

An Overview of Aggregating Vertical Results into Web Search Results

Suhel Mustajab
Department of Computer
Science, A.M.U., Aligarh,
U.P., India.

Mohd. Kashif Adhami
Department of Computer
Science, A.M.U., Aligarh,
U.P., India.

Rashid Ali
College of Computers and
Information Technology, Taif
University, Taif, Saudi Arabia

ABSTRACT

Vertical aggregation is the task of integrating results from specialized search services or verticals into the web search results. Aggregating verticals into the core web results helps in achieving diversity in information search. In this paper various efforts made for selecting relevant vertical and presenting the aggregated results to the users are reviewed. Various vertical selection approaches and design and evaluation of aggregated search interfaces have been discussed which has been a less focused area as compared to the most prior research work in conventional web search interfaces.

General Terms

Aggregation, rule-based triggers, Sampling, Query Classification, machine learned framework.

Keywords

Verticals, resource selection, aggregated search, vertical selection, web-page ranking.

1. INTRODUCTION

In traditional web search engines, whenever user issues a query, a ranked list of relevant URLs is retrieved and displayed to the users. However modern web search engines include contents from specialized sub-collections, called *verticals*, which focus on specific

domains (e.g., images, news, blogs and travel).

Thus general web search engines maintains or have access to vertical search engines and they present the search result page to the user after integrating, if required, relevant verticals. This process of selecting and presenting relevant vertical results is known as *aggregated search* (figure 1). A user can access vertical content in two ways. A user can issue a query directly to a vertical search engine or if he don't know the relevant vertical or prefers to use web search engine or may want results from different verticals at once, he can use directly the portal search engine. Various vertical blocks are shown in figure 2. Integrating or merging heterogeneous results or verticals into a single result page is a challenging task. There is a big difference between the core web searching and aggregated searching. In the core ranking problem, objects of the same nature are compared, i.e. web pages, and ranked according to their estimated relevance to the query while in aggregated search type of ranking problem objects of the different nature has to be compared which cannot be represented using similar features, e.g. web pages to videos, and thus cannot be ranked using same method used to rank web documents. This paper presents an overview of aggregating vertical results into the web results.

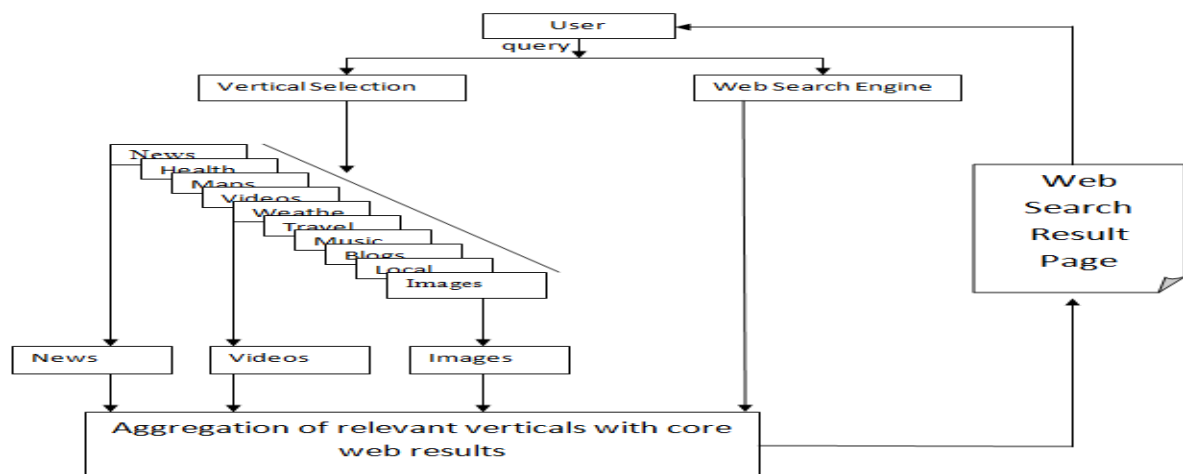




Fig. 1 A High level overview of aggregated search.


News for **kolkata fire**

 **Kolkata fire: Shopkeepers allege nexus, government refuses to take responsibility**
IBNLive - 1 hour ago
The recent deadly blaze at the Surya Sen market with 19 casualties is not the first fire tragedy in Kolkata. But no lessons seem to have been ...


