Something about IR

in the age of the web and social media

Jaap Kamps
University of Amsterdam

16 October 2014
Fall EARIA School, Notre Dame de Parménie, France

Outline

- IR in a changing world
  - Classic IR and revolutionary changes
  - Challenges to IR and how to tackle them
- Case study: Social Book Search
  - Changing collections, tasks, users, …
  - Toward rich task/user context

IR at age 50-something…

Part I

bing

notre dame de parmenie

Web Images Videos Maps News More

25 Jaap

20,700,000 RESULTS Any time Near Engins, Rhone-Alpes · Change

Images of **notre dame de parmenie**

bing.com/images

Related searches

Chateau **de Notre Dame**

Hotel **de Notre Dame**

What State is **Notre Dame** in

Where is **Notre Dame** College

Notre Dame Fighting Irish

"Modern search" is pervasive but has a long history

Google

notre dame de parmenie

+Jaap Share

Web Maps Images News Shopping More Search tools

About 3,640 results (0.32 seconds)

Tip: Search for **English** results only. You can specify your search language in Preferences

Izeaux, **Notre-Dame de Parménie** - LaSalle, France // Frères ...

www.lasallefrance.fr/-**Parmenie**- Translate this page

Notre-Dame de Parménie à Izeaux est un haut-lieu lasallien, qui accueille des jeunes pour un temps de réflexion et de partage ; des groupes internationaux ...

AMERICAN DOCUMENTATION

1955

A QUARTERLY REVIEW OF IDEAS, TECHNIQUES, PROBLEMS AND ACHIEVEMENTS IN DOCUMENTATION

VOL. VI, NOS. 1-4 JANUARY-OCTOBER 1955

Editorial

"The Truth, the Whole Truth...."

Interpreted literally this time-honored oath of the courts confronts the witness with a humanly impossible task. No one could be expected to perceive all the elements — all the minute detail, important or otherwise — that comprise a given incident, event, or situation. The best that one may expect is that the really important will be recognized, recorded in the memory, and recalled with accuracy when the occasion demands. As every lawyer knows, even the most honest and scrupulous witnesses will swear to contradictory accounts and disagree among themselves concerning relatively obvious detail.

So too, honest differences of opinion respecting the relative importance of the components of subject content have initiated extensive and bitter controversy among those who have devised or espoused classification schemes that follow a traditional pattern. The division and subdivision of areas of knowledge into rigid compartmentation is possible only by the identification of lines of demarcation which must, in their definition and location, be, in the last analysis, little more than arbitrary. Failure to recognize this basic fact has led logicians into error and classifiers into strife.

But classification unencumbered by the arbitrary delineation of knowledge into rigid compartmentation is now possible through the use of automatic, or semi-automatic mechanisms — machines which give to classification a new flexibility and elasticity through the ease with which entirely new categories, classes, or composite groupings may be generated whenever a particular situation or need so demands. Selection based on combinations, any combinations, of characteristic attributes is a fundamental property of documentation systems utilizing automatic or semi-automatic mechanisms (mechanical aids). Thus with properly designed equipment the full range of combinations permitted by the logical theory of class definitions becomes possible in making selections.

Jesse H. Shera comments on the indexing debate

the attributes and properties (characteristics) of the subject content of a given document or store of documents might well outweigh any anticipated future advantage. Thus does the documentalist become as much economist as logician, and the laws of logical division yield before the law of diminishing returns.

categories, classes, or composite groupings may be generated whenever a particular situation or need so demands. Selection based on combinations, any combinations, of characteristic attributes is a fundamental property of documentation systems utilizing automatic or semi-automatic mechanisms (mechanical aids). Thus with properly designed equipment the full range of combinations permitted by the logical theory of class definitions becomes possible in making selections.

Nevertheless, documentation systems employing such mechanical aids are also subject to the limitations imposed by human inability to tell "the *whole* truth." But even were the machines able to transcend this human limitation the expense involved in recording for subsequent searching *all* the attributes and properties (characteristics) of the subject content of a given document or store of documents might well outweigh any anticipated future advantage. Thus does the documentalist become as much economist as logician, and the laws of logical division yield before the law of diminishing returns.

One example may make our point clear. A mechanized system may prove entirely satisfactory when applied to a small collection of documents but break down completely under the burden of a more extensive and complex load. Conversely, a more elaborate system, quite capable of handling efficiently large masses of diverse materials, may be disappointingly ineffective when tested upon a small collection. One does not need a wheel to break a butterfly! Thus the most exact and careful of measurements can prove misleading when the efficiency of a system is strongly influenced, as it so often is, by the volume of documents that it is called upon to handle. Cautious and searching evaluation of all experimental results is essential in rating the efficiency of documentation systems.

May the age-old controversies that arose from the conventional concepts of classification not be reborn in the mechanized searching systems of the future. There is hope for the avoidance of such error if we will but regard documentation systems as useful devices the benefits of which must be determined, not by polemics but by the intelligent measurement of such benefits in relation to needs and costs. The machines of the future *can* make us free, but only if we are willing to subject them, and ourselves, to the most rigid intellectual discipline.

Shera asks for proper scientific evaluation (1955)

Picked up by a Librarian at Cranfield

ASLIB

CRANFIELD RESEARCH PROJECT

Cleverdon develops his evaluation method (1960-)

REPORT ON THE TESTING AND ANALYSIS OF AN INVESTIGATION INTO THE COMPARATIVE EFFICIENCY OF INDEXING SYSTEMS

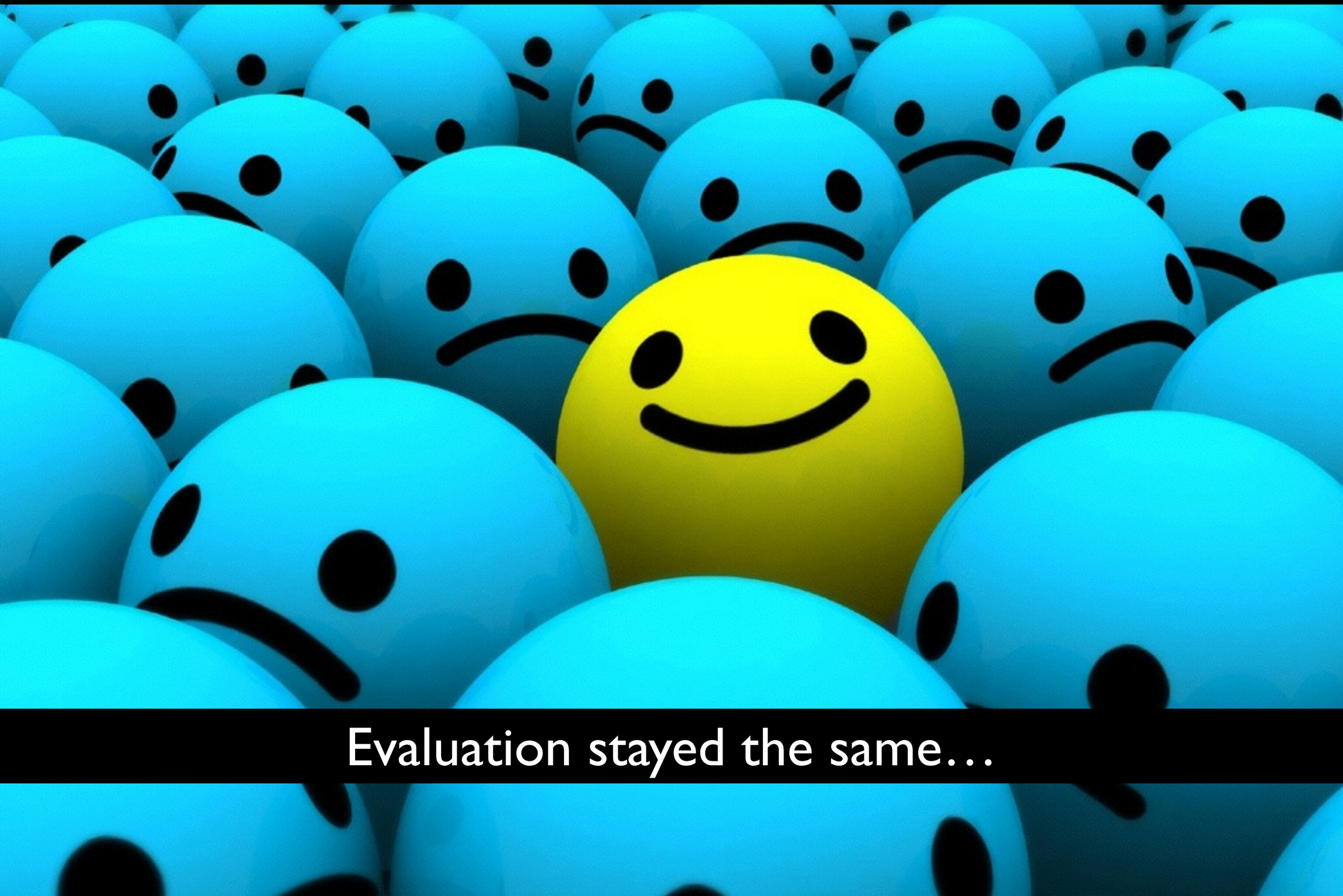
by

Cyril W. Cleverdon

An investigation supported by a grant from
The National Science Foundation

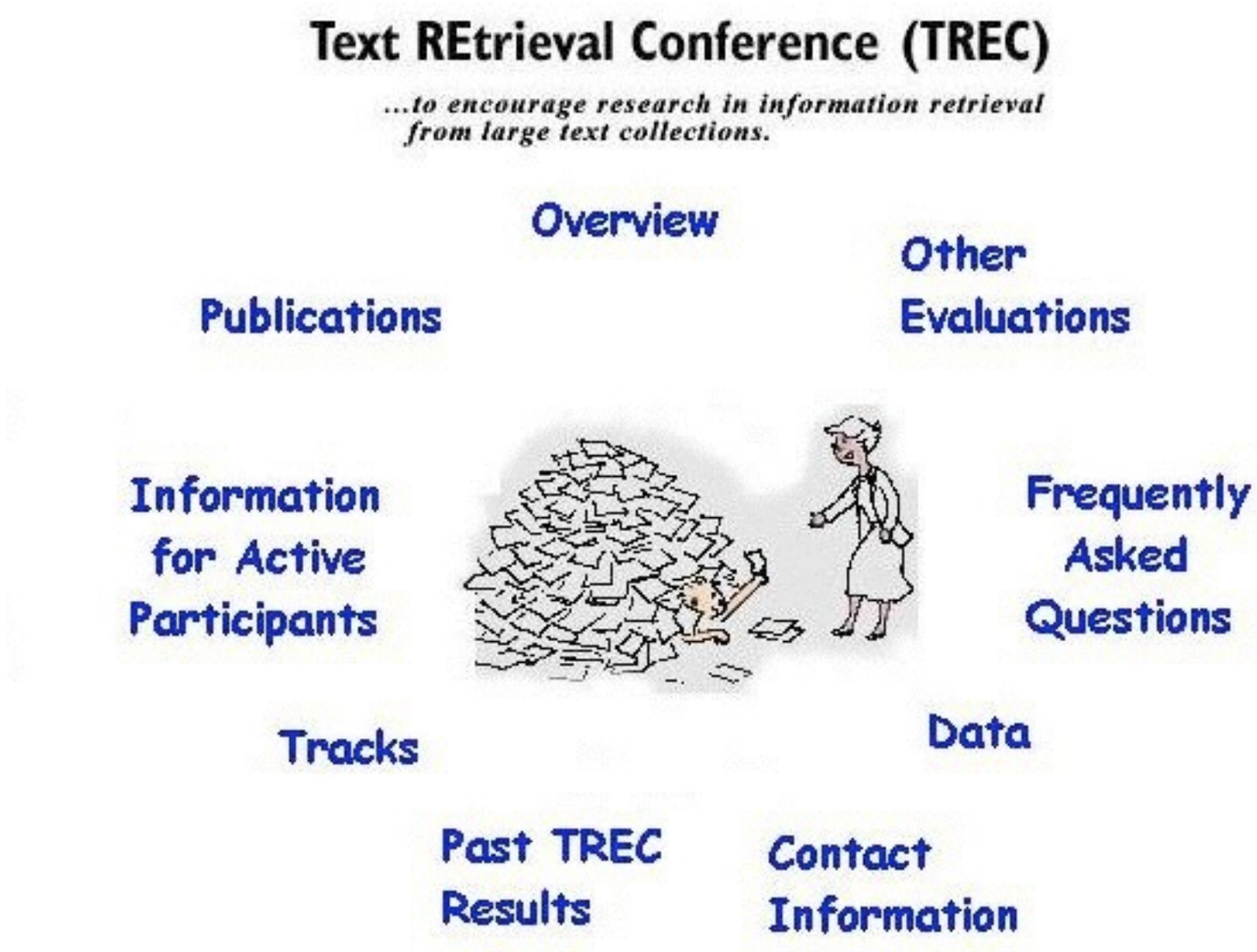
1960s: Automatic Indexing can be as effective as Manual

Since 1960 much development of search technology



Evaluation stayed the same…

Just need a corpus, a bunch of topics, some judgements



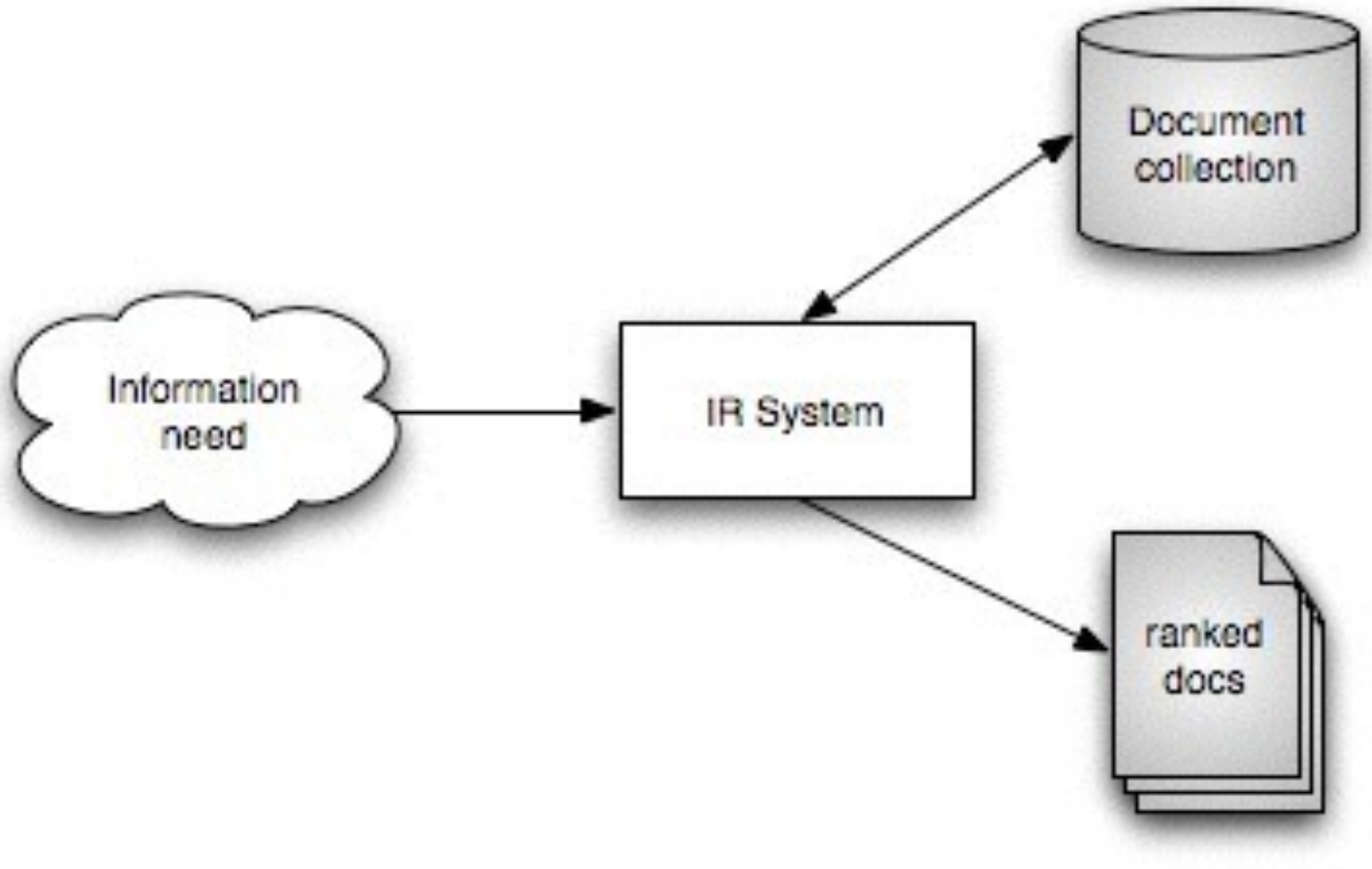
Still dominates academia *and* industry

What's the problem?

