

# Creating NIMAS Files

By AEM Center at CAST

Published: 2021

The contents of this document were developed under a cooperative agreement with the US Department of Education, #H327Z190004. However, those contents do not necessarily represent the policy of the US Department of Education and you should not assume endorsement by the Federal Government. Project Officer, Rebecca Sheffield, Ph.D.



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International license.  
<http://creativecommons.org/licenses/by-sa/4.0/>



National Center on Accessible Educational Materials (2021). *Creating NIMAS Files*. Wakefield, MA: National Center on Accessible Educational Materials.

## **Table of Contents**

|   |    |
|---|----|
| 1. The NIMAS Fileset .....  | 3  |
| 2. Creating a NIMAS-conformant XML Content File.....  | 4  |
| NIMAS 1.1 DTD Required Elements.....  | 4  |
| NIMAS 1.1 DTD Required Attributes.....  | 4  |
| DTD and OPF References.....   | 5  |
| 3. Preparing Images for a NIMAS Fileset.....  | 5  |
| Image files .....   | 5  |
| Organization of images .....  | 5  |
| Examples:.....  | 6  |
| Value-added components: mark-up of images using previously-published (print)<br>materials ..... | 7  |
| Example: .....  | 8  |
| Value-added components: writing alternative text and long descriptions.....                     | 8  |
| Markup of math notation with MathML .....   | 9  |
| 4. Preparing a Package File (OPF) .....   | 9  |
| NIMAS filesets and MIME types.....  | 10 |

# 1. The NIMAS Fileset

A NIMAS fileset consists of the following:

- an XML content file,
- a package file (OPF),
- a PDF-format copy of title page *and* ISBN and copyright page(s), and
- a full set of content images in SVG, JPG, or PNG format.

NIMAS is a subset of DAISY and files must validate to the DAISY 2005-3 DTD. All fileset components must be provided in a zip file. The naming convention for the NIMAS zip file is the Identifier. (See [NIMAC Metadata Guidelines](#) for more information.)

A print work's **XML content file** includes the complete non-graphic content of the textbook, including the frontmatter, such as title page information, tables of contents and prefaces; backmatter, such as epilogues and indices; front and back covers, if they include essential bibliographic information or other content that is not provided elsewhere in the book; and bodymatter, such as units, chapters and sections, text, charts, tables, etc. The XML must also include image references wherever graphics appear in the text.

The **OPF file** for a file set must include required metadata for the textbook, as well as a manifest section which lists all of the files supplied in the NIMAS zip file except the OPF file itself. When files are uploaded to the NIMAC system, metadata provided in the OPF is automatically extracted to create the system record for the title. For more information about metadata requirements for NIMAS, please see [NIMAC Metadata Guidelines](#).

The **PDF supplied in the NIMAS file set** is not the complete book, but only the pages and/or covers where essential bibliographic metadata is found. The PDF is not used in the production of accessible formats and does not have to be a production quality PDF. The file is used for file set verification in the NIMAC file certification process and can also be used by NIMAC users to confirm which file set they have downloaded.

Content which is supplied as an image in the textbook and is not captured as text in the XML file must be supplied as images in JPG, PNG, or SVG formats. If images are supplied as PNG or JPG, they must be a minimum of 300 dpi.

Fallback images must also be provided for math equations or other content provided as MathML in the NIMAS XML file. These images must also be provided as JPG, PNG or SVG and meet the 300 dpi requirement.

## **2. Creating a NIMAS-conformant XML Content File**

NIMAS XML files are required to use the Baseline Element Set found in the NIMAS Technical Specification and files must validate to the [DAISY DTD](#) (DAISY 2005-3). However, this is not the only requirement for valid NIMAS. Content in the print book must be tagged in accordance with the [DAISY Structure Guidelines](#). It is recommended that all publishers and vendors review the DAISY Structure Guidelines and refer to this guidance as needed when preparing NIMAS files.

The NIMAC provides a Validation Wizard as a stand-alone Windows program that provides DTD validation for NIMAS file sets. It also ensures that the XML is well-formed and that all referenced images are present in the images folder. NIMAC requires that all NIMAS files be validated using the tool before they are submitted to the NIMAC. The tool is available as a free download from the NIMAC system to all publishers and conversion vendors.

