

Formati di fruizione



Obiettivo della lezione

- La gestione dei documenti digitali
- Formati di fruizione di documenti digitali
- PDF: un formato universale per documenti digitali. Storia ed uso
- Le applicazioni di publishing per documenti PDF: Acrobat & C.

Le fasi di elaborazione dei documenti digitali

- **Authoring:** fase (e relativi strumenti di *editing*) in cui vengono creati i **contenuti** di un documento digitale
- **Transformation:** fase (e relativi strumenti di *presentazione*) in cui vengono elaborati i **formati di fruizione** di un documento digitale
- **Delivery:** fase (e relativi strumenti di *publishing*) in cui un documento digitale viene **trasmesso** e fruito da un utente mediante qualche **dispositivo**

Authoring

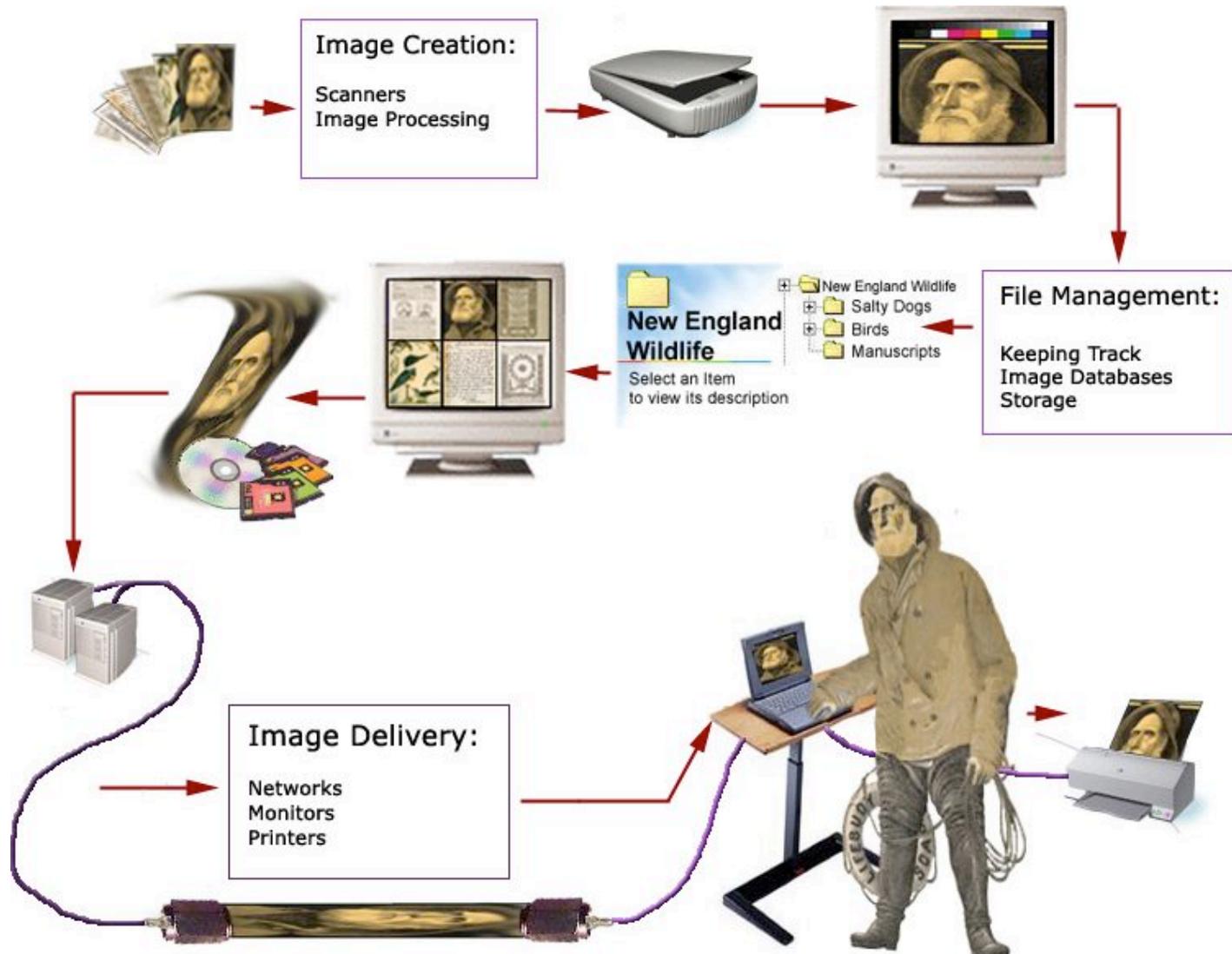
- Supponiamo che un autore sottometta per la pubblicazione un documento scritto in
 - Microsoft Word, o
 - LaTeX, o
 - Altro strumento di word processing
- L'autore, di solito,
 - Non definisce la struttura dei documenti
 - Non lavora linearmente
 - Non è interessato alla catena editoriale

Esempio:

Workflow di articoli scientifici

- Tradizionale
 1. Manoscritto cartaceo
 2. Articolo stampato su rivista
- Digitale
 1. Manoscritto digitale
 2. Articolo stampato
 3. Versione PDF
 4. Versione XML

Esempio: digitalizzazione



Cosa crea la fase di authoring

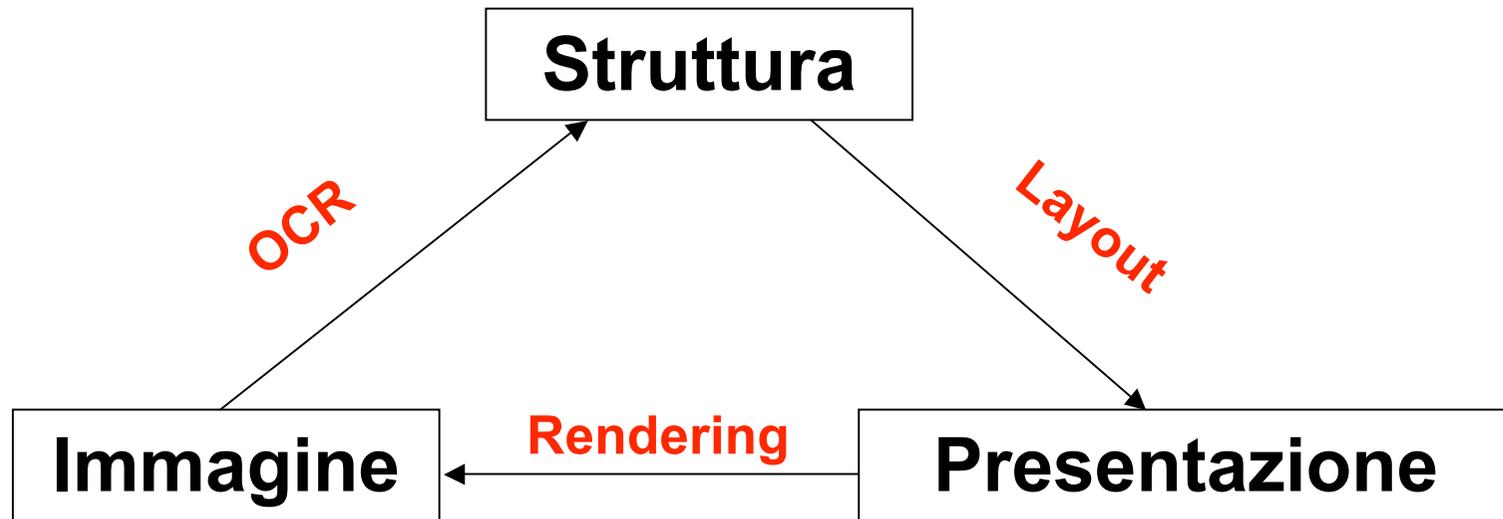
Gli strumenti di authoring (Es. Word) si basano sulla metafora della "pagina vuota": aiutano l'autore a riempirla di *contenuti*

I documenti digitali contengono principalmente

- Testo
- Grafica vettoriale
- Grafica bitmap

Chi deve gestire le fasi successive preferisce questi elementi nell'ordine di cui sopra

Conversione di documenti digitali



Transformation e Delivery

Mentre gli strumenti di authoring aiutano a "riempire pagine vuote",

- gli strumenti di *transformation* ridefiniscono la struttura o il layout di un documento
- gli strumenti di *delivery* "conservano le pagine" e le fanno arrivare al lettore così come l'autore vuole che arrivino

Qualità dei documenti digitali

- Fedeltà di riproduzione
- Rapidità di scaricamento
- Modificabilità (editabilità)
- Protezione del contenuto

Queste coppie di qualità sono antitetiche

Formato di fruizione

- **Formato di fruizione** (o *delivery format*): tipo del file che riceve l'utente che accede un documento digitale
- E' rilevante non solo per la miglior conservazione del contenuto e della sua forma, ma ad esempio anche per i motori di ricerca
- Nessun formato di fruizione esistente è superiore agli altri in ogni situazione
- Aspetti critici:
 - Formati aperti o chiusi
 - Usabilità e accessibilità
 - Aspetti sensibili al dominio applicativo dei documenti.
Esempio: documenti da archiviare

Formati

- Immagine: TIFF, JPEG, PNG, SVG
- Formati “aperti”: .txt, XML e HTML
- Formati ibridi: PDF, DjVu
- Formati ebook: .lit, .epub, .mobi
- Formati proprietari: .doc, .rtf
- Animazioni: Flash

Immagini pure: TIFF, GIF, JPEG

- TIFF miglior scelta per archivio immagini da scanner
- GIF scelta popolare per grafica su Web (ma non foto)
- JPEG formato compresso per immagini digitali

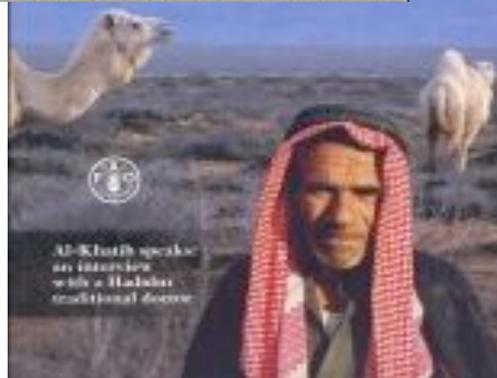
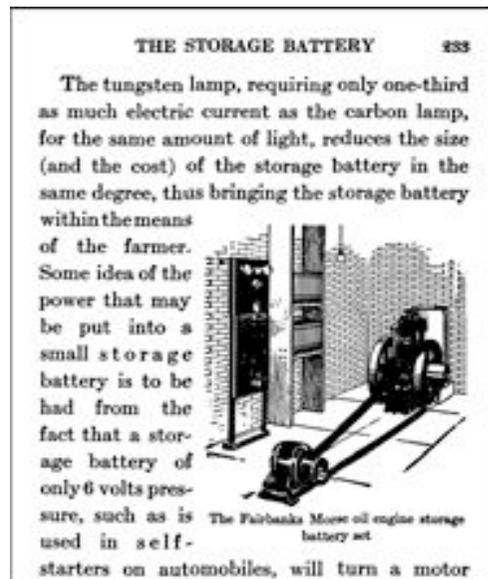
Pro

- Visualizzazione nativa nei web browser
- Formati aperti

Contro

- TIFF: Dimensione gigante dei file
- Testo disponibile solo via OCR (se il tasso di errore arriva al 20%, questo influenza i motori di ricerca)
- Supporto debole per documenti multipagina
- Quando un JPEG è fortemente compresso, la visualizzazione del testo è pessima
- Metadati solo sul file fisico, non sui contenuti

TIFF (Tagged Image File Format)



TIFF progettato da Aldus (oggi Adobe) nel 1987, l'ultima versione è del 1992. E' oggi il più comune formato bitmap, adatto per ogni profondità cromatica.

