

# Link detection in Political Networks



Fall 2018, CS 388 NLP final project

**Diego Garcia-Olano** & **Eric Holgate**

# DATA: Article Sets for 220 Texas congressional politicians collected from NYTimes, Houston Chronicle, etc..



WHO YOU ELECT .COM

TX 2015

## SEARCH RESULTS:

FOR "EDDIE RODRIGUEZ" FOUND:

362 ARTICLES FROM  
6 DIFFERENT SOURCES AND  
6226 ENTITIES DISCOVERED

[ LOCATION: 882, POLITICIAN: 270, MISC: 411, ORGANIZATION: 1945, PERSON:  
2631, BILL: 87]

OF WHICH 57 ARE CURRENTLY SHOWING

## FILTER BY DATE:

FROM

TO

APPLY DATE FILTER

## FILTER BY SOURCE:

AUSTIN AUSTIN STATESMAN (133)

DALLAS MORNING NEWS (22)

HOUSTON CHRONICLE (136)

NEW YORK TIMES (4)

TEXAS OBSERVER (6)

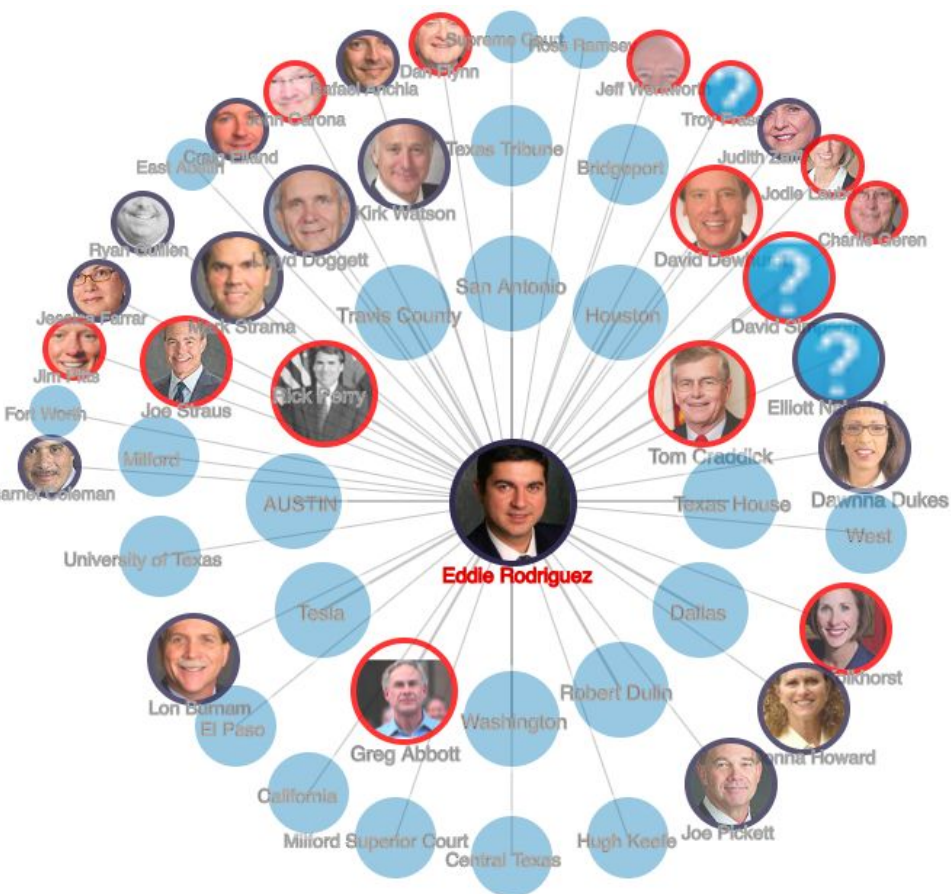
TEXAS TRIBUNE (60)

