

Specifications for the Digital Talking Book

Abstract: This standard defines the format and content of the electronic file set that comprises a digital talking book (DTB) and establishes a limited set of requirements for DTB playback devices. It uses established and new specifications to delineate the structure of DTBs whose content can range from XML text only, to text with corresponding spoken audio, to audio with little or no text. DTBs are designed to make print material accessible and navigable for blind or otherwise print-disabled persons.

An American National Standard
Developed by the
National Information Standards Organization

Approved April 21, 2005
by the
American National Standards Institute
Reaffirmed April 16, 2012

Published by the National Information Standards Organization
Baltimore, Maryland, U.S.A.

About NISO Standards

NISO standards are developed by the Committees and Working Groups of the National Information Standards Organization. The development process is a strenuous one that includes a rigorous peer review of proposed standards open to each NISO Voting Member and any other interested party. Final approval of the standard involves verification by the American National Standards Institute that its requirements for due process, consensus, and other approval criteria have been met by NISO. Once verified and approved, NISO Standards also become American National Standards.

This standard may be revised or withdrawn at any time. For current information on the status of this standard contact the NISO office or visit the NISO website at:

<http://www.niso.org>

Published by

NISO
3600 Clipper Mill Road
Suite 302
Baltimore, MD 21211

Copyright © 2010 by the National Information Standards Organization

All rights reserved under International and Pan-American Copyright Conventions. For noncommercial purposes only, this publication may be reproduced or transmitted in any form or by any means without prior permission in writing from the publisher, provided it is reproduced accurately, the source of the material is identified, and the NISO copyright status is acknowledged. All inquiries regarding translations into other languages or commercial reproduction or distribution should be addressed to: NISO 3600 Clipper Mill Road, Suite 302, Baltimore, MD 21211.

ISSN: 1041-5653 National Information Standards series
ISBN: 978-1-937522-35-3

Contents

	Page
Preface to 2005 Edition	iv
Foreword	v
1 General Information	1
1.1 Purpose and Scope of Standard	1
1.2 Definitions	1
1.3 Strategy	4
1.4 Accessibility Issues.....	4
1.5 Relationship to Other Specifications	5
1.5.1 Relationship to Unicode.....	5
1.6 Patent Rights	5
1.7 Maintenance Agency	5
2 Overview	5
3 The DTB Package File	7
3.1 Package Identity	8
3.2 Publication Metadata	9
3.2.1 Dublin Core Metadata.....	9
3.2.2 DTB ID Scheme.....	12
3.2.3 X-Metadata	12
3.3 Manifest	15
3.3.1 Allowed Characters in File Names	17
3.3.2 Case Sensitivity of URIs.....	18
3.4 Spine.....	18
3.5 Tours.....	19
3.6 Guide	19
4 Content Format for Text	19
4.1 Introduction	19
4.2 Using the DTBook Element Set.....	20
4.2.1 DTBook Markup Related to SMIL.....	20
4.2.2 Modular Extension of the DTD	21
5 Audio File Formats	22
5.1 Distribution Formats	22
5.2 Formats for Audio Notes.....	23
6 Image File Formats	23

7	Synchronization of Media Files	23
7.1	Introduction	23
7.1.1	Background	23
7.1.2	SMIL Modules	24
7.2	Application of SMIL to DTBs	25
7.3	SMIL Elements	26
7.3.1	Core Attributes	32
7.3.2	xml:lang Attribute	32
7.4	SMIL Requirements for DTBs	32
7.4.1	"Escapable" Structures	32
7.4.2	Automatic Invocation of Special Navigation Modes	33
7.4.3	"Skippable" Structures	33
7.4.4	Packaging Files across Several Media Units	34
7.4.5	Links	34
7.4.6	Layout Syntax	34
7.4.7	Content of <par>s	34
7.4.8	Notes and Annotations in SMIL	34
7.4.9	Images in SMIL	35
7.4.10	Text-Only DTBs	35
7.4.11	Producer Pauses	35
7.5	SMIL Metadata	35
7.6	Examples	36
7.7	Media Clipping and Clock Values	40
7.8	End Attribute Values	41
7.8.1	Allowed End Values	42
7.8.2	Computing the Active Duration	42
7.8.3	Processing Nested Structures	42
8	Navigation Control File (NCX)	43
8.1	Introduction	43
8.2	Key NCX Requirements	43
8.3	NCX Elements	43
8.4	Other File Requirements	50
8.4.1	Navigation Metadata	50
8.4.2	DTBs Spanning Multiple Media Units	51
8.4.3	playOrder Attribute	51
8.4.4	smilCustomTest Element	51
8.4.5	Enabling Page Navigation	52
8.5	How the NCX Works	52
8.6	Example	53
9	Portable Bookmarks and Highlights	56
9.1	Introduction	56
9.2	Bookmark/Highlight Elements	57
9.3	Examples	61

10 Resource File	63
10.1 Introduction	63
10.2 Resource Elements	64
10.3 Resource File Requirements	66
10.4 Examples	68
11 Packaging Files for Distribution	70
11.1 Introduction	70
11.2 Distribution Requirements	70
11.3 DistInfo Elements	71
11.4 Examples	75
12 Presentation Styles	77
12.1 Introduction	77
12.2 Implementing Style Sheets for DTBs	78
13 Content Rendering	78
14 Digital Rights Management	79
15 Time-Scale Modification	79
16 Conformance	79
16.1 General File Conformance Requirements	79
17 References to Other Specifications/Documents	80
17.1 Normative References	80
17.2 Informative References	81
Appendices	83
Appendix 1 : Document Type Definitions (DTDs)	83
Appendix 2 : Designation of Maintenance Agency	83

Preface to 2005 Edition

(This preface is not a part of ANSI/ Z39.86-2005, *Specifications for the Digital Talking Book*. It is included for information only.)

ANSI/NISO Z39.86 was first released in 2002. Preparation of this version of the standard was prompted by several factors: first, a desire to stay abreast of current practices in related standards and specifications; second, a need to clarify a few ambiguities present in the first release; and third, the wish to add enhancements requested by those using, or planning to use, the standard.

A variety of changes stemming from all of the above reasons were made. Some of the more significant enhancements are listed below:

- Producer-controlled pauses. These allow a DTB author to cause the presentation to pause at specified points and wait for a user action. This function could be used in workbooks when the user needs to, for example, explore a model before continuing.
- A more robust method of ensuring that players can maintain their location when navigation methods are mixed
- A cleaner mechanism for enabling direct access to pages
- New elements for marking up a textual content file
- A more efficient method of identifying individual books when more than one is present on a single piece of media

Individual changes made to the standard are described in comments that can be viewed by reading the HTML version in an HTML authoring tool or with a text editor.

Foreword

(This foreword is not a part of ANSI/ Z39.86-2005 (R2012), *Specifications for the Digital Talking Book*. It is included for information only.)