Buona scelta per l'archiviazione di documenti, ma non per publishing su Web.

partners.adobe.com/public/developer/en/tiff/TIFF6.pdf

TIFF

- Può memorizzare molti tipi diversi di immagini (monocrome, grigie, 8-bit & 24-bit RGB)
- Progettato per interoperabilità di applicazioni che manipolano immagini
- Diffuso per desktop publishing, scanning, e elaborazione delle immagini

GIF Graphics Interchange Format

- Standard del 1987, rivisto 1989
- Immagini a colori su 8-bit di profondità cromatica
- Dunque un'immagine GIF ha 256 colori (o 256 punti di grigio)
- Supporta trasparenza e animazione
- Molto diffuso a causa della sua efficienza
- Tutti i browser lo supportano efficacemente senza bisogno di plug-in

www.mwasoftware.co.uk/gif/gif89a.pdf

JPEG Joint **P**hotographics **E**xperts **G**roup

- **JPEG/JPG** (Joint Photographic Experts Group) non è un formato, quanto un *metodo* di codifica e compressione e si usa spesso insieme a TIFF
- Il formato prodotto dal metodo si chiama JPEG FIF (File Interchange Format) ed è standard dal 1992
- Progettato per comprimere immagini fotografiche a colori o B&W
- Ottimo per foto su Web; supporta milioni di colori con ottima comprimibilità

www.w3.org/Graphics/JPEG/

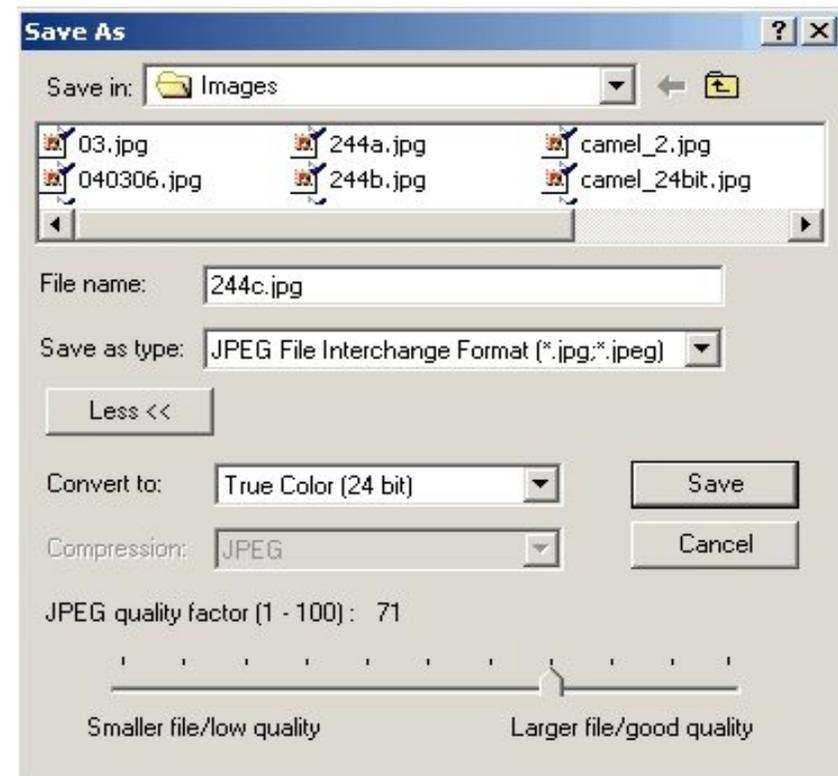
JPEG

JPEG offre un metodo di compressione per immagini fotografiche con una profondità da 6 a 24 bit.

La compressione è **lossy**, ma JPEG è considerato **visualmente lossless**

E' possibile scegliere quanta compressione applicare. In genere, un file JPEG comprime una foto da 2 a 3 volte in più rispetto a GIF

La compressione lossy rende JPG una **cattiva scelta** per scopi di archiviazione o comunque se in seguito occorre **piena qualità dell'immagine**



La profondità cromatica



1 bit (black & white)



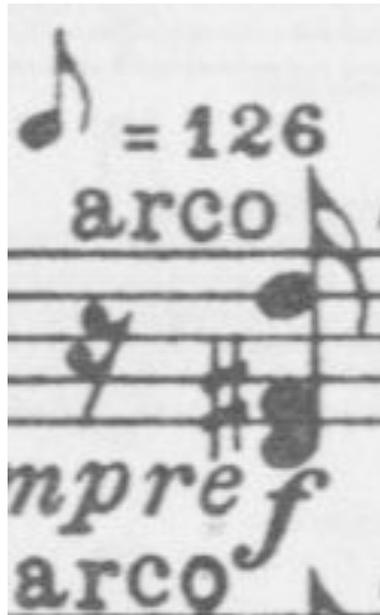
2 bit (4 colors)



4 bit (16 colors)

8 bit (256 colors)

Compressione in JPEG



Bassa compressione



Alta compressione

PNG (Portable Network Graphics)

- PNG è uno standard W3C per rimpiazzare GIF; profondità 16 (BW) o 48 (colore) bit
- Vantaggi su GIF: trasparenza variabile, controllo della luminosità, correzione del colore e visualizzazione progressiva
- Compressione lossless



SVG (Scalable Vector Graphics)

- SVG: standard W3C basato su XML
- Descrive grafica vettoriale a due dimensioni
- Occorre plugin (Adobe)



Tabella dei formati

	Progettato per...	Uso sul Web
TIFF	Immagini alta risoluzione da stampare o archiviare.	Non adatto perché un TIFF può essere gigantesco
GIF	Immagini con grosse aree omogenee di colore (e.g. logo, diagrammi, grafici)	Adatto, supportato da tutti i Web browsers
JPEG	Immagini con più di 256 colori (e.g. foto)	Adatto, supportato da tutti i Web browsers
PNG	Rimpiazzare GIF e TIFF	Adatto, supportato da molti Web browser
WMF EMF	Interscambio per MS Office di immagini vettoriali	Non adatto per publishing Web
EPS	Importazione di immagini PostScript	Non adatto per publishing Web
SVG	Immagini vettoriali, linguaggio basato su XML	Non supportato pienamente dai browser, occorre plugin

Formati aperti o chiusi

- Chi controlla il formato dei documenti?
 - Un comitato di standardizzazione? Il formato si dice *aperto*
 - Un'azienda specifica? Il formato si dice *chiuso*, o *proprietario*
- Vantaggi dei formati aperti (cioè standard)
 - Supporto nel futuro
 - Interoperabilità
 - Buona integrazione con software open source
 - Buona diffusione tra utenti e web designers

Formati proprietari

- MS Word, WordPerfect

Pro:

- Formati molto diffusi per authoring
- Integrazione con varie applicazioni (es. MS Office)

Contro:

- Formati chiusi, non facilmente condivisibili con altri programmi
- Strumenti di authoring di solito costosi
- Non adatti a fruizione su Web
- Mal supportati su piattaforme multiple
- Indicizzazione problematica con strumenti open source

RTF Rich Text Format

- RTF formato pubblicato da Microsoft: Versione 1.6 del 1999
- Specifica informazioni di layout su documenti testuali
- Basato su markup procedurale
- Utile come formato di interscambio anche al di fuori di applicazioni Microsoft
- <http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnrtfspec/html/rtfspec.asp>

Formati aperti: XML e HTML

Si creano in tre modi:

- Documenti digitali "nativi"
- Documenti creati a mano
- Da scanner e OCR, eventualmente con correzioni

Pro

- Ricerca facilitata (sono facilmente indicizzati dai motori di ricerca)
- Standard internazionali (sono raccomandazioni W3C)
- Accessibili per tutti i browser
- Metadati facili da aggiungere
- Editabili con qualsiasi editor (disponibilità di strumenti gratuiti)

Contro

- Costosi da creare a mano o correggere da OCR
- Pienamente sfruttabili solo da programmatori
- Non preservano il layout

Formati ibridi: PDF, DjVu

- Tecnologie proprietarie, tuttavia molto diffuse nelle comunità Open Source
- Li definiamo "ibridi" perché contengono entrambi un "layer" per il testo e un'immagine (*thumbnail*) per ciascuna pagina
- Disponibili nei browser via plugin
- Entrambi preservano l'aspetto dei documenti stampati
- Molti strumenti di conversione per entrambi i formati

DjVu

- Tecnologia proprietaria (Lizard) per publishing di documenti
- www.djvuzone.org

Pro:

- Compressione ottimale per il Web
- Documenti anche molto lunghi scaricabili e visualizzabili velocemente
- Plug-in DjVu multiplatforma e integrabile in applicazioni sw
- Server [Any2DjVu](#) disponibile gratuitamente per creare djvu

Contro:

- Poco diffuso in confronto a PDF
- Nessuno standard per i metadati nei documenti

PDF Portable Document Format

Pro:

- Molto diffuso, perché versatile e ricco di funzioni
- Base di PDF-Archive, uno standard di archiviazione di AIIM (Association for Information and Image Management)
- Manuale e reader software disponibili gratuitamente
- Offre un metodo standard per metadati: **XMP** (Extensible Metadata Platform), compatibile con Semantic Web

Contro:

- Troppe versioni dei reader, con funzioni diverse: gli utenti si possono confondere
- Prestazioni basate su plug-in lente per documenti lunghi
- Un file ottenuto da scanner può essere molto grande, anche per poche pagine (dipende dalla risoluzione)
- Adobe Acrobat è uno strumento costoso

PDF vs Djvu

- DjVu is a [raster](#) format, PDF is a [vector](#) format
- The [maximum](#) resolution of a DjVu file must be specified at creation time while a vector image represented by a PDF file can usually be magnified at arbitrary resolution
- DjVu files render characters as images, without using fonts. PDF files usually render characters using fonts. Many PDF files do not embed the full representation of the necessary fonts, but simply specify their names and properties. The PDF viewer uses the exact same font if it is available. Otherwise it transforms an available font to compute an approximation with the metrics of the desired font.
- Both DjVu and PDF files can be enriched with text, table of contents, [hyperlinks](#), and [metadata](#). The PDF format allows sounds, interactive forms, [JavaScript](#) program, and some DRM policies

