A page from a medieval manuscript, likely a Bible or a book of wisdom, featuring Latin text in a Gothic script. The page is decorated with a large, ornate initial 'P' on the left side, a small illustration of a bird perched on a branch in the upper right, and another decorative element at the bottom right. The text is arranged in two columns, with a large initial 'P' at the start of a section. The background is a light, aged paper color.

La creazione di documenti digitali

Paolo Ciancarini

Dipartimento di Scienze dell'Informazione

Università di Bologna

Obiettivi della lezione

- Introduzione ai documenti digitali
- I documenti digitali nei nuovi canali di comunicazione
- Il ciclo di vita dei documenti digitali
- Standard tecnologici per affermare la proprietà intellettuale

Rivoluzioni culturali ?

- La scrittura su pietra e su argilla
- I papiri e la carta
- La stampa
- L'informazione digitale e il software

Libri

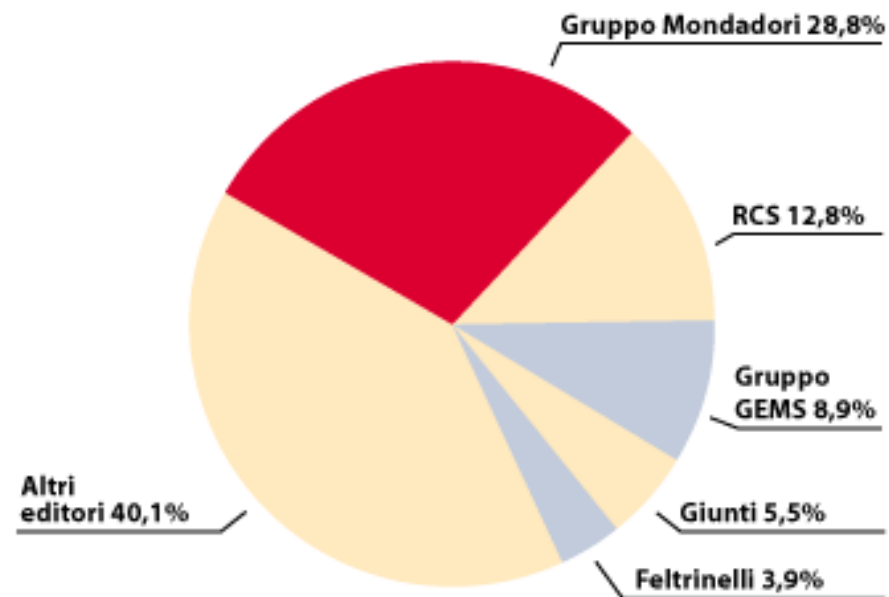
anno	1450	1950	2000
Libri pubblicati per anno	100	250.000	1.000.000
Totale cumulativo	100	16.000.000	36.000.000

Se leggessi un libro al giorno dai 5 agli 85 anni, alla fine avrei letto circa 30.000 libri

Il mercato italiano dei libri

2008 Libri: quote di mercato a valore

Fonte: Nielsen Bookscan



Fatturato 2008
434,3 milioni di euro

Produzione 2008
2.695 novità
5.225 ristampe

Totale copie prodotte
53,4 milioni

First published in hardback in Great Britain in 2007
by Atlantic Books, an imprint of Grove Atlantic, Inc.

Copyright © Daniel Johnson 2007

The moral right of Daniel Johnson to be identified as the
author of this work has been asserted in accordance with
the Copyright, Designs and Patents Act of 1988.

All rights reserved. No part of this publication may be
reproduced, stored in a retrieval system, or transmitted in
any form or by any means, electronic, mechanical, photo-
copying, recording, or otherwise, without the prior
permission of both the copyright owner and the above
publisher of this book.

Every effort has been made to trace or contact all copyright-
holders of material used in this book. The publishers will be
pleased to make good any omissions or rectify any mistakes
brought to their attention at the earliest opportunity.

1 3 5 7 9 8 6 4 2

A CIP catalogue record for this book is
available from the British Library.

ISBN 978 1 84354 609 2

Copyright
sui libri

Copyright su siti web

Credits: Art

General Credit Notice About Photos: Unless otherwise noted, all photographs of Thomas E. Kennedy are by Alice Maud Guldbrandsen.

All other photos at this website are by Thomas E. Kennedy, unless otherwise noted.

Photographs have been cropped and/or resized by Clare MacQueen, with permission, from materials provided by Kennedy.

Art Featured on Book Covers

The cover of *A Passion in the Desert* is by Andi Olsen.

The cover of *Bluett's Blue Hours* is a manipulation by Roger Derham of a photograph of Kennedy by Alice Maud Guldbrandsen and a photograph by David Applefield of the neon sign on a wall over Copenhagen's Sortedam Lake. The sign advertises the fresh eggs of Irma's supermarket — something of an iconic Copenhagen symbol since the 1950s. The image is described in the novel and its implications riffed on.

The cover of *Cast Upon the Day* is by William Skotte Olsen (used by kind permission of Henrik Kampmann and Galleri Kampmann). The oil-on-canvas painting is from Kennedy's collection of original art.

The cover of *Danish Fall* is a collage of a photo of Kennedy by Alice Maud Guldbrandsen and a photo by Kennedy of "Pussy Galore's Flying Circus," a bar on St. Hans Square in north Copenhagen, where several scenes from the novel are set. The collage is by Roger Derham, publisher of *Danish Fall*.

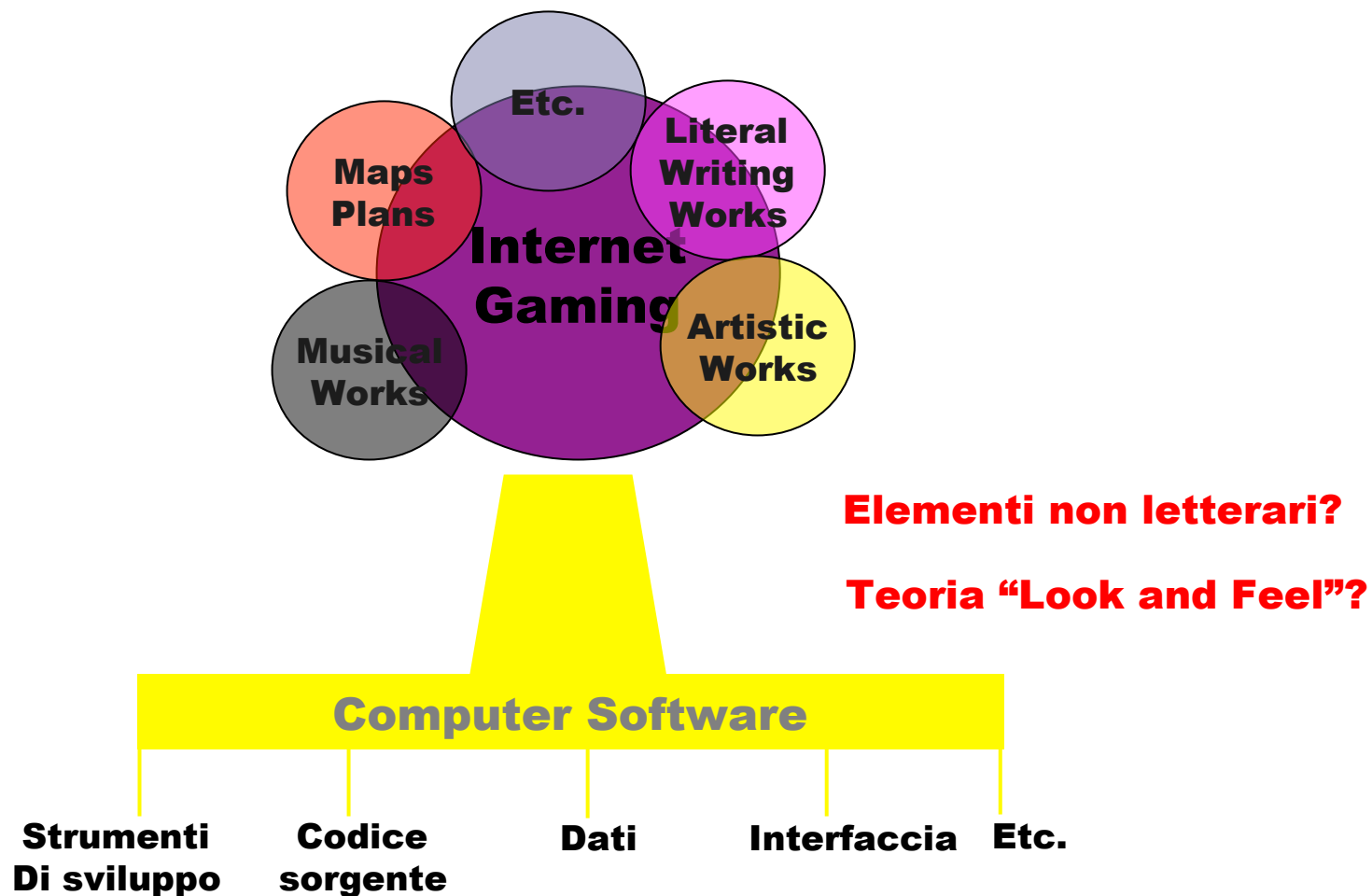
The cover of *Greene's Summer* includes a photo of Café Bopa on BOPA Place in east Copenhagen. Photo originally by Thomas E. Kennedy and manipulated by Roger Derham.

