3I export functions

http://ctap.inhs.uiuc.edu/dmitriev/
(Last updated December 15, 2011)

The new version of 3I provides export functionality to several global initiatives like Encyclopedia of Life, Catalog of Life, Discover Life, and Darwin Core Archive. DwCArchive is a new standard for data sharing from GBIF, which allows providing information not only on distribution, but also nomenclature, descriptions, illustrations, etc. If you have older version of 3I, you need to send your database to Dmitry Dmitriev for update.

NOTE: Export may take several hours for large database.

Before making any export, please make sure, that you clicked **Update for the Internet** button of the **Main Form** to refresh the hierarchical index for taxa. Some additional information also needs to be provided before the export to the tables **Export** and **Authors**:

Table **Export** (list of fields):

TaxonURL: This is a template for taxonomic pages. This will provide return links from the external web site (e.g. EoL) to your resource. Go to your web site (accessible through the Internet, not local copy on your computer). Open any taxonomic page. Copy the address line for this page and paste it to the TaxonURL field. Replace the number which follows "**hc=**" with "**{key}**". You should have the address line similar to:
http://ctap.inhs.uiuc.edu/dmitriev/taxahelp.asp?hc={key}&key=Erythroneura&lng=En
For the taxonomic page:
http://ctap.inhs.uiuc.edu/dmitriev/taxahelp.asp?hc=1234&key=Erythroneura&lng=En
This refers to a taxonomic page from the Typhlocybinae database.

MapURL (optional field): This is also a template similar to TaxonURL, but it provides link to the maps on your web site to Encyclopedia of Life. See an example to create links to your resource.
http://ctap.inhs.uiuc.edu/dmitriev/map.asp?key={key}&key1=Erythroneura&pric=3
You can try your template using browser window and replacing "{key}" with unique ID for any of your species.

EoLExport: name and physical location of the generated XML file for Encyclopedia of Life on your hard drive. For example, "C:\EoL.xml". The program will generate EoL.xml file on the C: drive.

CoLExportFolder: Catalog of Life export consists of several files which need to be archived with some software, for example WinZip. All generated files will be placed in the same folder (e.g. "C:\CoL").

DiscoverExport and **DiscoverTaxa**: these two fields provide names for two text files for Discover Life export. For example "C:\discover.txt" and "C:\discovertaxa.txt". The program will generate two files on the C: drive.

DwCArchiveFolder: Darwing Core Archive consists of several files which need to be archived with some software, for example WinZip. All generated files will be placed in the same folder (e.g. "C:\DwCArchive").

Key: Name of the highest taxon you want to export. For example, you have database for Hemiptera, but you want to export only one family Cicadellidae. Put Cicadellidae in this field.

Specialist: Put the name(s) of the people responsible for the database.

NomenclaturalCode: Code of nomenclature followed in the database. At the present time, 3I has full support to ICZN only. Put "ICZN" in the field.

DatabaseName: Name of your database.

DatabaseLongName (optional): Longer, more descriptive name of the database.

Abstract: Brief description of your database.

AdditionalInfo (optional): Any additional information.

IntelectualRights: Copyright restrictions; select one from the drop down box.
Difference of different licenses, as it is explained on the Encyclopedia of Life web site:

- Public Domain Dedication: Your work may be freely reproduced, distributed, transmitted, used, modified, built upon, or otherwise exploited by anyone for any purpose, commercial or non commercial, and in any way, including by methods that have not yet been invented or conceived.
- CC-BY (Attribution) This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered, in terms of what others can do with your works licensed under Attribution. (This is a default value in 3I, but you can change it).
- CC-BY-SA (Attribution Share Alike) This license lets others remix, tweak, and build upon your work even for commercial reasons, as long as they credit you and license their new creations under the identical terms. This license is often compared to open source software licenses. All new works based on yours will carry the same license, so any derivatives will also allow commercial use.
- CC-BY-NC (Attribution Non commercial) This license lets others remix, tweak, and build upon your work non commercially, and although their new works must also acknowledge you and be non commercial, they don't have to license their derivative works on the same terms.
- CC-BY-NC-SA (Attribution Non commercial Share Alike) This license lets others remix, tweak, and build upon your work non commercially, as long as they credit you and license their new creations under the identical terms. Others can download and redistribute your work as well as translate, make remixes, and produce new stories based on your work. All new work based on yours will carry the same license, so any derivatives will also be non commercial in nature.

InformationURL. This is a link to your web site with information about the resource.

LogoURL: If you use logo on your web site. You can provide a hyperlink to this file.

