- CIKM'12 Main Conference Program: p2-p36

- CIKM'12 Workshop Program: p37-52

CIKM'12 Main Conference Program

Tuesday, October 30, 2012

Conference Opening (8:10 – 8:30)
Room: Maui Ballroom
Welcome: Conference Chair
Program introduction: PC Chairs

Keynote Speech (8:30 – 9:30)
Chair:
Room: Maui Ballroom
Title: User Engagement: The Network Effect Matters!
Speaker: Ricardo Baeza-Yates, Yahoo! Research

Coffee Break (9:30 – 10:15)

Session 1 (10:15 – 12:20)
KM Track: Recommender Systems
Chair:
Room: Wailuku
- LogUCB: An Explore-Exploit Algorithm For Comments Recommendation
  Dhruv Kumar Mahajan, Rajeev Rastogi, Charu Tiwari, Adway Mitra
- DQR: A Probabilistic Approach to Diversified Query Recommendation
  Ruirui Li, Ben Kao, Bin Bi, Reynold Cheng, Eric Lo
- Dynamic Covering for Recommendation Systems
  Ioannis Antonellis, Anish Das Sarma, Shaddin Dughmi
- MEET: A Generalized Framework for Reciprocal Recommender Systems
  Lei Li, Tao Li
- Social Contextual Recommendation
  Meng Jiang, Peng Cui, Rui Liu, Qiang Yang, Fei Wang, Wenwu Zhu, Shiqiang Yang

Session 2 (10:15 – 12:20)
KM Track: Pattern Mining
Chair:
Room: Kahului
- Mining High Utility Itemsets without Candidate Generation
  Mengchi Liu, Junfeng Qu
- A General Framework to Encode Heterogeneous Information Sources for Contextual Pattern Mining
  Weishan Dong, Wei Fan, Lei Shi, Changjin Zhou, Xifeng Yan
- Incorporating Occupancy into Frequent Pattern Mining for High Quality Pattern Recommendation
  Linpeng Tang, Lei Zhang, Ping Luo, Min Wang
- PARMA: A Parallel Randomized Algorithm for Approximate Association Rules Mining in MapReduce

Matteo Riondato, Justin A DeBrabant, Rodrigo Fonseca, Eli Upfal
- Interactive Pattern Mining on Hidden Data: A Sampling-based Solution
  Mansurul Bhuiyan, Snehasis Mukhopadhyay, Mohammad Al Hasan

Session 3 (10:15 – 12:20)
IR Track: Evaluation Methodologies
Chair:
Room: Kihei and Wailea
- An Analysis of Systematic Judging Errors in Information Retrieval
  Gabriella Kazai, Nick Craswell, Emine Yilmaz, S.M.M Tahaghoghi
- On Caption Bias in Interleaving Experiments
  Katja Hofmann, Fritz Behr, Filip Radlinski
- Alternative assessor disagreement and retrieval depth
  William Webber, Praveen Chandar, Ben Carterette
- Incorporating Variability in User Behavior into Systems Based Evaluation
  Ben Carterette, Evangelos Kanoulas, Emine Yilmaz
- Constructing Test Collections by Inferring Document Relevance via Extracted Relevant Information
  Shahzad Rajput, Matthew Ekstrand-Abueg, Virgil Pavlu, Javed A. Aslam

Session 4 (10:15 – 12:20)
IR Track: Social Media Search
Chair:
Room: Kapalua
- Twevent: Segment-based Event Detection from Tweets
  Chenliang Li, Aixin Sun, Anwitaman Datta
- Making Your Interests Follow You on Twitter
  Marco Pennacchiotti, Fabrizio Silvestri, Hossein Vahabi, Rossano Venturini
- Generating Event Storylines from Microblogs
  Chen Lin, Chun Lin, Jingxuan Li, Dingding Wang, Yang Chen, Tao Li
- Social Book Search: Comparings Topical Relevance Judgements and Book Suggestions for Evaluation
  Marijn Koolen, Jaap Kamps, Gabriella Kazai
- Content-Based Crowd Retrieval on the Real-Time Web
  Krishna Y Kamath, James Caverlee

Short Paper Session S1 (10:15 – 12:20)
KM Track: Text/Web Mining
Chair:
Room: Napili
- Automatically Embedding Newsworthy Links to Articles
  Hakan Ceylan, Ioannis Arapakis, Pinar Donmez, Mounia Lalmas
- Feature Selection Based on Term Frequency and T-Test for Text Categorization
  Deqing Wang, Hui Zhang, Rui Liu, Weifeng Lv
- Extraction of Topic Evolutions from References in Scientific Articles and Its GPU Acceleration
  Tomonari Masada, Atsuhiro Takasu
- Reconciling Ontologies and the Web of Data

Ziawasch Abedjan, Johannes Lorey, Felix Naumann
- Exploiting Latent Relevance for Relational Learning of Ubiquitous Things
  Lina Yao, Quan Z. Sheng
- Mining Coherent Anomaly Collections On Web Data
  Hanbo Dai, Feida Zhu, Ee-Peng Lim, HweeHwa Pang
- Mining Topic-level Opinion Influence in Microblog
  Daifeng Li, Xin Shuai, Guozheng Sun, Jie Tang, Ying Ding, Zhipeng Luo
- Exploiting Enriched Contextual Information for Mobile App Classification
  Hengshu Zhu, Huanhuan Cao, Enhong Chen, Hui Xiong, Jilei Tian
- Incorporating Word Correlation into Tag-Topic Model for Semantic Knowledge Acquisition
  Fang Li, Tingting He, Xinhui Tu, Xiaohua Hu
- PriSM: Discovering and Prioritizing Severe Technical Issues from Product Discussion Forums
  Rashmi Gangadharaiah, Rose Catherine
- Community-Based Classification of Noun Phrases in Twitter
  Freddy Chong Tat Chua, William W Cohen, Justin Betteridge, Ee-Peng Lim
- Joint Bilingual Name Tagging for Parallel Corpora
  Qi Li, Haibo Li, Heng Ji, Wen Wang, Jing Zheng, Fei Huang

Short Paper Session S2 (10:15 – 12:20)
KM Track: Networks and Graphs
Chair:
Room: Kula and Hana
- Influence and Similarity on Heterogeneous Networks
  Guan Wang, Qingbo Hu, Philip S. Yu
- GRAFT: An Approximate Graphlet Counting Algorithm for Large Graph Analysis
  Mahmudur Rahman, Mansurul Bhuiyan, Mohammad Al Hasan
- Fast Approximation of Steiner Trees in Large Graphs
  Andrey Gubichev, Thomas Neumann
- Measuring Robustness of Complex Networks under MVC Attack
  Rong-Hua Li, Jeffrey Xu Yu, Xin Huang, Hong Cheng, Zechao Shang
- Meta Path-Based Collective Classification in Heterogeneous Information Networks
  Xiangnan Kong, Philip S. Yu, Ying Ding, David J. Wild
- Discretionary Social Network Data Revelation with a User-Centric Utility Guarantee
  Yi Song, Panagiotis Karras, Sadegh Nobari, Giorgos Cheliotis, Mingqiang Xue, Stephane Bressan
- Empirical Validation of the Buckley--Osthus Model for the Web Host Graph: Degree and Edge Distributions
  Maxim Zhukovskiy, Dmitry Vinogradov, Yuri Pritykin, Liudmila Ostroumova, Evgeniy Grechnikov, Gleb Gusev, Pavel Serdyukov, Andrei Raigorodskii
- gSCorr: Modeling Geo-Social Correlations for New Check-ins on Location-Based Social Networks
  Huiji Gao, Jiliang Tang, Huan Liu
- Unsupervised Discovery of Opposing Opinion Networks From Forum Discussions
  Yue Lu, Hongning Wang, ChengXiang Zhai, Dan Roth
- WiSeNet: Building a Wikipedia-based Semantic Network with Ontologized Relations
  Andrea Moro, Roberto Navigli

- Shaping Communities out of Triangles
  Arnau Prat-PÈrez, David Dominguez-Sal, Josep M Brunat, Josep-Lluis Larriba-Pey
- Degree Relations of Triangles in Real-world Networks and Graph Models
  Nurcan Durak, Ali Pinar, Tamara G. Kolda, C. Seshadhri

Lunch Provided by the Conference (12:20 – 01:30)

Session 5 (1:30 – 3:35)
KM Track: Link and Graph Mining
Chair:
Room: Wailuku
- Graph Classification: A Diversified Discriminative FeatureSelection Approach
  Yuanyuan Zhu, Jeffrey Xu Yu, Hong Cheng, Lu Qin
- Multi-Scale Link Prediction
  Donghyuk Shin, Si Si, Inderjit S Dhillon
- An Analysis of How Ensembles of Collective Classifiers Improve Predictions in Graphs
  Hoda Eldardiry, Jennifer Neville
- Density Index and Proximity Search in Large Graphs
  Nan Li, Xifeng Yan, Zhen Wen, Arijit Khan
- Gelling, and Melting, Large Graphs by Edge Manipulation
  Hanghang Tong, B. Aditya Prakash, Tina Eliassi-Rad, Michalis Faloutsos, Christos Faloutsos

Session 6 (1:30 – 3:35)
IR Track: Language Technologies
Chair:
Room: Kahului
- One Seed to Find Them All: Mining Opinion Features via Association
  Zhen Hai, Kuiyu Chang, Gao Cong
- Topic-Driven Reader Comments Summarization
  Zongyang Ma, Aixin Sun, Quan Yuan, Gao Cong
- Visualizing Timelines: Evolutionary Summarization via Iterative Reinforcement between Text and Image Streams
  Rui Yan, Xiaojun Wan, Mirella Lapata, Wayne Xin Zhao, Pu-Jen Cheng, Xiaoming Li
- Fast Multi-task Learning for Query Spelling Correction
  Xu Sun, Anshumali Shrivastava, Ping Li
- Cross-Argument Inference for Implicit Discourse Relation Recognition
  Yu Hong, Xiaopei Zhou, Tingting Che, Jianmin Yao, Qiaoming Zhu, Guodong Zhou

Session 7 (1:30 – 3:35)
DB Track: Graph and Knowledge Base
Chair:
Room: Kihei and Wailea
- Interpreting Keyword Queries over Web Knowledge Bases
  Jeffrey Pound, Alexander K Hudek, Ihab F Ilyas, Grant Weddell
- RDF Pattern Matching using Sortable Views

Zhihong Chong, He Chen, Zhenjie Zhang, Hu Shu, Guilin Qi, Aoying Zhou
- Efficient Algorithms for Generalized Subgraph Query Processing
  Wenqing Lin, Xiaokui Xiao, James Cheng, Sourav S Bhowmick
- G-SPARQL: A Hybrid Engine for Querying Large Attributed Graphs
  Sherif Sakr, Sameh Elnikety, Yuxiong He
- A Graph-Based Approach for Ontology Population with Named Entities
  Wei Shen, Jianyong Wang, Ping Luo, Min Wang

Session 8 (1:30 – 3:35)
DB Track: Temporal, Spatial, and Multimedia Databases
Chair:
Room: Kapalua
- Decomposition-by-Normalization (DBN): Leveraging Approximate Functional Dependencies for Efficient Tensor Decomposition
  Mijung Kim, K. SelÁuk Candan
- A Filter-based Protocol for Continuous Queries over Imprecise Location Data
  Yifan Jin, Reynold Cheng, Ben Kao, Kam-Yiu Lam, Yinuo Zhang
- Leveraging Read Rates of Passive RFID Tags for Real-Time Indoor Location Tracking
  Da Yan, Zhou Zhao, Wilfred Ng
- Location-Aware Instant Search
  Ruicheng Zhong, Ju Fan, Guoliang Li, Kian-Lee Tan, Lizhu Zhou
- Indexing Uncertain Spatio-Temporal Data
  Tobias Emrich, Hans-Peter Kriegel, Nikos Mamoulis, Matthias Renz, Andreas Z˛fle

Short Paper Session S3 (1:30 – 3:35)
KM Track: Recommendation and Summary
Chair:
Room: Napili
- A Simple Approach to the Design of Site-Level Extractors Using Domain-Centric Principles
  Chong Long, Xiubo Geng, Chang Xu, Sathiya Keerthi
- Graph-Based Workflow Recommendation: On Improving Business Process Modeling
  Bin Cao, Jianwei Yin, Shuiguang Deng, Dongjing Wang, Zhaohui Wu
- What is Happening Right Now ... That Interests Me?
  Ernesto Diaz-Aviles, Lucas Drumond, Zeno Gantner, Lars Schmidt-Thieme, Wolfgang Nejdl
- PRemiSE: Personalized News Recommendation via Implicit Social Experts
  Chen Lin, Runquan Xie, Lei Li, Zhenhua Huang, Tao Li
- Time-aware Topic Recommendation Based on Micro-blogs
  Huizhi Liang, Yue Xu, Dian Tjondronegoro, Peter Christen
- Topic-Sensitive Probabilistic Model for Expert Finding in Question Answer Communities
  Guangyou Zhou, Siwei Lai, Kang Liu, Jun Zhao
- The Early-Adopter Graph and its Application to Web-Page Recommendation
  Ida Mele, Francesco Bonchi, Aristides Gionis
- Real-Time Bid Optimization for Group-Buying Ads
  Raju Balakrishnan, Rushi P Bhatt

- A Probabilistic Approach to Mining Geospatial Knowledge from Social Annotations
  Suradej Intagorn, Kristina Lerman
- Providing Grades and Feedback for Student Summaries by Ontology-based Information Extraction
  Fernando Gutierrez, Dejing Dou, Stephen Fickas, Gina Griffiths
- Using Program Synthesis for Social Recommendations
  Alvin Cheung, Armando Solar-Lezama, Samuel Madden
- Web-Scale Multi-Task Feature Selection for Behavioral
  Targeting Amr Ahmed, Mohamed Aly, Abhimanyu Das, Alexander J Smola, Tasos Anastasakos
- Dynamic Effects of Ad Impressions on Commercial Actions in Display Advertising
  Joel Barajas, Ram Akella, Marius Holtan, Jaimie Kwon, Aaron Flores, Victor Andrei

Short Paper Session S4 (1:30 – 3:35)
IR Track: Web Search
Chair:
Room: Kula and Hana

