

Jake M. Hofman

| | | |
|---------------------|--|--|
| CONTACT INFORMATION | Microsoft Research 641 Avenue of the Americas, 7th Floor New York, NY 10011 | Last updated: Mar 2020 jhofman@jakehofman.com jakehofman.com |
| POSITIONS HELD | Microsoft Research Senior Principal Researcher, March 2020 – <i>Present</i> Principal Researcher, September 2015 – February 2020 Senior Researcher, May 2012 – September 2015 Computational Social Science Group Manager: Susan Dumais Columbia University Adjunct Assistant Professor, September 2009 – <i>Present</i> Department of Applied Mathematics Courses: Modeling Social Data, Computational Social Science, Data-Driven Modeling Yahoo! Research Research Scientist, September 2010 – April 2012 Postdoctoral Researcher, September 2008 – September 2010 Human and Social Dynamics Group Manager: Duncan Watts | |
| EDUCATION | Columbia University Ph.D. Physics, October 2008 Advisor: Chris Wiggins Thesis: Statistical Inference for Biophysical Image and Network Data Columbia University M.S. Applied Physics, June 2004 Boston University B.S. Electrical Engineering, June 2002 | |
| PUBLICATIONS | Manipulating and Measuring Model Interpretability F. Poursabzi-Sangdeh, D. G. Goldstein, J. M. Hofman, J. Wortman Vaughan, H. Wallach <i>Revise and resubmit at Management Science</i> Explanation, Prediction, and Causality: Three Sides of the Same Coin? E. Beck, E. J. Bienenstock, J. Bowers, A. B. Frank, T. Grubestic, J. M. Hofman, J. Rohrer, M. Salganik, D. J. Watts <i>Under review at Harvard Data Science Review</i> How visualizing inferential uncertainty can mislead readers about treatment effects in scientific results J. M. Hofman, D. G. Goldstein, J. Hullman <i>38th ACM Conference on Human Factors in Computing (CHI) 2020</i> To Put That in Perspective: Generating Analogies That Make Numbers Easier to Understand C. Riederer, J. M. Hofman, D. G. Goldstein <i>36th ACM Conference on Human Factors in Computing (CHI) 2018</i> The Split-Door Criterion: Identification of Causal Effects Through Auxiliary Outcomes A. Sharma, J. M. Hofman, and D. J. Watts <i>Annals of Applied Statistics</i> 2018 | |

Prediction and Explanation in Social Systems

J. M. Hofman, A. Sharma, and D. J. Watts
Science 2017

What's Happening and What Happened: Searching the Social Web

O. Alonso, V. Kandylas, S.E. Tremblay, J. M. Hofman, and S. Sen
9th ACM Web Science Conference (WebSci) 2017

Exploring the Limits to Prediction in Complex Social Systems

T. Martin, J. M. Hofman, A. Sharma, A. Anderson, D. J. Watts
25th International World Wide Web Conference (WWW) 2016

Improving Comprehension of Numbers in the News

P. J. Barrio, D. G. Goldstein, and J. M. Hofman
34th ACM Conference on Human Factors in Computing (CHI) 2016

Online and Social Media Data as a Flawed Continuous Panel Survey

F. Diaz, M. Gamon, J. M. Hofman, E. Kiciman, D. Rothschild
PLoS ONE 2016

The Structural Virality of Online Diffusion

S. Goel, A. Anderson, J. M. Hofman, and D. J. Watts
Management Science 2015

Scalable Recommendation with Hierarchical Poisson Factorization

P. Gopalan, J. M. Hofman, D. M. Blei
31st Conference on Uncertainty in Artificial Intelligence (UAI) 2015

Estimating the Causal Impact of Recommendation Systems from Observational Data

A. Sharma, J. M. Hofman, D. J. Watts
16th ACM Conference on Economics and Computation (EC) 2015

A Large-Scale Exploration of Group Viewing Patterns

A. Chaney, M. Gartrell, J. M. Hofman, J. Guiver, N. Koenigstein, P. Kohli, and U. Paquet
1st ACM Conference on Interactive Experiences for TV and Online Video (TVX) 2014
Honorable mention for Best Paper Award

Sharding Social Networks

Q. Duong, S. Goel, J. M. Hofman, and S. Vassilvitskii
6th International Conference on Web Search and Data Mining (WSDM) 2013

