

MANUAL PARA CRIAÇÃO DE ARQUIVOS D GTA S **NO PADRÃO** PDF/A ISO 19005-1

1° versão 4-2007 - Bruno Mortara





PREFÁCIO

O grande desafio de nossa era – a era da informação – é conviver de maneira significativa com uma enormidade de informações. Para dar significado ao grande número de informações que atingem o cidadão comum, as empresas e os governos se criaram mecanismos de busca, hierarquização e catalogação. Também são cada vez mais comuns os grupos de pessoas na Internet que partilham suas preferências em termos de sites, músicas, gastronomia, vinhos, carros e documentos.

O desenvolvimento de tecnologias e meios digitais permitiu ao homem criar e disseminar informação de maneiras inéditas e em volumes nunca antes vistos. Essas inovações trouxeram numerosas facilidades, como por exemplo a capacidade de qualquer cidadão comum ser o autor de livros, revistas, Blogs, Sites na Web, etc. Estas tarefas, antes custosas e restritas aos profissionais da informação (jornalistas, desenhistas, diretores de cinema, escritores), agora estão nas mãos dos cidadãos comuns. Isso provocou a ccriação de um volume enorme de conteúdos seja em meios tradicionais – livros, vídeos e revistas – como nos meios novos – sites na Web, blogs, música na rede, vídeo na rede etc. Cabe a cada cidadão a tarefa de selecionar e dar significação às informações que considerar úteis, separando-as do "lixo" informático. Há entretanto ainda mais uma tarefa: escolher a maneira de como armazenará as informações de forma que possam ser recuperadas num futuro indefinido. Em poucos anos de informática já vivemos diversas "gerações" de programas que criaram arquivos proprietários e aplicativos de algumas versões posteriores já não conseguiam ler. Paralelamente, temos o desafio a vencer: a obsolescência dos suportes – mídia sobre as quais foram gravados os dados digitais. Isso requer políticas de renovação e qualidade.

Este manual se destina a divulgar uma solução para uma boa parte desses desafios através da adoção de um padrão internacional para arquivos digitais: o PDF/A.

O PDF/A é um PDF – formato de arquivo criado pela Adobe Systems – que neste momento se encontra na versão 1.7. O PDF pode ser entendido como um contêiner de informações diversas, todas digitais: sons, imagens, vídeos, textos, anotações e símbolos. A adoção desse formato, que é uma especificação aberta e divulgada publicamente, fez com que a ISO – International Standards Organization–o escolhesse para ser a base do formato de arquivamento de longo prazo, o PDF/A, ou ISO 19005-1.

O PDF/A está sendo foco de enorme interesse em todo o mundo. No Brasil a ABNT está iniciando a tradução da norma ISO 19005-1 com o auxílio dos agentes do mercado, governos e meioa acadêmicos. Todos estão se estão mobilizando em direção à definição de um formato sólido, estável, não proprietário e aceito internacionalmente: o PDF/A.

Este manual se propõe a ser uma ferramenta de auxílio inicial para que pessoas, governo e instituições possam ter subsídios para iniciar a adoção do formato PDF/A na criação de acervos digitais de documentos com vistas à conservação de longo prazo.

É fundamental ressaltar o apoio recebido da Adobe do Brasil suportando as ações que levaram à confecção deste manual.

São Paulo, abril de 2007.

INTRODUÇÃO

Este manual se destina a pessoas e instituições que desejam conhecer ou implementar fluxos de trabalho baseados na norma internacional PDF/A, com a intenção de conservação de arquivos digitais a longo prazo. Portanto ele se baseia na norma PDF/A-1 ou ISO 19005-1. A norma tem dois níveis de conformidade: o PDF/A-1a e o PDF/A-1b. É recomendável que os documentos PDF assim gerados sejam arquivos digitais conformes ao menos ao padrão PDF/A-1b.

O Adobe Acrobat Professional, na sua versão mais disseminada –7.0.8, suporta o padrão oficial PDF/ A-1b fornecendo dois ajustes de configuração: PDF/A-1b:2005(RGB) e PDF/A-1b:2005(CMYK). Ao se utilizar a ferramenta Adobe Acrobat Professional para converter documentos para o formato PDF, é recomendavel que se use um destes dois ajustes padrão de PDF/A-1b sem modificação.



