OpenAIR@RGU

The Open Access Institutional Repository
at The Robert Gordon University

This is an author produced version of a paper published in

This version may not include final proof corrections and does not include published layout or pagination.

Citation Details

Citation for the version of the work held in 'OpenAIR@RGU':

Citation for the publisher's version:

DYNIQX: A Novel Meta-Search Engine for the Web

Jianhan Zhu, Dawei Song, Marc Eisenstadt, Cristi Barladeanu, Stefan Rüger

Knowledge Media Institute, The Open University, United Kingdom

{j.zhu, d.song, m.eisenstadt, s.rueger} @ open.ac.uk; cristi.barladeanu@gmail.com

Abstract

The effect of metadata in collection fusion has not been sufficiently studied. In response to this, we present a novel meta-search engine called Dyniqx for metadata based search. Dyniqx integrates search results from search services of documents, images, and videos for generating a unified list of ranked search results. Dyniqx exploits the availability of metadata in search services such as PubMed, Google Scholar, Google Image Search, and Google Video Search etc for fusing search results from heterogeneous search engines. In addition, metadata from these search engines are used for generating dynamic query controls such as sliders and tick boxes etc which are used by users to filter search results. Our preliminary user evaluation shows that Dyniqx can help users complete information search tasks more efficiently and successfully than three well known search engines respectively. We also carried out one controlled user evaluation of the integration of six document/image/video based search engines (Google Scholar, PubMed, Intute, Google Image, Yahoo Image, and Google Video) in Dyniqx. We designed a questionnaire for evaluating different aspect of Dyniqx in assisting users complete search tasks. Each user used Dyniqx to perform a number of search tasks before completing the questionnaire. Our evaluation results confirm the effectiveness of the meta-search of Dyniqx in assisting user search tasks, and provide insights into better designs of the Dyniqx' interface.

1. Introduction and Motivation

Large search engines such as Google have achieved tremendous success in recent years, thanks to their effective use of the PageRank algorithm [5], smart indexing, and efficiency in searching terabytes of data [6]. Search engines like Google are now moving into the area of searching professional repositories as evidenced by Google Scholar (http://scholar.google.com) and Google Patent Search (http://www.google.com/patents) etc.

In the light of these large scale powerful search engines, how can traditional professional, academic and library repositories survive and keep their successes within their specific domain? Even given the success of the big search engines, in fact it is still very difficult for them to work effectively with repositories that belong to specific professional or proprietary domains. We think there are two main reasons for this.

First, due to legal/proprietary constraints, sometimes search engines cannot get hold of full content of information and may provide only the link to the place where the information can ultimately be found.

Second, big search engines work on the whole World Wide Web, consisting of many resources of a heterogeneous nature and domain context, and thus it is hard for search engines to perform as well as some domain or context specific search services (for example, in the context of arranging air travel between London and New York, the British Airways website will provide much better search services than Google).

We think that the key for successful domain specific specialized search services is to fully utilize the domain context and metadata which describes the domain context. For example, articles in the PubMed (http://www.ncbi.nlm.nih.gov/pubmed/) databases often have rich metadata information such as title, authors, citations, publication date, and publication journal names etc.

However, a limitation of current domain search services has been identified as the wide existence of *information islands* where the integration is difficult, resulting in a contextual "jump" for users when they are searching different repositories [7]. We think that it is important to give users a unified search interface to get access to multiple information repositories of different natures in terms of their specialized domains and information types including documents, images, and videos etc., so that they won't get frustrated in finding where to start with.

We treat the problem of building a meta search engine on top of a number of search engines as a collection fusion problem as defined by Voorhees et al. [3, 4]. The research questions we would like to answer are: How to generate a single ranked search result list based on a number of ranked lists from search engines? How to take into account relevance of each result to the query and the original rankings of the search results in the integrated ranked list? How to integrate metadata in ranking? How to integrate documents, images, and videos in a single ranked list?

After reviewing existing work, we found the necessity for a meta-search system that can seamlessly integrate multiple search engines of different natures. Therefore, we propose a novel dynamic query meta-search system called DYNIQX that integrates multiple evidences, namely, search results' relevance to the query, original rankings, and metadata, in collection fusion, and provides a unified search interface on top of multiple search engines. DYNIQX provides plug-in interfaces for new search engines. DYNIQX can help facilitate our investigation of current cross-search and metadata-based search services, identification of resources suitable for cross-search or metadata-based search, and comparison of single source search, cross-search, and metadata-based search. DYNIQX has also been released as open source under GPL as a Google code project at http://code.google.com/p/dyniqx/ .

In the remainder of this paper, we present our novel dynamic query interface system called Dyniqx in Section 2. We report the results of a task-based user evaluation in Section 3, and a questionnaire-based user evaluation in Section 4. Finally, we conclude in Section 5.

2. DYNIQX

Currently many domain specific search engines have adopted what we call a linear/top-down/hierarchical approach. For example, in the Intute search (http://www.intute.ac.uk), a popular search engine among students for finding high quality educational websites, a searcher may select from a list of subject areas and/or resource types for his/her search, and he/she is then taken to the result page. We think the rigidity of this linear/top-down/hierarchical approach may limit the user to search within the classification of the resources. Additionally, there are many forms of metadata which have not been fully exploited during the search process.

To overcome the rigidity of linear/top-down/hierarchical search, we propose to experiment with the dynamic query approach used to great effect by Shneiderman [1] in other contexts. Dynamic queries help users search and explore large amounts of information by presenting them with an overview of the datasets, and then allow them quickly to filter out unwanted information. "Users fly through information spaces by incrementally adjusting a query (with sliders, buttons, and other filters) while continuously viewing the changing results." A popular example of this approach is that of Kayak.co.uk, a meta-search engine which searches over 100 travel sites to find flights. Kayak uses a dynamic query interface that allow users to change many kinds of filters, such as tick boxes for airlines, and sliding bars for flight departing and arrival times etc., in order to find flights matching these filters. It is our conjecture that a dynamic query interface will dramatically outperform the linear/top-down/hierarchical approach.

In DYNIQX, search results from a number of search engines are fused into a single list by both the relevance of each result to the search query based on our indexing of top results returned from these search engine, and the rankings of the result provided by one or more search engines as below:

$$p_{fuse}(q \mid d) \propto (1-\lambda)p(q \mid d) + \frac{\lambda}{\log(Rank_{average}(d)+1)}$$

where q is the query, d is the content of a document or the context of an image or video object in documents, $p_{fuse}(q/d)$ is the fused conditional probability of document d used to rank it in the final list, $p(q/d)$ is the conditional probability of d based on our index, λ is a parameter adjusting the effect of the two components in the final probability, and $Rank_{average}(d)$ is the average ranking of document d given by search engines. In the equation we take the log of the average ranking in order to transform the linear distribution of the rankings of d for integrating with the document conditional probability.

DYNIQX provides a novel way of meta-searching a number of search engines in terms that high quality search results from a number of search engines are integrated, metadata from heterogeneous sources are unified for filtering and searching these high quality search results, high quality results based on a number of queries covering a topic are all integrated in DYNIQX, and features such as metadata-driven controls and term clouds are used for facilitating search.

