Alicia L. Nobles

Summary

I am a systems engineer, trained in information science, data science, and public health, with 10 years of industry consulting and research experience. Currently, I apply natural language processing and machine learning to a variety of digital data sources to monitor human behavior and provide data-driven insights to inform interventions.

Education

PhD, Systems & Information Engineering   Dec 2018
University of Virginia, GPA: 3.90

MS, Civil Engineering   Aug 2014
University of Virginia, GPA: 3.93

BS, Civil Engineering, Summa Cum Laude   Dec 2007
Georgia Tech, GPA: 4.0

Select Skills

Data Analysis: machine learning (regression, classification, multilevel modeling, topic modeling), natural language processing, text mining, time series modeling (ARIMA), visualization, qualitative analysis

Programming & Software: 4+ years: Python (scikit-learn, gensim, pandas, numpy, spacy), R (ggplot2, forecast, lubridate, lme4, glmnet, caret) | 2+ years: GIS | 1 year: Matlab, SQL, Tableau

Data: structured/unstructured data: social media, text messages, search engine queries, electronic health records, surveys, geospatial data, cost data

Professional: systems engineering, interfacing with stakeholders (client, regulatory agencies, public), proposal development, developing findings and recommendations, presenting, deploying engineered systems

Select Research Experience

Fellow   2018 - Current
Division of Infectious Diseases & Global Public Health, Dept of Medicine, UC San Diego
Research Group: Steffanie Strathdee, John Ayers, Mark Dredze
- Developed a rule-based approach to label demographics of over 200k Reddit posts. Wrangled 500k comments to identify comments authored by a physician. Found that common disparities in access to healthcare typically present across gender, sex, racial, and ethnic minorities did not appear to be present in physician-patient interactions on Reddit.
- Interviewed by CBS and NPR to translate both my and others' research to the public.
- Identified a discrepancy of including helplines for celebrity opioid overdose reporting using rule-based NLP. Worked with major news media to include 1-800-662-HELP in their reporting of drug overdoses.

Visiting Scholar   Summer 2018
Center for Language & Speech Processing, Dept of Computer Science, Johns Hopkins University
Research Group: Mark Dredze
- First to combine computer vision with topic modeling to identify themes present in 50k Instagram images to inform HIV health communication.
- Identified unprecedented Google searches (86% higher; 40-54 million more) for sexual violence in response to #MeToo using an ARIMA model to manufacture a counterfactual scenario.

Graduate Research Fellow   2014 - 2018
Dept of Systems & Information Engineering, University of Virginia
Research Group: Laura Barnes

- Track record of pitching ideas resulting in over $300k of funding from 6 fellowships and publications with 22 multi-disciplinary collaborators during my PhD.
- First to identify digital markers of imminent suicide risk in text messages by building a supervised n-gram model from over 1+ million text messages of suicide attempters.
- Identified unprecedented Google searches (15% higher; 10-21 million more) for intrauterine devices correlated with discussion of US health reform using an ARIMA time series model identifying for the first time that discussion of health insurance policy by Congress influences widespread health behavior.

Professional Experience

Civil Engineer   2008 - 2012
Environmental Services, CH2M (now Jacobs)
- Managed large-scale, multi-million dollar groundwater remediation projects including design, permitting, staffing, report preparation, and public relations.
- Managed budget, staffing, and schedule for 44 simultaneous environmental investigation sites in US and Caribbean. Authored 44 reports in less than 3 weeks saving $20,000 of a $60,000 budget coordinating with personnel in 10+ offices. Authored proposal that was awarded $1.2 million for follow-up investigation.

Select Publications

1. **Nobles**, Leas, Althouse, Dredze, Longhurst, Smith, Ayers. (forthcoming). Requests for Diagnoses of Sexually Transmitted Diseases on a Social Media Platform. JAMA.
2. **Nobles**, Dredze, Ayers. 2019. "Repeal and Replace": Increased Demand for Intrauterine Devices Following the 2016 Presidential Election. Contraception.
3. Ayers, **Nobles**, Dredze. 2019. Media Trends for SAMHSA's 800-662-HELP Addiction Treatment Referral Services Following a Celebrity Overdose. JAMA Intern Med.
4. **Nobles**, Curtis, Ngo, Vardell, Holstege. 2018. Health Insurance Literacy of College Students: A Mixed Methods Analysis. J Am Coll Health.
5. Caputi, **Nobles**, Ayers. 2018. Internet Searches for Sexual Harassment and Assault, Reporting, and Training Following #MeToo. JAMA Intern Med.
6. **Nobles**, Glenn, Kowsari, Teachman, Barnes. 2018. Identification of Imminent Suicide Risk Among Young Adults using Text Messages. Proc of ACM Conference on Human Factors in Computing Systems (CHI).
7. **Nobles**, Vilankar, Wu, Barnes. 2015. Evaluation of Data Quality of Multisite Electronic Health Record Data for Secondary Analysis. Proc of IEEE Conference on Big Data.

Select Honors & Awards

Research: Altmetric Top 5% of All Research Outputs | UVA Presidential Fellowship In Data Science

Professional: New Faces of Engineering (nomination) | Beyond Performance from CH2M 4x

Teaching: All University Teaching Award (nomination) for instructor of linear statistical models

Relevant Training

Courses: big data in health, computational research, data mining, epidemiology, environmental epidemiology, GIS, US health policy & management, linear statistical models, machine learning, optimization, research ethics, stochastic systems, text mining

Workshops: Big Data Coursework for Computational Medicine (Weill Cornell Medicine) | Healthcare Data Analytics (Oregon Health & Science University)

NIH T32 Training Grants: Transdisciplinary Big Data Science | Substance Use, HIV, and Related Infections

Volunteer: graduate student representative for UVA student health insurance committee for 6 years