Os usuários do MS Office – Word e Excel, devem usar o plug-in PDF Maker do Word e Excel para criar arquivos PDF a partir de arquivos originais, e não através do Acrobat (fluxo PostScript, Distiller e Acrobat). Nos arquivos PDF/A-1b convertidos de Word e Excel pelo Acrobat Profissional falta o campo requerido "ID" no dicionário Trailer e por isso não passarão pelo teste de validação do PDF/A. Independentemente da maneira que seja criado o arquivo PDF/A, os usuários devem usar o preflight disponível no Adobe Acrobat Professional para validar o PDF contra o padrão PDF/A-1b e remover todas as não conformidades antes da submissão.



Ferramenta	Suporte a Conversão	Ajustes de PDF recomendados
Acrobat Professional	MS Office, MS Project, MS Visio, MS Publisher e Postscript para PDF/A-1b	PDF/A-1b:2005(RGB) ou PDF/A-1b:2005(CMYK)
Acrobat Distiller	Postscript para PDF/A-1b	PDF/A-1b:2005(RGB) ou PDF/A-1b:2005(CMYK)
Adobe PDF Maker plug-in Microsoft Word	MS Word para PDF/A-1b	PDF/A-1b:2005(RGB) ou PDF/A-1b:2005(CMYK)

Na plataforma Macintosh, a regulagem do plug-in se dá através da abertura do Acrobat Distiller e seleção do perfil desejado – PDF/A-1b:2005 (RGB) ou PDF/A-1b:2005(CMYK). Na plataforma Windows, a seleção se dá no painel de controle de impressoras, Adobe PDF Writer, propriedades.

Default Settings	✓ PDF/A-1b:2005 (RGB)
Compatibility Use these set conform to P electronic do documents, p Progress Status: Ready Pause OF File	PDF/X-1a:2001 PDF/X-3:2002 PR1520 Press Quality Prinergy Pages Globo Prinergy ViverBem Print_Luescher Smallest File Size sonopress Standard Texto Com-baixissima VISUALE_ALTA YazioiTela

Ajuste no Mac: Distiller

🗞 Impressoras e aparelhos de fax			<u>_ </u>
Arquivo Editar Exibir Eavoritos Ferr	amentas Ajuda		A
🕝 · 🕑 · 🎓 🔎 🕞	迩 🗙 🍤 📖 ·		
Endereço 🂫 Impressoras e aparelhos de fax			💌 🄁 Ir
Adicionar. Impressora	Adobe PDF 0 Pronto	EP50N StylusCOLOR1520 v3010.106 0	Auto Epson Stylus COLOR 777 ESC/P 2 em P2000
Generic PostScript Printer	Epson Stylus Pro 5000 ESC/P 2 0 Pronto		
Aiveta no DC 1.	Calaajamar aa mr	anniadadaa da F	

Ajuste no PC 1: Selecionar as propriedades do PDF Writer

Adobe PDF Pr	opriedades			?)
Gerenciar	nento de cores	1	Configurações d	lo dispositivo
Geral	Compartilhan	nento	Portas	Avançado
	lobe PDF			
L <u>o</u> cal:	feus documentos			
<u>C</u> omentário:				
<u>M</u> odelo: Ad	obe PDF Converte	er		
Recursos				
Cor: Sim		Papel dis	ponível:	
Frente e vers	o: Não	A4		<u> </u>
Grampeamer	to: Não			
Veloc.: 400 p	pm			
Resolução m	áxima: 4000 dpi			T
Pr	eferências de imp	essão	Imprimir página	a de teste
		OK	Cancela	r Apjicar

Ajuste no PC 2: Selecionar o ajuste adequado do PDF Writer

pector	🝓 Adobe PDF Padrões de impressão	? ×
as e aparelhos de fax ar Exibir Favoritos Ferramo bill Pavoritos Ferramo bill Pavoritos Ferramo bill Pavoritos Ferramo bill Pavoritos Ferramo bill Pavoritos Por Porpriedades derenciamento de core Geral Compar c Sempre disponível c Disponível de c Dis	Layout Papel/qualidade	d printing Adobe Edit Idit Add
Percent Reter os documentos per Imprimir primeiro os docu Manter documentos impr Ativar recursos de impres Padrões de impressão	rocessador de imp Página separadora OK Cancelar	Aplicar

OS AJUSTES Específicos

Os ajustes do Acrobat Professional podem ser customizados para cada caso específico. Porém é preciso que se tome cuidado para não se criar um arquivo não conforme ou com perda de conteúdo, o que comprometeria sua leitura no futuro. Portanto, é preciso estar dentro dos requerimentos da norma durante o processo da criação de PDF/A.