The architecture of our DYNIQX system is shown in Figure 1. In Figure 1, first, a user sends a query to the DYNIQX system. The query is processed and translated into the appropriate form for each search service, e.g., PubMed. For each query, each search engine, e.g., Intute, PubMed, or Google Scholar, returns a ranked list of search results. Results from all these ranked lists are pooled and indexed by Lucene [8]. Unlike typical search engines where the user can only specify one query at a time, in DYNIQX, the user can specify a number of queries on different aspects of a search topic, e.g., "bird flu", "avian influenza", and "H5N1" etc. in order to find documents relevant to "bird flu". The search results for the number of queries are all pooled and indexed. The user can further refine the search results based on the pooled data. This is illustrated in the DYNIQX search interface shown in Figure 2.
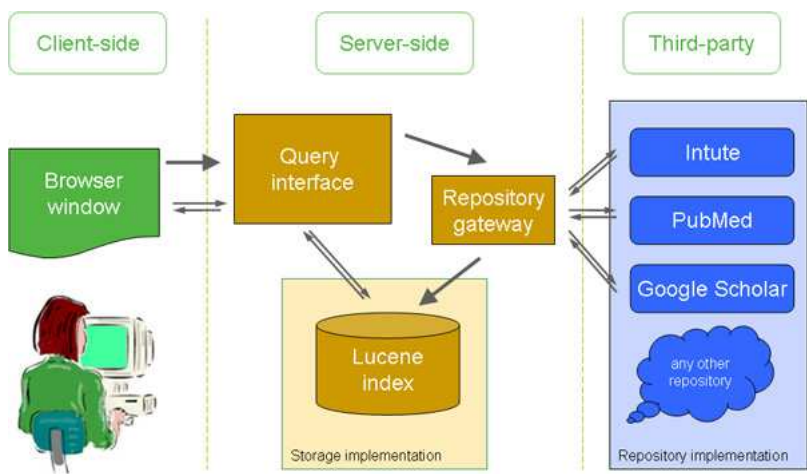

Figure 1: Architecture of DYNIQX

In Figure 2, the user can add a number of search queries to the pool. The user can reset pool to remove all search results cached. Statistics of search results from different search engines are shown in a table in the top right corner. The user can select search engines specialized in different types of information in the bottom right corner. Once search results are retrieved from these search engines, they are ranked by our collection fusion algorithm, and the user can view a single ranked list consisting of documents, images, and videos in the bottom left window. When more new results are obtained from these search services, the user can click a button to display an updated ranked list. Based on the significance of terms in these search results, a term cloud is displayed above the search results. The user can filter the research results by selecting some terms from the term cloud. On the right side pane, the user can further exclude some queries from the pool. Metadata has been effectively used in DYNIQX for cross-search since that metadata such as article title, author name, journal name, and publication date etc. can be used to rank search results.

Figure 2: DYNIQX Search Interface

3. Task-based User Evaluation

The aim of our user task-based evaluation is to measure the effectiveness, efficiency, and user satisfaction of DYNIQX. Effectiveness includes at least whether the task was completed successfully. Efficiency includes performance measures such as the elapsed time used for each search and number of viewed pages and mouse clicks etc. We also collect searcher background and satisfaction information.

We have carried out a controlled user evaluations of three search engines (Google Scholar, PubMed, Intute), and DYNIQX. In this comparative evaluation, users were given tasks. We used a Latin square design

to counterbalance order effects [9]. Based on the comparison, we qualitatively evaluated the usefulness of each search engine.

The four tasks where each consists of a group of related questions are designed as follows which reflect users' real world information needs.

A 'SARS' domain

Q1. Who sequenced the first SARS genome? (if many co-authors, then first two will be sufficient)

Q2. What was the exact publication (journal, date, title of paper)?

Q 3. How long was the genome sequence (typically this means the number of 'bases' or 'base pairs')?

B 'Bird Flu' domain

Q1. When was the first (or second... doesn't matter exactly) officially recorded outbreak of bird flu ('avian flu') in the UK?

Q2. What was the exact publication describing that outbreak [mentioned in 1] (journal, date, title of paper... may not be a scientific paper, but that's OK)?

Q3. What is the name, affiliation (institute) and email address of the lead researcher (don't spend more than a few minutes on this part)?

C 'Foot and Mouth' domain

Q1. When and where was the latest officially recorded outbreak of foot and mouth disease in the UK?

Q2. What was the exact publication describing that outbreak [mentioned in 1] (journal, date, title of paper... may not be a scientific paper, but that's OK)?

Q3. Will foot and mouth disease affect humans? Justify your answer with a journal reference.

D 'Breast Feeding' domain

Q1. What are the pros and cons of breast feeding vs bottle feeding for the baby and the mother (according to a peer-reviewed journal)?

Q2. What is the exact peer-reviewed journal article that has a satisfactory explanation of [1]?

Q3. Is there any connection between breast feeding and breast cancer? Justify your answer with a journal reference.

12 users participated in our evaluation according to the Latin square in Table 1.

Table 1. A Latin square for 12 searchers performing four tasks with four search engines

Searcher	Task Order			
	SARS	Bird Flu	Foot&Mouth	Breast Feeding
A1,A2,A3	Intute (I)	PubMed (P)	GS (G)	DYNIQX(D)
B1,B2,B3	PubMed (P)	Intute (I)	DYNIQX(D)	GS (G)
C1,C2,C3	GS (G)	DYNIQX(D)	Intute (I)	PubMed (P)

D1,D2,D3	DYNIQX(D)	GS (G)	PubMed (P)	Intute (I)

Average age of the 12 evaluators is 27. Among them, there are 6 males and 6 females, 6 PhD students, 4 research fellows, and two university staff representing a range of experience using search engines. While 9 of them are experienced search engines users, 10 of them used Google Scholar (GS) only occasionally.

The user followed the following steps in the evaluation:

Step 1: Entry questionnaire

Step 2: System and task familiarization of four search engines under supervision (10 minutes) and practice with a sample task: "find five researchers working on breast cancer treatment"

Step 3: Complete each task with a search engine, and fill out task questionnaire

Step 4: Complete exit questionnaire

All these questionnaires are online via SurveyMonkey (http://www.surveymonkey.com/).

3.1 Evaluation Results

We have used a tool called Slogger (http://www.kenschutte.com/slogger/) to automatically log searchers' activities during their entire evaluation process with their consent. The logged data help us to understand more about the searchers' behaviors during evaluation. Based on the logged user data, we can reconstruct each user's search history such as in Table 2.

Table 2. Example of reconstructed search history for user A1 on the "SARS" domain using PubMed, where the number in brackets shows the number of hits for the user's query on its left.

	'SARS' domain
A1	**PubMed:**
	who sequenced the first SARS genome (1)
	SARS genome (449)
	SARS genome sequence (280)
	The Genome sequence of the SARS-associated coronavirus. (208)
	"Marra MA"[Author]
	Got two answers where one is the right one and the other is by Chinese researchers
	The Genome sequence of the SARS-associated coronavirus
	sort by publication date
	p10
	Find right answers to Q1-4 in PubMed, fail on Q5

The average time spent by three searchers on each domain using each search engine is summarized in Table 3.

