

## **Psych 626: Topics in Computational Social Sciences**

**Fall—TH—2-6**

**Location:** BCI 266

**Instructor:** Morteza Dehghani

**Office:** SGM 607

**Office Hours:** M 10-12, or by appt.

**Contact Info:** mdehghan@usc.edu

### **Course Description**

Topics in computational social sciences focuses on applications of big-data methods (machine learning techniques, natural language processing and network analysis), guided by psychological theories, for identifying various social and cognitive properties evident in human related big data. In this course, we will survey state-of-the-art techniques, and applications of such techniques, for investigating various aspects of human cognition. The intended audience for this course is psychology and computer science PhD students, and more broadly graduate students in social sciences, who are interested in using machine-learning techniques for analysis of data. Also, this course may be of interest to PhD students in communications and the business school.

### **Learning Objectives**

This course is designed to survey current state of research in computational social sciences. In order to achieve this objective, each week several papers/books will be read and presented by students. Also, there will be a final project and written report.

**Prerequisite(s):** Instructor permission

**Recommended Preparation:** Psych 625 or a similar course

### **Course Notes**

Students are not allowed to use laptops or smartphones during class, unless used for class presentations. Homework assignments will be posted on Blackboard. Students are also highly encouraged to use the course forum on Blackboard.

### **Technological Proficiency and Hardware/Software Required**

The project for the class involves programming. Students are recommended to have taken a programming course before enrolling in this course.

### **Description and Assessment of Assignments**

- a. Paper presentation. Each student will present a set of papers related to one of the topics discussed in class.
- b. Reaction paragraphs. Students are asked to write a short note, one or two paragraphs in length, about their reaction to the reading assignments of the week. These can be a quick summary of the material, comments about the subject area, or a critique of a particular theory or experiment. I will read these paragraphs before each class, and will use them to guide the discussion in class.
- c. Class Project. This class is project oriented. The goal of the project is for students to get experience in applying big-data methods for analyzing behavioral data. This will include a project proposal presentation, three project update presentations, final project presentation, and a report. For project proposals, students will present a problem and a data collection method and/or dataset for which they want to analyze using the methods discussed in class. Each presentation should be about 10-15mins. The goal of the project update presentations

is to inform the class about the state of the project and brainstorm with other students on how to solve the remaining issues. Each update presentation should be around 10 minutes. For the final project presentation, each student/group will give a 15-20min presentation on their project. Students are expected to spend at least 80 hours working on their final project. The project report will be 10-15 pages.

- d. Review Paragraphs: Students will provide reviews for other students' projects. The reviews for the project proposal and project updates can each be about a paragraph in length and should address the strength and weaknesses of the discussed projects and other potential sources of related work. Each student is asked to read the final report of two other groups and provide comprehensive reviews of the final projects as well. These reviews will be sent to the authors, and the authors will have a chance to respond to the reviews.

## Grading Breakdown

Assignment	% of Grade
Project Proposal	10
Project Update 1	5
Project Update 2	5
Project Update 3	5
Project Presentation	10
Final Project Report	15
Reviews	10
Participation	15
Paper Presentations	15
Reaction Paragraphs	10
<b>TOTAL</b>	<b>100</b>

## Assignment Submission Policy

The reaction paragraphs need to be submitted through Blackboard before 9:00am of the day the corresponding readings are due. The final project report and the reviews need to be submitted through Blackboard on the day of the final.

## Course Schedule:

The following schedule is tentative and may change during the semester (\* indicates optional readings).

- Week 1: Introduction to Computational Social Sciences
  - Ruths, D., & Pfeffer, J. (2014). Social media for large studies of behavior. *Science*, 346(6213), 1063-1064.
  - Lazer D, et al. (2009) Computational social science. *Science*, 323(5915):721–723.
  - Yarkoni, T. (2012). Psychoinformatics New Horizons at the Interface of the Psychological and Computing Sciences. *Current Directions in Psychological Science*, 21(6), 391-397.
  - King, G. (2011). Ensuring the data-rich future of the social sciences. *Science*, 331(6018), 719:721.
- Week 2: Overview of Text Analysis Methods
  - Johnson, K. M., Garten, J., Dehghani, M. & Graham, J. (in revision). Advancements in language-based assessments: A practical guide for psychologists
  - liev, R., Dehghani, M., & Sagi, E. (2014). Automated text analysis in psychology: methods, applications, and future developments. *Language and Cognition*, 1-26.

- Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, mps028.
- Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. *Science*, 349(6245), 261-266.
- Chen, E. E., & Wojcik, S. P. (2016). A practical guide to big data research in psychology. *Psychological Methods*, 21(4), 458.
- \*Dehghani, M., Johnson, K. M., Garten, J., Boghrati, R., Hoover, J., Balasubramanian, V., ... & Parmar, N. J. (2016). TACIT: An open-source text analysis, crawling, and interpretation tool. *Behavior research methods*, 1-10.
  
- Week 3: Text Analysis in Social Sciences – Word Count 1
  - Pennebaker, J. W. (2011). *The secret life of pronouns: How our words reflect who we are*. New York, NY: Bloomsbury Press.
  - Boyd, R. L., & Pennebaker, J. W. (2015). Did Shakespeare write Double Falsehood? Identifying individuals by creating psychological signatures with text analysis. *Psychological science*, 0956797614566658.
  
- Week 4: Text Analysis in Social Sciences – Word Count 2, Project Proposals
  - Back, M. D., Küfner, A. C., & Egloff, B. (2010). The emotional timeline of September 11, 2001. *Psychological Science*, 21(10), 1417:1419.
  - Pury, C. L. (2011). Automation can lead to confounds in text analysis Back, Küfner, and Egloff (2010) and the Not-So-Angry Americans. *Psychological science*.
  - Back, M. D., Küfner, A. C., & Egloff, B. (2011). “Automatic or the people?” anger on September 11, 2001, and lessons learned for the analysis of large digital data sets. *Psychological Science*, 22(6), 837:838.
  - Frimer, J. A., Aquino, K., Gebauer, J. E., Zhu, L. & Oakes, H. (2015). A decline in prosocial language helps explain public disapproval of the U.S. Congress. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 6591-6594.
  - Dehghani, M., Bang, M., Medin, D., Marin, A., Leddon, E., Waxman, S. (2013). Epistemologies in the Text of Children’s Books: Native and Non-Native Authored Books. *International Journal of Science Education*, 35, 13.
  - Iliev, R., Hoover, J., Dehghani, M., & Axelrod, R. (2016). Linguistic positivity in historical texts reflects dynamic environmental and psychological factors. *Proceedings of the National Academy of Sciences*, 113(49), E7871-E7879.
  - Lansdall-Welfare, T., Sudhakar, S., Thompson, J., Lewis, J., Team, F. N., & Cristianini, N. (2017). Content analysis of 150 years of British periodicals. *Proceedings of the National Academy of Sciences*, 201606380.
  - \*Wojcik, S., Hovasapian, A., Graham, J., Motyl, M., & Ditto, P. H. (2015). Conservatives report, but liberals display, greater happiness. *Science*, 347, 1243-1246.
  - \*Decter-Frain, A., & Frimer, J. A. (2016). Impressive Words: Linguistic Predictors of Public Approval of the US Congress. *Frontiers in psychology*, 7.
  - \*Iliev, R., & Axelrod, R. (2016). Does causality matter more now? Increase in the proportion of causal language in English texts. *Psychological science*, 0956797616630540.
  
- Week 5: Text Analysis in Social Sciences – Modern Approaches 1, Project Update 1
  - Chae, D. H., Clouston, S., Hatzenbuehler, M. L., Kramer, M. R., Cooper, H. L., Wilson, S. M., ... & Link, B. G. (2015). Association between an Internet-Based Measure of Area Racism and Black Mortality. *PLOS ONE*
  - Eichstaedt, J. C., Schwartz, H. A., Kern, M. L., Park, G., Labarthe, D. R., Merchant, R. M., ... & Seligman, M. E. (2015). Psychological Language on Twitter Predicts County-Level Heart Disease Mortality. *Psychological science*

- Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., ... & Seligman, M. E. (2014). Automatic Personality Assessment Through Social Media Language. *Journal of personality and social psychology*, 108(6), 934.
  - Garcia, D., & Sikström, S. (2014). The dark side of Facebook: Semantic representations of status updates predict the Dark Triad of personality. *Personality and Individual Differences*, 67, 92-96.
  - Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., ... & Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one*, 8(9), e73791
- Week 6: Text Analysis in Social Sciences – Modern Approaches 2, Project Proposals
  - Sagi, E. & Dehghani, M. (2014). Measuring Moral Rhetoric in Text. *Social Science Computing Review*, 32, 2, 132-144.
  - Dehghani, M., Johnson, K., Hoover, J., Sagi, E., Garten, J., Parmar, N. J., ... & Graham, J. (2016). Purity homophily in social networks. *Journal of Experimental Psychology: General*, 145(3), 366.
  - Garten, J., Hoover, J., Johnson, K. M., Boghrati, R., Iskiwitch, C., & Dehghani, M. (2017). Dictionaries and distributions: Combining expert knowledge and large scale textual data content analysis. *Behavior Research Methods*, 1-18.
  - Hoover, J., Johnson, K., Boghrati, R., Graham, J., Dehghani, D. (2017). Charitable Donation, Moral Framing, and Social Media: Associations between moral framing and charitable donation
  - Lin, Y., Hoover, K., Moojiman, M., Dehghani, M. Ji, H. (2017) Acquiring Background Knowledge to Improve Moral Value Prediction
  - \*Boyd, R. L., Wilson, S. R., Pennebaker, J. W., Kosinski, M., Stillwell, D. J., & Mihalcea, R. (2015, April). Values in Words: Using Language to Evaluate and Understand Personal Values. *In Ninth International AAAI Conference on Web and Social Media*.
- Week7: Text Analysis in Social Sciences – Modern Approaches 3, Project Update 1
  - Rule, A., Cointet, J. P., & Bearman, P. S. (2015). Lexical shifts, substantive changes, and continuity in State of the Union discourse, 1790–2014. *Proceedings of the National Academy of Sciences*, 112(35), 10837-10844.
  - Kim, S. M., & Hovy, E. (2006, July). Extracting opinions, opinion holders, and topics expressed in online news media text. *In Proceedings of the Workshop on Sentiment and Subjectivity in Text*. Association for Computational Linguistics.
  - Wienberg, C., Roemmele, M., and Gordon, A. (2013) Content-Based Similarity Measures of Weblog Authors. *The 4th Annual ACM Web Science Conference (WebSci-13)*, May 2-4, 2013, Paris, France.
  - Dam, G., & Kaufmann, S. (2008). Computer assessment of interview data using latent semantic analysis. *Behavior Research Methods*, 40(1), 8:20.
  - Sagae, K., Gordon, A. S., Dehghani, M., Metke, M., Kim, J. S., Gimbel, S. I., Tipper, C., Kaplan, J., Immordino-Yang, M. H. (2013). A Data-Driven Approach for Classification of Subjectivity in Personal Narratives. *Proceedings of the 2013 Workshop on Computational Models of Narrative*.
- Week 8: Text Analysis in Social Sciences – Modern Approaches 4, Project Update 1
  - Sagae, K., Gordon, A. S., Dehghani, M., Metke, M., Kim, J. S., Gimbel, S. I., Tipper, C., Kaplan, J., Immordino-Yang, M. H. (2013). A Data-Driven Approach for Classification of Subjectivity in Personal Narratives. *Proceedings of the 2013 Workshop on Computational Models of Narrative*.
  - Iliev, R., & Smirnova, A. (2014). Revealing Word Order: Using Serial Position in Binomials to Predict Properties of the Speaker. *Journal of psycholinguistic research*, 1-31.
  - Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334), 183-186.
  - Peng, Y., & Jiang, H. (2015). Leverage financial news to predict stock price movements using word embeddings and deep neural networks. *arXiv preprint arXiv:1506.07220*.

