Keynotes

1  Information at Your Fingertips: Only a Dream for Enterprises?
   (Surajit Chaudhuri)

5  Data Crowdsourcing: Is it for Real?
   (Hector Garcia-Molina)

Research Session 1: Data Integration

6  Cleaning Uncertain Data with a Noisy Crowd
   (Chen Jason Zhang, Lei Chen, Yongxin Tong, Zheng Liu)

18 Proof Positive and Negative in Data Cleaning
   (Matteo Interlandi, Nan Tang)

30 Cleaning Structured Event Logs: A Graph Repair Approach
   (Jianmin Wang, Shaoxu Song, Xuemin Lin, Xiaochen Zhu, Jian Pei)

42 Query-Time Record Linkage and Fusion over Web Databases
   (El Kindi Rezig, Eduard C. Dragut, Mourad Ouzzani, Ahmed K. Elmagarmid)

Research Session 2: Data Privacy and Security 1

54 Preserving Privacy in Social Networks Against Connection Fingerprint Attacks
   (Yazhe Wang, Baihua Zheng)

66 Efficient Secure Similarity Computation on Encrypted Trajectory Data
   (An Liu, Kai Zheng, Lu Li, Guanfeng Liu, Lei Zhao, Xiaofang Zhou)

78 Privacy-Aware Dynamic Feature Selection
   (Erman Pattuk, Murat Kantarcioglu, Huseyin Ulusoy, Bradley Malin)

89 Scaling up Copy Detection
   (Xian Li, Xin Luna Dong, Kenneth B. Lyons, Wei yi Meng, Divesh Srivastava)

[Search]
Research Session 3: Distributed Storage and Processing

101 Chronos: An Elastic Parallel Framework for Stream Benchmark Generation and Simulation
   (Ling Gu, Minqi Zhou, Zhenjie Zhang, Ming-Chien Shan, Aoying Zhou, Marianne Winslett)

113 PABIRS: A Data Access Middleware for Distributed File Systems
   (Sai Wu, Gang Chen, Xianke Zhou, Zhenjie Zhang, Anthony K.H. Tung, Marianne Winslett)

125 Scalable Distributed Transactions Across Heterogeneous Stores
   (Akon Dey, Alan Fekete, Uwe Röhm)

137 The Power of Both Choices: Practical Load Balancing for Distributed Stream Processing Engines
   (Muhammad Anis Uddin Nasir, Gianmarco De Francisci Morales, David García-Soriano, Nicolaos Kourtellis, Marco Serafini)

Research Session 4: Graph and Timeseries Mining

149 Multicore Triangle Computations Without Tuning
   (Julian Shun, Kanat Tangwongsan)

161 Making Pattern Queries Bounded in Big Graphs
   (Yang Cao, Wenfei Fan, Jinpeng Huai, Ruizhe Huang)

173 Piecewise Linear Approximation of Streaming Time Series Data with Max-Error Guarantees
   (Ge Luo, Ke Yi, Siu-Wing Cheng, Zhenguo Li, Wei Fan, Cheng He, Yadong Mu)

185 On Historical Diagnosis of Sensor Streams
   (Charu C. Aggarwal, Philip S. Yu)
Research Session 5: Web Search and Crowdsourcing

195  Comprehensive and Reliable Crowd Assessment Algorithms
     (Manas Joglekar, Hector Garcia-Molina, Aditya Parameswaran)

207  Diversified Caching for Replicated Web Search Engines
     (Chuanfei Xu, Bo Tang, Man Lung Yiu)

219  Entity Resolution with Crowd Errors
     (Vasilis Verroios, Hector Garcia-Molina)

231  Result Selection and Summarization for Web Table Search
     (Thanh Tam Nguyen, Quoc Viet Hung Nguyen, Matthias Weidlich, Karl Aberer)

