



## Report to National Geospatial Digital Archive Regarding Geospatial Data Treatment in Data Format Registry Efforts

Content Innovations, LLC

*Revised*  
December 16, 2009

*Previous Revision*  
Monday, May 04, 2009  
Thursday, November 20, 2008

*Originally submitted*  
Tuesday, September 30, 2008

## Contents

Overview.....	3
Research: Methodology, data collected, and description of sources.....	4
Geospatial Formats Researched.....	4
Methodology.....	5
Sources & Web Resources:.....	7
Findings.....	7
Presence of Geospatial data formats in current registry efforts.....	7
Format Registry Model & Schema Research:.....	8
Registry Models' Accommodation of Geospatial Data.....	9
Comparing registry models:.....	10
Authoring File Format Definitions for Geospatial Data.....	10
Recommendations.....	11
About this Project.....	13
<i>About NGDA</i> .....	13
<i>About Content Innovations, LLC</i> .....	13
APPENDIX A: NGDA FORMAT REGISTRY RESEARCH BIBLIOGRAPHY AND RESOURCES	14
Bibliography.....	14
Utilities: Format Identification and Conversion Tools, Metadata Tools.....	15
Lists of Formats:.....	16
Appendix B: NGDA Registry Survey.....	17
Appendix C: NGDA Registry Field Map Research.....	18
Appendix D: Library of Congress Definition TIFF Sample.....	19
Appendix E: GDFR and PRONOM Format Registry Definitions' comparison.....	29
Appendix F: Geospatial File Format Definitions.....	30

## Overview

Content Innovations, LLC was retained to support Stanford's National Geospatial Digital Archive (NGDA) effort by researching treatment of geospatial data in select data format registry efforts. We researched treatment of 40 geospatial data formats or format subtypes in key format registries such as Library of Congress (LOC)'s Digital Formats Web site, The United Kingdom's National Archives' online registry of technical information (PRONOM), and the Global Digital Format Registry (GDFR). We examined how likely target formats for ingest into NGDA were represented in key registries.

We also compared format registry data models and mapped common fields and features across registry efforts such as NGDA, PRONOM, and GDFR, and Library of Congress.

We examined the data structure of format registry efforts and consulted with the Library of Congress about using their Format Data Definition structure to author data format definitions for geospatial data. We then authored 26 LOC compliant data format definitions for select geospatial data formats and updated three geospatial related formats pertinent to the NGDA effort.

Links to the format registry efforts described above are included in APPENDIX A: NGDA FORMAT REGISTRY RESEARCH BIBLIOGRAPHY AND RESOURCES.

The data format definitions we authored are detailed in APPENDIX F: Geospatial File Format Definitions.

## Research: Methodology, data collected, and description of sources.

### Geospatial Formats Researched

We examined presence of registry entries and researched format information for the following spatial data formats:

1. Band Interleaved by Line, Component File (BIL) , fdd000283
2. Band Interleaved by Pixel, Component File (BIP), fdd000288
3. BLW ESRI Arc View World file for BIL, fdd000287
4. Band SeQuential (BSQ), Component File, fdd000289
5. Digital Elevation Model (DEM), fdd000285
6. Digital Orthophoto Quadrangle (DOQ) in native DOQ , fdd000290
7. Digital Orthophoto Quadrangle (DOQ) as Geotiff
8. Digital Raster Graphic (DRG) as TIFF (6), fdd000292
9. ESRI ArcInfo Interchange File (e00), fdd000291
10. ESRI Arc/View ShapeFile, fdd000280
11. ESRI ArcInfo Coverage, fdd000284
12. ESRI Geodatabase, fdd000293
  - a. ESRI Geodatabase (ArcSDE)
  - b. ESRI Geodatabase (File-based), fdd000294
  - c. ESRI Geodatabase (MDB)
  - d. ESRI Geodatabase (XML), fdd000295
13. ESRI/GRID, fdd000281
14. GeoTIFF, fdd000279
15. GML for JPG2000, fdd000303
16. Hierarchical Data Format HDF (5), fdd000229
  - a. HDF EOS Hierarchical Data Format-Earth Observing System, fdd000297
17. Landsat
  - a. National Landsat Archive Production System (NLAPS) NDF, fdd000299
  - b. Landsat 4 /5: Geotiff
  - c. Landsat TM: Geometrically corrected NDF product (BIL) aka Landsat 4 /5 BIL
  - d. Landsat TM: Geometrically corrected NDF product (BSQ) aka Landsat 4/5 BSQ
  - e. Landsat 7 ETM+ off gap-filled products: Geotiff
  - f. Landsat 7 ETM+ SLC-on mode: Geotiff
18. MrSid Multi-resolution Seamless Image Database, fdd000031
19. National Aerial Photography Program (NAPP) in ESRI GRID format
20. National Elevation Dataset NED in ESRI GRID format
21. Navigational Charts as ARC Digitized Raster Graphics (ADRG) -see fdd000292
  - a. JNC as ADRG
  - b. ONC Operational Navigational Chart as ADRG
  - c. TPC as ADRG
22. Shuttle Radar Topo Mission (SRTM) as TIFF
23. TIFF (6) , fdd000022
24. SDTS Spatial Data Transfer Standard, fdd000286
  - a. SDTS-TVP Topological Vector Profile, fdd000301
25. Vector Product Format (VPF ), fdd000302
26. World Vector Shoreline Plus (ESRI Shapefiles or VPF)

## Methodology

For each of the formats in the format research set we documented its full name, short name (if any) and file extension convention(if any). We researched a vernacular description of the spatial data format and prepared format notes about application of the format & local notes regarding Stanford's sample data set or use of the format in California and the West where appropriate.

For each of the formats in the format research set we determined whether the format was a file format or spatial data container. If a container (rather than a file format type) we noted what the container's minimal file components were if known and any additional container components and their likely file extensions.

For each format examined we located the format's defacto home page/URL, located the format's published specification URL if public, and URL's for select whitepaper(s). We noted the version of the specification and/or format where known/germane.

Discriminating between containers and container elements, and isolating container element file formats were first steps in our research. Geospatial Data is frequently comprised of data sets which may contain multiple files in several different file formats. Format registries, while sometimes accommodating containers, are more typically geared toward defining data types at the file format level.

Our research was informed by the abstract notion of a container as "a class, a data structure, or an abstract data type (ADT) whose instances are collections of other objects. They are used to store objects in an organized way following specific access rules".<sup>1</sup> This is a similar but slightly narrower definition compared to JHOVE's where a "format is the set of syntactic and semantic rules for serializing an abstract information model, an expression of exchangeable knowledge. "

Our research was informed by a GIS Specific definition of data format as: "a specific, possibly proprietary, set of data structures within a software system."<sup>2</sup> rather than as a file format specific definition, hence we were able to examine both geospatial datasets as containers as well as specific file formats (often as container elements).

---

<sup>1</sup> *Survey and assessment of sources of information on file formats and software documentation* Final Report" The Representation and Rendering Project University of Leeds. 5/20/2003.

<sup>2</sup> *Handbook on Geographic Information Systems and Digital Mapping, Studies in Methods, Series F, No. 79*, United Nations Department of Economic and Social Affairs, Statistics Division, New York, 2000, Annex VI - Glossary.

Our research was also directed by the importance of properly documenting data formats at the file format level (e.g. TIFF, or for container element' file types) where a file format is defined as: "a particular way to encode information for storage in a computer file".<sup>3</sup> This is a more strict definition of a data format at the file format level. Common Methods for Identifying file formats include:

- filename extension (e.g. .doc, .xls, .ppt, etc.)
- internal & external signatures
- magic number
- explicit metadata
- Mac OS type-codes (superseded by Mac OSX Uniform Type Identifiers (UTI))
- OS/2 extended attributes (".TYPE")
- MIME types

We recorded file name extensions, linked to specifications, and noted where explicit metadata existed in container element files or headers. We also noted type and MIME type data fields in our format registry model field mapping effort.

Over the life of the project we struggled with web permanence. Linking to specifications and recording specification URLs was part of our mandate. Over the life of the project many specification URLs were updated/revised, or removed/disappeared from public access.

For example, we researched JNC, ONC and other chart formats. In November 18, 2004 National Geospatial-Intelligence Agency (NGA), Department of Defense announced its intent to remove aeronautical information from public sale and distribution by October 1, 2005<sup>4</sup>. Some specifications and data related to this action are no longer available to the public via their previously published URLs. Other data and publications related to this action appear to have proliferated across the internet on 3<sup>rd</sup> party sites when they became unavailable directly from nga.ml.

This experience leads us to recommend locally linked storage of specification documentation and/or web archiving for format registry efforts that support long-range archiving aims.

---

<sup>3</sup> Wikipedia

<sup>4</sup> Federal Register Federal Register / Vol. 69, No. 222 pg 67546-67547

## Sources & Web Resources:

For each format we examined select format registry efforts (NGDA<sup>5</sup>, PRONOM, GDFR & LOC) to determine if the format was present in the registry. For each format we also researched treatment at the format level by JHOVE/JSTORE –Harvard’s Object Validation Environment, FGDC and OGC.

For each format we also examined four commonly used GIS conversion tools/utilities (ESRI, GDAL, Manifold, & SAFE) to determine whether the format was supported for import & export and/or direct read & direct write.

A bibliography of our research targets and sources is presented in **Appendix A: NGDA FORMAT REGISTRY RESEARCH BIBLIOGRAPHY AND RESOURCES** (NGDAFormatRegResources.doc).

## Findings

### Presence of Geospatial data formats in current registry efforts

The Geospatial formats we researched were well represented in popular industry software and conversion utility packages. More than 50% of the formats we examined were accommodated types in GDAL, Manifold & SAFE’s conversion utilities. At least 70% of the formats we examined were directly read/writable by ESRI software or accommodated by ESRI’s Interoperability Extension.

However, while registry efforts at PRONOM and GDFR have been setting the standard for modeling and publishing data format definitions, these particular registries have impoverished and incomplete coverage of geospatial data formats. Less than 1/3 of the data formats we examined were present in PRONOM, LOC & GDFR had even fewer –as few as two or three formats represented in each.

Our complete geospatial data format registry research survey findings are recorded in *Appendix B: NGDA Registry Survey*

---

<sup>5</sup> NGDA Format Definition content is no longer online at <http://www.ngda.org/> and NGDA’s format definitions have been submitted for distribution by Library of Congress, See <http://www.digitalpreservation.gov/formats/fdd/descriptions.shtml>

## Format Registry Model & Schema Research:

We compared format registry data models and mapped common fields and features across registry efforts such as NGDA, PRONOM, and GDFR, as well as Library of Congress. We examined the published data models & data dictionaries for these registries (where available).

