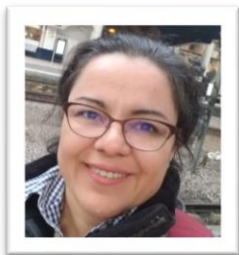




CS/Bioinformatics Seminar Myths and Misconceptions: Social Media Mining for Health Research



**Friday October 11, 2019, 2-3pm.
Chemistry and Computer Science Building 1.0202**

**Dr. Graciela Gonzalez-Hernandez
Perelman School of Medicine, Univ. of Pennsylvania**

Abstract: The total number of users of social media continues to grow worldwide, resulting in the generation of vast amounts of data. Popular social networking sites such as Facebook, Twitter and Instagram dominate this sphere. According to estimates, 500 million tweets and 4.3 billion Facebook messages are posted every day. A Pew Research Report on Social Media estimates that nearly half of adults worldwide and two-thirds of all American adults (65%) use social networking. The report states that of the total users, 26% have discussed health information, and, of those, 30% changed behavior based on this information and 42% discussed current medical conditions. Advances in automated data processing, machine learning and NLP present the possibility of utilizing this massive data source for biomedical and public health applications, if researchers adequately address the methodological challenges unique to this media. Despite numerous published studies, however, myths and misconceptions persist about the suitability and adequate use of these data, impacting the perception of researchers, institutional review boards, and the general public on the validity of the studies for health research. In this talk, we will discuss (and hopefully debunk!) some of the more poignant myths and misconceptions

Bio: Dr. Gonzalez Hernandez is a recognized expert and leader in natural language processing (NLP) applied to bioinformatics, medical/clinical informatics, and public-health informatics. She is an Associate Professor of Informatics in Biostatistics and Epidemiology at the University of Pennsylvania where she established the Health Language Processing Lab within the Institute of Biomedical Informatics. Her recent work focuses on NLP applications for public-health monitoring and surveillance and is funded by R01 grants from the National Library of Medicine and the National Institute of Allergy and Infectious Diseases. Her work on social media mining for pharmacovigilance has resulted in 25 publications in prestigious conferences and journals. Her work on enriching geospatial information for phylogeography uses NLP for the automatic extraction of relevant geospatial data from the literature and for linkage to GenBank records.

Contact: Natalia Villanueva Rosales (Computer Science), nvillanuevarosales@utep.edu