Research Session 6: Top-k and Pattern Mining

243  Mining Maximal Cliques from an Uncertain Graph
     (Arko Provo Mukherjee, Pan Xu, Srikanta Tirthapura)

255  Temporal Spatial-Keyword Top-k Publish/Subscribe
     (Lisi Chen, Gao Cong, Xin Cao, Kian-Lee Tan)

267  Finding Top-k Local Users in Geo-Tagged Social Media Data
     (Jinling Jiang, Hua Lu, Bin Yang, Bin Cui)

279  Answering Why-Not Questions on Spatial Keyword Top-k Queries
     (Lei Chen, Xin Lin, Haibo Hu, Christian S. Jensen, Jianliang Xu)
### Research Session 7: Query Processing 1

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>291</td>
<td><strong>Accelerating Aggregation Using Intra-Cycle Parallelism</strong></td>
<td>Zi Qiang Feng, Eric Lo</td>
</tr>
<tr>
<td>303</td>
<td><strong>Efficient and Scalable Trie-Based Algorithms for Computing Set Containment Relations</strong></td>
<td>Yongming Luo, George H.L. Fletcher, Jan Hidders, Paul De Bra</td>
</tr>
<tr>
<td>315</td>
<td><strong>Smooth Scan: Statistics-Oblivious Access Paths</strong></td>
<td>Renata Borovica-Gajic, Stratos Idreos, Anastasia Ailamaki, Marcin Zukowski, Campbell Fraser</td>
</tr>
<tr>
<td>327</td>
<td><strong>Progressive Diversification for Column-Based Data Exploration Platforms</strong></td>
<td>Hina A. Khan, Mohamed A. Sharaf</td>
</tr>
</tbody>
</table>

### Research Session 8: Graph Query Processing and Systems

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>339</td>
<td><strong>Asymmetric Structure-Preserving Subgraph Queries for Large Graphs</strong></td>
<td>Zhe Fan, Byron Choi, Jianliang Xu, Sourav S. Bhowmick</td>
</tr>
<tr>
<td>351</td>
<td><strong>Multi-Constrained Graph Pattern Matching in Large-Scale Contextual Social Graphs</strong></td>
<td>Guanfeng Liu, Kai Zheng, Yan Wang, Mehmet A. Orgun, An Liu, Lei Zhao, Xiao Fang Zhou</td>
</tr>
<tr>
<td>363</td>
<td><strong>LLAMA: Efficient Graph Analytics Using Large Multiversioned Arrays</strong></td>
<td>Peter Macko, Virendra J. Marathe, Daniel W. Margo, Margo I. Seltzer</td>
</tr>
<tr>
<td>375</td>
<td><strong>Answering Regular Path Queries on Workflow Provenance</strong></td>
<td>Xiaocheng Huang, Zhuowei Bao, Susan B. Davidson, Tova Milo, Xiaojie Yuan</td>
</tr>
</tbody>
</table>
Research Session 9: Keyword and Top-K Processing

387  Diversified Top-K Clique Search
     (Long Yuan, Lu Qin, Xuemin Lin, Lijun Chang, Wenjie Zhang)

399  Ranking Candidate Networks of Relations to Improve Keyword Search over Relational Databases
     (Pericles de Oliveira, Altigran da Silva, Edleno de Moura)

411  Meaningful Keyword Search in Relational Databases with Large and Complex Schema
     (Mehdi Kargar, Aijun An, Nick Cercone, Parke Godfrey, Jaroslaw Szlichta, Xiaohui Yu)

423  Interactive Top-k Spatial Keyword Queries
     (Kai Zheng, Han Su, Bolong Zheng, Shuo Shang, Jiajie Xu, Jiajun Liu, Xiaofang Zhou)

Research Session 10: Query Processing 2

435  Transaction Processing on Confidential Data Using Cipherbase
     (Arvind Arasu, Ken Eguro, Manas Joglekar, Raghav Kaushik, Donald Kossmann, Ravi Ramamurthy)

447  Efficient Structural Bulk Updates on the Pre/Dist/Size XML Encoding
     (Lukas Kircher, Michael Grossniklaus, Christian Grün, Marc H. Scholl)

