OSP Network Graphs Documentation

Data Extraction:

- Query using 03-top_texts_loc_inst.ipynb
- Change parameters (field, location, count)
- Set file name
- Run query (restart kernel and run all)

Gephi

- Load graphml file into Gephi (select undirected)
- Remove edge (bottom bar)
- Optional: change node color by rank, easier to work with layouts later
- Assign node size to 20 for all nodes
- OpenOrd layout
    - Num Iterations = 1900 – 2100 (typically 2000)
    - Edge cut = .8 – 1 (default is .8, this doesn't make a significant difference)
    - After OpenOrd is done running, move around clusters if necessary. You typically want to isolate the largest cluster (move the neighboring clusters away) because we don't want the clusters to mesh together when we expand with Noverlap
    - Make sure you aren't moving multiple clusters at the same time and that all the clusters stay intact
    - If shape of graph is not desirable, modify the random seed parameter (change the number, randomly it really doesn't matter) to get a different placed layout
- Noverlap layout
    - Run Noverlap with default conditions (the nodes will be very small)
    - Change node size to ranking by freq variable (range 50 – 200)
    - Move around clusters if necessary (noticeable crowding)
    - Run Noverlap again (change speed to something slower(1.5), just to track cluster crowding)
    - If clusters seem very tightly spaced, increase ratio parameter (~1.5 – 1.7)
- You may have to repeat OpenOrd and Noverlap many times before getting a desirable graph
    - You want the graph to roughly fit the shape of a square without a lot of white space (poster will be 4 x 4)
- Run Modularity function (right hand column titled "Statistics")
    - Keep at 1.0 or modify to a lower number (~.8) to get 4-5 groups (this is the most ideal for color visualizations)
- Change node coloring to illustrate modularity
    - Paint palette icon → partition → Modularity Class Attribute
    - The colors on the graph will now illustrate the modularity classes
- Color Gradation
    - The colors used (4 colors): http://www.paletton.com/#uid=72q0R0klZwHc9O2hCEEqCtWtFo4
    - If there is a 5th color, use the purple in: http://www.paletton.com/#uid=72p2i0koQKcebTejXOwubFTBVu5
    - I typically use the lightest shade, main shade, and darkest shade for the gradation
    - Go to Data Laboratory
    - Organize data by Modularity

    - Select all the entries in the first modularity class (class 0)
    - Right click → settle nodes → right click → copy data → new worksheet
    - Go to new worksheet and deselect edges
    - Go to color attribute → rank → freq
    - Double click arrows on color scale, change the hex values to the ones on Paletton (copy and paste)
    - Repeat for all the Modularity groups
    - One all modularity groups are colorized in different worksheets, merge them back together (highlight all the data in each worksheet → right click → move to → which ever worksheet you designate to contain all the final nodes)
    - Do this for all the groups until all the nodes are in one worksheet
- Add labels
    - Go to data laboratory
    - Merge columns (title, author) with comma delimiter, title first then author, name new column "title_auth"
    - Replace that Label column with new column (title_auth)
    - Go to Overview
    - Show labels, adjust size
    - Change label size attribute → ranking → freq
    - Range: 1 – 2.5
- Preview Window
    - View graph in preview layout (deselect show edges)
    - Remove node border, select show labels, deselect proportional size option
    - Font format: Avenir, 18, plain or bold
- Manually move around nodes with overlapping labels (done in Overview window), make sure you're not moving multiple nodes at once
- Export SVG

Adobe Illustrator

- 6000 x 6000 dimension (whichever dimension works best, as long as it is square)
- Place graph svg file
- Resize to reduce as much white space as possible
- Add title and logo on the bottom (font: ~100, Avenir bold)