When a Librarian's not there to Ask: Creating an Information Resource Advisory Tool

Tracy A. Gabridge, Nicole Hennig, Rebecca Lubas, Sarah G. Wenzel

It is 2am. A professor wakes up with a new direction for her research; she must immediately learn about bioethics. In a dorm a student is finally ready to begin a paper on Cuba. Where do they turn? The library web site presents them with a bewildering array of resources and no librarian on hand to serve as intermediary. How can librarians facilitate research in their absence? What interfaces can be designed to educate users in their search? What metadata is needed to enable accurate retrieval?   What is the librarian's role in the increasingly indirectly-mediated information-seeking environment? Can the reference interview be effectively translated into a search interface? This paper describes a step towards resolving these issues by creating an on-line tool to assist users in selecting the database(s) most germane to their research needs.

The MIT Libraries, a physically distributed system of five primary libraries and four branches, are rich in electronic resources, including over 300 databases. However, currently there is no good on-line method to steer patrons to the most appropriate resources. While patrons who ask for help are given the benefit of our expertise, those who do not – a large majority – must use their own judgment. A study of graduating seniors given in March 2004 indicates clearly that most students rely on library resources only occasionally. Forty-four percent had never asked a librarian for help finding information. At the same time, sixty-six percent searched the Internet for course material "very often."[1] In comparison to other institutions, a 2001 College Student Experiences Questionnaire reports that MIT students are less likely to use the library to study or ask a librarian for help than are their peers.[2] We seldom see graduate students or faculty, who tend to rely on their colleagues for information. Anecdotally, however, while this population is often very familiar with traditional resources in a field, it is much less sure of resources in new or interdisciplinary areas.

At MIT, users are fortunate to have a resource (Virtual Electronic Resource Access, or Vera[3]) that categorizes databases by subject. Nonetheless, these subjects are painted with broad strokes. A patron looking in "Biology & Neuroscience" must choose from an alphabetical list of twenty-two databases. The brief descriptions provided are little help to even a sophisticated user and usability tests show that our patrons do not read descriptions.[4] Alphabetical order privileges databases whose names begin with the initial letters of the alphabet. It is surmised that some of the volume of interlibrary loan dissertation requests is due to nearly every category containing "Dissertation Abstracts," which happens to begin with one of the first letters of the alphabet.

Another attempt to meet the challenge of communicating to users the complex research environment on-line has been the creation of subject guides covering fifty-five areas taught or researched at MIT.[5] Although the result of intense librarian efforts, these guides are underutilized. In usability testing, users are invariably astonished to discover that the pages exist and appreciate the resources on them. However, the guides are not generally well known; users

---

[1] Institutional Research, Office of the Provost, MIT, "Senior Survey," Consortium on Financing Higher Education, March 2004.

[2] Institutional Research, Office of the Provost, MIT, "CSEQ Report," C. Robert Pace and George D. Kuh, College Student Experiences Questionnaire Research Program, Indiana University Center for Postsecondary Research and Planning, 5th Ed, 2001.

[3] MIT Libraries, "Virtual Electronic Resource Access," http://libraries.mit.edu/vera.

[4] MIT Libraries' Web Advisory Group, "The 'Big Test' Usability Results," MIT Libraries, November, 2002, http://macfadden.mit.edu/webgroup/usability2002/big-test/summary.html

[5] MIT Libraries, "Subject Guides," http://libraries.mit.edu/help/subjects.html.

admit they would not think to look for such a resource on the library site.[6] Many print reference resources are not being used even when they are the appropriate, if not unique, source. The subject guides vary widely in quality and quantity of resources featured. Even more importantly, the relationship between databases and other library resources (including subject librarians) is not always clear to the user.

An issue for both Vera and the subject guides is the lack of a consistently applied metadata. Vera does have a limited controlled vocabulary, but the application of even basic subject terms can be haphazard. Vera is a reasonably good guide to our electronic databases and journals; however, outside of the usual technical services workflow, it fails to provide universally reliable or sufficiently deep subject access.

With issues of informed user selection of and access to resources increasingly prominent, a task force investigated the feasibility of translating the reference interview into the on-line environment. At the same time, a parallel task force elaborated requirements for a federated search project. The initial goal was to produce requirements for a Database Discovery Tool that would offer a simple, self-teaching interface to the collection of databases. In the process, we realized that no effective tool can limit itself to resources available only electronically; we needed to design a tool that will direct users beyond databases when appropriate, including reference books, relevant web sites and librarians. The final report focused on this concept: an Information Research Advisory Tool (iRAT).

A prototype of a similar system is proposed by Tim Cole and Wei Ma in their Smart Database Selector, using downloads of the controlled vocabularies of the resources within the Selector system.[7] The Smart Database Selector system offers a method of automatically narrowing down a list of databases without requiring additional steps from the user, but also without requiring additional, non-productive searches in a library's licensed resources.

The task force explored the necessary elements involved: knowledge management, metadata and interface design. We examined models and available metadata schemes to determine which would be most effective. This examination also required outlining key elements of a reference interview and methods by which librarian questions facilitate patrons' elaboration of their own needs.

This paper describes the process used to create the iRAT concept and discusses implications for reference services. Subsequently we elaborate on the faceted browsing model and articulate the metadata scheme. Finally, we explore implementation possibilities and imagine future enhancements.

