

# Denis Savenkov

Research Scientist  
Facebook Inc.

Phone: (404) 988-8472  
Email: denissavenkov@gmail.com  
Homepage: <http://denxx.github.io/>

## Research Interests

natural language understanding, dialog, natural language processing, information retrieval, machine learning

## Professional Experience

- **Research Scientist** - Facebook May 2017 - current days
  - Working on Natural Language Understanding in the Messenger team
- **Software Engineering Intern** - Google May - August 2013, 2014, 2015
  - 2015: Developed a component for coreference resolution using web term and entity statistics
  - 2014: Implemented a method for information extraction from question-answer pairs
  - 2013: Simplified a part of entity linking pipeline, improved performance of the internal entity linking evaluation pipeline
- **Research Intern** - Microsoft Research/Bing May - August 2012
  - During my internship I was working on reranking algorithms for diversifying search results using user behavior data
- **Software developer, team leader** - Yandex, Moscow, Russia Sep 2008 - Aug 2011
  - During internship followed by a job offer I was responsible for applying machine learning to the problem of query-biased document summarisation. This included the whole process from data collection and processing, testing different hypotheses and implementing all these code in production system.
  - From September 2010 the leader of a small research team working on snippets.

## Education

- **Ph.D.** in Computer Science, **Emory University**, Atlanta, GA, USA 2011-2017
  - Focused on information retrieval, natural language processing and machine learning.
  - Dissertation: "Question Answering with User Generated Content"
- **Yandex school of data analysis**, Moscow Institute of Physics and Technology, Russia 2008 - 2010
  - Two-year MSc program on data analysis, machine learning, algorithms, etc.
- **M.S.** in Computer Science, **Tula state university**, Tula, Russia 2005-2007
  - Thesis title: "Combining acyclic Markov models in the problem of data analysis in arrays of interrelated objects. Two best-paper awards on local conferences"
- **B.S.** in Computer Science, **Tula state university**, Tula, Russia 2001-2005
  - Thesis title: "Software for representing adjacency relations of array elements as a tree"

## Selected publications

- D. Savenkov, E. Agichtein "EviNets: Neural Networks for Combining Evidence Signals for Factoid Question Answering", *ACL 2017*
- A. Vtyurina, D. Savenkov, E. Agichtein and C. Clarke "Exploring Conversational Search With Humans, Search Engines, and Wizards", *CHI 2017*
- P. Braslavski, D. Savenkov, E. Agichtein and A. Dubatovka "What Do You Mean Exactly? Analyzing Clarification Questions in CQA", *CHIIR 2017*
- D.Savenkov and E.Agichtein "CRQA: Crowd-powered Real-time Automated Question Answering System", *HCOMP 2016*
- D.Savenkov and E.Agichtein "When a Knowledge Base is not Enough: Question Answering over Knowledge Bases with External Text Data", *SIGIR 2016*

- D.Savenkov “Ranking Answers and Web Passages for Non-factoid Question Answering: Emory IR Lab @ TREC 2015 LiveQA”, *TREC 2015*
- D.Savenkov, Wei-Lwun Lu, J. Dalton and E.Agichtein “Relation Extraction from Community Generated Question-Answer Pairs”, *NAACL 2015 Student Research Workshop*
- D.Savenkov, E.Agichtein “To Hint or Not: Exploring the Effectiveness of Search Hints for Complex Informational Tasks”, *SIGIR 2014*
- D.Savenkov, D.Lagun, and Q.Liu “Search Engine Switching Detection Based on User Personal Preferences and Behavior Patterns”, *SIGIR 2013*
- D.Pelleg, D.Savenkov, and E.Agichtein, “Touch Screens for Touchy Issues: Analysis of Accessing Sensitive Information from Mobile Devices”, *ICWSM 2013*
- Q. Guo, D. Lagun, D. Savenkov, and Q. Liu “Improving Relevance Prediction by Addressing Biases and Sparsity in Web Search Click Data”, *WSDM workshop on Web Search Click Data, WSCD 2012*
- Savenkov D., Braslavski P., Lebedev M. “Search Snippet Evaluation at Yandex: Lessons Learned and Future Directions”, *Multilingual and Multimodal Information Access Evaluation, 2011, Springer, pp.14-25*

## Technologies

- Programming Languages: C/C ++, Java, Python, Scala. Previous experience with C#, Delphi, Prolog, LISP.
- Big Data: MapReduce (Apache Hadoop), Apache Spark, MPI.
- Data analysis tools and libraries: tensorflow, pandas, scikit-learn, Rapid Miner, Weka, R.

## Teaching experience

- Emory University , Atlanta, USA 2013 - 2016  
Teaching assistant for CS534 - Machine Learning (Fall 2016)  
Instructor for CS170 - Introduction to computer science (Spring 2013, Fall 2013)  
Author and instructor for CS375 - Competitive programming (Spring 2013, Spring 2014)
- Coaching Emory programming competitions team 2011-2017
- Yandex school of Data Analysis (Moscow institute of Physics and Technology), Russia 2010 - 2011  
Teaching assistant; course: Machine learning
- Tula state university, Tula, Russia 2005 - 2007  
Teaching assistant for Functional and logical programming, Algorithms and data structures, System analysis and decision making

## Awards

- The Chris Schoettle Graduate Research Award, Emory University, 2017.
- 1st place on TREC LiveQA 2016 shared task (<https://sites.google.com/site/trecliveqa2016/>).
- 1st place on HackATL 2014 hackathon (in a team). Project: Scenter (geo-fence based messaging platform)
- 2nd place on Atlanta Big Data Analytics Team Challenge by Booze, Allen and Hamilton in 2014
- 1st place on the Yandex Switch Detection Challenge 2012 ([switchdetect.yandex.ru/en](http://switchdetect.yandex.ru/en))

## Other activities

- Conference Program Committee member: SIGIR 2017, CIKM 2017, WSDM 2018
- Competitive programming: TopCoder, Codeforces (username: DenXX)
- Data Science: Kaggle (username: DenXX)
- Ironman triathlete and runner (pacer at Tunnel Light Marathon 2017 and Seattle Half Marathon 2017)

Last updated: May 21, 2018

<http://www.mathcs.emory.edu/~dsavenk/>