

Integral: Link-based Integration of Library Digital Collections and Resources



NJLA Conference: NJLA CUS/ACRL-NJ
Technology Innovation Forum
April 27, 2010

Ann D. Hoang
Assistant University Librarian
New Jersey Institute of Technology
ahoang@njit.edu
973-596-5798

Members of the Team ...



Min Song

Assistant Professor
Information Systems Dept. @ NJIT
<http://is.njit.edu/people/song.php>

IntegraL Project Manager and Co-Principal Investigator

Research interests:

- Text mining
- Dynamic document clustering
- Semantic query expansion
- Bioinformatics
- Information extraction & retrieval

Ann D. Hoang

Assistant University Librarian
Van Houten Library @ NJIT
<http://web.njit.edu/~ahoang>

IntegraL Library Technical Support and System Evaluation

Responsible for:

- Staff, services related to technology including web services, ILS system and digital technology
- Acquisitions of print collections & cataloging

– **Improve quality of services across all library functions**

Members of the Team ...



Michael Bieber
Principal Investigator



Richard Sweeney
NJIT University
Librarian

- ☑ **Vahid Hamidullah,**
Undergraduate Student
- ☑ **Eric Koppel,**
Undergraduate Student
- ☑ **Pawel Bokota,**
Undergraduate Student
- ☑ **Said Bleik, Ph.D. Student**
- ☑ **NJIT Librarians**

The Question ...

Why do we need libraries, library resources or librarians when we have Google?



A typical question often asked by inexperienced library users ...
not by serious information seekers.

The Answer ...

Invisible Web aka the “Deep Web”

Robert Lackie spent most of the last decade educating libraries users and librarians about

“**Those Dark Hiding Places:
The Invisible Web ...**”

<http://www.robertlackie.com/>

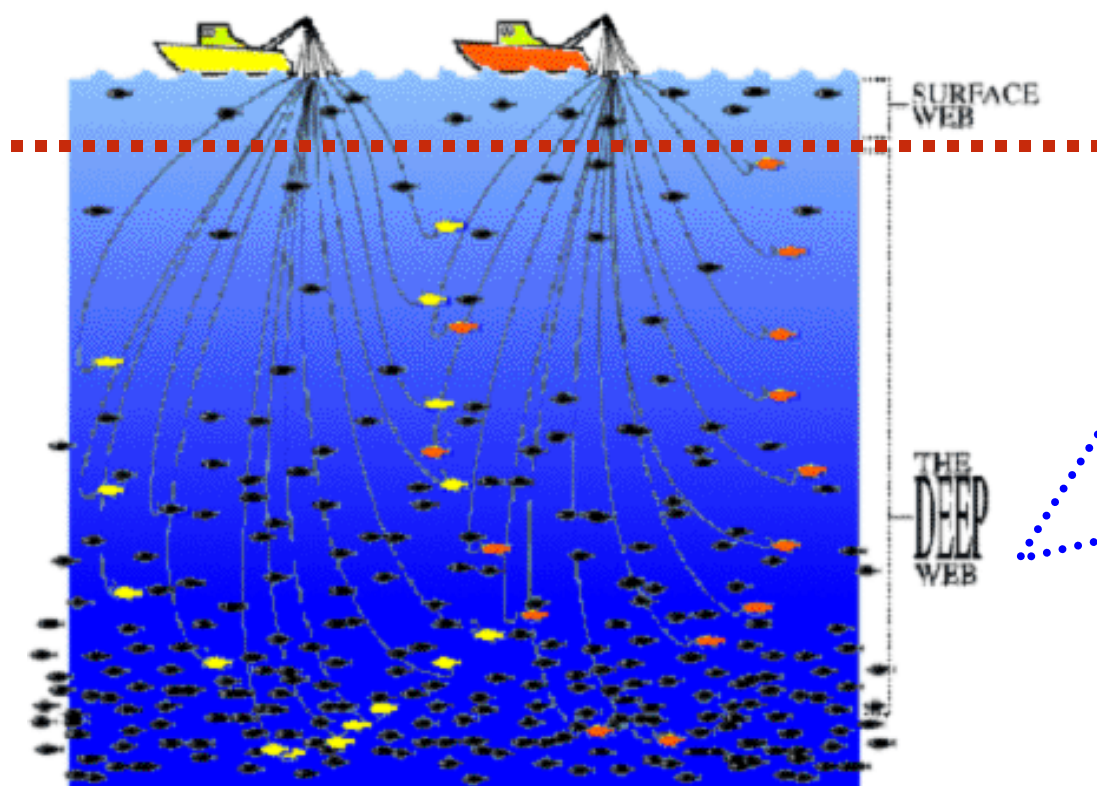
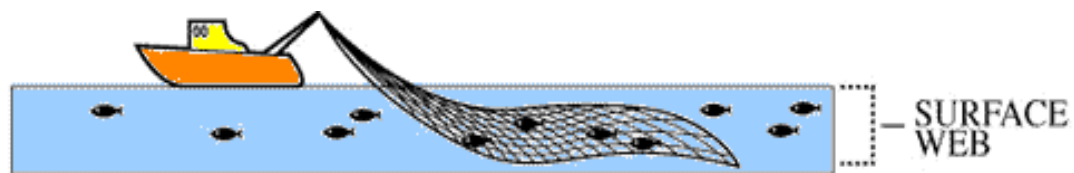
Michael Bergman, a web scientist, first coined the phrase

“**Deep Web**”

<http://www.mkbergman.com/>

Deep Web refers to web content that are created dynamically as a result of a specific search executed by a searcher

The Deep Web ...



- *“Serious information seekers can [not] avoid the importance or quality of deep Web information.”*
- *“ Deep Web information is only a component of total information available.”*
- *“Searching must evolve to encompass the complete Web.”*

The Deep Web ...



NEWJERSEY
Library Association



*Google does not
bring relevant
resources directly to
the users.*

The New York Times

This copy is for your personal, noncommercial use only. You can order presentation-ready copies for distribution to your colleagues, clients or customers [here](#) or use the "Reprints" tool that appears next to any article. Visit www.nytreprints.com for samples and additional information. Order a reprint of this article now.

February 23, 2009

Exploring a 'Deep Web' That Google Can't Grasp

By ALEX WRIGHT

One day last summer, [Google's](#) search engine trundled the list of Web pages it knows about. But as impossible as it seems, the entire Web.

Beyond those trillion pages lies an even vaster Web of hidden data, schedules, medical research and all kinds of other materials that search engines.

The challenges that the major search engines face in explaining why they still can't provide satisfactory answers to "What's in London next Thursday?" The answers are complex.

Now a new breed of technologies is taking over the corners. When that happens, it will do more than just change the way many companies do business online.

