

Generazione di file PDF conformi alle norme PDF/X e PDF/A con pdfTeX—pdfx*

C.V. Radhakrishnan, Hàn Thế Thành, Ross Moore, Peter Selinger

1 Introduzione

Il pacchetto pdfx (`pdfx.sty`) al momento permette di compilare documenti conformi alle norme PDF/X e PDF/A usando pdfTeX, rispettando alcune varianti di questi standard. Sono disponibili anche le funzionalità per gli standard PDF/E e PDF/VT; la lista completa è nel paragrafo 2.1. Con l'espressione "permette di compilare" intendiamo che il pacchetto mette a disposizione i mezzi corretti e sufficienti per far sì che un documento possa essere dichiarato conforme a una particolare variante di PDF (PDF/X, PDF/A, PDF/E, PDF/VT, eccetera) insieme ad un particolare livello di conformità. Il pacchetto consente anche di inserire i metadati e il profilo di colore secondo le specifiche della particolare variante PDF scelta.

Gli elementi che formano i metadati, la maggior parte dei quali alla fine vanno inseriti in XML usando la codifica UTF-8, sono forniti mediante un file col nome `\jobname.xmpdata` per ciascun documento L^AT_EX. Senza questo file, che fornisce alcune informazioni necessarie e diverse informazioni accessorie, non si può ottenere un file PDF completamente conforme ai requisiti. Il PDF può essere creato con l'aspetto corretto di ogni pagina, ma non passerà i controlli di validazione. Il paragrafo 2.2 spiega come costruire questo file.

Quello che questo pacchetto *non fa* è quello di controllare tutti i dettagli della struttura del documento e del suo contenuto che può essere richiesto (o proibito) secondo la specifica variante PDF. Per esempio, PDF/VT (ISO 16612) richiede che il documento sia ben strutturato mediante sezioni del tipo "Form XObject" etichettate come `'/DPart'`. Nella stessa maniera PDF/A-1a (e anche 2a e 3a) (ISO 19005-1; ISO 19005-2; ISO 19005-3) richiedono che il file PDF sia completamente "Tagged PDF", inclusa una struttura di marcature che qualifica completamente i contenuti del documento. Questo non si può ancora ottenere con la versione corrente di pdfTeX, come lo si trova normalmente nelle distribuzioni del sistema T_EX. Perciò, mentre questo pacchetto è in grado di fornire dati sufficienti per soddisfare la dichiarazione, i metadati e la gestione dei font per tutte le varianti PDF/A, non è ancora in grado di produrre in ogni circostanza

file completamente conformi. Però con altro software esterno a pdfTeX che *sia* capace di produrre file di tipo "Tagged PDF", questo pacchetto può essere usato come parte del normale ciclo di lavoro per produrre file completamente compatibili.

1.1 Gli standard PDF

PDF/X e PDF/A sono termini generici per denotare diversi standard ISO (ISO 19005-1; ISO 19005-2; ISO 19005-3; ISO 15930-1; ISO 15930-3; ISO 15930-4; ISO 15930-7; ISO 15930-8) che definiscono sottoinsiemi diversi dello standard PDF (ADOBE SYSTEMS INC., 2006; ISO 32000-1; ISO 32000-2). Lo scopo di PDF/X è quello di facilitare lo scambio di elementi grafici tra il creatore del documento e il mezzo di stampa e perciò ha tutti i requisiti legati al processo di stampa. Per esempio con PDF/X tutti i font devono essere incorporati e tutte le immagini devono essere in colore CMYK oppure in colori spot. PDF/X-2 e PDF/X-3 accettano colori RGB calibrati e colori CIELAB oltre alle altre specifiche di PDF/X. Dal 2005 sono emerse altre varianti di PDF/X, visto che altri effetti (come gli strati (*layering*) e le trasparenze) sono gestiti dallo stesso standard PDF. La serie completa di versioni e di requisiti di conformità è discusso nel paragrafo 2.1.4.

PDF/A definisce un profilo per archiviare documenti PDF, che assicuri che i documenti possano essere riprodotti esattamente negli anni futuri. Un elemento chiave per ottenere questo scopo è che un documento PDF/A sia completamente autocontenuto. Tutta l'informazione necessaria per mostrare sempre in modo identico il documento deve essere incorporata nel file. Un documento PDF/A non deve appoggiarsi assolutamente su informazioni esterne al documento stesso. Altre restrizioni richiedono che il file non contenga effetti multimediali (come audio e video), né JavaScript né crittografia. L'inclusione dei font e la specificazione del profilo di colore è obbligatoria per un file PDF/A, come lo sono i metadati da inserire secondo uno standard ben definito. Versioni successive consentono la compressione delle immagini e che vi siano allegati altri file.

