DDAP, TIFF/IT-P1, PDF/X-1

What's It All About?
Where Do We Stand in Terms of Data Exchange?

by David McDowell
Eastman Kodak Company/Kodak Polychrome Graphics

The 1997 March/April issue of *The Prepress Bulletin* carried an IPA Standards Update titled "Digital Data Exchange, A Report Card on CGATS Efforts to Meet DDAP Requirements." It is time to provide an update on the progress that has been made in the last year, and to also suggest some future directions that are anticipated.

Some Background

The DDAP (Digital Distribution of Advertising for Publications) Association, which represents the advertisers, separators, publishers, and printers involved in the publication arena, has prepared a User Requirements Specification for the Digital Distribution of Advertising for Publications. This has been a key input to the standards community, and specifically SC6 of CGATS (Committee for Graphic Arts Technologies Standards), as it has worked to put in place the formal standards necessary to enable exchange of graphic arts materials.

Although the standards activities have not been restricted to the DDAP requirements, the committee has limited their concerns to the exchange of material in a form ready for final print production (print-ready). That simply recognized that much of the material in earlier stages of production often was, of necessity, either less than fully defined and/or in vendor specific proprietary data formats.

Even though the focus has been on print-ready material, the standards community still faces the need to handle both object-based and raster-based data structures—both to meet the DDAP requirements and to meet the needs of the industry. Raster- and object-based are the terms that are being used to differentiate between the file formats needed to support systems of the type we typically have called CEPS (Color Electronic Prepress Systems) and those typical of the DTP (Desk Top Publishing) world.

Both of these acronyms are inaccurate but do serve to identify the two different ways that data are handled within our industry.

Raster-based systems deal exclusively with data that are organized in rows of pixel data at varying resolutions and encodings. The most common of these are picture data (often called contone or CT data), line work data, and high resolution contone data used to define edges, trapping, etc.

Object-oriented data on the other hand are made up of contone raster data for pictures, vector lines and fills for line art, and ASCI character strings and font data for text.

Both data types must eventually be converted to the on/off signals needed by the image or plate setter before the final printing materials can be prepared. However, that usually occurs in the RIP (raster image processor) associated with the output device. In addition, in some workflows it is convenient to convert object-oriented data into raster data at some point prior to final production. (More about RIP-IT-TO-TIFF/IT later.)

The initial standards activities focused on handling raster data—first on magnetic tape and later in media independent TIFF formats. The current raster standard is ISO 12639, *Graphic technology—Pre-*

Standards articles published by the IPA Standards Committee. Editors: Larry Warter, Fuji Photo Film, & David McDowell, Eastman Kodak.

When Adobe made the PDF format publicly available, it became feasible to work with them to extend the format to develop a standard that defined appropriate usage for graphic arts.

press digital data exchange—Tag image file format for image technology (TIFF/IT).

Because each DTP program initially seemed to have its own proprietary file format, work on standard file formats for object-based data followed much more slowly. Early work, by the IT8 and CGATS standards committees, led to the conclusion that development of an object-oriented file format by the standards community itself was impractical.

However, when Adobe made the PDF file format publicly available, it became feasible to work with Adobe to extend the format to meet the needs of the graphic arts and to then develop a standard that defined appropriate usage of the format in graphic arts applications. This is the path that was taken in developing *CGATS.12, Graphic technology—Prepress data exchange—Use of PDF for composite data*.

What About TIFF/IT?

The goals of the standards community, in developing TIFF/IT, were to carry forward the original IT8 magnetic tape formats into a media independent version. These, in turn, were an amalgamation of the formats used by the initial CEPS vendors. It is based on the Adobe TIFF 6.0 file specification and both extends TIFF 6 by adding additional tags and restricts it by limiting the tags that may be used and the values within tags to those appropriate to the graphic arts industry. It has also added a new file structure called FP (final page), which allows the various files needed to define a complete page to be grouped together. Currently the FP structure is included as an informative (recommended but not required) part of the standard.

The file types supported by TIFF/IT include CT (continuous tone picture data), LW (color line art data), HC (high resolution continuous tone image data), MP (monochrome picture image data), BP (binary picture image data), and BL (binary line art image data). The initial version of TIFF/IT was documented in an ANSI standard known as IT8.8 (which had the same title as the later ISO standard). IT8.8 was used as input to the work in ISO/TC130 which resulted in ISO 12639.

TIFF/IT-P1 is used widely for the exchange of advertising and editorial material. It has been endorsed by many groups worldwide as the preferred format for the exchange of raster data.

Within TIFF/IT, as defined in ISO 12639, provision is made for a simplified conformance level called profile 1 (P1). By limiting the options for most tag values (variables) to those most commonly used by the DTP industry, P1 maximizes the compatibility between the CEPS and DTP worlds. TIFF/IT-P1 is the version that has seen the most widespread acceptance by both users and equipment vendors. In virtually all applications the FP capability has been included even though it is only informative.

Is More Work Needed?

