



**EBriefPro Inc. White Paper**

# **PDF/A Allows External Hyperlinks**

**September 9, 2011**

**<http://www.ebriefpro.com>**



---

# PDF/A Allows External Hyperlinks

Joshua Neil Rubin, CEO, EBriefPro Inc.

## INTRODUCTION

The authoritative international standard for creating PDF/A files (archival PDF files) was promulgated by the International Organization for Standardization ("ISO") in 2005 and 2011.<sup>1</sup> It will be phased into the federal Case Management/Electronic Case Filing ("CM/ECF") system as a mandatory standard on a court-by-court basis.<sup>2</sup>

Some authorities erroneously state that PDF/A-compliant files cannot include external hyperlinks, i.e., hyperlinks to information outside of the PDF file itself. For example, Wikipedia's PDF/A entry<sup>3</sup> says that "[a] PDF/A document is not permitted to be reliant on information from external sources (e.g. font programs and hyperlinks)."<sup>4</sup>

---

<sup>1</sup> International Organization for Standardization, *ISO 19005-1:2005 - Document management -- Electronic document file format for long-term preservation -- Part 1: Use of PDF 1.4 (PDF/A-1)*, copyrighted, available for purchase at [http://www.iso.org/iso/catalogue\\_detail?csnumber=38920](http://www.iso.org/iso/catalogue_detail?csnumber=38920) (referred to herein as "ISO 19005-1"), and *ISO 19005-2:2011 - Document management -- Electronic document file format for long-term preservation -- Part 2: Use of ISO 32000-1 (PDF/A-2)*, copyrighted, available for purchase at [http://www.iso.org/iso/catalogue\\_detail?csnumber=50655](http://www.iso.org/iso/catalogue_detail?csnumber=50655) (referred to herein as "ISO 19005-2").

<sup>2</sup> Administrative Office of the U.S. Courts, PACER Service Center, *CM/ECF to Transition to PDF/A*, <http://www.pacer.gov/announcements/general/pdfa.html> (accessed August 31, 2011).

<sup>3</sup> Wikipedia, *PDF/A*, <http://en.wikipedia.org/wiki/PDF/A> (accessed August 31, 2011).

<sup>4</sup> See also, e.g., U.S. Court of Appeals for the Ninth Circuit, *Using PDF/A Format in CM/ECF* (Feb. 15, 2011), [http://www.ca9.uscourts.gov/datastore/uploads/cmecf/PDF-A\\_generic\\_transcript.pdf](http://www.ca9.uscourts.gov/datastore/uploads/cmecf/PDF-A_generic_transcript.pdf), at Slide 6 ("Hyperlinks contained within PDF/A documents will not be active. To access the referenced hyperlink, users must copy and paste the URL into a browser window."); New York State Educ. Dept., *Preserving Electronic Records in Colleges and Universities*, at p. 18 (hyperlinks to external information are "not allowed in a PDF/A file"),



---

Electronic legal briefs ("ebriefs") increasingly include hyperlinks to external information such as legal authority and evidence. The CM/ECF system, as implemented in many courts, already permits active hyperlinks in PDF ebriefs to certain external information. Under the view of Wikipedia and some other authorities, no PDF/A-compliant ebrief could include such hyperlinks. The purpose of this White Paper is to definitively refute that view.

## SUMMARY

The PDF/A standard expressly provides for compliant PDFs that contain usable hyperlinks to external information. The contrary view is probably based on an overgeneralization of various specific prohibitions in the standard that do not relate to hyperlinks. The only purpose of those prohibitions is to ensure that certain designated information about the appearance, metadata, and structure of the PDF/A file itself will be physically included within the file and not merely referenced.

## ANALYSIS

### A. PDF/A Permits Hyperlinks to External Information.

The first version of the PDF/A standard, PDF/A-1, was promulgated in 2005 in *ISO 19005-1*. It provides that the PDF/A-1 format permits any feature permitted by the "*PDF Reference*," subject only to the explicit prohibitions in *ISO 19005-1*.<sup>5</sup> "*PDF Reference*" refers to Adobe Systems Inc.'s *PDF Reference* in effect as of January 2004, i.e., the Third Edition dated December 2001, setting forth Version 1.4 of the Adobe Portable Document Format, as modified by the Errata dated June 18, 2003.<sup>6</sup> Those documents are copyrighted by Adobe, and Adobe makes them

---

[http://www.archives.nysed.gov/a/workshops/workshops\\_handouts\\_erecords\\_preservation\\_college.pdf](http://www.archives.nysed.gov/a/workshops/workshops_handouts_erecords_preservation_college.pdf) (accessed August 31, 2011).