We recorded common fields used by all or most of the registries we examined and mapped those fields to one another where appropriate. NGDA's format registry wiki fields were similarly mapped and where the NGDA registry effort had not yet modeled a particular field we noted the discrepancies. In a head to head comparison PRONOM and GDFR were most similar, with LOC and NGDA sharing some fields in common. PRONOM, GDFR both supported registry entries at the file format level which recorded format specific internal and external file signatures and compression. PRONOM, GDFR, and LOC each supported Intellectual Property Rights attached to a given format.

We examined the following format registry fields across NGDA, PRONOM, GDFR, & LOC:

1. System ID | Internal Unique Identifier|
2. External Identifier
3. Name
4. Version
5. Alias
6. Family
7. Format Type aka Classification
8. Description
9. Filename Extension
10. Assessment
11. Orientation
12. Byte Order
13. Grammar
14. Related File Formats
15. Internal Signature
16. External Signature
17. File type signifiers
18. Compression Type
19. Character Encoding
20. Format Disclosure
21. Release Date
22. Withdrawn Date
23. Developer aka Developed By, Created By
24. Support aka Supported By, Maintained by
25. Documentation
26. IPR
27. Caveats

28. Notes: General
29. Notes: History
30. Reference File
31. Local use
32. Production phase
33. Transparency
34. Self-documentation
35. External dependencies
36. Technical protection considerations
37. Internet Media Type
38. File type signifiers

Each Examined Model has some structure to accommodate Relationship to other formats

39. *Has subtype*
40. *Subtype of*
41. *Contains*
42. *May contain*
43. *Used by*
44. *Based on*
45. *Defined via*

Each Examined Registry Model has some structure to accommodate Relationship to other versions

46. *Has earlier version*
47. *Has later version*
48. *Has version*
49. *Version of*

Our complete Registry Field Map research findings are presented *in Appendix C: NGDA Registry Field Map Research*

## **Registry Models' Accommodation of Geospatial Data**

PRONOM and GDFR's registry data models handle a high enough level of abstraction to accommodate the challenges of geospatial data, but geospatial data are not well represented in these registries yet.

The intense use of containers in Geospatial data formats requires a format registry that will support complex relationships between containers, container elements, related data formats, and format versions.

## Comparing registry models:

The Library of Congress (LOC)'s Format Data Definition schema modeled according to GDFR concepts accommodates relationships between formats and has sufficient complexity to record most elements of a geospatial data format.

See *Appendix D: Library of Congress Definition TIFF Sample* for an example of a TIFF format definition in XML.

GDFR's naming conventions and the level of abstraction of the GDFR registry model combine to make it more difficult to decipher their registry entries out of context of the larger GDFR data structure than PRONOM's. Both GDFR and PRONOM's data models are valid and either can accommodate data format registry entries for geospatial data types.

An informative exercise is to compare two XML definitions for the same format side-by-side. Comparing definitions is one way to get a feel for the way the underlying data schemas influence format level definition entries in GDFR and PRONOM.

See *Appendix E: GDFR and PRONOM Format Registry Definitions' comparison* for a side-by-side comparison of TIFF format registry definition in XML for PRONOM and GDFR's registry data models.

## Authoring File Format Definitions for Geospatial Data

We finally elected to author our format definitions in according to the Library of Congress (LOC)'s Format Data Definition schema because it accommodated relationships between formats and had sufficient complexity to record most elements of a geospatial data format.

In our File Format Definitions, see *Appendix F: Geospatial File Format Definitions*, we populated the relationship element of each definition indicating the relationships between the defined format and other formats, using library of congress' authority controlled enumeration terms to describe the relationship type between formats in such a way that subtypes, versions and contains concepts could be recorded.

Support for automated ingest and data format validation demand that wherever possible it is important to author geospatial data format definitions down to the container element aka file format level so that all likely container element file types are accurately described in the registry for any given registry entry.

The LOC fdd schema also supports through its "fdd:relatedTo" element directly linking to other fdd's by their attribute ID (fdd:id element value) when defining a relationship between formats. We used this schema feature to strongly cross reference geospatial component files to one another and their master container formats/file types.

LOC's schema also supports multi-entry of specification and reference links, so we were able to transfer our research into identifying specification documents and supporting references directly into our format definition output using the LOC schema.

## Recommendations

- 1) Rich Definition of Container elements and relationships: It is particularly important for Geospatial Data format registry efforts like NGDA's to adopt data models for their registries that accommodate parent child relationships between containers and container elements, as well as relationships between format versions and related data types. NGDA and other efforts that maintain a format registry data structure should adopt and continue to develop registry models that support a rich definition of containers and parent/child relationships.
- 2) Sample Datasets: Having ready access to conforming sample datasets and files are an important part of understanding data format and can be essential to understanding geospatial file format relationships . Easy definition and linking to reference/sample files for each file format entry would enhance any data format registry effort that seeks to accommodate geospatial data.
- 3) XML import and export: Import and export should be format registry feature priorities. Good import and export features support easy self- export & registry entry portability and enhancement.
  - a) Treatment of empty tags in XML export is an important export consideration. An export to XML feature that allows the user to select whether to write empty tags for undefined attributes might be a useful registry enhancement. For example, GDFR's "save to XML" appears to write tags for populated attributes from GDFR's base, format, and product tables. If there has been no value entered for a particular registry attribute in GDFR, the empty tags for that attribute don't appear in GDFR's "save to XML". Since GDFR doesn't output empty tags so the default export/display of GDFR's minimal level entries in XML does not allow the casual user an immediate way to determine which elements of a given registry definition have not yet been completed.
- 4) Web UI: As currently implemented without a direct web to format registry data entry UI, there is a strong gate keeping element to the PRONOM and GDFR registry implementations that likely discourages format contribution and enhancements from knowledgeable scientific government, and industry parties who may be active in developing and enhancing formats, but are likely too far removed from exposure and participation in format registry contribution efforts. Opening enhancement and contribution up via a publically accessible web UI that allows the user community to

referee and contribute to format definitions would be one way to encourage greater participation.

- 5) Web Permanence: Over the life of the project many data format specification URLs were updated/revised, or removed/disappeared from public access. This experience leads us to recommend locally linked storage of specification documentation in conjunction with web archiving to providing for long-term storage and discovery of specifications and other reference files.

## About this Project

### *About NGDA*

From the first colonial maps to the satellite imagery of the 21st century, cartographic information has helped define and frame our view of the United States. Geospatial imagery is now a critical tool for state and federal researchers pursuing greater understandings of complex matters including: Disaster Relief, Census Data, and Land Use Management.

Concerned that millions of nationally important digital information resources are in danger of being lost or corrupted, the Library of Congress has formed The National Digital Information Infrastructure and Preservation Program (NDIIPP).

As partners in the project the university libraries of UCSB and Stanford are leading the formation of the National Geospatial Digital Archive (NGDA), a collecting network for the archiving of geospatial images and data.

### *About Content Innovations, LLC*

Content Innovations, LLC is a woman owned, certified small business based in San Francisco, California. Content Innovations, LLC is a provider of information services for research, utility, business, legal, and technology sectors. The company is led by information industry veterans providing systems analysis, project management, database services, information evaluation, workflow improvement, website management and training with specialties in geographic information systems, image databases, and integrated library system consulting. .

## APPENDIX A: NGDA FORMAT REGISTRY RESEARCH BIBLIOGRAPHY AND RESOURCES

### Bibliography

*Content standard for digital geospatial metadata (revised June 1998)*. Federal Geographic Data Committee. FGDC-STD-001-1998. Federal Geographic Data Committee. Washington, D.C.

*Content standard for digital geospatial metadata: Extensions for Remote Sensing Metadata*. Federal Geographic Data Committee. FGDC-STD-012-2002. Federal Geographic Data Committee. Washington, D.C.

### The Federal Geographic Data Committee (FGDC)

URL=<http://www.fgdc.gov/metadata/geospatial-metadata-standards>

FGDC Metadata Publications. URL=<http://www.fgdc.gov/metadata/metadata-publications-list>

### *Guide to FGDC Compliant Metadata*

URL=[http://sedac.ciesin.columbia.edu/metadata/guide/guide\\_toc.html](http://sedac.ciesin.columbia.edu/metadata/guide/guide_toc.html)

Global Digital Format Registry (GDFR) URL= <http://www.gdfr.info/>

*Note: GDFR Performs best in Firefox.*

Global Digital Format Registry *Data Model Specification 5\_0\_14*

[http://gdfr.info/docs/GDFR-data-model-5\\_0\\_14.pdf](http://gdfr.info/docs/GDFR-data-model-5_0_14.pdf)

Global Digital Format Registry Documentation URL=

[http://gdfr.info/docs.html#data\\_model](http://gdfr.info/docs.html#data_model)

JHOVE – JSTOR/Harvard Object Validation Environment URL=

<http://hul.harvard.edu/jhove/>

Library of Congress Related Resources for Digital Format Sustainability. URL=

<http://www.digitalpreservation.gov/formats/intro/resources.shtml>

Sustainability of Digital Formats: Planning for Library of Congress Collections, Format Descriptions URL =

<http://www.digitalpreservation.gov/formats/fdd/descriptions.shtml>

Also listed at the LOC site are selected resources that pertain to :

- the assessment of digital formats
- efforts to provide gateways and/or listings of format documentation and transcoding tools
- aspects of other digital preservation activities that include recommendations relating to digital content in particular formats.

*Metadata and GIS*, An ESRI ® White Paper , October 2002. URL=

[http://downloads2.esri.com/support/whitepapers/ao\\_/metadata-and-gis.pdf](http://downloads2.esri.com/support/whitepapers/ao_/metadata-and-gis.pdf)

Content Innovations, LLC

Page 14

655 Montgomery Street, 5<sup>th</sup> Floor, San Francisco, CA 94111 · (415) 550-0650

email: [info@contentinnovations.com](mailto:info@contentinnovations.com)

**National Geodata Archive Format Registry URL=**

<http://ngda.library.ucsb.edu/format/index.php/Category:Formats>

Format Definition content now superseded by definitions held at Library of Congress:

(<http://www.digitalpreservation.gov/formats/fdd/descriptions.shtml>)

**The National Software Reference Library (NSRL) URL=**

<http://www.nsrl.nist.gov/index.html>

Is described on the NSRL web site as aiming to "...collect software from various sources and incorporate file profiles computed from this software into a Reference Data Set (RDS) of information. The RDS can be used by law enforcement, government, and industry organizations to review files on a computer by matching file profiles in the RDS. This will help alleviate much of the effort involved in determining which files are important as evidence on computers or file systems that have been seized as part of criminal investigations."

