

Work

Pew Research Center, Data Labs

Washington, DC

Managing Director

2015 - Present

- Founded new data science research group—methodological innovation, machine learning for text and images, complex experimental designs, R&D, advising and consulting on data science efforts around the Center.
- Led publication of collaborative data science-oriented projects with national media attention. Served as public spokesperson for Data Lab's research. See <http://www.pewresearch.org/topics/data-labs/>.
- Supervised development of Python-oriented data science infrastructure with organization-wide adoption: Linux-based AWS cloud servers; custom Django dashboards/labeling/harmonization tools; large scale scraping/API query data collection efforts; PostgreSQL, with S3, BigTable, Hive for larger data sets; Jupyter and RStudio Server web interfaces for analysis; GitHub version control.
- Developed protocols for data science at Pew: machine learning standards, security, data management.
- Managing/hiring top data scientists, researchers, engineers for Data Labs and other groups around the center.

Facebook, Core Data Science Team

Menlo Park, CA

Researcher, Computational Social Scientist

2012 - 2015

- Civic engagement/politics: Research on echo chambers, ideology, media consumption, top tier publications. Predictive modeling for political campaigns; strong positive impacts on revenue, CTRs, client feedback. Groundwork for [Town Hall](#) constituent-communication features.
- Methodological/data science work: estimating heterogenous treatment effects in experiments while limiting false-positives (R); feature modeling using large scale, high dimensional network data (Python/Giraph); adaptive experimentation for online decision-making systems (contextual multi-armed bandits in Python).
- Misc: consultation/feedback on experimental design; simulation and data driven power analysis (R); third-party data integration and harmonization.
- Taught DataCamp statistics and large datasets course, instructed joint [Facebook-Udacity Exploratory Data Analysis](#) course.

Stanford University, Social Science Data and Software (SSDS)

Stanford, CA

Consultant

2011 - 2013

- Guided Stanford clients through experimental and observational study design, including web experiments, scraping, text analysis, network analysis, statistical modeling, visualization in R.
- Led workshops introducing R, data visualization, scraping & analysis of text data.

Science Applications International

Washington, DC

Media Analytics Developer, Team Lead

2005 - 2011

- Multi-language media analysis framework: sentiment, attribution of responsibility, economic framing; GUI; database; visualizations. Human-coder training, calibration/reliability, report production.
- Network visualizations showing spread of themes across media sources, identification of key originators and distributors of content on the web. Visualizations integrated into web-based analytic system.
- More than 50 reports using primary language internet sources, with readership across the U.S. Public Diplomacy community. Citations, follow-up research requests from government officials including Undersecretary Hughes, Congressman Markey, and then-Senator Clinton.

Law-on-line, Inc.

Tucson, AZ

Information Technology Associate/Consultant

1996 - 2005 (non-continuously)

- Developed and helped bring to market multiple security software products for applications related to digital signature technology. Java, PGP key infrastructure, web dev.

Education

Stanford University <i>Ph.D. Communication (2013), M.S. Statistics (2012)</i>	Stanford, CA 2008 - 2013
University of California, San Diego <i>Masters, School of Global Policy and Strategy (Int'l MBA-curriculum)</i>	La Jolla, CA 2003 - 2005
University of California, Santa Barbara <i>Bachelor of Arts with Honors in Political Science and Global Studies, Japanese minor</i>	Santa Barbara, CA
Sophia University (上智大学, 市谷キャンパス) <i>Study abroad</i>	Tokyo, Japan

Books

- The Impression of Influence: Legislator Communication, Representation, and Democratic Accountability* with Justin Grimmer & Sean Westwood. 2014. *Princeton University Press*.
- Reviewed in: [Journal of Politics](#), [Congress & The Presidency](#), [Political Communication](#), [Political Science Quarterly](#).
 - Subject of roundtable at Western Political Science Association, 2015.
 - Media coverage: [Vox \(Mischiefs of Faction\)](#).

Publications

- “[Partisan Conflict and Congressional Outreach](#)” with Patrick van Kessel, Adam Hughes, Nick Judd, Rachel Blum. *Pew Research Center*, 2017.
- Media coverage: [Washington Post](#), [Mother Jones](#).
- “[Estimating Heterogeneous Treatment Effects and the Effects of Heterogeneous Treatments with Ensemble Methods](#)” with Justin Grimmer and Sean J. Westwood. *Political Analysis*, 2017.
- [Replication materials](#).
 - Software implementation (under development): <https://github.com/SolomonMg/HetSL>
- “[Exposure to ideologically diverse news and opinion on Facebook](#)” with Eytan Bakshy and Lada A. Adamic. *Science*, 2015.
- [Review by David Lazer](#).
 - [Supplementary materials](#).
 - [Replication materials](#).
 - Media coverage: [New York Times](#), [Washington Post](#), [Ars Technica](#), [BBC](#), [CBS News](#).
- “[Bias in the Flesh: Skin Complexion and Stereotype Consistency in Political Campaigns](#)” with Ethan Plaut & Maria Jabon. *Public Opinion Quarterly*, 2015.
- [Replication materials](#).
 - Media coverage: [CBS News](#), [Washington Post](#).
- “[Quantifying Social Media’s Political Space: Estimating Ideology from Publicly Revealed Preferences on Facebook](#)” with Robert Bond. *American Political Science Review*, 2015.
- “[Selective Exposure in the Age of Social Media: Endorsements Trump Partisan Source Affiliation when Selecting News Online](#)” with Sean Westwood. *Communication Research*, 2013.
- “[Role of Diffusion Weighted Imaging in the Diagnosis of Pediatric Abdominal Tumors](#)” with Rakhee Gawande, Gabriel Gonzalez, Aman Khurana, & Heike Daldrup-Link. *Pediatric Radiology*, 2013.
- “[How Words and Money Cultivate a Personal Vote: The Effect of Legislator Credit Claiming on Constituent Credit Allocation](#)” with Justin Grimmer & Sean Westwood. *American Political Science Review*, 2012.

