

Internet federated search engines for bookseller databases: a comparative evaluation

Paul.Nieuwenhuysen @ vub.ac.be

Pleinlaan 2, Brussels, Belgium



Prepared for a live presentation at the

Fourth Shanghai International Library Forum = SILF2008

20-22 October 2008 <http://www.libnet.sh.cn/silf2008>



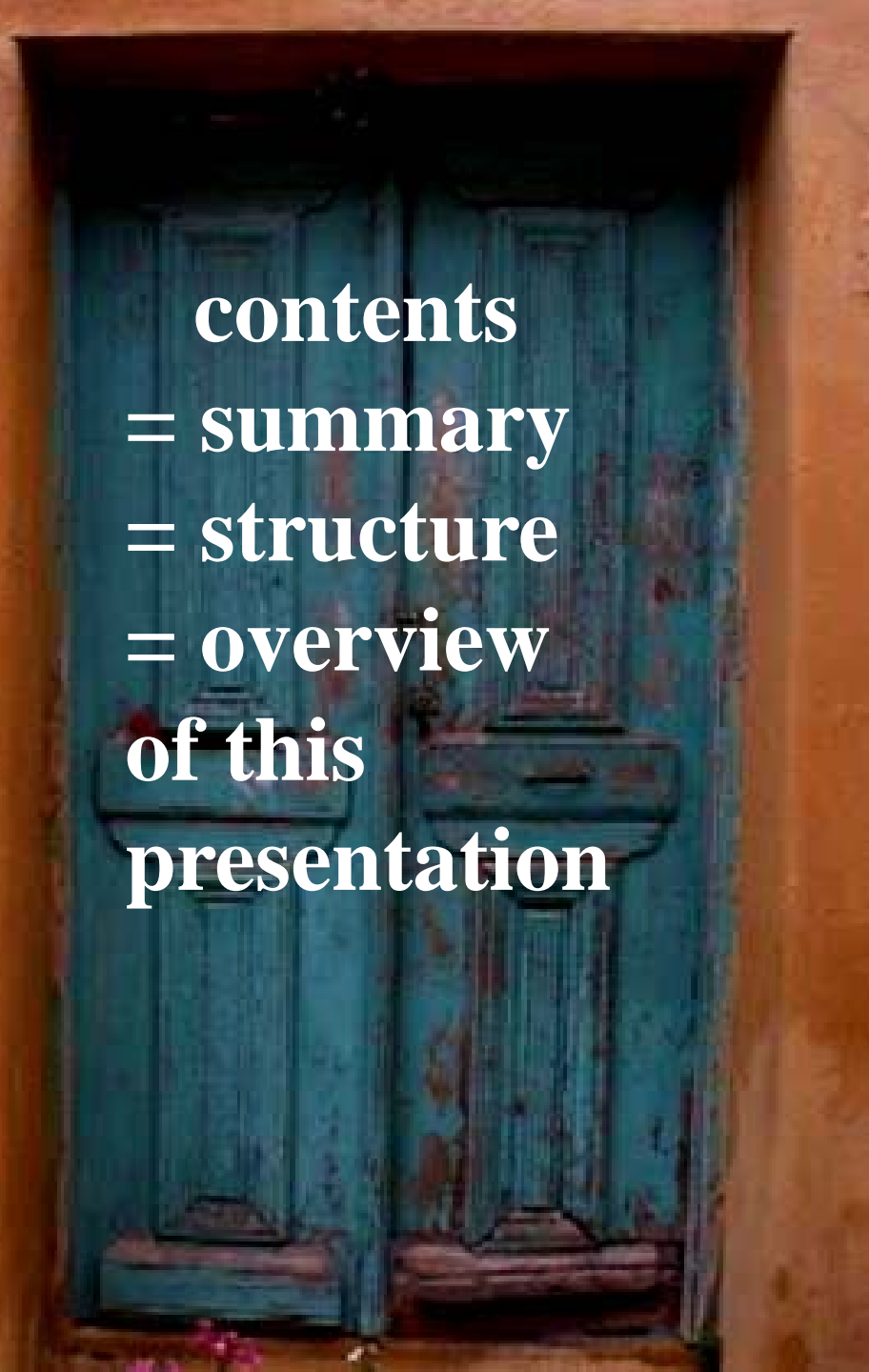
SILF 2008
The 4Th Shanghai International Library Forum



**Full text published as a paper in the
Proceedings of the
Fourth Shanghai International Library Forum = SILF2008
20-22 October 2008**

**These slides should be available from the WWW site
<http://www.vub.ac.be/BIBLIO/nieuwenhuysen/presentations/>
(note: BIBLIO *and not* biblio)**





contents
= summary
= structure
= overview
of this
presentation

- **Introduction (long)**
- **Problem statement**
- **Methodology (briefly)**
- **Results**
- **Discussion**
- **Conclusion !**



Internet federated search engines for bookseller databases

Introduction



Internet with WWW and printed books



- Since a few years, Internet with the WWW have become the primary information source for many people.
- However:
 - » A lot of information is still distributed only in the form of printed books
 - » The content of old printed books can still be interesting.
 - » The content of most printed books is (still) not available on the Internet.



Public access book search systems: introduction



- **Most general WWW search engines do NOT allow you to find out about the existence of books that may be interesting for you, at least not in a systematic and efficient way.**
- **So, specific search tools to find books can be useful.**



Public access book search systems: which one to use?



- **For years, information on books was mainly offered by expensive subscription-based providers of bibliographic information.**
- **Since a few years, the Internet with the WWW provide a key to unlock many possibilities to find bibliographic information.**
- **Which book search system should be preferred for particular applications is not clear for most end-users and not even for librarians and other information experts.**



Public access book search systems: an overview



- (Databases by publishers.)
- Fee-based databases by commercial providers
- Online public access catalogue databases of libraries
- Databases of scanned printed book pages
- **Catalogue databases by bookshops**
- **Meta-systems = systems that exploit combinations of the databases provided by bookshops !**



Public access book databases by commercial producers



- To find currently available books, some databases assembled by commercial producers can be interesting.
- Example: *Global Books in Print*
- These databases offer formal descriptions of books, prices of the books, short descriptions of the contents with subject terms...
- However, access to such a database is not free of charge and can be expensive (in comparison with alternatives).



Public access book databases provided by bookshops



- **To find currently available books, the bibliographic databases assembled by big bookshops are interesting.**
- **Several offer a good coverage.**
- **Many are accessible free of charge.**
- **The added price information can be useful for the acquisition and accounting department of a library or if an individual user wants to buy a book.**
- **Some provide a current awareness service, also free of charge.**
- **Take into account delivery costs: postage + tax**

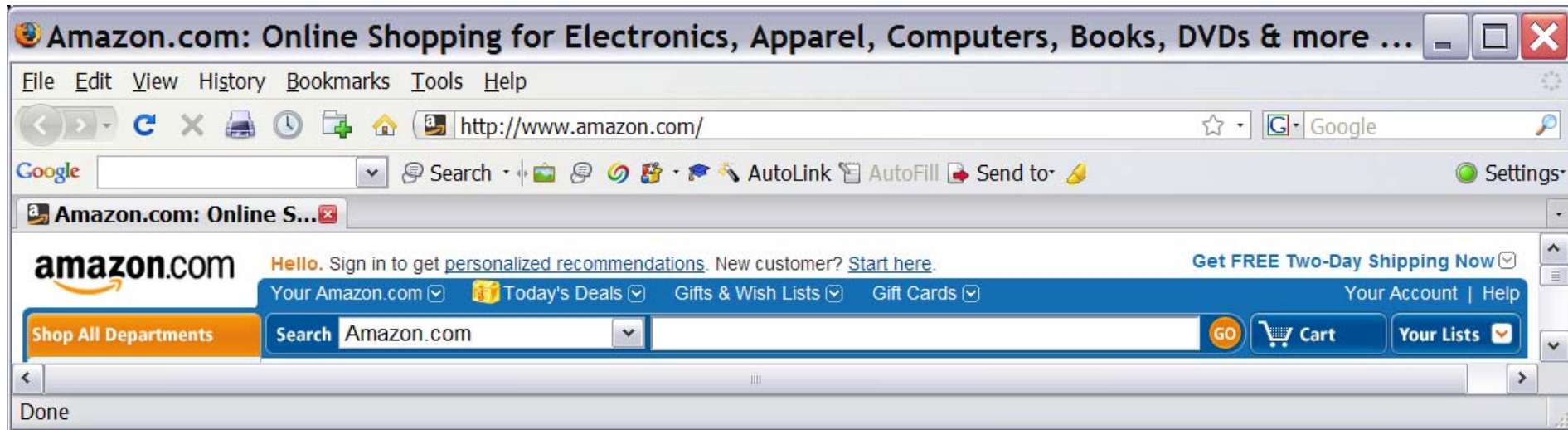


Book databases accessible free of charge: examples



- Amazon.com (US):
<http://www.amazon.com/>
[accessed 2008]
- This company offers also different, more local versions that offer books in other languages, such as
<http://www.amazon.co.uk/>
<http://www.amazon.fr/>
- note: amazon, NOT amazone

