

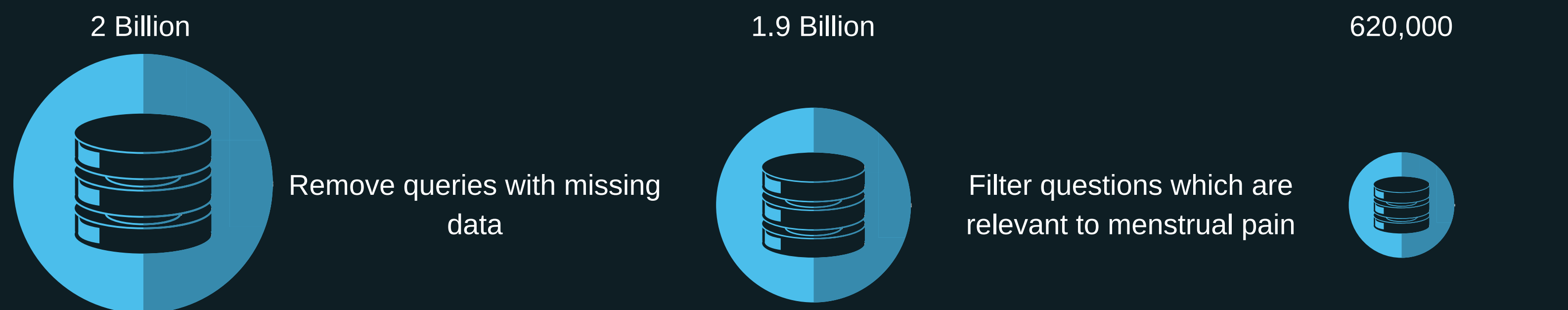
Text-Mining of User-Generated Queries on Menstrual Pain

What You Need to Know!

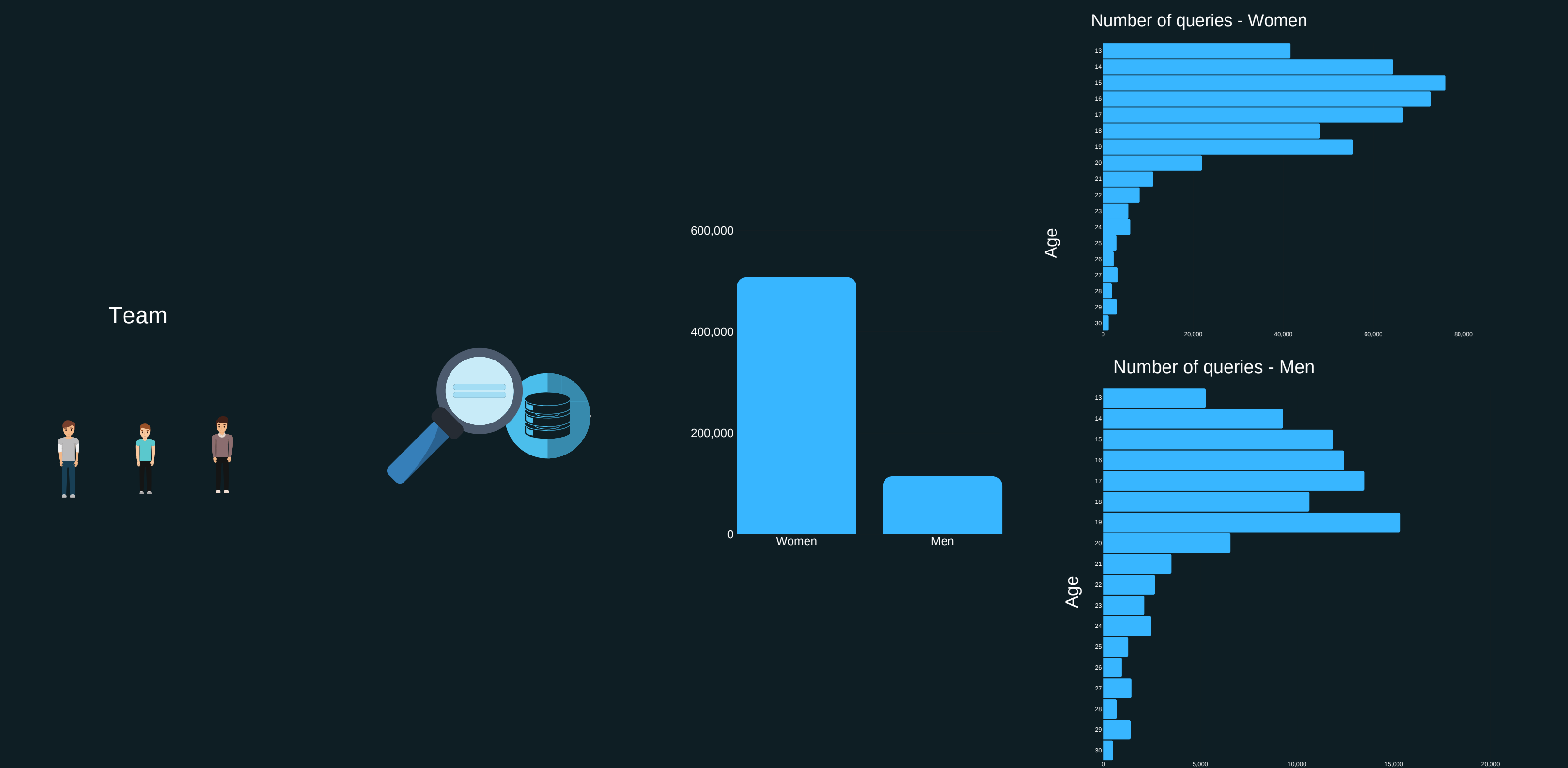
ChaCha was a question-answer service which provided answers to questions posed by users over phone, web or voice. Between Jan 2009 to Nov 2012, there were close to 2 Billion queries posed by people in the US. Menstrual pain is highly prevalent among women of reproductive age. By text-mining of user-generated queries, we try to understand the public's information needs and concerns related to menstrual pain. The information gained will be used to develop interventions to support menstrual pain management. Specific goals are to (1) visualize the text data; (2) cluster and categorize questions; (3) summarize patterns and themes