Process

The task force consisted of four librarians: two reference and subject librarians; the Libraries' web manager and expert in usability testing; and the head of cataloguing and metadata services. The report was completed after four months. We explored what other organizations in and outside of the library community were currently doing in terms of electronic pathfinders and knowledge management, evaluated the usefulness of existing metadata schemata, and drew from usability studies and the collected knowledge of our colleagues. This exploration included asking

---

[6] MIT Libraries' Web Advisory Group, "The 'Big Test' Usability Results," MIT Libraries, November, 2002, http://macfadden.mit.edu/webgroup/usability2002/big-test/summary.html.
[7] Wei Ma, "A Database Selection Expert System Based on Reference Librarians' Database Selection Strategy: A Usability and Empirical Evaluation," *Journal of the American Society for Information Science and Technology*, 53, 7 (2002): 567-580.

colleagues, over a period of a month, to fill out a form when they went through the process of recommending a database to a patron.[8]

The overall philosophy we developed for our tool is that of a self-teaching interface simulating the reference interview and providing a single starting point for all audiences, with different paths. We also assumed that implementation will occur while Vera and the on-line catalog exist, and include more than electronic resources. Our target audience is inexperienced researchers, such as undergraduates, who may not know where to begin; users doing research outside of their area of expertise; or researchers needing starting points for interdisciplinary research.

Next we created a list of the necessary metadata, and chose a scheme. We also created a prototype of the interface, focusing on faceted browsing and a highly refined search screen to create a self-teaching interface and simulate a reference interview. Our last step was to create a list of requirements we submitted to vendors in order to choose a product with which to work.

This task force and the federated searching task force combined their reports and recommended that both be implemented, with the understanding that federated searching would be implemented first.[9] However, we believe strongly that in order for federated searching to be useful, patrons must have a sense of which databases it makes sense to search in tandem.

Reference implications

One of the chief motivations for the creation of an Information Research Advisory Tool (iRAT) is to respond to the fact our patrons are not asking us for assistance in locating information. It should come as no surprise in the current climate of declining reference statistics and increasing on-line access that MIT is among the institutions struggling to assist patrons whom we never see or even know of (save when the server fails).

A skillful reference librarian provides a valuable service not only by assisting patrons in articulating their research concept, but also in explaining how information can be organized and directing them to appropriate resources. The important set of interactions that make up a reference interview are lost in an on-line list of databases. Subject guides frequently provide more guidance, but they require patience and a grasp of the organization of knowledge in a field to be useful. Live, on-line, reference services have been suggested as an alternative. In our experience with such a service, when asked, "You are not in the library and you need help finding information," our patrons responded that their preferred mode of assistance is "use a print or online guide."[10] Since our patrons prefer not to ask for help, we believe that creating a resource advisory tool is a better use of librarian time and knowledge.

The challenge lies in transposing or simulating the reference transaction into an on-line asynchronous environment. Our project focuses on elements of organization that we thought relevant to most fields, the inclusion of important resources independent of their format, and on the design of the interface. The iRAT envisioned is a web-based interface to a database of resources that allows a user to query the system and narrow the results; the set of criteria may include:

> • Subject, and narrower aspects of a subject (e.g. Business and Management Companies – News)

---

[8] DDT Requirements Group, "Librarians' Thought Process Survey," MIT Libraries, http://macfadden.mit.edu:9500/dmag/form.html.
[9] DDT/Metasearch Analysis Group, "Resource Discovery and Metasearch Requirements," MIT Libraries, http://macfadden.mit.edu:9500/dmag/MIT%20Libraries%20requirements%2007292004.pdf.
[10] Reference Vision Task Force, "Research Help Survey," MIT Libraries, http://macfadden.mit.edu:9500/refcomm/RefVision/documents/survey.pdf.

- Scope of project (a short paper versus thesis)
- Types of information (e.g. articles, books, facts, reviews, etc.)
- Dates of coverage

A user will be able to manipulate the criteria in two ways to retrieve a manageable number of results: via a traditional query through a web form with pull-down menus or via a faceted browse (such as that used at epicurious' recipe browse search).[11]

If a user has already identified a particular resource, two additional navigation methods will be available. One is to search by title, much like the current Vera interface. Alternatively, the user can create a list of relevant resources (which may be saved throughout multiple sessions). "My iRAT" functionality could be available via a personalized login/password.

However successful we are in introducing people to recommended resources and key concepts in the organization of knowledge, we will not be able to substitute for person-to-person interactions or capture the art of the creative reference interview. We can only hope that the inclusion of the subject librarian in the list of resources and the ready availability of the link to our mediated reference services as well as other marketing efforts influence patrons to seek assistance from our expert staff.

Model

The extent of our success relies on developing a model that patrons find convenient and reliable. Traditional search interfaces did not seem to serve to educate patrons as well as a model we developed that relies on the benefits of faceting browsing.[12] For those unfamiliar with this search concept, faceted browsing may be defined in a number of ways. A relatively clear definition may be paraphrased from the work of Keith Instone.

> Faceted browsing is an interaction style where users filter a set of items by progressively selecting from valid values of a faceted classification system. Instead of clicking on categories in a hierarchy and finally arriving at results at the ends (leaf nodes), one clicks, gets a sample, then either browses the results or clicks again to narrow further.