Who Does What on the Web: A Large-scale Study of Browsing Behavior

M. I. Sirer, J. M. Hofman, and S. Goel
6th International Conference on Weblogs and Social Media (ICWSM) 2012

Who Says What to Whom on Twitter

S. Wu, J. M. Hofman, W. A. Mason, and D. J. Watts
20th International World Wide Web Conference (WWW) 2011

Everyone's an Influencer: Quantifying Influence on Twitter

E. Bakshy, J. M. Hofman, W. A. Mason, and D. J. Watts
4th International Conference on Web Search and Data Mining (WSDM) 2011

Predicting Consumer Activity with Web Search

S. Goel, J. M. Hofman, S. Lahaie, D. M. Pennock, and D. J. Watts
Proceedings of the National Academy of Sciences (PNAS) 2010

Inferring Relevant Social Networks from Interpersonal Communication

M. De Choudry, W. A. Mason, J. M. Hofman, and D. J. Watts
19th International World Wide Web Conference (WWW) 2010

Filtering: A Technique for Counting Triangles Computing MSTs and Clustering in MapReduce

J. M. Hofman, S. Suri, and S. Vassilvitskii
Second Workshop on Massive Data Algorithmics 2010

Graphical Models for Inferring Single Molecule Dynamics

J. E. Bronson, J. M. Hofman, J. Fei, R. L. Gonzalez Jr., and C. H. Wiggins
BMC Bioinformatics 2010

Characterizing Individual Communication Patterns

R. Dean Malmgren, J. M. Hofman, L. A. N. Amaral, and D. J. Watts
15th ACM SIGKDD Conference (KDD) 2009

CentMail: Rate Limiting via Certified Micro-Donations

S. Goel, J. M. Hofman, J. Langford, D. Pennock and D. Reeves
18th World Wide Web Conference Developers Track (WWW) 2009

Learning Rates and States from Biophysical Time Series: A Bayesian Approach to Model Selection and Single-Molecule FRET Data

J. E. Bronson, J. Fei, J. M. Hofman, R. L. Gonzalez, Jr, and C. H. Wiggins
Biophysical Journal 2009

Allosteric Collaboration Between Elongation Factor G and the Ribosomal L1 Stalk Directs tRNA Movements During Translation

J. Fei, J. E. Bronson, J. M. Hofman, R. L. Srinivas, C. H. Wiggins and R. L. Gonzalez, Jr
Proceedings of the National Academy of Sciences (PNAS) 2009

A Bayesian Approach to Network Modularity

J. M. Hofman and C. H. Wiggins
Physical Review Letters (PRL) 2008

Quantification of Edge Velocities and Traction Forces Reveals Distinct Motility Modules During Cell Spreading

B. J. Dubin-Thaler, J. M. Hofman, H. S. Xenias, I. Spielman, A. V. Shneidman,
L. A. David, H. G. Döbereiner, and C. H. Wiggins, and M. P. Sheetz
PLoS ONE 2008

Opposing Effects of PKC θ and WASp on Symmetry Breaking and Relocation of the Immunological Synapse

T. N. Sims, T. J. Soos, H. S. Xenias, B. Dubin-Thaler, J. M. Hofman, J. C. Waite,
T. O. Cameron, V. K. Thomas, R. Varma, C. H. Wiggins, M. P. Sheetz, and D. R. Littman,
and M. L. Dustin
Cell 2007

Lateral Membrane Waves Constitute a Universal Dynamic Pattern of Motile Cells

H.G. Döbereiner, B. J. Dubin-Thaler, J. M. Hofman, H. S. Xenias, T. N. Sims, G. Giannone,
M. L. Dustin, and C. H. Wiggins, and M. P. Sheetz
Physical Review Letters (PRL) 2006

The Small GTPase R-Ras Regulates Organization of Actin and Drives Membrane Protrusions Through the Activity of PLC- ϵ

A. S. Ada-Nguema, H. Xenias, J. M. Hofman, C. H. Wiggins, M. P. Sheetz, and P. J. Keely
Journal of Cell Science 2006

Jake M. Hofman

Nonmuscle Myosin IIA-dependent Force Inhibits Cell Spreading and Drives F-actin Flow

Y. Cai, N. Biais, G. Giannone, M. Tanase, G. Jiang, J. M. Hofman, C. H. Wiggins, P. Silberzan, A. Buguin, and B. Ladoux, and M. P. Sheetz
Biophysical Journal. 2006

