

**Ayah Zirikly**  
[ayah.zirikly@gmail.com](mailto:ayah.zirikly@gmail.com)

## EDUCATION

- George Washington University** - Washington, DC 2012 – 2017  
*Ph.D.* - Computer Science  
Focus: Cross-lingual, Arabic/Dialectal Arabic, and Low Resource Languages Named Entity Recognition, Entity Linking, Named Entity Aliasing Detection, Entity clustering, Information Extraction  
Advisor: Dr. Mona Diab
- George Washington University** - Washington, DC 2009  
*Masters of Science* - Computer Science
- Damascus University** - Damascus, Syria 2006  
*Bachelors of Science* - Computer Science with focus on Artificial Intelligence
- 

## EXPERIENCE

### **NLP Post-Doctoral fellow**

**National Institutes of Health** - Bethesda, MD November 2017 -  
Research, develop, implement prototypes and apply methods to tackle some of the challenges faced by SSA's disability programs by tapping into information from medical records and other text documents. The project aims to identify list of entities and their attributes that play major role in determining disability with a special focus on mobility, self-care, and mental health domains. Additionally, we are exploring employing methods, such as transfer learning, that can help in low-resources domain in order to avert the need to create annotated data for every targeted domain.

### **Research Assistant**

**The George Washington University (GWU)** - Washington, DC Jan 2013 – Oct 2017  
Research Assistant and developer with focus on Arabic and low resource languages (e.g. Dialectal Arabic, Uyghur) in the domains of Named Entity Recognition, Entity Linking and Clustering and Alias detection in Arabic and English. In addition to transfer learning techniques from high to low resource languages.

### **Additional projects:**

- Computational Sociology:
  - CLPsych ReachOut shared task (ranked 2<sup>nd</sup> & 3<sup>rd</sup> in CLPsych2017), in collaboration with Dr. Philip Resnik: identify crisis posts that require immediate attention from ReachOut administrators and moderators
  - The effect of syntactic packaging on sympathy measurement, in collaboration with Dr. Philip Resnik
- Similar languages and dialects identification
- Dialectal Arabic NER using deep sequence models and investigating new types of embeddings for Arabic

Ayah Zirikly  
[ayahzirikly.wordpress.com](http://ayahzirikly.wordpress.com)

**Ayah Zirikly**  
[ayah.zirikly@gmail.com](mailto:ayah.zirikly@gmail.com)

**Research Scientist** -intern-

**Rakuten Institute of Technology (RIT) - New York NY**

May 2014 – Aug 2014

Research scientist intern in the area of cross-lingual of transfer Named Entity Recognition without the use of parallel data in E-commerce

Supervisors: Masato Hagiwara and Satoshi Sekine

**Graduate Teaching Assistant**

**Department of Computer Science in the School of Engineering and Applied Sciences**

**The George Washington University (GWU) - Washington, DC**

Sep 2008 - Jun 2013

Served as teaching assistant for Statistical Natural Language Processing, Algorithms and Data Structures, Java Software Development, and Discrete Mathematics

**Research Assistant**

**Columbia University - New York, NY**

May 2012 - Aug 2012

Research under the supervision of Dr. Mona Diab in the area of Named Entity Aliasing Recognition, Disambiguation and Transliteration in Arabic and English using Unsupervised Machine Learning methodologies and Latent Dirichlet Allocation (LDA)

**Consultant**

**Clark and Brody - Alexandria, VA**

Jun 2011 – Aug 2011

Performed research into the early history and development of certain language features of the HTML language.

**Programmer/Researcher**

**The Food and Drug Administration (FDA) - Silver Spring, MD**

Jun 2008 – Dec 2008

Programmer/Researcher in the area of enhancing medical display information systems to increase the cancer diagnosis accuracy. This research included the application of multiple algorithms to provide better contrast in the display of cancer and abnormality detection screening. The system was tested by clinical practitioners that provided feedback for creation and expansion of the project scope and complexity

