

# Sabina Tomkins

## PERSONAL DATA

---

CURRENT: **PhD Candidate** Technology Management | UC Santa Cruz  
EXPECTED GRADUATION: **June 2018**  
PHONE: +1 (206) 390-3931  
EMAIL: [sabina.tomkins@gmail.com](mailto:sabina.tomkins@gmail.com)  
WEBSITE: [travellingscholar.com](http://travellingscholar.com)

## EDUCATION

---

MAY 2017 Proposed Thesis: *Probabilistic Methods for Data-Driven Social Good*  
**UC Santa Cruz**, Santa Cruz  
*Qualifying Exam Committee:*  
*Lise Getoor, Brent Haddad, John Musacchio, Rayid Ghani*

JANUARY 2013 Bachelor of Arts in COMPUTER SCIENCE,  
**New York University**, New York  
*magna cum laude*  
GPA: 3.8/4.0

JANUARY 2011 Associate of Science , Associate of Arts,  
**North Seattle Community College**, Seattle  
*President's List*  
GPA: 3.8/4.0

## CONFERENCE PUBLICATIONS

---

*LLD 2017* | *Detecting Cyber-bullying from Sparse Data and Inconsistent Labels*  
**Sabina Tomkins**, Lise Getoor, Yunfei Chen, Yi Zhang  
POSTER PRESENTATION  
**Neural Information Processing Systems (NIPS)** Workshop  
Learning with Limited Labeled Data

*IJCAI 2017* | *Disambiguating Energy Disaggregation: A Collective Probabilistic Approach*  
**Sabina Tomkins**, Jay Pujara, Lise Getoor  
ORAL & POSTER PRESENTATION  
**International Joint Conference on Artificial Intelligence**

*WiNLP 2017* | *An Unsupervised Method for Dialog Act Detection*  
**Sabina Tomkins**, Anbang Xu, Zhe Lui, Yufan Guo  
POSTER PRESENTATION  
**Association for Computational Linguistics (ACL)** Workshop  
Women in Natural Language Processing

*UrbComp 2016* | *A Probabilistic Disaggregation Framework*  
**Sabina Tomkins**, Lise Getoor  
ORAL PRESENTATION  
**Knowledge Discovery and Data Mining (KDD)** Workshop  
Urban Computing

*EDM 2016* | *Predicting Student Success: A High School MOOC Case Study*  
**Sabina Tomkins**, Arti Ramesh, Lise Getoor  
ORAL PRESENTATION  
**International Conference of Educational Data Mining**

*BuildSys 2015* | *Poster Abstract: Contextual Air Conditioning Disaggregation with Probabilistic Soft Logic*  
**Sabina Tomkins**, Lise Getoor  
POSTER PRESENTATION  
**Energy-Efficient Built Environments**

## WORK EXPERIENCE

---

<i>Current</i>	Graduate Student Researcher UC SANTA CRUZ, Santa Cruz <i>Probabilistic Methods for Data-Driven Social Good</i> I am currently working on a socio-linguistic framework for cyber-bullying detection.
	Research Scientist Intern AMAZON, Seattle <i>Machine Learning and Music Science</i> Conducted empirical analysis and evaluated numerous methods for sequencing music.
SUMMER 2016	Watson Research Scientist Intern IBM RESEARCH, Almaden <i>Natural Language Processing and User Interfaces</i> Proposed and implemented an unsupervised framework for detecting dialog acts in novel domains.
SUMMER 2014	Fellow DATA SCIENCE FOR SOCIAL GOOD UNIVERSITY OF CHICAGO, Chicago <i>Energy Analytics</i> I co-created an energy analytics website to provide users actionable insight about their energy consumption. I also co-created the open Python package Energy Disaggregation, with tools for preprocessing smart meter data, and performing disaggregation tasks with machine learning techniques. Under the supervision of <a href="#">Professor Varun Chandola</a>
JULY 2013 - AUGUST 2015	Instructor and Course Developer EDHESIVE, New York and Remote <i>Advanced Placement Java Course</i> Students worldwide can learn Java with our open online AP computer science course. I answered student questions on an online forum and developed course material such as coding activities.
SEPTEMBER 2012 - AUGUST 2013	Research Assistant COMPUTATION AND COGNITION LAB, NEW YORK UNIVERSITY, New York <i>Cognitive Science</i> Assisted with research on cognitive science topics, such as humans' ability to learn actively and to generalize. Under the supervision of <a href="#">Professor Todd Gureckis</a> and John McDonell.
SUMMER 2012	Student Researcher DEPT. OF COMPUTER SCIENCE, UNIVERSITY OF MONTANA, Bozeman <i>Particle Swarm Optimization</i> Implemented several novel algorithms for constrained particle swarm optimization. Under the supervision of <a href="#">Professor Hashem Nehrir</a> .

## FELLOWSHIPS AND GRANTS

---

JULY 2017	IJCAI Student Travel Grant
SEPTEMBER 2016	BUILDSys Student Travel Grant
JANUARY 2016	Dept. of Technology & Information Management Fellowship
SEPTEMBER 2015	Dept. of Technology & Information Management Fellowship
JULY 2015	Dept. of Technology & Information Management Fellowship
SEPTEMBER 2013	NSF Mathematics Research Grant
FEBRUARY 2013	NSF Mathematics Research Grant
SEPTEMBER 2012	NSF Mathematics Research Grant

## ACADEMIC SERVICE

---

PROGRAM COMMITTEE 2018	International Workshop on Non-Intrusive Load Monitoring
EXTERNAL REVIEWER 2017	International Conference on Educational Data Mining
EXTERNAL REVIEWER 2016	SIGKDD Conference on Knowledge Discovery and Data Mining
VOLUNTEER 2016 & 2017	Women in Machine Learning Workshop
VOLUNTEER 2015	Grace Hopper Celebration
VOLUNTEER 2014	SIGKDD Conference on Knowledge Discovery and Data Mining

## SELECTED CODING PROJECTS

---

### **Appliance Disambiguation** Software Pack for Energy Disaggregation

This package employs a collective probabilistic approach to disambiguate appliances in an aggregate energy signal.

[Appliance Disambiguation Code](#)

### **AmmiTips** First Place in PAKATHON, A HACKATHON FOR PAKISTAN

The goal of AmmiTips is to distribute WHO certified health information freely to new and expecting mothers in Pakistan in English and Urdu, and via text and voice messages

## DATA SCIENCE SANTA CRUZ EVENTS

---

*October 2015 - 2017*

Organized AN EVENING WITH THE INTERNS, UC Santa Cruz

*Data Science Event*

Organized and hosted this event for three years, with between 30-70 attendees each year. The center-piece of the evening is a panel of graduate students who discuss their summer internship experiences. Each year I find the panelists, write the discussion topics, and moderate the panel.

*October 2015*

Organized COMPUTATIONAL CREATIVITY, UC Santa Cruz

*Data Science Event*

Organized and hosted this event, with approximately 50 attendees. The purpose of the event was to investigate recent advances in machine learning (such as convolutional or *deep* networks) as applied to creativity. The event had several lectures from machine learning scientists, followed by a debate between professors of different backgrounds. I moderated the debate, which centered on the question of whether computers can be creative. [Video from the event](#)

Invited Speakers: Ahmed Elgammal, Prof., Rutgers University; Chris Smith, Adatao

Panelists: Arnav Jhala (Computational Media), Samantha Matherne (Philosophy), David Cope (Music), Albert Narath (History of Art and Visual Culture)