Formati ebook

- .lit: per sw Microsoft
- .prc e .mobi: formati equivalenti, per Mobipocket, nativo per Kindle, basato su Open Ebook Standard, non stampabili
- .epub: basato su Open Ebook Standard

epub

- Uno standard per ebook definito da IDPF (International Digital Publishing Forum)
- Equivalente a .zip, contiene tre tipi di file:
 - XHTML, per i contenuti
 - File di impacchettamento (opf per metadati, ncx per la struttura dei file XHTML)
 - File contenitore DRM
- Principali applicazioni che leggono epub:
 - Adobe Digital Editions, Calibre, Stanza

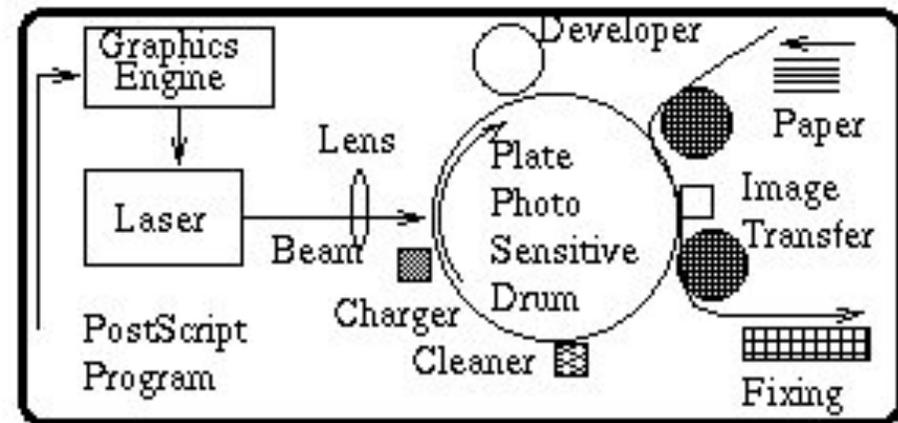
Storia e uso di PDF

Documenti e stampanti

- La qualità di stampa dipende dalla risoluzione della stampante
- La prima Apple Laserwriter nel 1985 aprì la strada al Desktop Publishing
- Adobe produsse il linguaggio Postscript per le Laserwriter
- Postscript è un linguaggio di descrizione di pagine da stampa

Stampanti laser

- Le stampanti laser sono computer che eseguono programmi PostScript
- La tecnologia di stampa è analoga a quella delle fotocopiatrici



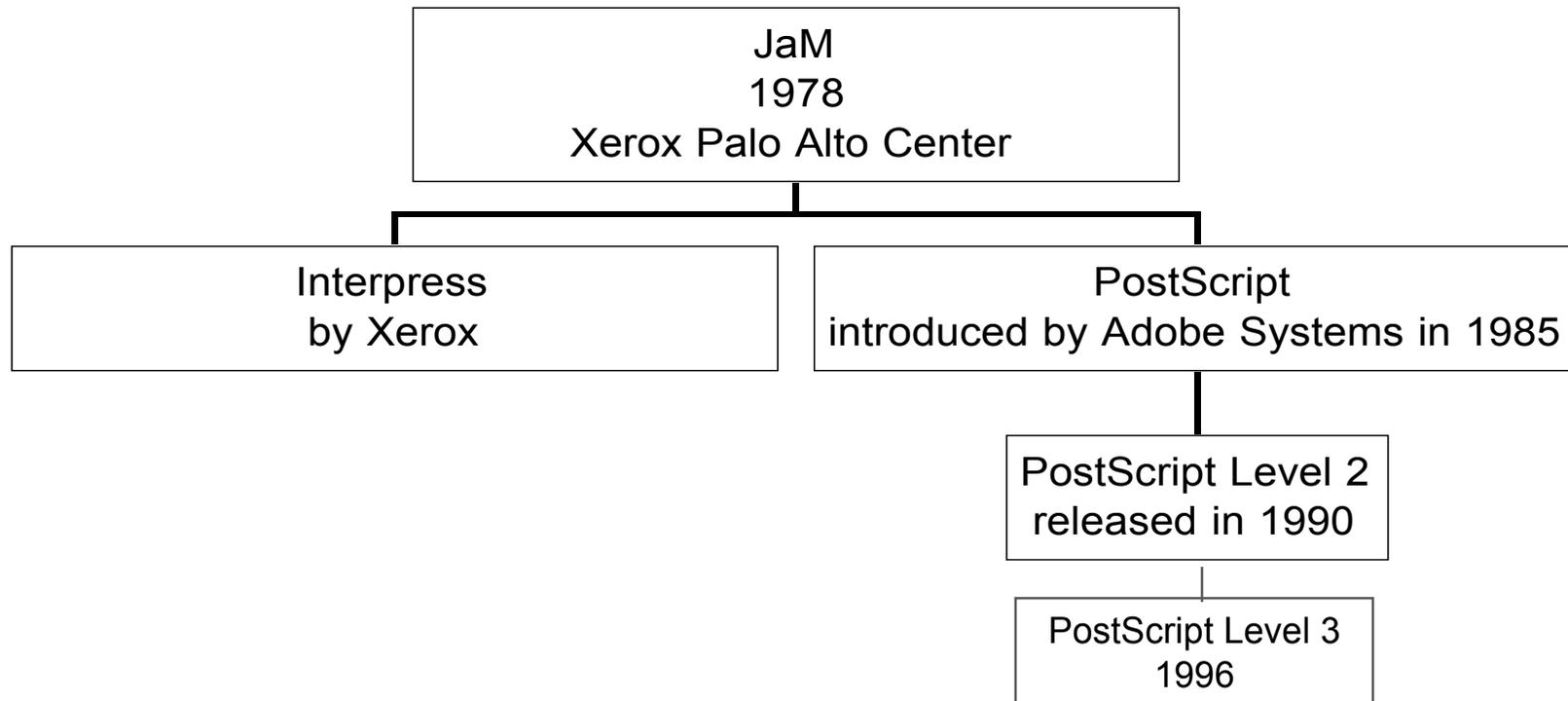
Prima di Postscript

- Prima del 1985 chi comprava una macchina tipografica si rivolgeva ad un'azienda X, che avrebbe venduto un sistema Y con dispositivo di stampa Z. Y e Z erano quasi sempre incompatibili con dispositivi di altri fabbricanti.
- Era impossibile scambiare dati con altri dispositivi
- Nella fascia inferiore del mercato, invece, i possessori di personal computer potevano utilizzare solo stampanti ad aghi capaci di stampare solo caratteri di bassa qualità

Linguaggi per descrivere pagine di stampa

- In Postscript una pagina si costruisce con comandi che creano
 - Oggetti geometrici di base (linee, rettangoli, cerchi, ecc.) disegnati da penne con punte di dimensione variabile
 - Caratteri in varie forme (fonti), di qualsiasi colore, dimensione e orientamento
 - Bitmap, ovvero figure descritte da griglie rettangolari di pixel
- La pagina costruita in Postscript può anche essere visualizzata su schermo!

Storia di Postscript



- Inventori di PostScript :
Warnock, Brotz, Paxton e Taft.

Postscript

- I documenti Postscript non sono fatti per essere editati
- Sono sequenze di istruzioni di stampa! Quindi sono fatti per essere eseguiti da una stampante abilitata al Postscript
- Il modello grafico di Postscript è simile a quello che si usa quando si disegna a mano: il programma Postscript guida un "pennello" che compone la pagina

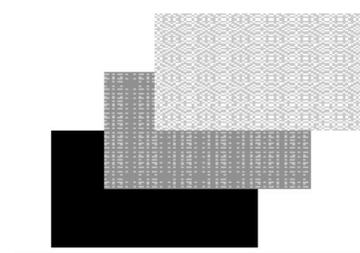
Un programma Postscript

```
%                BOX PROGRAM
%    ---- procedures ----
/inch {72 mul} def

/box {           % stack: x, y => _
newpath moveto
1 inch 0 rlineto
0 1 inch rlineto
-1 inch 0 rlineto
closepath } def

/fillbox {      % stack: grayvalue => _
setgray fill } def

%    -----main program -----
3.5 inch 4.5 inch box
0 fillbox
3.75 inch 5 inch box
0.4 fillbox
4 inch 5.5 inch box
0.8 fillbox
showpage
```



- **Le immagini Postscript sono opache**

Storia di Postscript

- La storia del PostScript inizia a XEROX PARC (Palo Alto Research Center). All'inizio degli '80 J.Warnock e C.Geschke svilupparono un linguaggio chiamato "Interpress", per controllare le stampanti laser Xerox
- Cercarono di convincere Xerox a trasformare Interpress in un prodotto commerciale
- Non ci riuscirono; lasciarono Xerox e fondarono Adobe con l'idea di sviluppare strumenti software per produttori di stampanti
- Con uno sforzo di circa 20 anni/uomo venne sviluppato PostScript, un linguaggio per dispositivi quali le stampanti laser

Postscript Level 1

- PostScript venne rilasciato al pubblico nel 1984; quella versione venne poi detta PostScript Level 1 per distinguerla dalla successiva
- PostScript necessita di hardware potente: nei primi anni le stampanti Postscript erano più potenti dei Macintosh cui erano collegate
- I vantaggi di PostScript all'epoca della sua invenzione:
 - PostScript rendeva i documenti indipendenti dal dispositivo. Su una Laserprinter di prima generazione si avevano 300 dpi, mentre con una imagesetter si arrivava a 2400 dpi: dunque gli utenti non erano più vincolati ad un dispositivo specifico per la stampa: potevano usare i dispositivi più adatti ai loro scopi
 - Ogni fabbricante di stampanti poteva comprare (sotto licenza) un interprete PostScript, da usare coi propri dispositivi di output
 - Le specifiche del PostScript furono pubblicate, quindi chiunque poté creare software di supporto a documenti PostScript

Encapsulated Postscript

- Il formato EPS venne introdotto per risolvere il problema dello scambio di documenti grafici *stampabili* tra diversi sistemi operativi.
- EPS si basa sul PostScript, quindi è indipendente dai dispositivi
- Anche se contiene comandi PostScript, EPS è un formato parzialmente indipendente dall'interprete PostScript, perché oltre al programma PostScript include di solito una bitmap del suo contenuto (usata per preview).
- Un terzo componente di ogni file EPS è una serie di *metadati*, soprattutto la dimensione dell'immagine, che verrà utilizzata dal programma che riceve il documento per definire lo spazio concesso all'immagine stessa entro un documento contenitore

eps in LaTeX

Package grafico:

- `\usepackage{epsfig}`

Comando LaTeX per inserire un file EPS:

- `\epsfig{figure=house.eps}`

Che cos'è un file PDF?