Underlying "User Model" did not change

Still *library consultation*: static collection, dynamic requests

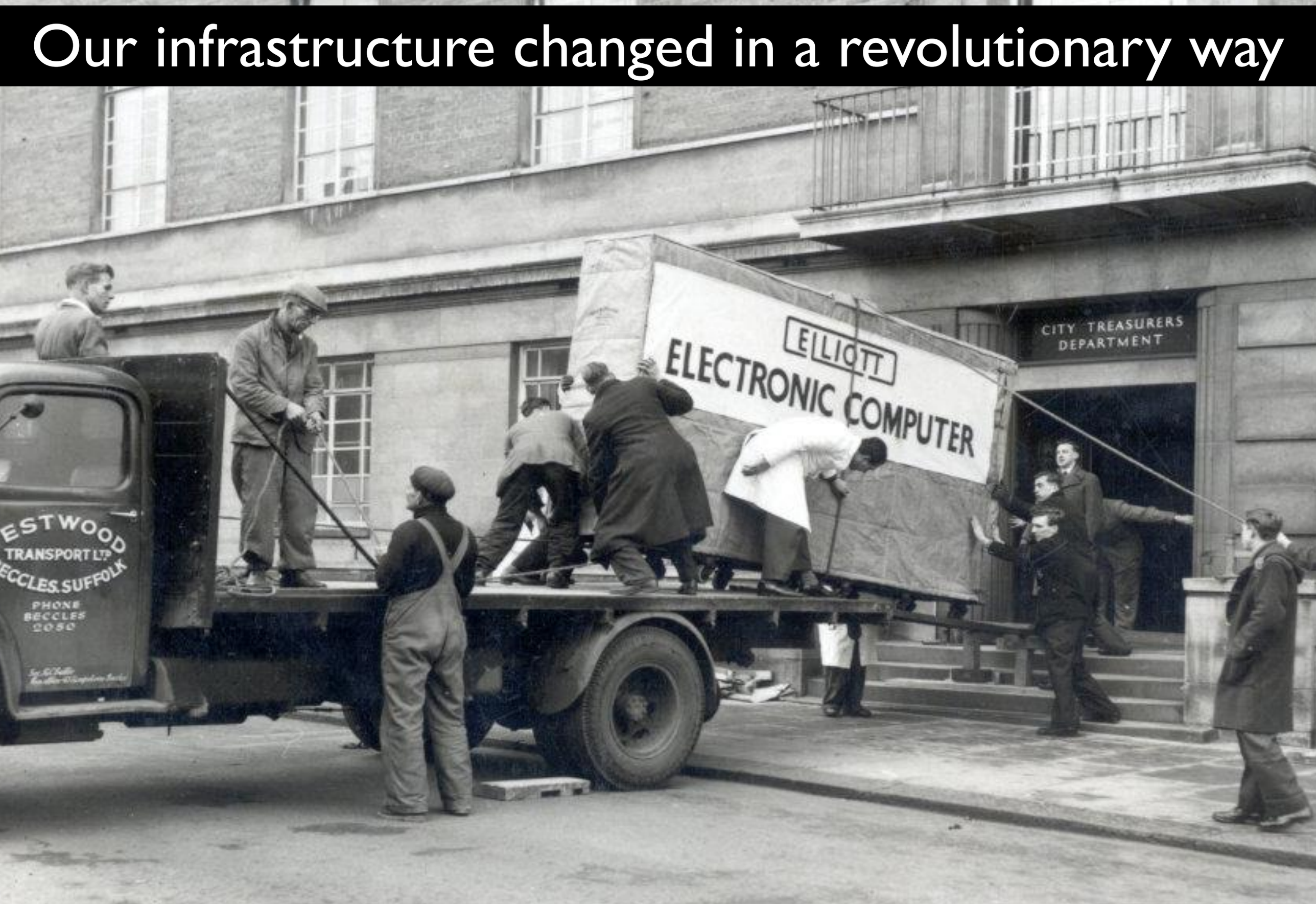
Predominantly one-step-search (batch, no interaction)



Isolated search ignoring the actual *use of the information*

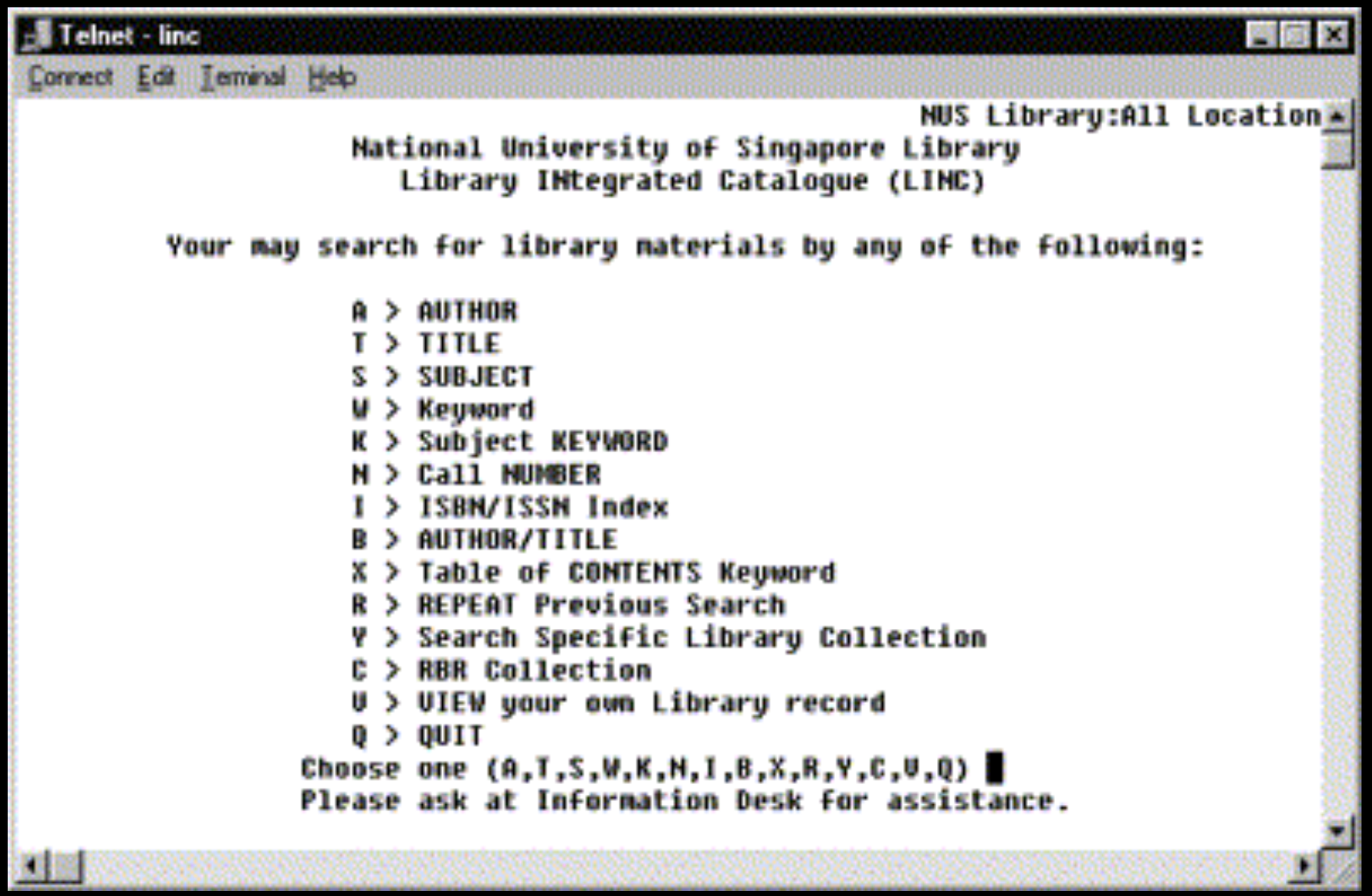
Is this the right model in the “always online” world?

Our infrastructure changed in a revolutionary way

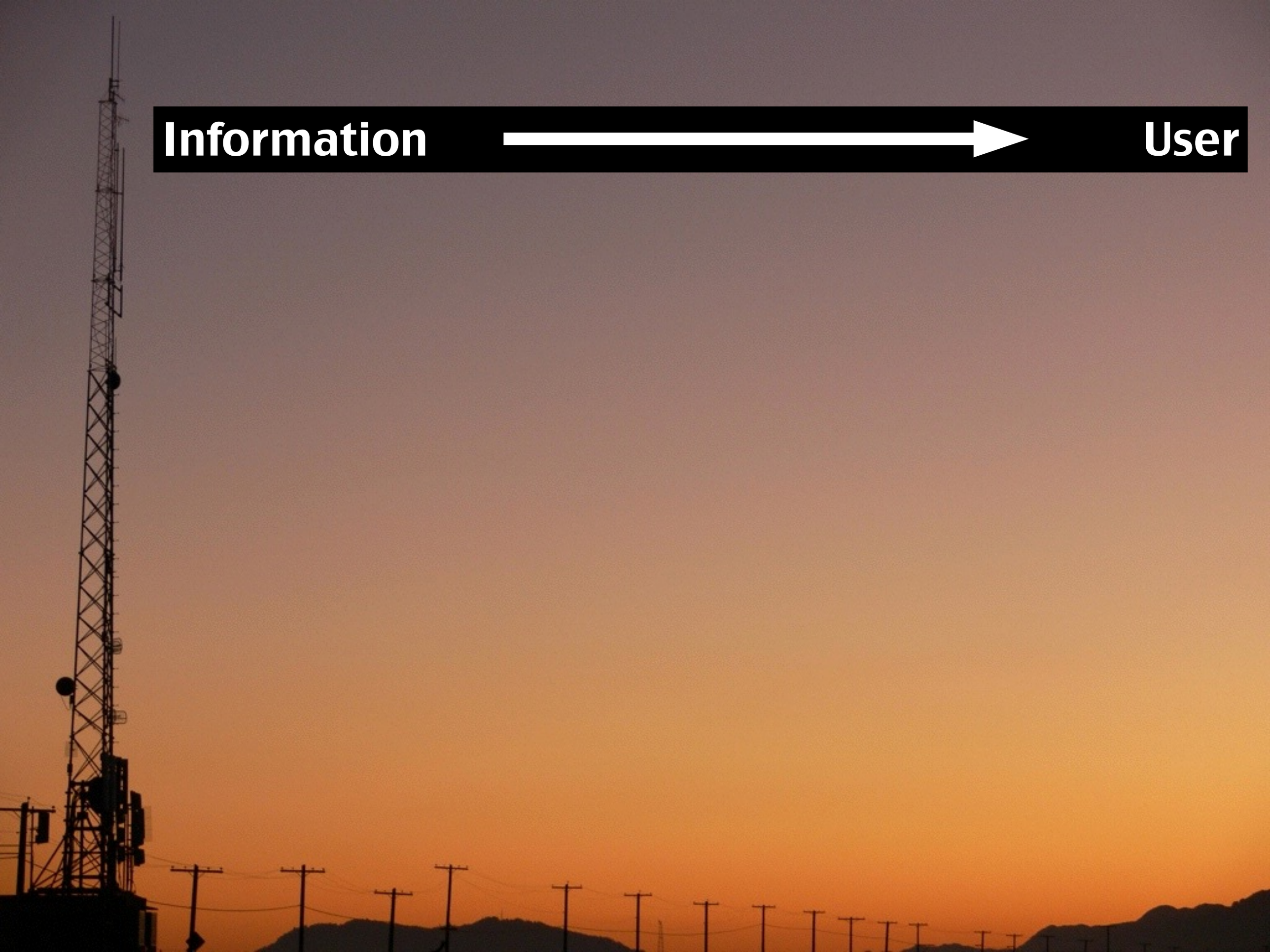


Our technology changed in a revolutionary way

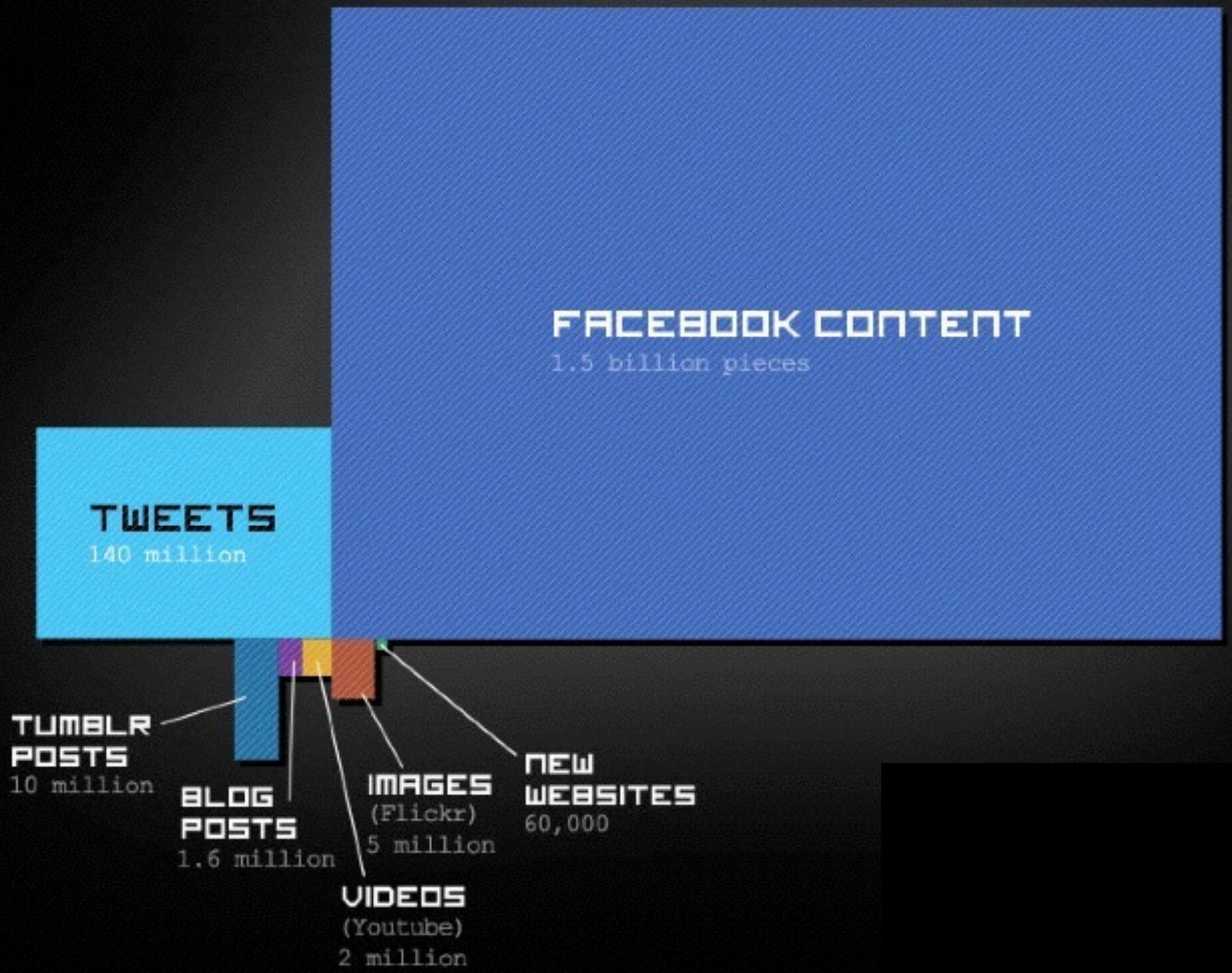
How radical did information access methods change?