Espaço de Cor

Os originais submetidos devem usar somente espaços de cor device-independent, tais como CalGray, CalRGB, Lab ou ICCBased. Todo o espaço de cor ICCBased deve ser embutido conforme a especificação do ICC. Alternativamente, a independência de dispositivo de saída pode ser conseguida usando um espaço de cor uncalibrated (como, DeviceRGB ou DeviceCMYK) que contenha também um OuputIntent para especificar as características da cor do dispositivo para o qual se pretende dar saída.

Dica: Para assegurar-se de que toda a cor device-dependent esteja convertida usando o espaco de cor device-independent ou em uma maneira device-independent, recomendase que os ajustes do Distiller para o padrão de PDF/A-1b, em relação a colorspace, não sejam alterados. Entretanto, se for necessário alterar os ajustes do Adobe PDF, selecione um dos seguintes ajustes presentes no Acrobat: "Tag Everything for Color Management" ou "Convert All Colors to RGB," ou "Convert Everything to CMYK". Não ajuste o espaço de trabalho para a política de gerência selecionada (color management policy) de cor a "none".

Compressão

Os documentos convertidos devem usar so-

mente os algoritmos de compressão lossless (sem perda) que não estejam sujeitos a limitações de propriedade intelectual. O uso da compressão LZW é proibido.

Dica: Para a compressão de imagens monocromáticas, escolha no Adobe Acrobat uma das opções a seguir: "Off," "Run Length," "CCITT Group 3" ou "CCITT Group4". Para imagens de Cor ou Grayscale, escolha nos ajustes da compressão qualquer uma das seguintes opções: "Off," "Zip," "JPEG" ou "automatic (JPEG)". Para o uso de imagens GrayScale ou coloridas (RGB, Lab ou CMYK) é recomendável a compressão "automatic (JPEG)" com a qualidade "maximum".

Referências Externas

Para a finalidade de preservação a longo prazo é imperativo que o original submetido seja autocontido (completo) com nenhuma dependência externa (links ou vínculos). Conseqüentemente, aquelas características do PDF que permitem referências externas não devem ser usadas. Estas características incluem:

- Actions para lançar ações externas
- Actions para executar Java scripts. O uso de Java scripts pode criar dependências externas e interferir na aparência do documento.
- Actions para importar dados de formulário de arquivos externos (ImportData action)
- Actions com referências a arquivos externos (FileAttachment annotation)

Dica: Quando se configurar os ajustes do Adobe PDF para criar documentos tipo PDF/A no Acrobat, Distiller ou PDF Maker, não acionar a opção "Add bookmarks and links to Adobe PDF file". Isso pode incluir os Links a arquivos externos.

Fontes

Todas as fontes usadas em um documento, inclusive as fontes padrão, devem ser legalmente embutidas para permitir uma reprodução ilimitada e universal do arquivo. Ao embutir as fontes assegura-se que o original submetido seja visualizado no futuro com a mesma aparência que foi originalmente criado. Para reduzir o tamanho do arquivo PDF, as fontes embutidas podem também ser subsetted contendo assim somente os caracteres utilizados no arquivo e os glifos necessários para visualizá-lo adequadamente.

Dica: Nos ajustes do Adobe PDF, selecione as opções the "Embed all fonts" e "Subset embedded font when percent of characters used is less than 100%". Não selecione "Optimize for fast web view" porque fará com que o Acrobat não embuta as fontes padrão no arquivo PDF resultante.

Formulários (Forms)

De forma a assegurar a aparência consistente do arquivo PDF nos campos de formulário é proibida qualquer ação que possa potencialmente mudar a aparência visual do arquivo.

Dica: Remova todas as form actions do documento antes de convertê-lo para o formato PDF.

Imagens

Imagens alternativas (Alternate Images), que permitem respresentações variáveis de uma imagem, não devem ser usadas. O uso de transparência em uma imagem também é proibido. Deve-se "achatar" arquivos com múltiplas camadas antes de convertêlo para PDF. Deve-se evitar o downsampling das imagens durante o processo de criação do PDF porque pode resultar na perda da qualidade (isto não é uma exigência formal do PDF/A).

Metadata (Metadados)

Para elencar de maneira uniforme Metadata descritiva, administrativa ou técnica, o arquivo PDF deve conter um stream Metadata que seja conforme a especificação XMP. As seguintes propriedades Metadata XMP devem ser fornecidas no arquivo PDF:

Metadata	XMP Property
Document Title	dc:title
Document Author	dc:creator
Subject	dc:subject
Keywords	pdf:keywords
Creator Program	xmp:CreatorTool
Producer	pdf:Producer
Creation Date	xmp:CreateDate
Modification Date	xmp:ModifyDate
File Identifier	ou através do uso de xmp:Identifier, xmpMM:DocumentID, mpMM:VersionID, ou xmpMM:RenditionClass
File Provenance Information	xmpMM:History

Cada property de Metadata XMP que tiver uma entrada correspondente no dicionário document information dictionary deve ser definida com o mesmo valor que tem neste dicionário. O stream Metadata XMP também deve conter property para indicar o nível da versão e do conformidade PDF/A do documento. Estas propriedades devem ser especificadas usando o esquema de extensão da Identificação do PDF/A.

