

In a case of non-periodical, documents are stored according to an Olive unique code in a three layered hierarchal fashion (similar to that of the periodical)

- The third level in the hierarchy contains the following information packaged in a zip file:
 - A Table of Contents XML file named TOC.xml
 - A Font Style Gallery file named StyleGallery.txt
 - A copy of the source PDF file, given the name of the original file
 - Per page and IMG directories are identical to the structure in the previously mentioned periodical repository

3.3 Standards

File Formats

The PrXML ~~image~~ repository is a coherent set of files of different formats. These currently consist of XML and XSL files as well as of corresponding media files such as PDF, GIF, JPG, PNG etc. that are used in conjunction with the XML files (see also Chapter 3.5 “Auxiliary Data”). All files are generated according to the respective format standards.

XSL Language

XSL scripts allow for automatic transformation of Olive XML content into formats like HTML, ASCII, WAP and more. These files vary according to application demands, and may allow either batch conversion of XML or on-the-fly content display in various formats and customized views. The XSL language used in conversion of PrXML XML is XSLT. XSL is currently supported by MS Internet Explorer, as well as by a majority of XML processors. XSLT is becoming a de-facto standard for the XML industry.

3.4 Auxiliary Files and Data

3.4.1 Images

Images play an important role in the PrXML Repository. Page image snapshots for example, are created and used not only to preserve and present historical scanned documents but also to preserve and present the original view of new documents, minimizing the dependency on special font's availability at the client side.

Olive can store every image in a document in a multiple resolutions and formats. By default all page snapshot images are stored as 256-color GIF or PNG image format of various resolutions, while color photos are stored in JPEG format with various compression levels.

The decision on how images are stored is based on a smart algorithm that calculates the best method of generating images based on the image