Case Study on Using Federated Search to Enable Science and Engineering Research

Presented by: Naveen Maddali
ACS 238th National Meeting
August 17, 2009
Scitopia Basics

- A free federated “vertical” search engine
- A collaboration of 20+ leading science and technology societies
- A gateway to 4 million articles and 350 years of historical content
- Open to the general public, but designed for researchers
- Focused on the fields of technology, including engineering, physics, mathematics and computer science
- Searches on patents and Government documents
- Search provided by Deep Web Technologies

Features

- Robust advanced search capability
- Clustered results
- Search alerts
- Citation downloading
- Multiple sort features
- Source tool
Important Milestones

FEB 2007  Scitopia.org concept and business model developed by 8 founding partners
APRIL 2007  Prototype developed
JUNE 2007  Beta version of Scitopia launched at SLA conference
OCT 2007  Scitopia comes out of Beta
JUNE 2008  Clustering functionality released
OCT 2008  Alerting functionality released
MARCH 2009  Software upgrade
Who are the partners?

- Acoustical Society of America
- American Geophysical Union (AGU)
- American Institute of Aeronautics and Astronautics (AIAA)
- American Institute of Physics (AIP)
- American Physical Society (APS)
- American Society of Civil Engineers (ASCE)
- American Society of Mechanical Engineers (ASME)
- Association for Facilities Engineering (AFE)
- Audio Engineering Society (AES)
- AVS
- ECS
- IEEE
- Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering (ICST)
- Institute of Physics Publishing (IOP)
- Institution of Mechanical Engineers (IMechE);
- International Union of Crystallography (IUCr)
- National Association for Corrosion Engineering (NACE)
- Optical Society of America (OSA)
- Royal Society (RS)
- Royal Society of Chemistry (RSC)
- Society for Industrial and Applied Mathematics (SIAM)
- Society for Information Display (SID)
- SPIE
Features

Alerting Service

- One source to get email and RSS alerts for any or all of the partners
- Frequency of alerts is customizable
- Fields that alerts can be based off of include title, author, affiliation, abstract and full record
- Can be filtered by source

Search Results

- Search results grouped into 6 categories
  - Topics
  - Authors
  - Publications
  - Publishers
  - Affiliation
  - Date

Clustered Search Results allow for easier navigation and retrieval of relevant documents
Goals and Strategies

- Scitopia’s overall goal is to drive traffic to each of the partner’s websites.
- Scitopia’s secondary goal is to raise awareness of the quality and value of society publishers’ content.
- Scitopia is first step towards deeper society collaboration.
- Scitopia’s focus is on increasing brand awareness and site traffic.
Federated Search

Federated search is the process of performing a simultaneous real-time search of multiple diverse and distributed sources from a single search page, with the federated search engine acting as intermediary.

- Typically include result aggregation, ranking and de-duplication
- Able to find content on the “deep web”
- Results are typically more current, relevant and of higher quality

Scitopia chose the federated search model because:

- Most current research is available
- More relevant content is available
- Quicker to launch federated search product
How it works

User conducts a search from scitopia.org home.

Partner's digital library access to full-text document determined by individual partner.

User accesses full-text based on subscription rights with that partner. Abstract records will be visible from within scitopia.org.

Search results are shown in tabbed format, clearly indicating results from partners, patent offices and DOE.
Why should societies participate with Scitopia?

- Increasing society site usage
- Positive public perception of collaboration of society publishers
- Filling a demand
- Discovery of relevant content from previously unfamiliar sources and add these new sources to their consideration set for research.
- Potentially acquire new institutional or member subscriptions to the publisher’s digital library.
- Front end interface for platform
Why should researchers use Scitopia?