"...[added](#) the [one trillionth](#) address to the list of web pages... beyond those trillion pages lies...web of hidden data...all kinds...materials stored in databases that remain...invisible to search engines."

"...deep web technologies hold the promise of achieving [web data integration], by automating the process of analyzing database structures and cross-referencing the results."

The Web + The User...

- The Web is ...
 - incredibly successful at delivering information to users and searchers;
 - not yet a well organized repository of presentable structured HTML pages;
 - a conglomerate of volatile HTML pages.
- The User expects ...
 - an urgent need to go beyond browsing;
 - response dynamically to users' queries;
 - wants to get there faster.

Web of Organized Repository of Structured Content

- ... health site could cross-reference data from pharmaceutical companies with the latest findings from medical researchers, or
- ... a local news site could extend its coverage by letting users tap into public records stored in government databases.

The Challenge ...

You mean all that relevant information I need is hidden. I'm unable to access it and therefore preventing me from getting an "A" on my research paper? What will I tell my parents?



Challenges for searchers

- need to know what resources to use before they can access them,
- find related information from other resources,
- leave current resource to search related information from other resources.

What is Integral ?

*Fear not, information seekers.
Integral is here to save you
from a lifetime of drudgery.
The "A" that you are expecting
is in your future.*




Background

- research funded by National Science Foundation and Institute of Museum and Library Services;
 - develop a system to integrate library resources to make searching easier + faster;
 - bring relevant + related content directly to the users;
 - improve the effectiveness of both users and library resources

What is Integral? ...


*Step right up, folks!
Witness amazing feats
of bravura and derring-
do! Watch as search
results transition before
your very eyes!*













- ... is an open source software search engine
- ... is designed to allow users to efficiently search online resources
 - authenticated + non-authenticated
- ... is browsable across online resources through “linking” of element of interest  icons

What is Integral? ...

Element of Interest

-  Icons highlight Element of Interest (EOIs)
 - allow ease of integration between various resources
 - supports drill-down for more detailed information on Elements of Interest (EOIs)
 - empowers user's deep search via NER (name entity recognition) plug-ins

- [1 Semantic deep web: automatic attribute extraction from the deep web data sources](#)
 [Yoo Jung An](#), [James Geller](#), [Yi-Ta Wu](#), [Soon Ae Chun](#)
March 2007  **SAC '07**: Proceedings of the 2007 ACM symposium on Applied computing 
Publisher: ACM  [Request Permissions](#)
Full text available:  [Pdf](#) (218.34 KB) Additional Information: [full citation](#), [abstract](#), [re](#)

What is Integral? ...



[Subscribe](#) (Full Service) [Register](#) (Limited Service, Free) [Login](#)

Search: The ACM Digital Library The Guide
"deep web"

ing within **The Guide** for: "deep web" ([start a new search](#))
2 of 1,408,899

YOUR SEARCH

Search by Keywords

Search box containing "web"

SEARCH

Related Terms

Search by People

Search by Journals

Search by Publishers

Search by Subjects

Search by Conferences

Search by Publications

Publication Year

Publication Names

Publications

Publications

Publication Formats

Publications

Search by Conferences

Search by Publishers

Search Results

Related Journals

Results 1 - 20 of 442

[Save results to a Binder](#)

1 [Semantic deep web: automatic attribute](#)

[Yoo Jung An](#), [James Geller](#), [Yi-Ta](#)

March 2007 **SAC '07: Proceeding**

Publisher: ACM [Request Permissions](#)

Full text available: [Pdf](#) (218.34 KB)

2 [Experiences in crawling deep web in the](#)

[Dheerendranath Mundluru](#), [Xiongwu Xia](#)

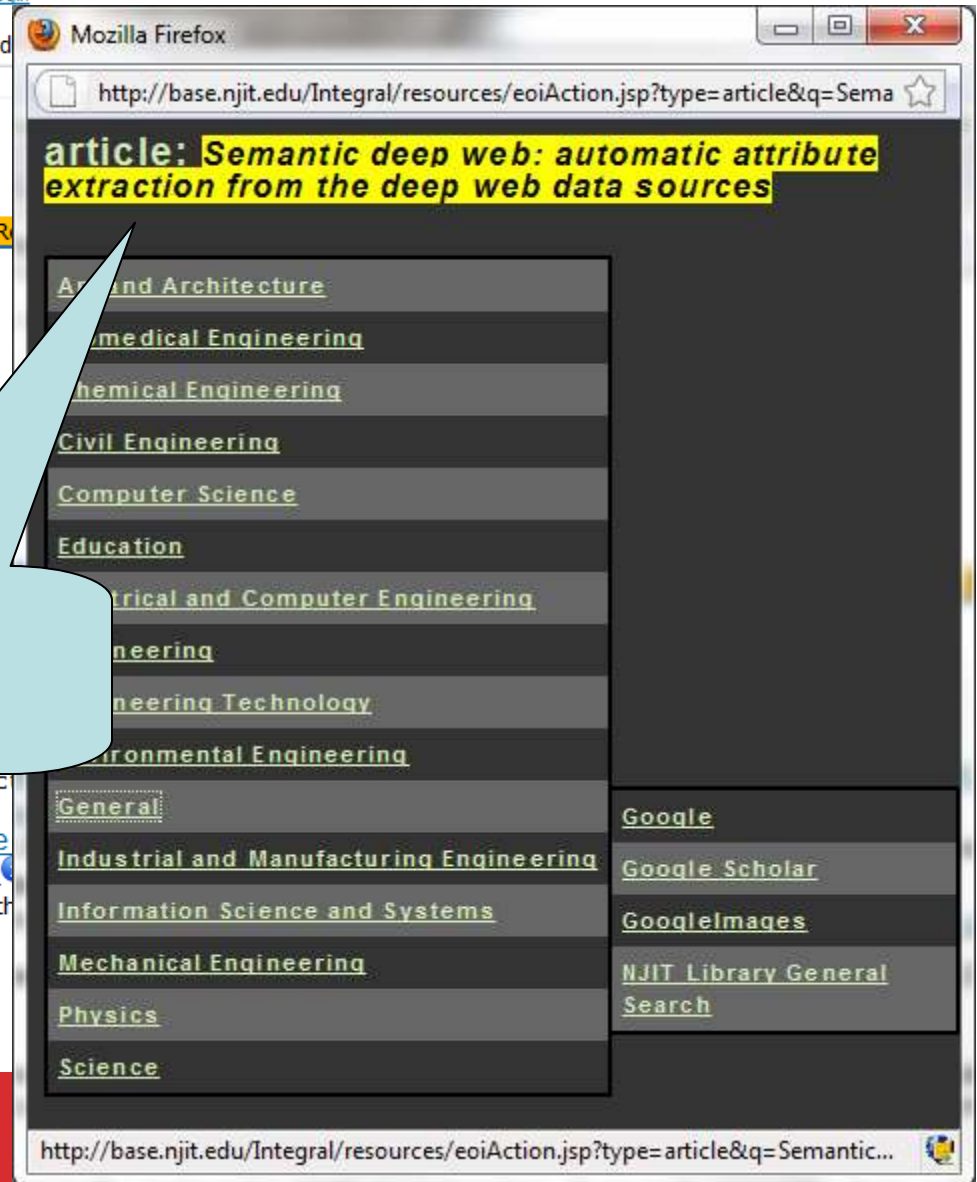
October 2008 **GIR '08: Proceeding of th**