459  Linear Path Skylines in Multicriteria Networks
     (Michael Shekelyan, Gregor Jossé, Matthias Schubert)

471  Bi-Temporal Timeline Index: A Data Structure for Processing Queries on Bi-Temporal Data
     (Martin Kaufmann, Peter M. Fischer, Norman May, Chang Ge, Anil K. Goel, Donald Kossmann)

[Search]
Research Session 11: Strings and Texts

483  Dish Comment Summarization Based on Bilateral Topic Analysis
     (Rong Zhang, Zhenjie Zhang, Xiaofeng He, Aoying Zhou)

495  Short Text Understanding Through Lexical-Semantic Analysis
     (Wen Hua, Zhongyang Wang, Haixun Wang, Kai Zheng, Xiaofang Zhou)

507  Generating Reading Orders over Document Collections
     (Georgia Koutrika, Lei Liu, Steve Simske)

519  Two Birds with One Stone: An Efficient Hierarchical Framework for Top-k and
     Threshold-Based String Similarity Search
     (Jin Wang, Guoliang Li, Dong Deng, Yong Zhang, Jianhua Feng)

Research Session 12: Recommender Systems

531  The Safest Path via Safe Zones
     (Saad Aljubayrin, Jianzhong Qi, Christian S. Jensen, Rui Zhang, Zhen He, Zeyi Wen)

543  Personalized Route Recommendation Using Big Trajectory Data
     (Jian Dai, Bin Yang, Chenjuan Guo, Zhiming Ding)

555  DiSCern: A Diversified Citation Recommendation System for Scientific Queries
     (Tanmoy Chakraborty, Natwar Modani, Ramasuri Narayaman, Seema Nagar)

567  A General Graph-Based Model for Recommendation in Event-Based Social Networks
     (Tuan-Anh Nguyen Pham, Xutao Li, Gao Cong, Zhenjie Zhang)
Research Session 13: Similarity Search and Join

579 Quick-Motif: An Efficient and Scalable Framework for Exact Motif Discovery  
(Yuhong Li, Leong Hou U, Man Lung Yiu, Zhiguo Gong)

591 Efficient Metric Indexing for Similarity Search  
(Lu Chen, Yunjun Gao, Xinhan Li, Christian S. Jensen, Gang Chen)

603 Scalable SimRank Join Algorithm  
(Takanori Maehara, Mitsuru Kusumoto, Ken-ichi Kawarabayashi)

615 Robust Clustering of Multi-Type Relational Data via a Heterogeneous Manifold Ensemble  
(Jun Hou, Richi Nayak)

Research Session 14: Social Network

627 Identifying Users in Social Networks with Limited Information  
(Norases Vesdapunt, Hector Garcia-Molina)

639 Real Time Personalized Search on Social Networks  
(Yuchen Li, Zhifeng Bao, Guoliang Li, Kian-Lee Tan)

651 False Rumors Detection on Sina Weibo by Propagation Structures  
(Ke Wu, Song Yang, Kenny Q. Zhu)

663 Convex Risk Minimization to Infer Networks from Probabilistic Diffusion Data at Multiple Scales  
(Emre Sefer, Carl Kingsford)

[Search]
Research Session 15: Spatial Query Processing

675 SAR: A Sentiment-Aspect-Region Model for User Preference Analysis in Geo-Tagged Reviews
(Kaiqi Zhao, Gao Cong, Quan Yuan, Kenny Q. Zhu)

687 Searchlight: Context-Aware Predictive Continuous Querying of Moving Objects in Symbolic Space
(Kenneth Fuglsang Christensen, Lasse Linnerup Christiansen, Torben Bach Pedersen, Jeppe Pihl)

699 Nearest Neighborhood Search in Spatial Databases
(Dong-Wan Choi, Chin-Wan Chung)

711 A Location-Aware Publish/Subscribe Framework for Parameterized Spatio-Textual Subscriptions
(Huiqi Hu, Yiqun Liu, Guoliang Li, Jianhua Feng, Kian-Lee Tan)

