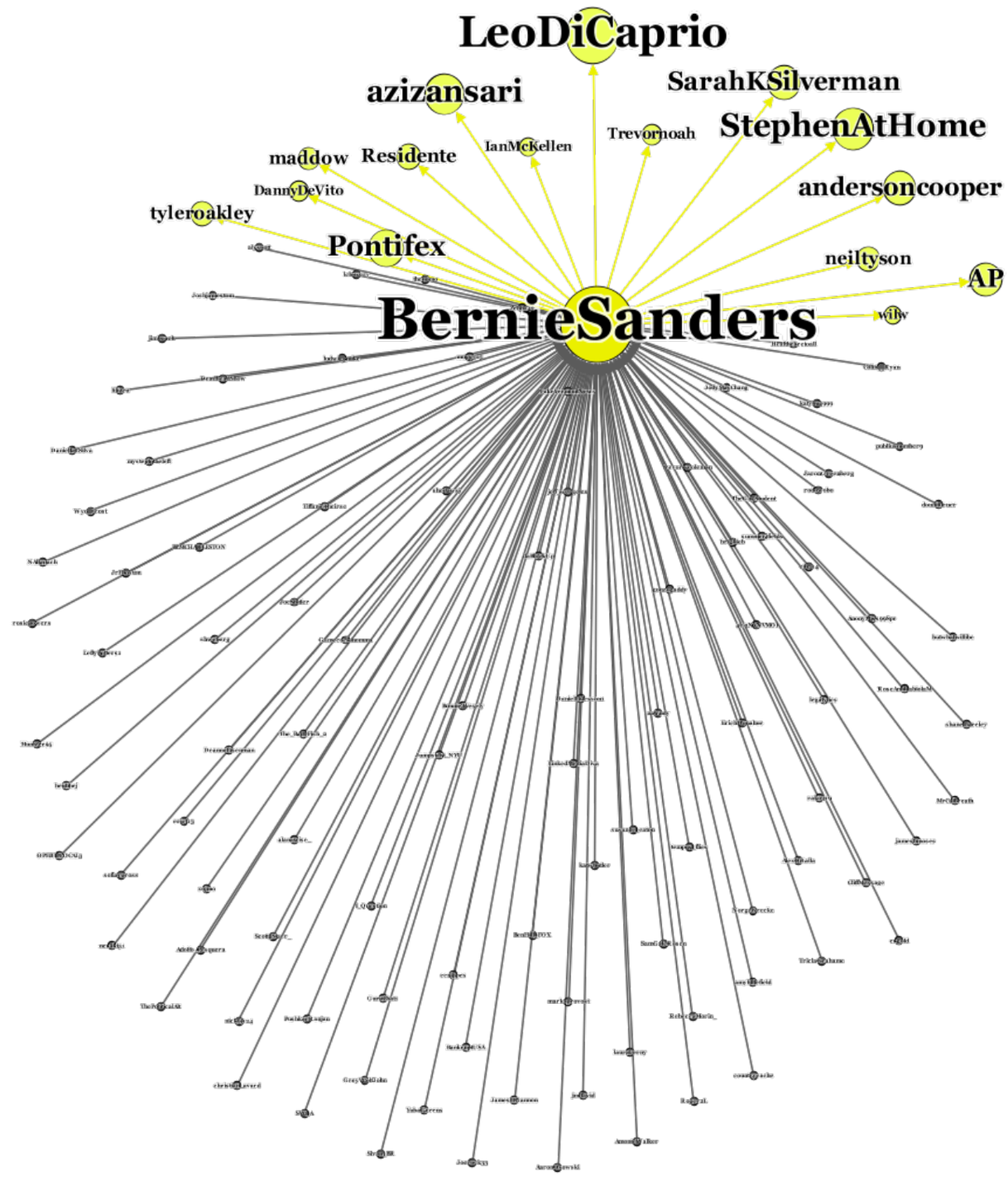
Follow the Bern

Kayla Andersen

Motivation

- Information Cascade
- Graphing relationships to better predict future sentiment and behavior
- Sanders as a use case: Are there influencers in the network that could sway Sanders' supporters to vote for her?