A descrição de todos os esquemas de extensão usados no original devem ser embutidas. As descrições devem ser especificadas utilizando-se o extension schema description do PDF/A. Além disso, o stream Metadata XMP não deve ser filtrado. **Dica:** Usando-se um gravador conforme PDF/A se terá certeza que este embutirá um stream Metadata XMP de acordo com o padrão de PDF/A. Não use gravador de PDF não conforme.

Multimedia

É proibida qualque Action para embutir conteúdos multimedia, ou seja, Sound Annotation, Sound Action, Movie Annotation e Movie Action.

Dica: Remova todos os conteúdos multimedia de um documento antes de convertê-lo para o formato PDF.

Segurança

O documento não deve ser protegido com senhas ou ser cifrado. A proteção com criptografia ou senha impede a acessibilidade aos arquivos, impedindo sua futura leitura ou processamento.

Sugestão: Ao configurarar ajustes de PDF para criar originais do PDF no Acrobat, Distiller ou PDF Maker, selecionam "None" no item "Adobe PDF Security" a fim de evitar a proteção com senha no arquivo a ser criado ou convertido.

Conformidade com o PDF/A-1a

A conformidade com o padrão PDF/A-1a inclui todas as características possíveis da norma PDF/A. Um documento PDF/A-1a, além das características do PDF/A-1b, tem que ter as seguintes características ou restrições:

• o uso de Tags para especificar a estrutura lógica do documento (o uso de Tags pode aumentar significativamente o tamanho do documento PDF).

 o uso do mapa de caráter Unicode (Unicode Character Map) a fim de mapear a relação entre os caracteres e seus valores Unicode. No final de 2006, a Adobe anunciou o lançamento do Adobe Acrobat Professional versão 8. Para a finalidadedecriação de arquivos PDF/Acompatíveis, o programa traz uma série de recursos que tornam a vida do usuário muito mais fácil. Vamos analisar os recursos existentes na nova versão a partir de um exemplo de geração de arquivo PDF/A, iniciando o processo com um documento em papel.

USANDO O ADOBE Acrobat Professional Versão 8

1º- Escaneamento ou digitalização do original:

De dentro do Acrobat iniciamos o processo com um escaner conectado ao computador. Selecionamos o menú Document > Scan to PDF...



Antes do Acrobat passar o controle para o software do escaner são solicitados alguns ajustes. Os mais importantes são o "Make PDF/ A Compliant" o "Make Searchable..." e o "Make Accessible". A combinação destes permitirá que o arquivo resultante seja um PDF/A-1a.

		Acrobat Scan
Input		
Scanner: (EPSON Perfection 167	0
Sides:	Front Sides	•
Output		
New PDF	Document	
O Append		÷ Browse
🗹 Make PD	F/A Compliant	
Document		
Optimizat	ion	
Small Size		High Quality Options
Text Reco	gnition and Metadata	
The second second	Searchable (Run OCR)	Options
Make :		

2°- OCR ou reconhecimento de caracteres:

Devemos selecionar a língua do texto a ser digitalizado, de acordo com a língua do original.

Primary OCR Language	Brazilian Portuguese	\$
PDF Output Style	Searchable Image (Exact)	÷
Downsample Images	Lowest (600 dpi)	4

Somos apresentados então ao diálogo de controle do escaner onde temos que fazer alguns ajustes importantes. O primeiro é em relação ao tipo de original, que no nosso caso é perto e branco e para isso utilizaremos o ajuste "8 bit GrayScale". O segundo é em relação à resolução de digitalização e para isso escolheremos, como regra geral, 300 dpi. Observe que o livro está rotacionado em relação à mesa do escaner, porém o Acrobat ajustará automaticamente isso. Selecione a área desejada e solicite o escaneamento.