Publisher: ACM [Request Permissions](#)

Full text available: [Pdf](#) (374.55 KB)

Keywords: automatic attribute extrac

Link-based integration



article: **Semantic deep web: automatic attribute extraction from the deep web data sources**

- Architecture
- Medical Engineering
- Chemical Engineering
- Civil Engineering
- Computer Science
- Education
- Electrical and Computer Engineering
- Engineering
- Engineering Technology
- Environmental Engineering
- General
- Industrial and Manufacturing Engineering
- Information Science and Systems
- Mechanical Engineering
- Physics
- Science

Google
Google Scholar
Google Images
NJIT Library General Search

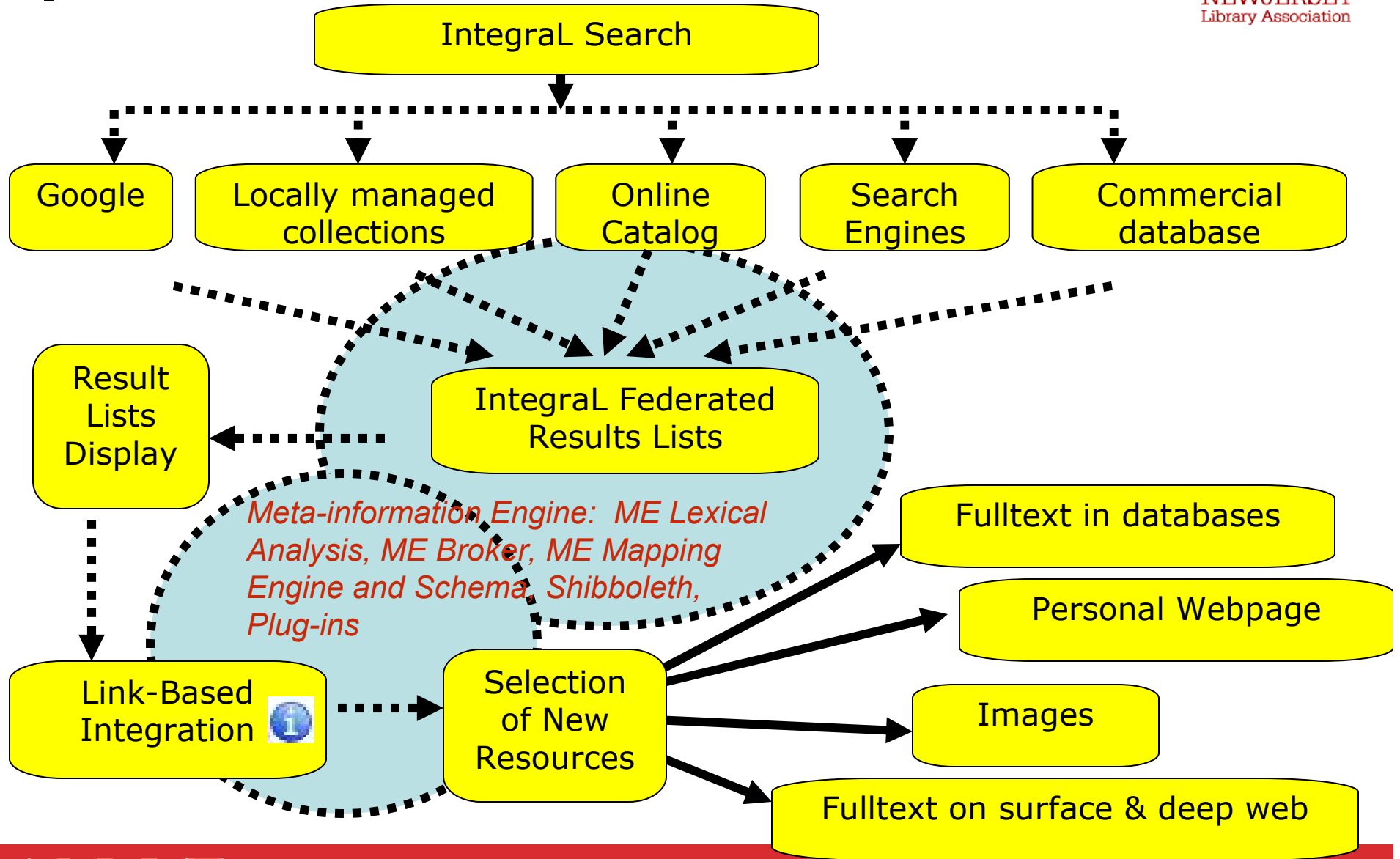
http://base.njit.edu/Integral/resources/eoiAction.jsp?type=article&q=Semantic...

What is IntegraL ? ...

IntegraL = **Integrating** + **Libraries**

- IntegraL search engine focuses on:
 - linking online resources selected by librarians
 - providing customized user preferences to the types of linked information.
- IntegraL system architecture:
 - Provide searchers with multiple access points
 - enables searchers to navigate unlinked web content
 - seamlessly links various resources to make browsing across resources easier
 - SSO with Shibboleth
 - your credentials is recognize across various services including email servers, database subscriptions, chat, etc. on your network.

System Architecture





NEWJERSEY
Library Association

*All right, folks. Stop
and relax. We are
entering the demo
zone.*



IntegraL Demo

Integral Demo

You only have to login once. Your system passes your credentials onto various services on the network.