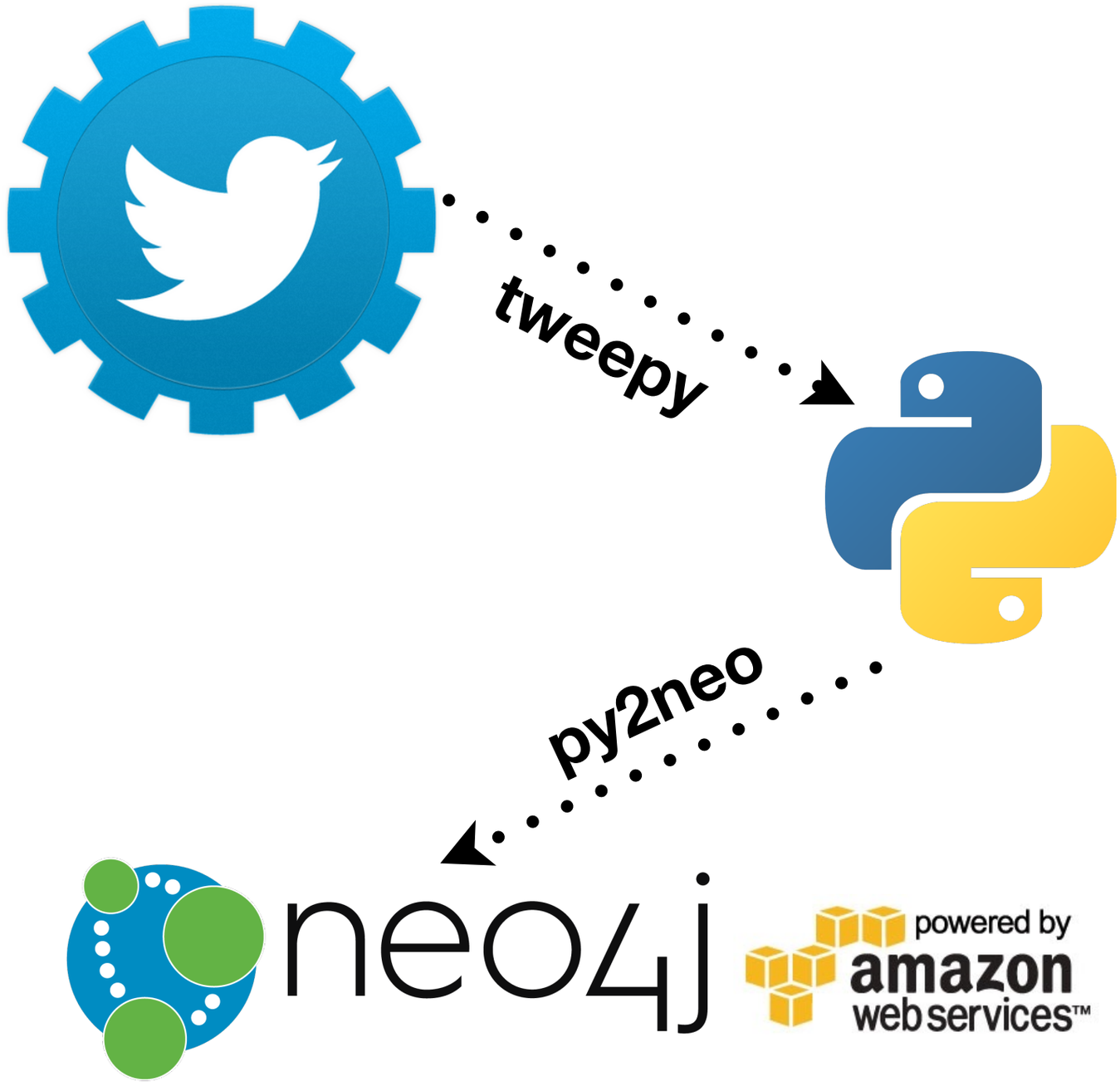


Process: Data Collection

- Tweepy library to pull user data and follower relationships from Twitter REST API

