Mona T. Diab, PhD
Associate Professor
Department of Computer Science
School of Engineering and Applied Science
George Washington University
mtdiab@gwu.edu
http://www.seas.gwu.edu/~mtdiab
Office: +1(202) 994.8109

RESEARCH FOCUS & INTERESTS

Computational lexical semantics, multilingual processing, computational sociolinguistics, computational pragmatics, social media analytics, health analytics, low resource language processing, resource building, applied machine learning techniques, text analytics, information extraction, sentiment and emotion analysis, Arabic computational linguistics.

PROFESSIONAL EXPERIENCE

01.2013-*present*	**Associate Professor**, Department of Computer Science, The George Washington University, Washington DC, USA
01.2013-*present*	**Director**, GW NLP Lab (CARE4Lang), The George Washington University, Washington DC, USA *(~20 active members)*
06.2005-*present*	**Co-Director**, Computational Approaches for Arabic Dialect Modeling (CADIM) Group, Columbia University, The George Washington University, NYU-Abu Dhabi *(~10 active members)*
09.2009-12.2012	**Research Scientist (Principal Investigator)**, Center for Computational Learning Systems (CCLS), Columbia University, New York NY, USA
09.2009-12.2012	**Adjunct Associate Professor**, Department of Computer Science, Columbia University, New York NY, USA
09.2007-08.2009	**Adjunct Assistant Professor**, Department of Computer Science, Columbia University, New York NY, USA
02.2005-08.2009	**Associate Research Scientist (Principal Investigator)**, Center for Computational Learning Systems (CCLS), Columbia University, New York NY, USA

EDUCATION & ACADEMIC TRAINING

01.2004-01.2005	**Postdoctoral Research Scientist**, Linguistics Department and Natural Language Processing Group, Stanford University, Stanford CA, USA. (Mentor: Daniel Jurafsky, MacArthur Grant Recipient 2002-2003)
07.2003-12.2003	**Research Associate**, Center for Spoken Language Understanding, University of Colorado at Boulder, Boulder CO, USA. (Mentor: Daniel Jurafsky, MacArthur Grant Recipient 2002-2003)
06.1997-08.2003	**Ph.D.**, *Computational Linguistics*, Linguistics Department, University of Maryland Institute for Advanced Computer Studies (UMIACS), University of Maryland at College Park, College Park MD, USA. (Advisor: Philip Resnik)
08.1995-05.1997	**M.Sc.**, *Computer Science (Major: Machine Learning and Artificial Intelligence)*, Department of Computer Science, The George Washington University, Washington DC, USA. (Advisor: Peter Bock).

| 01.1988-05.1992 | **B.Sc.**, *Computer Science*, Department of Computer Science, The American University in Cairo, Cairo, Egypt. |
| 09.1986-06.1990 | **B.Sc.**, *Tourist Guidance (Major: Egyptology & Archaeology)*, Faculty of Tourism and Hotel Management, Helwan University, Cairo, Egypt. |

EDITORIAL BOARD MEMBERSHIP

04.2016-*present*	Board Member, <u>Journal of Computer Speech and Language</u>
07.2013-*present*	Board Member, <u>Journal of Computational Linguistics</u>
10.2011-*present*	Board Member, <u>Journal of Language Resources and Evaluation</u>
07.2011-*present*	Co-Chair, <u>NLP Briefs</u>, Springer Publishers

ELECTED POSITIONS

11.2011-*present*	**President**, Association for Computational Linguistics Special Interest Group for Semitic Language Processing (SIG-Semitic)
10.2013-*present*	**President**, Association for Computational Linguistics Special Interest Group on the Lexicon (SIGLEX)
01.2012-06.2014	**Executive Board Member**, North American Association for Computational Linguistics (NAACL)
06.2010-09.2013	**Secretary**, Association for Computational Linguistics Special Interest Group on the Lexicon (SIGLEX)
08.2009-12.2012	**Senator**, University Senate, one of six senators representing >3000 Research Officers, Columbia University
06.2009-10.2010	**Secretary**, Association for Computational Linguistics Special Interest Group for Semitic Language Processing (SIG-Semitic)
10.2007-07.2010	**Executive Board Member**, Association for Computational Linguistics Special Interest Group on the Lexicon (SIGLEX)

ADVISORY BOARD MEMBERSHIP

09.2011-09.2013	Qordoba Inc., Dubai, UAE
09.2011-09.2012	(ISO) Semantic Annotation Framework –Named Entities
05.2010-09.2011	Cairo Microsoft Research Innovation Center (CMIC), Microsoft Research, Cairo Egypt
06.2009-08.2013	Human Language Technology Center of Excellence, Johns Hopkins University, Baltimore MD, USA

PROFESSIONAL CONSULTING

11.2015-*present*	Providing Expert Advice on Computational approaches to Readability Complexity Measures in Modern Standard Arabic for Patients' Pharmaceutical Leaflets. *Contact: Hend Alkhalifa*, **King Saud University**, Riyadh, Kingdom of Saudi Arabia.
04.2010-12.2012	Arabic Propbank, Expert advice predicate argument structures for Arabic. *Contact: Martha Palmer*, **University of Colorado at Boulder**, Boulder CO, USA
10.2007-06.2008	Providing Expert Advice on Arabic Statistical Natural Language Processing. Contact: *Tarek Abbady*. **Cairo Microsoft Research Innovation Center**, Cairo, Egypt.
04.2006-12.2007	Providing Expert Advice on Statistical Natural Language Processing.

Contact: Aly Fahmy and Nayer Wanas. **Egyptian Center of Excellence for Data Mining, Ministry of Technology and Information**, Cairo Egypt

09.2005-12.2007   Providing Expert Advice on Arabic Lexical Semantics for building an Arabic WordNet Repository. *Contact: Christiane Fellbaum*, **Princeton University,** New Jersey, USA

04.2006-12.2007   Providing Expert advice on Arabic Lexical Semantics for Machine Translation, *Contact: Jaime Carbonnell*. **Meaningful Machines Inc.**, New York, USA.

ACADEMIC INITIATIVES

[Total: 5]

1. [Spr.2016] **Co-creator and co-designer of MSc in Data Analytics** *(with Bhagi Narahari, Claire Monteleoni, Tim Wood from CS and Faculty from EMSE and Business School)*. SEAS & Business School, The George Washington University, Washington DC, USA.
2. [05.2014] **Co-initiator Data Science Track** *(with Evan Drumwright, Claire Monteleoni, Nan Zhang, Abdou Youssef)*. Department of Computer Science, The George Washington University, Washington DC, USA.
3. [07.2011] **Founding member, *SEM Conference** *(with Eneko Agirre, Johan Bos)*. A unified consolidated top tier computational semantics scientific venue.
4. [07.2007] **Founding member, Columbia Arabic Language Institute for Modeling and Technology (CALIMAT)** *(with Hatim Diab, Nizar Habash, Owen Rambow, David Waltz)*. Seeking joint institutional collaborations with the Middle East and North African countries. Center for Computational Learning Systems (CCLS), Columbia University, New York, NY, USA.
5. [06.2005] **Founding member, Columbia Arabic Dialect Modeling Group, currently Compuational Approaches for Arabic Dialect Modeling (CADIM)** *(with Nizar Habash and Owen Rambow)*, Center for Computational Learning Systems (CCLS), Columbia University, New York, NY, USA.

KEYNOTE ADDRESSES

[Total: 12]

1. [04.2015] **16th International Conference on Intelligent Text Processing and Computational Linguistics (CICLING 15),** Cairo, Egypt.
2. [04.2015] **1st Arabic Computational Linguistics Conference (ARCLing15),** Cairo, Egypt.
3. [12.2014] **9th IEEE International Conference on Computer Engineering and Systems (ICCES 14),** Rabat, Morocco.
4. [03.2014] **3rd Palestinian Symposium on Computational Linguistics and Arabic Content,** Birzayt University, West Bank, Palestine.
5. [11.2013] **Joint Symposium on Semantic Processing, Organized by the European Union funded projects EXCITEMENT and B-CROCE hosted by FBKI**, Trento, Italy.
6. [03.2013] **27th Arabic Linguistic Symposium (ALS 13)**, Indiana IN, USA.
7. [12.2011] **11th International Language Engineering Conference**, Cairo, Egypt.
8. [09.2011] **IEEE 5th International Conference on Semantic Computing (ICSC),** Palo Alto CA, USA.
9. [04.2007] **IEEE Information and Communication Technologies International Symposium (ICTIS07)**, Fez, Morocco.
10. [11.2006] **6th Conference on Language Engineering**, Cairo, Egypt

11. [04.2006] **EACL Workshop on Cross Language Knowledge Induction,** Trento, Italy
12. [03.2006] **Arabic Linguistics Society (ALS) Conference,** Kalamazoo MI, USA

INVITED TALKS

[Total: 45]

1. [11.2016] Department of Computer Science, NYU-AD, Abu Dhabi, UAE
2. [10.2016] Department of Computer Science, Georgetown University, Washington DC, USA.
3. [09.2016] Department of Computer Science, CLIP Lab, UMIACS, University of Maryland College Park, Maryland MD, USA.
4. [09.2016] Department of Computer Science, **Johns Hopkins University**, Baltimore MD, USA.
5. [04.2016] Department of Computer Science, **King Saud University**, Riyadh, Kingdom of Saudi Arabia.
6. [04.2016] Department of Computer Science, **Al Imam University**, Riyadh, Kingdom of Saudi Arabia.
7. [04.2016] Languages & Linguistics Department, **Princess Noura University**, Riyadh, Kingdom of Saudi Arabia.
8. [03.2016] School of Engineering and Applied Science, **King Abdullah University for Science & Technology**, Thuwal, Kingdom of Saudi Arabia.
9. [05.2015] **IBM Watson and Cognitive Computing Institute**, New York NY.
10. [03.2013] **MITRE Corporation**, Tyson's Corner VA, USA.
11. [03.2013] Computer Science Symposium, **Simon Fraser University,** British Columbia, Vancouver BC, Canada.
12. [02.2013] Computer Science Symposium, **CLIP Lab, UMIACS, University of Maryland College Park**, College Park MD, USA.
13. [09.2011] Upper Administration, **King Abdullah University for Science and Technology (KAUST)**, Thuwal, Saudi Arabia.
14. [09.2011] Mathematical & Computer Sciences School, **King Abdullah University for Science and Technology (KAUST),** Thuwal, Saudi Arabia.
15. [10.2011] **NSF-sponsored Workshop on Collaborative Resource Development and Delivery.** Organized by Nancy Ide, Vassar College, New York NY, USA.
16. [04.2011] **Distinguished Lecture Symposium Series, University of Washington,** Seattle WA, USA.
17. [04.2011] **Microsoft Research,** Redmond WA, USA.
18. [12.2010] Computer Science Department**, University of Alabama**, Birmingham AL, USA.
19. [07.2010] **Workshop on the Future of Arabic Language Development,** organized by the **UAE Crown Prince Court for Strategic Planning**, Abu Dhabi, UAE.
20. [06.2010] **IBM TJ Watson**, Hawthorne NY, USA.
21. [05.2010] School of Computing and Mathematical Sciences**, King Abdullah University for Science and Technology, KAUST**, Thuwal, Saudi Arabia.
22. [04.2010] Computer Science Department**, University of Trento**, Trento, Italy.
23. [06.2009] **NAACL Workshop on Semantic Evaluations (SEMEVAL 09),** Boulder CO, USA.
24. [05.2009] Departmento de Sistemas Informaticos y Computacion**, Universidad Politechnica de Valencia**, Valencia, Spain.
25. [04.2009] **IRST & Computer Science Department, University of Trento**, Trento, Italy.
26. [08.2008] Department of Informatics**, University of Sussex**, Brighton, England.
27. [05.2008] Computer Science Department**, Faculty of Science, University of Mohamed V**, Rabat, Morocco.

