Cover Story: Relevance Gone?
### Google Legacy News Search

**Google's plan for a computing platform to destroy Microsoft**

*In stitches: "The Google Legacy: How Google's Internet Search is Transforming Application Software," Arndt says Google is about much more than search...

**Google and the Internet Age**

*Arts: Sep 21, 2005

"Google and the Internet Age by ZDNet's Dan Farber, - Einer Mibe of ZDNet writes about Stephen Amir's The Google Legacy: How Google's Internet Search..."

**Google's grand ambition**

*Arts: Sep 21, 2005

"...a computing platform and could be about to unseat Microsoft from its throne," Amir writes in a summary of his book, "The Google Legacy: How Google's Internet..."

**Google: Time for world domination?**

*Arts: Sep 21, 2005

"...a computing platform and could be about to unseat Microsoft from its throne," Amir writes in a summary of his book, "The Google Legacy: How Google's Internet..."

### Google Legacy Web Search

**Amid: The Google Legacy (ebook)**

Google search, internet future, ebook, search engine, email, office software, mapping software.

[www.infonetics.com/publications/google/googlelegacy.html](http://www.infonetics.com/publications/google/googlelegacy.html) - 1k - Cached - Similar pages - Remove result

**Google Effect Search Engine Meeting 2005 Tutorial**

Tusday, 11 April 10, 2006 (afternoon). The Google Effect: Redefining the Landscape of Search. Stephen E. Amir, President...

[www.infonetics.com/searchengine/110406/05tutorial-amir.html](http://www.infonetics.com/searchengine/110406/05tutorial-amir.html) - 3k - Cached - Similar pages - Remove result

**The Google Legacy: Ebook Released**


[www.threadwatch.org/node/5602](http://www.threadwatch.org/node/5602) - 13k - Cached - Similar pages - Remove result

**Google Legacy order form**

The Google Legacy: How Google's Internet Search is Transforming Application Software. Stephen E. Amir... I wish to order access to The Google Legacy...

[https://site544.myk8now.com/asanid?https://google-order.html](https://site544.myk8now.com/asanid?https://google-order.html) - 6k - Cached - Similar pages - Remove result
What We’ll Cover

• Situational relevance
• Interfaces... growing importance
• Internet vs Intranet relevance
• You can cheat

Forces for Relevance

• Concerns about time wasted looking for information
• Under pressure, users accept “good enough”
• Web search
  – Voting
  – Popularity
• Intranet search
  – Algorithms that “read”
  – Human intervention
Relevance… Underseige

- *Dot Net* reported in August 2005:
  - Overlap among Ask Jeeves, Google, and Yahoo on 10,000 queries was as low as three percent
  - Coverage of search engines and their query matching algorithms are diverging
- 90 percent of search users
  - Cannot differentiate between “organic” search results and ads (“inorganic search results”)
  - Click on “what seems to be the best answer”
  - Do not recognize when searches run automatically and display results in a work-related system
- What is “relevance”? – What is beauty?

Relevance... A Fuzzy Black Box

Precision and recall give way to “situational relevance”
The Black Cube’s Needs

- Need for users
- Need for partners
- Need to respond to competition
- Need for traffic (clicks)
- Need for revenue
- Need for buzz

Intranet Search

Content must be in the system

+ Suggestions in the interface

+ Yahoo style plus work flow hooks
Internet Search
Web sites **must** be in Google

+ 

Appear in first page of results
and
In ads on the first page of results
but
Tricky and resource intensive

Web Search

- Best results:
  - Click a hard wired “topic” like News, Weather, etc.
  - Use Boolean syntax
- Google / MSN / Yahoo improving deduping
- Becoming more alike even though overlap among results lists range from five percent to 50 percent in our tests
- Conclusion…
The Obvious

What we’ve known all along...

If you know what you are looking for, a trained searcher can find relevant information

The Less Obvious

What we’ve known all along...

Users who take what is provided accept “good enough” information

Good enough is now dominant in search
The User Interface Gets Training Wheels

- Search takes time
- People have a hard time coming up with a query
- Long results lists are ignored
  - Clustering: FirstGov.gov will use MSN spider and Vivisimo federation engine with clustering
  - Canned queries: Push results so no search is needed
  - Appliance solution: Google Appliance, Index Engines, or EPI Thunderstone appliance

Mondosoft “Hybrid” Interface
Hybrid Search: Facets, Hard Coding, Synonym Expansion

What’s Needed for “Faceting”

- Human involvement
  - Stop word list essential
  - Synonym list
- Software tools
  - Data Harmony’s controlled vocabulary editor
  - Tools included in ISYS 7, Stratify, Endeca, Mondosoft and others for interacting with the search system
- Metrics and usage analysis
- Time
- Resources
“Relevance” via Cheating

- Linkfarms
- Fool the Googlebot with a doorway page
- Blog spam
- Duplicate content
- Etc.

Summary

- Many options in search… more coming
- Focus on basics
- Remember the principles of good librarianship:
  - Collections
  - Accuracy
  - Currency
  - Selective depth
- Read authoritative material; short cuts will:
  - Get you in the sandbox or blacklisted
  - Lower your ranking
  - Waste your money unless you are informed
Management consulting and strategy
“The Enterprise Search Report”
2nd edition now out. Order at www.cmswatch.com
“The Google Legacy” became
Available earlier this month
Order at www.infonortics.com

Additional information at
www.arnoldit.com/sitemap.html
Contact: sa@arnoldit.com