Welcome to Integral

Please select your institution



[Login with your NJIT UCID](#)



[Login as a guest](#)



Shibboleth Identity Provider Login

Not a NJIT student or faculty member? [Register as a guest!](#)

Username:

Password:

Login

Integral Demo



IntegralL Federated Search

[Your Preferences](#)

[Digital Libraries](#)

[Logout](#)



[Logout](#)
[Register](#)

March 18, 2010

[Home](#)

[About Us](#)

[List Of Digital Libraries](#)

[Get Integral](#)

Project Overview

Integrating Libraries Research (IntLib Project and Integral System)

Mission Statement

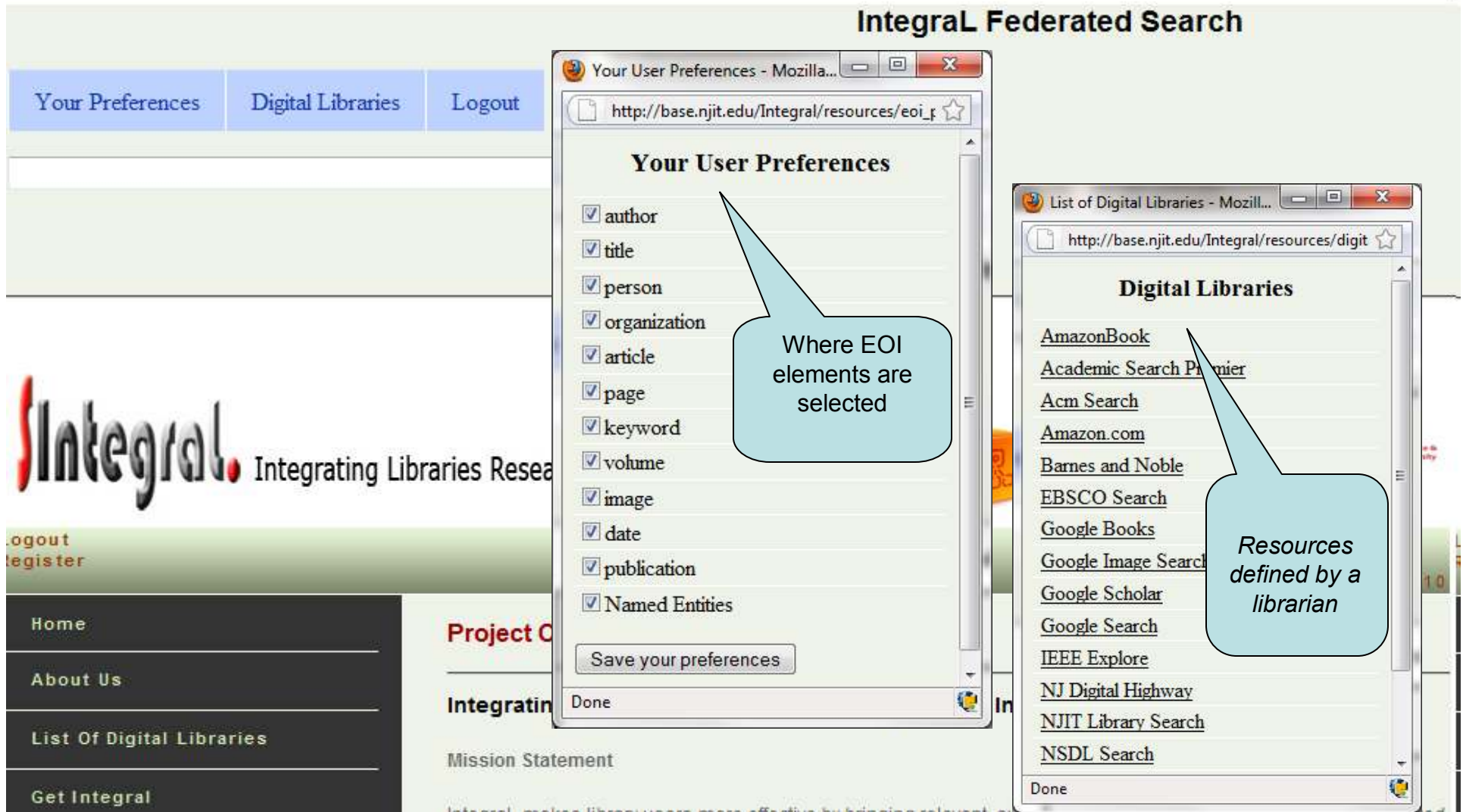
Integral makes library users more effective by bringing relevant, authenticated resources directly to them, customized



New Jersey's Science & Technology University

THE EDGE IN KNOWLEDGE

Integral Demo ...



Integral Federated Search

Your Preferences Digital Libraries Logout

Your User Preferences

- author
- title
- person
- organization
- article
- page
- keyword
- volume
- image
- date
- publication
- Named Entities

Save your preferences

Done

Digital Libraries

- [AmazonBook](#)
- [Academic Search Premier](#)
- [Acm Search](#)
- [Amazon.com](#)
- [Barnes and Noble](#)
- [EBSCO Search](#)
- [Google Books](#)
- [Google Image Search](#)
- [Google Scholar](#)
- [Google Search](#)
- [IEEE Explore](#)
- [NJ Digital Highway](#)
- [NJIT Library Search](#)
- [NSDL Search](#)

Done

Integral Integrating Libraries Research

Logout
Register

Home
About Us
List Of Digital Libraries
Get Integral

Project C
Integrating
Mission Statement

Integral makes library users more effective by bringing relevant


Where EOI elements are selected

Resources defined by a librarian

Integral Demo ...

Your Preferences Digital Libraries Administration Panel

“michael bergman” “deep web” Search



Federated Search Results

Google	<u>4200 results</u>
Ask Search	<u>132 results</u>
NJIT Library Author Search	<u>0 results</u>
NSDL Search	<u>0 results</u>
ACM Search	<u>0 results</u>
Google Book Search	<u>0 results</u>
GoogleImages	<u>93 results</u>
Scopus Affiliation Search	<u>0 results</u>
Google Scholar	<u>40 results</u>
IEEE Book Search	<u>0 results</u>
ScienceDirect	<u>1 results</u>

Integral Demo ...

Google scholar "michael bergman" and "deep web" Search [Advanced Scholar Search](#) [Scholar Preferences](#)

Scholar Articles and patents anytime include

[\[HTML\] Search engine technology and digital libraries](#)

N Lossau - D-Lib Magazine, 2004 - [bieson.uib.uni-bielefeld.de](#)

... Although there are no reliable figures on the overall volume of web c studies that give estimations. Already published in 2001, a white paper on t ...

Google Scholar and Science Direct Results

[Cited by 8](#) - [Related articles](#) - [All 21 versions](#)

... Search engine robots that are text-only and support HTML 2.0 cann sites or sites that hide content in images. In his study on the " deep W found that traditional surface Web sites are fraught with many quality p [Cited by 8](#) - [Related articles](#) - [All 21 versions](#)

[\[PDF\] Untapped Assets: The \\$3 Trillion Value of US Enterpris](#)

MK Bergman - BrightPlanet Corporation White Paper, July, 2005 - brig Page 1. July 2005 by Michael K. Bergman Untapped Assets: The \$3 T of US Enterprise Documents Page 2 - BrightPlanet Corporation, Sioux Washington, DC, and New York City, is a private venture-backed ...