Formato proprietario di Adobe

- Congela la forma ed il layout dei documenti
- Multipiattaforma (Windows, Mac, ecc.)
- Può contenere immagini in altri formati
- Può supportare DRM

Varianti

- Tagged PDF (reflowable, dalla versione pdf1.5)
- Accessible PDF (tagged + altri vincoli)
- PDF/X (ISO standard, per pre-press, es. grafica trasmessa a stampatori di riviste)
- PDF/A (ISO standard, pdf1.4 per archiviazione)

PDF deriva da EPS

- Successore di Postscript, è complementare a Postscript 3, e basato su EPS
- Postscript descrive sequenze di pagine, PDF descrive insiemi di pagine
- Un documento PDF contiene testo, grafica, audio/video, metadati, link, programmi, in formato **compresso**
- Quando usare PDF?
 - Per preservare il layout
 - Per la compatibilità multiplatforma
 - Per difendere il contenuto da modifiche
 - Per documenti lunghi

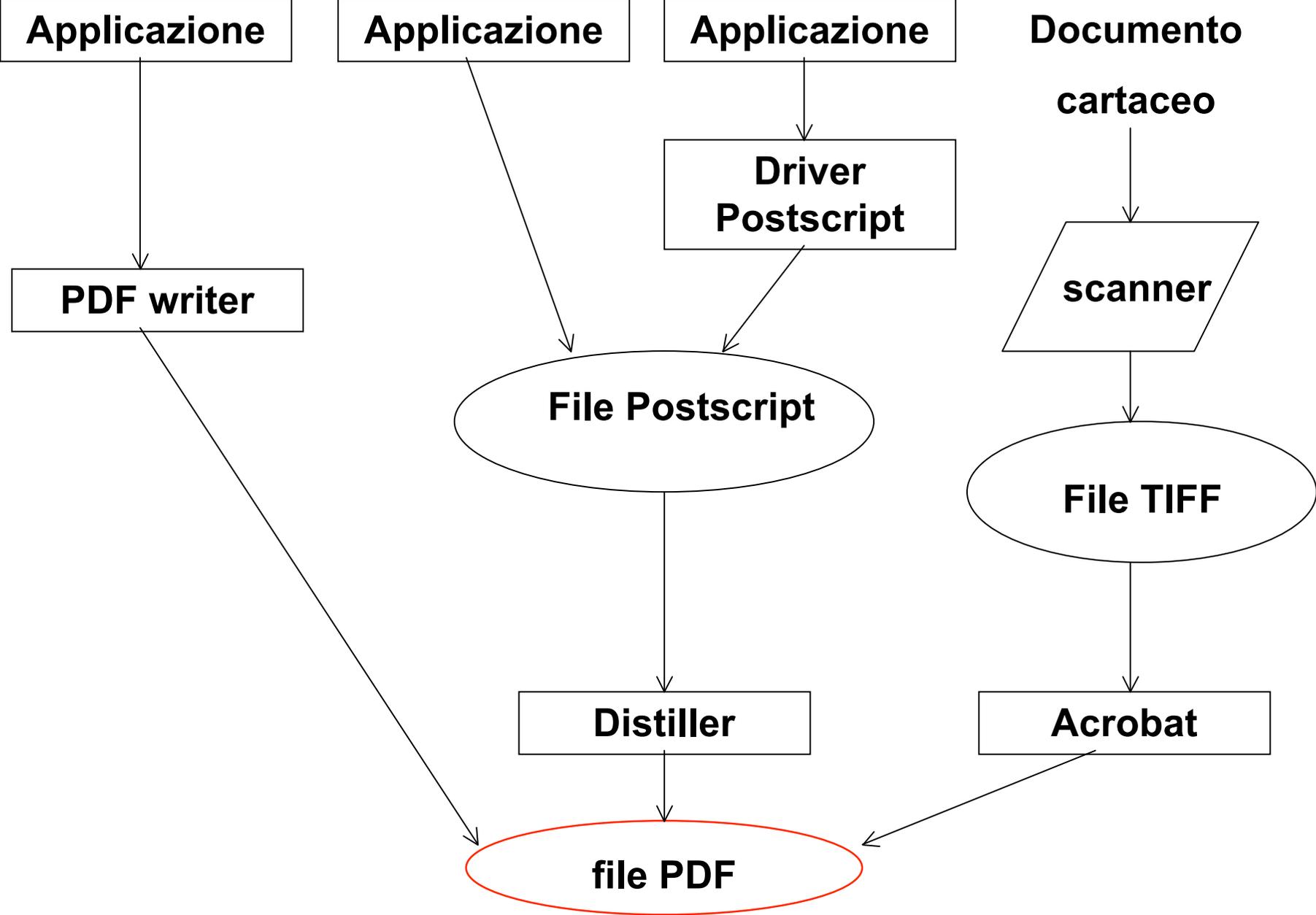
PDF vs Postscript

- Postscript è un linguaggio di programmazione per guidare i RIP (Raster Image Processor) delle stampanti
- EPS è un derivato del Postscript che definisce una singola pagina e contiene una figura di preview della pagina
- PDF deriva da EPS: gestisce insiemi di pagine "interpretate" da un RIP

Come si crea un PDF

Quattro modi:

- Esportazione diretta da applicazione
- Distiller, via Postscript
- PDFWriter (driver stampa virtuale)
- Da scanner, con Acrobat

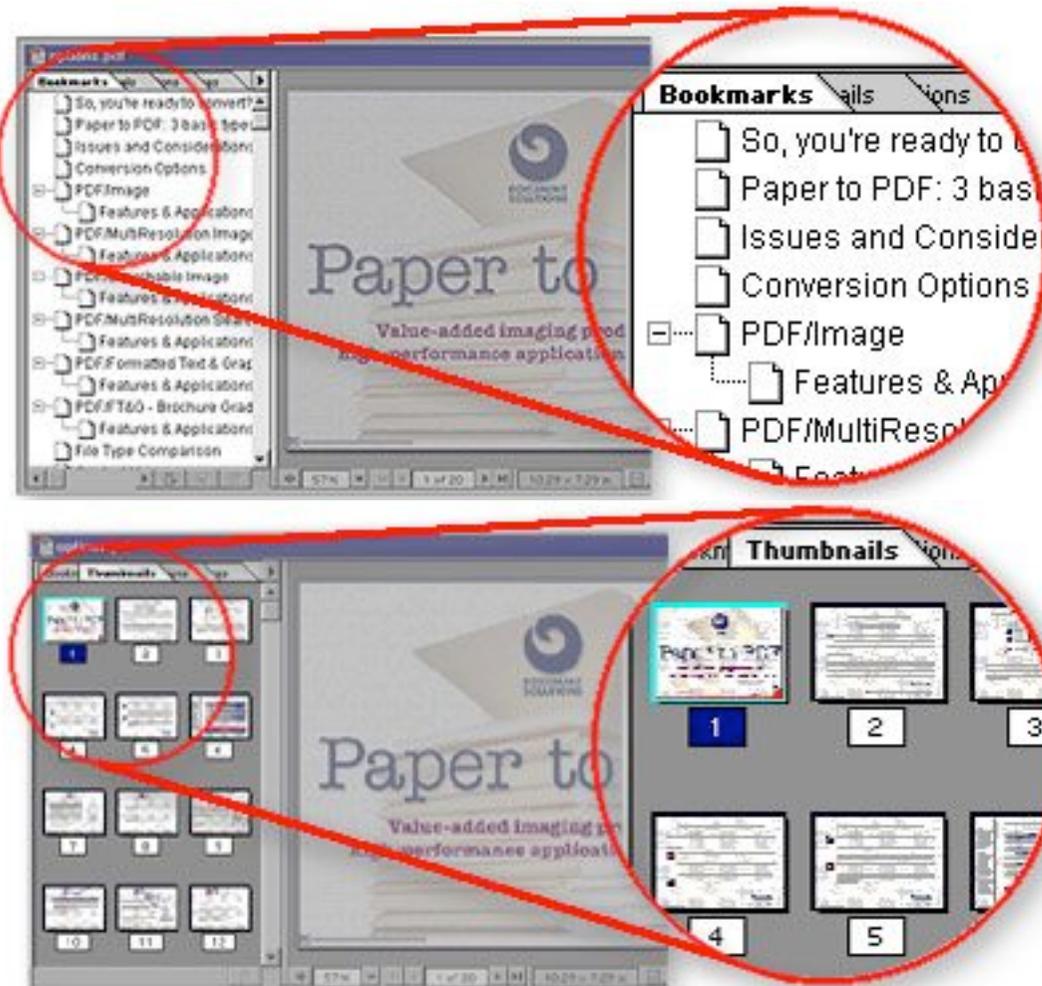


Creare un PDF con LaTeX

- Compilare con pdftex
- Per inserire i servizi, usare package hyperref
- Scegliere le opzioni. Esempio:

```
\usepackage[pdftex=true,  
  pdfauthor={Paolo Ciancarini},  
  pdftitle={La scacchiera invisibile},  
  colorlinks=true,  
  linkcolor=blue,  
  citecolor=red,  
  urlcolor=cyan]{hyperref}
```

Accessori del file PDF



Servizi accessori

- **Bookmarks:** indice ipertestuale
- **Thumbnails:** immagini miniature delle pagine
- **Hyperlinks:** link intra- o inter-testuali
- **Article Thread:** ordine di lettura alternativo
- **Named Destinations:** notazione per ancore su documenti da collegare in forma ipertestuale

Storia di PDF e Acrobat

- 1991: Adobe annuncia Interchange Postscript
- 1992: PDF1.0
- 1993: Acrobat 1.0
- 1994: Acrobat 2.0 e PDF1.1
- 1996: Acrobat 3.0 e PDF1.2
- 1998: Standard PDF/X
- 1999: Acrobat 4.0 e PDF1.3 (696 pagine)
- 2000: Illustrator e PDF1.4 (978 pagine)
- 2001: Acrobat 5.0
- 2003: Acrobat 6.0 e PDF1.5 (1172 pagine)
- 2004: Acrobat 7.0 e PDF1.6 (1213 pagine)
- 2005: PDF/A diventa standard ISO
- 2006: Acrobat 8.0 e PDF1.7
- 2007: PDF1.7 diventa ISO 32000
- 2008: Acrobat 9 e PDF1.7 extension 3

L'idea iniziale

Our vision for Camelot (=Acrobat) is to provide a collection of utilities, applications, and system software so that a corporation can effectively capture documents from any application, send electronic versions of these documents anywhere, and view and print these documents on any machines.