**Open Geospatial Consortium, Inc. URL=** <http://www.opengeospatial.org/>

**PRONOM 4 INFORMATION MODEL. PRONOM 4-IM-1. The National Archives. 4**

January 2005. URL=

[http://www.nationalarchives.gov.uk/aboutapps/fileformat/pdf/pronom\\_4\\_info\\_model.pdf](http://www.nationalarchives.gov.uk/aboutapps/fileformat/pdf/pronom_4_info_model.pdf)

**PRONOM Technical Registry Search by Format URL=**

<http://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=new>

***Shoreline Metadata Profile of the Content Standards for Digital Geospatial Metadata***, FGDC, June 2001, Reston, Virginia.

A [Metadata Profile for Shoreline Data](#) - The profile addresses variability in the definition and mapping of shorelines by providing a standardized set of terms and data elements required to support metadata for shoreline and coastal data sets. The profile also includes a glossary and bibliography.

**Utilities: Format Identification and Conversion Tools, Metadata Tools**

**ArcGIS Data Interoperability extension** enables [ArcView](#), [ArcEditor](#), and [ArcInfo](#) users to directly read and import more than 75 spatial data formats and export more than 50. <http://www.esri.com/software/arcgis/extensions/datainteroperability/index.html>

**CAMiLEON Migration on Request Tool**

CAMiLEON recognizes that Migration from format to format over time is a widely recognised method of performing preservation work, but is far from ideal. Conventional Migration can be costly and potentially inaccurate, as errors during transformation can propagate through subsequent transformations. CAMiLEON has developed an alternative approach, called Migration on Request. By maintaining the original bytestream along with a tool to migrate the object at the point of use, our preservation strategy can be more accurate and cost effective.

**CRiB Conversion and Recommendation of Digital Object Formats for Cultural Heritage Data includes format detector.** URL = <http://crib.dsi.uminho.pt/>

**DROID** National Archives Digital Record Object Identification tool that performs automated batch identification of file formats URL=  
<http://droid.sourceforge.net/wiki/index.php/Introduction>

**FGDC List of Geospatial Metadata Tools**

URL=<http://www.fgdc.gov/metadata/geospatial-metadata-tools>

**GDAL** - Geospatial Data Abstraction Library URL= <http://www.gdal.org/>

GDAL is a translator library for raster geospatial data formats that is released under an [X/MIT](#) style [Open Source](#) license by the [Open Source Geospatial Foundation](#). As a library, it presents a [single abstract data model](#) to the calling application for all supported formats. It also comes with a variety of useful [command line utilities](#) for data translation and processing.

**GDAL supported raster formats** are listed at URL=

[http://www.gdal.org/formats\\_list.html](http://www.gdal.org/formats_list.html)

The related [OGR](#) library (which lives within the GDAL source tree) provides a similar capability for simple features vector data.

**SAFE FME Desktop Translator/Converter:** Formats Supported by FME Desktop. URL =

<http://www.safe.com/products/desktop/formats.php>

**Lists of Formats:**

Directories of GIS and other file formats and what reads and writes those formats

**Grime** - <http://www.grime.net/gistools/a-z.htm>

**Wotsit.org** - <http://www.wotsit.org/default.asp>

**GIS Files and Formats** - <http://www.wotsit.org/list.asp?fc=14>

**FILEExt** (database of file extensions and the various programs that use them) -

<http://filext.com/>

**Ace.net** - <http://www.ace.net.nz/tech/TechFileFormat.html>

**Open Information Interchange initiative** (specific information on file formats suitable for interchange and sharing.) - <http://www.diffuse.org/oii/en/oiistand.html#oiistand>

## Appendix B: NGDA Registry Survey

*For more information regarding availability of the content of this appendix (NGDARegistrySurvey12.xls) contact Julie Sweetkind-Singer ([sweetkind@stanford.edu](mailto:sweetkind@stanford.edu)).*

## Appendix C: NGDA Registry Field Map Research

SEE NEXT PAGE FOR COMPLETE APPENDIX C

Registry Format Definition Record Fields Include:	Description of the Field	PRONOM <a href="http://www.nationalarchives.gov.uk/aboutapps/fileformat/pdf/pronom_4_info_model.pdf">http://www.nationalarchives.gov.uk/aboutapps/fileformat/pdf/pronom_4_info_model.pdf</a>	GDFR <a href="http://gdfp.info/docs.html#data_model">http://gdfp.info/docs.html#data_model</a>	LOC <a href="http://www.digitalpreservation.gov/formats/fdd/fdd_explanation.shtml#sustainability">http://www.digitalpreservation.gov/formats/fdd/fdd_explanation.shtml#sustainability</a>	NGDA Format Reg Research Includes for Ea Format where Known
<b>PRONOM System ID   GDFR Identifier  </b>	Internal system-generated ID	x	x		THIS WILL BE THE NGDA ID
<b>External Identifier</b>	External identifier which applies to the format	x	X (Product.identifier)		partial -pronom recorded, need to record for GDFR?
<b>Name</b>	Preferred name of the format	x	x	x	x
<b>Version</b>				⌘	partially recorded, need to add structure to accommodate iterative versions
<b>Alias</b>	Alias by which the format is known	x	X (Product.alias)	⌘	Short Name
<b>Family</b>	Family group to which the format is assigned	x	X (format.roleFacet=role:family)	⌘	Field Not Present
<b>Format Type aka Classification</b>	Generic format type	x	x	⌘	Field Not Present
<b>Description</b>	Description of the main characteristics of the format	x		x	x
<b>Filename Extension</b>	Name extensions generally used in Windows, UNIX, and other environments.	X	X	X	x
<b>Assesment</b>	Optional format assessment expressed in a typed formal notation		x		Field Not Present
<b>Orientation</b>	Indicates whether the format is text or binary based	x	X (format.FormFacet)	⌘	partial in notes -need to accommodate with registry structure
<b>Byte Order</b>	Allowable byte orders for the format	x	x	⌘	Field Not Present
<b>Grammar</b>	Optional format grammar expressed in a typed formal notation.		x		Field Not Present

<b>Related File Formats</b>		x	x	x	partial -we recorded parent formats
<b>Internal Signature</b>	Internal signature which may be used to identify the file format	x	x	⌘	Field Not Present
<b>External Signature</b>	External signature which may be used to identify the file format	x	x	⌘	Field Not Present
<b>File type signifiers</b>	signifiers that may be used by automated systems to identify a format or the data it contains	x	x	x	Field Not Present
<b>Compression Type</b>	Compression method employed by the file format	x	X (format relationship to a format with format.transformFacet=transform:compression)	⌘	partial in notes -need to accommodate with registry structure
<b>Character Encoding</b>	Character encoding employed by the file format	x	X (format relationship to a format with GenreFacet=genre:text and format.FormFacet=role:serialization)	⌘	partial in notes -need to accommodate with registry structure
<b>Format Disclosure</b>	Level of public disclosure provided for the format specification	x		x	partial at spec and in notes - need to accommodate in registry structure
<b>Release Date</b>	Date on which the format was released	x	x	⌘	partial in desc -need to accommodate in registry structure
<b>Withdrawn Date</b>	Date on which support for the format was, or is due to be, withdrawn	x	x	⌘	partial in desc & notes -need to accommodate in registry structure
<b>Developer aka Developed By, Created By</b>	Actor responsible for developing the format	x	X (format hasCreator relationship to Agent)	⌘	x
<b>Support aka Supported By, Maintained by</b>	Actor currently responsible for supporting or maintaining the format	x	x		x
<b>Documentation</b>	Documentation about the format	x	x	x	x

<b>IPR</b>	Intellectual property rights which attach to the format	x	x	x	Field Not Present
<b>Caveats</b>	Optional known caveat(s) associated, but not clearly formerly documented, with the format.		x		Field Not Present
<b>Notes: General</b>	Informative note about the format	x	X (Base.notes)	x	x
<b>Notes: History</b>	Information about the history of the format.			x	Field Not Present
<b>Reference File</b>		x	x		Field Not Present, should add
<b>Local use</b>	aka Report of actual practice at the Library of Congress.			x	x -partially recorded in local notes
<b>Production phase</b>	Indication of how the format is generally used during the content life cycle, e.g., by creators or authors (initial state), by distributors, publishers, or archives (middle state), or as delivered to endusers (final state).			x	Field Not Present
<b>Transparency</b>	Statement regarding the nature of the encoding and/or bitstream, suggestive of the ease with which rendering tools may be obtained or built.			x	Field Not Present
<b>Self-documentation</b>	Statement regarding the degree to which the format supports the inclusion of metadata (descriptive, administrative, and structural).			x	Field Not Present

<b>External dependencies</b>	Statement regarding the need for external software or hardware.	x	x	x	Field Not Present
<b>Technical protection considerations</b>	Support by this format for elements that protect intellectual (or other) property. More discussion: technical protection.			x	Field Not Present
<b>Internet Media Type</b>	When possible, from the IANA MIME Media Types site; otherwise (and in addition) from other sources, e.g., Filext.com.	X (external identifier of type MIME)	X (Format.MimeType)	x	Field Not Present
<b>File type signifiers</b>				x	Field Not Present
Each Examined Model has some structure to accommodate <b>Relationship to other formats</b>					
<b>Has subtype</b>		X (Supertype of relationship)		x	
<b>Subtype of</b>		X (Subtype of relationship)	X (format to format relationship of type isRestrictionOf)	x	
<b>Contains</b>			X (format to format relationship of type mustContain)	x	x
<b>May contain</b>		X (Can contain relationship)	X (format to format relationship of type canContain)	x	x
<b>Used by</b>				x	
<b>Based on</b>			X (format to format relationship of type isExtensionOf, isRestrictionOf, isModificationOf)	x	
<b>Defined via</b>			X (format to format relationship of type isDefinedBy or isRequisiteFor?)	x	

Each Examined Registry Model has some structure to accommodate <b>Relationship to other versions</b>					
<b>Has earlier version</b>	<i>Preferred to Has Version</i>	X (Subsequent version of relationship)		x	
<b>Has later version</b>	<i>Preferred to Has Version</i>	X (Previous version of relationship)	X (format to format relationship of type isPreviousVersion)	x	
<b>Has version</b>	<i>Use when version is non-sequential (e.g., geographical) or when sequence is unknown</i>			x	
<b>Version of</b>	<i>Use when version is non-sequential (e.g., geographical) or when sequence is unknown</i>			x	

