Diverse Relevance Ranking in Web Scrapping for Multimedia Answering

R. Priyambiga and D. Shanthi

Abstract--- *With exponential growth of media padding on the web, the skill to hunt for media entities not just based on text footnote, but also visual content has become important. The main problem is the lack of diversity of generated media data. To enhance the above said process, a scheme is proposed to enrich text answers with image and video and able to take relevance and diversity into account by exploring the content of images. Our Scheme investigates a rich set of techniques: question/answer categorization, query creation, image and video hunt re-ranking. This approach automatically determines which type of media information should be added for textual answers with high precision and to make enrich media data more diverse.*

Keywords--- *Re-Ranking, Question/Answer Classification, Query Generation*

I. INTRODUCTION

QUESTION answering (QA) is defined as the task of automatically providing a accurate answer to a natural language question posed by users [1]. Compared to keyword-based search systems, it greatly facilitates the message between humans and computers by naturally stating users' intention in plain sentences. It also avoids the scrupulous browsing of a vast quantity of information contents returned by search engines for the correct answers. In spite of great progress and encouraging results have been reported, traditional automated QA still faces challenges that are not easy to tackle, such as the profound understanding of complex questions and the tricky syntactic, semantic and relative processing to produce answers. It is found that, for the most part , computerized approach cannot achieve outcome that are as good as those generated by blue-collar processing [2].

In conjunction with the propagation and enhancement of fundamental communication technologies, community question answering (cQA) has emerged as an enormously popular alternative to finding information online, owning to the following facts. Initially information seekers are able to post their specific questions on any topic and obtain answers provided by additional participants. By leveraging group of people efforts, they are able to get better answers than simply using search engine to find them. Second, in comparison with robotic QA systems, cQA usually receives answers with better quality as they are generated based on human intelligence. Third, over times, a remarkable number of QA pairs have been accumulated in their repositories, and it smooth the progress of the preservation and salvage of answered questions.[3] Preservation and search of answered questions. For example, Wiki Answer, for the most part well-known cQA systems, hosts more than 15 million answered questions distributed in 9,000 categories (as of April 2012).

In spite of great success has been achieved; existing cQA forums mostly provide only textual answers, as shown in Fig 1. Yet a portrait is worth a thousand words[1]. In many cases, the questions cannot be well explained using only texts, and it will be much better to envisage the answers with images and videos. Fig 1 illustrates such an example: for the question "What is the procedure to cook beef gravy?" the answer is described by several long sentences. However, users still can scarcely seize the procedure. Evidently, it will be much enhanced if there is an accompanying video describing the process. Therefore, the textual answers in cQA can be significantly enhanced by adding multimedia contents, and it will provide answer seekers with better experience.

Fig. 1: An Example on Yahoo! Answer

In fact in cQA corpuses, there are already many answers that directly embed hyperlinks to images or videos from which the users can get supplementary information in media outline. For example, for the query "What are the best steps to take in organize to regret", the most excellent answer on Y!A is a URL that guide information seekers to YouTube. This point toward that numerous answers can be enhanced by leveraging multimedia information[2]. However, existing cQA forums do not provide adequate support on using media information.

In this work, a scheme is proposed that is able to find high relevant image or video information to complement the community-contributed textual answers in cQA. It explores a rich set of techniques, including question/answering categorization, query extraction and categorization, image and video hunt reranking, relevance ranking, etc.

