Contents

Talks

Information Extraction by Convergent Boundary Classification / 1
  Aidan Finn and Nicholas Kushmerick

IE Evaluation: Criticisms and Recommendations / 7
  A. Lavelli, M. E. Califf, F. Ciravegna, D. Freitag, C. Giuliano, N. Kushmerick, and L. Romano

A Comparison of Keyword- and Keyterm-based Methods for Automatic Web Site Summarization / 15
  Yongzheng Zhang, Evangelos Milios, and Nur Zincir-Heywood

The Use of Web-based Statistics to Validate Information Extraction / 21
  Stephen Soderland, Oren Etzioni, Tal Shaked, and Daniel S. Weld

Using Soft-Matching Mined Rules to Improve Information Extraction / 27
  Un Yong Nahm and Raymond J. Mooney

Populating the Semantic Web / 33
  Kristina Lerman, Cenk Gazen, Steven Minton, and Craig Knoblock

Handling Irregularities in ROADRUNNER / 39
  Valter Crescenzi, Giansalvatore Mecca, and Paolo Merialdo

Posters

A Model for Graded Levels of Generalizations in Intensional Query Answering / 45
  Farah Benamara

Learning Text Patterns for Web Information Extraction and Assessment / 50
  Doug Downey, Oren Etzioni, Stephen Soderland, and Daniel S. Weld

Using Statistical Techniques and WordNet to Reason with Noisy Data / 56
  Rakesh Gupta and Mykel J. Kochenderfer

A Bootstrapping Approach to Information Extraction Domain Porting / 62
  Cheng Niu, Wei Li, and Rohini K. Srinivasa

Class Extraction from the World Wide Web / 68
  Ana-Maria Popescu, Alexander Yates, and Oren Etzioni

Automatic Model Structuring from Text using BioMedical Ontology / 74
  Rohit Joshi, Xiaoli Li, Sreeram Ramachandran, and Tze Yun Leong
Unsupervised Induction of IE Domain Knowledge using an Ontology / 80
Mark Stevenson

Lexical Semantics Domain Model for Information Extraction / 86
Patricia Lutsky

A Graphical Model for Shallow Parsing Sequences / 89
Adrian Silvescu and Vasant Honavar