Table 3. Average time spent by three searchers on each domain using each search engine

Average time (mins)	Task Order			
	SARS	Bird Flu	Foot &Mouth	Breast Feeding
A1,A2,A3	10.3(I)	16(P)	23(G)	17(D)

B1,B2,B3	15(P)	11(I)	13(D)	20(G)
C1,C2,C3	10(G)	11(D)	15(I)	12(P)
D1,D2,D3	3.5(D)	16(G)	16(P)	12(I)

In Table 3, we did a *t-test* [10] based on the average time for each system, and Dyniqx is the most efficient system for the users to search for answers with statistical significance. Surprisingly, for three out of four domains, GS is the most inefficient. We think the reason might be that Dyniqx provides efficient ways for users to filter search results, and users spent lots of time reading large amount of search results returned by GS.

The average number of queries issued by three searchers to each search engine on each domain is summarized in Table 4.

Table 4. Average number of queries by three searchers to each search engine on each domain

Average num of page views	Task Order			
	SARS	Bird Flu	Foot &Mouth	Breast Feeding
A1,A2,A3	11.3(I)	4.67(P)	4.33(G)	3.33(D)
B1,B2,B3	3.33(P)	6.67(I)	4.67(D)	4.67(G)
C1,C2,C3	1.33(G)	3(D)	6.67(I)	3.33(P)
D1,D2,D3	4(D)	2.33(G)	4(P)	4.33(I)

In Table 4, for three out of four domains, users issued the least number of queries to GS than the other three search engines with statistical significance judged by *t-test*. This reflects that GS returns more content for each issued query than the other search engines, therefore, users tend to issue less number of queries.

When a user issues a new query or changes the filtering options of a query, e.g., rank the results by publication date etc., the user will get a new page view. We summarize the average number of page views by three searchers using each search engine on each domain in Table 5.

Table 5. Average number of page views by three searchers using each search engine on each domain

Average num of page views	Task Order			
	SARS	Bird Flu	Foot &Mouth	Breast Feeding
A1,A2,A3	21.3(I)	32(P)	22.33(G)	15.33(D)
B1,B2,B3	37(P)	24.67(I)	17.33(D)	25.67(G)
C1,C2,C3	7(G)	12.67(D)	16.67(I)	47.33(P)
D1,D2,D3	9.33(D)	16.67(G)	43(P)	33.33(I)

In Table 5, for all four domains, users viewed the most number of pages using PubMed among the four search engines with statistical significance judged by t-test.

Based on Table 3 and 5, we can calculate the average time spent by three searchers using each search engine on each page view as summarized in Table 6.

Table 6. Average time spent by three searchers using each search engine on each page view for each domain

Average time per page view	Task Order

	SARS	Bird Flu	Foot &Mouth	Breast Feeding
A1,A2,A3	0.4843(I)	0.5(P)	1.03(G)	1.1089(D)
B1,B2,B3	0.4054(P)	0.4459(I)	0.7501(D)	0.7791(G)
C1,C2,C3	1.4286(G)	0.8682(D)	0.8998(I)	0.2535(P)
D1,D2,D3	0.3751(D)	0.9598(G)	0.3721(P)	0.3600(I)

Users spent most amount of time per page view using Goolge scholar among all four search engines with statistical significance judged by *t-test*. This matches our observation that each page view returned by GS tends to have more contents than any of the other three search engines, therefore, the users had more to read using GS. However, users spent least amount of time per page view using PubMed among all four engines with statistical significance judged by t-test. This is due to the reason that users are generally having difficulty finding answers using PubMed, therefore, they tend to change the queries or filtering options more often and read less per page view. Our observation is that sufficient amount of time spent for each page view is an important indicator of the quality of search results, i.e., short amount of time spent reading search results indicates that the users are getting frustrated and tend to change the queries or filtering options more often.

Each user rated each search engine on each domain by choosing from very ineffective (-2), ineffective (-1), neutral (0), effective (1), or very effective (2). We average the ratings given to each engine on each domain by three searchers and summarize the results in Table 7.

Table 7. Average rating given by three searchers for each search engine on each domain

Average rating	Task Order			
	SARS	Bird Flu	Foot&Mouth	Breast Feeding
A1,A2,A3	-1.33(I)	0.67(P)	1.33(G)	1.67(D)
B1,B2,B3	-1.67(P)	-1.33(I)	1.33(D)	1.33(G)
C1,C2,C3	0.33(G)	1(D)	-1.33(I)	0.67(P)
D1,D2,D3	1.33(D)	0.33(G)	0.33(P)	-1.33(I)

Dyniqx is the best rated search engine by users, and GS is the second best rated search engines with statistical significance respectively.

We rate the quality of the answers given by each searcher to questions in each domain by choosing from very poor (-2), poor (-1), neutral (0), good (1), or very good (2). We average the quality ratings for three searchers' answers using each search engine on each domain and summarize the results in Table 8.

Table 8. Average quality ratings of search results given by three searchers for each search engine on each domain

Average answer quality rating	Task Order			
	SARS	Bird Flu	Foot &Mouth	Breast Feeding
A1,A2,A3	0.33(I)	0.67(P)	1.33(G)	1.33(D)
B1,B2,B3	-1.33(P)	-1.33(I)	1.33(D)	1(G)

C1,C2,C3	0.33(G)	0.67(D)	-1.33(I)	1(P)
D1,D2,D3	0.67(D)	-0.67(G)	0.67(P)	0.67(I)

Users gave the highest quality answers to questions using Dyniqx among all four engines with statistical significance. We think the reason is that Dyniqx successfully fuses search results from the three engines and the dynamic query interface is effective for filtering and searching.

Overall, based on the quality of the answers found, user ratings for each search engine, and time spent for finding answers, we judge Dyniqx as the most effective, and GS as the second best. Users can use Dyniqx to find better answers more efficiently than the other three search engines. The users also gave Dyniqx the best ratings overall. We think the best performance of Dyniqx is due to its effective use of metadata for filtering, term cloud, pooling of high quality results based on a number of queries, and collection fusion of a number of search engines. GS's good performance is due to its large coverage of information, ranking mechanism, and use of citation information.

3.2 Discussions

Searchers with different background tend to have different behavior in searching for information.

Some searchers seem to be more familiar with search and they are able to issue more complex search queries, such as using complex syntax in query formulation in PubMed.

English speakers can be more able to find answers than non-English speakers, and experienced users can more easily adjust to new search engines and find answers more effectively.

Many people tend to use Google more often than any other search engines and Google has an effect on them when they start using other search engines. For example, some searchers are used to natural language (NL) type of queries while using Goolge. However, other search engines such as Intute and PubMed cannot handle NL type of queries very well.

Due to its domain specific nature, Intute does not have as large a dataset as GS and PubMed. Therefore, users need to choose search query keywords carefully in searching, which create additional difficulty for novice users.

Users' familiarity of a particular domain can affect their search on the domain. For example, questions in the SARS domain tend to be more difficult for searchers with little medical knowledge. Therefore, the quality of answers shown in Table 8 for this domain is relatively lower than that for the other domains.

We also found that it is easy for searchers to find information relevant to a domain, but it can be very difficult for them to confirm whether the information is the most relevant to a question. For example, many users spent lots of time trying to find out whether a paper is the first report on SARS genome sequencing since there are a number of papers published around that time.