Fig. 1. Hierarchy browse vs. faceted browse. Reprinted by permission from Instone, Keith, "Faceted Browsing - How User Interfaces Represent and Benefit from a Faceted Classification

---

[11] epicurious.com, "Recipe Browse," http://www.epicurious.com/recipes/find/browse/ (accessed 7 January 2005).

[12] Ka-Ping Yee et. al., "Faceted Metadata for Image Search and Browsing," in *Proceedings of the ACM Conference on Computer-Human Interaction*, 2003, http://bailando.sims.berkeley.edu/papers/flamenco-chi03.pdf (accessed January 7, 2005).

System," presented at SOASIS&T, July 8, 2004, http://user-experience.org/uefiles/facetedbrowse/.

What is faceted classification?

Faceted browsing relies on a faceted classification scheme, which may be described as follows:

> A faceted classification differs from a traditional one in that it does not assign fixed slots to subjects in sequence, but uses clearly defined, mutually exclusive, and collectively exhaustive aspects, properties, or characteristics of a class or specific subject.

[Wynar, Bohdan S. Introduction to cataloging and classification. 8th edition. p. 320 http://kmconnection.com/DOC100100.htm]

Examples of faceted browse

Perhaps faceted browsing can best be understood by looking at some examples. This site for Boston restaurant reviews is a good example: http://www.local-i.com/bostonrestaurants/restaurants/
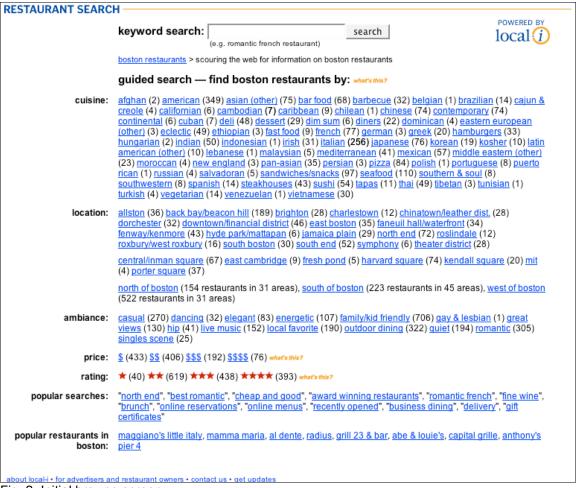

Fig. 2. Initial browse screen

In this example, all facets show on the first screen, (cuisine, location, price, etc.). Notice also that the number of items in each facet shows in parentheses after each link.

Fig. 3. Number of items in each category is shown

Fig. 4. More facets in left margin

If "chinese" is selected on the first screen, and "award-winning" is selected on the second screen, this results screen appears. Results are always shown in the main body of the page; the left side of the screen is used to show remaining facets that can be selected to narrow the search.

A user may choose to narrow and combine facets in different ways. This provides an interesting way to explore the contents of a database (in this case restaurant descriptions), all the while learning about the items in the database and possible ways to describe the contents.

Notice that the top of the screen may also contain "breadcrumb" links to indicate which choice the user has already selected.

Fig. 5. Breadcrumb links

This example also makes use of links to related topics, i.e. "similar cuisines."

A user may choose to combine as many facets as possible until no more choices are possible. One useful feature of this example is that if there are no results for a particular combination of facets, that choice is not shown in the list; it is impossible to get zero hits. Some systems choose to show facets with zero hits as "dimmed out," so that the user will know that the other facets exist although they don't apply to the choices made so far.

Examples of a potential system for browsing information resources in the MIT Libraries

We applied this model to a database of information resources that might exist in a research library to create a scenario demonstrating how facets might guide a user. One example uses a business school student looking for company information.
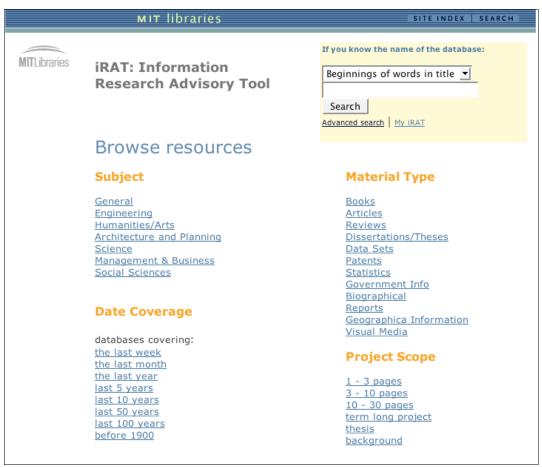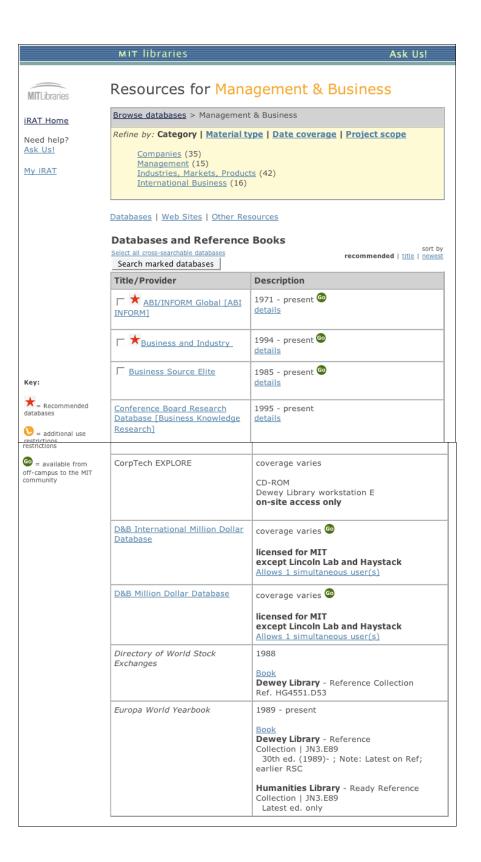

Fig. 6. Mockup of a possible faceted browse system

The screens used here are mockups using fictional data. The facets used include subject (and sub-categories), material type, date coverage, and project scope.