PDF/E è uno standard ISO (ISO 24517) finalizzato ai documenti usati nel processo di lavoro ingegneristico. PDF/VT (ISO 16612) è finalizzato alla produzione di documenti in grandi quantità personalizzati per esempio per stampare le bollette

*Col permesso degli autori questo loro articolo apparso su *TUGboat* 36(2015):2 è stato tradotto da C. Beccari; eventuali errori o travisamenti di significato sono dovuti solamente al traduttore. *N.d.T.*

delle utenze. PDF/UA (Universal Accessibility) è in fase di sviluppo come un nuovo standard (ISO 14289; ISO PDFUA) per contenere aspetti tecnologici di assistenza (ai disabili) che incorporano le linee guida (WCAG) per l'accesso in rete ai documenti elettronici. Nel futuro potrebbe venire sviluppato PDF/H per registrazioni relative alla salute e per documenti sanitari. Si possono immaginare ulteriori applicazioni. Le dichiarazioni e i metadati necessari sono disponibili per i primi due tipi. Le altre sono oggetto di lavoro futuro; si può prevedere che nei prossimi anni possano venire prodotte versioni modificate ed estese di questo pacchetto.

2 L'uso

Il pacchetto viene caricato col solito comando:

```
\usepackage[<opzioni>]{pdfx}
```

Le opzioni disponibili sono descritte qui di seguito¹.

2.1 Opzioni

2.1.1 Opzioni per PDF/A

PDF/A è uno standard ISO previsto per l'archiviazione di documenti elettronici nel lungo periodo. Perciò le proprietà di essere autocontenuto e sempre esattamente riproducibile sono tenute in particolare attenzione; richiede inoltre metadati leggibili direttamente su ogni macchina. Lo standard PDF/A ha tre livelli di conformità “a”, “b” e “u”. Il livello “a” è quello che richiede i requisiti più stringenti e non è ancora completamente sviluppato nel pacchetto pdfx. Il livello “u” ha i medesimi requisiti del livello “b”, con il requisito ulteriore che tutti i font usati nel documento siano codificati in UNICODE. Va notato che il pacchetto pdfx produce la mappatura UNICODE anche nei file di livello “b”. Lo standard ha inoltre tre diverse versioni, 1, 2, e 3, che sono state definite rispettivamente nel 2005, 2011 e nel 2012. Le versioni precedenti contengono un sottoinsieme delle prerogative delle versioni successive così per garantire la massima portabilità è preferibile usare una versione di numero più basso. Non esiste un livello di conformità “u” nella versione 1. Per gli usi più tipici del PDF/A è sufficiente usare il PDF/A-1b. Le opzioni sono le seguenti:

- **a-1a**: genera un PDF/A-1a; sperimentale, non ancora completamente implementato;
- **a-1b**: genera un PDF/A-1b;
- **a-2a**: genera un PDF/A-2a; sperimentale, non ancora completamente implementato;
- **a-2b**: genera un PDF/A-2b;
- **a-2u**: genera un PDF/A-2u;
- **a-3a**: genera un PDF/A-3a; sperimentale, non ancora completamente implementato;
- **a-3b**: genera un PDF/A-3b;
- **a-3u**: genera un PDF/A-3u.

Con l'espressione: ‘sperimentale, non ancora completamente implementato’, intendiamo sostanzialmente che la struttura ‘Tagged PDF’ non è gestita con questo pacchetto. Usando altro software basato su pdfTEX che *sia* capace di produrre questa marcatura completa, si possono davvero produrre documenti conformi².

2.1.2 Opzioni per PDF/E

PDF/E è uno standard ISO destinato ai documenti usati nel processo produttivo ingegneristico. A tutt'oggi c'è una sola versione dello standard PDF/E e si chiama PDF/E-1. L'opzione

- **e-1**: genera un PDF/E-1.

2.1.3 Opzioni per PDF/VT

PDF/VT è uno standard ISO destinato ai documenti in un formato di scambio per la stampa con dati variabili al fini di produrre documenti legati a certe forme di transazioni ed è una estensione dello standard PDF/X-4. Lo standard PDF/VT prevede tre livelli di conformità; il livello 1 è per file singoli; il livello 2 è per file multipli; e il livello 3 è per la consegna dei documenti in modo streamed. Al momento nessuno dei livelli di conformità PDF/VT è completamente ottenut con il pacchetto pdfx. Le opzioni sono le seguenti:

- **vt-1**: genera un PDF/VT-1; sperimentale, non ancora completamente implementato;
- **vt-2**: genera un PDF/VT-2; sperimentale, non ancora completamente implementato;
- **vt-3**: genera un PDF/VT-3; sperimentale, non ancora completamente implementato.

Con l'espressione: ‘sperimentale, non ancora completamente implementato’, intendiamo sostanzialmente che la struttura del documento in sezioni ‘DPart’ di ‘Form XObjects’, non è gestita da questo

2. Essendo in contatto con il team che si occupa di queste funzionalità per il sistema TEX, sono al corrente che la futura versione 1.50 di pdfTEX (non ancora distribuita) potrà produrre file di tipo Tagged PDF. (N.d.T.)

1. In Europa è necessario caricare il pacchetto dopo aver specificato la codifica in entrata, chiamando con l'opportuna opzione il pacchetto inputenc, e prima di chiamare il pacchetto hyperref senza opzioni; se invece di pdfLATEX si usa LuaLATEX bisogna comunque caricare il pacchetto prima di hyperref, sempre senza opzioni; se si vuole personalizzare il comportamento di hyperref, bisogna usare il comando \hypersetup. Con LuaLATEX, apportando un piccolo patch a pdfx e usando solo font OpenType, si riesce ad ottenere un file conforme alle norme anche se la documentazione di pdfx, al momento di tradurre questo file, non dice nulla in proposito.. (N.d.T.)

pacchetto. Questo è già possibile con il software pdfTEX, ma non nella maniera che si presta per una realizzazione completamente automatica, perché bisogna conoscere l'*Object number* di alcuni costrutti interni del file PDF. Tutti gli altri aspetti — la dichiarazione per il dizionario PDF `/Info`, i metadati e il profilo di colore — delle varianti PDF/VT sono gestite correttamente.

