Stirling Diversity

- Variety
- Disparity
- Balance

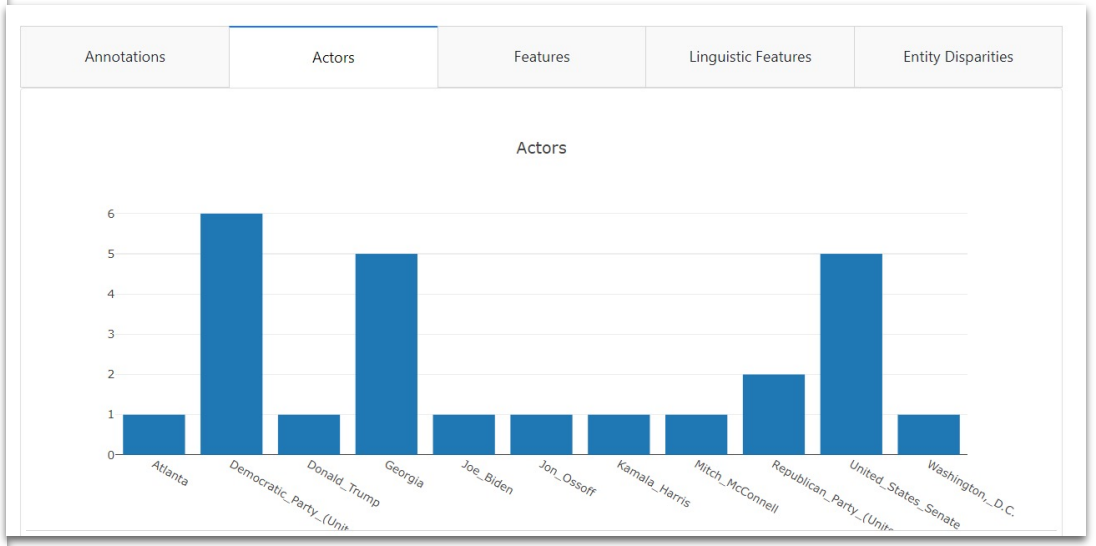
$$\Delta = \sum_{i \neq j} (d_{ij})^{\alpha} (p_i p_j)^{\beta}$$

Andy Stirling. 2007. A general framework for analysing diversity in science, technology and society. J. R. Soc. Interface (2007) 4, 707–719. DOI:10.1098/rsif.2007.0213

Diversity Searcher

use cases:

- For journalists (during text production)
- For editors etc. (“analytics” of their medium + others)
- For stakeholder groups (media monitoring)
- For news readers (critical reading, analysis)



diversitysearcher.net:4000
Logins available on request:
oguzozgur.karadeniz@kuleuven.be

How to determine how “diverse” (well, “disparate”) two entities are?

- Some common approaches
  - A matrix of pairwise disparity values
  - Path length in an ontology
  - Feature-based

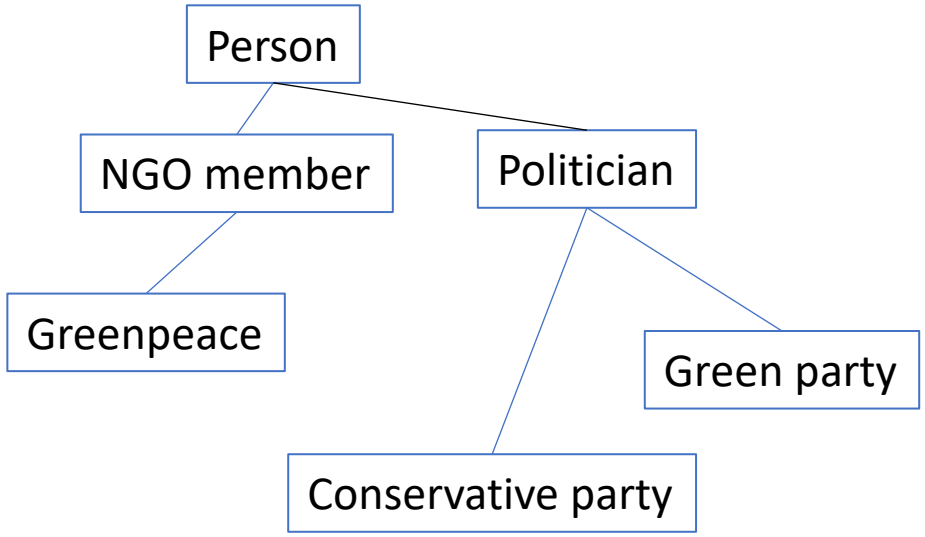
How to determine how “diverse” (well, “disparate”) two entities are?