This standard presents specifications for digital talking books (DTBs) for blind, visually impaired, physically handicapped, learning-disabled, or otherwise print-disabled readers. For many years, "talking books" have been made available to print-disabled readers on analog media such as phonograph records and audiocassettes. These media serve their users well in providing human-speech recordings of a wide array of print material in increasingly robust and cost-effective formats. However, analog media are limited in several respects when compared to a print book. First, they are by their nature linear presentations, which leave much to be desired when reading reference works, textbooks, magazines, and other materials that are often accessed randomly. In contrast, digital media offer readers the ability to move around in a book or magazine as freely as (and more efficiently than) a sighted reader flips through a print book. Second, analog recordings do not allow users to interact with the book by placing bookmarks or highlighting material. A DTB offers this capability, storing the bookmarks and highlights separate from, but associated with, the DTB itself. Third, talking book users have long complained that they do not have access to the spelling of the words they hear. As will be explained below, some DTBs will include a file containing the full text of the work, synchronized with the audio presentation, thereby allowing readers to locate specific words and hear them spelled. Finally, analog audio offers readers only one version of the document. If, for example, a book contains footnotes, they are either read where referenced, which burdens the casual reader with unwanted interruptions, or grouped at a location out of the flow of the text, making them difficult for interested readers to access. A DTB allows the user to easily skip over or read footnotes. The Digital Talking Book offers the print-disabled user a significantly enhanced reading experience -- one that is much closer to that of the sighted reader using a print book.

The DTB goes far beyond the limits imposed on analog audio books because it can include not just the audio rendition of the work, but the full textual content and images as well. Because the textual content file is synchronized with the audio file, a DTB offers multiple sensory inputs to readers, a great benefit to, for example, learning-disabled readers. Some visually impaired readers may choose to listen to most of the book, but find that inspecting the images provides information not available in the narrative flow. Others may opt to skip the audio presentation altogether and instead view the text file via screen-enlarging software. Braille readers may prefer to read some or all of the document via a refreshable Braille display device connected to their DTB player and accessing the textual content file. DTBs containing a textual content file but no audio material might be accessed via synthetic speech, screen-enlarging software, or a Braille device.

Digital Talking Books are not tied to a single distribution medium. CD-ROMs will be used first but DTBs will be portable to any digital distribution medium capable of handling the large files associated with digital audio recordings. Regardless of how a DTB is distributed, however, it will normally be in the context of an intellectual property protection system.

Reaffirmation

At the time this standard was up for its periodic review, it was determined that a major revision was needed to update the standard for a modular, extensible XML format. Following a draft for trial use of the revised standard, feedback was received that the scope had been broadened so substantially and the revisions were so significant that implementers needed a transition period of several years. As a result, it was decided that the revision would receive a new standard number and title (ANSI/NISO Z39.98-2012, *Authoring and Interchange Framework for Adaptive XML Publishing Specification*) and the existing Z39.86-2005 standard would be reaffirmed to remain active during the transition period.

The NISO consensus body approved the reaffirmation on March 18, 2012 and it was approved by ANSI on April 16, 2012.

Suggestions for improving this standard are welcome. They should be sent to the National Information Standards Organization, 3600 Clipper Mill Road, Suite 302, Baltimore, MD 21211, telephone (301) 654-2512.

NISO Voting Members

This Standard (2005 edition) was processed and approved for submittal to ANSI by the National Information Standards Organization. NISO approval of this Standard does not necessarily imply that all Voting Members voted for its approval. At the time it approved this Standard, NISO had the following members:

3M

Susan Boettcher
Roger D. Larson, Alt

American Association of Law Libraries

Robert L. Oakley
Mary Alice Baish, Alt

American Chemical Society

Matthew Toussant

American Library Association

Betty Landesman

American Society for Information Science and Technology (ASIS&T)

Gail Thornburg

American Society of Indexers

Judith Gibbs

American Theological Library Association

Myron Chace

ARMA International

Diane Carlisle

Armed Forces Medical Library

Diane Zehnpfennig
Emily Court, Alt

Art Libraries Society of North America (ARLIS/NA)

Mark Bresnan

Association for Information and Image Management (AIIM)

Betsy A. Fanning

Association of Information and Dissemination Centers (ASIDIC)

Margie Hlava

Association of Jewish Libraries

Caroline R. Miller
Elizabeth Vernon, Alt

Association of Research Libraries

Duane E. Webster
Julia Blixrud, Alt

Auto-Graphics, Inc.

Paul Cope

Barnes & Noble, Inc.

Douglas Cheney

Book Industry Communication

Brian Green

California Digital Library

Daniel Greenstein
John Kunze, Alt

Cambridge Information Group

Michael Cairns
Matthew Dunie, Alt

The Cherry Hill Company

Cary Gordon

College Center for Library Automation (CCLA)

Richard Madaus
Ann Armbrister, Alt

Colorado State Library

Brenda Bailey-Hainer
Steve Wrede, Alt

CrossRef

Edward Pentz
Amy Brand, Alt

Davandy, L.L.C.

Michael J. Mellinger

DYNIX Corporation

Ed Riding
Gail Wanner, Alt

EBSCO Information Services

Gary Coker
Oliver Pesch, Alt

Elsevier

Paul Mostert

Endeavor Information Systems, Inc.

Sara Randall
Shelley Hostettler, Alt

Entopia, Inc.

Igor Perisic

Ex Libris, Inc

James Steenberg

Fretwell-Downing Informatics

Robin Murray

Gale Group

Katherine Gruber
Justine Carson, Alt

Geac Library Solutions

Eric Conderaerts
Eloise Sullivan, Alt

GIS Information Systems, Inc.

Candy Zemon
Paul Huf, Alt

H.W. Wilson Company

Ann Case
Patricia Kuhr, Alt

NISO Voting Members (continued)

Helsinki University Library

Juha Hakala

Index Data

Sebastian Hammer
David Dorman, Alt

INFLIBNET Centre

T A V Murthy
Rajesh Chandrakar, Alt

Infotrieve

Jan Peterson

Innovative Interfaces, Inc.

Gerald M. Kline
Betsy Graham, Alt

The International DOI Foundation

Norman Paskin

Ithaka/JSTOR/ARTstor

David Yakimischak
Bruce Heterick, Alt

John Wiley & Sons, Inc.

Eric Swanson

Library Binding Institute

Debra Nolan

The Library Corporation

Mark Wilson
Ted Koppel, Alt

Library of Congress

Sally H. McCallum, Alt

Los Alamos National Laboratory

Richard E. Luce

Lucent Technologies

M.E. Brennan

Medical Library Association

Nadine P. Ellero
Carla J. Funk, Alt

MINITEX

Cecelia Boone
William DeJohn, Alt

Modern Language Association

Daniel Bokser
B. Chen, Alt

MuseGlobal, Inc.

Kate Noerr
Clifford Hammond, Alt

Music Library Association

Mark McKnight
David Sommerfield, Alt

National Agricultural Library

Eleanor G. Frierson
Gary K. McCone, Alt

National Archives and Records Administration

Nancy Allard

National Library of Medicine

Betsy L. Humphreys

National Security Agency

Kathleen Dolan

NFAIS

Marjorie Hlava

Nylink

Mary-Alice Lynch
Jane Neale, Alt

OCLC Online Computer Library Center

Thomas Hickey

Openly Informatics, Inc.

Eric Hellman

ProQuest Information and Learning

Thomas Hamilton
Carol Brent, Alt

Random House, Inc.

Laurie Stark

Recording Industry Association of America

Bruce Block
Carlos Garza, Alt

RLG

Lennie Stovel
Joan Aliprand, Alt

Sage Publications

Carol Richman
Richard Fidczuk, Alt

Serials Solutions, Inc.

Mike McCracken

SIRSI Corporation

Greg Hathorn
Slavko Manojlovich, Alt

Society for Technical Communication (STC)

Frederick O'Hara
Annette D. Reilly, Alt

Society of American Archivists

Lisa Weber

Special Libraries Association (SLA)

Foster J. Zhang

Synapse Corporation

Trish Yancey
Dave Clarke, Alt

TAGSYS, Inc.

