Tropical Diversity (2024) 4(1): 1-12.
ISSN: 2596-2388
DOI: 10.5281/zenodo.10863712

RESEARCH ARTICLE

Free DELTA: An open-source system for processing taxonomic descriptions

Free DELTA: Um sistema de código aberto para o processamento de descrições taxonômicas

Mauro J. Cavalcanti[1] * https://orcid.org/0000-0003-2389-1902, Edinaldo Nelson dos Santos-Silva[2] https://orcid.org/0000-0002-3340-4541, Luis Geraldes-Primeiro[2] https://orcid.org/0000-0002-3892-8969

[1]Ecoinformatics Studio, Caixa Postal 18123, CEP 20720-970 Rio de Janeiro, RJ, Brazil.

[2]Laboratório de Plâncton, Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia, Av. André Araújo, 2936, Aleixo, CEP 69060-000, Manaus, AM, Brazil.

*Email: maurobio@gmail.com

Received: 28, November 2023 / Accepted: 31, January 2024 / Published: 2, February 2024

Resumo O sistema DELTA (DEscription Language for TAxonomy) consiste de um formato flexível para a codificação de descrições taxonômicas e um conjunto de programas para o manejo e organização da informação taxonômica, como construção de chaves dicotômicas, produção de descrições em linguagem natural e identificação interativa. Todavia, o sistema DELTA original foi baseado em um modelo de software proprietário, que coloca limites à utilização e expansão do sistema. O Projeto Free DELTA é uma alternativa de software livre e código aberto para um sistema genérico de processamento de descrições taxonômicas, que proporcione uma ferramenta de suporte acessível para inventários sistemáticos de biodiversidade.

Palavras-Chave: identificação biológica, chaves taaxonômicas, informática para biodiversidade, formato DELTA

Abstract The DELTA (DEscription Language for TAxonomy) system consists in a flexible format for coding taxonomic descriptions and a set of programs to handling and organizing taxonomic information, as construction of dichotomous keys, production of natural language descriptions, and interactive identification. However, the original DELTA system was based on a proprietary software model, that puts limits to the use and expansion of the system. The Free DELTA Project is a free, open-source, software alternative for a generic system to the processing of taxonomic descriptions, which offers an accessible tool to support systematic biodiversity inventories.

Keywords: biological identification, taxonomic keys, biodiversity informatics, DELTA format

Introduction

A critical issue for the entire world is decreasing biological diversity, with detrimental impacts for food production, pharmaceuticals research, and the climate, which directly impair economic systems and human societies (Wilson, 1992). A better understanding of biodiversity is needed in order to decrease its current rate of loss and then revert it in the future. The first fundamental step is to know what exists and how to identify it, what is of primary concern to the science of biological systematics (Forey et al., 1994; Vane-Wright, 1996). As biodiversity losses increase all over the world, the need to increase the efficiency of systematists and the usefulness of taxonomic data, with better means of organizing them (Watson, 1971) has also become evident. The application of information technology to the storage, retrieval, and analysis of systematic data provides the primary means of achieving this (Edwards & Morse, 1995).

The DELTA format (Dallwitz, 1980; Dallwitz & Paine, 1992) is the most comprehensive format developed for the electronic storage and retrieval of taxonomic descriptive data which aimed to be universally accessible to the biological systematics research community. Software packages based on the DELTA format, in turn, provide a flexible and powerful set of tools for efficiently handling complex and variable taxonomic descriptive data and generating essential taxonomic products like dichotomous keys and natural-language descriptions (Watson et al., 1988; Dallwitz, 1993; Dallwitz, Paine, and Zurcher, 1993; Dallwitz, Paine, and Zurcher, 1999; Askevold & O'Brien, 1994; Coleman et al., 2010). The DELTA system has been demonstrated to be a valuable tool for biodiversity surveys, where the precise and efficient identification of organisms is of major importance (Ellis, 1988; Allkin et al. 1992; Jarvie & Stevens, 1998; Sharkey, 2001).

However, a major problem hampering the more widespread use of the DELTA system is that much of the existing software for handling data in DELTA format is closed source, available only in binary executable form, and so not amenable to be modified, upgraded, or ported to different operating systems, for any parties outside the original development team. This invariably leads to the end of support for such software, when funding ceases and the development team disband. In this paper, an open source, cross-platform software package for handling taxonomic descriptive data based on the DELTA format is presented and described. This package is developed after the GNU Project and the Free Software Movement (Stallman, 1999), and therefore assures the user community complete freedom not only to run the package, but to modify it to suit any special needs.

The DELTA System

DELTA stands for *DEscription Language for TAxonomy*. The DELTA system consists of a flexible format for coding descriptive taxonomic information and a set of associated programs which manipulate the data to produce natural-language descriptions and keys, interactive identification and information retrieval, and to convert data into formats required by programs for phenetic and cladistic analysis.

The DELTA format allows all kinds of characters, both qualitative (binary or multi-state, ordered or unordered) or quantitative (integer or real, with units if needed). Comments are allowed anywhere, and character dependency can be described. Directives are used to control computer processing of DELTA-coded data.

Although other generic formats for the representation of biological descriptive data have been proposed (White & Allkin, 1992; Maddison et al., 1997), the DELTA format was especially designed to provide a generalized system capable of handling all the different kinds of data used by taxonomists, as well as for facilities to directly generate taxonomic products as dichotomous keys and natural-language descriptions, and translate data into many external formats required by programs for classification and interactive identification (Dallwitz, 1993).