Book databases accessible free of charge: examples

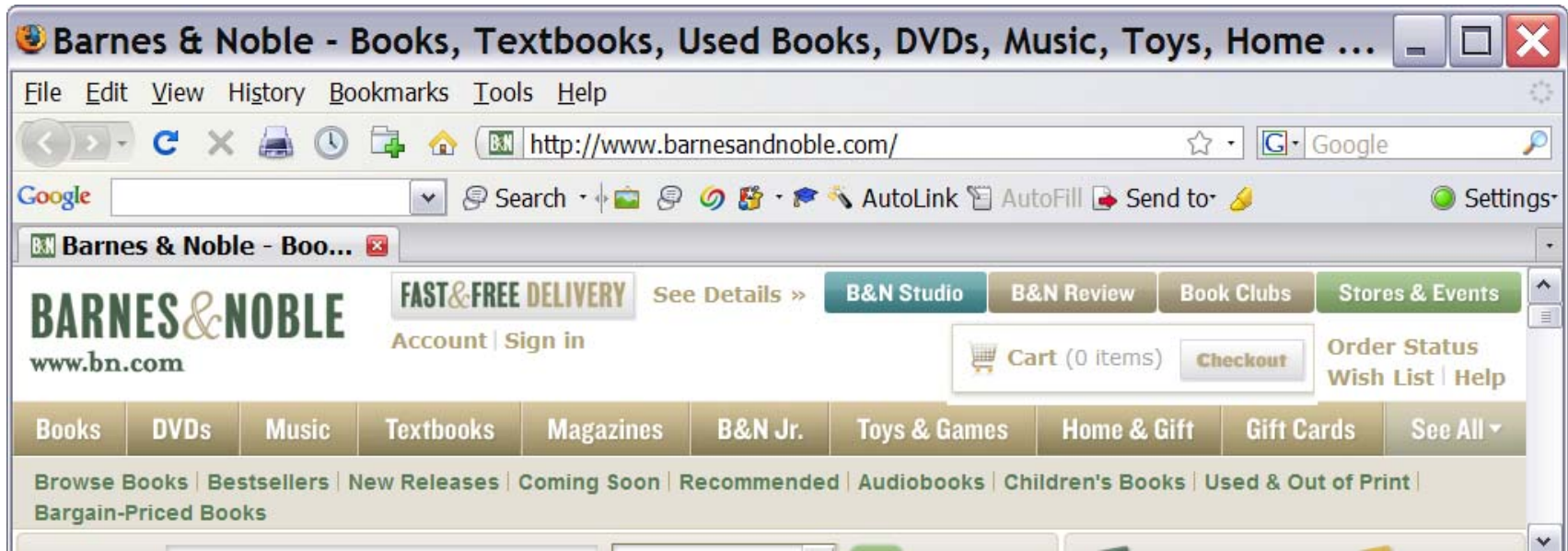


Book databases accessible free of charge: examples



- *Barnes and Noble (US):*

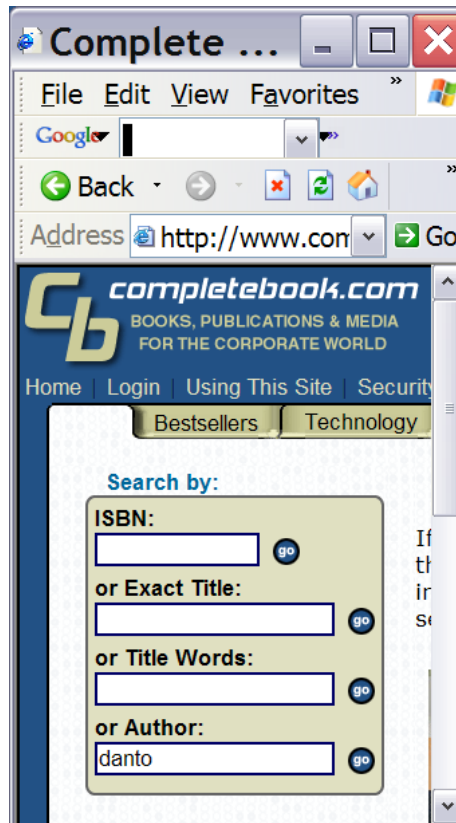
<http://www.barnesandnoble.com/> or <http://www.bn.com/>
[accessed 2008]



Book databases accessible free of charge: examples



- <http://www.completebook.com/cbmsi/bookaction.do>
[accessed 2008]



Free public access meta-search systems: introduction



- **Some systems allow searching through the catalogue databases of several dealers/shops, in one and the same action, simultaneously.**
- **Furthermore, the results from the various shops are presented in such a way that the user can easily compare some aspects of the books (for instance: used or new, price set by the owner/dealer, delivery costs to the user's country)**
- **We can call these “meta-catalogues” = “meta-search-systems”**



Free public access meta-search systems: types



We can make the following distinction between various types of meta-systems for searching:

- 1. Database resulting from merging several existing smaller databases = aggregator database**

In this case of books:

multi-dealer database = “listing service”

- 2. Federated search system
= cross-database search system**



Free public access meta-search systems: multi-dealer databases



- **Multi-dealer database = database obtained after merging of several existing catalogue / inventory databases, which are managed and updated by individual dealers / shops / sellers.**
- **Such a system can include from a few to more than 10000 shops / dealers.**
- **Some of these systems offer a potential buyer a link to the real bookshops that have the book for sale. Others allow a potential buyer to buy immediately on their site and communicate afterwards with the real dealer.**



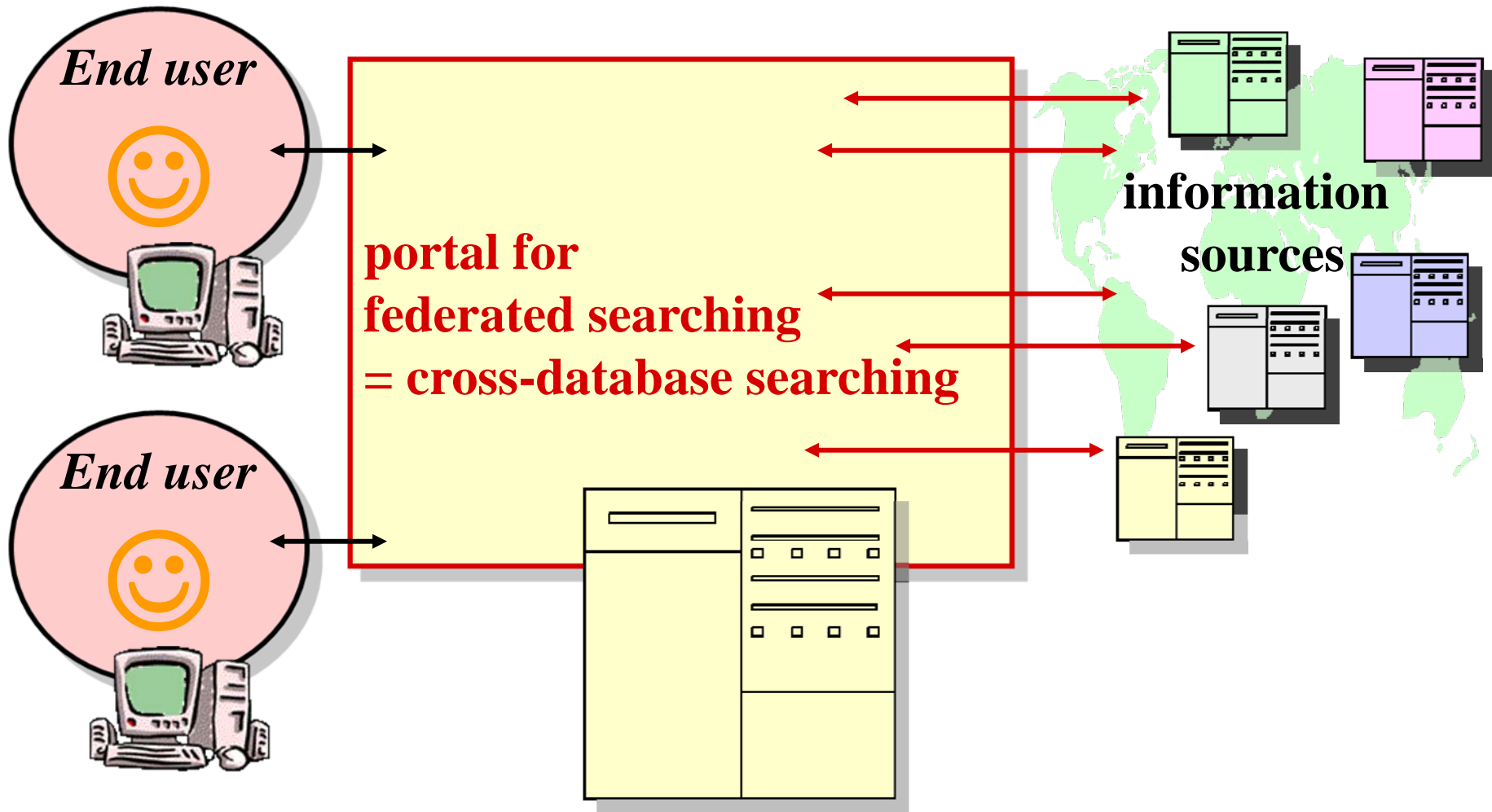
Free public access search systems: federated search systems



- **Federated search system
= cross-database search system**
- **These are gateway systems that allow you to search in one action through several existing book databases.**
- **They transmit your query to the real databases, they collect and merge the results, and show them finally in a useful way.**



Federated searching: general scheme



Free public access search systems: federated search systems



- **Each of the searched databases can be**
 - » a catalogue database managed by the owner/dealer/shop/seller, as well as
 - » a multi-dealer database



Free public access meta-search systems: comparison



- **Multi-dealer databases react faster than federated search systems (in most cases).**

» Explanation:

They do not need several simultaneous Internet connections &

they do not have to merge raw intermediate results into the result that is finally shown to the user.



Free public access federated search systems: comparison



- **Federated search systems offer a higher coverage than multi-dealer databases (in most cases).**

» Explanation:

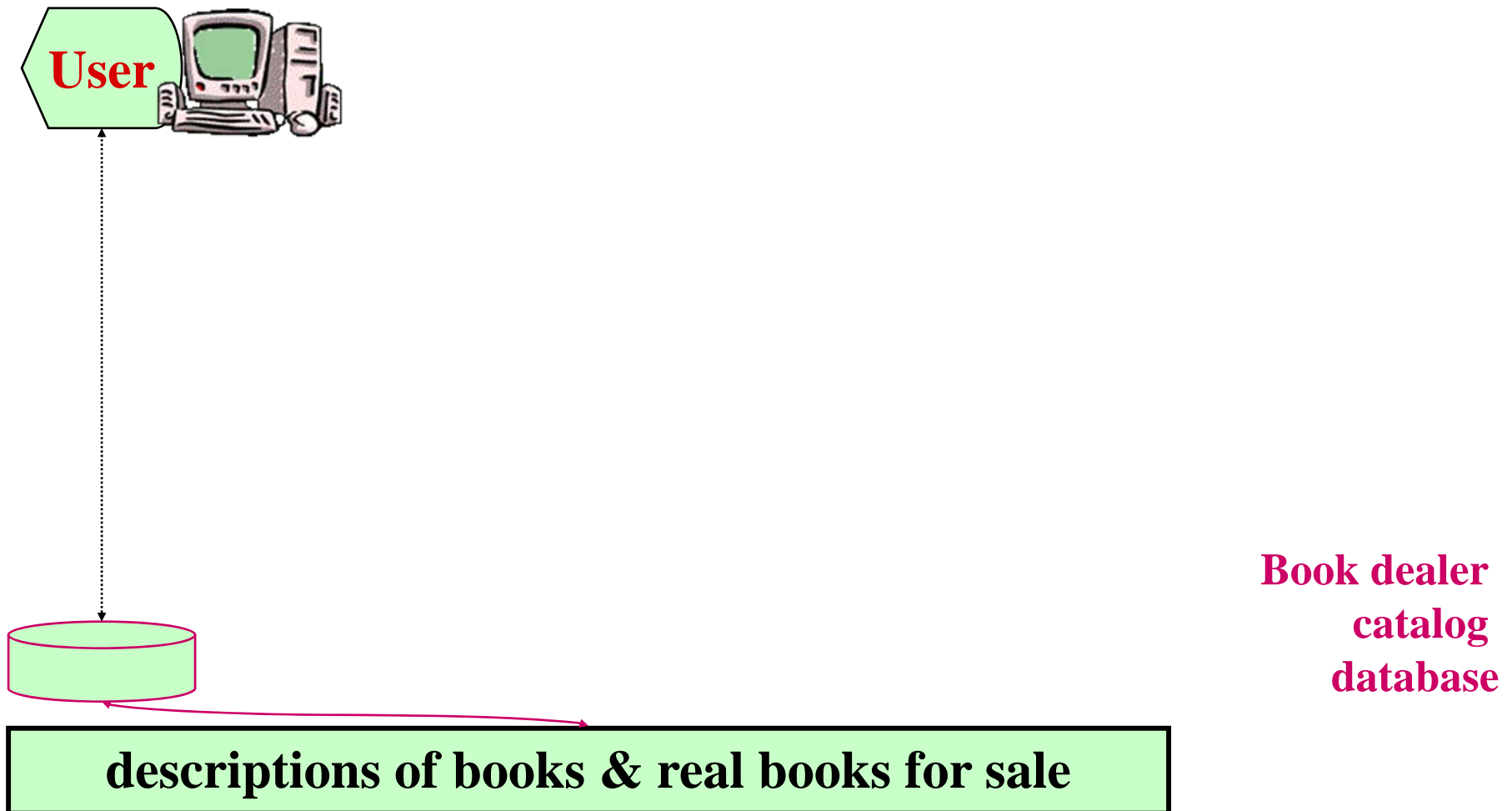
They can exploit several multi-dealer databases in one search action.