John Jordon
Anne Salado, Alt

Talis Information Ltd

Terry Willan
Katie Anstock, Alt

Thomson ISI

Carolyn Finn

Triangle Research Libraries Network

Mona C. Coutts, Alt

ANSI/NISO Z39.86-2005 (R2012)

NISO Voting Members (continued)

U.S. Department of Commerce, NIST, Office of Information Services

Mary-Deirdre Coraggio

U.S. Department of Defense, DTIC (Defense Technical Information Center)

Richard Evans

Jane L. Cohen, Alt

U.S. DOE, Office of Scientific & Technical Information

Ralph Scott

Karen Spence, Alt

U.S. Government Printing Office

Judith Russell

T.C. Evans, Alt

U.S. National Commission on Libraries and Information Science (NCLIS)

Robert Molyneux

VTLS, Inc.

Carl Grant, Alt

WebFeat

Todd Miller

Paul Duncan, Alt

ANSI/NISO Z39.86-2005 Review Voting Pool

At the time this standard was approved for reaffirmation, the following were members of the ANSI/NISO Z39.86-2005 Review Voting Pool. NISO approval of this Standard does not necessarily imply that all Voting Members voted for its approval.

American Library Association (ALA)

Nancy Kraft

American Psychological Association

Linda Beebe

American Society for Indexing

Judith Gibbs

American Society for Information Science & Technology (ASIS&T)

Mark Needleman

Association of Research Libraries (ARL)

Julia Blixrud

College Center for Library Automation (CCLA)

David Brightbill

DAISY Consortium

George Kerscher

Inera Inc.

Bruce Rosenblum

ITHAKA/JSTOR/Portico

Amy Kirchhoff

Library of Congress

John Zagas

Lyrasis

Peter Murray

Music Library Association

Mark McKnight

National Archives and Records Administration (NARA)

Marilyn Redman

National Library of Medicine (NLM)

Barbara Rapp

National Security Agency

Kate Dolan

Polaris Library Systems

Eric Graham

Recording Industry Association of America

Paul Jessop

NISO Board of Directors

At the time NISO approved this standard (2005), the following individuals served on its Board of Directors:

Jan Peterson, Chair
Infotrieve

Carl Grant, Vice Chair/Chair-Elect
VTLS, Inc.

Beverly C. Lynch, Immediate Past Chair
UCLA Graduate School of Education & Information Studies

Michael J. Mellinger, Treasurer
Davandy, L.L.C.

Patricia R. Harris, Executive Director/Secretary
National Information Standards Organization

Directors:

Douglas Cheney
Barnes & Noble, Inc.

Brian Green
BIC/EDITEUR

Daniel Greenstein
California Digital Library

Deborah Loeding
The H.W. Wilson Company

Richard E. Luce
Los Alamos National Laboratory

Robin Murray
Fretwell- Downing Informatics

James Neal
Columbia University

Oliver Pesch
EBSCO Information Services

Patricia Stevens (SDC Chair)
OCLC, Inc.

Eric Swanson
John Wiley & Sons, Inc.

Advisory Committee Members

Following approval of this standard in 2002, an advisory committee was formed to oversee the maintenance and enhancement of the document. The following individuals comprised the committee during the preparation of this (2005) version of the standard.

Thomas Kjellberg Christensen
The Danish National Library for the Blind

Markus Gylling
The DAISY Consortium and the Swedish Library of Talking Books and Braille

Markku Hakkinen
The DAISY for All Project/, DAISY Consortium

George Kerscher, co-Chair
The DAISY Consortium and Recording for the Blind & Dyslexic

Tom McLaughlin
National Library Service for the Blind and Physically Handicapped

Michael Moodie, co-Chair
National Library Service for the Blind and Physically Handicapped

David Pawson
The Royal National Institute of the Blind

James Pritchett
Recording for the Blind & Dyslexic

Lloyd Rasmussen
National Library Service for the Blind and Physically Handicapped

Jennifer Sutton, scribe

Acknowledgements

The Advisory Committee gratefully acknowledges the substantial contributions made by the following individuals to the continuing development of the standard. In particular, Ole Holst Andersen gave greatly of his time and expertise to the effort. The DAISY Consortium's XML-Techniques working group also provided excellent feedback on the use of the DTBook DTD and identified aspects needing improvement.

Ole Holst Andersen, Danish National Library for the Blind; Jon Beatty, Minnetonka Software, Inc.; Harvey Bingham; Don Breda, American Council of the Blind; Sean Brooks, Canadian National institute for the Blind; John Bryant, National Library Service for the Blind and Physically Handicapped; Curtis Chong, National Federation of the Blind; John Cookson, National Library Service for the Blind and Physically Handicapped; Keith Creasy, American Printing House for the Blind; Tim Curtin, gh;

ANSI/NISO Z39.86-2005 (R2012)

Marisa de Meglio, the DAISY for All Project; DAISY Consortium; Guillaume du Bourguet, BrailleNet Association; Jim Dust, Telex Communications Corporation; Daniel Farrington, Dolphin Audio Publishing; Dan Germann, LR Sound; Al Gilman; Luis Gutierrez, American Foundation for the Blind; Diana Hiorth Persson, Dolphin Audio Publishing; John Kibitlewski, gh; Jesper Klein, Swedish Library of Talking Books and Braille; Johan Knol, IDUNA Electronics BV; Brad Kormann, National Library Service for the Blind and Physically Handicapped; Kathy Kopolinski, Recording for the Blind & Dyslexic; Dominic Labbé, VisuAide, Inc.; Chris Lehn, Telex Communications Corporation; Lynn Leith, Canadian National Institute for the Blind; Olaf Mittelstaedt, Swiss Library for the Blind and Visually Handicapped; Brandon Nelson, Canadian National Institute for the Blind; Laust Skat Nielsen, Danish National Library for the Blind; Tatsuo Nishizawa, Plextor; Joe Said, gh; Janina Sajka, American Foundation for the Blind; Gregg Savage, Talking Book Publishers, Inc.; Dave Schleppenbach, gh; Per Sennels, National Resource Center for Special Education of the Visually Impaired (Norway); Sheela Sethuraman, CAST; Charles Steaderman; Jeff Suttor, SUN Microsystems; Niels Thögersen, Danish Institute for the Blind; Chris von See, TechAdapt; Christian Wallin, Danish National Library for the Blind; Chris Wilder-Smith, CAST.

Specifications for the Digital Talking Book

1 General Information

1.1 Purpose and Scope of Standard

(This section is [informative](#).)

This standard establishes specifications for digital talking books ([DTBs](#)) for blind, visually impaired, physically handicapped, learning-disabled, or otherwise print-disabled [readers](#). Its purpose is to ensure interoperability across service organizations and vendors providing content and [playback systems](#) to the [target population](#).

This standard provides specifications primarily for DTB files and their interrelationships. It also includes specifications for DTB [playback](#) devices in two areas: player performance related to file requirements and player behavior in areas defined in user requirements.

1.2 Definitions

(This section is [normative](#).)

The following abbreviations, acronyms, phrases, and terms are used in this standard as defined below. In the following definitions and throughout the standard, bracketed items correspond to entries in section 17, "[References to Other Specifications/Documents](#)," where the full URL is provided for each reference.

Accessible

Fully usable by the target population.

CSS

[Cascading Style Sheets \[CSS\]](#) is a mechanism for adding style (e.g. fonts, colors, spacing, formatting) to HTML or [XML](#) documents.

DRM

Digital Rights Management is a system of tools and processes that protect intellectual property when it is encoded and distributed in digital form.

DTB

The *Digital Talking Book* content data set that complies with the specifications in this standard.