## Appendix D: Library of Congress Definition TIFF Sample

```
<?xml version="1.0" encoding="UTF-8"?>
<?altova_sps ..\locfiles\fddCreation.sps?>
<fdd:FDD id="fdd000022" titleName="TIFF, Revision 6.0" shortName="TIFF_6"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:fdd=".fdd.xsd"
xsi:schemaLocation=".fdd.xsd ..\locfiles\fdd.xsd">
  <fdd:properties>
<fdd:gdfGenreSelection>
  <fdd:gdfGenre>still-image</fdd:gdfGenre>
</fdd:gdfGenreSelection>
<fdd:formatCategories>
  <fdd:category>file-format</fdd:category>
</fdd:formatCategories>
<!--<fdd:gdfComposition>container-wrapper</fdd:gdfComposition-->
<fdd:gdfForm>binary</fdd:gdfForm>
<fdd:gdfConstraint>unstructured</fdd:gdfConstraint>
<fdd:gdfBasis>sampled</fdd:gdfBasis>
<fdd:gdfDomains>
  <fdd:gdfDomain>
<fdd:value>gis</fdd:value>
  </fdd:gdfDomain>
</fdd:gdfDomains>
<fdd:updates>
  <fdd:date>2006-10-04</fdd:date>
  <fdd:date>2009-12-14</fdd:date>
</fdd:updates>
<fdd:draftStatus>Full</fdd:draftStatus>
  <fdd:properties>
  <fdd:identificationAndDescription>
<fdd:fullName>TIFF (Tagged Image File Format), Revision 6.0</fdd:fullName>
<fdd:description>
  <p>A tag-based file format for storing and interchanging raster images. TIFF serves as a
wrapper for different bitstream encodings for bit-mapped (raster) images. The different encodings
may represent different compression schemes and different schemes for color representation
(photometric interpretation). See also <a href="#notes">Notes</a>.</p>
  <p>The most recent version of TIFF is 6.0, published in 1992. Since TIFF images
conforming to earlier versions are valid TIFF 6.0 files, the information in this description is also
pertinent to earlier versions of the TIFF standard. Many TIFF files with uncompressed image
data are still being created as TIFF 5.0 files.</p>
</fdd:description>
<fdd:shortDescription>A tag-based file format for storing and interchanging raster images. TIFF
serves as a wrapper for different bitstream encodings for bit-mapped (raster) images. The
different encodings may represent different compression schemes and different schemes for
color representation (photometric interpretation).</fdd:shortDescription>
<fdd:productionPhase>Most often an initial-state or middle-state format; may serve as final-state
format.
</fdd:productionPhase>
<fdd:relationships>
  <fdd:relationship>
<fdd:typeOfRelationship>Has earlier version</fdd:typeOfRelationship>
<fdd:comment>TIFF, Revision 5.0, not separately described</fdd:comment>
  </fdd:relationship>
```

```

    <fdd:relationship>
<fdd:typeOfRelationship>Has subtype</fdd:typeOfRelationship>
<fdd:relatedTo>
    <fdd:id>fdd000023</fdd:id>
    <fdd:shortName>TIFF_UNC</fdd:shortName>
    <fdd:titleName>TIFF, Uncompressed Bitmap</fdd:titleName>
</fdd:relatedTo>
    </fdd:relationship>
    <fdd:relationship>
<fdd:typeOfRelationship>Has subtype</fdd:typeOfRelationship>
<fdd:relatedTo>
    <fdd:id>fdd000024</fdd:id>
    <fdd:shortName>TIFF_G4</fdd:shortName>
    <fdd:titleName>TIFF Bitmap with Group 4 Compression</fdd:titleName>
</fdd:relatedTo>
    </fdd:relationship>
    <fdd:relationship>
<fdd:typeOfRelationship>Has subtype</fdd:typeOfRelationship>
<fdd:relatedTo>
    <fdd:id>fdd000074</fdd:id>
    <fdd:shortName>TIFF_LZW</fdd:shortName>
    <fdd:titleName>TIFF Bitmap with LZW Compression</fdd:titleName>
</fdd:relatedTo>
    </fdd:relationship>
    <fdd:relationship>
<fdd:typeOfRelationship>Has subtype</fdd:typeOfRelationship>
<fdd:relatedTo>
    <fdd:id>fdd000237</fdd:id>
    <fdd:shortName>TIFF_PYR</fdd:shortName>
    <fdd:titleName>TIFF, Pyramid</fdd:titleName>
</fdd:relatedTo>
    </fdd:relationship>
    <fdd:relationship>
<fdd:typeOfRelationship>Has subtype</fdd:typeOfRelationship>
<fdd:relatedTo>
    <fdd:id>fdd000072</fdd:id>
    <fdd:shortName>TIFF/IT</fdd:shortName>
    <fdd:titleName>TIFF/IT, for Image Technology</fdd:titleName>
</fdd:relatedTo>
    </fdd:relationship>
    <fdd:relationship>
<fdd:typeOfRelationship>Has subtype</fdd:typeOfRelationship>
<fdd:relatedTo>
    <fdd:id>fdd000073</fdd:id>
    <fdd:shortName>TIFF/EP</fdd:shortName>
    <fdd:titleName>TIFF/EP, for Digital Photography</fdd:titleName>
</fdd:relatedTo>
    </fdd:relationship>
    <fdd:relationship>
<fdd:typeOfRelationship>Has subtype</fdd:typeOfRelationship>
<fdd:relatedTo>
    <fdd:id>fdd000188</fdd:id>
    <fdd:shortName>DNG_1_1</fdd:shortName>
    <fdd:titleName>Adobe Digital Negative (DNG), Version 1.1</fdd:titleName>
</fdd:relatedTo>

```

</fdd:relationship>  
 <fdd:relationship>  
 <fdd:typeOfRelationship>May contain</fdd:typeOfRelationship>  
 <fdd:comment>Bitstream encodings for other compression schemes, not documented at this time.</fdd:comment>  
 </fdd:relationship>  
 <fdd:relationship>  
 <fdd:typeOfRelationship>Has subtype</fdd:typeOfRelationship>  
 <fdd:relatedTo>  
   <fdd:id>fdd000279</fdd:id>  
   <fdd:shortName>GeoTIFF</fdd:shortName>  
   <fdd:titleName>GeoTIFF, Revision 1.0</fdd:titleName>  
 </fdd:relatedTo>  
 <fdd:comment>GeoTiff is the standard for storing georeference and geocoding information in a TIFF 6.0 compliant raster file.  
 </fdd:comment>  
 </fdd:relationship>  
 </fdd:relationships>  
   </fdd:identificationAndDescription>  
   <fdd:localUse>  
 <fdd:experience>  
 TIFF 5.0 and 6.0 have been used by the American Memory project as the master format for most digital reproductions from paper and photographic media such as negatives. For the first several years of the project, most images of typographic pages (no illustrations) were formatted as <fddLink id="fdd000024">TIFF\_G4</fddLink>; in more recent years, such originals have been captured as grayscale images, generally formatted as as <fddLink id="fdd000023">TIFF\_UNC</fddLink> uncompressed bitmaps. Pictorial items from printed matter, original photographs, and other graphic arts are generally reproduced as <fddLink id="fdd000023">TIFF\_UNC</fddLink>. For TIFF images from paper originals, the spatial resolution tends to be from 300 to 400 dpi (based on dimension of original on paper). When reformatting photographic negatives, the Library of Congress Prints and Photographs Division (P&P) often uses 4,000 to 5,000 pixels on long side and a bit depth of 8 bits (grayscale) or 24 bits (RGB color). Certain high-value content has warranted higher spatial resolution or bit-depth; in some of its black-and-white negative reformatting projects, P&P has produced grayscale files with 16 bits per pixel.  
 </fdd:experience>  
 <fdd:preference>One of the preferred formats for bitmapped images.</fdd:preference>  
   </fdd:localUse>  
   <fdd:sustainabilityFactors>  
 <fdd:disclosure>  
 Fully documented. TIFF was developed by the Aldus and Microsoft Corporations, and the specification is owned by Aldus (now absorbed into the Adobe Corporation). The TIFF tag set is extensible through a registry maintained by Adobe; the list of registered extensions is not available from Adobe; see <a href="http://partners.adobe.com/public/developer/en/tiff\_tags.shtml">Tags for TIFF and Related Specifications</a>.  
 </fdd:disclosure>  
 <fdd:documentation>  
   <p>  
   <a href="http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf">TIFF, Revision 6.0, Final -- June 3, 1992</a> (at http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf) </p>  
   <p>  
   <a href="http://cool-palimpsest.stanford.edu/bytopic/imaging/std/tiff5.html">TIFF 5.0 Aldus/Microsoft Technical Memorandum: 8/8/88</a> (at http://cool-palimpsest.stanford.edu/bytopic/imaging/std/tiff5.html)</p>  
   </fdd:documentation>

<fdd:adoption>

<p>Particular subtypes are very widely deployed as master formats for scanned images. Not supported by all browsers in native format but, as of early 2004, new PC configurations tend to include a viewer.</p>

<p>The format is widely supported by image-manipulation applications (Adobe Photoshop and many others), by desktop publishing and page layout applications (QuarkXPress, Adobe InDesign, and others), and by scanning, faxing, word processing, optical character recognition, and other applications.</p>

<p>Particular TIFF subtypes are preferred data formats in the list of <a href="http://www.fcla.edu/digitalArchive/pdfs/recFormats.pdf">FCLA recommended formats</a>. Library and Archives Canada has also adopted TIFF as a <a href="http://www.collectionscanada.gc.ca/government/products-services/007002-3017-e.html">recommended format</a> for still images.</p>

<p>EXIF metadata (see <fddLink id="fdd000145">TIFF\_UNC\_EXIF</fddLink> and <fddLink id="fdd000073">TIFF/EP</fddLink>), a desirable form of self-documentation, is very often provided by digital cameras.</p>

</p>

</fdd:adoption>

<fdd:licensingAndPatents>

Not exploited for the TIFF wrapper format. The only widely used compression scheme for the embedded image data that has been subject to patent claims in recent years is <fddLink id="fdd000135">LZW</fddLink> (Lempel-Ziv-Welch) compression, i.e., in <fddLink id="fdd000074">TIFF\_LZW</fddLink>. The LZW patent is generally reported as having expired in 2003 (U.S.) and 2004 (Europe and Japan).</p>

</fdd:licensingAndPatents>

<fdd:transparency>Depends on bitstream encoding.</p>

</fdd:transparency>

<fdd:selfDocumentation>

The TIFF specification defines a framework for an Image File Header (IFH), Image File Directories (IFDs), and associated bitmaps. Each IFD and its associated bitmap are sometimes called a <i>TIFF subfile</i>. There is no limit to the number of subfiles a TIFF image file may contain. Each IFD contains one or more data structures called <i>tags</i>, each one of which is a 12-byte record that contains a specific piece of information about the bitmapped data. The TIFF specification defines a number of tags and a set of rules for extensibility; see <a href="http://content/tiff\_tags.shtml">Tags for TIFF and Related Specifications</a>. Tags are always found in contiguous groups within each IFD.</p>

</fdd:selfDocumentation>

<fdd:externalDependencies>None</p>

</fdd:externalDependencies>

<fdd:techProtection>None</p>

</fdd:techProtection>

</fdd:sustainabilityFactors>

<fdd:qualityAndFunctionalityFactors>

<fdd:stillImageQF>

<fdd:normalImage>Good support.</fdd:normalImage>

<fdd:clarity>

Excellent support for images with very high spatial resolution. The standard is flexible as to color space and bit depth. In practice, 8-bit grayscale and 24-bit RGB color are common; some activities create files with greater than 8 bits per channel (color or grayscale).</p>

</fdd:clarity>

<fdd:colorMaint>

<p>The TIFF tag for the ICC profile (tag 34675, <i>InterColourProfile</i>) for a capture device has been added as a &quot;private&quot; extension in the <fddLink id="fdd000072">TIFF/IT</fddLink> and <fddLink id="fdd000073">TIFF/EP</fddLink> standards.<footNote id="1"/> Extended tags of this kind may be used in any TIFF\_6 file, although</p>

they may not be recognized by all readers. ICC [Profile version 4.2.0.0](http://www.color.org/icc1V42.pdf) (Specification ICC.1:2004-10, page 69) provides guidance for embedding ICC profiles in TIFF files: "as a single TIFF field or Image File Directory (IFD)." Meanwhile, Adobe Photoshop software appears to provide an alternate means to embed an ICC profile in a TIFF file; the compilers of this Web site seek explanatory [comments](mailto:..../contact_format.shtml) from readers: how proprietary or interoperable is PhotoShop embedding of ICC profiles?

