



La distribuzione e l'accesso ai contenuti

Paolo Ciancarini
Dipartimento di Scienze dell'Informazione
Università di Bologna



Scopo della lezione

- La distribuzione e l'accesso ai documenti digitali
- I motori di ricerca
- Problematiche di PI nei motori di ricerca
- Problematiche di PI nelle biblioteche digitali

I canali digitali offrono molte informazioni

- Internet contiene oltre 10×10^{18} bytes
- Google indicizza oltre 10×10^{11} documenti Web
- Google Books digitalizza oltre un milione di libri l'anno

- Molte biblioteche offrono pubblicazioni digitali on line
- Moltissime pubblicazioni di ricerca accademica vengono pubblicate on line
- **Molto contenuto pregiato nasce e resta digitale**

Table 1.2: Worldwide production of original information, if stored digitally, in terabytes circa 2002. Upper estimates assume information is digitally scanned, lower estimates assume digital content has been compressed.

Storage Medium	2002 Terabytes Upper Estimate	2002 Terabytes Lower Estimate	1999- 2000 Upper Estimate	1999- 2000 Lower Estimate	% Change Upper Estimates
Paper	1,634	327	1,200	240	36%
Film	420,254	76,69	431,690	58,209	-3%
Magnetic	4,999,230	3,416,230	2,779,760	2,073,760	80%
Optical	103	51	81	29	28%
TOTAL:	5,421,221	3,416,281	3,212,731	2,132,238	69%

Source: How much information 2003

www.sims.berkeley.edu/research/projects/how-much-info-2003/execsum.htm#summary

Nuovi bisogni

Utenti

Complessità e ricchezza dell'informazione

Interdisciplinarietà

Collaborazione anche con distribuzione geografica

Informazioni scientifico-tecnologiche che raddoppiano ogni lustro

Grande interesse per i dati

Risposte

Personalizzazione

Integrazione

Multicanalità

Funzioni di collaborazione

Servizi di notifica e raccomandazioni

Data mining

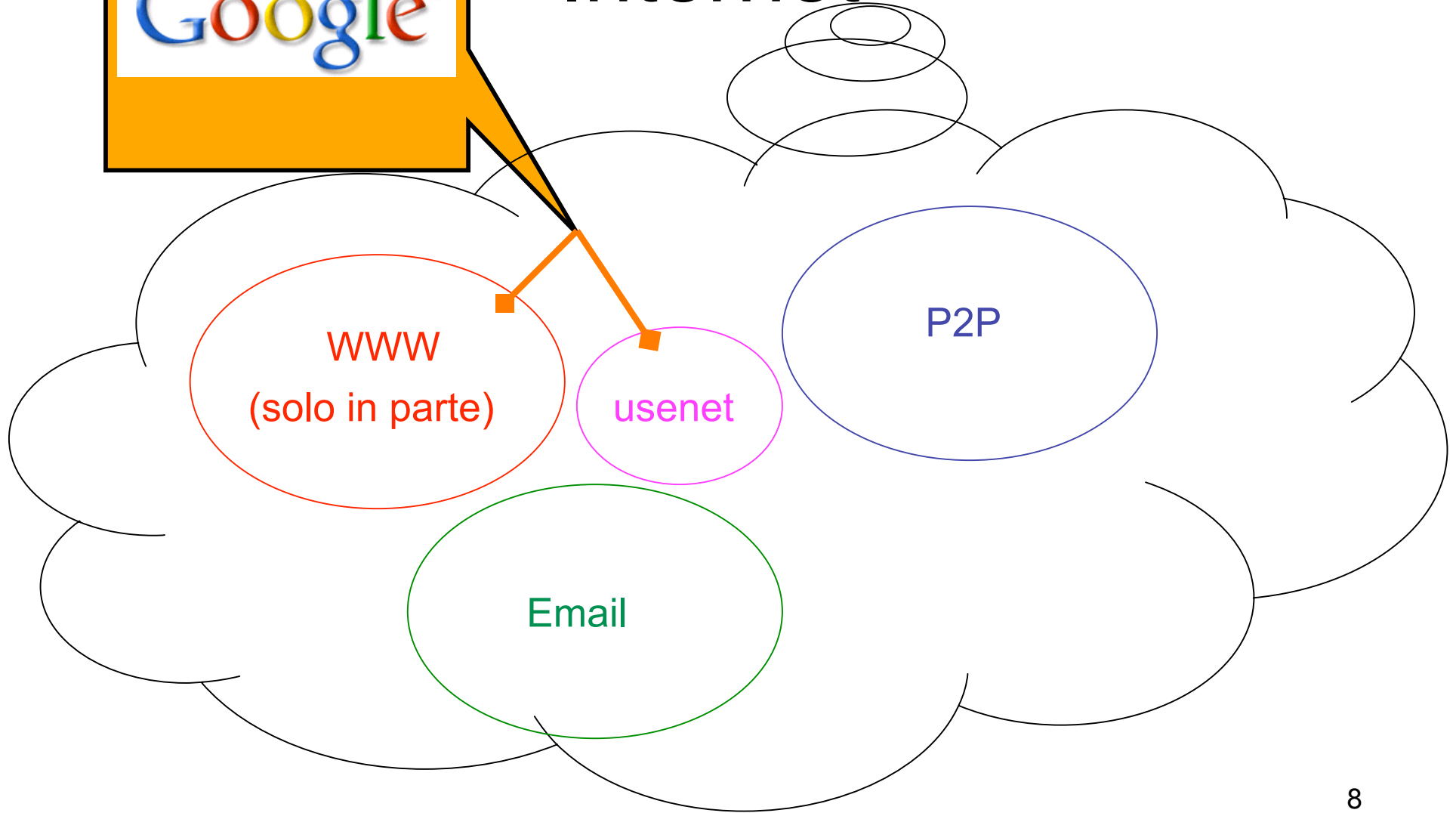
La condivisione delle informazioni

Le informazioni digitali possono essere distribuite in molti modi su Internet

- Email
- Siti Web, server e repository
- Servizi Peer to Peer
- Altri servizi Internet (es. Usenet)



Internet



Cos'è Google?

- “Un motore di ricerca perfetto è paragonabile alla mente di Dio”,
Sergey Brin, co-inventore di Google
www.technologyreview.com/articles/05/01/issue/ferguson0105.asp?p=0
- “Google oggi è un sistema per fare pubblicità”. Eric Schmidt, Google CEO,
Intervista a Wired, Aprile 2007



L'accesso alle informazioni

- Tutti i distributori di informazioni (es. giornali, oppure università) su Internet si preoccupano di offrire accesso via motori di ricerca specializzati o generalisti
- Il più famoso motore generalista al momento (2008) è Google

U.S. Expanded Search Entities by Search Queries, July 2008			
Expanded Search Entity	June 2008 (M)	July 2008 (M)	Percent Change (%)
Total expanded search	16,668	17,158	3.0
Google sites	9,601	9,945	4.0
Google	7,277	7,463	3.0
YouTube/all other	2,324	2,482	7.0
Yahoo sites	2,570	2,546	-1.0
Yahoo	2,530	2,510	-1.0
All other	40	36	-10.0
Microsoft sites	1,102	1,090	-1.0
MSN-Windows Live	1,069	1,058	1.0
Microsoft/all other	33	32	-3.0
AOL LLC	792	814	3.0
AOL Search Network	430	452	5.0
Mapquest/all other	362	362	0.0
Fox Interactive Media	457	547	20.0
MySpace	448	539	20.0
All other	9	8	-11.0
Ask Network	506	535	6.0
Ask.com	341	364	7.0
MyWebSearch.com/all other	165	171	4.0
eBay	444	435	-2.0
craigslist.org	342	340	-1.0
Facebook	157	173	10.0
Amazon sites	152	166	9.0

Note: locations include home, work, and universities.

Source: comScore, 2008

Come fa soldi Google?

Con la pubblicità!

- AdWords
- AdSense
- Google Search Appliance
- Google Answers
- Content-Specific Ads

HITTING THE BRAKES

GM's U.S. ad spending by medium (\$ in millions)

MEDIA	JAN.-NOV.'06	JAN.-NOV.'05	%CHANGE
TV	\$1,228.9	\$1,448.0	-15.1
Magazine	364.5	441.5	-17.4
Newspaper	196.3	527.6	-62.8
Internet	110.5	92.6	19.4
Radio	93.8	106.5	-11.8
Outdoor	32.0	37.2	-14.0
Total	2,026.0	2,653.3	-23.6

Source: TNS Media Intelligence

AdWords

- **AdWords** è il servizio di pubblicità prodotto da Google
- E' la sua maggior fonte di reddito (16.4 miliardi \$ nel 2007)
- AdWords offre pubblicità pay-per-click (PPC) e di posizionamento di siti sia testuali che grafici (banner)
- Modello dei prezzi di servizio basato su aste