The original DELTA system has been under development since the mid 1970's by Mike Dallwitz, Toni Paine, and Eric Zurcher at CSIRO Division of Entomology, Canberra, Australia (Dallwitz, 1980). Similar packages for processing of DELTA data were also developed in the United Kingdom by Richard Pankhurst of the Royal Botanic Gardens at Edinburgh (Pankhurst, 1986), in The Netherlands by Eric Gouda at the Botanic Gardens of Utrecht University (Gouda, 1996), in Australia by Nicholas Lander at the Western Australian Herbarium (Lander, 1993), in Spain by Antonio Valdecasas at the Museo Nacional de Ciencias Naturales (Bello et al. 1996), and in Brazil by Mauro J. Cavalcanti at Museu Nacional/Universidade Federal do Rio de Janeiro (Cavalcanti, 1996). All these packages are closed source, proprietary software, available only in binary executables and restricted to a single operating system. Later on, DELTA programs based on the open source model were developed in Germany by Gregor Hagedorn at the Institute of Microbiology, Federal Biological Research Center (Hagedorn 2001), in the United States by Michael Bartley and Noel Cross at the Arnold Arboretum of the Harvard University and then in Germany by Dieter Neubacher and Gerhard Rambold at Mycology Department, University of Bayreuth (Neubacher & Rambold, 2005), in Italy by Claudio Rivetti and Riccardo Percudani at University of Parma (Rivetti & Percudani, 1999), and in Australia by the Atlas of Living Australia (2014).

In 1988, DELTA was adopted by the International Working Group on Taxonomic Databases for Plant Sciences (TDWG) as an internationally recognized standard for data encoding and exchange. Combining traditional taxonomic practices with modern information technology, the DELTA system has since became one of the most valuable and widely used computerized tools for biodiversity research.

Many datasets have already been prepared with DELTA, for flowering plant families, grasses, sedges, legumes, beetles (adults and larvae), ants, corals, plant viruses, etc. Several of these datasets incorporate line, gray-scale and color illustrations, and some are freely available on the Internet; at the time of writing, there are 61 datasets in the DELTA website (https://www.delta-intkey.com/www/data.htm) and 28 datasets in the NaviKey website (http://www.navikey.net/webprojects.html).

The Free DELTA Project

The Free DELTA project was launched in April, 2000 with the aim of creating a complete cross-platform, free, open-source software system for processing taxonomic descriptions coded in DELTA format, following the decision by CSIRO Division of Entomology to stop funding the development of the original DELTA programs.

The Free DELTA package was designed from to start to include a library for parsing DELTA-format files plus all the utilities needed to process taxonomic data coded in DELTA: format-conversion, key-construction, distance-matrix generation, and interactive identification programs, as well as a specialized editor for handling taxonomic descriptive data. However, although the Free DELTA programs should be able to process regular DELTA-coded files, they would not be identical to the original DELTA system. Free DELTA software would incorporate several improvements, based on our experience with other taxonomic computing systems. In particular, it should offer a better treatment of quantitative data (with automatic breakdown of real numeric characters into classes and computing of descriptive statistics for each), as well as more analytical and format-conversion procedures.

The Free DELTA parser libraries should provide a standard, versatile and extensible collection of routines for handling DELTA datasets. This set of routines is currently available in C++, Python, and Object Pascal programming languages and provide the basic support for the Free DELTA software; furthermore, they are ready for incorporation within individual application programs or database systems as they are created by different software developers. These routines are released as complete, open source code. So, a user who needs a different system for taxonomic data management is able to develop it himself, or hire any available programmer or company to develop it for him or her, while retaining the benefits of a standard format for encoding taxonomic data. The availability of full source code also enables a user to tailor these routines according to his or her more specific needs. It is hoped that this will help to further expand the DELTA user community, as well as to strengthen the role of DELTA as a standard taxonomic data coding and transfer system.

In July, 2000 the first DELTA parser library, written in C++ by Denis Ziegler, with additional contributions by Guillaume Rousse and Bastiaan Wakkie, has been developed for the Free DELTA Project.

In May, 2005 the Free DELTA project was registered at SourceForge, the world's largest open-source software development website, with the largest repository of open-source code and applications available on the Internet, and where the Free DELTA website is now to be located on a permanent basis.

In February, 2010 the PyDELTA parser library, written in Python by the Mauro Cavalcanti and Thomas Kluyver was released as yet another contribution to the software infrastructure of the Free DELTA Project; this library would later be used as the basis for the development of the first prototype of the Free DELTA Editor.

In February, 2013 the former CSIRO DELTA Development Team released the

FORTRAN source code of the main "classic" DELTA programs (CONFOR, KEY, and DIST), which were then incorporated into the core of the Free DELTA Project; the source code of these programs was not modified by the Free DELTA project, and new executable files were not generated.