Searchers' habits also have effect in evaluation. Some people are more cautious in deciding the right answers than the others. For example, for the SARS domain, searcher A found a paper returned by GS as the

first result. He judged the paper as the first report on SARS genome sequencing based on the paper's high ranking and citation counts. However, searcher B spent lots of time investigating whether this paper is the first paper on the subject by comparing its publication date with many other papers. Clearly that searcher B ended up spending a lot more time than searcher A. In our evaluation, some searchers may even have found the right answer without knowing it.

On the other hand, this shows that search engines are typically much better at finding relevant information than providing proof of the authenticity of the information. Therefore, the process of finding proof can be time consuming.

Since PubMed is for people with medical background looking for academic publications, searchers have some difficulty in using it. Giustini and Barsky [2] have shown that PubMed has better coverage on scientific papers than GS, since GS tends to favor older publications which have attracted more citations. However, GS has features such as citations, abstract, keyword highlighting, and PageRank based ranking algorithms to outweigh the benefits of PubMed in our evaluation.

Most users tried to use metadata as soon as they are available, e.g., extracted author information in GS. Our observation is consistent with Kazai and Trotman's findings [11].

3.3 Comments

GS, PubMed, and Intute are built on different datasets. GS has the widest coverage of resources among the three by tapping into a large number of publication information sources on the web. PubMed searches proprietary medical publications. Intute is based on a database of 120365 manually constructed records of high quality descriptions of web resources written by subject specialists from a network of UK universities and partners. Therefore, the three search engines are affected by the scope of the information they search. In order to counterbalance the effect of the scope of information, we have designed the tasks by making sure that all three search engines have a good coverage of all four tasks.

4. Questionnaire-based User Evaluation

The aim of the questionnaire-based evaluation is to get the users' opinions about the interface of Dyniqx, and whether the interface design is effective for user search tasks.

7 people (P1 to P7) evaluated our DYNIQX interface which integrates six search engines, i.e., Google Scholar, Intute, PubMed, Google Image, Yahoo Image, and Google Video. The questionnaire and evaluation results are shown in the Appendix.

The user followed the following steps in the evaluation:

Step 1: Entry questionnaire

Step 2: System and task familiarization of Dyniqx under supervision (10 minutes)

Sample task: finding information on "England football"

Step 3: Choose two or more tasks from four suggested domains, and perform other tasks created by the users themselves

Step 4: Complete evaluation questionnaire

Average age of the seven evaluators is 31, 3 males and 4 females, 4 PhD students and 3 research fellows, 2 are native English speaker, and 5 are not, and all are experienced search engine users. They are generally satisfied with search engines. However, they all use meta-search engine infrequently, and if they do, their experiences with meta-search engines are less satisfying than ordinary search engines, showing that it is hard to produce good meta-search engines. They are all quite happy with Google Scholar, Google Image Search, but less happy with Google Video Search. They do not know much about PubMed, Intute, and Yahoo Image Search.

In addition to the four tasks created by us, the users carried out other search tasks of their own interest including search topics such as "Lymphoma", "alcohol and cancer", "liver cirrhosis", "sinus air pollution", "food poisoning", and "osteoporosis" etc.

In completing the evaluation questionnaire, the users mainly compared Dyniqx against Google scholar.

We divided the evaluation questionnaire into nine parts, namely, "search tasks", "overall search experience", "query component", "tag cloud", "result display", "query status table", "result filtering by queries", and "result filtering by sources", which correspond to different parts of the DYNIQX interface shown in Figure 2.

4.1 Evaluation Results on Search Tasks

Six users agree that the four tasks can reflect typical user information needs, and one is neutral. All users agree or strongly agree that Dyniqx can help them find answers to search tasks, showing that Dyniqx has been successful in meta-search. When comparing with Google Scholar, two strongly agree, two agree, two neutral, and only one disagree that Dyniqx is better in finding answers to search tasks. Since users were given the opportunity to try their self designed tasks, we have more confidence of their answers in the questionnaire.

4.2 Overall Search Experience

One user is very satisfied, four are satisfied, and two are neutral with the search experience using Dyniqx. When comparing with other search engines, two rated Dyniqx much better, three rated better, and two rated similar. One user commented that Dyniqx has more types of information than Google Scholar which searches mainly papers.

When comparing Dyniqx's user interface with the other search engines', one rated much better, one rated better, two rated similar, and three rated worse. Therefore, there is plenty of space for improvement of Dyniqx's user interface. In fact, it is quite a challenge to design an easy to use interface for a meta-search engine like Dyniqx which integrates search results from a number of heterogeneous search engines, has a variety of functionalities, and uses metadata for search and filtering.

One user said that "Dyniqx is very useful for exploration tasks because you can do a single search and then look at different aspects". As an experienced programmer himself, he also said that "Dyniqx's interface is definitely over complex – there are many different options and things like the search engines whose results are shown get adjusted automatically when other parameters are changed."

The other major comments by the users are based on their experience of using Google, such as that the query input is not clear, no need to press one button to get search results, and results needs paging etc. The users also commented that more space needs to be kept for results display.

When users are asked to compare Dyniqx's search result ranking with the other search engines', two rated Dyniqx better and five rated neutral. This is quite encouraging since Dyniqx uses less information than each underlying search engines for result ranking. Our algorithm for result ranking seems quite effective.

When asked whether it is easy to learn to use Dyniqx, two users agree, four neutral, and one disagree. Users agree that the tutorial we gave them is quite useful for starting using the system. We still need to improve the interface for Dyniqx for making it easier for users to learn to use the system.

4.3 Query Component

Three users strongly agree, one agrees, one keeps neutral, and two disagree that the search box is useful in search. One user commented that the search box should be put on top of the search component, which is what Google does. One user commented that the search box needs to be made more obvious, certainly the complex user interface of Dyniqx makes where to put search query difficult for the first time of a user.

One unique feature of Dyniqx is its ability to pool search results for a number of queries on a topic. The user can then work with the pool by filtering and searching high quality results in this pool. However, this novel idea can create some confusion for users. For example, they are not used to send multiple queries on a topic, and they need time to learn to filter and search the pool.

Once the user is done with a topic, i.e., move on to a new topic, the user needs to manually reset the pool. When asked about this new feature, two users strongly agree, four agree, and one keeps neutral that the reset pool button is useful in search. However, one user is not familiar with the button and almost forgot to press it every time.

A practical issue is that Dyniqx forwards search queries to a number of search engines, and these search engines may return search results with different speed and quantities. Users generally expect to get search results very quickly. Therefore, Dyniqx pools search results on the fly, and can update its result list dynamically. So while the user is interacting with Dyniqx, there may be always new results coming in. Currently, the user needs to press a "Get them" button in Dyniqx in order to get updated results, since we don't want the users to be presented with automatically updated results which may confuse them.

When asked about the "Get them" button, two users agree, three keep neutral, and two disagree that the "Get them" button is useful in search. Since users are more used to automatically displayed results like in Google, the users' feedback is not surprising. Two users commented that the "Get them" button is not helpful and she would like automatic display of results. We fully understand the users' expectations. However, as the practical issue we mentioned earlier, we may choose to implement this feature better, such as alert the user when new results are coming like in an email application or instant messenger.