Color space is indicated in *Photometric Interpretation* (tag 262); in TIFF\_6, this tag does not include sRGB as a value, although sRGB images may be delivered tagged as RGB.

**Color Maintenance**  
No support for vector graphics.  
Multi-page files supported for a sequence of images.

**Still Image Quality Factors**  
Multi-page files supported for a sequence of images.  
**Signifiers**  
Filename Extension  
**Signatures**  
tif, tiff

The TIFF 6.0 standard recommends the use of tif (or TIF) as extension. Macintosh Filetype is TIFF and tiff is sometimes found as a filename extension.

**Filename Extension**  
**InternetMediaType**  
**Signatures**  
image/tif  
From the [IETF](http://www.rfc-editor.org/rfc/rfc2302.txt) (RFC2302).

**InternetMediaType**  
**Signatures**  
image/tif  
image/x-tif  
image/x-tiff  
application/tif  
application/x-tif  
application/tiff  
application/x-tiff

**Signatures**  
Selected from [The File Extension Source](http://filext.com/)

**InternetMediaType**  
**MagicNumbers**  
**Signatures**  
Hex: 49 49 2A 00

**Signatures**  
For little endian byte order; from the [JHOVE page](http://hul.harvard.edu/jhove/tiff-hul.html) for TIFF.

```

    </fdd:magicNumbers>
    <fdd:magicNumbers>
<fdd:sigValues>
    <fdd:sigValue>Hex: 49 49</fdd:sigValue>
    <fdd:sigValue>ASCII: II</fdd:sigValue>
</fdd:sigValues>
<fdd:note>
For little endian byte order; from Gary Kessler's <a
href="http://www.garykessler.net/library/file_sigs.html">File Signatures</a>.
</fdd:note>
    </fdd:magicNumbers>
    <fdd:magicNumbers>
<fdd:sigValues>
    <fdd:sigValue>Hex: 4D 4D 00 2A</fdd:sigValue>
    <fdd:sigValue>ASCII: MM</fdd:sigValue>
</fdd:sigValues>
<fdd:note>
For big endian byte order; from the <a href="http://hul.harvard.edu/jhove/tiff-hul.html">JHOVE
page</a> for TIFF and <a
href="http://local.wasp.uwa.edu.au/~pbourke/dataformats/tiff/">http://local.wasp.uwa.edu.au/~pbo
urke/dataformats/tiff/</a>.
</fdd:note>
    </fdd:magicNumbers>
</fdd:signifiersGroup>
    </fdd:fileTypeSignifiers>
    <fdd:notes>
<fdd:general>
    <p>The acronym TIFF is variously glossed as <i>Tagged</i> Image File Format and
<i>Tag</i> Image File Format; the title page of the 1992 specification does not spell out the
abbreviation. According to the Wikipedia article <a
href="http://en.wikipedia.org/wiki/TIFF">Tagged Image File Format</a> (consulted on August 30,
2006), earlier versions of the specification used <i>Tag Image File Format</i>.</p>
    <p>
<a href="http://www.fileformat.info/format/tiff/">FileFormatInfo</a> offers an overview of the
format: files are organized into three sections: the Image File Header (IFH), the Image File
Directory (IFD), and the bitmap data. Of these three sections, only the IFH and IFD are required.
It is possible to have a TIFF file with no bitmapped data, although such a file would be highly
unusual. A TIFF file that contains multiple images has one IFD and one bitmap per image
stored.</p>
</fdd:general>
<fdd:history>
The Wikipedia article <a href="http://en.wikipedia.org/wiki/TIFF">Tagged Image File Format</a>
(consulted on August 30, 2006) reports that the format was originally created by the company
Aldus, jointly with Microsoft, for use with PostScript printing. <a
href="http://www.fileformat.info/format/tiff/">FileFormatInfo</a> reports that Aldus first published a
TIFF specification in 1986 and many consider this to be &#34;version 3&#34;. Another motive to
create the specification was to encourage desktop scanner vendors of the mid-1980s to agree on
a common scanned image file format, rather than have each company promulgate its own
proprietary format. In the beginning, TIFF was only a bilevel image format, since that was all that
desktop scanners could handle. As scanners became more powerful, and as desktop computer
disk space became more plentiful, TIFF grew to accommodate grayscale images, then color
images. Today, TIFF is a popular format for high-color-depth images, along with JPEG and PNG.
Adobe Systems, which acquired the PageMaker publishing program from Aldus, now controls the
TIFF specification.
</fdd:history>

```

```

    </fdd:notes>
    <fdd:formatSpecifications>
<fdd:urls>
    <fdd:url>
<fdd:urlReference>
    <link>http://partners.adobe.com/public/developer/tiff/index.html</link>
    <tag/>
</fdd:urlReference>
    </fdd:url>
    <fdd:url>
<fdd:urlReference>
    <link>http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf</link>
    <tag/>
</fdd:urlReference>
    </fdd:url>
    <fdd:url>
<fdd:urlReference>
    <link>http://www.npes.org/pdf/TIFF-v6.pdf</link>
    <tag/>
</fdd:urlReference>
    </fdd:url>
    <fdd:url>
<fdd:urlReference>
    <link>http://cool-palimpsest.stanford.edu/bytopic/imaging/std/tiff5.html</link>
    <tag/>
    <comment>(specification for version 5)</comment>
</fdd:urlReference>
    </fdd:url>
    <fdd:url>
<fdd:urlReference>
    <link>http://partners.adobe.com/public/developer/en/tiff/TIFFPM6.pdf</link>
    <tag>Adobe PageMaker 6.0: TIFF Technical Notes, 1995</tag>
    <comment>
    Also known as TIFF Specification Supplement 1. Adobe technical notes relating to TIFF
    files created by PageMaker.
    </comment>
</fdd:urlReference>
    </fdd:url>
    <fdd:url>
<fdd:urlReference>
    <link>http://partners.adobe.com/public/developer/en/tiff/TIFFphotoshop.pdf</link>
    <tag>Adobe Photoshop: TIFF Technical Notes, 2002</tag>
    <comment>
    Also known as TIFF Specification Supplement 2. This document describes use of JPEG
    compression within a TIFF image, correcting an earlier, flawed specification.
    </comment>
</fdd:urlReference>
    </fdd:url>
</fdd:urls>
    </fdd:formatSpecifications>
    <fdd:usefulReferences>
<fdd:urls>
    <fdd:url>
<fdd:urlReference>
    <link>http://hul.harvard.edu/jhove/tiff-hul.html</link>

```

```

    <tag>JHOVE TIFF-hul module</tag>
</fdd:urlReference>
    </fdd:url>
    <fdd:url>
<fdd:urlReference>
    <link>http://www.fileformat.info/format/tiff/</link>
    <tag>TIFF</tag>
    <comment>
        From a licensed version of the O'Reilly Encyclopedia of Graphics File Formats; see also
        print version below
    </comment>
</fdd:urlReference>
    </fdd:url>
    <fdd:url>
<fdd:urlReference>
    <link>http://netghost.narod.ru/gff/</link>
    <tag>Graphics File Formats</tag>
    <comment>
        Possibly unlicensed version of the O'Reilly Encyclopedia of Graphics File Formats from a
        Russian Web site; see also print version below.
    </comment>
</fdd:urlReference>
    </fdd:url>
    <fdd:url>
<fdd:urlReference>
    <link>http://www.nilesritter.com/webdocs/tiff/</link>
    <tag>Niles Ritter's Unofficial TIFF Home Page</tag>
</fdd:urlReference>
    </fdd:url>
    <fdd:url>
<fdd:urlReference>
    <link>http://preserve.harvard.edu/resources/imageformats.html</link>
    <tag>Image Digitization: File Formats, Digital Still Images</tag>
</fdd:urlReference>
    </fdd:url>
    <fdd:url>
<fdd:urlReference>
    <link>http://www.awaresystems.be/imaging/tiff/faq.html</link>
    <tag>AWare Systems TIFF FAQ</tag>
</fdd:urlReference>
    </fdd:url>
    <fdd:url>
<fdd:urlReference>
    <link>http://www.awaresystems.be/imaging/tiff/tifftags.html</link>
    <tag>AWare Systems info about TIFF tags</tag>
</fdd:urlReference>
    </fdd:url>
    <fdd:url>
<fdd:urlReference>
    <link>http://www.digitalpreservation.gov/formats/content/tiff_tags.shtml</link>
    <tag>Tags for TIFF and Related Specifications</tag>
    <comment>From Library of Congress resource: Sustainability of Digital
    Formats.</comment>
</fdd:urlReference>
    </fdd:url>

