# Julian McAuley

Stanford University 353 Serra Mall Stanford, CA 94305 U.S.A.

Phone: 650 521 3166

email: jmcauley@cs.stanford.edu URL: http://i.stanford.edu/~julian/

## Current position

*Postdoctoral Scholar*, Computer Science Department, Stanford University Advised by Jure Leskovec

## Areas of specialization

Machine Learning • Data Mining • Social Networks • Graphical Models

### Education

2008-2011	РнD, Australian National University
	Advised by Tibério Caetano
	Thesis: Graphical Models for Inference and Learning in Computer Vision

2003-2007 BENG, BSc, University of New South Wales Software Engineering and Pure Mathematics (first-class honours and the University Medal)

### Professional experience

- 2010 INTERN, Google (3 months) I worked as an intern on the Google's *docs* platform.
- INTERN, Xerox Research Centre Europe (4 months)
   I worked in the *Textual and Visual Pattern Analysis* group under Teofilo de Campos. We studied applications of machine learning to image classification and segmentation, leading to two patents, and publications in *BMVC* and *CVPR*.
- 2006 RESEARCHER, NICTA (2 years, part time) I worked in NICTA's *Statistical Machine Learning* group, under Tibério Caetano and Alex Smola. Primarily, I developed open-source software for inference and learning in graphical models, and also produced several publications during this time.
- INTERN, Australian National University (3 months)
   I worked on higher-order color models of natural images, leading to state-of-the-art performance on image denoising problems. This led to my first publication, in *ICML*.

### Honors & awards

2008	UNSW University Medal for Software Engineering
2007	UNSW CSE Performance Award (1st place)
2006	Canon Information Systems Research Australia Prize
2006	Skillsearch Computing Prize

2004-2007 UNSW Dean's Award

## Career-best publications

McAuley, J. and Leskovec, J. (2013), "Hidden factors and hidden topics: understanding rating dimensions with review text," *RecSys*Showed how the text of users' reviews can be used to discover their preferences, identify skilled reviewers, and automatically categorize products. We analyzed over 40 million reviews, including the complete *Amazon* corpus, which to our knowledge is the largest-scale study of online reviews to date.
McAuley, J. and Leskovec, J. (2013), "From amateurs to connoisseurs: modeling the evolution of

user expertise through online reviews," *WWW* Studied how users gain expertise as they rate and review products. We studied over ten million reviews where users have "acquired tastes," including beers, wines, gourmet foods, and movies. Featured on Business Insider, CBS Chicago, and The Beer Street Journal.

- Lakkaraju, H. and McAuley, J. and Leskovec, J. (2013), "What's in a name? Understanding the interplay between titles, content, and communities in social media," *ICWSM* One of the first articles to study *reddit.com*, we studied the factors that determine the success or failure of social media submissions. This article received substantial media attention, from Business Insider, New Scientist, Time, Forbes, Phys.org, Gizmodo, and many others.
- 2013 McAuley, J. and Leskovec, J. (2013), "Discovering social circles in ego networks," *TKDD*
- 2012 McAuley, J. and Leskovec, J. (2012), "Learning to discover social circles in ego networks," *NIPS* These two papers studied *social circles* on *Facebook*, *Google+*, and *Twitter*. We examined how people organize and categorize their friends, and developed algorithms to help users perform this task automatically. Following this, we are collaborating with *Kaggle* to obtain thousands of hand-labeled social circles, which we hope will form the basis of an online competition. Featured on Wired and the MIT Technology Review.
- 2012 McAuley, J. and Jurafsky, J. and Leskovec, J. (2012), "Learning attitudes and attributes from multiaspect reviews," *ICDM* We studied the *aspects* of users' ratings, and developed algorithms to determine which parts of reviews relate to different aspects of an opinion. For example, we learned how beer reviews can be separated into language about the feel, look, smell, taste, and overall impression of the beer, and developed sentiment lexicons for each of these aspects. This helped us to better model users' opinions, to recommend better products, and to summarize users' reviews.
- 2011 McAuley, J. and Caetano, T. (2011), "Faster algorithms for max-product message-passing," *JMLR*
- 2011 Felzenszwalb, P. and McAuley, J. (2011), "Fast inference with min-sum matrix product," PAMI
- 2010 McAuley, J. and Caetano, T. (2010), "Exploiting data-independence for fast belief-propagation," ICML

Here, we showed that message passing algorithms can be made faster in graphical models that exhibit a certain type of *factorization*, which we showed to be ubiquitous in a variety of applications ranging from computer vision, to protein design, and natural language processing. This series of papers would eventually form the core of my graduate thesis.

