

Mossaab Bagdouri

Résumé

October 2017

Address: Walmart Labs
860 W California Ave
Sunnyvale, CA 94086, USA
Email: mossaab.bagdouri@walmartlabs.com
Phone: +1 240 281 3671

Research Interests

Information Retrieval, Text Classification, Natural Language Processing.

Education

2017 **Ph.D. in Computer Science** University of Maryland at College Park
2013 Master's in Computer Science University of Maryland at College Park
2011 Master's in Computer Science University of Colorado Boulder
2007 Ingénieur d'Etat (CS) ENSIAS - Rabat, Morocco
2004 Undergraduate Degree Classes Préparatoire aux Grandes Ecoles d'Ingénieurs - Tangier, Morocco

Professional Experiences

2017 - Present **Data Scientist** @WalmartLabs, Sunnyvale, CA
2011 - 2017 **Research Assistant** CLIP Lab, UMIACS, University of Maryland, College Park

- Thesis topic: Cross-platform question answering in social networking services.
- Optimal allocation of annotated data for training and testing certified text classifiers.
- Finding professionals in microblogging services and analyzing their user behavior.
- Use of CLIR to answer non-factoid questions from multilingual social media.

2016 (4 months) **Internship** IBM TJ Watson Research Center

- Deep neural networks for ranking user-generated answers to non-factoid questions. Tools: Keras.

2007 - 2009 **Project Engineer** Maroc Telecom, Morocco

- Project management of offers and promotions for enterprise products.
- Integration of the Internet and Landline Information Systems (BSCS).

2007 (5 months) **Internship** SIS-Consultants, Morocco

- Design and implementation of an E-gov web site for government tendering. Tools: JEE, JSF, MySQL.

Selected Project

- **A multilingual news aggregator for the Arab world.** The applications (Maghress.com and 6 others) crawl hundreds of online newspapers to get the latest news, classify and rank them before presenting the top stories on the main page (**120K+ daily unique visitors**). Tools: Java, Lucene, JSF, MySQL.

Relevant Coursework

Information Retrieval Systems. Data-Intensive Computing with MapReduce.
Natural Language Processing. CL Seminar in Bayesian Nonparametrics.
Machine Learning. Probabilistic Models of Human and Machine Intelligence.

Fellowships and Awards

2016 - 2017 **IBM Ph.D. Fellowship.** 2013 - 2017 **SIGIR Student Travel Grant x 4.**
2016 - 2017 **Future Faculty Fellowship, UMD.** 2013 - 2016 **Goldhaber Travel Grant, UMD x 2.**
2009 - 2011 **Fulbright Foreign Student Scholarship.** 2011 - 2012 **Dean's Fellowship, UMD.**

Publications

Papers in Refereed Conference Proceedings

1. **Bagdouri, Mossaab** and D. W. Oard (2017). Building bridges across social platforms: answering Twitter questions with Yahoo! Answers. In: *SIGIR*. Tokyo, Japan, pp.1181–1184.
2. Gao, N., **Mossaab Bagdouri**, and D. W. Oard (2016). Pearson rank: A head-weighted gap-sensitive score-based correlation coefficient. In: *SIGIR*. Pisa, Italy, pp.941–944.
3. **Bagdouri, Mossaab** (2016). Journalists and Twitter: A multidimensional quantitative description of usage patterns. In: *ICWSM*. Cologne, Germany, pp.22–31.
4. **Bagdouri, Mossaab** and D. W. Oard (2015). On predicting deletions of microblog posts. In: *CIKM*, pp.1707–1710.
5. **Bagdouri, Mossaab** and D. W. Oard (2015). Profession-based person search in microblogs: Using seed sets to find journalists. In: *CIKM*. Melbourne, Australia, pp.593–602.
6. **Bagdouri, Mossaab**, D. W. Oard, and V. Castelli (2014). CLIR for informal content in Arabic forum posts. In: *CIKM*. Shanghai, China, pp.1811–1814.
7. **Bagdouri, Mossaab**, W. Webber, D. D. Lewis, and D. W. Oard (2013). Towards minimizing the annotation cost of certified text classification. In: *CIKM*. San Francisco, CA, USA, pp.933–936.
8. Webber, W., **Mossaab Bagdouri**, D. Lewis, and D. Oard (2013). Sequential testing in classifier evaluation yields biased estimates of effectiveness. In: *SIGIR*. Dublin, Ireland, pp.989–998.
9. Mark, G., **Mossaab Bagdouri**, L. Palen, J. Martin, B. Al-Ani, and K. Anderson (2012). Blogs as a collective war diary. In: *CSCW*. Seattle, WA, USA, pp.37–46.
10. Sarcevic, A., L. Palen, J. White, K. Starbird, **Mossaab Bagdouri**, and K. Anderson (2012). “Beacons of hope” in decentralized coordination: Learning from on-the-ground medical twitterers during the 2010 Haiti earthquake. In: *CSCW*. Seattle, WA, USA, pp.47–56.

Workshop Papers and Technical Reports

1. **Bagdouri, Mossaab** and D. W. Oard (2016). CLIP at TREC 2016: LiveQA and RTS. In: *TREC*. Gaithersburg, MD.
2. Hasanain, M., **Mossaab Bagdouri**, T. Elsayed, and D. W. Oard (2016). What questions do journalists ask on Twitter? In: *ICWSM SMnews*. Cologne, Germany, pp.127–134.
3. **Bagdouri, Mossaab** and D. W. Oard (2015). CLIP at TREC 2015: Microblog and LiveQA. In: *TREC*.
4. **Bagdouri, Mossaab** (2015). Cross-platform question routing for better question answering. In: *SIGIR*, pp.1053.
5. Bergsma, S., P. McNamee, **Mossaab Bagdouri**, C. Fink, and T. Wilson (2012). Language identification for creating language-specific Twitter collections. In: *LSM*. Montreal, Canada, pp.65–74.

Book Chapters

1. Starbird, K., L. Palen, S. B. Liu, S. Vieweg, A. Hughes, A. Schram, K. M. Anderson, **Mossaab Bagdouri**, J. White, C. McTaggart, and C. Schenk (2011). *Promoting structured data in citizen communications during disaster response: An account of strategies for diffusion of the ‘Tweak the Tweet Syntax’*. Chandos.

Technical Skills

Data Analysis : Spark, Hadoop, SPSS.
Databases: MySQL, Oracle.

Java EE Architecture: Java, JSF, PrimeFaces.
Programming languages: Python, PHP.

Service

- 2013-2016 **Reviewer** for journals (CSUR, TWEB, FnTIR) and conferences (SIGIR, CHI, ICWSM).
 2015 **Co-chair** of the 2015 Mid-Atlantic Student Colloquium on Speech, Language and Learning.
 2014 **Student volunteer** at the 52nd Annual Meeting of the Association for Computational Linguistics.

Languages

- Arabic, English and French.