J.Warnock, 1991

Le versioni PDF

- Il formato PDF nasce nel 1992 con la versione 1.0
- Nel 2007 è uscita la versione 1.7
- Ogni versione aggiunge capacità grafiche nuove. Es.: nella 1.4 venne aggiunta la grafica trasparente

Il problema delle versioni PDF

Nota bene: a causa della continua evoluzione tecnologica, sulla stessa macchina programmi diversi creeranno diversi PDF. Ad esempio, sul mio MacOS

- Tex2pdf crea PDF1.4
- Ghostscript più visualizzatore PDF di MacOS creano PDF 1.3
- Distiller 7.0 crea PDF1.4
- Acrobat Professional 7.0 crea PDF1.6
- Acrobat Professional 8.0 crea PDF1.7

La specifica PDF 1.6

- Sintassi
- Grafica
- Testo
- Rendering
- Trasparenza
- Funzioni interattive
- Funzioni multimediali
- Interscambio: metadati, accessibilità e prepress

PDF/X

- PDF/X sottoinsieme della specifica PDF
- Molti editori preferiscono questo formato al TIFF quando devono ricevere materiale pubblicitario da pubblicare
- Molto usato per trasmettere grafica da stampare in riviste o brochure
- PDF/X restringe il PDF proibendo elementi non necessari alla stampa di alta qualità, quali Javascript e multimedia.
- PDF/X aiuta ad eliminare errori di trasformazione

PDF/A

- **ISO 19005-1:2005** è uno Standard ISO
- Scopo: conservazione su lungo periodo di documenti digitali
- Lo standard definisce un formato (**PDF/A**) per l'archiviazione di documenti digitali; si basa su PDF 1.4
- PDF/A è un sottoinsieme di PDF, che trascura tutte le funzioni inadatte all'archiviazione su lungo periodo

PDF/A

- PDF/A cerca di ottenere:
 - Massima indipendenza dall'hardware
 - Autocontenuto
 - Autodocumentato
- Vincoli:
 - Audio e video sono proibiti
 - Javascript e invocazioni di file eseguibili sono proibiti
 - Tutte le fonti necessarie sono incluse; l'inclusione dev'essere legale, per permettere la stampa ovunque ed per sempre
 - Colori specificati in modo indipendente dall'hardware
 - Crittazione proibita
 - Obbligatorie metadati standardizzati

Standard futuri (secondo Adobe)

- **PDF/E:** accepted as an ISO work item, PDF/E provides guidelines for the reliable exchange of documentation and drawings. Architects, engineers, construction professionals, and manufacturing product teams will be able to exchange project documents, including complex 3D content
- **PDF/UA:** provides guidance for making PDF information universally accessible. Businesses and government agencies can create PDF/UA-compliant documents that can be accessed anytime, on any hardware or software platform, by anyone even people with disabilities

Acrobat

The screenshot displays the Adobe Acrobat application window. The title bar shows the file name 'promessi sposi.pdf'. The top menu bar includes options like 'Apri', 'Registra', 'Stampa', 'Email', 'Ricerca', 'Crea PDF', 'Rivedi e commenta', 'Protezione', 'Firma', and 'Modifiche avanzate'. Below the menu bar is a toolbar with icons for text selection, zoom (set to 153%), and other navigation tools. The main content area shows the first page of the PDF, which is the title page of 'I Promessi sposi' by Alessandro Manzoni. The page features the title 'I Promessi sposi' in a large serif font, followed by the author's name 'di Alessandro Manzoni' in a smaller italicized font. On the left side, there is a vertical sidebar with tabs for 'Segnalibri', 'Firme', 'Livelli', 'Pagine', and 'Commenti'. The 'Pagine' tab is active, showing a thumbnail view of the document's pages, with page 1 highlighted. On the right side, there is a 'Come fare per..' (How to) panel with a 'Nascondi' (Hide) button. This panel contains several action buttons: 'Crea PDF', 'Rivedi e commenta', 'Protezione', 'Firma', 'Modifiche avanzate', 'Stampa produzione', 'Altri argomenti...', and 'Acrobat 6.0-Guida completa'. At the bottom of this panel, there is a checked checkbox labeled 'Mostra la finestra Come fare per all'avvio'. The bottom status bar shows the page dimensions '109,9 x 177 mm' and the page number '1 di 711'.

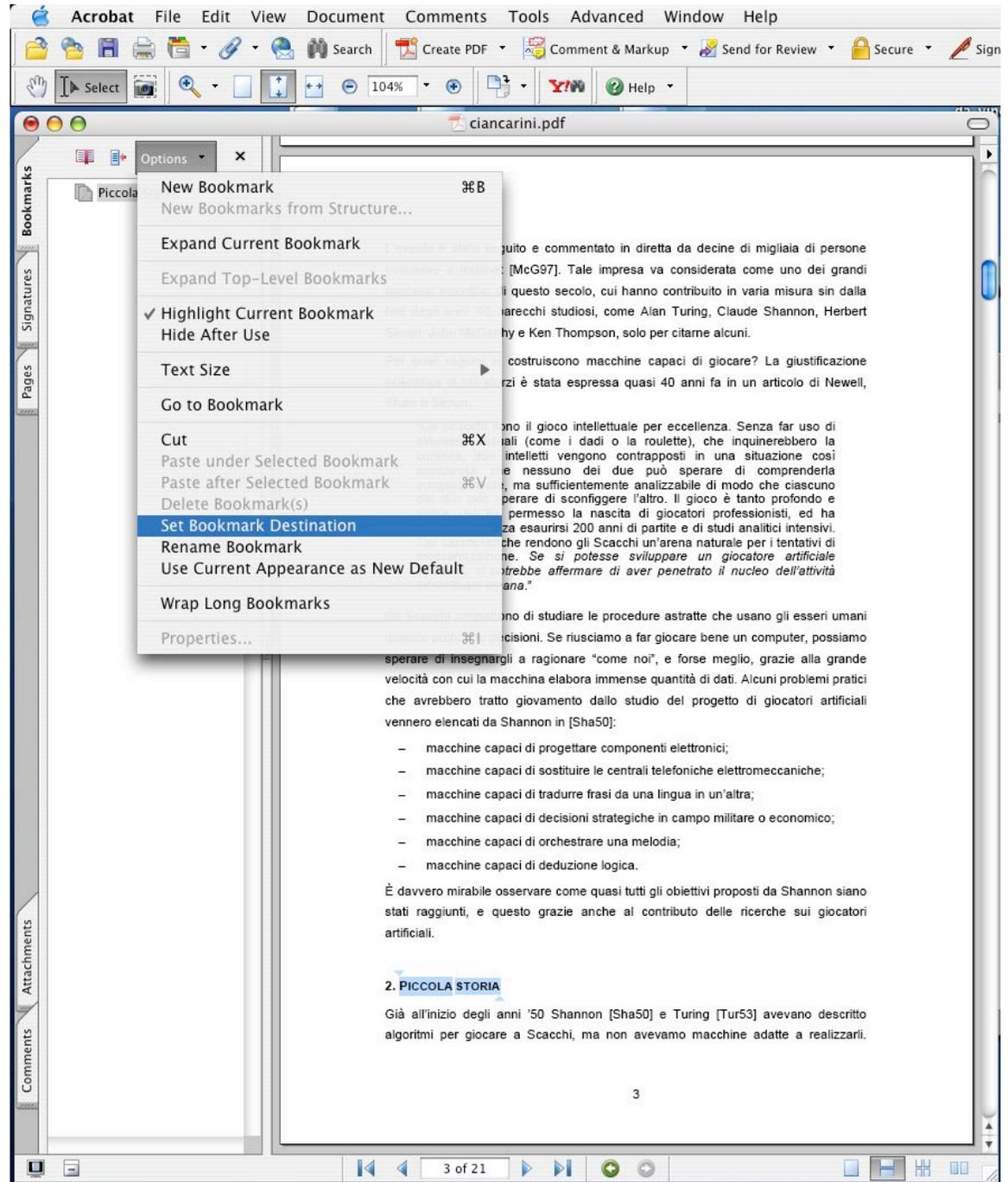
Acrobat

- Adobe Acrobat è l'applicazione di riferimento per manipolare file PDF
- Esistono al momento (2009)
 - Acrobat Reader (gratuita)
 - Acrobat Elements (a pagamento)
 - Acrobat Standard (a pagamento)
 - Acrobat Professional (a pagamento)
 - Acrobat 3D (a pagamento)

Classi di funzioni di Acrobat

- Creazione di PDF
- Publishing
- Revisioni e commenti
- Lettura e navigazione
- Accessibilità
- Sicurezza e autenticazione
- Digital Rights Management (Ebook)

Esempio: Usare Acrobat per aggiungere un bookmark



Esempio: funzioni interattive

- PDF supporta modulistica (es. tasse)
- E' possibile sfruttare transizioni e altri elementi di grafica animata
- Ma non solo: vedere demo

Proprietà dei documenti PDF

Document Properties

Description Security Fonts Initial View Custom Advanced

Description

File: ciancarini

Title:

Author:

Subject:

Keywords:

Created: 19/04/05 23:12 Additional Metadata...

Modified: 19/04/05 23:14

Application: Word

Advanced

PDF Producer: Mac OS X 10.3.9 Quartz PDFContext

PDF Version: 1.3 (Acrobat 4.x)

Location: MacintoshHD:Users:paolo:Desktop:sistema:

File Size: 4,44 MB (4.656.733 Bytes)

Page Size: 8,27 x 11,69 in Number of Pages: 21

Tagged PDF: No Fast Web View: No

Help Cancel OK

Proprietà: sicurezza

Password Security - Settings

Compatibility: Acrobat 7.0 and later

Encryption Level: High (128-bit AES)

Select Document Components to Encrypt

- Encrypt all document contents
- Encrypt all document contents except metadata (Acrobat 6 and later compatible)
- Encrypt only file attachments (Acrobat 7 and later compatible)

 All contents of the document will be encrypted and search engines will not be able to access the document's metadata.

Require a password to open the document

Document Open Password:

 When set, this password is required to open the document.

Permissions

Use a password to restrict printing and editing of the document and its security settings

Permissions Password:

Printing Allowed: High Resolution

Changes Allowed: Commenting, filling in form fields, and signing existing signature fields

Enable copying of text, images, and other content

Enable text access for screen reader devices for the visually impaired

Accessibilità

- Quale software occorre per vedere il formato?
- Un browser web lo gestisce direttamente?
- Se il formato richiede al browser un plug-in:
 - È gratis?
 - L'utente medio lo conosce già?
 - Funziona su tutti i sistemi operativi?
- I motori di ricerca indicizzano il formato?
- Il formato si presta alla trasmissione su linee di comunicazione lente?

Documenti accessibili

- Struttura logica e ordine di lettura
- Testo alternativo (descrizioni) per figure, link e campi da riempire
- Supporto alla navigazione.
- Funzioni di sicurezza che non interferiscano con la tecnologia che assiste i lettori disabili
- Fonti che permettono di accedere facilmente al testo

www.adobe.com/enterprise/accessibility

PDF: livelli di accessibilità

1. Unstructured
2. Structured
3. Tagged

Questi tipi di documento differiscono nell'informazione strutturale ed in come possono essere ristrutturati (**repurposed**)

PDF non strutturato

- PDF è un file immagine
 - non contiene testo
 - documento senza tag
 - Reader e motori di ricerca non riescono ad accedere il testo
- Un file PDF **unstructured** contiene "contenuto" che può essere salvato in altri formati, come RTF
- Gli utenti disabili non accedono

PDF non strutturato

What is a PDF?

PDF files are universally used because it preserves all the fonts, formatting, colors, and graphics of any source document, regardless of the application and the platform used to create it.

When is it appropriate to use PDF?

- It is essential to preserve exactly the visual the formatting and layout of the original document.
- No electronic version exists and you need to publish a paper document
- You want to ensure people cannot modify the document. (Copyright issues)

Learning how to create accessible PDF is easy!



