

THE IMPACT OF PDF 2.0 ON PRINT PRODUCTION

Contents

Introduction	3
Standardization of PDF	4
The first “post-Adobe” PDF specification	5
Increased clarity; reduced variation	5
Evolutionary, not revolutionary	6
Risks and recommendations	6
Overview of new features in PDF 2.0	7
Drivers for update	8
PDF 2.0 for print	8
Color and output intents	8
Transparency	9
Inheritance of transparency color spaces	9
Clarifications on when object colors must be transformed to the blend color space	11
ColorDodge & Burn	11
Halftones	11
Lists of spot functions in halftones	11
Halftone Origin (HTO)	12
Halftone selection in transparent areas	13
Document parts	13
PDF 2.0 headers	13
UFT-8 text strings	14
Security	14
Deprecation and what this means in PDF	15

Introduction

In August 2017 the first new version of the Portable Document Format (PDF) since 2008 was published.

This white paper provides a summary of the changes that will affect print production and is designed to give the industry an understanding of the implications for their workflows.

It provides background information on the new standard, insight into what it means for print, and recommendations around implementation planning.

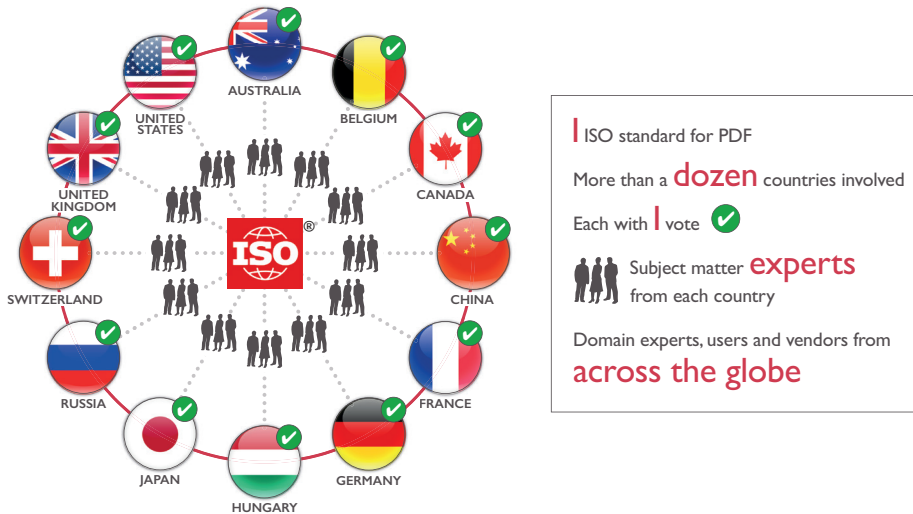
The White Paper's author, Global Graphics' CTO Martin Bailey, is the primary UK expert to the ISO committees working on PDF, PDF/X and PDF/VT.

Standardization of PDF

In early 2007 Adobe Systems started working with the International Standards Organization (ISO), requesting them to take over stewardship of the PDF specification, recognizing that the broader review from many different vendors that ISO would bring would be a good thing.

An ISO standard is developed in a committee (in this case Working Group 8 of Subcommittee 2 of Technical Committee 171 – there are a lot of groups working on a lot of standards in ISO). That committee is composed of subject matter experts from many countries. As in most other ISO committees those subject matter experts are largely employed by vendors and academic institutions who are experts because they work in the relevant field every day.

When the committee believes that they have completed the work for a new milestone of the standard ... or for the final text ... the standard is put to a ballot where each nation who has signed up to participate in the development of that standard gets a single vote. There's a slightly complex formula for assessing the result, but in essence it needs a 2/3 vote in favor to pass. In all ballots but the last one, nations are also invited to submit comments to the current text which are then reviewed in preparing for the next milestone.



This very wide and thorough review is an excellent way of teasing out any areas where people may read the standard differently, or where there is insufficient detail for multiple implementers to achieve consistent results. But it also obviously has an impact on the time taken to bring out the new version.