28. [05.2008] Departmento de Sistemas Informaticos y Computacion, **Universidad Politecnica de Valencia**, Valencia, Spain.
29. [04.2008] Linguistics Department, **Georgetown University**, Washington DC, USA.
30. [12.2006] **Cairo Microsoft Innovation Center (CMIC), MSR**, Cairo, Egypt.
31. [09.2006] Computer Science Department, **University of North Texas**, Dallas TX, USA.
32. [12.2005] Linguistic Data Consortium, **University of Pennsylvania,** Philadelphia PA, USA.
33. [10.2005] Computer Science Department, **New York University**, New York NY, USA.
34. [03.2005] Computer Science Department, **Columbia University**, New York NY, USA.
35. [02.2005] Linguistics Department, **Ohio State University**, Columbus OH, USA.
36. [02.2005] Linguistics Department, **University of Texas Austin**, Austin TX, USA.
37. [01.2005] **Yahoo Research Labs, Overture**, Pasadena CA, USA.
38. [12.2004] **Language Computer Corporation** (**LCC),** Dallas TX, USA.
39. [12.2004] Natural Language Group, **Microsoft Research (MSR),** Redmond WA, USA.
40. [08.2004] Department of Informatics, **University of Edinburgh**, Edinburgh, Scotland.
41. [05.2004] Linguistic Data Consortium, **University of Pennsylvania**, Philadelphia PA, USA.
42. [05.2004] **IBM T. J. Watson Research Center**, Hawthorne NY, USA.
43. [11.2003] Linguistics Dept. & Institute for Cognitive Science, **University of Colorado, Boulder**, Boulder CO, USA.
44. [11.2002] Laboratoire Perception, Systèmes, Information (PSI), **Université de Rouen**, Rouen, France.
45. [11.2002] NLP Lab, Information Sciences Institute (ISI), **University of Southern California**, Marina Del Rey CA, USA.

INVITED PANELIST/PARTICIPANT

[Total: 13]

1. [11.2016] Wardat Workshop, by invitation only, NYU-AD, Abu Dhabi, UAE
2. [05.2012] **Computational lexical resources and interoperability,** *Language Resources and Evaluation Conference (LREC),* Istanbul, Turkey.
3. [10.2012] **Arbiter/Referee, One of 6 Arab Expat Scientists to decide on Best Scientific Research in Qatar,** *Qatar Foundation, Arab Expat Society & Scientific Research Forum,* Doha, Qatar**.**
4. [11.2011] **One of 50 Arab Expat Scientists to discuss and advise on Future of Scientific Research Strategy in Qatar***, Qatar Foundation, Arab Expat Society & Scientific Research Forum,* Doha, Qatar.
5. [10.2011] **Future of Semantics in Machine Translation**, *DARPA BOLT PI Kick-off Meeting,* Tampa FL, USA.
6. [09.2011] **Adjectives Workshop**, NSF, Organized by Christiane Fellbaum. (Princeton University), Arlington VA, USA.
7. [07.2011] **EMNLP Workshop on Geometrical Methods in Semantics (GEMS).** *Empirical Methods For Natural Language Processing (EMNLP 11),* Edinburgh, UK.
8. [04.2011] **NSF-funded SILT Workshop Interoperable Semantic Annotation.** Organized by Nancy Ide. (Vassar College) and James Pustejovsky. (Brandeis University), Waltham MA, USA.
9. [04.2011] **QCRI Social Networks Roundtable**, Qatar Computing Research Institute, Qatar Foundation, Doha, Qatar.
10. [01.2011] **QCRI Arabic Language Technologies Roundtable,** *Qatar Computing Research Institute, Qatar Foundation,* Doha, Qatar**.**
11. [08.2010] **COLING Workshop on Multiword Expressions (MWE 10),** Beijing, China.
12. [07.2010] **ACL Workshop on Semantic Evaluations (SEMEVAL),** Upsala, Sweden.

13. [07.2010] **Workshop on the Future of Arabic Language Development,** *organized by the UAE Crown Prince Court for Strategic Planning*, Abu Dhabi, UAE.

RESEARCH EXPERIENCE

CURRENT RESEARCH GRANTS

[Total: 6 grants amounting to $8.785M, Diab's portion $2.849M]

1. [**Google Faculty Research Award**] *Tharwa++: Building a multidialectal Arabic Lexical Repository*, **(PI),** GW portion $75K, Sep 2015 –Dec 2016.
2. [**NSF CRI-P**] *A Unified Universal Multiword Expression Repository*, **(PI),** GW portion $100K, Jul 2015 –Jun 2017.
3. [**DARPA-LORELEI**] *Sentiment and Emotion Detection in Low Resource Languages (SEEM)* **(GW-PI;** collaborative with Columbia University Julia Hirschberg & Kathy Mckeown), GW portion $240K, Jul 2015 – Aug 2019.
4. [**Qatar National Research Fund (QNRF)**] *Optdiac: An optimal Diacritization Scheme for Arabic*, **(Lead PI;** collaborative research with Carnegie Mellon University Qatar Co-PI Kemal Oflazer) Total $1.1M. (GW portion $367K), Sep 2013 – Apr 2017.
5. [**DARPA-DEFT**] *DRATS: Detecting Relations and Anomalies in Text and Speech*, **(GW PI;** subaward from Columbia University as Prime institution) Total $5.5M, Total for PI $934,107. (GW portion $795K), May 2013 – Nov 2017.
6. [**NSF-Large CRI**] *Creating Resources for Multilingual Linguistic Code Switching*, **(PI;** with Julia Hirschberg from Columbia University, collaborative with Thamar Solorio, UAB/UH, GW Prime institution), GW portion $295K, Apr 2013 –Dec 2016.

PAST RESEARCH GRANTS

[Total: 18 grants amounting to $16.432M, Diab's portion $5.790M]

1. [**Google Faculty Research Award**] *Nuanced Sentiment and Perspective Analysis for Arabic Social Media Text*, (**PI**), GW portion $72K, Dec 2014 –Dec 2015
2. [**DARPA–BOLT**] Subcontract to GW from BBN – *Dialectal Tool and Resource Creation,* (**PI**; with Columbia University Nizar Habash and Owen Rambow), Total $5.5M, GW portion $241,570, Oct 2011 –May 2015
3. [**TSWG-DOD/ODNI–NADIA**] *Natural Language Processing for Dialectal Arabic– Industrialization and Insertion*, (**PI**; with Owen Rambow and Nizar Habash), Total $898K, Apr 2012 –Jun 2013
4. [**DARPA–MADCAT**] Subcontract to BBN–*Multilingual Automatic Document Classification Analysis and Translation for Arabic OCR*, (**Co-PI**; with Nizar Habash) Total $750K, Apr 2008 –Aug 2013
5. [**DARPA–BOLT**] Subcontract to SRI–*Semantic Machine Translation for Arabic Dialects,* (**CCLS**-**PI**; with Kathy Mckeown, Michael Collins, Julia Hirschberg), Total $1.7M, Nov 2011 – Aug 2013
6. [**DARPA–BOLT**] Subcontract to LDC–*Resource Guideline Creation for Arabic Dialects*, (**Co-PI**; with Nizar Habash and Owen Rambow), Total $500K, Nov 2011 –Dec 2012
7. [**DARPA–GALE**] Subcontract to IBM–*Pilot Unified Framework for Semantic Textual Similarity*, (**PI**), Total $25K, Sep 2011 –Apr 2012
8. [**NSA–TTO8**] Subcontract to JHU-COE–*Power and Modality Analysis and Prediction in Social Communication, (***Co-PI**; with Owen Rambow) Total $345K, Sep 2009 –Dec 2012
9. [**IARPA–SCIL**] *Power Rifts and Shifts in Social Communication*. (**Co-PI**; with Kathy

McKeown and Owen Rambow) Total $3.210M, Aug 2009 –Oct 2012

10. [**NSF–CRI Planning**] *Linguistic Code Switching Pilot Annotation*. (**PI**; with Julia Hirschberg) Total $78K, Mar 2010–Mar 2011
11. [**TSWG-DOD/ODNI–COLABA**] Subcontract to ACXIOM–*Cross language blog alerts–Arabic Blog Processing*. (**PI**; with Nizar Habash and Owen Rambow) Total $820K, (CAT/CAIM support), Aug 2009–Apr 2011
12. [**DARPA–GALE**] Subcontract to IBM-Rosetta–*Arabic to English Machine Translation*. (**Co-PI**; with Owen Rambow and Nizar Habash) Total $1.170M, May 2008–Apr 2011
13. [**NSA–TTO8**] Subcontract to JHU-COE–*Committed Belief and Modality in Multigenres Text*. (**Co-PI**; with Owen Rambow) Total $36K, Aug 2007 –Aug 2010
14. [**DARPA–GALE**] Subcontract to SRI-Nightingale–*Arabic to English Machine Translation*. (**Co-PI**; with Owen Rambow and Nizar Habash) Total $1.5M, Sep 2005–Apr 2008
15. [**NSF–SGER**] *Automatic Processing of Natural Language Code-Switching*. (**PI**; with Owen Rambow) Total $41K, Sep 2007–Aug 2008
16. [**NSF–SGER**] *Inducing Multiword Expressions from Comparable Corpora*. (**PI**) Total $19K, Sep 2006–Feb 2007
17. [**NSF–JHU Summer Workshop**] *Arabic Dialect Parsing*. (**Co-PI**) Total $15K, Jun 2005–Jul 2005
18. [**DARPA–GALE**] Subcontract to University of Colorado–OntoNotes. *Building an Arabic PropBank.* (**PI**) Total $15K, Sep 2005–Feb 2006

TEACHING EXPERIENCE

FORMAL SEMESTER LONG CLASSES

[Total: 9 courses]

1. [Fall.2016] **CSCI6709/3709: Introduction to Statistical Natural Language Processing,** Department of Computer Science, GW
2. [Spring.2014] **CSCI6709/3709: Introduction to Statistical Natural Language Processing,** Department of Computer Science, GW
3. [Fall.2013] **CSCI1112: Data Structures and Algorithms I in Java,** Department of Computer Science, GW
4. [Spring.2013] **CSCI 6709/3709: Introduction to Statistical Natural Language Processing,** Department of Computer Science, GW
5. [Fall.2009] **Arabic Computational Linguistics,** Linguistics and Computer Science Departments, Georgetown University.
6. [Spring.1997] **C as a 2nd Language,** Department of Computer Science, GW
7. [Fall.1996] **C as a 2nd Language,** Department of Computer Science, GW
8. [Spring.1996] **C as a 2nd Language,** Department of Computer Science, GW
9. [Fall.1995] **C as a 2nd Language,** Department of Computer Science, GW

TUTORIALS

[Total: 10 Tutorials]

1. [10.2014] **Arabic Dialect Processing (5 hrs),** *co-designed and co-taught with Nizar Habash,* Empirical Methods for Natural Language Processing (EMNLP), Doha, Qatar. Attended by 35 people.
2. [06.2012] **Arabic Dialect Processing (5 hrs),** *co-designed and co-taught with Nizar*

Habash, North American Association for Computational Linguistics (NAACL). Montreal, Canada. Attended by 40 people.

3. [04.2010] **Arabic Dialect Processing (5 hrs),** *co-designed with Nizar Habash,* Georgetown University Round Table (GURT), Washington DC. Attended by 60 people.
4. [04.2009] **Arabic Dialect Processing (5 hrs),** *co-designed and co-taught with Nizar Habash,* 2nd International *Conference on Arabic language Resources and Tools. (MEDAR*), Cairo, Egypt. Attended by over 100 people.
5. [05.2008] **Arabic Dialect Processing (5 hrs),** *co-designed and co-taught with Nizar Habash, Language Resources & Evaluation Conference (LREC),* Marrakech, Morocco. Attended by 45 people.
6. [06.2007] **Arabic Dialect Processing (5 hrs),** *co-designed and co-taught with Nizar Habash,* Human Language Technologies-North American Association for Computational Linguistics (HLT-NAACL). Rochester NY, USA. Attended by 45 people.
7. [03.2007] **An A-Z Statistical NLP Crash Course (6 hrs),** IEEE Information and Communication Technologies International Symposium (ICTIS07). Fez, Morocco. Attended by 40 people.
8. [09.2006] **Statistical Natural Language Processing: An A-Z Crash Course.** Faculty of Engineering and Computer Studies, Cairo University, Cairo, Egypt. Attended by over 45 people.
9. [08.2008] **Arabic Dialect Processing (5 hrs),** *co-designed and co-taught with Nizar Habash, American Machine Translation Association (AMTA), Boston MA, USA. Attended by over 30 people*
10. [02.2005] **Introduction to Computational Linguistics (2 hrs),** *Linguistics Department, University of Texas at Austin, Austin TX, USA. Attended by 100 people.*

RESEARCH SUPERVISION

RESEARCH SCIENTIST MENTORSHIP

[Total: 3 RS]

1. [11.2015-*present*] **Abdelati Hawwari, PhD**
   Received PhD Jul 2012: Faculty of Literature & Linguistics, Cairo University, Egypt
   Research Focus: Building Lexical Resources and Computational Lexicography.
2. [04.2015-*present*] **Mahmoud Ghoneim, PhD**
   Received PhD Jul 2012: Faculty of Engineering, Cairo University, Egypt
   Research Focus: Building shallow parsing technology
3. [02.2014-11-2014] **Mohammed Attia, PhD**
   Received PhD Jul 2008: University of Manchester, UK.
   Research Focus: Building Lexical Resources and Spelling error correction
   Current Position: Research Scientist, Google Inc., New York NY, USA

POSTDOCTORAL MENTORSHIP

[Total: 6 Postdoctoral Research Scientists]

1. [09.2015-*present*] **Efsun Kayi, PhD**
   Received PhD Jul 2014: Department of Computer Science, George Washington University, Washington DC, USA

Research Focus: <u>Automatic Detection of Psychological Disorders (Schizophrenia and PTSD) in Social Media</u>.