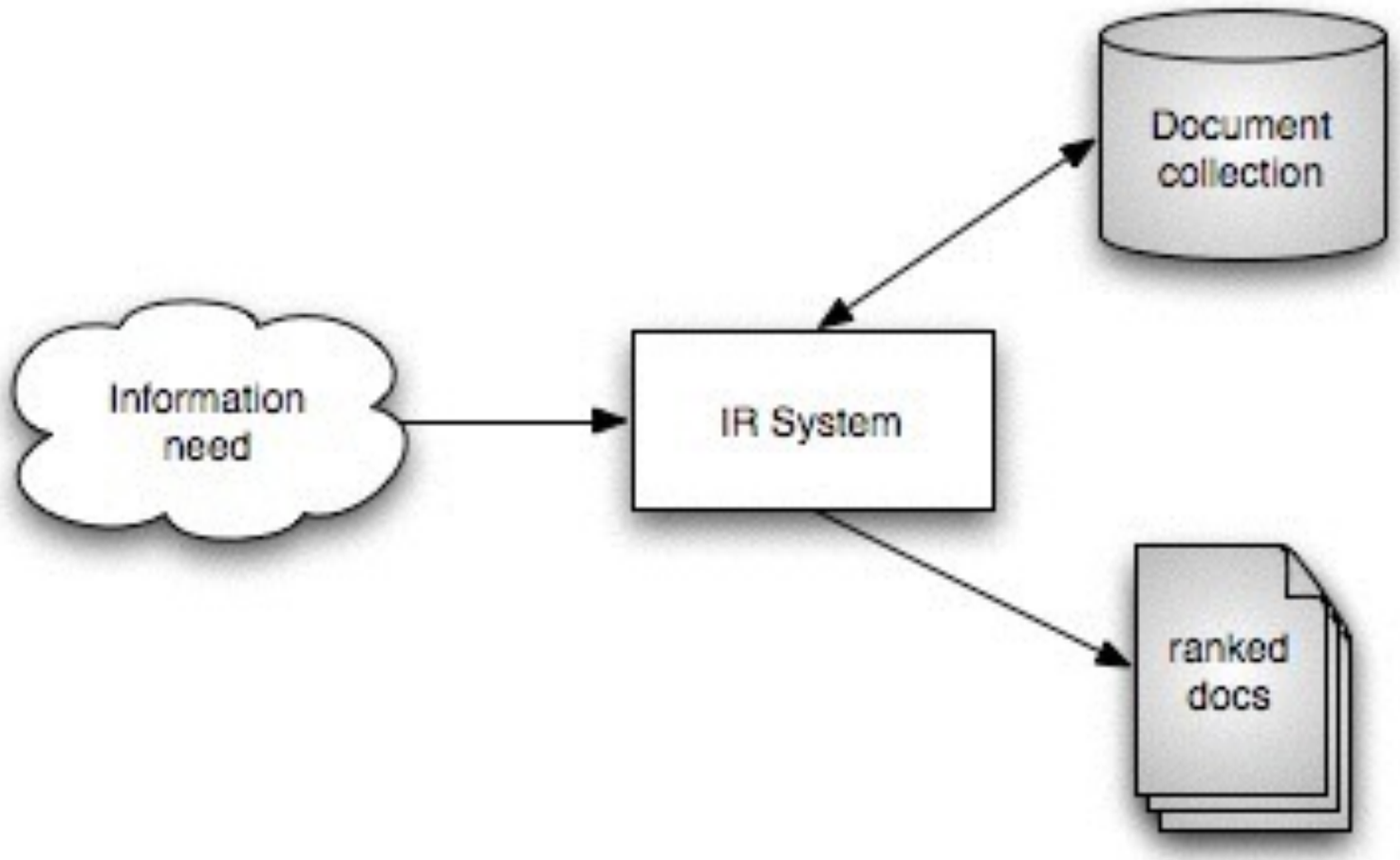
Information
User


Information
User

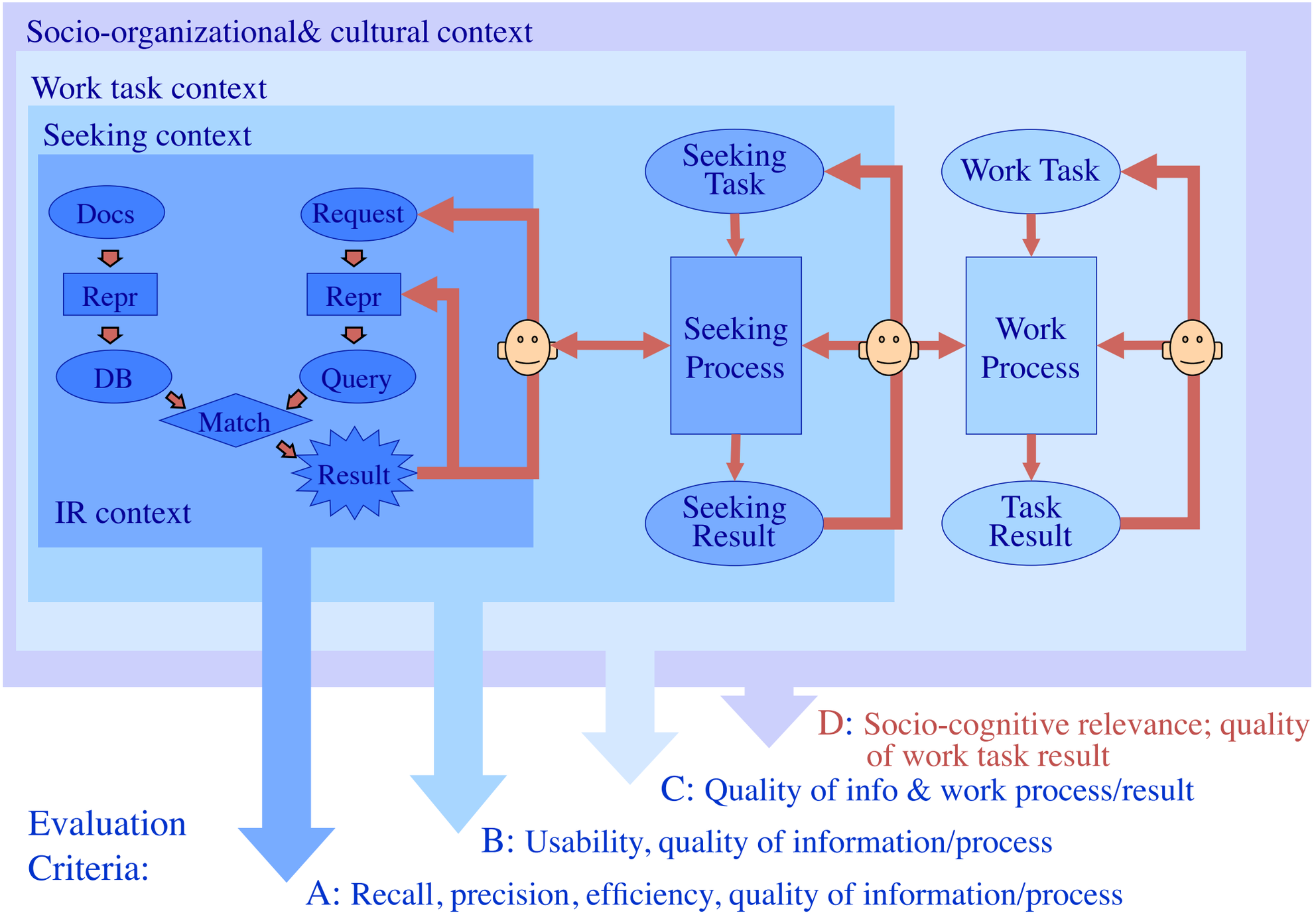
DAILY CONTENT PUBLISHED, BY NUMBER OF PIECES



Information
User



Classic IR: Query → Doc


Socio-organizational& cultural context
Work task context
Seeking context
Docs
Repr
DB
Request
Repr
Query
Match
Result
IR context
Seeking Task
Seeking Process
Seeking Result
Work Task
Work Process
Task Result
Evaluation Criteria:
A: Recall, precision, efficiency, quality of information/process
B: Usability, quality of information/process
C: Quality of info & work process/result
D: Socio-cognitive relevance; quality of work task result

Query++ → Doc

Zeitgeist Overview

More information than you require.

Vital Statistics

Members 1,746,913
Books cataloged 85,804,743
Tags added 102,854,851
Unique works 7,886,744

Reviews 2,005,318
Works reviewed 747,354
Ratings 14,274,940

Member-contributed covers 3,687,595
Author pictures 74,142
Venue pictures 10,608

Groups 9,666
Talk topics 152,131
Talk messages 4,311,760
Talk touchstones 3,038,567

Free books given out

Early Reviewer books given out 166,074
Member Giveaway books given out 341,245
Total free books given out 507,319

50 largest libraries (see more)

eandino2012 (65,535 books), jonathan_rigby (50,225 books), kamran3 (45,604 books, private), Kaethe (35,152 books), ohclibrary (28,402 books), diana.n (26,155 books), lyzard (25,757 books), Nedrin (24,498 books), credo (24,001 books, private), stansteiner (23,817 books), BlueTysonSS (23,544 books), gangleri (19,755 books), mvuijlst (18,744 books), JeanLittleLibrary (18,737 books), zenosbooks (18,677 books), arjaygee (18,426 books), michael.brodesky.1 (18,422 books), meftung (18,275 books, private), sljmbll (18,099 books), kday_working (17,952 books), shearrob (17,765 books), tglovell (17,520 books),

25 most reviewed books (more)

The Hunger Games (2,536 reviews), Twilight (1,883 reviews), The Girl With the Dragon Tattoo (1,388 reviews), The Book Thief (1,274 reviews), Catching Fire (1,217 reviews), Harry Potter and the Deathly Hallows (1,187 reviews), The Help (1,123 reviews), Mockingjay (1,107 reviews), The Road (1,041 reviews), The Time Traveler's Wife (1,028 reviews), The Da Vinci Code (1,010 reviews), Harry Potter and the Sorcerer's Stone (Book 1) (1,008 reviews), Water for Elephants (963 reviews), The Kite Runner (959 reviews), The Curious Incident of the Dog in the Night-Time (951 reviews), Breaking Dawn (872 reviews), New Moon (870 reviews), The Guernsey Literary and Potato Peel Pie Society (853 reviews), The Graveyard Book (836 reviews), Pride and Prejudice (810 reviews), The Giver (801 reviews), To Kill a Mockingbird (793 reviews), The Lovely Bones (753 reviews), Life of Pi (722 reviews), The Lightning Thief (719 reviews)

Top 25 books (more)

Harry Potter and the Sorcerer's Stone (Book 1) (74,757), Harry Potter and the Half-Blood Prince (65,785), Harry Potter and the Chamber of Secrets (65,119), Harry Potter and the Prisoner of Azkaban (64,631), Harry Potter and the Order of the Phoenix (64,053), Harry Potter and the Goblet of Fire (63,285), Harry Potter and the Deathly Hallows (60,080), The Da Vinci Code (49,697), The Hobbit (49,530), 1984 (47,587), Pride and Prejudice (47,201), The Catcher in the Rye (45,317), To Kill a Mockingbird (44,234), The Great Gatsby (41,304), Twilight (40,596), The Kite Runner (36,134), Jane Eyre (33,920), Animal Farm (33,594), The Lord of the Rings (32,621), New Moon (32,150), Brave New World (31,838), The Hunger

Top 25 works by rating

Rated by at least 25 users

Queen Elizabeth's Wardrobe Unlocked (4.97), The Complete Calvin and Hobbes (4.9), Patterns of Fashion 4: The Cut and Construction of Linen Shirts, Smocks,

Authors who LibraryThing (complete list)

Diana Gabaldon, Lee Child, Don Maitz, Sandra Brown, David Brin, Laurie R. King, John Green, Jude Deveraux, Diane Duane, Ann Brashares, Brandon Sanderson, Jon Scieszka, Laurie Halse Anderson, James Rollins, Naomi Novik, Lisa See, Katie MacAlister, Andrew Gross, Janny Wurts, Chris Bohjalian, Dan Chaon, Elizabeth Bear, Nancy Holder, Melissa Marr, Jo Beverley, Tite Kubo, Sharon Kay Penman, Susan Mallery, Patrick Rothfuss, Edward R. Tufte, Susan Wittig Albert, Adriana Trigiani, Alan Furst, Susan Wiggs, Matthew Pearl, Lynn Viehl, Randy Alcorn, Elisabeth Waters, Bryan Lee O'Malley, Sharon Lee, James Dashner, Tatiana Holway, Zachary Sholem Berger, Sarah Addison Allen, Neal Asher, Sabrina Jeffries, Andrew Vachss, Charlotte Hughes, Chitra Divakaruni, Janet Morris — see the full list.

Top 75 authors (author cloud)

By number of copies

J. K. Rowling (498,000), Stephen King (366,414), Terry Pratchett (311,057), J. R. R. Tolkien (237,113), C. S. Lewis (221,456), Neil Gaiman (201,379), William Shakespeare (196,648), Nora Roberts (185,599), Agatha Christie (171,705), Jane Austen (154,517), James Patterson (149,884), Stephenie Meyer (148,462), Isaac Asimov (146,640), Charles Dickens (139,253), John Grisham (130,053), Anne McCaffrey (123,347), Dan Brown (122,806), Orson Scott Card (121,297), Janet Evanovich (118,378), Kurt Vonnegut (115,000), Anne Rice (112,090), Douglas Adams (110,282), Dean Koontz (108,855), George Orwell (107,360), Robert A. Heinlein (104,498), Dr. Seuss

Query++ → Doc++

(87,273), Ernest Hemingway (86,016), George R. R. Martin (85,857), Piers Anthony (85,699), Lemony Snicket (83,747), Fyodor Dostoevsky (83,307), Suzanne Collins (81,953), Philip Pullman (80,529), Haruki Murakami (80,102), Ray Bradbury (79,239), Gabriel García Márquez (77,184), David Eddings

XML Find: GNU GPL

http://berk.science.uva.nl:8080/xmlfind/result.php?srcQuery=GNU+GPL

Getting Started Latest Headlines news mail search uva blog amsterdam vefrit Wikiii

xmlfind

Search: GNU GPL Search

Display results 1 to 10 of 24 (Lucene took 0.335 sec. to return results, XML Find took 2.591 sec. to display results)

1 2 3 next

Open Source Software Development: An Overview

Ming-Wei Wu, Ying-Dar Lin

article[1]
bdy[1]
sec[2] (OPEN SOURCE BACKGROUND)
... for work on the **GNU** Project (http://www.**gnu**.org/**gnu**/the**gnu**project.html). **GNU** is a recursive acronym for "**GNU**'s Not Unix" and a homophone for "new." The **GNU** Project seeks to develop ... principal author of the **GNU** C Compiler (GCC), **GNU** symbolic debugger (GDB), **GNU** ... in essence, copyrights with **GPL** regulations. Open source software, essentially ...
sec[4] (LICENSING MODELS)
p[5]
... code falls under the **GPL** model, nVidia had to remove all **GPL** code, then re-release the ... Softway Systems, makers of **GPL**-regulated software, and repackaged its ... work, thereby skirting the **GPL** regulation.
p[6]
... multilicensing models to avoid **GPL** violations. For example, Sun's ... adapts three licensing models: **GPL**, L**GPL**, and SISSL. Ordinary users who can fulfill **GPL** regulations can use StarOffice under **GPL**. Proprietary developers and companies can use L**GPL** or SISSL, which both ...
sec[5] (BUSINESS MODELS)
... software company cannot fulfill **GPL** regulations. Thus, a private ... its reputation. For example, **GNU** has

Query++ → Hitlist → Doc++

Larry Graham

article[1]
bdy[1]
sec[4] (DEVELOPING LINUX)
p[5]

https://www.facebook.com/search/me/friends/friends/employers/present/111777152182...