In 1 search, they can cover more than 100 million descriptions of physical books

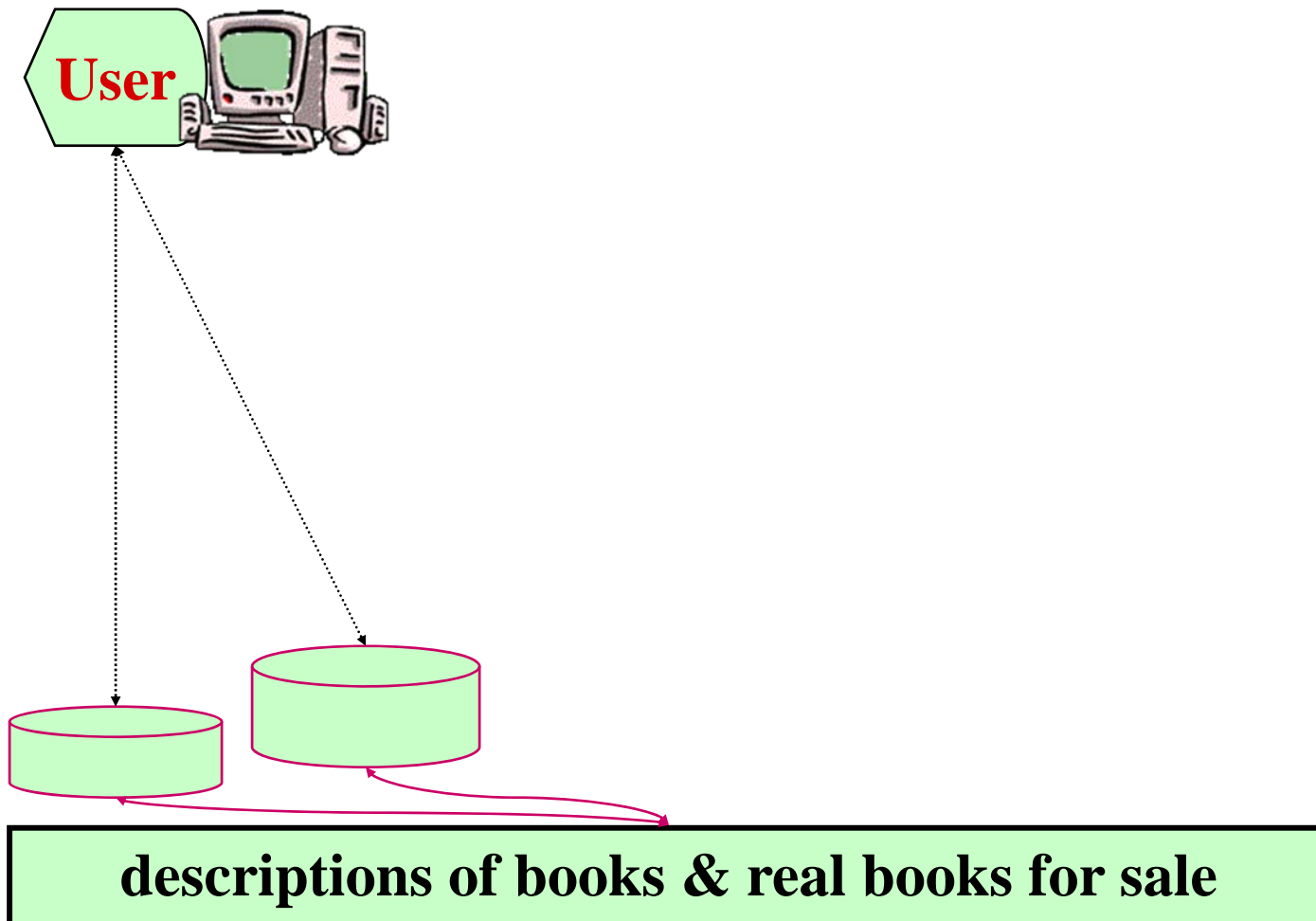
= couples of book and dealer (not book titles).