Research Session 16: Stream

723 ChronoStream: Elastic Stateful Stream Computation in the Cloud
(Yingjun Wu, Kian-Lee Tan)

735 Conflict-Aware Event-Participant Arrangement
(Jieying She, Yongxin Tong, Lei Chen, Caleb Chen Cao)

747 PIE: Approximate Interleaving Event Matching over Sequences
(Zheng Li, Tingjian Ge)

759 Window-Chained Longest Common Subsequence: Common Event Matching in Sequences
(Chunyao Song, Tingjian Ge)
Research Session 17: RDF-Processing

771 CliqueSquare: Flat Plans for Massively Parallel RDF Queries  
(François Goasdoué, Zoï Kaoudi, Ioana Manolescu, Jorge-Arnulfo Quiané-Ruiz, Stamatis Zampetakis)

783 INSURE: An Integrated Load Reduction Framework for XML Stream Processing  
(Mingzhu Wei, Elke A. Rundensteiner, Murali Mani)

795 Scalable SPARQL Querying Using Path Partitioning  
(Buwen Wu, Yonglun Zhou, Pingpeng Yuan, Ling Liu, Hai Jin)

807 Executing Queries over Schemaless RDF Databases  
(Güneş Aluç, M. Tamer Özsür, Khuzaima Daudjee, Olaf Hartig)

Research Session 18: System-level Techniques

819 Configurable Hardware-Based Streaming Architecture Using Online Programmable-Blocks  
(Mohammadreza Najafi, Mohammad Sadoghi, Hans-Arno Jacobsen)

831 Network Motif Discovery: A GPU Approach  
(Wenqing Lin, Xiaokui Xiao, Xing Xie, Xiaoli Li)

843 Automatic Tuning of Bag-of-Tasks Applications  
(Majed Sahli, Essam Mansour, Tariq Alturkestani, Panos Kalnis)

855 Cache-Oblivious Scheduling of Shared Workloads  
(Arian Bär, Łukasz Golab, Stefan Ruehrup, Mirko Schiavone, Pedro Casas)

[Search]
## Research Session 19: Foundations of Query Processing

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>867</td>
<td>Optimizing Recursive Queries with Monotonic Aggregates in DeALS</td>
<td>Alexander Shkapsky, Mohan Yang, Carlo Zaniolo</td>
</tr>
<tr>
<td>879</td>
<td>Size-Constrained Weighted Set Cover</td>
<td>Lukasz Golab, Flip Korn, Feng Li, Barna Saha, Divesh Srivastava</td>
</tr>
<tr>
<td>891</td>
<td>Online Frequent Episode Mining</td>
<td>Xiang Ao, Ping Luo, Chengkai Li, Fuzhen Zhuang, Qing He</td>
</tr>
<tr>
<td>903</td>
<td>Dynamic Programming: The Next Step</td>
<td>Marius Eich, Guido Moerkotte</td>
</tr>
</tbody>
</table>

## Research Session 20: Graph Sampling and Matching

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>915</td>
<td>Finding Dense and Connected Subgraphs in Dual Networks</td>
<td>Yubao Wu, Ruoming Jin, Xiaofeng Zhu, Xiang Zhang</td>
</tr>
<tr>
<td>927</td>
<td>On Random Walk Based Graph Sampling</td>
<td>Rong-Hua Li, Jeffrey Xu Yu, Lu Qin, Rui Mao, Tan Jin</td>
</tr>
<tr>
<td>939</td>
<td>A Tale of Three Graphs: Sampling Design on Hybrid Social-Affiliation Networks</td>
<td>Junzhou Zhao, John C.S. Lui, Don Towsley, Pinghui Wang, Xiaohong Guan</td>
</tr>
<tr>
<td>951</td>
<td>Hierarchical In-Network Attribute Compression via Importance Sampling</td>
<td>Arlei Silva, Petko Bogdanov, Ambuj K. Singh</td>
</tr>
</tbody>
</table>
Research Session 21: Trajectories

