



# Text Parsing and Text Search Demo

---

Dr. Goutam Chakraborty



# Case: Analyzing SAS Global Forum Abstracts

---

- Data: Abstracts of all papers published between 1976 and 2012 in SAS Global forums. Four data sets:
  - 1976-1980, 1981-1990, 1991-2000, 2000-2012
- Data Extraction and Cleaning: %TMFILTER macro used to prepare data from a repository of papers in .pdf and .rxt format (for details see paper titled “SAS since 1976: An Application of Text Mining to Reveal Trends” by Shaik, Garla, and Chakraborty)
- Primary objective of case: understand research topics represented in global forum abstracts from 1976-1980
- Primary goals of this demonstration:
  - Explore **1976-1980 data** to learn about terms in the data, search for specific terms, view concept links of terms, create synonym list by combining similar terms, and so on.



# Procedure

---

- Follow handout titled “Demo with SAS GF 1976-1980 abstracts”
- Note: I will show several features of SAS EM as we go through this demo. While the settings and process I will use here is specific to the SAS GF case study, you should be able to apply these settings (with appropriate modifications as needed) for your own project/case studies etc.