FITS – The File Information Tool Set

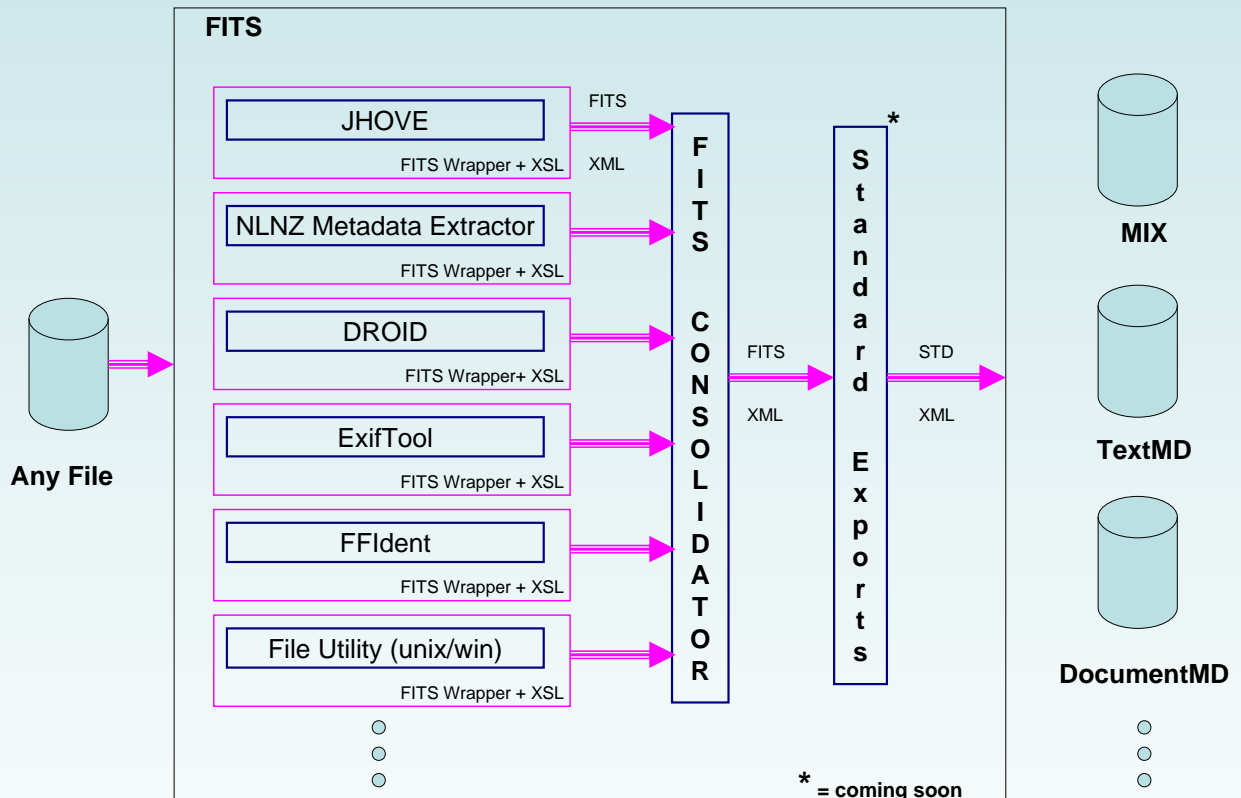## File Characterization and Metadata Extraction

Why:
- Digital preservation requires the best possible technical metadata for files so that preservation monitoring can be performed and actions taken
- Automated format identification and metadata extraction reduces the amount of metadata that must be created manually and takes full advantage of the metadata embedded within files
- No single file analysis tool is best for all file types and metadata extraction tasks

What:
- An open source Java toolkit
- A wrapper for any file tool that can run on your host
- Produces canonical FITS-XML
- Includes Java crosswalks to MIX, TextMD, DocumentMD, and will soon include Audio
- Is extensible

How:

Features:
- Conflict management (configuration, prioritization, error handling if a tool fails)
- Value normalization (e.g. "inches" vs "2")
- Tool order preference
- Format specific tool exclusion
- Format tree for matching more specific format identities

More info:

FITS is part of the second generation **Harvard University Library Digital Repository Service** (DRS2), which supports content models and METS/PREMIS object descriptors. An overview of DRS2 can be found at http://hul.harvard.edu/ois/systems/drs/enhancements.html

For more information on FITS, see http://code.google.com/p/FITS , or contact Spencer McEwen at spencer_mcewen@harvard.edu. FITS currently integrates with the following tools:
- JHOVE - http://hul.harvard.edu/jhove/
- National Library of New Zealand Metadata Extractor - http://meta-extractor.sourceforge.net/
- DROID - http://droid.sourceforge.net/wiki/index.php/Introduction
- ExifTool - http://www.sno.phy.queensu.ca/~phil/exiftool/
- FFIdent - http://www.devlib.org/ffident/index.html
- File Utility - http://unixhelp.ed.ac.uk/CGI/man-cgi?file