```

```

    <fdd:url>
<fdd:urlReference>
    <link>http://www.digitizationguidelines.gov/stillimages/documents/tiff_metadata.html</link>
>
    <tag>TIFF IMage Metadata</tag>
    <comment>From the Federal Agencies Still Image Digitization Working
Group.</comment>
</fdd:urlReference>
    </fdd:url>
    <fdd:url>
<fdd:urlReference>
    <link>http://www.color.org/icc1V42.pdf</link>
    <tag>ICC Profile version 4.2.0.0</tag>
</fdd:urlReference>
    </fdd:url>
    <fdd:url>
<fdd:urlReference>
    <link>http://web.archive.org/web/20070627164620/http://www.tiki-
lounge.com/~raf/tiff/fields.html</link>
    <tag>Ross Finlayson's TIFF Web page (from Internet Archive)</tag>
</fdd:urlReference>
    </fdd:url>
    <fdd:url>
<fdd:urlReference>
    <link>http://local.wasp.uwa.edu.au/~pbourke/dataformats/tiff/</link>
    <tag>Paul Bourke's TIFF Image Creation instructions</tag>
</fdd:urlReference>
    </fdd:url>
    <fdd:url>
<fdd:urlReference>
    <link>http://www.fcla.edu/digitalArchive/pdfs/recFormats.pdf</link>
    <tag>FCLA recommended formats</tag>
</fdd:urlReference>
    </fdd:url>
    <fdd:url>
<fdd:urlReference>
    <link>http://www.collectionscanada.gc.ca/government/products-services/007002-3017-
e.html</link>
    <tag>Library and Archives Canada recommended formats</tag>
</fdd:urlReference>
    </fdd:url>
</fdd:urls>
<fdd:citations>
    <fdd:citation>
<fdd:bookReference>
    <bookRefDetail rel="au">Murray, James D.</bookRefDetail> and <bookRefDetail
rel="au">William vanRyper
    </bookRefDetail>.
    <bookRefDetail rel="btitle">
    Encyclopedia of Graphics File Formats, Second Edition. Sebastopol, CA. : O'Reilly
& Associates,
    </bookRefDetail>
<bookRefDetail rel="date">1994</bookRefDetail>.
    Includes CD-ROM with complete text of book, and copies of several file format specifications.
</fdd:bookReference>

```

</fdd:citation>  
</fdd:citations>  
</fdd:usefulReferences>  
</fdd:footNotes>  
<fdd:footNote id="1">  
<fdd:text>  
The most effective color maintenance systems rely on the existence of an ICC  
(International Color Consortium) profile of the capture device, which can then be compared to  
profiles for output devices, permitting appropriate adjustments of image color.  
</fdd:text>  
</fdd:footNote>  
<fdd:footNote id="2">  
<fdd:text>The color space sRGB, standardized as IEC 61966-2-1, establishes an image  
viewing environment with a known color temperature (6500 degrees Kelvin) and gamma (2.2),  
thus increasing the user's ability to maintain color.</fdd:text>  
</fdd:footNote>  
</fdd:footNotes>  
</fdd:FDD>

## Appendix E: GDFR and PRONOM Format Registry Definitions' comparison

SEE NEXT PAGE

## GDFR Format Registry Entry for TIFF in XML

```
<n:format xsi:schemaLocation="info:gdf/xmlSchemas/format
http://www.formatregistry.org/webservices/registry/content/Sche
mas/format info:gdf/xmlSchemas/gdfrbase
http://www.formatregistry.org/webservices/registry/content/Sche
mas/gdfrbase">
-
<gdfrbase:gdfrbase>
<gdfrbase:adminStatus>Active</gdfrbase:adminStatus>
<gdfrbase:source>Andrea Goethals</gdfrbase:source>
<gdfrbase:creationDate>2008-06-29T16:52:55-
04:00</gdfrbase:creationDate>
<gdfrbase:modificationDate>2008-06-29T16:52:55-
04:00</gdfrbase:modificationDate>
<gdfrbase:systemID>info:rfa/localhost/Formats/1017</gdfrbase:sy
stemID>
</gdfrbase:gdfrbase>
-
<gdfrbase:product
xsi:schemaLocation="info:gdf/xmlSchemas/gdfrbase
http://www.formatregistry.org/webservices/registry/content/Sche
mas/gdfrbase">
<gdfrbase:name>TIFF image data, big-endian</gdfrbase:name>
<gdfrbase:identifiers/>
</gdfrbase:product>
-
<n:gdfIdentifier>
<n:type>GFID_URI</n:type>
<n:value>info:gdf/localhost/Formats/17</n:value>
</n:gdfIdentifier>
<n:mimeType>image/tiff</n:mimeType>
<n:classification/>
<n:byteOrder>Unknown</n:byteOrder>
-
<n:internalSignature>
-
<n:signature>
<n:obligation>Required</n:obligation>
<n:value>MM\x00\x2a</n:value>
</n:signature>
<n:type>BOF</n:type>
<n:offset>0</n:offset>
</n:internalSignature>
</n:format>
```

## PRONOM Format Registry Entry For TIFF in XML

```
<FileFormat>
  <FormatID>612</FormatID>
  <FormatName>Tagged Image File Format</FormatName>
  <FormatVersion>6</FormatVersion>
  <FormatAliases>TIFF (6)</FormatAliases>
  <FormatFamilies>
  </FormatFamilies>
  <FormatTypes>Image (Raster)</FormatTypes>
  <FormatDisclosure>Full</FormatDisclosure>
  <FormatDescription>The Tagged Image File Format
(TIFF) is a raster image format originally developed by
the Aldus Corporation, primarily for use in scanning and
desk-top publishing. When Adobe Systems Incorporated
purchased Aldus in 1994, they acquired the rights to the
TIFF format and have maintained it since then. TIFF
files comprise three sections: an Image File Header
(IFH), an Image File Directory (IFD), and the image
data. TIFF files can contain multiple images (multi-page
TIFF), and each image has a separate IFD. The IFH always
appears at the beginning of the file, and is immediately
followed by a pointer to the first IFD. The IFD contains
metadata which describes the associated image, stored as
a series of tags. The IFD also contains a pointer to the
actual image data. TIFF 6.0 supports colour depths from
1 bit to 24 bit (e.g. monochrome to true colour), and a
wide range of compression types (RLE, LZW, CCITT Group 3
and Group 4, and JPEG), as well as uncompressed
data.</FormatDescription>
  <BinaryFileFormat>Binary</BinaryFileFormat>
  <ByteOrders>Little-endian (Intel) and Big-endian
(Motorola)</ByteOrders>
  <ReleaseDate>03 Jun 1992</ReleaseDate>
  <WithdrawnDate>
  </WithdrawnDate>
  <ProvenanceSourceID>1</ProvenanceSourceID>
  <ProvenanceName>Digital Preservation Department /
The National Archives</ProvenanceName>
  <ProvenanceSourceDate>11 Mar
2005</ProvenanceSourceDate>
  <ProvenanceDescription>
  </ProvenanceDescription>
  <LastUpdatedDate>02 Aug 2005</LastUpdatedDate>
  <FormatNote>
  </FormatNote>
```

PRONOM Format Registry Entry for TIFF in XML *(continued)*

```
<FormatRisk>
</FormatRisk>
<TechnicalEnvironment>
</TechnicalEnvironment>
<FileFormatIdentifier>
  <Identifier>fmt/10</Identifier>
  <IdentifierType>PUID</IdentifierType>
</FileFormatIdentifier>
<FileFormatIdentifier>
  <Identifier>image/tiff</Identifier>
  <IdentifierType>MIME</IdentifierType>
</FileFormatIdentifier>
<FileFormatIdentifier>
  <Identifier>public.tiff</Identifier>
  <IdentifierType>Apple Uniform Type
Identifier</IdentifierType>
</FileFormatIdentifier>
<Developers>
  <DeveloperID>82</DeveloperID>
  <DeveloperName>
</DeveloperName>
  <OrganisationName>Aldus Corporation</OrganisationName>
  <DeveloperCompoundName>Aldus
Corporation</DeveloperCompoundName>
</Developers>
<Document>
  <DocumentID>52</DocumentID>
  <DisplayText>Adobe Systems Incorporated, 1992, TIFF
Revision 6.0</DisplayText>
  <DocumentType>Authoritative</DocumentType>
  <AvailabilityDescription>Public</AvailabilityDescription>
  <AvailabilityNote>
</AvailabilityNote>
  <PublicationDate>03 Jun 1992</PublicationDate>
  <TitleText>TIFF Revision 6.0</TitleText>
  <DocumentIPR>
</DocumentIPR>
  <DocumentNote>
</DocumentNote>

<DocumentIdentifier><Identifier>partners.adobe.com/public/develop
er/en/tiff/TIFF6.pdf</Identifier>
```

PRONOM Format Registry Entry for TIFF in XML (continued)

```
<IdentifierType>URL</IdentifierType>
</DocumentIdentifier>
<Author>
  <AuthorID>7</AuthorID>
  <AuthorName>
    </AuthorName>
    <OrganisationName>Adobe Systems
    Incorporated</OrganisationName>
    <AuthorCompoundName>Adobe Systems
    Incorporated</AuthorCompoundName>
  </Author>
  <Publisher>
    <PublisherID>7</PublisherID>
    <PublisherName>
      </PublisherName>
      <OrganisationName>Adobe Systems
      Incorporated</OrganisationName>
      <PublisherCompoundName>Adobe Systems
      Incorporated</PublisherCompoundName>
    </Publisher>
  </Document>
<Document>
  <DocumentID>53</DocumentID>
  <DisplayText>Adobe Systems Incorporated, 1995, TIFF
  Technical Note #2 (draft)</DisplayText>
  <DocumentType>Authoritative</DocumentType>
  <AvailabilityDescription>Public</AvailabilityDescription>
  <AvailabilityNote>
    </AvailabilityNote>
  <PublicationDate>01 Jan 1995</PublicationDate>
  <TitleText>TIFF Technical Note #2 (draft)</TitleText>
  <DocumentIPR>
    </DocumentIPR>
    <DocumentNote>The implementation of JPEG compression
    described in Revision 6.0 is flawed: a revised design was issued
    as TIFF Technical Note #2, which should be followed
    instead.</DocumentNote>
  <DocumentIdentifier>

  <Identifier>www.remotesensing.org/libtiff/TIFFTechNote2.html</Identifier>
  <IdentifierType>URL</IdentifierType>
</DocumentIdentifier>
<Author>
```

PRONOM Format Registry Entry for TIFF in XML (continued)

```
<AuthorID>7</AuthorID>
<AuthorName>
</AuthorName>
<OrganisationName>Adobe Systems
Incorporated</OrganisationName>
<AuthorCompoundName>Adobe Systems
Incorporated</AuthorCompoundName>
</Author>
<Publisher>
<PublisherID>7</PublisherID>
<PublisherName>
</PublisherName>
<OrganisationName>Adobe Systems
Incorporated</OrganisationName>
<PublisherCompoundName>Adobe Systems
Incorporated</PublisherCompoundName>
</Publisher>
</Document>
<Document>
<DocumentID>54</DocumentID>
<DisplayText>Adobe Systems Incorporated, 1995, Adobe
PageMaker 6.0: TIFF Technical Notes</DisplayText>
<DocumentType>Informative</DocumentType>
<AvailabilityDescription>Public</AvailabilityDescription>
<AvailabilityNote>
</AvailabilityNote>
<PublicationDate>01 Jan 1995</PublicationDate>
<TitleText>Adobe PageMaker 6.0: TIFF Technical
Notes</TitleText>
<DocumentIPR>
</DocumentIPR>
<DocumentNote>
</DocumentNote>
<DocumentIdentifier>

<Identifier>partners.adobe.com/public/developer/en/tiff/TIFFPM6.p
df</Identifier>
<IdentifierType>URL</IdentifierType>
</DocumentIdentifier>
<Author>
<AuthorID>7</AuthorID>
<AuthorName>
</AuthorName>
```