Search systems for books that are made available by dealers: scheme

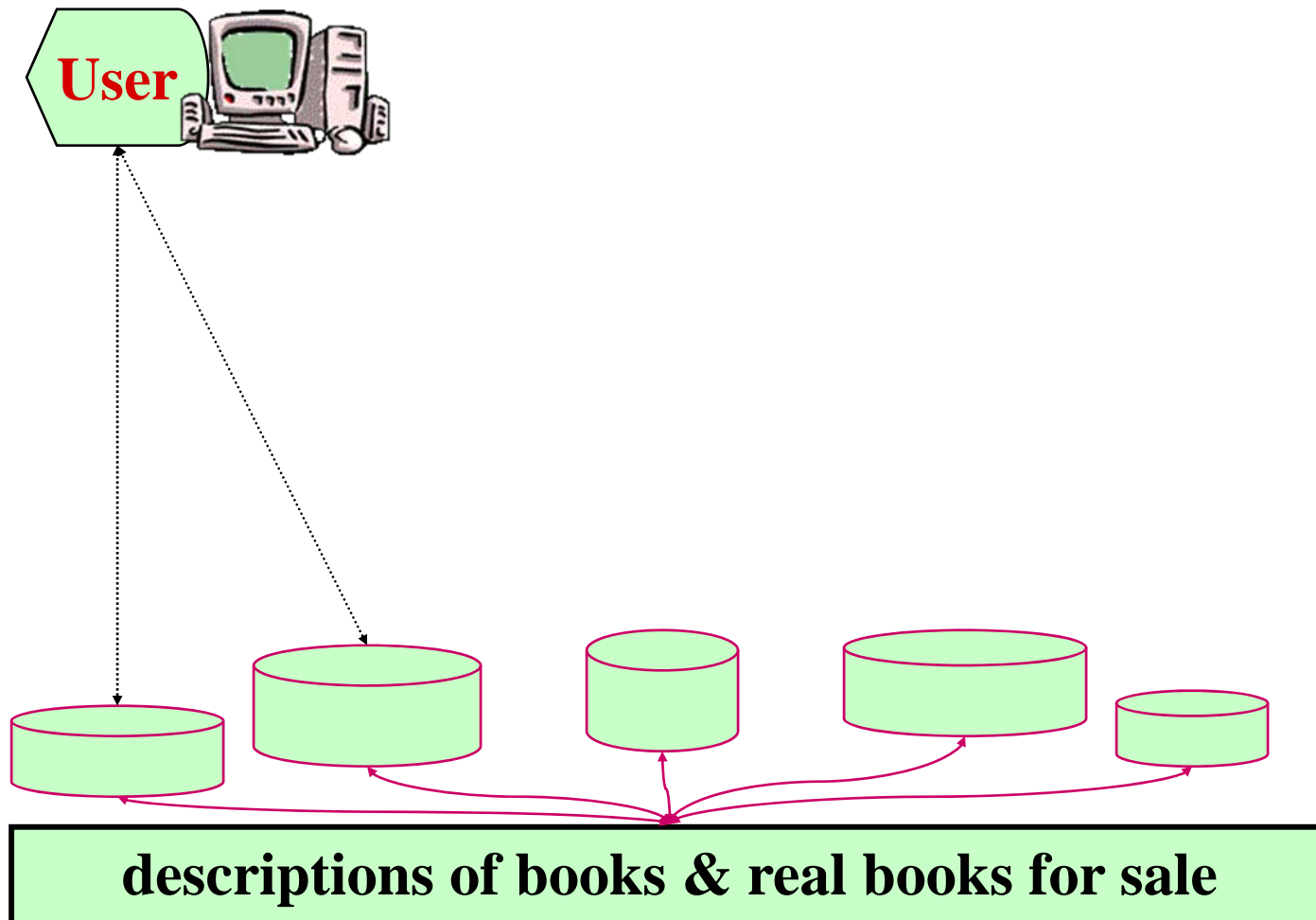


Search systems for books that are made available by dealers: scheme



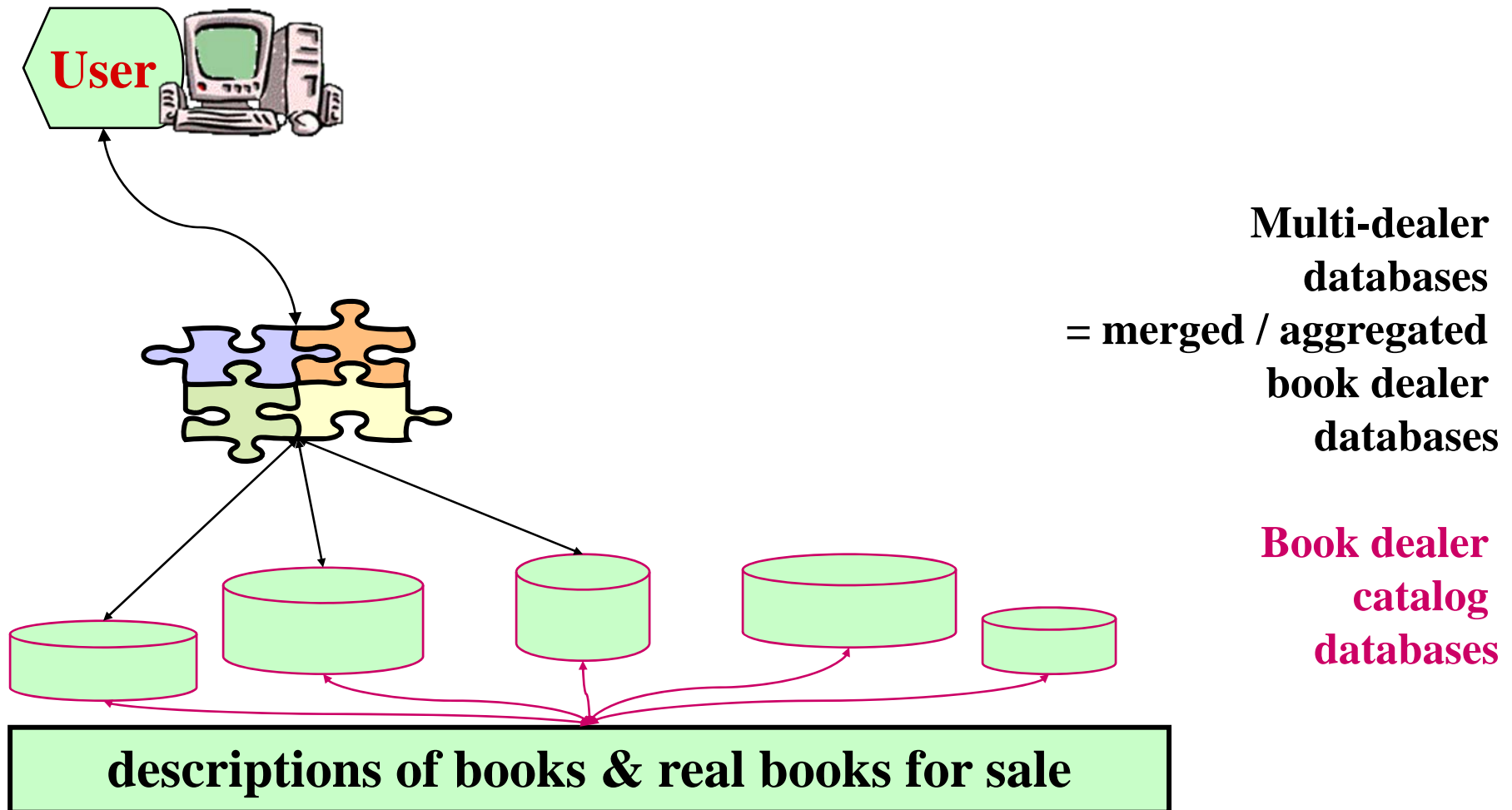
**Book dealer
catalog
databases**

Search systems for books that are made available by dealers: scheme

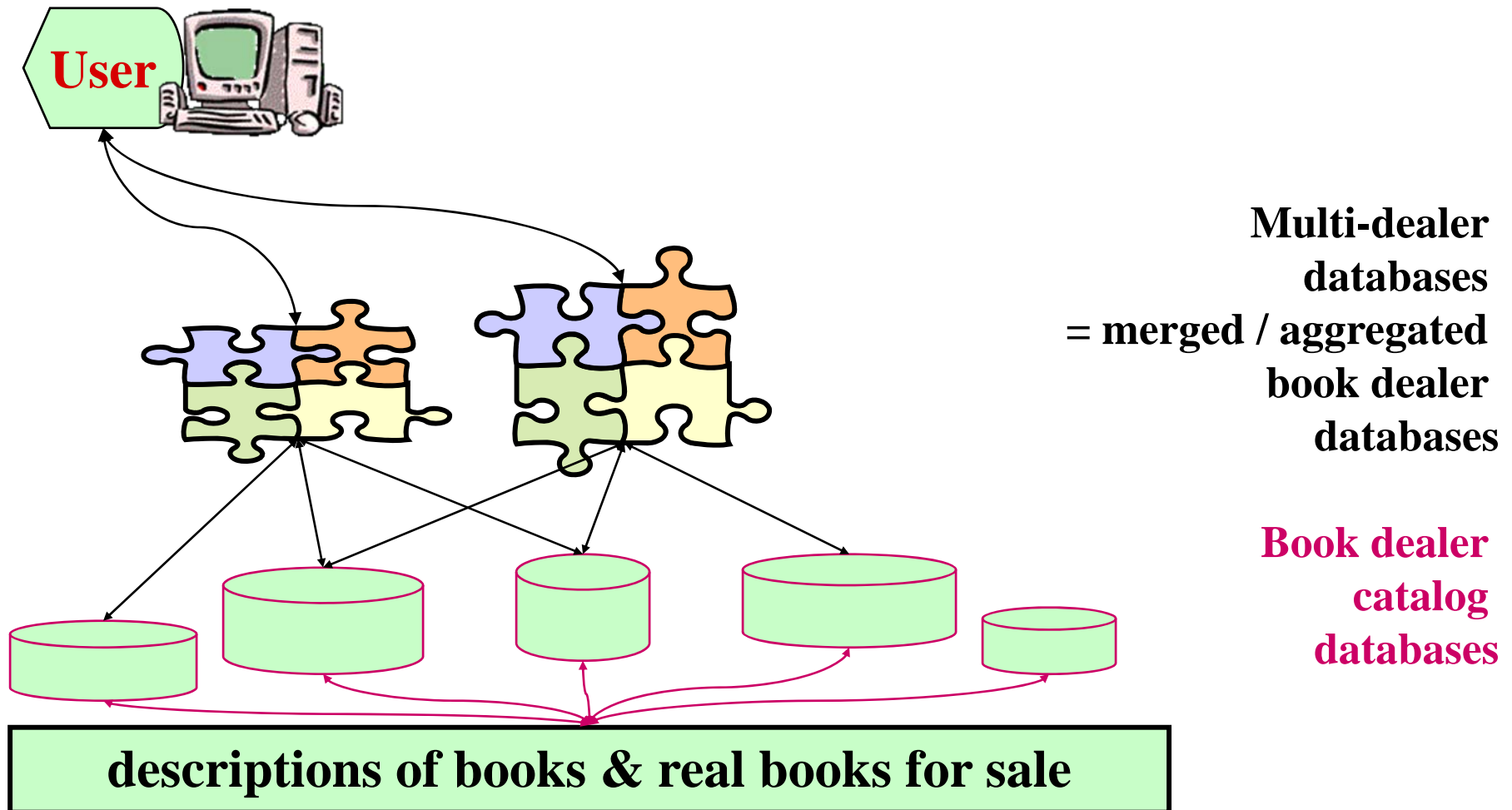


**Book dealer
catalog
databases**

Search systems for books that are made available by dealers: scheme



Search systems for books that are made available by dealers: scheme



Free public access multi-dealer book databases: examples



- <http://www.abebooks.com/>
[accessed 2008]
- Covers > 10 000 bookshops.
- The company has been acquired by *Amazon* in 2008.

The screenshot shows the AbeBooks.com website. The browser window title is "New & Used Books, Textbooks, Rare & Out of Print Books...". The address bar shows "http://www.abebooks.com/". The website header includes the AbeBooks.com logo and navigation links: "sign on", "my account", "basket", and "help?". Below the header is a red navigation bar with links: "Advanced Search", "Browse", "Bookstores", "Community", "Sell Books", and "Textbooks". A blue banner below the navigation bar reads "Textbooks at AbeBooks" and "CHEAP FAST EASY" with a "Learn More" link. The main content area features a search form with fields for "Author:", "Title:", "Keyword:", "what's this?", and "ISBN:". A "Find Book" button and a "More Search Options" link are at the bottom of the form. On the left side, there is a "Browse for Books" section with a "CURRENTLY FEATURED" banner for "Creative Juices" and a "Most Popular Subjects" list: 1. Fiction, 2. History, 3. Antiques.

Free public access multi-dealer book databases: examples



- <http://www.alibris.com/>
[accessed 2008]



Free public access multi-dealer book databases: examples



- <http://www.antiqbook.com/books/>
[accessed 2008]
(NOT www.antiqbooks.com)

“ANTIQBOOK unites more than 400 independent booksellers from all over the world. You can use our search pages for a free search of over 3.8 million books, and order them directly from your bookseller. Strong areas in our database are books from European booksellers, many of them specialist antiquarian booksellers.

While ANTIQBOOK takes care that you can order safely from our booksellers we do not take part in their sales. We just bring you in touch with some of the finest booksellers in the world. You can order your books straight from the source, at their original prices and no hidden costs or markup fees.”

Free public access multi-dealer book databases: examples




Antiquarian books, used books, out of print books: Antiqbook - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.antiqbook.com/books/index.phtml?Language=en

Google Search

Antiquarian books, us...

 **ANTIQBOOK** Europe's Premier Antiquarian Booksite

Author Name: Title: >>

SPEEDSEARCH; AUTHOR AND/OR BEGINNING OF TITLE ONLY

Choose a language
English
Français
Deutsch
Nederlands

Smart Search Guide

Search Bookseller
By name
By specialization

Friendly advice: You can also [find our books at](#) *mareLibri*

Book search form

Author's last name:

Title:

Other keywords: ☐ Exact phrase

Author's first name (if needed):

... and to refine your search: ☐ First edition ☐ Signed
☒ All books ☐ Last 7 days

from booksellers in: first shown

Binding: ☐ Hardcover ☐ Softcover ☒ Any binding

Dust jacket: ☐ With dust jacket
☐ Books, but no ex library copies

Year of publication: OR century:

Nobel Prize for Literature 2007
Doris Lessing
[previous winners...](#)

Man Booker Prize 2007:
Anne Enright for *The Gathering*
[previous winners...](#)

Done

Free public access multi-dealer book databases: examples



- *Biblio*
- <http://biblio.com/>
[accessed 2008]



Free public access multi-dealer book databases: examples



- *Choosebooks*
- <http://www.choosebooks.com/>
[accessed 2008]

ChooseBooks.com - Advanced Search - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.choosebooks.com/showAdvancedSearch

ChooseBooks.com - A...

choosebooks.com MY ACCOUNT SHOPPING CART HELP DEUTSCH

HOMEPAGE ADVANCED SEARCH BROWSE ABOUT US

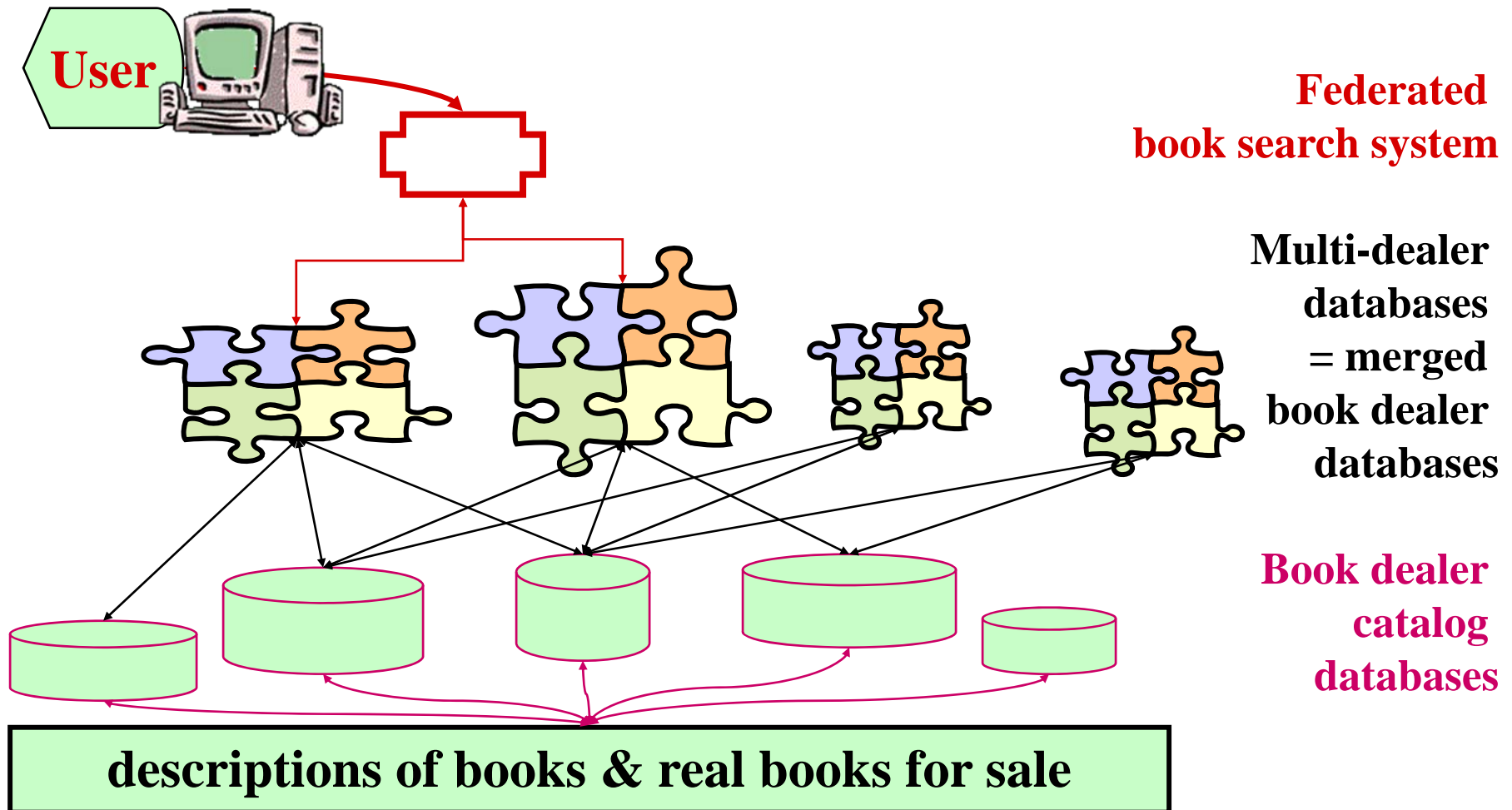
Advanced Search

Author	<input type="text"/>	Publisher	<input type="text"/>
Title	<input type="text"/>	Keyword	<input type="text"/>
Any Search Words	<input type="text"/>	ISBN	<input type="text"/>