R. Priyambiga, Student, Department of CSE,PSNA College of Engineering and Tech, Tamil Nadu.E-mail:priyambiga@gmail.com

D. Shanthi, Professor, Department of CSE, PSNA College of Engineering and Tech, Tamil Nadu. E-mail:dshan71@gmail.com

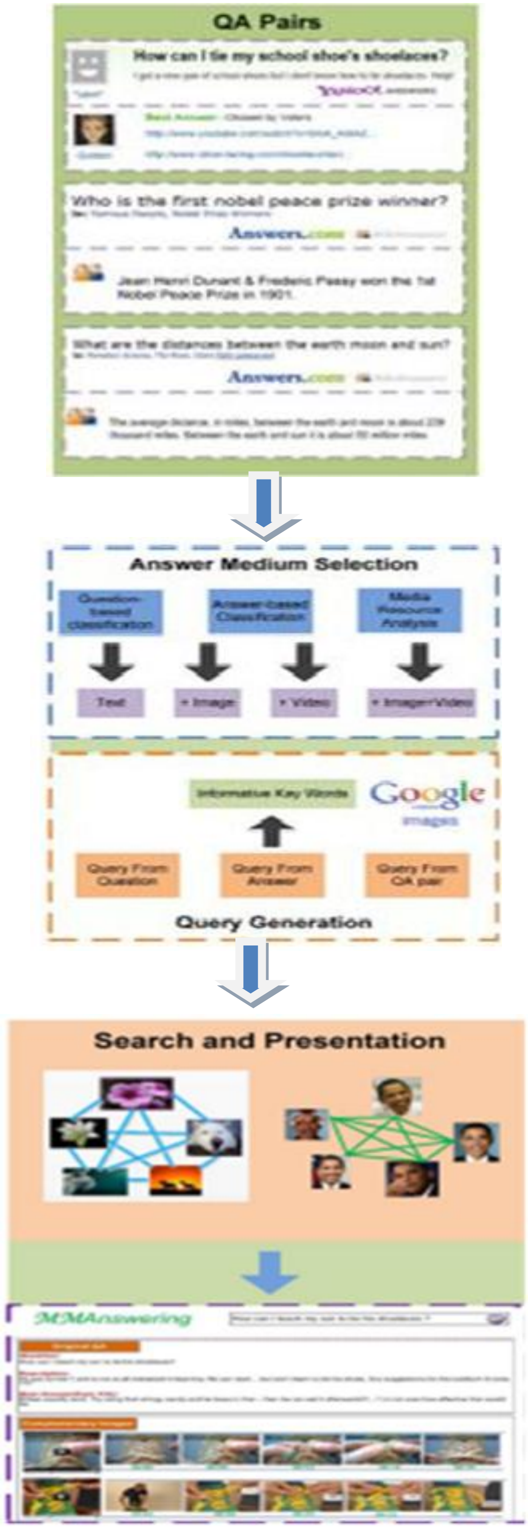

Fig. 2: Multimedia Answering Scheme

As shown in Fig 2, the scheme consists of four main components: (1) Answer medium selection. In this work, we consider the subsequent four cases for answer media: (a) only texts, i.e., the original textual answers are sufficient; (b) text +image, i.e., image information needs to be added; (c) text+ video, i.e., only video information is to be added; and (d)text + image + video, i.e., we add both image and video information. We observe it as a QA pair classification dilemma, that is, given a question and its community-contributed answer in cQA corpus, we categorize it into one of the exceeding four module. (2) Multimedia query generation. In order to pull together multimedia records from the web, we generate queries from each QA pair. Here we generate three types of inquiry from (a) question, (b) answer, and (c) both question and answer. We then choose one from the three queries by hearing a classification model. (3) Multimedia data collection and arrangement. (4)We then perform relevance ranking and duplicate removal to obtain a set of accurate and representative samples for presentation together with the textual answers.

The aid of this work can be reviewed as follows:

- Disambiguating entities by supporting fast and accurate mappings of textual phrases onto named entities in the knowledge base
- An enabler for entity-relation- ship-oriented semantic hunt on the Web, for detecting entities and relations in Web pages and reasoning about them in expressive logics
- A backbone for natural-language question answering that would aid in dealing with entities and their relationships in answering when/where/who etc. questions
- A method for acquisition of supplementary knowledge and largely automated maintenance and growth of the knowledge base.

II. Related Work

One of the major problems in question answering (QA) is that the queries are either too brief or often do not contain most relevant terms in the target corpus. In order to defeat this problem, integrate external knowledge extracted from the Web and Word Net to perform Event-based QA on the TREC-11 task [2]. It extends our approach to perform event-based QA by detection the structure within the external knowledge. The knowledge structure loosely models different facets of QA events, and is used in conjunction with consecutive constraint relaxation algorithm to achieve effective QA. Our results obtained on TREC-11 QA corpus indicate that the new approach is more effective and able to attain a confidence-weighted score of above 80% [11]. It concentrates only on retrieving textual answers.

Yahoo Answers (YA) is a large and diverse question-answer forum, acting not only as a medium for sharing technical knowledge, but as a place where one can try to find advice, gather opinions, and satisfy one's curiosity about a countless number of things.[3] Here, we seek to understand YA's knowledge sharing activity. We analyze the forum categories and cluster them according to content characteristics and patterns of interaction among the users. While interactions in some categories resemble expertise sharing forums, others integrate discussion, everyday advice, and support. With such a diversity of categories in which one can participate, we find that some users focus narrowly on specific topics, while others participate across categories [12]. This not only allows us to map related categories, but to characterize the entropy of the users' interests. We find that lower entropy correlates with receiving higher answer ratings, but only for categories where factual proficiency is primarily sought after. We combine both user attributes and answer characteristics to predict, within a given category, whether a particular answer will be chosen as the best answer by the asker. It also concentrates only on retrieving textual answers.