- Content-Based Relevance Estimation on the Web Using Inter-Document Similarities
  Fiana Raiber, Oren Kurland, Moshe Tennenholtz
- Estimating Interleaved Comparison Outcomes from Historical Click Data
  Katja Hofmann, Shimon Whiteson, Maarten de Rijke
- Ranking News Events by Influence Decay and Information Fusion for Media and Users
  Liang Kong, Shan Jiang, Rui Yan, Shize Xu, Yan Zhang
- Leveraging Tagging for Neighborhood-aware Probabilistic Matrix Factorization
  Le Wu, Enhong Chen, Qi Liu, Linli Xu, Tengfei Bao, Lei Zhang
- Federated Search in the Wild
  Dong Nguyen, Thomas Demeester, Dolf Trieschnigg, Djoerd Hiemstra
- Task Tours: Helping Users Tackle Complex Search Tasks
  Ahmed Hassan, Ryen W White
- Structured Query Reformulations in Commerce Search
  Sreenivas Gollapudi, Samuel Ieong, Anitha Kannan
- Characterizing Web Search Queries that Match Very Few or No Results
  Ismail Sengor Altingovde, Roi Blanco, Berkant Barla Cambazoglu, Rifat Ozcan, Erdem Sarigil, ÷zg,r Ulusoy
- The Downside of Markup: Examining the Harmful Effects of CSS and Javascript on Indexing Today's Web
  Karl Gyllstrom, Carsten Eickhoff, Arjen P. de Vries, Marie-Francine Moens
- A Unified Optimization Framework for Auction and Guaranteed Delivery in Online Advertising
  Konstantin Salomatin, Tie-Yan Liu, Yiming Yang
- Sentiment-Focused Web Crawling
  Gural Vural, B. Barla Cambazoglu, Pinar Senkul
- GTE: A Distributional Second-Order Co-Occurrence Approach to Improve the Identification of Top Relevant Dates in Web Snippets
  Ricardo Campos, GaÎl Dias, AlÌpio Jorge, CÈlia Nunes

Coffee Break (3:34 – 4:00)

Session 9 (4:00 – 5:40)
KM Track: Matrix Methods and Anomaly Detection
Chair:
Room: Wailuku

- Local Anomaly Descriptor: A Robust Unsupervised Algorithm for Anomaly Detection based on Diffusion Space
  Hao Huang, Hong Qin, Shinjae Yoo, Dantong Yu
- Fast and Reliable Anomaly Detection in Categorical Data
  Leman Akoglu, Hanghang Tong, Jilles Vreeken, Christos Faloutsos
- TALMUD Transfer Learning for Multiple Domains
  Orly Moreno, Bracha Shapira, Lior Rokach, Guy Shani
- Utilizing Common Substructures to Speedup Tensor Factorization for Mining Dynamic Graphs
  Wei Liu, Jeffrey Chan, James Bailey, Christopher Leckie, Ramamohanarao Kotagiri

Session 10 (4:00 – 5:40)
KM Track: Social Networks
Chair:
Room: Kahului

- Predicting Emerging Social Conventions in Online Social Networks
  Farshad Kooti, Winter A. Mason, Krishna P. Gummadi, Meeyoung Cha
- Collective Intelligence in the Online Social Network of Yahoo!Answers and Its Implications
  Ze Li, Haiying Shen, Joseph Edward Grant
- From Face-to-Face Gathering To Social Structure
  Chunyan Wang, Mao Ye, Wang-chien Lee
- Delineating Social Network Data Anonymization via Random Edge Perturbation
  Mingqiang Xue, Panagiotis Karras, Raissi Chedy, Panos Kalnis, Hung Keng Pung

Session 11 (4:00 – 5:40)
IR Track: Advertising
Chair:
Room: Kihei and Wailea

- Multiview Hierarchical Bayesian Regression Model andApplication to Online Advertising
  Tianbing Xu, Ruofei Zhang, Zhen Guo
- Visual Appearance of Display Ads and Its Effect on Click Through Rate
  Javad Azimi, Ruofei Zhang, Yang Zhou, Vidhya Navalpakkam, Jianchang Mao, Xiaoli Fern
- The Wisdom of Advertisers: Mining Subgoals via Query Clustering
  Takehiro Yamamoto, Tetsuya Sakai, Mayu Iwata, Chen Yu, Ji-Rong Wen, Katsumi Tanaka
- Sequential Selection of Correlated Ads by POMDPs
  Shuai Yuan, Jun Wang

Session 12 (4:00 – 5:40)
IR Track: System Architecture, Distributed IR, Scalability
Chair:

Room: Kapalua
- Diversity in Blog Feed Retrieval
  Mostafa Keikha, Fabio Crestani, Bruce W Croft
- Efficient Retrieval of Recommendations in a Matrix Factorization Framework
  Noam Koenigstein, Parikshit Ram, Yuval Shavitt
- KORE: Keyphrase Overlap Relatedness for Entity Disambiguation
  Johannes Hoffart, Stephan Seufert, Dat Ba Nguyen, Martin Theobald, Gerhard Weikum
- Shard Ranking and Cutoff Estimation for Topically Partitioned Collections
  Anagha Kulkarni, Almer S. Tigelaar, Djoerd Hiemstra, Jamie Callan

Short Paper Session S5 (4:00 – 5:40)
DB Track: Search, Retrieval and Big Data
Chair:
Room: Napili
- Top-k Retrieval Using Conditional Preference Networks
  Hongbing Wang, Xuan Zhou, Wujin Chen, Peisheng Ma
- LINDA: Distributed Web-of-Data-Scale Entity Matching
  Christoph Bˆhm, Gerard de Melo, Felix Naumann, Gerhard Weikum
- Finding the Optimal Path over Multi-Cost Graphs
  Yajun Yang, Jeffrey Xu Yu, Hong Gao, Jianzhong Li
- CloST: A Hadoop-based Storage System for Big Spatio-Temporal Data Analytics
  Haoyu Tan, Wuman Luo, Lionel M. Ni
- Loyalty-based Selection: Retrieving Objects That Persistently Satisfy Criteria
  Zhitao Shen, Muhammad Aamir Cheema, Xuemin Lin
- Optimizing Data Migration for Cloud-based Key-value Stores
  Xiulei Qin, Wenbo Zhang, Wei Wang, Jun Wei, Xin Zhao, and Tao Huang
- A New Tool for Multi-Level Partitioning in Teradata
  Young-Kyoon Suh, Ahmad Ghazal, Alain Crolotte, Pekka Kostamaa

Short Paper Session S6 (4:00 – 5:40)
DB Track: Query and Indexing
Chair:
Room: Kula and Hana
- Sort-based Query-adaptive Loading of R-trees
  Daniar Achakeev, Bernhard Seeger, Peter Widmayer
- Diversifying Query Results on Semi-Structured Data
  Mahbub Hasan, Abdullah Mueen, Vassilis Tsotras, Eamonn Keogh
- An Efficient Index for Massive IOT Data in Cloud Environment
  Youzhong Ma, Jia Rao, Weisong Hu, Xiaofeng Meng, Xu Han, Yu Zhang, Yunpeng Chai, Chunqiu Liu
- Impact Neighborhood Indexing (INI) in Diffusion Graphs
  Jung Hyun Kim, K. Selcuk Candan, Maria Luisa Sapino
- Applying Weighted Queries on Probabilistic Databases
  Sebastian Lehrack
- Fast PCA Computation in a DBMS with Aggregate UDFs and LAPACK

Carlos Ordonez, Naveen Mohanam, Carlos Garcia-Alvarado, Predrag T. Tosic, Edgar Martinez
- A Probabilistic Approach to Correlation Queries in Uncertain Time Series Data
  Mahsa Orang, Nematollaah Shiri

Short-paper Posters (6:40 – 9:00)
Room: Maui Ballroom

KM Track
- Influence and Similarity on Heterogeneous Networks
  Guan Wang, Qingbo Hu, Philip S. Yu
- GRAFT: An Approximate Graphlet Counting Algorithm for Large Graph Analysis
  Mahmudur Rahman, Mansurul Bhuiyan, Mohammad Al Hasan
- Mining Long-lasting Exploratory User Interests from Search History
  Bin Tan, Yuanhua Lv, ChengXiang Zhai
- Fast Approximation of Steiner Trees in Large Graphs
  Andrey Gubichev, Thomas Neumann
- Automatically Embedding Newsworthy Links to Articles
  Hakan Ceylan, Ioannis Arapakis, Pinar Donmez, Mounia Lalmas
- A Simple Approach to the Design of Site-Level Extractors Using Domain-Centric Principles
  Chong Long, Xiubo Geng, Chang Xu, Sathiya Keerthi
- Reconciling Ontologies and the Web of Data
  Ziawasch Abedjan, Johannes Lorey, Felix Naumann
- Efficient Extraction of Ontologies from Domain Specific Text Corpora
  Tianyu Li, Pirooz Chubak, Laks V.S. Lakshmanan, Rachel Pottinger
- Effective and Efficient? Bilingual Sentiment Lexicon Extraction using Collocation Alignment
  Zheng Lin, Songbo Tan, Xueqi Cheng, Xueke Xu, Weisong Shi
- Exploiting Latent Relevance for Relational Learning of Ubiquitous Things
  Lina Yao, Quan Z. Sheng
- Discovering Personally Semantic Places from GPS Trajectories
  Mingqi Lv, Ling Chen, Gencai Chen
- Mining Coherent Anomaly Collections On Web Data
  Hanbo Dai, Feida Zhu, Ee-Peng Lim, HweeHwa Pang
- Meta Path-Based Collective Classification in Heterogeneous Information Networks
  Xiangnan Kong, Philip S. Yu, Ying Ding, David J. Wild
- Discretionary Social Network Data Revelation with a User-Centric Utility Guarantee
  Yi Song, Panagiotis Karras, Sadegh Nobari, Giorgos Cheliotis, Mingqiang Xue, Stephane Bressan
- Empirical Validation of the Buckley--Osthus Model for the Web Host Graph: Degree and Edge Distributions
  Maxim Zhukovskiy, Dmitry Vinogradov, Yuri Pritykin, Liudmila Ostroumova, Evgeniy Grechnikov, Gleb Gusev, Pavel Serdyukov, Andrei Raigorodskii
- gSCorr: Modeling Geo-Social Correlations for New Check-ins on Location-Based Social Networks
  Huiji Gao, Jiliang Tang, Huan Liu
- Swimming against the Streamz: Search and Analytics over the Enterprise Activity Stream
  Ido Guy, Tal Steier, Maya Barnea, Inbal Ronen, Tal Daniel

- What is Happening Right Now ... That Interests Me?
  Ernesto Diaz-Aviles, Lucas Drumond, Zeno Gantner, Lars Schmidt-Thieme, Wolfgang Nejdl
- Frequent grams based Embedding for Privacy Preserving Record
  Linkage Luca Bonomi, Li Xiong, Rui Chen, Benjamin C. M. Fung
- If You are Happy and You Know It... Tweet
  Amir Asiaee T., Mariano Tepper, Arindam Banerjee, Guillermo Sapiro
- Hierarchical Topic Integration Through Semi-supervised Hierarchical Topic Modeling
  Xian-Ling Mao, Jing He, Hongfei Yan, Xiaoming Li
- PriSM: Discovering and Prioritizing Severe Technical Issues from Product Discussion Forums
  Rashmi Gangadharaiah, Rose Catherine
- Preprocessing of Informal Mathematical Discourse in Context ofControlled Natural Language
  Ra˙l Ernesto GutiÈrrez de PiÒerez Reyes, Juan Francisco DÌaz FrÌas
- PathRank: A Novel Node Ranking Measure on a Heterogeneous Graph for Recommender Systems
  Sangkeun Lee, Sungchan Park, Minsuk Kahng, Sang-goo Lee
- Exploring the Existing Category Hierarchy to Automatically Label the Newly-arising Topics in cQA
  Guangyou Zhou, Li Cai, Kang Liu, Jun Zhao
- Query-Focused Multi-document Summarization Based on Query-Sensitive Feature Space
  Wenpeng Yin, Yulong Pei, Fan Zhang, Lian'en Huang
- Time-aware Topic Recommendation Based on Micro-blogs
  Huizhi Liang, Yue Xu, Dian Tjondronegoro, Peter Christen
- iSampling: Framework for Developing Sampling Methods Considering User's Interest
  Jinoh Oh, Hwanjo Yu
- WiSeNet: Building a Wikipedia-based Semantic Network with Ontologized Relations
  Andrea Moro, Roberto Navigli
- Shaping Communities out of Triangles
  Arnau Prat-PÈrez, David Dominguez-Sal, Josep M Brunat, Josep-Lluis Larriba-Pey
- Relational Co-Clustering via Manifold Ensemble Learning
  Ping Li, Jiajun Bu, Chun Chen, Zhanying He
- SemaFor: Semantic Document Indexing using Semantic Forests
  George Tsatsaronis, Iraklis Varlamis, Kjetil N¯rvÂg
- Measuring Website Similarity using an Entity-Aware Click Graph
  Pablo N Mendes, Peter Mika, Hugo Zaragoza, Roi Blanco
- Community-Based Classification of Noun Phrases in Twitter
  Freddy Chong Tat Chua, William W Cohen, Justin Betteridge, Ee-Peng Lim
- Real-Time Bid Optimization for Group-Buying Ads
  Raju Balakrishnan, Rushi P Bhatt
- A Probabilistic Approach to Mining Geospatial Knowledge from Social Annotations
  Suradej Intagorn, Kristina Lerman
- Providing Grades and Feedback for Student Summaries by Ontology-based Information Extraction
  Fernando Gutierrez, Dejing Dou, Stephen Fickas, Gina Griffiths
- Joint Bilingual Name Tagging for Parallel Corpora
  Qi Li, Haibo Li, Heng Ji, Wen Wang, Jing Zheng, Fei Huang
- Using Program Synthesis for Social Recommendations
  Alvin Cheung, Armando Solar-Lezama, Samuel Madden

- Web-Scale Multi-Task Feature Selection for Behavioral
  Targeting Amr Ahmed, Mohamed Aly, Abhimanyu Das, Alexander J Smola, Tasos Anastasakos
- Balanced Coverage of Aspects for Text Summarization
  Takuya Makino, Hiroya Takamura, Manabu Okumura
- Dynamic Effects of Ad Impressions on Commercial Actions in Display Advertising
  Joel Barajas, Ram Akella, Marius Holtan, Jaimie Kwon, Aaron Flores, Victor Andrei
- A Hybrid Approach for Efficient Provenance Storage
  Yulai Xie, Dan Feng, Zhipeng Tan, Lei Chen, Kiran-Kumar Muniswamy-Reddy, Yan Li, Darrell D. E. Long