Restaurants in **Amsterdam, Netherlands** that friends of my friends work at

Home Find Friends Jaap

- **Restaurants in Amsterdam, Netherlands** that friends of my friends work at
- Cafes in **Amsterdam, Netherlands** that friends of my friends work at
- Asian restaurants in **Amsterdam, Netherlands** that friends of my friends work at
- Restaurants in **Amsterdam, Netherlands** that friends of my friends who work ...
- American restaurants in **Amsterdam, Netherlands** that friends of my friends ...
- Restaurants in **Amsterdam, Netherlands** that friends of my friends who work ...
- Pizza Places in **Amsterdam, Netherlands** that friends of my friends work at
- Bars & grills in **Amsterdam, Netherlands** that friends of my friends work at

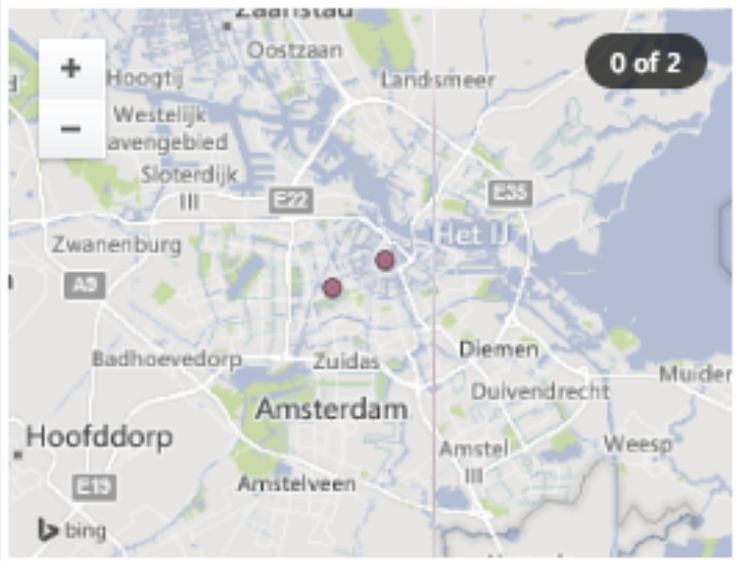
See more results for "Restaurants in Amsterdam, Netherlands that friends of my friends work at"

Restaurants in Amsterdam, Netherlands that friends of friends of my friends work at — DoubleT...

Amsterdam is the capital and most populous city of the Netherlands. Its status as the Dutch capital is mandated by the C...
SEE MORE

www.amsterdam.nl

2 Places Share



REFINE THIS SEARCH

Place Type: Restaurant
Category...
Liked by: Add...
Name: Add...
Places in: Amsterdam, Netherla...

EXTEND THIS SEARCH

Chat (Off)

Query++ → Query Suggest → Hitlist → Doc++

From Text to Data

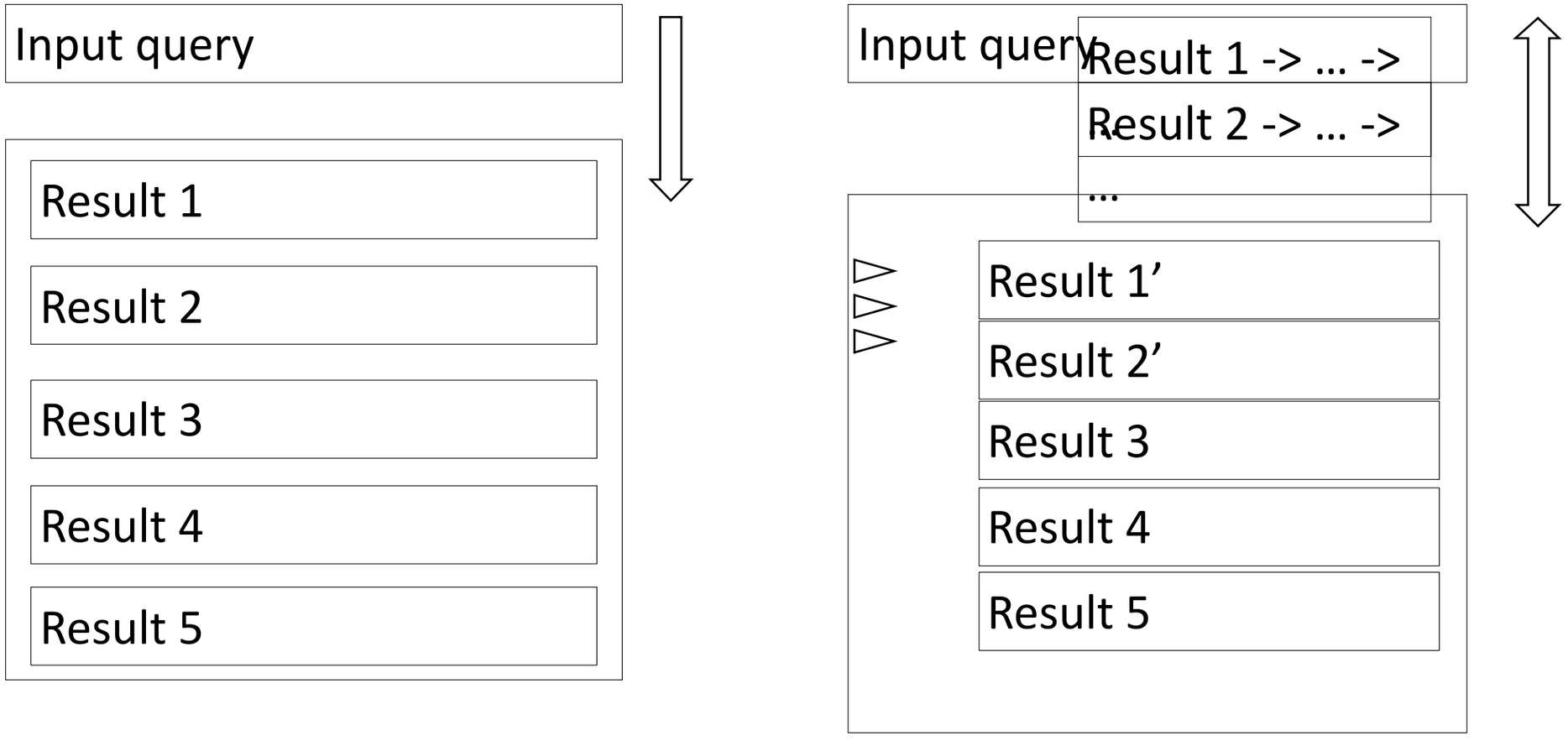
- Modern Web data is highly structured resulting in a massive multidimensional graph
  - Entities from massive knowledge resources, geo-temporal references, social network structure,…
  - E.g. Facebook: entities, relations between entities, locations and timestamps, social network, https://www.facebook.com/about/graphsearch
  - E.g., Hansards: speeches, speakers, role, party, full bio, …, http://translate.google.com/translate?u=http%3A%2F%2Fpoliticalmashup.nl%2F2013%2F10%2Fhilariteit-in-de-kamer%2F

Related Work: Facebook Graph Search

Graph Search

- Unifies the searcher **and** the information resources that played fundamentally different roles in classic IR
  - Searching information from a personal point of view — you are the query! — extreme personalization
- Highly structured and curated information space
- Beyond 2-term queries, beyond 10 blue links
- More interactive session

SERP vs IQEP: Incremental Query Exploration Page



Interaction Stages (1)

- Incremental Structured Query Formulation
  - Dynamic incremental query suggestion
  - Previews and surrogates from SERP moving to query suggestion
  - Generic entity types and instances (based on your activity)

Interaction Stages (II)

- Dynamic Structured Result Set Exploration
  - Unique for you and a point in time
  - Not a top-10 anymore; faceted exploration based on your interests
  - Far more dynamic structure than classic facet hierarchies

Evaluation

- How to evaluate the overall process? How strict are both stages divided?
- How to evaluate the first stage (query formulation)? How well the query expresses the need? How well the system promotes verbose queries? …
- How to evaluate the second stage (explore and exploit the result set)? How many filters are used? What is the engagement? …

Wrap Up (I)

- IR in a changing world
  - Core IR problems as relevant as ever!
  - Classic IR has value, but didn't keep up…
- Exciting new opportunities!
  - User/task context; Structured/heterogeneous data; Novel access tools (search stage sensitive)
  - Classic dichotomy of user and information broken
  - User-centered and system-centered IR reunited

Richer theories, richer experiments

Stephen Robertson
Microsoft Research Cambridge and
City University
ser@microsoft.com

A caricature

On the one hand we have the Cranfield / TREC tradition of experimental evaluation in IR

- a powerful paradigm for laboratory experimentation, but of limited scope

On the other hand, we have observational studies with real users

- realistic but of limited scale

[please do not take this dichotomy too literally!]

Richer models, richer experiments

Why develop richer models?

- because we want richer understanding of the phenomena
- as well as other useful predictions

Why design richer experiments?

- because we want to believe in our models
- and to enrich them further

A rich theory should have something to say *both* to lab experiments in the Cranfield/TREC tradition, *and* to observational studies

Social Book Search

Part II

Motivation

- Web gives access to a wealth of information different in quantity but also in character
  - Traditional IR focuses on ad hoc search & topical relevance
  - Social media have a different character:
    - different data (volume, subjective, opiniated)
    - different tasks (views/interpretations rather than facts)
    - different notions of relevance?
- What are such new tasks? and do these require new IR models?

amazon

Today's Deals | Gift Cards | Help

FREE Two-Day Shipping: See details

Shop by Department

Books

Hello, Jaap
Your Amazon

0

Wish List

Books | Advanced Search | Browse Subjects | New Releases | Best Sellers | The New York Times® Bestsellers | Libros en español | Bargain Books | Textbooks

Professional metadata

Browse Books

Amazon Kindle

- Kindle (Wi-Fi, 6")
- Kindle 3G (Free 3G+Wi-Fi, 6")
- Kindle DX (Free 3G, 9.7", Graphite)
- Kindle Books

Books

Browse the **best books of 2011 so far**, best-selling books, new and used textbooks, and Kindle books

Fall Reading

What's hot in Books this fall

Amazon Omnivoracious

Our Books editors' blog — Read new posts

Sell Back Your Copy

Receive an Amazon.com Gift Card for this item:

Modern Information Retrieval: The Concepts and ...

Social Book Search

LibraryThing

English | Help

Home | Zeitgeist | Talk | Groups | Local

Search site

Social content

Enter what you're reading or your whole library. It's an easy, library-quality catalog.

A community of 1,400,000 book lovers.

LibraryThing connects you to people who read what you do.

Become a Member?

name

password

SIGN IN

or sign in with...

amazon

Your Amazon.com | Today's Deals | Gift Cards | Help

Shop by Department

Search | Books | politics multiculturalism

Now more than traditional metadata

Last 90 days (62)
Coming Soon (22)

Department
Any Department
Books
Politics & Social

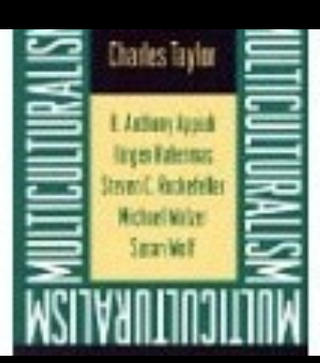
Showing 1 - 12 of 3,852 Results

Format

Paperback (2,133)	Hardcover (1,629)	Kindle Edition (11)	HTML (28)

UGC allows new search requests

Professional & Technical (1,041)
Gay & Lesbian (112)
Humor & Entertainment (216)
Law (324)
Literature & Fiction (388)
Parenting & Relationships (30)

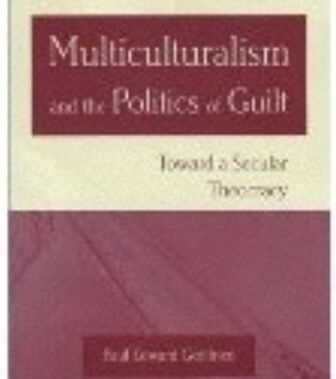


(5 customer reviews)

Formats	Price
Paperback Order in the next **6 hours** to get it by **Thursday, Sep 13.** Eligible for **FREE** Super Saver Shipping and **1 more promotion**	~~$27.95~~ **$17.84**
Kindle Edition	**$16.79**

Yet book search ignores reviews and tags!

Science & Math (197)
Medical Books (35)
Health, Fitness & Dieting (104)
Biographies & Memoirs (58)
Christian Books & Bibles (81)
Sports & Outdoors (16)
Computers & Technology (13)
Teens (20)



(16 customer reviews)

Formats	Price	Ne
Hardcover		$24.9

Sell this back for an Amazon.com Gift Card

Introduction

- Directly compare standard IR & social search
- Context: LibraryThing (LT) discussion forums
  - book requests and suggestions from LT forums
  - compare with familiar search tasks (ad hoc, known-item)
- Relevance in book search is very complex
  - readers have personal preferences, knowledge/reading level
  - want books that are engaging, funny, well-written, recent, educational, comprehensive, ...

Research Questions

- How does social book search compare to traditional search tasks?
  - Can we use forum suggestions for evaluation?
  - How is social book search related to traditional IR tasks?
  - How effective are professional metadata and UGC for book suggestion?
  - How do standard IR models cope with UGC?
  - Do users/judges prefer professional metadata or UGC for determining relevance and recommendation?

Social Book Search

- INEX Social Book Search Track
  - user query is NL statement, system returns ranked list
- Collection: 2.8 million book descriptions
  - ISBN is document ID
  - Mix of Amazon and LT data, fiction & non-fiction
  - crawled by Uni. Duisburg-Essen (for INEX iTrack)

Book Descriptions

Table 1: Statistics on the Amazon/LibraryThing collection

type	*min*	*max*	*median*	*mean*	*std. dev.*
Professional					
Dewey	0	1	1	0.61	0.49
Subject	0	29	1	0.66	0.72
BrowseNode	0	213	18	19.84	10.21
User-generated					
Tag	0	50	5	11.45	14.55
Rating/Review	0	100	0	5.05	14.98
Automatic					
Similar product	0	15	1	2.37	2.40

- Professional metadata is balanced, but low quantity (except Amazon browsing categories)
- UGC is skewed (popularity), higher mean (max cutoff for tag & review)

SBS 2011

LT Recommendations

- LT members discuss and ask for books on the LT discussion forums
  - groups for specific genres/subjects
  - member starts topic to ask for recommendations, others suggest books
  - requests are real needs and often very detailed
- Topics covers broad range of needs
  - subject, genre, author, known-item, `similar to book X'

Topic title

Narrative

Group name

Recommended books

LibraryThing
All topics
Hot topics

Your world
Grou
Y
Y

Book
All discussions
Your books

Post
Post a new topic

More options »

Politics of Multiculturalism Recommendations?

Political Philosophy

11 messages | Star this topic | Ignore topic | Jump to bottom (0 unread)

1 steve.clason Sep 26, 2010, 11:32

m new, and would appreciate any recommended reading on the politics of multiculturalism. Parekh's Rethinking Multiculturalism: Cultural Diversity and Political Theory (which I just finished) in the end left me unconvinced, though I did find much of value I thought he depended way too much on being able to talk out the details later. It may be that I found his writing style really irritating so adopted a defiant skepticism, but still...