963  📚 Making Sense of Trajectory Data: A Partition-and-Summarization Approach  
     (Han Su, Kai Zheng, Kai Zeng, Jiamin Huang, Shazia Sadiq, Nicholas Jing Yuan,  
     Xiaofang Zhou)

975  📚 Approximate Keyword Search in Semantic Trajectory Database  
     (Bolong Zheng, Nicholas Jing Yuan, Kai Zheng, Xing Xie, Shazia Sadiq, Xiaofang Zhou)

987  📚 Bounded Quadrant System: Error-Bounded Trajectory Compression on the Go  
     (Jiajun Liu, Kun Zhao, Philipp Sommer, Shuo Shang, Brano Kusy, Raja Jurdak)

999  📚 Indexing and Matching Trajectories Under Inconsistent Sampling Rates  
     (Sayan Ranu, Deepak P., Aditya D. Telang, Prasad Deshpande, Sriram Raghavan)

Research Session 22: Data Privacy and Security 2

1011 📚 A Hybrid Private Record Linkage Scheme: Separating Differentially Private Synopses  
       from Matching Records  
       (Jianneng Cao, Fang-Yu Rao, Elisa Bertino, Murat Kantarcioglu)

1023 📚 Conservative or Liberal? Personalized Differential Privacy  
       (Zach Jorgensen, Ting Yu, Graham Cormode)

1035 📚 Differentially Private Frequent Sequence Mining via Sampling-Based Candidate Pruning  
       (Shengzhi Xu, Sen Su, Xiang Cheng, Zhengyi Li, Li Xiong)

[Search]
<table>
<thead>
<tr>
<th>Research Session 23: MapReduce</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1047</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>1059</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>1071</strong></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Research Session 24: Query Processing 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1083</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>1095</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>1107</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>1119</strong></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Research Session 25: Graph-based Data Management</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1131</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>1143</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>1155</strong></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Research Session 26: Cloud Computing

1167 PerfAugur: Robust Diagnostics for Performance Anomalies in Cloud Services
(Sudip Roy, Arnd Christian König, Igor Dvorkin, Manish Kumar)

1179 LDV: Light-Weight Database Virtualization
(Quan Pham, Tanu Malik, Boris Glavic, Ian Foster)

1191 Efficient Sample Generation for Scalable Meta Learning
(Sebastian Schelter, Juan Soto, Volker Markl, Douglas Burdick, Berthold Reinwald, Alexandre Evfimievski)

Research Session 27: Indexing

1203 High Performance Temporal Indexing on Modern Hardware
(David B. Lomet, Faisal Nawab)

1215 Predictive Tree: An Efficient Index for Predictive Queries on Road Networks
(Abdeltawab M. Hendawi, Jie Bao, Mohamed F. Mokbel, Mohamed Ali)

1227 A Comparison of Adaptive Radix Trees and Hash Tables
(Victor Alvarez, Stefan Richter, Xiao Chen, Jens Dittrich)
# Industry Session 1: Main Memory Databases

1239  **Evolving the Architecture of SQL Server for Modern Hardware Trends**  
*(Per-Ake Larson, Eric N. Hanson, Mike Zwilling)*

1246  **In-Memory BLU Acceleration in IBM’s DB2 and dashDB: Optimized for Modern Workloads and Hardware Architectures**  
*(Ronald Barber, Guy Lohman, Vijayshankar Raman, Richard Sidle, Sam Lightstone, Berni Schiefer)*

1253  **Oracle Database In-Memory: A Dual Format In-Memory Database**  
*(Tirthankar Lahiri, Shasank Chavan, Maria Colgan, Dinesh Das, Amit Ganesh, Mike Gleeson, Sanket Hase, Allison Holloway, Jesse Kamp, Teck-Hua Lee, Juan Loaiza, Neil Macnaughton, Vineet Marwah, Niloy Mukherjee, Atrayee Mullick, Sujatha Muthulingam, Vivekanandhan Raja, Marty Roth, Ekrem Soylemez, Mohamed Zait)*