---

## PUBLICATIONS & PATENTS

### Publications & Presentations:

- [1] Denis Newman-Griffis and [Ayah Zirikly](#). Embedding Transfer for Low-Resource Medical Named Entity Recognition: A Case Study on Patient Mobility. In the Proceedings of the Association for Computational Linguistics (ACL) Workshop on BioNLP, Melbourne, Australia, 2018
- [2] [Ayah Zirikly](#) et al. Mobility Dictionaries: A Step Towards a Standardized Ontology for Functioning Information. Poster in the Mobilizing Computable Biomedical Knowledge, Bethesda, Maryland, 2018
- [3] Han-Chin Shing, Suraj Nair, [Ayah Zirikly](#), Meir Friedenberg, Hal Daumé III, Philip Resnik. Expert, Crowdsourced, and Machine Assessment of Suicide Risk via Online Postings. In the Proceedings of

Ayah Zirikly

[ayahzirikly.wordpress.com](http://ayahzirikly.wordpress.com)

## Ayah Zirikly

[ayah.zirikly@gmail.com](mailto:ayah.zirikly@gmail.com)

North American Chapter of the Association for Computational Linguistics (NAACL) Workshop on Computational Linguistics and Clinical Psychology (CLPsych), New Orleans, Louisiana, 2018

- [4] Sean MacAvaney, Bart Desmet, Arman Cohan, Luca Soldaini, Andrew Yates, [Ayah Zirikly](#), Nazli Goharian. RSDD-Time: Temporal Annotation of Self-Reported Mental Health Diagnoses. In the Proceedings of North American Chapter of the Association for Computational Linguistics (NAACL) Workshop on Computational Linguistics and Clinical Psychology (CLPsych), New Orleans, Louisiana, 2018
- [5] [Ayah Zirikly](#). Detecting Mobility Functions in Free Text as an Indicator of Disability. Deep Learning in Health Care Summit. Boston, Massachusetts, 2018 [presentation]
- [6] [Ayah Zirikly](#). From Preprocessing to Named Entity Recognition, Linking and Clustering in Multilingual, Cross-Lingual, High-Low Resources Settings [dissertation]
- [7] [Ayah Zirikly](#), Bart Desmet and Mona Diab. Dialects and Similar Languages Detection: Artificial Neural Networks vs. Conventional Supervised Classifiers. Oral presentation in Mid Atlantic Student Colloquium (MASC), Washington DC, 2017
- [8] [Ayah Zirikly](#), Bart Desmet and Mona Diab. The GW/LT3 VarDial 2016 Shared Task System for Dialects and Similar Languages Detection. In the Proceedings of Conference on Computational Linguistics (COLING) Workshop on NLP for Similar Languages, Varieties and Dialects, Osaka, Japan, 2016 [Ranked 1st in two subtracks]
- [9] Mohammed Attia, [Ayah Zirikly](#) and Mona Diab. The Power of Language Music: Arabic Lemmatization through Pattern. In the Proceedings of Conference on Computational Linguistics (COLING) Workshop on Cognitive Aspects of the Lexicon, Osaka, Japan, 2016
- [10] [Ayah Zirikly](#), Varun Kumar and Philip Resnik. GW/UMD CLPsych 2016 Shared Task System. In the Proceedings of North American Chapter of the Association for Computational Linguistics (NAACL) Workshop on Computational Linguistics and Clinical Psychology, San Diego, California, 2016
- [11] [Ayah Zirikly](#), Yassine Benajiba, and Mona Diab. GWU English TAC-KBP EL Diagnostic Task with Name Mention. In the Proceedings of Text Analysis Conference (TAC), Gaithersburg, MD, 2015
- [12] [Ayah Zirikly](#) and Masato Hagiwara. Cross-Lingual Transfer of Named Entity Recognizers without Parallel Corpora. In Proceeding of the Association for Computational Linguistics (ACL), Beijing, China, 2015
- [13] [Ayah Zirikly](#) and Mona Diab. Named Entity Recognition for Arabic Social Media. In the Proceedings of North American Chapter of the Association for Computational Linguistics (NAACL) Workshop on Vector Space Modeling for NLP, Denver, Colorado, 2015
- [14] [Ayah Zirikly](#) and Mona Diab. Named Entity Recognition for Dialectal Arabic. In the Proceedings of the Empirical Methods of Natural Language Processing (EMNLP) Workshop on Arabic Natural Language Processing, Doha, Qatar, 2014
- [15] [Ayah Zirikly](#) and Mona Diab, Named Entity Aliasing Resolution, SEAS R&D Showcase, The George Washington University, 2014
- [16] [Ayah Zirikly](#) and Mona Diab. ANEAR: Automatic Named Entity Aliasing Resolution. In the Proceedings of the 18th International Conference on Application of Natural Language to Information Systems (NLDB), University of Salford, UK, 2013