For question answering on lecture videos a multifaceted approach is introduced [4]. Text extracted from PowerPoint slides associated with the lecture videos is used as a source of domain knowledge to boost the answer taking out performance on these domain specific videos. The three steps of this approach are described and the evaluation plan is discussed. It focuses only on domain specific videos.

For image classification an integrated multilabel multi-instance learning (MLMIL) approach is introduced based on hidden conditional random fields (HCRFs), which simultaneously captures both the connections between semantic labels and regions, and the correlations among the labels in a single formulation. We apply this MLMIL framework to image classification and report superior performance compared to key existing approaches over the MSR Cambridge (MSRC) and Corel data sets[5]. It only increases the performance for classifying image.

Photo-based question answering is a useful way of finding information about physical objects. Current question answering (QA) systems are text-based and can be difficult to use when a question involves an object with distinct visual features. A photo-based QA system allows direct use of a photo to refer to the object. We develop three-layer system architecture for photo-based QA that brings together recent technical achievements in question answering and image toning [6]. The first, template-based QA layer matches a query photo to online images and extracts structured data from multimedia databases to answer questions about the photo. To simplify image matching, it exploits the question text to filter images based on categories and keywords. The second, information retrieval QA layer searches an internal repository of resolved photo-based questions to retrieve relevant answers. The third, human-computation QA layer leverages community experts to handle the most difficult cases. A series of experiments performed on a pilot dataset of 30,000 images of books, movie DVD covers, grocery items, and landmarks demonstrate the technical feasibility of this architecture. We present three prototypes to show how photo-based QA can be built into an online album, a text-based QA, and a mobile application[13]. It retrieves only textual and image answers.

Insufficiency of labeled training data is a major obstacle for automatic video annotation. Semi-supervised learning is an effective approach to this problem by leveraging a large amount of unlabeled data. However, existing semi-supervised learning algorithms have not verified promising results in large-scale video annotation due to several difficulties, such as large variation of video content and intractable computational cost. Here, a novel semi-supervised learning algorithm named semi supervised kernel density estimation (SSKDE) is proposed based on kernel density estimation (KDE) approach [7]. While only labeled data are utilized in classical KDE, in SSKDE both labeled and unlabeled data are leveraged to estimate class conditional probability densities based on an extended form of KDE. It is a non-parametric method, and it thus naturally avoids the model assumption problem that exists in many parametric semi-supervised methods. Meanwhile, it can be implemented with an efficient iterative solution process. So, this method is appropriate for video annotation. Furthermore, motivated by existing adaptive KDE approach, we propose an improved algorithm named semi-supervised adaptive kernel density estimation (SSAKDE). It employs local adaptive kernels rather than a fixed kernel, such that broader kernels can be applied in the regions with low density. In this way, more accurate density estimates can be obtained[14]. Widespread experiments have demonstrated the effectiveness of the proposed methods.

The use of set expansion (SE) is explored to improve question answering (QA) when the expected answer is a list of entities belonging to a certain class. Given a small set of seeds, SE algorithms extract textual resources to produce an extended list including additional members of the class represented by the seeds [8]. We explore the hypothesis that a noise-resistant SE algorithm can be used to extend candidate answers produced by a QA system and generate a new list of answers that is better than the original list produced by the QA system. We further introduce a hybrid approach which combines the original answers from the QA system with the output from the SE algorithm. Investigational results for several state-of-the-art QA systems show that the hybrid system performs better than the QA systems alone when tested on list question data from past TREC evaluations.

Existing community question-answering forums usually provide only textual answers. However, for many questions, pure texts cannot provide spontaneous information, while image or video contents are more appropriate. Here, we introduce a scheme that is able to enrich text answers with image and video information [9]. Our scheme investigates a rich set of techniques including question/answer categorization, query creation, image and video hunt re-ranking, etc. Given a question and the community-contributed answer, our approach is able to determine which type of media information should be added, and then automatically collects data from Internet to enrich the textual answer [15]. Different from some efforts that attempt to directly answer questions with image and video data, our approach is built based on the community-contributed textual answers and thus it is more feasible and able to deal with more complex questions. We have conducted empirical study on more than 3,000 QA pairs and the results demonstrate the effectiveness of our approach. The problem is lack of diversity.

