

Proceedings

ICSC 2007

**International Conference
on Semantic Computing**

***17-19 September 2007
Irvine, California***



**Los Alamitos, California
Washington • Tokyo**



TABLE OF CONTENTS

PARALLEL SESSION 1

TUNLP (1): TEXT UNDERSTANDING AND NATURAL LANGUAGE PROCESSING

DR. RADU FLORIAN, IBM RESEARCH, USA

| | |
|---|----|
| Robust Classification of Dialog Acts from the Transcription of Utterances | 1 |
| <i>M.S. Sorower, M. Yeasin</i> | |
| Timelines from Text: Identification of Syntactic Temporal Relations | 9 |
| <i>S. Bethard, J.H. Martin, S. Klingenstein</i> | |
| Lexical and Discourse Analysis of Online Chat Dialog | 17 |
| <i>E.N. Forsyth and C.H. Martell</i> | |
| Automatic Generation of Multi-Modal Dialogue from Text Based on Discourse Structure Analysis | 25 |
| <i>H. Prendinger, P. Piwek, M. Ishizuka</i> | |

SSE (1): SERVICES AND SOFTWARE ENGINEERING

DR. RAMAZAN SAVAS AYGUN, UNIVERSITY OF ALABAMA, HUNTSVILLE, USA

| | |
|--|----|
| A Software Birthmark Based on Dynamic Opcode n-gram | 33 |
| <i>B. Lu, F. Liu, X. Ge, B. Liu, X. Luo</i> | |
| A Framework to Build an Object Oriented Mathematical Tool with Computer Algebra System (CAS) Capability | 41 |
| <i>D.M.A. Hussain, O.T. Buss, F. Kiros, N. Wichmann, B. Selvarajah, Z. Ahmad</i> | |
| Specifying and Verifying Cases Retrieval System Combining Event B and Spin | 49 |
| <i>H. Gao, Z. Qin, L. Shao, X. Heng</i> | |
| Procuring Requirements for ERP Software Based on Semantic | 57 |
| <i>J. Gao, L. Zhang, W. Jiang</i> | |

SS (1): CONTENT-BASED RETRIEVAL AND SEMANTIC ANNOTATION OF MULTIMEDIA CONTENTS

DR. BALAKRISHNAN PRABHAKARAN, UNIVERSITY OF TEXAS AT DALLAS, USA

| | |
|---|----|
| Segmenting Photo Streams in Events Based on Optical Metadata | 65 |
| <i>B. Gong, R. Jain</i> | |
| A Hybrid Approach to Improving Semantic Extraction of News Video | 73 |
| <i>A.G. Hauptmann, M.Y. Chen, M. Christel, W.H. Lin, J. Yang</i> | |
| CDIP: Collection-Driven, yet Individuality-Preserving Automated Blog Tagging | 81 |
| <i>J.W. Kim, K.S. Candan, J. Tatemura</i> | |

| | |
|---|-----------|
| Eventory – An Event Based Media Repository | 89 |
| <i>X. Wang, S. Mamadgi, A. Thekdi, A. Kelliher, H. Sundaram</i> | |

PARALLEL SESSION 2

KEDM (1): KNOWLEDGE ENGINEERING AND DATA MINING

DR. GERALD FRIEDLAND, ICSI BERKELEY, USA

| | |
|--|------------|
| Knowledge Integration in OpenWorlds: Utilizing the Mathematics of Hierarchical Structure | 97 |
| <i>C.A. Joslyn, K.M. Verspoor, D.D.G. Gessler</i> | |
| An Interactive Diagnosis and Repair of OWL Ontology | 105 |
| <i>L. Lu, N. Parameswaran, P.K. Ray</i> | |
| Sub-Symbolic Semantic Layer in Cyc for Intuitive Chat-Bots | 113 |
| <i>G. Pilato, A. Augello, G. Vassallo, S. Gaglio</i> | |
| A Method for the Construction of a Probabilistic Hierarchical Structure Based on a Statistical Analysis of a Large-scale Corpus | 121 |
| <i>A. Terai, B. Liu, M. Nakagawa</i> | |
| Reasoning within the Extended Fuzzy Description Logics with Restricted Terminological Boxes..... | 129 |
| <i>D. Kang, B. Xu, Y. Li</i> | |

