



ELSEVIER

# Science-Specific Search: Bridging the Gap between Dissemination & Access to Information

Presented by:

Joris van Rossum, Head of Scirus

Event:

IATUL 2007

- How has content provision changed?
- How has information retrieval changed?
  - Lessons from Elsevier products
  - How can librarians improve information retrieval?
- Future Trends



# How did scientific content change with Web?



ELSEVIER

arXiv.org e-Prints



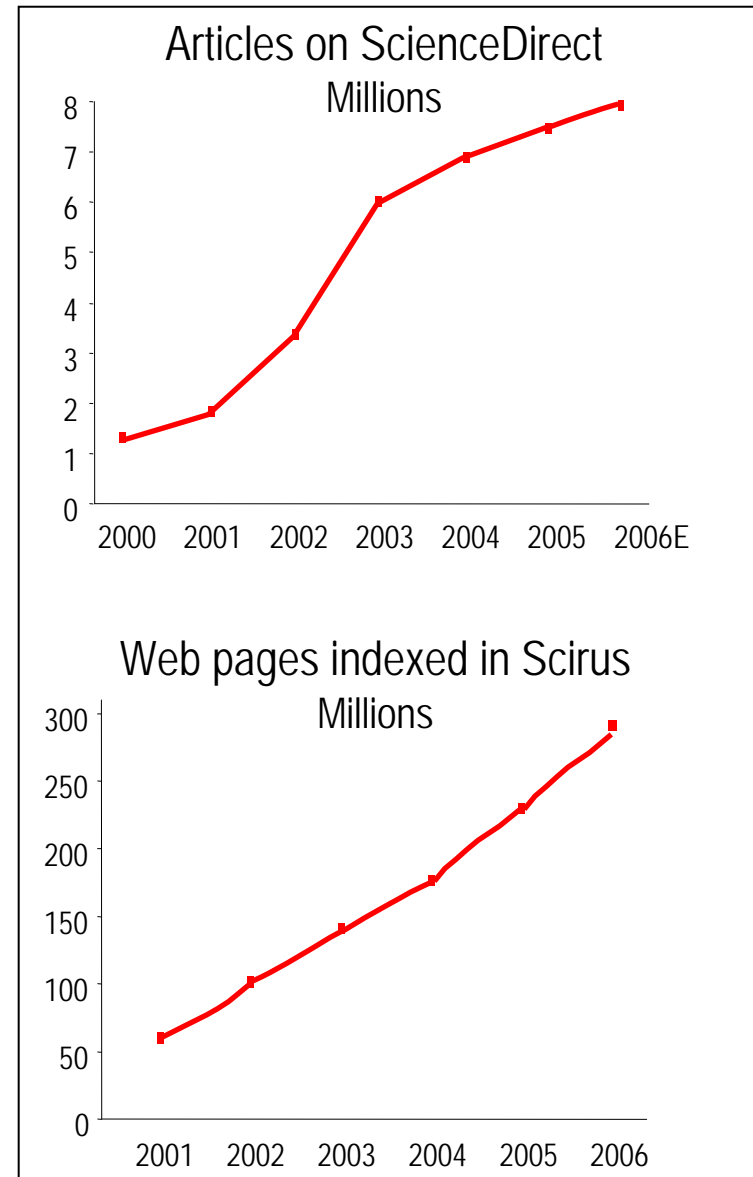
WIKIPEDIA



Internet has made journal publishing just one of many options

## Increasing amount of content available on-line

- High amount of published content
  - Scopus has 30 million abstracts
  - ScienceDirect has 8 million articles
  - # of articles published per year by Elsevier increased from 160K in 2000 to 250K+ in 2005
- Amount of scholarly Web content even higher
  - Scirus currently indexes over 400 million scholarly Web pages
- Size of general Web has exploded
  - Aug '05, Yahoo indexed over 19 Billion pages
  - Google says it indexes 3x more than closest competitor\*



\*<http://battellemedia.com/archives/001954.php>

## Different content discovery methods

---

- Browsing
- Linking
- Alerting
- Searching
- User collaboration/sharing

Covered in Following Slides

## Different content discovery methods

---

- Browsing
- Linking
- Alerting
- Searching
- User collaboration/sharing

Covered in Following Slides

## Browsing remains effective type of content discovery

---

- Journal browsing remains important for content discovery
  - Journal browsing is used to keep abreast of latest developments in subject area
- 31% of all full text article use on ScienceDirect is a result of journal browsing
- Users that start on a journal home page download, on average, 1.9 articles in a single session
- Users are provided with the option to list favourite journals and receive Journal Issue alerts (Table of Contents)

