TDWG Standards Documentation Specification

Date: (created) 6th January 2007; (last modified) 5th November 2007

Contributors: Roger Hyam (TDWG Infrastructure Project) <roger@tdwg>

Abstract: This document defines how TDWG standards should be presented. Each standard is a logical directory or folder containing two or more files - a cover page outlining basic metadata for the standard and one or more normative files specifying the standard itself. Rules are specified for the naming of standards and files. Human readable files should be in English, follow basic layout principles and be marked up in XHTML. The legal statements that all documents must contain are defined.

Part of TDWG Standard: *** permanent URL here ***

Legal: This document is governed by the standard legal, copyright, licensing provisions and disclaimers issued by the Taxonomic Databases Working Group.

Table of Contents

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 (http://www.ietf.org/rfc/rfc2119.txt).

1 Motivation

Historically TDWG Standards have taken diverse forms. They have included controlled vocabularies (such as lists of approved abbreviations), transfer formats specified as XML Schemas

and geographic regions specified in hard copy. TDWG Standards can therefore be made up of multiple documents in many files of different formats. This standard specifies how these documents should be presented.

2 Rationale

Users see a standard as a single 'product'. It is therefore important that, although TDWG standards may contain multiple files, they are presented to the user as a single entity and can be downloaded as a single archive. File names need to be tightly constrained to be robust across multiple environments. A Cover Page is required for all standards to supply a uniform metadata interface.

3 Documents and Files

A document (a logical entity) may be composed of several files (physical entities) such as an HTML page containing images. This specification recognises 3 types of documents:

Type 1 documents are the normative parts of a standard.

Type 2 documents are parts of the standard that are non-normative (informative)

Type 3 documents are not part of the standard and may contain tutorials, introductory overviews. They are not governed by this document and are not stored in the standards repository.

The three different types are further described in Table 1.

	Type 1	**Type 2**	**Type 3**
Normative	Yes	No	No
Part of a standard	Yes	Yes	No
Function	Defines	Explains and justifies	Helps and Supports
Versioned with Standard	Yes	Yes	No
Controlled by TDWG Process	Yes	Yes	No
Document format	Tightly Controlled	Tightly Controlled	Not Controlled
Language	English	English + translations	Any

Table 1: Comparison of document types

This documentation standard specifies how Type 1 and Type 2 documents should be named and structured. It does not govern Type 3 documents.

4 Contents of Standards

At a minimum, each standard must contain:

- the normative (prescriptive) form of the standard itself (e.g., an XML Schema or human readable text); and

- a 'Cover Page' document that summarizes the content of the standard (see below).

A standard should also contain information on:

- the 'Motivation' for the existence of the standard; and

- the 'Rationale' for why the standard takes the form it does.

Standards may contain any number of files provided they are of an appropriate format (see below).

5   Packaging of Standards

Standards take the form of a logical folder or directory, but may be distributed as a zip or tar archive file. The name of the archive must be the file name of the standard followed by a period and the appropriate suffix if the standard is compressed.

6   Naming of Standards

Every standard must have two names, a full name and a file name.

The **full name** is used in the cover page and as the title of the main document if it is human readable. The full name must be in English and must be unique within the scope of TDWG standards, but otherwise is not controlled and may contain any combination of characters.

The **file name** should be created as a shortened version of the full name and is used for the logical folder or directory and any distribution archive. Rules for constructing file names are given below. The file name must also be unique among TDWG standards.

7   File Naming

The following rules govern file names used in standards and for standards archives:

- File names must start with 'tdwg_' followed by a shortened version of the full name of the standard.

- File names must consist of only lowercase letters, numbers, underscores and hyphens.

- File names must not contain spaces (replace with underscore).

- File names must not contain non ASCII characters.

- File names must not contain periods other than the one used to separate the suffix.

- Suffix should be the commonly used suffix for file type and the three letter version where possible; for example, jpg rather than jpeg and xsl rather than xslt.

8   Versioning of Standards

TDWG provides no formal versioning mechanism for standards. Once a standard has been ratified it cannot be changed in any substantive way; it must be superseded by a standard with a different name. Relationships between related standards (old standards and their replacements) should be indicated in the Cover Page document.