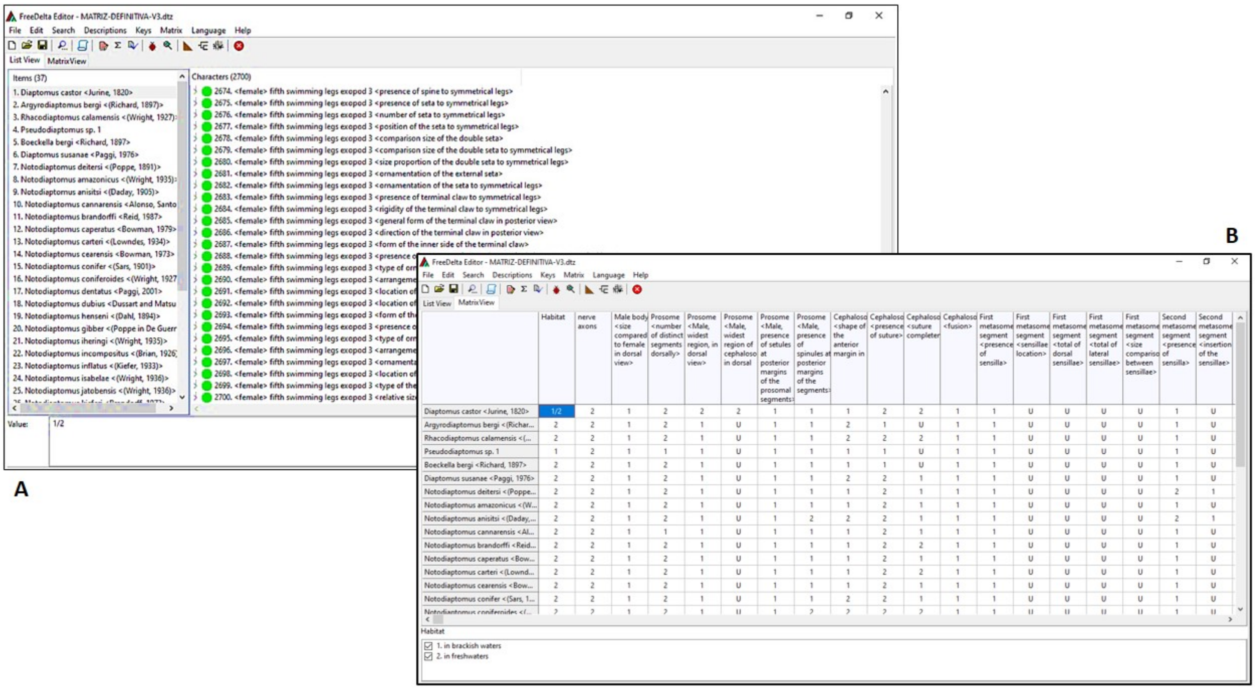
In July, 2016 a prototype of the Free DELTA Editor was developed as a cross-platform desktop application written in the Python programming language and using the PyDELTA parser library. This tool was released on an experimental basis in December, 2017.

Finally, in January, 2020 a production version of the Free DELTA Editor was released to the public, That is a much-improved version of the experimental software previously developed, entirely re-implemented in FreePascal/Lazarus as a full cross-platform desktop application.

The Free Delta Editor

The Free Delta Editor (FDE) is a cross-platform software package for editing DELTA data files and running DELTA based programs (CONFOR, KEY, DIST, INTKEY).

This tool provides an interactive, user-friendly interface to DELTA-format files as a real DELTA data entry/maintenance program. With it, it is possible to insert or delete items, characters, and states, without having to worry about the consistency of the DELTA files. A database-like interface makes it easy to setup and maintain the main DELTA data files (ITEMS, CHARS and SPECS), as well as several of the DELTA directives files (TONAT, TOKEY, TODIS). A spreadsheet-like visualization facility for item and character data is also available (Figure 1).



Items (37)

1. Diaptomus castor <Jurine, 1820>
2. Argyrodiaptomus bergi <(Richard, 1897)>
3. Rhacodiaptomus calamensis <(Wright, 1927)
4. Pseudodiaptomus sp. 1
5. Boeckella bergi <Richard, 1897>
6. Diaptomus susanae <Paggi, 1976>
7. Notodiaptomus deitersi <(Poppe, 1891)>
8. Notodiaptomus amazonicus <(Wright, 1935)
9. Notodiaptomus anisitsi <(Daday, 1905)>
10. Notodiaptomus cannarensis <Alonso, Santo
11. Notodiaptomus brandorffi <Reid, 1987>
12. Notodiaptomus caperatus <Bowman, 1979>
13. Notodiaptomus carteri <(Lowndes, 1934)>
14. Notodiaptomus cearensis <Bowman, 1973>
15. Notodiaptomus conifer <(Sars, 1901)>
16. Notodiaptomus coniferoides <(Wright, 1927
17. Notodiaptomus dentatus <Paggi, 2001>
18. Notodiaptomus dubius <Dussart and Matsu
19. Notodiaptomus henseni <(Dahl, 1894)>
20. Notodiaptomus gibber <(Poppe in De Guerr
21. Notodiaptomus iheringi <(Wright, 1935)>
22. Notodiaptomus incompositus <(Brian, 1926
23. Notodiaptomus inflatus <(Kiefer, 1933)>
24. Notodiaptomus isabelae <(Wright, 1936)>
25. Notodiaptomus jatobensis <(Wright, 1936)>

Characters (2700)

2674. <female> fifth swimming legs exopod 3 <presence of spine to symmetrical legs>
2675. <female> fifth swimming legs exopod 3 <presence of seta to symmetrical legs>
2676. <female> fifth swimming legs exopod 3 <number of seta to symmetrical legs>
2677. <female> fifth swimming legs exopod 3 <position of the seta to symmetrical legs>
2678. <female> fifth swimming legs exopod 3 <comparison size of the double seta>
2679. <female> fifth swimming legs exopod 3 <comparison size of the double seta to symmetrical legs>
2680. <female> fifth swimming legs exopod 3 <size proportion of the double seta to symmetrical legs>
2681. <female> fifth swimming legs exopod 3 <ornamentation of the external seta>
2682. <female> fifth swimming legs exopod 3 <ornamentation of the seta to symmetrical legs>
2683. <female> fifth swimming legs exopod 3 <presence of terminal claw to symmetrical legs>
2684. <female> fifth swimming legs exopod 3 <rigidity of the terminal claw to symmetrical legs>
2685. <female> fifth swimming legs exopod 3 <general form of the terminal claw in posterior view>
2686. <female> fifth swimming legs exopod 3 <direction of the terminal claw in posterior view>
2687. <female> fifth swimming legs exopod 3 <form of the inner side of the terminal claw>
2688. <female> fifth swimming legs exopod 3 <presence o
2689. <female> fifth swimming legs exopod 3 <type of orn
2690. <female> fifth swimming legs exopod 3 <arrangeme
2691. <female> fifth swimming legs exopod 3 <location of
2692. <female> fifth swimming legs exopod 3 <location of
2693. <female> fifth swimming legs exopod 3 <form of th
2694. <female> fifth swimming legs exopod 3 <presence o
2695. <female> fifth swimming legs exopod 3 <type of orn
2696. <female> fifth swimming legs exopod 3 <arrangeme
2697. <female> fifth swimming legs exopod 3 <ornamenta
2698. <female> fifth swimming legs exopod 3 <location of
2699. <female> fifth swimming legs exopod 3 <type of the
2700. <female> fifth swimming legs exopod 3 <relative siz