Vuoto

Aspetto visuale

Testo letto da reader

PDF strutturato

- PDF è un'immagine con testo "incastonato"
- Documenti parzialmente accessibili da tecnologia per disabili
- I PDF strutturati contengono un albero di struttura logica che definisce un ordine di lettura primario sul testo incastonato
- Si creano con alcune applicazioni, quale per esempio Framemaker
- Il salvataggio in altri formati riconosce parte della formattazione, ma non liste o tabelle
- Un programma reader PDF può leggere un documento strutturato "semplice"

PDF strutturato

What is a PDF?

PDF files are universally used because it preserves all the fonts, formatting, colors, and graphics of any source document, regardless of the application and the platform used to create it.

When is it appropriate to use PDF?

- It is essential to preserve exactly the visual the formatting and layout of the original document.
- No electronic version exists and you need to publish a paper document
- You want to ensure people cannot modify the document. (Copyright issues)

Learning how to create accessible PDF is easy!



Aspetto visuale

What is a PDF? PDF files are universally used because it preserves all the fonts, formatting, colors, and graphics of any source document, regardless of the application and the platform used to create it. When is it appropriate to use PDF? It is essential to preserve exactly the visual the formatting and layout of the original document. No electronic version exists and you need to publish a paper document. You want to ensure people cannot modify the document. (Copyright issues. Learning how to create accessible PDF is easy)

Testo letto da reader

PDF con tag (tagged)

- Il file PDF più completo, che contiene
 - Immagine del documento
 - testo
 - Tags
- Si crea con Acrobat ma anche con applicazioni come Word, se il documento è ben strutturato e predisposto per l'accessibilità
- Questo tipo di file PDF è il più accessibile

Tagged PDF

Accessibility of PDFs

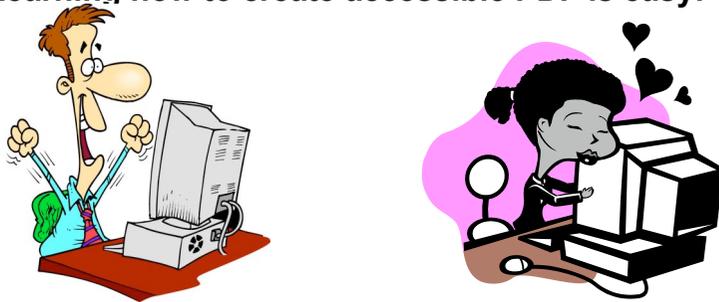
What is a PDF?

PDF files are universally used because it preserves all the fonts, formatting, colors, and graphics of any source document, regardless of the application and the platform used to create it.

When is it appropriate to use PDF?

- It is essential to preserve exactly the visual the formatting and layout of the original document.
- No electronic version exists and you need to publish a paper document
- You want to ensure people cannot modify the document. (Copyright issues)

Learning how to create accessible PDF is easy!



Aspetto visuale

H1 Accessibility of PDF's

H2 What is a PDF?

PDF files are universally used because it preserves all the fonts, formatting, colors, and graphics of any source document, regardless of the application and the platform used to create it.

H2 When is it appropriate to use PDF?

Bullet It is essential to preserve exactly the visual the formatting and layout of the original document.

Bullet No electronic version exists and you need to publish a paper document

Bullet You want to ensure people cannot modify the document. (Copyright issues)

H2 Learning how to create accessible PDF is easy!

Image Cartoon of a person in front of a computer raising his hands

Image Cartoon of a lady kissing her computer

Testo letto da reader

Come creare un PDF tagged

Solo Office su Windows 2000 o XP (no Mac)

1) in Word, dare formato a:

- Intestazioni
- Liste
- Immagini (aggiungere testo alternativo)
- Hyperlinks

2) “Convert to PDF”

Controllo accessibilità con Acrobat

Accessibility Full Check

Report and Comment Options

Create Accessibility Report

Browse...

Folder: MacintoshHD:Users:paolo:Documents:

Include repair hints in Accessibility Report

Create comments in document

Page Range

All pages in document

Only those pages selected in Pages panel

Pages from to

Checking Options

Alternative descriptions are provided

Text language is specified

Reliable character encoding is provided

All content is contained in the document structure

All form fields have descriptions

List and table structure is correct

Tab order is consistent with the structure order

Help Cancel Start Checking

Standard per creare PDF

- **PDF Style Guide**

www.eere.energy.gov/communicationstandards/web/pdfs.html

- **Ottimizzare PDF per i motori di ricerca**

www.eere.energy.gov/communicationstandards/web/pdfs_opt.html

Publishing

- Il publishing consiste nella produzione e distribuzione di documenti
- Caratteristiche di fruizione dei documenti digitali
- Caratteristiche di produzione dei documenti digitali (workflow)

Publishing tradizionale

- Stabilità del documento
- Alta fedeltà: il documento appare esattamente come l'originale
- Documento non interattivo
- Modello di colore importante
- Integrazione nei workflow esistenti

Publishing digitale

- Molte funzioni di interazione
- Documenti da completare (moduli)
- Interazione con database
- Reflow del testo per piccoli monitor
- Riutilizzo delle informazioni
- Sicurezza: autenticazione e firme digitali

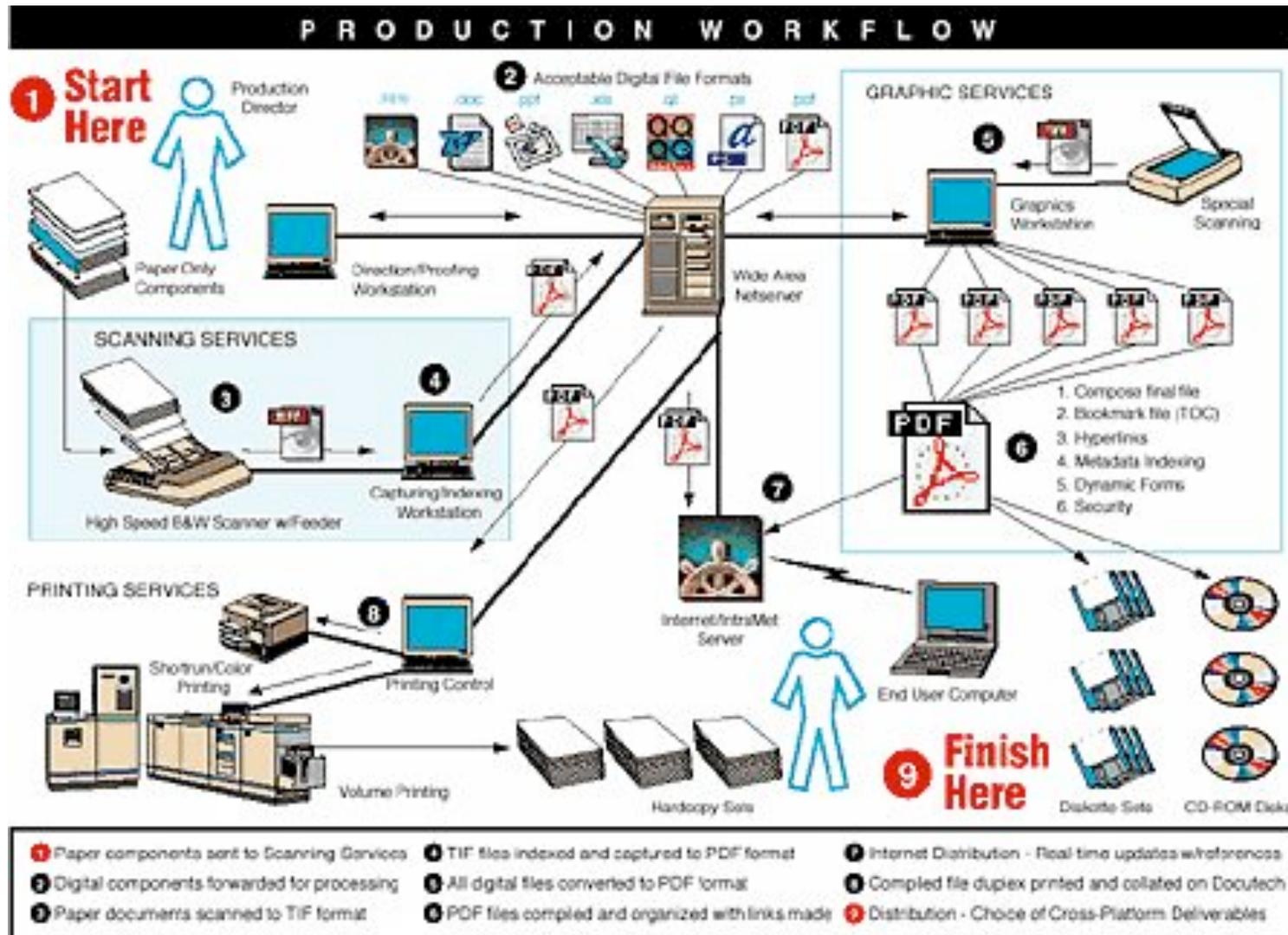
PDF per documenti digitali

- Struttura logica adatta al reflow (adattamento del layout al dispositivo) e repurposing (adattamento del contenuto all'applicazione)
- Funzioni di sicurezza
- Dati dei moduli esportati in XML
- Collaborazione integrata con Web
- Accessibilità per persone con difetti visivi
- Supporto all'integrazione con database