The cover of *Kerrigan's Copenhagen* is a manipulation by Roger Derham of a photo by Thomas E. Kennedy of the sculpture, "A Drunken Faun," by Andreas Kolbert, 1857. Sculpture stands on the bank of St. John's Lake in west Copenhagen.

The cover of *Poems and Sources* features "Island Cities," painted by Hysse Forchhammer in 1999 especially for the anthology, and inspired by the John Updike sonnet in the issue. Original painting now hangs in Kennedy's collection.

The cover of *Stories and Sources* features "Carousel," an oil painting by Gladys Swan.

Il software per giocare su Internet

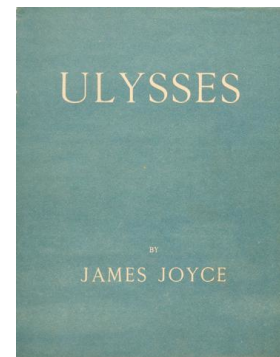




Quali tra queste opere
sono sotto copyright?



1571



1922



1928

Opere “orfane”



<http://orphanworks.ssl.co.uk/>

Opere orfane

- Opere sotto copyright per cui non si conosce il detentore dei diritti, spesso come risultato di modifiche legislative
 - Nel 1989 venne rimossa la condizione le opere dovessero contenere una nota di copyright
 - Nel 1992 venne cancellato il requisito del rinnovo del copyright
 - Nel 1994 molte opere straniere uscirono dal public domain
- Si possono usare solo dopo una ricerca “ragionevolmente diligente”
- Report 2006 al Congresso USA
- Sistema di determinazione www.copyright.gov/orphan/

Google Books

- Google Books permette di accedere
 - Opere in commercio
 - Opere fuori commercio
- Acquisisce le opere in commercio dagli editori che danno il permesso
- Digitalizza (con OCR) oltre un milione di libri l'anno (vedi prodotti di scanning automatico come Bookdrive)

Wikipedia

7 milioni di articoli

Oltre 200 lingue

WIKIPEDIA

English

The Free Encyclopedia
2 561 000+ articles

Español

La enciclopedia libre
400 000+ artículos

Français

L'encyclopédie libre
707 000+ articles

Polski

Wolna encyklopedia
538 000+ haseł

Deutsch

Die freie Enzyklopädie
806 000+ Artikel

日本語

フリー百科事典
522 000+ 記事

Português

A enciclopédia livre
428 000+ artigos

Русский

Свободная энциклопедия
317 000+ статей

中文

自由的百科全書
204 000+ 條目

Italiano

L'enciclopedia libera
497 000+ voci



Wikipedia e la proprietà intellettuale

Roma 2.7.2007- Secondo quanto deciso da amministratori ed utenti di [Wikimedia Commons](#) e della [Wikipedia italiana](#) le fotografie di opere in territorio italiano di architetti morti dopo il 1937 vanno **eliminate dalla nota enciclopedia** a causa del diritto d'autore.

L'intera architettura contemporanea e moderna italiana, perciò, rischia di non poter essere raffigurata nella più diffusa enciclopedia del mondo.

Questo perché la legislazione italiana, a differenza di molti altri paesi, non contemplerebbe il cosiddetto [panorama freedom](#) (libertà di panorama), che permette a chiunque di fotografare e riprodurre quanto pubblicamente visibile senza preoccuparsi di dover trovare il progettista e pagargli i diritti d'autore.

Qual è il problema?

- Tutte le opere dell'ingegno sono rappresentabili in almeno due forme: su supporto fisico (analogico) e in forma di documento digitale
- Cos'è un documento digitale e in cosa differisce da un documento analogico?

Documento

Documento:

Contenitore (o *supporto*) di dati
(testo, numeri, figure, ecc.) strutturati
per essere usati come informazioni

Un documento



**Sao ko kelle terre, per kelle fini que ki contene,
trenta anni le possette parte sancti Benedicti...**

*Placito Capuano: atto giudiziario custodito nel
monastero di Montecassino, primo scritto ufficiale in
lingua volgare (Capua, marzo 960)*

Un documento

Cognome	Zucchetto
Nome	Zucchini
nato il	20/06/1980
(atto n. 1234 p. 5 s. U)	
a	Verdulandia (VE)
Cittadinanza	Verduresca
Residenza	Riccione
Via	delle Melanzane n°8
Stato civile	celibe
Professione	Clown
CONNOTATI E CONTRASSEGNI SALIENTI	
Statura	1,74
Capelli	castani
Occhi	belli
Segni particolari	il pizzetto



Firma del titolare *Zucchetto Zucchini*
Riccione n. 30/02/2004

Il Sindaco
[Signature]
Ufficiale d'Anagrafe



Un documento

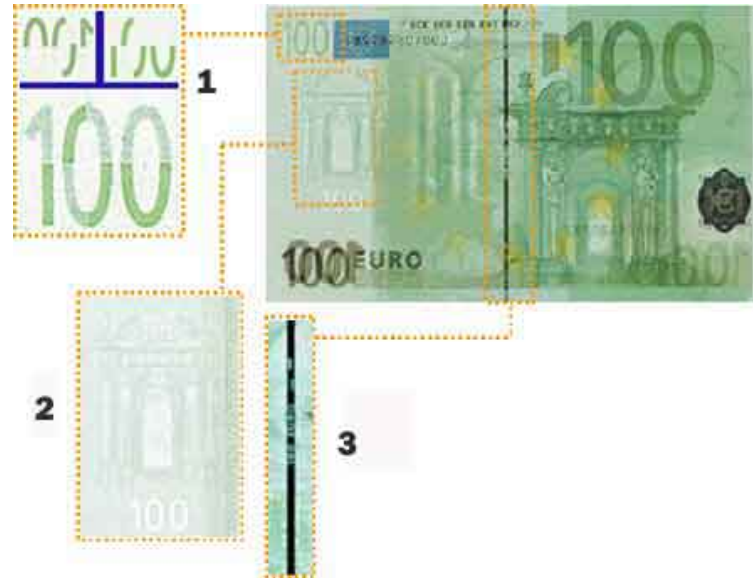


Watermarking

- La tecnologia più comune di autenticazione
- Inserisce informazione nel documento originale