Value: 1/2

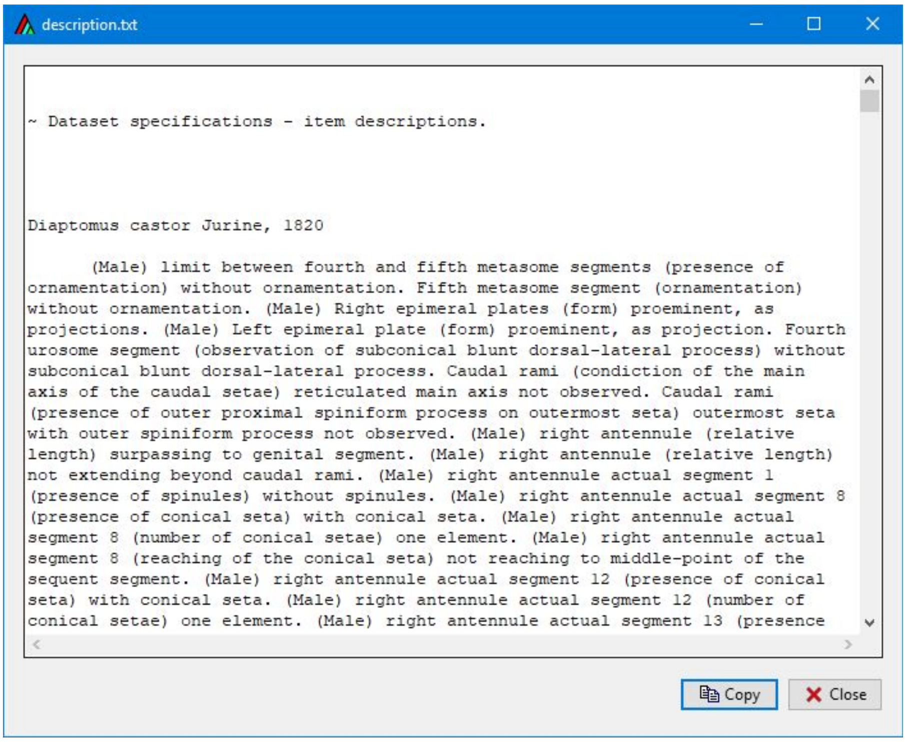
	Habitat	nerve axons	Male body <size compared to female in dorsal view>	Prosome <number of distinct segments dorsally>	Prosome <Male, widest region, in dorsal view>	Prosome <Male, widest region of cephaloso in dorsal	Prosome <Male, presence of setules at posterior margins of the prosomal segments>	Prosome <Male, presence of spinules at posterior margins of the segments>	Cephaloso <shape of the anterior margin in	Cephaloso <presence of suture>	Cephaloso <suture complete	Cephaloso <fusion>	First metasome segment <presence of sensilla>	First metasome segment <sensillae location>	First metasome segment <total of dorsal sensillae>	First metasome segment <total of lateral sensillae>	First metasome segment <size compariso between sensillae>	Second metasome segment <presence of sensilla>	Second metasome segment <insertion of the sensillae>
Diaptomus castor <Jurine, 1820>	1/2	2	1	2	2	2	1	1	1	2	2	1	1	U	U	U	U	1	U
Argyrodiaptomus bergi <(Richar...	2	2	1	2	1	U	1	1	2	1	U	1	1	U	U	U	U	1	U
Rhacodiaptomus calamensis <(...	2	2	1	2	1	U	1	1	2	2	2	1	1	U	U	U	U	1	U
Pseudodiaptomus sp. 1	1	2	1	1	1	U	1	1	1	1	U	1	1	U	U	U	U	1	U
Boeckella bergi <Richard, 1897>	2	2	1	2	1	U	1	1	1	1	U	1	1	U	U	U	U	1	U
Diaptomus susanae <Paggi, 1976>	2	2	1	2	1	U	1	1	2	2	1	1	1	U	U	U	U	1	U
Notodiaptomus deitersi <(Poppe...	2	2	1	2	1	U	1	1	1	2	1	1	1	U	U	U	U	2	1
Notodiaptomus amazonicus <(W...	2	2	1	2	1	U	1	1	1	2	1	1	1	U	U	U	U	1	U
Notodiaptomus anisitsi <(Daday,...	2	2	1	2	1	U	1	2	2	2	1	1	1	U	U	U	U	2	1
Notodiaptomus cannarensis <Al...	2	2	1	1	1	U	1	1	1	2	1	1	1	U	U	U	U	1	U
Notodiaptomus brandorffi <Reid...	2	2	1	2	1	U	1	1	1	2	2	1	1	U	U	U	U	1	U
Notodiaptomus caperatus <Bow...	2	2	1	2	1	U	1	1	1	2	1	1	1	U	U	U	U	1	U
Notodiaptomus carteri <(Lownd...	2	2	1	2	1	U	1	1	1	2	2	1	1	U	U	U	U	1	U
Notodiaptomus cearensis <Bow...	2	2	1	2	1	U	1	1	1	2	1	1	1	U	U	U	U	1	U
Notodiaptomus conifer <(Sars, 1...	2	2	1	2	1	U	1	1	2	2	1	1	1	U	U	U	U	1	U
Notodiaptomus coniferoides <(...	2	2	1	2	1	U	1	2	2	2	2	1	1	U	U	U	U	1	U

Habitat

☑ 1. in brackish waters
☑ 2. in freshwaters

Figure 1 - The Free Delta Editor database viewer (A) and data matrix viewer (B).