III. MATERIALS AND METHODS

A Diverse relevance ranking scheme is proposed which is able to take relevance and diversity into account. It takes benefit of both the content of images and their correlated tags. Initially it estimates the relevance scores of images with respect to the query term based on both the visual information of images and the semantic information of related tags. Then we mine the semantic similarities of images based on their tags. With the relevance scores and the similarities, the position list is generated by a greedy ordering algorithm which optimizes Average Diverse Precision (ADP), a novel measure that is extended from the straight Average Precision (AP). Fig 3 illustrates the proposed multimedia answering scheme.

A. Posting the Opinion

In this module, we get the opinions from various people about medical through online. The opinions may be of two types. Direct opinion is to post a comment about the components and attributes of products directly. Comparative opinion is to post a comment based on comparison of two or more products. The comments may be positive or negative.

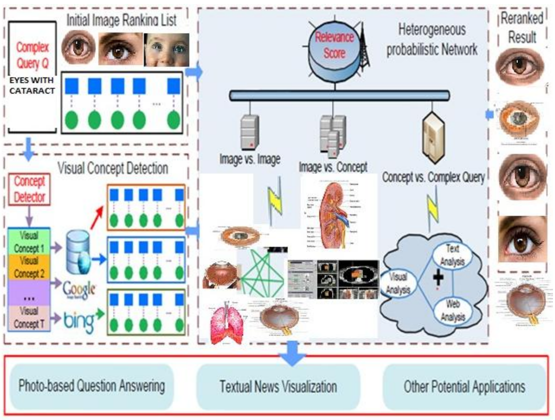

Fig. 3: The schematic Illustration of the Proposed Multimedia Anwering Scheme

Description of the Architectural Diagram
Input: Complex Query
First Step: Retrieval of media data for complex query
Second Step: Explores the content of media data by using Diverse Relevance Ranking Algorithm.
Output: Produce multimedia data with relevance and diversity in cQA.

B. Diverse Relevance Ranking

However, the quality of recommendations can be evaluated along a number of dimensions, and relying on the accuracy of recommendations alone may not be enough to find the most relevant items for each User, these studies argue that one of the goals of recommender systems is to provide a user with highly personalized items, and more diverse recommendations result in more opportunities for users to get recommended such items. With this motivation, some studies proposed new recommendation methods that can increase the diversity of recommendation sets for a given individual user. They can give the feedback of such items.

C. Rating Predictions

First, the ratings of unrated items are estimated based on the available information (typically using known user ratings and possibly also information about item content) using some recommendation algorithm. Heuristic techniques typically calculate recommendations based directly on the previous user activities (e.g., transactional data or rating values). For each user, ranks all the predicted items according to the predicted rating value ranking the candidate (highly predicted) items based on their predicted rating value, from lowest to highest (as a result choosing less popular items).

D. Ranking Approach

Ranking items according to the rating variance of neighbors of a particular user for a particular item. There exist a number of different ranking approaches that can improve recommendation diversity by recommending items other than the ones with topmost predicted rating values to a user. A comprehensive set of experiments was performed using every rating prediction technique in conjunction with every recommendation ranking function on every dataset for different number of top-N recommendations.

E. Diverse Relevance Ranking Algorithm

There exist multiple variations of neighborhood-based CF techniques. In this paper, to estimate R*(u, i), i.e., the rating that user u would give to item i, we first compute the similarity between user u and other users u' using a cosine similarity metric.

Formula: for finding prediction by using recommendation techniques:

Sim(u,u’)=sum[R(u,i).R(u’,i)]/

sqrt[sum[R(u,i)2].sqrt[sum[R(u’,i)]

Where I (u, u') represents the set of all items rated by both user u and user u'. Based on the similarity calculation, set N (u) of nearest neighbors of user u is obtained. The size of set N (u) can range anywhere from 1 to |U|-1, i.e., all other users in the dataset.

Then, R*(u, i) is calculated as the adjusted weighted sum of all known ratings R (u', i) Here R (u) represents the average rating of user u. A neighborhood-based CF technique can be user-based or item-based, depending on whether the similarity is calculated between users or items, the user-based approach, but they can be straightforwardly rewritten for the item-based approach because of the symmetry between users and items in all neighborhood-based CF calculations. In our experiments we used both user-based and item-based approaches for rating estimation.

