
Diversity Preferent Ranking Improves Query Results

What can Search Engine Designers Learn about Users
from the Reference Interview?

John G. Dove
CEO, Xrefer Ltd.

Today's Talk

- Just a small part of what's in the paper
- This is part of a larger inquiry into what reference systems designers can learn from reference librarians
- So we'll:
 - See why reference librarians are better at understanding the queries of their patrons than query systems are
 - Take a brief look at how we at Xrefer model users' needs
 - Take a simple example of looking for an acronym through Google to see how they seem to be using diversity to improve results
 - Look at Diversity Preferent Ranking in Xreferplus
- End with the bigger question: How can systems be customizable by librarians to apply their knowledge of how information needs of their users/patrons map to available resources?

Audio.

Vision from Google's Director of Research

“In 50 years the scene will be transformed. Instead of typing a few words into a search engine, people will discuss their needs with a digital intermediary, which will offer suggestions and refinements. The result will not be a list of links, but an annotated report (or a simple conversation) that synthesizes the important points, with references to the original literature. People won't think of "search" as a separate category - it will all be part of living.”

Peter Norvig, Google's Director of Research, From issue 2578 of New Scientist magazine, 18 November 2006, page 50.

Reference Librarians vs. Query Systems

- Mingfang Wu, Royal Melbourne Institute of Technology—private correspondence
60% of queries end up being reformulated by the user
- Ragnar Nordie from Oslo College, Library and Information Studies:

“User revelation” - a comparison of initial queries and ensuing question development in online searching and in human reference interactions.

Reference interviews: 90% effective

Query systems: 25 to 50% effective

“...topical searches fail most frequently (between $\frac{1}{2}$ and $\frac{3}{4}$ of the time in different studies); and users have difficulties both in choosing search terms to represent their problem and in redefining search terms in case of failure.”

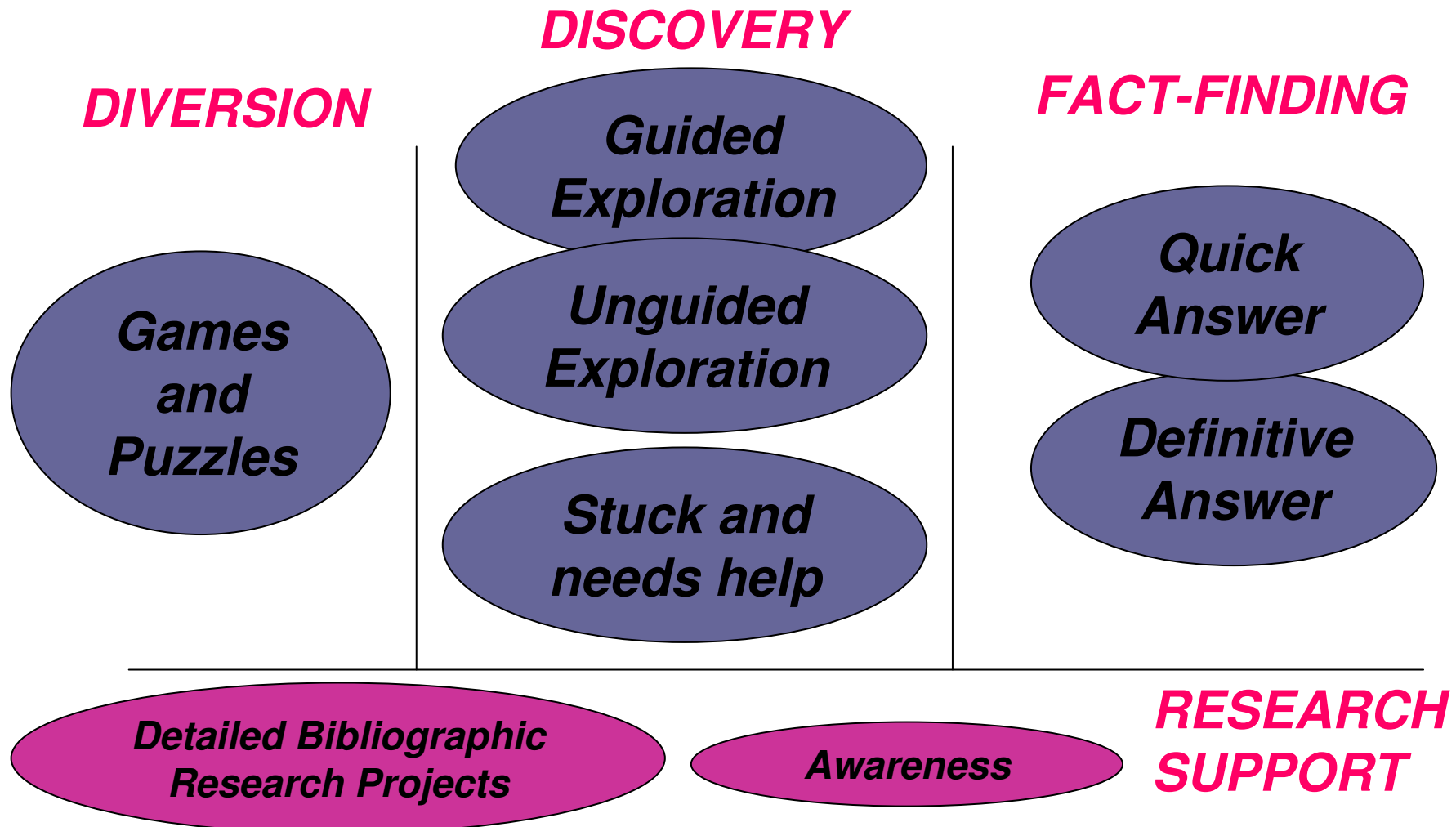
Why are reference librarians skilled in the Reference Interview so much better?