### **NIMAS 1.1 DTD Required Elements**

Following is a list of elements that are required (taken from the NIMAS Technical Specification)—other elements may be required if applicable to content.

| <b>required elements</b> | <b>description</b>   |
|--------------------------|--|
| <dtbook>                 | The root element in a Digital Talking Book DTD. For a NIMAS file, <dtbook> contains an empty <head> and the contents itself in <book>.     |
| <head>                   | Should be empty for a NIMAS file.  |
| <book>                   | Surrounds the actual content of the document, which is divided into <frontmatter>, <bodymatter>, and <rearmatter>. <head> precedes <book>. |

### **NIMAS 1.1 DTD Required Attributes**

Following is a list of attributes that are required if the noted elements are used.

| <b>required attributes</b>         | <b>examples</b>   |
|------------------------------------|---|
| version on <dtbook>                | <dtbook version="2005-3">   |
| src and alt on <img>               |  |
| type on <list>                     | <list type="ol">  |
| render on <prodnote> and <sidebar> | <prodnote render="optional">                                      |

| required attributes                       | examples                   |
|---|----------------------------|
| id on <pagenum>, <note>, and <annotation> | <note id="U01C04.002">     |
| idref on <noteref>, <annoref>             | <noteref idref="0114-012"> |

## **DTD and OPF References**

The [DAISY DTD](#) (DAISY 2005-3) should be used as the DTD reference for all NIMAS files. Files that include MathML should also include the appropriate DTD reference for the version of MathML supplied.

The DTD that should be referenced for all NIMAS OPF files is "http://openebook.org/dtds/oeb-1.2/oebpkg12.dtd." Further information regarding the OPF format can be found at [Open eBook Publication Structure 1.2 package file specification](#).

## **3. Preparing Images for a NIMAS Fileset**

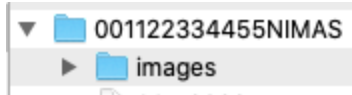
### **Image files**

- All of the images included within a work should be provided using a folder structure within the images folder of the NIMAS fileset.
- XML content files must include an "img" element corresponding to each use of an image in the location found in the textbook.
- "src" attributes of img tags should contain a reference to the appropriate image's filename. This reference is a relative pathname (example: 
- Preferred image type is PNG, JPG or SVG format
- PNG and JPG images should be provided at a resolution of 300 dots per inch (dpi), based on their actual size in the printed work, so a 2" x 1" image would be 600 x 300 pixels.
- Text should not be supplied only as an image file.

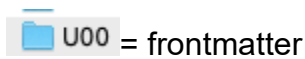
### **Organization of images**

To simplify images organization and to recommend an efficient way to handle images for use with NIMAS filesets, the following is recommended:

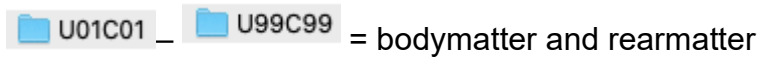
All images should be saved in an images folder:



Within this parent folder, images should be saved as follows:



The zeroes allow for sequential ordering.



Naming here provides information about image location to unit and chapter level.



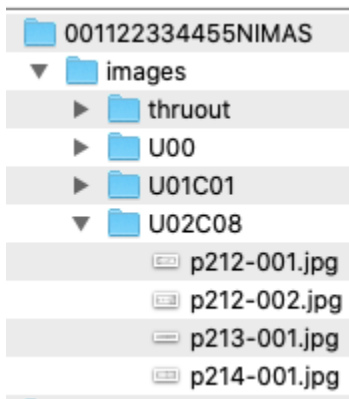
Naming here provides for the fact that many works contain images that recur throughout content at all levels.

### Examples:

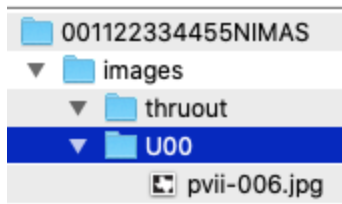
An icon that occurs several times in frontmatter and twice per chapter in bodymatter:



A folder of images occurring in the middle of a work:\*



An image that occurs in the frontmatter of a work:



\*Note that “U02C08” subsequent to “U01C01” in the second example above illustrates the fact that not all pages/chapters/units of all works will have images. The example demonstrates a numerical sequence where not every step in a possible sequence is necessary.

## **Value-added components: mark-up of images using previously-published (print) materials**

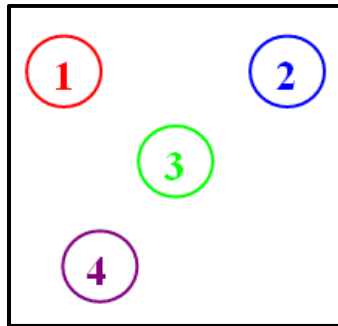
When marking up images from a previously-published, print-based source, it is recommended that images be named according to their placement within original source content. Create filenames based on an image’s page location and sequential order.

Example: p004-002.jpg. An image with this filename appears on page 4 of the original work and is the second image on that page.

Do not use spaces in image filenames.

Choose an image’s sequential order number according to its position: name image files from top to bottom and from left to right.

## Example:



Occasionally it will not be entirely obvious which layout position an image holds. In such cases, simply choose a logical sequence number and make a note of it for production.