From the 2 billion questions, available 1.9 billion questions had complete information. These queries were from both females and males across a wide range of topics. From these 1.9 billion, only queries related to menstrual pain / dysmenorrhea were filtered. This resulted in a data set of 620,000 questions. There were 507,000 questions from females and 114,000 were from males.



While it makes sense to have large number of queries (500k) on menstrual pain from females, there seems to be a good number of queries from males (100k – 20%). Based on the number of queries raised by age, it seems that teenagers – both girls and boys – have the maximum number of queries related to the topic of menstrual cramps. The number of queries drops of significantly as people get out of the teenage years.



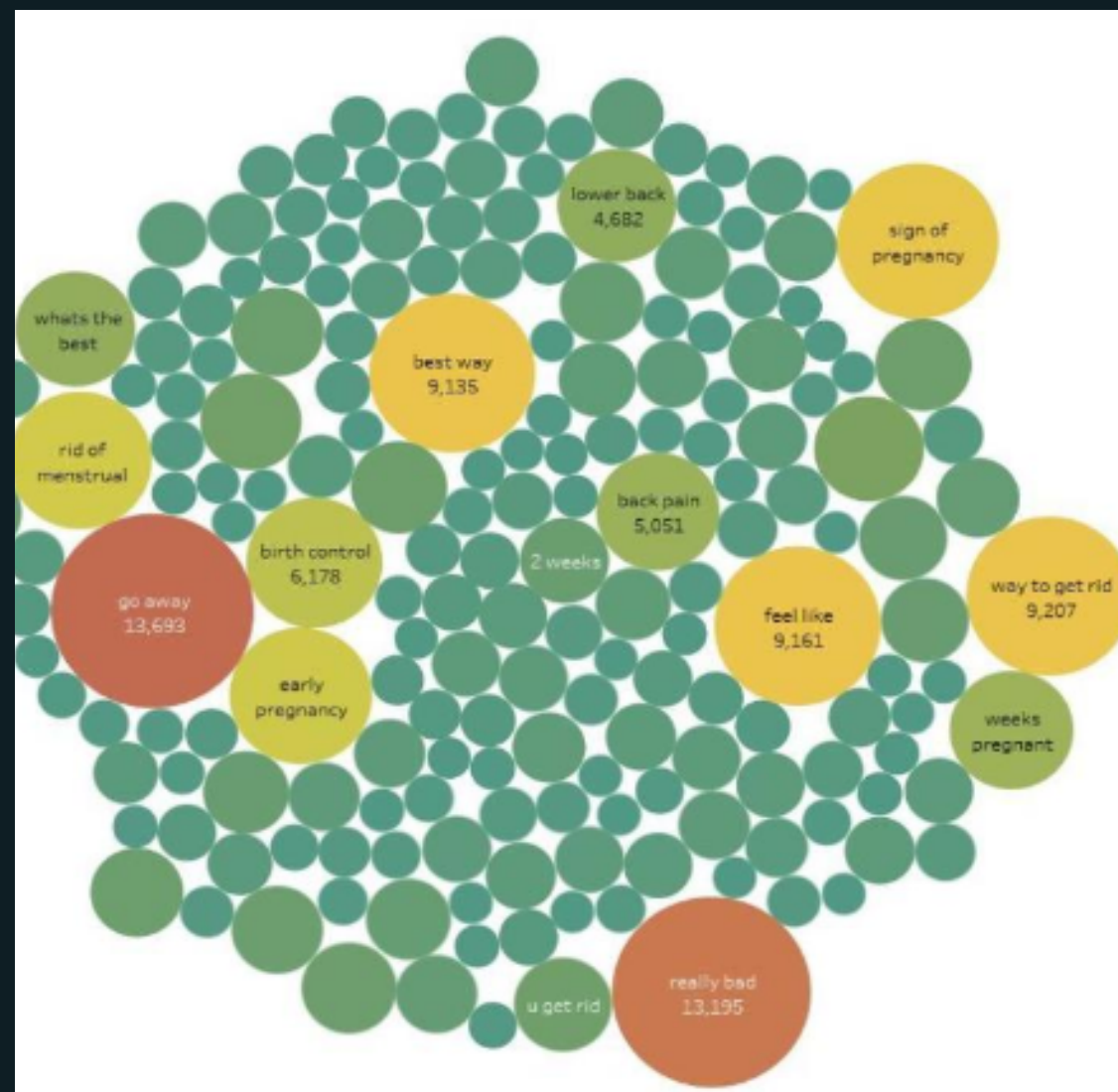
What You Need to Know!



