

# Report on the Symposium for the Adoption of JPEG 2000 in Archives and Libraries

<http://j2kArcLib.info/>

**Peter E. Murray**

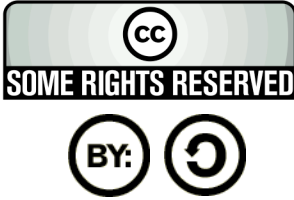
University of Connecticut Libraries  
369 Fairfield Rd, Unit 2005-A  
Storrs, CT 06095-2005  
Phone: 860-486-6771  
Peter.Murray@uconn.edu  
<http://www.lib.uconn.edu/>

*After December 17, 2004*

OhioLINK  
2455 North Star Rd, Suite 300  
Columbus, OH 43221  
Phone: 614-728-3600 x388  
peter@OhioLINK.edu  
<http://www.ohiolink.edu/>

<http://www.pandc.org/peter/presentations/cni-fall2004-j2karclib/>

December 6, 2004  
Version 1.0.1



Copyright © 2004 by Peter E. Murray; some rights reserved

All original material in this work is licensed under the Creative Commons Attribution -ShareAlike License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/2.0/> or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

```
<rdf:RDF xmlns="http://web.resource.org/cc/"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
<Work rdf:about="http://www.pandc.org/peter/presentations/jpeg2000-cni-fall2004/">
  <dc:title>Report on the Symposium for the Adoption of JPEG 2000 in Archives and
Libraries</dc:title>
  <dc:description>Presentation slides from a briefing on the outcomes of the Sym-
posium on the Adoption of JPEG 2000 in Archives and Libraries. Presented at the
CNI Fall 2004 Task Force meeting on December 6, 2004, Portland, Oregon.
</dc:description>
  <dc:creator><Agent>
    <dc:title>Peter E. Murray</dc:title>
  </Agent></dc:creator>
  <dc:rights><Agent>
    <dc:title>Peter E. Murray</dc:title>
  </Agent></dc:rights>
  <dc:date>2004</dc:date>
  <dc:format>application/pdf</dc:format>
  <dc:type rdf:resource="http://purl.org/dc/dcmitype/Text" />
  <license rdf:resource="http://creativecommons.org/licenses/by-sa/2.0/" />
</Work>

<License rdf:about="http://creativecommons.org/licenses/by-sa/2.0/">
  <permits rdf:resource="http://web.resource.org/cc/Reproduction" />
  <permits rdf:resource="http://web.resource.org/cc/Distribution" />
  <requires rdf:resource="http://web.resource.org/cc/Notice" />
  <requires rdf:resource="http://web.resource.org/cc/Attribution" />
  <permits rdf:resource="http://web.resource.org/cc/CommercialUse" />
  <permits rdf:resource="http://web.resource.org/cc/DerivativeWorks" />
  <requires rdf:resource="http://web.resource.org/cc/ShareAlike" />
</License>

</rdf:RDF>
```

# Report on the Symposium for the Adoption of JPEG 2000 in Archives and Libraries

<b>What is JPEG 2000?</b> .....	<b>1</b>
Reasons for Creation.....	1
<b>Impact on Libraries and Archives</b> .....	<b>2</b>
JPEG 2000 as Evolutionary.....	2
Open Standard .....	2
Replacement for TIFF, JPEG and SID .....	2
JPEG 2000 as Revolutionary.....	2
Bundling of Metadata with Image.....	2
Catalyst for Advancement of Imaging Practice .....	3
<b>Coordination and Acceleration of Adoption</b> .....	<b>3</b>
Education and Persuasion .....	3
Information Required.....	3
Conditions Required to Ensure Adoption.....	3
Questions from Our Communities .....	4
Inquiry and Research.....	4
Coordination of Derivative Standards .....	4
Creation of Usage Profiles.....	4
Effective use of Metadata Boxes.....	4
Interoperability Testbeds .....	4
Models for Migration .....	4
Creation of Tools.....	5
Tools for Practitioners.....	5
Tools for Users .....	5
Viewer Attributes of Interest to the Library and Archive Communities .....	5
Tool Architecture .....	5
Tools for Agents.....	5
<b>Status of Efforts</b> .....	<b>6</b>
<b>Acknowledgements</b> .....	<b>6</b>



