



BIG DATA

Interdisciplinary Data Sciences Consortium



* **IDSC Seminar Series** *

October 28, 2016 1:30pm-2:30pm

Location: CUTR 102

Presents

Dr. Loni Hagen,
School of Information
University of South Florida

Title: E-petition popularity: Do linguistic and semantic factors matter?

Abstract: E-petitioning technology platforms elicit the participation of citizens in the policy-making process but at the same time create large volumes of unstructured textual data that are difficult to analyze. Fortunately, computational tools can assist policy analysts in uncovering latent patterns from these large textual datasets. This study uses such computational tools to explore e-petitions, viewing them as persuasive texts with linguistic and semantic features that may be related to the popularity of petitions, as indexed by the number of signatures they attract. Using We the People website data, we analyzed linguistic features, such as extremity and repetition, and semantic features, such as named entities and topics, to determine whether and to what extent they are related to petition popularity. The results show that each block of variables independently explains statistically significant variation in signature accumulation, and that 1) language extremity is persistently and negatively associated with petition popularity, 2) petitions with many names tend not to become popular, and 3) petition popularity is associated with petitions that include topics familiar to the public or about important social events. We believe explorations along these lines will yield useful strategies to address the wicked problem of too much text data and to facilitate the enhancement of public participation in policy-making.



Biography: **Loni Hagen** is an Assistant Professor of Data Science at the School of Information at University of South Florida. She uses natural language processing and data mining techniques to automatically extract information from e-petition texts and further provides interpretation of them in context. Her current research interests include developing new ways of organizing opinions expressed by the public online, and introducing computational techniques to analyze big textual data for policy analysis and policy decision-making.

IDSC Contact:
Dr. K. Ramachandran
University of South Florida
4202 E Fowler Ave, CMC317
Tampa, FL 33620-5700
E-mail: ram@usf.edu
Telephone: (813)-974-1270
Fax: (813)-974-2700

Sponsors: USF Statistics Club