- “Intravenous Ferumoxytol Allows Noninvasive MR Imaging of Stem Cell Transplants” with Aman Khurana, Rakhee Gawande, Guiting Lin, Sungmin Lee, Rosalinda Castaneda, Nikita Derugin, Laura Pisani, Tom F. Lue, & Heike Daldrup-Link. *Radiology*, 2012.
- “Do Attitudes about Immigration Predict Willingness to Admit Individual Immigrants? A Cross-National Test of the Person-Positivity Bias” with Shanto Iyengar, Simon Jackman, Nicholas Valentino, Toril Aalberg, Raymond Duch, Kyu S. Hahn, Stuart Soroka, Allison Harell, & Tetsuro Kobayashi. *Public Opinion Quarterly*, 2012.
- “Differentiation of the Normal Thymus from Anterior Mediastinal Lymphoma on Pediatric PET-CTs” with Rakhee Gawande, Aman Khurana, Dong Zhang, Rosalinda Castaneda, Robert Goldsby, Randal Hawkins, & Heike Daldrup-Link. *Radiology*, 2011.
- “Who is a ‘Deserving’ Immigrant? An Experimental Study of Norwegian Attitudes” with Toril Aalberg & Shanto Iyengar. *Scandinavian Political Studies*, 2011.

Honors and Awards

- 2010** \$60,000 Google Research Award (with Sean Westwood, supervised by Dan McFarland & Jeremy Bailenson) “Harnessing the social graph: Social cues and reliability in content selection and evaluation”
- 2010** \$1,000 A-REP Research Grant (with Sean Westwood), “The impact of social news ratings on media consumption heuristics and agenda setting”
- 2008** Department of Communication Fellowship
- 2007** Annual Achievement Award for Excellence in Science and Technology - Technical Collaboration, Science Applications International
- 2007** Selected for Future Leaders program, Science Applications International
- 2004–2005** Dean’s Fellow, University of California, San Diego IR/PS
- 2001** Phi Beta Kappa, University of California, Santa Barbara
- 2000** Golden Key, University of California, Santa Barbara

Invited Talks and Workshops

- USC Media Exposure and Opinion Formation workshop, Los Angeles, November 9, 2017.
- Social Media and Political Participation Global Conference, New York, November 3, 2017.
- World Bank Big Data in Government, Washington, May 16, 2017.
- Data and Society, New York, March 31, 2017.
- Social Media and Political Participation Global Conference, Abu Dhabi, February 9, 2017.
- Analyst Institute Social Networks Roundtable, March 9, 2016.
- GAO CG Forum on 21st Century Data and Analytics, January 28, 2016.
- Duke Network Analysis Center (DNAC), October 20, 2015.
- Spatial Computational Social Science at Stanford, Oct 20, 2014.
- UCLA Institute for Politics in Math, April 22, 2014.

Stanford Political Science American Politics Workshop, April 22, 2013.

Stanford Computational Social Science, Analyzing large data sets, September 20, 2012.

Stanford Political Science Methods Workshop, May 2, 2012.

Software

HetEffects with Justin Grimmer. R package utilizing ensemble machine learning methods to estimate high-dimensional causal effects that vary by treatment-group combination. Development version at <https://github.com/SolomonMg/HetSL>. See [Estimating Heterogeneous Treatment Effects and the Effects of Heterogeneous Treatments with Ensemble Methods](#).

ImageMetrics R package to facilitate the analysis of image data, designed to facilitate analysis described in “Bias in the flesh” (see above). See <https://github.com/SolomonMg/ImageMetrics>.

NetCluster, **triads**, **NetData** with Mike Nowak, Sean Westwood & Dan McFarland. R packages to determine node-level triad type membership for the 16 types of triads that occur in a directed network, facilitate network clustering and evaluation of cluster configurations, and provide example data. Examples in “[Social network analysis labs in R.](#)”

Computing

R Statistical Programming Language Data visualization, conventional statistical analysis/econometrics, Bayesian statistics, text-as-data, image, GIS, and network data, including authoring multiple software packages.

Python Data analysis (Pandas, NumPy, SciPy), Machine Learning (SciKit Learn, Gensim), parsing data from the web, computer vision, text-as-data, deep-learning (text data), data management.

SQL Experience utilizing Hive (HQL) with petabyte scale data sets, PostGreSQL, database design with MySQL, SQLite.

Java Undergraduate & natural language processing courses taken; some instrustry coding experience.

Web Dev Working knowledge of JavaScript, PHP, JSON, RegEx, HTTP, Linux shell scripts, C-sharp/.NET.

Journal Service

Reviewer, *American Political Science Review*

Reviewer, *American Journal of Political Science*

Reviewer, *Political Analysis*

Reviewer, *Public Opinion Quarterly*

Reviewer, *Political Behavior*

Assistant Editor, Reviewer, *Political Communication*

Founding Editor, Reviewer, *Journal of International Policy Solutions*

Guest Editor, Reviewer, *Journal of Public and International Affairs*

Languages

Indonesian/Malay Bahasa 2 years professional experience as Indonesian analyst/linguist; 2 years intensive Indonesian graduate study; Certified 3/2 Written/Spoken on the Interagency Language Roundtable (ILR) scale used by the U.S. Foreign Service.

Japanese 3 years undergraduate language study, 1 year study abroad in Tokyo, Japan.

Arabic Digital courses taken; introductory-level translation experience.