19 people die in Kolkata blaze
Hindustan Times - 12 hours ago

Kolkata fire was no Jatugriha, it was far worse
Firstpost - by Sandip Roy - 15 hours ago

Video – **Volcano Lava** – National Geographic
 video.nationalgeographic.com/.../volcanoes/volcano-l...
Oct 18, 2011
Get immersed in a volcanic landscape with bubbling lava, spewing eruptions, and colliding rivers of fire.

BBC News - Ecuador's Tungurahua **volcano** erupts
 www.bbc.co.uk/news/world-latin-america-21738184
Mar 11, 2013
Ecuador's Tungurahua volcano continues to spew lava for a second day.

More videos for **volcano** >

Images for **volcano** - Report images


Volcano-EnchantedLearning.com
www.enchantedlearning.com/subjects/volcano/
A **volcano** is a place on the Earth's surface (or any other planet's or moon's surface)

Fig. 2 Vertical blocks: news, video and image.

2. BACKGROUND

Aggregated search is analogous to federated search, in which diverse information can be collected from various resources such as databases, web-based search engines, library catalogues and corporate data collections. A federated search system maintains several resources and when user issues a query to the federated search system, it passes on the query to the selected relevant resources, which returns the result to the federated search system. The results returned by the selected resources are integrated and merged into a single list and presented to the user. Some of the examples of federated search systems are *Westlaw*, providing federated search for legal professionals covering thousands of databases [6,44,45,46] and *Fedstats*, an online portal for statistical information. In Federated search cooperative environment a central section or the broker has the extensive information about the contents of each collection, collection sizes and lexicon statistics whereas in uncooperative environment the broker has to use query-based sampling [7] to know the lexicon statistics for each collection [6]. Shokouhi and Si [6] presented a comprehensive study on the federated search. Metasearch is another technique available for searching diverse information sources. They provide a single interface for searching the results from multiple search engines [30]. Metasearch engine architecture has been discussed in [31-40]. Metasearch engines have broader coverage of the web and better search scalability than centralized search engines [41]. Some of the existing metasearch engines are Metacrawler, Dogpile, and search.com. Earlier web search engines used to show only text results and users have to separately type queries to the specialized *verticals* for other types of information such as images, news, maps and videos. The highly heterogeneous nature of the web data,

diversity in information needs and search tasks and difficulty in keeping a fresh index of real-time data led to the introduction of comprehensive search engines like the Korean search engine (Naver), which started blending multimedia answers in their core web results [42]. Another topic closely related to the federated search was introduced by Lu [43] known as peer-to-peer network, consists of three types of objects, i.e., information providers, information consumers, and a search mechanism, which addresses problems similar to federated search.

3. VERTICAL SELECTION

3.1 Constructing Representative Vertical

Vertical aggregation is the task of assembling or merging results from specialized search engines or verticals into web search results [1,2,3,4]. For selecting the relevant vertical in response to a user query and aggregating it with the web result, an aggregated search system first needs proper representation of its verticals. An aggregated search system should know the comprehensive statistics of the documents in the vertical content such as term frequency, vertical size etc. Vertical representation in aggregated search can be compared to resource representation in federated search [5,6,8,26,19]. Shokouhi and Si [6] proposed various techniques for resource representation in federated search. Gravano et al. [5] proposed methods for manually generating the resource representation, but automatically generated representation are more common for dynamic resources where new documents are added in large number [6].

Vertical representation can be generated using the similar approach as used for federated search working with cooperative resources, where resources provide

various statistics (e.g. term frequency, size) about the documents to the federated search systems. Arguello et al. [2] also proposed techniques for vertical representation using sampling approaches. Two main sampling approaches were given by them. In the first one, documents were sampled directly from the vertical and in the second approach documents were sampled using external resources or surrogate corpora such as wikipedia.

Query-based sampling technique, proposed by Callan and Conell [7] was used for directly sampling from the vertical. A single term query is used to retrieve document from the resource, then new sampling query is selected. The evaluating resource description and new sampling queries are derived from retrieved documents. Another work from Shokouhi et al [9] shows that better performance was obtained when high-frequency queries were used for sampling the documents as compared to deriving sampled documents themselves. Arguello et al. [2] however used high-frequency queries, extracted from vertical query logs, for sampling documents.