- The librarian can draw on a rich set of contextual information even before the patron or student appears with a question.
 - Robert Taylor (1968) identifies 5 ‘filters’ narrowing the librarians’ attempt to fit resources to need:
 - likely subjects
 - likely motives and objectives of inquirer
 - personal characteristics of inquirer
 - relationship of the inquirer to the institution
 - probable level or characteristics of acceptable answers
- Question for Query System Developers: can any of these factors be used to allow reference librarians to customize the self-service reference experience of their patrons?

A Quick Look at Xrefer's model of Reference System Users

- 8 “modes”
- Each represents a different set of goals and elements of satisfaction.
- In particular, modes reveal a different tolerance for false negatives and false positives, ranging from:
 - extreme annoyance to
 - utter delight and fascination (serendipity)
- An individual reference ‘experience’ may have the user shift from one mode to another

User Modes for Reference



Where does diversity come in?

- Users under-specify the desired results for several reasons:
 1. economy (time and effort) of specifying
 2. want to see a whole set of choices so that they can direct the ‘drill-down’
 3. lack of the vocabulary to effectively navigate the focusing choices
- In all three of these cases having a diverse set of distinct choices at each level of the search helps the user.

A simple example from the “Quick Answer” mode

- An acronym you don’t know.

You’re reading a paper on events during the 2003 Iraq war and read:

“ATMs were first deployed by the US Army and Marine Corps and Australian Special Forces during Operation Iraqi Freedom in March/April 2003.”

It’s clearly not the two acronyms you know for ATM, “Automatic Teller Machine” nor “Asynchronous Transfer Mode”

What you'd really like at this point:

- A simple list sufficiently long to find the acronym that fits and get on with what you were doing:

- Able Toastmaster, a post-nominal for recipients of Advanced Toastmaster awards
- Adobe Type Manager
- Advanced Technical Materials
- Advanced Traffic Management
- Aggregate Trailer Mass
- Air Traffic Management
- Amateur Telescope Maker
- Anti-Tank Missile
- Association of Teachers of Mathematics
- Asynchronous Transfer Mode
- At the Money, designation for options priced at approximately the same price as the underlying security
- Ataxia Telangiectasia Mutation
- Automatic Teller Machine
- Unit of pressure for water-resistant watches equal to atmospheric pressure at sea-level

- Let's see what Google gets you searching for "ATM"

Look at Diversity Preferent Ranking in Xreferplus

- Xreferplus has, on average, 10 links for every entry, linking an entry from one reference source to other sources available to the same user. These give the user the value of a single reference resource that librarians like to call: “a one-stop shop for online reference”
- We call these resource-to-resource links, “xreferences”

whistler mother and ch

SEARCH

Search All

Search The Bridgeman Art Library Archive

[Browse topics](#) | [Concept Map](#) | [Advanced search](#)

[Home](#) > [All topics](#) > [Art](#) > [The Bridgeman Art Library Archive](#) > [m](#)