DTBook

An [XML](#) element set (dtbook.dtd) that defines the markup for the textual content of a [DTB](#).

DTD

The *Document Type Definition* file contains machine- and human-readable rules that define allowable [XML](#) markup for a particular application.

FIXED

When used in definitions of XML element attributes, means that the attribute has a single, fixed value specified in the [DTD](#). See [IMPLIED](#) and [REQUIRED](#).

Fragment Identifier

A means to address a named place in a document. For reference within the current document, the reference part is to a named target and begins with "#". See [URI](#) for addressing into another document.

Global navigation

Movement to user-selected portions of a document, with that movement enabled by the [NCX](#). Navigation targets [may](#) be headings representing the hierarchical structure of the document or specific points such as pages, notes, sidebars, etc.

IMPLIED

When used in definitions of XML element attributes, means that the attribute is optional and that no default value is supplied. See [FIXED](#) and [REQUIRED](#).

Informative

Supplying background or explanation. Contrast with [Normative](#).

Local navigation

Movement within a document at a granularity finer than that provided by the [NCX](#). For example, navigation by paragraph or sentence, or within a table or nested list. Precise local navigation can be controlled by the [textual content file](#) or the [SMIL](#) file(s); the granularity is limited by the degree to which the textual content file has been marked up or the level to which synchronization has been applied in the [SMIL](#) file(s). Time-based movement through a document (e.g., fast-forward and rewind as on an analog cassette, or time jumps by specified intervals) may also be implemented.

Manifest

A component of the Package File, the *Manifest* lists all files included in the [DTB](#).

May

In normative sections, the word *may* means that a course of action is optional.

Media Unit

A single object on which a [DTB](#) is stored for distribution to the [reader](#). For example, a single CD-ROM disk.

Must

In normative sections, the word *must* is to be interpreted as a mandatory requirement on the content or implementation. The term [shall](#) has the same definition as *must*.

NCX

The *Navigation Control file for XML applications (NCX)* provides the reader efficient and flexible access to the hierarchical structure of a [DTB](#) as well as direct access to selected elements such as page numbers, notes, figures, etc.

Normative

Setting forth requirements that must be met to establish conformance with this standard; or providing recommendations or optional courses of action. For recommended or optional features, conformance is not dependent on the fact of implementation, but, if implemented, that implementation is as prescribed in this standard. Contrast with [Informative](#). Notes within a normative section may be informative.

OEBF

The Open eBook Forum™ [\[OEBF\]](#) is an organization formed to create and maintain standards and promote the successful adoption of electronic books. The *Open eBook Publication Structure Version 1.2* provides a specification for representing the content of a book when it is converted from print to electronic form. This [DTB](#) standard uses a subset (the Package File) of that specification.

OPF

Open eBook Forum Package File. See [Package File](#).

Package File

The *Open eBook Forum Package File* (OPF) is an [XML](#) file conforming to the oebpkg12.dtd that contains administrative information about the [DTB](#), the files that comprise it, and how these files interrelate.

Playback

With regard to implementations, *playback* refers to the methods used to render the [DTB](#) content. *Playback* may include audio, Braille, large print, and synthetic speech as appropriate for the content and as supported by the playback system.

Playback System

The hardware/software platform that renders the contents of a [DTB](#) to a [reader](#). Synonymous with *Player*.

Player

See [Playback System](#).

Reader

The person reading the digital talking book. Synonymous with *User*.

REQUIRED

When used in definitions of XML element attributes, means that the attribute must always be provided. See [FIXED](#) and [IMPLIED](#).

Shall

See [Must](#)

Should

In normative sections, the word *should* means that a course of action is recommended but not required.

SMIL

The Synchronized Multimedia Integration Language [\[SMIL\]](#) is a W3C recommendation (SMIL 2.0) used in this standard to control the synchronized presentation of content in multiple media.

Spine

A component of the Package File, the *Spine* lists in default reading order the [SMIL](#) files included in the [DTB](#).

Target population

The *target population* consists of blind, visually impaired, physically handicapped, learning-disabled, and otherwise print-disabled readers.

Textual Content File

The content of the subject document in a character set specified by ISO/IEC 10646 [\[ISO 10646\]](#) to which [XML](#) markup valid to the [DTBook DTD](#) has been applied.

TSM

Time-scale modification varies playback rate (both slower and faster than real time) while maintaining constant pitch.

URI

A *Uniform Resource Identifier* is a compact string of characters for identifying resources: documents, images, audio files, etc. Within a DTB, URIs are most likely to appear as attribute values for various XML elements, used as a way of identifying other documents or files either in whole or part. For the purposes of this specification, URIs must adhere to the syntax defined in RFC 2396 [\[RFC 2396\]](#). A [URI](#) may include a [fragment identifier](#) suffix beginning with "#" that matches some named anchor in the target document. See [Fragment Identifier](#).

User

See [Reader](#).

XML

The Extensible Markup Language [\[XML\]](#) is a standardized language for marking up files containing structured information.

XSL

The Extensible Stylesheet Language is a series of recommendations by the Worldwide Web Consortium that describes how [XML](#) documents can be transformed and rearranged [\[XSLT\]](#), then formatted [\[XSL\]](#) for screen, handheld device, paper, or audio presentation.

XSLT

A language for transforming [XML](#) documents into other XML documents. [\[XSLT\]](#) is designed for use as part of XSL. See [XSL](#).

1.3 Strategy

(This section is informative.)

This standard is based primarily on a variety of widely used standards and specifications, including several from the World Wide Web Consortium and the Open eBook Forum™. Wherever applicable and appropriate standards or specifications existed they were used. The use of these specifications and technologies is intended to promote a fast and consistent adoption of this standard for the target population, while encouraging its extension into mainstream use.

1.4 Accessibility Issues

(This section is informative.)

Digital Talking Book files, streams, transformation processes, and [player](#)s have been designed to present their content to people with a wide range of abilities and disabilities. They are designed to allow presentation in forms other than conventional print, due to the inaccessibility of printed documents to these [users](#). It is in the best interest of users that, to the greatest extent possible, files, streams, transformation processes, and players make information available in as many presentation modes as practical, including human-narrated audio, Braille, synthesized speech, large print with [user](#)-specifiable size and text re-wrapping for players with visual display, and text and audio synchronization and other enhancements for persons with learning disabilities. Users will also be greatly benefited if controls on players are readily usable by people with a wide range of manual dexterity.

During the development of this standard, an advisory document, [DTB Playback Device Features List](#), was created. Although it is not a [normative](#) part of this standard, player developers will find useful accessibility concepts embodied in it.

In addition to the provisions of this standard, valuable supplemental information is available from the guidelines and techniques produced by the Worldwide Web Consortium's Web Accessibility Initiative. At this time, these documents include:

[Web Content Accessibility Guidelines](#),

[Authoring Tool Accessibility Guidelines](#), and

[User Agent Accessibility Guidelines](#).

(This section is normative.)

Not all modes of presentation will be available in all players and documents, but it is strongly recommended that multiple equivalent presentations be made available to users whenever possible. Historically, products marketed to specific user groups with disabilities have sometimes proven unusable. Not all players need to be [accessible](#) to all target groups, but any device compliant with this standard [must](#) be accessible to the target group for which it is

advertised. It is also strongly recommended that DTB production tools and processes be made accessible to persons with disabilities.

1.5 Relationship to Other Specifications

(This section is normative.)