3.2 Prior Approaches

3.2.1 Resource Selection in Federated Search

Prior work related to resource selection in federated search can be considered analogous to vertical selection in aggregated search. Some of the prior approaches such as UUM [61], CVV [10], CORI [8] and KL-divergence [11] consider sampled documents or collections as “large documents”. The document scoring techniques were used by these methods for scoring these sampled documents or collections. These techniques make no distinction between documents and do not model the number of relevant documents in a collection. CVV-based methods [10] are used for ranking and selecting index servers with respect to a query and merging the results returned by the index servers. Index servers are the keyword-based search systems which use index databases for mapping keywords to documents or we can say information resources. CORI [8] referred to as collection retrieval inference network uses the TFxIDF document ranking method as an analogy for collection ranking. Another method KL-divergence [11] is a clustering algorithm on a set of documents. These methods adapt document scoring techniques for scoring collections. These techniques make no distinction between documents and do not model the number of relevant documents in the collection [26]. For explicitly modeling the distribution of relevant documents in a collection, approaches such as ReDDE [19], GIOSS [15], Sampling and Selection algorithm [60] and gGIOSS [25] are used. The ReDDE algorithm estimates the database sizes as well as its contents for the uncooperative environments in federated search which is an extension of query-based sampling [7] method used for acquiring resource description. Si and Callan [19] showed experimentally that the ReDDE algorithm results in better document rankings than the CORI algorithm. The gGIOSS

represented the database by a vector containing document frequencies of different words in the databases. Ipeirotis and Gravano [60] uses document frequency information as a part of the database description to build the hierarchical structure for the databases and showed it to have better retrieval performance than CORI algorithm.

3.2.2 Query-classification into Topical Categories

Some prior work in query-classification into topical categories is also relevant to vertical selection. Query classification is an important task in information retrieval. The search engines can be more effective and can return more relevant and representative web pages to the users if the category information for a query is known. Query classification approaches can be used to augment the query with features other than query string. These other features are contingently obtained from query-logs or corpora of documents associated with the target classes. Shen et al. [27] and Li et al. [28] used corpus-based evidence for mapping documents to target categories and Bietzel et al. [13,14] derive evidence from query-logs. Another work from Shen et al. [12] derived a soft mapping from documents to target categories using term similarity.

3.3 Vertical Selection Approaches

For selecting relevant verticals, given a user query Arguello et al. [2] proposed three sources of evidences. These three were the query string, logs of queries previously issued to the vertical and corpora representation of vertical content.

Some prior work in vertical selection was also done by Li et al. [20], focusing on *shopping* and *job* verticals. They focused on query lexical features and used a query-click graph to propagate category labels to unlabeled queries. Prior work from Diaz [1] investigated vertical selection with respect to *news* vertical, where he derived features from news collection, web and vertical query-logs and incorporated click-feedback into the model. Perhaps we can say that the lowest effort approach to vertical selection is the query string alone. Arguello et al. [2] defined two types of query string features, rule-based vertical triggers and geographic features. These triggers map conceptually one-to-one to a target vertical or maps many to one. Some triggers may do not map directly to a vertical, but may provide positive or negative evidence in a supervised classification framework. (e.g., patents, weather, events)[2].

For the set of geographic entities appearing in the query, Arguello et al. [2] defined geographic features which were produced using a rule-based geographic annotation tool. The second source of evidence for vertical selection used evidence from queries previously issued to the vertical, reflecting topics in vertical that are of interest to the users. For each vertical, they computed the query likelihood given by a unigram language model constructed from the vertical’s query-

log. For the third source of evidence, that is, corpora representation of vertical content, they used corpus features, derived from document rankings obtained by issuing the query to different collections. They constructed smaller representative corpora of vertical content local to the vertical selector as discussed previously in section 3.