Proprietà del watermarking



- L'informazione non è rimovibile senza distruggere il documento
- L'informazione è costantemente verificabile

Documento digitale

Documento digitale:

Documento rappresentato in forma **binaria** e **memorizzato** in un sistema informatico

Ha un **nome**, un **contenuto**, una **struttura**, alcuni **comportamenti**, alcune **relazioni** con altri documenti

Il contenuto è di solito **codificato** (anche più volte, da codici sovrapposti)

Gli si associano alcune informazioni dette **metadati**

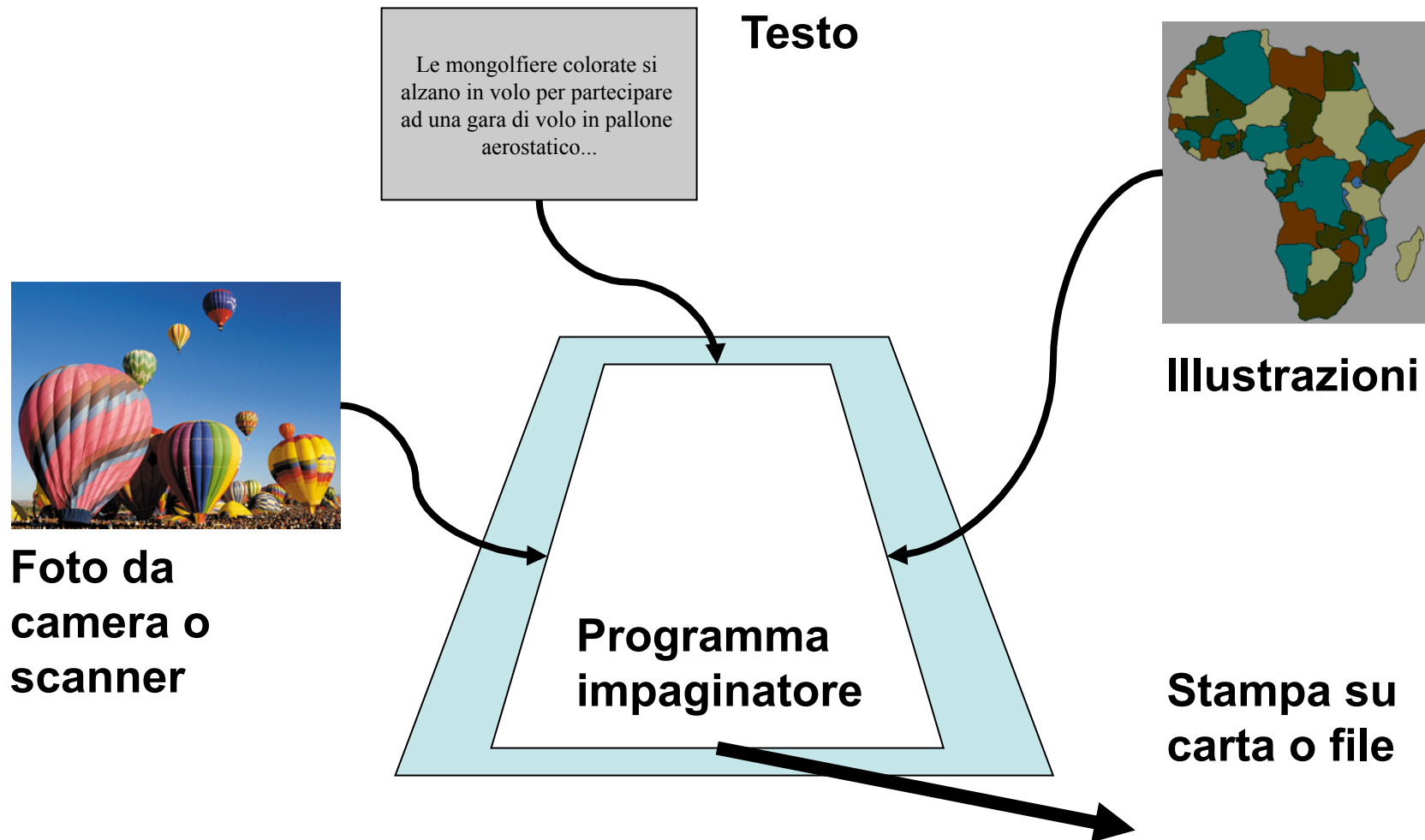
Creazione e Proprietà

- L'autore di un documento è il proprietario dei diritti d'autore
- Il documento può essere ceduto, dunque il proprietario – non l'autore - cambia
- Autore e proprietario di un documento digitale sono di solito indicati da appositi e distinti *metadati*

Esempio

- Sono al lavoro, in azienda
- Uso il computer per mandare un email
- Chi è l'autore dell'email?
- Chi è il proprietario dell'email?

I documenti digitali sono oggetti "costruibili"

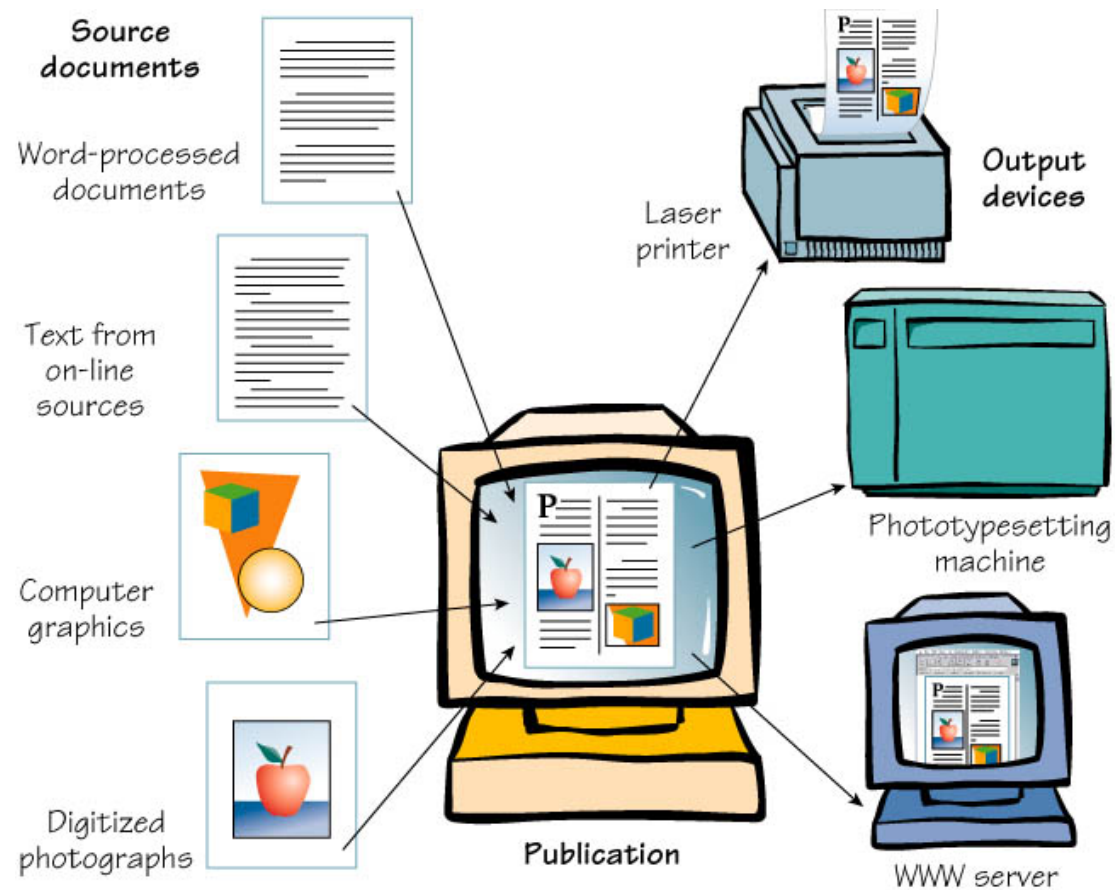


Risoluzione

- I documenti che rappresentano immagini hanno una qualità che si chiama **risoluzione** (numero di pixel per unità di superficie)



