

October 28, 2013
San Francisco, CA, USA



Association for
Computing Machinery

Advancing Computing as a Science & Profession



DUBMOD'13

Proceedings of the 2013 Workshop on
**Data-driven User Behavioral Modelling
and Mining from Social Media**

Co-located with:

CIKM 2013

Sponsored by:

ACM SIGIR & ACM SIGWEB



**Association for
Computing Machinery**

Advancing Computing as a Science & Profession

**The Association for Computing Machinery
2 Penn Plaza, Suite 701
New York, New York 10121-0701**

Copyright © 2013 by the Association for Computing Machinery, Inc. (ACM). Permission to make digital or hard copies of portions 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. Copyright for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permission to republish from: permissions@acm.org or Fax +1 (212) 869-0481.

For other copying of articles that carry a code at the bottom of the first or last page, copying is permitted provided that the per-copy fee indicated in the code is paid through www.copyright.com.

Notice to Past Authors of ACM-Published Articles

ACM intends to create a complete electronic archive of all articles and/or other material previously published by ACM. If you have written a work that has been previously published by ACM in any journal or conference proceedings prior to 1978, or any SIG Newsletter at any time, and you do NOT want this work to appear in the ACM Digital Library, please inform permissions@acm.org, stating the title of the work, the author(s), and where and when published.

ISBN: 978-1-4503-2417-5

Additional copies may be ordered prepaid from:

ACM Order Department

PO Box 30777

New York, NY 10087-0777, USA

Phone: 1-800-342-6626 (USA and Canada)

+1-212-626-0500 (Global)

Fax: +1-212-944-1318

E-mail: acmhelp@acm.org

Hours of Operation: 8:30 am – 4:30 pm ET

Printed in the USA

DUBMOD 2013 Chairs' Welcome

It is our great pleasure to welcome you to the workshop on *Data-driven user behavioral modeling and mining from social media (DUBMOD)*, co-located with 22nd *ACM Conference on Information and Knowledge Management – CIKM'13*. This workshop builds upon the success of our previously held workshops *User modeling from social media (UMSOCIAL)*, in conjunction with ACM IUI 2012 and *Data-driven user behavioral modeling and mining from social media (DUBMMSM)*, in conjunction with ACM CIKM 2012. The call for papers attracted 12 submissions. We accepted 8 papers. The papers represent various topics such as reaction time for user behavior model for social network, user profile for answer quality, predicting requests for Wikipedia articles, linking Pinterest pins to online webshops, scalable mining of social data, predicting online buying behavior, flickr group recommendation. We hope that these workshop proceedings will serve as a valuable reference for researchers and practitioners from diverse areas, such as user modeling, social media analysis, natural language processing, data mining and machine learning.

Putting together *DUBMOD'13* was a team effort. We thank the authors for providing the content of the program. We thank to our program committee who put lot of effort in reviewing papers and providing feedback to the authors. We hope that you will find this program 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.

**Jalal Mahmud, James Caverlee, Jeffrey Nichols,
John O'Donovan, Michelle Zhou**
DUBMOD'13 Chairs

Table of Contents

DUBMOD 2013 Workshop Organization	vi
--	----

Paper Session 1 (Ranking and Recommendations)

Session Chair: Jalal Mahmud (*IBM Research - Almaden*)

• Session Modeling to Predict Online Buyer Behavior	1
Andrii Cherniak (<i>University of Pittsburgh</i>), Jesse Bridgewater (<i>eBay Inc.</i>)	
• Twitter Anticipates Bursts of Requests for Wikipedia Articles	5
Gabriele Tolomei, Salvatore Orlando (<i>Università Ca' Foscari Venezia</i>), Diego Ceccarelli, Claudio Lucchese (<i>I.S.T.I. "A. Faedo" & CNR</i>)	
• I Pinned It. Where Can I Buy One Like It? Automatically Linking Pinterest Pins to Online Webshops	9
Susana Zoghbi, Ivan Vulić, Marie-Francine Moens (<i>KU Leuven</i>)	
• Content-Boosted Maximum Margin Matrix Factorization for Flickr Group Recommendation	13
Yilun Wang, Liang Chen, Jian Wu (<i>Zhejiang University</i>)	

Paper Session 2 (Mining & Analysis)

Session Chair: James Caverlee (*Texas A&M University*)

• Reaction Times for User Behavior Models in Microblogging Online Social Networks	17
Samuel Martins Barbosa Neto, Maira Athanzio De Cerqueira Gatti, Paulo Rodrigo Cavalin, Claudio Santos Pinhanez, Cicero Nogueira Dos Santos, Ana Paula Appel (<i>IBM Research Brazil</i>)	
• Scalable Mining of Social Data Using Stochastic Gradient Fisher Scoring	21
Jeon-Hyung Kang, Kristina Lerman (<i>University of Southern California</i>)	
• User Profiling for Answer Quality Assessment in Q&A Communities	25
Alexandru Lucian Ginsca , Adrian Popescu (<i>CEA, LIST, Vision & Content Engineering Laboratory</i>)	
• An Analysis of Question Quality and User Performance in Crowdsourced Exams	29
Sarah K. K. Luger (<i>The University of Edinburgh</i>), Jeff Bowles (<i>University of New Mexico</i>)	

Author Index	33
---------------------------	----

DUBMOD 2013 Workshop Organization

Chairs: Jalal Mahmud, IBM Research – Almaden
James Caverlee, Texas A&M University
Jeffrey Nichols, IBM Research – Almaden
John O’ Donovan, University of California, Santa Barbara
Michelle Zhou, IBM Research – Almaden

Program Committee: Jilin Chen, IBM Research, Almaden
Kyumin Lee, Utah State University
Krishna Kamath, Twitter
Jianqiang Shen, PARC
Aditya Pal, IBM Research, Almaden
Jesse Vig, PARC
Byungkyu Kang, University of California, Santa Barbara
Munmun De Choudhury, Microsoft Research
Huahai Yang, IBM Research, Almaden
Eduardo Graells Garrido, Universitat Pompeu Fabra
Ke Tao, TU Delft

Sponsors:

SIGIR
Special Interest Group
on Information Retrieval