The findings from Arguello et al. [2] showed that ranking verticals by the query likelihood estimated from the vertical query log language model was the best single evidence to select a vertical. Using rule-based triggers for mapping query strings to verticals led to significant improvement in vertical selection. However this approach does not explicitly model user feedback. Diaz and Arguello [3] proposed algorithms for vertical selection which adapts to user feedback. The objective of their work was to maximize the user satisfaction by presenting the appropriate vertical display. They presented several algorithms for combining the feedback with offline classifier information. They demonstrated that combining results from feedback-only models and offline-only models results in significant improvements. Later work from Arguello et al. [4] proposed methods for reusing training data from the set of existing verticals to learn a predictive model for a target vertical.

3.4 Evaluation

The evaluation of aggregated results is mostly based on implicitly or explicitly collected user feedback. User feedback can be implicitly collected by observing click-through behavior [1,28], that is clicks and skips, or explicitly asking user's preference [29]. The table 1 shows the comparison of various vertical selection approaches on the basis of evaluation measures and metrics used for evaluating aggregation performance. Arguello et al. [4] and Li et al. [20] used precision and recall based evaluation. Arguello et al. [4] adopted two rank-based metrics, average precision and normalized discounted cumulative gain for evaluating model's target vertical predictions. Diaz [1] integrated news vertical always above the first web result using predictions based on clicks and skips [28]. Arguello et al. [2] evaluated single vertical selection in terms of precision, defined as percentage of queries for which we either correctly predict a relevant vertical or correctly predict 'no relevant vertical'.

4. AGGREGATED SEARCH INTERFACES

There are mainly two types of result page design in aggregated search, that is, *blended* and *non-blended* design. Section below discusses both of them.

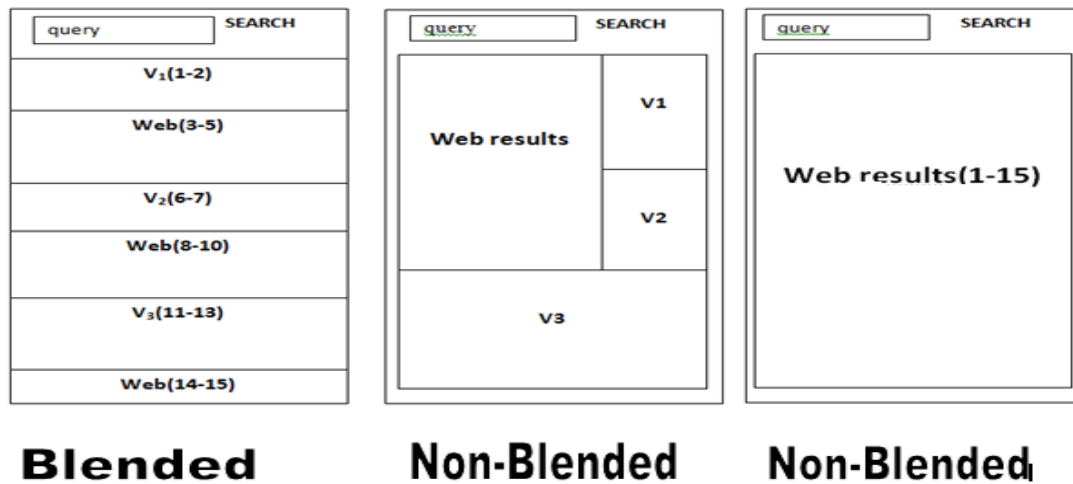
4.1 Blended Design

In blended design, results from different verticals are mingled into a single list. The blended design is mostly used to present results in aggregated search and are

adapted by most of the search engines. In this type of design slotted results from the same vertical are embedded into the result page after ranking them with relevance criteria, in most of the cases, but can also be ranked with some other criteria like newsworthiness, when results come from the news vertical. The earliest work [1] simply integrated the news content from a single vertical above the first web result. Liu et al. [21] proposed a probabilistic model based approach for blended search. Their experimental findings showed that, in generic web search, large number of queries has vertical search intentions. The identification of vertical search intention is proposed in another work from Liu et al. [62], where they used method to propagate social annotation to both queries and vertical search engines. Liu et al. [21] proposed algorithm to model the blended search problem using the conditional dependencies among queries, vertical search engines and search results. The Vertical search engine's query log through unigram language model was used to learn the probability distribution of the model. Their proposed algorithm could effectively blend vertical search result into generic web search, improving the mean average precision by as much as 16% compared to traditional web search without blending. More recent work from Ponnuswami et al. [17] focused on blended result page composition. They presented a machine-learned framework for result page composition in the presence of multiple relevant verticals. Earlier work [1,2,4,3] focused on relevant vertical selection, but this work [17] focused on ranking and placing components or vertical results into different slots in the web search result page, according to the algorithmically computed threshold values. All these work in aggregated search uses machine learning methods requiring training data which includes editorial judgments [20,22,2], indirect labels [4], and click information [1,23,17]. Another recent work from Arguello et al. [18] elicited methods of learning orderings of generic web and vertical content using pairwise labels [24].