- Some common approaches
  - A matrix of pairwise disparity values *)
  - Path length in an ontology
  - Feature-based

*) more about this below

How to determine how “diverse” (well, “disparate”) two entities are?

- Some common approaches
  - A matrix of pairwise disparity values
  - Path length in an ontology
  - Feature-based



How to determine how “diverse” (well, “disparate”) two entities are?

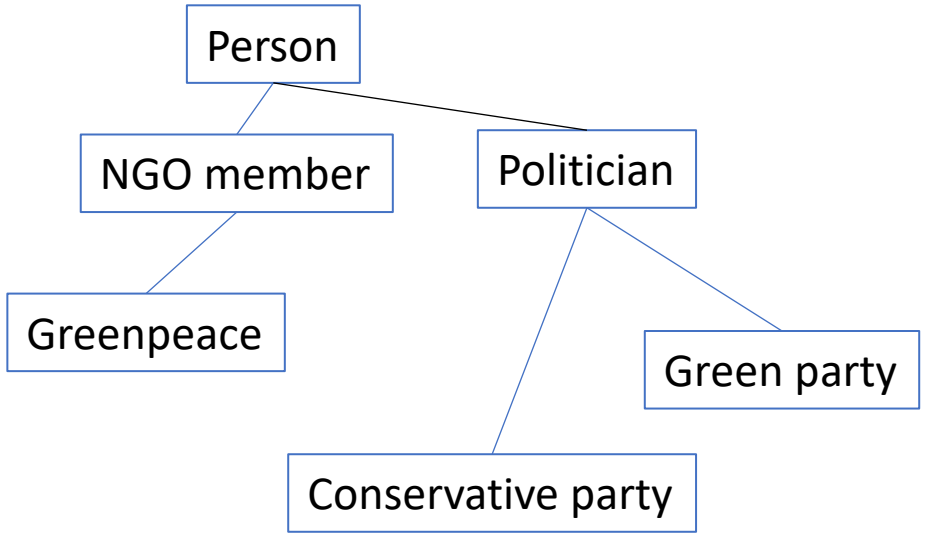
- Some common approaches
  - A matrix of pairwise disparity values
  - Path length in an ontology
  - Feature-based *)

*) more about this below

How to determine how “diverse” (well, “disparate”) two entities are?

- Some common approaches
  - A matrix of pairwise disparity values
  - Path length in an ontology
  - Feature-based

- What could be pros and cons of these approaches?