1259  **Towards a Web-Scale Data Management Ecosystem Demonstrated by SAP HANA**  
*(Franz Faerber, Jonathan Dees, Martin Weidner, Stefan Baeuerle, Wolfgang Lehner)*

# Industry Session 2: Main Memory and Stream Databases

1268  **“Anti-Caching”-based Elastic Memory Management for Big Data**  
*(Hao Zhang, Gang Chen, Beng Chin Ooi, Weng-Fai Wong, Shensen Wu, Yubin Xia)*

1280  **Supporting Hierarchical Data in SAP HANA**  
*(Robert Brunel, Jan Finis, Gerald Franz, Norman May, Alfons Kemper, Thomas Neumann, Franz Faerber)*

1292  **AEDSMS: Automotive Embedded Data Stream Management System**  
*(Akihiro Yamauchi, Yukikazu Nakamoto, Kenya Sato, Yoshiharu Ishikawa, Yousuke Watanabe, Shinya Honda, Hiroaki Takada)*

[Search]
Industry Session 3: Big Data

1304  Accelerating Big Data Analytics with Collaborative Planning in Teradata Aster 6  
(Aditi Pandit, Derrick Kondo, David Simmen, Anjali Norwood, Tongxin Bai)

1316  FCCE: Highly Scalable Distributed Feature Collection and Correlation Engine for Low  
Latency Big Data Analytics  
(Douglas L. Schales, Xin Hu, Jiyong Jang, Reiner Sailer, Marc Ph. Stoecklin, Ting Wang)

1328  Seamlessly Integrating Disk and Tape in a Multi-Tiered Distributed File System  
(Ioannis Koltsidas, Slavisa Sarafijanovic, Martin Petermann, Nils Haustein, Harald Seipp,  
Robert Haas, Jens Jelitto, Thomas Weigold, Edwin Childers, David Pease, Evangelos Eleftheriou)

1340  DualTable: A Hybrid Storage Model for Update Optimization in Hive  
(Songlin Hu, Wantao Liu, Tilmann Rabl, Shuo Huang, Ying Liang, Zheng Xiao,  
Hans-Arno Jacobsen, Xubin Pei, Jiye Wang)

Industry Session 4: Spatial and Temporal Data

1352  SpatialHadoop: A MapReduce Framework for Spatial Data  
(Ahmed Eldawy, Mohamed F. Mokbel)

1364  Time Dependent Transportation Network Models  
(Petko Bakalov, Erik Hoel, Wee-Liang Heng)

1376  Growing the Charging Station Network for Electric Vehicles with Trajectory Data  
Analytics  
(Yanhua Li, Jun Luo, Chi-Yin Chow, Kam-Lam Chan, Ye Ding, Fan Zhang)

1388  Clustering to Forecast Sparse Time-Series Data  
(Abhay Jha, Shubhankar Ray, Brian Seaman, Inderjit S. Dhillon)

[Search]
Demo 1

1400  🎥 Advanced Analytics on SAP HANA: Churn Risk Scoring Using Call Network Analysis
   (Shane Bracher, Mark Holmes, Liam Mischewski, Asadul Islam, Michael McClenaghan, Daniel Ricketts, Glenn Neuber, Hoyoung Jeung, Priya Vijayarajendran)

1404  🎥 PrivGeoCrowd: A Toolbox for Studying Private Spatial Crowdsourcing
   (Hien To, Gabriel Ghinita, Cyrus Shahabi)

1408  🎥 PandaSearch: A Fine-Grained Academic Search Engine for Research Documents
   (Feiran Huang, Jia Li, Jiaheng Lu, Tok Wang Ling, Zhaoan Dong)

1412  🎥 EcoSky: Reducing Vehicular Environmental Impact Through Eco-Routing
   (Chenjuan Guo, Bin Yang, Ove Andersen, Christian S. Jensen, Kristian Torp)

