

## **Samuel H. Carton**

### **Curriculum Vitae**

**Contact:** University of Michigan School of Information  
105 S State Street  
Room 3349  
Ann Arbor, MI 48104  
[scarton@umich.edu](mailto:scarton@umich.edu)

### **Education:**

**Ph.D. Information**, 2012-2017 (expected)  
University of Michigan School of Information, Ann Arbor MI USA  
GPA 3.6/4.0  
Advisors Qiaozhu Mei, Paul Resnick

**B.S. Computer Science**, 2008-2012  
Northwestern University, Evanston IL USA  
GPA 3.3/4.0  
Major GPA 3.5/4.0

### **Research Interests:**

I am interested in using machine learning and natural language processing to answer questions in the social sciences and humanities by leveraging collective sources of written information. I am also interested in building interactive systems that allow ordinary people to explore those information sources through the lenses that ML and NLP can impose on them. My past/ongoing projects include creating arbitrary concept maps using Wikipedia, recommending scientific literature using survey articles, visualizing the spread of rumors across Twitter, and predicting police misconduct from event logs. Future projects I have planned include learning to recognize narrative tropes in plot summaries, a cross-domain study of online deception, and analyzing the HathiTrust literature collection to understand how societal attitudes toward traumatic events change as time passes.

### **Research Experience:**

**Research Assistant**, September 2012 – present  
University of Michigan School of Information  
Advisors: Professors Qiaozhu Mei and Paul Resnick  
Current projects:

- **Prediction of Narrative Tropes in Plot Summary Text**: Design a model of narrative trope appearance in plot summary text. Work is in preliminary stages.
- **Visualization of Rumors on Twitter**: Design and test a technique for visualizing the interplay of rumors and debunking corrections on Twitter in a way that overcomes problems of scale and data complexity.  
Demo: <http://rumorlens.org/demo/>
- **Improved Literature Reviews**: Improve support for literature reviews in academic search engines by analyzing the qualities of sets of papers cited by existing

literature reviews (e.g. diversity across time and topic, impact) and developing retrieval and ranking techniques to optimize these qualities.

Demo: <http://104.196.32.239:8080/SurveyUpdate/> (very slow research prototype)

**Research Assistant**, June 2010 – August 2012

**CollabLab, Northwestern University School of Communication**

Advisors: Dr. Doug Downey (Northwestern University), Dr. Brent Hecht (University of Minnesota)

**Research Intern**, June 2009 - Sep. 2009

**Laboratory of Neurobiology, National Institute for Neural Disorders and Stroke (NINDS), NIH**

Advisor: Dr. Xiaobing Chen

### **Work Experience:**

**Fellow**, June 2015 – August 2015

**University of Chicago Data Science for Social Good Fellowship**

Predicting Adverse Police-Public Interactions: Use Charlotte-Mecklenburg Police Department data to predict adverse interactions between officers and the public.

Overview: <http://dssg.uchicago.edu/project/early-intervention-system-for-adverse-police-interactions/>

**Programmer**, Jan. 2010 – June 2010

ArticuLab, Northwestern University School of Communication

### **Teaching Experience**

**Teaching Assistant**, Sep. 2014 – December 2015

School of Information, University of Michigan

SI 601/618 Exploratory Data Analysis – Dr. Kevyn Collins-Thompson

SI 301 Models of Social Information Processing – Dr. Qiaozhu Mei

SI 106 Introduction to Python Programming – Dr. Paul Resnick

### **Publications:**

- 2015 **Carton, S.**, Park, S., Zeffer, N., Adar, E., Mei, Q., Resnick, P. RumorLens: Visualizing the Audience of Competing Memes in Social Media. *Eighth International AAAI Conference on Weblogs and Social Media (2015)*.
- 2014 Resnick, P., **Carton, S.**, Park, S., Shen, Y., Zeffer, N. RumorLens: A System for Analyzing the Impact of Rumors and Corrections in Social Media. *3<sup>rd</sup> Computation + Journalism Symposium (2014)*. (short paper)
- 2012 Hecht, B., Quaderi, M., **Carton, S.**, Schöning, J., Raubal, M., Gergle, D. and Downey, D. 2012. Explanatory Semantic Relatedness and Explicit Spatialization for Exploratory Search. *SIGIR '12: 35th International conference on Research and Development in Information Retrieval (2012)*.
- 2012 Bao, P., Hecht, B., **Carton, S.**, Quaderi, M., Horn, M. and Gergle, D. 2012. Omnipedia: Bridging the Wikipedia Language Gap. *CHI '12: 30th International Conference on Human Factors in Computing Systems (2012)*.
- \*Nominated for Best Paper Award.**

### **Other work:**

2011 **Carton, S.**, Park, K., Olalde, M., Lieberman, D.  
HamTracker: Automated tracking of government earmark spending  
<http://projects.knightlab.com/projects/hamtracker>

### **Honors and Awards:**

2012 Rackham Regents' Fellowship

### **Skills:**

- **Programming:** Java, MATLAB, C, C++, PHP, Javascript, R, Python, Pig
- **Languages:** English (native), Spanish (proficient)
- **Tools:** Hadoop, Eclipse, JetBrains, Git, Subversion, MySQL, MongoDB, Google Web Toolkit
- **Miscellaneous:** Familiar with development in Unix/Linux, Windows, MacOS

### **Graduate Coursework:**

Probabilistic Graphical Models (2011), Information Retrieval (2011), Kolmogorov Complexity (2011), Data Mining (2012), Research Methods for Social Science (2012), Corpus Linguistics (2013), Computational Models of Cognition (2013), Natural Language Processing (2013), Linear Models (2013), Network Theory (2013), Machine Learning (2014), Analysis of Multivariate Data (2014).

### **Selected Undergraduate Coursework:**

Data Structures (2009), Networking (2010), Machine Learning (2010), Computer Graphics (2010), Design & Analysis of Algorithms (2011), Theory of Computation (2011), Cryptography (2011), Natural Language Processing (2012), Linguistics: Syntax (2012)