MAVS (1): MULTIMEDIA, AUDIO, AND VISUAL SEMANTICS

PROF. ALEX HAUPTMANN, CARNEGIE MELLON UNIVERSITY, USA

| | |
|---|------------|
| A Novel Probabilistic Approach to Modeling the Pleasure-Arousal-Dominance Content of the Video based on “Working Memory” | 137 |
| <i>S. Arifin, P.Y.K. Cheung</i> | |
| LDA-Based Retrieval Framework for Semantic News Video Retrieval..... | 145 |
| <i>J. Cao, J. Li, Y. Zhang, S. Tang</i> | |
| Semantic Analysis for Topical Segmentation of Videos..... | 151 |
| <i>Y. Park, Y. Li</i> | |
| Automated Summarization of Narrative Video on a Semantic Level..... | 159 |
| <i>T. Tsoneva, M. Barbieri, H. Weda</i> | |
| Who Needs Video Summarization Anyway? | 167 |
| <i>S. Santini</i> | |

SW (1): SEMANTIC WEB

DR. HASIDA KOITI, AIST, JAPAN

PROF. EURO HYVONEN, HELSINKI UNIVERSITY OF TECHNOLOGY, FINLAND

| | |
|--|------------|
| Ontology Evaluation and Ranking using OntoQA..... | 173 |
| <i>S. Tartir, I.B. Arpinar</i> | |
| A Trust Based Methodology for Web Service Selection | 181 |
| <i>S. Galizia, A. Gugliotta, J. Domingue</i> | |

| | |
|---|-----|
| Convergence of Web 2.0 and Semantic Web: A Semantic Tagging and Searching System for Creating and Searching Blogs | 189 |
| <i>G. Hope, T. Wang, S. Barkataki</i> | |
| Semantic Web Infrastructure | 197 |
| <i>J. Dokulil, J. Tykal, J. Yaghob, F. Zavoral</i> | |
| Elements of a National Semantic Web Infrastructure— Case Study Finland on the Semantic Web | 204 |
| <i>E. Hyvonen, K. Viljanen, E. Makela, T. Kauppinen, T. Ruotsalo, O. Valkeapaa, K. Seppala, O. Suominen, O. Alm, R. Lindroos, T. Kansala, R. Henriksson, M. Frosterus, J. Tuominen, R. Sinkkila, J. Kurki</i> | |
| Semantic Locations in Online Communities | 212 |
| <i>M. Gruhn, S. Foll, J. Pontow, D. Linner, I. Radusch</i> | |

PARALLEL SESSION 3:

TUNLP (2): TEXT UNDERSTANDING AND NATURAL LANGUAGE PROCESSING

PROF. CRAIG MARTELL, NAVAL POSTGRADUATE SCHOOL, USA

| | |
|--|-----|
| Toward Spotting the Pedophile: Telling Victim from Predator in Text Chats | 220 |
| <i>N. Pendar</i> | |
| Random-Walk Term Weighting for Improved Text Classification | 227 |
| <i>S. Hassan, R. Mihalcea, C. Banea</i> | |
| Sentence Similarity based on Dynamic Time Warping | 235 |
| <i>X. Liu, Y. Zhou, R. Zheng</i> | |
| Integrating Semantic Knowledge into Text Similarity and Information Retrieval | 242 |
| <i>C. Muller, I. Gurevych, M. Muhlhauser</i> | |
| Cross-Genre Feature Comparisons for Spoken Sentence Segmentation | 250 |
| <i>S. Cuendet, D. Hakkani-Tur, E. Shriberg, J. Fung, B. Favre</i> | |

SSE (2): MODELING, SERVICES AND SOFTWARE ENGINEERING

PROF. ILJA RADUSCH, TECHNISCHE UNIVERSITAT BERLIN, GERMANY

| | |
|--|-----|
| Context-Aware Architecture for Intelligent Application Services in Ubiquitous Computing | 257 |
| <i>J. Chang, H. Lee</i> | |
| Thematic Role Based Generation of UML Models from Real World Requirements | 264 |
| <i>T. Gelhausen, W.F. Tichy</i> | |
| Challenges and Directions for Semantic Communication | 272 |
| <i>V. Rodoplu, S.S. Vadvalkar</i> | |
| Towards Light-weight Semantic Descriptions for Decentralized Service-oriented Systems | 277 |
| <i>H. Pfeiffer, D. Linner, C. Jacob, I. Radusch, S. Steglich</i> | |
| Restraining False Feedbacks in Peer-to-Peer Reputation Systems | 286 |
| <i>Y. Jin, Z. Gu, Z. Ban</i> | |