2.1.4 Opzioni per PDF/X

Lo standard ISO PDF/X è previsto per l'interscambio di file con contenuto grafico. Esso sottolinea gli aspetti che riguardano la stampa, come il fatto che i font siano incorporati insieme ai profili di colore. Lo standard PDF/X ha diversi livelli di conformità e di varianti. Le varianti di base sono X-1, X-1a, X-3, X-4 e X-5. (La versione riveduta e corretta X-2 è stata pubblicata nel 2003, ma è stata ritirata nel 2011, sostanzialmente per mancanza di interesse nell'usarla.) Lo standard PDF/X-1a è presente nelle revisioni del 2002 e del 2003, lo standard PDF/X-3 esiste nelle revisioni del 2001 e del 2003, gli standard PDF/X-4 e PDF/X1-5 esistono nelle revisioni del 2008 e del 2010. Alcune di queste varianti hanno una versione 'p' che permette l'uso di un profilo di colore esterno (invece di usare un profilo incluso), e/o una versione 'g' che consente l'uso di file grafici esterni. Inoltre PDF/X-5 ha una versione 'n' che estende lo standard PDF/X-4p, permettendo ulteriori spazi di colore oltre alla scala di grigi, lo spazio RGB e quello CMYK. Per molti usi tipici di PDF/X è sufficiente usare la variante PDF/X-1a.

- **x-1**: genera un file PDF/X-1;
- **x-1a**: genera un file PDF/X-1a; le ulteriori opzioni **x-1a1** e **x-1a3** sono disponibili per specificare esplicitamente le varianti approvate nel 2001 o nel 2003;
- **x-3**: genera un file PDF/X-3; le ulteriori opzioni **x-302** e **x-303** sono disponibili per specificare esplicitamente le varianti approvate nel 2002 o nel 2003;
- **x-4**: genera un file PDF/X-4; le ulteriori opzioni **x-408** e **x-410** sono disponibili per specificare esplicitamente le varianti approvate nel 2008 o nel 2010;
- **x-4p**: genera un file PDF/X-4p; le ulteriori opzioni **x-4p08** e **x-4p10** sono disponibili per specificare esplicitamente le varianti approvate nel 2008 o nel 2010;
- **x-5g**: genera un file PDF/X-5g; le ulteriori opzioni **x-5g08** e **x-5g10** sono disponibili per specificare esplicitamente le varianti approvate nel 2008 o nel 2010;
- **x-5n**: genera un file PDF/X-5n; le ulteriori opzioni **x-5n08** e **x-5n10** sono disponibili per

specificare esplicitamente le varianti approvate nel 2008 o nel 2010; sperimentale, non ancora completamente implementato.

- **x-5pg**: genera un file PDF/X-5pg; le ulteriori opzioni **x-5pg08** e **x-5pg10** sono disponibili per specificare esplicitamente le varianti approvate nel 2008 o nel 2010;

2.1.5 Altre opzioni

Le seguenti opzioni per il pacchetto pdfx sono sperimentali e normalmente non dovrebbero venire usate.

- **useBOM**: genera esplicitamente un marcatore di ordine di byte nella codifica UTF-8 (*UTF-8 byte-order marker*) nei metadati XMP inclusi e rende il pacchetto XMP scrivibile. Nessuna di queste funzionalità è richiesta per lo standard PDF/A, ma sembra che ci siano alcuni validatori di file PDF/A che li richiedono. Nota: l'implementazione di questa funzionalità è sperimentale e potrebbe cessare di funzionare con i futuri aggiornamenti del pacchetto xmpincl.
- **noBOM**: non genera il *byte-order marker*; questa è l'impostazione di default.
- **pdf13**: usa PDF nella versione 1.3, scavalcando la versione specificata dallo standard richiesto. Questo può impedire la validazione del file PDF.
- **pdf14**: usa PDF nella versione 1.4, scavalcando la versione specificata dallo standard richiesto. Questo può impedire la validazione del file PDF.
- **pdf15**: usa PDF nella versione 1.5, scavalcando la versione specificata dallo standard richiesto. Questo può impedire la validazione del file PDF.
- **pdf16**: usa PDF nella versione 1.6, scavalcando la versione specificata dallo standard richiesto. Questo può impedire la validazione del file PDF.
- **pdf17**: usa PDF nella versione 1.7, scavalcando la versione specificata dallo standard richiesto. Questo può impedire la validazione del file PDF.