# What is JPEG 2000?

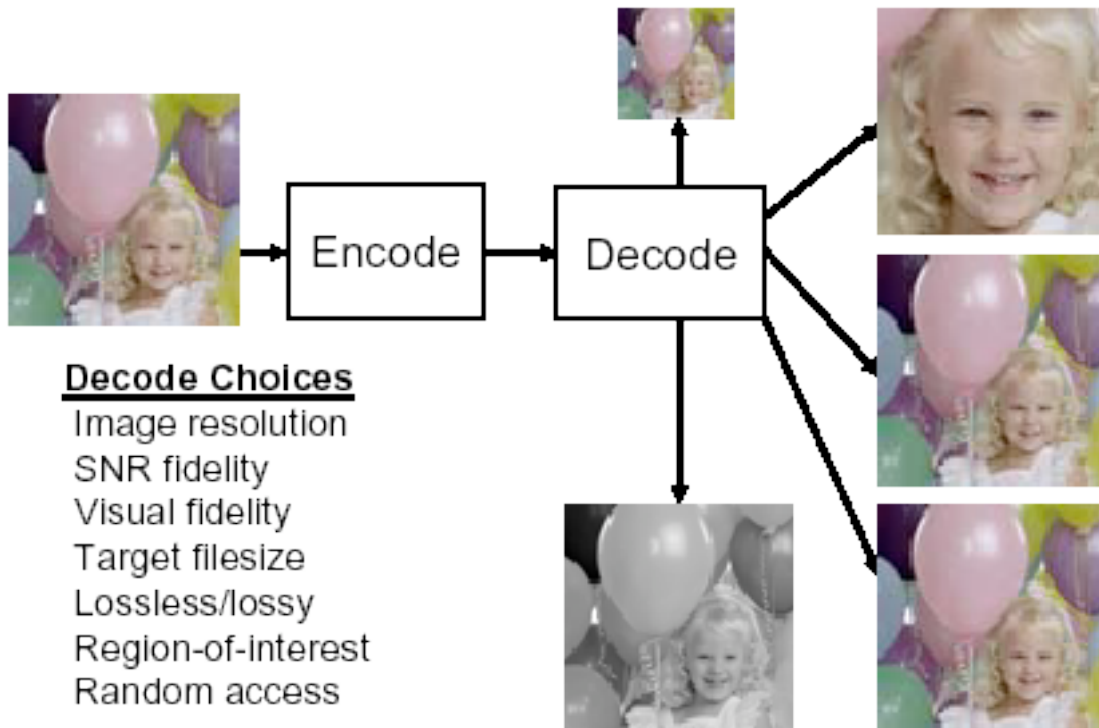
JPEG 2000 is a standard for image compression that defines "a set of lossless (bit-preserving) and lossy compression methods for coding continuous-tone, bi-level, grey-scale, or colour images."

- Named for Joint Photographic Experts Group and adoption year of the first part of the standard
- Under the auspices of the International Standards Organization and the standards section of the International Telecommunication Union, the Joint Photographic Experts Group

## Reasons for Creation

From the committee:

- Better compression than JPEG
- One algorithm for both lossless and lossy compression
- Supports multiple decompression options
  - Image resolution
  - SNR fidelity
  - Visual fidelity
  - Target file size
  - Region-of-interest
  - Random access
- One asset supports multiple derivatives
- Can specify exact compression ratio or bit rate
- Can process images in compressed form
- Allows Region-of-Interest (ROI) Coding
- Easily handles large images and high bit-depth images
- Defines file formats with generous metadata support



## Impact on Libraries and Archives

On November 4-5, 2004, the University convened the Symposium on the Adoption of JPEG 2000 by Archives and Libraries. Invited were:

- Policy makers and practitioners
- Digital imaging specialists and software engineers
- Image/signal processing scientists and end users of the standards
- Vendors and users of products

## JPEG 2000 as Evolutionary

### Open Standard

- Implementations of specification are (assumed to be) free from intellectual property claims
- Encourages competitive creation of products while maintaining interoperability
- Software can be created/maintained on new hardware or after commercial advantage exhausted
- *TIFF*: Open standard, but compression algorithms are not
- *SID*: Wavelet compression, but in a patented form

### Replacement for TIFF, JPEG and SID

- True bit-for-bit lossless compression; also includes adjustable lossy compression
- Multiple resolutions of an image stored in the same file; multiple "page" images in the same file
- *JPEG*: Compression algorithm sacrifices image detail, not as efficient
- *TIFF*: Many competing compression schemes, most popular is subject to royalties
- *SID*: Non-proprietary file format and compression algorithm

## JPEG 2000 as Revolutionary

### Bundling of Metadata with Image

- Basic file format includes metadata boxes ("UUID" for arbitrary data, or XML)
- Permanently associate metadata with image(s); serve as an interchange format
- Reduced reliance on complex systems; increased confidence in preservation

## Catalyst for Advancement of Imaging Practice

- TIFF's origins from desktop publishing and pre-press workflows in 1980s
- Image processing evolved with signal processing; represents images as mathematical algorithms
- Reduction of "noise" in image "signal" is key to good image capture
- Noise from hardware, software, and process for capturing image (e.g., misalignment of lenses, proper lighting, construction of sensors)

## Coordination and Acceleration of Adoption

- Open standard of growing importance to other vertical markets
- Improved -- and lossless -- compression
- Reduce system complexity: eliminate separate thumbnails, imbed metadata in image
- Royalty free encoding, storage, and decoding mechanism
- [Advance profession's best practices]