Anyway, I've read Sen, Rawls, Habermas, and Nussbaum, still don't feel like I've wrapped my little brain around the issue very well and would appreciate any suggestions for further anyone might offer.

Reply | More

2 rsterling Edited: Sep 27, 2010, 1:31am

Will Kymlicka's Multicultural Citizenship is one of the key works within this literature, and his later work has built on but also modified his argument there. See his author page here. I think his latest ones are Multicultural Odysseys and Politics in the Vernacular.

Group: Political Philosophy

About

Th opic is not marked as pri rily about any work, au r or other topic.

Touchstones

Works

Rethinking Multiculturalism: Cultural Diversity and Political Theory by Bhikhu Parekh

Multicultural Citizenship by Will Kymlicka

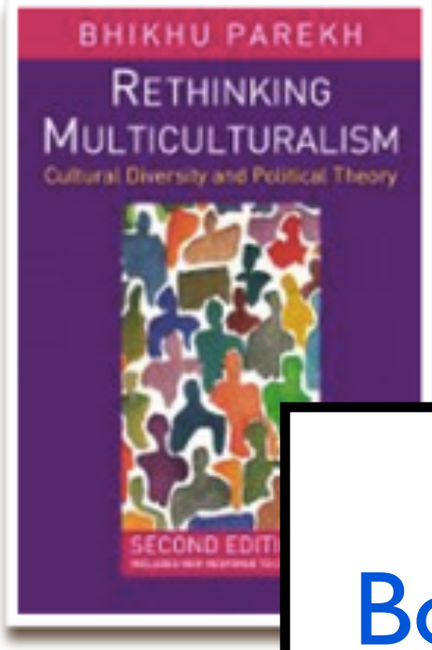
Multicultural Odysseys by Will Kymlicka

www.librarything.com/work/164382

Wayback Blogs Digital Humanities

LibraryThing

Home Profile Your books Add books Talk Groups Local More Zeitgeist



Rethinking Multiculturalism: Cultural Diversity and Political Theory

by Bhikhu Parekh

Members	Reviews	Popularity	Average rating	Mentions
				1

Book suggestion => LT work ID => ISBN

Main page

Work details

Reviews (1)
Recommendations
Members
Descriptions (3)
Conversations (1)
Common Knowledge
Popularity
Cover images

all m

kettenblume,

Tags

numbers

British literature contemporary theory cultural studies culture diversity E3 Educational Politics **identity** immigrati
liberalism Mread **multiculturalism** multiculturalism fall 2005 non-fiction philoso
Philosophy and Theory **pluralism** political philosophy political science **political theory** politi
race read in 2012 religion & secularism religion and politics

Creating Topics

- Can we use forum threads for evaluation?
  - Crawled 18,000 threads
  - Manually Selected ~1,000 threads with request for books
    - Used title of thread as topic/query on full-text index
    - Selected 211 threads with 50% recall of suggestions in top 1,000 (to filter non-descriptive titles)
- 211 topics used at INEX'11 SBS track
  - another 43 topics for training to get good pools

LT Forum Suggestions

- Most topic have few relevant books (suggestions): median 7, mean 11
- Incomplete judgments?
  - depends on number of members of discussion group
  - members only suggest books they know or have read
  - but small number may also indicate only best books are suggested (=recommendation rather than relevance?)

Mechanical Turk

- Judge top 10 results of submissions, at least on topical relevance
  - we ran MTurk experiment, selected 24 topics (12 fiction, 12 non-fiction)
  - workers read topic narrative, get 10 book descriptions, answer few questions per book
  - 272 HITs, 3 workers per HIT, $0.50 per HIT ($408 total)

Instructions

For this HIT you're shown a message from the LibraryThing discussions forums, of someone requesting recommendations for books in a certain genre or on a certain topic. We ask you to look at book descriptions of 10 books and decide whether you would recommend these books to this person. Each book description has 2 parts: 1) official description and 2) user-generated description.

1. the official description has information like title, author, publisher and subject classification information.
2. the user-generated description has user reviews and tags from Amazon and LibraryThing

We want to find out which part of the book description is helpful in determining whether you recommend a book or not.

Book request

Would you recommend the books below to a person with the following request:

Topic: Help: WWII pacific subs

LibraryThing group: Second World War History

Request: Can anyone recommend a good strategic level study of us sub campaign in pacific? All I seem to scare up is exploits of individual subs. I have ordered clay blairs big study but I would like something from this decade if it exists

Important: Please note that some of the pages below have been pre-judged by experts. Your answers will need to match at least 60% of the experts' answers to qualify for payment.

How familiar are you with the topic of the request? **Very unfamiliar** ○ ○ ○ ○ ○ **Very familiar**

MTurk Quality Control

- We checked MTurk judgments for quality
  - LT Agreement: at least one forum suggestion in each HIT
  - Relevance contradiction: answering book is rel. for Q1 then say it's non-rel. for Q3 (suggests random clicking)
  - Type contradiction: preferring UGC when there is none
  - Qualification: only US workers, at least 50 assignments, 95% approval rate (previous MTurk work of high quality)
- 7 HITs rejected
  - 3 low agreement, 4 for skipping questions

MTurk Stats

- Quality check:
  - 816 assignments, 133 workers (30 did 3 or more, half of assignments by 7 workers)
  - LT Agreement: over workers 0.52, over assignments 0.84, workers who did many have high agreement
  - 18 relevance contradictions (judgments were discarded)
  - no type contradictions
- Suggests workers were conscientious

System Centered Evaluation

- INEX'11 SBS: 4 teams submitted 22 runs
- Submissions were deduplicated
  - ISBN mapped to LT work ID
  - multiple ISBN can map to single LT work ID
  - we keep highest ranked and ignore the rest
- Official evaluation measure is nDCG@10

System Rank Correlation

Table 2: Kendall's τ and τ_{AP} system ranking correlations on nDCG@10 between the three sets of judgements (τ/τ_{AP})

	LT-24	AMT-24-Rel
LT-211	0.90/0.83	0.39/0.20
LT-24	–	0.36/0.19

- 3 sets of judgments
  - LT-211: all 211 forum topics, forum suggestions as judgments
  - LT-24: 24 selected topics, same judgments
  - AMT-24: 24 selected topics, MTurk topical relevance judgments
- MTurk judgments give different rankings
  - LT suggestions stable across topic selection
  - So different judgments? Or incomplete suggestions?

Incompleteness & MRR

- Best score of 0.481 on MRR for LT suggestions
  - if suggestions are incomplete, chance of getting high MRR is low
  - high MRR for a few topics could be accidental
  - over 211 topics, high MRR with incomplete judgments is unlikely
- Indicates that suggestions are relatively complete
  - LT suggestions can be used for evaluation, resulting in a test collection with high fidelity to real-world applications

Importance of Fields

- We created indexes of metadata fields:
  - Title (book title only)
  - Dewey (Dewey classification code)
  - Subject (LoC Subject heading)
  - BrowseNode (Amazon browse categories)
  - Review (all Amazon reviews for a book)
  - Tag set (all distinct LT tags)
  - Tag bag (LT tags, tag count = term frequency)
- Indri LM, dirichlet (mu=2,500), Krovetz, stopwords removed

Suggestions vs. MTurk

Table 4: MTurk and LT Forum evaluation of runs over different index fields

Field	AMT-Rel			AMT-Rec			AMT-Rel&Rec			LT-Sug		
	nDCG10	MAP	R@1000	nDCG10	MAP	R@1000	nDCG10	MAP	R@1000	nDCG10	MAP	R@1000
Title (field)	0.212	0.105	0.601	0.260	0.107	0.545	0.172	0.088	0.591	0.055	0.040	0.350
BrowseNode	0.096	0.052	0.322	0.142	0.056	0.321	0.083	0.046	0.328	0.043	0.031	0.261
Dewey	0.000	0.000	0.009	0.003	0.000	0.007	0.000	0.000	0.005	0.001	0.001	0.022
Subject	0.016	0.002	0.008	0.021	0.002	0.010	0.016	0.003	0.009	0.003	0.002	0.009
Review	**0.579**	**0.309**	0.720	**0.786**	**0.389**	**0.756**	**0.542**	**0.333**	**0.783**	**0.251**	**0.174**	**0.680**
Tag (set)	0.337	0.173	**0.744**	0.422	0.199	0.711	0.288	0.158	0.754	0.125	0.097	0.616
Tag (bag)	0.368	0.182	0.694	0.435	0.197	0.665	0.320	0.176	0.718	0.216	0.154	0.602

- Reviews most effective for all types of judgments (and currently ignored when searching Amazon or LT)
- Titles work for AMT relevance/rec. not for LT suggestions (title bias?)
- Standard LM is effective for UCG -- both for AMT and LT
- AMT-Rel&Rec closest to LT-Sug, but not identical judgment

Suggestions vs. Known-Item

Table 3: Known-item and forum suggestion evaluation of runs over different index fields

	Known-item			Forum suggestions		
Field	MRR	R@10	R@1000	MRR	R@10	R@1000
Title	0.414	0.540	**0.820**	0.118	0.048	0.350
BrowseNode	0.004	0.000	0.240	0.083	0.028	0.261
Dewey	0.000	0.000	0.000	0.002	0.000	0.022
Subject	0.010	0.020	0.020	0.012	0.002	0.009
Review	**0.480**	**0.680**	0.800	**0.382**	**0.227**	**0.680**
Tag (set)	0.118	0.220	0.540	0.213	0.125	0.616
Tag (bag)	0.227	0.400	0.560	0.342	0.178	0.602

- We created 50 Known-Item topics based on sampled suggestions from 211 topics
  - For *KI*: topics, Review and Title indexes are effective
  - For *LT forum* suggestions: Title index is much less effective

User Centered Analysis

Table 5: Impact of presence of reviews and tags on judgements

	Reviews		Tags	
	0 rev.	≥1 rev.	0 tags	≥10 tags
Top. Rel. (Q1)				
Not enough info.	0.37	0.01	0.09	0.09
Relevant	0.30	0.54	0.49	0.48
Recommend. (Q3)				
Not enough info.	0.53	0.01	0.14	0.12
Rel. + Rec.	0.22	0.51	0.46	0.45

- MTurk workers need at least one review, otherwise they don't have enough info for judgment
- Reviews are important for *both* relevance and recommendation
- Presence of tags has little impact (similar to subject headings?)

Metadata Preference

Table 6: Impact of the presence of reviews on metadata preference

	Q2. Relevance			Q4. Recommendation		
	all	0 rev.	≥1 rev.	all	0 rev.	≥1 rev.
Prof.	0.29	0.51	0.20	0.16	0.33	0.10
Equal	0.27	0.40	0.21	0.18	0.22	0.17
UGC	0.43	0.06	0.57	0.53	0.08	0.71
Skip	0.02	0.02	0.02	0.12	0.37	0.03

- When there are no reviews, workers use professional metadata for relevance, but can't make recommendation
- With at least one review, they prefer UGC and can make a recommendation

Wrap Up SBS'11

- Social Book Search is a new area of research
  - Book search over professional metadata and UGC
  - LT forums as window on real book search needs and suggestions
  - Suggestions are complete enough for evaluation!
    - ...but different from topical rel. judgments or known-item
- Standard retrieval models deal well with skewed distribution of UGC across descriptions
  - Although currently Amazon and LT ignore UGC in their search, it improves *both* ad hoc search and book suggestion
  - Reviews are important for both relevance and recommendation: provide information of a different nature than other metadata

SBS 2012

Continue with SBS

- Recall the motivation: Web data is different in quantity but also in character
  - Understand the differences in task and relevance judgments
  - Differentiate suggestions based on personal catalogue of topic creator
  - Analyze relevance dimensions: recommendation versus relevance

SBS Changes

- Collection: 2.8M book records
  - Extended with library records from BL and LOC
- Topics: forum requests from LibraryThing (LT)
  - User Profiles: LT catalogues of topic creators
- Judgments: suggestions from forum threads
  - graded by what's in catalogue of topic creator
- Task: submit ranked list of books that user wants to catalogue

Creating Topic Set

- We use the forum topics for evaluation
  - Crawled 60,000 topics, discarded topics without suggestions
  - Ran remaining topic titles as queries on full-text index
  - Discarded topics with no suggestion in top 1,000
  - Labelled 89 new topics that have request for books
  - Added them to 211 topics of last year
- That's this year's topic set (300 topics)
  - suggestions are used as relevance judgments

Suggestion? Relevant?

- Each suggestion is a human judgment, so relevant in some sense...
- Topic creator may not like every suggestion
- We compare suggestions to what topic creator has in her catalogue

User Profile/Catalogue

- Book selection is very personal
  - others cannot make good judgments (so traditional topical rel. judgments are out)
    - requires knowledge of user preference, likes/dislikes, mood, reading level, etc.
- Can we use profile information to reflect personal preference in judgments?

Topic title

Narrative

Group name

Recommended books

LibraryThing
All topics
Hot topics

Your world
Grou
Y
Y

Book
All discussions
Your books

Post
Post a new topic

More options »

Politics of Multiculturalism Recommendations?

Political Philosophy

11 messages | Star this topic | Ignore topic | Jump to bottom (0 unread)

1 steve.clason Sep 26, 2010, 11:32

m new, and would appreciate any recommended reading on the politics of multiculturalism. Parekh's Rethinking Multiculturalism: Cultural Diversity and Political Theory (which I just finished) in the end left me unconvinced, though I did find much of value I thought he depended way too much on being able to talk out the details later. It may be that I found his writing style really irritating so adopted a defiant skepticism, but still...

Anyway, I've read Sen, Rawls, Habermas, and Nussbaum, still don't feel like I've wrapped my little brain around the issue very well and would appreciate any suggestions for further anyone might offer.

Reply | More

2 rsterling Edited: Sep 27, 2010, 1:31am

Will Kymlicka's Multicultural Citizenship is one of the key works within this literature, and his later work has built on but also modified his argument there. See his author page here. I think his latest ones are Multicultural Odysseys and Politics in the Vernacular.

Group: Political Philosophy

About

Th opic is not marked as pri rily about any work, au r or other topic.