Diversity Searcher Pipeline

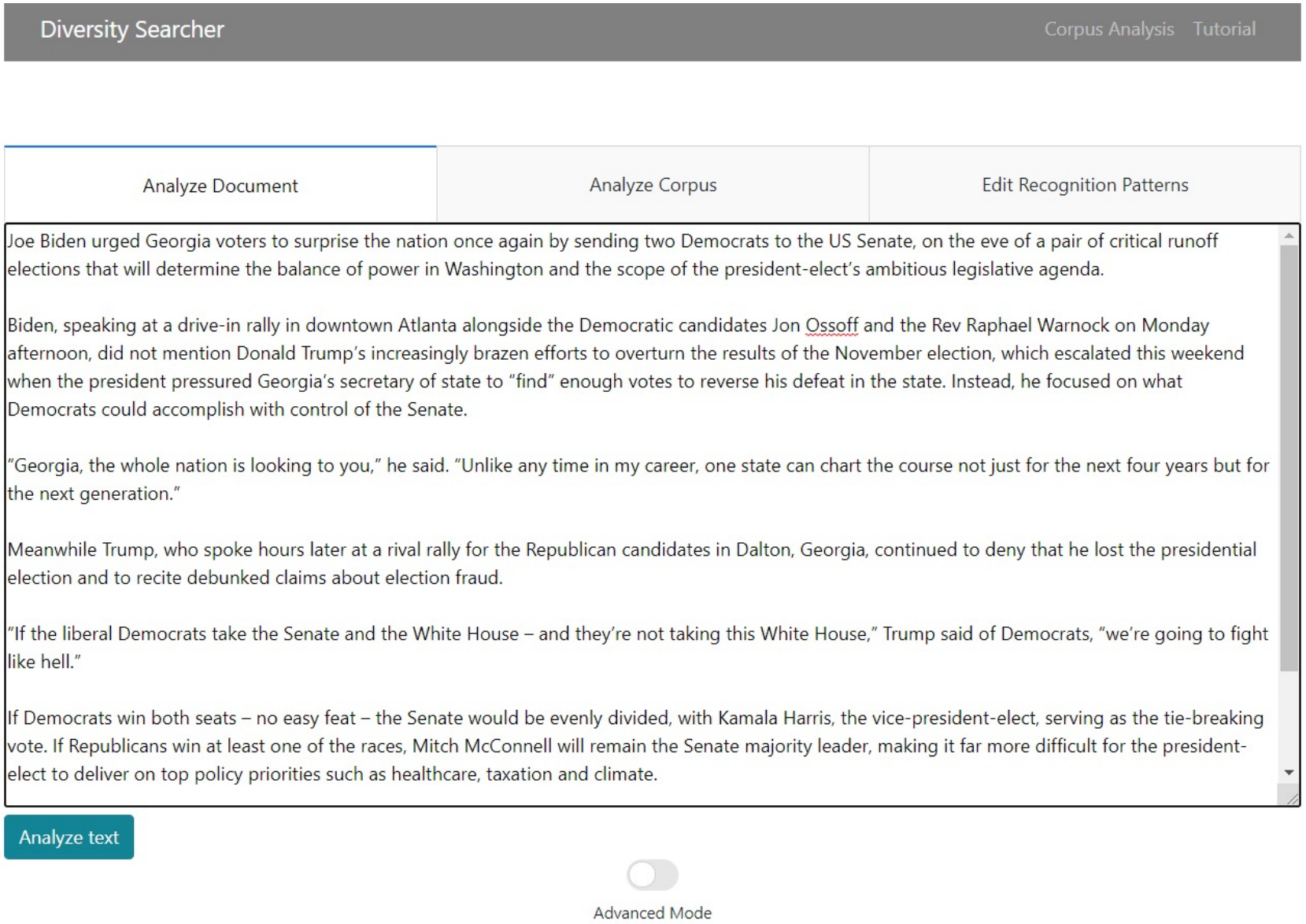
Text / Corpus

Actor Detection

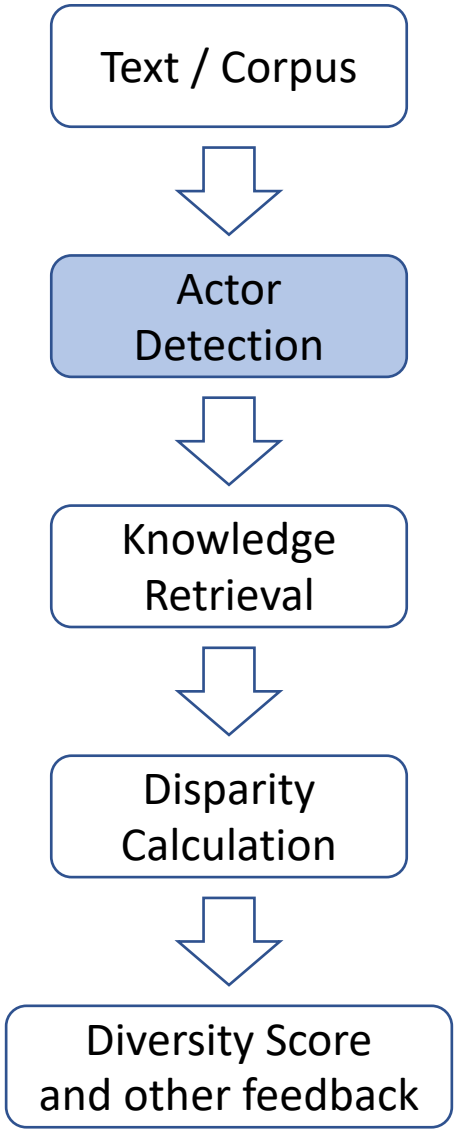
Knowledge Retrieval

Disparity Calculation

Diversity Score and other feedback



Diversity Searcher Pipeline



DBpedia Spotlight

spaCy

Annotations | Actors | Features | Linguistic Features | Entity Disparities

Joe Biden PERSON urged Georgia GPE voters to surprise the nation once again by sending two Democrats ORG to the US Senate ORG on the eve of a pair of critical runoff elections that will determine the balance of power in Washington GPE and the scope of the president-elect's ambitious legislative agenda.