EPSON Scan Professional Mode	C Zoom	Preview Zoom
Settings	Marquee	
Name: Current Setting Save Delete		I IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
Driginal Document Type: Reflective Document Source: Document Table Auto Exposure Type: Document		Adde jan ver enterhalde spenner strateging das de angement strateging das de angement strateging das de angement spenner de formeren et angement spenner de formeren et angement spenner de ange
Destination Image Type: B-bit Grayscale Resolution: Document Size: V 5.62 H 8.91 in. Target Size: Original V 5.62 H 8.91 in. Scale: 100 Go f Go On Adjustments Reset Reset		<text><text><text><text></text></text></text></text>

3º- Metadados: informação para acesso futuro:

Terminada a digitalização o Acrobat nos apresenta a caixa de diálogo para que possamos inserir os metadados. É muito importante a padronização das palavras-chave para que se encontrem as informações nos arquivos digitalizados de maneira mais previsível. A Adobe aconselha a utilização dos campos default, não obstante seja possível em "Additional Metadata..." inserir pares de chaves/ valores particulares.

		C	Oocumer	nt Properties			
	Description	Security	Fonts	Initial View	Custom	Advance	d
Description							
File:	Acr274663.t	mp					
Title:	Dicionário d	e termos gr	áficos				
Author:	ABTG						
Subject:	termos gráfi	cos					
Keywords:	Dicionário t	ermos gráfie	cos nome	clatura ABTG			
Created:	05/01/2006	9:07:58 AN	1		A	dditional N	Metadata
Modified:							
Application:	EPSON Perfe	ction 1670					
Advanced							
PDF Produce	r: Adobe Ad	crobat 8.0 P	aper Capt	ture Plug-in			
PDF Version:	1.4 (Acrob	at 5.x)					
Location:	iMacG5:pri	ivate:var:tm	p:folders.	502:Temporar	yltems:		
File Size:	10.50 KB (10,749 Byte	s)				
Page Size:	226.1 x 14	12.5 mm		N	Number of Pa	ages: 1	

3º- Conversão e verificação do PDF/A1-a

Em seguida, executamos o Preflight do Acrobat através do menú Advanced e o perfil selecionado será "Convert to PDF/A1a (RGB)".

0		Preflight	
🔎 Profiles 🔌 Results	🕄 Report 🤤	🖞 Comments 🥖 Ed	it Options -
Preflight profiles:			<u> </u>
Convert to PD So Convert to PD So Convert to PD D So Convert to PD D So Convert to PD	F/A-1a (sRGB) F/A-1a (SWOP F/A-1b (Coate) ?) ed EOGRA27)	
D Convert to PD	F/A-1b (Japan	Color Coated)	
D Convert to PD	F/A-1b (SWOF	P)	U
😰 🖉 Verify complia	ance with PDF/	/A-la	¥
Purpose of the selected	Preflight pr	rofile:	Lat-
Converts the current de	ocument to	PDF/A-1a with an	sRGB OutputIntent.
Run Preflight checks	for visible	layers only	
🗌 Run Preflight profile	without app	plying fixups	
Preflight only pages	1 t	0 1	Execute
Not a PDF/X file.			
PDF/A-1b:2005 IEC 61966-2-1:	(not yet veri 1999 (sRGB	ified) IEC61966-2.1)	-

E o resultado vem como positivo!!

0	Preflight
Profile	es % Results 🔀 Report 🖑 Comments 🥒 Edit Options 🔹
Pre wa	flight profile "Convert to PDF/A-1a (sRGB)" did not find any errors or rnings:
📜 Page	1 from "teste _A1a.pdf"
S Co	nvert PDF document to PDF/A-1a:2005 (1 object)
S Ad	d "_A1a" to PDF file name after processing (1 object)
S Re	move all alternate images (1 object)
S Re	compress LZW as ZIP (1 object)
V No	problems found
NIL OV	erview
🕨 🔊 Pre	flight information



E o Acrobat gera um novo arquivo, um PDF/A1a.



Referências

[Adobe] "Create Adobe PDF Online," http://createpdf.adobe.com/cgi-feeder.pl/help_color?BP=&LOC=en_US

[Adobe 2004] "XMP Specification,

"Adobe Systems Incorporated, January 2004, http://www.aiim.org/ documents/standards/xmpspecification.PDF.

[ISO 19005-1] "Document management – Electronic document file format for long-term preservation – Use of PDF 1.4 (PDF/A-1)," International Organization for Standardization, 2004.

[ICC] "ICC.1: 1998-09, File Format for Color Profiles," International Color Consortium, http://www.color.org/ICC-1_1998-09.PDF.

[PDF Tools] "White Paper PDF/A – The Basics," PDF Tools AG, February 7, 2006, http://www.PDF-tools.com/public/downloads/whitepapers/ whitepaper-pdfa.PDF.

Design SK2 Design - Suli Edel

Prata da Casa Consultoria

11 3825 7690 www.pdfa.com.br bmortara@pratadacasa.com.br