Ayah Zirikly

[ayahzirikly.wordpress.com](http://ayahzirikly.wordpress.com)

## Ayah Zirikly

[ayah.zirikly@gmail.com](mailto:ayah.zirikly@gmail.com)

- [17] [Ayah Zirikly](#) and Ammar Joukhadar, Software Quality Assurance, Faculty of Information Technology, Damascus University, July 2006, Volume: 1
- [18] [Ayah Zirikly](#) and Oumayma Al-Dakkak, Aircraft Movement Modeling and Simulation in 3D Environment, Faculty of Information Technology, Damascus University, July 2005, Volume: 1
- [19] [Ayah Zirikly](#) and Firas Saegh, TAPI Based IVR System for Banking Services, Faculty of Information Technology, Damascus University, July 2004, Volume: 1

### Patents:

Computer System and Method for Performing Multilingual Named Entity Recognition Transfer  
U.S. Provisional Application No. 62/288,109 [Pending]  
Japan, granted application JP6077727

### Tools:

MADAMIRA: Toolset for Morphological Analysis, Disambiguation, Base Phrase Chunking and Named Entity Recognition in Arabic [Responsible for Named Entity Recognition component in the tool]

---

## Committee Service

### Conferences

- Program committee member for Conference on Empirical Methods in Natural Language Processing (EMNLP), 2018
- Program committee member for the Association for Computational Linguistics (ACL), 2018
- Program committee member for the Computational Linguistics and Clinical Psychology (CLPsych) colloated with NAACL, 2018
- Program committee member for the Open-Source Arabic Corpora and Processing Tools workshop co-located with the International Conference on Language Resources and Evaluation (LREC), 2018
- Program committee member for the International Conference on Computer Systems and Applications AICCSA 2017
- Program committee member for the Association for Computational Linguistics (ACL), 2017
- Reviewer for the Third Arabic Natural Language Processing Workshop (WANLP), co-located with the European Chapter of the Association for Computational Linguistics (EACL), 2017
- Program committee member for the Second Annual Mid-Atlantic Student Colloquium (MASC) on Speech, Language and Learning, 2012

### Journals

- Reviewer for the journal of Natural Language Engineering, 2018
- Reviewer for the Physical Therapy journal, 2018
- Program committee member for the Traitement Automatique des Langues (TAL), 2017
- Reviewer for the Information Processing and Management journal
- Reviewer for the Transactions on Asian and Low-Resource Language Information Processing (TALLIP) journal

Ayah Zirikly

[ayahzirikly.wordpress.com](http://ayahzirikly.wordpress.com)

**Ayah Zirikly**  
[ayah.zirikly@gmail.com](mailto:ayah.zirikly@gmail.com)

---

## **SPECIALTIES**

### **Technical:**

Languages - Java, Python, CLIPS, Fuzzy CLIPS, LaTeX, XML, HTML, Git, Apache Jena and Fuseki, basic knowledge in MongoDB, TensorFlow, beginner knowledge in R, Keras and PyTorch.

Scientific Applications - MATLAB, Octave

Operating Systems - Linux, Windows, Mac OS

Testing tools - Traceview, Emma, JUnit, Android JUnit

Applications & Tools - Weka, Stanford CoreNLP, CRFSuite, Mallet, Dreamweaver, Microsoft Office

Productivity tools

### **Certificates:**

2008 - Computer Security and Information Assurance - Graduate Certificate

### **Awards:**

2011 - The George Washington University Chapter of the Association of Computing Machinery (ACM) Graduate Teaching Assistant of the year Award

### **Languages:**

Arabic: Native Proficiency

English: Full professional proficiency

French: Limited working proficiency, earned Diplome d'Etude en Langue Française (DELF)

Dutch: Beginner level