This standard is based on the specific versions of the standards and specifications referenced herein, which are used as defined except as noted by this document. Any refinement or replacement of a referenced specification by a newer or different version is not directly applicable to this standard. Conformance to this standard is based on the versions of the standards and specifications in effect at the time of this writing.

1.5.1 Relationship to Unicode

(This section is normative.)

Playback systems must support at least UTF-8 and UTF-16 encodings. See section 2.2 of the XML specification [\[XML\]](#).

1.6 Patent Rights

(This section is informative.)

Implementation of this standard may involve the use of one or more inventions covered by patent rights. It is believed that all companies claiming such rights have agreed to grant a license under such rights as they hold on reasonable and nondiscriminatory terms and conditions to any applicant.

Producers of DTB systems or any component thereof are responsible for obtaining the appropriate licenses for any and all technology they use that is defined by the relevant standards and specifications referenced by this standard. There may be applicable patents of which this standards committee is unaware; it is the responsibility of the implementer to ensure that the implementation is non-infringing.

Issues surrounding the protection of intellectual property embodied in the works distributed as digital talking books are discussed in [section 14, "Digital Rights Management"](#).

1.7 Maintenance Agency

(This section is informative.)

The maintenance agency designated in [Appendix 2](#) will be responsible for reviewing and acting upon suggestions for modifications to this standard. Questions concerning the implementation of this standard and requests for information should be sent to the maintenance agency.

A list of errata, proposed changes, and maintenance activities related to this standard will be maintained at <http://www.daisy.org/z3986/2005/errata.html>.

2 Overview

(This section is informative.)

A digital talking book (DTB) is a collection of electronic files arranged to present information to the target population via alternative media, namely, human or synthetic speech, refreshable Braille, or visual display, e.g., large print. When these files are created and assembled into a DTB in accordance with this standard, they make possible a wide range of

features such as rapid, flexible navigation; bookmarking and highlighting; keyword searching; spelling of words on demand; and user control over the presentation of selected items (e.g., footnotes, page numbers, etc.). Such features enable readers with visual and physical disabilities to access the information in DTBs flexibly and efficiently, and allow sighted users with learning or reading disabilities to receive the information through multiple senses. For a full discussion of these capabilities, see the "[Document Navigation Features List](#)" [[Navigation Features](#)], the user requirements document on which this standard was based. A document written during the development of this standard, [Theory Behind the DTBook DTD](#) [[DTBook Theory](#)], also describes the navigational capabilities of a DTB in some detail. The content of DTBs will range from audio alone, through a combination of audio, text, and images, to text alone.

DTB players will also be produced with a variety of capabilities. The simplest might be portable devices with audio-only capabilities. More complex portable players could include text-to-speech capabilities as well as audio output for recorded human speech. The most comprehensive playback systems are expected to be PC-based, supporting visual and audio output, text-to-speech capability, and output to a Braille display. The [Playback Device Features List](#) [[Player Features](#)] mentioned above presents the committee's priorities for a range of functions across three types of playback devices.

The files comprising a DTB fall into ten categories, as described below:

[Package File](#)

The Package File, drawn from the Open eBook Publication Structure 1.2, contains administrative information about the DTB and the files that comprise it. A valid [XML](#) version 1.0 file, it contains a set of metadata describing the DTB, a list (the [manifest](#)) of the files that make up the DTB, and a [spine](#) that defines the default reading order of the document. See section 3, "[Package File](#)."

[Textual Content File](#)

A DTB can contain part or all of the text of the document as an XML 1.0 file marked up in accordance with the document type definition ([DTD](#)) defined for this standard, [dtbook.dtd](#). (See [Appendix 1](#).) The textual content file enables properly-configured playback devices to spell words on demand, carry out keyword searches, and permit finely-grained navigation. It can also be accessed directly via refreshable Braille display, synthetic speech, or screen-enlarging software. See section 4, "[Content Format for Text](#)."

Audio Files

A DTB can include human or synthetic speech recordings of the document embodied in audio files encoded in one of a specified group of audio formats. Section 5, "[Audio File Formats](#)," presents the formats specified by this standard.

Image Files

In addition to text and audio, DTBs can include images that can be presented on players with visual displays. Section 6, "[Image File Formats](#)," lists the formats specified by this standard.

Synchronization Files

To synchronize the different media files of a DTB during playback, this standard specifies the use of the World Wide Web Consortium's (W3C) Synchronized Multimedia Integration Language ([SMIL](#)), [SMIL 2.0](#) version, an XML 1.0 application. The DTB SMIL files define a sequence of media events. During each event, text elements and corresponding audio clips as well as any additional visual elements are presented simultaneously. DTB players use the synchronization information to both access points in the audio presentation and to track, during audio playback, the corresponding position in the textual content file. This standard uses a subset of the full [SMIL 2.0](#) specification. See section 7, "[Synchronization of Media Files](#)," for discussion of these issues, and [Appendix 1](#) for the DTD information.

Navigation Control File

The DTB system supports two modes of navigation, global and local. [Global navigation](#) -- movement by structure (chapter, section, subsection) and by other selected points such as

pages, figures, or notes -- is effected through the Navigation Control file for XML applications ([NCX](#)). The NCX presents a dynamic view of the document's hierarchical structure, allowing the user to move through the document in large steps corresponding to its major divisions or in progressively smaller steps down to a limit set by the document's detail. Text, audio, and image elements present to the user the document's headings, and id-based links point to the SMIL presentation at the corresponding locations. [Appendix 1](#) contains information about the DTD for the NCX. Local, more finely-grained, navigation is not handled by the NCX but is enabled through the textual content file or SMIL file(s) or through time-based movement through the audio presentation, depending on the document and on the player. See section 8, "[Navigation Control File \(NCX\)](#)," and [Appendix 1](#), "NCX DTD" for specifications related to the NCX.

Bookmark/Highlight File

This standard supports user-set, exportable bookmarks and highlights to which text and audio notes can be applied. Specifications for the XML 1.0 file for portable bookmarks and highlights are presented in section 9, "[Portable Bookmarks and Highlights](#)" and [Appendix 1](#), "DTD for Portable Bookmarks/Highlights."

Resource File

The resource file contains or references various text segments, audio clips, and/or images that provide alternative representations of navigational information -- for example, feedback on the user's current location in the document. It supplies information normally presented in a print book via typographical clues. See section 10, "[Resource File](#)," and [Appendix 1](#), "DTD for Resource File" for file specifications.

Distribution Information File

Given the great size of audio files even when heavily compressed, it will be common for large books to span several [media units](#). Section 11, "[Packaging Files for Distribution](#)," describes how the "distInfo" file maps the location of each SMIL file to a specific media unit, e.g., disk 1 of 3. It also explains how, when several books are distributed on the same media unit, the distInfo file stores information about each book for presentation to the reader. [Appendix 1](#), "Distribution Information DTD," provides information on the document type definition for "distInfo" files.

Presentation Styles

Section 12, "[Presentation Styles](#)," discusses how the presentation of a DTB in various media can be controlled through the use of optional style sheets.

3 The DTB Package File

(This section is informative.)

The Package File, drawn from the Open eBook Forum™ ([OEBF](#)) Publication Structure 1.2, contains administrative information about the DTB, the files that comprise it, and how these files interrelate. This section, drawn largely from the Publication Structure, provides only a brief summary of the function of each section with an example illustrating how it is applied to the DTB. See section 2 of the full [OEBF](#) Publication Structure 1.2 for complete details on the Package File.

The Publication Structure describes the major parts of the Package File as follows:

PACKAGE IDENTITY – a unique identifier for the OEB publication as a whole.

METADATA – Publication metadata (title, author, publisher, etc.).