2. [07.2012-03.2015] **Mahmoud Ghoneim, PhD**
    Received PhD Jul 2012: Faculty of Engineering, Cairo University, Egypt
    Research Focus: <u>Integrating Semantic Knowledge in Statistical Machine Translation</u>
    Current Position: Research Scientist, Department of Computer Science, George Washington University, Washington DC, USA.

3. [03.2012-10.2015] **Abdelati Hawwari, PhD**
    Received PhD Jul 2012: Faculty of Literature & Linguistics, Cairo University, Egypt
    Research Focus: <u>Building Lexical Resources and Computational Lexicography.</u>
    Current Position: Research Scientist, Department of Computer Science, George Washington University, Washington DC, USA.

4. [05.2010-04.2011] **Davide Picca, PhD**
    Received PhD Apr 2010: University of Lausanne, Switzerland
    Research Focus: <u>Unsupervised semantic modeling for comparable corpora</u>
    Current Position: Research Technical Laison, University of Lausanne, Switzerland

5. [05.2009-04.2011] **Marine Carpuat, PhD**
    Received PhD Apr 2008: HKUST, Hong Kong
    Research Focus: <u>Incorporating Lexical Semantics in Statistical Machine Translation</u>
    Current Position: Director of Research Symanto Group, New York NY, USA

6. [09.2009-09.2010] **Yassine Benajiba, PhD**
    Received PhD May 2009: Polytechnical University of Valencia, Spain
    Research Focus: <u>Shallow syntactic parsing in Arabic and integration in SMT and IE</u>
    Current Position: Director of Research Symanto Group, New York NY, USA

THESIS SUPERVISION

Graduated

[Total: 6]

1. [07.2015] **Vinodkumar Prabkaharan, PhD**
    Joint Supervision with Owen Rambow
    Thesis: Social Power in Interactions: Computational Analysis and Detection of Power Relations
    Department of Computer Science, Columbia University, New York NY, USA
    Current Position: Postdoctoral Scientist, Stanford University. (Mentor Daniel Jurafsky)

2. [04.2015] **Weiwei Guo, PhD**
    Thesis: Dimension Reduction for Short Text Similarity and its Applications
    Department of Computer Science, Columbia University, New York NY, USA
    Current Position: Research Scientist, Linked In, CA USA

3. [12.2014] **Muhammad Abdul-Mageed, PhD**
    Joint Supervision with Sandra Kuebler
    Thesis: Sentiment Analysis in Arabic Social Media Text
    Linguistics Department, Information Sciences Department
    Indiana University, Indiana IN, USA
    Current Position: Assistant Professor, University of British Columbia, Canada

4. [12.2012] **Pradeep Dasigi, MSc**
    Thesis: Dialectal Arabic Orthography Normalization
    Department of Computer Science, Columbia University, New York NY, USA
    Current Position: PhD Candidate at Carnegie Mellon University working with Ed Hovy

5. [03.2012] **Mahmoud Ghoneim, PhD**
   Thesis: Explicit Modeling of Semantics in Statistical Machine Translation
   Joint Supervision with Mohsen Rashwan
   Computer Engineering Department, Faculty of Engineering, Cairo University, Egypt
   Current Position: Research Scientist, Department of Computer Science, George Washington University
6. [07.2009] **Yassine Benajiba, PhD**
   Thesis: Arabic Named Entity Recognition
   Joint Supervision with Paolo Rosso
   Polytechnical University of Valencia, Spain
   Current Position: Director of Research Symanto Group, New York NY, USA

Current PhD Students (Date Started)

[Total: 15 Students]

1. [Fall.2010] **Hebatullah ElFardy**
   Thesis Topic: Perspective Identification in multilingual and multigenre contexts
   Department of Computer Science, Columbia University, New York NY, USA
2. [Spring.2012] **Ali Elkahky**
   Thesis Topic: Unified models for dialectal processing
   Department of Computer Science, Columbia University, New York NY, USA
3. [Spring.2013] **Aya Zirikly**
   Thesis Topic: Tracking Named Entities in Multilingual and Multigenre Settings
   Department of Computer Science, George Washington University, DC, USA
4. [Fall.2013] **Hanan AlDarmaki**
   Thesis Topic: Cross Language Semantic Textual Similarity
   Department of Computer Science, George Washington University, DC, USA
5. [Fall.2013] **Maryam Aminian**
   Thesis Topic: Domain Independent Semantic Role Labeling
   Department of Computer Science, George Washington University, DC, USA
6. [Spring.2014] **Shabnam Tafreshi**
   Thesis Topic: Affect and Emotion Detection
   Department of Computer Science, George Washington University, DC, USA
7. [Fall.2014] **Fahad Alghamdi**
   Thesis Topic: Computational Approaches to Linguistic Code Switching
   Department of Computer Science, George Washington University, DC, USA
8. [Fall.2014] **Nada Almarwani**
   Thesis Topic: Cross lingual Textual Entailment
   Department of Computer Science, George Washington University, DC, USA
9. [Fall.2014] **Sawsan Alqahtani**
   Thesis Topic: Optimal Diacritization for Morphologically Rich Languages
   Department of Computer Science, George Washington University, DC, USA
10. [Fall.2014] **Sardar Hamidian**
    Thesis Topic: Modeling of trustworthiness: automatic identification of misinformation
    Department of Computer Science, George Washington University, DC, USA
11. [Fall.2015] **Ali Seyfi**
    Thesis Topic: Multilingual Event Detection and Tracking
    Department of Computer Science, George Washington University, DC, USA
12. [Spring.2015] **Mohammed AlBadrashiny**
    Thesis Topic: Language Identification and Spelling error correction

Department of Computer Science, George Washington University, DC, USA
13. [Fall.2016] **Pedram Hosseini**
    Thesis Topic: Cross language named entity linking
    Department of Computer Science, George Washington University, DC, USA
14. [Fall.2016] **Jared Mowery**
    Thesis Topic: Event Detection and Tracking in social media
    Department of Computer Science, George Washington University, DC, USA
15. [Fall.2016] **Rana Ahmed Aref**
    Thesis Topic: Bias Detection in Media
    Department of Computer Science, Cairo University, Cairo, Egypt

PHD THESIS COMMITTEE MEMBERSHIP

[Total: 10]

1. [Fall.2016] **Gilbert Badaro**
   Thesis Proposal: Resources and Analytics for Opinion Mining and Recommender Systems, with Application to Arabic
   Faculty of Engineering and Architecture, American University in Beirut (AUB), Lebanon
2. [Spring.2016] **Alawya Alawami**
   Thesis Proposal: Aspect Extraction for Sentiment Analysis in Arabic Dialects
   School of Information Sciences, University of Pittsburg, PA, USA
3. [Fall.2015] **Aous Mansouri**
   Thesis: Stative and Stativizing Constructions in Arabic News Reports: A corpus based study
   Department of Linguistics, University of Colorado, Boulder CO, USA
4. [Fall.2014] **Walaa Medhat**
   Thesis: Text Mining on Social Networking using NLP Techniques
   Faculty of Engineering, Ain Shams University, Cairo, Egypt
5. [Fall.2013] **John Sipple**
   Thesis Proposal: Developing a language-learning model to incrementally infer syntactic structures from a continuous stream of English-language discourse.
   Department of Computer Science, George Washington University, Washington DC, USA
6. [Spring.2012] **Aous Mansouri**
   Thesis Proposal: Imperfective Constructions in Arabic.
   Department of Linguistics, University of Colorado, Boulder CO, USA
7. [Spring.2010] **Daniele Pighin**
   Thesis: Reverse Feature Engineering for SVMs
   Computer Science Department, University of Trento, Italy
8. [Fall.2009] **Kapil Thadani**
   Thesis Proposal: Semantic Fusion and Generation.
   Department of Computer Science, Columbia University, New York NY, USA
9. [Spring.2009] **Bonaventura Coppola**
   Thesis: Automatic Processing of Language Frame Semantics.
   ICT Division, University of Trento, Italy
10. [Spring.2008] ***Fadoua Ataa Allah***
    Thesis: Information Retrieval: English and Arabic Documents.
    Faculty of Science, University of Mohamed V, Rabat, Morocco

PUBLICATIONS

BOOKS

1. **Diab, Mona** and Christiane Fellbaum. (*forthcoming*) Computational Lexical Semantics: A Multilingual Perspective. *Springer Publishers*, under contract, expected Spring 17.
2. **Diab, Mona,** Nizar Habash, Owen Rambow. (*forthcoming*) Processing Arabic Dialects. *Cambridge Press*, under contract, expected Fall 17.
3. **Diab, Mona,** Aline Villavicencio, Anna Korhonen, Marianna Apidianaki, Valia Kordoni, Mark Stevenson, Preslav Nakov, *Editors*. (*forthcoming*) Essays Computational Lexical Semantics in Honor of Adam Kilgarriff. *Springer Publishers,* under contract, expected Fall 16**.**

THESIS

1. **Diab, Mona**. (2003) Word Sense Disambiguation within a multilingual framework. **PhD Thesis**, *University of Maryland,* College Park MD, USA.

EDITED CONFERENCE/WORKSHOP PROCEEDINGS

[Total: 8]

1. **Diab, Mona,** Pascale Fung, Julia Hirschberg, Thamar Solorio, *Editors.* (2016) 2nd Workshop on Computational Approaches to Linguistic Code Switching. *In Proceedings of Empirical Methods of Natural Language Processing (EMNLP).*
2. Ghoneim, Mahmoud, **Mona Diab,** Houda Bouamor, Ahmed ElKholy, Yuval Marton, *Editors*. (2016) Workshop on Machine Translation for Semitic Languages (SEMAT). *In Proceedings of Automatic Machine Translation in the Americas (AMTA).*
3. **Diab, Mona,** Pascale Fung, Julia Hirschberg, Thamar Solorio, *Editors*. (2014) 1st Workshop on Computational Approaches to Linguistic Code Switching. *In Proceedings of Empirical Methods of Natural Language Processing (EMNLP).*
4. **Diab, Mona,** Timothy Baldwin, Marco Baroni. *Editors. (2013)* 2nd International Joint Conference on Semantics (*SEM). *Proceedings of *SEM 2013.*
5. Agirre, Eneko, Johan Bos, **Mona Diab.** *Editors.* (2012) 1st International Joint Conference on Semantics (*SEM). *Proceedings of *SEM 2012.*
6. Matveeva, Irina, Chris Biemann, Monojit Choudhury, **Mona Diab.** *Editors*. (2008) Proceedings of the 3rd Textgraphs workshop on Graph-based Algorithms for Natural Language Processing. *COLING 2008.*
7. Choukri, Khalid, *Mona Diab*, Bente Maegaard, Paolo Rosso, Abdelhadi Soudi, Ali Farghaly. *Editors*. (2008) Proceedings of LREC Workshop on HLT & NLP within the Arabic world: Arabic Language and local languages processing: Status Updates and Prospects. *LREC 2008.*
8. Darwish, Kareem, **Mona Diab**, Nizar Habash, *Editors*. (2005) Proceedings of Computational Approaches to Semitic Languages. *Association for Computational Linguistics.*

JOURNAL ARTICLES

[Total: 8]
1. Al Aqeel S, Abanmy N, Abeer Aldayel, Hend S. Al-Khalifa, Maha Al-Yahya, **Mona Diab**. (Submitted). Readability of medication information materials in Saudi Arabia: expert and non-expert evaluation. *Journal of Methods of Medical Information.*

2. Zaghouani, Wajdi, Abdelati Hawwari, **Mona Diab.** (2016) AMPN: A Lexical Semantic Resource for Arabic Morphological Patterns. *International Journal of Speech Technologies*, Springer Publishers.
3. Abdul-Mageed, Mohammad, **Mona Diab** and Sandra Kuebler. (2014) SAMAR: A System for Subjectivity and Sentiment Analysis for Arabic Social Media. *Computer Speech and Language, 28. (201)*
4. Abdul-Mageed, Muhammad, and **Diab, Mona.** (2012) Linguistically-motivated subjectivity and sentiment annotation and tagging of Modern Standard Arabic. *International Journal on Social Media MMM: Monitoring, Measurement, and Mining.*
5. Katz, Graham, and **Mona Diab.** (2011) Introduction to the special issue on Arabic Computational Linguistics. *ACM Transactions on Asian Language Information Processing (TALIP). Volume 10 Issue 1, Mar. Article (1).*
6. BenAjiba, Yassine, **Mona Diab,** Paolo Rosso. (2009) Arabic Named Entity Recognition: A Feature-driven Study. *In the special issue on Processing Morphologically Rich Languages In IEEE Transactions on Audio, Speech and Language Processing. 17(5), pp.926-934.*
7. BenAjiba, Yassine, **Mona Diab,** Paolo Rosso. (2009) Using Language Independent and Language Specific Features to Enhance Arabic NER. *International Arabic Journal of Information Technology. 6(5), pp.464-473.*
8. Resnik, Philip, Mari B Olsen, **Mona Diab.** (1999) The Bible as a Parallel Corpus: Annotating the `Book of 00 Tongues' *Computers and the Humanities*, 33. (201-2).

