**Why political actor representation learning?**

Analyzing political text with language models and NLP models has become a thriving research field at the intersection of natural language processing and computational political science.

These political texts, such as political op-eds, legislator statements, and political news articles frequently mention various political actors. While the names of political actors are just tokens to a language model, real-world political analysis often goes beyond tokens and requires much more domain knowledge, such as the referenced legislator’s party affiliation, elected office, voting records, ideological position, and more.

However, language models often could not achieve such fine-grained understanding of external knowledge on their own, hence the need for political actor representation learning to improve political text analysis.

**Why social context and expert knowledge?**

While existing works on political actor analysis achieved relative success, they fall short of incorporating the social context and expert knowledge that are integral to the understanding of political actors.

**Social context information** such as home state and party affiliation helps connect, compare, and contrast different political actors.

**Expert knowledge** from political think tanks helps to guide and ground the political actor representation process in the political reality.

---

### PAR Methodology

PAR follows a three-step approach to learn and use legislator representations:

1. Construct **heterogeneous information networks** to jointly represent political actors and the affiliated social context.
2. Learn legislator representation with **three training objectives**: expert knowledge, stance consistency, and echo chamber.
3. Apply PAR representations to political text analysis tasks as **features**

---

### Analyzing Political Text with PAR representations

Instead of knowledge base embeddings and ideal point vectors, we use PAR as features to represent political actors in two political text analysis tasks.

**PAR and Blue/Red/Swing States**

PAR learns state ideological preferences that correlate well with the 2020 election.

**Finding 1:** certain traditional swing states are no longer as competitive? (PA, NC)

**Finding 2:** Georgia is the most electorally competitive state in the United States?

---

**PAR Representation Learning**

PAR learns political actor representations that correlate well with various socio-political factors, such as home state, congressional caucuses, and voting records.


**Code:** [https://github.com/BunsenFeng/PAR](https://github.com/BunsenFeng/PAR)