Resources for Management & Business

Refine by: **Category | Material type | Date coverage | Project scope**

Companies (35)
Management (15)
Industries, Markets, Products (42)
International Business (16)

Databases | Web Sites | Other Resources

Databases and Reference Books

Select all cross-searchable databases

[Search marked databases]

sort by
recommended | title | newest

Title/Provider	Description
☐ ★ ABI/INFORM Global [ABI INFORM]	1971 - present **Go** details
☐ ★ Business and Industry	1994 - present **Go** details
☐ Business Source Elite	1985 - present **Go** details
Conference Board Research Database [Business Knowledge Research]	1995 - present details
CorpTech EXPLORE	coverage varies CD-ROM Dewey Library workstation E **on-site access only**
D&B International Million Dollar Database	coverage varies **Go** **licensed for MIT** **except Lincoln Lab and Haystack** Allows 1 simultaneous user(s)
D&B Million Dollar Database	coverage varies **Go** **licensed for MIT** **except Lincoln Lab and Haystack** Allows 1 simultaneous user(s)
Directory of World Stock Exchanges	1988 Book **Dewey Library** - Reference Collection Ref. HG4551.D53

Europa World Yearbook	1989 - present
	
	 Book
	 Dewey Library - Reference Collection | JN3.E89
	   30th ed. (1989)- ; Note: Latest on Ref; earlier RSC
	
	 Humanities Library - Ready Reference Collection | JN3.E89
	   Latest ed. only

Key:

★ = Recommended databases

= additional use restrictions

Go = available from off-campus to the MIT community

Selected Web Sites return to top

Title
★ Market Places and Quotations Publisher of the web site | details | http://www.website.com
Industry Link IndustryLink, Ontario, Canada | details | http://www.industrylink.com/
Forbes Annual Report on American Industry Forbes | details | http://www.forbes.com/
North American Industry Classification System (NAICS) U.S. Census Bureau | details | http://www.census.gov/epcd/www/naicstab.htm
Standard Industrial Classification Manual U.S. Dept. of Labor | details | http://www.osha.gov/oshstats/sicser.html

Other Resources return to top

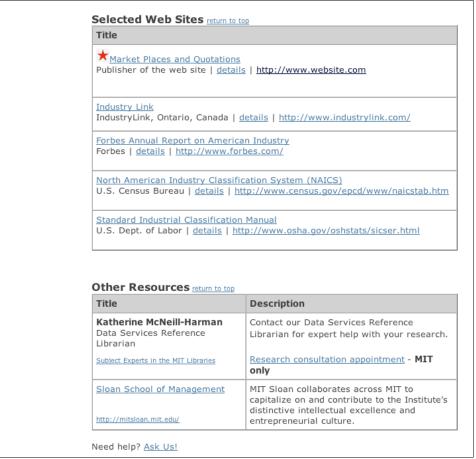| Title	Description
Katherine McNeill-Harman Data Services Reference Librarian Subject Experts in the MIT Libraries	Contact our Data Services Reference Librarian for expert help with your research. Research consultation appointment - **MIT only**
Sloan School of Management http://mitsloan.mit.edu/	MIT Sloan collaborates across MIT to capitalize on and contribute to the Institute's distinctive intellectual excellence and entrepreneurial culture.

Need help? Ask Us!

Fig. 7. Initial results screen

In this example, the user selects Management and Business. The sub-topics displayed are "Companies," etc. The box at the top of the display shows the other facets as well: material type, date coverage, and project scope. At this point the user may continue narrowing by topic, selecting "Companies" – or may choose another facet, such as "material type."

At every step, results display in the body of the page. Databases and reference books display first, then selected web sites, then other resources, such as the relevant librarian (yes, we catalog people!), and relevant university departments.
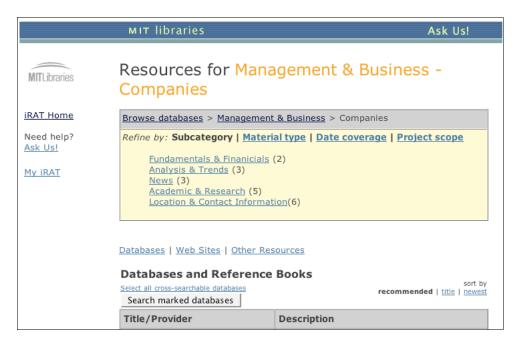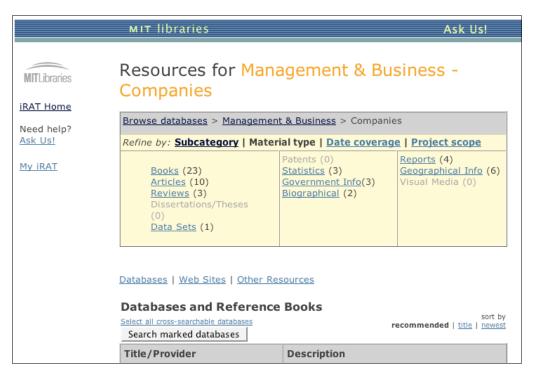
Fig. 8. More subcategory choices

Fig. 9. Material type choices

To continue the example, the user selects "Companies," then "Material type." Note the different material types displayed (e.g., article, books, etc.). Numbers after each link indicate how many items are in that facet. We chose to show facets with zero hits as "dimmed out," in order to demonstrate that those facets exist and might be relevant had he followed a different path. At any point he may choose to go back up a level and make different choices; results will be displayed at every step of the way.