## Different content discovery methods

---

- Browsing
- Linking
- Alerting
- Searching
- User collaboration/sharing

Covered in Following Slides



## Linking articles

---

- Reference linking and 'cited-by' links are very effective content discovery methods
- Publishers are collaborating to ensure correct reference linking (CrossRef)
- 8% of all full text article use on ScienceDirect comes from reference linking. Same is expected from 'cited-by' links
- Next to reference and cited-by links in official literature there is
  - Web references and cited-bys
  - Patent references and cited-bys
  - Clustering
  - Author linking

## Different content discovery methods

---

- Browsing
- Linking
- Alerting
- Searching
- User collaboration/sharing

Covered in Following Slides

## Different content discovery methods

---

- Alerting
  - Journal Issue alerts (RSS)
  - Top articles alert
  - Citation alert
  - Search alert

## Different content discovery methods

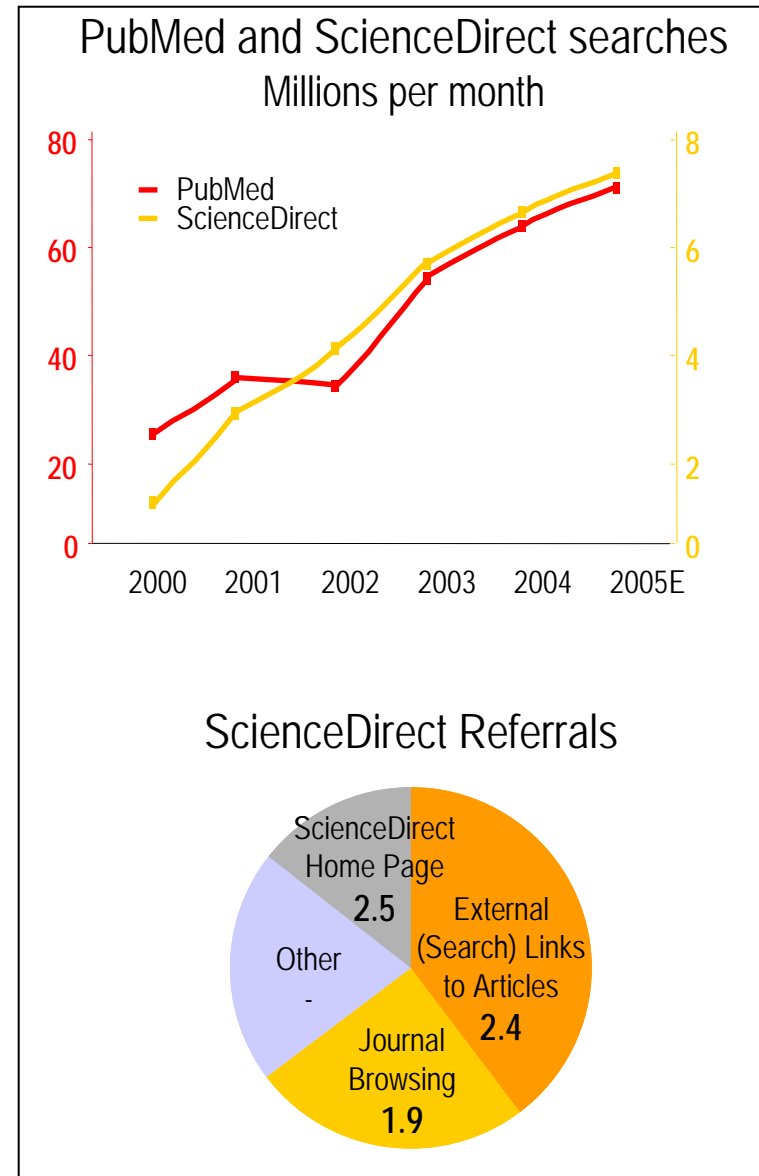
---

- Browsing
- Linking
- Alerting
- Searching
- User collaboration/sharing

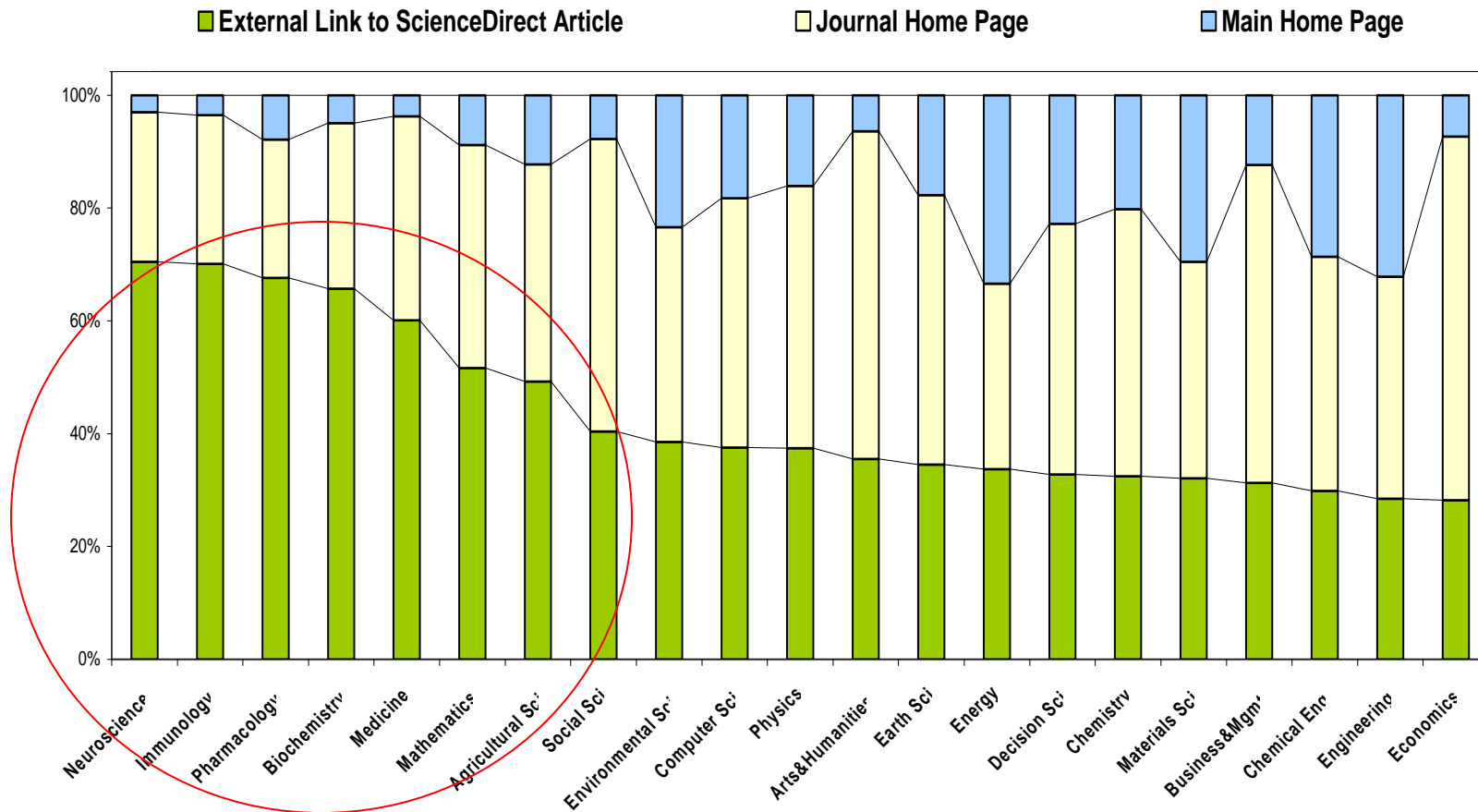
Covered in Following Slides

## Search is main driver of journal article use

- Exponential growth of PubMed and ScienceDirect searches
  - Growth rates between 20 – 110% from 2001 to 2005
- Search has overtaken Browsing
  - In 2005 39% of Full Text Article Downloads on ScienceDirect resulted from sessions starting in External Search platforms
- Search proven more effective than browsing
  - 2.4 full text articles downloads per session starting from external search platform vs. 1.9 per session starting from journal homepage