2009 Caetano, T. and McAuley, J. and Cheng, L. and Le, Q. and Smola, A. (2009), "Learning graph matching," *PAMI*  We showed how *structured learning* can be applied to matching problems in computer vision, by using a small amount of human-labeled data to tune algorithms to a specific matching setting. We showed that with learning, even simple, scalable algorithms can outperform highly sophisticated and computationally costly alternatives.

### Publications (as first or second author)

JOURNAL PAPERS

- 2013 McAuley, J. and Leskovec, J. (2013), "Discovering social circles in ego networks," *TKDD*
- 2012 McAuley, J. and Ramisa, A. and Caetano, T. (2012) "Optimization of robust loss functions for weakly-labeled image taxonomies," *IJCV*
- 2012 McAuley, J. and Caetano, T. (2012) "Fast matching of large point sets under occlusions," *Pattern Recognition*
- 2011 McAuley, J. and Caetano, T. (2011), "Faster algorithms for max-product message-passing," *JMLR*
- 2011 Felzenszwalb, P. and McAuley, J. (2011), "Fast inference with min-sum matrix product," PAMI
- 2009 Caetano, T. and McAuley, J. and Cheng, L. and Le, Q. and Smola, A. (2009), "Learning graph matching," *PAMI*
- 2009 Caetano, T. and McAuley, J. (2009), "Faster graphical models for point-pattern matching," *Spatial Vision*
- 2008 McAuley, J. and Caetano, T. and Barbosa, M. (2008), "Graph rigidity, cyclic belief propagation and point pattern matching," *PAMI*
- 2006 McAuley, J. and Caetano, T. and da Fontoura Costa, L. (2007), "The rich-club phenomenon across complex network hierarchies," *Applied Physics Letters*

**CONFERENCE PAPERS** 

- 2013 McAuley, J. and Leskovec, J. (2013), "Hidden factors and hidden topics: understanding rating dimensions with review text," *RecSys*
- 2013 McAuley, J. and Leskovec, J. (2013), "From amateurs to connoisseurs: modeling the evolution of user expertise through online reviews," *WWW*
- Lakkaraju, H. and McAuley, J. and Leskovec, J. (2013), "What's in a name? Understanding the interplay between titles, content, and communities in social media," *ICWSM*
- 2013 Yang, J. and McAuley, J. and Leskovec, J. (2013), "Community detection in networks with node attributes," ICDM
- 2012 McAuley, J. and Leskovec, J. (2012), "Learning to discover social circles in ego networks," NIPS
- 2012 McAuley, J. and Jurafsky, J. and Leskovec, J. (2012), "Learning attitudes and attributes from multiaspect reviews," *ICDM*
- 2012 McAuley, J. and Leskovec, J. (2012), "Image labeling on a network: using social-network metadata for image classification," ECCV
- 2011 McAuley, J. and Ramisa, A. and Caetano, T. (2011) "Optimization of robust loss functions for weakly-labeled image taxonomies: an ImageNet case study," EMMCVPR
- 2010 McAuley, J. and Caetano, T. (2010), "Exploiting data-independence for fast belief-propagation," ICML

2010	McAuley, J. and de Campos, T. and Caetano, T. (2010), "Unified graph matching in Euclidean
	spaces," CVPR

- 2010 McAuley, J. and Caetano, T. (2010), "Exploiting within-clique factorizations in junction-tree algorithms," *AISTATS*
- 2009 McAuley, J. and de Campos, T. and Csurka, G. and Perronnin, F. (2009), "Hierarchical image-region labeling via structured learning," *BMVC*
- 2009 Chen, L. and McAuley, J. and Feris, R. and Caetano, T. and Turk, M. (2009), "Shape classification through structured learning of matching measures," *CVPR*
- 2008 McAuley, J. and Caetano, T. and Smola, A. (2008), "Robust near-isometric matching via structured learning of graphical models," *NIPS*
- 2006 McAuley, J. and Caetano, T. and Smola, A. and Franz, M. (2006), "Learning high-order MRF priors of color images," *ICML*

For a complete list, including workshop papers, book chapters, and papers in which I was a contributing author, see my homepage and Google Scholar profile.