- Boghrati, R., Hoover, J., Johnson, K., Garten J., Dehghani, M. (2017). Conversation Level Syntax Similarity Metric
- Dehghani, M., Sagae K., Sachdeva, S. & Gratch, J. (2014). Linguistic Analysis of the debate over the Construction of the 'Ground Zero Mosque'. *Journal of Information Technology & Politics*, 11, 1-14.
- Week 9: Text Analysis in Social Sciences – Clinical Applications, Project Update 1
  - Coppersmith, G., Dredze, M., Harman, C., & Hollingshead, K. (2015). From ADHD to SAD: Analyzing the Language of Mental Health on Twitter through Self-Reported Diagnoses. In *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, Denver, Colorado, USA, June. North American Chapter of the Association for Computational Linguistics*.
  - Walsh, C. G., Ribeiro, J. D., & Franklin, J. C. (2017). Predicting Risk of Suicide Attempts Over Time Through Machine Learning. *Clinical Psychological Science*, 2167702617691560.
  - Mitchell, M., Hollingshead, K., & Coppersmith, G. Quantifying the Language of Schizophrenia in Social Media. *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*. Denver, CO: North American Association for Computational Linguistics; 2015.
  - Kumar M, Dredze M, Coppersmith G, De Choudhury M. Detecting changes in suicide content manifested in social media following celebrity suicides. In: *Proceedings of the 26th ACM Conference on Hypertext and Hypermedia*. New York, NY: Association for Computing Machinery; 2015
  - Resnik, P., Armstrong, W., Claudino, L., Nguyen, T., Nguyen, V. A., & Boyd-Graber, J. (2015). Beyond LDA: Exploring Supervised Topic Modeling for Depression-Related Language in Twitter. In *Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology (CLPsych)*.
  - \*Flekova, L., Ruppert, E., & Preotiuc-Pietro, D. (2015, September). Analysing domain suitability of a sentiment lexicon by identifying distributionally bipolar words. In *WASSA 2015 : 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis* (p. 77).
  - \*Schwartz, H. A., Eichstaedt, J., Kern, M. L., Park, G., Sap, M., Stillwell, D., ... & Ungar, L. (2014, June). Towards assessing changes in degree of depression through facebook. In *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality* (pp. 118-125).
- Week 10: Applications of Network/Geo-Spatial Analysis in Social Sciences
  - Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting From Left to Right Is Online Political Communication More Than an Echo Chamber? *Psychological science*, 0956797615594620.
  - Lerman, K., & Ghosh, R. (2010, May). Information contagion: An empirical study of the spread of news on Digg and Twitter social networks. In *Proceedings of 4th International Conference on Weblogs and Social Media (ICWSM)*.
  - Phan, T. Q., & Airoidi, E. M. (2015). A natural experiment of social network formation and dynamics. *Proceedings of the National Academy of Sciences*, 112(21), 6595-6600.
  - Rentfrow, P. J., Gosling, S. D., Jokela, M., Stillwell, D. J., Kosinski, M., & Potter, J. (2013). Divided we stand: Three psychological regions of the United States and their political, economic, social, and health correlates. *Journal of Personality and Social Psychology*, 105(6), 996.
  - Jokela, M., Bleidorn, W., Lamb, M. E., Gosling, S. D., & Rentfrow, P. J. (2015). Geographically varying associations between personality and life satisfaction in the London metropolitan area. *Proceedings of the National Academy of Sciences*, 112(3), 725-730.
  - \* Hodas, N. O., Kooti, F., & Lerman, K. (2013). Friendship paradox redux: Your friends are more interesting than you. arXiv preprint arXiv:1304.3480.
- Week 11: Audio/Video Analysis and their applications – Manual Analysis

- Sbarra, D. A., Smith, H. L., & Mehl, M. R. (2012). When leaving your ex, love yourself: Observational ratings of self-compassion predict the course of emotional recovery following marital separation. *Psychological Science*, 23, 261-269.
- Mehl, M. R., Vazire, S., Holleran, S. E., & Clark, C. S. (2010). Eavesdropping on happiness: Well-being is related to having less small talk and more substantive conversations. *Psychological Science*, 21, 539-541.
- Ramirez-Esparza, N., Mehl, M. R., Alvarez Bermudez, J., & Pennebaker, J. W. (2009). Are Mexicans more or less sociable than Americans? Insights from a naturalistic observation study. *Journal of Research in Personality*, 43, 1-7.
- Mehl, M. R., Vazire, S., Ramirez-Esparza, N., Slatcher, R. B., & Pennebaker, J. W. (2007). Are women really more talkative than men? *Science*, 317, 82.
- Mehl, M. R. & Robbins, M. L. (2012). Naturalistic observation sampling: The Electronically Activated Recorder (EAR). In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life*. New York, NY: Guilford Press
  