4.2 Non-Blended Design

In the non-blended design, the core web results are presented to the user without mingling the vertical results. The vertical results could also be shown but in separate panels [16], and the separate vertical results are not necessarily relevant to the core web results or have any relationship with each other. Mostly, the placement of different verticals is predefined. Naver [47] and kosmix [49] shows results in this type of pattern. Main web results are generally shown on the left side of the search result page, within the largest panel, but not necessarily as in [28] Sushmita et al tested the non-blended interface with image vertical, placing it right-top, right-bottom, and left. Figure 3 shows both design layouts. In blended design, verticals (V_1 , V_2 , V_3) are embedded within the web results whereas non-blended design shows either no vertical results or vertical results in separate panels.



4.3 Viewing Results

M.Lalmas [16] gave an overview of some of the eye-tracking studies ([48],[50],[51]) which proved to be useful in designing and evaluating aggregated search interfaces. When compared with the studies related to design and evaluation of conventional web search interfaces, much less work has been done till date for aggregated search interfaces. Hotchkiss [51] showed in conventional web search interfaces, user starts viewing or scanning the result from upper left corner and scan results first vertically downwards and then across. In Nielsen[50], scanning results in conventional web interface is referred to as F pattern, but in aggregated search interface scanning depends on how verticals, say an image vertical, is blended in the result page. User starts scanning from the image or thumbnail, first vertically than across referred to as E pattern [51]. Aula and Rodden [48] however studied that returning aggregated or blended results have no effect in order of scanning the results.

4.4 User Behavior for Result Presentation

In order to find out user's behavior for search result display many factors has to be taken in consideration [1,52,28,53,54,55], such as vertical types, position of relevant verticals and search task orientation strength towards a particular vertical. Sushmita et al. [28] tested both blended and non-blended aggregated search interfaces using 1,296 search sessions performed by 48 participants. Their analysis suggested that position of search results is significant only in blended design and not in non-blended design. The position of augmented elements in the blended design affects the user's click-through behavior. Some of the previous studies [56-59] also elicited this. Also they discovered that search task's orientation towards a particular source, that is news, images and videos in their case, can affect user's click-through behavior. They discovered that videos resulted in different click-through pattern than news and images. According to Goodman and Feldblum [55]

video results with thumbnails generate higher click-through rates across all positions and especially helped results in the lower positions.

An eye tracking study from Gaun and Cutrell [58] showed that when results were placed relatively low in the result page, people spent more time in searching and were less successful in their search task.

5. CONCLUSION AND FUTURE WORK

Vertical aggregation refers to assembling and merging heterogeneous results from specialized search services called *verticals* into the core web results. Vertical selection is the sub-task of selecting relevant verticals in response to a user's query. In this paper we reviewed various vertical selection approaches and aggregated result presentation. Major problems associated with vertical aggregation include sources for predictive evidence and training data, relevance modeling and evaluation. In terms of future work, efforts should be made to increase diversity in vertical presentation as verticals can be redundant as well as complementary and should evaluate whole page relevance while aggregating. More work can be done on attention modeling as some verticals can visually attract a user's attention without being relevant but it needs a better understanding of user's attention in response to page layout decisions. In terms of layout constraints for aggregation, majority of work has focused on conservative ranked list layouts, in future we can think about arbitrary layouts. Finally work can be done for automatically detecting when to begin modeling a new vertical as currently set of candidate verticals are manually detected.