IR Track

- Content-Based Relevance Estimation on the Web Using Inter-Document Similarities
  Fiana Raiber, Oren Kurland, Moshe Tennenholtz
- Trust Prediction via Aggregating Heterogeneous Social Networks
  Jin Huang, Feiping Nie, Heng Huang, Yi-Cheng Tu
- Estimating Interleaved Comparison Outcomes from Historical Click Data
  Katja Hofmann, Shimon Whiteson, Maarten de Rijke
- Automatic Image Annotation Using Tag-Related Random Search over Visual Neighbors
  Zijia Lin, Guiguang Ding, Mingqing Hu, Jianmin Wang, Jiaguang Sun
- Diversionary Comments under Political Blog Posts
  Jing Wang, Clement T Yu, Philip S Yu, Bing Liu, Weiyi Meng
- Discover Breaking Events with Popular Hashtags in Twitter
  Anqi Cui, Min Zhang, Yiqun Liu, Shaoping Ma, Kuo Zhang
- Query Likelihood with Negative Query Generation
  Yuanhua Lv, ChengXiang Zhai
- On the Connections between Explicit Semantic Analysis and Latent Semantic Analysis
  Chao Liu, Yi-Min Wang
- Variance Maximization via Noise Injection for Active Sampling in Learning to Rank
  Wenbin Cai, Ya Zhang
- More Than Relevance: High Utility Query Recommendation By Mining Users' Search Behaviors
  Xiaofei Zhu, Jiafeng Guo, Xueqi Cheng, Yanyan Lan
- Finding Nuggets in IP Portfolios: Core Patent Mining through Textual Temporal Analysis
  Po Hu, Minlie Huang, Peng Xu, Weichang Li, Adam K Usadi, Xiaoyan Zhu
- Customizing Search Results for Non-Native Speakers
  Theodoros Lappas, Michail Vlachos
- Do Ads Compete or Collaborate? Designing Click Models with Full Relationship Incorporated
  Xin Xin, Irwin King, Ritesh Agrawal, Michael R. Lyu, Heyan Huang
- Ranking News Events by Influence Decay and Information Fusion for Media and Users
  Liang Kong, Shan Jiang, Rui Yan, Shize Xu, Yan Zhang
- Leveraging Tagging for Neighborhood-aware Probabilistic Matrix Factorization
  Le Wu, Enhong Chen, Qi Liu, Linli Xu, Tengfei Bao, Lei Zhang
- Semantic Context Learning with Large-Scale Weakly-Labeled Image Set
  Yao Lu, Wei Zhang, Ke Zhang, Xiangyang Xue
- Sketch-based Indexing of n-Words

Samuel Huston, J. Shane Culpepper, W. Bruce Croft

- Interactive and Context-Aware Tag Spell Check and Correction
  Francesco Bonchi, Ophir Frieder, Franco Maria Nardini, Fabrizio Silvestri, Hossein Vahabi
- Federated Search in the Wild
  Dong Nguyen, Thomas Demeester, Dolf Trieschnigg, Djoerd Hiemstra
- From sBoW to dCoT Marginalized Encoders for Text Representation
  Zhixiang (Eddie) Xu, Minmin Chen, Kilian Q. Weinberger
- Structured Query Reformulations in Commerce Search
  Sreenivas Gollapudi, Samuel Ieong, Anitha Kannan
- Towards Jointly Extracting Aspects and Aspect-Specific Sentiment Knowledge
  Xueke Xu, Songbo Tan, Yue Liu, Xueqi Cheng, Zheng Lin
- Collaborative Ranking: Improving the Relevance for Tail Queries
  Ke Zhou, Xin Li, Hongyuan Zha
- BiasTrust: Teaching Biased Users About Controversial Topics
  V.G.Vinod Vydiswaran, ChengXiang Zhai, Dan Roth, Peter Pirolli
- Recommending Citations: Translating Papers into References
  Wenyi Huang, Saurabh Kataria, Cornelia Caragea, Prasenjit Mitra, C. Lee Giles, Lior Rokach
- Discovering Logical Knowledge for Deep Question Answering
  Zhao Liu, Xipeng Qiu, Ling Cao, Xuanjing Huang
- Mining Noisy Tagging from Multi-label Space
  Zhongang Qi, Ming Yang, Zhongfei (Mark) Zhang, Zhengyou Zhang
- Learning from Mistakes: Towards a Correctable Learning Algorithm
  Karthik Raman, Krysta M Svore, Ran Gilad-Bachrach, Chris J.C. Burges
- Consento: A New Framework for Opinion Based Entity Search and Summarization
  Jaehoon Choi, Donghyeon Kim, Seongsoon Kim, Junkyu Lee, Sangrak Lim, Sunwon Lee, Jaewoo Kang
- Search Result Presentation Based on Faceted Clustering
  Benno Stein, Tim Gollub, Dennis Hoppe
- PolariCQ: Polarity Classification of Political Quotations
  Rawia Awadallah, Maya Ramanath, Gerhard Weikum
- A Comprehensive Analysis of Parameter Settings for Novelty-Biased Cumulative Gain
  Teerapong Leelanupab, Guido Zuccon, Joemon M. Jose
- Map to Humans and Reduce Error - Crowdsourcing for Deduplication Applied to Digital Libraries
  Mihai Georgescu, Dang Duc Pham, Claudiu S. Firan, Wolfgang Nejdl, Julien Gaugaz
- Full-Text Citation Analysis: Enhancing Bibliometric and Scientific Publication Ranking
  Xiaozhong Liu, Jinsong Zhang, Chun Guo
- Detecting Offensive Tweets via Topical Feature Discovery over a Large Scale Twitter Corpus
  Guang Xiang, Bin Fan, Ling Wang, Jason Hong, Carolyn Rose
- The Downside of Markup: Examining the Harmful Effects of CSS and Javascript on Indexing Today's Web
  Karl Gyllstrom, Carsten Eickhoff, Arjen P. de Vries, Marie-Francine Moens
- You Should Read This! Let Me Explain You Why
  Roi Blanco, Diego Ceccarelli, Claudio Lucchese, Raffaele Perego, Fabrizio Silvestri
- Characterizing Web Search Queries that Match Very Few or No Results

Ismail Sengor Altingovde, Roi Blanco, Berkant Barla Cambazoglu, Rifat Ozcan, Erdem Sarigil, ÷zg‚r Ulusoy

- A Unified Optimization Framework for Auction and Guaranteed Delivery in Online Advertising
  Konstantin Salomatin, Tie-Yan Liu, Yiming Yang
- Modeling Browsing Behavior for Click Analysis in Sponsored Search
  Azin Ashkan, Charles L. A. Clarke
- Sentiment-Focused Web Crawling
  Gural Vural, B. Barla Cambazoglu, Pinar Senkul
- User Guided Entity Similarity Search Using Meta-Path Selection in Heterogeneous Information Networks
  Xiao Yu, Yizhou Sun, Brandon Norick, Tiancheng Mao, Jiawei Han
- GTE: A Distributional Second-Order Co-Occurrence Approach to Improve the Identification of Top Relevant Dates in Web Snippets
  Ricardo Campos, GaÎl Dias, AlÌpio Jorge, CÈlia Nunes
- Stochastic Simulation of Time-Biased Gain
  Mark D. Smucker, Charles L. A. Clarke
- SonetRank: Leveraging Social Networks to Personalize Search
  Abhijith Kashyap, Reza Amini, Vagelis Hristidis
- Predicting Web Search Success with Fine-grained Interaction Data
  Qi Guo, Dmitry Lagun, Eugene Agichtein
- Multi-Session Re-Search: In Pursuit of Repetition and Diversification
  Sarah K Tyler, Yi Zhang
- Theme Chronicle Model: Chronicle Consists of Timestamp and Topical Words over Each Theme
  Noriaki Kawamae
- Fast Top-K Similarity Queries Via Matrix Compression
  Yucheng Low, Alice X Zheng

DB Track

- Sort-based Query-adaptive Loading of R-trees
  Daniar Achakeev, Bernhard Seeger, Peter Widmayer
- Schema-Free Structured Querying of DBpedia Data
  Lushan Han, Tim Finin, Anupam Joshi
- Discovering Conditional Inclusion Dependencies
  Jana Bauckmann, Ziawasch Abedjan, Ulf Leser, Heiko M¸ller, Felix Naumann
- Diversifying Query Results on Semi-Structured Data
  Mahbub Hasan, Abdullah Mueen, Vassilis Tsotras, Eamonn Keogh
- SliceSort: Efficient Sorting of Hierarchical Data
  Quoc Trung Tran, Chee-Yong Chan
- Efficient Buffer Management for Piecewise Linear Representation of Multiple Data Streams
  Qing Xie, Jia Zhu, Mohamed A. Sharaf, xiaofang zhou, Chaoyi Pang
- On Skyline Groups
  Chengkai Li, Nan Zhang, Naeemul Hassan, Sundaresan Rajasekaran, Gautam Das
- Finding the Optimal Path over Multi-Cost Graphs
  Yajun Yang, Jeffrey Xu Yu, Hong Gao, Jianzhong Li

- An Efficient Index for Massive IOT Data in Cloud Environment
  Youzhong Ma, Jia Rao, Weisong Hu, Xiaofeng Meng, Xu Han, Yu Zhang, Yunpeng Chai, Chunqiu Liu
- Clustering Wikipedia Infoboxes to Discover their Types
  Thanh Hoang Nguyen, Huong Dieu Nguyen, Viviane Moreira, Juliana Freire
- Keyword-based k-Nearest Neighbor Search in Spatial Databases
  Guoliang Li, Jing Xu, Jianhua Feng
- Credibility-Based Product Ranking for C2C Transactions
  Rong Zhang, Chao Feng Sha, Min qi Zhou, Ao ying Zhou
- Location Selection for Utility Maximization with Capacity Constraints
  Yu Sun, Jin Huang, Yueguo Chen, Rui Zhang, Xiaoyong Du
- Efficient Estimation of Dynamic Density Functions with an Application to Outlier Detection
  Abdulhakim Ali Qahtan, Xiangliang Zhang, Suojin Wang
- A Positional Access Method for Relational Databases
  Dongzhe Ma, Jianhua Feng, Guoliang Li
- Real-Time Aggregate Monitoring with Differential Privacy
  Liyue Fan, Li Xiong
- Efficient Distributed Locality Sensitive Hashing
  Bahman Bahmani, Ashish Goel, Rajendra Shinde
- Author-Conference Topic-Connection Model for Academic Network Search
  Jianwen Wang, Xiaohua Hu, Xinhui Tu, Tingting He
- Impact Neighborhood Indexing (INI) in Diffusion Graphs
  Jung Hyun Kim, K. Selcuk Candan, Maria Luisa Sapino
- Loyalty-based Selection: Retrieving Objects That Persistently Satisfy Criteria
  Zhitao Shen, Muhammad Aamir Cheema, Xuemin Lin
- Star-Join: Spatio-Textual Similarity Join
  Sitong Liu, Guoliang Li, Jianhua Feng
- Adapt: Adaptive Database Schema Design for Multi-Tenant Applications
  Jiacai Ni, Guoliang Li, Jun Zhang, Lei Li, Jianhua Feng
- A New Tool for Multi-Level Partitioning in Teradata
  Young-Kyoon Suh, Ahmad Ghazal, Alain Crolotte, Pekka Kostamaa
- Fast PCA Computation in a DBMS with Aggregate UDFs and LAPACK
  Carlos Ordonez, Naveen Mohanam, Carlos Garcia-Alvarado, Predrag T. Tosic, Edgar Martinez
- Scaling Multiple-Source Entity Resolution using Statistically Efficient Transfer Learning
  Sahand N Negahban, Benjamin I. P. Rubinstein, Jim Gemmell
- A Probabilistic Approach to Correlation Queries in Uncertain Time Series Data
  Mahsa Orang, Nematollaah Shiri
- On Bundle Configuration for Viral Marketing in Social Networks
  De-Nian Yang, Wang-Chien Lee, Nai-Hui Chia, Mao Ye, Hui-Ju Hung

Conference Reception (6:40 – 9:00)
Room: Maui Ballroom

Wednesday, October 31, 2012

Keynote Speech (8:30 – 9:30)
Chair:
Room: Maui Ballroom
Title: Learning Similarity Measure based on Random Walks
Speaker: William Cohen, Carnegie Mellon University

Coffee Break (9:30 – 10:15)

Session 13 (10:15 – 12:20)
KM Track: Advertisement and Products
Chair:
Room: Wailuku

- Daily-Deal Selection for Revenue Maximization
  Theodoros Lappas, Evimaria Terzi
- Enabling Direct Interest-Aware Audience Selection
  Ariel Fuxman, Anitha Kannan, Zhenhui Li, Panayiotis Tsaparas
- Influence Propagation in Adversarial Setting: How to Defeat Competition with Least Amount of Investment
  Shahrzad Shirazipourazad, Brian Bogard, Harsh Vachhani, Arunabha Sen, Paul Horn
- Large-scale Item Categorization for e-Commerce
  Dan Shen, Jean-David Ruvini, Badrul Sarwar
- Matching Product Titles using Web-based Enrichment Vishrawas
  Gopalakrishnan, Suresh Parthasarathy Iyengar, Amit Madaan, Rajeev Rastogi, Srinivasan Sengamedu

Session 14 (10:15 – 12:20)
KM Track: Clustering
Chair:
Room: Kahului

- Scalable Clustering of Signed Networks Using Balance Normalized Cut
  Kai-Yang Chiang, Joyce Jiyoung Whang, Inderjit S. Dhillon
- Maximum Margin Clustering on Evolutionary Data
  Xuhui Fan, Lin Zhu, Longbing Cao, Xia Cui, Yew-Soon Ong
- Document-Topic Hierarchies from Document Graphs
  Tim Weninger, Yonatan Bisk, Jiawei Han
- Improving Document Clustering Using Automated Machine Translation
  Xiang Wang, Buyue Qian, Ian Davidson
- Right-Protected Data Publishing with Hierarchical Clustering Preservation
  Michail Vlachos, Aleksander Wieczorek, Johannes Schneider

Session 15 (10:15 – 12:20)
IR Track: Recommendation Systems
Chair:
Room: Kihei and Wailea

- Metaphor: A System for Related Search Recommendations
  Azarias Reda, Yubin Park, Mitul Tiwari, Christian Posse, Sam Shah
- Exploring Personal Impact for Group Recommendation
  Xingjie Liu, Yuan Tian, Mao Ye, Wang-Chien Lee
- The Efficient Imputation Method for Neighborhood-based Collaborative Filtering
  Yongli Ren, Gang Li, Jun Zhang, Wanlei Zhou
- Multi-Faceted Ranking of News Articles using Post-Read Actions
  Deepak Agarwal, Bee-Chung Chen, Xuanhui Wang
- A Decentralized Recommender System for Effective Web Credibility Assessment
  Thanasis G. Papaioannou, Jean-Eudes Ranvier, Alexandra Olteanu, Karl Aberer

Session 16 (10:15 – 12:20)
IR Track: Digital Libraries and Citation Analysis
Chair:
Room: Kapalua

- Towards an Effective and Unbiased Ranking of Scientific Literature through Mutual Reinforcement
  Xiaorui Jiang, Xiaoping Sun, Hai Zhuge
- A Math-Aware Search Engine for Math Question Answering System
  Tam T. Nguyen, Kuiyu Chang, Siu Cheung Hui
- Contextualization using Hyperlinks and Internal Hierarchical Structure of Wikipedia Documents
  Muhammad Ali Norozi, Paavo Arvola, Arjen P. de Vries
- Understanding Book Search Behavior on the Web
  Jinyoung Kim, Henry Feild, Marc Cartright
- Temporal Corpus Summarization Using Submodular Word Coverage
  Ruben Sipos, Adith Swaminathan, Pannaga Shivaswamy, Thorsten Joachims