The first “post-Adobe” PDF specification

The first ISO PDF standard, ISO 32000-1:2008, was very much a rewording of the last Adobe PDF Reference, 1.7, and was intended to be technically identical to PDF 1.7.

ISO standards enforce use of terminology, e.g. so that requirements and recommendations can be clearly distinguished, and the majority of the differences between PDF 1.7 and ISO 32000-1 were down to that initial clarification. In other words, it was an improved PDF 1.7 rather than a truly ISO version.

That means that ISO 32000-2:2017, or PDF 2.0 is the first real “post Adobe” version of PDF.

But it's also worth noting that, even when Adobe were managing the specification, there were thousands, probably tens of thousands, of other vendors implementing PDF from the ground up, without using any Adobe code and without paying any licenses to Adobe¹. So in that regard ISO 32000-2 makes no difference at all. Anybody with the skills and need to implement PDF in their own code can still do so without any legal impediment or additional licensing fees to pay

Increased clarity; reduced variation

But the text in some areas of PDF 1.7 was slightly incomplete in a number of other areas as well, and, inevitably, products from different vendors ended up producing slightly different results. To be fair, so did different versions of Adobe applications from time to time.

To illustrate the level of these differences, imagine a specification that says that you should drive west out of a city on a specific highway. Most people around the world would drive out on the right side of the highway, because they drive on the right in their own countries. A few would drive out on the left side of the highway. Both would be correct according to the specification.

In PDF 2.0 many of the cases where the wording was too vague have been identified and more detail added. For some people the biggest real value of the new version will be that many differences between products or vendors will disappear as a result, giving more consistency across implementations.

Those differences were not deliberate in the first place, and they weren't areas that vendors competed on. So increasing uniformity in processing PDF files will allow vendors to focus more energy on areas where they are competing and providing unique value to their users.

Evolutionary, not revolutionary

One of the guiding principles in developing PDF 2.0 was compatibility; that the billions of existing PDF files should still be usable and that the hundreds of thousands² of existing PDF writers and readers should be extensible to support the new standard without huge development efforts.

So a PDF 1.7 *something* file is very close in many ways to being PDF 2.0 compatible.

The PDF 2.0 standard doesn't explicitly require products that read PDF 2.0 files to be able to read anything earlier than PDF 2.0 ... but the structural changes from PDF 1.7 to PDF 2.0 are so small, and the commercial benefits to supporting earlier versions as well are so obvious that it would be very surprising to find a product that *only* read 2.0³.

The same will sort of work in the other direction. A reader written for PDF 1.7, say, will consume PDF 2.0 files that don't happen to use any of the new features just fine, which provides obvious benefits for users.

¹ Note that there may be a need to pay licence fees for specific technologies used by PDF, e.g. around patents for encryption.

Risks and recommendations

If a file has used some of the new features in PDF 2.0 those will usually be silently ignored by an older reader. PDF was designed to be very flexible, and to allow custom and proprietary data to be embedded virtually anywhere in the file structure. It does that by saying that a reader should simply ignore anything it doesn't recognize. To a PDF 1.7 reader, most new PDF 2.0 features are just objects that it won't recognize and should therefore ignore⁴.

This is the biggest risk area for people considering when and how to roll out PDF 2.0 support. Some readers will emit a warning that the file you're opening has a PDF version number that is not explicitly supported. That's helpful, but it can never be more than a hint to take care because that older reader doesn't know anything about any new features in the file; it cannot possibly know if they're important to you or to your workflow.

This means that the safest approach to adoption of PDF 2.0 is to ensure that all applications and tools that consume PDF are upgraded to support PDF 2.0 before you start thinking about upgrading the file creators.

Some tools are both readers and writers. If you use a stand-alone imposition tool that reads a PDF file and re-saves it as an imposed PDF file, for instance, that does both. So start at the back end of your workflow (probably the RIP, or DFE, or an integrated prepress workflow in a printing situation) and work upstream. That way you'll never be trying to consume PDF 2.0 in a product that doesn't really know what to do with it.