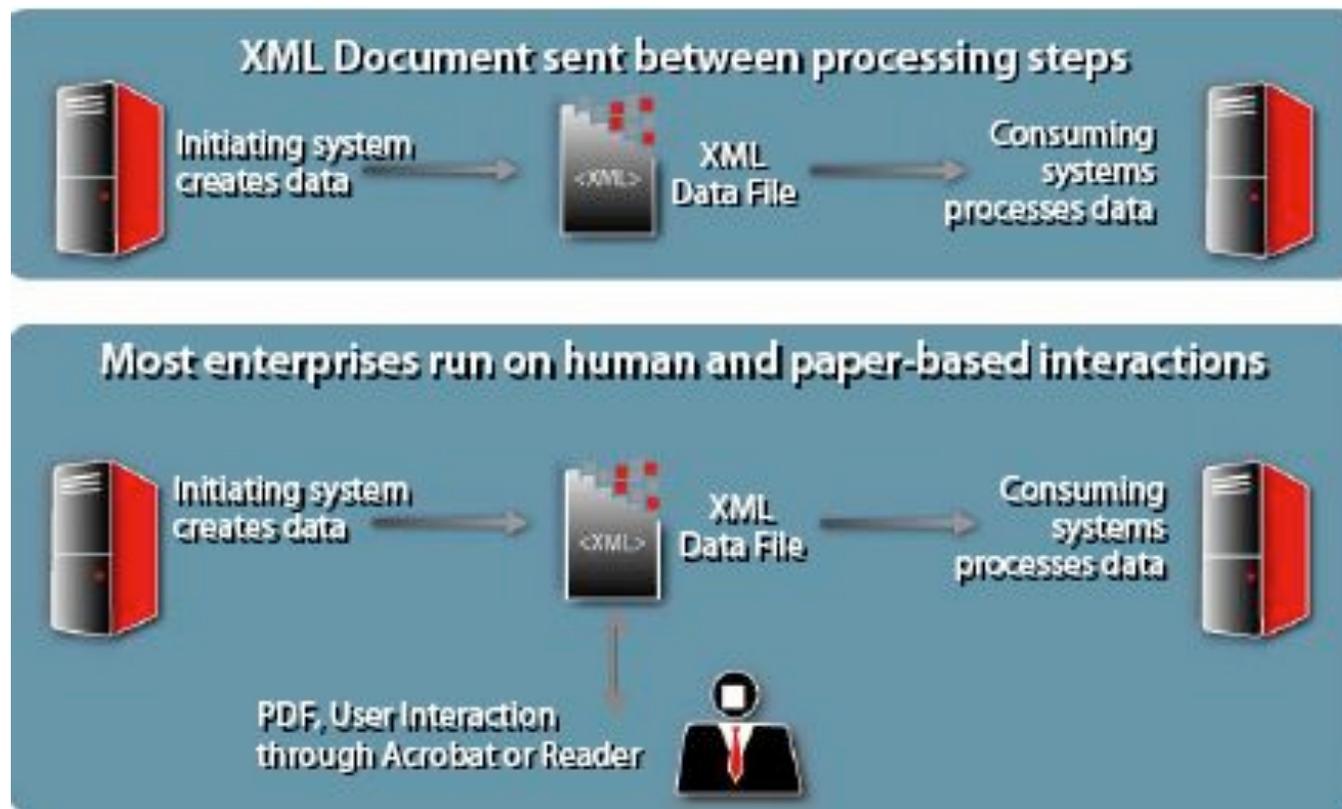
PDF per publishing tradizionale

- Trasparenza
- Esistenza di formato apposito: PDF-X
- Editabilità (limitata) della bozza
- Modello di colore comune a tutte le applicazioni Adobe

Workflow basati su PDF



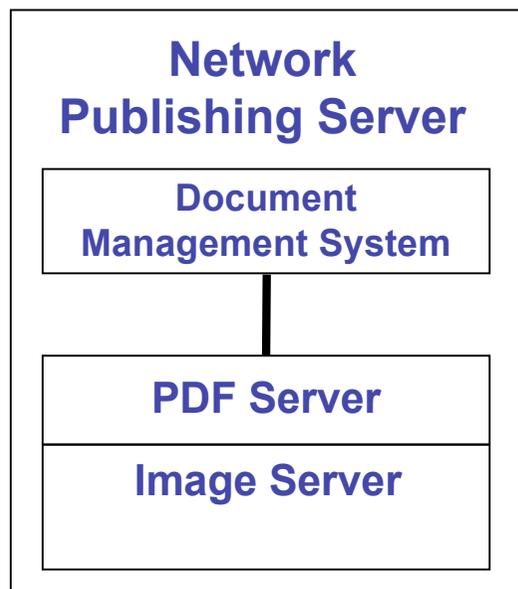
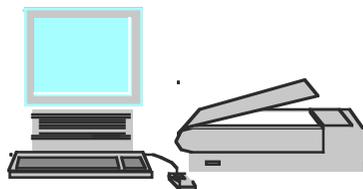
Workflows con XML e PDF



Publishing

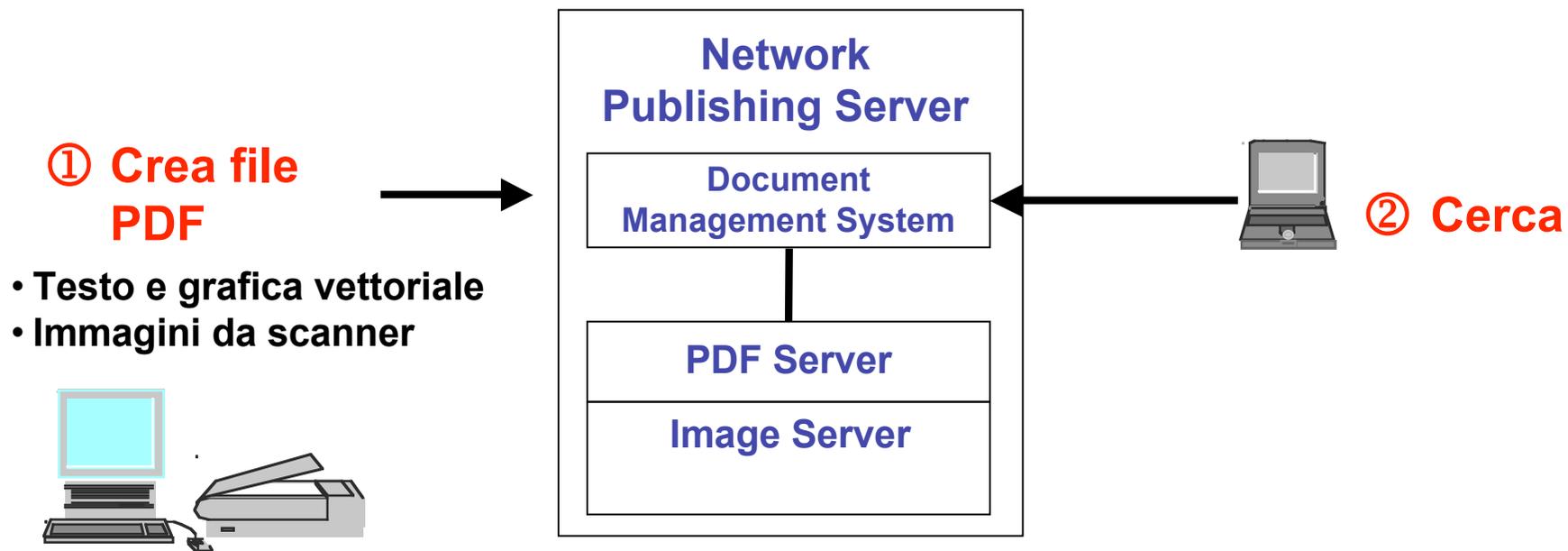
① Crea file PDF

- Testo e grafica vettoriale
- Immagini da scanner



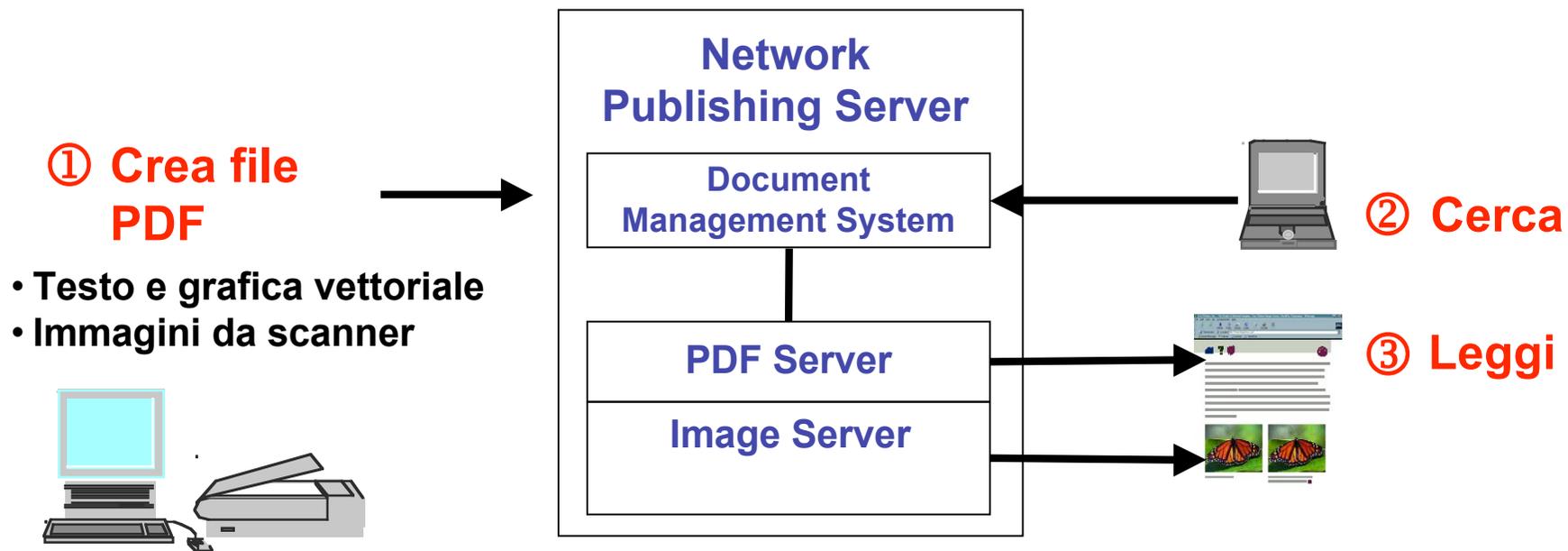
- Immagini salvate alla risoluzione massima
- La creazione di documenti è normale

Publishing



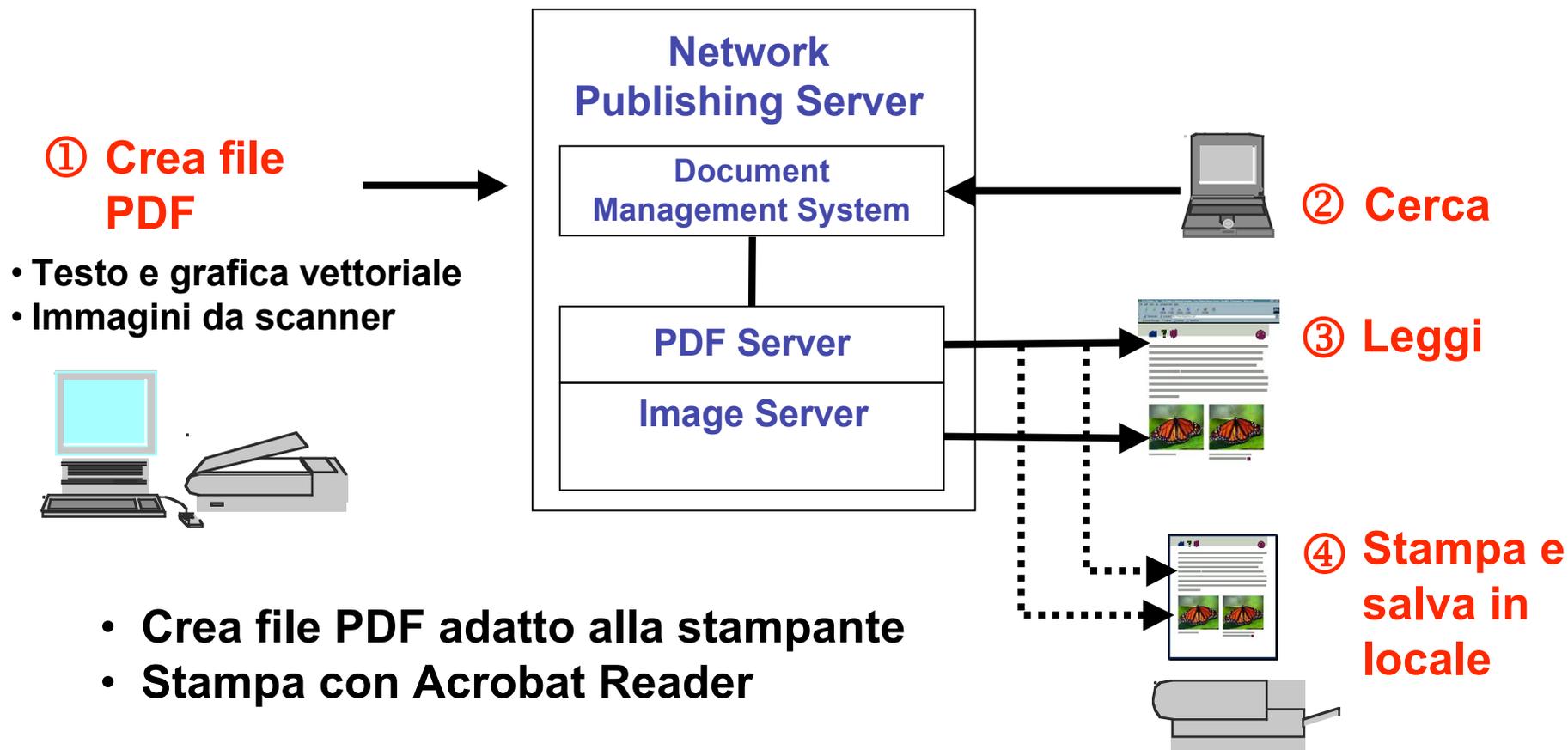
- Document management system opzionale
- Ricerca di documenti per lettura, stampa o archiviazione

