

# Twitter-Follower-Analyzer

Andrew Dougherty  
FRDCSA/POSI/OSBS/OI  
For FVCP August 10, 2009

# Categorize Users

- Twitter Follower Analyzer (TFA)
- Uses Twitter API to extract follower data
- Extracts information such as
  - Last 1000 posts
  - Number of:
    - Tweets
    - Followers
    - Retweets

# Categorize Users

- Classifies users into any appropriate categories from a **user-defined** set of categories.
- Sample Hierarchical Category Set:
  - People
    - Multi-Level Marketers, Programmers, Web Designers
  - Bots
  - Topics
    - Products, News, Real Estate, Adult, Other Topics
  - Types
    - Multi-Level Marketers, Retweeters, Spammers

# Training of Classifiers

- In order to perform this analysis it applies:
  - Text Clustering of tweets
  - WEKA Machine Learning over all extracted features:
    - Num of followers, num of tweets, cluster names
- In order to train/test, we must make a dataset of hand labelled followers
  - Developing GUI interface to with **user-defined** categories for browsing and labelling followers

# Actions on Groups of Followers

- Can perform operations on labelled groups, such as:
  - Blocking
  - Following
  - etc

# Potential Features

- Offered as a web service or stand-alone application
- Classification and extraction based on tweets (using POSI code) of Goals, Interests, and Abilities
- Estimation of services you can offer your followers (and other feedback)
- Analysis of users you are following
- Social Network Analysis

# Additional Notes

- The idea for Twitter Follower Analyzer came from Pete DuMelle and Eric Michalsen
- If you are in need of work, see Job-Search at <http://intranet.posithon.org>
- If you need software development services, please contact me (or take a flier):
  - [andrewdo@frdcusa.org](mailto:andrewdo@frdcusa.org)
  - (630) 300-5565
- Questions about the lightning talk?