2.2 Il file dei metadati

Come detto precedentemente, i documenti conformi allo standard PDF/A richiedono che vi siano inclusi alcuni metadati composti secondo precise direttive (DUBLIN CORE, 2010). Il pacchetto pdfx si aspetta che i metadati siano forniti tramite un file speciale dal nome `\jobname.xmpdata`. Qui `\jobname` è il nome del file `.tex` principale per

la composizione del documento. Per esempio se il vostro file principale si chiama `main.tex`, allora il file dei metadati deve chiamarsi `main.xmpdata`. Nessuno dei singoli campi dei metadati è obbligatorio, ma per la maggior parte dei documenti è ragionevole definire almeno il titolo e l'autore. Qui c'è un esempio di un breve file di metadati:

```
\Title{Storia della cottura al forno}
\Author{U. Fornaio\sep A. Pastaio}
\Keywords{Biscotti\sep crescenti}
\Publisher{Casa editrice I Forni}
```

Si noti che una lista di più autori e di più parole-chiave ha gli elementi separati con il comando `\sep`. L'uso della macro `\sep` è consentito solo nei campi `\Author`, `\Keywords` e `\Publisher`.

Dopo l'elaborazione dei metadati, la cartella di lavoro contiene un file del tipo `pdfa.xmpi` o `pdfx.xmpi` a seconda della variante PDF richiesta. Questo file costituisce il pacchetto dei metadati. Può essere controllato per verificarne la validità mediante un validatore on-line, come ad esempio: <http://www.pdflib.com/knowledge-base/xmp-metadata/free-xmp-validator>

2.3 Elenco dei campi per i metadati

Qui viene esposta la lista completa dei metadati attualmente definibili dall'utente corredati del loro significato. Se ne potranno aggiungere altri in futuro. I comandi che seguono possono essere usati solo nei file `.xmpdata`.

2.3.1 Informazioni generali

- `\Author`: contiene il nome e il cognome dell'autore del documento. Più autori vanno separati con `\sep`. Il nome può essere abbreviato con l'iniziale puntata.
- `\Title`: contiene il titolo del documento.
- `\Keywords`: contiene l'elenco delle parole chiave separate con `\sep`.
- `\Subject`: contiene un breve sommario del documento.
- `\Publisher`: Contiene la ragione sociale della casa editrice o dell'organizzazione che ha prodotto il documento.

2.3.2 Informazioni sul copyright

- `\Copyright`: contiene la dichiarazione del copyright.
- `\CopyrightURL`: contiene l'indirizzo web che descrive il possessore del copyright e/o i diritti di copyright relativi al documento.
- `\Copyrighted`: contiene "True" se il documento è soggetto a copyright o "False" altrimenti. Questo metadato è impostato automaticamente a "True" se sono stati specificati i campi

`\Copyright` o `\CopyrightURL`, ma può essere scavalcato. Per esempio, se la dichiarazione di copyright è "Di pubblico dominio", questo campo dovrebbe essere impostato a "False".

2.3.3 Informazioni sulla pubblicazione

- `\PublicationType`: contiene il nome del tipo di pubblicazione. Se questo campo viene specificato, esso può contenere solo uno dei nomi seguenti: book, catalog, feed, journal, magazine, manual, newsletter, pamphlet. Questo campo è automaticamente impostato a "journal" se si specifica il campo `\Journaltitle`.
- `\Journaltitle`: contiene il titolo della rivista dove è stato pubblicato il documento.
- `\Volume`: contiene il numero del volume della rivista.
- `\Issue`: contiene il numero della rivista.
- `\Firstpage`: contiene il numero della prima pagina dove inizia il documento pubblicato.
- `\Lastpage`: contiene il numero dell'ultima pagina dove finisce il documento pubblicato.
- `\Doi`: contiene il Digital Object Identifier (DOI) del documento senza il prefisso "doi".
- `\CoverDisplayDate`: contiene la data che compare sulla copertina del fascicolo della rivista come data da leggere da parte delle persone.
- `\CoverDate`: contiene la data che compare sulla copertina del fascicolo della rivista in un formato adatto per essere memorizzato in un campo di un database avente la tipologia "date".

2.4 Simboli consentiti nei metadati

Nei metadati sono usabili tutti i caratteri ASCII, tranne `\`, `char123`, `}` e `%`, i quali rappresentano se stessi. Inoltre sono usabili tutti i caratteri e i simboli UNICODE del piano di base (cioè fino al code point U+FFFF) ma devono essere rappresentati con la transcodifica UTF-8. (Attualmente non sono accettabili nei metadati codifiche diverse da UTF-8). Le sequenze di spazi vengono contratti in un solo spazio. Lo spazio dopo una macro come `\copyright`, `\backslash` o `\sep` viene ignorato. In più si può usare questo mark-up:

- `"\_:"` uno spazio reale, per esempio dopo una macro;
- `\%`: un % letterale;
- `\{`: una { letterale;
- `\}`: una } letterale;

- `\backslash`: un `\` letterale
- `\copyright`: il simbolo di copyright ©.

La macro `\sep` è consentita solo negli argomenti di `\Author`, `\Keywords` e `\Publisher`. Il suo scopo è proprio quello di separare una molteplicità di autori, parole chiave, eccetera. Però, ai fini della validazione, gli autori e le parole chiave non devono essere veramente separati, ma il pacchetto ne tiene conto anche quando si usa `\sep`.

Risulta che si possono usare anche altre macro TEX, purché l'autore sia molto attento ad evitare macro TEX troppo complicate o sviluppi di macro LATEX nei comandi di base o primitive diverse da caratteri; in sostanza sono accettabili solo gli accenti, le macro per caratteri latini speciali, e semplici sostituzioni testuali, talvolta con un semplice argomento. Viene fornita la macro speciale

```
\cs{pdfxEnableCommands{...}
```

per aiutare a risolvere situazioni difficili che potrebbero presentarsi.