IV. Experiments

A. Experimental Settings

The empirical evaluation of the proposed scheme is introduced. We first introduce the experimental settings, such as dataset and ground truth labeling. Then, we present two kinds of evaluation. One is local evaluation which tests the effectiveness of the components in the scheme, such as answer medium selection, query selection, and multimedia search re-ranking. The other one is global evaluation which tests the usefulness of the enrichment of media data for question answering with relevance and diversity.

B. Evaluation of Posting the Opinion

We first are collecting the opinions from various people about medical through online. Opinions may be direct or indirect.

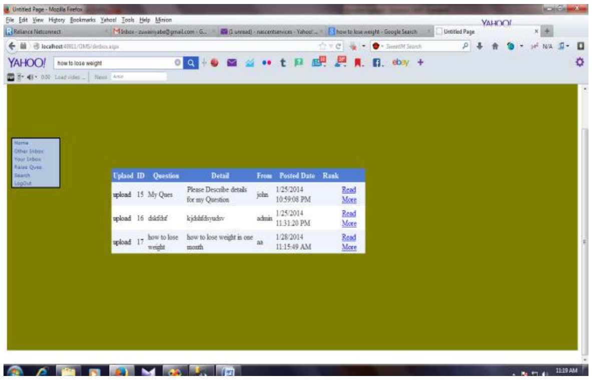

Fig. 4: Questions Asked by the User

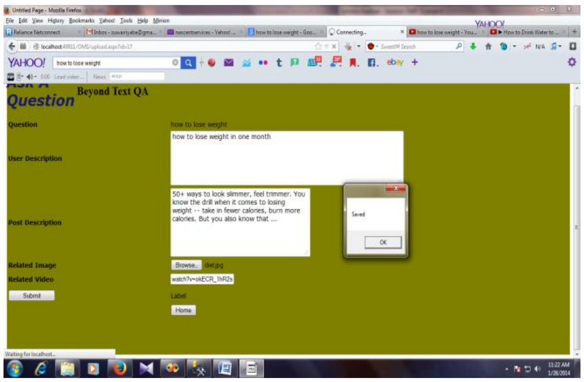

Fig. 5: Posting the Opinion by another User

C. Evaluation of Diverse Relevance Ranking

Similar to Fig.5 various people post the opinion about the question asked by the user. The next process is to find the relevance and diversity of the opinion post by various people. In order to accomplish this task rating and ranking approach is established. Based on the available information (typically using known user ratings such as (like and dislike) and possibly also information about item content) rating is predicted.

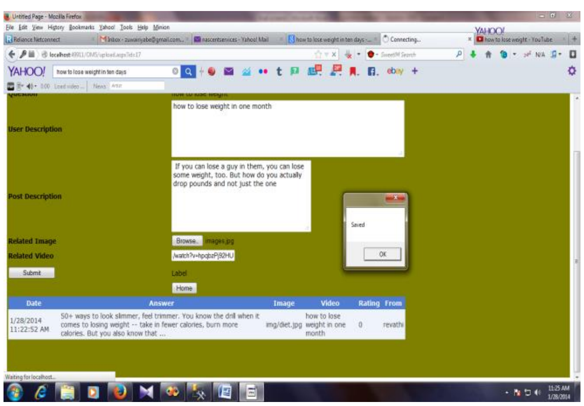

Fig. 6:Rating Prediction

After prediction of rating the next process is to rank items according to the rating variance of neighbors of a particular user for an exacting item.

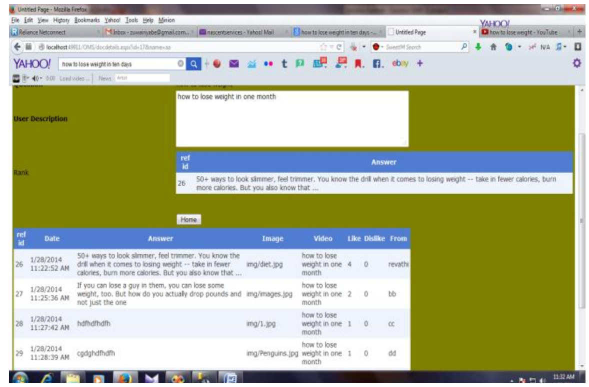

Fig. 7: Ranking Approach

Finally,textual answers are enriched with media answers with high relevance and diversity using DRAWMA scheme.