INVITED ARTICLES

[Total: 9]

1. **Diab, Mona**. (2016) Overview of Arabic Computational Linguistics, *Routledge Handbook on Arabic Linguistics. Editors Reem Bassiouney and Abbas Benmamoun*.
2. **Diab, Mona.** (2015) Tharawat: A Vision for a Comprehensive Resource for Arabic Computational Processing. *Journal for Computational Linguistics and Intelligent Text Processing,* pp. 85-97.
3. **Diab, Mona** and Yuval Marton. (2014) Semitic Semantics. *Book Chapter in Natural Language Processing for Semitic Languages, Editor Imed Zitouni, Springer Publishers, pp. 129-159.*
4. ElFardy, Hebatullah, Mohamed AlBadrashiny, **Mona Diab**. (2014) A Hybrid System for Code Switch Point Detection in Informal Arabic Text. *XRDS: Crossroads, The ACM Magazine for Students* 21. (201), pp. 52-57.
5. Bar, Kfir, **Mona Diab**, Abdelati Hawwari**. (201**3) Arabic Multiword Expressions: Resource and Tool Creation. *Book Chapter in Natural Language Processing for Semitic Languages, Editors Naachum Derschowitz and Ephraim Nissim, Vol.3, in honor of Yaacov Choueka, Springer Publishers.*
6. **Diab, Mona,** Nizar Habash, Owen Rambow, Ryan Roth. (2011) CADIM Arabic Tools: Morphological Analysis, Disambiguation and Generation, Tokenization, Diacritization, Lemmatization, POS Tagging and Base Phrase Chunking. *Book Chapter in the Handbook of Natural Language Processing and Machine Translation, Editors Joseph Olive, Caitlin Christianson and John McCary. Springer Publishers.*
7. Katz, Graham and **Mona Diab**. (2010) Special Issue on Georgetown Round Table on Arabic Computational Linguistics Conference. (Guest Editors), *ACM Transactions on Asian Language Information Processing. (TALIP).*
8. **Diab, Mona. (200**8) It's About Time. *CMIC Newsletter, Volume 1, issue #2. Feb 15th. http://www.microsoft.com/middleeast/Egypt/CMIC/Newsletter.aspx*

9. **Diab, Mona,** Kadri Hacioglu, Daniel Jurafsky. (2007) <u>Automated Methods for Processing Arabic Text: From Tokenization to Base Phrase Chunking.</u> *In Arabic Computational Morphology: Knowledge-based and Empirical Methods. Editors Antal van den Bosch and Abdelhadi Soudi. Kluwer/Springer Publications.*

PEER REVIEWED CONFERENCE PAPERS

[Total: 75]

[2016]
1. AlBadrashiny, Mohamed and **Mona Diab.** (2016) LILI: <u>A Simple Efficient Language Independent Framework for Linguistic Code Switch Point Detection.</u> *Proceedings of COLING 2016, Japan*, Dec.
2. AlQahtani, Sawsan, Mahmoud Ghoneim, **Mona Diab**. (2016) <u>Impact of Explicit encoding of vowelization in Arabic Machine translation.</u> *Proceedings of Automatic Machine Translation Association (AMTA) 2016, Texas Austin, USA, Nov.*
3. Hawwari, Abdelati, Mohammed Attia, Mahmoud Ghoneim, **Mona Diab**. (2016) <u>Explicit Fine grained Syntactic and Semantic Annotation of the Idafa Construction in Arabic.</u> *In Proceedings of LREC 2016, Slovenia, May.*
4. **Diab, Mona**, Mahmoud Ghoneim, Abdelati Hawwari, Fahad AlGhamdi, Nada AlMarwani, Mohamed Al-Badrashiny. (2016) <u>Creating a Large Multi-Layered Representational Repository of Linguistic Code Switched Arabic Data.</u> *In Proceedings of LREC 2016, Slovenia, May.*
5. Zaghouani, Wajdi, Houda Bouamor, Abdelati Hawwari, **Mona Diab**, Ossama Obeid, Mahmoud Ghoneim, Sawsan Alqahtani, Kemal Oflazer. (2016) <u>Large Scale Arabic Diacritized Corpus: Guidelines and Framework.</u> *In Proceedings of LREC 2016, Slovenia*, May.
6. Al-Badrashiny, Mohamed, Arfath Pasha, **Mona Diab**, Nizar Habash, Owen Rambow, Wael Salloum, and Ramy Eskander, <u>SPLIT: Smart Preprocessing (Quasi) Language Independent Tool,</u> *In Proceedings of LREC 2016, Slovenia*, May.

[2015]

7. Prabhakaran, Vinodkumar, and Tomas By, Julia Hirschberg, Owen Rambow, Samira Shaikh, Tomek Strzalkowski, Jennifer Tracey, Michael Arrigo, Rupayan Basu, Micah Clark, Adam Dalton, **Mona Diab**, Louise Guthrie, Anna Prokofieva, Stephanie Strassel, Gregory Werner, Yorick Wilks and Janyce Wiebe. (2015) <u>A New Dataset and Evaluation for Belief/Factuality</u>. *In Proceedings of 4th *SEM Conference*, Denver CO, USA.
8. Elfardy, Heba, and **Mona Diab** and Chris Callison-Burch. (2015) <u>Ideological Perspective Detection Using Semantic Features.</u> *In Proceedings of 4th *SEM Conference*, Denver CO, USA.
9. Al-Badrashiny, Mohamed, Heba Elfardy, **Mona Diab**. (2015) <u>AIDA2: A Hybrid Approach for Token and Sentence Level Dialect Identification in Arabic.</u> In Proceedings *of the 9th Conference on Natural Language Learning (CoNLL),* Beijing, China.
10. Hamidian, Sardar, and **Mona Diab**. (2015) <u>Improved Automatic Rumor Detection.</u> *In Proceedings of The Fifth International Conference on Social Media Technologies, Communication, and Informatics (SOTICS),* Barcelona, Spain.
11. Zirikly, Aya, Yassine Benajiba, **Mona Diab**. *(2015*) <u>GW-EDL English Diagnostic Short System Description.</u> KPB Task Evaluation, *Proceedings of the TAC Conference*, NIST, Maryland MD, USA.

[2014]

12. Abdul-Mageed, Muhammad, **Mona Diab**. (2014) <u>SANA: A Large Scale Multi-Genre, Multi-Dialect Lexicon for Arabic Subjectivity and Sentiment Analysis.</u> *In Proceedings of Language Resources and Evaluation Conference (LREC)*, Reykjavik, Iceland.

13. **Diab, Mona**, Mohamed AlBadrashiny, Maryam Aminian, Mohamed Attia, Hebatullah ElFardy, Nizar Habash, Abdelati Hawwari, Wael Salloum, Pradeep Dasigi, Ramy Eskander. (2014) <u>Tharwa: A Large Scale Dialectal Arabic –Standard Arabic – English Lexicon.</u> *In Proceedings of Language Resources and Evaluation Conference (LREC)*, Reykjavik, Iceland.

14. Guo, Weiwei, Wei Liu, **Mona Diab**. (2014) <u>Fast Tweet Retrieval with Compact Binary Codes.</u> *In Proceedings of Conference of Computational Linguistics (COLING)*, Dublin, Ireland.

15. Pasha, Arfath, Mohamed AlBadrashiny, **Mona Diab,** Ahmed ElKholy, Ramy Eskander, Nizar Habash, Manoj Poolery, Owen Rambow, Ryan Roth. (2014) <u>MADAMIRA: A Fast, Comprehensive Tool for Morphological Analysis and Disambiguation of Arabic.</u> *In Proceedings of Language Resources and Evaluation Conference (LREC)*, Reykjavik, Iceland.

16. Salloum, Wael, Hebatullah ElFardy, Linda Alamir-Salloum, Nizar Habash, **Mona Diab**. (2014) <u>Sentence Level Dialect Identification for Machine Translation System Selection</u>. *In Proceedings of the Annual Meeting of the Association for Computational Linguistics. (ACL)*, Baltimore MD, USA.

[2013]

17. Abdul-Mageed, Muhammad, **Mona Diab**, Sandra Kubler. (2013) <u>ASMA: A system for Automatic Segmentation and Morpho-syntactic Disambiguation of Modern Standard Arabic.</u> *In Proceedings of Recent Advances in Natural Language Processing (RANLP)*, Sophia, Bulgaria.

18. Abu-Jbara, Amjad, Ben King, **Mona Diab**, Dragomir Radev. (2013) <u>Identifying Opinion Subgroups in Arabic Online Discussions.</u> *In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL), Sofia, Bulgaria.*

19. Agirre, Eneko, Daniel Cer, **Mona Diab**, Aitor Gonzalez-Agirre and Weiwei Guo. (2013) <u>*SEM 13 shared task: Semantic Textual Similarity</u>. *In Proceedings of *SEM,* Atlanta GA, USA.

20. ElFardy, Hebatullah, and **Mona Diab**. (2013) <u>Sentence-Level Dialect Identification in Arabic</u>*, In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL),* Sofia, Bulgaria.

21. ElFardy, Hebatullah, Mohamed Al-Badrashiny, **Mona Diab**. (2013) <u>Code-Switch Point Detection in Arabic</u>, *In Proceedings of the 18th International Conference on Application of Natural Language to Information Systems. (NLDB),* MediaCity, Manchester, UK.

22. Ghoneim, Mahmoud and **Mona Diab**. (2013) <u>Multiword Expressions in the context of Statistical Machine Translation.</u> *In Proceedings of IJCNLP*, Nagoya, Japan.

23. Guo, Weiwei, and **Mona Diab**. (2013) <u>Improving Lexical Semantics for Sentential Semantics: Modeling Selectional Preference and Similar Words in a Latent Variable Model.</u> *In Proceedings of North American Association for Computational Linguistics (NAACL),* Atlanta GA, USA.

24. Guo, Weiwei, Hao Li, Heng Ji and **Mona Diab**. (2013) <u>Linking Tweets to News: A Framework to Enrich Online Short Text Data in Social Media.</u> *In Proceedings of Association for Computational Linguistics (ACL),* Sofia, Bulgaria.

25. Hawwari, Abdelati, Wajdi Zaghouani, Tim O'Gorman, **Mona Diab**, and Ahmed Badran. (2013) Building a Lexical Semantic Resource for Arabic Morphological Patterns. *In Proceedings of ICCSPA13,* Sharjeh, UAE.
26. Tomeh, Nadi, Nizar Habash, Ryan Roth, Noura Farra, Pradeep Dasigi and **Mona Diab**. (2013) Ensemble Reranking with Linguistic and Semantic Features for Arabic Character Recognition, *In Proceedings of Association for Computational Linguistics (ACL),* Sofia, Bulgaria.
27. Zirikly, Aya and **Mona Diab**. (2013) ANEAR: Automatic Named Entity Aliasing Resolution. *In Proceedings of the 18th International Conference on Application of Natural Language to Information Systems (NLDB),* MediaCity, Manchester, UK.