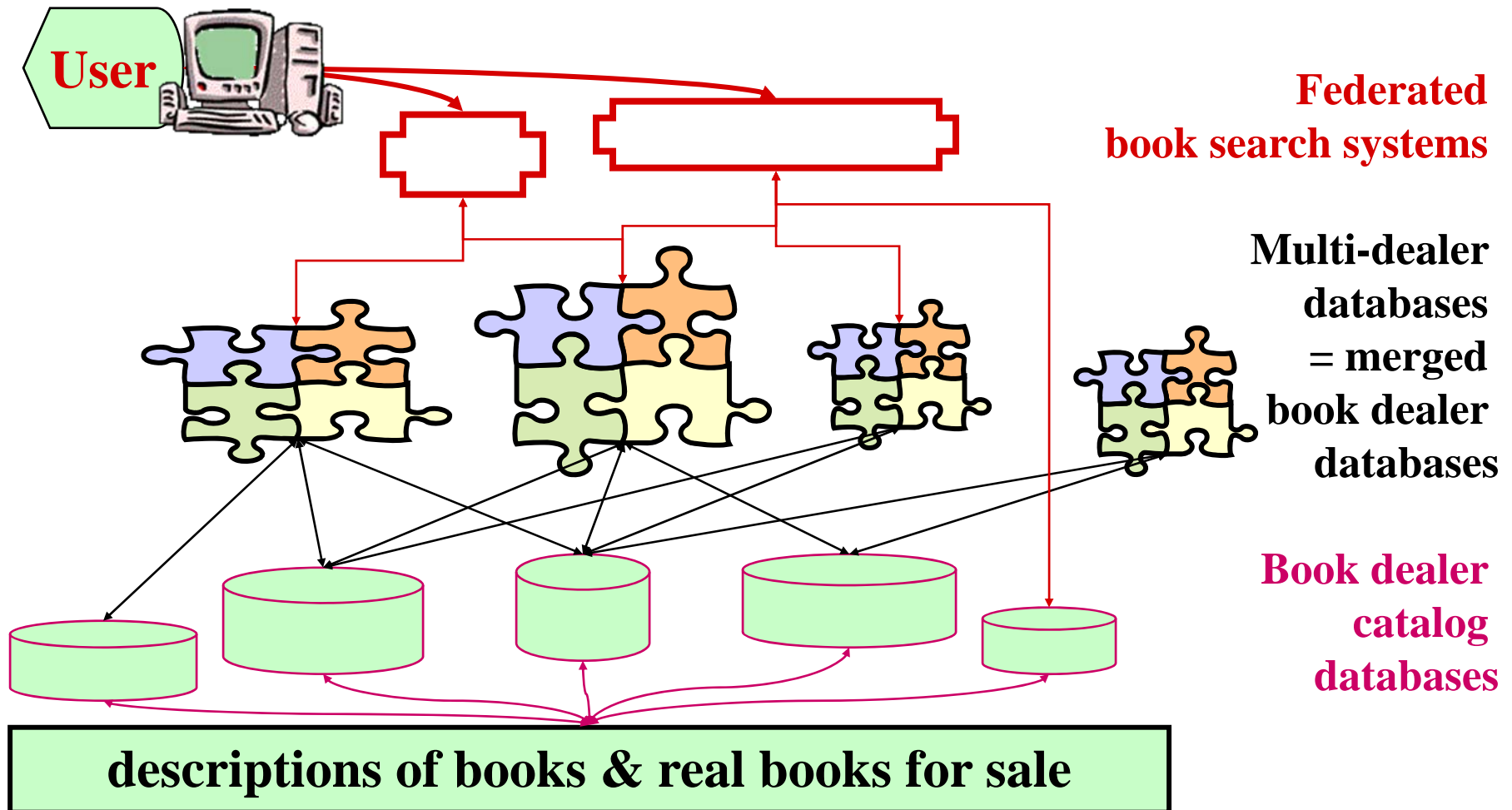
Year of publication	from <input type="text"/> to <input type="text"/>	Price	from <input type="text"/> to <input type="text"/>
Medium	All <input type="text"/>	Also show prices in	US Dollar <input type="text"/>
Country of seller	All <input type="text"/>	Only display entries	from the last -- <input type="text"/> days
Language of book	All <input type="text"/>	Sort:	-- <input type="text"/>

<input type="checkbox"/> First Edition	<input type="checkbox"/> Signed	Show	10 <input type="text"/> results per page
<input type="checkbox"/> Hard cover	<input type="checkbox"/> Soft cover	Retrieve up to	200 <input type="text"/> matches

Search systems for books that are made available by dealers: scheme



Search systems for books that are made available by dealers: scheme



Free public access federated search systems for books: examples



- **<http://www.addall.com/>**
[accessed 2008]
- **Covers many book dealer databases and multi-dealer databases, including unique databases that are not covered by competing search systems.**
- **Searches only new books;
to find used books, a companion system should be used.
This is inconvenient if the user is interested in both types of books.**

Free public access federated search systems for books: examples



AddALL book search and price comparison - Microsoft Internet Ex...

File Edit View Favorites Tools Help

Back Search Address http://www.addall.com/ Go

AddALL Book Search and Price Comparison
*Be smart: don't buy any book without comparing the price.
Save your TIME and MONEY with AddALL.com in your book shopping.*

[In Print Books](#) [Used & Out of Print Books](#) [Magazine](#) [Music](#) [Movie](#) [Help](#)

[FAQ/About us](#) [Recommend us](#) [Browse](#) [Memo](#) [Book Reviews](#) [Random Quotes](#)

HELP Topics

- [new books](#)
- [used books](#)
- [use a coupon](#)
- [Memo](#)
- [Printer friendly](#)
- [Customized search result](#)
- [No cookie option](#)
- [Match this title](#)
- [Refine it](#)
- [Search](#)

Looking for a book? Want a deal?
No problem **AddALL!**
Search and Compare among 40+ sites, 20,000 sellers, millions of books!

Find it

Search by: ISBN

Shipping Destination: == Belgium

U.S. State: (for sales tax) Select One

Display in: EURO

Find it Start Over

Other Comparison Sites

- [Out of Print Books](#)
- [Magazine](#)
- [Music CD/Tape](#)
- [Movie DVD/VHS](#)

Internet

Free public access federated search systems for books: examples



- ***Addall.com* allows searching by title OR author OR ISBN, but NOT by a combination of those fields.**
- **Can calculate the cost to ship/send a book to you, taking into account your country and currency.**

Free public access federated search systems for books: examples



- *Allbookstores*
- <http://www.allbookstores.com/>
[accessed 2008]

Free public access federated search systems for books: examples



- *BookButler*
- <http://nl.bookbutler.com/> and other URLs
[accessed 2007]
- Includes new and used books in 1 search action
- Can calculate the cost to ship/send a book to you, taking into account your country and currency.
- In our experience, *BookButler* gives on average less results than *AddAll*, *BookFinder*, *Campusi*.

Free public access federated search systems for books: examples



BookButler - Book Price Comparison - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://nl.bookbutler.com/ Amazon.com

Google Search AutoLink AutoFill Send to Settings

BookButler - Book Pri... Kopen in Mijn eBay: Mijn ...

BookButler

We try hard to find the best prices for you!

Search [Advanced Search](#)

Search by: ☒ **Keyword** ☐ Title ☐ Author ☐ ISBN / EAN ☐ Publisher

Catalog: ☐ Germany ☒ **Great Britain** ☐ United States ☐ France ☐ Netherlands ☐ Canada

Shipping to:

Currency:

[Add to Firefox™ Search](#) [?]

This site in: [Deutsch](#) - [English](#) - [Français](#) - [Nederlands](#) - [Español](#)

[Homepage](#) | [Advanced Search](#) | [Help & FAQ](#) | [About us & Contact](#) | [Search Tools](#) | **112** Bookstores

Done

Free public access federated search systems for books: examples



- *BookFinder*
- <http://www.bookfinder.com/>
[accessed 2008]
- Covers many book dealer databases and multi-dealer databases, including unique databases that are not covered by competing search systems.
- It is efficient that new and used books are searched in 1 action;
the results are presented in 2 columns: new | used.

Free public access federated search systems for books: examples



BookFinder.com: Search for New & Used Books, Tex...

File Edit View History Bookmarks Tools Help

BF http://www.bookfinder.com/?i Amazon.com

Google Search Settings

BF BookFinder.com: Sear...

BookFinder.com Search | About | Preferences | Interact | Help
1997-2007

125 million books. 1 search engine.