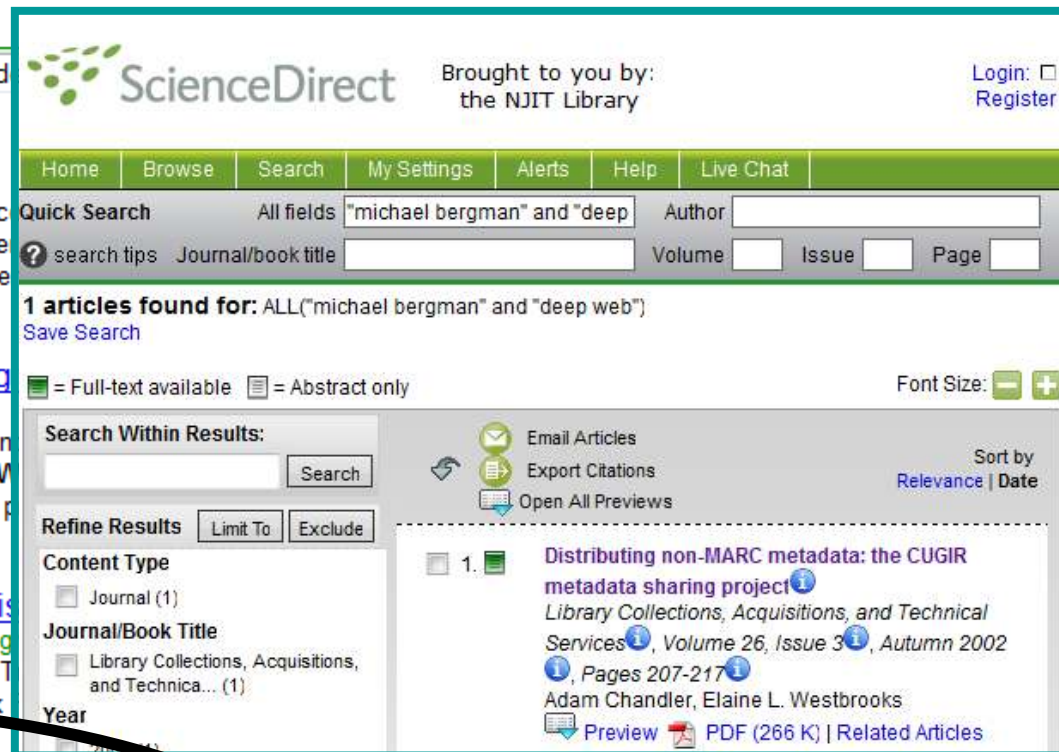
[Cited by 2](#) - [Related articles](#) - [View as HTML](#) - [All 6 versions](#)

[Distributing non-MARC metadata: the CUGIR metadata sharing project](#)

A Chandler, EL Westbrooks - Library Collections, Acquisitions and ..., 2002 - Elsevier

... The first problem was that individual metadata records are not accessible in CUGIR because they reside in the "Deep Web." In the white paper "The Deep Web: Surfacing Hidden Value," Michael Bergman of BrightPlanet Inc. defines Deep-Web as being: ...

[Cited by 5](#) - [Related articles](#) - [All 5 versions](#)



ScienceDirect Brought to you by: the NJIT Library Login: Register

Home Browse Search My Settings Alerts Help Live Chat

Quick Search All fields "michael bergman" and "deep" Author

Journal/book title Volume Issue Page

1 articles found for: ALL("michael bergman" and "deep web") Save Search

= Full-text available = Abstract only Font Size:

Search Within Results: Search

Refine Results Limit To Exclude

Content Type Journal (1)

Journal/Book Title Library Collections, Acquisitions, and Technica... (1)

Year

Email Articles Export Citations Open All Previews Sort by Relevance | Date

1. [Distributing non-MARC metadata: the CUGIR metadata sharing project](#)
Library Collections, Acquisitions, and Technical Services, Volume 26, Issue 3, Autumn 2002, Pages 207-217
Adam Chandler, Elaine L. Westbrooks
[Preview](#) [PDF \(266 K\)](#) | [Related Articles](#)

[unl.edu](#) [PDF]

Integral Demo ...



Brought to you by:
the NJIT Library

Home Browse Search My Settings Alerts Help Live Chat

Quick Search All fields "michael bergman" and "deep" Author

search tips Journal/book title Volume

Library Collections, Acquisitions, and Technical Services
Volume 26, Issue 3, Autumn 2002, Pages 207-217

Article Figures/Tables References PDF (266 K)

doi:10.1016/S1464-9055(02)00247-6

[Cite or Link Using DOI](#)

Copyright © 2002 Elsevier Science Inc. All rights reserved.

Special section on metadata

Distributing non-MARC metadata: the CUGIR metadata

Adam Chandler^a and Elaine L. Westbrooks^b

^a Cornell University, Central Technical Services, 110c Olin Library, Ithaca, NY 14853

^b Cornell University, Albert R. Mann Library, Technical Services, Ithaca, NY 14853-5000

Available online 1 October 2002.

address two access-related problems. The first problem was that individual metadata records are not accessible in CUGIR because they reside in the "Deep Web." In the white paper "The Deep Web: Surfacing Hidden Value," Michael Bergman of BrightPlanet Inc. defines Deep-Web as being:

... content that resides in searchable databases, the results from which can only be discovered by a direct query. Without the directed query, the database does not publish the result. When queried, deep Web sites post their results as dynamic Web pages in real-time. Though these dynamic pages have a unique URL address that allows them to be retrieved again later, they are not persistent [27].

Integral Demo

Google and Google Images Results

Google "michael bergman" and "deep web"

Web Search results for "michael bergman" and "deep web".

Results 1 - 10 of about 4,200 for "michael bergman" and "deep web".

AI3::Adaptive Information » Mike Bergman on the semantic Web and ...
... Bioinformatics · Blogs and Blogging · Book Reviews · Brown Bag Lunch · Deep Web · Description Logics · Document Assets · Information Automation · irON ...
[www.mkbergman.com/](#) - [Cached](#) - [Similar](#)


About Mike » AI3::Adaptive Information
Nov 24, 2008 ... He was the author of BrightPlanet's award-winning Internet search tutorial and deep Web and document assets white papers. ...
[www.mkbergman.com/about-mike/](#) - [Cached](#) - [Similar](#)

Google Images results for "michael bergman" and "deep web".

Search results for "michael bergman" and "deep web".

Custom Web Crawler: [www.mozzila.com/And-Deep+Web/](#) Looking for And Deep Web? Scrape anything fast. See how!