9   File Formats

For archival reasons all files in a standard must be in an open format for which parsers are commonly available. For this purpose an open format is defined as being one for which it would be possible to write a parser on the basis of a published specification without having to rely on code libraries for which the source code is not available or to pay a license fee. The definition of 'commonly available' is left to the discretion of the TDWG Executive Committee.

Formats that are currently considered to be acceptable include:

- ASCII Text
- XML
- HTML
- Portable Document Format (PDF)
- Portable Network Graphic (PNG)
- JPEG
- OASIS OpenDocument

10 Referencing of Standards

Each standard must have the following URIs (Uniform Resource Identifiers)

- A URI for the standard itself. This must resolve to human readable rendering of the information in the standard's cover page.
- A URI that resolves to an archive file containing all the files in the standard.
- A URI for each of the files within the standard.

The URIs are issued as part of the standards process by the administrator of the standards repository and remain constant from the point at which they are issued.

Some standards may involve technologies that make use of XML namespaces. The URIs of the standard and its constituent parts should not be used for this purpose. TDWG provides a facility for reserving URIs associated with namespaces with the base namespace http://rs.tdwg.org/.

11 Human Readable Documents

This section specifies the layout and style that should be followed for any human readable documents that make up part of a TDWG standard with the exception of the cover page document which may be automatically generated. Normative human readable documents should be in XHTML format. Documents should be divided into three main sections:

- A header section
- Table of contents
- A body section

11.1 Header Section

This section must contain the following parts:

Title: The official title of the document.

Date: The date of publication in the format "day Month year" (for example, 28 February 2007). This must not be later than the date the standard has been accepted/ratified as given in the Cover Page document.

Abstract: A single paragraph summary of the document.

Contributors: A list of the people who are responsible for the document's creation. The first person in the list should be the principal contact.

Legal: A reference or link to a document containing the Copyright, Licensing and Disclaimer statements that govern this document. This will usually be the standard one supplied by

TDWG.

Part of TDWG Standard: The URL of the standard this document belongs to. This should be a link to the standard.

Contributors' names must be quoted in the following format: "FirstName Initials LastName (Affiliation) <email>" where:

Either one or both of FirstName and Initials must be present.

Affiliation is the organisation or group the contributor is working with in relation to this document and is OPTIONAL.

Examples:

- Roger D. Hyam (TDWG) <roger@tdwg.org>

- Roger Hyam (TDWG Infrastructure Project) <roger@tdwg.org>

- R. Hyam <roger@hyam.net>

11.2 Table of Contents Section

The table of contents section should contain an ordered list of all the headings (along with their numbers) that occur in the body section and nothing else. The items in the list should be hypertext links to the headings in the body.

11.3 Body Section

The body text should be divided by a hierarchy of subheadings. The subheadings must be numbered consecutively using simple decimal system. The final number must not be followed by a decimal point. An example:

1 First Main Section

1.1 First subsection within first main section.

1.12.5 The fifth sub-subsection within the twelfth subsection of the first section.

2 Second Main Section

12   Language

All normative (Type 1) documents must be in English. Translations of normative documents may be included in the standard but the translations must be treated as informative (Type 2) documents. Note: Type 3 documents may be produced in any language with no requirement for translation, but they do not form part of a standard and are not governed by this document.

12.1 Linguistic Style

Text should be written with the assumption that it will be read from the screen rather than paper. Current best practise should be followed (for example, http://www.sun.com/980713/webwriting/).

12.2 Key Words Governing Compliance

It is highly RECOMMENDED that the policy of the IETF be followed as regards the use of the key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL". This policy is specified in RFC 2119 (http://www.ietf.org/rfc/rfc2119.txt).

13   Cover Page Contents

This section specifies content for the Cover Page documents that must be contained in all TDWG standards. The table below shows the components that must be contained in the cover page of all TDWG standards. These are based on a subset of the Dublin Core metadata initiative elements and terms.

It is envisaged that the contents and format of cover page documents will be controlled by the collaborative environment used to develop and host TDWG standards and that the data will be made available in RDF or similar formats.

Table 2: Fields for Cover Page documents.