- ▶ [email to a friend](#)
- ▶ [printer-friendly version](#)
- ▶ [your search results](#)

xreferences

- [Whistler, James Abbott McNeill \(1834 - 1903\)](#)
Bloomsbury Biographical Dictionary of Quotations
- [Whistler, James Abbott McNeill \(1834 - 1903\)](#)
The Bloomsbury Guide to Art
- [Whistler, James Abbott McNeill \(1834 - 1903\)](#)
The Hutchinson Encyclopedia, Helicon
- [Whistler, James \(Abbott\) McNeill 1834 – 1903](#)
Chambers Biographical Dictionary
- [Artists](#)
Bloomsbury Thematic Dictionary of Quotations

[view all xreferences](#) (22)

adjacent entries

- [Mother and Child, 1897 \(pastel on paper\)](#)
- [Mother and child \(pastel on paper\)](#)
- [Mother and Child in the Park, 1914 \(oil on canvas\)](#)
- **Mother and Child on a Couch (w/c)**
- [Mother and Child Watching the Ducks](#)
- [Mother and Child with a](#)

Trial tip: try the [xreferplus starter quiz](#)



About [The Bridgeman Art Library Archive](#) from [Bridgeman](#)

Mother and Child on a Couch (w/c)



Diversity of xreferences

xreferences

- [Whistler, James Abbott McNeill \(1834 - 1903\)](#)
Bloomsbury Biographical Dictionary of Quotations
- [Whistler, James Abbott McNeill \(1834 - 1903\)](#)
The Bloomsbury Guide to Art
- [Whistler, James Abbott McNeill \(1834 - 1903\)](#)
The Hutchinson Encyclopedia, Helicon
- [Whistler, James \(Abbott\) McNeill 1834 – 1903](#)
Chambers Biographical Dictionary
- [Artists](#)
Bloomsbury Thematic Dictionary of Quotations

[view all xreferences](#) (22)

Referring to:

- 2 Quotation resources for quotes by or about James Whistler
- 1 Subject Guide for Art
- 1 Biographical Dictionary
- 1 General Encyclopedia

And finally the question we began with:

How can systems be customizable by librarians to put some of their knowledge of how information needs of their users/patrons map to available resources?

And, yes, in case you were wondering, there is a 'mind reading machine' available on eBAY:

« [2005-04-29 Spike activity](#) | [Main](#) | [Mental health and human rights](#) »

April 30, 2005

'Mind reading machine' for sale on eBay:

More futuristic eBay tomfoolery: "hello, i am selling what i believe to be a [mind reading machine](#) built by Dr. J. S. Strauss in the year 2282".

It is difficult to write anything about the auction page that even partially captures its kooky brilliance.



Although you may be interested to know that apart from getting a mind reading machine from the future, you also get a "picture of a young lady in a waterfall" thrown in.

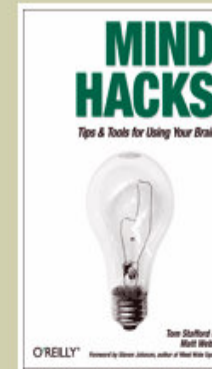
Something tells me that despite solving some fundamental problems in cognitive science by the year 2282, neuroscientists may still be spending a little too much time in the lab.

[Link](#) to eBay page *MIND READING MACHINE ?for minds? like time machine: i found it in my attic wrapped in a bed sheet (via [anomalist](#))*

—[Vaughan](#).

Posted at April 30, 2005 08:00 AM

Neuroscience and psychology tricks to find out what's going on inside your brain.



Mind Hacks is a book by Tom Stafford and Matt Webb. [Find out more](#), or buy it:

→ [at Amazon](#) (34% off)

→ [at Amazon UK](#) (30% off)

SECTIONS

- [Authors](#)
- Chapters
 - [Attention](#)
 - [Hearing and Language](#)
 - [Inside the Brain](#)
 - [Integrating](#)
 - [Moving](#)
 - [Other People](#)
 - [Reasoning](#)
 - [Remembering](#)
 - [Seeing](#)

My favorite historical
example of a query
system available to the
public at no charge.

Query System of 1850?



http://commons.wikimedia.org/wiki/Image:British_Museum_Reading_Room_Panorama_Feb_2006.jpg

75,000 volumes all within line-of-sight on 30 ft. three story bookshelves
1,200,000 volumes down in the stacks
Open freely to the public—no sign in or members required
Staffed by scores of reference librarians