This is even more important if your business model requires that you take pre-created PDF files from your customers or agencies. In that situation it's always tempting to just tell your customers that you won't take PDF 2.0, but experience shows that kind of instruction is often ignored. The best thing to do is to upgrade to add PDF 2.0 support in your workflow as soon as your vendors can provide it.

The PDF 2.0 standard will be published in May 2017, but text that is stable enough to develop product from has been available since Q3 2016, so many vendors will already be well advanced in bringing their products up to speed. Even so, introducing support for a new PDF version does take time; it may well be late 2018 before all of the major vendors are ready to claim PDF 2.0 support.

² I don't think anyone really knows how many tools are available around the world for writing and manipulating PDF files, both commercial and in-house.

³ The one exception to this may well be around security, where supporting some older versions may incur additional licensing fees.

Overview of new features in PDF 2.0

Given the potential risks outlined in the previous section, what was important enough to add new features and introduce that risk? Clarifying areas and removing the potential for differences between multiple vendors doesn't necessarily mean adding new features.

Amongst the changes in PDF 2.0 are:

- Unencrypted wrapper document; allowing most of a document to be encrypted, but also to include unencrypted content that can be used to explain what is in, and how to access the encrypted part.
- 256-bit AES encryption and ECC-based certificates.
- CAdES signatures and long-term validation of signatures; building on ETSI standards and best practice around digital signatures.
- Support for Unicode in passwords and UTF-8 for strings used in metadata, etc.
- 3D annotations and support for PRC-format 3D models.
- Geospatial features; enriching interaction with documents containing maps and other geo data.
- Greatly expanded structure tagging, primarily in support of accessibility for disabled users.
- Associated files; allowing other files to be attached in appropriate structures within a document, e.g. for computer-readable copies, or for access to the source document for the PDF.
- Keys to request the use of black point compensation.
- Document part structures; carrying a tree structure of metadata associated with ranges of pages.
- Page-level Output Intents, and addition of spectral data in output intents.
- Halftone Origin (HTO), and rules around halftone selection in transparency regions.

Drivers for update

Probably the most significant individual area of work for PDF 2.0 has been around accessibility for disabled users, in support of assistive technologies for reading aloud, transformation to braille, re-formatting to view in large text sizes, etc.

This isn't directly relevant for most print environments, but it may well be important in driving adoption. Consumer advocates, government agencies and large enterprises will realize the value of tools using the new functionality in PDF 2.0 to help in achieving and exceeding the requirements for publication of accessible documents, e.g. to meet obligations around Section 508, and the equivalents in other jurisdictions.

PDF 2.0 for print

You'll see a lot of the new features in PDF 2.0 are not specifically relevant for print, which is as it should be. PDF is used for many different activities where reliable delivery of a digital document is valuable. But this white paper is specifically addressing those issues that are relevant for printing.

Color and output intents

The printing condition for which a job was created can be encapsulated in professional print production jobs by specifying an "output intent" in the PDF file. The output intent structure was invented for the PDF/X standards, at first in support of pre-flight, and later to enable color management at the print site to match that used in proofing at the design stage.

But the PDF/X standards only allow a single output intent to be specified for all pages in a job.

PDF 2.0 allows separate output intents to be included for every page individually. The goal is to support jobs where different media are used for various pages, e.g. for the first sheet for each recipient of a transactional print job, or for the cover of a saddle-stitched book. The output intents in PDF 2.0 are an extension of those described in PDF/X, and the support for multiple output intents will almost certainly be adopted back into PDF/X-6 and into the next PDF/VT standard over the next year.

But of course, like many improvements, this one does demand a little bit of care. A PDF 1.7 or existing PDF/X reader will ignore the new page level output intents and could therefore produce the wrong colors for a job that contains them.

In PDF 2.0 it's also now possible to embed spectral measurement data for spot colors in the form of CxF/X-4 files (ISO 17972-4), but work to standardize how that data is used for accurate emulation of spot colors, e.g. on a digital press, is ongoing. In other words this is a useful piece of forward thinking in the standard, but not yet as useful as one might hope in a real production workflow.