**SS (2): CLOSING THE SEMANTIC GAP: CONCEPT-BASED MULTIMEDIA
MINING AND RETRIEVAL**
DR. RONG YAN, DR. APOSTOL (PAUL) NATSEV, AND DR. YING LI, IBM, USA

| | |
|--|-----|
| Mapping Query to Semantic Concepts: Leveraging Semantic Indices for Automatic and Interactive Video Retrieval | 293 |
| <i>D. Wang, Z. Wang, X. Li, X. Liu, J. Li, B. Zhang</i> | |
| Multi-Concept Multi-Modality Active Learning for Interactive Video Annotation | 301 |
| <i>M. Wang, X. Hua, Y. Song, J. Tang, L. Dai</i> | |
| Searching Visual Semantic Spaces with Concept Filters | 309 |
| <i>E. Zavesky, Z. Liu, D. Gibbon, B. Shahraray</i> | |
| Collective Media Annotation using Undirected Random Field Models | 317 |
| <i>M. Cooper</i> | |
| Exploring Concept Selection Strategies for Interactive Video Search | 324 |
| <i>M.G. Christel, A.G. Hauptmann</i> | |

PARALLEL SESSION 4
**TUNLP (3): TEXT UNDERSTANDING AND NATURAL LANGUAGE
PROCESSING**
DR. DANIEL BIKEL, IBM RESEARCH, USA

| | |
|--|-----|
| A Graph Modeling of Semantic Similarity Between Words | 332 |
| <i>M.A. Alvarez, S. Lim</i> | |
| Unsupervised Graph-based Word Sense Disambiguation Using Measures of Word Semantic Similarity | 340 |
| <i>R. Sinha, R. Mihalcea</i> | |
| Clustering Using Feature Domain Similarity to Discover Word Senses for Adjectives | 347 |
| <i>N. Tomuro, S.L. Lytinen, K. Kanzaki, H. Isahara</i> | |
| Towards Large-scale High-Performance English Verb Sense Disambiguation by Using Linguistically Motivated Features | 355 |
| <i>J. Chen, D. Dligach, M. Palmer</i> | |

MAVS (2): MULTIMEDIA, AUDIO, AND VISUAL SEMANTICS
DR. ZHU LIU, AT&T LABS, USA

| | |
|---|-----|
| An Efficient Multimodal Language Processor for Parallel Input Strings in Multimodal Input Fusion | 363 |
| <i>Y. Sun, Y. Shi, F. Chen, V. Chung</i> | |
| Smoothing Algorithm for N-Gram Model Using Agglutinative Characteristic of Korean | 371 |
| <i>J. Park, Y. Song, H. Rim</i> | |
| Adding Semantics to Formal Data Specifications to Automatically Generate Corresponding Voice Data-Input Applications | 379 |
| <i>F. Hanna, R. Frost</i> | |
| Multimedia Reasoning with Natural Language Support | 387 |
| <i>S. Dasiopoulou, J. Heinecke, C. Saathoff, M.G. Strintzis</i> | |

| | |
|--|-----|
| Harnessing Language in Mobile Environments | 395 |
| <i>B. Katz, G. Borchardt, S. Felshin, F. Mora</i> | |
| Large-Margin Discriminative Training of Hidden Markov Models for Speech Recognition | 403 |
| <i>D. Yu, L. Deng</i> | |

SS (3): RECOGNIZING ENTITIES AND CONCEPTS IN WEB PAGES AND THE WORLD OF SEMANTICS BEYOND THE SEMANTIC WEB
DR. PEI CAO, GOOGLE, USA

| | |
|--|-----|
| Answering English Questions using Foreign-Language, Semi-Structured Sources | 411 |
| <i>B. Katz, G. Borchardt, S. Felshin, Y. Shen, G. Zaccak</i> | |
| Unrestricted Coreference: Identifying Entities and Events in OntoNotes | 418 |
| <i>S.S. Pradhan, L. Ramshaw, R. Weischedel, J. MacBride, L. Micciulla</i> | |
| N-gram Statistics in English and Chinese: Similarities and Differences | 426 |
| <i>S. Yang, H. Zhu, A. Apostoli, P. Cao</i> | |
| Semantic Interpretation of the Web without the Semantic Web: Toward Business-Aware Web Processors | 433 |
| <i>P.Z. Yeh, D.R. Farina, A. Kass</i> | |
| SA-REST and (S)mashups : Adding Semantics to RESTful Services | 441 |
| <i>J. Lathem, K. Gomadam, A.P. Sheth</i> | |
| The Programmable Web: Agile, Social, and Grassroot Computing | 449 |
| <i>E.M. Maximilien, A. Ranabahu</i> | |
| Khmer POS Tagger: A Transformation-based Approach with Hybrid Unknown Word Handling | 454 |
| <i>C. Nou, W. Kameyama</i> | |