Per esempio, `\pdfxEnableCommands` è necessaria per il nome di uno degli autori Hàn Thế Thành, a causa della 'e' con due accenti 'ê'. Si è soliti definire una macro

```
\def\thanh{H\`an Th\`{\'e} Th\`anh}
```

Nelle versioni precedenti del pacchetto pdfx l'uso di questa macro all'interno del file `.xmpdata`, per esempio nella dichiarazione di copyright eseguita con `\Copyright`, produceva lo sviluppo delle macro per gli accenti in termini di comandi primitivi, del tipo di

```
H\unhbox \voidb@x \bgroup
\let unhbox \voidb@x \setbox ...
```

e così di seguito per diverse righe. Questo evidentemente è fuori luogo fra i metadati XMP. Per aggirare questo inconveniente si sarebbero potute usare delle definizioni semplificate:

```
\pdfxEnableCommands{
\def\`#1{#1\`cc\`80}
\def\'#1{#1\'cc\`81}
\def\^#1{#1\^cc\`82}}
```

dove le sequenze `\`cc\`80`, `\`cc\`81`, `\`cc\`82`, permettono a TEX di inserire i byte in codifica UTF-8 per i caratteri definiti "combining accent". Questo funzionerebbe per i metadati che compaiono solo nel pacchetto XMP. Tuttavia non è sufficiente per il campo `/Author` del file PDF, che deve corrispondere esattamente all'elemento dei metadati `dc:creator`. Quello che serve, invece è:

```
\pdfxEnableCommands{
\def\thanh{H\`c3\`a0n Th\`eee Th\`c3\`a0nh}
\def\eee{\`c3\`aa\`cc\`81 }}
```

o la stessa cosa con 'à' inserito direttamente in UTF-8 al posto di `\`cc\`a0` ed 'ê' al posto di `\`c3\`aa`. Il motivo per fare ciò nasce dal comando `\pdfstringdef` che costruisce direttamente le lettere latine come singoli caratteri a ed ê, senza ricorrere ai "combining accents", almeno quando è possibile. Se i metadati non contengono esattamente le stesse cose, nonostante la stessa normalizzazione UNICODE, la validazione fallisce.

Con l'ultima versione del pacchetto pdfx, queste difficoltà sono state superate almeno per le lingue dell'Europa occidentale basate sull'alfabeto latino. La codifica d'entrata usata durante la lettura del file `.xmpdata` ora include l'interpretazione dei soliti comandi TEX per gli accenti al fine di produrre le giuste sequenze di byte per la codifica UTF-8. Stiamo lavorando per estendere questa codifica d'entrata per altre lingue scritte con altri alfabeti (per esempio, latino esteso, cirillico, greco, eccetera). Una parte significativa dei caratteri del piano di base Unicode può essere trattato nello stesso modo. Si potrebbero fornire anche i moduli per i caratteri CJK (cinese, giapponese, coreano), per i simboli matematici, ed altro. Questo però diventerebbe un forte carico per la memoria, quindi sono ancora necessari molti test significativi prima che queste estensioni possano far parte del pacchetto pdfx.

2.5 Profili di colore

La maggior parte dei documenti conformi agli standard PDF richiedono che un *profilo di colore* sia inserito nel file. In breve, questo profilo determina con precisione come i colori usati nel documento abbiano lo stesso aspetto anche quando vengano riprodotti su mezzi fisici diversi. Questo deve assicurare che il documento mantenga esattamente il suo aspetto anche quando viene stampato con stampanti diverse, con tipi diversi di carta, o vengano mostrati su schermi diversi. L'inclusione di un profilo di colore è necessario affinché il documento sia autocontenuto.

Siccome la maggior parte degli utenti di LATEX non sono dei professionisti grafici e spesso non sono particolarmente attenti alle sfumature di colore, il pacchetto pdfx è provvisto di un certo numero di profili di default che vengono inclusi nel documento finale se non si specifica qualcosa d'altro. Perciò l'utente medio non ha di che preoccuparsi per ciò che riguarda il colore. Gli utenti che desiderano usare un profilo di colore specifico, lo possono fare inserendo fra i metadati un comando `\setRGBcolorprofile` oppure `\setCMYKcolorprofile` nel file `.xmpdata`. Si noti che gli standard PDF/A e PDF/E accettano solo un profilo di colore RGB, mentre gli standard PDF/X e PDF/VT richiedono un profilo di colore CNYK³. Si usino quindi i comandi per specificare

3. salvo alcune varianti che accettano anche altri profili. (N.d.T.)