APPLY NEW SOURCE FILTER

+ **Named Entity Recognition over articles**

**to construct Political Networks.**

# DATA: sample Political Network and Article Distribution

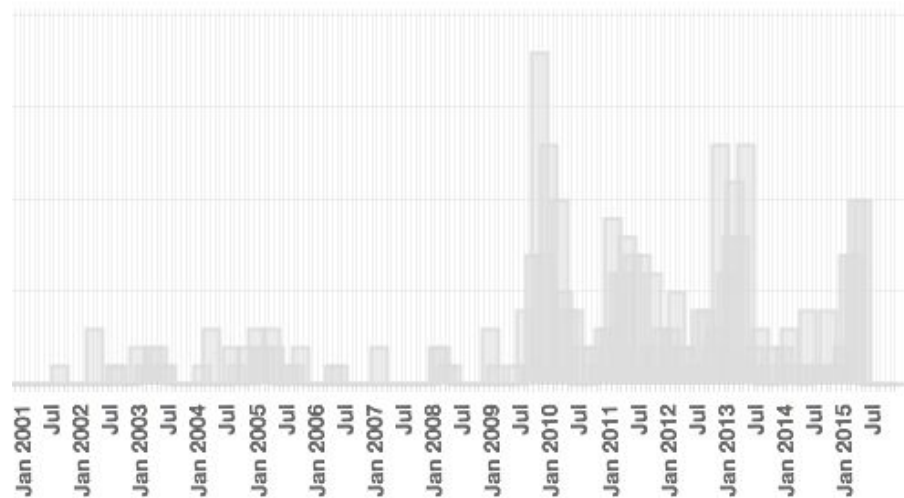


Eddie Rodriguez has most combined same, near, article occurrences with

- Politician **Tom Craddick** - (77)
- Person **Rick Perry** - (96)
- Organization **Houston** - (83)
- Location **AUSTIN** - (241)
- Bill **House Bill 506** - (3)
- Misc **Travis County** - (91)

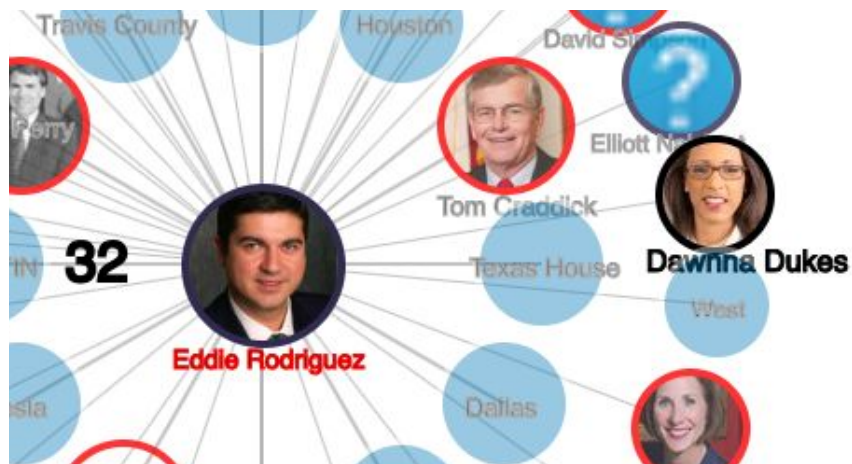
Eddie Rodriguez has most same sentence occurrences with

- Politician **Dawnna Dukes** - (17)
- Person **Hugh Keefe** - (6)
- Organization **Tesla** - (8)
- Location **AUSTIN** - (52)
- Bill **House Bill 506** - (2)
- Misc **Travis County** - (27)



# DATA:

## Link between 2 Politicians



Relationship between **Eddie Rodriguez** and **Dawnna Dukes**

32 co-occurrences found in 31 articles.

( [AAS: 15](#) | [DMN: 0](#) | [HC: 4](#) | [NYT: 0](#) | [TOB: 1](#) | [TXR: 11](#) | [ALL: 31](#) )



*Dawnna Dukes is a Democratic member of the Texas House of Representatives. First elected in 1994, Dukes represents District 46 (East Austin) within Travis County. In 2017, Dukes was indicted by a grand jury on 13 felony charges and two misdemeanor charges related to abuse of public office. Later that year, all criminal ... [wiki](#)*

[show map](#)

[Austin faces gathering storm over rates 1](#)

2012-06-09 | AUSTIN AUSTIN STATESMAN

*They are the six members of the Austin legislative delegation: State Senator Kirk Watson and Representatives, Elliott Naishtat, Donna Howard, **Eddie Rodriguez** and **Dawnna Dukes**, all Democrats, and Paul Workman, a Republican.*