Currently TIFF/IT-P1 is being used relatively widely for the exchange of advertising and editorial material in the publication industry. It has been endorsed, and used successfully, by many groups worldwide as the preferred format for the exchange of raster data. However, as it is being used, additional needed features are being identified. Currently ISO/TC130/WG2 has an active task force (under the leadership of Bruce Shifrin), which is proposing and defining the capabilities to be included in an addendum. An important ground rule being followed is that the current P1 profile will not be modified. Additional capabilities may be added to the basic TIFF/IT standard and/or additional profiles (P2, etc.) may be defined.

The current topics under consideration, and a brief report on their status, include:

•The FP capability will be both enhanced and made normative. The key enhancements being discussed include the requirement for preview images and provision to allow multiple instances of file types within an FP.

•Copy-dot files are a major concern in the direct-to workflow and the task force has determined that a TIFF/IT-CT file can be configured

as a planar interleaved file with 1 bit per sample. Although not part of P1, this will accommodate copy-dot files. Consideration is also being given to identifying this as a unique file type.

•Data compression has also been raised as a concern in the transmission of ads. While there are no restrictions (or guidelines) on the use of compression when applied to a complete file, TIFF/IT makes no provision for compression within the file structure itself. As both the LW and HC file types use run length encoding as an integral part of the file structure, additional compression of these file types is not appropriate. CT files are therefore the only logical candidates. Consideration is being given to adding FLATE (ZIP) as the lossless compression for traditional CT files, or ITU Group 4 for planar interleaved copy-dot files.

•A need has been expressed to be able to allow multiple FPs to be grouped (several individual ads each represented by an FP). A super FP or group FP structure is being investigated to answer this need. It would be similar in some ways to the FP structure, and as with the FP structure, a key consideration is that it not "break" a traditional TIFF reader.

•There is a need to have a flag that indicates the state of trapping of a particular file or group of files. This is being investigated with the intent of providing the capability through a TIFF 6 private tag assigned to CGATS for TIFF/IT use.

•Consideration is also being given to incorporating an ICC compatible CMYK input profile into tag 34029 (color characterization) to specify the exact CMYK characteristics involved. This would allow TIFF/IT and PDF/X to use the same technique for color data definition.

•In the area of color, three additional issues are under discussion: ways to handle spot color, the definition of more than 256 colors in an LW file, and the use of color spaces other than CMYK. The discussion of these issues is still in the preliminary stages.

Where Do We Stand with PDF?
The primary focus of CGATS/SC6 has been to work with the DDAP User Requirements Document, users, and representatives from Adobe to ensure all the features required to allow exchange of graphic arts print-ready material were available in the PDF format. The reference for this work is the *Portable Document Format Reference Manual*, Version 1.2 dated November 27, 1997. As a result of these discussions, a set of extensions to PDF version 1.2 have been documented in Adobe Technical Note #5188, *PDF features to facilitate ANSI CGATS.12, PDF/X* dated February 4, 1998.

The new features added by this Technical Note include a "trapped" key in the PDF Info dictionary, an "ICCBased" color space, a "file specification table," and an "EmbeddedFile" named tree. The committee believes the addition of these features allows the PDF file format to carry all of the information identified as required for graphic arts data exchange.

While the committee was working to ensure all requirements were being identified, SC6 was also developing a standard to specify the methods for the use of the Portable Document Format (PDF) for dissemination of object-based composite digital data, in a single exchange, that are complete and ready for final print reproduction.

That standard is *CGATS.12/1, Graphic technology—Prepress digital data exchange—Use of PDF for composite data—Part 1: Complete exchange (PDF/X-1).* Early in their work SC6 determined two separate specifications needed to be written, one for the exchange of complete composite digital data

with tight restrictions to assure successful complete exchanges and another for the exchange of partial composite digital data allowing more lenient exchanges but requiring greater communication between the parties in the exchange for successful rendering of the data. As its title suggests, this first standard covers complete (sometimes referred to as "blind") exchanges. Part 2, which will follow shortly, will cover exchange of partial data.

The term "composite" may be confusing. Unfortunately, the graphic arts has no consistent way to refer to the various parts of the job during the typical workflow. Therefore, CGATS/SC6 had to come up with some definitions to allow them to communicate—particularly with the computer folks. One of these is "composite entity," which is defined as:

> *A unit of work with all text, graphics and image elements prepared for final print reproduction. A composite entity can represent a single page for printing, a portion of a page or a combination of pages.*

What Does CGATS.12/1 Say?

First, and maybe most important, CGATS.12/1 requires that "all components of a composite entity shall be contained in the body of a single PDF/X-1 file." The new features added to the PDF allow external files, identified by OPI references, to be incorporated into the PDF file for exchange, but later separated back out for ease of workflow management.

It also requires that all PDF resources used in the file—including all fonts, font metrics, and font encodings—be included. It further requires that all process color data be CMYK data properly prepared for a single intended printing condition (for example, a PDF/X-1 file shall not have two objects based on different characterized printing conditions). An ICC color input profile is used to carry both a description of the intended printing condition and the data to allow proper display of the images.