Michael Bergman at Amazon: [www.Amazon.com/Books/](#) Low Free 2-Day Shipping w/ Amazon



with Michael Bergman
400 x 370 - 18k - jpg
[tedata.com/searchblog.com](#)

Michael Bergman's Pictures
480 x 640 - 37k - jpg
[123people.com](#)

Cited by Mike Bergman



AI3

Adaptive Information
Adaptive Innovation
Adaptive Infrastructure

About Mike

Michael K. Bergman is the CEO and co-founder of Structured Dynamics LLC since November 2008, a provider of services and products for the semantic enterprise, and editor of the lightweight UMBEL subject concept ontology. He began this blog in July 2005.

For the past twenty years Mike has been an entrepreneur, Web scientist, and independent consultant. Mike was CEO of Ztgist LLC, the predecessor company to Structured Dynamics. For six years prior to that up to late-2006, Mike was the co-founder, chief technology officer, and chairman of BrightPlanet Corporation. Prior to that he was founder and CEO of VisualMetrics Corporation, a data warehousing software firm with an emphasis on large-scale Internet databases and bioinformatics, and TheWebTools Company, an early pioneer in meta-search tools.

Mike began his professional career in the mid-1970s as a project director for the U.S. EPA for a major energy study called the Coal Technology Assessment (CTA). He later taught in the Graduate School of Engineering at the University of Virginia and was a fellow in the Energy Policies Study Center. He then joined the American Public Power Association in 1982 and was soon promoted to director of energy research. He quadrupled that trade association's voluntarily supported R&D program to 500 members.


Main Links

- Home
- Structured Dynamics
- UMBEL
- SweetSearch
- Sweet Tools (Sem Web)
- SWEETpedia
- A Timeline of Information History
- Pro Blogging Guide
- Chronological Listing
- Acronyms & Glossary
- This Blogged
- About Mike

Integral Demo ...

Your Preferences Digital Libraries Administration Panel Logout

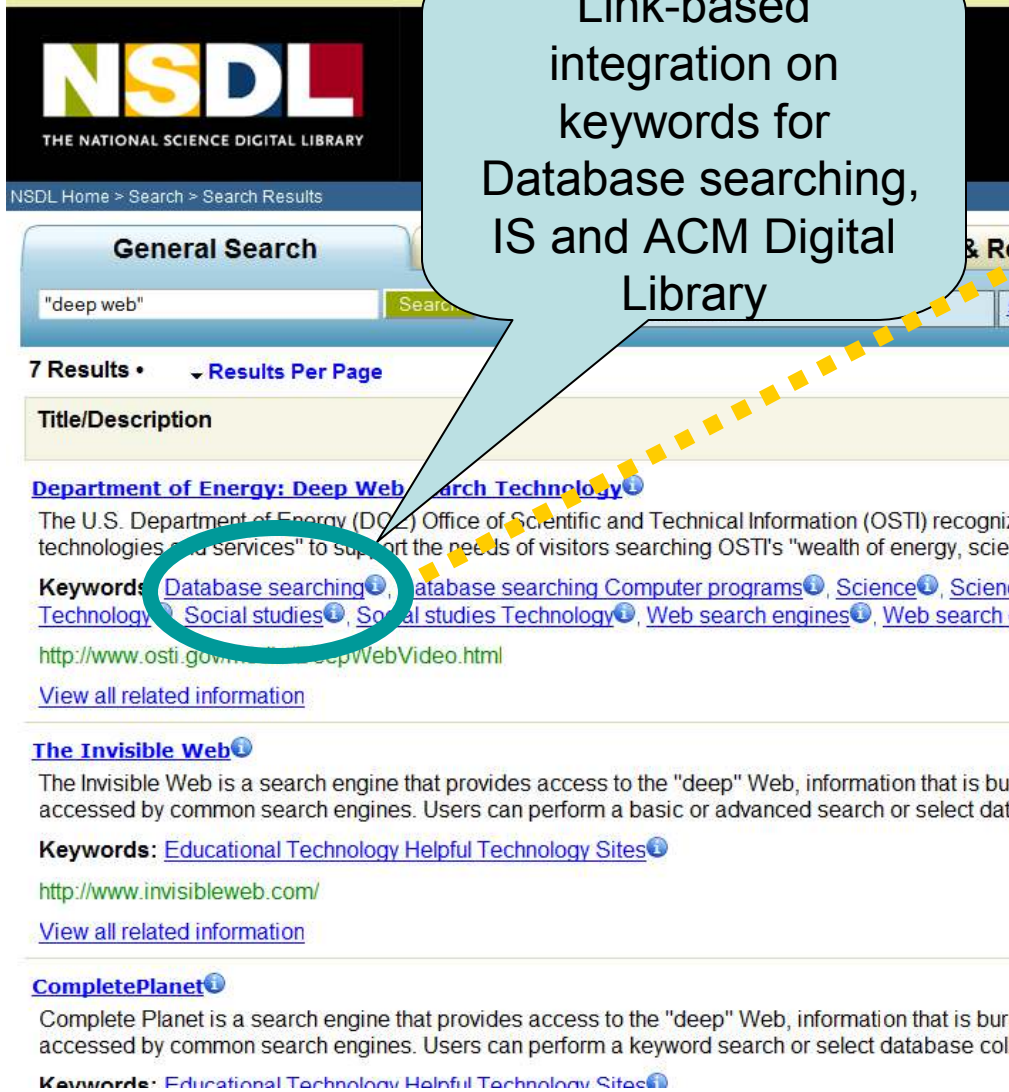
“deep web” Search



Federated Search Results

Google	<u>245000 results</u>
Ask Search	<u>41900 results</u>
NJIT Library Author Search	<u>5599 results</u>
NSDL Search	<u>7 results</u>
ACM Search	<u>442 results</u>
Google Book Search	<u>0 results</u>
GoogleImages	<u>15100 results</u>
Scopus Affiliation Search	<u>0 results</u>
Google Scholar	<u>4280 results</u>
IEEE Book Search	<u>0 results</u>
ScienceDirect	<u>163 results</u>

Integral Demo ...



NSDL
THE NATIONAL SCIENCE DIGITAL LIBRARY

NSDL Home > Search > Search Results

General Search

"deep web" Search

7 Results • Results Per Page

Title/Description

Department of Energy: Deep Web Search Technology
The U.S. Department of Energy (DOE) Office of Scientific and Technical Information (OSTI) recognizes technologies and services "to support the needs of visitors searching OSTI's "wealth of energy, science and technology information" for the "deep web".

Keywords: Database searching, Database searching Computer programs, Science, Science Technology, Social studies, Social studies Technology, Web search engines, Web search engines

<http://www.osti.gov/patents/2006022822/deepWebVideo.html>

[View all related information](#)

The Invisible Web
The Invisible Web is a search engine that provides access to the "deep" Web, information that is buried and not accessed by common search engines. Users can perform a basic or advanced search or select data from a list of categories.