Table 1: Comparison of various vertical selection approaches on the basis of evaluation measures and metric

Author	Method	Evaluation Measures	Evaluation Metric	In terms of Relevance	Significance/Objective
Li et al. [20]	Classified queries into two classes of vertical intent. (<i>product</i> and <i>job</i>)	Precision and Recall	Accuracy	A vertical is <i>relevant</i> if it satisfies user's job or product intent	Predict appropriate vertical when relevant; otherwise, predict no relevant vertical
Diaz [1]	Predicting when to display news result (always displayed above web results)	Correctly predicted clicks and skips.	Preference	Presentation is good if the user can easily find more relevant content before less relevant content.	Predict appropriate vertical preferences.
Arguello et al. [2]	Single vertical is predicted for each query	No. of correctly predicted queries	Accuracy	Vertical is <i>relevant</i> if it satisfies some possible intent	Predict appropriate vertical when relevant; otherwise, predict no relevant vertical
Diaz and Arguello [3]	Vertical selection in presence of user feedback.	Simulated stream of queries.	Utility of whole page layout.	A vertical is <i>relevant</i> if it satisfies the intent of a particular user at a particular time.	Predict appropriate vertical when relevant; otherwise, predict no relevant vertical.
Arguello et al. [4]	Model adaptation for vertical selection.	Precision and Recall	Accuracy	Vertical is <i>relevant</i> if it satisfies some possible intent	Predict new (target) vertical reusing training data from existing (source) vertical.

6. REFERENCES

- [1] F. Diaz. Integration of news content into web results. In WSDM 2009, 2009.
- [2] J. Arguello, F. Diaz, J. Callan, and J.-F. Crespo.: Sources of evidence for vertical selection. In SIGIR09, pages 315-322, 2009.
- [3] F. Diaz and J. Arguello. Adaptation of online vertical selection predictions in the presence of user feedback. In SIGIR 2009, pages 323-330. ACM, 2009.
- [4] J. Arguello, F. Diaz, and J.-F. Paiement. Vertical selection in the presence of unlabeled verticals. In *SIGIR 2010*, Pages 691-698. ACM, 2010.
- [5] L. Gravano, C. Chang, H. Garca-Molina and A. Paepcke. STARTS: Stanford proposai for internet metasearching. SIGMOID, pages 207-218, T1997.
- [6] M. Shokouhi and L. Si.: Federated search. Foundations and Trends in Information Retrieval, 5(1):1-102, 2011.
- [7] J. Callan and M. Connell.: Query-based sampling of text databases. ACM Trans. Inf. Syst., 19:97-130, 2001.
- [8] J. P. Callan, Z. Lu, and W. B. Croft.: Searching distributed collections with inference networks. In SIGIR95, pages 21-28, 1995.
- [9] M. Shokouhi, J.Zobel, S. Tahaghoghi and F. Scholer. Using query logs to establish vocabularies in distributed information retrieval. IP&M, 43(1):169180, 2007.
- [10] B.Yuwono and D.L.Lee. Server Ranking for Distributed text Retrieval Systems on the internet. In DASFAA 1997, pages 1-50. World Scientific Press, 1997.
- [11] J. Xu and W. B. Croft. Cluster-based language models for distributed retrieval. In SIGIR 1999, pages 254-261. ACM, 1999.
- [12] D. Shen, J.-T. Sun, Q. Yang, and Z. Chen. Building bridges for web query classification. In SIGIR 2006, pages 131-138, 2006.
- [13] S. M. Beitzel, E. C. Jensen, D. D. Lewis, A. Chowdhury, and O. Frieder. Automatic classification of web queries using very large unlabeled query logs. TOIS, 25(2):9, 2007.
- [14] S. M. Beitzel, E. C. Jensen, O. Frieder, D. D. Lewis, A. Chowdhury and A. Kolcz. Improving automatic query classification via semi-supervised learning. In ICDM 2005, pages 42-49, 2005.
- [15] L. Gravano, H. Garca_molina, A. Tomasic, I. Rocquencourt and N. L. Gravano. Gloss: Text_source discovery over the internet. Transactions on Database systems, 24: pages 229-264, 1999.

