Big data: identifying hotspots of radicalization in space and time

* Primary data sources
  + Twitter
    - Raw text
    - Location by source (explicit/derived)
    - Relations (followed/followers)
  + News media
    - Raw text
    - Geolocation (state/market region)
    - Political/economic affiliation by source
* Defining a “hotspot”
  + Content
    - Linguistic patterns – bigrams
  + Spatial
    - Physical space
    - Relational space
  + Temporal
    - Rate of change
* Detecting a hotspot
  + Change in bigram usage
    - New bigram emergence
    - Frequency change of existing bigram
  + Change in location of conversation
    - Relocation
    - Diffusion
    - Concentration
  + Rate of change in usage or location
    - Linear(ish)
    - Monotonic/exponential
    - Threshold effects
* Substantive goals
  + Identify early warning signs
  + Predict
    - Magnitude (size)
    - Duration (length)
    - Direction (population or region)