WORKING PAPERS

Effectively Communicating Effect Sizes

Y. S. Kim, J. M. Hofman, D. G. Goldstein

Improving Comprehension of Small Risks: Estimation, Recall, and Error Detection

J. M. Hofman, D. G. Goldstein, M. A. Jenny, S. M. Herzog

Predictive modeling needs pre-registration too

J. M. Hofman, A. Sharma, D. J. Watts

PATENTS

Awareness engine

O. Alonso, J. M. Hofman, V. Kandylas, S. Sen, S. E. Tremblay
US Patent 10,176,265 2019

Interactive graphical system for estimating body measurements

J. M. Hofman, D. G. Goldstein
US Patent 10,521,429 B2 2019

Perspective Annotation for Numerical Representations

J. M. Hofman, M. Dudik, D. G. Goldstein
US Patent 10,146,756 2018

Content Virality Determination and Visualization

J. M. Hofman, R. Gruen, S. Goel, D. J. Watts
US Patent 9,491,038 2016

System and Method to Enable Communication Group Identification

J. M. Hofman, W. Mason, D. J. Watts
US Patent 9,430,755 2016

Media Recommendations for a Social-Software Website

K. Elliott-McCrea, W. Mason, S. Suri, J. M. Hofman
US Patent 8,484,226 2013

Apparatus, System, and Method for Expert Identification to Answer User Queries

J. M. Hofman, S. Goel, D. J. Watts
US Patent App. 13/563,638 2012

COURSES

Modeling Social Data

Columbia University, Departments of Applied Mathematics and Computer Science
Spring 2015, Spring 2017, Spring 2019
modelingsocialdata.org

Microsoft Research Data Science Summer School

Director and co-founder
Summer 2013 – Present
ds3.research.microsoft.com

Jake M. Hofman

Computational Social Science

Columbia University, Department of Applied Mathematics
Spring 2013
compsocialscience.org

Data-Driven Modeling

Columbia University, Department of Applied Mathematics
Fall 2009, Spring 2012
jakehofman.com/ddm

TUTORIALS

Applied Machine Learning for Engineers

Microsoft, Redmond, WA, 2013 – Present
Microsoft, Herzliya, Israel, 2013

Machine Learning for Social Scientists

University of Michigan, Center for Complex Systems, 2018

Data-Driven Modeling

Math for America, New York, NY, 2012

Data Bootcamp

Strata, Santa Clara, CA, 2011

Large-Scale Social Media Analysis with Hadoop

ICWSM, Washington, D.C., 2010

Machine Learning Tutorial

Boulder Biophysics Summer School, Boulder, CO, 2007

EDITORIAL BOARDS

Associate Editor

Proceedings of the Internal Conference on Web and Social Media (ICWSM)
2018 – Present

JOURNAL REVIEWING

Management Science, 2018, 2019

Proceedings of the National Academy of Sciences (PNAS), 2009, 2015, 2016, 2017, 2019