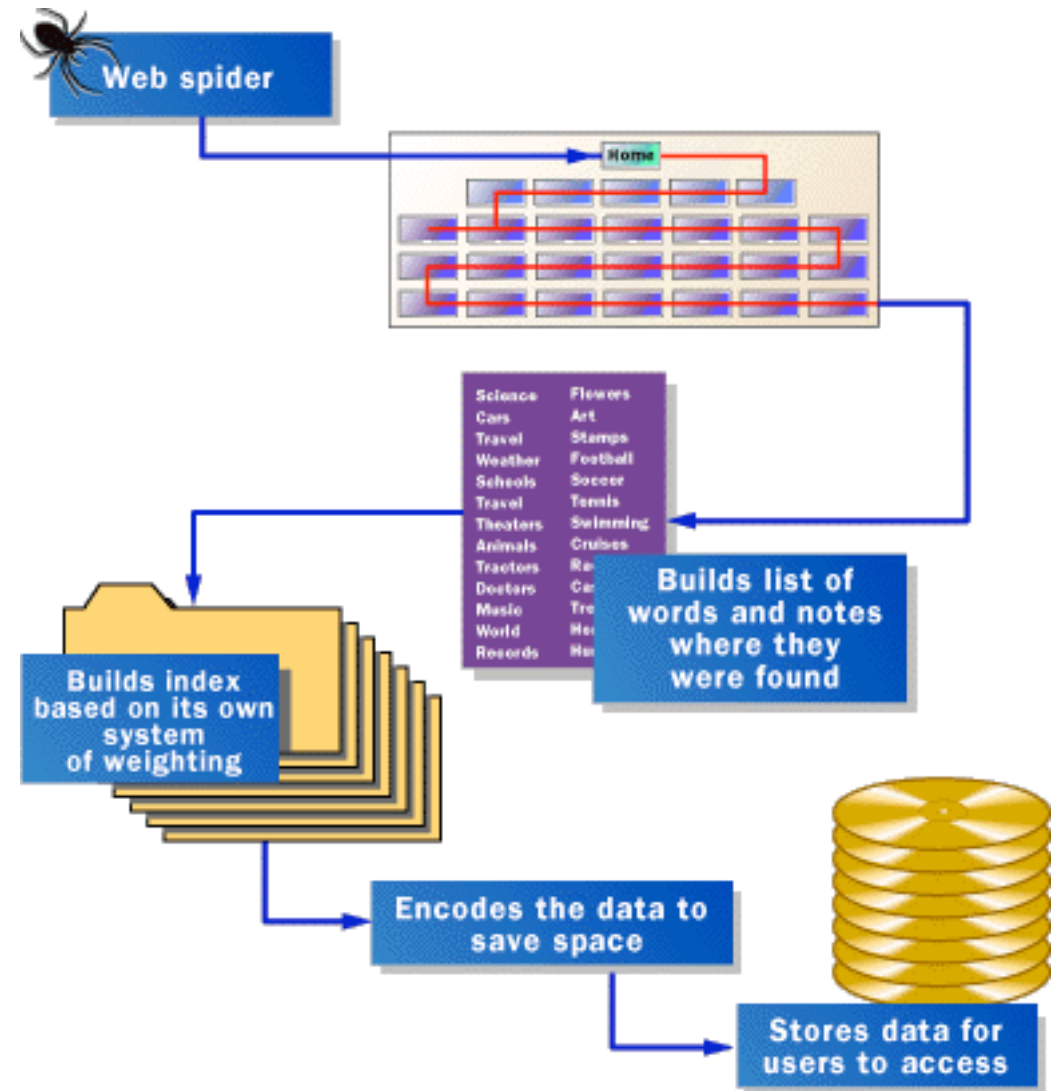
Adsense

- AdSense è un software pubblicitario gestito da Google
- Chi gestisce un sito può inglobarlo nelle proprie pagine, e Google gli riconosce un compenso pay-per-click

Come funziona Google

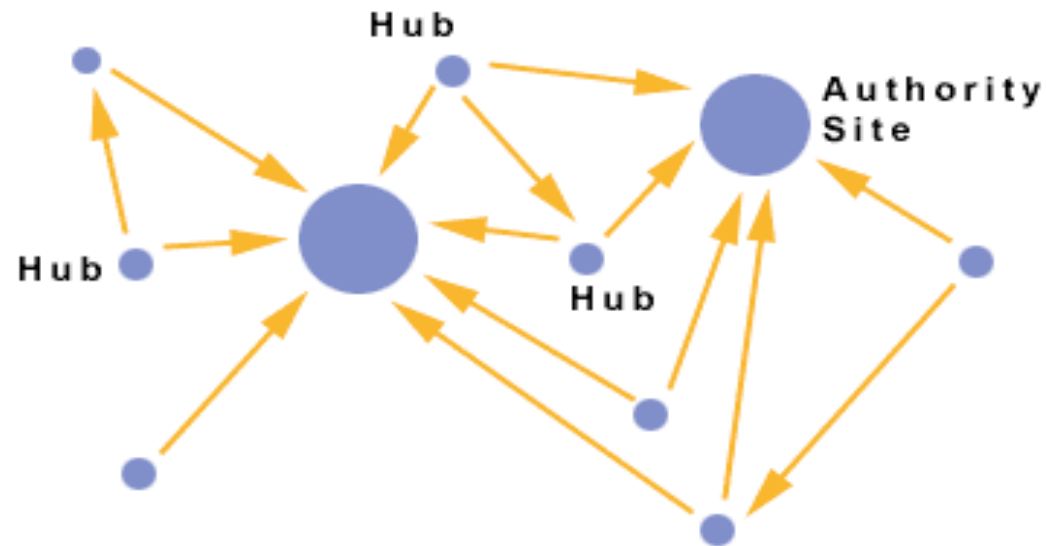
Un motore di ricerca su Web

1. Naviga il Web (**crawling**) e legge e copia le pagine che trova
2. Classifica le pagine trovate (**indexing**)
3. Risponde alle interrogazioni (**answering**)



La forma del Web

- Il Web è un "grafo non connesso"
- **Hub**: pagina che punta a molte autorità
- **Autorità**: pagina puntata da molti hub



PageRank

- Idea: le pagine cui puntano molti link sono migliori
 - Ma non tutti i link sono uguali
 - I link di una pagina importante debbono contare di più di quelli di una pagina poco importante
- PageRank base $R(A)$ della pagina A:

$$R(A) = \sum_{(B,A) \in G} \frac{R(B)}{\text{outDegree}(B)}$$

- $\text{outDegree}(B)$ = numero di link su B
 - La pagina B contribuisce al voto di tutte le pagine cui punta
- Modello del navigatore casuale
 - Una persona naviga casualmente seguendo i link di pagina in pagina
 - Il PageRank di A è proporzionale alla frequenza con cui il navigatore casuale incontra A

Oltre Google

- Google Scholar
- Google Books
- Google Patents
- Google News
- Google Immagini e Video



Google Scholar



[Web](#) [Images](#) [Video](#) [News](#) [Maps](#) [more »](#)

the hospital anxiety depression scale

Search

[Advanced Scholar Search](#)
[Scholar Preferences](#)
[Scholar Help](#)

"the" is a very common word and was not included in your search. [\[details\]](#)

Scholar All articles [Recent articles](#) Results 1 - 10 of about 136,000 for the [hospital anxiety depression scale](#). (0.23 seconds)

All Results

[A Zigmond](#)
[M Hamilton](#)
[R Snaith](#)
[C Herrmann](#)
[A Dahl](#)

[The hospital anxiety and depression scale. - group of 2 »](#)

AS Zigmond, RP Snaith - Acta Psychiatr Scand, 1983 - ncbi.nlm.nih.gov
The **hospital anxiety and depression scale**. Zigmond AS, Snaith RP. A self-assessment **scale** has been developed and found to be a reliable ...
[Cited by 3432](#) - [Related Articles](#) - [Web Search](#) - [A-Link](#)

[International experiences with the Hospital Anxiety and Depression Scale--a review of validation ... - group of 3 »](#)

C Herrmann - J Psychosom Res, 1997 - ncbi.nlm.nih.gov
Click here to read International experiences with the **Hospital Anxiety and Depression Scale**--a review of validation data and clinical results. ...
[Cited by 447](#) - [Related Articles](#) - [Web Search](#) - [A-Link](#) - [ACNP Holdings](#) - [BL Direct](#)

[The validity of the Hospital Anxiety and Depression Scale. An updated literature review. - group of 3 »](#)

I Bjelland, AA Dahl, TT Haug, D Neckelmann - J Psychosom Res, 2002 - ncbi.nlm.nih.gov
Click here to read The validity of the **Hospital Anxiety and Depression Scale**. An updated literature review. Bjelland I, Dahl AA, Haug ...
[Cited by 278](#) - [Related Articles](#) - [Web Search](#) - [A-Link](#) - [ACNP Holdings](#) - [BL Direct](#)

[A validation study of the Hospital Anxiety and Depression Scale \(HADS\) in different groups of Dutch ... - group of 5 »](#)

PH SPINHOVEN, J ORMEL, PPA SLOEKERS, G KEMPEN, AEM ... - Psychological Medicine, 1997 - Cambridge Univ Press
... Research on the dimensional structure and reliability of the **Hospital Anxiety and Depression Scale (HADS)** and its relationship with age is scarce. ...
[Cited by 194](#) - [Related Articles](#) - [Web Search](#) - [A-Link](#) - [ACNP Holdings](#) - [BL Direct](#)