I documenti digitali sono distribuibili su più canali



iPad

- Disponibile da 2010
- Schermo 9.7" touch retroilluminato
- Fino a 64GB
- Connessioni WIFI, 3G, Bluetooth



Il futuro dell'editoria

“Il futuro è già qui,
distribuito in modo diseguale”

William Gibson, scrittore

Canali di comunicazione digitale

- Siti Web e Portali
- Really Simple Syndacation (RSS)
- Podcast e Internet Radio
- Motori di ricerca
- Blog, Wiki, Forum e chat
- Servizi di instant messaging
- Newsfeeds (es. USENET)
- Massive Multiplayer Games (es. Second Life)
- Reti peer 2 peer (es. Gnutella)

La filiera digitale

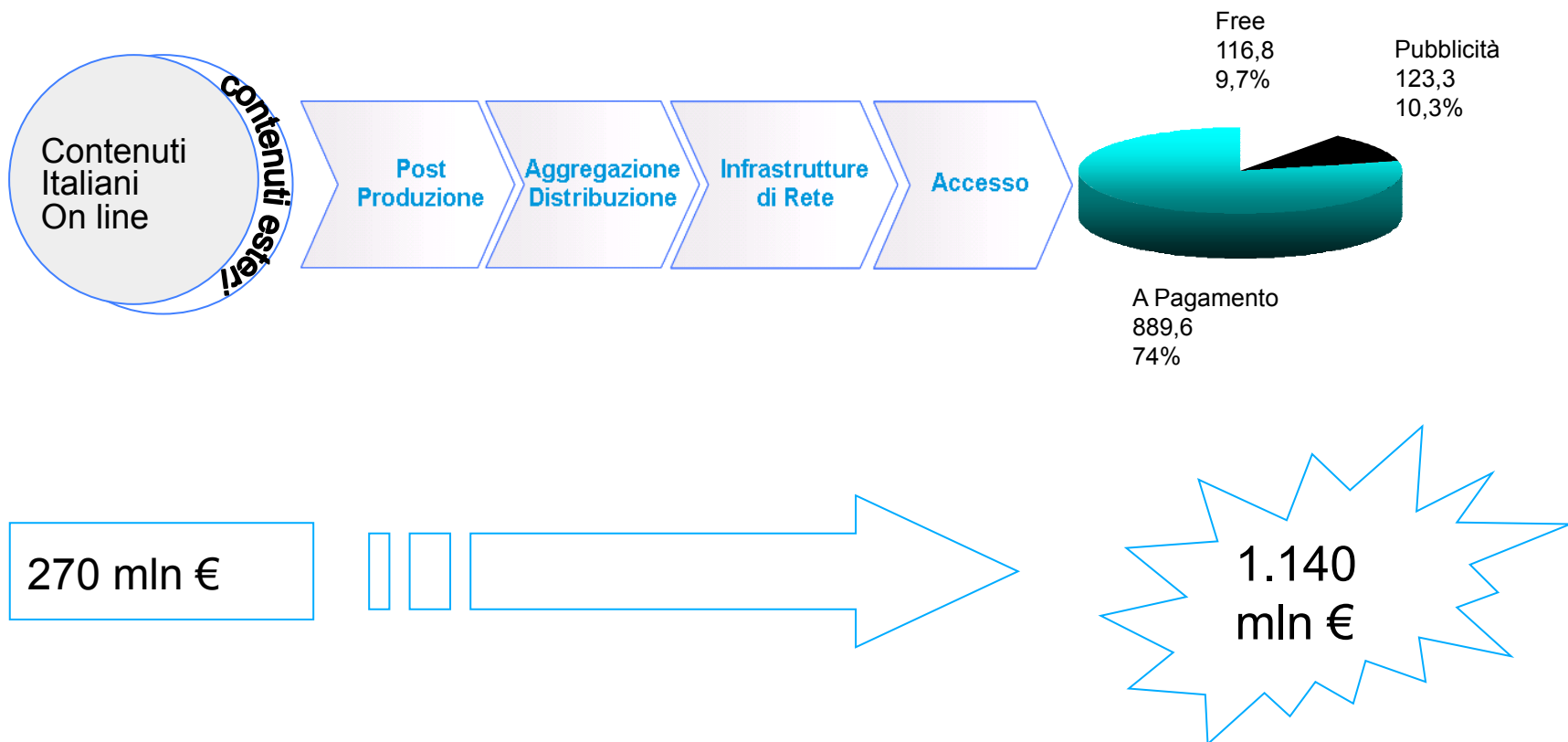
La filiera dell'eContent è costituita da attori specializzati in aree molto diverse, accumulati dalla missione di realizzare e rendere fruibili contenuti creati o trasformati mediante le tecnologie ICT sui canali digitali di distribuzione

La filiera digitale



La filiera aggiunge valore

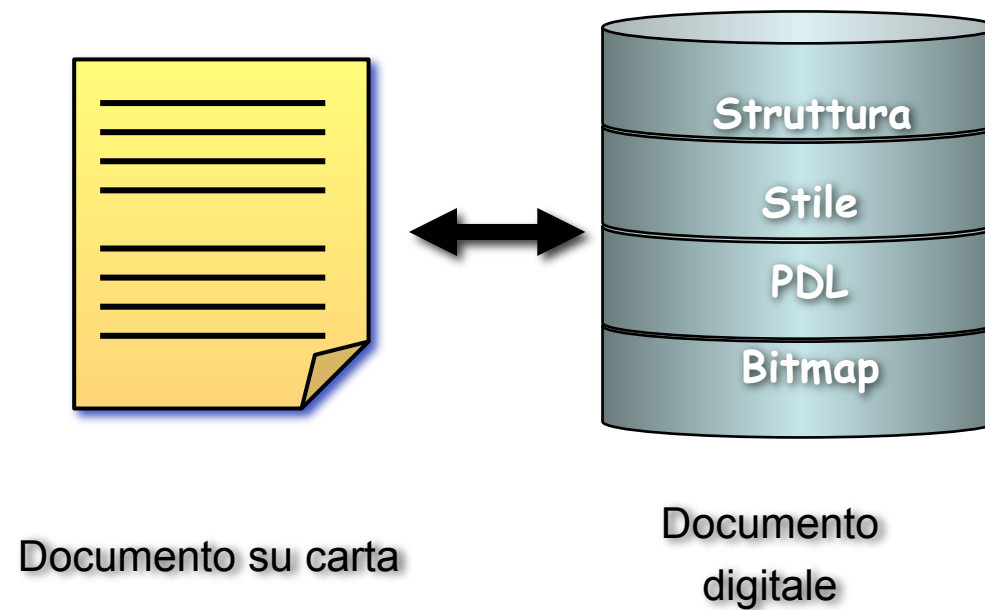
La filiera rende fruibili attraverso diversi canali contenuti disponibili in rete
Nel ciclo di trasformazione il valore dei contenuti si accresce circa 5 volte



User generated content

- ✓ Il "contenuto generato dagli utenti" (*User generated content*) è qualsiasi documento creato e messo on line direttamente dal suo creatore – testo, grafica, audio, video, tag, categorie o rank
- ✓ 35% degli utenti adulti di internet hanno messo online dei contenuti
- ✓ 26% condividono qualcosa online che hanno creato loro stessi
 - 8% hanno un blog
 - 14% gestiscono la propria homepage
 - 13% creano o gestiscono pagine per altre persone