[2012]

28. Abdul-Mageed, Muhammad and **Mona Diab**. (2012) Toward Building a large scale Arabic Sentiment Lexicon. *In Proceedings of the 6th International Global WordNet Conference*, Matsue, Japan.
29. Abdul-Mageed, Muhammad and **Mona Diab**. (2012) AWATIF: A multigenre corpus for MSA subjectivity and sentiment Analysis. *In Proceedings of Language Resources and Evaluation Conference (LREC),* Istanbul, Turkey.
30. Abu Jabara, Amjad, Pradeep Dasigi, **Mona Diab**, Dragomir Radev. (2012) Subgroup Detection in Ideological Discussions. *In Proceedings of Association for Computational Linguistics (ACL)*, Jeju Island, South Korea.
31. Dasigi, Pradeep, Weiwei Guo, and **Diab, Mona (2012)** Cross-genre subgroup detection in online discussion threads: A study of the implicit attitude using textual latent semantics. *In Proceedings of Association for Computational Linguistics (ACL),* Jeju Island, South Korea.
32. ElFardy, Hebatullah and **Mona Diab.** (2012) Simplified guidelines for the creation of Large Scale Dialectal Arabic Annotations. *In Proceedings of Language Resources and Evaluation Conference (LREC),* Istanbul, Turkey.
33. ElFardy, Hebatullah and **Mona Diab**. (2012) AIDA: Automatic Identification & Glossing of Dialectal Arabic. *Poster Session, In Proceedings of European Association for Machine Translation (EAMT),* Trento, Italy.
34. Guo, Weiwei and **Mona Diab**. (2012) Modeling Semantic Textual Similarity in the Latent Space. *In Proceedings of Association for Computational Linguistics (ACL),* Jeju Island, South Korea.
35. Guo, W. & **Diab, Mona.** (2012) Learning the latent semantics of a concept from its definition. *In Proceedings of Association for Computational Linguistics (ACL),* Jeju Island, South Korea.
36. Habash, Nizar, **Mona Diab,** Owen Rambow. (2012) CODA: A Conventional Orthography for Dialectal Arabic. *In Proceedings of Language Resources and Evaluation Conference (LREC),* Istanbul, Turkey.
37. Prabhakaran, Vinodkumar, Owen Rambow, **Mona Diab.** (2012) Who's. (Really) the Boss? Perception of Situational Power in Written Interactions. *In Proceedings of Conference on Computational Linguistics (COLING),* Mumbai, India.
38. Prabhakaran, Vinodkumar, Owen Rambow, **Mona Diab.** (2012) Annotations for Power Relations on Email Threads. *In Proceedings of Language Resources and Evaluation Conference (LREC),* Istanbul Turkey.
39. Prabhakaran, Vinodkumar, Owen Rambow, **Mona Diab.** (2012) Predicting Overt Display of Power in Written Dialogs. *In Proceedings of North American Association for Computational Linguistics (NAACL), Short Papers,* Montreal, Canada.

[2011]

40. Abdul-Mageed, Muhammad, **Mona Diab**, Mohammed Korayem. (2011) Subjectivity And Sentiment Analysis Of Modern Standard Arabic*. In Proceedings of Association for Computational Linguistics (ACL),* Portland OR, USA*.*
41. **Diab, Mona** and Pradeep Dasigi. (2011) CODACT: Towards Identifying Orthographic Variants in Dialectal Arabic. In Proceedings *of IJCNLP,* Chiang Mai, Thailand*.*
42. Ghoneim, Mahmoud, Marine Carpuat and **Mona Diab**. (2011) Linguistically motivated reordering constraints for phrase-based SMT: base phrase chunks and predicate argument structures*. In Proceedings of Eleventh Conference on Language Engineering (ESOLEC11),* Cairo, Egypt*.*
43. Guo, Weiwei and **Mona Diab**. (2011) Semantic Topic Models: Combining Word Distributional Statistics and Dictionary Definitions*. In Proceedings of Empirical Methods For Natural Language Processing (EMNLP),* Edinburgh, Scotland, UK*.*

[2010]

44. Benajiba Yassine, Imed Zitouni, **Mona Diab**, Paolo Rosso. (2010) Arabic Named Entity Recognition: Using Features Extracted from Noisy Data*. In Proceedings of Association for Computational Linguistics (ACL),* Uppsala, Sweden*.*
45. Carpuat, Marine, and **Mona Diab**. (2010) Task-based Evaluation of Multiword Expressions: a Pilot Study in Statistical Machine Translation*. In Proceedings of North American Association for Computational Linguistics (NAACL),* Los Angeles CA, USA*.*
46. Guo, Weiwei and **Mona Diab**. (2010) Combining Orthogonal Monolingual and Multilingual Sources of Evidence for All Words WSD. *In Proceedings of Association for Computational Linguistics (ACL*), Uppsala, Sweden*.*
47. Prabhakaran, Vinodkumar, Owen Rambow, **Mona Diab**. (2010) Automatic Committed Belief Tagging. *In Proceedings of COLING*, Beijing, China*.*

[2009]

48. **Diab, Mona. (200**9) Second Generation Tools. (AMIRA 2.0): Fast and Robust Tokenization, POS tagging, and Base Phrase Chunking. *MEDAR 2nd International Conference on Arabic Language Resources and Tools,* Cairo, Egypt.
49. **Diab, Mona** and Madhav Krishna. (2009) Unsupervised classification of Verb Noun Multiword Expression Tokens. *CICLING,* Mexico City, Mexico*.*
50. Parton, Kristen, Kathleen R. McKeown, Robert Coyne, **Mona T. Diab**, Ralph Grishman, Dilek Hakkani-Tür, Mary Harper, Heng Ji, Wei Yun Ma, Adam Meyers, Sara Stolbach, Ang Sun, Gokhan Tur, Wei Xu and Sibel Yaman Who. (2009) What, When, Where, Why? Comparing Multiple Approaches to the Cross-Lingual 5W Task*. In Proceedings of ACL-IJCNLP,* Singapore.

[2008]

51. BenAjiba, Yassine, **Mona Diab,** Paolo Rosso. (2008) Arabic Named Entity Recognition: An SVM Based Approach*. International Arab Conference on Information Technology (ACIT'08),* University of Sfax*,* Sfax, Tunisia*.*
52. BenAjiba, Yassine, **Mona Diab,** Paolo Rosso,. (2008) Arabic Named Entity Recognition using Optimized Feature Sets. *In Proceedings of Empirical Methods for Natural Language Processing (EMNLP*), Hawaii HI, USA.

53. **Diab, Mona,** Alessandro Moschitti, Daniele Pighin,. (2008) Semantic Role Labeling Systems for Arabic Language using Kernel Methods *In Proceedings of Association for Computational Linguistics (ACL),* Columbus OH, USA.

54. Palmer, Martha, Ann Bies, Olga Babko-Malaya, **Mona Diab**, Mohamed Maamouri, Aous Mansouri, Wajdi Zaghouni. (2008) A pilot Arabic Propbank. *In Proceedings of Language Resources and Evaluation Conference (LREC),* Marrakech, Morocco.

55. Roth, Ryan, Owen Rambow, Nizar Habash, **Mona Diab**, Cynthia Rudin. (2008) Arabic Morphological Tagging, Diacritization, and Lemmatization Using Lexeme Models and Feature Ranking. *In Proceedings of Association for Computational Linguistics (ACL),* Columbus OH, USA.

[2007]

56. **Diab, Mona**. (2007) Towards an optimal POS tag set for Modern Standard Arabic Processing. *In Proceedings of Recent Advances in Natural Language Processing (RANLP),* Borovets, Bulgaria.

57. **Diab, Mona** and Alessandro Moschitti. (2007) Semantic Parsing for Modern Standard Arabic. *In Proceedings of Recent Advances in Natural Language Processing (RANLP),* Borovets, Bulgaria.

58. **Diab, Mona**, Mahmoud Ghoneim, Nizar Habash. (2007) Arabic Diacritization in the Context of Statistical Machine Translation. *In Proceedings of Machine Translation Summit. (MT-Summit),* Copenhagen, Denmark.

59. Kirchhoff, Katrin, Owen Rambow, Nizar Habash, **Mona Diab**. (2007) Semi-Automatic Error Analysis for Large-Scale Statistical Machine Translation Systems. *In Proceedings of Machine Translation Summit. (MT-Summit),* Copenhagen, Denmark.

[2006]

60. Chiang, David, **Mona Diab**, Nizar Habash, Owen Rambow, Safi-Allah Sharif. (2006) Parsing Arabic Dialects. *In Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics (EACL),* Trento, Italy.

61. **Diab, Mona**. (2006) Developing a Mentoring Culture in Egypt. *Position paper at the Conference for the Association of Egyptian American Scholars (AEAS), National Research Council of Egypt,* Cairo, Egypt.

62. **Diab, Mona**, and Neal Snider. (2006) Modern Standard Arabic Verb Alternations. *In Proceedings of Annual Meeting of the Arabic Linguistics Society (ALS),* Kalamazoo MI, USA.

63. Maamouri, Mohamed, Ann Bies, Tim Buckwalter, **Mona Diab**, Nizar Habash, Owen Rambow, Dalila Tabessi (2006) Developing and Using a Pilot Dialectal Arabic Treebank. *In Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC),* Genoa, Italy.

64. Snider, Neal and **Mona Diab**. (2006) Automatic Discovery of Verb Classes in Modern Standard Arabic. *In Proceedings of the Joint Conference of the International Committee on Computational Linguistics and the Association for Computational Linguistics. (ACL-COLING),* Sydney, Australia.

65. Snider, Neal and, **Mona Diab**. (2006) Unsupervised Induction of Arabic Verb Classes. *In Proceedings of Human Language Technology Conference /North American chapter of the Association for Computational Linguistics Annual Meeting (HLT-NAACL),* New York, NY, USA.

[2004]

66. **Diab, Mona**. (2004) <u>Relieving the Data Acquisition Bottleneck for Word Sense Disambiguation</u>. *In Proceedings of Association for Computational Linguistics (ACL).*
67. **Diab, Mona**, Kadri Hacioglu, Daniel Jurafsky. (2004) <u>Automatic Tagging of Arabic Text: From Raw Text to Base Phrase Chunks</u>. *In Proceedings of Human Language Technology-North American Association for Computational Linguistics (HLT-NAACL).*
68. **Diab, Mona**. (2004) <u>Bootstrapping a WordNet Taxonomy for Arabic</u>. *Proceedings of First Arabic Language Technology Conference. (NEMLAR*), Cairo, Egypt.

[2002]

69. **Diab, Mona** and Philip Resnik. (2002) <u>An Unsupervised Method for Word Sense Tagging using Parallel Corpora. </u>*In Proceedings of Association for Computational Linguistics. (ACL),* Philadelphia PA, USA.

[2000]

70. **Diab, Mona** and Steven Finch. (2000) <u>A Statistical Word-Level Translation Model for Comparable Corpora</u>*. In Proceedings of Conference on Content-based Multimedia Information Access (RIAO00),* Paris, France.
71. Resnik, Philip and **Mona Diab**. (2000) <u>Measuring Verb Similarity</u>*. In Proceedings of the conference of International Cognitive Science Society (COGSCI00),* Philadelphia PA, USA.
72. Dorr, Bonnie, Gina Levow, Douglas Oard, Philip Resnik, Amy Weinberg, **Mona Diab**, Maria Katsova. (2000) <u>MADLIBS: An Event Translingual Lexical Conceptual Structure Based Information Retrieval System</u>. *In Proceedings of North American Association for Computational Linguistics. (NAACL),* Seattle, WA, USA.

[1998]

73. **Diab, Mona**, John Schuster, Peter Bock. (1998) <u>A Preliminary Statistical Investigation into the impact of an N-Gram Analysis Approach based on Word Syntactic Categories toward Text Author Classification</u>. *In Proceedings of 6th International Conference on Artificial Intelligence & Applications,* Cairo, Egypt.
74. Riopka, Terry, **Mona Diab**, Peter Bock. (1998) <u>Quantifying and Interpreting the Effect of Intelligent Information Exchange</u>. *In Proceedings of 6th International Conference on Artificial Intelligence & Applications,* Cairo, Egypt.

[1997]

75. Resnik, Philip, Mari Broman Olsen, **Mona Diab**. (1997) <u>Creating a Parallel Corpus from the Book of 00 Tongues</u>, *Text Encoding Initiative 10th Anniversary User Conference. (TEI-10),* Providence RI, USA.

PEER REVIEWED WORKSHOP PAPERS

[Total: 49]

[2016]

1. Molina, Giovanni, Fahad Algamdi, Mahmoud Ghoneim, Abdelati Hawwari, **Mona Diab**, Solorio, Thamar. (2016) An overview of the shared task on multilingual linguistic code switching. *Proceedings of EMNLP Workshop on Computational Approaches to Linguistic Code Switching (CALCS 2016), EMNLP,* Austin TX, USA, Nov.

2. Alghamdi, Fahad, Giovanni Molina, **Mona Diab,** Thamar Solorio**.** (2016) Part of Speech Taggers for Code Switched Data. *Proceedings of EMNLP Workshop on Computational Approaches to Linguistic Code Switching (CALCS 2016), EMNLP*, Austin TX, USA, Nov.