- [16]] M. Lalmas, *Advanced topics in information retrieval*, The Information Retrieval (Eds) M. Melucci, R. Baeza-Yates 2011, XXX, 306 p. 87 illus., Hardcover ISBN: 978-3-642-29945-1
- [17] A. K. Ponnuswami, K. Pattabiraman, Q. Wu, R. Gilad-Bachrach, and T. Kanungo. On composition of a federated web search result page: Using online users to provide pairwise preference for heterogeneous verticals. In WSDM 2011, pages 715{724. ACM, 2011.
- [18] J. Arguello, F. Diaz, and J. Callan. Learning to aggregate vertical results into web search results. *CIKM*, 2011.
- [19] L. Si and J. Callan. Relevant document distribution estimation method for resource selection. In SIGIR 2003, pages 298-305, 2003.
- [20] X. Li, Y. -Y. Wang and A. Acero. Learning query intent from regularized click graphs. In SIGIR 2008, pages 339-346. ACM 2008.
- [21] N. Liu, J. Yan, and Z. Chen. A Probabilistic Model based Approach for Blended Search. *WWW 2009*, pages 20-24. ACM 2009.
- [22] J. Arguello, F. Diaz, J. Callan, and B. Carterette.: A methodology for evaluating aggregated search results. In ECIR11, pages 141-152, 2011.
- [23] A. C. Konig, M. Gamon, and Q. Wu. Click-through production for news queries. In SIGIR 2009, pages 347-354. ACM 2009.
- [24] T. -Y. Liu. Learning to rank for information retrieval. *Foundations and Trends in Information Retrieval*, 3:pages 225-331, 2009.
- [25] L. Gravano, H. Garcia_Molina, and A. Tomasic. The effectiveness of GIOSS for the text database discovery problem. In SIGMOID 1994, ACM 1994.
- [26] L. Si. Federated search of text search engines in uncooperative environments. Phd. Thesis, Carnegie Mellon University, Pittsbiurgh, PA, 2006.
- [27] D. Shen, R. Pan, J. -T. Sun, J. J. Pan, K. Wu, J. Yin, and Q. Yang. Q2c@ust: our winning solution to query classification in kddcup 2005. In SIGKDD Explor. Newsl., 7(2): pages 100-110, 2005.
- [28] S. Sushmita, H. Joho, M. Lalmas and R. Villa. Factors Affecting Click-Through Behavior in Aggregated Search Interfaces. In CIKM, pages 519-528, 2010.
- [29] D. Zhu and B. Carterette. An analysis of assessor behavior in crowdsourced preference judgements. In SIGIR , pages 21-26, ACM 2010.
- [30] W. Y. Meng and C. Yu, *Advanced Metasearch Engine Technology*. Morgan & Claypool Publishers, 2010. ISBN 1608451925.
- [31] E. Selberg and O. Etzioni, "Multi-service search and comparison using the metacrawler," in *Proceedings of the Fourth International Conference on World Wide Web*, Boston, MA: Oreilly, 1995. ISBN 978-1-56592-169-6.
- [32] E. Selberg and O. Etzioni, "The MetaCrawler architecture for resource aggregation on the web," *IEEE Expert*, vol. 12, no. 1, pages. 8–14, ISSN 0885-9000, 1997.
- [33] S. Gauch and G. Wang, "Information fusion with ProFusion," in *Proceedings of the First World Conference of the Web Society*, pages. 174–179, San Francisco, CA, 1996.
- [34] S. Gauch, G. Wang, and M. Gomez, "ProFusion: Intelligent fusion from multiple distributed search engines," *Journal of Universal Computer Science*, vol. 2, no. 9, pages. 637–649, ISSN 0948-695X, 1996.
- [35] K. Liu, W. Meng, J. Qiu, C. Yu, V. Raghavan, Z. Wu, Y. Lu, H. He, and H. Zhao, "Allinonenews: development and evaluation of a large-scale news metasearch engine," in *Proceedings of the ACM SIGMOD International Conference on Management of Data*, (C. Y. Chan, B. C. Ooi, and A. Zhou, eds.), pages. 1017–1028, Beijing, China, 2007. ISBN 978-1-59593-686-8.
- [36] E. Han, G. Karypis, D. Mewhort, and K. Hatchard, "Intelligent metasearch engine for knowledge management," in *Kraft et al. [152]*, pages 492–495, 2003. ISBN 1-58113-723-0.
- [37] Z. Bar-Yossef and M. Gurevich, "Efficient search engine measurements," in *Williamson et al. [269]*, pages 401–410. ISBN 978-1-59593-654-7.
- [38] D. Dreilinger and A. Howe, "Experiences with selecting search engines using metasearch," *ACM Transaction on Information Systems*, vol. 15, no. 3, pages 195–222, ISSN 1046-8188, 1997.
- [39] A. Smeaton and F. Crimmins, "Using a data fusion agent for searching the WWW," in *selected papers from the sixth international conference on world wide web*. CA Elsevier, 1997.
- [40] E. Glover, S. Lawrence, W. Birmingham, and C. Giles, "Architecture of a metasearch engine that supports user information needs," in *Gauch [102]*, pages 210–216, 1999. ISBN 1-58113-1461
- [41] W. Meng, C. Yu, and K. Liu, "Building efficient and effective metasearch engines," *ACM Computing Surveys*, vol. 34, no. 1, pages 48–89, ISSN 0360-0300, 2002.
- [42] J. Arguello, F. Diaz, and M. Shokouhi. Integrating and Ranking Aggregated Content on the Web. IUNC Chapel Hill2Yhoo! Labs 3Microsoft Research, 2012.
- [43] J. Lu. "Full-text federated search in peer-to-peer networks," Phd thesis, Carnegie Mellon University, 2007.
- [44] J. Conrad and J. Claussen, "Early user — system interaction for database selection in massive domain-specific online environments," *ACM Transactions on Information Systems*, vol. 21, no. 1, pages 94–131, ISSN 1046-8188, 2003.
- [45] J. Conrad, X. Guo, P. Jackson, and M. Meziou. Database selection using actual physical and acquired logical collection resources in a massive domain specific operational environment, in *Bernstein et al.[29]*, pages 71-82.
- [46] J. Conrad, C. Yang, and J. Claussen, "Effective collection metasearch in a hierarchical environment: global vs. localized retrieval performance," in *J'arvelin t al. [142]*. ISBN 1-58113-561-0.
- [47] Naver Search Engine. <http://www.naver.com/>.