Name	Notes
Title	**Required** The official TDWG title for the standard (see above)
Description	**Required** A brief (<=250 words) description of the standard
Subject	**Required** Key words, key phrases and classification codes that describe the topic covered by the standard.
Creator	**Required** The primary author of the standard following the conventions outlined above. Should only occur once.
Contributor	A contributor to the standard following the conventions outlined above. This assertion may be repeated.
Date Modified	The date the standard was last modified. Will be prior to Date Accepted if that is present.
Date Accepted	**Required for accepted standards** The date the standard was officially ratified by TDWG and fixed. If a modified date is present it must be prior to this date.
Identifier	The GUID of the standard. An LSID.
Publisher	**Required** Should always contain the text "Biodiversity Information Standards TDWG."
Rights	**Required** One or more URLs linking to documents describing the copyright, licensing provisions and disclaimers associated with the standard. Typically this will be contained within another TDWG standard. (see Legal Statements below)
Access Rights	**Required** Should always contain the single word 'public'. All TDWG standards should be freely accessible to all via the Internet.
Bibliographic Citation	**Required** How this standard should be cited.
Has Part	**Required** The full URL to a file within the standard. This assertion should be repeated for each file within the standard (other than cover.xml itself).
Replaced By	**Required if this standard has been replaced by another.** If the standard has been replaced then this assertion should be present and contain the full URL to the new standard.
Replaces	**Required if this standard replaces another.** If this standard replaces a previous one then this assertion should be present and contain the full URI of the standard replaced.

14   Legal Statements

TDWG standards must contain explicit statements of copyright (the assertion of intellectual property rights), the licensing provisions for the standard, and a disclaimer of liability. This section

defines the texts that should be included in TDWG standards and actions that should be taken if these texts are deemed not to be suitable for a particular standard.

14.1 Copyright

The usual practice will be for participants in TDWG activities to contribute their intellectual works to TDWG freely, and to combine them with the contributions of others to create a TDWG standard. TDWG strongly recommends that TDWG standards be placed under TDWG copyright.

The following copyright notice should be included in all TDWG documents

Copyright © TDWG *(year).* All Rights Reserved.

14.2 Licensing Provisions

TDWG intends its standards to be used as broadly as possible and should impose no royalties or licensing fees for their use. TDWG will normally publish its standards under the Creative Commons "By Attribution" license. Under this license users may copy, completely or in part, or create derivative works for any purpose, whether commercial or public benefit, subject only to the requirement that the original standard must be cited.

The following licensing provision should be included or referenced in all TDWG standards:

```
Creative Commons License Deed: Attribution 3.0
You are free:
    ●  to Share — to copy, distribute, display, and perform the work
    ●  to Remix — to make derivative works under the following conditions:
        ●  Attribution. You must attribute the work to TDWG by citing the standard
           by name and by providing the URL to the original document, but not in any
           way that suggests that TDWG endorses you or your use of the work.
For any reuse or distribution, you must make clear to others the license terms of
this work. The best way to do this is with a link to this web page.
Any of the above conditions can be waived if you get permission from TDWG.
Apart from the remix rights granted under this license, nothing in this license
impairs or restricts the author's moral rights.
Your fair use and other rights are in no way affected by the above.
This is a human-readable summary of the Legal Code (the full license).
```

The text above is derived from the Creative Commons License Deed: Attribution 3.0 and represents a TDWG specific application of the original (generic) license.

14.3 Disclaimer

The following text must be included or referenced in all TDWG standards documents:

Disclaimer: This document and the information contained herein are provided on an "AS IS" basis. TDWG MAKES NO WARRANTIES REGARDING THE INFORMATION PROVIDED, AND DISCLAIMS LIABILITY FOR DAMAGES RESULTING FROM ITS USE.

14.4 Convenience Document

As a convenience this standard includes a document that just contains these texts which can be referenced (linked) from any other document. The link should bear the following text: This document is governed by the standard legal, copyright, licensing provisions and disclaimers issued by the Taxonomic Databases Working Group.

14.5 Departure from Usual Practice

If the participants in a TDWG activity intend to depart from the legal framework provided above, the alternative statements of intellectual property rights and licensing must be described fully in the

charter of that activity as well as the standard itself.