Biden, speaking at a drive-in rally in downtown Atlanta GPE alongside the Democratic ORG candidates Jon Ossoff PERSON and the Rev Raphael Warnock on Monday afternoon, did not mention Donald Trump PERSON s increasingly brazen efforts to overturn the results of the November election, which escalated this weekend when the president pressured Georgia GPE s secretary of state to "find" enough votes to reverse his defeat in the state. Instead, he focused on what Democrats ORG could accomplish with control of the Senate ORG

" Georgia GPE the whole nation is looking to you," he said. "Unlike any time in my career, one state can chart the course not just for the next four years but for the next generation."

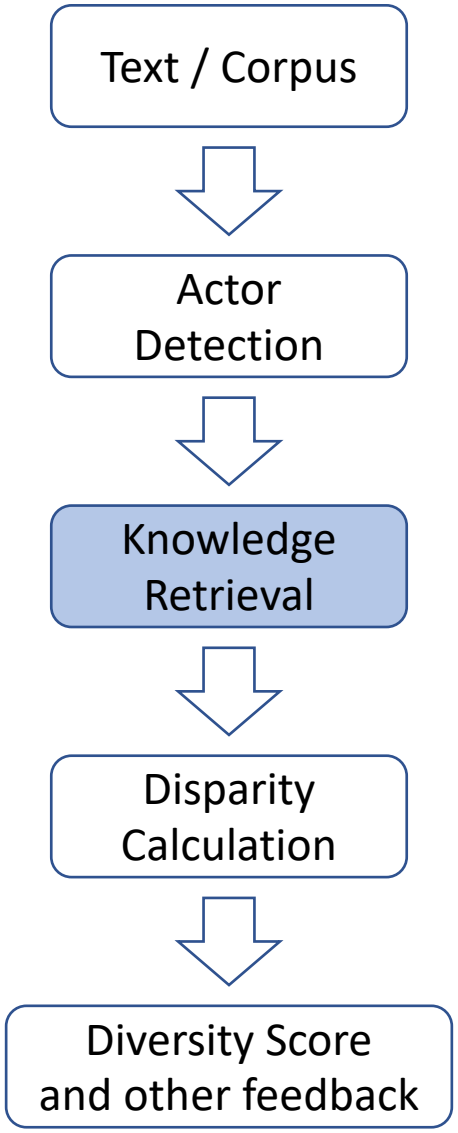
Meanwhile Trump, who spoke hours later at a rival rally for the Republican ORG candidates in Dalton, Georgia GPE continued to deny that he lost the presidential election and to recite debunked claims about election fraud.

"If the liberal Democrats ORG take the Senate ORG and the White House – and they're not taking this White House," Trump said of Democrats ORG "we're going to fight like hell."

If Democrats ORG win both seats – no easy feat – the Senate ORG would be evenly divided, with Kamala Harris PERSON the vice-president-elect, serving as the tie-breaking vote. If Republicans ORG win at least one of the races, Mitch McConnell PERSON will remain the Senate ORG majority leader, making it far more difficult for the president-elect to deliver on top policy priorities such as healthcare, taxation and climate.

Three million Georgia GPE voters cast ballots during the early voting period, which ended on Thursday – a record for runoff elections in the state. Nearly half a billion dollars has been spent on the twin races, as residents are bombarded with political ads and messaging urging them to vote in Tuesday's elections.

Diversity Searcher Pipeline



Annotations | Actors | Features | Linguistic Features | Entity Disparities

Joe Biden PERSON urged Georgia GPE voters to surprise the nation once again by sending two Democrats ORG to the US Senate ORG on the eve of a pair of critical runoff elections that will determine the balance of power in Washington GPE and the scope of the president-elect's ambitious legislative agenda.