- Week 12: Project Update 2
  
- Week 13: Audio/Video Analysis and their applications – Automated Analysis
  - Scherer, S., Stratou, G., Gratch, J., & Morency, L. P. (2013). Investigating voice quality as a speaker-independent indicator of depression and ptsd. *Proceedings of Interspeech 2013*.
  - Scherer, S., Pestian, J., & Morency, L. P. (2013, May). Investigating the speech characteristics of suicidal adolescents. In *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on* (pp. 709-713). IEEE.
  - Nasir, M., Xia, W., Xiao, B., Baucom, B., Narayanan, S. S., & Georgiou, P. (2015). Still Together?: The Role of Acoustic Features in Predicting Marital Outcome. In *Proceedings of Interspeech*.
  - Gibson, J., Malandrakis, N., Romero, F., Atkins, D. C., & Narayanan, S. (2015) Predicting Therapist Empathy in Motivational Interviews using Language Features Inspired by Psycholinguistic Norms. In *Proceedings of Interspeech*.
  - Timmons, A. C., Chaspari, T., Han, S. C., Perrone, L., Narayanan, S. S., & Margolin, G. (2017). Using Multimodal Wearable Technology to Detect Conflict among Couples. *Computer*, 50(3), 50-59.
  - Wienberg, C. and Gordon, A. (2015) Insights on Privacy and Ethics from the Web's Most Prolific Storytellers. *Proceedings of ACM Web Science 2015*, June 28 - July 1, 2015, Oxford, UK.
  - Watch in class: Friends You Haven't Met Yet
  
- Week 14 & 15: Final Project Presentations

## Statement on Academic Conduct and Support Systems

### Academic Conduct

Plagiarism – presenting someone else’s ideas as your own, either verbatim or recast in your own words – is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in *SCampus* in Section 11, *Behavior Violating University Standards*<https://scampus.usc.edu/1100-behavior-violating-university-standards-and-appropriate-sanctions/>. Other forms of academic dishonesty are equally unacceptable. See additional information in *SCampus* and university policies on scientific misconduct, <http://policy.usc.edu/scientific-misconduct/>.

Discrimination, sexual assault, and harassment are not tolerated by the university. You are encouraged to report any incidents to the *Office of Equity and Diversity* <http://equity.usc.edu/> or to the *Department of Public Safety* <http://capsnet.usc.edu/department/department-public-safety/online-forms/contact-us>. This is important for the safety whole USC community. Another member of the university community – such as a friend, classmate, advisor, or faculty member – can help initiate the report, or can initiate the report on behalf of another person. *The Center for Women and Men* <http://www.usc.edu/student-affairs/cwm/> provides 24/7 confidential support, and the sexual assault resource center webpage [sarc@usc.edu](mailto:sarc@usc.edu) describes reporting options and other resources.

### Support Systems

A number of USC’s schools provide support for students who need help with scholarly writing. Check with your advisor or program staff to find out more. Students whose primary language is not English should check with the *American Language Institute* <http://dornsife.usc.edu/ali>, which sponsors courses and workshops specifically for international graduate students. *The Office of Disability Services and Programs* [http://sait.usc.edu/academicsupport/centerprograms/dsp/home\\_index.html](http://sait.usc.edu/academicsupport/centerprograms/dsp/home_index.html) provides certification for students with disabilities and helps arrange the relevant accommodations. If an officially declared emergency makes travel to campus infeasible, *USC Emergency Information* <http://emergency.usc.edu/> will provide safety and other updates, including ways in which instruction will be continued by means of blackboard, teleconferencing, and other technology.