Compare prices on over 125 million books for sale
new • used • rare • out of print • international • textbooks

Search BookFinder.com

Author

Title

Language Book is written in

Begin search [Reset fields](#)

Options:

Type ☒ Any ☐ New ☐ Used/Out of Print

Features ☐ First edition ☐ Signed copy

Destination

Currency

[Show more options...](#)

Done

Free public access federated search systems for books: examples

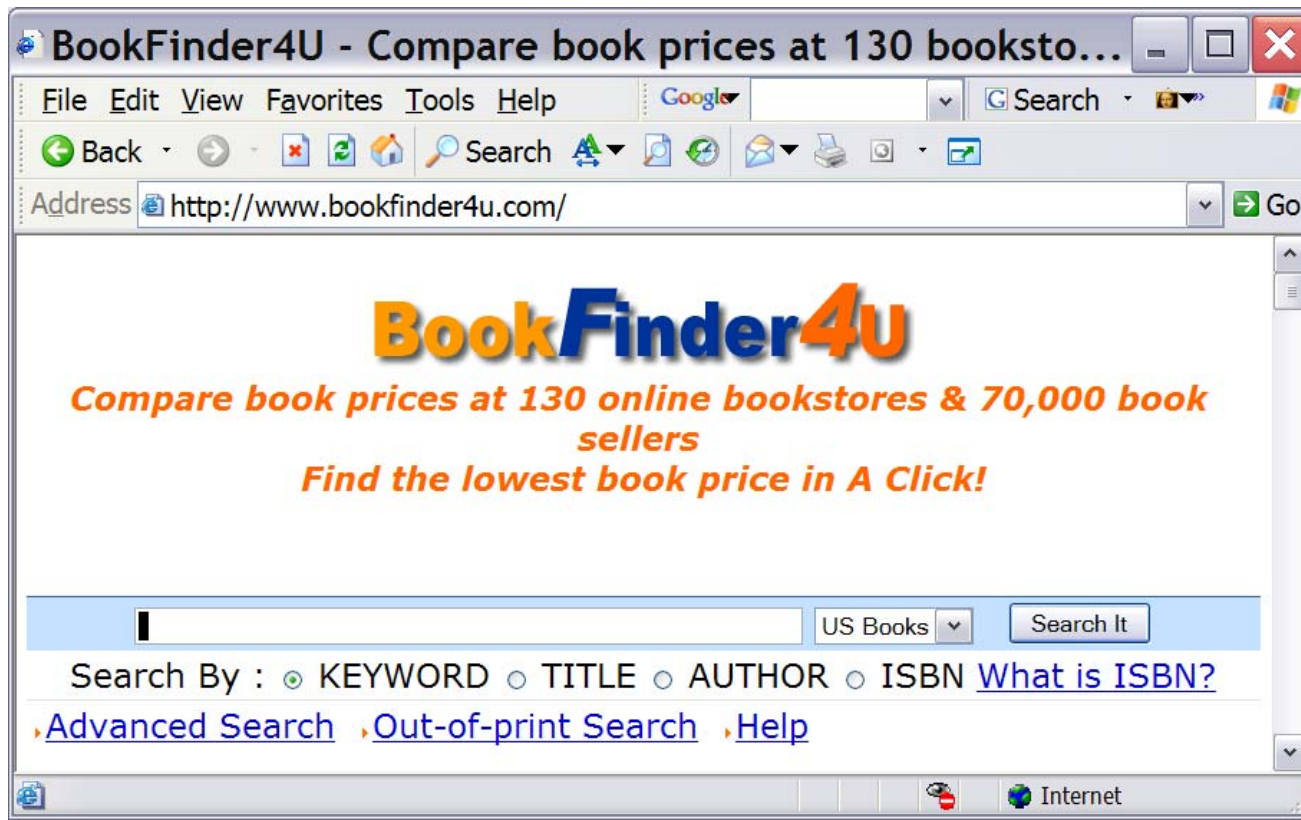


- *BookFinder* can calculate the cost to ship/send a book to you, taking into account your country and currency.

Free public access federated search systems for books: examples



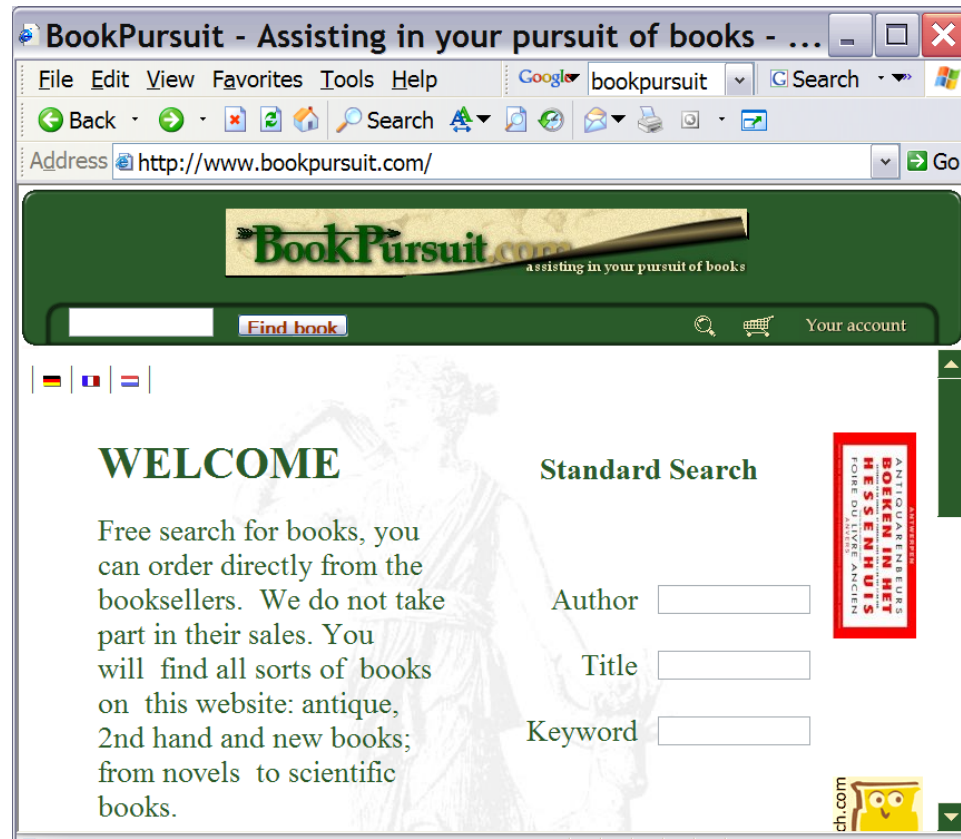
- <http://www.bookfinder4u.com/> [accessed 2007]



Free public access federated search systems for books: examples



- <http://www.bookpursuit.com/>
[accessed 2006]



Free public access federated search systems for books: examples



- <http://www.campusi.com/>
[accessed 2007]
- Searches for new books;
used books should be searched with an accompanying system.

Free public access federated search systems for books: examples



Campusi: Find the best price on books - Microsoft Internet Ex...

File Edit View Favorites Tools Help

Back Forward Stop Home Search

Address <http://www.campusi.com/> Go

CAMPUS!

[Book](#) [Computer / Electronics](#) [New!](#) [Email](#) [Rare Book](#) [Movie](#) [Music](#) [Video Game](#) [Credit Card](#)

Find the best price on books. (with coupons)

BUY BOOKS

ISBN:

Title:

Author:

Keyword:

Search

[Rare Book Search](#)

• Fill in **fewer fields** to broaden the search.

Narrow your search by selecting one of the options below. (optional)

Format:

Internet

Free public access federated search systems for books: examples



- <http://www.marelibri.com/>
[accessed 2008]

The screenshot shows the Marelibri search interface within a Mozilla Firefox browser window. The browser's address bar displays the URL <http://www.marelibri.com/search/current?maximumPrice=0&keywords=&firstResult=0&ISBN=>. The page features a search bar at the top with a "Search" button and a language dropdown set to "English". Navigation links include "Thematic catalogue", "Advanced search", "Home page", "Register", and "Login". The "Search" section includes a "New" link and a search box with radio buttons for "All words", "Any word", "Match exact sentence", and "Boolean". The "Advanced search" section contains fields for "Author", "Title", "Description", "Topic(s)", "ISBN", "Bookseller", and "Location", along with a "Match exact sentence" checkbox and an "Allow extended syntax" checkbox. A "Search filters" sidebar on the left allows users to sort by "Relevance", filter by "Year" and "Price", select "All books" for the latest results, and show "50 Results". The "Source" section lists several libraries with checkboxes: Antikvariat.net, Antiqbook, Livre Rare Book, MareMagnum, Prolibri, and Uniliber.

Free public access federated search systems for books: examples



- <http://www.usedbooksearch.co.uk/books.htm>
[accessed 2008]
- Specialised in used books, not in new books.

Free public access federated search systems for books: examples



- <http://www.vialibri.net/> [accessed 2008]

viaLibri - - Resources For Bibliophiles. - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.vialibri.net/

Google vialibri Search AutoLink AutoFill Send to vialibri Settings

viaLibri - - Resources ...

Home Search Manager Libraries Links 553 Years More... Login / Register

viaLibri
Resources for Bibliophiles

Use this form to search for books from more than 20,000 antiquarian booksellers world-wide:

What To Search For...

Author: Printed: after: before: Title: Price: above: below: Imprint: Curr.: US\$ Prefer: Rare Books Keyword: First ed. Signed Dust Jacket: Sort by: Year Ascending Filter: Strict Wait: 30 sec.

Reset Search Now Help

Where To Search...

<input checked="" type="checkbox"/> Abebooks	<input checked="" type="checkbox"/> Alibris
<input checked="" type="checkbox"/> Amazon	<input checked="" type="checkbox"/> Antikbuch24
<input checked="" type="checkbox"/> antikvariat.net	<input checked="" type="checkbox"/> Antiqubook
<input checked="" type="checkbox"/> Biblio	<input checked="" type="checkbox"/> Biblion.co.uk
<input checked="" type="checkbox"/> Bibliophile	<input checked="" type="checkbox"/> Bibliopoly
<input checked="" type="checkbox"/> BooksatPBFA	<input checked="" type="checkbox"/> Choosebooks
<input checked="" type="checkbox"/> ILAB-LILA	<input checked="" type="checkbox"/> IOBABooks
<input checked="" type="checkbox"/> livre-rare-book	<input checked="" type="checkbox"/> Maremagnum
<input checked="" type="checkbox"/> prolibri	<input checked="" type="checkbox"/> Uniliber

Latest Online Catalogues [LIST FOR FREE]

Waverley Books List 154

Modern First Editions / Mystery / Science Fiction / Art & Photography /

Done

Free public access search systems for books: confusion



- **It is not always clear for a user in which category that a book search system belongs, because**
 - » Systems can evolve over time, for instance from a bookshop catalog to a multi-dealer catalog.
 - » It is not always explained in the user interface if a system is an aggregated database or a federated search system.
 - » The term “meta-search system” is sometimes used to indicate an aggregated database system or a federated search system or both.



Internet federated search engines for bookseller databases

Problem statement



Free public access federated search systems for books are interesting

- **Knowledge about the quality of this type of systems is interesting**
 - » for end users as well as for librarians who buy books,
 - » for librarians who serve their users by performing searches for books,
 - » for librarians who propose databases to their users, for instance on their library WWW site or who want to include one or several book search engines in their own local system for federated searching through several targets in one action.



Free public access federated search systems for books: problem statement

- In spite of the popularity and usefulness of these systems,
a comparative evaluation of their features and performance these days has not been available, as far as we know.