Nature Physics, 2019

Transactions on Computer-Human Interaction, 2019

Transactions on Social Computing, 2019

Socius, 2017

Transactions on Network Science and Engineering (TSNE), 2016

Journal of Complex Networks, 2014

Transactions on Information Systems (TIS), 2012

Transactions on Knowledge and Data Engineering (TKDE), 2010, 2011, 2012

Transactions on Knowledge Discovery from Data (TKDD), 2011, 2012

Journal of Statistical Computing, 2010, 2012

Very Large Databases (VLDB), 2011

Physical Review Letters (PRL), 2009, 2011

Statistical Modeling, 2010

Physical Review E (PRE), 2009

Annals of Applied Statistics (AOAS), 2009

Europhysics Letters, 2009

Bioinformatics, 2009

SENIOR PROGRAM COMMITTEES

International World Wide Web Conference (WWW/TheWebConf) 2017 - Present

International Conference on Web and Social Media (ICWSM) 2017

Jake M. Hofman

| | |
|-----------------|---|
| PROGRAM | Neural Information Processing (NIPS/NeurIPS), 2010, 2015, 2018 |
| COMMITTEES | Mining and Learning with Graphs (MLG), 2013, 2016, 2017, 2018 Computer-Supported Cooperative Work and Social Computing (CSCW), 2018 Data Science, Journalism, and Media workshop at KDD (DSJM), 2018 International Conference on Web Search and Data Mining (WSDM), 2012 – 2017 International World Wide Web Conference (WWW/TheWebconf) 2010 – 2016 International Conference on Machine Learning (ICML), 2012 – 2015 International Conference on Knowledge, Discovery, and Data mining (KDD), 2011, 2015 International Conference on Weblogs and Social Media (ICWSM), 2011, 2014 International Conference on Information and Knowledge Management (CIKM), 2009 |
| ORGANIZING | Co-organizer for the New York City Data Science Seminar Series, 2017 – Present Board of advisors, Bronx Academy of Software Engineering High School, 2013 – Present Sponsorship chair for Web Search and Data Mining (WSDM), 2014 Workshop on Computational and Online Social Science (CAOSS), 2012 Workshop on Machine Learning Careers in NYC Startups, 2012 National Academy of Sciences Japanese-American Frontiers of Science (JAFOS), 2011 Workshop on Analyzing Graphs: Theory and Applications (part of NIPS), 2008 |
| HONORS & AWARDS | Fellow at the Columbia University Center for Applied Statistics, 2014 – Present Outstanding Reviewer Award, WSDM 2015 Honorable mention for Best Paper Award at TVX 2014 Best Student Presentation Award, MLG 2008 Selvaggi Award in Physics, Columbia University 2007 Summa Cum Laude, Boston University 2002 Tau Beta Pi Engineering Honor Society Nominee, Boston University 2002 |
| INVITED TALKS | How Predictable is the Spread of Information? Columbia University Business School, Marketing Department Seminar, 2019 Annual Meeting of the Society for Mathematical Psychology, Keynote, 2019 Duke University, Fuqua School of Business Strategy Department Seminar, 2018 UC Boulder, Summer Institute in Computational Social Science, 2018 BIRS Workshop on Statistical Inference and Machine Learning, 2018 University of Pennsylvania, OID Department Seminar, 2017 Amazon's Radical Social Sciences Workshop, 2017 Credit Suisse, 2017 Department of Information Systems and Statistics Seminar at Baruch College, 2017 National Science Foundation Data Science Seminar, 2017 ForecastNY, 2017 Columbia University, Applied Mathematics Seminar, 2016 UT Austin Workshop on Social and Business Analytics, 2016 Columbia University, ISDE Seminar, 2016 Prediction and Explanation in the Social Sciences: The Good, the Bad, and the Ugly National Academies of Sciences, Predictive Analytics Colloquium, 2019 National Institutes of Health, OBSSR Methods Seminar, 2018 To Put That in Perspective: Generating Analogies That Make Numbers Easier to Understand Columbia University, Becoming Better Choice Architects, 2016 – 2018 Data Science and Journalism Workshop at KDD, 2017 |

Microsoft Research Data Science Summer School

Machine Eatable, Civic Hall, 2018

Lehman College, Department of Computer Science, 2017

Prediction and Causation

Yale University, Designing the Digital Economy, 2016, 2017

Rules to Hack By: What Social Science Taught Me About Coding

Columbia University, Data Science Hackathon Keynote, 2016

A Data-Driven Approach to Understanding Organizations

The World Bank, Visual Knowledge Conference, 2016

Improving Comprehension of Numbers in the News

IBM Watson Yorktown Heights, 2016

The Structural Virality of Online Diffusion

Datapoint NYC, 2015

Connective Media Seminar, Cornell Tech, 2014

Joint Statistical Meetings, 2013

Santa Fe Institute, 2013

Estimating the Causal Impact of Recommendation Systems from Observational Data

RAIN Seminar, Stanford University, 2015

Computational Social Science

NYC Data Science Meetup, 2014

Modeling Social Data

IGERT Distinguished Speaker Series, Columbia University, 2014

University of Minnesota, Institute for Mathematics and its Applications, 2012

University of Massachusetts, Computational Social Science Seminar, 2012

Learning from Web Activity

Joint Statistical Meetings, 2011

UCLA, Department of Statistics, 2011

Inferring Relevant Networks

Duke University, Statistical and Applied Mathematical Sciences Institute, 2011

Inferring the Structure and Scale of Modular Networks

Rutgers University, Department of Computer Science, 2011

International Conference on Information Theory and Applications, 2010

International Workshop on Mining and Learning with Graphs (MLG), 2008

Bayesian Inference: Principles and Practice

NYC Machine Learning Meetup, 2010

NYC R Meetup, 2009

A Bayesian Approach to Network Modularity

Columbia University, Department of Statistics Student Seminar, 2009

Workshop on Statistical Models of Networks, NIPS, 2008

American Physical Society March Meeting, 2008

Northeastern University, Center for Complex Network Research, 2008

Cambridge University, Department of Engineering, 2008

Princeton University, Department of Bioinformatics, 2007

Columbia University, Department of Applied Mathematics, 2007

Jake M. Hofman

Bayesian Statistics and Massive Data Streams

11th Annual Japanese-American Kavli Frontiers of Science Symposium, 2008

Community detection: Model Fitting, Comparison, and Utility

SFI Workshop on Statistical Inference for Complex Networks, 2008

Social Network Analysis with Hadoop

Hadoop World, 2009

Simple Math for a Complex World: Random Walks in Biology and Finance

Dalton High School, 2007