

# The 2nd workshop on Vertical Search Relevance at WSDM 2015

Dawei Yin  
Yahoo Labs  
daweiy@yahoo-inc.com

Chih-Chieh Hung  
Rakuten Inc.  
smallosin@gmail.com

Rui Li, Yi Chang  
Yahoo Labs  
{ruililab,yichang}@yahoo-inc.com

It is our great pleasure to welcome you to the VSR 2015, the second International Workshop on Vertical Search Relevance, held as part of the WSDM 2015 conference in Shanghai, China.

**Background:** As the popularity of search engines has grown, the information needs of end users continue being refined. One of the emerging trends is using vertical search intent. For example, a user may want to find a restaurant near her current location; another user may want to follow the recent development of a breaking event such as the earthquake in Japan. Some recent studies show that at least 20% of Web queries have some local intent [3]. As a result, vertical search engines start attracting more and more attention. For example, many search engines provide specialized vertical search results for local search<sup>1</sup> and for real-time search<sup>2</sup>. Furthermore, vertical search results are often slotted into general Web search results [1, 2]. Thus, designing effective ranking functions for vertical search has become practically important to improve users' search experience.

A natural way to build a vertical search engine is to apply the existing ranking techniques on a vertical. In the TREC conference, several specialized tracks such as blog and chemical tracks have been introduced to provide a test bed to study retrieval tasks on these special text collections. The main focus of these tracks is on content-based relevance and most participants extend traditional IR techniques to consider a few task-specific ranking signals. Recently, learning to rank approaches have been studied extensively and shown to be effective to combine many useful signals into a ranking function. To adapt such a technique on a vertical, an intuitive approach is to construct a training data set by collecting a set of queries and documents which belong to the vertical and asking human editors to give a single relevance label between a query and a document. A ranking function thus can be learned for the vertical.

<sup>1</sup>Google Local. <http://local.google.com/>.

<sup>2</sup>Cinicalkey. <https://www.clinicalkey.com/>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).

WSDM '15, February 2–6, 2015, Shanghai, China.

ACM 978-1-4503-3317-7/15/02.

<http://dx.doi.org/10.1145/2684822.2697031>.

However, we observed that in many verticals, the meaning of relevance is domain-specific and usually consists of multiple well-defined aspects. Blindly applying the conventional learning to rank approaches by ignoring vertical-specific domain knowledge may not be effective. Thus, we have identified a list of challenging research issues in the field of relevance for vertical search, which mainly fall into two categories. First category includes how to learn an effective ranking model considering multi-facets relevance: (1) Since there are several pertinent aspects in a vertical, human editors naturally need to consider and tradeoff the relevance from different aspects before making the overall relevance judgment. Thus assessing aspect relevance is a necessary step. (2) Trading off multiple aspects is not trivial since such a tradeoff can vary for different queries or in different contexts. (3) For different verticals, different aspects are involved and the tradeoff among aspects is vertical-dependent. Collecting training data with overall relevance for a new vertical requires human editors learn how to appropriately tradeoff different aspects. Second category focuses on building effective business model in the context of specific vertical search systems.

**Objectives:** This workshop is dedicated to presentations and discussions on relevance for vertical search. For the workshop, we aim to bring together researchers from IR, ML, NLP, and other areas of computer and information science who are working on or interested in this area, and provide a forum for them to identify the issues and the challenges, to share their latest research results, to express a diverse range of opinions about this topic, and to discuss future directions.

**Review Process:** The call for papers solicited submissions in the area of vertical search. Themes of interests were diverse and spanned technologies and applications, models, algorithms, and evaluation methods. Each submission was reviewed by at least two Program Committee members and final decisions were made by the workshop chairs. PC members are listed below and include many experts in the field. We are grateful to all of them for their thorough and insightful reviews.

**Program:** In this year's workshop, we are delighted to have several invited talks in both academic and industrial tracks as part of the workshop. Min Zhang from Tsinghua University, Jun Xu from Chinese Academy of Sciences, Ya Zhang from Shanghai Jiaotong University will give the invited talks in the academic track. Toru Shimizu from Yahoo Japan, Hao Ma from Microsoft Research, Tzu-Chiang Liou from Yahoo Taiwan and Jie Luo from Yahoo Labs will give invited talks in the industrial track.

We hope that you will find this workshop interesting and thought-provoking and that the workshop will provide you with a valuable opportunity to share ideas with other researchers and practitioners from institutions around the world.

## 1. REFERENCES

- [1] J. Arguello, F. Diaz, J. Callan, and J.-F. Crespo. Sources of evidence for vertical selection. In *Proceedings of the 32Nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '09, pages 315–322, New York, NY, USA, 2009. ACM.
- [2] J. Arguello, F. Diaz, and J.-F. Paiement. Vertical selection in the presence of unlabeled verticals. In *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '10, pages 691–698, New York, NY, USA, 2010. ACM.
- [3] P. Venetis, H. Gonzalez, C. S. Jensen, and A. Halevy. Hyper-local, directions-based ranking of places. *Proc. VLDB Endow.*, 4(5):290–301, Feb. 2011.