# Contribution Search differs per subject area



- In Life Sciences 60-70% of users link directly to a ScienceDirect article via an External Search platform
- % lower for subject areas lacking strong subject-specific search platform

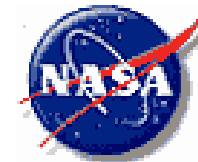
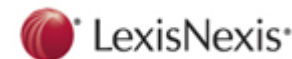
# Search yields more than just journal results!



arXiv.org e-Prints



WIKIPEDIA



## Subject Specific Search platforms remain important

---

- General Web search engines often 1st choice information tool for scientists (66%) and physicians (55%)\*
- Subject-specific search platforms remain important
  - Professionals are becoming increasingly dissatisfied with general search\*\*
  - Average # of full text article downloads per session initiated from PubMed is 3. With alternative platforms this is as low as 1.5



## Important elements in Searching

---

- Indexing
  - Ensuring completeness (including only relevant sources)
  - Capturing existing structures and adding automatic structures and information
- Ranking
  - Date, term frequency and location (title, reference, abstract, full text)
    - » Abstract (A&I Databases) vs. full text searching (Web search engines)
  - Link analysis
    - » Formal citations (Web of Science, Scopus – journal and patents)
    - » Web citations (Google Scholar, Scopus)
    - » Web links and use of anchor text (most Web Search engines)
  - Clustering
    - » Results from one source (most Web Search engines)
    - » Versions (Google Scholar)
- Result classification
  - Keyword or Thesauri (specialised A&I)
  - Subject classification (e.g. math or physics); journal; author etc.
  - Sources and document type classification (e.g. journal article, patent, pre-print or web document) (Scopus, Scirus)
  - Similar results (Amazon, Scirus)

# Scirus search results



**SCIRUS**  
for scientific information only

**FEBS Letters** VIEW Back issues, Special issues, Minireviews, Editorials and much more for FREE on FEBS Letters

- About Us
- Newsroom
- Advisory Board
- Submit Web Site
- Help
- Contact Us

Basic Search [Advanced Search](#) [Search Preferences](#)

neural networks

Journal sources  Preferred Web sources  Other Web sources  Exact phrase

|               |  |
|---------------|--|
| Searched for: | All of the words <b>neural AND networks</b>  |
| Found:        | <b>682,545 total</b>   <a href="#">63,827 journal results</a>   <a href="#">42,680 preferred web results</a>   <a href="#">576,038 other web results</a> |
| Sort by:      | <b>relevance</b>   <a href="#">date</a>  |

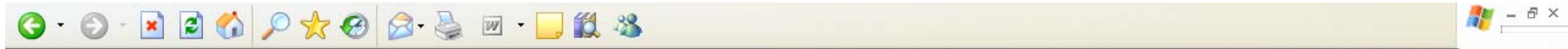
[Save checked results](#) [Email checked results](#) [Export checked results](#)

- [1. RECURRENT NEURAL NETWORKS - FEEDBACK NETWORKS - LSTM RECURRENT NETWORK - FEEDBACK NEURAL NETWORK - RECURRENT NETS - FEEDBACK ... \[30K\]](#)  
Jul 2006  
...Schmidhuber 's page on Recurrent **Neural Networks** (updated June 2006) Why use...**neural** network architectures. **Neural Networks** 18:5-6, pp. 602-610, 2005...traditional recurrent nets. **Neural Networks** 16(2):241-250, 2003. PDF PS...  
[more hits from](#) [http://www.idsia.ch/~juergen/rnn.html]  
[similar results](#)
- [2. NeuroDimension - Neural Network Software, Neural Net Software, Neural Networks, Neural Nets \[14K\]](#)  
Dec 2006  
...Problem Types Resources Intro to **Neural Networks** Intro to Genetic Algorithms...NeuroSolutions.com TradingSolutions.com **Neural networks** and genetic algorithms

Refine your search using these keywords found in the results:

- [activation function](#)
- [artificial life](#)
- [artificial neural network](#)
- [artificial neural networks](#)
- [backpropagation](#)
- [back-propagation](#)
- [cognitive science](#)
- [feedforward networks](#)
- [feed-forward neural networks](#)
- [multi-layer](#)