Touchstones

Works

Rethinking Multiculturalism: Cultural Diversity and Political Theory by Bhikhu Parekh

Multicultural Citizenship by Will Kymlicka

Multicultural Odysseys by Will Kymlicka

Pre/Post Cataloguing

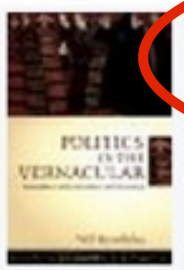
LibraryThing

Sign in / Join | En

Search site

All collections | List | Covers | Tags | Authors | Style A B C D E | Search this library

1 – 11 of 11

	Politics in the Vernacular: Nationalism, Multiculturalism, and Citizenship	Will Kymlicka	2001	democracy, political theory, immigration, multiculturalism	★★★	Oct 25, 2010
	Politics of Piety: The Islamic Revival and the Feminist Subject	Saba Mahmood	2005	feminist theory, feminism, political theory, multiculturalism		Aug 20, 2012
	Constitutional Patriotism	Jan-Werner Muller	2007	political theory, multiculturalism, european union, germany		Oct 31, 2010
	Rethinking Multiculturalism: Cultural Diversity and Political Theory	Bhikhu Parekh	2002	political theory, immigration, multiculturalism	★★	Sep 1, 2010

Post-Cataloguing

- 94 out of 300 topics have at least one Post-Catalogued Suggestion (PCS)
  - these are used for evaluation (graded judgments)
- Forum suggestions relevant (rv=1) except:
  - post-catalogued suggestions (rv=4)
  - creator-provided suggestions (rv=0)
  - avg. 16.2 suggestions per topic, 2 provided by creator, 1.7 post-catalogued, 12.5 have rv=1

Results (LT Forums)

Table 5. Evaluation results for the official submissions. Best scores are in bold

Run	MRR	nDCG@10	P@10	R@10	R@1000
p54.run2.all-topic-fields.all-doc-fields	0.3069	**0.1492**	0.1198	0.1527	0.5736
p54.run3.all-topic-fields.QIT.alpha0.99	0.3066	0.1488	0.1198	0.1527	0.5736
p4.inex2012SBS.xml_social.fb.10.50	**0.3616**	0.1437	0.1219	0.1494	**0.5775**
p62.B_IT30_30	0.3410	0.1339	**0.1260**	**0.1659**	0.5130
p4.inex2012SBS.xml_social	0.3256	0.1297	0.1135	0.1476	0.5588
p62.mrf-booklike	0.3584	0.1295	0.1250	0.1514	0.5242
p54.run5.title.II.alpha0.94	0.2558	0.1173	0.1073	0.1289	0.4891
p62.IOT30	0.2933	0.1141	0.1240	0.1503	0.5864
p62.IT30	0.2999	0.1082	0.1187	0.1426	0.5864
p54.run6.title.II.alpha0.97	0.2392	0.0958	0.0823	0.0941	0.4891
p62.lcm-2	0.2149	0.0901	0.0667	0.1026	0.5054
p100.sb_g0	0.2394	0.0884	0.0844	0.1145	0.5524
p54.run4.title.QIT.alpha0.65	0.1762	0.0875	0.0719	0.0949	0.4891
p100.sb_g_ttl_nar0	0.1581	0.0740	0.0594	0.0939	0.4634
p54.run1.title.all-doc-fields	0.1341	0.0678	0.0583	0.0729	0.4891
p100.sb_2xsh_ttl_nar0	0.0157	0.0057	0.0021	0.0022	0.0393
p100.sb_2xsh0	0.0199	0.0042	0.0021	0.0020	0.0647

Results (LT Forums)

- RSLIS Copenhagen has best run: all topic fields as query, all document fields in index
- RSLIS Copenhagen & LIA Avignon used user profile information for some submissions...
- ... but not for their best runs
- Official submissions do not show added value of user profiles (yet)

Mechanical Turk

- LT suggestions are relatively complete
  - high MRR scores over last year's 211 topics means suggestions can't be highly incomplete
- Analyze suggestions with MTurk
  - complex relevance judgments for top 10 results of submissions, covering different aspects of relevance
  - workers read topic narrative, get 10 book descriptions, answer few questions per book

MTurk Judgments (1/2)

- Judge top-k results of submitted runs
  - Add at least 1 suggestion from forum (so one known good book)
- We want to know:
  - is book relevant/recommendable?
  - what information is used for the judgment? (metadata fields: title, reviews, ...)
  - what relevance aspects are important? (subject, genre, recency, comprehensiveness, engagement, etc.)

MTurk Judgments (2/2)

- 37 runs from 5 groups (2 yrs), top 10 pool, 25 topics
- 10 books/HIT for $0.50
- 326 HITs, 3 workers/HIT (US only, 95% approval)
- Cost: 326 x 3 x $0.50 = $489 + 10% = $537
- Mean time/HIT: 11:59 (median 9:17)
- 55% of workers spent 6-15 minutes

MTurk Design

Topic: Tudor Fiction

LibraryThing group: English History - Tudor through Edwardian

Request: Any Tudor fiction suggestions? I'm already completely wrapped up in Philippa Gregory's books about the women of Henry VIII's court: The Constant Princess, The Other Boleyn Girl, and The Boleyn Inheritance. What others do people recommend??

Important: Please note that some of the books have been pre-judged by experts. Your answers will need to match at least 60% of the experts' answers for payment.

How familiar are you with the topic of the request? **Very unfamiliar** ○ ○ ○ ○ ○ **Very familiar**

Which aspects of books are relevant to this request:

- ☐ Subject
- ☐ Genre
- ☐ Recency
- ☐ Comprehensiveness
- ☐ Reading level (difficulty)
- ☐ Objectivity
- ☐ Engagement/fun factor
- ☐ Novelty / familiarity

Request:

Title information | Catalogue information | Reviews | Tags

Topic: Tudor Fiction

LibraryThing group: English History - Tudor through Edwardian

Request: Any Tudor fiction suggestions? I'm already completely wrapped up in Philippa Gregory's books about the women of Henry VIII's court: The Constant Princess, The Other Boleyn Girl, and The Boleyn Inheritance. What others do people recommend??

With the forum request in mind, please look at the different parts of the book description on the right and answer the questions below.

Q1. Which of the requested aspects does this book satisfy?

- ☐ Subject
- ☐ Genre
- ☐ Recency
- ☐ Comprehensiveness
- ☐ Reading level
- ☐ Objectivity
- ☐ Engaging/fun
- ☐ Novel/familiar

Q2. What book information did you find useful determine this?

Type	Very	Somewhat	Not	Didn't use
Title information	○	○	○	○
Catalogue information	○	○	○	○
User reviews	○	○	○	○
User tags	○	○	○	○

Q3. Would you recommend this book to the requestor?

Yes, it perfectly fits the request and looks like a great book.

Title: The Last Wife of Henry VIII: A Novel

Author:	Carolly Erickson
Reader:	Terry Donnelly
Publication date:	2006-10-03
Publisher:	Macmillan Audio
Edition:	Unabridged
Price:	$39.95
ISBN:	1593979622
EAN:	9781593979621
Dimension:	160 x 510 x 580 mm

Worker Judgments

Choice	Rel	Rec
Yes, it perfectly fits the request and looks like a great book		
Probably, it doesn't fit the request completely, but still looks worth reading		
No, it fits the request, but doesn't look like it's worth reading		
No, it doesn't match the request		
No, the requester already has (read) this book		
I don't know, the book description is too incomplete or too vague		

Quality Control

- Some quality control measures:
  - Agreement: 1 forum suggestion/HIT (should be relevant). Agreement > 60% once worker did 3 HITs or more.
  - Consistency: 1) if book aspect is relevant (Q1), some description parts must be relevant (Q2), 2) if topic. rel. (Q3) -> subject or genre aspects must be ticked (Q1)
- Unfortunately, over 600 HITs approved automatically (too late to check work)
  - But in first 350 assignments only 8 had to be rejected

Agreement

	min	max	med.	mean	std.dev
rel+rec	0	1	0,6	0,56	0,21
rel	0,1	1	0,7	0,72	0,19
rec	0,2	1	0,7	0,73	0,17

- Q3: would you recommend this book?
- 1) rec+rel, 2) rec+nonrel, 3) nonrec+rel, 4) nonrec+nonrel
- pairwise per HIT (3 workers = 3 pairs)

Agreement with Suggestions

	All books		LT Suggestions	
total	9780	100%	1247	100%
unknown	522	5%	24	2%
unrelated	3491	36%	68	4%
related	1952	20%	213	17%
relevant	3865	41%	942	76%
recommended	5212	56%	976	78%

- Agreement with suggestions is high
- Suggestions more often relevant and recommended than other books

Relevance Aspects

- Subject is relevant for almost all requests (93%)
  - and a necessary condition for a book's relevance.
- Other important aspects
  - genre (71%), comprehensiveness (39%), engagement (21%)
  - these aspects are usually ignored in trad. IR evaluation
- Workers select relevance aspects for books that they don't select for the forum request
  - perhaps to point out recommendable qualities of a book, regardless of whether requester asked for it or not

Relevant or Recommended?

- Difference between relevance and recommendation:
  - 22% of recommended books is not relevant but related (recommended on other qualities)
  - 5% is relevant but not recommended
- 92% of LT suggestions is relevant or related, 76% is topically relevant. Request could be about genre, not subject.

Browsing Behavior

- We log clicks and track which parts of description are browsed by workers
  - title information always shown, other parts require click.
- How often is book part clicked?
  - catalogue (76%), reviews (72%), and tags (63%)
  - For 1 in 4 books judges only use title information

Usefulness of Metadata

	Very	Some	Not	Not used	Not present
Title	0,65	0,25	0,07	0,02	0
Catalogue	0,37	0,27	0,18	0,18	0,07
Reviews	0,29	0,15	0,24	0,32	0,4
Tags	0,26	0,24	0,18	0,32	0,07

- How useful is metadata type for judgment?
- Title information is very important (book judged by cover)
- No reviews for 40% of books. When present they're useful.

Conclusions

- Social Book Search is a new area of research
  - Social media integration raises many open questions
  - Ideal playground for studying professional metadata and UGC
  - LT forums provide real book search needs and suggestions
- Studied the character of topic-based recommendation
  - Relevance aspects: Subject (93%), genre (73%), comprehensiveness (39%), engagement (21%)
  - Information: Title, catalogue (76%), reviews (72%), and tags (63%)
  - Relevant/Recommended: 5% is rel+not rec.22% is rec+not rel.

SBS 2013/2014

Social Book Search (SBS)

- Book Search with social data:
  - book information from trained professionals and users
  - user-generated content
    - almost always ignored in search index
    - uncontrolled, inconsistent, unlike library catalogues (well...)
    - unbalanced, skewed towards popular books
- but:
  - “The user-generated review was so enthusiastic, I would recommend it just based on that.”

Search or Recommendation?

- Two paradigms:
  - Retrieval (trad. catalogue)
  - Recommendation (user data)
- Amazon, GoodReads, LibraryThing
  - Offer both paradigms for search and discovery
  - Retrieval ignores user profile and user-gen. content
  - Recommendation ignores specific information need
  - Book search requires both?

Suggestions and Relevance?

- Discussion participants form crowd of recommenders
  - Thread is discussion of "best" books?
  - Each suggestion is a human judgement, so relevant in some sense (2011)
  - Topic creator may not like every suggestion: check if topic creator adds to catalogue (2012)
  - Not all books mentioned may be intended as suggestions: analyse and label suggestions (2013/2014)

Creating Topic Set

- We use the forum topics for evaluation
  - we built a new, simple interface to gather annotations
  - 8 students from RSLIS, OUC & AAU annotated topics and suggestions
  - many thanks to Toine Bogers, Birgir Larsen and Michael Preminger for recruiting and paying them!
- 2,646 topics annotated
  - 944 are book search topics, 680 with suggestions
  - annotated suggestions are used as relevance judgements

Example Topic

<topic id="99309">

<title>Politics of Multiculturalism Recommendations?</title>

<query>Politics of Multiculturalism</query>

<group>Political Philosophy</group>

<narrative>I'm new, and would appreciate any recommended …</narrative>

<catalog>

<book><LT_id>9036</LT_id><entry_date>2007-09</entry_date>…</book>…

</catalog>

</topic>

Distribution of Aspects

# aspects	# topics	%
Accessibility	106	16
Content	523	77
Engagement	154	23
Familiarity	261	38
Known-item	97	14
Metadata	177	26
Novelty	29	4
Socio-cultural	108	16

- Content is most important relevance aspect (topical relevance!)
- Familiarity is important (recommendation!)
- Engagement is hard to express in query (reason for social search?)

Number of Aspects

# aspects	# topics	%
1	191	28
2	260	38
3	183	27
4	37	5
5	7	1
6	2	0

- Topics have multiple relevance aspects
- Suggests multiple metadata sources are needed

From Suggestions To Relevance Judgements

- We used a complex decision tree to map annotated suggestions to relevance values
  - opinion of topic creator overrules others
  - opinions of users who read book overrule opinions of those who didn't
  - opinion of multiple people weights stronger than opinion of single person

User Profile/Catalogue

- Book selection is very personal
  - others cannot make good judgements for user,
  - traditional topical rel. judgements are problematic (Koolen & Kamps, CIKM 2012)
- Requires knowledge of user
  - preference, knowledge, reading level, ...
  - Can we use personal catalogue to reflect personal preference in suggestions?

Topic Creator Profiles

- We distributed user profiles for 680 topics
  - plus anonymised profiles of 93,976 other users
- Allows
  - content based recommendation (topical, temporal, ...)
  - collaborative filtering

Example Profiles

# user-id	book-id	entry-date	rating	tags
u8218518	952822	2012-02	0.0	
u8218518	3344349	2012-02	0.0	
u8218518	2317257	2012-02	10.0	
u8218518	6415999	2012-02	8.0	
u8218518	5525956	2012-09	0.0	
u8218518	842432	2012-02	8.0	
u8218518	3171103	2012-02	4.0	
u8780837	542201	2009-05	10.0	tarot
u9054475	5403381	2010-11	10.0	adventure, potter, philosopher, Harry, stone

…

Profile Statistics

Table 4. User profile statistics of the topic creators and all other users.

Type	N	total	min	max	median	mean	stdev
Topic Creators							
Pre-catalogued	680	399,147	1	5884	239	587	927
Post-catalogued	680	209,289	1	5619	114	308	499
Total catalogue	680	608,436	2	8563	432	895	1202
All users							
Others	93,976	33,503,999	1	41,792	134	357	704
Total	94,656	34,112,435	1	41,792	135	360	710

Evaluation Results

Table 5. Evaluation results for the official submissions. Best scores are in bold. Runs marked with * are manual runs.

Group	Run	nDCG@10	P@10	MRR	MAP	Profiles
USTB	run6.SimQuery1000.rerank_all.L2R_RandomForest	**0.303**	**0.464**	**0.232**	0.390	No
USTB	run4.newXml.rerank_all.L2R_RandomForest	0.142	0.258	0.102	0.390	No
UJM	326	0.142	0.275	0.107	0.426	No
USTB	run3.newXml.rerank_all.L2R_Coordinate	0.138	0.256	0.101	0.390	No
USTB	run5.newXml.rerank_all.L2R_RankNet	0.133	0.246	0.098	0.390	No
USTB	run2.newXml.rerank_T	0.131	0.246	0.096	0.390	No
USTB	run1.newXml.feedback	0.128	0.246	0.095	0.390	No
LSIS	InL2	0.128	0.236	0.101	0.441	No
AAU	run1.all-plus-query.all-doc-fields	0.127	0.239	0.097	**0.444**	No
AAU	run3.all-plus-query.all-doc-fields	0.120	0.227	0.090	0.425	No
CYUT	Type2QTGN	0.119	0.246	0.086	0.340	No
CYUT	0.95AverageType2QTGN	0.119	0.243	0.085	0.332	No
UJM	328	0.117	0.226	0.088	0.392	Yes
UJM	329	0.116	0.217	0.087	0.392	Yes
UJM	325	0.115	0.214	0.087	0.392	Yes
LSIS	InL2Feedback	0.114	0.230	0.094	0.434	No
UJM	324	0.112	0.214	0.086	0.392	No
LSIS	InL2tagFeedback	0.102	0.212	0.075	0.388	No
UvA	inex14.ti_qu.fb.10.50.5000	0.097	0.179	0.073	0.421	No
UMD	Full_TQG_fb.10.50_0.0000227_50	0.097	0.188	0.069	0.328	Yes
UMD	Social_TQG_fb.10.50_0.0000222_50	0.096	0.184	0.067	0.327	Yes

Analysis

- All top runs use User-Generated Content
  - best run: USTB (Beijing), hybrid learning to rank
  - no run uses prof. metadata different from user content
- MRR score are low
  - Many hard topics, all systems score zero
- Exploiting profile information
  - Some runs use profile information, but none in top ranks

Tasks & Relevance Aspects

- 8 aspects:
  - accessibility, content, engagement, familiarity, known-item, metadata, novelty, socio-cultural
- Aspects reveal relation between retrieval and recommendation?
  - Content aspect typical for search, familiarity for retrieval?
  - many topics contain both aspects

Topic Categories

- Known-item (KI) contains all known-item topics (202 topics)
- Search (S): contains topics with content but not familiarity (338 topics)
- Search and Recommendation (SR) contains topics with both content and familiarity topics (260 topics)
- Recommendation (R) contains topics with familiarity, but not content (66 topics)
- Context (C) contains all topics without content, familiarity, and known-item (78 topics)

Known-Item

3 cooks baking a cake

Librarians who LibraryThing

3 messages | ★ Star this topic | × Ignore topic | Jump to first unread (3 unread)

This topic is currently marked as "dormant"—the last message is more than 90 days old. You can revive it by posting a reply.