•Link all'articolo
•Metadati
•Snippet

• accesso full text privilegiato

• cerca nel Web la stessa query

• Archivio Collettivo Nazionale Periodici

• accesso a pagamento

• Articoli citanti

• Articoli correlati

Google Scholar è un sottoinsieme di Google



Google Index



AlCensus - USGenWeb Archives - Rootsweb

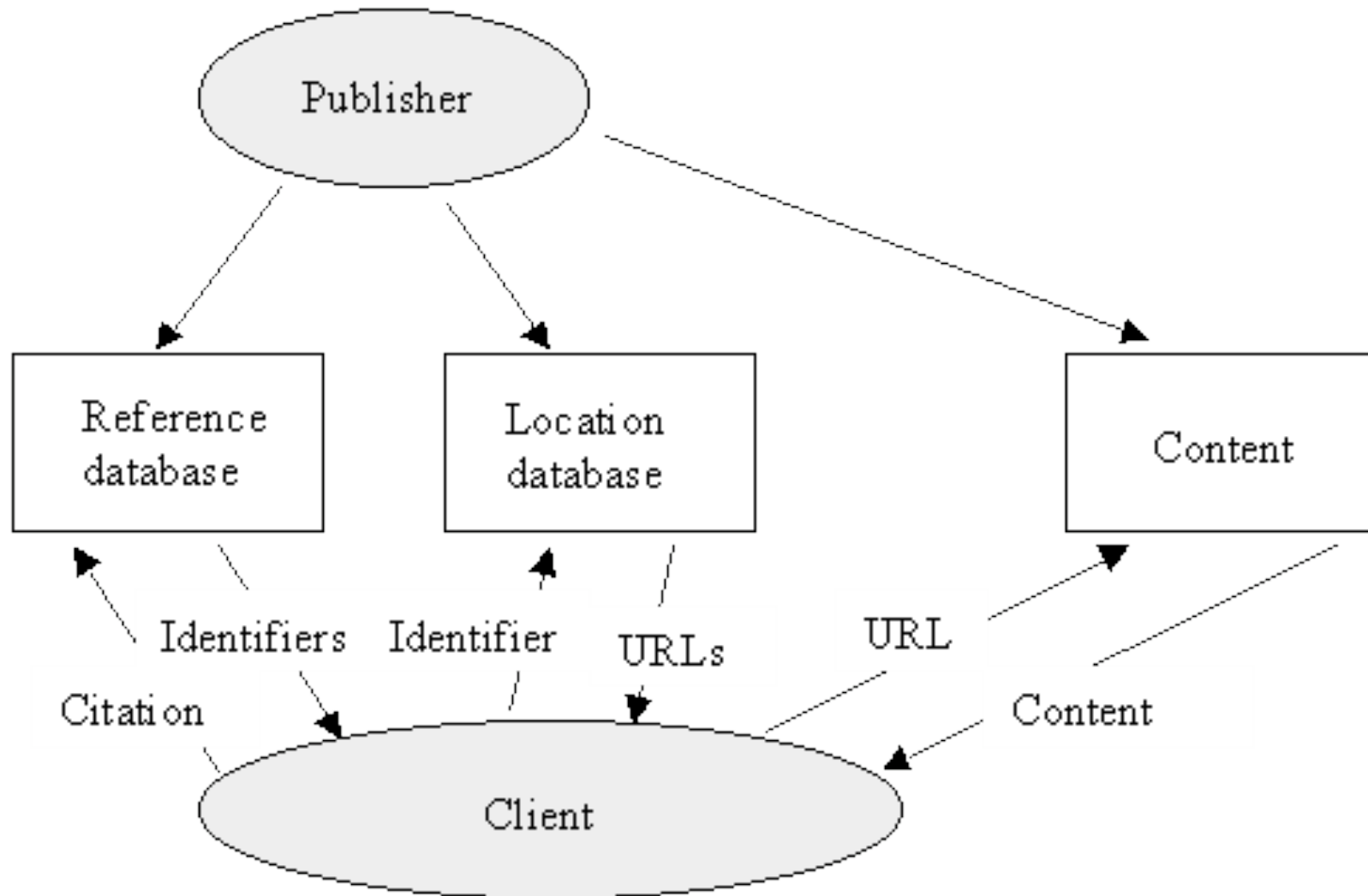
INDEX.

M-Daniel, Daniel, 25	M-Daniel, John, 27	M-Daniel, John, 28	M-Daniel, John, 29	M-Daniel, John, 30	M-Daniel, John, 31	M-Daniel, John, 32	M-Daniel, John, 33	M-Daniel, John, 34	M-Daniel, John, 35	M-Daniel, John, 36	M-Daniel, John, 37	M-Daniel, John, 38	M-Daniel, John, 39	M-Daniel, John, 40	M-Daniel, John, 41	M-Daniel, John, 42	M-Daniel, John, 43	M-Daniel, John, 44	M-Daniel, John, 45	M-Daniel, John, 46	M-Daniel, John, 47	M-Daniel, John, 48	M-Daniel, John, 49	M-Daniel, John, 50	M-Daniel, John, 51	M-Daniel, John, 52	M-Daniel, John, 53	M-Daniel, John, 54	M-Daniel, John, 55	M-Daniel, John, 56	M-Daniel, John, 57	M-Daniel, John, 58	M-Daniel, John, 59	M-Daniel, John, 60	M-Daniel, John, 61	M-Daniel, John, 62	M-Daniel, John, 63	M-Daniel, John, 64	M-Daniel, John, 65	M-Daniel, John, 66	M-Daniel, John, 67	M-Daniel, John, 68	M-Daniel, John, 69	M-Daniel, John, 70	M-Daniel, John, 71	M-Daniel, John, 72	M-Daniel, John, 73	M-Daniel, John, 74	M-Daniel, John, 75	M-Daniel, John, 76	M-Daniel, John, 77	M-Daniel, John, 78	M-Daniel, John, 79	M-Daniel, John, 80	M-Daniel, John, 81	M-Daniel, John, 82	M-Daniel, John, 83	M-Daniel, John, 84	M-Daniel, John, 85	M-Daniel, John, 86	M-Daniel, John, 87	M-Daniel, John, 88	M-Daniel, John, 89	M-Daniel, John, 90	M-Daniel, John, 91	M-Daniel, John, 92	M-Daniel, John, 93	M-Daniel, John, 94	M-Daniel, John, 95	M-Daniel, John, 96	M-Daniel, John, 97	M-Daniel, John, 98	M-Daniel, John, 99	M-Daniel, John, 100
----------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	---------------------



= Google Scholar

Il problema del reference linking



Google Scholar Preferences



Preferences

[About Google Schola](#)

Save your preferences when finished and **return to search**.

[Save Preferences](#)

Scholar Preferences

Interface Language

Display Google tips and messages in:

Search Language

Search for pages written in any language (Recommended).

Search only for pages written in these language(s):

- Chinese (Simplified) English German Portuguese
 Chinese (Traditional) French Japanese Spanish

Library Links

[\(what's this?\)](#)

[Find Library](#)
e.g., Harvard

Show library access links for (choose up to three libraries):

- Universita' degli Studi di Bologna (A-Link)
 ACNP Italian Union Catalogue of Serials (ACNP Full Text)
 Scuola Normale Superiore (Full Text@IngentaConnect)

Online access to library subscriptions is usually restricted to patrons of that library. You may need to login with your library password, use a campus computer, or configure your browser to use a library proxy. Please visit your library's website or ask a local librarian for assistance.

Motori specializzati per ricercatori

- Scopus: 265 milioni di "riferimenti" (orientato alle Scienze "dure")
- WebofScience: oltre il 50% di Scopus (orientato alle Scienze umane)
- Springerlink: oltre 3 milioni di documenti
- PUBMED: circa 16 milioni di documenti
- IngentaConnect: 22 milioni di documenti

- Scirus: circa 300 milioni di pagine
- Google Scholar: circa 500 milioni di pagine

Cosa ci dicono queste cifre? (cifre al maggio 2007)

Un confronto

Table 1

Comparison of features in Web of Science, Scopus, and Google Scholar

	Web of Science	Scopus	Google Scholar
Indexing and abstracting	Yes	Yes	No
Years covered-journals	1900 to present (Science) 1956-present (Social Science) 1975-present (Arts and Humanities)	1966 to present for some journals, but many date back to 1996 to present	Not revealed
Years covered-citations	1900 to present	1996 to present	Not revealed
Fee-based	Yes	Yes	No
Contents	9300 journals (Science, Social Science and Arts and Humanities)	15,000 journals (Science and Social Science)	Not revealed

Springerlink

English



Athens Authentication Point

Welcome!

To use the personalized features of this site, please **log in** or **register**.

If you have forgotten your username or password, we can **help**.

My SpringerLink

[Marked Items](#)

[Alerts](#)

[Order History](#)

Saved Items

[All](#)

[Favorites](#)



Welcome to SpringerLink!

SpringerLink is one of the world's leading interactive databases for high-quality STM journals, book series, books, reference works and the Online Archives Collection. SpringerLink is a powerful central access point for researchers and scientists.

Find content by keyword

[more options](#)

Content type

All (3,248,481)
[Publications \(18,847\)](#)
[Journals \(1,847\)](#)
[Book Series \(761\)](#)
[Books \(16,239\)](#)
[Reference Works \(90\)](#)

Featured library

[Chinese Library of Science \(13,718\)](#)
[Russian Library of Science \(435,326\)](#)

Subject

[Architecture, Design and Arts \(37\)](#)
[Behavioral Science \(76,548\)](#)
[Biomedical and Life Sciences \(839,537\)](#)
[Business and Economics \(72,310\)](#)
[Chemistry and Materials Science \(600,189\)](#)
[Computer Science \(275,563\)](#)
[Earth and Environmental Science \(309,278\)](#)
[Engineering \(366,423\)](#)
[Humanities, Social Sciences and Law \(188,296\)](#)
[Mathematics and Statistics \(262,421\)](#)
[Medicine \(680,695\)](#)
[Physics and Astronomy \(440,872\)](#)

[Frequently asked questions](#) | [General information on journals and books](#) | [Send us your feedback](#) | [Impressum](#)

© Springer. Part of Springer Science+Business Media

[Privacy](#), [Disclaimer](#), [Terms and Conditions](#), © Copyright Information

Google Books

- Google fece un accordo nel 2003 con 5 grandi biblioteche
 - **Stanford**: tutto
 - **Harvard**: 40K libri estratti da una collezione di 15M
 - **U. Michigan**: tutto
 - **New York Public Library**: sottoinsieme estratto da 20M volumi
 - **Oxford**
- Criteri di selezione = fuori diritto autore (1923), interessante, non troppo fragile

Google Books

- Digitalizza 3,000 libri al giorno
- Scanning manuale
 - <http://books.google.com/books?vid=OCLC03812955&id=1GB1kuY5-pkC&pg=PA3&lpg=PA3>
 - http://books.google.com/books?vid=0sVgqoZH8_0vk2uEA6uPPZ&id=n-28bvRNoroC&pg=RA1-PR1000
 - <http://books.google.com/books?vid=OCLC03812955&id=1GB1kuY5-pkC&pg=PR32>
- Le biblioteche consorziate li prestano e poi li riottengono indietro