When asked about whether it is easy to learn to issue multiple queries on a topic, six users agree and one disagrees. Our tutorial has been helpful to users in learning this feature. Multiple queries do have some advantages over single query based approach. The users' response matches our expectation. When asked to compare multiple queries against one query at a time, five agree and two keep neutral that issuing multiple queries per topic is more effective than one query at a time in search.

Overall, two users agree, two keep neutral, and three disagree that the issuing query component is well designed for search. One user commented that the component is too complex compared with Google, and needs time to get used to. Another user said that "the interface is 'nearly there', few small-ish changes would make a big difference to the user experience".

4.4 Tag Cloud

Two users strongly agree, and five agree that the tag cloud is useful in search. Three users strongly agree, three agree, and one keeps neutral that tag cloud is useful in result filtering. Two users strongly agree, four agree, and one keeps neutral that the tag cloud shows terms that are helpful to query reformulation. This shows a very attractive feature of Dyniqx that users may not know which terms to search on a topic, but when they come up with some initial terms, our tag cloud could suggest other good terms for them to search.

When asked about whether term highlighting in tag cloud is helpful to search, two users strongly agree and five agree.

Overall, two strongly agree, four agree, and one keeps neutral that the tag cloud is well designed for search. One user commented that tag cloud is one of the best designed parts of Dyniqx.

4.5 Result Display

Five users agree and two users disagree that results are displayed properly for finding information in Dyniqx. One user commented that results should be divided into pages and the number of results should be shown clearer. Another user commented that there should be more space reserved for result display, which we will work on.

As for the meta-search feature of Dyniqx, three users strongly agree and four agree that displaying results from different sources is helpful to finding information. This shows the advantage of meta search over single source search.

A novel feature of Dyniqx is that documents, images, and videos are ranked together in a single list. Users seem a bit divided on this new feature since many of them are more used to search specific type of information in a search engine. When asked, one strongly agree, two agree, three keep neutral, and one disagrees that showing images together with documents is helpful to search, three agree, three keep neutral, and one disagrees that showing videos together with documents is helpful to search, and three agree, three keep neutral, and one disagrees that showing images, videos together with documents is helpful to search. When comparing with document based search alone, one strongly agree, two agree, two keep neutral, and two disagree that showing images, videos together with documents is more helpful than showing documents alone in finding information.

But users can find the advantages of integrating different types of information in search as one user commented that "Depend on the tasks. It would be very helpful for some topics and intuitive users, e.g., finding out the cell structure, or some dieses (SARS, bird flu). But it doesn't help when the users are experienced and try to find deeper information."

Three users agree and four keep neutral that the ranking of results are accurate.

In terms of scrolling of search results, five agree, one keeps neutral, one disagree, and one strongly disagree that the results scrolling is well designed for search

Three strongly agree and four agree that display of metadata associated with search results is helpful to search. This shows the importance of metadata when they are available. The metadata include meta information from search engines such as author, publication date, and information sources, and also search engines names, the rankings by search engines etc.

Overall, three agree, one keeps neutral, and three disagree that the result display component is well designed for search. In fact, it is due to some details of the component that users are not happy about including that they would like the result display area to be larger, and scrolling more effective etc.

4.6 Query Status Table

Four users keep neutral and three disagree that the query status table is help to search, i.e., the query status table mainly shows the number of results returned from each search engine, and the users typically do not pay much attention to such statistics, and one user suggested more compact way show such info. However, five users agree that information in the table is easy to understand. Overall, the table is not very helpful to user search tasks, and most of them keep neutral about its usefulness.

4.7 Results Filtering by Queries

The filtering search results by queries function is designed for the multiple queries on a topic in Dyniqx. Three users strongly agree and four agree that the filtering search results by queries function is helpful to finding information. Overall, two strongly agree, four agree, and one keeps neutral that the filtering search results by queries component is well designed for search.

4.8 Results Filtering by Sources

Since Dyniqx carries out meta search based on a number of search engines, the filtering search results by search engines function enables the users to toggle/select information from these underlying search engines. Four users strongly agree and three agree that the filtering search results by search engines function is helpful to search.

When asked about which type of information is most helpful to search, four users rated documents as the most useful, three users rated images as the secondly useful, and three users rated videos as the thirdly useful. Two other users rated all three information types as equally useful, and they argued that the usefulness is entirely dependant on the context of the search. However, a user commented that "documents are the one I could not live without and videos are what I use least".

When asked about which search engine is the most helpful to search, two users rated Google scholar the most useful, two users rated Pubmed the most useful, and one user rated Intute the most useful. Overall, image and video based search engines are rated less useful than document based search engines. Three users commented that the usefulness "depends on the information needs", and it is hard to compare search engines in this way.

Overall, two users strongly agree, four agree, and one keeps neutral that the filtering search results by search engines component is well designed for search. The users suggested adding separate checkboxes for search engines for a certain media type

5. Conclusions and Future Work

In this paper, we propose a novel metadata based search engine called DYNIQX which fuses information from data collections of heterogeneous natures such as their specialized domains and information types including documents, images and videos. Metadata from multiple sources are integrated for generating dynamic controls in the forms of sliders and tick boxes etc for the users to further filter and rank search results. Since the effect of metadata in IR has not been sufficiently studied previously, our work provides insights into how to integrate metadata with mostly content based information retrieval systems.

Our task-based evaluation with 12 users has shown that DYNIQX can help users to complete real world information search tasks more effectively and efficiently with statistical significance than three well known search engines, namely, Google scholar, Intute, and PubMed, respectively. DYNIQX was compared with the three systems in terms of average time spent, average number of queries issued, average number of pages viewed, average time spent on each page view, average rating of effectiveness given, and average quality of search results given, by users, respectively.

We evaluated the meta-search interface of DYNIQX via a questionnaire with 7 users. The users completed the questionnaire after completing a range of real search tasks. The evaluation has covered DYNIQX's overall search experience, query component, tag cloud, result display, query status table, and query filtering, search engine filtering components. Overall, the users have rated DYNIQX more effective than current state-of-the-art search engines such as Google Scholar in helping complete search tasks. They agreed that DYNIQX has the advantage of effectively integrating information from different sources. Users also agreed that DYNIQX has effectively used meta-data for filtering and searching information. Users also raised concerns about aspects of the DYNIQX interface that need to be improved. We will address these in future work.

Our other future work includes the integration of other search engines in DYNIQX, evaluation of our approach on standard TREC datasets, and study of the effect of different ranking algorithms in collection fusion.

6. Acknowledgements

7. References

[1] Shneiderman, B. (1994) Dynamic Queries for Visual Information Seeking. IEEE Software 11(6): 70-77.

[2] Giustini, D., and Barsky, E. (2005) A Look at Google Scholar, PubMed, and Scirus: Comparisons and Recommendations. The Journal of the Canadian Health Libraries Association 26(3).

[3] Voorhees, E.M. et al. (1994) The Collection Fusion Problem. In Proc of Text REtrieval Conference (TREC).

[4] Voorhees, E.M. et al. (1995) Learning Collection Fusion Strategies. In Proc. of SIGIR: 172-179.