Internet federated search engines for bookseller databases

Methodology



Federated search systems for books: methodology for evaluation

Testing and comparing various federated search engines for books is complicated due to several facts:

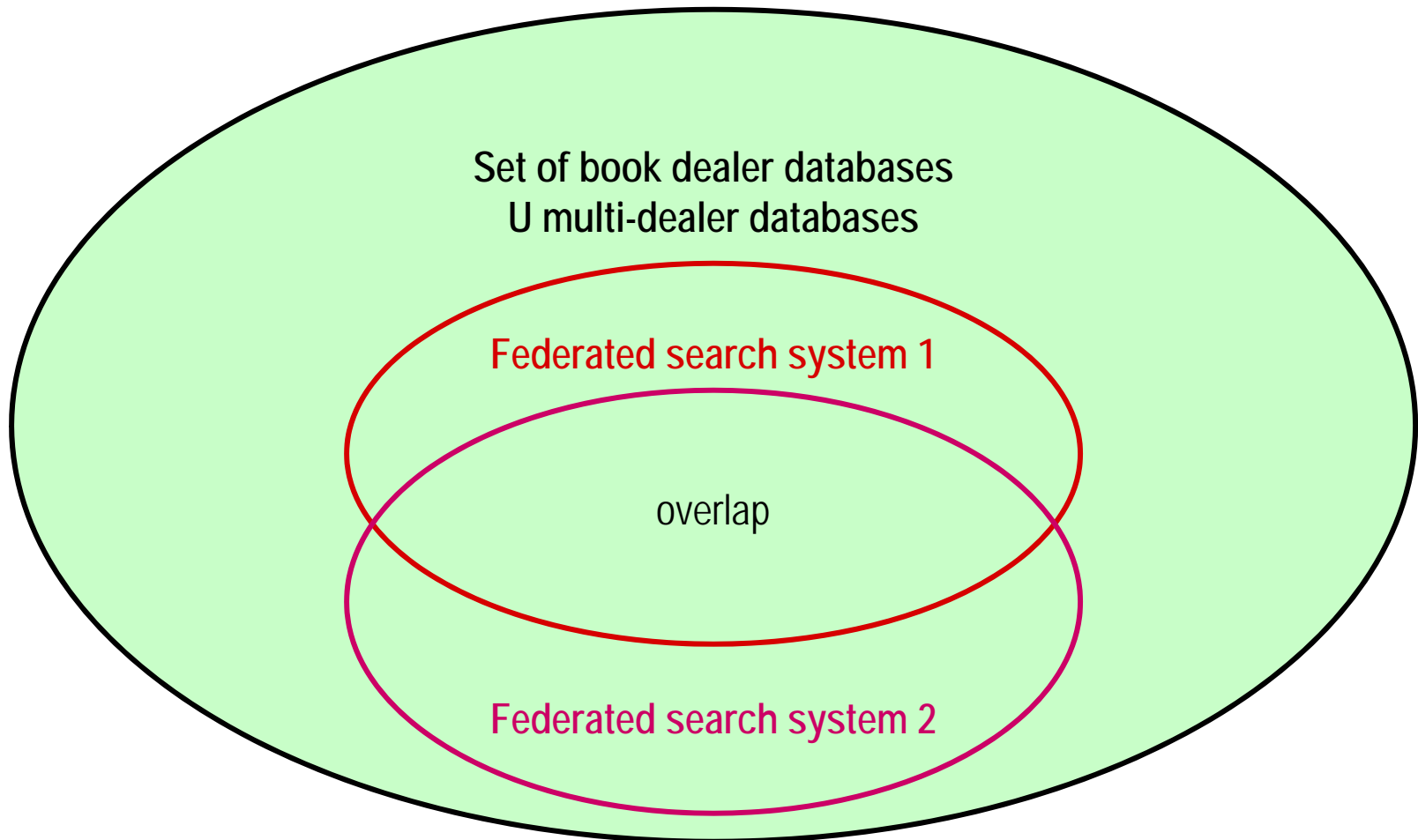
- **One meta-search engine 1 is NOT always simply smaller than another meta-search engine 2, in the sense that all book dealer databases covered by 1 are also covered by 2.**

Instead, the big, important systems have a great overlapping coverage.

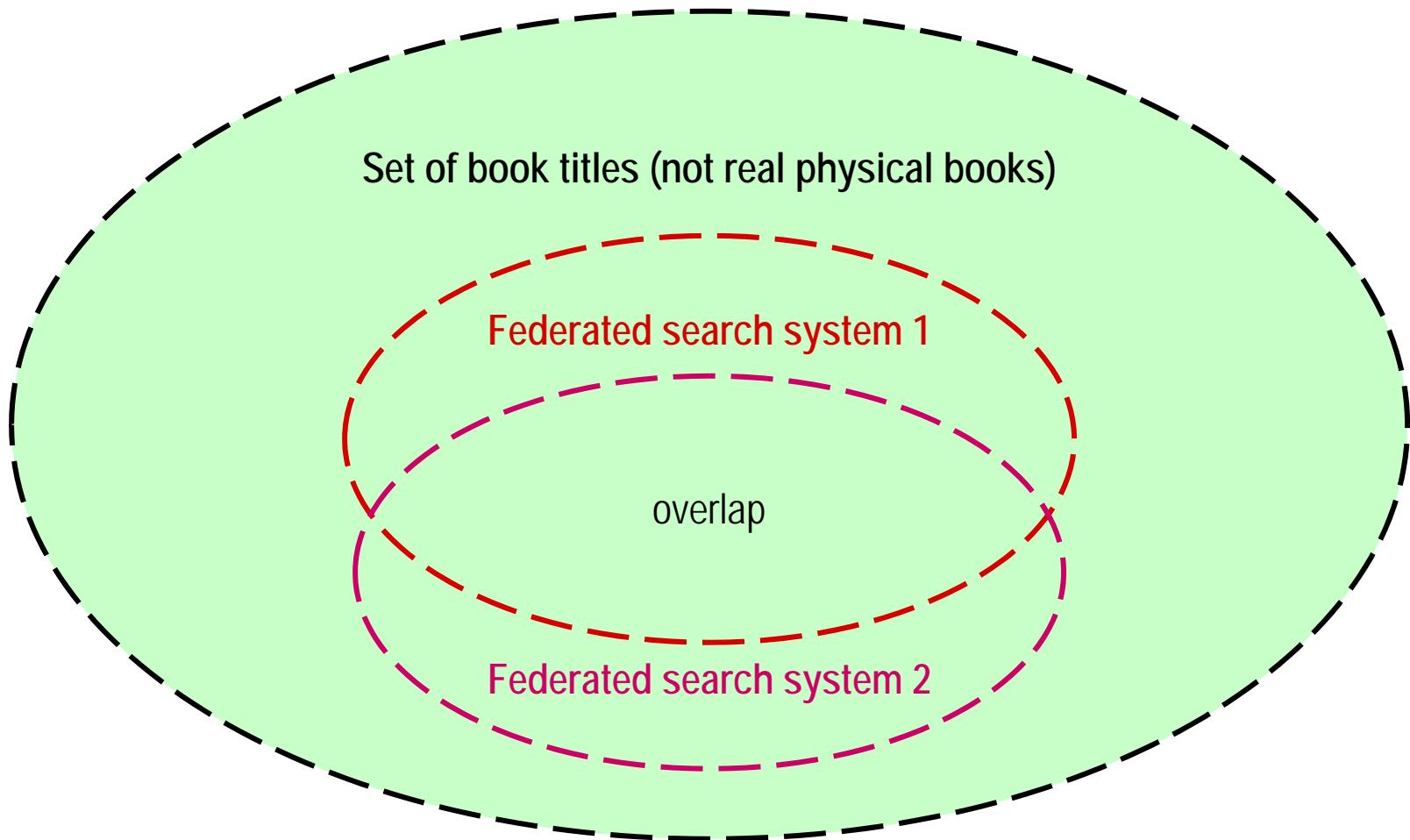
Most of the important multi-dealer databases are covered by the important federated search systems.