1416  🎥 Demonstration of Taghreed: A System for Querying, Analyzing, and Visualizing Geotagged Microblogs
   (Amr Magdy, Louai Alarabi, Saif Al-Harthi, Mashaal Musleh, Thanaa M. Ghanem, Sohaib Ghani, Saleh Basalamah, Mohamed F. Mokbel)

1420  🎥 CROWN: A Context-Aware Recommender for Web News
   (Shaoming Wang, Benyou Zou, Cuiping Li, Kankan Zhao, Qiang Liu, Hong Chen)

1424  🎥 LearningAssistant: A Novel Learning Resource Recommendation System
   (Lei Liu, Georgia Koutrika, Shanchan Wu)

1428  🎥 A Framework for Computation of Popular Paths from Crowdsourced Data
   (Gregor Jossé, Maximilian Franzke, Georgios Skoumas, Andreas Züfle, Mario A. Nascimento, Matthias Renz)

1432  🎥 CliqueSquare in Action: Flat Plans for Massively Parallel RDF Queries
   (Benjamin Djahandideh, François Goasdoué, Zoi Kaoudi, Ioana Manolescu, Jorge-Arnulfo Quijano-Ruíz, Stamatis Zampetakis)

1436  🎥 MARS: A Multi-Aspect Recommender System for Point-of-Interest
   (Xin Li, Guandong Xu, Enhong Chen, Lin Li)

[Search]  Demo 1 continues next page …
Demo 1 continued …

1440 AllegatorTrack: Combining and Reporting Results of Truth Discovery from Multi-Source Data  
(Dalia Attia Waguih, Naman Goel, Hossam M. Hammady, Laure Berti-Equille)

1444 A Demonstration of Shahed: A MapReduce-Based System for Querying and Visualizing Satellite Data  
(Ahmed Eldawy, Saif Al-Harthi, Abdulhadi Alzaidy, Anas Daghistani, Sohaib Ghani,  
Saleh Basalamah, Mohamed F. Mokbel)

1448 PGWinFunc: Optimizing Window Aggregate Functions in PostgreSQL and its Application for Trajectory Data  
(Jiansong Ma, Yu Cao, Xiaoling Wang, Chaoyong Wang, Cheqing Jin, Aoying Zhou)

1452 CDR-To-MoVis: Developing a Mobility Visualization System from CDR Data  
(Manoranjan Dash, Kee Kiat Koo, James Decraene, Ghim-Eng Yap,  
Shonali Priyadarsini Krishnaswamy, Wei Wu, Joao Bartolo Gomes, Amy Shi-Nash, Xiaoli Li)

1456 Understanding Business Trends from Data Evolution with Tornado  
(Azhar Ait Ouassarah, Nicolas Averseng, Xavier Fournet, Jean-Marc Petit, Romain Revol,  
Vasile-Marian Scuturici)

Demo 2

1460 Enjoy FRDM — Play with a Schema-Flexible RDBMS  
(Hannes Voigt, Patrick Damme, Wolfgang Lehner)

1464 ControVol: A Framework for Controlled Schema Evolution in NoSQL Application Development  
(Stefanie Scherzinger, Thomas Cerqueus, Eduardo Cunha de Almeida)

1468 The XDa-TA System for Automated Grading of SQL Query Assignments  
(Amol Bhangdiya, Bikash Chandra, Biplab Kar, Bharath Radhakrishnan,  
K.V. Maheshwara Reddy, Shetal Shah, S. Sudarshan)

[Search] Demo 2 continues next page …
Demo 2 continued …

1472 Querying Databases by Snapping Blocks
(Yasin N. Silva, Jaime Chon)

1476 SMAS: A Smart Meter Data Analytics System
(Xiufeng Liu, Lukasz Golab, Ihab F. Ilyas)

1480 ViSUAL: An HCI-inspired Simulator for Blending Visual Subgraph Query Construction and Processing
(Sourav S. Bhowmick, Huey Eng Chua, Benji Thian, Byron Choi)

1484 selP: Selective Tracking and Presentation of Data Provenance
(Daniel Deutch, Amir Gilad, Yuval Moskovitch)