[5] Brin, S., and Page, L. (1998) The Anatomy of a Large-Scale Hypertextual Web Search Engine. Computer Networks 30(1-7): 107-117

[6] Ghemawat, S. et al. (2003) The Google File System. In Proc. of ACM Symp. on Operating Systems Principles.

[7] Awre, C. et al. (2005) The CREE Project: investigating user requirements for searching within institutional environments. D-Lib Magazine, October 2005, 11(10).

[8] Hatcher, E., and Gospodnetic, O. (2004) Lucene in Action. Manning Publications Co, ISBN: 1932394281.

[9] MacKenzie, S. (2002) Research Note: Within-subjects vs. Between-subjects Designs: Which to Use? Toronto, Ontario, Canada, http://www.yorku.ca/mack/RN-Counterbalancing.html

[10] Press, W.H. et al. (1992) Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press

[11] Kazai, G., and Trotman, A. (2007) Users' perspectives on the Usefulness of Structure for XML Information Retrieval. In Proc. of ICTIR.

8. Appendix

Entry Questionnaire:
Complete the following questionnaire based on your previous search experience. (P1 to P7 are seven users)

Part 1: Background
1. Name: P1, P2, P3, P4, P5, P6, P7
2. Sex: F, F, M, M, F, M, F
3. Age: 30, 26, 35, 27, 26, 34, 40
4. Education: PhD, PhD, PhD, MSc, MSc, Postgrad, PhD
5. Area of study: CBIR, Business/Finance, Computer Science, CS, CBIR, Maths/Software development, Knowledge management
6. English speaker (P6 P7)   Non-English speaker   (P1,P2, P3, P4 P5)
7. How often do you use search engines?
   Very often (P1, P2, P3 P4 P6 P7)
   Often (P5)
   Neutral ( )
   Less often ( )
   Seldom ( )
   Never ( )

8. You would rate your experience using these search engines?
Very satisfying (P3)
Satisfying (P2 P4 P5 P6 P7)
Borderline (P1)
Not satisfying ( )
Very much not satisfying ( )

9. How often do you use meta-search engine?
Very often ( )

Often ( )
Neutral ( )
Less often (P2 P3 P7)
Seldom (P4 P6)
Never (P1 P5)

10. You would rate your experience using meta-search engines?
Very satisfying ( )
Satisfying (P4)
Borderline (P2 P3 P6 P7)
Not satisfying ( )
Very much not satisfying ( )

11. How often do you use Google scholar?
Very often ( )
Often (P1, P1 P5)
Neutral (P3)
Less often ( )
Seldom (P4 P6 P7)
Never ( )

12. You would rate your experience using Google scholar?
Very satisfying (P5)
Satisfying (P1, P1 P3)
Borderline (P4 P6 P7)
Not satisfying ( )
Very much not satisfying ( )

13. How often do you use Google Image Search?
Very often ( )
Often (P2 P7)
Neutral (P1 P3 P6)
Less often ( )
Seldom (P4 P5)
Never ( )

14. You would rate your experience using Google Image Search?
Very satisfying ( )
Satisfying (P2 P3 P6 P7)
Borderline (P1 P4)
Not satisfying (P5)
Very much not satisfying ( )

15. How often do you use Google Video Search?
Very often ( )
Often ( )
Neutral ( )
Less often (P3)
Seldom (P5)
Never (P1, P2 P4 P6 P7)

16. You would rate your experience using Google Video Search?
Very satisfying ( )
Satisfying ( )
Borderline (P2 P3)
Not satisfying (P5)
Very much not satisfying ( )

17. How often do you use Pubmed?
Very often ( )
Often ( )
Neutral ( )
Less often ( )
Seldom (P2)
Never (P1 P3 P4 P5 P6 P7)

18. You would rate your experience using Pubmed?
Very satisfying ( )
Satisfying ( )
Borderline (P2 P3)
Not satisfying ( )
Very much not satisfying ( )

19. How often do you use Yahoo Image Search?
Very often ( )
Often ( )
Neutral ( )
Less often ( )
Seldom (P1 P5)
Never (P2 P3 P4 P6 P7)

20. You would rate your experience using Yahoo Image Search?
Very satisfying ( )
Satisfying ( )
Borderline (P1 P3)
Not satisfying (P5)
Very much not satisfying ( )

21. How often do you use Intute?
Very often ( )
Often ( )
Neutral ( )
Less often ( )
Seldom ( )
Never (P1, P2 P3 P4 P5 P6 P7)

22. You would rate your experience using Intute?
Very satisfying ( )
Satisfying ( )
Borderline (P3)
Not satisfying ( )
Very much not satisfying ( )

Other comments about your background relevant to search
P5: Should combine text based multimedia search with content based search
P7: I do research on semantic search systems

Evaluation Questionnaire:
After carrying out the search tasks, please complete the following questionnaire.

Part 1: Search tasks
1. The four search tasks reflect typical user information needs?
Strongly agree ( )
Agree (P1, P2 P3 P4 P6 P7)
Neutral (P5)
Disagree ( )
Strongly disagree ( )
Don't' know ( )
Comments:
P1: I agree the question on my interested tasks including my own tasks.

2. Dyniqx can help you find answers to the four search tasks?
Strongly agree (P3)
Agree (P1, P2 P4 P5 P6 P7)
Neutral ( )
Disagree ( )
Strongly disagree ( )
Don't' know ( )
Comments:

3. Dyniqx is better than other search engines in helping you find answers to the four search tasks?
Strongly agree (P1 P3)
Agree (P5 P6)
Neutral (P2 P4)
Disagree (P7)
Strongly disagree ( )
Don't' know ( )
Comments:
P1: Yes, comparing Dy and GS, the information from Dy is multiple, that is, not only papers.

P2: Google Scholar focuses on the search for research papers and literatures. The output of Intute can be different types including web pages and news. It is hard to compare the relevance level of them.

P6: if (eg) I was specifically looking for a video then I would go directly to (eg) Google Video/YouTube search, rather than a meta-engine.
Comments to these search tasks:
P1: Complicated. Don't know the area very well. Better to use my own tasks.

P4: Ranking of results sometimes was not adequate: irrelevant answers ranked higher than the correct ones.

P7: I liked the tasks – they made me explore a range of functions in the system and illustrated its possibilities quite well.

Part 2: Dyniqx overall search experience
1. You would rate your experience using Dyniqx?
Very satisfying (P1)
Satisfying (P3 P4 P5 P7)
Borderline (P2 P6)
Not satisfying ( )
Very much not satisfying ( )

2. In completing search tasks, how would you compare Dyniqx with other search engines you used?
Much better (P1 P3)
Better (P5 P6 P7)
Similar (P2 P4)
Worse ( )
Much worse ( )

P1: Comparing Dy and GS, the information from Dy is multiple, that is, not only papers

3. How would you compare Dyniqx's user interface with other search engines you used?
Much better (P1)
Better (P3)
Similar (P2 P4)
Worse (P5 P6 P7)
Much worse ( )

P1: But some places still need to improve. Eg. Show result without click the button; divide results into pages; submit button is rarely used.
P2: Too complex.