Poster Session (10:15 – 12:20)
KM Track
Chair:
Room: Napili

- Learning to Rank for Hybrid Recommendation
  Jiankai Sun, Shuaiqiang Wang, Byron J. Gao, Jun Ma
- Importance Weighted Passive Learning
  Shuaiqiang Wang, Xiaoming Xi, Yilong Yin
- A Tag-Centric Discriminative Model for Web ObjectsClassification
  Lina Yao, Quan Z. Sheng
- Outlier Detection using Centrality and Center-Proximity
  Duck-Ho Bae, Seo Jeong, Sang-Wook Kim, Minsoo Lee
- An Effective Category Classification Method Based on a Language Model for Question Category
  Recommendation on a cQA service

Kyoungman Bae, Youngjoong Ko
- Clustering Short Text Using Ncut-weighted Non-negative Matrix Factorization
  Xiaohui Yan, Jiafeng Guo, Shenghua Liu, Xue-qi Cheng, Yanfeng Wang
- Polygene-based Evolution: A Novel Framework for Evolutionary Algorithms
  Shuaiqiang Wang, Byron J. Gao, Shuangling Wang, Guibao Cao, Yilong Yin
- A Tensor Encoding Model for Semantic Processing
  Michael Symonds, Peter D Bruza, Laurianne Sitbon, Ian Turner
- Accelerating Locality Preserving Nonnegative Matrix Factorization
  Guanhong Yao, Cai Deng
- The Twitaholic Next Door.
  Patrick Bamba, Julien Subercaze, Christophe Gravier, Nabil Benmira, Jimi Fontaine
- Information Propagation in Social Rating Networks
  Priyanka Garg, Irwin King, Michael R. Lyu
- Maximizing Revenue from Strategic Recommendations under Decaying Trust
  Paul D¸tting, Monika Henzinger, Ingmar Weber
- Weighted Linear Kernel with Tree Transformed Features For Malware Detection.
  Prakash Mandayam Comar, Lei Liu, Sabyasachi Saha, Antonio Nucci, Pang-Ning Tan
- Learning to Predict the Cost-Per-Click for Your Ad Words
  Chieh-Jen Wang, Hsin-Hsi Chen
- Dual Word and Document Seed Selection for Semi-supervised Sentiment Classification
  Shengfeng Ju, Shoushan Li, Yan Su, Guodong Zhou, Yu Hong, Xiaojun Li
- On Empirical Tradeoffs in Large Scale Hierarchical Classification
  Rohit Babbar, Ioannis Partalas, Eric Gaussier, Cecile Amblard
- An Interaction Framework of Service-oriented Ontology Learning
  Jingsong Zhang, Yinglin Wang, Hao Wei
- Infobox Suggestion for Wikipedia Entities
  Afroza Sultana, Quazi Mainul Hasan, Ashis Kumer Biswas, Soumyava Das, Habibur Rahman, Chris Ding, Chengkai Li
- Time Feature Selection for Identifying Active Household Members
  Pedro G. Campos, Alejandro Bellogin, Fernando Diez, Ivan Cantador
- Text Classification with Relatively Small Positive Documents and Unlabeled Data
  Fumiyo Fukumoto, Takeshi Yamamoto, Suguru Matsuyoshi, Yoshimi Suzuki
- On Compressing Weighted Time-evolving Graphs
  Wei Liu, Andrey Kan, Jeffrey Chan, James Bailey, Christopher Leckie, Jian Pei, Ramamohanarao Kotagiri
- Graph-based Collective Classification for Tweets
  Yajuan Duan, Furu Wei, Ming Zhou, Heung-Yeung Shum
- A Word-Order Based Graph Representation For Relevance Identification
  Lakshmi Ramachandran, Edward F Gehringer
- Tracing Clusters in Evolving Graphs with Node Attributes
  Brigitte Boden, Stephan G¸nnemann, Thomas Seidl
- Prediction of Retweet Cascade Size over Time
  Andrey Kupavskii, Liudmila Ostroumova, Alexey Umnov, Svyatoslav Usachev, Pavel Serdyukov, Gleb Gusev, Andrey Kustarev

- An Efficient and Simple Under-sampling Technique for Imbalanced Time Series Classification
  Guohua Liang, Chengqi Zhang
- Top-N Recommendation through Belief Propagation
  Jiwoon Ha, Soon-Hyoung Kwon, Sang-Wook Kim, Christos Faloutsos, Sunju Park
- Mining Advices from Weblogs
  Alfan Farizki Wicaksono, Sung-Hyon Myaeng
- Parallel Proximal Support Vector Machine for High-Dimensional Pattern Classification
  Zhenfeng Zhu, Xingquan Zhu, Yangdong Ye, Yue-Fei Guo, Xiangyang Xue
- On Using Category Experts for Improving the Performance and Accuracy in Recommender Systems
  Won-Seok Hwang, Ho-Jong Lee, Sang-Wook Kim, Minsoo Lee
- Finding Influential Products on Social Domination Game
  Jinyoung Yeo, Jin-woo Park, Seung-won Hwang
- Entity Resolution using Search Engine Results
  Madian Khabsa, Pucktada Treeratpituk, C. Lee Giles
- Tweet Classification Based on Their Lifetime Duration
  Hikaru Takemura, Keishi Tajima
- Scalable Collaborative Filtering Using Incremental Update and Local Link Prediction
  Xiao Yang, Zhaoxin Zhang, Ke Wang
- Composing Activity Groups in Social Networks
  Cheng-Te Li, Man-Kwan Shan
- A Co-training based Method for Chinese Patent Semantic Annotation
  Xu Chen, Zhiyong Peng, Cheng Zeng
- Automatic Labeling Hierarchical Topics
  Xian-Ling Mao, Zhao-Yan Ming, Zheng-Jun Zha, Tat-Seng Chua, Hongfei Yan, Xiaoming Li
- An Unsupervised Method for Author Extraction from Web Pages Containing User-Generated Content
  Jing Liu, Xinying Song, Jingtian Jiang, Chin-Yew Lin
- Hierarchical Target Type Identification for Entity-oriented Queries
  Krisztian Balog, Robert Neumayer
- Dictionary based Sparse Representation for Domain Adaptation
  Rishabh Mehrotra, Rushabh Agrawal, Syed Aqueel Haider

Demo Session (10:15 – 12:20)
All three tracks
Chair:
Room: Kula and Hana

KM Track
- LUKe and MIKE:Learning from User Knowledge andManaging Interactive Knowledge Extraction
  Steffen Metzger, Michael Stoll, Katja Hose, Ralf Schenkel
- PRAVDA-live: Interactive Knowledge Harvesting
  Yafang Wang, Maximilian Dylla, Zhaochun Ren, Marc Spaniol, Gerhard Weikum
- 4Is of Social Bully Filtering: Identity, Inference, Influence, and Intervention
  Yunfei Chen, Lanbo Zhang, Aaron Michelony, Yi Zhang
- lonomics Atlas - A Tool To Explore Interconnected Ionomic, Genomic and Environmental Data

Eduard C. Dragut, Mourad Ouzzani, Amgad Madkour, Nabeel Mohamed, Peter Baker, David E. Salt
- CarbonDB: a Semantic Life Cycle Inventory Database
  Benjamin Bertin, Vasile-Marian Scuturici, Jean-Marie Pinon, Emmanuel Risler
- Supporting Temporal Analytics for Health-Related Events in Microblogs
  Nattiya Kanhabua, Sara Romano, AvarÈ Stewart, Wolfgang Nejdl
- InCaToMi: Integrative Causal Topic Miner Between Textual and Non-textual Time Series Data
  Hyun Duk Kim, ChengXiang Zhai, Thomas A. Rietz, Daniel Diermeier, Meichun Hsu, Malu Castellanos, Carlos A. Ceja Limon
- A Tool for Automated Evaluation of Algorithms
  Philipp Kranen, Stephan Wels, Tim Rohlfs, Sebastian Raubach, Thomas Seidl

IR Track
- A Summarization Tool for Time-Sensitive Social Media Walid
  Magdy, Ahmed Ali, Kareem Darwish
- CrowdTiles: Presenting Crowd-based Information for Event-driven Information Needs
  Stewart Whiting, Ke Zhou, Joemon Jose, Omar Alonso, Teerapon Leelanupab
- ESA: Emergency Situation Awareness via Microbloggers
  Jie Yin, Sarvnaz Karimi, Bella Robinson, Mark Cameron
- Cager: A Framework for Cross-page Search
  Zhumin Chen, Byron J. Gao, Qi Kang
- Mixed-Initiative Conversational System using Question-Answer Pairs Mined from the Web
  Wilson Wong, Lawrence Cavedon, John Thangarajah, Lin Padgham
- PicAlert!: A System for Privacy-Aware Image Classification and Retrieval
  Sergej Zerr, Stefan Siersdorfer, Jonathon Hare
- TASE: A Time-Aware Search Engine
  Sheng Lin, Peiquan Jin, Xujian Zhao, Lihua Yue
- Gumshoe Quality Toolkit: Administering Programmable Search
  Zhuowei Bao, Benny Kimelfeld, Yunyao Li, Sriram Raghavan, Huahai Yang
- Simultaneous Realization of Page-centric Communication and Search
  Yuhki Shiraishi, Jianwei Zhang, Yukiko Kawai, Toyokazu Akiyama
- MOUNA: Mining Opinions to Unveil Neglected Arguments
  Mouna Kacimi, Johann Gamper

DB Track
- MAGIK: Managing Completeness of Data
  Ognjen Savkovic, Mirza Paramita, Sergey Paramonov, Werner Nutt
- Exploration of Monte-Carlo based Probabilistic Query Processing in Uncertain Graphs
  Tobias Emrich, Hans-Peter Kriegel, Johannes Niedermayer, Matthias Renz, AndrÈ Suhartha, Andreas Z̦fle
- The Nautilus Analyzer: Understanding and Debugging Data Transformations
  Melanie Herschel, Hanno Eichelberger
- Demonstrating ProApproX 2.0: A Predictive Query Engine for Probabilistic XML
  Asma Souihli, Pierre Senellart
- HadoopXML: A Suite for Parallel Processing of Massive XML Data with Multiple Twig Pattern Queries
  Hyebong Choi, Kyong-Ha Lee, Soo-Hyong Kim, Yoon-Joon Lee, Bongki Moon
- MADden: Query-Driven Statistical Text Analytics

Christan Earl Grant, Joir-dan Gumbs, Kun Li, Daisy Zhe Wang, George Chitouras
- STFMap: Query- and Feature-Driven Visualization of Large Time Series Data Sets
  K. SelÁuk Candan, Rosaria Rossini, Maria Luisa Sapino, Xiaolan Wang
- Primates: A Privacy Management System for Social Networks
  Imen Ben Dhia, Talel Abdessalem, Mauro Sozio
- AMADA: Web Data Repositories in the Amazon Cloud
  AndrÈs Aranda-And˙jar, Francesca Bugiotti, Jes˙s Camacho-RodrÌguez, Dario Colazzo, FranÁois GoasdouÈ, Zoi Kaoudi, Ioana Manolescu

Lunch on you own (12:20 – 01:30)

Session 17 (1:30 – 3:35)
KM Track: Text Mining
Chair:
Room: Wailuku
- TCSST: Transfer Classification of Short & Sparse Text Using External Data
  Guodong Long, Ling Chen, Xingquan Zhu, Chengqi Zhang
- The Generalized Dirichlet Distribution in Enhanced Topic Detection
  Karla L Caballero, Joel Barajas, Ram Akella
- Modeling Topic Hierarchies with the Recursive Chinese Restaurant Process
  Joon Hee Kim, Dongwoo Kim, Suin Kim, Alice Oh
- Two-part Segmentation of Text Documents
  Deepak P, Karthik Visweswariah, Nirmalie Wiratunga, Sadiq Sani
- On the Design of LDA Models for Aspect-based Opinion Mining
  Samaneh Moghaddam, Martin Ester

Session 18 (1:30 – 3:35)
IR Track: Formal Retrieval Models and Learning to Rank
Chair:
Room: Kahului
- Predicting Query Performance for Fusion-Based Retrieval
  Gad Markovits, Anna Shtok, Oren Kurland, David Carmel
- Back to the Roots: A Probabilistic Framework for Query-Performance Prediction
  Oren Kurland, Anna Shtok, Shay Hummel, Fiana Raiber, David Carmel, Ofri Rom
- Learning to Rank for Robust Question Answering
  Arvind Agarwal, Hema Raghavan, Karthik Subbian, Prem Melville, Richard D Lawrence, David C Gondek, James Fan
- Learning to Rank By Aggregating Expert Preferences
  Maksims N Volkovs, Hugo Larochelle, Richard S Zemel
- Learning to Rank Duplicate Bug Reports
  Jian Zhou, Hongyu Zhang

Session 19 (1:30 – 3:35)
DB Track: Probabilistic and Uncertain Data
Chair:
Room: Kihei and Wailea

- A Model-based Approach for RFID Data Stream Cleansing
  zhou zhao, Wilfred Ng
- What is the IQ of your Data Transformation System?
  Giansalvatore Mecca, Paolo Papotti, Salvatore Raunich, Donatello Santoro
- On the Foundations of Probabilistic Information Integration
  Fereidoon Sadri
- GPU Acceleration of Probabilistic Frequent Itemset Mining from Uncertain Databases
  Yusuke Kozawa, Toshiyuki Amagasa, Hiroyuki Kitagawa
- Completeness of Queries over SQL Databases
  Werner Nutt, Simon Razniewski

Session 20 (1:30 – 3:35)
DB Track: Top-k and Nearest Neighbor Queries
Chair:
Room: Kapalua

- Being Picky-Processing Top-K Queries with Set-Defined Selections
  Aleksandar Stupar, Sebastian Michel
- Finding Top k Most Influential Spatial Facilities over Uncertain Objects
  Liming Zhan, Ying Zhang, Wenjie Zhang, Xuemin Lin
- Efficient Safe-Region Construction for Moving Top-K Spatial Keyword Queries
  Weihuang Huang, Guoliang Li, Kian-Lee Tan, Jianhua Feng
- Monochromatic and Bichromatic Reverse Nearest Neighbor Queries on Land Surfaces
  Da Yan, Zhou Zhao, Wilfred Ng
- Pay-as-you-go Maintenance of Precomputed Nearest Neighbors in Large Graphs
  Tom Crecelius, Ralf Schenkel

Poster Session (1:30 – 3:35)
IR Track
Chair:
Room: Napili

- Selecting Expansion Terms as a Set via Integer Linear Programming
  Qi Zhang, Yan Wu, Xuanjing Huang
- An Evaluation and Enhancement of Densitometric Fragmentation for Content Slicing Reuse
  Killian Levacher, Seamus Lawless, Vincent Wade
- Mathematical Equation Retrieval Using Plain Words as a Query
  Shinil Kim, Seon Yang, Youngjoong Ko
- Serial Position Effects of Clicking Behavior on Result Pages Returned by Search Engines
  Mingda Wu, Shan Jiang, Yan Zhang
- Towards Measruing the Visualness of a Concept
  Jin-Woo Jeong, Xin-Jing Wang, Dong-Ho Lee
- Fast Candidate Generation for Two-Phase Document Ranking: Postings List Intersection with Bloom Filters
  Nima Asadi, Jimmy Lin
- Semantically Coherent Image Annotation with a Learning-based Keyword Propagation Strategy
  Chaoran Cui, Jun Ma, Shuaiqiang Wang, Shuai Gao, Tao Lian