TABELLA 1: Comandi per diversi profili di colore

Comando	Nome del profilo
<code>\FOOGRAXXXIX</code>	Coated FOOGRA39 (ISO 12647-2:2004)
<code>\SWOPCGATSI</code>	U.S. Web Coated (SWOP) v2
<code>\JapanColorMMICoated</code>	Japan Color 2001 Coated
<code>\JapanColorMMIUncoated</code>	Japan Color 2001 Uncoated
<code>\JapanColorMMIINewspaper</code>	Japan Color 2002 Newspaper
<code>\JapanWebCoatedAd</code>	Japan Web Coated (Ad)
<code>\CoatedGRACoL</code>	Coated GRACoL 2006 (ISO 12647-2:2004)
<code>\SNAPCGATSII</code>	CGATS TR 002
<code>\SWOPCGATSIII</code>	CGATS TR 003
<code>\SWOPCGATSV</code>	CGATS TR 005
<code>\ISOWebCoated</code>	Web Coated FOGRA28 (ISO 12647-2:2004)
<code>\ISOCoatedECI</code>	ISO Coated v2 (ECI)
<code>\CoatedFOGRA</code>	Coated FOGRA27 (ISO 12647-2:2004)
<code>\WebcoatedFOGRA</code>	Web Coated FOGRA28 (ISO 12647-2:2004)
<code>\UncoatedFOGRA</code>	Uncoated FOGRA29 (ISO 12647-2:2004)
<code>\IFRAXXVI</code>	ISOnewspaper26v4 ISO/DIS 12647-3:2004
<code>\IFRAXXX</code>	ISOnewspaper30v4 ISO/DIS 12647-3:2004

i profili di colore RGB o CMYK secondo la sintassi seguente:

```
\setRGBcolorprofile<nome del file>
<identificatore><stringa info><URL di
registrazione>

\setCMYKcolorprofile<nome del file>
<output intent><identificatore><URL di
registrazione>
```

Negli argomenti di queste macro i caratteri `<`, `>`, `&`, `^`, `_`, `#`, `$`, `~`, rappresentano se stessi ma per il segno % bisogna usare il comando `\%`. Le impostazioni di default sono le seguenti:

```
\setRGBcolorprofile
{ sRGB_iEC61966-2-1_black_scaled.icc }
{ sRGB_iEC61966-2-1_black_scaled }
{ sRGB_iEC61966-2-1 with black scaling }
{ http://www.color.org }

\setCMYKcolorprofile
{ coated_FOGRA39L_arg1.icc }
{ Coated FOGRA39 }
{ FOGRA (ISO Coated v2 300% (ECI)) }
{ http://www.argyllcms.com/ }
```

Alcuni altri profili di colore si possono scaricare dall'International Color Consortium; si veda nel sito <http://www.color.org/iccprofile.xalter>.

In alternativa molti profili di colore sono allegati a molte applicazioni software della Adobe; questi diventano usabili anche con software non di origine Adobe. Ora il pacchetto `pdfx` include il codice per l'immissione diretta di questi profili nei documenti PDF, o per specificarli come profili esterni con le varianti PDF/X4p e PDF/X5pg. Due file `AdobeColorProfiles.tex` e

`AdobeExternalProfiles.tex` sono contenuti nella distribuzione del pacchetto `pdfx`. Il secondo file serve con le varianti PDF/X-4p e PDF/X5pg che non richiedono l'immissione del profilo di colore nel file del documento, mentre il primo serve con le altre varianti di PDF/X. Entrambi definiscono comandi per usare quei profili di colore come indicato nella tabella 1.

Al momento di scrivere solo i primi sei di quei profili di colore esterni (cioè usati con PDF/X-4p o PDF/X-5pg) portano a file PDF che risultano validati usando la versione corrente di Preflight del programma Adobe Acrobat Pro. Non è chiaro se gli altri profili (inclusi `\IFRAXXVI` e `\IFRAXXX`) falliscono a causa di dati non corretti o per problemi connessi con il software di validazione. Tutti, tranne gli ultimi due possono essere usati per profili inclusi validi, se i file corrispondenti sono effettivamente accessibili per l'inclusione. La macro seguente può essere usata per fissare il percorso (assoluto o relativo) nel sistema operativo locale per accedere ai file dei profili di colore.

```
\pdfxSetColorProfileDir
{ <percorso ai profili di colore Adobe> }
```

2.6 Note sulla rappresentazione interna dei metadati

All'interno del file PDF i metadati sono inseriti in due punti: alcuni dati vanno a finire nel dizionario `/Info` nativo del PDF, mentre alcuni altri dati vanno a finire in un pacchetto XMP memorizzato in una posizione separata all'interno del file. XMP è il formato dell'Extensible Metadata Platform dell'Adobe, ed è un formato basato su l'XML. Si veda l'Adobe Development Center (<http://www.adobe.com/devnet/xmp>) per maggiori informazioni in merito all'XMP. È anche

disponibile un XMP Toolkit SDK, usabile su GNU/Linux, Macintosh e Windows rilasciato con una licenza BSD modificata.

Alcuni dei metadati come l'autore, il titolo e le parole chiave sono memorizzati sia nel dizionario /Info sia nel pacchetto XMP. Per la validazione di conformità con gli standard le due copie dei dati devono essere identiche. Di ciò si occupa il pacchetto pdfx.

In linea di principio gli utenti possono ricorrere ad altri mezzi per includere un file XMP. In questo caso gli utenti dovrebbero creare un file pdfa.xmp o pdfx.xmp (o altro, a seconda del tipo di PDF) che contengono i dati predefiniti. Ma va osservato che questo modo di procedere si presta a facili errori e non lo si raccomanda alla maggior parte degli utenti. Se c'è un campo di metadati particolare di cui l'utente ha bisogno ma che non è ancora disponibile, l'utente è invitato a rivolgersi agli autori.

pdfx si serve del pacchetto xmpincl per inserire i dati XMP nel pdf. La documentazione del pacchetto xmpincl può essere utile per capire il meccanismo dei metadati XMP.