4. How would you compare Dyniqx's search result ranking with other search engines' result ranking?
Much better ( )
Better (P3 P6)
Similar (P1, P2 P4 P5 P7)
Worse ( )
Much worse ( )

5. It is easy to learn to use Dyniqx?
Strongly agree ( )
Agree (P1 P4)
Neutral (P2 P3 P5 P6)
Disagree (P7)
Strongly disagree ( )

Suggestions to improve/change **Dyniqx search engine**:
P1: Query input is not clear; add Google Web search to Dy; add copy and paste function on the interface—help query reformulation.

P2: categorize the six engines into three groups, which can make the filtering easier. Confused with the results of putting images, videos and documents together.

P4: 1. Necessary to clear the result pool is sometimes confusing.
2. It would be nice if the sources to search could be selected in advance.

3. Quick checkboxes/radio buttons to filter the media type in advance might be helpful.

P5: Leave more space for the results display; better logo; automatically display results;

P6: I think the user interface needs quite a lot more work to get right, I didn't like the scrolling box at the bottom with the results in - meaning I had to expand the size of my browser window – what would happen if I only had a small screen in the first place?
    It's non-obvious that you can click on some of the text – eg to filter, or to exclude/include search terms from the results.
    The selection of search resources kept resetting to include all when I clicked to filter – which was a little frustrating.
    The repository listing could be moved down the page, as this is not as important as the selection of repositories – which get a little hidden because of the repository listing images.

P7: Dyniqx is very useful for exploration tasks because you can do a single search and then look at different aspects. For searching a single medium the individual search engines seem to me to give a better result. For example finding a sensible video on breast feeding seemed easier in Google video– however this may have been due to ranking that put papers with the search term in the title at the top.
    The interface is definitely over complex – there are many different options and things like the search engines whose results are shown get adjusted automatically when other parameters are changed. Also a large amount of screen real estate is given over to the search functions leaving less room for the search results – you have to put it on full screen mode to see even one whole image.

Part 3: Issuing query component

1. The search box is useful in search?
Strongly agree (P2 P4 P6)
Agree (P5)
Neutral (P3)
Disagree (P1 P7)
Strongly disagree ( )
Comments:
P2: set the search box on top of the search component.
P5: not quite clear
P6: though it wasn't totally obvious to me where is was in dyniqx – I think it needs a label or to look more like a standard html text input box.

2. The Reset Pool button is useful in search?
Strongly agree (P1 P5)
Agree (P2 P3 P6 P7)
Neutral (P4)
Disagree ( )
Strongly disagree ( )
Comments:
P2: not familiar with the button and almost forget to reset every time.
P6: But maybe needs making more obvious how to exclude certain search queries from the current pool - the interface allows you to do it, but it's non-obvious.

3. The "40 new results, Get Them!" button is useful in search?
Strongly agree ( )
Agree (P3 P4)

Neutral (P1, P2 P7)
Disagree (P5 P6)
Strongly disagree ( )
Comments:
P2: the display of the number of results is not quite useful and the automotive display of results without clicking this button is better.
P6: I wasn't; really sure why this was needed – When I do a search I would expect the results to appear without me having to make another click to retrieve the results.

4. It is easy to learn issuing multiple queries on a search topic?
Strongly agree ( )
Agree (P1, P2 P3 P4 P5 P6)
Neutral ( )
Disagree (P7)
Strongly disagree ( )
Comments:

5. Issuing multiple queries is more effective than one query in search?
Strongly agree ( )
Agree (P1 P3 P5 P6 P7)
Neutral (P2 P4)
Disagree ( )
Strongly disagree ( )
Comments:

6. Overall, the issuing queries component is well designed for search?
Strongly agree ( )
Agree (P2 P3)
Neutral (P6 P7)
Disagree (P1 P4 P5)
Strongly disagree ( )
Comments:
P2: too complex, and need some time to get used to.
P6: I think the interface for this is "nearly there" few small-ish changes would make a big difference to the user experience.

Suggestions to improve/change the **Issuing queries component**:
P1: Hope the result can display automatically after enter the query. Don't need the submit button. Query enter place is not clear----like a button but not an input window.

P5: I don't quite understand why we need to click the button so that can get the result; and the query box is not quite clear for a fresh man

P7: A small but important thing is that the query input box is just a grey square that looks like a button – without a trainer present I would not have known where to enter my search. Putting the text "search" above it would help.
    I learnt how to use the "get results button. It wasn't intuitive at first but once I got used to the pool system I realized why it was there and began to like the fact that I did not have to add the results of unsuccessful searches. The only problem is that it can be hard to judge if a search is right just from the number of hits
    I wasn't always clear how Dyniqx combined multiple searches – I expect Boolean AND but think I got Boolean OR.

I think I could learn to use the query input but that less experience searchers have problems with it.

Part 4: Tag cloud

1. The tag cloud is useful in search?
Strongly agree (P1 P4)
Agree (P2 P3 P5 P6 P7)
Neutral ( )
Disagree ( )
Strongly disagree ( )
Comments:

2. The tag cloud is useful in result filtering?
Strongly agree (P1 P4 P5)
Agree (P2 P3 P7)
Neutral (P6)
Disagree ( )
Strongly disagree ( )
Comments:
P6: I was slightly confused why you would need a pool of queries (which you can include/exclude) plus a filter – they seem to me to offer quite similar functions, and so may be a little confusing that there are 2 functions that appear to do the same thing.

3. The tag cloud shows terms helpful to result filtering?
Strongly agree (P1 P4 P5)
Agree (P2 P3 P6 P7)
Neutral ( )
Disagree ( )
Strongly disagree ( )
Comments:

4. The tag cloud shows terms helpful to query reformulation?
Strongly agree (P1 P5)
Agree (P3 P4 P6 P7)
Neutral (P2)
Disagree ( )
Strongly disagree ( )
Comments:

5. Term highlighting in the tag cloud is helpful to search?
Strongly agree (P1 P5)
Agree (P2 P3 P4 P6 P7)
Neutral ( )
Disagree ( )
Strongly disagree ( )
Comments:

5. Overall, the tag cloud component is well designed for search?
Strongly agree (P1 P7)
Agree (P2 P3 P4 P6)
Neutral (P5)

Disagree ( )
Strongly disagree ( )
Comments:

Suggestions to improve/change the **Tag cloud component**:
P2: every time when a tag is added to the filtering, the search engines have to be reselected.

P5: The technique is very good, but it will be good if the display is smaller, so that the main results area can be bigger

P6: Make it slightly more obvious how the cloud relates to the query pool.

P7: I think this was one of the best designed parts of the system. The only problem is that it takes up quite a lot of space – I might make it smaller and more selective.

Part 5: Result display

1. The results are displayed properly for finding information?
Strongly agree ( )
Agree (P2 P3 P4 P5 P6)
Neutral ( )
Disagree (P1 P7)
Strongly disagree ( )
Comments:
P1: Better divide the results into pages and showing the number of results clearer.

P6: For the individual search result they are displayed well, but I didn't like the fact the result box was so small – it needs to be the main component on the screen, but it seems relegated in comparison to the others functions. I would much prefer a fully scrolling browser window – rather than separate components that scroll individually – I have experienced problems because of interfaces like these when trying to use on small screen resolutions

2. Displaying results from different sources is helpful to finding relevant information?
Strongly agree (P1 P3 P6)
Agree (P2 P4 P5 P7)
Neutral ( )
Disagree ( )
Strongly disagree ( )
Comments:

3. Showing Images together with documents is helpful to finding relevant information?
Strongly agree (P3)
Agree (P5 P6)
Neutral (P1 P4 P7)
Disagree (P2)
Strongly disagree ( )
Comments:
P2: images and documents are irrelevant in nature, and had better be categorized.