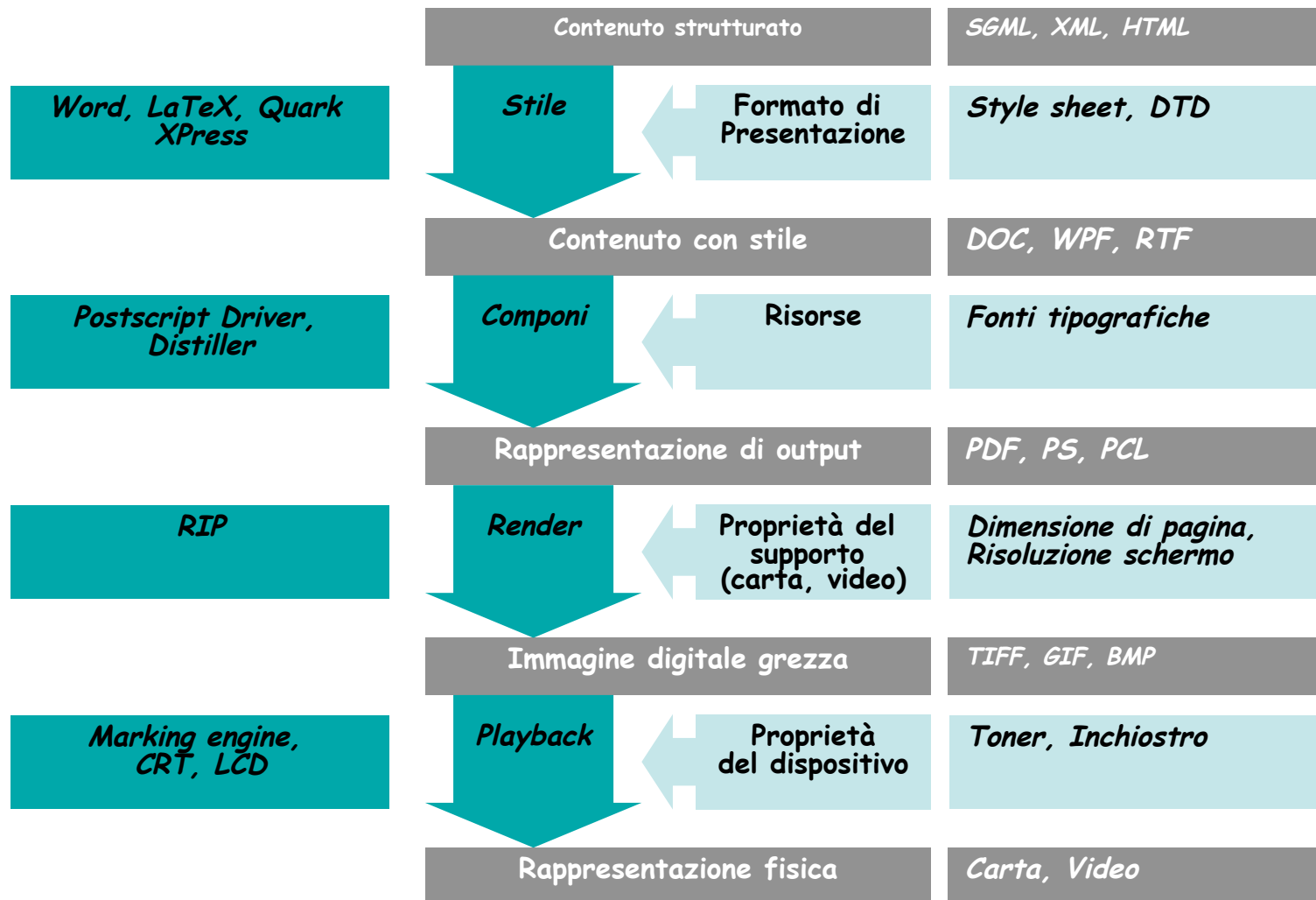
Documenti di carta vs documenti digitali



Dalla carta al digitale



Dal digitale alla carta



Esempi

- Documenti Word
- Presentazioni Powerpoint
- Archivi PDF
- E-book
- File HTML
- Sito Web
- ...

I documenti nelle organizzazioni

Strategie d'impiego dei documenti

- Strategie di acquisizione
- Strategie di memorizzazione e trasformazione dei contenuti
- Strategie di erogazione e trasmissione
- Architetture informative indipendenti dall'hw

The screenshot displays the official website of the University of Bologna (UNIVERSITÀ DI BOLOGNA). The header includes the university's name, the founding year 'A.D.1088', and the 'Multicampus' locations: Bologna, Cesena, Forlì, Ravenna, and Rimini. The date 'venerdì 9 ottobre 2009' is shown in the top right corner. A navigation bar contains links for 'Download', 'Supporto', 'Rubrica Persone', 'Rubrica Strutture', 'La mia e-mail', 'UniboSearch', and 'UniboMappe'. A search bar is present with the text 'cerca'. Below the header, there are several sections: 'Home' with a menu for 'Ateneo', 'Studenti', 'Offerta formativa', 'Ricerca', 'Relazioni Internazionali', 'Biblioteche e musei', and 'Personale'; 'SEMPRE UTILI' with links to 'AlmaZeta', 'Il Mio Portale', 'La mia e-mail', 'Cerca Docenti', 'Guida dello Studente', 'Autocertificazione online', 'Merchandising', and 'Biblioteche online'; 'SERVIZI ONLINE' with dropdown menus for 'Studenti', 'Docenti e ricercatori', and 'Tecnici-amministrativi'; 'STRUTTURE' with a dropdown menu; 'IN PRIMO PIANO' featuring articles like 'L'Unibo ti guida', 'Parti con il Portale giusto', and 'Piani di studio online'; 'MULTICAMPUS' with a map and 'Consigli'; 'Avvisi' with news about international seminars and Erasmus placements; 'Virus A/H1N1' with information on flu responses; 'Interruzione di alcuni servizi online dal 9 al 12 ottobre 2009'; 'Iscrizioni ai corsi di lingue straniere - C.I.L.T.A.'; and 'Bando Erasmus Placement 2009/2010'. On the right side, there is a 'Rivivi AlmaFest' section with a photo and a 'UNIBOMAGAZINE' section with a list of news items.

Applicazioni informatiche per documenti digitali

- Scrivania (desktop)
- Editori del testo o delle immagini
- Visualizzatori
- Catalogatori
- Compressori
- Codificatori e decodificatori
- Strumenti di protezione del contenuto
- Sistemi di gestione di biblioteche digitali
- Sistemi di condivisione

Ciclo di vita del documento

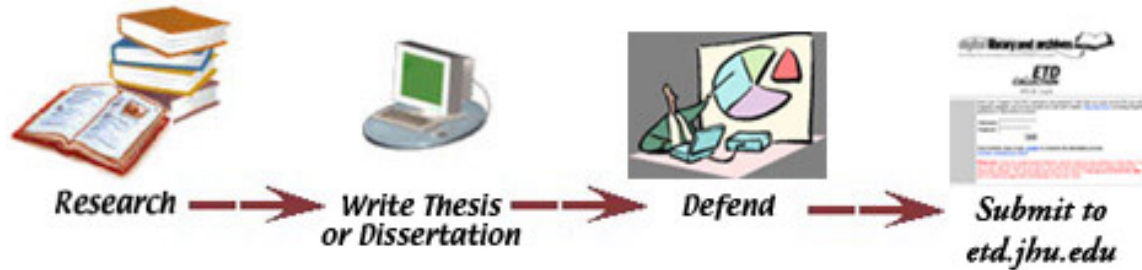
Descrizione organizzata delle fasi in cui viene