2.7 Tutorial e note tecniche

Un file di iniziazione con istruzioni passo-passo per generare file PDF/A si trova all'indirizzo:

<http://wwwmathstat.dal.ca/~selinger/pdfa>

Alcune note tecniche in merito ad alcuni problemi che gli autori hanno incontrato nel produrre e nel generare documenti PDF/A conformi allo standard si trovano all'indirizzo:

http://support.river-valley.com/wiki/index.php?title=Generating_PDF/A_compliant_PDFs_from_pdftex

3 Installazione

Il pacchetto pdfx è scaricabile da CTAN attraverso <http://ctan.org/pkg/pdfx>. Esso è già incluso nelle distribuzioni aggiornate e complete del sistema TeX come TeXLive e MiKTeX.

3.1 Limiti e dipendenze

pdfx.sty lavora con pdfTeX e anche con LuaTeX⁴. Inoltre esso dipende dai seguenti pacchetti:

4. Nella documentazione del pacchetto pdfx c'è scritto esplicitamente che esso lavora solo con pdfTeX (PDFLaTeX); in effetti lanciando la compilazione con LuaLaTeX su un file predisposto come indicato, *mutatis mutandis*, per lavorare con LuaLaTeX, la compilazione non va a buon fine, perché il motore LuaTeX (almeno al momento attuale in cui si sta eseguendo questa traduzione) manca di un comando nativo di pdfTeX; tuttavia si può richiamare il pacchetto di Heiko Oberdiek `pdfxcmds` che fornisce a LuaTeX i comandi specifici di pdfTeX ma non (più) inclusi in LuaTeX, e la compilazione va a buon fine; il file può risultare valido se sono verificate tutte le clausole richieste dallo standard. (N.d.T.)

1. `xmpincl` per l'inserimento dei metadati nel file PDF;
2. `hyperref` per eseguire i collegamenti ipertestuali, creare i bookmark, eccetera;
3. `glyptounicode.tex` per mappare i nomi dei glifi sui nomi UNICODE.

3.2 I file inclusi

I seguenti file fanno parte del pacchetto; alcuni sono generati da pdfx secondo le indicazioni di `Makefile`⁵.

3.2.1 I file del pacchetto

- `pdfx.sty`: è il file principale del pacchetto, (estratto da `pdfx.dtx`).
- `pdfa.xmp`: modello di file .xmp per produrre file PDF/A.
- `pdfe.xmp`: modello di file .xmp per produrre file PDF/E
- `pdfvt.xmp`: modello di file .xmp per produrre file PDF/VT.
- `pdfx.xmp`: modello di file .xmp per produrre file PDF/X.
- `8bit.def`: codifica d'entrata speciale.
- `l8uenc.def`: dichiarazione di macro per la codifica d'entrata.
- `glyptounicode-cmr.tex`: mappa i nomi dei glifi dei font Computer Modern e di altri font specifici di TeX sui nomi UNICODE.
- `coated_F00GRA39L_arg1.icc`: profilo di colore CMYK liberamente distribuibile.
- `sRGB_IEC61966-2-1_black_scaled.icc`: profilo di colore RGB, liberamente distribuibile.
- `ICC_LICENSE.txt`: licenza per i profili di colore.
- `AdobeColorPorfiles.tex`: macro per l'inclusione dei profili di colore della Adobe.
- `AdobeExternalProfiles.tex`: macro per usare profili di colore esterni.

5. L'utente che si serva di TeXLive o di MiKTeX non deve preoccuparsi di `Makefile` perché quelle distribuzioni del sistema TeX provvedono automaticamente all'installazione del pacchetto pdfx. (N.d.T.)

3.2.2 Documentazione

- **README:** è la solita informazione generale in formato `.txt` anche se manca di questa estensione.
- **manifest.txt:** elenco dei file del pacchetto.
- **sample.tex** e **sample.xmpdata:** un esempio di file con un esempio di file contenente i metadati.
- **small2e-pdfx.tex**, **small2e-pdfx.xmpdata:** un altro esempio con i corrispondenti metadati.

3.2.3 I file sorgente

- **src/pdfx.dtx:** “Documented TEX file” contenente il codice e la documentazione.
- **src/pdfx.ins:** il file batch per l’installazione.
- **src/rvdtx.sty:** usato dal file `pdfx.dtx`.
- **Makefile:** uno script Makefile per produrre la documentazione.

3.3 Licenza e contatti

Il pacchetto è soggetto alla LATEX Project Public Licence. Segnalazioni di errori, suggerimenti, richieste di ulteriori funzionalità, eccetera, possono essere inviate a qualunque degli autori indicati alla fine dell’articolo.

Riferimenti bibliografici

ADOBE SYSTEMS INC. (2006). «PDF Reference 1.7». URL http://www.adobe.com/devnet/pdf/pdf_reference.html.

DUBLIN CORE (2010). «Core Metadata Element Set, Version 1.1». URL <http://dublincore.org/documents/dces/>.