1 TheNovelWorld Dec 16, 2011, 7:15pm

That's all I have to go on for this picture book request. Any ideas on the title or author?

Thanks so much!

Reply | More

Search

Recommendations for novels about the build-up to World War II

Historical Fiction

6 messages | ★ Star this topic | × Ignore topic | Jump to bottom (0 unread)

This topic is currently marked as "dormant"—the last message is more than 90 days old. You can revive it by posting a reply.

1 sturlington Dec 9, 2011, 10:05am

I am interested in reading some fiction set immediately before WWII depicting the build-up to the war, particularly from a German perspective. Can anyone recommend a book? I know this is fairly specific...

Reply | More

Recommendation

Catholic Children's Books

Catholic Tradition

7 messages | ★ Star this topic | × Ignore topic | Jump to first unread (7 unread)

This topic is currently marked as "dormant"—the last message is more than 90 days old. You can revive it by posting a reply.

1 CaleeMLee Oct 18, 2011, 12:14am

I always get my kids (preschool age) a few books for Christmas and would love your suggestions. My daughter is just getting into listening to chapter books so I think we'll start the Narnia books. Any other books out there you'd recommend? She loved The Holy Twins by Kathleen Norris.

More

Search+Recommendation

Merchant/Trader SF

Science Fiction Fans

33 messages | ★ Star this topic | × Ignore topic | Jump to bottom (0 unread)

This topic is currently marked as "dormant"—the last message is more than 90 days old. You can revive it by posting a reply.

1 cosmicdolphin Jan 1, 2012, 8:52pm

I'd like suggestions for good Merchant/Trader SF? I've read Cherryh's Merchanter books, and I'm currently reading the Andre Norton Solar Queen books (both of which are good in their own ways)

Any other SF with Traders/Merchants as the primary focus of the story?

Thanks

Reply | More

Context

Great books for year 8 students

Librarians who LibraryThing

5 messages | ★ Star this topic | × Ignore topic | Jump to first unread (5 unread)

This topic is currently marked as "dormant"—the last message is more than 90 days old. You can revive it by posting a reply.

1 jothom Oct 12, 2011, 9:12pm

I have been asked to recommend a class text for year 8 English classes. I would really like to have something new that kids will love. Does anyone have any suggestions? Not too long, and fairly accessible as we have lots of kids with literacy problems.

Reply | More

Types & Performance

		nDCG@10	
	# topics	with aspect	w/out aspect
Context	56	0,120	0,100
Known-item	97	0,208	0,084
Recommend	53	0,126	0,100
Search	274	0,094	0,107
Search+Rec	200	0,050	0,123

- Note: some categories harder than others, search+rec hardest

User Catalogues

- For 589 of 680 topics, user has catalogue
  - For 91, we don't know (private vs. empty)
- Profiles give us:
  - catalogue size: how many books catalogued before posting request on forum
  - popularity of books in catalogue: some read only popular books, others obscure books or mix

Catalogue Size

	All	Known-Item	Context	S+R	Rec	Search
Pre-Topic	84	0	38	100	104	177
Post-Topic	65	4	80	65	81	108
Pre+Post	197	16	155	201	195	415

- Median catalogue size related to type of request
- Heavy readers can explain more precisely what they want?
  - specific aspects instead of “similar to X, Y and Z”
- Categories represent different stages in search process?

Impact of Catalogue Size

		nDCG@10		
	#tpcs	small cat.	big cat.	no cat.
Context	24 / 24 / 8	0,127	0,147	0,016
Known-item	24 / 24 / 9	0,258	0,146	0,262
Recommend	23 / 24 / 6	0,176	0,068	0,168
Search	119 / 119 / 36	0,068	0,120	0,094
Search+Rec	84 / 84 / 32	0,032	0,068	0,053
All	294 / 295 / 91	0,100	0,106	0,094

- Search topics of heavy readers better formulated? More specific?

Book Popularity

- Profiles provide popularity information
  - Some users catalogue only popular books
  - others only obscure books or a mix
  - popular books have more UGC
  - Do readers of obscure books have different needs or express them differently than readers of popular books?
  - we rank searchers by the median popularity of books in their catalogue

Median Book Popularity

		nDCG@10		
	#tpcs	least pop.	most pop.	no cat.
Context	24 / 24 / 8	0,119	0,155	0,016
Known-item	44 / 44 / 9	0,199	0,205	0,262
Recommend	23 / 24 / 6	0,094	0,146	0,168
Search	119 / 119 / 36	0,116	0,071	0,094
Search+Rec	84 / 84 / 32	0,046	0,053	0,053
All	294 / 295 / 91	0,111	0,095	0,094

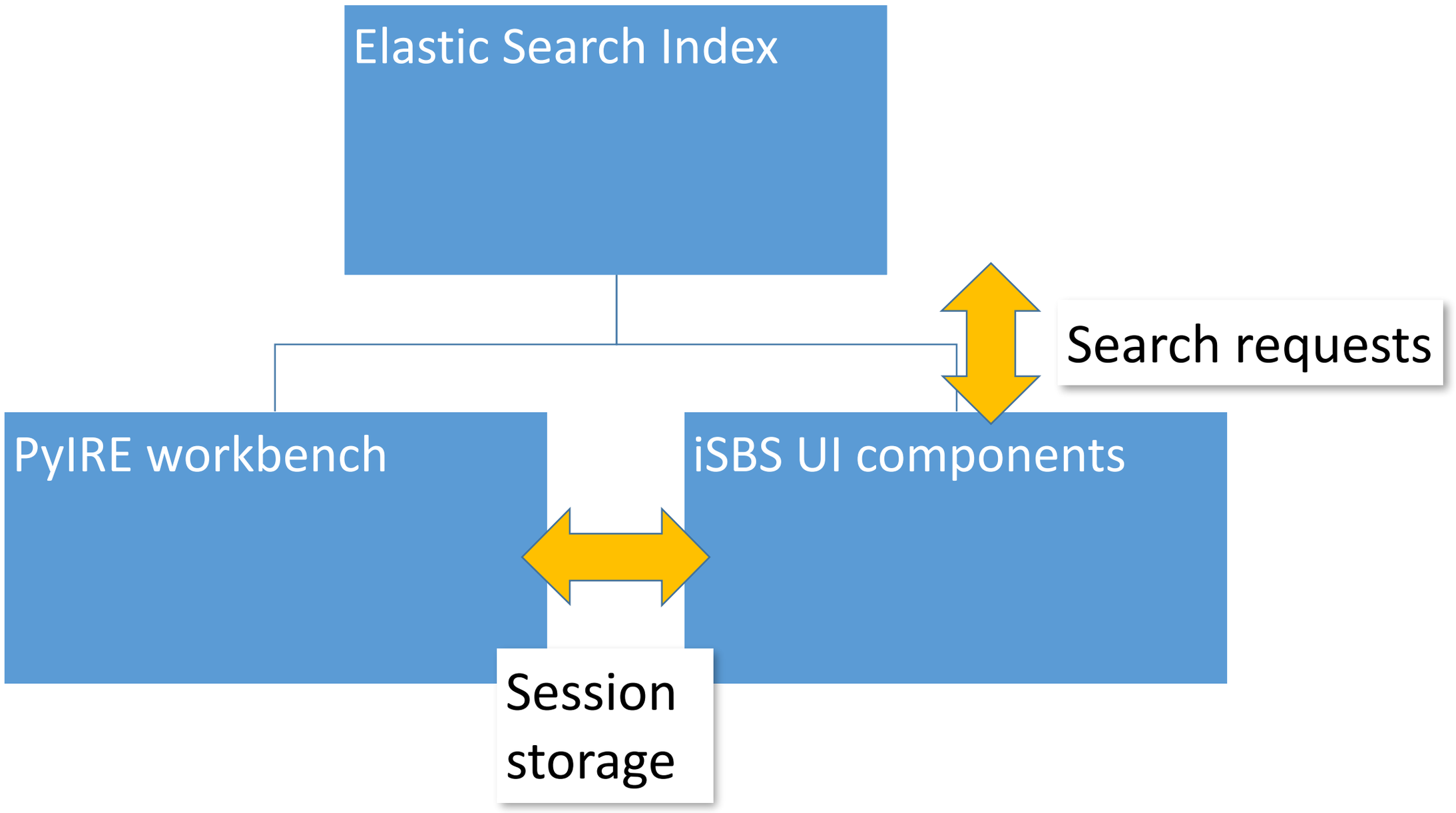
- Note: popularity affects context, recommend and search categories

SBS'14 Conclusions

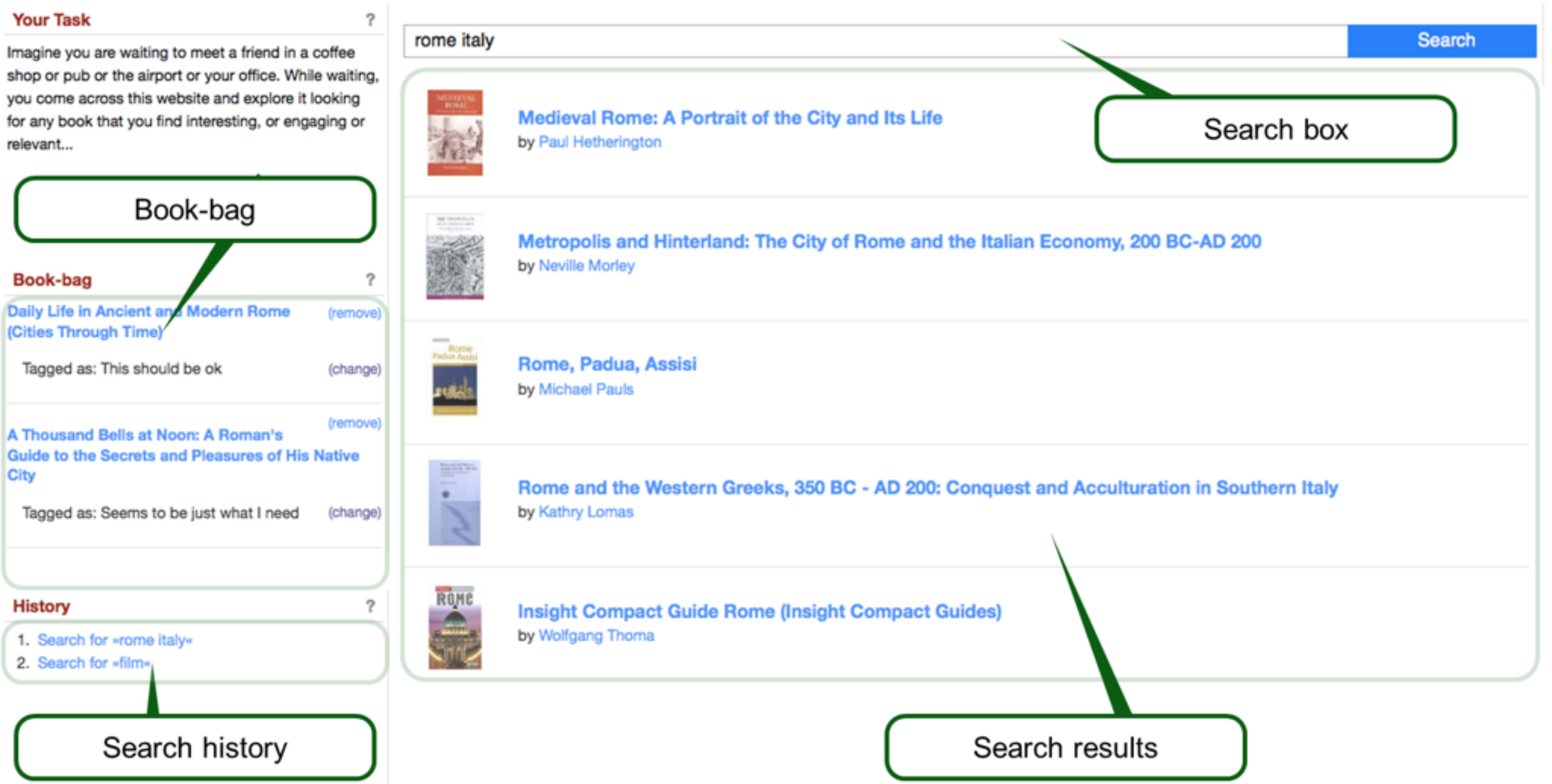
- Topic set covers many request types
  - mix of tasks: topic search, recommendation, known-item
- User profiles provide another perspective on requests
  - request type related to catalogue size and popularity of books
- Catalogue size and popularity of books affect performance on topics
  - Related to how specific/concrete or well-expressed it is?

Interactive Social Book Search

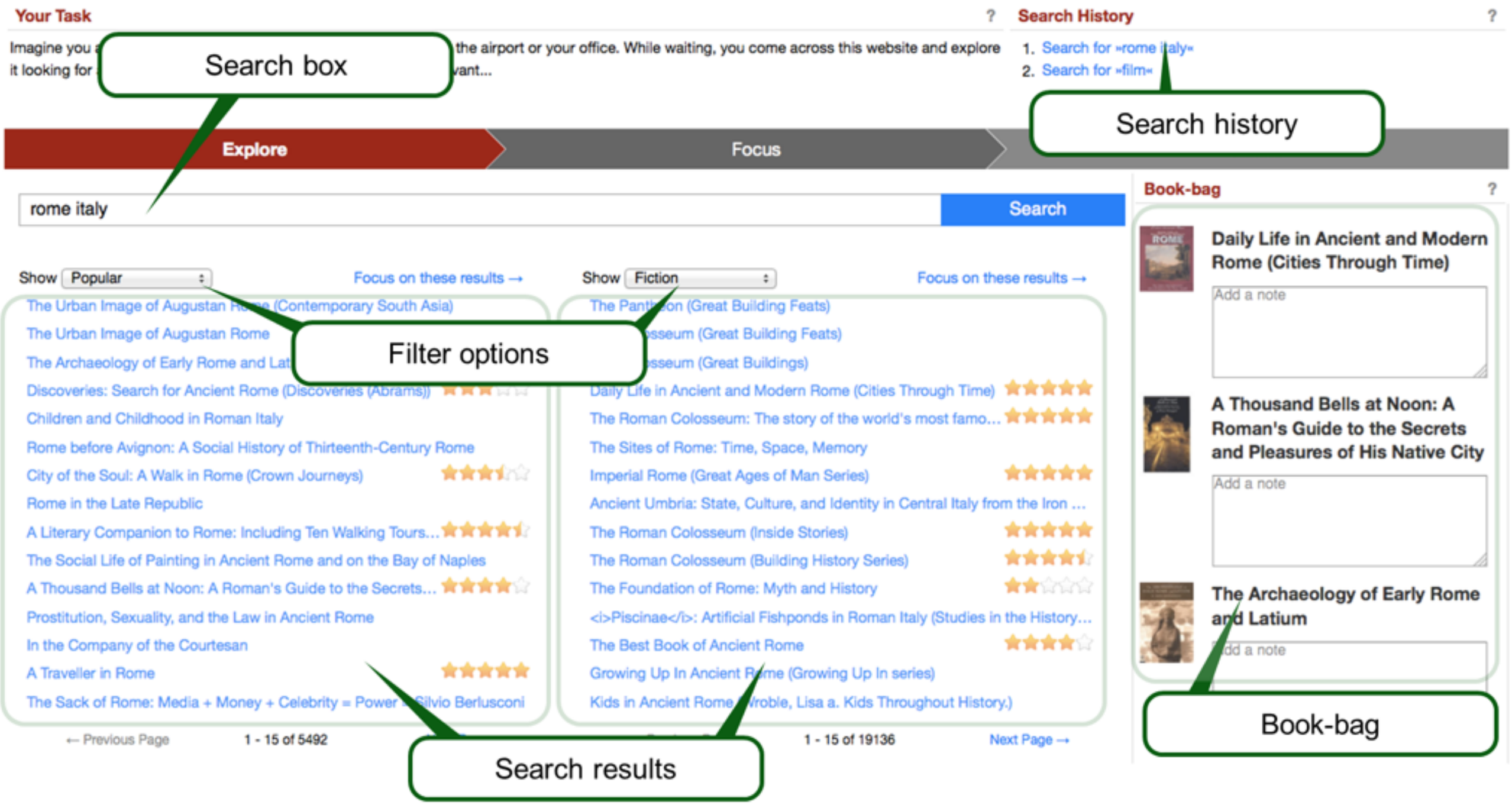
Information Seeking Stage Aware Systems



Baseline System



Multistage: Explore



Multistage: Focus

Your Task ?