- Language Processing for Arabic Microblog Retrieval
  Kareem Darwish, Walid Magdy, Ahmed Mourad
- Hierarchical Image Annotation Using Semantic Hierarchies
  Hichem Bannour, CÈline Hudelot
- On the Inference of Average Precision from ScoreDistributions
  Ronan Cummins
- An Evaluation of Corpus-driven Measures of Medical Concept Similarity for Information Retrieval
  Bevan Koopman, Guido Zuccon, Peter Bruza, Laurianne Sitbon, Michael Lawley
- A Constraint to Automatically Regulate Document-LengthNormalisation
  Ronan Cummins, Colm O'Riordan
- Bridging Offline and Online Social Graph Dynamics
  Manuel Gomez Rodriguez, Monica Rogati
- Predicting the Performance of Passage Retrieval for Question Answering
  Eyal Krikon, David Carmel, Oren Kurland
- Coarse-to-Fine Sentence-level Emotion Classification based on the Intra-sentence Features and Sentential Context
  Jun Xu, Ruifeng Xu, Qin Lu, Xiaolong Wang
- Query-Performance Prediction and Cluster Ranking: Two Sides of the Same Coin
  Oren Kurland, Fiana Raiber, Anna Shtok
- Learning to Rank Search Results for Time-Sensitive Queries
  Nattiya Kanhabua, Kjetil N⁻rvÂg
- On Active Learning in Hierarchical Classification
  Yu Cheng, Kunpeng Zhang, Yusheng Xie, Ankit Agrawal, Alok Choudhary
- Question-Answer Topic Model for Question Retrieval in Community Question Answering
  Zongcheng Ji, Fei Xu, Bin Wang, Ben He
- How Do Humans Distinguish Different People with Identical Names on the Web?
  Harumi Murakami, Yuki Miyake
- Enhancing Product Search by Best-Selling Prediction in E-Commerce
  Bo Long, Jiang Bian, Anlei Dong, Yi Chang
- Survival Analysis for Freshness in Microblogging Search
  Gianni Amati, Giuseppe Amodeo, Carlo Gaibisso
- Information Preservation in Static Index Pruning
  Ruey-Cheng Chen, Chia-Jung Lee, Chiung-Min Tsai, Jieh Hsiang
- Temporal Models for Microblogs
  Jaeho Choi, W. Bruce Croft
- I want what I need! Analyzing Subjectivity of Online Forum Threads
  Prakhar Biyani, Cornelia Caragea, Amit Singh, Prasenjit Mitra
- Improving the Performance of the Reinforcement Learning Model for Answering Complex Questions
  Yllias Chali, Sadid A. Hasan, Kaisar Imam
- Relation Regularized Subspace Recommending for Related Scientific Articles
  Qing Zhang, Jianwu Li, Zhiping Zhang, Li Wang
- Exploring the Cluster Hypothesis, and Cluster-Based Retrieval, over the Web
  Fiana Raiber, Oren Kurland
- A Picture Paints a Thousand Words: a Method of Generating Image-text Timelines

Shize Xu, Liang Kong, Yan Zhang
- Short-Text Domain Specific Key Terms/Phrases Extraction Using an n-gram Model with Wikipedia
  M. Atif Qureshi, Colm ORiordan, Gabriella Pasi
- A New Probabilistic Model for Top-k Ranking Problem
  Shuzi Niu, Yanyan Lan, Jiafeng Guo, Xueqi Cheng
- Large Scale Analysis of Changes in English Vocabulary over Recent Time
  Adam Jatowt, Katsumi Tanaka
- Climbing the App Wall: Enabling Mobile App Discovery through Context-Aware Recommendations
  Alexandros Karatzoglou, Linas Baltrunas, Karen Church, Matthias Bˆhmer
- TwiSent: A Multistage System for Analyzing Sentiment inTwitter
  Subhabrata Mukherjee, Akshat Malu, Balamurali A.R., Pushpak Bhattacharyya
- Twitter Hyperlink Recommendation with User-Tweet-Hyperlink Three-way Clustering
  Dehong Gao, Renxian Zhang, Wenjie Li, Yuexian Hou
- Concavity in IR Models
  StÈphane Clinchant
- Extracting Interesting Association Rules from Toolbar Data
  Ilaria Bordino, Debora Donato, Barbara Poblete
- Predicting CTR of New Ads via Click Prediction
  Alexander Kolesnikov, Yury Logachev, Valeriy Topinskiy
- An Examination of Content Farms in Web Search using Crowdsourcing
  Richard McCreadie, Craig Macdonald, Iadh Ounis, Jim Giles, Ferris Jabr
- Demographic Context in Web Search Re-ranking
  Eugene Kharitonov, Pavel Serdyukov

Demo Session S2 (1:30 – 3:35)
Chair:
Room: Kula and Hana
All three tracks (Papers are listed in Demo Session S1)

Coffee Break (3:35 – 4:00)

Session 21 (4:00 – 5:40)
KM Track: Spatial and Temporal Methods
Chair:
Room: Wailuku
- Spatial Influence vs. Community Influence: Modeling the Global Spread of Social Media
  Krishna Y Kamath, James Caverlee, Zhiyuan Cheng, Daniel Z Sui
- TUT: A Statistical Model for Detecting Trends, Topics and User Interests in Social Media
  Xuning Tang, Christopher C. Yang
- Predicting Aggregate Social Activities Using Continuous-Time Stochastic Process
  Shu Huang, Min Chen, Bo Luo, Dongwon Lee
- Acquiring Temporal Constraints between Relations
  Partha Pratim Talukdar, Derry Wijaya, Tom Mitchell

Session 22 (4:00 – 5:40)
IR Track: Web Search
Chair:
Room: Kahului
- Towards Optimum Query Segmentation: In Doubt Without
  Matthias Hagen, Martin Potthast, Anna Beyer, Benno Stein
- Leaving So Soon? Understanding and Predicting Web Search Abandonment Rationales
  Abdigani Diriye, Ryen White, Georg Buscher, Susan Dumais
- Click Patterns: An Empirical Representation of Complex Query Intents
  Huizhong Duan, Emre Kiciman, ChengXiang Zhai
- Domain Dependent Query Reformulation for Web Search
  Van Dang, Giridhar Kumaran, Adam Troy

Session 23 (4:00 – 5:40)
DB Track: Web Data Management
Chair:
Room: Kihei and Wailea
- An Automatic Blocking Mechanism for Large-Scale De-duplication Tasks
  Anish Das Sarma, Ankur Jain, Ashwin Machanavajjhala, Philip Bohannon
- Processing Continuous Text Queries Featuring Non-Homogeneous Scoring Functions
  Nelly Vouzoukidou, Bernd Amann, Vassilis Christophides
- Comprehension-Based Result Snippets
  Abhijith Kashyap, Vagelis Hristidis
- An Effective Rule Miner for Instance Matching in a Web of Data
  Xing Niu, Shu Rong, Haofen Wang, Yong Yu

Short Paper Session S7 (4:00 – 5:40)
IR Track: Ranking and Recommendation
Chair:
Room: Kapalua
- Variance Maximization via Noise Injection for Active Sampling in Learning to Rank
  Wenbin Cai, Ya Zhang
- More Than Relevance: High Utility Query Recommendation By Mining Users' Search Behaviors
  Xiaofei Zhu, Jiafeng Guo, Xueqi Cheng, Yanyan Lan
- Recommending Citations: Translating Papers into References
  Wenyi Huang, Saurabh Kataria, Cornelia Caragea, Prasenjit Mitra, C. Lee Giles, Lior Rokach
- Discovering Logical Knowledge for Deep Question Answering
  Zhao Liu, Xipeng Qiu, Ling Cao, Xuanjing Huang
- Consento: A New Framework for Opinion Based Entity Search and Summarization
  Jaehoon Choi, Donghyeon Kim, Seongsoon Kim, Junkyu Lee, Sangrak Lim, Sunwon Lee, Jaewoo Kang
- Search Result Presentation Based on Faceted Clustering
  Benno Stein, Tim Gollub, Dennis Hoppe
- Entity Centric Query Expansion for Enterprise Search
  Xitong Liu, Hui Fang, Fei Chen, Min Wang

- Automatic Query Expansion Based on Tag Recommendation
  Vitor Oliveira, Guilherme Gomes, Fabiano Belem, Wladmir Brandao, Jussara Almeida, Nivio Ziviani, Marcos GonÁalves
- Query Recommendation for Children
  Sergio Duarte Torres, Djoerd Hiemstra, Ingmar Weber, Pavel Serdyukov

Poster Session (4:00 – 5:40)
IR + DB Track
Chair:
Room: Napili

IR Track
- On the Usefulness of Query Features for Learning to Rank
  Craig Macdonald, Rodrygo L.T. Santos, Iadh Ounis
- Session-based Query Performance Prediction
  Andrey Kustarev, Yury Ustinovskiy, Anna Mazur, Pavel Serdyukov
- A Latent Pairwise Preference Learning Approach for Recommendation from Implicit Feedback
  Yi Fang, Luo Si
- Topic Based Pose Relevance Learning In Dance Archives
  Reede Ren, John Collomosse, Joemon Jose
- PhotoFall: Discovering Weblog Stories Through Photographs
  Christopher Wienberg, Andrew S. Gordon
- RESQ: Rank-Energy Selective Query Forwarding for Distributed Search Systems
  Amin Teymorian, Xiao Qin, Ophir Frieder
- The Face of Quality in Crowdsourcing Relevance Labels
  Gabriella Kazai, Jaap Kamps, Natasa Milic-Frayling
- Data Filtering in Humor Generation
  Pawel Dybala, Rafal Rzepka, Kenji Araki, Kohichi Sayama
- Predicting Primary Categories of Business Listings for Local Search
  Changsung Kang, Jeehaeng Lee, Yi Chang
- Where Do the Query Terms Come from? An Analysis of Query Reformulation in Collaborative Web Search
  Zhen Yue, Jiepu Jiang, Shuguang Han, Daqing He
- Learning to Recommend with Social Relation Ensemble
  Lei Guo, Jun Ma, Zhumin Chen, Haoran Jiang
- A Scalable Approach For Performing Proximal Search For Verbose Patent Search Queries Sumit Bhatia, Bin He, Qi He, Scott Spangler
- Is Wikipedia Too Difficult? Comparative Analysis of Readability of Wikipedia, Simple Wikipedia and Britannica
  Adam Jatowt, Katsumi Tanaka
- Finding Food Entity Relationships using User-generated Data in Recipe Service
  Young-joo Chung
- SRGSIS: A Novel Framework Based on Social Relationship Graph for Social Image Search
  Bo Lu, Ye Yuan, Guoren Wang

- Exploring Simultaneous Keyword and Key Sentence Extraction: Improve Graph-based Ranking Using Wikipedia

  Xun Wang, Lei Wang, Jiwei Li, Sujian Li
- Estimating Query Difficulty for News Prediction Retrieval

  Nattiya Kanhabua, Kjetil NrvÂg
- Recency-Sensitive Model of Web Page Authority

  Maxim Zhukovskiy, Dmitry Vinogradov, Gleb Gusev, Pavel Serdyukov, Andrei Raigorodskii
- Evaluating Reward and Risk for Vertical Selection

  Ke Zhou, Ronan Cummins, Mounia Lalmas, Joemon M. Jose
- Contextual Evaluation of Query Reformulations in a Search Session by User Simulation

  Jiepu Jiang, Daqing He, Shuguang Han, Zhen Yue, Chaoqun Ni

DB Track

- Information-complete and Redundancy-free Keyword Search Over Large Data Graphs

  Byron J. Gao, Zhumin Chen, Qi Kang
- Spatial-aware Interest Group Queries in Location-based Social Networks

  Yafei Li, Dingming Wu, Jianliang Xu, Byron Choi, Weifeng Su
- Probabilistic Ranking in Fuzzy Object Databases

  Thomas Bernecker, Tobias Emrich, Hans-Peter Kriegel, Matthias Renz, Andreas Z¸fle
- Enabling Ontology Based Semantic Queries in Biomedical Database Systems

  Shuai Zheng, Fusheng Wang, James Lu, Joel Saltz
- Similarity Search in 3D Object-Based Video Data

  Jakub Lokoc, Jurgen Wunschmann, Tomas Skopal, Albrecht Rothermel
- Continuous Top-k Query for Graph Streams

  Shirui Pan, Xingquan Zhu
- Latent Topics in Graph-Structured Data

  Christoph Bhm, Gjergji Kasneci, Felix Naumann

Demo Session S3 (4:00 – 5:40)
Chair:
Room: Kula and Hana
All three tracks (Papers are listed in Demo Session S1)

Conference Banquet (6:30 – 9:30)

Thursday, November 1, 2012

Conference Opening (8:10 – 8:30)
Room: Maui Ballroom
- **Best paper Award Announcement**
- **CIKM2013 Presentation**

Keynote Speech (8:30 – 9:30)
Chair:
Room: Maui Ballroom
Title: Compressed Data Structures with Relevance
Speaker: Jeffrey S. Vitter, University of Kansas

Coffee Break (9:30 – 10:15)

Industry Day Morning Session (10:15 – 12:20)
Chair:
Room: Wailuku
- 10:15 – 10:20: Opening Remarks
- 10:20 – 11:05: Keynote Talk, Having a Great Career in Research
  Eric Brill, eBay
- 11:05 – 11:35, Is This Entity Relevant to Your Needs?
  David Carmel, IBM Research
- 11:35 – 12:20, Keynote Talk, The Future of Information Diversity and Search: Content Optimization, Interactivity, Semantics, and Social Networks
  Raghu Ramakrishnan, Microsoft

Session 24 (10:15 – 12:20)
KM Track: Information Extraction
Chair:
Room: Kahului
- Non-stationary Bayesian Networks based on Perfect Simulation
  Yi Jia, Wenrong Zeng, Jun Huan
- Active Learning for Relation Type Extension with Local and Global Data Views
  Ang Sun, Ralph Grishman
- Segmenting Web-Domains and Hashtags using Length Specific Models
  Sriram Srinivasan, Sourangshu Bhattacharya, Rudrasis Chakraborty
- Crosslingual Distant Supervision for Extracting Relations of Different Complexity
  Andre Blessing, Hinrich Schütze
- Labeling by Landscaping: Classifying Tokens in Context by Pruning and Decorating Trees
  Siddharth Patwardhan, Branimir Boguraev, Apoorv Agarwal, Alessandro Moschitti, Jennifer Chu-Carroll