It also makes possible to start all DELTA programs and create dichotomous keys for identification, natural-language descriptions, statistical reports, distance matrices, cluster and ordination analysis, providing a viewer to look at and print the results (output files) (Figures 3 and 4).



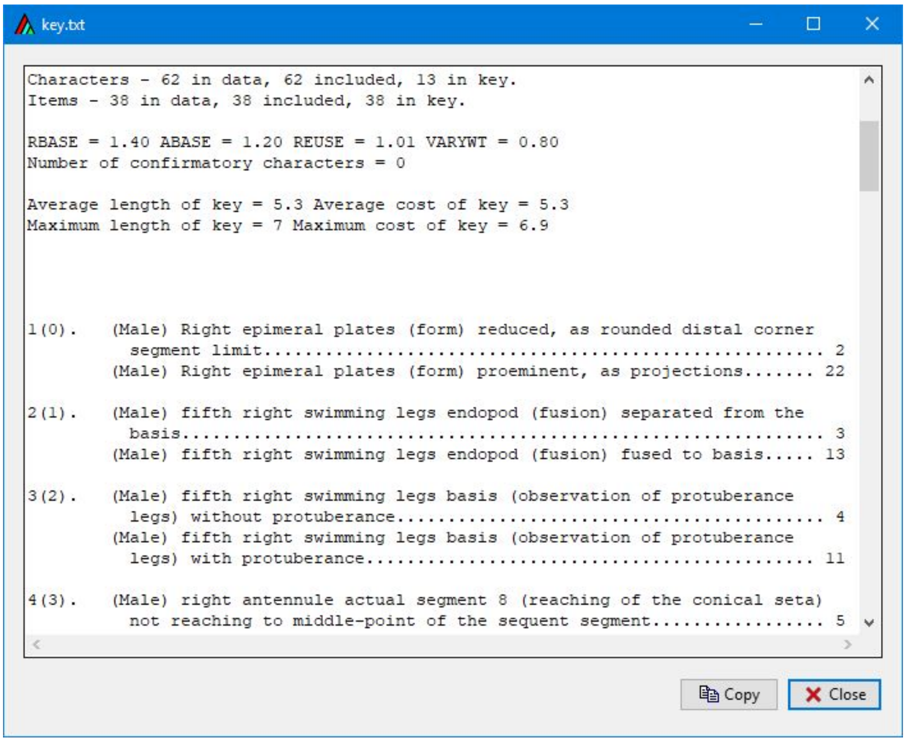
```
~ Dataset specifications - item descriptions.



Diaptomus castor Jurine, 1820

      (Male) limit between fourth and fifth metasome segments (presence of
ornamentation) without ornamentation. Fifth metasome segment (ornamentation)
without ornamentation. (Male) Right epimeral plates (form) proeminent, as
projections. (Male) Left epimeral plate (form) proeminent, as projection. Fourth
urosome segment (observation of subconical blunt dorsal-lateral process) without
subconical blunt dorsal-lateral process. Caudal rami (condition of the main
axis of the caudal setae) reticulated main axis not observed. Caudal rami
(presence of outer proximal spiniform process on outermost seta) outermost seta
with outer spiniform process not observed. (Male) right antennule (relative
length) surpassing to genital segment. (Male) right antennule (relative length)
not extending beyond caudal rami. (Male) right antennule actual segment 1
(presence of spinules) without spinules. (Male) right antennule actual segment 8
(presence of conical seta) with conical seta. (Male) right antennule actual
segment 8 (number of conical setae) one element. (Male) right antennule actual
segment 8 (reaching of the conical seta) not reaching to middle-point of the
sequent segment. (Male) right antennule actual segment 12 (presence of conical
seta) with conical seta. (Male) right antennule actual segment 12 (number of
conical setae) one element. (Male) right antennule actual segment 13 (presence
```

Figure 3 – Natural-language description viewer.



```
Characters - 62 in data, 62 included, 13 in key.
Items - 38 in data, 38 included, 38 in key.

RBASE = 1.40 ABASE = 1.20 REUSE = 1.01 VARYWT = 0.80
Number of confirmatory characters = 0

Average length of key = 5.3 Average cost of key = 5.3
Maximum length of key = 7 Maximum cost of key = 6.9




1(0).   (Male) Right epimeral plates (form) reduced, as rounded distal corner
          segment limit....................................................... 2
        (Male) Right epimeral plates (form) proeminent, as projections....... 22

2(1).   (Male) fifth right swimming legs endopod (fusion) separated from the
          basis............................................................... 3
        (Male) fifth right swimming legs endopod (fusion) fused to basis..... 13

3(2).   (Male) fifth right swimming legs basis (observation of protuberance
          legs) without protuberance.......................................... 4
        (Male) fifth right swimming legs basis (observation of protuberance
          legs) with protuberance............................................ 11

4(3).   (Male) right antennule actual segment 8 (reaching of the conical seta)
          not reaching to middle-point of the sequent segment................. 5
```

Figura 4 - Dichotomous key viewer.

The Free Delta Editor is written in FreePascal/Lazarus, a free, cross-platform object-oriented programming language for Rapid Application Development. Therefore, the program is platform-independent, and can run without modifications on any PC compatible with the x86 architecture, under GNU/Linux, MacOSX, and MS-Windows.