Google Books

In totale 18 milioni di libri (al 2006) in 430 linguaggi

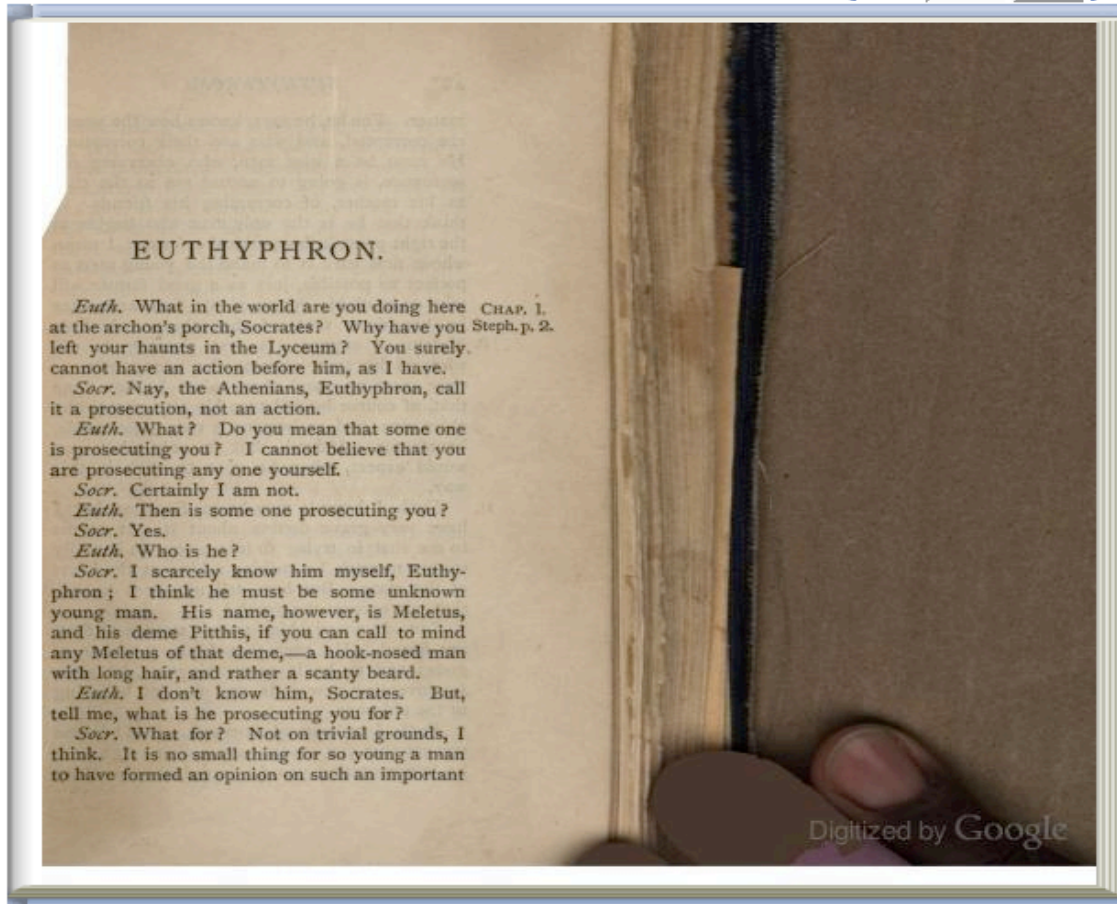
- 16% fuori © -> accesso completo
- 9% in print (e con ©) -> solo quel che consente l'editore
- 15% non in print ma sotto © -> snippet
- Dei restanti nulla (no preview)

Quanto costa?

- Se lo scanning time è di 1 pag/sec
- 20.000 pag/giorno
- 4M pag/anno
- 100 anni uomo per 1 milione di libri (1 libro=400 pag)
- Centri di scansionamento in India e Cina
- Costo medio per libro meno di 10 US \$

Bookdrive





The Trial and Death of Socrates: Being the Euthyphron, Apology, Crito and Phaedo of Plato

By Plato, F. J. (Frederick John) Church



[Download](#) PDF - 8.9M

[About this book](#)

[Title Page](#)
[Copyright](#)

Buy this book
[Abebooks](#)
[Alibris](#)
[Amazon](#)
[Froogle](#)

[Find this book in a library](#)

Search in this book

[Sign in](#)



[Web](#) [Images](#) [Video](#) ^{New!} [News](#) [Maps](#) [more »](#)

"stanford university libraries" OR "3 6105" OR "36105"

Search Books

[Advanced Book Search](#)
[Google Book Search Help](#)

Search: All books Full view books

Book Search

Books 1 - 100 with 160000 pages on "stanford university libraries" OR "3 6105" OR "36105". (0.00 seconds)

Did you mean: ["stanford university libraries" OR "36105" OR "36105"](#)



[Encyclopedia of Library and Information Science: Volume 29 - Stanford University Libraries to... - Page 1](#)

by Allen Kent, Kent Kent - Language Arts & Disciplines - 1980 - 480 pages

... **STANFORD UNIVERSITY LIBRARIES** Historical Introduction This section sketches the history of the **Stanford University Libraries**; it covers primarily the ...

Limited preview - [Index](#) - [About this book](#)

[Academic Data Service user's guide](#)

by Stanford University. Libraries - 1989 - 9 pages

Page 1

No preview available



[Subject Catalog - Page 973](#)

by Stanford University. Libraries. J. Henry Meyer Memorial Library - 1967

KNOWLEDGE, THEORY OF A new theory of vision; and other writings. By George Berkeley. Introd. by AD Lindsay. Dent; Dutton, 1963. 303 p. Contents. ...

Snippet view - [About this book](#)



[List of Serials in the Leland Stanford Junior University Library - Page 3](#)

by Stanford University. Libraries - 1916 - 169 pages

EXPLANATORY NOTE. This list includes the magazines, yearbooks, newspapers and publications of societies and institutions, which are in the library of Leland ...

Full view - [About this book](#)

[Economic Indicators](#)

by United States. Congress. Joint Committee on the Economic Report, Council of Economic Advisers (U.S.), United States.

Congress. Joint Economic Committee

No preview available

[Sign in](#)



"stanford university libraries" and "36105"

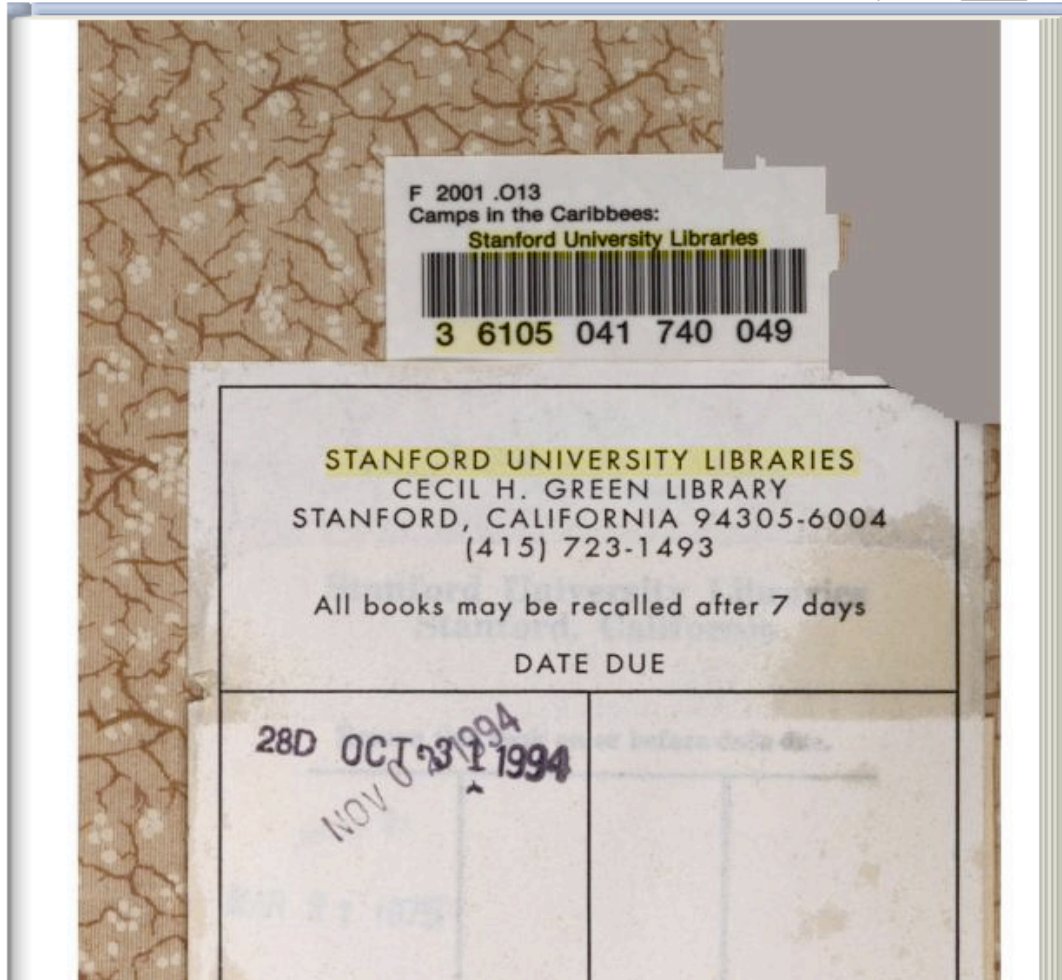
Search Books

[Advanced Book Search](#)
[Google Book Search Help](#)

[« Back to Search results](#)

Full view - [Learn more](#)

Page Go



Camps in the Caribbees: The Adventures of a Naturalist in the Lesser Antilles

By Frederick A. (Frederick Albion) Ober



[Download](#) PDF - 13.0M

[About this book](#)

[Table of Contents](#)

Buy this book

[Abebooks](#)

[Alibris](#)

[Amazon](#)

[Froogle](#)

[Find this book in a library](#)

Search in this book

"stanford univer: Go

Results 1-1 of 1

Page

... : **Stanford University Libraries**

...