MANIFEST – A list of files (documents, images, style sheets, etc.) that make up the publication. The manifest also includes fallback declarations for files of types not supported by this specification.

SPINE – An arrangement of documents providing a linear reading order.

TOURS – A set of alternate reading sequences through the publication, such as selective views for various reading purposes, reader expertise levels, etc.

GUIDE – A set of references to fundamental structural features of the publication, such as table of contents, foreword, bibliography, etc.

Here is an informal outline of the package file:

```
<?xml version="1.0"?>
<!DOCTYPE package PUBLIC "-//ISBN 0-9673008-1-9//DTD OEB 1.2
Package//EN"
    "oebpkg12.dtd">

<package xmlns="http://openebook.org/namespaces/oeb-package/1.0/"
    unique-identifier="foo">
    <metadata>...</metadata>
    <manifest>...</manifest>
    <spine>...</spine>
    <tours>...</tours>
    <guide>...</guide>
</package>
```

(This section is normative.)

A DTB conforming to this standard must include exactly one Package File which must be a valid XML 1.0 document conforming to the [OEBF](#) Publication Structure 1.2 package DTD (oebpkg12.dtd) and its associated entity reference (oeb12.ent). The full specification, DTD, and entity reference for the [OEBF](#) package file are available for download from the OEBF site [\[OEBF\]](#). The Package File must be named with the extension ".opf". If a DTB spans multiple media units, the identical Package File must be present on each media unit.

A Package File conforming to this standard must comply with all aspects of section 2 of the [OEBF](#) Publication Structure 1.2, with the following two exceptions:

- 1) Section 2.3.1 does not apply. Specifically, there is no requirement on DTB authors or playback devices to implement the fallback mechanism for files that are not of the [OEBF](#) core MIME media types.
- 2) Section 2.4 of the Publication Structure states that the spine element may refer only to item elements of media type *text/x-oeb1-document*. In DTB applications, the spine must only reference items of media type *application/smil*.

Namespace (xmlns) attributes and their values, although declared as #FIXED, must be explicitly specified in the document instance. Entity declarations must occur in the internal DTD subset. See further section 16.1 ["General File Conformance Requirements."](#)

3.1 Package Identity

(This section is normative.)

The package must include a value for its unique-identifier attribute. This is required because more than one dc:Identifier may be present in a DTB's Package File metadata and the unique-identifier specifies which dc:Identifier element provides the package's primary identifier. The value of unique-identifier must match the id attribute of one and only one dc:Identifier element, which is a descendant of the package element.

The primary identifier of the DTB must be globally unique.

(This example is informative.)

Example 3.1:

```
<?xml version="1.0"?>
<!DOCTYPE package PUBLIC "-//ISBN 0-9673008-1-9//DTD OEB 1.2 Package//EN"
    "oebpkg12.dtd">
<package xmlns="http://openebook.org/namespaces/oeb-package/1.0/" unique-
  identifier="uid">
  <metadata>
    <dc-metadata...>
      <dc:Identifier id="uid" scheme="DTB">uk-rnib-db02006
    </dc:Identifier>
    ....
  </package>
```

3.2 Publication Metadata

(This section is normative.)

This portion of the Package File contains the information about a DTB that would normally be found in a library catalog record. It includes data about the DTB itself (e.g., title, author, producer, format, and narrator) as well as information about the source publication (usually a print book) such as publisher, edition, copyright statement, etc.

The Package File must contain exactly one metadata element, which must contain one and only one dc-metadata element holding Dublin Core [\[DC\]](#) metadata and must contain supplemental metadata in an x-metadata element. The x-metadata element must contain at least one instance of the meta element, which uses name and content attributes to define its value. (See section 3.2.3, "[X-Metadatas](#).")

3.2.1 Dublin Core Metadata

(This section is normative.)

The use of Dublin Core metadata within a compliant DTB must conform to the following description from the [OEBF](#) Publication Structure 1.2:

The dc-metadata element contains specific publication-level metadata as defined by the Dublin Core initiative (<http://purl.org/dc/>). The descriptions below are included for convenience, and the Dublin Core's own definitions take precedence (see <http://www.ietf.org/rfc/rfc2413.txt>).

The dc-metadata element can contain any number of instances of any Dublin Core elements. Dublin Core element names begin with the "dc:" prefix followed by a leading uppercase letter. Dublin Core metadata elements may occur in any order; in fact, multiple instances of the same element type (multiple dc:Creator elements, for example) can be interspersed with other metadata elements without change of meaning.

For upwards compatibility, the element dc-metadata in an OEB package is required to have an attribute of

```
xmlns:dc=http://purl.org/dc/elements/1.1/
and
```

```
xmlns:oebpackage="http://openebook.org/namespaces/oeb-package/1.0/".
```

Following are brief definitions of the Dublin Core elements. See the Publication Structure and the [Dublin Core](#) itself for more complete descriptions. The attributes xml:lang and id can be applied to all "dc:..." elements. Additional attributes can be used with several elements as

detailed below. Note that all Dublin Core element types may be repeated (occur more than once) within dc-metadata.

- **dc:Title**
 - o Content: The title of the DTB, including any subtitles.
 - o Occurrence: Required
- **dc:Creator**
 - o Content: Names of primary author or creator of the intellectual content of the publication.
 - o Occurrence: Optional (not all documents have known creators) - recommended.
 - o Added attributes:
 - role -- (optional) The function performed by the creator (e.g., author, editor). See Publication Structure for details on normative list of values.
 - file-as -- (optional) A normalized form of the contents suitable for machine processing.
- **dc:Subject**
 - o Content: The topic of the content of the publication.
 - o Occurrence: Optional - recommended.
- **dc:Description**
 - o Content: Plain text describing the publication's content.
 - o Occurrence: Optional
- **dc:Publisher**
 - o Content: The agency responsible for making the DTB available. (Compare dtb:sourcePublisher and dtb:producer.)
 - o Occurrence: Required
- **dc:Contributor**
 - o Content: A party whose contribution to the publication is secondary to those named in dc:Creator.
 - o Occurrence: Optional
 - o Added attributes:
 - role -- (optional) The function performed by the contributor (e.g., translator, compiler). See Publication Structure for details on normative list of values.
 - file-as -- (optional) A normalized form of the contents suitable for machine processing.
- **dc>Date**
 - o Content: Date of publication of the DTB. (Compare dtb:sourceDate and dtb:producedDate.) In format from [\[ISO8601\]](#); the syntax is YYYY[-MM[-DD]] with a mandatory 4-digit year, an optional 2-digit month, and, if the month is present, an optional 2-digit day of month.
 - o Occurrence: Required

- o Added attributes:
 - event -- (optional) Significant occurrence related to publication of the DTB. Allows repetition of dc:Date to describe, for example, multiple revisions. Best practice is to use dtb:revision and dtb:revisionDate instead.
- **dc:Type**
 - o Content: The nature of the content of the DTB (i.e., sound, text, image). Best practice is to draw from the Dublin Core's enumerated list [\[DC-Type\]](#).
 - o Occurrence: Optional
- **dc:Format**
 - o Content: The standard or specification to which the DTB was produced. Values of dc:Format in a DTB conforming to this standard are valid only if they read "ANSI/NISO Z39.86-2005".
 - o Occurrence: Required
- **dc:Identifier**
 - o Content: A string or number identifying the DTB. One instance of this element, that which is referenced from the package unique-identifier attribute, must include an id.
 - o Occurrence: Required
 - o Added attributes:
 - scheme – (optional) The name of the system or authority that generated or assigned the identifier. For example, "DOI", "ISBN", or "DTB".
- **dc:Source**
 - o Content: A reference to a resource (e.g., a print original, ebook, etc.) from which the DTB is derived. Best practice is to use the ISBN when available.
 - o Occurrence: Optional - recommended.
- **dc:Language**
 - o Content: Language of the content of the publication. An [\[RFC 3066\]](#) language code. For Sweden: "sv" or "sv-SE"; for UK: "en" or "en-GB"; for US: "en" or "en-US"; etc.
 - o Occurrence: Required
- **dc:Relation**
 - o Content: A reference to a related resource.
 - o Occurrence: Optional
- **dc:Coverage**
 - o Content: The extent or scope of the content of the resource. Not expected to be used for DTBs.
 - o Occurrence: Optional
- **dc:Rights**
 - o Content: Information about rights held in and over the DTB. (Compare dtb:sourceRights.)
 - o Occurrence: Optional

3.2.2 DTB ID Scheme

(This section is informative.)