ISO 14289 (2012). «ISO 14289-1:2012; Document management applications — Electronic document file format enhancement for accessibility — Part 1: Use of ISO 32000-1 (PDFUA-1)». URL http://www.iso.org/iso/catalogue_detail?csnumber=54564. Technical Committee ISO/TC 171/SC2 Riveduto col nome ISO 14289-1:2104 all’indirizzo web http://www.iso.org.iso/catalogue_details?csnumber=64599.

ISO 15930-1 (2001). «ISO 15930-1:2001; Graphic Technology — Prepress digital data exchange — Use of PDF — Part 1: Complete exchange using CMYK data (PDF/X-1 and PDF/X-1a)». URL http://www.iso.org/iso/catalogue_details.htm?csnumber=29061. Technical Committee ISO/TC 130.

ISO 15930-7 (2010). «ISO 15930-7:2010; Graphic technology — Prepress digital data exchange — Use of PDF — Part 7: Complete exchange of printing data (PDF/X-4) and partial exchange of printing data with external colour profile reference (PDF/X-4p) using PDF 1.6». URL http://www.iso.org/iso/catalogue_detail.htm?csnumber=55843. Technical Committee ISO/TC 130.

ISO 16612 (2005). «ISO 16612-2:2010; Graphic Technology — Variable data exchange — Part 2: Using PDF/X-4 and PDF/X-5 (PDFVT-1 and PDFVT-2)». URL http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=38013. Technical Committee ISO/TC 130.

ISO 19005-1 (2005). «ISO 19005-1:200; Document Management — Electronic document file format for long term preservation — Part 1: Use of PDF 1.4 (PDF/A-1)». URL http://www.iso.org/iso/catalogue_detail?csnumber=38920. Technical Committee ISO/TC171/SC2. Revisions via corrigenda: ISO 19005-1:2005/Corr 1:2007 (marzo 2007); ISO 19005-1:2005/Corr 2:2011 (dicembre 2011).

ISO 19005-2 (2011). «ISO 19005-2:2011; Document Management — Electronic document file format for long term preservation; Part 2: use of ISO 32000-1 (PDF/A-2)». URL http://www.iso.org/iso/catalogue_detail?csnumber=50655. ISO/TC 171/SC2.

ISO 19005-3 (2012). «ISO 19005-3:2012; Document Management — Electronic document file format for long term preservation; Part 3: use of ISO 32000-1 with support for embedded files (PDF/A-3)». URL http://www.iso.org/iso/catalogue_detail?csnumber=57229. Technical Committee ISO/TC 171/SC2.

ISO 24517 (2008). «ISO 24517-1:2008; Document management — Engineering document format using PDF — Part 1: Use of PDF 1.6 (PDF/E)». URL http://www.iso.org/iso/catalogue_detail?csnumber=42274. Technical Committee ISO/TC 171/SC2.

ISO 32000-1 (2008). «ISO 32000-1:2008; Document Management — Portable document format (PDF 1.7)». URL http://www.iso.org/iso/catalogue_detail?csnumber=51502. Technical Committee ISO/TC 171/SC2.

ISO 32000-2 (2014). «ISO 32000-2:20140220; Document management — Portable document format — Part 2: PDF2.0». Technical committee ISO/TC 171/SC2 (in bozza febbraio 2014).

- ISO PDFUA (2015). «PDFUA Technical Implementation Guide: Understanding ISO 14289-1 (PDFUA-1)». URL <http://www.aiim.org/Research-and-Publications/standards/committees/PDFUA/Technical-Implementation-Guide>. AIIM Global Community of Information Technologies.
- ISO15930-3 (2002). «ISO 15930-3:2002; Graphic technology — Prepress digital data exchange — Use of PDF — Part 3: Complete exchange suitable for colour-managed workflows (PDF/X-3)». URL http://www.iso.org/iso/catalogue_detail.htm?csnumber=34941. Technical Committee ISO/TC 130.
- ISO15930-4 (2003). «ISO 15930-4:2003; Graphic technology — Prepress digital data exchange — Use of PDF — Part 4: Complete exchange of CMYK and spot colour printing data using PDF 1.4 (PDF/X-1a)». URL http://www.iso.org/iso/catalogue_detail.htm?csnumber=39938. Technical Committee ISO/TC 130.
- ISO15930-8 (2010). «ISO 15930-8:2010; Graphic technology — Prepress digital data exchange — Use of PDF — Part 8: Partial exchange of printing data using PDF 1.6 (PDF/X-5)». URL http://www.iso.org/iso/catalogue_detail.htm?csnumber=55844. Technical Committee ISO/TC 130. Rivisto via Corrigendum: ISO 15930-8:2010/Cor:2011 (agosto 2011).
- ▷ C.V. Radhakrishnan
River Valley Technologies
Trivandrum, India 695014
cvr (at) river-valley dot org
<http://cvr.cc/>
- ▷ Hàn Thế Thành
River Valley Technologies
Trivandrum, India 695014
thanh (at) river-valley dot org
- ▷ Ross Moore
Mathematics Department
Macquarie University
Australia 2109
ross (at) mq dot edu dot au
<http://maths.mq.edu.au/~ross/>
- ▷ Peter Selinger
Department of Mathematics and Statistics
Dalhousie University
Halifax, Nova Scotia B3H 4R2
Canada
selinger (at) mathstat dot dal dot ca
<http://www.mathstat.dal.ca/~selinger>