PDF 2.0 also now provides a way for the job (or even individual graphical elements within a job) to request that black-point compensation be used in color management. In many print shops black-point compensation is already turned on for all jobs already, so the new key probably won't be required.

In some cases, however, it can be valuable to turn black-point compensation on for graphics such as images to achieve the most attractive rendering, but off for vector graphics, and especially brand colors in logos, for the very best color match.

If the print shop isn't already using black-point compensation it's possible that their output profiles won't have been tested with it, which means requesting it in a PDF supplied for print may be slightly risky without some prior discussion.

Transparency

PDF 2.0 includes a number of changes around transparency, driven by what we've learned in the last few years about where the previous PDF standards could trip people up in real-world jobs.

Inheritance of transparency color spaces

Under certain circumstances a RIP will now automatically apply a color-managed (CIEBased) color space when a device color space (such as DeviceCMYK) is used in a transparent object. It will do that by inheriting it from a containing Form XObject or the current page.

That sounds very technical, but the bottom line is that it will now be much easier to get the correct color when imposing multiple PDF files from different sources together. That's especially the case when you're imposing PDF/X files that use different profiles in their output intents, even though they may all be intended for the same target printing condition. The obvious examples of this kind of use case are placing display advertising for publications or newspaper, or imposing for gang-printing.



The committee has tried hard to minimize the impact on existing workflows in making these improvements, but there will inevitably be some cases where a PDF 2.0 workflow will produce different results from at least some existing solutions, and this is one case where that could happen. But the kinds of construct where PDF 2.0 will produce different output are very uncommon in PDF files apart from in the cases where it will provide a benefit by allowing a much closer color match to the designer/advertiser's goal than could be achieved easily before.

Clarifications on when object colors must be transformed to the blend color space

The ISO PDF 1.7 standard, and all previous PDF specifications, were slightly vague about exactly when the color space of a graphical object involved with PDF transparency needed to be transformed into the blending color space. The uncertainty meant that implementations from different vendors could (and sometimes did) produce very different results.

Those statements have been greatly clarified in PDF 2.0.

This is another area where an upgrade to a PDF 2.0 workflow may mean that your jobs render slightly differently ... but the up-side is that if you run prepress systems or digital presses from multiple vendors they should now all produce output that is more similar to each other.

As a note to Harlequin RIP users, the new rules are in line with the way that Harlequin has always behaved; in other words, you won't see any changes in this area when you upgrade.

You may also find that vendors choose to apply these newly clarified processing requirements to older versions of PDF as well. Since they are just clarifications, and don't contradict the earlier standards and specifications that's a perfectly valid thing to do.

ColorDodge & Burn

It tends to be taken for granted that the older PDF specifications must match what Adobe® Acrobat® does, but that's not always correct. As an example, the implementation in Acrobat has never matched the formulae for the ColorDodge and ColorBurn transparency blending modes in the PDF specification.

In the interests of compatibility Harlequin was changed to match Acrobat rather than the specification many years ago. In PDF 2.0 the standard is finally catching up with reality and now both Acrobat and Harlequin will be formally 'correct'!

Halftones

Halftoning and screening are often thought of as completely understood and stable, but there were still a number of issues identified in real-world print environments that will benefit from changes in PDF 2.0.

Lists of spot functions in halftones

The PDF format allows a PDF file to specify the halftone to be used for screening output in a variety of ways. The simplest is to identify a spot function by name, but that method was constrained in versions of the PDF standard up to PDF 1.7, to use only names that were explicitly listed in the specification itself. This has been a significant limitation in some print sectors where custom halftones are common, such as flexography, gravure ... and pretty much everywhere apart from offset plate-making!

PDF 2.0 allows the PDF file to specify the halftone dot shape as a list of spot function names, and those names no longer need to be picked from the ones specified in the standard. The renderer should use the first named spot function in the list that it supports. This allows a single file to be created that can be used in a variety of RIPs that support different sets of proprietary halftones and to select the best one available in each RIP for that specific object.