Publishing

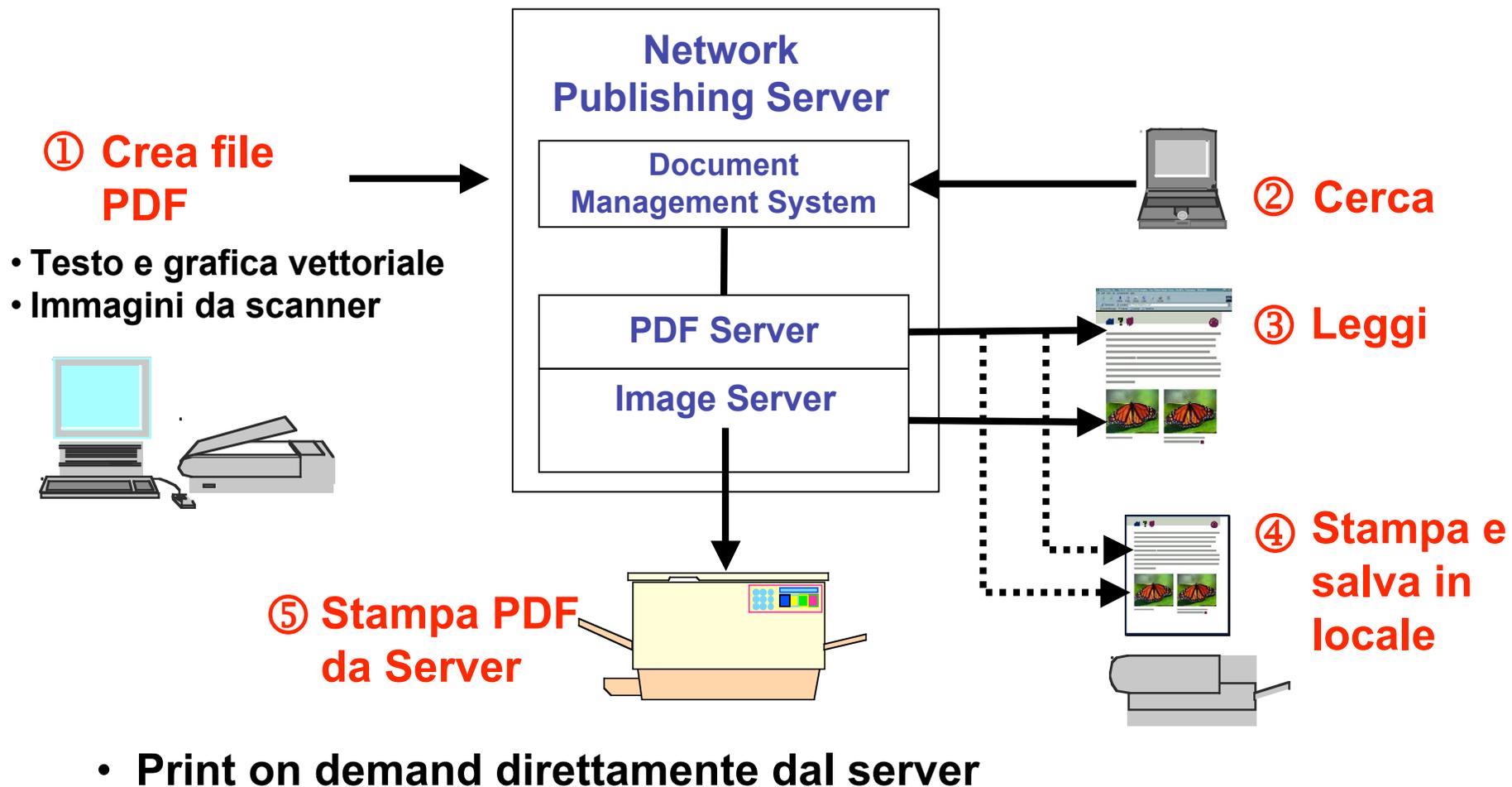


- Leggi documenti PDF e ingrandisci immagini senza plugin
- Correzione on line
- Funziona con tutti i browser

Publishing



Publishing



Sommario

- La creazione di documenti digitali si chiama **authoring**
- I documenti digitali vanno **trasformati** per i diversi canali di fruizione
- La fase di **delivery** si occupa di trasmettere il documento nel formato di fruizione prescelto
- Il **formato PDF** è uno standard di fruizione molto diffuso e ricco di servizi

Riferimenti

- J. Warnock, *The Camelot Project*, manuscript, 1990
- T.Merz, *Postscript & Acrobat/PDF*, Springer, 1996
- T.Merz, *Web Publishing with Acrobat/PDF*, Springer, 1998
- Adobe, *PDF Reference*, version 1.7, 2006 www.adobe.com

- Adobe, *PDF as a Standard for Archiving*, White Paper 2003
- Comuserve, *GIF version 89a*, 1990
- Adobe Developers Association, *TIFF 6.0*, 1992

Siti

Siti sulla conservazione di documenti digitali:

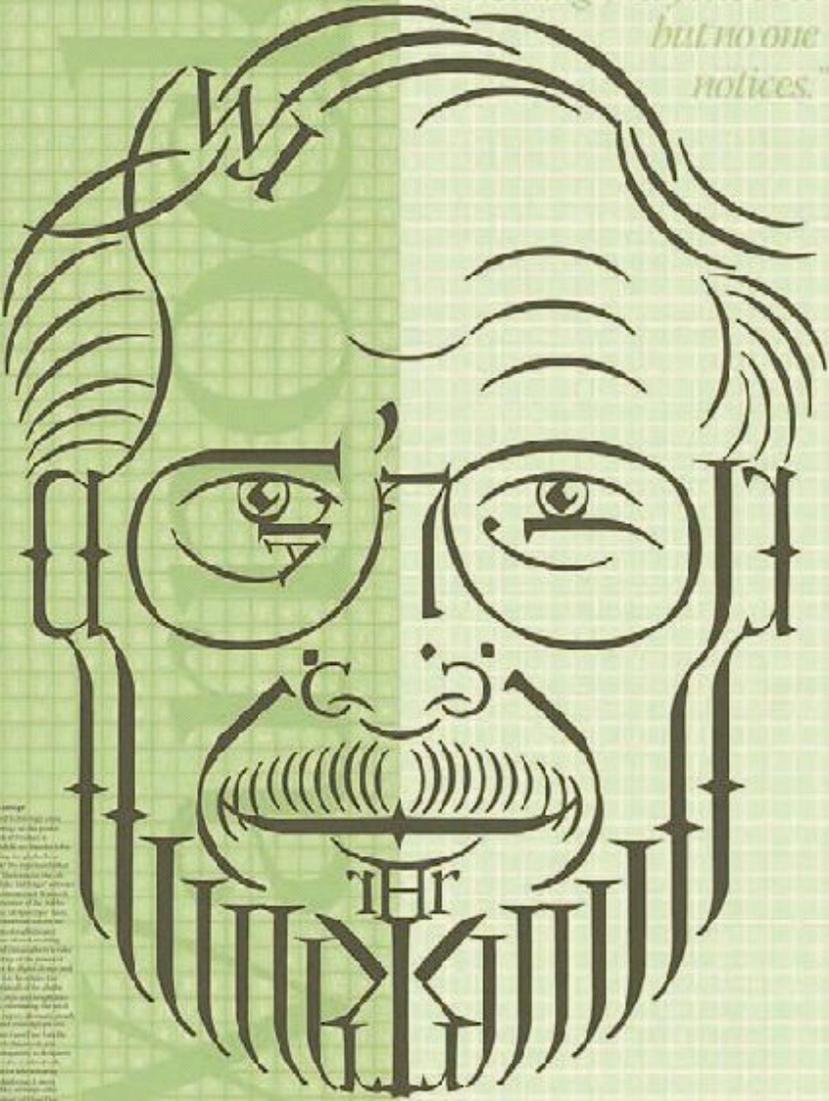
- www.library.cornell.edu/preservation/tutorial/contents.html
- www.digitalpreservation.gov/formats/

Siti sul PDF:

- www.acrobat-services.com
Uso di PDF in pubblicità
- www.math.uakron.edu/~dpstory/acrotex.html
Un sito sui rapporti TeX/PDF
- www.prepressure.com
Sull'uso di PDF nella catena editoriale

Domande?

"Good typography is something everyone sees but no one notices."



A better message
is often needed to
communicate the
value of a product
or service. In the
case of John Warnock
Pro, the message
is that the font is
the most beautiful
and versatile
font ever created.
It is a font that
can be used for
any purpose, in
any medium, and
it is a font that
will last for
many years to
come. It is a
font that is
designed to be
used by everyone,
from the most
casual user to the
most professional
designer. It is a
font that is
designed to be
used in any
medium, from
print to digital,
and it is a font
that is designed
to be used in
any language.
It is a font that
is designed to be
used by everyone,
from the most
casual user to the
most professional
designer. It is a
font that is
designed to be
used in any
medium, from
print to digital,
and it is a font
that is designed
to be used in
any language.

John Warnock

Font Bundles Ltd. 2011

Mucca Design created this poster entirely from characters of the font Warnock Pro.