- Creato
- Manipolato
- Memorizzato
- Collegato
- Catalogato
- Trasmesso
- Distrutto

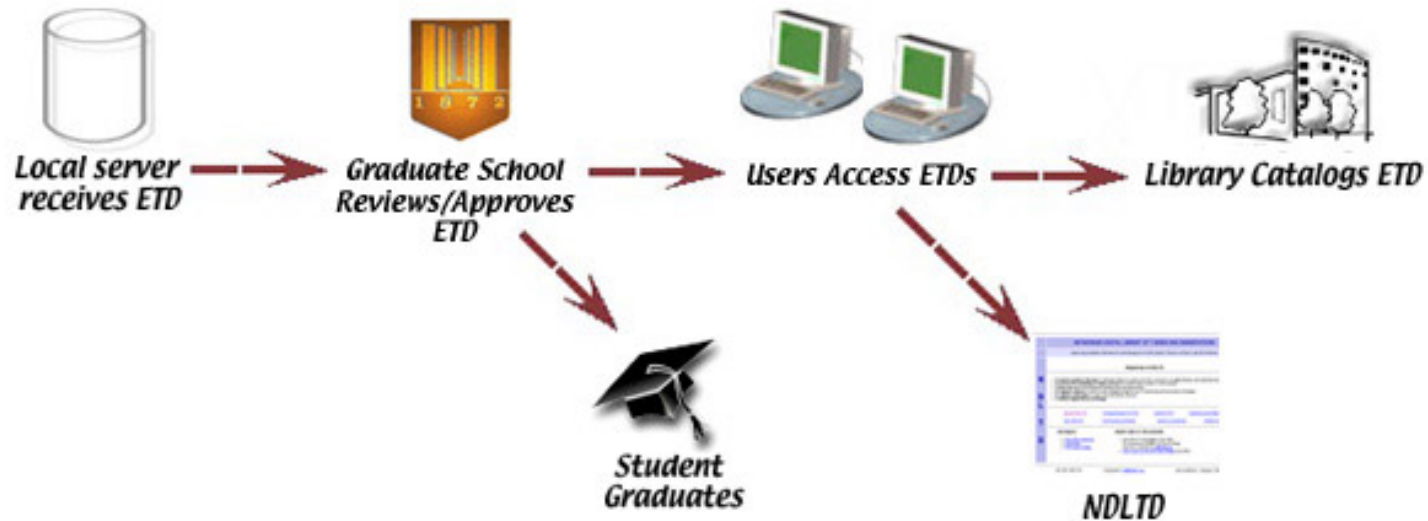
Esempio di ciclo di vita

Life Cycle of an ETD

Student thesis work



Steps after submitting an ETD



Nomi di documenti

- Nomi all'interno di un computer
- URL: nomi su Web
- URI: identificatori unici
- DOI: Digital Object Identifiers
- PUBMED ID

- Metadati

DOI e PUBMED ID

- DOI = Digital Object Identifier. Identificatore permanente di un documento digitale, assegnato a cura dell'editore (registrant). Usato per accedere i metadati completi
- PubMed ID – identificatore usato dal database PubMed database (noto a Google)

Cosa sono i DOI

I DOI sono uno standard di identificazione di un'opera digitale ai fini della proprietà intellettuale

Applicabile a qualsiasi formato di contenuto: testo, musica, film, video, foto, software, elemento di database, certificato ...

- Identificazione **univoca**, indipendente dal luogo di memorizzazione
- Identificazione **permanente**, indipendente dal proprietario
- Identificazione "**actionable**", cioè utilizzabile da servizi informatici

Sintassi DOI

Ogni DOI è una stringa alfanumerica

Ha due parti, chiamate prefisso e suffisso



Esempio d'uso DOI

THE GUIDE TO COMPUTING LITERATURE



Coordinating rule-based software processes with ESP

Full text  [Pdf](#) (1.71 MB)

Source **ACM Transactions on Software Engineering and Methodology (TOSEM)** [archive](#)
Volume 2 , Issue 3 (July 1993) [table of contents](#)
Pages: 203 - 227
Year of Publication: 1993
ISSN:1049-331X

Author [Paolo Ciancarini](#) Univ. of Bologna, Bologna, Italy

Publisher [ACM](#) New York, NY, USA

Bibliometrics Downloads (6 Weeks): 4, Downloads (12 Months): 43, Citation Count: 0

Additional Information: [abstract](#) [references](#) [index terms](#) [review](#) [collaborative colleagues](#) [peer to peer](#)

Tools and Actions:

[Review this Article](#)

[Save this Article to a Binder](#)

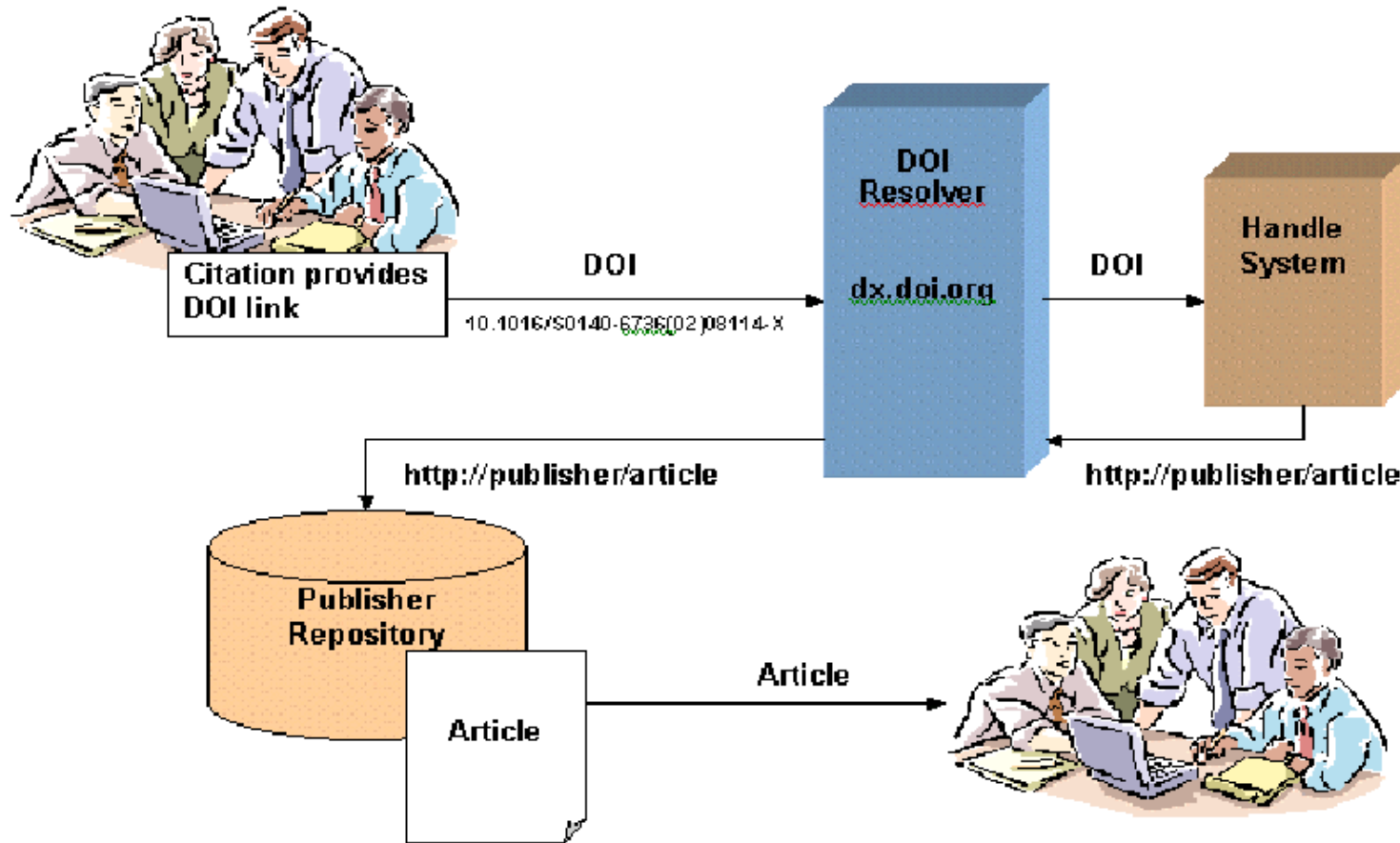
Display Formats: [BibTex](#) [EndNote](#) [ACM Ref](#)