Its user interface resembles, but is not a clone of, that of the CSIRO and ALA DELTA Editors. A major difference is that in the Free Delta Editor all views are read-only and there is only one way of describing a taxon; the ‘Items’ and ‘Characters’ panes are read-only, and attributes may be edited only in the ‘Value’ pane This is an attempt to make the process of creating a descriptive database more orderly and less confusing than it seems to be in the case of the CSIRO/ALA Editors; these editors also have a ‘Values’ pane, but, in addition, attributes can be edited in the ‘Characters’ pane, which means that there is more than a way of editing the character attributes, therefore making the process prone to confusion for the inexperienced user.

Another important difference concerns the data file format: the Free Delta Editor uses conventional zip-compressed files (with the .dtz extension) to store all DELTA-formatted data. At a minimum, these files must include the three basic DELTA files (CHARS, ITEMS, and SPECS) and can be created either by exporting an already existing dataset from the CSIRO Delta or ALA Delta programs and compressing them into a zip file (changing its extension to .dtz), or by using the Free Delta Environment Import option from the File menu.

The Free Delta Editor provides for clustering and ordination (using principal coordinates) of the distance matrix computed by DIST, using for that the powerful R free software environment for statistical computing and graphics (Ihaka & Gentleman, 1996). Therefore, R must be installed on a users machine in order for performing cluster and principal coordinates analyses with the Free Delta Editor (Figure 5).

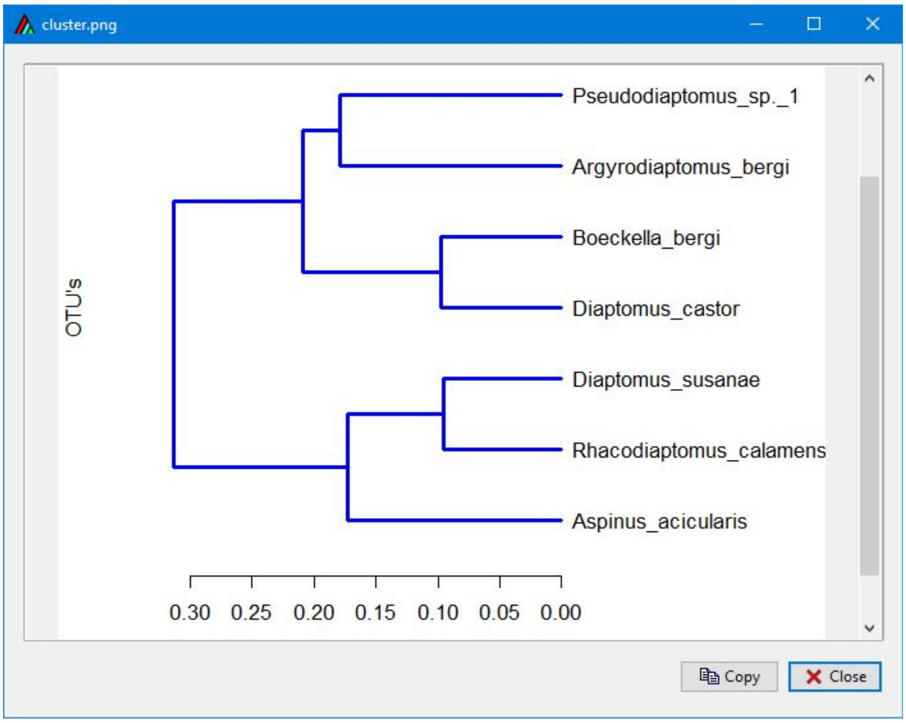


Figure 5 – Cluster analysis viewer.

The Free Delta Editor is initially available in English, French, Spanish, and Brazilian Portuguese, but any interested person can provide a translation for her/his native language. No knowledge of programming is required, just a freely available and easy to use free software tools like POEdit or Virtaal.

The current emphasis of the Free Delta Editor is on the production of natural language descriptions, dichotomous keys, and phenetic analyses. Support for interactive identification is quite limited and restricted to the MS-Windows version, which simply calls the CSIRO DELTA Intkey program, if installed. Support for Intkey is not available in the GNU/Linux version. The primary goal of the Free Delta Project in the near future is the development of a new, open source, cross platform interactive identification program (‘Freekey’), similar, but not identical to, the current CSIRO DELTA Intkey. This will be implemented as a separate application, which will use the same data files created and edited using the Free Delta Editor (therefore no 'format conversion' will be necessary).

Currently, binary executables are provided for MS-Windows (XP, 7, 8, 10) and GNU/Linux systems (Debian); a Mac OS version is not available; however, since the development tool now used for the implementation of the Free Delta software (FreePascal/Lazarus) is fully cross-platform, it should be fairly simple to compile a version for Mac's.

All Free DELTA software is released as free software, under the GNU General Public License (Stallman, 1999).

The Free Delta Editor applied to biodiversity research

In a systematic study on Neotropical microcrustaceans, the Free Delta Editor was applied to create a large database of 2700 characters and 37 species in the copepod genus *Notodiaptomus* Kiefer, 1936(Copepoda, Calanoida, Diaptomidae). The ability to store and process this volume of taxonomic data has made the Free Delta Editor a unique multiplatform software package among DELTA-based programs.

Using the DELTA programs directives (commands), it was possible to obtain automatic description in natural language, dichotomous identification keys for males and females, an assisted interactive key with images, annotations, and synonyms (Figure 6), as well as a character matrix for the analysis of phylogenetic relationships using the PAUP program (Swofford, 2003). Undoubtedly, the possibility of managing these types of data, in the scale and complexity, was decisive for the systematic study of such a morphologically diverse group of organisms as microcrustaceans (see Geraldes-Primeiro, 2023).