Imagine you a… the airport or your office. While waiting, you come across this website and explore it looking for … vant...

Search History ?

1. Search for »rome italy«
2. Search for »film«

Search box

Search history

Explore | Focus

rome italy | Search

Book-bag ?

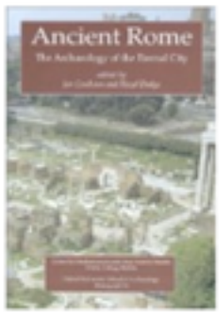
Filter by Topics

History	19631
Fiction	19136
Literature	17875
World	13178
Nonfiction	12011
Europe	10678
Reference	8218
Contemporary	7263
Science	6811
Social	6640

Filter by User tags

History	16857
Non-fiction	15508
Unread	15180
Fiction	14[illegible]2
Read	[illegible]
Own	[illegible]
Tbr	[illegible]
Italy	9190
Novel	8183

Filter options



Save to book bag

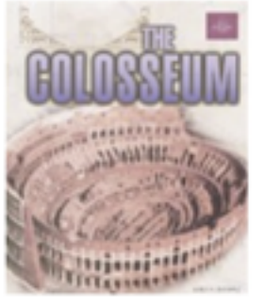
Ancient Rome: The Archaeology of the Eternal City (Monograph)

by Jon Coulston and Hazel Dodge

Description | Publication | 1 Review | 10 Tags

Product Description

A major new book on the archaeology of Rome. The chapters, by an impressive list of contributors, are written to be as up-to-date and useful as possible, detailing a plethora of new research. There are new maps for the topography and monuments of Rome, a huge research bibliography containing 1,700 titles, and numerous rich illustrations. Essential for all s Show more



The Colosseum (Great Building Feats) Unrated

by Lesley A. Dutemple

Description | Publication | No review

Type	Hardcover	*Published by*	Lerner Publishing Group
Pages	96	*ISBN*	0822546930

Search results

Daily Life in Ancient and Modern Rome (Cities Through Time)

Add a note

A Thousand Bells at Noon: A Roman's Guide to the Secrets and Pleasures of His Native City

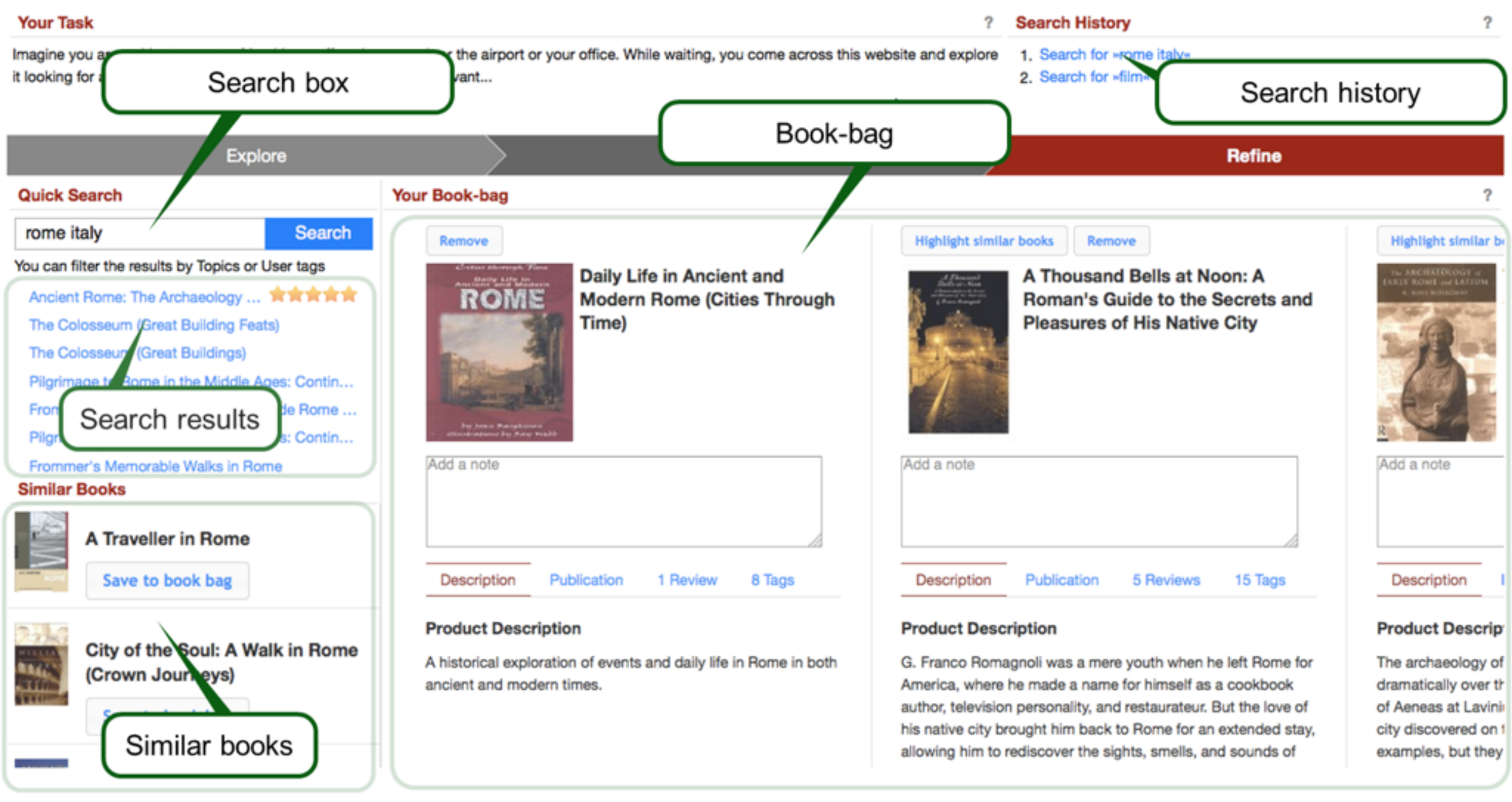
Add a note

The Archaeology of Early Rome and Latium

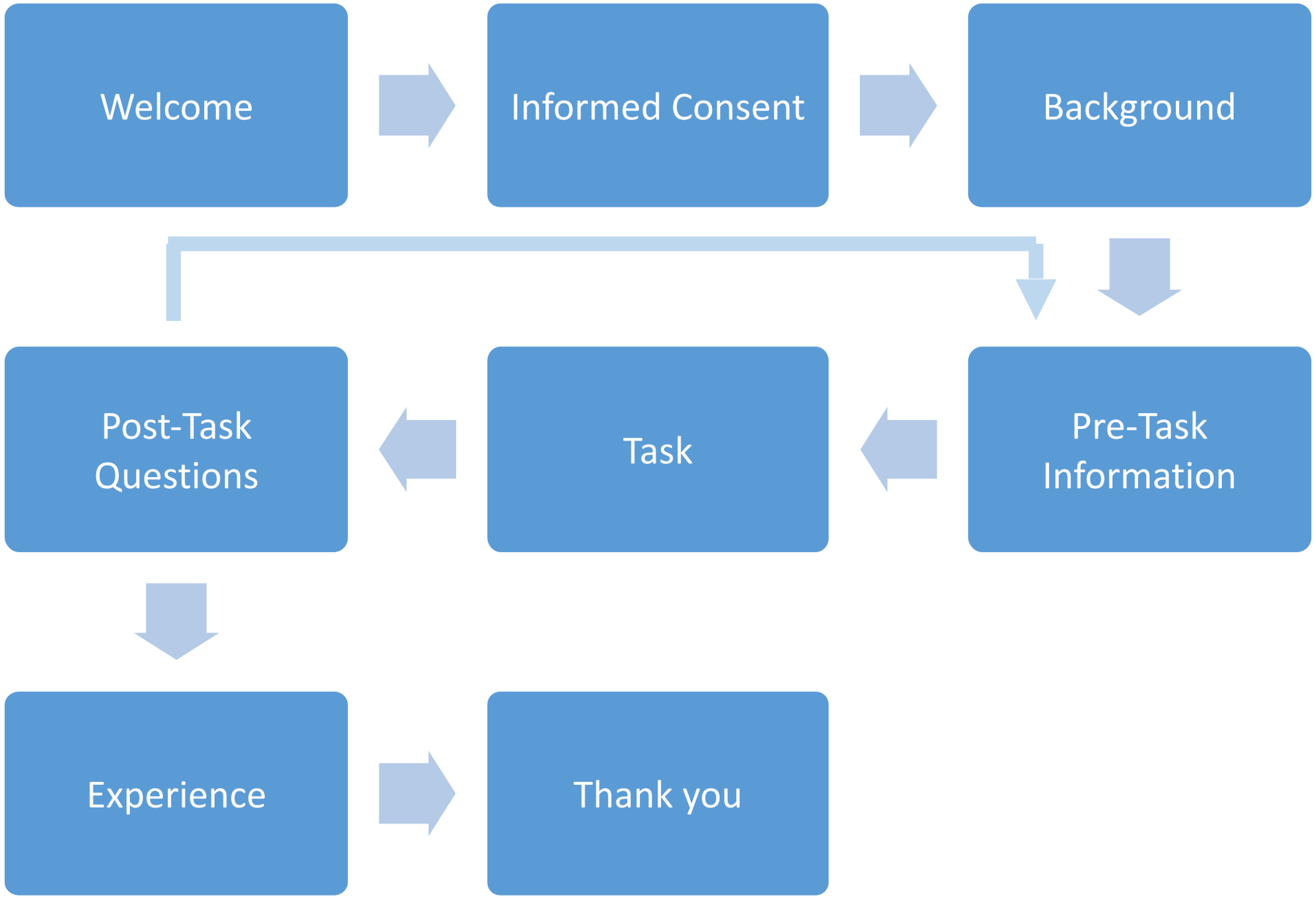
Add a note

Book-bag

Multistage: Refine



User Study with goal/non-goal oriented tasks



41 Test Persons in "pilot" run

	Goal-oriented		Non-goal	
Session Length				
Baseline	6:25	(3:42)	3:42	(3:45)
Multi-Stage	3:35	(4:24)	2:40	(6:21)
Number of Queries				
Baseline	4	(5.5)	2	(4.5)
Multi-Stage	3	(2.75)	2	(3)
Number of Books Viewed				
Baseline	4	(5.5)	2	(4.5)
Multi-Stage	3	(2.75)	2	(3)
Number of Books Collected				
Baseline	3	(3)	1	(2)
Multi-Stage	3.5	(3)	2	(3)

Overall: less queries, more exploration, more books collected in *multi-stage UI*

Wrap Up (II)

- SBS as playground to study (aspects of) non-classic IR
  - Traditional vs. UGC
  - Rich context of user and request
  - Mixing search and recommendation
- Bridge system-centric and user-centric research (iSBS)
  - Running at CLEF, ECIR WS, RecSys WS in 2015 (tbc)

Take Home Messages

- Information access problems are more relevant than ever
  - Classic abstraction important, but limited
  - Current systems framed by past: changes have just begon!
  - Many new opportunities: best time ever to do a PhD in IR!
- IR revolution with the impact of ‘cranfield’ is happening
- Not “*anything goes*”
  - Scientific understanding requires generalization of results
  - Abstraction to research task, focus on 1 aspect crucial
  - Science should lead and not follow industry…

CLEF 2015 Conference and Labs of the Evaluation Forum

Information Access Evaluation meets Multilinguality, Multimodality, and Interaction

8-11 September 2015, Toulouse - France

CLEF 2015 Toulouse

Home
Programme
Keynote Talks
Conference
Accepted Papers
Call for Papers
Labs
Venue
About Toulouse
Accommodation
Publications
WN Instructions
Labs Overview
Important dates
Organization
Lab Organising
Committee

Welcome

CLEF 2015 is the sixth CLEF conference continuing the popular CLEF campaigns which have run since 2000 contributing to the systematic evaluation of information access systems, primarily through experimentation on shared tasks.

Building on the format first introduced in 2010, CLEF 2015 consists of an independent peer-reviewed conference on a broad range of issues in the fields of multilingual and multimodal information access evaluation, and a set of labs and workshops designed to test different aspects of mono and cross-language Information retrieval systems. Together, the conference and the lab series will maintain and expand upon the CLEF tradition of community-based evaluation and discussion on evaluation issues.

CLEF 2015 will be hosted by the Institut de Recherche en Informatique de Toulouse (IRIT), University of Toulouse, France, 8-11 September 2015.

Tweets Follow

CLEF 2015 @clef2015 — 19 Sep
Labs accepted for #clef2015: eHealth, ImageCLEF, LifeCLEF, LL4IR, NEWSREEL, PAN, QA, SBS. Congrats! @livinglabsnet @CLEFNEWSREEL @imageclef.
Expand

Iadh Ounis @iadh — 18 Sep
8 labs will run in #clef2015 - RepLab is temporarily retired, INEX is morphing into Social Book Search #clef2014
pic.twitter.com/c78uuF5DL8

Expand

Iadh Ounis @iadh — 18 Sep
CLEF 2016 will be in Evora

Tweet to @clef2015

SBS/iSBS continues @ CLEF in Toulouse

Nov 7: Join the discussion @ ESAIR/CIKM

ESAIR'14: Seventh International Workshop on Exploiting Semantic Annotations in Information Retrieval

Introduction

There is an increasing amount of structure on the Web as a result of modern Web languages, user tagging and annotation, emerging robust NLP tools, and an ever growing volume of linked data. These meaningful, semantic, annotations hold the promise to significantly enhance information access, by enhancing the depth of analysis of today's systems. Currently, we have only started exploring the possibilities and only begin to understand how these valuable semantic cues can be put to fruitful use. To complicate matters, standard text search excels at shallow information needs expressed by short keyword queries, and here semantic annotation contributes very little, if anything.

Articulate Queries and Query Auto Suggest

The goal of the ESAIR'14 is to advance the general research agenda on this core problem, with an explicit focus on two of the most challenging aspects to address in the coming years.

- First, there is a need to explore more articulate queries, with concepts and relations linking their statement of request to existing semantic models as offered by emerging knowledge bases (DBpedia, Freebase).
- Second, there is a need to extend the query suggestion paradigm to dynamically negotiate longer queries exploring powerful new aspects or facets of the underlying information need.

We Need Help!

The Workshop will bring together researchers working with semantic annotations, its use cases, its sources (authoring to NLP tools), its users, and its use in DB, IR, KM, or Web research, and work together on one of the greatest challenges in the years to come. We envision a lively and interactive workshop, with the explicit aim to push the boundaries and think outside the box.

ESAIR history

Previous ESAIR editions are found at:

Nov 19-21: TREC Contextual Suggestion

TREC Contextual Suggestion

TREC 2014
TREC 2013
TREC 2012

TREC 2014

As part of the National Institute of Standards and Technology (NIST) Text REtrieval Conference (TREC), the Contextual Suggestion Track deals with complex information needs which are highly dependent on context and user interests.

Track Organizers

- Charles L A Clarke, Waterloo
- Adriel Dean-Hall, Waterloo
- Jaap Kamps, Amsterdam
- Paul Thomas, CSIRO

TREC 2014 Contextual Suggestion Track Guidelines

Submission Validation and Evaluation Scripts

Contexts, Profiles, and Example Suggestions

Become ‘Friend’ event of ACM SIGIR

Looking for a Postdoc!