identification system for lichenized and lichenicolous fungi". In: Anonymous (ed.), *The Third Symposium IAL 3. Progress and problems in Lichenology in the Nineties. Abstracts. - 9. Salzburg*, 1996.

[12] G. Hagedorn, G. Weber, A. Plank, M. Giurgiu, A. Homodi, C. Veja, G. Schmidt, P. Mihnev, M. Roujinov, D. Triebel, R. A. Morris, B. Zelazny, E. van Spronsen, P. Schalk, C. Kittl, R. Brandner, S. Martellos, and P. L. Nimis, "An online authoring and publishing platform for field guides and identification tools". In: P. L. Nimis and R. Vignes Lebbe (eds.)*, Tools for Identifying Biodiversity. Progress and Problems,* pp. 13-18, 2010.

[13] S. Martellos, "Multi-authored interactive identification keys: the FRIDA (FRiendly IDentificAtion) package". *Taxon*, vol. 59 (3), pp. 922-929, 2010.

[14] M.J. Dallwitz, T.A. Paine and E. J. Zurcher, "*User's guide to the DELTA System: a general system for processing taxonomic descriptions*". 4th edition. (http://delta-intkey.com), 1993 (onwards).

[15] S. Martellos, E. van Spronsen, D. Seijts, N. Torrescasana Aloy, P. Schalk and P. L. Nimis, "Digital identification keys to organisms and user-generated content. The *KeyToNature* approach". In: M. Muñoz and F. Ferreira (eds.), *Proceedings of the IASK International Conference Teaching and Learning 2009*, pp. 96-102, 2009.

[16] S. Martellos, E. van Spronsen, D. Seijts, N. Torrescasana Aloy, P. Schalk and P. L. Nimis, "User-generated content in the digital identification of organisms: the *KeyToNature* approach". *Int. J. Information and Operations Management Education*, vol. 3, 3, pp. 272-83, 2010.

[17] E. van Spronsen, S. Martellos, D. Seijts, P. Schalk and P. L. Nimis, "Modifiable digital identification keys". In: P. L. Nimis and R. Vignes Lebbe (eds.), *Tools for Identifying Biodiversity: Progress and Problems,* pp. 127-131, 2010.

[18] M. Giurgiu, A. Homodi, E. van Spronsen, S. Martellos and P. L. Nimis, "The Open Key Player:

A new approach for online interaction and user-tracking in identification keys". In: P. L. Nimis and R. Vignes Lebbe (eds.), *Tools for Identifying Biodiversity: Progress and Problems,* pp. 133-136, 2010.

[19] G. Hagedorn, P. L. Nimis and P. Schalk, "*KeyToNature*: Software, data formats, and communities". *Biodiversity Informatics Symposium 2008. The Book of Abstracts*, Swedish Museum of Natural History, Stockholm, Sweden, 27, 2008.