Google Books: problematiche di copyright

- Visualizza un pezzettino di testo (snippet) per i libri sotto copyright
 - Snippet = tre righe
 - Ogni ricerca risulta in tre "pezzettini", ma dice anche il numero di volte che la parola cercata appare nel libro

Copyright negli USA

- pubblicato prima del 1923 - Public Domain
- pubblicato dal 1923 al 1963
 - Rinnovo dopo 14 anni
 - Solo il 15% fu rinnovato (200,000 opere)
 - Il restante 85% entrò nel Public Domain
- Pubblicato dopo il 1963: copyright
 - Fino a 70 anni dalla morte dell'autore

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 usano solo dopo una ricerca “ragionevolmente diligente”
- Report 2006 al Congresso USA
- Sistema di determinazione www.copyright.gov/orphan/

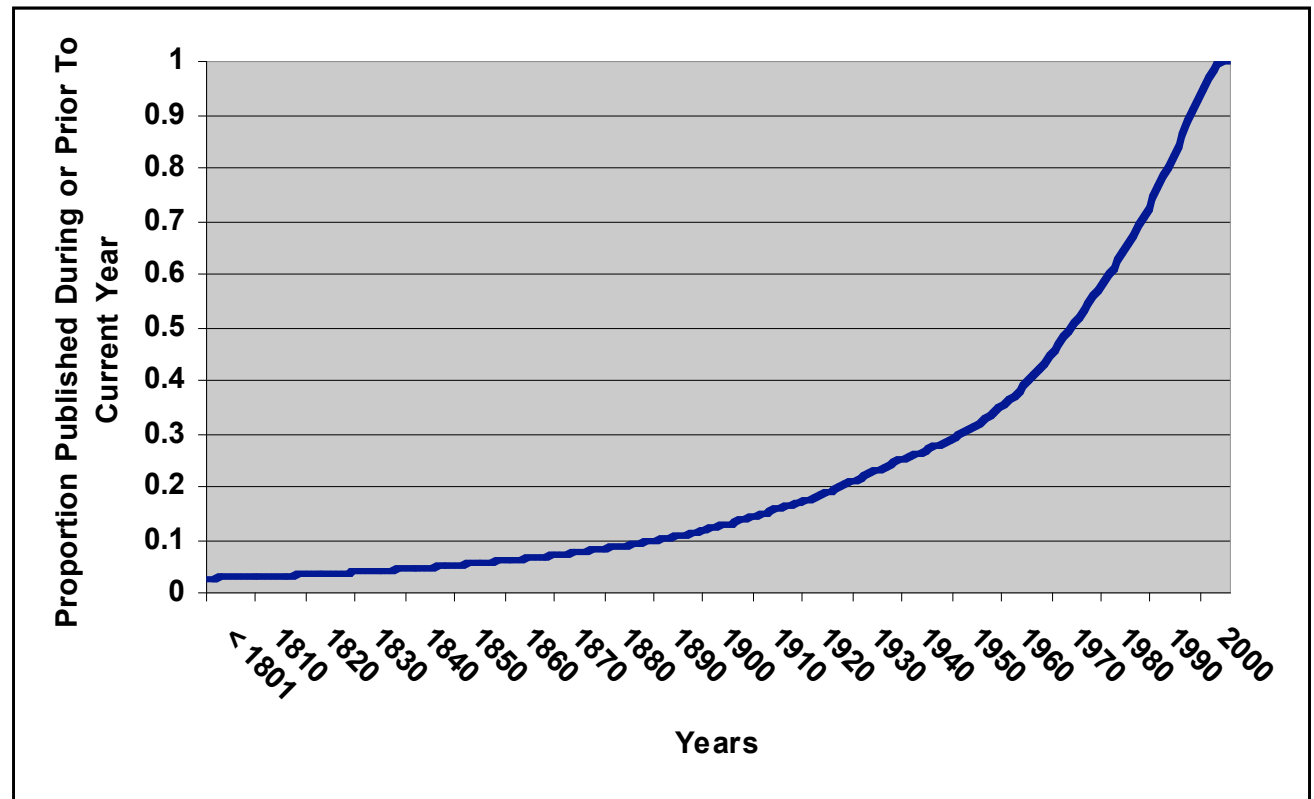
Google Books: problemi

R. Litwin, “On Google’s Monetization of Libraries”

1. Privacy [cookies]
2. Commercializzazione
3. Democrazia e parità di accesso
4. Disintermediazione
5. Decontestualizzazione della conoscenza
6. Chiusura (superamento) delle biblioteche?

Google Books: problemi

- Se si fissa lo spartiacque al 1923, l'80% dei libri offerti a Google da digitalizzare è sotto diritto di autore



Azioni contro Google

- Google wins a nude copyright infringement case in the U.S. Federal Court of Appeals (5-2007)

Google wins and is now able to show thumbnail sized photographs from sexually explicit sites. The Federal court of Appeals lifted the preliminary injunction against Google for showing such images. Though the court did say that the company might be liable for allowing links to such sites displaying unauthorized photographs. A lower court found that Google's thumbnail images violated copyright of that adult magazine and Web publisher, Perfect 10. Google used "fair use" as its defense. The court said Google was probably not responsible for displays of underlying images from Perfect 10's Web site. But the Federal court said that they concluded that Perfect 10 is unlikely to overcome Google's "fair use" defense and so then ordered the preliminary injunction be vacated, allowing Google to again be able to use the thumbnail images. This is not the first time that Perfect 10 has taken some exception to Google's showing their thumbnail images. Back in 2001, the publisher alleged that Google links to Web sites that republish images of its nude models without authorization. Then again in 2004, it sued Google alleging copyright infringement, and in 2005, filed similar claims against Amazon.com and its A9.com subsidiary, saying they provide links to Google search results. Others claim that this is just more publicity for Perfect 10 and its subsidiaries.

Azioni contro Google Books

- L'associazione delle case editrici americane (AAP) immediatamente protestò contro Google notificando l'infrazione alle regole di copyright
- Risposta di Google: offrire agli editori un meccanismo di opt-out

Azioni contro Google Books

- Cause intentate da:
 - Associazione degli autori + autori come individui
 - 5 case editrici per conto della loro associazione AAP
- Accusa: Google Books viola il *Copyright Act 1976*
- Giustificazione di Google: 'Fair Use'

Sec. 106. Exclusive rights in copyrighted works

- Subject to sections 107 through 122, the owner of copyright under this title has the exclusive rights to do and to authorize any of the following:
 1. To reproduce the copyrighted work in copies or phonorecords;
 2. To prepare derivative works based upon the copyrighted work;
 3. To distribute copies or phonorecords of the copyrighted work to the public by sale or other transfer of ownership, or by rental, lease, or lending;
 4. In the case of literary, musical, dramatic, and choreographic works, pantomimes, and motion pictures and other audiovisual works, to perform the copyrighted work publicly;
 5. In the case of literary, musical, dramatic, and choreographic works, pantomimes, and pictorial, graphic, or sculptural works, including the individual images of a motion picture or other audiovisual work, to display the copyrighted work publicly; and
 6. In the case of sound recordings, to perform the copyrighted work publicly by means of a digital audio transmission.

'Fair use'

Sec. 107. Limitations on exclusive rights: Fair use:

Notwithstanding the provisions of sections 106 and 106A, the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright.”

Fair use

In sostanza la dottrina del "*fair use*" prevede quattro eccezioni al "copyright":

- Motivi educativi
- Motivi editoriali (es. recensioni)
- Motivi industriali (es. thumbnails)
- Copia personale di prodotto "acquistato"

Valore del Fair Use

- Un rapporto del sett. 2007 preparato per la Computer & Communication Industry Association degli USA sostiene che l'industria basata sul Fair Use negli USA vale il 18% del PIL (4.500 G\$)
- Le industrie interessate:
 - Istituzioni educative
 - Motori di ricerca e fornitori di "Web Hosting"
 - Sviluppatori di software
 - Fabbricanti di dispositivi che permettono di duplicare informazioni

Esempi di Casi basati su "Fair Use"

- **Universal City Studios vs. Sony Corp., 464 U.S. 417 (1984)**
 - La Corte Suprema USA decise che la registrazione casalinga di una trasmissione TV ricade sotto il "fair use".
 - Uno dei pochi casi ove copiare interamente un'opera è considerato "fair use"
 - Le prove indicarono che la maggior parte degli utenti volevano spostare nel tempo la fruizione ("time-shifting") e non costruire una videoteca
 - *Importante*: La Corte arguì che lo spostare nel tempo la fruizione non priva i proprietari dei diritti del giusto compenso
- **Kelly vs. Arriba-Soft, 03 C.D.O.S. 5888 (9th Cir. 2003)**
 - Un motore di ricerca può creare piccole riproduzioni ("thumbnails") per metterle sul proprio sito: non danneggia il mercato potenziale di tali immagini
 - *Importante*. I thumbnail sono di peggior qualità rispetto agli originali e servono solo per denotare le immagini ed aiutare il pubblico ad accedere gli originali
- **Religious Technology Center vs. Lerma, 40 U.S.P.Q. 2d 1569 (E.D. Va. 1996)**
 - Intere pubblicazioni di Scientology furono messe su Internet da vari individui senza permesso. La Corte negò il "fair use" perché le opere erano state copiate interamente (sarebbe invece stata consentita la pubblicazione di estratti)
- Vedi anche cyberlaw.stanford.edu

Google Books: 'fair use'?