DOI Bookmark:

Use this link to bookmark this Article: <http://doi.acm.org/10.1145/152388.152393>

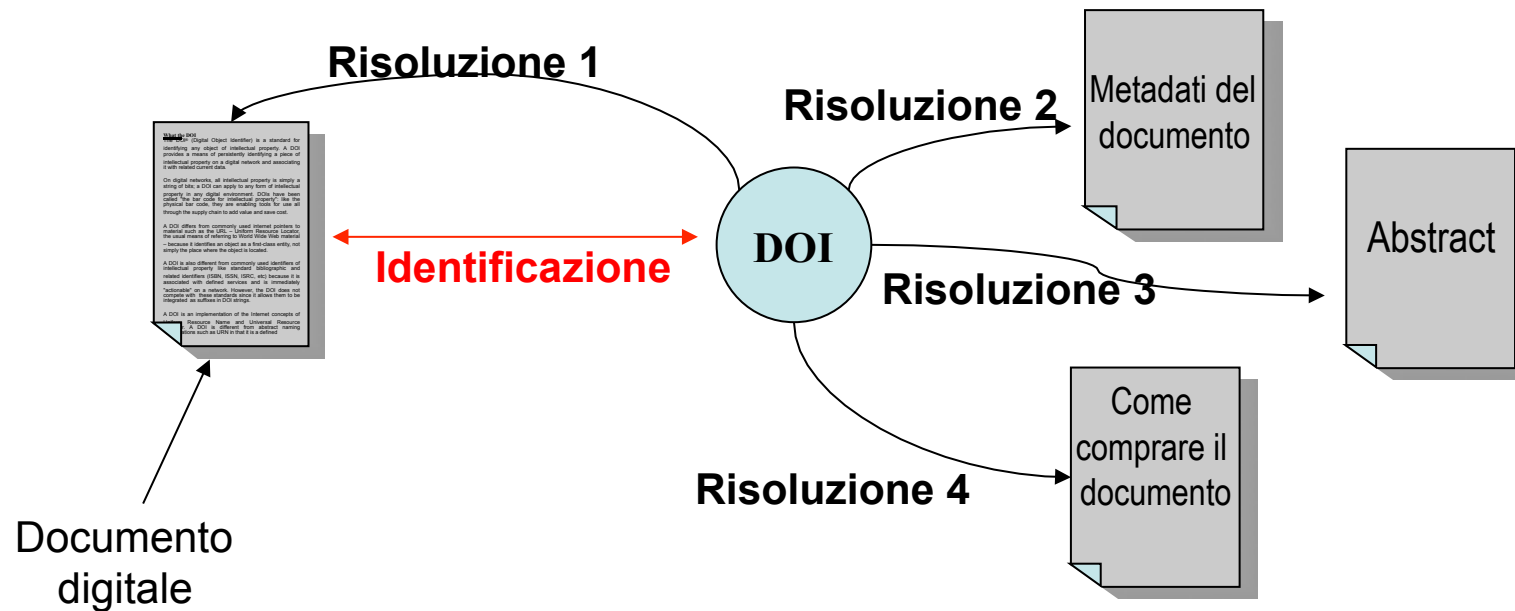
[What is a DOI?](#)

Risolvere un DOI



Identificare vs risolvere

“what the DOI identifies” e “what the DOI resolves to”
sono operazioni diverse



Benefici dei DOI

- persistenza: il DOI funziona anche se il documento cambia ubicazione
- cooperazione con altri dati da altre fonti
- estensibilità
- indipendenza dalla piattaforma
- aggiornamenti dinamici

I metadati

- I *metadati* sono un insieme di dati e informazioni che descrivono una risorsa o un documento digitale
- *Metadata is data about data. The term refers to any data used to aid the identification, description and location of networked electronic resources. Many different metadata formats exist, some quite simple in their description, others quite complex and rich.*
- I metadati si dividono in tre classi:
 - metadati descrittivi
 - metadati amministrativi
 - metadati strutturali

I metadati descrittivi

- Descrizione bibliografica di risorsa digitale
- Descrizione di risorse correlate: ad esempio dell'origine cartacea di un documento digitale
- Documentano gli attributi principali di proprietà intellettuale e di classificazione
- **Esempio di metadati in XML:**

```
<libro>  
  <titolo>Promessi Sposi</titolo>  
  <autore>Manzoni, Alessandro</autore>  
  <data>1840</data>  
  <ISBN>123-456-789</ISBN>  
  <editore>Mondadori</editore>  
</libro>
```

Esempio: metadati in pdf

The image shows a PDF viewer window displaying a document titled "EXTREME CHAOS". The document content includes a title, a quote by Paul Marz, a main text block, a pie chart titled "Project Resolution (2000)", and several columns of text. The "Document Properties" panel on the right shows metadata for the file "extreme_chaos.pdf", including title, author, creation/modification dates, application, PDF producer, version, location, file size, page size, and tagging status.

EXTREME CHAOS

"The world acquires value only through its extremes and endures only through moderation; extremists make the world great, the moderates give it stability."
—Paul Marz, French poet, critic

A project gold rush is in full swing. In 1998 Corporate America issued 200,000 new-start application software development projects. By the end of the Y2K "non-event," new starts went up by almost 100,000 projects to 300,000 this year and more than half a million will be resolved in the next 12 months. Most of these new projects are well within The Standish Group's criteria established in "CHAOS: A Recipe for Success" (1999), which limits the size of the projects to six months and six people. This newest paper is a continuation of our CHAOS research updated with our most recent findings.

The timeline of The Standish Group's seven years of CHAOS research shows decided improvement in IT project management. This year project success rates, while modest, are up again across the board, while cost and time overruns are uniformly down. Time overruns have significantly decreased from 222% over the original time estimates in 1994 down to 63% in this latest study.

Cost overruns have gone from 189% over the original cost estimates in 1994 down to 45% in the 2000 study. In 1994 required features comprised 61% of the final product. This year's research shows 67% of the required features and functions. This notably increases end-user satisfaction in terms of time, cost and features.

The best news is that we as researchers and you our

readers are learning how to succeed more often. According to 1994 results, only 28,000 application development projects met the criteria for success — completed on time, on budget and with all the features and functions originally specified. This year's results show that 78,000 U.S. projects were successful.

The reasons for the increase in successful projects vary. First, the average cost of a project has been cut more than half. Better tools have been created to monitor and control progress and better skilled project managers with better management processes are being used. The fact that there are processes is significant in itself.

However, Nirvana is still a long way off — 137,000 projects were late and/or over budget, while another 65,000 failed outright. The reason most of these projects failed was not for lack of money or technology, most failed for lack of skilled project management and executive support.

Lack of executive support has replaced user involvement as the number one cause of project failure. Without a staunch project champion with a solid business vision, projects can drift into a technological or political abyss. Project stakeholders must create business value by improving customer service, communicating a clear business plan and delivering a competitive advantage.