3. AlBadrashiny, Mohamed and **Mona Diab.** (2016) The George Washington University System for the Code-Switching Workshop Shared Task 2016. *Proceedings of EMNLP Workshop on Computational Approaches to Linguistic Code Switching (CALCS 2016), EMNLP*, Austin TX, USA, Nov.

4. Ossama Obeid, Houda Bouamor, Wajdi Zaghouani, Mahmoud Ghoneim, Abdelati Hawwari, **Mona Diab**, Kemal Oflazer. (2016) MANDIAC: A Web-based Annotation System For Manual Arabic Diacritization. *Proceedings of the 2nd Workshop on Arabic Corpora and Processing Tools*, LREC 2016. **BEST POSTER AWARD FROM QNRF**

5. Abdul-Mageed, Muhammad, Hassan AlHuzliy, Duaa' Abu Elhija, **Mona Diab.** (2016) DINA: A Multi-Dialect Dataset for Arabic Emotion Analysis. *Proceedings of the 2nd Workshop on Arabic Corpora and Processing Tools,* LREC 2016.

6. Elfardy, Heba and **Mona Diab.** (2016) CU-GW Perspective at SemEval-2016 Task 6: Ideological Stance Detection in Informal Text. *In Proceedings of the International Workshop on Semantic Evaluation (SemEval 2016), NAACL 2016.* San Diego, CA, USA.

7. Elfardy, Heba and **Mona Diab.** (2016) Annotation Complexity: The Case of Annotating Ideological Perspective in Egyptian Social Media. *In Proceedings of the 10th Linguistic Annotation Workshop (LAW X), ACL 2016***.** Berlin, Germany.

8. AlDarmaki, Hanan, and **Mona Diab.** (2016) GW-NLP at SemEval Task 1: Matrix Factorization for Cross Lingual STS. *In Proceedings of the International Workshop on Semantic Evaluation (SemEval 2016), NAACL 2016.* San Diego, CA, USA.

9. AlDarmaki, Hanan, and **Mona Diab.** (2016) Learning Cross-Lingual Representations with Matrix Factorization. *Proceedings of Multilingual and Cross-Lingual Methods in Computational Linguistics (MLCL), NAACL*, San Diego CA, Jun.

10. Hamidian, Sardar, and **Mona Diab.** (2016) Rumor Identification and Belief Investigation on Twitter. *In Proceedings of Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis (WASSA), NAACL 2016*, San Diego CA, Jun*.*

11. Agirre, Eneko, Carmen Banea, Daniel Cer, **Mona Diab,** Aitor Gonzalez-Agirre, Rada Mihalcea, German Rigau, Janyce Wiebe. (2016) SemEval-16 Task 1: Semantic Textual Similarity, Monolingual and Cross Lingual Evaluation*. In Proceedings of SEMEVAL, North American Association for Computational Linguistics, NAACL 2016*, San Diego CA, Jun*.*

[2015]

12. Agirre, Eneko, Carmen Banea, Claire Cardie, Daniel Cer, **Mona Diab**, Aitor Gonzalez-Agirre, Weiwei Guo, Inigo Lopez-Gazpio, Montse Maritxalar, Rada Mihalcea, German Rigau, Larraitz Uria, Janyce Wiebe. (2015) SemEval-15 Task 2: Semantic Textual Similarity, English, Spanish and Pilot on Interpretability. *In Proceedings of SEMEVAL, North American Association for Computational Linguistics (NAACL),* Denver CO, USA*.*

13. Aminian, Maryam, Mahmoud Ghoneim, **Mona Diab**. (2015) Unsupervised False Friend Disambiguation Using Contextual Word Clusters and Parallel Word Alignments. In Proceedings *of Workshop 9th Semantics Syntax Statistical Translation (SSST), North American Association for Computational Linguistics (NAACL),* Denver CO, USA*. BEST PAPER AWARD.*

14. Werner, Gregory, and Vinodkumar Prabhakaran, **Mona Diab** and Owen Rambow. (2015) Committed Belief Tagging on the Factbank and LU Corpora: A Comparative Study. In Proceedings *of Workshop EXPROM*, *North American Association for Computational Linguistics (NAACL),* Denver CO, USA.
15. Zirikly, Ayah, **Mona Diab**. (2015) Named Entity Recognition for Arabic Social Media. In Proceedings *of Workshop on Vector Space Models for NLP*, *North American Association for Computational Linguistics (NAACL),* Denver CO, USA.
16. Bouamor, Houda, Wajdi Zaghouani, **Mona Diab**, Ossama Obeid, Kemal Oflazer, Mahmoud Ghoneim, Abdelati Hawwari. (2015) A Pilot Study on Arabic Multi-Genre Corpus Diacritization. *In Proceedings* of Second Workshop on Arabic Natural Language Processing (ANLP), *Association for Computational Linguistics (ACL),* Beijing, China.
17. Attia, Mohammed, Mohamed Al-Badrashiny, **Mona Diab**. (2015) GW-HASP-15$@$QALB-15 Shared Task: Priming Spelling Candidates with Probability*. In Proceedings of Second Workshop on Arabic Natural Language Processing (ANLP), Association for Computational Linguistics (ACL),* Beijing, China.
18. Aldarmaki, Hanan and **Mona Diab**. (2015) Robust Part-of-speech Tagging of Arabic Text. *In Proceedings of Second Workshop on Arabic Natural Language Processing (ANLP), Association for Computational Linguistics (ACL),* Beijing, China.

[2014]

19. Agirre, Eneko, Carmen Banea, Claire Cardie, Daniel Cer, **Mona Diab**, Aitor Gonzalez-Agirre, Weiwei Guo, Rada Mihalcea, German Rigau, Janyce Wiebe. (2014) SemEval-14 Task 10: Multilingual Semantic Textual Similarity. *In Proceedings of SEMEVAL*, *Conference of Computational Linguistics (COLING),* Dublin, Ireland.
20. Aminian, Maryam, Mahmoud Ghoneim, **Mona Diab**. (2014) Handling OOV Words in Dialectal Arabic to English Machine Translation. *In Proceedings of Workshop on Language Technology for Closely Related Languages and Language Variants (LT4CloseLang), Empirical Methods For Natural Language Processing (EMNLP),* Doha, Qatar.
21. Attia, Mohamed, Mohamed AlBadrashiny, **Mona Diab**. (2014) GW-HASP: Hybrid Arabic Spelling and Punctuation Correction*. In Proceedings of Workshop on Arabic Natural Language Processing (ANLP), Empirical Methods For Natural Language Processing (EMNLP),* Doha, Qatar.
22. ElFardy, Hebatullahtullah, Mohamed AlBadrashiny, **Mona Diab**. (2014) AIDA: Identifying Code Switching in Informal Arabic Text. *In Proceedings of First Workshop on Computational Approaches to Linguistic Code Switching (CodeSwitch), Empirical Methods For Natural Language Processing* (EMNLP), Doha, Qatar.
23. Hawwari, Abdelati, Mohamed Attia, **Mona Diab**. (2014) A Framework for the Classification and Annotation of Multiword Expressions in Dialectal Arabic. *In Proceedings of Workshop on Arabic Natural Language Processing (ANLP), Empirical Methods For Natural Language Processing (EMNLP),* Doha, Qatar.
24. Solorio, Thamar, E. Blair, S. Maharjan, S. Bethard, M. Diab, M. Ghoneim, A. Hawwari, F. AlGhamdi, J. Hirschberg, A. Chang, P. Fung. (2014) Overview for the First Shared Task on Language Identification in Code Switched Data. *In Proceedings of First Workshop on Computational Approaches to Linguistic Code Switching (CodeSwitch)*, *Empirical Methods For Natural Language Processing (EMNLP),* Doha, Qatar.
25. Zirikly, Ayah, **Mona Diab**. (2014) Named Entity Recognition for Dialectal Arabic. *In Proceedings of Workshop on Arabic Natural Language Processing. (ANLP*), Empirical Methods For Natural Language Processing (EMNLP), Doha, Qatar.

[2012]

26. Abdul-Mageed, M., Kuebler, S., Mona Diab. (2012) <u>SAMAR: A system for subjectivity and sentiment analysis of social media Arabic.</u> *In Proceedings of the 3rd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis. (WASSA), held in conjunction with the 50th Annual Meeting on Association for Computational Linguistics (ACL),* Jeju Island, South Korea.

27. Agirre, Eneko, Daniel Cer, **Mona Diab**; Aitor Gonzalez-Agirre. (2012) <u>SemEval-12 Task 6: A Pilot on Semantic Textual Similarity.</u> *In Proceedings of SEMEVAL Workshop, North American Association for Computational Linguistics (NAACL),* Montreal, Canada*.*

28. Guo, W. & Diab, Mona. (2012) <u>A simple unsupervised latent semantics based approach for sentence similarity.</u> *In Proceedings of the 6th International Workshop on Semantic Evaluation. (SEMEVAL), North American Association for Computational Linguistics (NAACL),* Montreal, Canada.

29. Hawwari, Abdelati; Kfir Bar; **Mona Diab**. (2012) <u>Building an Arabic Multiword Expressions Repository</u>*. In Proceedings of Workshop on Computational Linguistics for Literature, North American Association for Computational Linguistics (NAACL),* Montreal, Canada.

30. Prabhakaran, Vinodkumar, Michael Bloodgood, **Mona Diab**, Bonnie J. Dorr, Lori Levin, Christine Piatko, Owen Rambow, Benjamin Van Durme. (2012) <u>Statistical Modality Tagging from Rule-based Annotations and Crowdsourcing.</u> *In Proceedings of Workshop on Extra-propositional aspects of meaning in computational linguistics. (ExProM)* Jeju Island, South Korea.

31. Zaghouani, Wajdi, Abdelati Hawwari, **Mona Diab**. (2012) <u>A Pilot PropBank Annotation for Quranic Arabic.</u> *In Proceedings of the computational linguistics for literature workshop, North American Association for Computational Linguistics (NAACL)*, Montreal, Canada.

32. Abdul-Mageed, Muhammad and **Mona Diab**. (2011) <u>Subjectivity and Sentiment Annotation of Modern Standard Arabic Newswire</u>*. In Proceedings of workshop LAW V, Association for Computational Linguistics (ACL*), Portland OR, USA*.*

[2011]

33. Dasigi, Pradeep and **Mona Diab**. (2011) <u>Noisy Channel Modeling for Named Entity Transliteration</u>. In Proceedings *of NEWS 11 Workshop, IJCNLP,* Chiang Mai, Thailand.

34. **Diab, Mona** and Ankit Kamboj. (2011) <u>An Annotation Scheme Leveraging Crowd Sourcing for Linguistic Code Switching in Hinglish.</u> *In Proceedings of ALR9 Workshop, IJCNLP*, Chiang Mai, Thailand*.*

[2010]

35. **Diab, Mona,** Nizar Habash, Owen Rambow, Mohamed Al Tantawy, Yassine Benajiba. (2010) <u>COLABA: Arabic Dialect Annotation and Processing</u>*. In Proceedings of the Workshop on Semitic Language Processing, Language Resources and Evaluation Conference (LREC),* Malta*.*

36. Benajiba, Yassine and **Mona Diab**. (2010) <u>A Web Application for Dialectal Arabic Text Annotation</u>*. In Proceedings of the Workshop on Semitic Language Processing, Language Resources and Evaluation Conference (LREC),* Malta*.*

37. Guo, Weiwei and **Mona Diab**. (2010) <u>COLEPL and COLSLM: An Unsupervised WSD Approach to Multilingual Lexical Substitution, Tasks 2 and 3 SemEval 10</u>*. In Proceedings of SEMEVAL, Association for Computational Linguistics (ACL),* Uppsala, Sweden.

38. Zaghouani, Wajdi, **Mona Diab**, Aous Mansouri, Sameer Pradhan, Martha Palmer. (2010) <u>The</u>

Revised Arabic PropBank. *In Proceedings of Linguistic Annotation Workshop. (LAW) IV, Association for Computational Linguistics (ACL),* Uppsala, Sweden.