### Patents

- 2011 McAuley, J. and de Campos, T., "Unified graph matching in Euclidean spaces and applications to image comparison and retrieval," *U.S. Patent Application* 12/571,630
- 2011 McAuley, J. and de Campos, T. and Csurka, G. and Perronnin, F., "Consistent hierarchical labeling of images and image regions," *U.S. Patent Application* 12/546,948

#### Recent coverage

#### Media

BUSINESS INSIDER "The next time you're with a beer geek, order a strong ale" BUSINESS INSIDER "Scientists demonstrate that beer geeks have great taste" BUSINESS INSIDER "How to execute the perfect reddit submission" BUSINESS INSIDER "How a title can sink or float a piece of content" NEW SCIENTIST "Things that make a meme explode" TIME "How to succeed on reddit" FORBES "How to craft the perfect reddit posting" THE VERGE "The math behind successful reddit submissions" PHYS.ORG "Stanford trio explore success formula for reddit posts" GIZMODO "This equation can tell you how successful a reddit post will be" GEEKOSYSTEM "The recipe for reddit success" INTERNATIONAL BUSINESS TIMES "The secret to what makes something go viral" SLATE "Les titres qui marchent sur reddit" MIT TECH REVIEW "Algorithm predicts circles of friends using contacts data" WIRED "Algorithm can know your friendship circles better than you do"

Blogs etc.

BEER STREET JOURNAL "Helping beer geeks better understand themselves" 104.3 K-HITS CHICAGO "This is serious business, people: serious beer business" GIGAOM "How to maximize your reddit upvotes, by the numbers" GIRL GUIDES "Science, and the best ways to get on reddit's front page" ACM TECHNEWS "Stanford trio explore success formula for reddit posts" THE CONNECTIVIST "Social media: science not art" ADVENTURES IN CAPITALISM "Three scientific rules for writing viral headlines" PIILAAKSO "Mitä löytyy, kun analysoidaan viisi miljoonaa olutarviota?" EDLAB "Should we recommend right or recommend smart?"

See http://i.stanford.edu/~julian/ for a more complete list.

### Service & teaching experience

#### Reviewing

I have reviewed over a hundred articles for journals and conferences including CVPR, ICML, IJCAI, IJCV, ICWSM, JMLR, NIPS, PAMI, Pattern Recognition, TKDD, UAI, WWW, and many others.

#### Organization

I served a primary role in organizing the CVPR 2011 Workshop on Inference in Graphical Models with Structured Potentials, and the KDD 2013 Workshop on Mining and Learning with Graphs. Both workshops were very successful, and were among the most widely attended at their respective conferences.

#### TEACHING

I have worked as a teaching assistant for classes on concurrency, complexity theory, algorithms, software engineering, and introductory computer science. A typical course involved giving a series of 12 one-hour tutorials, supervising laboratory exercises, marking, and holding office hours.

#### Mentoring

As a postdoc, I have advised, instructed, and mentored several students. I mentored Himabindu Lakkaraju while she rotated in our group as a first-year graduate student. During three months working together, we wrote a paper for *ICWSM*, which generated significant media attention. I mentored Norases Vesdapunt on a rotation project studying users' ratings of articles on *Wikipedia*, and I mentored Max Bodoia on a project on socially-regularized recommender systems. I have also advised several undergraduate and masters students preparing deliverables for an industry collaboration with a major automobile manufacturer.

### Referees

#### Jure Leskovec

Assistant Professor Computer Science Department Stanford University 353 Serra Mall Stanford, CA 94305 U.S.A.

Phone: 650 725 3711 email: jure@cs.stanford.edu

#### Tibério Caetano

Principal Researcher Machine Learning Group NICTA Level 5, 13 Garden Street Eveleigh, NSW 2015 Australia

Phone: (+61) 2 9376 2076 email: tiberio.caetano@nicta.com.au

#### Alexander Smola

Professor Machine Learning Department Carnegie Mellon University 5000 Forbes Avenue Pittsburgh, PA 15213 U.S.A.

Phone: 408 759 1044 email: alex@smola.org

#### Dan Jurafsky

Professor Linguistics and Computer Science Departments Stanford University 450 Serra Mall Stanford, CA 94305 U.S.A.

Phone: 650 723 4284 email: jurafsky@stanford.edu