PRONOM Format Registry Entry for TIFF in XML *(continued)*

```
<OrganisationName>Adobe Systems
Incorporated</OrganisationName>
  <AuthorCompoundName>Adobe Systems
Incorporated</AuthorCompoundName>
  </Author>
  <Publisher>
    <PublisherID>7</PublisherID>
    <PublisherName>
      </PublisherName>
    <OrganisationName>Adobe Systems
Incorporated</OrganisationName>
    <PublisherCompoundName>Adobe Systems
Incorporated</PublisherCompoundName>
  </Publisher>
</Document>
<Document>
  <DocumentID>55</DocumentID>
  <DisplayText>Adobe Systems Incorporated, 2002, Adobe
Photoshop: TIFF Technical Notes</DisplayText>
  <DocumentType>Informative</DocumentType>
  <AvailabilityDescription>Public</AvailabilityDescription>
  <AvailabilityNote>
    </AvailabilityNote>
  <PublicationDate>01 Jan 2002</PublicationDate>
  <TitleText>Adobe Photoshop: TIFF Technical
Notes</TitleText>
  <DocumentIPR>
    </DocumentIPR>
  <DocumentNote>
    </DocumentNote>
  <DocumentIdentifier>

<Identifier>partners.adobe.com/public/developer/en/tiff/TIFFphoto
shop.pdf</Identifier>
  <IdentifierType>URL</IdentifierType>
</DocumentIdentifier>
<Author>
  <AuthorID>7</AuthorID>
  <AuthorName>
    </AuthorName>
  <OrganisationName>Adobe Systems
Incorporated</OrganisationName>
  <AuthorCompoundName>Adobe Systems
Incorporated</AuthorCompoundName>
```

PRONOM Format Registry Entry for TIFF in XML *(continued)*

```
</Author>
<Publisher>
  <PublisherID>7</PublisherID>
  <PublisherName>
    </PublisherName>
  <OrganisationName>Adobe Systems
Incorporated</OrganisationName>
  <PublisherCompoundName>Adobe Systems
Incorporated</PublisherCompoundName>
</Publisher>
</Document>
<ExternalSignature>
  <ExternalSignatureID>602</ExternalSignatureID>
  <Signature>tiff</Signature>
  <SignatureType>File extension</SignatureType>
</ExternalSignature>
<ExternalSignature>
  <ExternalSignatureID>685</ExternalSignatureID>
  <Signature>tif</Signature>
  <SignatureType>File extension</SignatureType>
</ExternalSignature>
<InternalSignature>
  <SignatureID>9</SignatureID>
  <SignatureName>TIFF generic (little-
endian)</SignatureName>
  <SignatureNote>Header</SignatureNote>
  <ByteSequence>
    <ByteSequenceID>183</ByteSequenceID>
    <PositionType>Absolute from BOF</PositionType>
    <Offset>0</Offset>
    <MaxOffset>
</MaxOffset>
    <IndirectOffsetLocation>
</IndirectOffsetLocation>
    <IndirectOffsetLength>
</IndirectOffsetLength>
    <Endianness>
</Endianness>
    <ByteSequenceValue>49492A00</ByteSequenceValue>
  </ByteSequence>
</InternalSignature>
<InternalSignature>
  <SignatureID>10</SignatureID>
  <SignatureName>TIFF generic (big-endian)</SignatureName>
```

PRONOM Format Registry Entry for TIFF in XML *(continued)*

```
<SignatureNote>Header</SignatureNote>
<ByteSequence>
  <ByteSequenceID>184</ByteSequenceID>
  <PositionType>Absolute from BOF</PositionType>
  <Offset>0</Offset>
  <MaxOffset>
</MaxOffset>
  <IndirectOffsetLocation>
</IndirectOffsetLocation>
  <IndirectOffsetLength>
</IndirectOffsetLength>
  <Endianness>
</Endianness>
  <ByteSequenceValue>4D4D002A</ByteSequenceValue>
</ByteSequence>
</InternalSignature>
<RelatedFormat>
  <RelationshipType>Has lower priority
than</RelationshipType>
  <RelatedFormatID>672</RelatedFormatID>
  <RelatedFormatName>Exchangeable Image File Format
(Uncompressed)</RelatedFormatName>
  <RelatedFormatVersion>2.2</RelatedFormatVersion>
</RelatedFormat>
<RelatedFormat>
  <RelationshipType>Has lower priority
than</RelationshipType>
  <RelatedFormatID>673</RelatedFormatID>
  <RelatedFormatName>Exchangeable Image File Format
(Uncompressed)</RelatedFormatName>
  <RelatedFormatVersion>2.1</RelatedFormatVersion>
</RelatedFormat>
<RelatedFormat>
  <RelationshipType>Has lower priority
than</RelationshipType>
  <RelatedFormatID>752</RelatedFormatID>
  <RelatedFormatName>Exchangeable Image File Format
(Uncompressed)</RelatedFormatName>
  <RelatedFormatVersion>2.0</RelatedFormatVersion>
</RelatedFormat>
<RelatedFormat>
  <RelationshipType>Is subsequent version
of</RelationshipType>
  <RelatedFormatID>611</RelatedFormatID>
```

PRONOM Format Registry Entry for TIFF in XML *(continued)*

```
<RelatedFormatName>Tagged Image File
Format</RelatedFormatName>
  <RelatedFormatVersion>5</RelatedFormatVersion>
</RelatedFormat>
<RelatedFormat>
  <RelationshipType>Is supertype of</RelationshipType>
  <RelatedFormatID>795</RelatedFormatID>
  <RelatedFormatName>Digital Negative Format
(DNG)</RelatedFormatName>
  <RelatedFormatVersion>1.1</RelatedFormatVersion>
</RelatedFormat>
<RelatedFormat>
  <RelationshipType>Is supertype of</RelationshipType>
  <RelatedFormatID>796</RelatedFormatID>
  <RelatedFormatName>Tagged Image File Format for Image
Technology (TIFF/IT)</RelatedFormatName>
  <RelatedFormatVersion>
</RelatedFormatVersion>
</RelatedFormat>
<RelatedFormat>
  <RelationshipType>Is supertype of</RelationshipType>
  <RelatedFormatID>797</RelatedFormatID>
  <RelatedFormatName>Tagged Image File Format for
Electronic Still Picture Imaging (TIFF/EP)</RelatedFormatName>
  <RelatedFormatVersion>
</RelatedFormatVersion>
</RelatedFormat>
<RelatedFormat>
  <RelationshipType>Is supertype of</RelationshipType>
  <RelatedFormatID>799</RelatedFormatID>
  <RelatedFormatName>Tagged Image File Format for Internet
Fax (TIFF-FX)</RelatedFormatName>
  <RelatedFormatVersion>
</RelatedFormatVersion>
</RelatedFormat>
<CompressionType>
  <CompressionID>11</CompressionID>
  <CompressionName>Baseline JPEG</CompressionName>
  <CompressionVersion>
</CompressionVersion>
  <CompressionAliases>Discrete Cosine Transform, DCT,
JPEG</CompressionAliases>
  <CompressionFamilies>
</CompressionFamilies>
```

## PRONOM Format Registry Entry for TIFF in XML *(continued)*

<Description>The JPEG compression algorithm was developed in 1990 by the Joint Photographic Experts Group of ISO and CCITT, for the transmission of colour and greyscale images. It is a lossy technique which provides best compression rates with complex 24-bit (True Colour) images. It achieves its effect by discarding image data which is imperceptible to the human eye, using a technique called Discrete Cosine Transform (DCT). It then applies Huffman encoding to achieve further compression. The JPEG specification allows users to set the degree of compression, using an abstract Quality Setting. This provides a trade-off between compression rate and image quality. It is important to note that the Quality Setting is not an absolute value, with different JPEG encoders use different scales, and that even the maximum quality setting for baseline JPEG involves some loss. JPEG compression is most commonly used in the JPEG File Interchange Format (JFIF), SPIFF and TIFF.</Description>

<Lossiness>Lossy</Lossiness>

<ReleaseDate>01 Jan 1994</ReleaseDate>

<WithdrawnDate>

</WithdrawnDate>

<CompressionDocumentation>ISO/IEC 10918-1: 1994, Information technology - Digital compression and coding of continuous-tone still images: Requirements and guidelines</CompressionDocumentation>

<CompressionIPR>

</CompressionIPR>

<CompressionNote>

</CompressionNote>

<CompressionIdentifier>

<Identifier>x-cmp/11</Identifier>

<IdentifierType>PUID</IdentifierType>

</CompressionIdentifier>

<Developer>

<DeveloperID>112</DeveloperID>

<DeveloperName>

</DeveloperName>

<OrganisationName>International Standards Organisation</OrganisationName>

<DeveloperCompoundName>International Standards Organisation</DeveloperCompoundName>

</Developer>

</CompressionType>

<CompressionType>

<CompressionID>12</CompressionID>

PRONOM Format Registry Entry for TIFF in XML *(continued)*

```
<CompressionName>Lempel-Ziff-Welch</CompressionName>
<CompressionVersion>
</CompressionVersion>
<CompressionAliases>LZW</CompressionAliases>
<CompressionFamilies>
</CompressionFamilies>
<Description>The Lempel-Ziff-Welch compression algorithm
was developed by Terry Welch in 1984, as a modification of the
LZ78 compressor. It is a lossless technique which can be applied
to almost any type of data, but is most commonly used for image
compression. LZW compression is effective on images with colour
depths from 1-bit (monochrome) to 24-bit (True Colour). LZW
compression is encountered in a range of common graphics file
formats, including TIFF and GIF.</Description>
<Lossiness>Lossless</Lossiness>
<ReleaseDate>01 Jan 1984</ReleaseDate>
<WithdrawnDate>
</WithdrawnDate>
<CompressionDocumentation>Welch, T A, 1984, A technique
for high performance data compression, IEEE Computer, 17:
6</CompressionDocumentation>
<CompressionIPR>
</CompressionIPR>
<CompressionNote>
</CompressionNote>
<CompressionIdentifier>
<Identifier>x-cmp/12</Identifier>
<IdentifierType>PUID</IdentifierType>
</CompressionIdentifier>
<Developer>
<DeveloperID>121</DeveloperID>
<DeveloperName>T A Welch</DeveloperName>
<OrganisationName>
</OrganisationName>
<DeveloperCompoundName>T A Welch / [No organisation
specified]</DeveloperCompoundName>
</Developer>
</CompressionType>
<CompressionType>
<CompressionID>13</CompressionID>
<CompressionName>Run Length Encoding</CompressionName>
<CompressionVersion>
</CompressionVersion>
<CompressionAliases>RLE</CompressionAliases>
```