<sup>5</sup> *ISO 19005-1* at ¶ 5.1 at pp. 4-5. **Important Note:** ISO standards are copyrighted and the ISO does not make them available to the general public. References to the official versions herein are therefore truncated, paraphrased, and presented without hyperlinks.

<sup>6</sup> *ISO 19005-1* at p. vi.



---

publicly available for download without registration at the Adobe website.<sup>7</sup> This White Paper refers to the specification set forth in the *PDF Reference* as the "*PDF 1.4 Specification*."

The *PDF 1.4 Specification* provides for hyperlinks that link to external information, both on the local file system and on the internet. It defines an "annotation" as an "interact[ive]" location on a PDF page that a user can "activate" by clicking it with a mouse.<sup>8</sup> "Link annotations" represent both links to destinations within the same file and "Actions."<sup>9</sup> "Actions" include "GoToR" actions and "URI" actions.<sup>10</sup> GoToR actions refer to destinations in external (i.e., "remote") PDF files.<sup>11</sup> URI ("Uniform Resource Identifier") actions refer to internet resources, "typically a file that is the destination of a hypertext link . . ."<sup>12</sup>

The corresponding part of *ISO 19005-1* is ¶ 6.6.3.<sup>13</sup> That Paragraph is entitled "Hypertext Links." It refers to "GoToR actions" and "URI actions" as "hypertext links" and "hyperlinks." It does not limit the nature of hyperlinks included in PDF/A files in any way. It provides that compliant PDF-reading software may make any hyperlinks usable or not. It also requires that, in either event, the software must allow the user to see "the F and D keys of a GoToR action dictionary

---

<sup>7</sup> Adobe Systems Inc., *PDF Reference, Third Ed.* (December 2001), [http://wwwimages.adobe.com/www.adobe.com/content/dam/Adobe/en/devnet/pdf/pdfs/pdf\\_reference\\_archives/PDFReference.pdf](http://wwwimages.adobe.com/www.adobe.com/content/dam/Adobe/en/devnet/pdf/pdfs/pdf_reference_archives/PDFReference.pdf), Adobe Systems Inc., *Errata for PDF Reference, Third Ed.* (June 18, 2003), [http://wwwimages.adobe.com/www.adobe.com/content/dam/Adobe/en/devnet/pdf/pdfs/pdf\\_reference\\_archives/PDF14errata.txt](http://wwwimages.adobe.com/www.adobe.com/content/dam/Adobe/en/devnet/pdf/pdfs/pdf_reference_archives/PDF14errata.txt).

<sup>8</sup> *PDF 1.4 Specification* at ¶ 8.4 at p. 488.

<sup>9</sup> *Id.* at ¶ 8.4.5 at p. 501.

<sup>10</sup> *Id.* at ¶ 8.5.3 at p. 518 (Table 8.34).

<sup>11</sup> *Id.* at p. 520.

<sup>12</sup> *Id.* at pp. 523-24.

<sup>13</sup> *ISO 19005-1* at ¶ 6.6.3 at pp. 13-14.



---

[and] the URI key of a URI action dictionary . . ." It also identifies these keys as the locations that the hyperlinks point to.<sup>14</sup> The *PDF 1.4 Specification* itself defines the "F and D keys of a GoToR action dictionary" as the specification of a file and a location within it. It defines the "URI key of a URI action dictionary" as "[t]he uniform resource identifier to resolve . . ."<sup>15</sup>

Other than in ¶ 6.6.3., the words "hyperlink" and "hypertext" do not appear in *ISO 19005-1*.

The second version of the PDF/A standard, PDF/A-2, was promulgated in 2011 in *ISO 19005-2*. It provides that the PDF/A-2 format permits any feature permitted by the specifications for PDF Version 1.7 (referred to herein as the "*PDF 1.7 Specification*"),<sup>16</sup> subject only to the explicit prohibitions in *ISO 19005-2*.<sup>17</sup>

The *PDF 1.7 Specification* provides for hyperlinks that link to external information, both on the local file system and on the internet, in the same terms as the *PDF Reference*. It defines an "annotation" as an "interact[ive]" location on a PDF page that a user can "activate" by clicking it with a mouse.<sup>18</sup> "Link annotations" represent both links to destinations within the same file and "Actions."<sup>19</sup> "Actions"

---

<sup>14</sup> *ISO 19005-1* at ¶ 6.6.3 at pp. 13-14.