La "Section 107" definisce il "fair use" via quattro questioni:

1. Scopo e tipo di uso, in particolare se l'uso sia di natura commerciale o per scopi educativi no-profit;
2. La natura dell'opera protetta;
3. La quantità e la qualità della parte di opera usata, confrontata con l'opera intera;
4. L'effetto dell'uso sul valore potenziale di mercato dell'opera protetta

Si pensa che Google abbia buone possibilità di far passare la sua condotta nel progetto "Google Books" nell'ambito del 'fair use' dello *US Copyright Act*

Open Access (OA)

- Il movimento Open Access supporta la pubblicazione su Web di opere scientifiche, allo scopo di massimizzarne l'impatto
- Nel 2003 la Dichiarazione di Berlino definì la nozione di Open Access
- Due approcci principali:
 - Autoarchiviazione
 - Pubblicazione in riviste OA

Biblioteche digitali

- Biblioteca digitale personale
- Biblioteca digitale di comunità
- Biblioteca digitale di Google

Biblioteca digitale personale

- di testi
- di immagini
- di film
- di bookmarks

- Interfacce speciali (es. `www.netvibes.com`)
- Condivisibilità (Web 2.0)
- Funzioni semantiche (Semantic Web)

Cos'è una biblioteca digitale di comunità?

Una biblioteca digitale di questo tipo è un'**organizzazione** che offre ad una **comunità** servizi di lettura, ricerca e prestito; inoltre valuta i contenuti che offre (a volte li produce: in tal caso è molto simile ad una casa editrice)

Esempi di contenuti:

- collezioni di testi e immagini digitalizzati
- pubblicazioni accademiche
- carteggi digitalizzati
- mappe
- corsi e-learning

Biblioteca digitale di Bologna (Alm@-DL)



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA



Alm@DL

Accessibilità | Mappa del sito | Motore di ricerca | Servizio Proxy |

Alma Mater Digital Library
A cura del CIB Centro Inter-Bibliotecario

Informazioni | Contatti & Supporto | Aiuto & Documentazione | Strumenti | Statistiche

Home

Organizzazione
Staff tecnico
Progetto

Collezioni

AMS Acta
AMS Campus
AMS Miscellanea
AMS Historica
Riviste AMS

Servizi

Publicare in rete
Deposito legale
Stampa a richiesta
Digitalizzazione



Portale delle
Biblioteche
Accesso alle
risorse
elettroniche in
abbonamento:
Banche dati
Periodici
Libri
Dizionari e Enciclopedie

Alm@-DL Home | Alm@-DL Search | Collezioni | Servizi | Aree di sviluppo

Sei in: Alm@-DL

Benvenuto su Alm@-DL

Alm@-DL è la Biblioteca Digitale dell'Università di Bologna. Raccoglie, conserva e rende disponibili in rete collezioni digitali a supporto della didattica e della ricerca. Offre un accesso integrato alle risorse elettroniche disponibili gratuitamente o in abbonamento. Fornisce assistenza tecnica ai progetti di editoria elettronica e di digitalizzazione.

In tutte le collezioni con Alm@-DL Search | Nel sito web di Alm@-DL

Tutti i campi e Testo Pieno con il Catalogo Alm@-DL Search

[Ricerca avanzata](#) | [Preferenze](#) | [Aiuto](#)

In evidenza

Cerca in tutte le collezioni di Alm@-DL con Alm@-DL Search

Scopri le funzioni di ricerca avanzata di Alm@-DL Search che ti consentono di selezionare le collezioni di interesse e organizzare i risultati della ricerca. Con Alm@-DL Search puoi salvare i risultati delle tue ricerche in formati standard per aggiungerli alle tue bibliografie.

Publicare con Alm@-DL

Alm@-DL offre a Docenti e Ricercatori dell'Ateneo la possibilità di pubblicare in elettronico e diffondere via Internet:

- **Contributi di ricerca** (per massimizzarne la diffusione)
- **Materiali didattici** (per renderli disponibili agli studenti)

[Diventa Utente Autore!](#)

Notizie

Chiedi al bibliotecario!

Servizio centralizzato di assistenza on-line fornito dalle biblioteche dell'Università di Bologna

[Continua](#)

06/03/2007

AMS Campus di Facoltà

E' stato attivato un nuovo deposito campus per i materiali didattici della Facoltà di Giurisprudenza.

[Continua](#)

05/03/2007

AMS Acta

Contributi di ricerca dell'Ateneo →

AMS Campus

Materiali didattici on-line →

AMS Miscellanea

Contributi culturali dell'Ateneo →

AMS Historica

Tesori delle biblioteche →

Riviste AMS

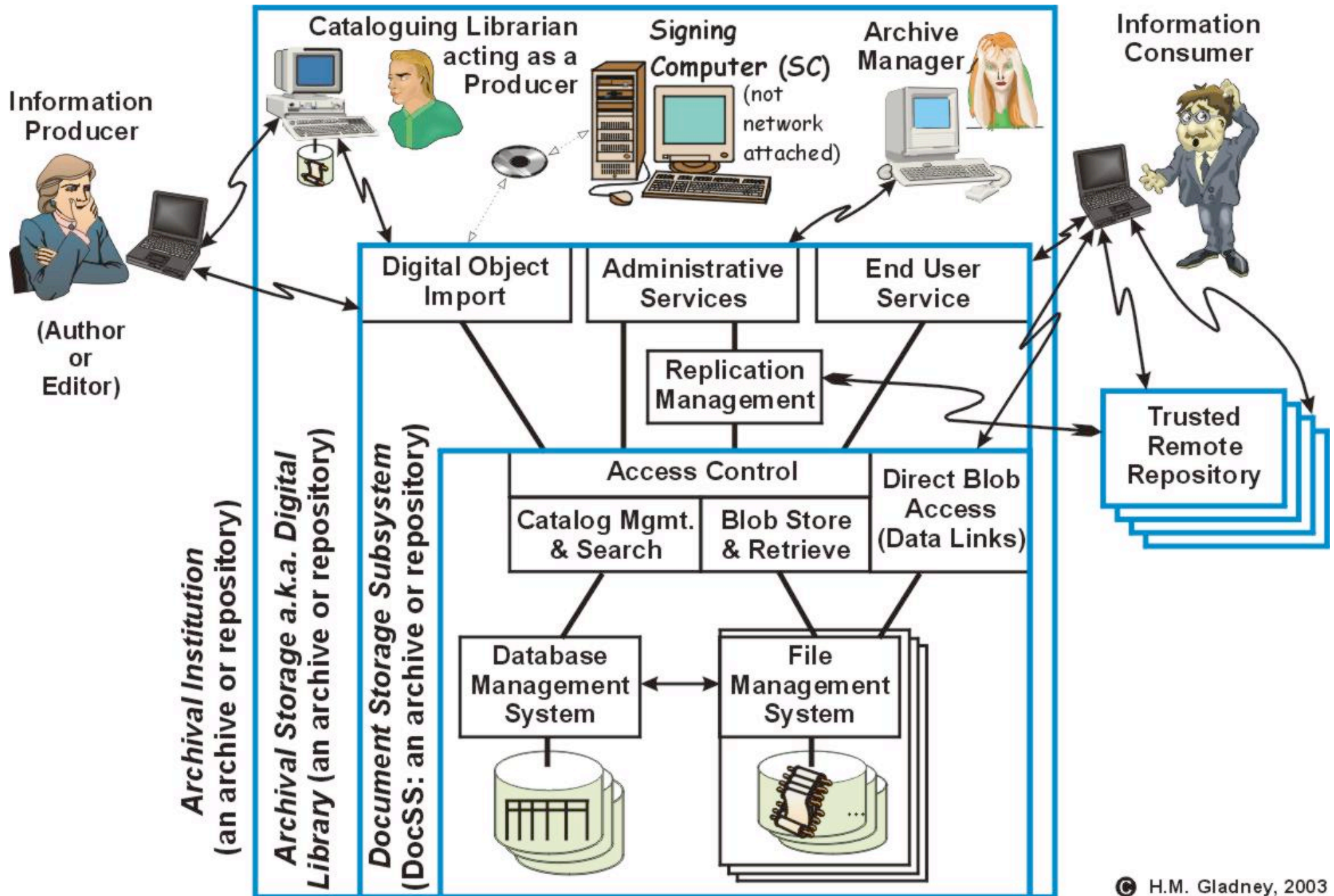
Riviste elettroniche di Alm@-DL →

Come archiviare i documenti digitali?

- Come rappresentare gli archivi on line?
 - **Modello mostra**: non adatto alle interrogazioni ma fornisce contestualizzazione
 - **Modello database**: adatto alle interrogazioni, poco contestualizzato
- Come classificare i documenti digitali?
- Come proteggere la proprietà dei contenuti?

La biblioteca digitale

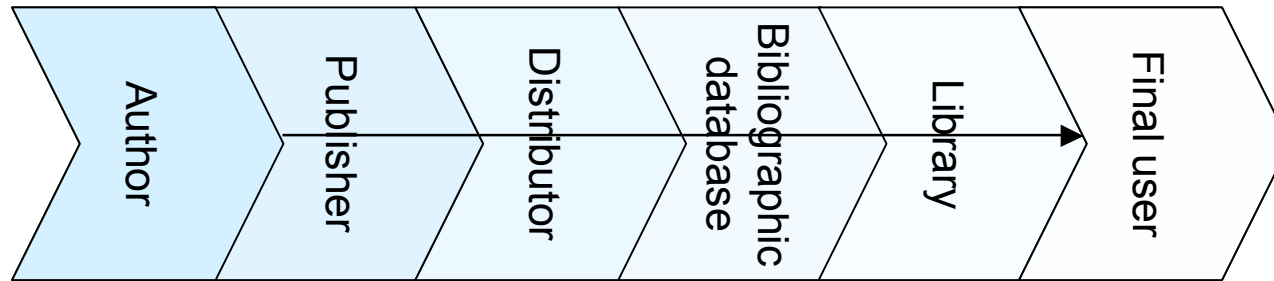
- **Biblioteca digitale** = sistema informativo on-line caratterizzato da
 - una collezione organizzata di *documenti* digitali
 - un apparato di *metadati* descrittivi e amministrativi associati a tali documenti
 - un insieme di strumenti e *servizi* di gestione, ricerca e disseminazione dei documenti e dei metadati
- Una biblioteca digitale si basa dunque su un **sistema** informatico che permette l'accesso *ubiquo* a contenuti in forma digitale: l'accesso è diretto (conoscendo l'URL di una risorsa digitale), oppure via catalogo, oppure via motore di ricerca



La filiera della biblioteca

- Scelta (con l'aiuto della comunità di utenti)
- Acquisizione
- Catalogazione
- Stoccaggio (e accesso alle opere)
- Prestito (e circolazione interlibraria)
- Riconsegna (e inventario)
- Monitoraggio della circolazione
- Conservazione (e restauro)

La filiera



Internet semplifica la filiera?