[2009]

39. **Diab, Mona**, and Pravin Bhutada. (2009) Verb Noun Construction MWE Supervised Token Classification. *In Proceedings of ACL-IJCNLP 09 Workshop on Multiword Expressions: Identification, Interpretation, Disambiguation and Applications, ACL-IJCNLP,* Singapore.
40. **Diab, Mona**, Lori Levin, Teruko Mitamura, Owen Rambow, Vinodkumar Prabhakaran, Weiwei Guo. (2009) Committed Belief Annotation and Tagging. *In Proceedings of workshop on Third Linguistic Annotation Workshop (LAW III), ACL-IJCNLP,* Singapore.
41. **Diab, Mona** and Madhav Krishna. (2009) Handling Sparsity for Unsupervised MWE Classification. *GEMS Workshop, European Association for Computational Linguistics (EACL),* Athens, Greece.
42. Guo, Weiwei, and **Mona Diab**. (2009) Improvements to Monolingual English Word Sense Disambiguation. *In Proceedings of the NAACL 09 workshop: Semantic Evaluations: Recent Achievements and Future Directions. (SEW-09), North American Association for Computational Linguistics (NAACL),* Boulder CO, USA.

[2008]

43. BenAjiba, Yassine, **Mona Diab**, Paolo Rosso. (2008) Arabic Named Entity Recognition using Conditional Random Fields. *In Proceedings of Workshop on HLT & NLP within the Arabic world. Arabic Language and local languages processing: Status Updates and Prospects, 6th Int. Conf. on Language Resources and Evaluation Conference (LREC),* Marrakech, Morocco.
44. Habash, Nizar, Owen Rambow, **Mona Diab**, Reem Farraj. (2008) Guidelines for Annotating Arabic Dialect, *Proceedings of Workshop on Arabic and its local languages, Language Resources and Evaluation Conference (LREC),* Marrakech, Morocco.

[2007]

45. **Diab, Mona**. (2007) Improved Arabic Base Phrase Chunking with a new enriched POS tag set. *In Proceedings of workshop on Computational Approaches to Semitic Languages (CASL), Association for Computational Linguistics (ACL),* Prague, Czech Republic.
46. **Diab, Mona**, Musa Alkhalifa, Sabry ElKateb, Christiane Fellbaum, Aous Mansouri, Martha Palmer. (2007) SemEval-07 Task 18: Arabic Semantic Labeling. *In Proceedings of SEMEVAL, Association for Computational Linguistics (ACL),* Prague, Czech Republic.
47. **Diab, Mona**, Alessandro Moschitti, Daniele Pighin. (2007) CUNIT: A Semantic Role Labeling System for Modern Standard Arabic. *In Proceedings of SEMEVAL, Association for Computational Linguistics (ACL),* Prague, Czech Republic.

[2004]

48. **Diab, Mona**. (2004) An Unsupervised Approach for Bootstrapping Arabic Sense Tagging. *In Proceedings of Arabic Workshop, Computational Linguistics (COLING).*

[2000]

49. **Diab, Mona**. (2000) An Unsupervised Method for Word Sense Tagging using Parallel Corpora: A Preliminary Investigation. *In Proceedings of Special Interest Group in Lexical*

Semantics (SIGLEX) Workshop, Association for Computational Linguistics (ACL), Hong Kong.

TECHNICAL REPORTS

[Total: 10]

1. Habash, Nizar, **Mona Diab**, Owen Rambow. (2014) Conventional Orthoraphy for Dialectal Arabic (CODA) Principles and Guidelines- Egyptian Arabic – Version 0.7- March 14. *Technical Report,* http://dx.doi.org/10.7916/D83X8562, *Center for Computational Learning Systems. (CCLS), Columbia University*
2. **Diab, Mona**, Nizar Habash, Owen Rambow, Ryan Roth. (2013) LDC Arabic Treebanks and Associated Corpora: Data Divisions Manual. *Technical Report,* http://hdl.handle.net/10022/AC:P:21761, *Center for Computational Learning Systems. (CCLS), Columbia University*
3. Habash, Nizar, **Mona Diab**, Owen Rambow. (2011) Conventional Orthography for Dialectal Arabic. (CODA) V.1.0. *Technical Report 137382, http://academiccommons.columbia.edu/catalog/ac:137382, Center for Computational Learning Systems. (CCLS), Columbia University*
4. Rambow, Owen, David Chiang, **Mona Diab**, Nizar Habash, Rebecca Hwa, Khalil Sima'an, Vincent Lacey, Roger Levy, Carol Nichols, and Safi-Allah Shareef. (2005) Parsing Arabic Dialects. *Final Report, JHU Summer Workshop*
5. **Diab, Mona**. (2000) Exploiting Translations for Semantic Annotation. *Candidacy Exam Thesis, Linguistics Department, University of Maryland,* College Park MD, USA.
6. Riopka, Terry, **Mona Diab**, Peter Bock. (2000) Quantifying and Interpreting the Effect of Intelligent Information Exchange Between Chromosomes in a Human Simulation of a Genetic Algorithm. *Technical Report: LAMP-TR-045/UMIACS-TR-00-38/CS-TR-4147, University of Maryland College Park*
7. Resnik, Philip and **Mona Diab**. (2000) Measuring Verb Similarity. *Technical Report: LAMP-TR-047/ UMIACS-TR-00-40/CS-TR-4149, University of Maryland College Park*
8. **Diab, Mona** and Steve Finch. (2000) A Statistical Word-Level Translation Model for Comparable Corpora. *Technical Report: LAMP-TR-048/ UMIACS-TR-00-41/CS-TR-4150, University of Maryland College Park*
9. **Diab, Mona**, John Schuster, Peter Bock. (2000) A Preliminary Statistical Investigation into the Impact of an N-Gram Analysis Approach Based on Word Syntactic Categories Toward Text Author Classification. *Technical Report: LAMP-TR-046/UMIACS-TR-00-39/CS-TR-4148, University of Maryland College Park*
10. **Diab, Mona**, Douglas W. Oard, Tapas Kanungo, Mark Ames, Russell Ball. (2000) Using Arabic Examples to find English Documents. *Technical Report, University of Maryland College Park*

TUTORIALS

1. **Diab, Mona** and Nizar Habash. (2006) Arabic Dialect Processing.
2. **Diab, Mona.** (2005) Statistical NLP: An A-Z Crash Course.

PRESENTATIONS & POSTERS (Not Associated with Publications)

[Total: 39]

1. *Emotion, Intensity, and Urgency.* (2016) **DARPA LORELEI Site Visit**, Columbia University, New York NY, USA, Jul.
2. *Committed Belief Tagging.* (2016) **DARPA DEFT Site Visit,** Columbia University, New York NY, USA, Jul.
3. *Named Entity Linking & Event Detection.* (2016) **DARPA DEFT Site Visit**, Columbia University, New York NY, USA, Jul.
4. *Arabic Dialect Tools.* (2014) **DARPA BOLT Site Visit**, March, Columbia University, New York NY, USA, Mar.
5. *Perspective, Novelty and Belief in Streaming data and informal genres.* (2014) **DARPA DEFT Site visit**, George Washington University, Washington DC, USA, Feb.
6. *Arabic Dialect Processing under BOLT*. Invited Presentation. (2011) **DARPA BOLT Kick-off Meeting**, Tampa FL, USA, Oct.
7. *Arabic Language Modeling for OCR.* (With Nizar Habash) Invited Presentation. (2011) **DARPA MADCAT PI Meeting**. Boston MA, USA, Jul.
8. *Arabic Subjectivity and Sentiment Analysis: System & Resources*. Invited Poster. (2011) DARPA **BOLT Kick-off Meeting**, Tampa FL, USA, Oct.
9. *CADIM Arabic MSA NLP Tools*. (With Nizar Habash and Owen Rambow) Invited Poster, (2011) DARPA **BOLT Kick-off Meeting**, Tampa FL, USA, Oct.
10. *COLABA Tools and Resources*. (With Nizar Habash and Owen Rambow) Invited Presentation (2011) DARPA **BOLT Kick-off Meeting**, Tampa FL, USA, Oct.
11. *COLABA Tools and Resources*. (With Nizar Habash and Owen Rambow) Invited Presentation. (2011) **Brigham Young University**, Salt Lake City UT, USA, Aug 11.
12. *Conventionalized Dialectal Arabic. (CODA): Guidelines and Tools*. (With Nizar Habash and Owen Rambow) Invited Poster. (2011) DARPA **BOLT Kick-off Meeting**, Tampa FL, USA, Oct.
13. *Multilayered Semantic Annotation and challenges of Interoperability*. Invited Presentation. (2011) **NSF-funded SILT/FLaReNet Forum on Sustainable Interoperability for Language Technologies.** Brandeis University, Boston MA, USA, Apr.
14. *Power Shifts & Rifts in Social Communications*. (With Kathleen McKeown, Owen Rambow, Dragomir Radev) Invited Presentation. (2011) **IARPA SCIL PI Meeting**. Washington DC, USA, May.
15. *Power Shifts & Rifts in Social Communications*. (With Kathleen McKeown, Owen Rambow, Dragomir Radev) Invited Presentation. (2011) **IARPA SCIL PI Meeting**. Washington DC, USA, Nov 11.
16. *Arabic Dialect Processing within the Context of Social Media*. (With Nizar Habash and Owen Rambow) Invited Presentation. (2010) **COLABA Review Day, MITRE Corporation.** McLean VA, USA, Apr 10.
17. *Arabic Language Modeling for OCR*. (With Nizar Habash) Invited Presentation. (2010) **DARPA MADCAT PI Meeting.** Boston MA, USA, Sep.
18. *COLABA Annotation, Tools, Resources*. (With Nizar Habash and Owen Rambow) Invited Presentation. (2010) **Government Day AMTA 10**. Denver CO, USA, Oct.
19. *Power Shifts & Rifts*. (With Kathleen McKeown, Owen Rambow, Dragomir Radev) Invited Presentation. (2010) **IARPA SCIL PI Meeting**. Washington DC, USA, May.
20. *Arabic Language Modeling for OCR*. (With Nizar Habash) Invited Presentation. (2009) **DARPA MADCAT PI Meeting**. Boston MA, USA, Sep.
21. *Power Shifts & Rifts*. (With Kathleen McKeown, Owen Rambow, Dragomir Radev) Invited Presentation. (2009) **IARPA SCIL PI Meeting**. Washington DC, USA, Oct.
22. *CADIM: Natural Language Processing for Arabic and its dialects*. (With Nizar Habash and Owen Rambow) Invited Presentation. (2008) **CASL, University of Maryland.** College Park MD, USA, Jun.

23. *Columbia CCLS: Committed Belief and Dialog Acts.* (With Owen Rambow and Becky Passaneau) Invited Presentation. (2008) **Human Language Technology Center of Excellence, University of Maryland.** College Park MD, USA**,** Sep.
24. *Factivity Annotations.* (With Owen Rambow, Lori Levin, and Teriku Mitamura) Invited Presentation. (2008) **JHU Human Language Technology Center of Excellence, University of Maryland.** College Park MD, USA, May.
25. *SALAMAT: Sense Assignment Leveraging Alignments, Monolingual And Translation.* (With Weiwei Guo) Invited Demo/Poster. (2008) **NSF sponsored symposium on Semantic Knowledge Discovery, Organization and Use**, **NYU**. New York NY, USA, Nov.
26. *Tools for Arabic Morphological Analysis, Disambiguation and Tokenization.* (With Nizar Habash and Owen Rambow) Invited Presentation. (2008) **Technology Forum, DIA and TSWG**. McLean VA, Jul.
27. *Impact of Partial Arabic Diacritization on Statistical Machine Translation.* (With Mahmoud Ghoneim and Nizar Habash) Invited Presentation. (2007) **DARPA GALE PI Meeting**. San Francisco CA, USA, May.
28. *Semi-Automatic Error Analysis of the NIGHTINGALE Machine Translation System.* (With Katrin Kirchhoff, Nizar Habash, Owen Rambow, Evgeny Matusov) Invited Presentation. (2007) **DARPA GALE PI Meeting**. San Francisco CA, USA, May.
29. *Modern Standard Arabic Verb Alternations.* (With Neal Snider) (2006) Paper presented at the **Annual Meeting of the Arabic Linguistics Society. (ALS) Conference**. Kalamazoo MI, USA, Mar.
30. *NLP Tools for Arabic.* (With Nizar Habash and Owen Rambow) Invited Presentation (2006) **DARPA GALE PI Meeting**. Boston MA, USA, May.
31. *Treebank Transduction. Arabic Dialect Parsing**.** (With Owen Rambow, Nizar Habash, Khalil Simaan, David Chiang, Rebecca Hwa, Roger Levy, Vincent Lacey, Saif-Allah Shareef) (2005) **Johns Hopkins Summer Workshop Final Presentation Symposium.** Baltimore MD, USA, Aug.
32. *An Unsupervised method for Word Sense Tagging using Parallel corpora. (2001)* Poster Session: **UMIACS Research Day, University of Maryland.** College Park MD, USA, Mar.
33. *A Statistical Model of Word-Level Mapping for Comparable Corpora.* (2000) **UMIACS Computational Linguistics Colloquium, University of Maryland.** College Park MD, USA*,* Spr.
34. *An Unsupervised Method for Word Sense Tagging using Parallel corpora.* (2000) **North East Scholars NLP workshop: IBM T.J. Watson Research Center,** Hawthorne NY, USA*,* Oct.
35. *MADLIBS: An LCS based Translingual Information Retrieval System.* (With Maria Katsova) (2000) **Demonstration: NAACL2000.** Seattle WA, USA, May.
36. *Statistical Methods and Linguistics.* (With Elixabete Murguia) (2000) **LingFest**, **Linguistics Department University of Maryland.** College Park MD, USA, Nov.
37. *Cross Language Information Retrieval: An LCS based approach for Spanish queries and English Documents.* (With Maria Katsova) (1999) **Poster Session: UMIACS Research Day, University of Maryland.** College Park MD, USA, Feb.
38. *Word Sense Disambiguation: Lexical Ambiguity Resolution within a Selection Preference Model.* (1999) **Linguistics Student Conference, Linguistics Dept., University of Maryland.** College Park MD, USA, Fall.
39. *A Preliminary Statistical Investigation into the Impact of Using an N-Gram Analysis for the Purposes of Text Author Identification/Classification. (1997)* **Linguistics Student Conference, Linguistics Dept., University of Maryland.** College Park MD, USA. Fall.