Session 25 (10:15 – 12:20)
IR Track: Topic Modeling and Content and Sentiment Analysis
Chair:
Room: Kihei and Wailea
- G-WSTD: A Framework For Geographic Web Search Topic Discovery
  Di Jiang, Jan Vosecky, Kenneth Wai-Ting Leung, Wilfred Ng
- Supporting Factual Statements with Evidence from the Web
  Chee Wee Leong, Silviu Cucerzan
- Role-explicit Query Identification and Intent Role Annotation
  Haitao Yu, Fuji Ren
- Understanding Book Search Behavior on the Web
  Jinyoung Kim, Henry Feild, Marc Cartright
- Joint Topic Modeling for Event Summarization across News and Social Media Streams
  Wei Gao, Peng Li, Kareem Darwish

Session 26 (10:15 – 12:20)
DB Track: Query Processing, Optimization and Performance
Chair:
Room: Kapalua
- CGStream: Continuous Correlated Graph Query for Data Streams
  Shirui Pan, Xingquan Zhu
- Efficient Influence-Based Processing of Market Research Queries
  Anastasios Arvanitis, Antonios Deligiannakis, Yannis Vassiliou
- Deco: Declarative Crowdsourcing Aditya Ganesh Parameswaran,
  Hyunjung Park, Hector Garcia-Molina, Neoklis Polyzotis, Jennifer Widom
- Predicting the Effectiveness of Keyword Queries on Databases
  Shiwen Cheng, Arash Termehchy, Vagelis Hristidis
- You Can Stop Early with COLA: Online Processing of Aggregate Queries in the Cloud
  Yingjie Shi, Xiaofeng Meng, Fusheng Wang, Yantao Gan

Short Paper Session S8 (10:15 – 12:20)
KM Track: Learning and Knowledge Discovery
Chair:
Room: Napili
- Hierarchical Co-Clustering Based on Entropy Splitting
  Wei Cheng, Xiang Zhang, Feng Pan, Wei Wang
- Adapting Vector Space Model to Ranking-based Collaborative Filtering
  Shuaiqiang Wang, Jiankai Sun, Byron J Gao, Jun Ma
- Joint Relevance and Answer Quality Learning for Question Routing in Community QA
  Guangyou Zhou, Kang Liu, Jun Zhao
- Learning Spectral Embedding via Iterative Eigenvalue Thresholding
  Fanhua Shang, L. C. Jiao, Yuanyuan Liu, Fei Wang
- Discovering Personally Semantic Places from GPS Trajectories
  Mingqi Lv, Ling Chen, Gencai Chen
- Swimming against the Streamz: Search and Analytics over the Enterprise Activity Stream

Ido Guy, Tal Steier, Maya Barnea, Inbal Ronen, Tal Daniel
- Frequent grams based Embedding for Privacy Preserving Record
  Linkage Luca Bonomi, Li Xiong, Rui Chen, Benjamin C. M. Fung
- If You are Happy and You Know It... Tweet
  Amir Asiaee T., Mariano Tepper, Arindam Banerjee, Guillermo Sapiro
- Hierarchical Topic Integration Through Semi-supervised Hierarchical Topic Modeling
  Xian-Ling Mao, Jing He, Hongfei Yan, Xiaoming Li
- iSampling: Framework for Developing Sampling Methods Considering User's Interest
  Jinoh Oh, Hwanjo Yu
- Relational Co-Clustering via Manifold Ensemble Learning
  Ping Li, Jiajun Bu, Chun Chen, Zhanying He
- Measuring Website Similarity using an Entity-Aware Click Graph
  Pablo N Mendes, Peter Mika, Hugo Zaragoza, Roi Blanco
- A Hybrid Approach for Efficient Provenance Storage
  Yulai Xie, Dan Feng, Zhipeng Tan, Lei Chen, Kiran-Kumar Muniswamy-Reddy, Yan Li, Darrell D. E. Long

Short Paper Session S9 (10:15 – 12:20)
IR Track: Search and Advanced IR
Chair:
Room: Kula and Hana
- Customizing Search Results for Non-Native Speakers
  Theodoros Lappas, Michail Vlachos
- Sketch-based Indexing of n-Words
  Samuel Huston, J. Shane Culpepper, W. Bruce Croft
- Interactive and Context-Aware Tag Spell Check and Correction
  Francesco Bonchi, Ophir Frieder, Franco Maria Nardini, Fabrizio Silvestri, Hossein Vahabi
- From sBoW to dCoT Marginalized Encoders for Text Representation
  Zhixiang (Eddie) Xu, Minmin Chen, Kilian Q. Weinberger
- BiasTrust: Teaching Biased Users About Controversial Topics
  V.G.Vinod Vydiswaran, ChengXiang Zhai, Dan Roth, Peter Pirolli
- A Comprehensive Analysis of Parameter Settings for Novelty-Biased Cumulative Gain
  Teerapong Leelanupab, Guido Zuccon, Joemon M. Jose
- Differences in Effectiveness Across Sub-collections
  Mark Sanderson, Andrew Turpin, Ying Zhang, Falk Scholer
- Map to Humans and Reduce Error - Crowdsourcing for Deduplication Applied to Digital Libraries
  Mihai Georgescu, Dang Duc Pham, Claudiu S. Firan, Wolfgang Nejdl, Julien Gaugaz
- You Should Read This! Let Me Explain You Why
  Roi Blanco, Diego Ceccarelli, Claudio Lucchese, Raffaele Perego, Fabrizio Silvestri
- User Guided Entity Similarity Search Using Meta-Path Selection in Heterogeneous Information Networks
  Xiao Yu, Yizhou Sun, Brandon Norick, Tiancheng Mao, Jiawei Han
- Multi-Session Re-Search: In Pursuit of Repetition and Diversification
  Sarah K Tyler, Yi Zhang

- Fast Top-K Similarity Queries Via Matrix Compression
  Yucheng Low, Alice X Zheng

Lunch Provided by the Conference (12:20 – 01:30)

Industry Day Midday Session (1:30 – 3:35)
Chair:
Room: Wailuku
- 1:30 – 2:00, Social Media, Data Integration, and Human Computation
  AnHai Doan, WalmartLabs and UW-Madison
- 2:00 – 2:30, Question Answering Through Tencent Open Platform
  Chao Liu, Tencent
- 2:30 – 3:00, Data by the People, for the People
  Daniel Tunkelang, Linkedln
- 3:00 – 3:30, Leveraging Data to Power Local Commerce
  Rajesh Parekh, Groupon

Session 27 (1:30 – 3:35)
KM Track: Classification and Semantic Methods
Chair:
Room: Kahului
- A Novel Local Patch Framework for Fixing Supervised Learning Models
  Yilei Wang, Bingzheng Wei, Jun Yan, Yang Hu, Zhi-Hong Deng, Zheng Chen
- Automated Feature Weighting in Naive Bayes for High-dimensional Data Classification
  Lifei Chen, Shengrui Wang
- Learning to Discover Complex Mappings from Web Forms to Ontologies
  Yuan An, Xiaohua Hu, Il-Yeol Song
- Modeling Semantic Relations between Visual Attributes and Object Categories via Dirichlet Forest Prior
  Xin Chen, Xiaohua Hu, Zhongna Zhou, Yuan An, Tingting He, E.K. Park
- CoNet: Feature Generation for Multi-View Semi-Supervised Learning with Partially Observed Views
  Brian Quanz, Jun Huan

Session 28 (1:30 – 3:35)
IR Track: Multimedia and User Feedback
Chair:
Room: Kihei and Wailea
- Generating Facets for Phone-based Navigation of Structured Data
  Krishna Kummamuru, Ajith Jujjuru, Mayuri Duggirala
- The Effect of Aggregated Search Coherence on Search Behavior
  Jaime Arguello, Robert Capra
- Improving Bag-of-visual-Words Model with Spatial-Temporal Correlation for Video Retrieval
  Lei Wang, Dawei Song, Eyad Elyan
- Exploring and Predicting Search Task Difficulty

Jingjing Liu, Chang Liu, Michael Cole, Nicholas J. Belkin, Xiangmin Zhang
- Iterative Relevance Feedback with Adaptive Exploration/Exploitation Trade-off
  Nicolae Suditu, Fran,äö√Öois Fleuret

Session 29 (1:30 – 3:35)
DB Track: Emerging and Advanced Topics
Chair:
Room: Kapalua
- A Practical Concurrent Index for Solid-State Drives
  Risi Thonangi, Shivnath Babu, Jun Yang
- Robust Distributed Indexing for Locality-Skewed Workloads
  Mu-Woong Lee, Seung-won Hwang
- Efficient Provenance Storage For Relational Queries
  ZHIFENG BAO, Henning Kˆhler, Liwei Wang, Xiaofang Zhou, Shazia Sadiq
- Generically Extending Anonymization Algorithms to Deal with Successive Queries
  Manuel Barbosa, Alexandre Pinto, Bruno Gomes
- Authentication of Moving Range Queries
  Duncan Yung, Eric Lo, Man Lung Yiu

Short Paper Session S10 (1:30 – 3:35)
IR Track: Click Models, Learning and Mining
Chair:
Room: Napili
- Do Ads Compete or Collaborate? Designing Click Models with Full Relationship Incorporated
  Xin Xin, Irwin King, Ritesh Agrawal, Michael R. Lyu, Heyan Huang
- Finding Nuggets in IP Portfolios: Core Patent Mining through Textual Temporal Analysis
  Po Hu, Minlie Huang, Peng Xu, Weichang Li, Adam K Usadi, Xiaoyan Zhu
- Exploiting Concept Hierarchy for Result Diversification
  Wei Zheng, Hui Fang, Conglei Yao
- Semantic Context Learning with Large-Scale Weakly-Labeled Image Set
  Yao Lu, Wei Zhang, Ke Zhang, Xiangyang Xue
- Mining Noisy Tagging from Multi-label Space
  Zhongang Qi, Ming Yang, Zhongfei (Mark) Zhang, Zhengyou Zhang
- Learning from Mistakes: Towards a Correctable Learning Algorithm
  Karthik Raman, Krysta M Svore, Ran Gilad-Bachrach, Chris J.C. Burges
- PolariCQ: Polarity Classification of Political Quotations
  Rawia Awadallah, Maya Ramanath, Gerhard Weikum
- Modeling Browsing Behavior for Click Analysis in Sponsored Search
  Azin Ashkan, Charles L. A. Clarke
- User Activity Profiling with Multi-Layer Analysis
  Hongxia Jin
- Stochastic Simulation of Time-Biased Gain
  Mark D. Smucker, Charles L. A. Clarke
- Predicting Web Search Success with Fine-grained Interaction Data
  Qi Guo, Dmitry Lagun, Eugene Agichtein

- Mining Sentiment Terminology Through Time
  Hadi Amiri, Tat-Seng Chua

Short Paper Session S11 (1:30 – 3:35)
DB Track: Advanced DB Topics
Chair:
Room: Kula and Hana

- Efficient Logging for Enterprise Workloads on Column-Oriented In-Memory Databases
  Johannes Wust, Joos-Hendrick Boese, Frank Renkes, Sebastian Blessing, Jens Krueger, Hasso Plattner
- Discovering Conditional Inclusion Dependencies
  Jana Bauckmann, Ziawasch Abedjan, Ulf Leser, Heiko M¸ller, Felix Naumann
- Efficient Buffer Management for Piecewise Linear Representation of Multiple Data Streams
  Qing Xie, Jia Zhu, Mohamed A. Sharaf, xiaofang zhou, Chaoyi Pang
- On Skyline Groups
  Chengkai Li, Nan Zhang, Naeemul Hassan, Sundaresan Rajasekaran, Gautam Das
- Clustering Wikipedia Infoboxes to Discover their Types
  Thanh Hoang Nguyen, Huong Dieu Nguyen, Viviane Moreira, Juliana Freire
- Efficient Estimation of Dynamic Density Functions with an Application to Outlier Detection
  Abdulhakim Ali Qahtan, Xiangliang Zhang, Suojin Wang
- Real-Time Aggregate Monitoring with Differential Privacy
  Liyue Fan, Li Xiong
- Efficient Distributed Locality Sensitive Hashing
  Bahman Bahmani, Ashish Goel, Rajendra Shinde
- Star-Join: Spatio-Textual Similarity Join
  Sitong Liu, Guoliang Li, Jianhua Feng
- Adapt: Adaptive Database Schema Design for Multi-Tenant Applications
  Jiacai Ni, Guoliang Li, Jun Zhang, Lei Li, Jianhua Feng
- Scaling Multiple-Source Entity Resolution using Statistically Efficient Transfer Learning
  Sahand N Negahban, Benjamin I. P. Rubinstein, Jim Gemmell
- On Bundle Configuration for Viral Marketing in Social Networks
  De-Nian Yang, Wang-Chien Lee, Nai-Hui Chia, Mao Ye, Hui-Ju Hung

Coffee Break (3:35 – 4:00)

Industry Day Afternoon Session (4:00 – 6:00)
Chair:
Room: Wailuku

- 4:00 – 4:45, Keynote Talk, Revolutionizing Digital Marketing with Big Data Analytics
  Tom Malloy, Adobe
- 4:45 – 5:15, Programming and Debugging Large-Scale Data Processing Workflows
  Christopher Olston, Google
- 5:15 – 6:00, Keynote Talk, From HyperText to HyperTEC
  Xuedong Huang, Microsoft

Session 30 (4:00 – 5:40)
KM Track: Novel Applications
Chair:
Room: Kahului
- Model the Complex Dependence Structures of Financial Variables by Using Canonical Vine
  Wei Wei, Xuhui Fan, Jinyan Li, Longbing Cao
- A Unified Learning Framework for Auto Face Annotation by Mining Web Facial Images
  Dayong Wang, Steven Chu Hong Hoi, Ying He
- Efficient Jaccard-based Diversity Analysis of Large Document Collections
  Fan Deng, Stefan Siersdorfer, Sergej Zerr
- Knowing Where and How Criminal Organizations Operate Using Web Content
  Michele Coscia, Viridiana Rios

Session 31 (4:00 – 5:40)
IR Track: Social Networks
Chair:
Room: Kihei and Wailea
- Social Recommendation Across Multiple Relational Domains
  Meng Jiang, Peng Cui, Fei Wang, Qiang Yang, Wenwu Zhu, Shiqiang Yang
- Mining Competitive Relationships by Learning across Heterogeneous Networks
  Yang Yang, Jie Tang, Jacklyne Keomany, Yanting Zhao, Juanzi Li, Ying Ding, Tian Li, Liangwei Wang
- Evaluating Geo-Social Influence in Location-Based Social Networks
  Chao Zhang, Lidan Shou, Ke Chen, Gang Chen, Yijun Bei
- The Walls Have Ears: Optimize Sharing for Visibility and Privacy in Online Social Networks
  Thang N. Dinh, Yilin Shen, My T. Thai