Keywords: Educational Technology Helpful Technology Sites

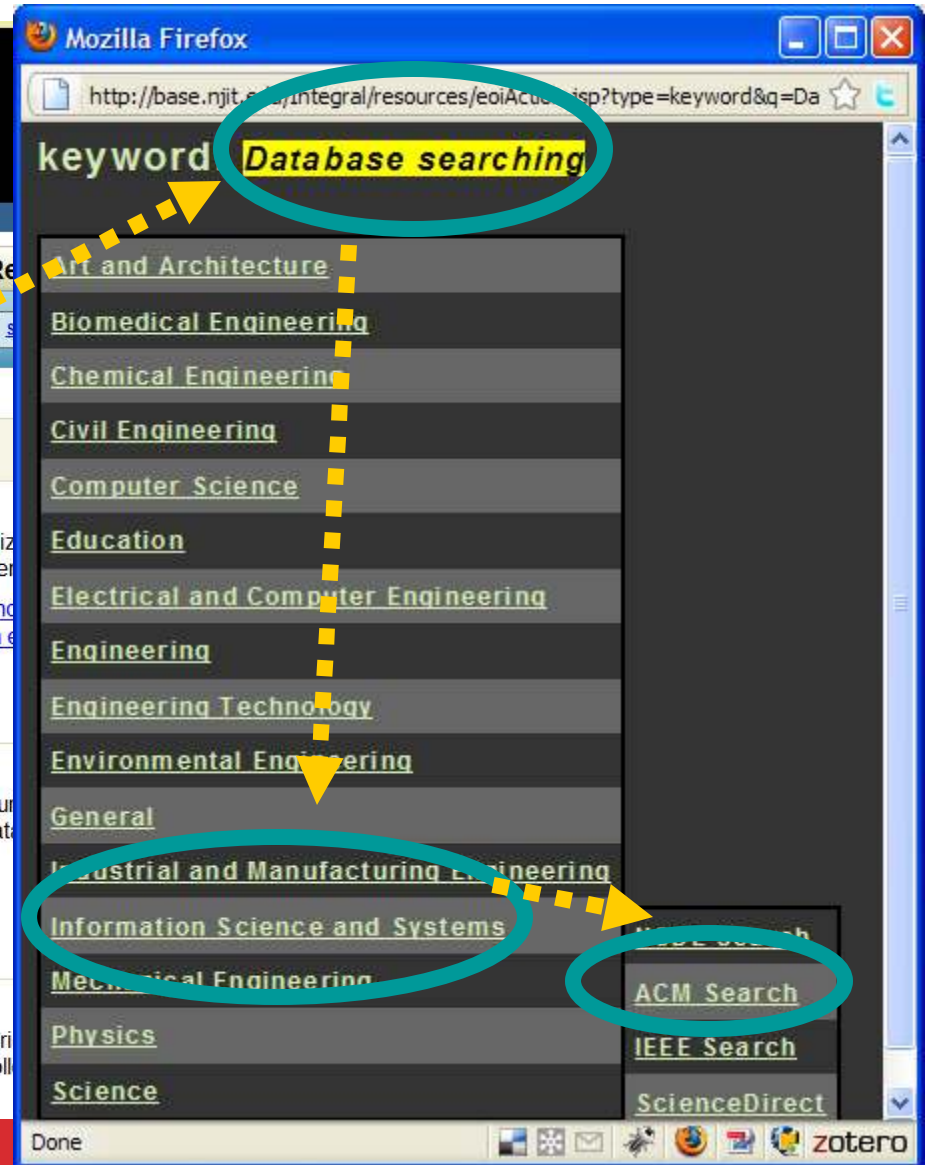
<http://www.invisibleweb.com/>

[View all related information](#)

CompletePlanet
Complete Planet is a search engine that provides access to the "deep" Web, information that is buried and not accessed by common search engines. Users can perform a keyword search or select database collections from a list of categories.

Keywords: Educational Technology Helpful Technology Sites

Link-based integration on keywords for Database searching, IS and ACM Digital Library



Mozilla Firefox

http://base.njit.edu/integral/resources/eoiACM...isp?type=keyword&q=Da

keyword **Database searching**

- Art and Architecture
- Biomedical Engineering
- Chemical Engineering
- Civil Engineering
- Computer Science
- Education
- Electrical and Computer Engineering
- Engineering
- Engineering Technology
- Environmental Engineering
- General
- Industrial and Manufacturing Engineering
- Information Science and Systems
- Mechanical Engineering
- Physics
- Science

ACM Search

IEEE Search

ScienceDirect

Integral Demo ...



[Subscribe](#) (Full Service) [Register](#) (Limited Service, Free) [Log Out](#)

Search: The ACM Digital Library The Guide

Database searching

Searching within **The Guide** for: Database searching ([start a new search](#))

Found 21,660 of 1,408,899

REFINE YOUR SEARCH

▼ Refine by Keywords

Database searching

SEARCH

[Discovered Terms](#)

▼ Refine by People

[Names](#)
[Institutions](#)
[Authors](#)
[Editors](#)
[Advisors](#)
[Reviewers](#)

▼ Refine by Publications

[Publication Year](#)
[Publication Names](#)
[ACM Publications](#)
[All Publications](#)
[Content Formats](#)
[Publishers](#)

▼ Refine by Conferences

[Sponsors](#)

Search Results

Related Journals

Related Conferences

Related Conferences

Results 1 - 20 of 21,660

 [Save results to a Binder](#)

Sort by

Result page

1 [Democratizing information: online databases and the rise of end-user searching](#)ⁱ

[Bryan Pfaffenberger](#)ⁱ

January 1990 ⁱ Democratizing information: online databases and the rise of end-user searching ⁱ

Publisher: G. K. Hall & Co.

Additional Information: [full citation](#), [cited by](#), [index terms](#), [review](#)

Bibliometrics: Downloads (6 Weeks): n/a, Downloads (12 Months): n/a, Downloads (Overall): n/a, Citation Count: 3

2 [MPI framework for parallel searching in large biological databases](#)ⁱ

[Dominic Battré](#)ⁱ, [David Sigfredo Angulo](#)ⁱ

December 2006 ⁱ **Journal of Parallel and Distributed Computing**, Volume 66 Issue 12 ⁱ

Publisher: Academic Press, Inc.

Additional Information: [full citation](#), [abstract](#), [references](#)

Bibliometrics: Downloads (6 Weeks): n/a, Downloads (12 Months): n/a, Downloads (Overall): n/a, Citation Count: 0


In this paper, we address the problem of searching huge biological databases on the scale of at least several terabytes. Biological databases storing DNA sequences, protein sequences, or mass spectra are growing exponentially.