Indeed, the versatility of character matrices generated by DELTA-based programs extends to their compatibility with various phylogenetic inference software, such as PAUP (Swofford, 2003) and TNT (Goloboff et al., 2008). This allows researchers to seamlessly integrate their taxonomic data into different analytical tools, enabling a broader spectrum of analyses and enhancing the utility of the generated datasets.

Best Characters (50)

[male] right antennule actual segment 13 [range of the modified seta]
[male] Right epimeral plates [form]
[male] right antennule actual segment 12 [size comparison of the conical seta to segment 8]
[male] fifth left swimming legs exopod 2 [reach of the seta]
[male] Epimeral plates [symmetry]
[male] right antennule actual segment 8 [reaching of the conical seta]
[male] fifth right swimming legs basis [observation of protuberance legs]
[male] fifth right swimming legs exopod 2 [presence of chitinous knobs]
[male] right antennule [relative length]
[male] limit between fourth and fifth metasome segments [presence of ornamentation]
[male] fifth right swimming legs exopod 2 [Spine size condition smaller than insertion segment]
[male] fifth right swimming legs coxa [reach of the projection of the process]
[male] fifth right swimming legs exopod 2 [insertion of spine legs]
[male] fifth right swimming legs exopod 1 [observation of internal prominence]
[male] fifth right swimming legs exopod 1 [size relation of the segment]
[male] fifth right swimming legs exopod 2 [Relative size of spine smaller beyond to 2 times]
[male] Fourth metasome segment [fusion or separation indicated by suture presence, absence or partially presence]
Left antennule actual segment 11 [number of seta]
[male] fifth right swimming legs endopod [segmentation]
[male] fifth right swimming legs exopod 2 [presence of inner-posterior process]
[male] right antennule [relative length]
[male] fifth right swimming legs coxa [projection of the process]

Used Characters (0)

Remaining Taxa (32)

Notodiaptomus deitersi
Notodiaptomus amazonicus
Notodiaptomus anisitsi
Notodiaptomus brandorffi
Notodiaptomus cannarensis
Notodiaptomus caperatus
Notodiaptomus carteri
Notodiaptomus cearensis
Notodiaptomus conifer
Notodiaptomus coniferoides
Notodiaptomus dentatus
Notodiaptomus dubius
Notodiaptomus gibber
Notodiaptomus henseni
Notodiaptomus iheringi
Notodiaptomus incompositus
Notodiaptomus inflatus
Notodiaptomus isabelae
Notodiaptomus jatobensis
Notodiaptomus kieferi
Notodiaptomus maracaibensis
Notodiaptomus nelsoni

Eliminated Taxa (0)

[male] right antennule [relative length]

extending beyond caudal rami
not extending beyond caudal rami

Remaining Taxa (2)

Notodiaptomus henseni
Notodiaptomus transitans

Used Characters (3)

fifth right swimming legs exopod 2 inner-posterior process observed
fifth right swimming legs exopod 2 subtriangular
right antennule actual segment 12 not smaller than to segment 8

Eliminated Taxa (30)

[1] Notodiaptomus deitersi
[1] Notodiaptomus amazonicus
[1] Notodiaptomus brandorffi
[1] Notodiaptomus cannarensis
[1] Notodiaptomus caperatus

Figure 6 – DELTA Intkey program running the interactive key for the species of *Notodiaptomus*. A. Initial screen with the characters used (left) and taxa for identification (right); B. Secondary screen with the selected character (top left) and remaining taxa (top right); C. Screen illustrating the states observed for the selected character.

Conclusion

Recent advances in the application of computer technology to taxonomy offer the prospect of fast, large-scale character extraction and automated identification of organisms using computer vision and neural networks (Weeks & Gaston, 1997; MacLeod, 2008; Gaston & O'Neill, 2004; La Salle et al., 2009). Coupled with the more conventional methods of computer-aided identification best represented by the DELTA system, such techniques can provide a much-needed boost in the productivity of taxonomists faced with the biodiversity crisis. In this context, we expect that the availability of a free, open source, cross-platform software tool for the storage, retrieval, and analysis of DELTA-coded taxonomic descriptive data will prove a useful contribution to the present and future generations of taxonomists worldwide. This further justifies the importance of such initiatives as the Free-DELTA software, which in addition to ensuring speed and consistency in the management of complex taxonomic data, provides greater precision to achieve one of the most fundamental logical precepts of science: reproducibility (Peng, 2011)..

Acknowledgments

We thank Drs. Mike Dallwitz (CSIRO Division of Entomology [retired], Australia), Sami Rabei (Mansoura University, Egypt), and Gleyson Desidério (Instituto Nacional de Pesquisas da Amazônia, Brazil) for their reviews of the manuscript, providing corrections and useful suggestions which much contributed to the improvement of this paper.

References

Allkin, R., Moreno, N. P., Gama Campillo, L. & Mejia, T. 1992. Multiple uses for computer-stored taxonomic descriptions: keys for Veracruz. Taxon 41: 413-435.

Askevold, I. S. & O'Brien, C. W. 1994. DELTA, an invaluable computer program for generation of taxonomic monographs. Annals of the Entomological of America 87: 1-16.

Atlas of Living Australia. 2014. Open DELTA – A port of the DELTA (DEscription Language for TAxonomy) suite of applications into Java. https://github.com/atlasoflivingaustralia/open-delta.

Bello, E., Valdecasas, A. G.- & Elvira, J. R. 1996. A Spanish and English DELTA Editor: EDEL 95. DELTA Newsletter 12: 24-26.

Cavalcanti, M. J. 1996. Designing a Graphical User Interface for DELTA: Some considerations. DELTA Newsletter 12: 2-4.