These examples help us to think about what aspects of our research materials could be informatively shown to users. This type of browse experience makes much more information available to users than other interfaces. It also encourages trying different combinations, thus learning more about what is available.

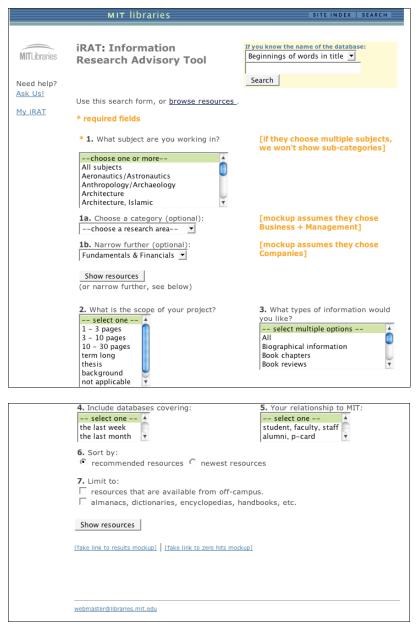Contrast that to a typical "advanced search" screen below.

Fig. 10. Advanced search screen (not a faceted browse)

A screen like this would be provided as an alternative interface, but imagine the typical scenario where a user carefully fills out all the fields and happens to select combinations that result in zero hits. Everything is invisible until the search is performed. Given zero hits, the user doesn't know which combination of items caused that. The only choice is to try again with different selections.

There are many advantages of faceted browsing, which

- helps users understand the scope of the database
- shows item characteristics from the beginning
- suggests what questions should be asked to select the best resources
- eliminates sets with zero hits

All of these combine to create a self-teaching interface that teaches students about the materials we own, just by using the system.[13] This mode of instruction is more and more necessary in a world where few people read "help" files, use "tutorials," or ask for help from librarians.

Relationship to metasearching

The self-teaching interface to select resources allows a user to define and focus intelligently on a well-selected set. However, it does not indicate which among those might contain the most information or how best to search within them. The user, ready to begin to search for specific items to ultimately meet an information need, requires further assistance. This is the point at which cross-database searching dovetails with resource discovery.

The resource discovery results page, illustrated below, shows the potential of a diverse set of resources; it may include print resources, CD-ROM databases and a variety of on-line resources, such as databases. The user may decide to retrieve a print item or CD-ROM from its indicated location, search an online resource via its native interface by clicking on the resource title, or search multiple online resources simultaneously with cross-database searching.

Cross-database searching allows a person to select multiple resources in which to execute a single search. The resource discovery results set illustrated in the mockup shows check boxes next to resources for which cross-database, or metasearching is available. Once the user is satisfied with a set of resources, she can choose to select those boxes and "Search marked databases". At that point she will be asked for keywords to be searched.

---

[13] Larry L. Constantine and Lucy A.D. Lockwood, "Instructive Interaction: Making Innovative Interfaces Self-Teaching, preprint (January 2002), http://www.foruse.com/articles/instructive.htm http://www.foruse.com/articles/instructive.pdf (accessed January 7, 2005).
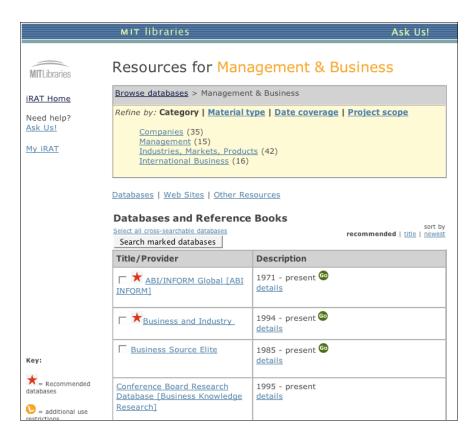
Fig. 11. Results screen

The metasearch system then sends out properly formatted queries to target resources, retrieves matching hits, and coherently presents the results list.

Using the iRAT before performing a cross-database search provides many benefits to searchers, libraries and database vendors. Without guidance to select appropriate databases, the potential list to search may be quite large. While cross-database searching is more efficient than searching individual resources, there are currently limits to the number of databases that can be simultaneously cross-searched. Additionally, the efficacy of such searching is lowered for a user if there are high numbers of irrelevant records retrieved from inappropriate databases. Lastly, database vendors do not welcome large numbers of irrelevant metasearches as searching is one of the more expensive aspects of providing on-line services. Irrelevant searches drive up the cost of doing business; perhaps ultimately the costs might be passed on to their customers, the libraries.