# Scirus search results



**SCIRUS**  
for scientific information only

- About Us
- Newsroom
- Advisory Board
- Submit Web Site
- Help
- Contact Us

**Advanced Search** [Basic Search](#) [Search Preferences](#)

All of the words  in The complete document

AND

All of the words  in The complete document

Search

|                          |   |
|--------------------------|---|
| <b>Dates</b>             | <b>Only show results published between</b><br>before 1900 and 2008  |
| <b>Information types</b> | <b>Only show results that are</b><br><input checked="" type="checkbox"/> Any information type<br><input type="checkbox"/> Abstracts<br><input type="checkbox"/> Articles<br><input type="checkbox"/> Books<br><input type="checkbox"/> Company homepages<br><input type="checkbox"/> Conferences<br><input type="checkbox"/> Patents<br><input type="checkbox"/> Preprints<br><input type="checkbox"/> Scientist homepages<br><input type="checkbox"/> Theses and Dissertations |
| <b>File formats</b>      | <b>Only show results that are</b><br><input checked="" type="checkbox"/> Any format<br><input type="checkbox"/> PDF<br><input type="checkbox"/> HTML<br><input type="checkbox"/> Word<br><a href="#">List more file types</a>   |
| <b>Content sources</b>   | <b>Only show results from</b><br><b>Journal sources</b><br><input checked="" type="checkbox"/> All<br><input type="checkbox"/> BioMed Central<br><input type="checkbox"/> Crystallography Journals Online<br><input type="checkbox"/> Institute of Physics Publishing<br><b>Preferred Web sources</b><br><input checked="" type="checkbox"/> All<br><input type="checkbox"/> E-Print ArXiv<br><input type="checkbox"/> CogPrints<br><input type="checkbox"/> NASA               |

# How can librarians improve search efficiency?

- Education
  - There are more options available than Google...
  - ...and they are more focused and superior
- Library Integration
  - Search on Library Homepage
  - Open URL integration

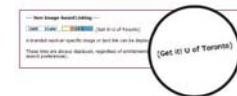
## Are you linked in to Scirus Library Partners?

As a librarian, you want your user community to find the shortest paths to as many full text links as possible. By doing so, you increase article usage and optimise your investment in journals content through your entitlements.

Now Scirus can help you achieve these objectives.

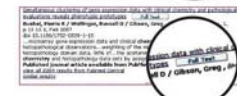
The free and most comprehensive science-specific search engine has integrated OpenURL through the Library Partners Program.

Scirus Library Partners Program offers librarians two methods of setting up OpenURL integration in Scirus. The method you choose depends on your institutions' link resolver settings.



### Non-Image-based linking

With the first method, a branded resolver-specific image or text link is displayed in Scirus search results. The link leads to the Link Resolver where users can see whether they have access to full text or not.



### Image-based linking

A full text icon is displayed in the search results list automatically if the user is entitled to access full text. When the user clicks on this icon-based link, they will be taken either directly to the full text or to a resolver-specific entitlements screen where they can choose the full text or an alternative library service. The latter depends on the institution's link resolver settings.



### HOW TO SET UP?

Set up is easy. You simply log on [www.scirus.com/srsapp/librarypartners/adminlibrary.jsp](http://www.scirus.com/srsapp/librarypartners/adminlibrary.jsp) and supply us with your institution's details and your link resolver base URL. We will respond within 5 working days with guidelines to help you set up the service for your users.

Log on to  
[www.scirus.com/srsapp/librarypartners](http://www.scirus.com/srsapp/librarypartners)  
 to sign up now.

**SCIRUS**  
 for scientific information only

Freely accessible, Scirus is the world's most comprehensive, science-specific search engine with 113 million indexed terms, and growing. Coverage includes journal sources, institutional repositories, preprints, science web pages and much more...

## Different content discovery methods

---

- Browsing
- Linking
- Alerting
- Searching
- User collaboration/sharing

Covered in Following Slides

## What will the future bring

---

- Discovery through communities and networks
  - Community-driven sources selection
  - Social network searching
  - Reviewing and commenting
  - Ranking & classification by users
    - » Active and inactively (through usage & behaviour)

Combining Browsing, Linking, Alerting and Search in  
a Community and Network-Driven system!

Thank you!