

# Curriculum Vitae

## Kyle Williams

University Park, PA.

kwilliams@psu.edu

<http://www.personal.psu.edu/kiw5209/>

Extended version of CV: [http://www.personal.psu.edu/kiw5209/kwilliams\\_cv.pdf](http://www.personal.psu.edu/kiw5209/kwilliams_cv.pdf)

Updated: 16 April 2016

### Personal Profile

I am Ph.D. candidate at The Pennsylvania State University advised by Dr. C. Lee Giles. Since beginning my Ph.D. I've interned at Microsoft and Oracle Labs; achieved the highest  $F_1$  score among all participants in the source retrieval task for plagiarism detection at PAN 2013 and PAN 2014; placed 2nd and 3rd at the Penn State Graduate Exhibition in successive years; and presented my work at several conferences.

**Research Areas:** Information retrieval, applied machine learning, digital libraries

### Education

- **The Pennsylvania State University (2012-present)**  
Ph.D. in Information Sciences and Technology, advised by Prof. Lee Giles.
- **University of Cape Town (2010-2012)**  
Masters in Computer Science by Dissertation, advised by A/Prof. Hussein Suleman.  
*Degree awarded with distinction*
- **University of Cape Town (2006-2009)**  
Bachelor of Business Science in Management Studies in the field of Computer Science.  
*Degree awarded with second class division two honours*

### Relevant Employment History

- **Microsoft, Bellevue, WA (May 2015 - August 2015)**  
**Position:** Research Intern  
**Research Project:** Identifying signals for detecting satisfaction in mobile search.
- **Oracle Labs, Burlington, MA (June 2014 - August 2014)**  
**Position:** Research Intern  
**Research Project:** Relevance ranking and result diversification for eCommerce search.
- **College of Information Sciences and Technology, Pennsylvania State University (August 2012 - Present)**  
**Position:** Graduate Research Assistant and Teaching Assistant  
**Responsibilities:** Research under the guidance of Prof. Lee Giles .
- **Digital Libraries Laboratory, Department of Computer Science, University of Cape Town (December 2009; May 2012 - August 2012)**  
**Position:** Research Assistant  
**Responsibilities:** Development, writing of articles, assisting honours students, assisting with day-to-day operations in the lab.

- **Department of Computer Science, University of Cape Town (February 2010 - June 2011)**

**Position:** Teaching assistant for CSC1015F, CSC2001F and CSC2020S

**Responsibilities:** Setting test-cases for practical tests, supervising marking, managing class tutors, handling student queries and other administrative aspects of the course.

## Research, Experience and Skills

- **Selected Recent Projects**

- 2012-Present: SimSeerX is a similarity search engine where documents are submitted as queries and similar documents are retrieved based on arbitrary similarity functions.
- 2012-Present: Source Retrieval for Plagiarism Detection  
This project involves investigating techniques for retrieving potential sources of plagiarism for a given suspicious document. The techniques developed as part of this project have gone on to achieve the highest F1 scores in the 2013 and 2014 PAN Source Retrieval Task.

- **Professional Activities.** *Reviewer:* IEEE Computer, South African Computer Journal (SACJ), iConference; *Sub-reviewer:* WWW, TPDL, ASONAM, DL, DSAA, JCDL. Judge in Penn State 2014 and 2015 Undergraduate Exhibitions
- **Teaching:** Teaching assistant for 1st, 2nd and 4th year University courses; 5 years of extra-curricular music teaching at high school level,
- **Programming Languages.** *High Experience:* Java, Python; *Medium Experience:* PHP, JavaScript (incl. jQuery), Perl, C++; *Low Experience:* XML + XSLT, SQL, Scala
- **Core Technologies:** Solr, Lucene, scikit-learn. Experience with Hadoop and Pig

## Selected Publications

For a full list of publications see [http://www.personal.psu.edu/kiw5209/kwilliams\\_pubs.pdf](http://www.personal.psu.edu/kiw5209/kwilliams_pubs.pdf)

- [1] Kyle Williams, Julia Kiseleva, Aidan C. Crook, Imed Zitouni, Ahmed Hassan Awadallah, Madian Khabsa. 2016. Detecting Good Abandonment in Mobile Search. To appear in: *Proceedings of the 2016 International World Wide Web Conference (WWW '16)*.
- [2] Kyle Williams, Julia Kiseleva, Aidan C. Crook, Imed Zitouni, Ahmed Hassan Awadallah, Madian Khabsa. 2016. Is This Your Final Answer? Evaluating the Effect of Answers on Good Abandonment in Mobile Search. To appear in: *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '16)*.
- [3] Julia Kiseleva, Kyle Williams, Jiepu Jiang, Ahmed Hassan Awadallah, Imed Zitouni, Aidan C. Crook, Tasos Anastasakos. 2016. Predicting User Satisfaction with Intelligent Assistants. To appear in: *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '16)*.
- [4] Jian Wu, Kyle Williams, Hung-Hsuan Chen, Madian Khabsa, Cornelia Caragea, Suppawong Taurob, Alexander Ororbia, Douglas Jordan, Prasenjit Mitra, C. Lee Giles. 2015. CiteSeerX: AI in a Digital Library Search Engin. In: *Artificial Intelligence Magazine (AI Magazine) 36(3)*, pages 35-48.
- [5] Kyle Williams, Jian Wu, C. Lee Giles. 2014. SimSeerX: A Similar Document Search Engine. In: *Proceedings of the 2014 ACM Symposium on Document Engineering (DocEng '14)*, pages 143-146.