Cooperative metadata

Sufficient and well-defined metadata to describe the resources included are crucial to the success of Information Research Advisory Tool (iRAT) searches. Therefore, we devised a list of metadata requirements to drive the desired searches. The metadata requirements use elements based on the Dublin Core, a highly flexible descriptive metadata standard. Dublin Core was selected for several reasons. It is widely used, and many maps from other schema to Dublin Core exist to make it a portable standard for migrating data to other systems. While other metadata schema may have strengths in focused areas, Dublin Core's flexibility allows the data creator to use it for multiple purposes. If a schema that better meets the iRAT's needs is available in the future, Dublin Core could be mapped to it. Where possible, standard Dublin Core qualifiers were used; however, we found it necessary to create some custom qualifiers. Some fields require a specific,

controlled vocabulary. At this time, the fields that will require controlled vocabularies have been identified but in many cases the actual vocabulary is not yet fully devised. The iRAT team created partial vocabularies as examples (see 12). It is widely acknowledged that the creation of a controlled vocabulary, even a limited one, is a time-consuming task that cannot be approached lightly. Should the iRAT be fully developed, this stage would be a research project, hopefully funded by a grant.

The faceted browsing interface and other web-based user forms will use much of the metadata in the iRAT records. The metadata will be either independent or hierarchical. For example, a facet such as Material Type is independent of other facets and repeatable; a resource could have more than one material type. Other facets, such as "Subject, Recommended" are hierarchical, repeatable and have dependent facets, such as "recommended," or "audience level" associated with them. For instance, a particular resource such as Web of Science may be highly recommended in Mathematics, but not be as significant (although still appropriate in some cases) in Electrical Engineering. In this case, faceted metadata for subjects might appear as:

Science - Mathematics: recommended
Engineering - Electrical Engineering

Additionally, a database might be a good general resource in one subject area and a specialized resource in another. For example, Lexis-Nexis is a good general resource in many subject areas, but can be highly specialized for legal topics. The subject metadata might appear as:

Business and Management - Companies - News: Level 1, recommended
Business and Management - Law: Level 4, recommended

Where the levels can be interpreted as:

Level 1: Primarily general level content
Level 2: Primarily general with some specialized content
Level 3: Primarily specialized with some general content
Level 4: Primarily specialized content

To create iRAT metadata requires a new workflow. Instead of simply running iRAT titles through the cataloging process in the manner in which metadata is usually created, we would rely on a cooperative process to distill the collective expertise of subject specialists and catalogers alike. Subject specialists, using judgment and experience about the reference nature and depth of the resource described, would create the first draft of the metadata record using the iRAT vocabulary lists. Cataloging and Metadata Services personnel would provide quality assurance: first training the subject specialists in the use of standardized elements and vocabularies, then methodically checking the formatting and validity of the iRAT metadata records. Ideally, with time and experience, the subject specialist would achieve a high level of independence in metadata creation and require little editing.

In combining subject expertise of public services staff with decades of standards experience of the cataloging staff, we maximize important resources of the Libraries. Such a combination should result in a powerful tool.

Requirement	DC Element	Qualifier	Controlled or Free Text
Title	Title		Free text
Subject, Recommended	Subject	*Local*	Subject scheme+level
SFX tag	Subject	Other	SFX or blank

Coverage dates	Coverage	Temporal	Formatted dates
Languages	Language		Use English form
Format of title described	Type		unique Type vocabulary list
Rights	Rights		Free text
Licensed or Free	Rights		Licensed or free
Material type	Description		Type vocabulary list
Audience	Description	Audience	Vocabulary list
URL	Identifier	URL	URL

Italics = not registered in Dublin Core

Example: Content level for Subject, Recommended

Level 1: Primarily general level content
Level 2: Primarily general with some specialized content
Level 3: Primarily specialized with some general content
Level 4: Primarily specialized content

Example: Vocabulary for Type field

website
database, online
database, CD-ROM
print
database, DVD-ROM
librarian
department
database, web

Fig. 12. Partial list of potential metadata

Requirement	DC Element	Qualifier	Controlled or Free Text
Title	Title		Web of Science
Subject, Recommended	Subject	Other	Engineering: level 3
			Science - Biology: level 4, recommended
			Science - Mathematics: level 4, recommended
			Science - Chemistry: level 3
			Science - Physics: level 3
			Science - Materials Science: level 3
			Humanities - General: level 4
			Humanities - Foreign Literature: level 4
			Humanities - Linguistics: level 3, recommended
			Humanities - Linguistics: level 2
			Humanities - History: level 4, recommended
			Humanities - History: level 3
			Social Sciences - History: level 4, recommended
			Social Sciences - History: level 3
SFX tag	Subject	Other	SFX

Requirement	DC Element	Qualifier	Controlled or Free Text
Coverage dates	Coverage	Temporal	1983-present
Languages	Language		English
Format of title described	Type		Database, online
Rights	Rights		
Licensed or Free	Rights		Licensed
Material type	Description		Journal articles
Audience	Description	Audience	All MIT

Fig. 13. Descriptive metadata for a database: Web of Science

Requirement	DC Element	Qualifier	Controlled or Free Text
Title	Title		Jane's all the world's aircraft
Subject, Recommended	Subject	Other	Engineering - Aeronautics and Astronautics: recommended, level 4
SFX tag	Subject	Other	
Coverage dates	Coverage	Temporal	1990-present
Languages	Language		English
Format of title described	Type		Print
Rights	Rights		
Licensed or Free	Rights		
Audience	Description	Audience	Library patrons
Material type	Description		Text