This functionality is expected to be used mainly for high-quality flexo press work, where it's a key part of the workflow to specify which halftone should be used for each graphical element.

A PDF 1.7 reader will probably either error or completely ignore the screening information embedded in the PDF if a file using the new list form is encountered. In the flexo space that could easily cause problems on-press, so take care that you've upgraded your RIPs before you start to try rendering PDF files using this new capability.

Halftone Origin (HTO)

Very old versions of PDF (up to PDF 1.3) included a partial definition of an entry named HTP, which was intended to allow the location of the origin or phase of a halftone to be specified. That entry was unfortunately useless because it did not specify the coordinate system to apply and it was removed many years ago.

PDF 2.0 adds a new entry called HTO to achieve the same goal, but this time fully specified. The use case is anywhere where precise specification of the halftone phase is valuable. Examples include:

- Pre-imposed sheets for VLF plate-setters, where specifying the halftone phase for each imposed page can reduce the misalignment of halftones that can occur over very long distances
- Setting the halftone phase of each of a set of step-and-repeat labels to ensure that the halftone dots are placed in exactly the same position relative to the design in each instance.

A PDF 1.7 reader will simply ignore the new key, so there's no danger of new files causing problems in an older workflow. On the other hand, those older RIPs will render as they always have, which would be a missed opportunity for the target use cases.

Halftone selection in transparent areas

Up to PDF 1.7 there was a requirement to apply the “default halftone” in all areas involved in live PDF transparency. This was problematic for those print technologies where different halftones must be used for different object types to achieve maximum quality, e.g. for flexo. Following older PDF specifications almost always means using the wrong halftone for drop shadows. Those use transparency to achieve the correct color, but are also one of those cases where using the right halftone is really important to reproduce the highlight end of the shadow without objectionable artifacts.

PDF 2.0 effectively gives complete freedom to renderers to apply the supplied screening parameters in whatever way they see fit, but two example implementations are described to encourage similarity between implementations. One of those matches the requirements from PDF 1.7, while the other applies the screen defined for the top-most graphical element in areas where transparency was applied, matching requirements for the flexo market.

Document parts

The PDF/VT standard for variable data printing defines a structure of Document parts (commonly called DPart) that can be used to associate hierarchical metadata with ranges of pages within the document. In PDF/VT the purpose is to enable embedding of data to guide the application of different processing to each page range.

PDF 2.0 has added the Document parts structure into baseline PDF, although no associated semantics or required processing for that data have yet been defined.

It is anticipated that a new ISO standard on workflow control (ISO 21812, expected to be published around the end of 2017) will make use of the DPart structure, as will the next version of PDF/VT. The specification in PDF 2.0 is therefore good forward planning, but not particularly useful until such time as products are written to work with those new standards.

PDF 2.0 headers

The first line of a PDF file identifies the version number of PDF in which the file was saved ... although it can then be overridden with the Version key in the Catalog. It's only a small thing, but a PDF reader must be prepared to encounter a value of 2.0 in those places.

```

1 %PDF-2.0 %âãîó
2 1007 0 obj <</Linearized 1/L 868179/0
. 1009/E 102691/N 40/T 866595/H [ 480 619]>>
. endobj
3 1017 0 obj <</DecodeParms<</Columns
. 4/Predictor
. 12>>/Filter/FlateDecode/ID[<7A5AAC7C95DD8BB
. 90533E94C3CE9A13A><90E7A59C5796664EBBAEB47E
. 5B89FDFB>]/Index[1007 18]/Info 1006 0
. R/Length 66/Prev 866596/Root 1008 0 R/Size
. 1025/Type/XRef/W[1 2 1]>>stream
4 hPbbd`b`±

```

PDF 1.7 readers will probably vary significantly in their handling of files marked as PDF 2.0. Some may error; others may warn that a future version of that product is required, while others may simply ignore the version completely and continue silently.