Project Resolution (2000)

Category	Percentage
Failed	23%
Succeeded	28%
Challenged	49%

1 of 12

Document Properties

Description Security Fonts Initial View

Description

File: extreme_chaos.pdf

Title: Extreme CHAOS 2001

Author: The Standish Group

Subject:

Keywords:

Created: 08/03/01 15:36

Modified: 16/03/01 11:53

Application: PageMaker 6.5

Advanced

PDF Producer: Acrobat Distiller 3.01 for Power Macintosh

PDF Version: 1.2 (Acrobat 3.x)

Location: Macintosh HD:Users:paolo:Desktop:

File Size: 224,35 KB (229.737 Bytes)

Page Size: 8,50 x 11,00 in

Tagged PDF: No

I metadati amministrativi

- Metadati tecnici: documentazione degli aspetti tecnici relativi ai formati di digitalizzazione
- Metadati di preservazione: documentazione storica delle procedure e delle tecnologie adottate nella digitalizzazione e nella manutenzione del documento digitale
- Metadati di gestione dei diritti: documentazione relativa alla compensazione della proprietà intellettuale, ai diritti di accesso e alle loro eventuali variazioni

I metadati strutturali

- Descrizione della struttura logica del contenuto di una risorsa digitale
- Descrizione della struttura fisica di una risorsa digitale (elenco dei file che la compongono, diverse versioni e manifestazioni, ecc.)
- Descrizioni dei rapporti tra struttura logica e struttura fisica

Dublin Core 1.0

- 15 elementi per metadati formano il nucleo
 - Raffinabili mediante qualificatori
 - Aumentabili con altri elementi per scopi locali
- Risorsa
 - “qualcosa che ha un identità”
 - Analoga a “entità” (in un sistema bibliografico)
- Non impone un vocabolario o file di autorità
 - Quindi persone diverse potrebbero generare descrizioni molto diverse della stessa risorsa

Esempio: metadati con DC

<http://www.lub.lu.se/cgi-bin/nmdc.pl>

- DC.Title Giocatori Artificiali
- DC.Title.Alternative Il computer gioca a Scacchi
- DC.Creator Ciancarini, Paolo
- DC.Creator.Address ciancarini@cs.unibo.it
- DC.Subject Scacchi
- DC.Subject Computer
- DC.Subject Science
- DC.Description Chess
- DC.Publisher Mursia
- DC.Contributor nessuno
- DC.Date ISO86011992DC.TypeText
- DC.FormatIMT application/msword
- DC.Identifier <http://www.cs.unibo.it>
- DC.Language ISO639-1it
- DC.Coverage 1800-1992
- DC.Rights Diritti normali
- DC.Date.X-Metadata LastModifiedISO86012005-05-09

Metadati per immagini e multimedia

- I metadati servono anche per i documenti multimediali
- Molti formati immagine non permettono di memorizzarli assieme al contenuto
- Alcuni formati immagine includono metadati sulla risoluzione

[REMBRANDT van Rijn](#)

[\(Dutch, 1606-1669\)](#)

Old Man in Velvet Cloak

Date unknown

Reproductive etching

Date unknown

6 7/8 x 5 3/4 in. (17.5 x 14.7
cm) plate; 18 1/2 x 13 in.

(47...

1985.2.264

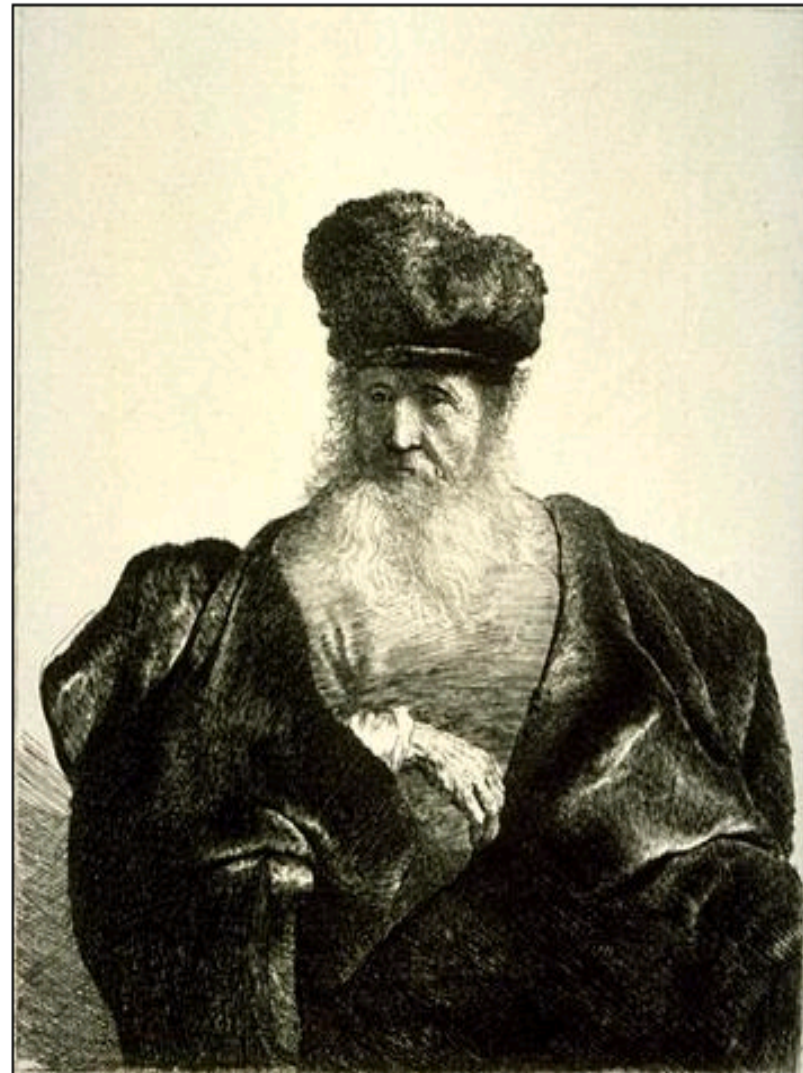
Topics:

[\[Movement - Baroque\]](#)

[\[Genre - Portraiture - Men\]](#)

[Return to Search Results](#)

Immagine con
metadati



Metadati per DRM

- **permessi** azioni ammesse: “*you can copy this*”
- **proibizioni** azioni NON ammesse, spesso come eccezioni di permessi “*you can copy this – but not for commercial purposes*”
- **requisiti** azioni che devono essere effettuate allo scopo di usufruire dei permessi “*you can copy this – but not for commercial purposes – but you must report the fact that you have done so*”

Sommario

- L'industria dei contenuti digitali è in continua espansione e genera nuove problematiche per le questioni di proprietà intellettuale
- I documenti digitali sono oggetti complessi, facilmente "costruibili" a partire da diverse fonti, e quindi pongono molte questioni di proprietà intellettuale
- Alcune tecnologie sono cruciali per proteggere la proprietà intellettuale: es. DOI

Domande di autotest

- Cos'è un documento digitale?
- Cos'è un metadato?
- Cosa sono i DOI e a che servono?

Riferimenti

- Vari Ministeri, *I contenuti digitali nell'era di Internet*, 2005
- A. Mathes, *Folksonomies - Cooperative Classification and Communication Through Shared Metadata*, 2004
www.adammathes.com/academic/computer-mediated-communication/folksonomies.html
- N. Paskin, *The DOI Handbook*, vers 4.4.1, 2006

Siti

- www.copyrightmyths.org
- www.doi.org
- dublincore.org
- www.ifla.org/II/metadata.htm
- www.rightscom.com
- www.openrightsgroup.org
- www.dmin.it/
- www.futureofthebook.org/
- www.flickr.org
- www.facebook.com/

Domande?