Fonte: Piero Attanasio (mEDRA, AIE – Associazione Italiana Editori)

Biblioteche digitali vs cartacee

Acquisizione: accesso a copia digitale vs possesso del cartaceo

Catalogazione: indicizzazione full text vs catalogazione umana

Stoccaggio: digitalizzazione+CMS vs scaffali e faldoni

Prestito: gestione DRM vs restituzione fisica

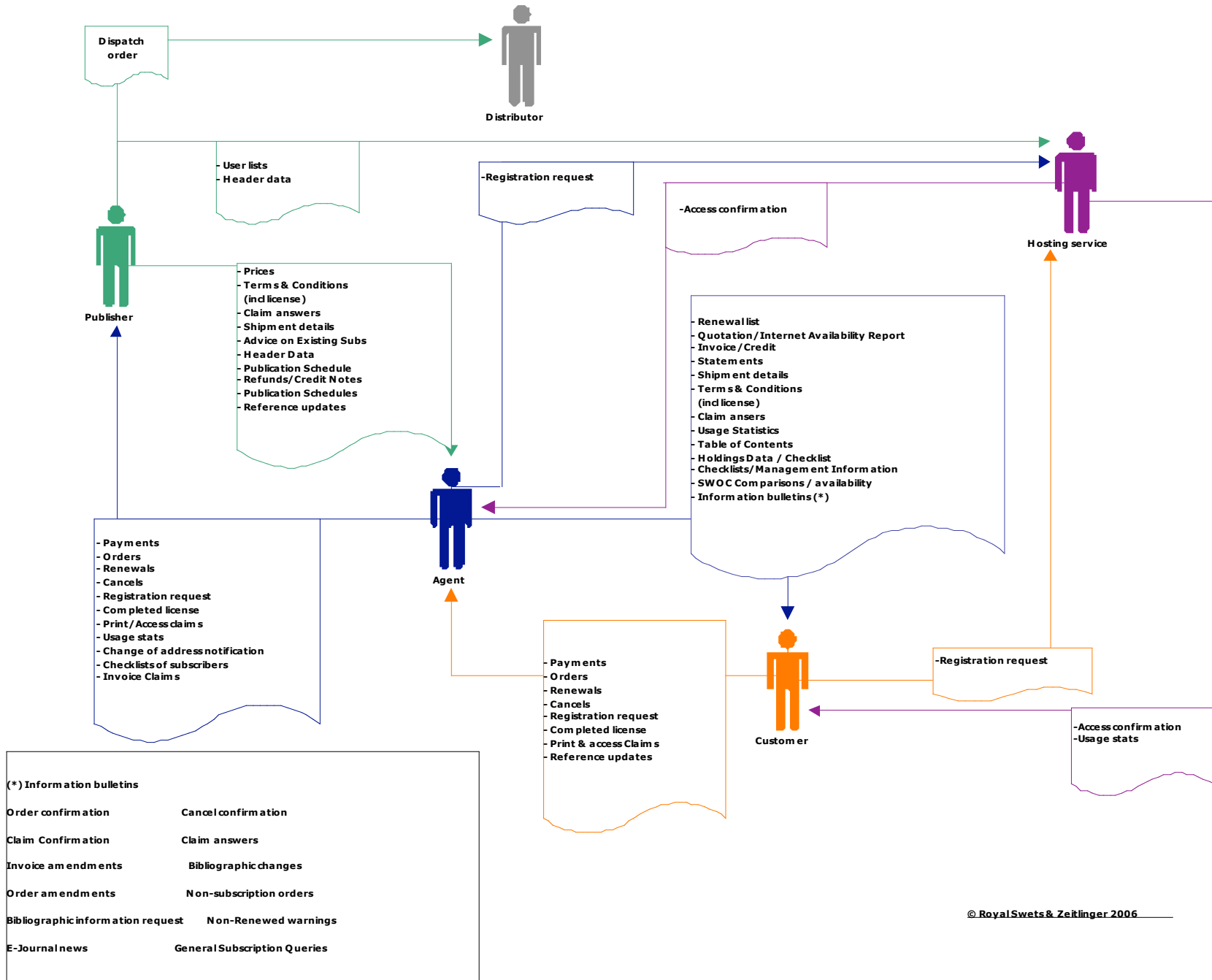
Prestito simultaneo: possibile (via DRM) vs impossibile

Prestito interlibrario: possibile in entrambi i casi

Monitoraggio: completo vs statistiche sui prestiti

Conservazione: backup e migrazione vs restauro

Accesso: 24x7 vs solo in orario d'apertura



Costi: biblioteche vs dischi



Per libro/anno: \$3-8

Per unità di personale: 7,000-20,000 libri



Per megabyte-anno: meno di un centesimo

Per unità di personale: terabytes (milioni di libri)

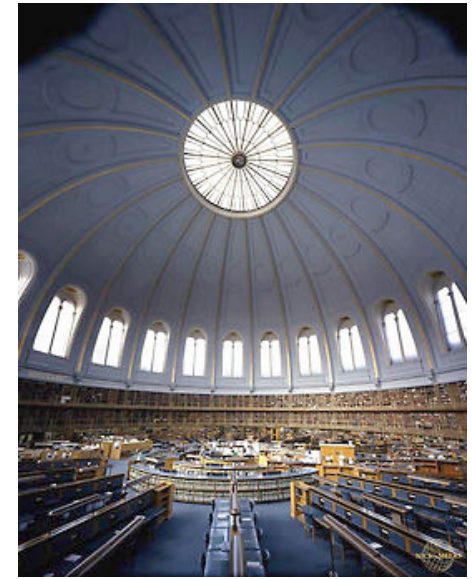
Biblioteche vecchie e nuove



gallica
la bibliothèque numérique



Bibliothèque
nationale
de France



Google
Scholar

ciancarini

Search

[Advanced Scholar Search](#)
[Scholar Preferences](#)
[Scholar Help](#)

Scholar

Results 1 - 10 of about 2,350 for ciancarini. (0.08 seconds)

[Agent-Oriented Software Engineering: The State of the Art](#)

M Wooldridge, P **Ciancarini** - [Cited by 178](#) - [Web Search](#)

Abstract. Software engineers continually strive to develop tools and techniques to manage the complexity that is inherent in software systems. In this ...

[AOSE, 2000](#) - [springerlink.com](#) - [kingston.ac.uk](#) - [science.unitn.it](#) - [cis.upenn.edu](#) - [all 30 versions »](#)

[The Concurrent Language, Shared Prolog](#)

A Brogi, P **Ciancarini** - [Cited by 108](#) - [Library Search](#) - [Web Search](#)

Shared Prolog is a new concurrent logic language. A Shared Prolog system is composed of a set of parallel agents that are Prolog programs extended by a ...

[ACM Transactions on Programming Languages and Systems, 1991](#) - [portal.acm.org](#) - [cs.unibo.it](#) - [csa.com](#) - [all 5 versions »](#)

[\[PDF\] Coordinating Multiagent Applications on the WWW: A Reference Architecture](#)

P **Ciancarini**, R Tolksdorf, F Vitali, D Rossi, A ... - [View as HTML](#) - [Cited by 88](#) - [Web Search](#)

Authors: Paolo **Ciancarini** Dept. of Computer Science, Univ. of Bologna, Via Mura A. Zamboni, 7, I-40127 Bologna, Italy. mailto:cianca@cs.unibo.it

[TSE, 1998](#) - [cs.unibo.it](#) - [cs.unibo.it](#) - [ieeexplore.ieee.org](#) - [portal.acm.org](#) - [all 7 versions »](#)

[Coordination models and languages as software integrators](#)

P **Ciancarini** ... - [Cited by 59](#) - [Web Search](#)

(1) Coordination entities. These are the entity types that are coordinated. These could be Unix-like processes, threads, concurrent objects and the ...

[ACM Computing Surveys, 1996](#) - [portal.acm.org](#) - [portal.acm.org](#)

[Interaction abstract machines](#)

JM Andreoli, P **Ciancarini**, R Pareschi - [Cited by 60](#) - [Web Search](#)

... Paolo **Ciancarini**, ... Paolo **Ciancarini**, Coordination models and languages as software integrators, [ACM Computing Surveys \(CSUR\)](#), v.28 n.2, p.300-302, June 1996. ...

[Research directions in concurrent object-oriented ...](#), MIT Press Cambridge, MA, USA, 1993 - [portal.acm.org](#) - [portal.acm.org](#)

[\[PDF\] Distributed Programming with Logic Tuple Spaces](#)

P **Ciancarini**, TR UBLCs, PPS Donato - [View as HTML](#) - [Cited by 69](#) - [Web Search](#)

93-26 AxiomatisingSTBisimulationEquivalence.N.Busi,R.vanGlabbeek,R.Gorrieri, December1993. 93-27 ATheoryofProcesseswithDurationalActions.R.Gorrieri,M. ...