PARALLEL SESSION 5
TUNLP (4): TEXT UNDERSTANDING AND NATURAL LANGUAGE PROCESSING
DR. IRYNA GUREVYCH, DARMSTADT UNIVERSITY OF TECHNOLOGY, GERMANY

| | |
|---|-----|
| If We Want Your Opinion | 462 |
| <i>D.M. Bikel, J. Sorensen</i> | |
| Reliability of Human Annotation of Semantic Roles in Noisy Text | 470 |
| <i>D. Higgins</i> | |
| Determining Modality and Factuality for Text Entailment | 478 |
| <i>R. Sauri, J. Pustejovsky</i> | |
| OntoNotes: A Unified Relational Semantic Representation | 486 |
| <i>S.S. Pradhan, E. Hovy, M. Marcus, M. Palmer, L. Ramshaw, R. Weischedel</i> | |

KEDM (2): KNOWLEDGE ENGINEERING AND DATA MINING
PROF. SHU-CHING CHEN, FLORIDA INTERNATIONAL UNIVERSITY, USA

| | |
|---|-----|
| ASIC: Supervised Multi-class Classification using Adaptive Selection of Information Components | 494 |
| <i>Z. Xie, T. Quirino, M. Shyu, S. Chen</i> | |
| Semantic Clustering of Information Systems' Users with Stochastic Techniques | 502 |
| <i>K. Raftopoulos, N. Papadakis, K. Ntalianis, S. Kollias</i> | |
| Management of Meta Knowledge for RDF Repositories | 510 |
| <i>B. Schueler, S. Sizov, S. Staab</i> | |
| The “Spree” Expert Finding System | 518 |
| <i>F. Metzke, C. Bauckhage, T. Alpcan, K. Dobbrott, C. Clemens</i> | |
| Intelligent Parsing of Scanned Volumes for Web Based Archives | 526 |
| <i>X. Lu, J.Z. Wang, C.L. Giles</i> | |

WS (1): MULTIMEDIA SEMANTICS
DR. MIN CHEN, UNIVERSITY OF MONTANA, USA

| | |
|--|-----|
| FSM-based Model for Spatio-temporal Event Recognition for HCS | 534 |
| <i>W.M. Ahmed, J. Liu, D. Lenz, A. Ghafoor, J.P. Robinson</i> | |
| Semantic Event Extraction Using Neural Network Ensembles | 540 |
| <i>M. Chen, C. Zhang, S. Chen</i> | |
| An Effective and Robust Method for Automatic Bacterial Colony Enumeration | 546 |
| <i>C. Zhang, W. Chen</i> | |

PARALLEL SESSION 6
TUNLP (5): TEXT UNDERSTANDING AND NATURAL LANGUAGE
PROCESSING
DR. PHILIPP CIMIANO, INSTITUT AIFB, UNIVERSITÄT KARLSRUHE,
GERMANY

| | |
|--|-----|
| Nested Named Entity Recognition in Historical Archive Text | 552 |
| <i>K. Byrne</i> | |
| Modeling Discriminative Global Inference | 560 |
| <i>N. Rizzolo, D. Roth</i> | |
| Toward Video Bookmarking Search: Search-Target Inference from TV Watching | 568 |
| <i>M. Okamoto, H. Umeki</i> | |

MAVS (3): MULTIMEDIA, AUDIO, AND VISUAL SEMANTICS
DR. DONG YU, MICROSOFT RESEARCH, USA

| | |
|--|-----|
| Automatic ROI Selection for JPEG2000 Compression of Remote Sensing Images | 576 |
| <i>F. Wang, J. Wei, N. Zheng, S. Du, B. Gao</i> | |
| Perceptual Shape-Based Natural Image Representation and Retrieval | 583 |
| <i>X. Zheng, S.A. Sherrill-Mix, Q. Gao</i> | |

| | |
|---|-----|
| Content-based Image Authentication: Current Status, Issues, and Challenges | 591 |
| <i>S. Han, C. Chu, S. Yang</i> | |
| Robust Invariant Descriptor for Symbol-Based Image Recognition and Retrieval | 598 |
| <i>A. Wong, W. Bishop</i> | |