Short Paper Session S12 (4:00 – 5:40)
IR Track: Social Media
Chair:
Room: Napili
- Trust Prediction via Aggregating Heterogeneous Social Networks
  Jin Huang, Feiping Nie, Heng Huang, Yi-Cheng Tu
- Diversionary Comments under Political Blog Posts
  Jing Wang, Clement T Yu, Philip S Yu, Bing Liu, Weiyi Meng
- Discover Breaking Events with Popular Hashtags in Twitter
  Anqi Cui, Min Zhang, Yiqun Liu, Shaoping Ma, Kuo Zhang
- Interest-Matching Information Propagation in Multiple Online Social Networks
  Yilin Shen, Thang N. Dinh, Huiyuan Zhang, My T. Thai
- Quality Models for Microblog Retrieval
  Jaeho Choi, W. Bruce Croft, Jin Young Kim
- Sentiment Analysis in Twitter with Lightweight Discourse Analysis
  Subhabrata Mukherjee, Pushpak Bhattacharyya, Balamurali A.R.
- Query-biased Learning to Rank for Real-time Twitter Search

Xin Zhang, Ben He, Tiejian Luo, Baobin Li

- Location-Sensitive Resources Recommendation in Social Tagging Systems
  Chang Wan, Ben Kao, David W. Cheung
- Detecting Offensive Tweets via Topical Feature Discovery over a Large Scale Twitter Corpus
  Guang Xiang, Bin Fan, Ling Wang, Jason Hong, Carolyn Rose

CIKM'12 Workshop Program

Monday, October 29, 2012

DUBMMSM - Data-driven User Behavioral Modelling and Mining from Social Media
Room: Wailuku

9:00 - 9:15 am - Welcome and Introduction

9:15 - 9:30 am - Madness session (5 madness papers, 3 min each).
- Analyzing Social Media Friendship for Personalization - Jonghyun Han, Hyunju Lee.
- A Collective Synchronous Behavior Model on Social Media - Victor Liang, Vincent Ng
- Probabilistic Macro Behavioral Targeting - Yusheng Xie
- Pinteresting: Towards a Better Understanding of User Interests -Ana-Maria Popescu
- The Framework of a People Recommender Based on a Time Series of User Preferences - Kosuke Takano, Kin Fun Li

9:30 - 10:30 am - Paper Session 1 (2 paper + discussions)
- Ranking and Combining Social Network Data for Web Personalization - Yi Zeng
- Please Spread: Recommending Tweets for Retweeting with Implicit Feedback - Sheng Wang, Xiaobo Zhou, Ziqi Wang, Ming Zhang

10:30 - 11:00 am - Coffee Break

11:00 - 12:00 pm - Paper Session 2 (2 paper + discussions)
- Identifying and Characterizing User Communities on Twitter during Crisis Events - Aditi Gupta, Anupam Joshi, Ponnurangam Kumaraguru
- Using Social Data for Resume Job Matching - David Hardtke, Jacob Bollinger, Ben Martin

12:00 - 2 pm - Lunch

2:00 - 3:30 pm - Paper Session 3 (3 paper + discussions)
- Twitter User Behavior Understanding with Mood Transition Prediction -Aditya Mogadala, Vasudeva Varma
- Analyzing Sentiments From Street Harassment Stories -Parvathi Chundi, April Corbet
- Modeling Online Collective Emotions -David Garcia, Frank Schweitzer

3:30 - 4:00 pm - Coffee Break

4:00 - 5:00 pm - Panel

5:00 - 5:15 pm - Closing remarks

Monday, October 29, 2012

CloudDB - 2012 The Third International Workshop on Cloud Data Management
Room: Kahulul

8:15-8:30 -- Welcome by the Chair

8:30-10:00 -- Keynote Session 1: OLTP
- 8:30-9:15 -- Keynote 1: Carlo Curino, Microsoft. Benchmarking OLTP/Web Databases in the Cloud: the OLTP-Bench Framework (Carlo Curino, Djellel Difallah, Andrew Pavlo, Phil Cudre-Mauroux)
- 9:15-10:00 -- Keynote 2: Prof. Mohamed Sharaf, The University of QueenslandData Freshness in Key-Value Data Stores

10:00-10:30 -- Coffee Break

10:30-12:00 -- Session 1: Workload-Aware Processing
- 10:30-11:00 -- Toward Non-Intrusive Elastic Query Processing in the Cloud. Ticiana Coelho Da Silva, Mário Nascimento, Jose Macedo, Flávio R. C. Sousa, Javam Machado
- 11:00-11:30 -- The Yahoo! Cloud Datastore Load Balancer. Markus Klems, Adam Silberstein, Jianjun Chen, Masood Mortazavi, Andrews Albert Sahaya,   P.P.S. Narayan
- 11:30-12:00 -- HEDC: A Histogram Estimator For Data in the Cloud. Yingjie Shi, Xiaofeng Meng, Fusheng Wang, Yantao Gan

12:00-1:30 -- Lunch

1:30-3:00 -- Keynote Session 2: Analytics and Social
- 1:30-2:15 -- Keynote 3: Prof. Geoffrey Fox, University of Indiana, Bloomington: Large Scale Data Analytics on Clouds
- 2:15-3:00 -- Keynote 4: Prof. Ashwin Machanavajjhala, Duke University: Challenges in Enabling Social Applications At Scale

3:00-3:30 -- Coffee Break

3:30-5:30 -- Session 2: Security, Privacy, Analytics
- 3:30-3:55 -- Cloud Computing for Environment-Friendly Data Centers. Michael Pawlish, Aparna S. Varde, Stefan A. Robila
- 3:55-4:20 -- A Security Aware Stream Data Processing Scheme on the Cloud and its Efficient Execution Methods. Katsuhiro Tomiyama, Hideyuki Kawashima, Hiroyuki Kitagawa
- 4:20-4:55 -- Differentially Private Top-k Query over MapReduce. Xu Han, Miao Wang, Xiaojian Zhang, Xiaofeng Meng
- 4:55-5:20 -- Facilitating Real-Time Graph Mining. Zhuhua Cai, Dionysios Logothetis, Georgos Siganos

5:20-5:30 -- Wrap-up and Summary

Monday, October 29, 2012

WKR/CDMW - The 2012 International Workshop on Web-scale Knowledge Representation Retrieval and Reasoning & City Data Management 2012 Workshop
Room: Kihei

Introduction [Spyros Kotoulas (IBM Research)]

Session 1   Web-scale Knowledge Representation, Retrieval and Reasoning. Session Chair: Yi Zeng (Chinese Academy of Sciences)
- A Distributed, Semiotic-Inductive, and Human-Oriented Approach to Web-Scale Knowledge Retrieval. Edy Portmann, Michael Alexander Kaufmann, Cédric Graf
- OmpiJava - A Tool For Development Of High-Performance Reasoning Applications For The Semantic Web. Alexey Cheptsov
- Efficient Mining of Correlated Sequential Patterns Based on Null Hypothesis. Cindy Xide Lin, Ming Ji, Marina Danilevsky, Jiawei Han

Session 2   City Data Management . Session Chair: Spyros Kotoulas (IBM Research)
- DataBridges: Data Integration for Digital Cities. Melanie Herschel, Ioana Manolescu
- U2STRA: High-Performance Data Management of Ubiquitous Urban Sensing Trajectories on GPGPUs. Fatiha Amanzougarene, Mohamed Chachoua, Karine Zeitouni
- Qualitative Representation of Building Sites Annoyance. Jianting Zhang, Simin You, Le Gruenwald

Discussion, Chair: Yi Zeng (Chinese Academy of Sciences)

Monday, October 29, 2012

SHB - International Workshop on Smart Health and Wellbeing
Room: Wailea

Opening ceremony

Keynote Speech

Session 1:
- Moving from Descriptive to Causal Analytics: Case Study of Discovering Knowledge from US Health Indicators Warehouse. Jack Schryver, Mallikarjun Shankar and Songhua Xu
- An Automated Data Utility Clustering Methodology using Data Constraint Rules. Stuart Morton, Malika Mahoui and P. Joseph Gibson
- Designing the Reconciled Schema for a Pharmacovigilance Data Warehouse Through a Temporally-Enhanced ER Model. Riccardo Lora, Alberto Sabaini, Carlo Combi and Ugo Moretti

Session 2:
- Towards Large-scale Twitter Mining for Drug-related Adverse Events. Jiang Bian, Umit Topaloglu and Fan Yu
- Social Media Mining for Drug Safety Signal Detection. Christopher C. Yang, Haodong Yang, Ling Jiang, and Mi Zhang

Session 3:
- An Architecture for Personalized Health Information Retrieval. Nikhil Yadav and Christian Poellabauer
- Combining Multi-level Evidence for Medical Record Retrieval. Dongqing Zhu and Ben Carterette
- Simulating Prosthetic Vision with Distortions for Retinal Prosthesis Design. Mahadevan Subramaniam, Parvathi Chundi and Eyal Margalit

Monday, October 29, 2012

BookOnline - Online Books, Complementary Social Media
Room: Kapalua

8.50-9.00   Welcome

9.00-10.10   Keynote
Session Chair: Gabriella Kazai (Microsoft Research)

Maribeth Back (FX Palo Alto)
Revisiting the Future of Reading: The Research and Design Behind XFR

10.10-10.30   Session: Search and Discovery
Session Chair: Monica Landoni (University of Lugano)

eBook meets Tabletop: Using Collaborative Visualization for Search and Serendipity in On-line Book Repositories
R. Rädle, A. Weiler, S. Huber, H.-C. Jetter, S. Mansmann, H. Reiterer, and M. Scholl

10.30-11.00   Coffee

11.00-11.40   Session cont'd: Search and Discovery
Session Chair: Monica Landoni (University of Lugano)

Spread Co-citation Relationship as a Measure for Document Retrieval
M. Eto

Search and Exploration of Scanned Books
M.-A. Cartright, J. Dalton, and J. Allan

11.45-12.25   Session: Personalization and Recommendation
Session Chair: Peter Brusilovsky (University of Pittsburgh)

Personalized Recommendations on Books for K-12 Readers
M. S. Pera and Y.-K. Ng

Stylometric Relevance-feedback towards a Hybrid Book Recommendation Algorithm
P. C. Vaz, D. M. de Matos, and B. Martins

12.30-13.30   Lunch

13.30-14.40   Keynote
Session Chair: Gabriella Kazai (Microsoft Research)

Natasa Milic-Frayling (Microsoft Research)
The Future of Digital

14.40-15.20   Session: Reading Experience Beyond Text
Session Chair: Carsten Eickhoff (Delft University of Technology)

Accessible, Large-Print, Listening & Talking E-book (ALLT)
A. Attarwala, R. Baecker, and C. Munteanu

Need for Automatically Generated Narration
D. A. Evans and J. Reichenbach

15.20-16.00   Open-Discussion

16.00-16.30   Coffee

16.30-17.00   Open-Discussion cont'd

17.00-17.30   Report back

17.30   Close

Monday, October 29, 2012

DTMBIO – ACM Sixth International Workshop on Data and Text Mining in Biomedical Informatics
Room: Napili

Session 1: Keynote Address. Session Chair: Doheon Lee (KAIST, Korea)
- Gwan-Su Yi (KAIST)

Session 2: Mining Clinical Data and Text. Session Chair: Hua Xu (Vanderbilt University, US)
- Lexicon-free and context-free drug names identification methods using Hidden Markov Models and Pointwise Mutual Information (Jacek Malyszko ; Agata Filipowska)
- Clinical Entity Recognition using Structural Support Vector Machines with Rich Features (Buzhou Tang; Hongxin Cao; Yonghui Wu; Min Jiang ; Hua Xu)

Coffee Break

- Inferring Appropriate Eligibility Criteria in Clinical Trial Protocols Without Labeled Data (Angelo Restificar ; Sophia Ananiadou)
- Predicting Baby Feeding Method from Unstructured Electron Health Record (Ashwani Rao; Kristin Maiden; Benjamin Carterette ; Deborah Ehrentha)
- Extracting Structured Information from Free-Text Medication Prescriptions Using Dependencies (Andrew MacKinlay ; Karin Verspoor)

Lunch Break

Session 3: Mining Biological Data and Text. Session Chair: Min Song (Yonsei University, Korea)
- Indexing Methods for Efficient Protein 3D Surface Search. Sungchul Kim; Sael Lee ; Hwanjo Yu
- Protein Complex Prediction via Bottleneck-Based Graph Partitioning. Jaegyoon Ahn; Dae Hyun Lee; Youngmi Yoon; Yunku Yeu ; Sanghyun Park
- Finding associations among SNPs for prostate cancer using collaborative filtering. Rohit Kugaonkar; Aryya Gangopadhyay; Yelena Yesha; Anupam Joshi; Yaacov Yesha; Michael Grasso; Mary Brady ; Napthali Rishe
- Prediction of E3-specific Substrates by Using Known E3-Substrate Network. Youngwoong Han ; Gwan-Su Yi
- Detecting Type 2 Diabetes Causal SNP Combinations from GWAS Dataset with Optimal Filtration. Chiyong Kang; Hyeji Yu ; Gwan-Su Yi

Coffee Break

- TNMCA: Generation and Application of Network Motif-Based Inference Models for Drug Repositioning. Jaejoon Choi; Kwangmin Kim; Min Song ; Doheon Lee
- High Precision Rule Based PPI Extraction and Per-Pair Basis Performance. Junkyu Lee; Seongsoon Kim; Sunwon Lee; Kyubum Lee ; Jaewoo Kang

- Rule-based whole body modeling for analyzing multi-compound effects Woochang Hwang; Yongdeuk Hwang; Sunjae Lee ; Doheon Lee

Monday, October 29, 2012

MIXHS - The 2nd International Workshop on Managing Interoperability and complexity in Health Systems
Room: Kula + Hana

09:00-9:10 Welcome

9:10-10:40 Session 1: Ontology-based Application on Clinical Data. Session Chair: Guoqian Jiang, (Mayo Clinic, USA)
- Clinical Clarity versus Terminological Order – The Readiness of SNOMED CT Concept Descriptors for Primary Care. Zhe He, Michael Halper, Yehoshua Perl and Gai Elhanan
- Extraction and analysis of the structure of labels in biomedical ontologies. Manuel Quesada-Martínez, Jesualdo Tomás Fernández-Breis and Robert Stevens
- Clinical Data Analysis using Ontology-guided Rule Learning. Hua Min and Janusz Wojtusiak

10:40-11:00 Coffee Break

11:00-12:30 Session 2: Electronic Health Systems Interoperability and Integration. Session Chair: Cui Tao (Mayo Clinic, USA)
- Harmonization of Detailed Clinical Models with Clinical Study Data Standards. Guoqian Jiang, Julie Evans, Tom Oniki, Joey Coyle, Landen Bain, Stan Huff, Rebecca Kush and Christopher Chute
- Modeling UIMA Type System Using Web Ontology Language – towards Interoperability among UIMA-based NLP Tools. Hongfang Liu, Stephen Wu, Cui Tao and Christopher Chute
- Quality Assessement of Electronic Health Information Management Systems. Matt-Mouley Bouamrane, Cui Tao and Frances Mair