Fig. 14. Descriptive metadata for a print resource: Jane's All the World's Aircraft (print version)

Requirement	DC Element	Qualifier	Controlled or Free Text
Title	Title		Sarah G. Wenzel
Alternative title	Title	Alternative	Reference Coordinator, Humanities Library
Subject, Recommended	Subject	Other	Humanities - General: level 1
			Humanities - General: level 2
			Humanities - General: level 3
			Humanities - General: level 4
			Humanities - Foreign Literature: level 1
			Humanities - Foreign Literature: level 2
			Humanities - Foreign Literature: level 3
			Humanities - Foreign Literature: level 4
			Social Sciences – Library Science: level 1
			Social Sciences – Library Science: level 2
			Social Sciences – Library Science: level 3
			Social Sciences – Library Science: level 4
Languages	Language		English
			French
Format of title described	Type		Librarian
Rights	Rights		
Licensed or Free	Rights		Free advice!
Contents notes	Description	Tableofcontents	M.S. Library & Information Science, University of

			Illinois at Urbana-Champaign
			M.A. French Studies, University of Illinois at Urbana-Champaign
Summary note	Description	Abstract	Databases sponsored: Humanities Abstracts, ISTA, Library Literature, ARTFL
			Interfaces sponsored: WilsonWeb, ProQuest, FirstSearch
			NERD member
			Departmental contact: SHASS, Foreign Languages & Literatures
			Web page responsibility: Virtual Ref. Collection, Table of Contents Services, Citing Sources
Audience	Description	Audience	Everyone

Fig. 15. Descriptive metadata for a librarian: Sarah

Potential systems

At the time of publication of this article, no vendor offers the magnitude of features the MIT Libraries wish to provide in one package as an information research advisory tool. There are many cross-database searching products that provide fairly rudimentary research advisory features, such as broad subject access. There are also many commercial and a few open source avenues for implementing faceted browse features as a database front-end. However, as of yet, no single package combines the two.

We investigated options for implementing the Information Research Advisory Tool (iRAT) in conjunction with a cross-database searching tool. We could acquire a cross-database searching tool:

- from a vendor who might be willing to develop advisory aspects according to the iRAT principles, or,
- along with a separate faceted browsing interface; create connections between them to form a seamless user interface.

The former was widely seen as the ideal for the MIT Libraries, with a willingness to pursue the latter if vendors were neither interested nor willing to improve on their resource discovery features. The latter option would entail obtaining funds through grants or other avenues in order to develop connecting mechanisms between a cross-database searching tool and a faceted browsing interface.

To communicate the desire for a research advisory tool to vendors, the task force wrote a requirements document outlining functionality and features needed. Accompanying the requirements was a series of web pages simulating the desired functionality and created metadata for a sample set of resources.

A small team approached several metasearch system vendors with an abridged version of the requirements document to narrow the field. Aside from standard questions about metasearch capabilities, initial interviews included a description of the iRAT and questions designed to assess a vendor's interest in creating the features.

Questions included:

- Can the user interface be customized? Can new questions be added to the search interface? To what degree? How difficult is it to customize? For example, is it possible to create a different interface to the metasearch portion (e.g., for faceted-browsing)?
- Can locally defined metadata or fields be created to further characterize resources?
- What standards or protocols does the vendor use to load content and data? (i.e., MARC records, XML format, batch uploads).
- Is it possible to export data from the system to other systems? What export protocols are supported?

After a series of phone interviews, the field narrowed to three potential vendors. One provides an open system, for which a faceted browse system could be customized. The other two vendors were already considering ways in which to incorporate faceted browsing into their interface; they voiced a willingness to discuss the development of research advisory features.

The MIT Libraries have since selected the Ex Libris Metasearch product and are determining the next steps in developing the iRAT. These include meeting with, or "lobbying", the vendor's other customers to assess interest in this direction. In addition, there is a desire to do a proof-of-concept by providing a series of mocked-up resource records to demonstrate how the faceted browse might work. This remains a work-in-progress; research advisory features are not yet developed. If the vendor does indeed develop the features, it is expected that it will be at least a year before they could become generally available.

One question for the proof-of-concept phase is whether the information domain is large enough to sustain effective faceted browsing. For example, is a set of 500 targets rich enough to provide interesting, adequate, and useful facets? Many faceted browsing systems that are in use today have far larger databases of items to draw from. The epicurious recipe database contains more than 15,000 recipes and the Boston restaurant guide more than 1000 restaurants.[14] Currently a simple broad subject search in Vera can yield from four to more than 60 hits. A set of resources that might be recommended by a librarian for any particular query is probably in the range of 1 to 5. Clearly, some facets can help a user narrow a list of 60 to a more manageable size. However, the proof-of-concept must discover the limits of good faceted browsing interface design in the context of information sources.

Usability testing will also provide an important mode of input. Lists of facets that are too long may cause users to not want to enter the system, but if facets are too broad and not interesting enough, it might be difficult to convince a user to use the interface.[15] Finding that balance will be an important challenge.