## PRONOM Format Registry Entry for TIFF in XML *(continued)*

```
<CompressionFamilies>
</CompressionFamilies>
<Description>Run length encoding (RLE) is perhaps the
simplest image compression technique in common use. RLE
algorithms are lossless, and work by searching for runs of bits,
bytes, or pixels of the same value, and encoding the length and
value of the run. As such, RLE achieves best results with images
containing large areas of contiguous colour, and especially
monochrome images. Complex colour images, such as photographs, do
not compress well – in some cases, RLE can actually increase
the file size. There are a number of RLE variants in common use,
which are encountered in the TIFF, PCX and BMP graphics
formats.</Description>
<Lossiness>Lossless</Lossiness>
<ReleaseDate>
</ReleaseDate>
<WithdrawnDate>
</WithdrawnDate>
<CompressionDocumentation>
</CompressionDocumentation>
<CompressionIPR>
</CompressionIPR>
<CompressionNote>
</CompressionNote>
<CompressionIdentifier>
  <Identifier>x-cmp/13</Identifier>
  <IdentifierType>PUID</IdentifierType>
</CompressionIdentifier>
</CompressionType>
<CompressionType>
  <CompressionID>14</CompressionID>
  <CompressionName>CCITT T.4</CompressionName>
  <CompressionVersion>
  </CompressionVersion>
  <CompressionAliases>CCITT Group 3</CompressionAliases>
  <CompressionFamilies>
  </CompressionFamilies>
  <Description>Officially known as CCITT T.4, Group 3 is a
compression algorithm developed by the International Telegraph
and Telephone Consultative Committee in 1985 for encoding and
compressing 1-bit (monochrome) image data. Its primary use has
been in fax transmission, and it is optimised for scanned printed
or handwritten documents. Group 3 is a lossless algorithm, of
which two forms exist: one-dimensional (which is a modified
```

## PRONOM Format Registry Entry for TIFF in XML *(continued)*

version of Huffman encoding) and two-dimensional, which offers superior compression rates. Due to its origin as a data transmission protocol, Group 3 encoding incorporates error detection codes. Group 3 compression is most commonly used in the TIFF file format.</Description>

```
<Lossiness>Lossless</Lossiness>
<ReleaseDate>01 Jan 1985</ReleaseDate>
<WithdrawnDate>
</WithdrawnDate>
<CompressionDocumentation>CCITT Blue Book, 1989, Volume
VII, Fascicle VII.3: Terminal equipment and protocols for
telematic services, recommendations T.0 -
T.63</CompressionDocumentation>
<CompressionIPR>
</CompressionIPR>
<CompressionNote>
</CompressionNote>
<CompressionIdentifier>
<Identifier>x-cmp/14</Identifier>
<IdentifierType>PUID</IdentifierType>
</CompressionIdentifier>
<Developer>
<DeveloperID>124</DeveloperID>
<DeveloperName>
</DeveloperName>
<OrganisationName>International Telecommunication
Union</OrganisationName>
<DeveloperCompoundName>International Telecommunication
Union</DeveloperCompoundName>
</Developer>
</CompressionType>
<CompressionType>
<CompressionID>15</CompressionID>
<CompressionName>CCITT T.6</CompressionName>
<CompressionVersion>
</CompressionVersion>
<CompressionAliases>CCITT Group 4</CompressionAliases>
<CompressionFamilies>
</CompressionFamilies>
<Description>Officially known as CCITT T.6, Group 4 is a
compression algorithm developed by the International Telegraph
and Telephone Consultative Committee as a development of the two-
dimensional Group 3 standard for encoding and compressing 1-bit
(monochrome) image data. It is faster, and offers compression
```

## PRONOM Format Registry Entry for TIFF in XML *(continued)*

rates which are typically double those of Group 3. Like Group 3, it is lossless and designed for 1-bit images. However, being designed as a storage rather than transmission format, it does not incorporate the error detection and correction functions of Group 3 compression. Group 4 compression is most commonly used in the TIFF file format.</Description>

```
<Lossiness>Lossless</Lossiness>
<ReleaseDate>01 Jan 1989</ReleaseDate>
<WithdrawnDate>
</WithdrawnDate>
<CompressionDocumentation>CCITT Blue Book, 1989, Volume
VII, Fascicle VII.3: Terminal equipment and protocols for
telematic services, recommendations T.0 -
T.63</CompressionDocumentation>
<CompressionIPR>
</CompressionIPR>
<CompressionNote>
</CompressionNote>
<CompressionIdentifier>
<Identifier>x-cmp/15</Identifier>
<IdentifierType>PUID</IdentifierType>
</CompressionIdentifier>
<Developer>
<DeveloperID>124</DeveloperID>
<DeveloperName>
</DeveloperName>
<OrganisationName>International Telecommunication
Union</OrganisationName>
<DeveloperCompoundName>International Telecommunication
Union</DeveloperCompoundName>
</Developer>
</CompressionType>
<FormatProperties>
<FormatProperty>
<PropertyName>Compression Type</PropertyName>
<PropertyDescription>
</PropertyDescription>
<PropertyType>Instance</PropertyType>
<PropertyRiskScore>
</PropertyRiskScore>
<PropertyHighRisk>
</PropertyHighRisk>
<PropertyValueDescription>
</PropertyValueDescription>
```

PRONOM Format Registry Entry for TIFF in XML *(continued)*

```
<PropertyRisks />
</FormatProperty>
<FormatProperty>
  <PropertyName>Compression Level</PropertyName>
  <PropertyDescription>
</PropertyDescription>
  <PropertyType>Instance</PropertyType>
  <PropertyRiskScore>
</PropertyRiskScore>
  <PropertyHighRisk>
</PropertyHighRisk>
  <PropertyValueDescription>
</PropertyValueDescription>
  <PropertyRisks />
</FormatProperty>
<FormatProperty>
  <PropertyName>Image Width</PropertyName>
  <PropertyDescription>
</PropertyDescription>
  <PropertyType>Instance</PropertyType>
  <PropertyRiskScore>
</PropertyRiskScore>
  <PropertyHighRisk>
</PropertyHighRisk>
  <PropertyValueDescription>
</PropertyValueDescription>
  <PropertyRisks />
</FormatProperty>
<FormatProperty>
  <PropertyName>Image Height</PropertyName>
  <PropertyDescription>
</PropertyDescription>
  <PropertyType>Instance</PropertyType>
  <PropertyRiskScore>
</PropertyRiskScore>
  <PropertyHighRisk>
</PropertyHighRisk>
  <PropertyValueDescription>
</PropertyValueDescription>
  <PropertyRisks />
</FormatProperty>
<FormatProperty>
  <PropertyName>Colour Space</PropertyName>
  <PropertyDescription>
```

PRONOM Format Registry Entry for TIFF in XML *(continued)*

```
</PropertyDescription>
<PropertyType>Instance</PropertyType>
<PropertyRiskScore>
</PropertyRiskScore>
<PropertyHighRisk>
</PropertyHighRisk>
<PropertyValueDescription>
</PropertyValueDescription>
<PropertyRisks />
</FormatProperty>
<FormatProperty>
  <PropertyName>Sampling Frequency Unit</PropertyName>
  <PropertyDescription>
  </PropertyDescription>
  <PropertyType>Instance</PropertyType>
  <PropertyRiskScore>
  </PropertyRiskScore>
  <PropertyHighRisk>
  </PropertyHighRisk>
  <PropertyValueDescription>
  </PropertyValueDescription>
  <PropertyRisks />
</FormatProperty>
<FormatProperty>
  <PropertyName>X Sampling Frequency</PropertyName>
  <PropertyDescription>
  </PropertyDescription>
  <PropertyType>Instance</PropertyType>
  <PropertyRiskScore>
  </PropertyRiskScore>
  <PropertyHighRisk>
  </PropertyHighRisk>
  <PropertyValueDescription>
  </PropertyValueDescription>
  <PropertyRisks />
</FormatProperty>
<FormatProperty>
  <PropertyName>Y Sampling Frequency</PropertyName>
  <PropertyDescription>
  </PropertyDescription>
  <PropertyType>Instance</PropertyType>
  <PropertyRiskScore>
  </PropertyRiskScore>
  <PropertyHighRisk>
```

PRONOM Format Registry Entry for TIFF in XML *(continued)*

```
</PropertyHighRisk>
<PropertyValueDescription>
</PropertyValueDescription>
<PropertyRisks />
</FormatProperty>
<FormatProperty>
  <PropertyName>Bits Per Sample</PropertyName>
  <PropertyDescription>
</PropertyDescription>
  <PropertyType>Instance</PropertyType>
  <PropertyRiskScore>
</PropertyRiskScore>
  <PropertyHighRisk>
</PropertyHighRisk>
  <PropertyValueDescription>
</PropertyValueDescription>
  <PropertyRisks />
</FormatProperty>
<FormatProperty>
  <PropertyName>Samples Per Pixel</PropertyName>
  <PropertyDescription>
</PropertyDescription>
  <PropertyType>Instance</PropertyType>
  <PropertyRiskScore>
</PropertyRiskScore>
  <PropertyHighRisk>
</PropertyHighRisk>
  <PropertyValueDescription>
</PropertyValueDescription>
  <PropertyRisks />
</FormatProperty>
<FormatProperty>
  <PropertyName>Byte Order</PropertyName>
  <PropertyDescription>
</PropertyDescription>
  <PropertyType>Instance</PropertyType>
  <PropertyRiskScore>
</PropertyRiskScore>
  <PropertyHighRisk>
</PropertyHighRisk>
  <PropertyValueDescription>
</PropertyValueDescription>
  <PropertyRisks />
</FormatProperty>
```

PRONOM Format Registry Entry for TIFF in XML *(continued)*

```
</FormatProperties>  
</FileFormat>
```

## Appendix F: Geospatial File Format Definitions

This appendix is available as an archive (ngdafdd.zip) and includes the format definitions listed with fdd\* xml suffixes on pages 3-4. The xml files were generated using Altova XML spy and the Library of Congress fdd.xsd schema .

An Altova XMLSpy project file(locngdaproject.spp) is included, as are the template fddcreation.xml, and the LOC schema LOC schema(fdd.xsd). Supporting files include: fddcreation.sps – An Altova StyleVision Power Stylesheet & htmltypes.xsd (a schema called by fdd.xsd).

*For more information regarding the content of this appendix contact Julie Sweetkind-Singer ([sweetkind@stanford.edu](mailto:sweetkind@stanford.edu)).*