## Value-added components: writing alternative text and long descriptions

Alt text and long descriptions for images provide an accessibility alternative for individuals who cannot access image content visually. Assistive technology such as screen-readers will describe the image by reading the text in the alt attribute. The specification requires that the alt text attribute be present on the img tag.(i.e., `<img alt=""/>`), and including actual alt text is strongly recommended to publishers, as is including long descriptions, when appropriate.

Alt text is a brief description of an image and should be no longer than 1-2 sentences; it should state the type of image and provide a brief summary. The goal is to capture the key information conveyed in the image and to express it in terms that make sense in the context of the text and in language that is appropriate for the grade level.

More complex images may require long descriptions. There is no limit to the length of long descriptions. They are context-specific and the details provided will depend on how the image supports or supplements the text. The `<prodnote>` element is used to provide long descriptions in NIMAS files. Remember that long descriptions vary according to learning goals. Try to create a balance between brevity and sufficient information so that every learner can access key content.

Many resources are available online that provide guidance on writing good image descriptions:

- [DIAGRAM Center: Image Description Guidelines](#)
- [Stanford Online Accessibility Program: Image Descriptions](#)



- [Perkins Paths to Literacy: How to Write Alt Text and Image Descriptions for the Visually Impaired](#)
- [Harvard University Digital Accessibility: Write Good Alt Text to Describe Images](#)

## **Markup of math notation with MathML**

A 2012 Dear Colleague letter from the Office of Special Education and Rehabilitative Services (OSERS) to all state directors of Special Education strongly recommended requiring MathML as a part of the NIMAS language in textbook adoption contracts and purchase agreements. MathML3 is a viable and robust technical standard for the inclusion of mathematical content in K-12 instructional materials. This specification significantly facilitates the creation of valid XML files containing mathematical content, which in turn will allow these files to be transformed into student-ready versions—such as digital text and braille—that are far more accessible to automated conversion processes than image-based math content in NIMAS filesets.

For additional information about MathML, see the following:

- [MathML in NIMAS](#)
- [NIMAS Exemplar 12](#)

## **4. Preparing a Package File (OPF)**

Package files created for use with NIMAS XML files must conform to the OEBPS 1.2 standard (the Open eBook Publication Structure Specification Version 1.2).

The current OEBPS 1.2 declaration for a NIMAS OPF file is as follows:

```
<!DOCTYPE package PUBLIC "+//ISBN 0-9673008-1-9//DTD OEB 1.2 Package//EN"
"http://openebook.org/dtds/oeb-1.2/oebpkg12.dtd">
```

To create valid package files, keep the following in mind:

- A package file should have metadata, manifest and spine sections.
- The package file must include a manifest that lists of all files contained in the NIMAS file set, except for the OPF file itself.
- Package files must have a unique-identifier attribute (example: <package unique-identifier="NIMAS9782222156542">).
- Please note that this convention is the opposite of the NIMAS file set Identifier, which must show "NIMAS" after the ISBN: <dc:Identifier scheme="NIMAS">9782222156542NIMAS</dc:Identifier>)

## NIMAS filesets and MIME types

Manifest list items in OPF files should include standard MIME types in order to be fully accurate. For example, `<item id="idname" href="filename.xml" media-type="text/xml"/>` where text in `media-type="text/xml"` is a standard MIME type and xml in `media-type="text/xml"` is a standard sub-type. Another example is `<item id="idname" href="path/path/filename.jpeg" media-type="image/jpeg"/>` where image in `media-type="image/jpeg"` is a standard MIME type and jpeg in `media-type="image/jpeg"` is a standard sub-type.

MIME types for use with NIMAS filesets reflect the NIMAS' alignment with DAISY and follow the [guidelines of Z39.86-2005](#) (section 3.3). They are as follows:

- **XML content files: `media-type="application/x-dtbook+xml"`**

A full reference/OPF manifest item would read as follows:

```
<item id="xmlexemplar" href="contentfilename.xml" media-type="application/x-dtbook+xml"/>
```

- **PDF files: `media-type="application/pdf"`**

A full reference/OPF manifest item would read as follows:

```
<item id="copyrightpagepdf" href="copyrightpage.pdf" media-type="application/pdf"/>
```

- **images: `media-type="image/jpeg"` or `media-type="image/svg+xml"` or `media-type="image/png"`**

A full reference/OPF manifest item would read as follows:

```
<item id="imgarrow" href="images/U00C00/arrow.jpg" media-type="image/jpeg"/>
```

*or*

```
<item id="imgarrow" href="images/U00C00/arrow.svg" media-type="image/svg+xml"/>
```

*or*

```
<item id="imgarrow" href="images/U00C00/arrow.png" media-type="image/png"/>
```

NOTE: The DAISY DTD on which the NIMAS technical specification is based does not include all possible MIME media types and sub-types because official types are not a static, permanent list but evolve through the coordination of the Internet Assigned Numbers Authority (IANA [<http://www.iana.org>]).