TaxonomicCoverage: Place of the group in hierarchy, e.g. "Animalia - Arthropoda - Insecta - Hemiptera - Membracoidea - Cicadellidae (subfamily Typhlocybinae)"

GroupNameInEnglish: common name of the group.

Organisation: Name of organization.

Coverage (values "Global" and "Regional". In the case of "Regional" I would suggest to specify the region in parentheses, for example "Regional (New Zealand)".

Completeness (values 0-100%; completeness should be stated by the custodian)

Checklist Confidence (quality of taxonomic checklist with values 1-5; quality should be stated by the custodian)

> 1 - Caution! This data set does not contain a scrutinised taxonomic checklist, and in parts may be a list of names only. However, as the only available resource at the time, it was included in the Catalogue of Life to fill temporarily a major gap. See database abstract for more details.

> 2 - Caution! This data set is a scrutinised taxonomic checklist, but is incomplete and at an early stage of its development. See database abstract for more details.

> 3 - This is a well scrutinised taxonomic checklist but is restricted to a subset of species by geography, or sector of biological discipline (conservation, medical and molecular oriented systems, etc). This data set was included in the Catalogue of Life to fill temporarily gaps at lower levels of the taxonomic classification (e.g. species, genera). See database abstract for more details.

> 4 - This is a nearly complete and fully scrutinised taxonomic checklist with a good quality of expertise at the current stage of its development.

> 5 - This is a complete and fully scrutinised taxonomic checklist for an entire taxon with a high quality of expertise and frequent updates, which covers nearly all known species diversity in the taxon worldwide.

ContactPesron: name of the contact person.

Table **Authors**

Role: role of the person in the database development. One person may have several roles. Each role for each person should be provided on a separate line. Roles include, but not restricted to: Author, Compiller, Composer, Director, Editor, Illustrator, Photographer, Publisher, etc. Two roles (could be the same person) need always to be in the list: **Creator** and **Metadata provider**.

Names of the other fields in this table are self explanatory. Provide as much information as possible.

Discover Life export

Click the button of the Main Form **Export file for DiscoverLife**. This will generate two files described in the field definitions to the table Export. The generated files could be placed on your web site and links provided to Discover Life for regular harvest (John Pickering, E-mail: pick@www.discoverlife.org). Update the files periodically. If files with the same names already exist in you computer, they will be replaced with newly generated ones.

Catalog of Life export

Click the button of the Main Form **Export file for Catalog of Life**. This will generate several CSV files in the folder designated in the table Export. Old data will be erased from this folder. The files need to be archived with any archiving program to make a single file with .zip, .gzip, or .tar extension.

Generated file could be sent to Catalog of Life (Yuri Roskov, E-mail: y.roskov@reading.ac.uk).

Encyclopedia of Life export

Each individual species needs to be selected for being included to the EoL export. To do this, go to the table **Taxon** and mark all the species in the Check Box **EoL**. You can also select license for each species in the **EoL_License** field. If no license is selected, the default value, **CC-BY**, will be applied. The license could also be selected for each individual image linked to the database.

Click the button of the Main Form **Export file for Encyclopedia of Life**. This will generate XML file defined in the table Export. The file could be placed on your web site and link registered at the EoL Web site. To do this, you need to register as a content partner. See http://www.eol.org/content/partners for more details.

Darwin Core Archive export

Each individual species needs to be selected for being included to the DwCArchive export. To do this, go to the table **Taxon** and mark all the species in the Check Box **DwC**. You can also select license for each species in the **EoL_License** field (the license is the same with EoL export). If no license is selected, the default value, **CC-BY**, will be applied.

Click the button of the Main Form **Export file for Darwin Core Archive**. This will generate several files in the folder defined in the table Export. The files need to be archived with any archiving program to make a single file with .zip, .gzip, or .tar extension. To register you DwCArchive with GBiF, go to http://tools.gbif.org/dwca-register/. Darwin Core Archive is also accepted by some journals (e.g. ZooKeys, PhyloKeys) as supplementary material, and could provide all associated information about revised group. The archive could also be provided to the Encyclopedia of Life and Global Name Index http://gni.globalnames.org/data_sources.

Nexus export

The data matrix could be exported in the format similar to Nexus and used for phylogenetic analysis or import into another application. Click **Data Matrix Export** button of the main form. Define characters used as line, character, and state separators, the data matrix to be exported, and name of the export file. The program will generate a TXT file.

SDD export

SDD export is not directly supported by 3i, but Gregor Hagedorn developed a converter from the 3i. See following web site for details:
http://wiki.tdwg.org/twiki/bin/view/SDD/RealWorldExamples_SDD1dot1