12:30-14:00 Lunch Break

14:00-16:00 Session 3: Bio-Medical Knowledge Representation & Engineering. Session Chair: Hua Min, (George Mason University, USA)
- Construction and Maintenance of Clinical Pathway using Data Mining Methods. Shusaku Tsumoto, Haruko Iwata and Shoji Hirano
- Optimizing Semantic MEDLINE for Translational Science Studies Using Semantic Web Technologies. Cui Tao, Yuji Zhang, Guoqian Jiang, Matt Mouley Bouamrane and Christopher Chute
- Bridging the Unstructured and Structured Worlds: an Adaptive Self Learning Medical Form Generating System. Shuai Zheng, Fusheng Wang and James Lu
- A Hybrid Approach to Finding Negated and Uncertain Expressions in Biomedical Documents. Kazuki Fujikawa, Kazuhiro Seki and Kuniaki Uehara

16:00 Workshop Concluding remarks

Friday, November 2, 2012

PIKM - The 5th Ph.D. Workshop in Information and Knowledge Management
Room: Wailuku

8:45   Opening: Aparna Varde

9:00 Session 1: Database Systems
- 9:00   When Big Data Leads to Lost Data. V.M. Megler and David Maier
- 9:30   Querying External Source Code Files of Programs Connecting to a Relational Database. Carlos Garcia-Alvarado and Carlos Ordonez
- 10:00   SciQL: A Query Language for Unified Scientific Data Processing and Management. Javad Chamanara and Birgitta König-Ries

10:30 Coffee Break

11:00 Session 2: Knowledge Management / Data Mining
- 11:00   Feature Selection for Link Prediction. Ye Xu and Dan Rockmore
- 11:30   Exploring and Analyzing Documents with Online Analytical Processing. Grzegorz Drzadzewski and Frank Tompa
- 12:00   Is That Scene Dangerous?: Transferring Knowledge Over a Video Stream. Omar U Florez and Curtis Dyreson

12:30 Lunch Break

14:00 Keynote: Advice for Young Jedi Knights and PhD Students (Invited talk) Ingmar Weber

15:00 Session 3: Information Retrieval
- 15:00   iTop: Interaction Based Topic Centric Community Discovery on Twitter. Denzil Correa, Ashish Sureka and Mayank Pundir

15:30 Coffee Break
- 16:00   Search Tactics as Means of Examining Search Processes in Collaborative Exploratory Web Search. Zhen Yue, Shuguang Han, Jiepu Jiang and Daqing He
- 16:30   Assessing the Relationship between Context, User Preferences, and Content in Search Behavior. Hanna Knäusl and Bernd Ludwig
- 17:00   Recommendation Using Linked Data. Rouzbeh Meymandpour and Joseph Davis

17:30-18:30 Session 4: Posters
- Intent-Aware Temporal Query Modeling for Keyword Suggestion. Fredrik Johansson, Tobias Färdig, Vinay Jethava and Svetoslav Marinov
- Towards an Advanced System for Real-Time Event Detection in High Volume Data Streams. Andreas Weiler, Svetlana Mansmann and Marc Scholl
- Multilevel Business Process Modeling: Motivation, Approach, Design Issues and Applications. Christoph Schütz, Michael Schrefl and Lois Delcambre
- Towards a More Efficient and Personalized Advertisement Content in On-line Social Networks. Patxi Galán-García, Carlos Laorden and Pablo G. Bringas

Friday, November 2, 2012

PLEAD - Politics, Elections and Data
Room: Kahului

Session 1:
- Invited talk: "The Diffusion of Political Memes in Social Media" by Filippo Menczer

Session 2:
- "From Twindex to PredictWise: A Quick Overview of Political Analysis Tools" by Ingmar Weber
- "Political Polarization and Popularity in Online Participatory Media: an Integrated Approach" by David Garcia, Fernando Mendez, Uwe Serdult and Frank

Session 3:
- "Party Cohesion in Presidential Races: Applying Social Network Theory to the 2011 Preprimary" by Andrew Dowdle, Song Yang, Scott Limbocker, Patrick Stewart and Karen Sebold
- "Opinions Network for Politically Controversial Topics" by Rawia Awadallah, Maya Ramanath and Gerhard Weikum
- "French Presidential Elections: What are the Most Efficient Measures for Tweets?" by Flavien Bouillot, Pascal Poncelet, Mathieu Roche, Dino Lenco, Elnaz Bigdeli and Stan Matwin

Session 4:
- "The Price of Precision: Voter Microtargeting and its Potential Harms to the Democratic Process" by Solon Barocas

Panel discussion with representatives from academia, industry and media; moderated by Ana-Maria Popescu

Friday, November 2, 2012

ESAIR - Fifth Workshop on Exploiting Semantic Annotations in Information Retrieval
Room: Kihei

9:15-10:00 Keynote Session I: Keynote Presentation [Chair: Jussi Karlgren]

10:00-10:30 Coffee Break

10:30-11:15 Keynote Session II: Keynote Presentation [Chair: Jaap Kamps]

11:15-12:30 Boaster and Poster Session [Chair: Jaap Kamps]
- Krisztian Balog and Kjetil Nørvåg / On the Use of Semantic Knowledge Bases for Temporally-aware Entity Retrieval
- Amitava Das and Björn Gambäck / Exploiting 5W Annotations for Opinion Tracking
- Ann-Marie Eklund / Why query annotations may help in providing accurate public health information
- Sumio Fujita, Georges Dupret and Ricardo Baeza-Yates / Semantics of Query Rewriting Patterns in Search Logs
- Arunav Mishra, Sairam Gurajada and Martin Theobald / Design and Evaluation of an IR-Benchmark for SPARQL Queries with Full-text Conditions
- Tadashi Nomoto and Noriko Kando / Conceptualizing Documents with Wikipedia
- Sana Sellami and Claudia Catalin Gutiérrez Rodríguez / Semantic Annotation: What About Quality?
- Petr Sojka / Exploiting Semantic Annotations in Math Information Retrieval
- Giovanni Yoko Kristianto, Goran Topic, Minh-Quoc Nghiem and Akiko Aizawa / Annotating Scientific Papers for Mathematical Formulae Search
- Masaharu Yoshioka and Noriko Kando / Multifaceted analysis of news articles by using semantic annotated information

12:30-14:00 Lunch

14:00-15:30 Breakout session: Two breakout groups in parallel
- semantic search [Chair/Reporter: Jussi Karlgren/Peter Mika]
- structured retrieval [Chair/Reporter: Vanessa Murdock/Jaap Kamps]

15:30-16:00 Coffee

16:00-17:30 Final session [Chair: tba]: Reporting from breakout groups and concluding remarks.

18:00++ Social program: Dinner and symposium drinks and continued discussion!

Friday, November 2, 2012

ClowdSens - 1st International Workshop on Multimodal Crowd Sensing
Room: Wailea

Keynote Address. Session Chair: Haggai Roitman (IBM Research)
- Invited Talk, Ido Guy: Crowdsourcing in the enterprise

Session 1. Session Chair: Haggai Roitman (IBM Research)
- Algorithm for Representative Democracy Voting in Social Network.   Zeinab Saeidi
- Conceptual Modeling Principles for Crowdsourcing. Roman Lukyanenko; Jeffrey Parsons
- Event Detection using Twitter and Structured Semantic Query Expansion. Heather S. Packer; Sina Samangooei; Jonathon S. Hare; Nicholas Gibbins; Paul Lewis

Invited Talk. Session Chair: Haggai Roitman (IBM Research)
- Invited Talk, Manuel Cebrian: Using Friends as Sensors to Detect Planetary-Scale Contagious Outbreaks

Session 2. Session Chair: Haggai Roitman (IBM Research)
- Harnessing the Crowds for Smart City Sensing. Haggai Roitman; Jonathan Mamou; Sameep Mehta; Aharon Satt; L. V. Subramaniam
- Greaaaat bargains starting from just 99p!!!! :-) Brand Perception in the Social Media. Michal Shmueli-Scheuer; Benjamin Sznajder; Doron Cohen; Ariel Raviv; David Konopnicki, Haggai Roitman

Session 3: Discussion. Session Chair: Haggai Roitman (IBM Research)

Friday, November 2, 2012

IKM2DR - Information and Knowledge Management for Developing Regions
Room: Kapalua

First AM Session: (chair: Nitendra Rajput)
- Keynote address [Ricardo Baeza--Yates, Yahoo]
- Direction Setting Panel (Ricardo Baeza--Yates, Doug Oard, Nitendra Rajput)

Second AM Session: (chair: Luz Quiroga)

- Domain--specific search in Indian Languages [Nikihil Pattisapu, SKYPE]
- Speech retrieval for India [Pekka Kallioniemi]
- Query by Babbling for speech retrieval [Doug Oard]
- Named entity recognition for Indian Languages (Mahathi Bhagavatula, SKYPE]

Lunch and Discussion Tables: Two discussion tables, led by Krishna Kummamuru, Luz Quiroga

First PM Session: (chair: William Webber)
- Invited talk (abstract uploaded) [Anitha Kannan, MSR]
- Two parallel breakout sessions to develop a research agenda (Facilitator: Rajput)

Second PM Session: (chair: Doug Oard)
- Report--outs from the three breakout sessions (by group-selected reporters)
- Future directions panel discussion (Luz Quiroga, Nitendra Rajput, William Webber)

No--Host Evening Events:
- Drinks: 5:30 PM Hula Grill (sunset is 5:54 PM)
- Dinner: 6:30 PM Hula Grill

Friday, November 2, 2012

WIDM - Twelfth International Workshop on Web Information and Data Management
Room: Napili

8:45-9:00 Welcome
9:00-10:00 Keynote
- Search Beyond the Web: Data from Social Networks and Native Apps. Maria Grineva.

10:00-10:30 Web Data I
- Modeling Topic Trends on the Social Web Using Temporal Signatures. Laura Christiansen, Thomas Schimoler, Robin Burke and Bamshad Mobasher.

10:30-11:00 Coffee break
11:00-12:30 Web Data II
- XPath satisfiability with downward and sibling axes is tractable under most of real-world DTDs. Yasunori Ishihara, Kenji Hashimoto, Shogo Shimizu and Toru Fujiwara.
- A Multi-layer Data Representation of Trajectories in Social Networks Based on Points of Interest. Reinaldo Braga, Ali Tahir, Michela Bertolotto and Hervé Martin.
- A Distributed Index for Efficient Parallel Top-k Keyword Search on Massive Graphs. Ming Zhong and Mengchi Liu.

12:30-14:00 Lunch break
14:00-15:30 Web Context
- Managing Analysis Context. Hua Li and Rafael Alonso.
- Using Social Tags to Infer Context in Hybrid Music Recommendation. Negar Hariri, Bamshad Mobasher and Robin Burke.
- SNOPS: a smart environment for Cultural Heritage applications. Vincenzo Moscato, Antonio Picariello, Angelo Chianese, Flora Amato and Giancarlo Sperlì.

15:30-16:00 Coffee break
16:00-17:30 Web Information Engineering
- Web Crawler Middleware for Search Engine Digital Libraries: A Case Study for CiteSeerX. Jian Wu, Pradeep Teregowda, Madian Khabsa, Douglas Jordan and C. Lee Giles.
- TitleFinder: Extracting the Headline of News Web Pages based on Cosine Similarity and Overlap Scoring Similarity. Hadi Mohammadzadeh, Thomas Gottron, Franz Schweiggert and Gerhard Heyer.
- M3D: A Tool for the Model Driven Development of Web Applications. Mario Luca Bernardi, Marta Cimitile, Giuseppe Di Lucca and Fabrizio Maria Maggi.

17:30-17:45 Closing remarks

Friday, November 2, 2012

DOLAP - Fifteenth International Workshop on Data Warehousing and OLAP
Room: Kula + Hana

8:30-8:45 Workshop Welcome and Introduction. Matteo Golfarelli
8:45-9:30 Invited Talk. Chair: Il-Yeol Song
- Kostamaa, Pekka – Teradata. Efficient Big Data Analytics using SQL and Map-Reduce

9:30-10:45 Session 1: OLAP Query processing and Trends. Chair: Alkis Simitsis
- Bernd Neumayr, Stefan Anderlik and Michael Schrefl. Towards Ontology-based OLAP: Datalog-based Reasoning over Multidimensional Ontologies (25 mins)
- Patrick Marcel, Rokia Missaoui and Stefano Rizzi. Towards Intensional Answers to OLAP Queries for Analytical Sessions (25 mins)
- Carlos Garcia-Alvarado and Carlos Ordonez. Query Processing on Cubes with Dimension Ontologies (25 mins)

11:05-13:00 Session 2: Data Warehouse Design and Maintainability. Chair: Alfredo Cuzzocrea
- Petar Jovanovic, Oscar Romero, Alkis Simitsis and Alberto Abello. ORE: An Iterative Approach to the Design and Evolution of Multi-Dimensional Schemas (25 mins)
- Svetlana Mansmann, Nafees Ur Rehman, Andreas Weiler and Marc H Scholl. Discovering OLAP dimensions in semi-structured data (25 mins)
- Nicolas Prat, Imen Megdiche and Jacky Akoka. Multidimensional Models Meet the Semantic Web: Defining and Reasoning on OWL-DL Ontologies for OLAP (25 mins)
- Alejandro Maté, Juan Trujillo, Elisa De Gregorio and Il-Yeol Song. Improving the Maintainability of Data Warehouse Designs: Modeling Relationships between Sources and Requirements (25 mins)
- Stefan Berger and Michael Schrefl. FedDW Global Schema Architect - UML-based Design Tool for the Integration of Logical Data Mart Schemas (16 mins)

13:00-14:00 Lunch break
14:00-15:45 Session 3: Performance and Benchmarking. Chair: Carlos Ordonez
- Stephan Mueller. An In-Depth Analysis of Data Aggregation Cost Factors in a Columnar In-Memory Database (25 mins)
- Chantola Kit, Marouane Hachicha and J. Darmont. Benchmarking Summarizability Processing in XML Warehouses with Complex Hierarchies(16 mins)
- Craig Stanfill. Type 2 Slowly Changing Dimensions: A Case Study Using the Co>Operating System. (16 mins)
- Jianting Zhang, Simin You and Le Gruenwald. High-Performance Online Spatial and Temporal Aggregations on Multi-core CPUs and Many-Core GPUs (16 mins)
- Doulkifli Boukraa, Omar Boussaid, Fadila Bentayeb and Djamel Eddine Zegour. Managing a Fragmented XML Data Cube with Oracle and Timesten (16 mins)
- Arian Baer and Lukasz Golab. Towards Benchmarking Stream Data Warehouses (16 mins)

16:05-17:30 Session 4: Warehousing complex data. Chair: Patrick Marcel

- Elio Masciari. Warehousing and Querying Trajectory Data Streams With Error Estimation (25 mins)
- Michel De Rougemont and Phuong Thao Cao. Approximate Answers to OLAP Queries on Streaming Data Warehouses (25 mins)
- Alfredo Cuzzocrea and Paolo Serafino. Enhanced Clustering of Complex Database Objects in the ClustCube Framework (16 mins)
- Mu Yin, Bin Wu and Zengfeng Zeng. HMGraph OLAP: A Novel Framework for Multi-dimensional Heterogeneous Network Analysis (16 mins)