4. Showing Videos together with documents is helpful to finding relevant information?
Strongly agree ( )

Agree (P3 P5 P6)
Neutral (P1 P4 P7)
Disagree (P1)
Strongly disagree ( )
Comments:
P2: similar to above.

5. Showing Images, Videos together with documents is helpful to finding relevant information?
Strongly agree ( )
Agree (P3 P5 P6)
Neutral (P1 P4 P7)
Disagree (P2)
Strongly disagree ( )
Comments:
P1: Depend on the tasks. It would be very helpful for some topics and intuitive users, eg, finding out the cell structure, or some dieses (SARS, bird flu). But it doesn't help when the users are experienced and try to find the deeper information.
P2: similar to above.

6. Showing Images, Videos together with documents is more helpful than showing documents alone in finding relevant information?
Strongly agree (P3)
Agree (P5 P7)
Neutral (P1 P6)
Disagree (P2 P4)
Strongly disagree ( )
Comments:
P6: Though as I said earlier, when I do a search I know what type of resource I'm looking for so may just go direct to a resource specific (eg image) search

7. The rankings of results are accurate?
Strongly agree ( )
Agree (P1 P3 P6)
Neutral (P2 P4 P5 P7)
Disagree ( )
Strongly disagree ( )
Comments:

8. The scrolling of search results is well designed for search?
Strongly agree ( )
Agree (P1, P2 P4 P5)
Neutral (P3)
Disagree (P7)
Strongly disagree (P6)
Comments:
P6: in my mind this must be changed so the browser window scrolls – not just the results window,

10. Display of metadata associated with search results is helpful to search?
Strongly agree (P1 P4 P7)
Agree (P2 P3 P5 P6)
Neutral ( )

Disagree ( )
Strongly disagree ( )
Comments:

11. Use of multiple queries as opposed to single query is helpful to search?
Strongly agree (P1)
Agree (P3 P5 P6 P7)
Neutral (P2 P4)
Disagree ( )
Strongly disagree ( )
Comments:

12. Overall, the result display component is well designed for search?
Strongly agree ( )
Agree (P2 P3 P4)
Neutral (P5)
Disagree (P1 P6 P7)
Strongly disagree ( )
Comments:
P6: mainly because of the scrolling problem.

Suggestions to improve/change the **Result display component**:
P5: Should be displayed in a bigger area

P6: I found the dynamic updating of the results display a little odd – I was often unsure if it was currently processing or whether it had completed. Something to tell me when the results were loading and then fully loaded would certainly help – plus refresh the screen when there are no results. Eg a I found it was still displaying: "Displaying 45 search results…" followed by no results when I've clicked on too many filters – in this case it should show me "There are no results for the search ABC filtered on X, Y & Z"

P7: The key problem of the results display is that it is too small – you have to scroll through results – it would be better to see 10 or so summaries at once to get an idea of the span of what is found

Part 6: Current active query status table

1. The active queries status table shows info helpful to your search tasks?
Strongly agree ( )
Agree ( )
Neutral (P3 P4 P5 P7)
Disagree (P1, P2 P6)
Strongly disagree ( )
Comments:
P2: it is not quite helpful to be informed of the number of results from each search engines.
P6: I didn't feel this really offered me anything – much better to me would be to put the number of results returned for each rerpository just as text next to the repository selection checkboxes, eg:
Pubmed (45)   Intute (100+)   GoogleScholar (56)

2. The active queries status table shows info easy to interpret?
Strongly agree ( )
Agree (P1, P2 P3 P4 P5)

Neutral (P6)
Disagree ( )
Strongly disagree (P7)
Comments:

3. Overall, the active queries status table is well designed for search?
Strongly agree ( )
Agree (P2 P5)
Neutral (P1 P3 P4 P6 P7)
Disagree ( )
Strongly disagree ( )
Comments:
P6: don't really feel it gave me any information I actually needed for the search tasks

Suggestions to improve/change the **Current active queries status table**:
P1: It feels more useful to the designers; users don't care too much where is from.

P7: I am not sure what is the difference between hits and results.

Part 7: Filtering search results by queries function

1. The excluding queries from search function is helpful to finding information?
Strongly agree (P5 P6 P7)
Agree (P1, P2 P3 P4)
Neutral ( )
Disagree ( )
Strongly disagree ( )
Comments:

2. Overall, the excluding queries from search function is well designed for search?
Strongly agree (P5 P7)
Agree (P1, P2 P3 P4)
Neutral (P6)
Disagree ( )
Strongly disagree ( )
Comments:
P1: But better add copy and past function for reformulating query use

Suggestions to improve/change the **Excluding queries from search function component**:
P5: Difficult for fresh man to discover
P6: It needs to be more obvious that you need to click to include/exclude – maybe use checkboxes like the repository selection -

Part 8: Filtering search results by search engines function

1. Filtering search results by their original search engine is helpful to search?
Strongly agree (P1 P4 P5 P6)
Agree (P2 P3 P7)
Neutral ( )
Disagree ( )
Strongly disagree ( )

Comments:

2. Sort the usefulness of types of sources in descending order?
Documents [P1:1] [P2: 1] [P3:1] [P4:1] [P5:1] [P7:1]
Images [P1:2] [P2: 3] [P3:2] [P4:1] [P5:1] [P7:2]
Videos [P1:3] [P2: 4] [P3:3] [P4:1] [P5:1] [P7:3]
Comments:
P5: depends on the information needs
P6: not sure I really have answer for this – would depend entirely on the context of the search.
P7: Well it depends what you are looking for – but documents is the one I could not live without and videos is what I use least.

3. Sort the usefulness of search engines of results' origin in descending order?
Pubmed [P1:3] [P2:4] [P3:1] [P4:1] [P7:3]
Intute [P1:1] [P2:2] [P3:2] [P4:6] [P7:2]
Google scholar [P1:2] [P2:1] [P3:3] [P4:2] [P7:1]
Google Image [P1:5] [P2:3] [P3:4] [P4:3] [P7:4]
Yahoo Image [P1:4] [P2:6] [P3:5] [P4:4] [P7:6]
Google video [P1:6] [P2:5] [P3:6] [P4:5] [P7:5]
Comments:
P5: depends on the information needs
P6: Again, not sure I can rate them in this way.
P7: I really can't judge this – it would be user dependent – someone else would have a very different view.

4. Overall, the display following engines function is well designed for search?
Strongly agree (P4 P5)
Agree (P1 P3 P6 P7)
Neutral (P2)
Disagree ( )
Strongly disagree ( )
Comments:

Suggestions to improve/change the **Displaying following engines function component**:
P4: Maybe a separate checkbox for all sources for a certain media type might be useful
P6: Only few things that need tidying up – eg the checkboxes seemed to reset themselves occasionally when I clicked to filter, which I found very confusing and annoying!
P7: Just 3 buttons for document, image or video would have done for me a lot of the time