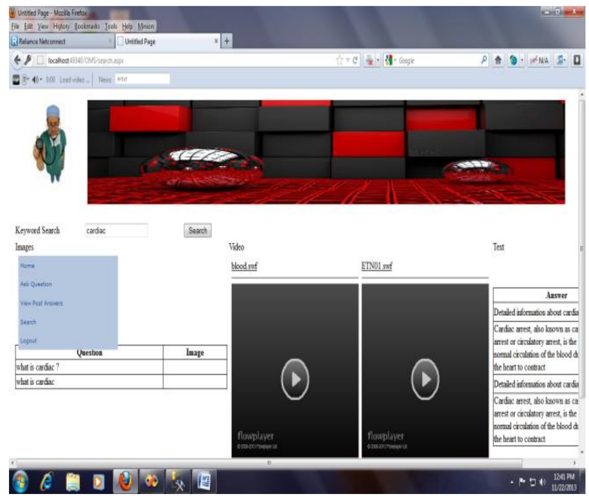

Fig. 8: Results of Multimedia Answering for the Question "What is Cardiac?"

V. CONCLUSION

In this work, a scheme diverse relevance ranking method is used to enrich text answers with image and video and able to take relevance and diversity into account by exploring the content of images. For a specified QA pair from cQA, our scheme initially it predicts which medium is suitable to improve the original textual answer. Then, it automatically generates a query based on the QA facts, and retrieves significant image and video from hunt engines. Lastly, query-dependent reranking and duplicate removal are performed using diverse relevance ranking to obtain a set of images and videos for presentation along with the original textual answer with high precision. Diverse relevance ranking scheme provide disambiguating entities by supporting fast and accurate mappings of textual phrases onto named entities in the knowledge base and also An enabler for entity-relation-ship-oriented semantic search on the Web, for detecting entities and relations in Web pages and reasoning about them in expressive logics and a backbone for natural-language question answering that would aid in dealing with entities and their relationships in answering when/where/who etc. questions.

REFERENCES

[1] Liqiang Nie,Meng Wang,Yue gao,Zheng-Jun Zha,Tat-Seng Chau,"Beyond Text QA:Multimedia Answer generation by Harvesting Web Information",in IEEE Transaction s on Multimedia ,VOL.15,No.2,Febraury 2013

[2] H. Yang, T.-S. Chua, S. Wang, and C.-K. Koh, "Structured use of external knowledge for event-based open domain question answering," in Proc. ACM Int. SIGIR Conf., 2003.

[3] G. Zoltan, K. Georgia, P. Jan, and G.-M. Hector, Questioning Yahoo! Answers, Stanford InfoLab, 2007, Tech. Rep.

[4] J. Cao, F. Jay, and J. Nunamaker, "Question answering on lecture videos: A multifaceted approach," in Proc. Int. Joint Conf. Digital Libraries, 2004.

[5] Zheng-Jun Zha,Xian-Sheng Hua, Tao Mei, Jingdong Wang, Guo-Jun Qi, Zengfu Wang, "Joint Multi-Label Multi-Instance Learning for Image Classification", Department of Automation ‡Internet Media Group ,University of Science and Technology of China ,Microsoft Research Asia.

[6] T. Yeh, J. J. Lee, and T. Darrell, "Photo-based question answering," in Proc. ACM Int. Conf. Multimedia, 2008.

[7] Wang,X.S.Hua,T.Mei,R.Hong,G.J.Qi,Y.Song,andL.R. Dai, "Semi-supervised kernel density estimation for video annotation," *Comput. Vision Image Understand.*, vol. 113, no. 3, pp. 384–396, 2009.

[8] R. C. Wang, N. Schlaefer, W. W. Cohen, and E. Nyberg, "Automatic set expansion for list question answering," in *Proc. Int. Conf. Empirical Methods in Natural Language Processing*, 2008.

[9] L. Nie, M. Wang, Z. Zha, G. Li, and T.-S. Chua, "Multimedia answering: Enriching text QA with media information," in *Proc. ACM Int. SIGIR Conf.*,2011.

[10] Trec: The text retrieval conference. see http://trec.nist.gov/.

[11] L. A. Adamic, J. Zhang, E. Bakshy, and M. S. Ackerman. Knowledge sharing and yahoo answers: everyone knows something. In Proceedings of International World Wide Web Conference, 2008.

[12] H. Cui, M.-Y. Kan, and T.-S. Chua. Soft pattern matching models for de_nitional question answering. ACM Transactions on Information Systems, 2007.

[13] Z. Gyongyi, G. Koutrika, J. Pedersen, and H. Garcia-Molina. Questioning yahoo! answers. Technical report, Stanford InfoLab, 2007.

[14] S. A. Quarteroni and S. Manandhar. Designing an interactive open domain question answering system. Journal of Natural Language Engineering, 2008.