Discussions are currently under way to make a CMYK input profile, based on SWOP characterization data, available through CGATS so that those users not presently using color management will have no problems meeting this requirement.

The allowable file types for OPI reference and embedding are TIFF/IT-FP/P1, TIFF/IT-CT/P1, TIFF/IT-LW/P1, TIFF/IT-HC/P1, TIFF/IT-BP/P1, TIFF/IT-BL/P1, TIFF/IT-MP/P1, TIFF 6.0, EPS, and Desktop Color Separation (DCS) Version 2.0.

TIFF 6.0 files are limited to CMYK formats conforming to the comparable TIFF/IT restrictions, and EPS files are limited to CMYK raster content consistent with the attributes of TIFF/IT-P1.

A preview image is required and the trapping key must be used. Files may either be fully trapped or untrapped. Files with some elements trapped and others not, are not permitted. Data compression is not required. If data compression for PDF objects is desired, only FLATE and RunLength lossless data compression methods are allowed, although extemally referenced files may have their own restrictions or extensions.

The Portable Document Format Reference Manual, Version 1.2, Adobe Technical Note #5188, and the current draft of CGATS.12/1 are available from the CGATS and TC130/WG2 secretariat, NPES The Association for Suppliers of Printing and Publishing Technologies, at http://www.npes.org/standards/workroom.htm#CGATS.

Where Does that Leave Us?

The current vision is that PDF/X-1 files can always be rendered to create the raster files used by the CEPS type systems and defined in TIFF/IT-P1. Further, TIFF/IT-P1 files can be incorporated within a PDF/X-1 file, with all of the accompanying information intact, using the appropriate operators. More succinctly, PDF can always be ripped to TIFF/IT and TIFF/IT can always have a PDF wrapper applied, and either can be used as a standardized exchange format. (See diagram on next page.)

PDF/X and TIFF/IT provide the file structure to accomplish most data exchange tasks required by the graphic arts. The additions to TIFF/IT described earlier should complete this capability.

Why Do We Still Need TIFF/IT?

A question that is often asked is, "Why do we still need TIFF/IT?" There is no single answer to this question. Of the many answers some of the key points relate to the use of existing equipment, the level of predictability and time involved in the PostScript interpretation task, and the experience base in the industry. We must remember that as part of the output process a PDF file is converted to PostScript for rendering and that rendering task involves conversion of all of the vector and font data to raster data.

This step, while getting more predictable, is still RIP dependent. TIFF/IT files require far less processing at output and are, therefore, generally considered more predictable between output sites. This predictability comes at the cost of editability, which can be viewed either as a benefit or liability, depending on the particular situation involved. Further, within a PDF file, raster data must be carried in some defined file format.

All contone picture data are raster and, for graphic arts applications, are logically stored as TIFF/IT-CT files. In addition, as suggested earlier, copy-dot files should be stored as planar interleaved TIFF/IT-CT files with 1 bit per sample. This suggests that, even in an object-oriented architecture, these TIFF/IT file types are an integral part of the graphic arts use of PDF/X.

It would therefore seem that there is a significant need for both PDF and TIFF/IT in the graphic arts workflow. Neither one can do the job alone.

When Can We Use It?

The definitions of the PDF/X and TIFF/IT formats are essentially ready now. However, file formats are not products that we can use to accomplish the data processing tasks required. Vendors need to create products that provide the tools to accomplish the data processing and manipulation required to implement these file formats. More important, vendors are not going to implement unless users clearly state their need for, and willingness to purchase, products that comply with these standards.

In their meeting in early March the DDAP Association listed many companies that had implemented support for TIFF/IT-P1 in their products. They also reported results of their testing of some of these products.

PDF/X is too new for anyone to have products available. It is expected that these will follow quickly and that the DDAP Association will also track and test the arrival of the offerings to the industry.

The Bottom Line!

Most important of all, the availability and implementation of these standards will allow graphic arts data, both object based and raster based, in final composited form, to be exchanged using accredited standards.

Simply being able to exchange data will not solve all of the problems of the electronic distribution of print-ready data. There are still many areas of process control, print characterization, color management, etc., that will keep the standards committees and the industry busy for some time to come. The file format issue, however, seems to be well in hand. **IPA**

As shown in the illustration above, a PDF file (PDF1) would typically contain text as ASCII strings (along with font data for rendering), line art as geometric drawing data and fills, and CT picture data. The CT picture data are raster data that are common to both PDF and TIFF/IT and can move easily between the two file structures. The rest of the object oriented data can be easily RIPped to the LW and HC files used by the raster based systems, as is routinely done today in a CEP system. It is important to note that a second PDF file (PDF2) can be used as a wrapper on the newly created family of TIFF/IT files, which are linked by an FP (final-page) file. Alternatively, this second PDF file can combine both TIFF/IT and/or PDF inputs to create a larger printing element for another level of data exchange.