<sup>15</sup> *PDF 1.4 Specification* at ¶ 8.5.3 at p. 520 (Table 8.36) (GoToR) and p. 523 (Table 8.40) (URI).

<sup>16</sup> The *PDF 1.7 Specification* is officially set forth in ISO 32000-1 (International Organization for Standardization, *ISO 32000-1:2008 - Document management -- Portable Document Format -- Part 1: PDF 1.7*), copyrighted, available for purchase at [http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=51502](http://www.iso.org/iso/catalogue_detail.htm?csnumber=51502). By agreement with the ISO, Adobe makes freely available an unofficial but exact copy of the *PDF 1.7 Specification* as set forth in ISO 32000-1, albeit without the ISO mark. That copy is at [http://www.adobe.com/content/dam/Adobe/en/devnet/acrobat/pdfs/PDF32000\\_2008.pdf](http://www.adobe.com/content/dam/Adobe/en/devnet/acrobat/pdfs/PDF32000_2008.pdf).

<sup>17</sup> *ISO 19005-2* at ¶ 5.1 at pp. 4-5.

<sup>18</sup> *PDF 1.7 Specification* at ¶ 12.5.1 at p. 381.

<sup>19</sup> *Id.* at ¶ 12.5.6.5 at p. 394.



---

include "GoToR" actions and "URI" actions.<sup>20</sup> GoToR actions refer to destinations in external (i.e., "remote") PDF files.<sup>21</sup> URI ("Uniform Resource Identifier") actions refer to internet resources, "typically a file that is the destination of a hypertext link . . ." <sup>22</sup>

The corresponding part of *ISO 19005-2* is ¶ 6.5.3.<sup>23</sup> That Paragraph restates ¶ 6.6.3 of *ISO 19005-1* so that it does not implicitly limit its scope to "Hypertext Links" but instead expressly encompasses all "GoToR actions" and "URI actions." Like ¶ 6.6.3 of *ISO 19005-1*, ¶ 6.5.3 of *ISO 19005-2* does not limit the nature of hyperlinks included in PDF/A files in any way. It provides that compliant PDF-reading software may make any "GoToR actions" and "URI actions" usable or not. It also requires that, in either event, the software must allow the user to see "the F and D keys of a GoToR action dictionary [and] the URI key of a URI action dictionary . . ." <sup>24</sup>

Like the *PDF 1.4 Specification*, the *PDF 1.7 Specification* defines the "F and D keys of a GoToR action dictionary" as the specification of a file and a location within it, and it defines the "URI key of a URI action dictionary" as "[t]he uniform resource identifier to resolve . . ." <sup>25</sup>

The words "hyperlink" and "hypertext" do not appear at all in *ISO 19005-2*.

Two things are therefore apparent from *ISO 19005-1* and *ISO 19005-2*:

---

<sup>20</sup> *Id.* at ¶ 12.6.4.1 at pp. 417-418 (Table 198).

<sup>21</sup> *Id.* at ¶ 12.6.4.3 at p. 419.

<sup>22</sup> *Id.* at ¶ 12.6.4.7 at pp. 423-424.

<sup>23</sup> *ISO 19005-2* at ¶ 6.5.3 at p. 20.

<sup>24</sup> *Id.*

<sup>25</sup> *PDF 1.7 Specification* at ¶ 12.6.4.3 at p. 419 (Table 200) (GoToR), ¶ 12.6.4.7 at pp. 423-424 (Table 206) (URI).



- 
1. The PDF/A standard permits compliant PDF-reading software to allow users to use any hyperlinks contained within PDF/A files, not only intra-file hyperlinks.
  2. The PDF/A standard permits a PDF/A file to include hyperlinks that link to external information. It merely requires software used for reading a PDF/A file to allow the user to discern the location that a hyperlink points to.

In short, both the *PDF 1.4 Specification* and the *PDF 1.7 Specification* provide for hyperlinks to external information, and neither *ISO 19005-1* nor *ISO 19005-2* contains any restriction on the hyperlinks permitted by the relevant PDF specifications. Therefore, any PDF/A file may include hyperlinks to external information.

## **B. The Probable Bases for the Contrary View Do Not Support It.**

What then accounts for the mistaken conclusions of several diverse authorities that PDF/A prohibits hyperlinks to external information?

Those conclusions are most likely overgeneralizations of many specific prohibitions in the standard, all of which support the archival purpose of the standard and none of which relate to hyperlinks.