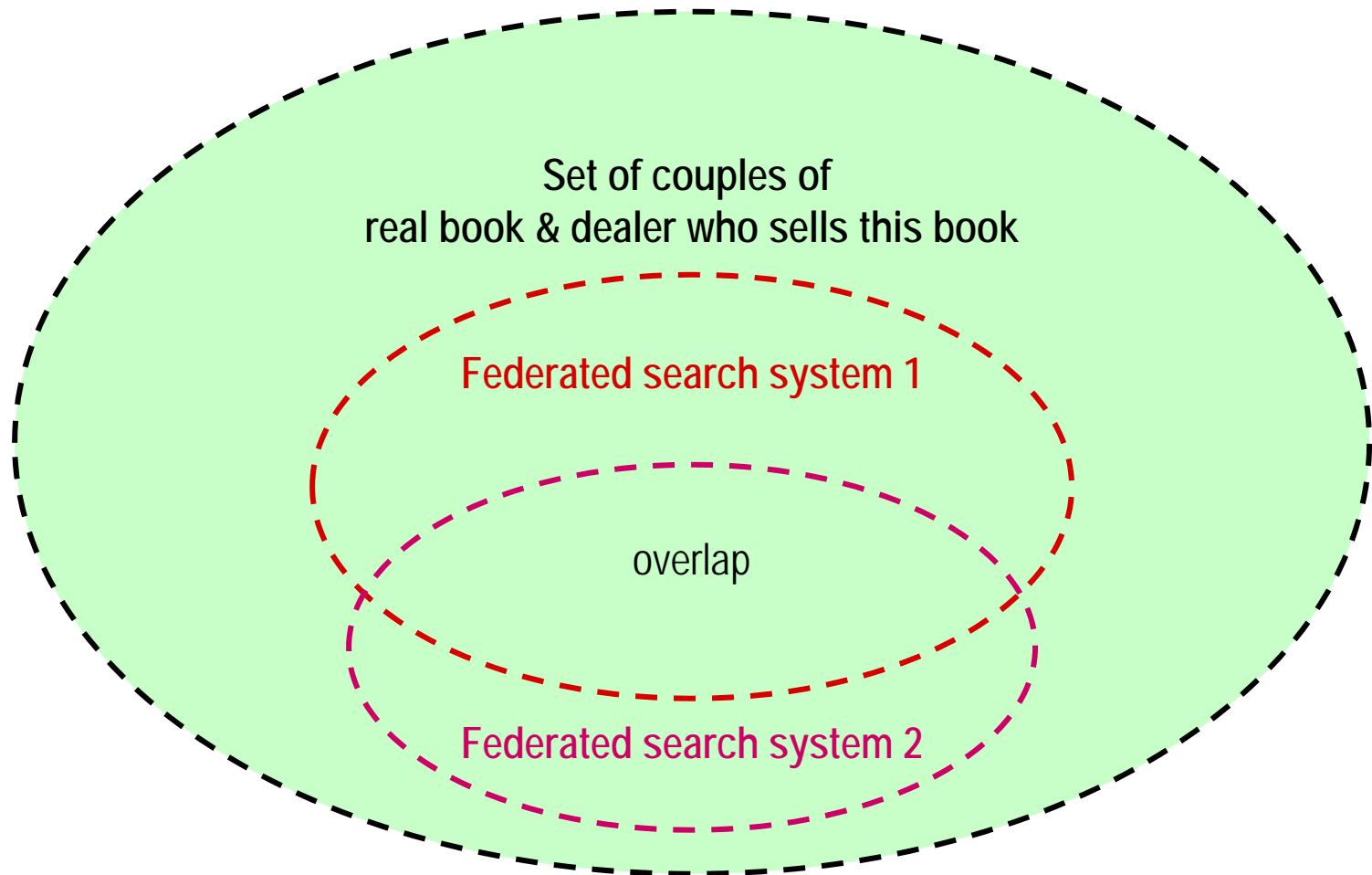
Federated search systems: relations



Federated search systems: relations



Federated search systems: relations



Metazoekmachines				Bookfinder	ViaLibri	Addall	BookFinder4U	KVK	UsedBookSearch	Eurobuch	Marelibri.com
Aantal doorzochte sites=> Doorzochte listing services	Land	Aantal Boekhandels	Aantal keren doorzocht	128 (= 25+103)	14	23 (= 14+9)	130 (= 119+11)	8 (= 7+1)	32 (= 4+28)	28 (= 13+15)	5
Books & Collectibles	AU	399	3	399		399	399				
AbeBooks.com	CA	13202	7	13202	13202	13202	13202	13202	13202	13202	
Amazon.com	CA	OK	6	OK		OK	OK	OK	OK	OK	
Guthschrift	DU	72	2	72						72	
prolibri.de	DU	200	4	200	200					200	200
Antikbuch24.de	DU	250	3	250	250					250	
Biblioman.de	DU	263	1							263	
Buchfreund.de	DU	350	2	350						350	
Antbo.com	DU	480	3	480			480			480	
Antiquario.de	DU	650	3	650				650		650	
Libri.de	DU	1000	1					1000			
ChooseBooks/ZVAB	DU	4100	5	4100	4100	4100	4100	4100			
booklooker	DU	OK	3	OK				OK		OK	
livre-rare-book	FR	607	4	607	607	607					607
Chapitre.com	FR	OK	0								
Maremagnum	IT	620	4	620	620	620					620
Antiqbook.com	NL	800	6	800	800	800	800			800	800
Libri.pl	PL	OK	1							OK	
antikvariat.net	Scan	97	1		97						
Uniliber	SP	787	1								787
ibooknet	VK	34	1	34							
Bibliopoly.com	VK	49	2		49	49					
Biblion.co.uk	VK	500	2	500		500					
PBFA	VK	700	2	700	700						
Barnes & Noble	VK	OK	3	OK		OK	OK				
ABAA Southern California	VS	45	0								
Americana Exchange	VS	94	0								
Abooksearch.com	VS	135	0								
TomFolio.com	VS	219	3	219		219	219				
Bookbase.com	VS	346	3	346	346	346					
IOBABooks.com	VS	514	1	514							
Biblio.com	VS	5500	5	5500	5500	5500	5500		5500		
Alibris.com	VS	10000	7	10000	10000	10000	10000	10000	10000	10000	
Used Book Central	VS	OK	2	OK			OK				
ILAB		590	3	590	590	590					
Totaal		42603	2,685714286	40133	37075	36932	34700	28952	28702	26267	3019
% Abebooks		24		22	26	26	28	46	46	50	0

From a presentation by Hanselaer (2008)

Federated search systems for books: methodology for evaluation

- **For a user of a book search engine it is not always clear in which of the categories mentioned above it fits. Many users are not aware of these technical differences or they do not care. Even if they do care, then the help or FAQ pages of the system do not always clearly and fully explain the workings of the system. Most systems are available free of charge for any user but they are commercial systems in the sense that the producer/owner has to make profit to stay in business; so they are not public open systems that reveal how they work to interested users.**



Federated search systems for books: methodology for evaluation

- **The results set of a search depends not only on the coverage (= the target databases), but also on how the search is executed, that is how and how well the search system is implemented for each target database = how well the search system has been adapted to each individual target database.**



Federated search systems for books: methodology for evaluation

- **A meta-search engine 1 may search through single-dealer databases 1, 2, etc...
as well as through multi-dealer databases A, B, etc...
in one action.**

This has some consequences that are time consuming and confusing for the user, as outlined hereafter.



Federated search systems for books: methodology for evaluation

- » Searching each of the underlying target databases can give results. If the federated search service presents these entries to the user without merging/deduplication of very similar book descriptions, then each of the entries obtained from a multi-dealer database may further lead to one OR MORE THAN ONE real shop/dealer. Furthermore each of these can correspond to an entry that was already directly presented as a result from a single-dealer database. Summarized, a federated search engine can give a result list in a first stage, which contains a number of relevant entries, but in a second stage, each of these relevant entries can eventually lead to more than one shop.



Federated search systems for books: methodology for evaluation

» Moreover, the contents covered by databases 1,2, etc... and A, B, etc... overlap in many cases, as database 1, 2 etc... can be included in database A, B, etc.

Thus searching all these in a single action may lead to several entries in the results that correspond to only one and the same real book available from one and the same dealer / shop.



Federated search systems for books: methodology for evaluation

- » Summarized, a federated search service does not always simply and directly provide a result set that consists only of descriptions of concrete books, each one available by a different dealer.
- » In other words, not all federated search systems implement full deduplication or merging of very similar results or this action is not working ideally.
This difficulty is important but not exceptional or remarkable; it is encountered with many federated search systems for many kinds of databases.



Federated search systems for books: methodology for evaluation

- Some federated search systems offer not just one but two search systems:

one for new books and one for used or rare books.

This may be inconvenient, if the user/searcher is not interested in the distinction between new and used books.

Furthermore, the difference between normal and rare is not clear.

(For instance the two related systems *addall.com* and *used.addall.com* take this approach.)



Federated search systems for books: methodology for evaluation

- Of course the alternative approach in which the user/searcher must NOT choose in advance between new and used books gives in many cases in the search results new as well as used books more or less mixed or separated on the display by the computer system. In that case a user who is only interested in for instance new books may find this inconvenient.
(For instance *bookfinder.com* takes this approach.)



Federated search systems for books: methodology for evaluation

- **Most federated search systems give many results in the case of well-known, popular books.**

The federated search systems tested gave no results in the case of many “obscure” books.

Therefore comparisons of search systems make sense mainly in the intermediate cases in which there exist only few couples of book & dealer, so that perhaps significant differences between the federated search systems can be investigated and revealed.



Federated search systems for books: methodology for evaluation

- **In view of the complexity observed, a “blind”, automated, computerised test was NOT developed.**
- **Instead, data have been collected by executing searches in the same way as normal users exploit these systems, interpreting the retrieved entries shown on the computer display as well as possible, knowing more or less how the systems work, being aware of the complications and taking these into account.**
- **Not only one single system was used, but at least 2 search systems were compared one after the other.**



Federated search systems for books: methodology for evaluation

- **When a first inspection of intermediate results showed that less than 9 couples book & dealer would be revealed, the results were inspected more closely:**
 1. Intermediate results from multi-dealer databases were expanded to reveal all couples of book & dealer.
 2. Results for couples of book & dealer were deduplicated.
 3. The finally obtained unique couples book & dealer were counted.
- **Immediately afterwards, another search system was tested.**



Federated search systems for books: methodology for evaluation

- **To avoid possible bias,
the order of search systems was random.**
- **This test / evaluation is carried out over a period of
about a year from 2007 up to May 2008.
This minimizes the influence of a temporary failure
or malfunctioning on the results.**



Federated search systems for books: methodology for evaluation

- **Books chosen were published mainly after 1940, in several countries and in several languages.**
- **Most were non-fiction in the arts and humanities; this domain is relevant and interesting for older books,
as they do not get outdated as fast as books in other domains like science, technology and medicine.**



Federated search systems for books: methodology for evaluation

- **Books were included that have been published by famous publishers for a wide audience, but also books published in fewer copies which are harder to find and thus more interesting in a comparison of search services (for instance at the occasion of a temporary exhibition as well as auction catalogs).**





Internet federated search engines for bookseller databases

Results of this investigation



Remarkable behavior of federated search systems observed



- **In most federated search systems, in the settings of the preferences, the user can choose the destination country. It turned out that not only the price estimated for shipping depends on this setting, but more surprisingly also the number and type of found shops/dealers.**



Remarkable behavior of federated search systems observed



- Searching by ISBN does not work identically in all systems.

For instance:

Searching with an ISBN that contains spaces was not accepted by *bookfinder.com*,
but was well interpreted by *used.addall.com*.



Comparative evaluation of several federated searching systems



- **Most federated search systems give many results in the case of well-known, popular books.**

The federated search systems tested gave no results in the case of many “obscure” books.

Therefore, comparisons of search systems make sense mainly in the intermediate cases in which there exist only few couples of book & dealer, so that perhaps significant differences between the federated search systems can be investigated and revealed.



Comparative evaluation of several federated searching systems



- To start with, 4 systems were compared.
 - » <http://nl.bookbutler.com/>
 - » <http://www.addall.com/>
with <http://used.addall.com/> for used books
 - » <http://www.bookfinder.com/>
 - » <http://www.campusi.com/>



Comparative evaluation of several federated searching systems



- **Meta-search systems react slower than “real” database search engines in most cases (as they need several simultaneous Internet connections from the search system to the target databases and they have to merge raw preliminary results into the result that is finally shown to the user).**

Nevertheless the tests showed that response times were in the order of a few seconds only, which should be acceptable for most users.



Comparative evaluation of several federated searching systems



- In a series of tests, 14 test cases showed a number of retrieved couples book & dealer, which was small enough to allow and motivate a more detailed quantitative comparison.



Comparative evaluation of several federated searching systems



- Of the search systems investigated only the *campusi* system for rare books referred occasionally to a book that was for sale through the famous and popular Internet auction system *Ebay* (by coincidence at the time of testing).

This occurred in 3 of the 14 test cases.

In 1 of the test cases this yielded the only couple book & dealer.

So this feature can be seen as a plus.



Comparative evaluation of several federated searching systems



- In the majority of cases (13 of 14) *addall / used.addall* or *BookFinder* gave more couples book & dealer than *BookButler* or *Campusi*.



Quantitative evaluation of federated search systems with a high coverage



- In view of the outcome of the first tests, subsequent evaluation was executed only with

» *addall / used.addall*

» *bookfinder.*



Quantitative evaluation of federated search systems with a high coverage



- For most books either no result is found or quite many; so in these cases the search systems offer equal value.
- However with all the book titles used in this test, more than 40 books fell between those extremes of untraceable and abundant books.
- These revealed differences in the couples book & dealer retrieved by the investigated search engines.



Quantitative evaluation of federated search systems with a high coverage



- After each test with a single book title, we can count a search system as “**winner**” when that system finds for that specific book title more couples of book & dealer. Then the number of times that a system could be called a winner is like this:

» *used.addall.com* 22

» *bookfinder.com* 7

» Comparisons without winner: 16



Quantitative evaluation of federated search systems with a high coverage



- In a more strict, restrictive way, after each investigation of a single book title, we can count a search system as a **“strict winner”** when that system finds at least one dealer (occurrence) for that book title, while the other system finds NO dealer at all.

Then the number of times that a system could be called a strict winner is:

» *used.addall.com* 7

» *bookfinder.com* 4





Internet federated search engines for bookseller databases

Discussion



Discussion: generalization



- **Our results should not be generalized as a reflection of the potential of the investigated search engines for EVERY book.**

A search engine may perform better for one type of book (older versus more recent, new versus used, subject domain, language, value and price...)

as the target databases are different and some of the dealers are specialized in a particular type of books.





Internet federated search engines for bookseller databases

Conclusion



Conclusion



- **This investigation has clarified to some extent the state of the art of Internet book search services for librarians as well as other users.**
- **For an abundant book title, most of the systems that search bookseller databases will satisfy the user.**



Conclusion



- **In the case of a scarce book title, it is wise to choose a federated search system with a wide coverage.**
- **More than 1 search system can serve as complementary search tools as**
 - » their user interface and features are far from identical
 - » their coverage is overlapping but has unique parts
- **So a book that cannot be found by one system can perhaps be retrieved by the other system.**



Questions?

Suggestions?

Topics for discussion?

- **You are free to copy, distribute, display this work under the following conditions:**



» Attribution:

You must mention the author.



» Noncommercial:

You may not use this work for commercial purposes.



» No Derivative Works:

You may not change, modify, alter, transform, or build upon this work.

- **For any reuse or distribution, you must make clear to others the license terms of this work.**