1488 SMART: A Tool for Analyzing and Reconciling Schema Matching Networks
(Quoc Viet Hung Nguyen, Thanh Tam Nguyen, Vinh Tuan Chau, Tri Kurniawan Wijaya, Zoltán Miklós, Karl Aberer, Avigdor Gal, Matthias Weidlich)

1492 PIGEON: Progress Indicator for Subgraph Queries
(Xiaojing Xie, Zhe Fan, Byron Choi, Peipei Yi, Sourav S. Bhowmick, Shuigeng Zhou)

1496 Enabling Generic Keyword Search over Raw XML Data
(Manoj K. Agarwal, Krithi Ramamritham)

1500 DAVINCI: Data-Driven Visual Interface Construction for Subgraph Search in Graph Databases
(Jinbo Zhang, Sourav S. Bhowmick, Hong H. Nguyen, Byron Choi, Feida Zhu)

1504 Elaps: An Efficient Location-Aware Pub/Sub System
(Long Guo, Lu Chen, Dongxiang Zhang, Guoliang Li, Kian-Lee Tan, Zhifeng Bao)

1508 A Graph-Based RDF Triple Store
(Xuchuan Shen, Lei Zou, M. Tamer Özsu, Lei Chen, Youhuan Li, Shuo Han, Dongyan Zhao)

1512 Interactive Preference-Aware Query Optimization
(N.R. Ong, S.E. Rojcewicz, N.L. Farnan, A.J. Lee, P.K. Chrysanthis, Ting Yu)

[Search]
Tutorials

1516  Open Data Challenges at Facebook
      (Nathan Bronson, Thomas Lento, Janet L. Wiener)
1520  The Solid-State Drive Technology, Today and Tomorrow
      (Sangyeun Cho, Sanghoan Chang, Insoon Jo)
1523  Blind Men and an Elephant — Coalescing Open-Source, Academic, and Industrial
      Perspectives on BigData
      (Chris Douglas, Carlo Curino)
1527  Data-Driven Crowdsourcing: Management, Mining, and Applications
      (Lei Chen, Dongwon Lee, Tova Milo)
1530  How to Stop Under-Utilization and Love Multicores
      (Anastasia Ailamaki, Erietta Liarou, Pınar Tözün, Danica Porobic, Iraklis Psaroudakis)
1534  Inferencing in Information Extraction: Techniques and Applications
      (Denilson Barbosa, Haixun Wang, Cong Yu)
1538  Goals in Social Media, Information Retrieval and Intelligent Agents
      (Dimitra Papadimitrou, Yannis Velegrakis, Georgia Koutrika, John Mylopoulos)
1541  Reasoning on Web Data: Algorithms and Performance
      (François Goasdoué, Ioana Manolescu, Alexandra Roatis)
1545  DBMS on Modern Storage Hardware
      (Ilia Petrov, Robert Gottstein, Sergej Hardock)
Application

1549  Understanding Computer Usage Evolution  
(David C. Anastasiu, Al M. Rashid, Andrea Tagarelli, George Karypis)

1561  STREAMCUBE: Hierarchical Spatio-Temporal Hashtag Clustering for Event Exploration over the Twitter Stream  
(Wei Feng, Chao Zhang, Wei Zhang, Jiawei Han, Jianyong Wang, Charu C. Aggarwal, Jianbin Huang)

(Siarhei Bykau, Flip Korn, Divesh Srivastava, Yannis Velegrakis)

1585  SHAHED: A MapReduce-Based System for Querying and Visualizing Spatio-Temporal Satellite Data  
(Ahmed Eldawy, Mohamed F. Mokbel, Saif Al-Harthi, Abdulhadi Alzaidy, Kareem Tarek, Sohaib Ghani)

Panels

1597  Big Data: Old Wine in New Bottle?  
(Rakesh Agrawal)

1600  Data Engineering in Asia: Unique Technical Challenges and Opportunities  
(Rakesh Agrawal)