[With filing over election season can begin](#)

2012-03-09 | AUSTIN AUSTIN STATESMAN

*Austin state Representatives, **Dawnna Dukes**, Donna Howard, Elliott Naishtat, **Eddie Rodriguez** and Mark Strama will all go unchallenged in the Democratic primary.*

[Many texas schoolchildren fail fitness evaluatio 1](#)

2011-04-14 | AUSTIN AUSTIN STATESMAN

*Austin-area state Representatives, Tim Kleinschmidt, **Dawnna Dukes**, **Eddie Rodriguez** and Lois Kolkhorst did not return calls or emails seeking comment.*

# TASK: LINK DETECTION

**Given:** Politician A



network + articles + meta

Politician B



network + articles + meta

**Predict:** who has more of a “link” or “affiliation” with Politician C



David Simpson

# TASK: LINK DETECTION

Pol A



Pol B



PolC



Pol A - Pol C = 38 co-occurrences

Pol B - Pol C = 10 co-occurrences

**So Pol A - Pol C**  
**is the correct prediction**  
**(ie, network gives us supervision )**

# TASK: LINK DETECTION

Pol A



Pol B



Pol C



Pol A - Pol C = 38 co-occurrences **correct prediction**

Championing a farm to table movement in texas.html

2012-09-16 | NEW YORK TIMES

**Simpson** and Mr. **Rodriguez** supported an unsuccessful bill last session to allow licensed farmers to sell unpasteurized milk at farmers' markets or approved locations.

**Rodriguez** has drafted a variety of ideas for the caucus to consider, like reducing barriers to tax exemptions for urban farms, allowing on-site processing of wild game that could be introduced into soup kitchens, and expanding the Cottage Food Law. "Right now, it's easier to get cigarettes than to get milk the way God made it," Representative **David Simpson**, a Republican from Longview and caucus member, said of restrictions on the sale of unpasteurized milk.

**Rodriguez** supported an unsuccessful bill last session to allow licensed farmers to sell unpasteurized milk at farmers' markets or approved locations. "I just support freedom," Mr. **Simpson** said, "and it's a shame that government is often in the way of farmers. "

American digest

2013-02-23 | AUSTIN AUSTIN STATESMAN

She and other members Representative **Eddie Rodriguez**, D-Austin, and tea party favorite Representative **David Simpson**, R-Longview are working on measures dealing with such issues as reducing taxes on small farms, making it legal for small producers to offer samples at farmers markets and limiting fees that farmers can be charged at those markets.

Caucus offers several small farm bills

2013-02-16 | AUSTIN AUSTIN STATESMAN

State Representative **Eddie Rodriguez**, D-Austin and chairman of the Legislature's new Farm to Table Caucus, has been working with the caucus vice chair, Representative Lois Kolkhorst, R-Brenham, and other members, including tea party sweetheart Representative **David Simpson**, R-Longview, on an assortment of food-related bills.

## **INITIAL RESULTS:** BASELINES over 125k test set

### **1. NLP only baseline = 63 % accuracy**

Each politician is represented as a normalized sum of pre-trained GloVe embeddings of each of their articles. Predict based on greater Cosine Similarity of Pol A and Pol C vs Pol B and Pol C.

### **2. Network Analysis = 72 % accuracy**

Each politician is represented as the set of their immediate neighbors. Predict whoever has greater set intersection with Pol C



## **INITIAL RESULTS:** BASELINES over 125k test set

### **3. Graph Embeddings - *Ongoing***

- Compare Factorization, Random Walk & Deep Learning methods as mentioned in *Goyal, Ferrara 2017*
- Poor initial results across the board with “one-hop” nets.
- Simple baseline based on full combined large networks with 100 mil edges to get a political adjacency matrix where each row is its representation = **65.9 % accuracy**
- Graph embedding methods on combined network with more extensive grid search over hyper params.

## NEXT STEPS:

### Combine and Conquer ( hopefully )



- Compare pretrained GloVe representations to doc2vec learned embeddings
- Use concatenation of NLP rep and Graph Emb rep as input to logistic regression or neural model to jointly learn task.
- (future) Voting prediction as sequence model task
- (future) Explainability of predictions