- [48] A. Aula and K. Rodden. Eye-tracking studies: more than meets the eye. June 2009.
- [49] Kosmix Search Engine. <http://www.kosmix.com/>.
- [50] J. Neilsen. *Eyetracking Web Usability*. Kara Pernice, 2009.
- [51] G. Hotchkiss. Eye Tracking on Universal and Personalized Search. <http://searchengineland.com/eye-trackink-on-universal-and-personalized-search-12233>, September 2007.
- [52] iProspect Blended Search Result Study. http://www.iprospect.com/about/researchstudy_2008_blendedsearchresults.htm, April 2008.
- [53] S. Sushmita, H. Joho and M. Lalmas. A Task-Based Evaluation of an Aggregated Search Interface. SPIRE, pages 322-333, 2009.
- [54] S. Sushmita, B. Piwowarski and M. Lalmas. Dynamics of Genre and Domain Intents, AIRS, 2010.
- [55] E. Goodman and E. Feldblum. Blended Search and the New Rules of Engagement. ComScore Report, October 2010.
- [56] T. Joachims, L. Granka, B. Pan, H. Hembrooke and G. Gay. Accurately interpreting clickthrough data as implicit feedback. ACM SIGIR, pages 154–161, 2005.
- [57] E. Agichtein and Z. Zheng. Identifying “best bet” web search results by mining past user behavior. ACM SIGKDD, pages 902–908, 2006.
- [58] Z. Guan and E. Cutrell. An eye tracking study of the effect of target rank on web search. ACM SIGCHI, pages 417–420, 2007.
- [59] M.T. Keane, M. O’Brien and B. Smyth. Are people biased in their use of search engines? CACM, 51(2):49–52, 2008.
- [60] P. Ipeirotis and L. Gravano. Distributed search over the hidden web: Hierarchical database sampling and selection. In *proceedings of the 28th International Conference on Very Large Databases (VLDB)*, 2002.
- [61] L. Si and J. Callan. Unified utility maximization framework for resource selection. In *proceedings of the ACM CIKM International Conference on Information and Knowledge Management*, pages 32-41, Washington, DC, 2004. ISBN 1-58113-874-1.
- [62] N. Liu, J. Yan, W. Fan, Q. Yang, and Z. Chen. Identifying Vertical Search Intention of Query through Social Tagging Propagation. WWW 2009, pages 20-24, ACM 2009.