- Scitopia is focused on science and technology topics only.
- Search results are only coming from the highest quality sources (leading Sci-tech societies), as most results are peer-reviewed.
- The focused set of sources prevents internet noise to show in the search results.
- Federated searching provides complete and up to date results.
- Advanced search fields such as affiliation and summary are available on Scitopia.
- Searches on patents and Government documents are available from scitopia.
- Site navigation and sorting through search results is easy, intuitive and user-friendly.
- Users can discover society publications that were previously unknown to them.
- Clustered search results provides quicker and easier navigation to the desired articles.
- Alerting on scitopia allows users to use one location to get alerts from multiple societies.
- Archive goes back as far as 1665.
Valuable Research

scitopia.org content is cited 6x as often as the closest competitor

Source: 1790 Analytics
Competitive Landscape

Competitors to Scitopia include:

1. Popular search engines
2. Commercial Publisher search products
3. Society Publisher search products
4. Abstracting and Indexing services

- Those who use search engines for research are primarily looking to get relevant research quickly and are not interested in gimmicks and “bells and whistles”
- Differentiation is primarily done through the amount of content, the effectiveness of search, and the features and services of the site.
- Features such as clustering, alerting and OpenURL do not give scitopia a competitive advantage, but are important to compete.
- Scitopia’s competitive advantage comes from the immediacy and relevance of its search.
Sources of Traffic

Direct Traffic
- Typing URL into browser or using favorites
- Most ideal source of traffic

Paid Search Sites
- “Accidental visits”, often low quality
- High # of visits will lead to good number of “good visits”

Referrals from Partner Sites
- Scitopia hyperlinks from web pages, search boxes, press releases, articles, etc
- High quality visits

Contextual Advertising
- users who affiliated with or interested in science and technology disciplines
- “moderate quality” users

Library Pages
- links, search boxes, press releases, etc on library pages.
- Quality traffic, but reliant on librarian’s efforts

Others
- viral marketing, blogs, wikis, etc
- Majority of these sources cannot be managed or controlled
Referral Traffic

- Most important metric
- Steadily increasing overall
- Referrals to partner sites dependent on several factors:
  - Overall scitopia.org traffic
  - Research field
  - Amount of content
  - Promotion within own society
### User Profile

- Use Google Analytics service to measure site traffic
- More academic and Government users than corporate
- Need additional user information such as age, area of study, user behavior, etc.

<table>
<thead>
<tr>
<th>Country/Territory</th>
<th>Visits</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>47.9%</td>
</tr>
<tr>
<td>India</td>
<td>8.9%</td>
</tr>
<tr>
<td>China</td>
<td>4.3%</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>3.8%</td>
</tr>
<tr>
<td>Canada</td>
<td>3.6%</td>
</tr>
<tr>
<td>France</td>
<td>2.5%</td>
</tr>
<tr>
<td>Germany</td>
<td>2.2%</td>
</tr>
<tr>
<td>Australia</td>
<td>2.0%</td>
</tr>
<tr>
<td>Spain</td>
<td>1.9%</td>
</tr>
<tr>
<td>Malaysia</td>
<td>1.6%</td>
</tr>
</tbody>
</table>
Promotions Strategy

- Paid Search Placement
- Contextual Advertising
- Search Boxes, toolbars, widgets
- Press Releases
- Conference and Tradeshow Promotion
- Promotional Events
- Sponsorships
- Marketing collateral (flyers, business cards, flashing pins, t-shirts)
Business Strategies

- Acquisition of new partners
  - Number of societies is a competitive advantage
  - Additional revenue for promotion, etc.
  - Partner acquisition is a focus in 2009
  - Continue to include only science and technology societies
  - More outreach, better “sales tools”

- Advertising on scitopia
- Practical and innovative new features
- OpenURL
- New “Metrics of Success” to assess performance
- Alternative business models
Self-Evaluation

- Scitopia has been successful in creating a practical and functional site
- Scitopia has been moderately successful in growing its usage
- Scitopia’s strength is in its content quality.
- Scitopia can improve in terms of amount of content, brand name and features
- Federated search of sources has worked well
- Current business model of operating as a self-sustaining business is working and can continue to work
- Competition is strong and is expanding
Lessons Learned

- Working with numerous stakeholders requires extensive and careful communication, as well as compromise
- Benefits of partnership can also be intangible
- Innovation and differentiation can overcome limited resources. However, flash will not overcome substance
- Appropriate metrics of success are very important to evaluate the current state of a product
- Anticipated events may not always happen and products need to be re-evaluated periodically to adjust for it