During the transition period to PDF 2.0 you will need to understand what your existing PDF reading tools do when presented with a PDF 2.0 file. If they don't warn at all, you won't get any indication that the files may not be processed exactly as their creators intended.

The current version of Harlequin reports "PDF Warning: Unexpected PDF version - 2.0" and then continues to process the job. Obviously that warning will disappear when we ship a new version that fully supports PDF 2.0.

UTF-8 text strings

Previous versions of PDF allowed certain strings in the file to be encoded in PDFDocEncoding or in 16-bit Unicode; these are usually used in metadata. PDF 2.0 adds support for UTF-8 as well, primarily to make it as easy as possible to copy data to and from XML representations.

Many PDF 1.7 readers may not recognize the UTF-8 string as UTF-8 and will therefore treat it as using PDFDocEncoding, resulting in those strings being treated as what looks like a random sequence of mainly accented characters.

Security

The encryption algorithms included in previous versions of PDF have fallen behind current best practices in security, so PDF adds AES-256-bit encryption⁵ and states that all passwords must be encoded in Unicode. While this isn't directly print-related, solutions for use in secure workflows may well need to take account of these changes.

A PDF 1.7 reader will almost certainly error and refuse to process any PDF files using the new AES-256 encryption.

Deprecation and what this means in PDF

PDF 2.0 has deprecated a number of implementation details and features that were defined in previous versions. In this context 'deprecation' means that tools writing PDF 2.0 are recommended not to include those features in a file; and that tools reading PDF 2.0 files are recommended to ignore those features if they find them.

⁵ Adobe's ExtensionLevel 3 to ISO 32000-1 defines a different AES-256 encryption algorithm, as used in Acrobat 9 (R=5). That implementation is now regarded as dangerously insecure it has been deprecated completely, to the extent that use of it is forbidden in PDF 2.0.

Global Graphics has taken the deliberate decision not to ignore relevant deprecated items in PDF files that are submitted and happen to be identified as PDF 2.0. This is because it is quite likely that some files will be created using an older version of PDF and using at least some of the features that are now deprecated. If those files are then pre-processed in some way before submitting to Harlequin (e.g. to impose or trap the files) the pre-processor may well tag them as now being PDF 2.0. It would not be appropriate in such cases to ignore anything in the PDF file simply because it is now tagged as PDF 2.0.

We expect most other PDF readers to take the same course, at least for the next few years, but this is an area where you may wish to seek clarity from your vendors.

About the author

Martin Bailey, CTO, Global Graphics Software



Martin Bailey works to analyze and understand current and future needs for workflows across many sectors of print. This enables him to guide Global Graphics' industry-leading printing technology. He represents Global Graphics on a number of industry bodies and standards committees including acting as the primary UK expert on the committees working on PDF, PDF/X and PDF/VT.

Martin has over 30 years of experience building, using, supporting and improving products for processing digital documents and the print industry in technical support, product management and programming as well as in consulting, and production environments.

About Global Graphics Software

Global Graphics Software <http://www.globalgraphics.com> is a leading developer of platforms for digital printing, including the Harlequin RIP®, ScreenPro, Fundamentals and Mako. Customers include HP, Canon, Durst, Roland, Kodak and Agfa. The roots of the company go back to 1986 and to the iconic university town of Cambridge, and, today the majority of the R&D team is still based near here. Global Graphics Software is a subsidiary of Hybrid Software Group (Euronext: GLOG).



Oct 2021 v3

Contact us.
info@globalgraphics.com

www.globalgraphics.com

Global Graphics Software Inc.

5996 Clark Center Avenue
Sarasota, FL 34238
United States of America
Tel: +1 (941) 925-1303

Global Graphics Software Ltd

Building 2030
Cambourne Business Park
Cambourne, Cambridge
CB23 6DW UK
Tel: +44 (0) 1954 283100

Global Graphics KK

610 AIOS Nagatacho Bldg.
2-17-17 Nagatacho, Chiyoda-ku,
Tokyo 100-0014
Japan
Tel: +81-3-6273-3198