RELEASED SOFTWARE

ROBUST RELEASED TECHNOLOGY

1. [DIRA] Dialectal Arabic Information Retrieval Assistant. (2013)
   Demo: http://nlp.ldeo.columbia.edu/dira
2. [AIDA] Automatic Identification of Dialectal Arabic. (2013)
   Demo: http://nlp.ldeo.columbia.edu/aida
3. [MADAMIRA] A system for morphological analysis and disambiguation, base phrase chunking and named entity recognition. (2013)
   Demo: http://nlp.ldeo.columbia.edu/madamira.

RESEARCH PROTOTYPE TECHNOLOGY

1. **Diab, Mona** and Pradeep Dasigi**.** (2011) ***CODACT*** A Tool for Arabic Dialect Orthographic Conventionalization. https://nlp.ccls.columbia.edu/CODACT
2. Benajiba, Yassine and **Mona Diab**. (2010) ***NIDA*** Arabic Named Entity Recognition. Current Version 2.0. https://nlp.ccls.columbia.edu/NIDA
3. **Diab, Mona,** Yassine Benajiba, Hebatullah ElFardy**.** (2011) ***AMIRA*** Automatic Segmentation/Tokenization, Lemmatization, Diacritization, POS Tagging, Base Phrase Chunking/Shallow Parsing of Arabic Text. Current Version 2.0.
   https://nlp.ccls.columbia.edu/AMIRA
4. Benajiba, Yassine**, Mona Diab,** Hebatullah ElFardy**.** (2011) ***COLANN*** COLABA Annotation Interface v.1.0: An annotation interface for Arabic Dialects. Current Version 1.0. https://nlp.ccls.columbia.edu/COLANN
5. **Diab, Mona** and Weiwei Guo. (2009) ***SALAAM-CONT*** Automatic Multilingual Online All words Word Sense Disambiguation. Current Version 1.0.
   https://nlp.ccls.columbia.edu/COLANN
6. Prabhakaran, Vinodkumar, **Mona Diab**, Owen Rambow. (2010) ***COMBEL*** Automatic Committed Belief Tagging for Running Text. Current Version 1.0.
   https://nlp.ccls.columbia.edu/COMBEL

PROFESSIONAL ACTIVITIES

ACADEMIC SERVICES WITHIN GW

CS Department Level
- [Sum.2016 -*present*] **Member,** *Dept. Promotion & Tenure Committee for Prof. Tim Wood*
- [Spring 2014] Established in collaboration with other CS faculty *Undergraduate track on Data Science*
- [Fall.2014- *present*] **Member**, *Dept. Committee Grad. Admissions & Support (GAS)*
- [Fall.2013 –Fall.2015] **Member**, *Dept. Committee on Awards*
- [Fall.2014-Spring.2015] **Co-chair**, *Big Data Chair Search Committee*
- [Fall.2014- Spring.2015] **Member**, *Dept. Committee on Budget*
- [Fall.2014- Spring.2015] **Member,** *Dept. Committee on Student Relations*
- [Fall.2013 – Spring.2014] **Member**, *Dept. Committee on Research*
- [Spring.2013 – Spring.2014] **Member**, *Dept. Committee on Graduate Student Admission*
- [Spring.2013] **Member**, *Dept. Committee on Curriculum*
- [Spring.2013] **Member**, *Dept. Committee on Faculty Retreat*
- [Spring.2013] **Member**, *Dept. Committee on Target of Opportunity Hires*

SEAS Level
- [Fall.2014 – *present*] **Member**, *Steering Committee overall IBM collaboration and Point Person for any work or education related to Watson*
- [Spring.2014 – Fall.2014] **Elected Senator**, representing *SEAS on Faculty Senate*
- [Fall.2013 – Spring.2014] **Co-chair,** *Big Data Initiative Task Force*

GW Level
- [Spring.2016] **Co-designer** of new MSc Degree in Data Analytics (SEAS, Business School)
- [Fall.2015-present] **Member,** *Vice President of Research Advisory Board*

ACADEMIC SERVICES WITHIN COLUMBIA UNIVERSITY

- [Fall.2005–Fall.2012] **Member** of *Columbia Senate Committee for Research Officers*,
  - The committee represents a ~3k constituency of Research Officers at Columbia University including: Professional Research Officers, Post-Docs, and Research Staff. It comprises 9 members.
  - Leading the effort on the creation of professorial titles for Professional Research Officers who serve as independent PIs and advisors of record for students.
- [Fall.2005–Fall.2012] **Representing** *Columbia University Bi-annual National Council on Women in Information Technology. (NCWIT) meetings*
- [Fall.2011–Fall.2012] **Member** of the *Columbia Senate Committee on the External Affairs and Public Relations*
- [Fall.2006–Fall.2012] **Member** of *CCLS Strategy Committee*
- [Spring.2010–Fall.2012] **Member** of the *Columbia Senate Commission on the Status of Women*
- [Fall.2009–Fall.2012] **Member** of *SEAS Dean's Committee for Research Scientists*
- [Spring.2010–Fall.2012] **Organizer** of *CCLS-Yahoo Distinguished Lecture Series*
- [Spring.2008–Fall.2009] **Member** of *Columbia Senate committee on University Rules of Conduct*
- [Spring.2008–Fall.2009] **Member** of *Columbia Senate committee on Alumni Relations*
- [Fall.2007–Spr.2008] **Organizer** of *CCLS Colloquium Series*
- [Fall.2007–Spring.2009] **Represented** CCLS in the *SEAS Executive Council Meetings*
- [Fall.2005–Spring.2006] **Member** of *Columbia Senate Committee on Physical Development*
- [Spring.2004–Fall.2004] **Organizer** of *Stanford Computational Linguistics Colloquium Series*

SERVICES WITHIN NLP COMMUNITY

- *[06.2016-11.2016] Chair Nominating committee for NAACL 2016-2018 Board Membership*
- *[04.2016-08.2016] Returning Officer for ACL SIGLEX Board Elections.*

Grant Proposal Reviewer

- *[12.2015] NSF RI CI Panel*
- *[01.2012-12.2015] KACST Research Fund*
- *[03.2013] NSF Semantics Panel*
- *[02.2012] NSF OISE Program*
- *[03.2010-09.2012] Qatar National Research Fund*
- *[04.2011] NSF Linguistics Program*
- *[01.2010] NSF SoCS Program Panel*
- *[05.2008] NSF HSD Awards Panel*
- *[10.2010] NSF Career Awards*

EVENT/ CONFERENCE/WORKSHOP ORGANIZATION

- [Fall.2016-Spring.2017] **Co-Chair** *Semeval 17 Task 1, Semantic Textual Similarity*
- [11.2016] **Co-Chair** *2nd Workshop on Computational Approaches to Linguistic Code Switching (CALCS), Proceedings of EMNLP 2016, Austin TX, USA*
- [11.2016] **Chair** *Workshop on Machine Translation for Semitic Languages (SeMAT), AMTA 2016, Austin TX, USA*
- [Fall.2015-Spring.2016] **Co-Chair** *Semeval 16 Task 1, Semantic Textual Similarity, NAACL 16, San Diego CA, USA.*
- [Fall.2014-Spring.2015] **Co-Chair** *Semeval 15 Task 2, Semantic Textual Similarity, NAACL 15, Boulder CO, USA.*
- [04.2015] **Organizer**, Mutliword Expressions Special Event, *CICLING 15,* Cairo, Egypt.
- [Spring.2014] **Co-Chair,** Town Hall Meeting for Big Data efforts at GW, Washington DC, USA
- [Fall.2014] **Co-Chair**, *First Workshop on Computational Approaches to Linguistic Code Switching, Empirical Methods For Natural Language Processing (EMNLP), Doha, Qatar.*
- [Spring.2014-Fall.2014] **Publicity Co-Chair**, *Empirical Methods For Natural Language Processing (EMNLP), Doha, Qatar.*
- [Spring.2014-Fall.2014] **Co-Chair** *Semeval 14 Task 10, Multilingual Semantic Textual Similarity**, International Conference of Computational Linguistics (COLING), Dublin, Ireland.*
- [Spring 2013] **General Chair**, **SEM 2nd Joint Semantics Conference*
- [02.2012] **Chair**, *DARPA Funded Unified Semantic Framework for Textual Similarity Workshop*
- [08.2011-07.2012] **Area Co-Chair** *for ACL 12 Multilinguality*
- [08.2011-06.2012] **Program Co-Chair** *of 1st International Joint Conference on Semantics *SEM 12*
- [06.2011-06.2012] **Workshops Co-Chair**, *NAACL 12*
- [07.2011-06.2012] **Co-Organizer** *SEMEVAL Task on Multilayer Semantic Annotation and STS*
- [12.2010] **Chair and Co-Organizer** *of 2nd NSF Funded Workshop on Linguistic Code Switching*
- [06.2010] **Chair and Co-Organizer** *of NSF Funded Workshop on Linguistic Code Switching*
- [02.2010] **Area Chair** *for NAACL Semantics and Machine Learning*
- [05.2009] **Area Chair** *for Empirical Methods For Natural Language Processing (EMNLP) Multilinguality*
- [08.2008] **Co-Chair** *COLING-HLT Workshop on TextGraphs*
- [05.2008] **Co-Chair** *LREC Workshop on Arabic and Local Languages*

- [04.2008] **Chair and Co-Organizer** *of NSF Funded Workshop on Linguistics Code Switching*
- [06.2007] **Organizer** *of SEMEVAL Task on Arabic Semantic Labeling*
- [06.2005] **Co-Chair** ACL Workshop on Computational Approaches to Semitic Languages

PROGRAM COMMITTEE MEMBER

Journals

Transactions on Computational Linguistics; Computational Linguistics; IEEE Speech Communication; Computers and Humanities; Computers, Speech and Language; Journal of Natural Language Engineering; Egyptian Journal of Language Engineering; Journal for Language Resources and Evaluation; ACM Transactions on Speech and Language Processing; Journal of Artificial Intelligence Research; Journal of Cognitive Science; ACM Transactions on Asian Language Information Processing (TALIP); Machine Translation

Conferences

Association for Computational Linguistics (ACL); Cognitive Science Society; Computational Linguistics (COLING); NEMLAR; Human Language technologies (HLT); North American Association for Computational Linguistics (NAACL); Empirical Methods in Natural Language Processing (EMNLP); Egyptian Conference on Language Engineering; European Association for Computational Linguistics. (EACL); IJCAI; IJCNLP; RANLP; LREC; MT Summit; AMTA; CICLING, ALTIC; AIRS

Workshops

MWE (ACL 11); Text Graphs (ACL 11); LAWS V. (ACL 11); LREC Workshop on HLT for Semitic Languages (LREC 10); Statistical Machine Translation. (ACL 08);TextGraphs. (COLING 08); Student Research Workshop (ACL 08); Arabic and Local Languages (LREC 08); Computational Approaches to Semitic Languages (ACL 07); SEMEVAL 07 (ACL 07); Computational Approaches to Arabic Script based Languages (2007); Student Research Workshop (HLT-NAACL 07); Cross Language knowledge Induction (EACL 06); Student Research Workshop (ACL 05); Parallel Text (ACL 05); Computational Approaches to Semitic Languages (ACL 05); Cross-Language Knowledge Induction Workshop (EUROLAN 05); Computational Approaches to Arabic Script based Languages (COLING 04)