Various schemes are available for identifying digital publications. In the DTB domain, the requirements for an identifier are simply to identify the publication in a manner that is highly likely to be globally unique. A major purpose of the uniqueness requirement is to prevent filename collisions among bookmark files.

To meet this base requirement, a simple DTB id scheme might be used. A DTB identifier under this scheme consists of a hyphen-separated string consisting of a two-letter country code drawn from [\[ISO 3166\]](#), an agency code unique within its country, and an identifier unique within the agency. For example, us-afb-x12345.

This scheme will provide a simple solution to the uniqueness requirement that will serve DTB-publishers' needs in the short term. In the longer term, as the requirements of a global library of alternative format materials become more important, other more sophisticated mechanisms will doubtless be employed.

3.2.3 X-Metadata

(This section is normative.)

The following names were developed for the DTB application to supply information that the Dublin Core element set does not cover. These names may only appear within the x-metadata containing element, as values of the name attribute on the meta element. Each x-metadata name below is shown as either "Repeatable" (it may be used more than once) or "Not repeatable". Content producers may introduce other metadata within x-metadata besides those listed below, if needed. However, metadata names shall not begin with the prefix "dtb:" unless defined in this standard. Players must not fail when encountering unknown metadata but must, at a minimum, ignore it.

- **dtb:sourceDate**
 - o Content: Date of publication of the resource (e.g., a print original, ebook, etc.) from which the DTB is derived. In format from [\[ISO8601\]](#); the syntax is YYYY[-MM[-DD]] with a mandatory 4-digit year, an optional 2-digit month, and, if the month is present, an optional 2-digit day of month.
 - o Occurrence: Optional - recommended. Not repeatable.
- **dtb:sourceEdition**
 - o Content: A string describing the edition of the resource (e.g., a print original, ebook, etc.) from which the DTB is derived.
 - o Occurrence: Optional - recommended. Not repeatable.
- **dtb:sourcePublisher**
 - o Content: The agency responsible for making available the resource (e.g., a print original, ebook, etc.) from which the DTB is derived. (Compare dc:Publisher.)
 - o Occurrence: Optional - recommended. Not repeatable.
- **dtb:sourceRights**
 - o Content: Information about rights held in and over the resource (e.g., a print original, ebook, etc.) from which the DTB is derived. (Compare dc:Rights.)
 - o Occurrence: Optional - recommended. Not repeatable.

- **dtb:sourceTitle**
 - o Content: The title of the resource (e.g., a print original, ebook, etc.) from which the DTB is derived. To be used only if different from `dc:Title`.
 - o Occurrence: Optional. Not repeatable.
- **dtb:multimediaType**
 - o Content: One of the six types of DTB defined in the Structure Guidelines [\[StructGuide\]](#). Values are: `audioOnly`, `audioNCX`, `audioPartText`, `audioFullText`, `textPartAudio`, `textNCX`.
 - o Occurrence: Required. Not repeatable.
- **dtb:multimediaContent**
 - o Content: Summary of the general types of media used in the content of this DTB. Value is a comma-delimited list of the top-level media types defined in [\[RFC2046\]](#). In the current version of this standard, the only applicable types will be `audio`, `text`, and `image`. Media types that are referenced only by NCX, Resource File, or `distInfo` must not be listed here.
 - o Occurrence: Required. Not repeatable.
- **dtb:narrator**
 - o Content: Name of the person whose recorded voice is embodied in the DTB.
 - o Occurrence: Optional - recommended. Repeatable.
- **dtb:producer**
 - o Content: Name of the organization/production unit that created the DTB. (Compare `dc:Publisher`.)
 - o Occurrence: Optional. Repeatable.
- **dtb:producedDate**
 - o Content: Date of first generation of the complete DTB, i.e. Production completion date. (Compare `dc:Date`.) In format from [\[ISO8601\]](#); the syntax is `YYYY[-MM[-DD]]` with a mandatory 4-digit year, an optional 2-digit month, and, if the month is present, an optional 2-digit day of month.
 - o Occurrence: Optional. Not repeatable.
- **dtb:revision**
 - o Content: Non-negative integer value of the specific version of the DTB. Incremented each time the DTB is revised.
 - o Occurrence: Optional. Not repeatable.
- **dtb:revisionDate**
 - o Content: Date associated with the specific `dtb:revision`. In format from [\[ISO8601\]](#); the syntax is `YYYY[-MM[-DD]]` with a mandatory 4-digit year, an optional 2-digit month, and, if the month is present, an optional 2-digit day of month.
 - o Occurrence: Optional. Not repeatable.
- **dtb:revisionDescription**
 - o Content: A string describing the changes introduced in a specific `dtb:revision`.
 - o Occurrence: Optional. Not repeatable.

- **dtb:totalTime**
 - o Content: Total playing time of all SMIL files comprising the content of the DTB. Value is a Clock Value from SMIL 2.0 Timing and Synchronization Module [\[SMIL\]](#). See section 7.7, "[Media Clipping and Clock Values](#)."
 - o Occurrence: Required. Not repeatable.
- **dtb:audioFormat**
 - o Content: A string describing the format in which the audio files in the DTB file set are written. If more than one audio format is used, this element may be repeated.
 - o Occurrence: Optional, recommended for audio DTBs. Repeatable.
 - o Formats specified in [section 5](#) are shown below, followed by the normative value for dtb:audioFormat:
 - o [MPEG-4 AAC](#): MP4-AAC
 - o [MPEG-1/2 Layer III](#): MP3
 - o Linear PCM,RIFF WAVE format: WAV
 - o (Values are not case-sensitive.)

(This example is informative.)