WS (2): MULTIMEDIA KNOWLEDGE MANAGEMENT AND PROCESSING
DR. CHENGCUI ZHANG, THE UNIVERSITY OF ALABAMA AT BIRMINGHAM,
USA

| | |
|--|-----|
| Knowledge Management in Virtual Communities of Practice: Experiments with a Helper Tool | 604 |
| <i>J. Cesar de Lima, C. Luiz de Carvalho</i> | |
| Notes for a Collaboration: On the Design of a Wiki-type Educational Video Lecture Annotation System | 610 |
| <i>S. Ferretti, S. Mirri, M. Rocchetti, P. Salomoni</i> | |
| The Study of the Classification of Chinese Folk Songs by Regional Style | 616 |
| <i>Y. Liu, J. Xu, L. Wei, Y. Tian</i> | |
| Multicast based Middleware for Parallel Stream Processing | 622 |
| <i>A. Hakro, M.C. Murphy</i> | |

PARALLEL SESSION 7
TUNLP (6): TEXT UNDERSTANDING AND NATURAL LANGUAGE
PROCESSING
DR. TAO WU, NOKIA, USA

| | |
|--|-----|
| Corpus-based Pattern Induction for a Knowledge-based Question Answering Approach | 628 |
| <i>P. Cimiano, M. Erdmann, G. Ladwig</i> | |
| Comparing the Contribution of Syntactic and Semantic Features in Closed versus Open Domain Question Answering | 636 |
| <i>A.K. Lamjiri, L. Kosseim, T. Radhakrishnan</i> | |
| Semantic Processing of Natural Language Queries in the OntoNL Framework | 643 |
| <i>A. Karanastasi, S. Christodoulakis</i> | |
| Learning by Reading by Learning to Read | 651 |
| <i>S. Nirenburg, T. Oates, J. English</i> | |
| CLDA: Feature Selection for Text Categorization Based on Constrained LDA | 659 |
| <i>Z. Cui, B. Xu, W. Zhang, D. Jiang, J. Xu</i> | |

MAVS (4): MULTIMEDIA, AUDIO, AND VISUAL SEMANTICS
PROF. WILLIAM BISHOP, UNIVERSITY OF WATERLOO, CANADA

| | |
|--|-----|
| Semantic Annotation and Retrieval of Video Events using Multimedia Ontologies | 667 |
| <i>A.D. Bagdanov, M. Bertini, A. Del Bimbo, G. Serra, C. Torniai</i> | |
| Semantic Information Extraction of Video Based on Ontology and Inference | 675 |
| <i>J. Ma, J. Zhang, H. Lu, X. Xue</i> | |

| | |
|--|-----|
| Video Object Segmentation by 2-D Mesh-Based Motion Analysis According to Motion Connectivity | 681 |
| <i>Y. Wang, Z. Wu</i> | |
| Ontology-Based Trajectory Analysis for Semantic Event Detection | 689 |
| <i>A. Briassouli, S. Dasiopoulou, I. Kompatsiaris</i> | |
| Current Multimedia Data Formats and Semantic Computing: A Practical Example and the Challenges for the Future | 697 |
| <i>G. Friedland</i> | |

SW (2): SEMANTIC WEB
DR. MILIND NAPHADE, IBM WATSON USA

| | |
|---|-----|
| A Semantic Matching Algorithm for Discovery in UDDI | 703 |
| <i>U. Aguilera, J. Abaitua, J. Diaz, D. Bujan, D.L. de Ipina</i> | |
| S-MOJOHON: Towards a Semantic Architecture for Message Exchanging | 711 |
| <i>D.C. Lobato, R.F Bulcao Neto, M.Q. Barbosa, M. Graca Pimentel, C.A. Teixeira</i> | |
| Improving Legal Ontologies through Semantic Representation of Adjectives | 719 |
| <i>A. Bertoldi, R. Chishman</i> | |
| Grouping and Aggregate Queries over Semantic Web Databases | 727 |
| <i>D. Seid, S. Mehrotra</i> | |
| Evolutionary Conceptual Clustering of Semantically Annotated Resources | 735 |
| <i>N. Fanizzi, C. d'Amato, F. Esposito</i> | |
| Remote Experiments as SemanticWeb Services | 743 |
| <i>C.V.S. Prazeres, M.G.C. Pimentel, C.A.C. Teixeira</i> | |

Author Index