- Py2Neo OGM library to convert user->nodes and follows->edges

- Inserted into Neo4j graphing database hosted on AWS EC2

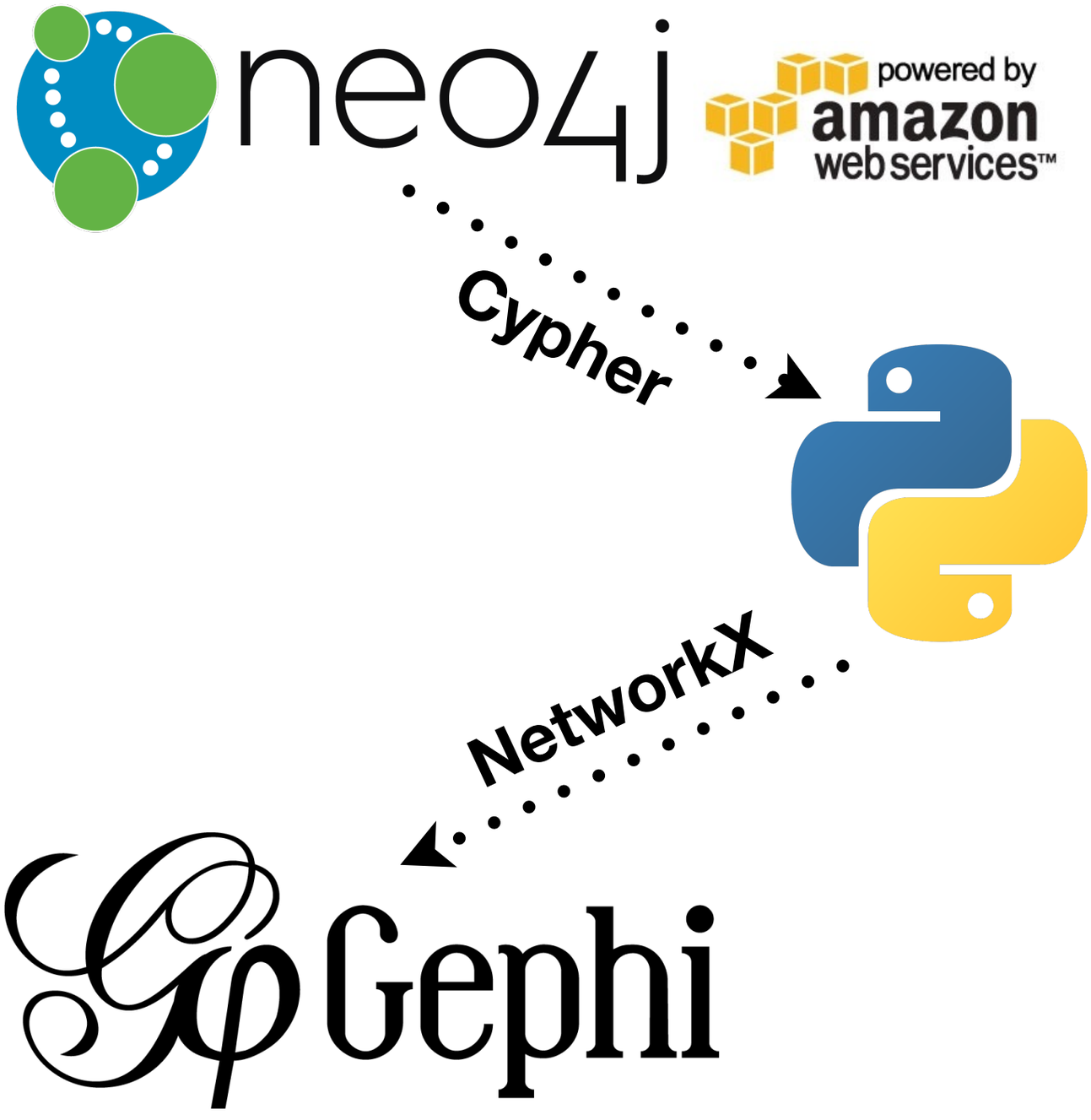


Process: Data Graphing

- Tweepy library to pull user data and follower relationships from Twitter REST API

- Py2Neo OGM library to convert user->nodes and follows->edges

- Inserted into Neo4j graphing database hosted on AWS EC2



What happens when everyone follows each other…

- Difficult to cluster
- A few groups of 200+, with many groups of 1
- Need weights to edges to make true neighborhoods

Moving Forward

- Use Twitter Stream API to track tweets in real time:
    - Track trend flow, use retweets and mentions to add weights to user relationships
- Sentiment analysis, unsupervised latent features of clusters

Questions?

(303) 842-2625

kaynandersen@gmail.com

in/kayla-n-andersen

kaylaandersen