[New Generation Computing, Ohmsha Tokyo, Japan, Japan, 1994](#) - [cs.unibo.it](#) - [cs.unibo.it](#) - [portal.acm.org](#) - [csa.com](#) - [all 5 versions »](#)

Gallica

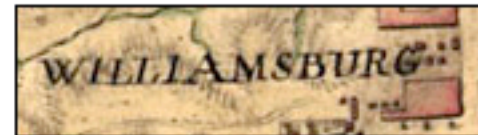
- Bibliothèque Numérique Nationale de France
- 800K documenti, di cui 25% in modo testo
- Formati TIFF e PDF
- Focalizzata sui testi francesi, ma non solo
- Oltre un milione di accessi al mese

Library of Congress - American Memory

Browse Collections by Topic

[Advertising](#)[African American History](#)[Architecture, Landscape](#)[Cities, Towns](#)[Culture, Folklife](#)[Environment, Conservation](#)[Government, Law](#)[Immigration, American Expansion](#)[Literature](#)[Maps](#)[Native American History](#)[Performing Arts, Music](#)[Presidents](#)[Religion](#)[Sports, Recreation](#)[Technology, Industry](#)[War, Military](#)[Women's History](#)[➔ More browse options](#)[➔ List all collections](#)

Collection Highlights



[The Rochambeau Map Collection](#)

Revolutionary War maps.



[The James Madison Papers](#)

Insight into the Father of the Constitution.

Today in History May 10



Today In History mines the American Memory historical collections to discover what happened in American history today...and every day.

Teachers

Use American Memory in the classroom.

Ask a Librarian

Get help from an expert.

ACM Digital Library

- Collezione di milioni di citazioni e di articoli di riviste e conferenze accademiche su scienze informatiche
- Una citazione consiste di
 - Titolo, autore, dati di pubblicazione
 - Abstract
 - Articoli citanti questo articolo
 - Riferimenti di questo articolo ad altri articoli
 - Parole chiave di classificazione
 - DOI, ovvero URL che denota il testo sul sito dell'editore

Articolo del 2002 trovato dopo una ricerca

Web of Science - Web of Science
Full Record
Title: Family environment of children and adolescents with bipolar parents
Author(s): Cheng TD, Blasey C, Lister TA, Stamer H
Subject Category: Clinical Neurology, Neuropsychology, Psychiatry
Article 2002

Citanti (futuro)

Web of Science - Web of Science
Full Record
Title: Family environment of children and adolescents with bipolar parents
Author(s): Cheng TD, Blasey C, Lister TA, Stamer H
Subject Category: Clinical Neurology, Neuropsychology, Psychiatry
Article 2003

Citazioni (passato)

Altri articoli correlati

← Citanti ←

Web of Science - Web of Science
Full Record
Title: Family environment of children and adolescents with bipolar parents
Author(s): Cheng TD, Blasey C, Lister TA, Stamer H
Subject Category: Clinical Neurology, Neuropsychology, Psychiatry
Article 2001
Article 2000
Article 1998
Article 1993

Web of Science - Web of Science
Full Record
Title: Family environment of children and adolescents with bipolar parents
Author(s): Cheng TD, Blasey C, Lister TA, Stamer H
Subject Category: Clinical Neurology, Neuropsychology, Psychiatry
Article 2004
Article 2002
Article 1999
Article 1994

Integratore di DL: ISI Web of Science

- Un **integratore** di DL offre accesso a pagamento a molteplici biblioteche digitali
- ISI è uno dei più grandi integratori
- Dà accesso a milioni di schede e relative pubblicazioni in molti campi scientifici e umanistici (migliore per questi ultimi)

Aziende per biblioteche digitali

Aziende e-book: atRandom, iPublish, MightyWords

Aziende DL (Digital Libraries): Questia, Netlibrary, e-brary.

Le aziende e-book vendevano singoli documenti digitali (e-book) da leggere su dispositivi speciali: mercato oggi fallito

Le aziende DL vendono l'accesso a contenuti specializzati, di solito su **abbonamento**

Esistono anche altri modelli di editoria digitale (es. *self publishing, print on demand, ecc.*)

q u e s t i a

1.500.000 titoli scienze umanistiche e sociali

67.000 libri completi da 250 case editrici

155.000 articoli scientifici

185.000 articoli di riviste

1.100.000 articoli di quotidiani

Target: studenti e studiosi \$19.95/mese, \$100/anno

Fondi di startup: \$135M

Quanto costano le biblioteche?

Reed Elsevier: offre 1200 riviste scientifiche per un prezzo "bundled". Ha raddoppiato i profitti sci-med a \$2.33B, 40% di profitto sul 25% del business.

Università Cornell: abbonata a 930 riviste, \$1.7M, 20% per 2% delle riviste, molte lette raramente

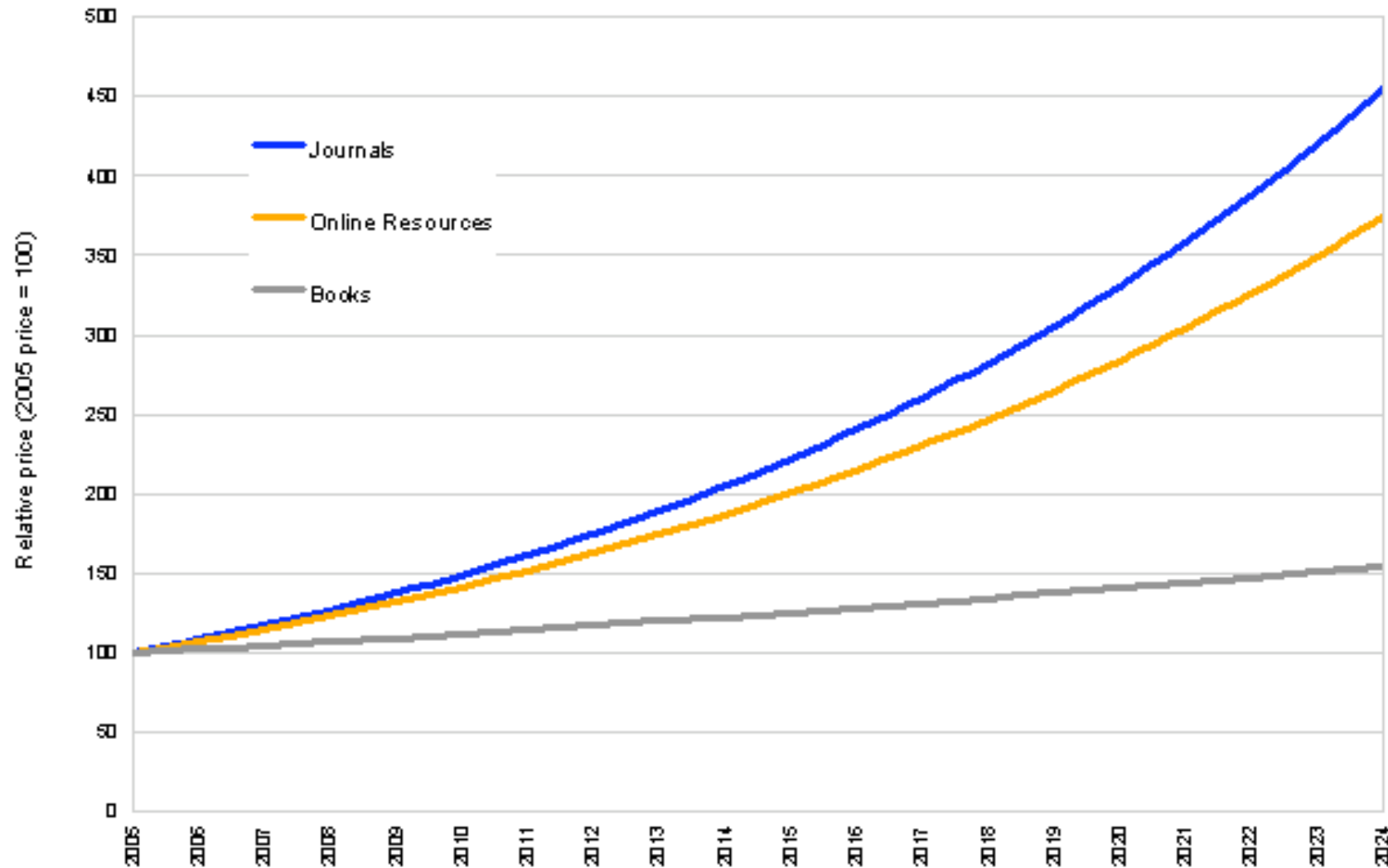
Università Oldenburg: solo 20 articoli letti all'anno in fisica nucleare, al costo di \$1.250 per articolo

Università Harvard taglia 100 abbonamenti

Fonte: C.Goldsmith: "Reed Elsevier Feels Resistance to Web Pricing", *Wall Street Journal*, 19/1/2004

Alcuni trend

Cost-increase projections for library materials



Costruire una biblioteca digitale

- Costruire una biblioteca digitale richiede:
 - Investire risorse umane e finanziarie
 - Classificare i documenti digitali
 - Proteggere i diritti degli autori
- La classificazione si basa su **metadati**
- La protezione si basa su **sistemi DRM**

Sfide

- Finanziare la digitalizzazione dei materiali
- Garantire l'integrità degli archivi
- Fare accordi con gli editori (es. Open Access)
- Sviluppo di strategie internazionali
- Sviluppo di cataloghi internazionali

Lecture

- Ippolita, *The Dark Side of Google*, 2007
- Jacso, Google Scholar revisited, 2008

Riferimenti

- Battelle, *The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture*, 2005
- Levoie et al., *Anatomy of Aggregate Collections: The Example of Google Print for Libraries*, 2005
- Bakkalbasi et al., *Three options for citation tracking: Google Scholar, Scopus and Web of Science*, 2006
- P.Cotoneschi e G.Pepeu, *Raccomandazioni per lo sviluppo dell'editoria elettronica negli Atenei Italiani*, Univ. Firenze, 2006
- USCO, *Report on Orphan Works*, 2006

Siti

- www.google-watch.org
- cyberlaw.stanford.edu
- battellemedia.com
- www.dlib.org/dlib
- Biblioteca dell'Università di Bologna almadl.cib.unibo.it
- Library of Congress www.loc.gov
- Digital Library Federation www.diglib.org