Link-based integration on keyword for Database searching, IS and ACM Digital Library

Integral Demo ...

Your Preferences Digital Libraries Administration Panel Logout

“deep web” Search



Federated Search Results

Google	<u>245000 results</u>
Ask Search	<u>41900 results</u>
NJIT Library Author Search	<u>5599 results</u>
NSDL Search	<u>7 results</u>
ACM Search	<u>442 results</u>
Google Book Search	<u>0 results</u>
GoogleImages	<u>15100 results</u>
Scopus Affiliation Search	<u>0 results</u>
Google Scholar	<u>4280 results</u>
IEEE Book Search	<u>0 results</u>
ScienceDirect	<u>163 results</u>

Integral Demo ...



PORTAL
WALDO Consortium: NJIT

Subscribe (Full Service) Register (Limited Service, Free)

Search: The ACM Digital Library The Guide
"deep web"

ing within **The Guide** for: "deep web" ([start a new search](#))
2 of 1,408,899

YOUR SEARCH

Search Results **Related Journals**

Results 1 - 20 of 442

[Save results to a Binder](#)

1 [Semantic deep web: automatic attribute](#)
Yoo Jung An, James Geller, Yi-Ta Wu
March 2007 **SAC '07: Proceedings of**
Publisher: ACM [Request Permissions](#)
Full text available: [Pdf](#) (218.34 KB)

author:
James Geller,
General, Google



Mozilla Firefox

http://base.njit.edu/integral/resources/eoiAction.jsp?type=author&q=James+Geller

author **James Geller**

- Art and Architecture
- Biomedical Engineering
- Chemical Engineering
- Civil Engineering
- Computer Science
- Education
- Electrical and Computer Engineering
- Engineering
- Engineering Technology
- Environmental Engineering
- General**
- Industrial and Manufacturing Engineering
- Information Science and Systems
- Mechanical Engineering
- Physics
- Science

- Google
- NJIT Library Online Catalog
- Google Scholar
- Google Images
- NJIT Library General Search

http://base.njit.edu/Integral/resources/eoiAction...

Integral Demo ...



The image shows a Google search for "james geller". The search results include a link to the "Home page of James Geller" at NJIT. A callout box highlights this link with the text: "James Geller is affiliated with NJIT and his research interest is in semantic web". Below the search results, there is a detailed profile page for James Geller, including a photo, his title as Professor and Lab Director, and his research interests in Semantic Web & Ontology, Medical Informatics, and SABOC. The profile also lists his department, office hours, mailing address, and office address. Three smaller images of James Geller are shown at the bottom, each with a caption: "Home page of James", "Home page of James", and "Home page of James".

Home page of James Geller

Google

Web Images Videos Maps News Shopping Gmail more ▾ Web History

Google

Web Images Videos Maps News Shopping Gmail more ▾ Web History

Home page of James Geller

Enter: Geller James in the "for" field, and click on the "Search" button. Ignore the other Geller J who publishes with web.njit.edu/~geller/ - Cached - Similar

DBLP: James Geller

Yan Chen, Huanying Helen Gu, Yehoshua F. Group-based auditing of missing hierarchic www.informatik.uni-trier.de/~ley/db/.../a.../Ge Cached - Similar

james geller - Directory | LinkedIn

View the profiles of professionals named ja professionals named james geller who use www.linkedin.com/pub/dir/james/geller - C

Scientific Commons: James Geller

4 Feature Article: Yugyung Lee, Soon Ae Ch James Geller. The challenge formulated in en.scientificcommons.org/james_geller - C

- Professor
- Lab Director, [Semantic Web & Ontology](#)
- Co-Director, [Medical Informatics](#)
- Co-Director, [SABOC](#)
- Former Associate Director, Bioinformatics

Department of Computer Science

Office Hours: [at my public Calendar](#)

Mailing Address:
Department of Computer Science
New Jersey Institute of Technology
323 Dr. King Blvd.
Newark, NJ 07102

Office Address:
4307 Information Technology Center (GI
telephone: +1 973 596-3383
messages: +1 973 596-3366
fax: +1 973 596-5777
Email: geller@njit.edu

Home page of James
432 x 669 - 70k - jpg
web.njit.edu

Home page of James
777 x 583 - 303k - jpg
web.njit.edu

Home page of James
427 x 640 - 57k - jpg
web.njit.edu

Benefits of Using Integral

Logical? What the heck are these bozos talking about? Clowns don't do logical.



Browsing cross resources

- multiple access points through linking of EOIs,
- ability to navigate unlinked web content e.g. from propriety databases to author's personal webpage,
- single sign-on authentication using Shibboleth, to seamlessly links to databases an resources that requires a password login.
- users awareness of additional resources

Benefits of Using Integral ...

- Direct access to system's features
 - access to admin panel via the web
 - Ability add/modify/delete resources with minimal technical expertise.
- Information Literacy
 - Integral automates a search strategy that experienced searchers might go through by manually and inexperienced searchers might not even be aware how to do.
- Bridging silos between surface web and deep web
 - This bridges the gap between the silos of the structured web and the deep web
 - Integral leverages metadata and structure from one silo into the another to create a better search toward getting better results

Links & Further Reading



- Wikipedia Definition of “Deep Web” http://en.wikipedia.org/wiki/Deep_web
- Bergman, M. (2001). White Paper: The Deep Web: Surfacing Hidden Value. *Journal of Electronic Publishing*, V 7(1). <http://dx.doi.org/10.3998/3336451.0007.104>
- Wright, A. (2009). Exploring a “Deep Web” that Google Can’t Grasp. *New York Times*. Feb 23, 2009. <http://www.nytimes.com/2009/02/23/technology/internet/23search.html>
- Bradley, B. (2007). Debate: Do we need librarians now we have Google? <http://www.slideshare.net/Philbradley/debate-do-we-need-librarians-now-we-have-google>
- Song, M. and Bieber, M. (2008). Integral: Lightweight Link-based Integration of Heterogeneous Digital Library Collections and Services in the Deep Web. E-Commerce Technology and the Fifth IEEE Conference on Enterprise Computing, E-Commerce and E-Services, 2008 10th IEEE Conference on Jul 21-24, 2008. <http://doi.ieeecomputersociety.org/10.1109/CECandEEE.2008.111>
- Song, M., Ho, S.M., Bieber, M., Koppel, E., Hamidullah, V., and Bokota, P. (2009). A Scalable Digital Library Infrastructure Expands Search and Beyond (2009) BookOnline’09 Workshop at ECDL 2009, October 2, Corfu, Greece



NEWJERSEY
Library Association

*Are we done
yet?*

Thanks.
Questions?



NJIT
New Jersey's Science & Technology University

THE EDGE IN KNOWLEDGE