Example 3.2:

```
<metadata>
  <dc-metadata xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:oebpackage="http://openbook.org/namespaces/oeb-package/1.0/">
    <dc>Title>Revised Standards and Guidelines of
      Service for the Library of Congress Network of Libraries for the
      Blind and Physically Handicapped 1995</dc>Title>
    <dc:Subject>library information networks</dc:Subject>
    <dc:Subject>libraries and the physically
      handicapped--standards--U.S.</dc:Subject>
    <dc:Subject>libraries and the
      blind--standards--U.S.</dc:Subject>
    <dc:Identifier id="uid"
      scheme="DTB">us-nls-db00001</dc:Identifier>
    <dc:Identifier
      scheme="DOI">10.1000/DX44998</dc:Identifier>
    <dc:Creator role="aut">American Library Association.
      Association of Specialized and Cooperative Library
      Agencies</dc:Creator>
    <dc:Publisher>National Library Service for the Blind
      and Physically Handicapped, Library of Congress</dc:Publisher>
    <dc>Date>2000-06-22</dc>Date>
    <dc:Source>0-8389-7797-9</dc:Source>
    <dc:Language>en</dc:Language>
    <dc:Format>ANSI/NISO Z39.86-2005</dc:Format>
    <dc>Description>A document developed to improve library
      service for blind and physically disabled persons by providing a
      tool for assessing the current status of those services
      and for developing long-range plans.</dc>Description>
  </dc-metadata>
  <x-metadata>
    <meta name="dtb:sourceDate" content="1995" />
    <meta name="dtb:sourcePublisher">
```

```

        content="American Library Association" />
        <meta name="dtb:sourceRights" content="copyright 1995,
        American Library Association" />
        <meta name="dtb:narrator" content="Lowenstein, Ralph" />
        <meta name="dtb:producer"
        content="American Foundation for the Blind" />
        <meta name="dtb:multipediaContent" content="audio" />
        <meta name="dtb:multipediaType" content="audioNCX" />
        <meta name="dtb:totalTime" content="06:22:34.143" />
    </x-metadata>

</metadata>
....

```

3.3 Manifest

(This section is normative.)

The manifest, which is a child of the package element, must contain a complete list of all of the files (documents, audio files, images, style sheets, etc.) that make up a given DTB, including the package file itself. The manifest shall list only files of types permitted by this standard. The manifest shall list only files that are part of the DTB. The distInfo file and any associated audio changeMsgs (or any other files listed *only* in the distInfo file) are not considered part of the DTB and thus shall not be listed (See section 11, "[Packaging Files for Distribution](#).") The Resource File and any associated media files are considered part of the DTB and thus shall be listed. Each file is referenced by an item element. Each item must have an href attribute that is the [URI](#) of the referenced file and is unique within the manifest. This URI must not include [fragment identifiers](#); if relative, it is interpreted as relative to the package file itself. Further, any relative URIs contained within an XML file listed in the manifest are considered to be relative to the referring file.

In addition, each item must have a media-type attribute containing the MIME media type of the file, and an id attribute. The id is used primarily when a manifest item is referenced by the spine. The manifest may also include fallback declarations to allow players to choose among alternative presentation formats. (See [OEBF Publication Structure](#) for details.) Support for the fallback mechanism is not required by this standard. The NCX entry in the Package File manifest must have an id value equal to "ncx". The Resource File entry in the Package File manifest must have an id value equal to "resource". The order of item elements within the manifest is not significant.

All media-type attribute values must conform to [RFC 2046](#). In addition, the media-type attribute must have the value of the IANA-registered MIME media type for that type of file, if one exists (see [RFC 2048](#) for information on the registration process for MIME media types). If no MIME media type has been registered for the file type, then the MIME media type recommended in the applicable standard for that file type must be used. If there is no standard MIME media type for the file, then an appropriate "x-" MIME media type must be used, following the rules of [RFC 2046](#). In the case of WAV files, this standard mandates the use of the media type value "audio/x-wav". For files defined by this standard that are XML documents, this standard defines mandatory, unique media types following [RFC 3023](#): "application/x-dtbnx+xml" for the NCX, "application/x-dtbresource+xml" for the resource file, and "application/x-dtbook+xml" for textual content files.

In addition, producers must use particular file name extensions for all the different kinds of files mentioned in this standard (see table below). These values are case-sensitive. In cases where the file name extension and MIME media type values do not agree, players should consider the MIME media type to take precedence.

The following table summarizes the required file name extensions and MIME media type values for all the different kinds of files that may appear in the package manifest:

Kind of file	File name extension	MIME media type
MPEG-4 AAC audio	.mp4	audio/mpeg4-generic
MPEG-1/2 Layer III (MP3) audio	.mp3	audio/mpeg
Linear PCM - RIFF WAVE format audio	.wav	audio/x-wav
JPEG image	.jpg	image/jpeg
PNG image	.png	image/png
Scalable Vector Graphics (SVG) image	.svg	image/svg+xml
Cascading Style Sheets (CSS)	.css	text/css
SMIL files	.smil	application/smil
Package file	.opf	text/xml
DTD and DTD fragments (entities or modules)	[no requirement]	application/xml-dtd
Navigation Control File (NCX)	.ncx	application/x-dtbnx+xml
Textual content files (dtbook)	.xml	application/x-dtbook+xml
Resource file	.res	application/x-dtbresource+xml

(This example is informative.)

A sample manifest for a DTB with audio, structure, and text follows:

Example 3.3:

```

<manifest>

  <item id="opf" href="rs.opf" media-type="text/xml" />
  <item id="text" href="rs.xml" media-type="text/x-dtbook+xml" />
  <item id="text_style" href="dtbbase.css" media-type="text/css" />
  <item id="ncx" href="rs.ncx" media-type="application/x-dtbnx+xml"
/>
  <item id="ncx_style" href="ncx16.css" media-type="text/css" />
  <item id="SMIL" href="rs.smil" media-type="application/smil" />
  <item id="foreword" href="rs_fwdx.mp3" media-type="audio/mpeg" />
  <item id="standards" href="rs_std.mp3" media-type="audio/mpeg" />
  <item id="appendices" href="rs_app.mp3" media-type="audio/mpeg" />
  <item id="index" href="rs_index.mp3" media-type="audio/mpeg" />
  <item id="fig_01" href="fig1.png" media-type="image/png" />
  <item id="resource" href="rs.res" media-type="application/x-
dtbresource+xml" />

```

```

        <item id="resource_audio" href="res.mp3" media-type="audio/mpeg"
    />

    </manifest>
    ....

```

Here is a manifest for an audio-only version of the above DTB where separate SMIL files were created for each segment of the book.

Example 3.4:

```

    ....
    <manifest>
        <item id="opf" href="rs.opf" media-type="text/xml" />
        <item id="ncx" href="rs.ncx" media-type="application/x-dtbncx+xml"
    />
        <item id="foreword" href="rs_fwdx.mp3" media-type="audio/mpeg" />
        <item id="standards" href="rs_std.mp3" media-type="audio/mpeg" />
        <item id="appendices" href="rs_app.mp3" media-type="audio/mpeg" />
        <item id="index" href="rs_index.mp3" media-type="audio/mpeg" />
        <item id="SMIL1" href="rsfwd.smil" media-type="application/smil"
    />
        <item id="SMIL3" href="rsapp.smil" media-type="application/smil"
    />
        <item id="SMIL4" href="rsind.smil" media-type="application/smil"
    />
        <item id="SMIL2" href="rsstd.smil" media-type="application/smil"
    />

    </manifest>
    ....

```

3.3.1 Allowed Characters in File Names

(This section is normative.)

To assure interoperability when transporting DTBs between various locales and platforms, characters in file names must be restricted to the *alphanum* definition of [RFC2396](#), and a subset of the RFC2396 *mark* definition. This effectively reduces allowed characters to [A-Za-z0-9._-].

The following definition "filename" identifies in EBNF form which characters are allowed in file names. The definition "foldername" identifies allowed characters in folder names, when folders are used in path specifications of URIs that reference members of the DTB file set.

```

digit ::= [#x0030-#x0039]
lowalpha ::= [#x0061-#x007A]
upalpha ::= [#x0041-#x005A]
alphanum ::= digit|lowalpha|upalpha
hyphen ::= [#x002D]
underscore ::= [#x005F]
period ::= [#x002E]
filename ::= (alphanum|hyphen|underscore|period)+
foldername ::= (alphanum|hyphen|underscore)+

```