Creating and adding metadata to describe the information resources that populate an information research advisory tool is viewed as a substantial amount of work for the library community. Some library staff question whether the value of the tool is worth the time that must be expended to build it. Staff time to create metadata is another aspect that must be evaluated as part of a proof-of-concept development. Nonetheless, the MIT Libraries feel that it is essential to create systems that allow our users to become increasingly self-sufficient. As the number of questions answered at library reference desks declines, one could imagine using some resources previously devoted to answering in-person reference questions instead redirected to creating such systems.

Creating common, standard metadata for the interface facets would be an additional strategy to reduce staff effort. Interestingly, there are several ongoing efforts that may impact this area. An

---

[14] See 12.
[15] Ibid.

XML format for sharing hierarchical faceted metadata has emerged: exchangeable faceted metadata language (XFML).[16] Additionally, the NISO Metasearch Initiative (http://www.lib.ncsu.edu/niso-mi/index.php/Main_Page) is identifying standards in the area of cross-database searching; one aspect of that work includes Collection Description. Part of that task group includes development of a metadata element set that can be used across systems to allow users to "**find/discover** collections that meet a specified set of criteria". Standardization of metadata associated with resources would permit sharing parts of the metadata. Libraries could then concentrate their efforts on customizing aspects of the tool unique to their communities. An example of metadata that could be standardized and shared would be that of material types held in a database (e.g., journal articles, conference articles, patents, book reviews, book chapters, etc.). It is not likely that these types would frequently change in a particular resource. On the other hand, how a resource is characterized by audience or subject area might vary grately depending on the institution. For example, a university research library might characterize resources quite differently than a community college library.

Conclusion

In the future, the iRAT in conjunction with federated searching may evolve into a "smart system," able to recognize and retain the interests of a user (based on his or her name, personal profile, and search keywords). Such a system could automatically choose appropriate resources to be cross-searched and execute the search for the patron. The system would also explain the reasons for the resources selected, as well as indicate relevant resources that are not cross-searchable – including print material and human experts. After the patron has established a profile, searching would involve only one step.

While regretting our patrons' preference to avoid working directly with those who desire to assist them, we must create the best self-help tools possible. As both technology and research into searching behavior develop, we will find ways to harness the expertise of professional librarians as catalogers, subject specialists and reference staff in conjunction with those whose expertise lies in creative and innovative technologies. This powerful combination of knowledge and skills can be brought to bear on the issues raised by new information-seeking behaviors and needs. We can continue to play a role of advisor and educator through self-teaching interfaces. The role of librarians in an increasingly unmediated information-seeking society becomes that of providing tools that enable our patrons to learn searching skills and to work efficiently – in a sense to believe that they do not need us, while at the same time communicating our availability for direct assistance.

---

[16] Peter Van Dijck, "eXchangeable Faceted Metadata Language," http://www.xfml.org (accessed January 7, 2005).

Works Cited

Constantine, Larry L. and Lucy A.D. Lockwood. Instructive Interaction: Making Innovative
     Interfaces Self-Teaching. Preprint (January 2002).
     http://www.foruse.com/articles/instructive.htm
     http://www.foruse.com/articles/instructive.pdf (accessed January 7, 2005).

Database Discovery Tool Metasearch Analysis Group. Resource Discovery and Metasearch
     Requirements. MIT Libraries.
     http://macfadden.mit.edu:9500/dmag/MIT%20Libraries%20requirements%2007292004.p
     df.

Database Discovery Tool Requirements Group. Librarians' Thought Process Survey. MIT
     Libraries. http://macfadden.mit.edu:9500/dmag/form.html.

epicurious.com. Recipe Browse. http://www.epicurious.com/recipes/find/browse/ (accessed 7
     January 2005).

Institutional Research. Office of the Provost. MIT. *Senior Survey*. Consortium on Financing
     Higher Education. March 2004.

Institutional Research. Office of the Provost. MIT. *CSEQ Report*. C. Robert Pace and George D.
     Kuh. College Student Experiences Questionnaire Research Program. Indiana University
     Center for Postsecondary Research and Planning. 5th Ed. 2001.

Instone, Keith. "Faceted Browsing: How User Interfaces Represent and Benefit from a Faceted
     Classification System." Presented at SOASIS&T. July 8, 2004. http://user-
     experience.org/uefiles/facetedbrowse/.

Ma, Wei. "A Database Selection Expert System Based on Reference Librarians' Database
     Selection Strategy: A Usability and Empirical Evaluation," *Journal of the American
     Society for Information Science and Technology*, 53, 7 (2002): 567-580.MIT Libraries.
     Subject Guides. http://libraries.mit.edu/help/subjects.html.

MIT Libraries. Virtual Electronic Resource Access. http://libraries.mit.edu/vera.

Reference Vision Task Force. Research Help Survey. MIT Libraries.
     http://macfadden.mit.edu:9500/refcomm/RefVision/documents/survey.pdf.

Web Advisory Group. The 'Big Test' Usability Results. MIT Libraries. November, 2002.
     http://macfadden.mit.edu/webgroup/usability2002/big-test/summary.html.

Van Dijck, Peter. eXchangeable Faceted Metadata Language. http://www.xfml.org (accessed
     January 7, 2005).

Yee, Ka-Ping. et. al. "Faceted Metadata for Image Search and Browsing." In *Proceedings of the
     ACM Conference on Computer-Human Interaction*, 2003.
     http://bailando.sims.berkeley.edu/papers/flamenco-chi03.pdf (accessed January 7, 2005).