## Education and Persuasion

### Information Required

- List of vendors and open source project
- List of institutions planning, developing, or using JPEG 2000
- Mailing list for questions from potential implementers and discussion of profession-specific issues
- Opportunities to view and use JPEG 2000 images
- Testbed of JPEG 2000 tools for testing software and image files
- Expanded and comprehensive bibliography
- Case studies describing projects, architectures, and associated workflows
  - Rationale for choosing JPEG 2000
  - Successes or problems that can be reported
  - Issues of risk (e.g. TIFF versus JPEG 2000 image standards)
  - Impact of introducing JPEG 2000 into existing systems and workflows
- Weblog for continued discussion and networking
- Presentations at SAA, ALA, CNI meetings

### Conditions Required to Ensure Adoption

- Acceptance of the standard by professional bodies
- Use of the standard by community leaders
- Inclusion in documents describing imaging best practices
- Establishment as a required element in contracts to system vendors

## Questions from Our Communities

- Does JPEG 2000 offer cheaper and/or more sophisticated tools for access, preservation and interchange?
- Is the promise of a single file that can be used as a preservation master and access mechanism achievable?
- Is it desirable to have a preservation master serving restorative and transformative (added value) uses?
- Is "lossless" compression *truly* lossless?

## Inquiry and Research

## Coordination of Derivative Standards

- Very flexible standard causes widely divergent applications until common threads are found
- Leverage existing standard parts and related developments, then build derivative standards for archives and libraries
- Identify rigorous mechanism for adoption of derivative standards: creation of new "part" through JPEG 2000 standards committee or stand-alone process from NISO

## Creation of Usage Profiles

- Very flexible standard also means higher barrier for effective use
- New best practices for variety of material types, capture mechanisms, and intended uses

## Effective use of Metadata Boxes

- Standards or best practices on the use of XML/UUID boxes for embedding metadata
- DIG-35 (image metadata standard from digital photography) -- already in JPEG 2000
- PREMIS (emerging preservation/digital provenance metadata standard)
- Embed full metadata and/or pointers to metadata
- Common practices for including metadata in files delivered to end-users

## Interoperability Testbeds

- Many permutations of software tools, image codestream settings, and metadata boxes
- Testbed of reference images ensures all are operating at a defined level
- Reference images also serve as examples for experimentation

## Models for Migration

- Experimentation with software and workflows for bulk conversion of TIFFs, JPEGs, SIDs, etc. to JPEG 2000
- Detail effects of new practices on workflows
- Migration of TIFF header metadata to JPEG 2000 boxes



slide 24

## Creation of Tools

slide 25

## Tools for Practitioners

- Software libraries for imbedding, validating, updating, and extracting metadata
- File editors for manipulating XML and UUID boxes
- Integration with existing digital library tools (e.g. software for batch-processing, cataloging, etc.)

slide 26

## Tools for Users

- As of late 2004, few freely available viewers/plugin-ins
- Existing viewers do not make use of metadata boxes

slide 27

## Viewer Attributes of Interest to the Library and Archive

### Communities

- Web distributable, browser compliant, broadly available
- Meet communities' image quality requirements
- Reasonable performance viewing all or parts of image over narrow-band connections
- Manipulation functions (pan, zoom, rotate, invert, mirror)
- Show image in context with imbedded metadata and/or multimedia file
- Dynamically retrieve contextual information
- Includes transformative tools

slide 28

### Tool Architecture

- Image annotation for a biological taxonomist
- Georeferenced, multi-spectral satellite imagery for a geologist
- Reviewing cultural heritage photographs for a middle school student
  
- Cross-platform Java classes built upon a common framework
- Allows specialized tools be created for each material type and user's needs

slide 29

## Tools for Agents

- Expose embedded metadata through OAI and other harvesting protocols
- Reconcile annotations from various sources

## Status of Efforts

- Website Active: <http://j2kArcLib.info/>
  - Available: Symposium Report, Bibliography, Directory of Resources, RSS Feeds, Participant Blogs
  - Forthcoming: Mailing List, Sample Files, Project Case Studies
- Informal Gathering at ALA Midwinter (January 2005); Possible Formation of LITA Interest Group at ALA Annual Conference (June 2005)

## Acknowledgements

- Financial support from the Gladys Kreible Delmas Foundation and the Connecticut State Library
- Robert Buckley, Research Fellow, Xerox and Ronald Murray, Digital Conversion Specialist, Library of Congress
- Invited Speakers and Participants of the Symposium on the Adoption of JPEG 2000 in Archives and Libraries
- Kristin Eshelman, Curator of Multimedia Collections, University of Connecticut and Rutherford Witthus, University of Connecticut (retired)

*Revised Contact Information:* As of December 17th, 2004, contact information is 'peter@OhioLINK.edu'; OhioLINK, 2433 North Star Rd Suite 300, Columbus, OH 43221; 614-728-3600 x338