The purpose of the PDF/A standard is to try to ensure that future computer programs running on future computers will be able to accurately reproduce the visual appearance, metadata, and technical structure of the PDF itself.<sup>26</sup>

Accordingly, the standard requires that archival PDFs must physically include certain specific information that computer programs may need to accurately reproduce those aspects of the PDF itself in the future. For example, ¶ 6.3.4 of *ISO 19005-1* requires that any PDF/A file must physically include each font used in the PDF (unless the font is converted to an image). Note 4 to that Paragraph explains

---

<sup>26</sup> *ISO 19005-1* at p. v.; *ISO 19005-2* at p. v.



---

that embedding fonts allows for accurate reproduction without reference to external information that may be unavailable in the future.<sup>27</sup>

This prohibition sounds a bit like Wikipedia's view that the PDF/A standard prohibits hyperlinks to external information. However, unlike information that a user would get by clicking a hyperlink in a PDF, the actual font used to create the PDF, and not a similar font or a later version, is necessary to ensure the precise reproduction of the appearance of the PDF itself in the future. Therefore, each font used needs to be embedded even if that font is currently ubiquitous (e.g., Times New Roman) and even if it currently seems highly likely that that font will be available in the future on any computer system that will be used to read the PDF/A file.

*See also ISO 19005-1*, ¶ 6.1.7 (prohibiting the use of references in the PDF file to "filters" that would otherwise be used to process data streams from external files and to present that data as part of the contents of the PDF, because the external data may be different or absent in the future, undermining the archival purposes of the standard)<sup>28</sup>; ¶ 6.1.11 (prohibiting embedding files containing arbitrary content for the same reasons)<sup>29</sup>; ¶ 6.2.3.1 (requiring the PDF file to specify information used to correctly reproduce colors for the same reasons)<sup>30</sup>; ¶ 6.2.6 (prohibiting inclusion of "reference XObjects" for the same reasons)<sup>31</sup>; ¶ 6.6.1

---

<sup>27</sup> *ISO 19005-1* at ¶ 6.3.4 at pp. 10-11; *see also ISO 19005-2* at ¶ 6.2.11.4.1 at pp. 15-16.

<sup>28</sup> *ISO 19005-1* at ¶ 6.1.7 at p. 6; *see also ISO 19005-2* at ¶ 6.1.7.1 at p. 7.

<sup>29</sup> *ISO 19005-1* at ¶ 6.1.11 at p. 7; but *see ISO 19005-2* at ¶ 6.8 at p. 27 (prohibiting embedded files unless they comply with *ISO 19005-1* or *ISO 19005-2*).

<sup>30</sup> *ISO 19005-1* at ¶ 6.2.3.1 at p. 8; *see also ISO 19005-2* at ¶ 6.2.4.1 at p. 10.

<sup>31</sup> *ISO 19005-1* at ¶ 6.2.6 at p. 9; *see also ISO 19005-2* at ¶ 6.2.9.2 at p. 13.



---

(prohibiting the "ImportData action" for the same reasons)<sup>32</sup>; ¶ 6.6.2 (prohibiting JavaScript actions for the same reasons)<sup>33</sup>.

Neither *ISO 19005-1* nor *ISO 19005-2* includes any similar prohibitions relating to GoToR actions or URI actions. The only references to external information that are prohibited by either standard are references to specific types of external information that would be used to reproduce the appearance, metadata, or structure of the PDF itself. No prohibition of any particular type of hyperlink furthers the archival purposes of the standard.

## CONCLUSION

The PDF/A standard permits hyperlinks to external information. Hyperlinks do not create the type of dependency on external information that undermines the ability to reliably reproduce the visual aspect, metadata, or technical or semantic structure of a PDF.

### About EBriefPro Inc.

[EBriefPro Inc.](http://www.ebriefpro.com) (<http://www.ebriefpro.com>) is an information technology company headquartered in Brooklyn, NY. It specializes in electronic document analysis, visualization, and presentation. Its mission is to create and facilitate fast, simple, inexpensive, appropriate, preferably paperless solutions.

---

<sup>32</sup> *ISO 19005-1* at ¶ 6.6.1 at p. 13; see also *ISO 19005-2* at ¶ 6.5.1 and Note 1 at p. 19.

<sup>33</sup> *ISO 19005-1* at ¶ 6.6.2 at p. 13; see also *ISO 19005-2* at ¶ 6.5.2 at p. 19.