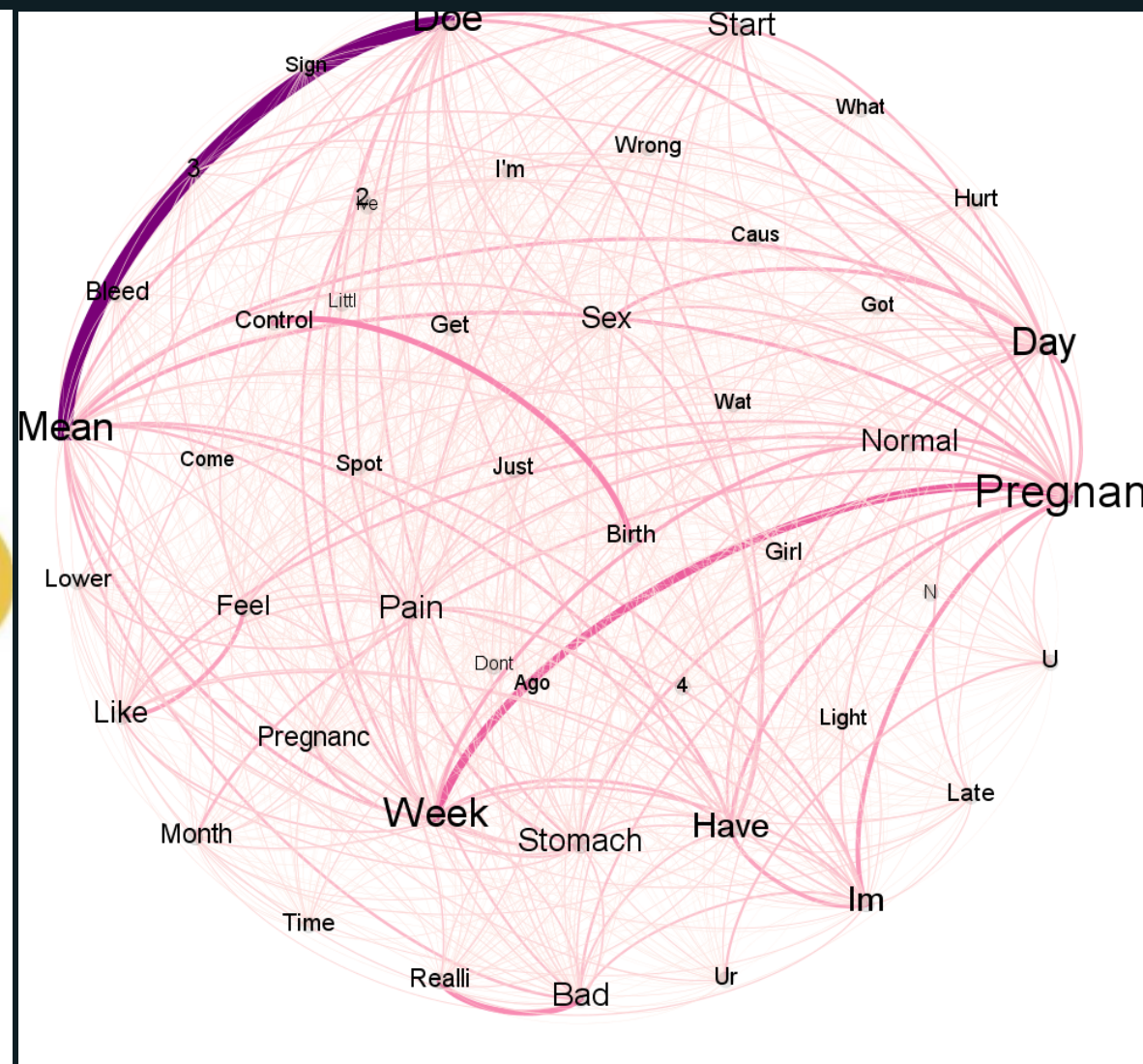
Word Cloud



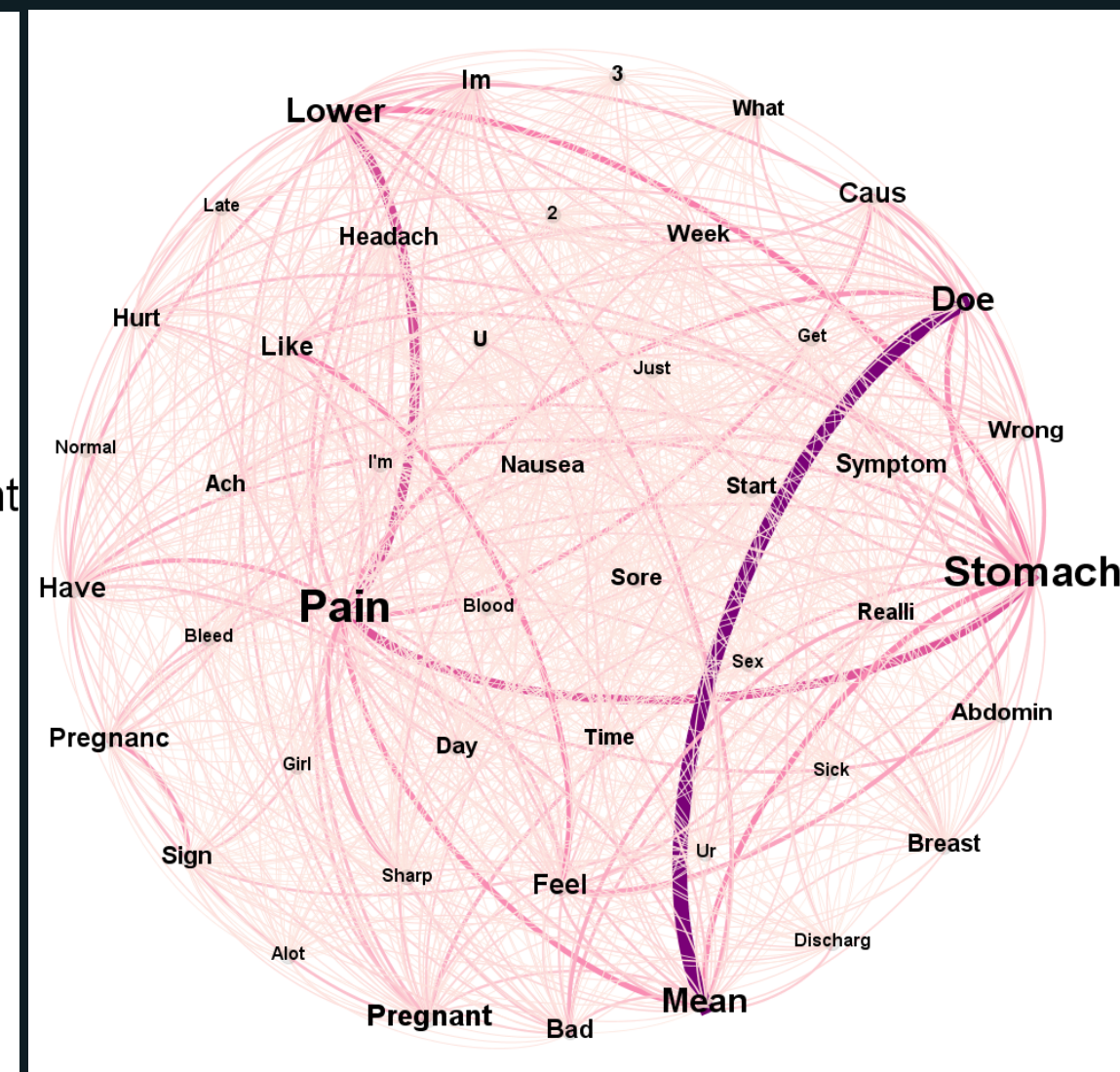
Phrase Count



Theme 1 - Pregnancy/Sex



Theme 1 - Symptoms



AUTHORS

NAIMESH CHAUDHARY, SIDDHARTH SHANKAR,
GAUTHAM NAGENDRA KAMATCHI

SPONSOR

Dr. CHEN X. CHEN