Coleman, C. O., Lowry, J. K. & Macfarlane, T. 2010. DELTA for Beginners: An introduction into the taxonomy software package DELTA. ZooKeys 45: 1-75.

Dallwitz, M. J. 1974. A flexible computer program for generating identification keys. Systematic Zoology 23: 50-57.

Dallwitz, M. J. 1980. A general system for coding taxonomic descriptions. Taxon 29: 41-46.

Dallwitz, M. J. 1993. DELTA and INTKEY. In: Fortuner, R. (ed.). Advances in Computer Methods for Systematic Biology: Artificial

Intelligence, Databases, Computer Vision. The Johns Hopkins University Press, Baltimore, p. 287-296.

Dallwitz, M.J. & Paine, T.A. 1992 onwards. Definition of the DELTA format. https://www.delta-intkey.com/www/standard.htm

Dallwitz, M.J., Paine, T.A. & Zurcher, E.J. 1993 onwards. User's guide to the DELTA System: a general system for processing taxonomic descriptions. https://www.delta-intkey.com/www/uguide.htm

Dallwitz, M.J., Paine, T.A. & Zurcher, E.J. 1999 onwards. User's guide to the DELTA Editor. https://www.delta-intkey.com/www/delta-ed.htm

Edwards, M. & Morse, D. R. 1995. The potential for computer-aided identification in biodiversity research. Trends in Ecology and Evolution 10: 153-158.

Ellis, D. V. 1988. Quality control of biological surveys. Marine Pollution Bulletin 19: 506-512.

Forey, P. L., Humphries, C. J., & Vane-Wright, R. I. (Eds.). 1994. Systematics and Conservation Evaluation. Clarendon Press, Oxford.

Gaston, K. J. & O'Neill, M. A. 2004. Automated species identification: why not? Philosophical Transactions of the Royal Society, Biological Sciences 359: 655-667.

Geraldes-Primeiro, L. J. O. 2023. Sistemática Biológica de *Notodiaptomus* (Copepoda, Calanoida, Diaptomidae). Tese de Doutorado. Instituto Nacional de Pesquisas da Amazônia – INPA, Manaus, Amazonas. Ministério da Ciência, Tecnologia e Informação – Governo Federal do Brasil. 662 pp.

Goloboff, P. A., Farris, J. S., & Nixon, K. C. 2008. TNT, a free program for phylogenetic analysis. Cladistics 24: 774-786.

Gouda, E. J. 1996. TAXASOFT DELTA Editor. DELTA Newsletter 12: 12-13.

Hagedorn, G. 2001. Making DELTA accessible: databasing descriptive information. Bocconea 13: 261-280.

Ihaka, R. & Gentleman, R. 1996. R: A language for data analysis and graphics. Journal of Computational and Graphical Statistics5: 299-314.

Jarvie, J. K. & Stevens, P. F. 1998. Interactive keys, inventory, and conservation. Conservation Biology 12: 222-224.

Lander, N. S. 1993. DELTA Menu System for Windows. DELTA Newsletter 8: 10-12.

La Salle, J., Wheeler, Q., Jackway, P., Winterton, S., Hobern, D. & Lovell, D. 2009. Accelerating taxonomic discovery through automated character extraction. Zootaxa 2217: 43-55.

MacLeod, N. (ed.) 2008. Automated Taxon Identification in Systematics: Theory, Approaches and Applications. CRC Press, Boca Raton, FL.

Maddison, D. R., Swofford, D. L. & Maddison, W. P. 1997. NEXUS: an extensible file format for systematic information. Systematic Biology 46: 590-621.

Neubacher, D. & Rambold, G. 2005. NaviKey – a Java applet and application for accessing descriptive data coded in DELTA format: http://www.navikey.net.

Pankhurst, R. J. 1986. A package of computer programs for handling taxonomic databases.

Computer Applications in the Biosciences 2: 33-39.

Pankhurst, R. J. 1991. Practical Taxonomic Computing. Cambridge University Press, Cambridge.

Peng, R. D. 2011. Reproducible research in computational science. Science 334: 1226-1227

Rivetti, C. & Percudani, R. 1999. WebDelta – A set of Perl scripts for Web entry of DELTA descriptions. https://web.archive.org/web/20050227052425/irisbioc.bio.unipr.it/webdelta/webdelta.html.

Sharkey, M. J. 2001. The All Taxa Biological Inventory of the Great Smoky Mountains National Park. Florida Entomologist 84: 556-564.

Stallman, R. 1999. The GNU Operating System and the Free Software Movement. In: DiBona, C., Ockman, S. & Stone, M. (eds.). Open Sources: Voices from the Open Source Revolution. O'Reilly Media, Boston, p. 31-38.

Swofford, D. L. 2003. PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Sinauer Associates, Sunderland.

Watson, L. 1971. Basic taxonomic data: the need for organisation over presentation and accumulation. Taxon 20: 131-136.

Watson, L., Dallwitz, M. J., Gibbs, A. J., & Pankhurst, R. J. 1988. Automated taxonomic descriptions. In: Hawksworth, D. L. (ed.). Prospects in Systematics. Clarendon Press, Oxford, p. 292–304.

Weeks, P. J. D. & Gaston, K. J. 1997. Image analysis, neural networks, and the taxonomic impediment to biodiversity studies. Biodiversity and Conservation 6, 263-274.

White, R. J. & Allkin, R. 1992. A language for the definition and exchange of biological data sets. Mathematical and Computer Modelling 16: 199-223.

Wilson, E. O. 1992. The Diversity of Life. Harvard University Press, Cambridge, MA.

Vane-Wright, R. I. 1996. Systematics and the conservation of biological diversity. Annals of the Missouri Botanical Garden 83: 47-57.