Biden, speaking at a drive-in rally in downtown Atlanta GPE alongside the Democratic ORG candidates Jon Ossoff PERSON and the Rev Raphael Warnock on Monday afternoon, did not mention Donald Trump PERSON s increasingly brazen efforts to overturn the results of the November election, which escalated this weekend when the president pressured Georgia GPE s secretary of state to "find" enough votes to reverse his defeat in the state. Instead, he focused on what Democrats ORG could accomplish with control of the Senate ORG

" Georgia GPE the whole nation is looking to you," he said. "Unlike any time in my career, one state can chart the course not just for the next four years but for the next generation."

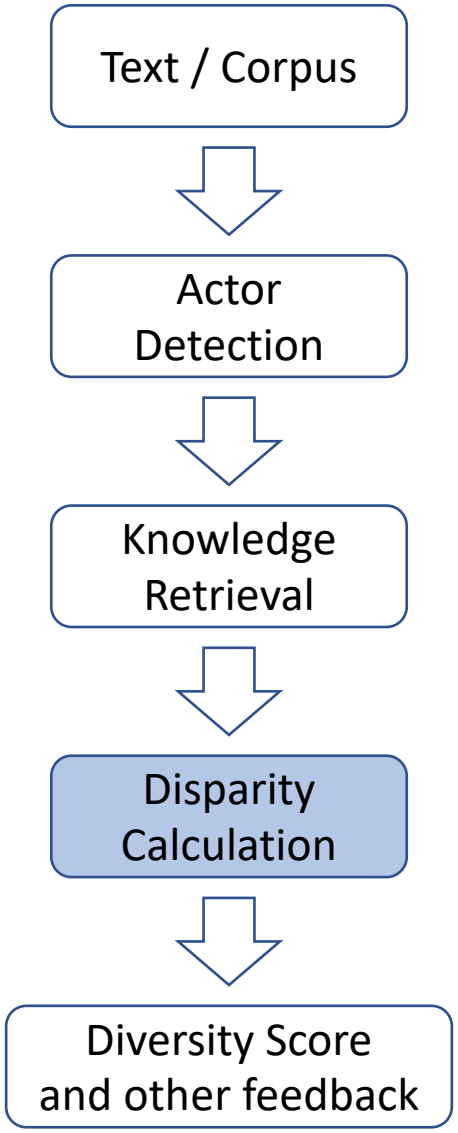
Meanwhile Trump, who spoke hours later at a rival rally for the Republican ORG candidates in Dalton, Georgia GPE continued to deny that he lost the presidential election and to recite debunked claims about election fraud.

"If the liberal Democrats ORG take the Senate ORG and the White House – and they're not taking this White House," Trump said of Democrats ORG "we're going to fight like hell."

If Democrats ORG win both seats – no easy feat – the Senate ORG would be evenly divided, with Kamala Harris PERSON the vice-president-elect, serving as the tie-breaking vote. If Republicans ORG win at least one of the races, Mitch McConnell PERSON will remain the Senate ORG majority leader, making it far more difficult for the president-elect to deliver on top policy priorities such as healthcare, taxation and climate.

Three million Georgia GPE voters cast ballots during the early voting period, which ended on Thursday – a record for runoff elections in the state. Nearly half a billion dollars has been spent on the twin races, as residents are bombarded with political ads and messaging urging them to vote in Tuesday's elections.

Diversity Searcher Pipeline



Annotations | Actors | Features | Linguistic Features | Entity Disparities

Joe Biden PERSON urged Georgia GPE voters to surprise the nation once again by sending two Democrats ORG to the US Senate ORG on the eve of a pair of critical runoff elections that will determine the balance of power in Washington GPE and the scope of the president-elect's ambitious legislative agenda.

Biden, speaking at a drive-in rally in downtown Atlanta GPE alongside the Democratic ORG candidates Jon Ossoff PERSON and the Rev Raphael Warnock on Monday afternoon, did not mention Donald Trump PERSON s increasingly brazen efforts to overturn the results of the November election, which escalated this weekend when the president pressured Georgia GPE s secretary of state to "find" enough votes to reverse his defeat in the state. Instead, he focused on what Democrats ORG could accomplish with control of the Senate ORG

" Georgia GPE the whole nation is looking to you," he said. "Unlike any time in my career, one state can chart the course not just for the next four years but for the next generation."

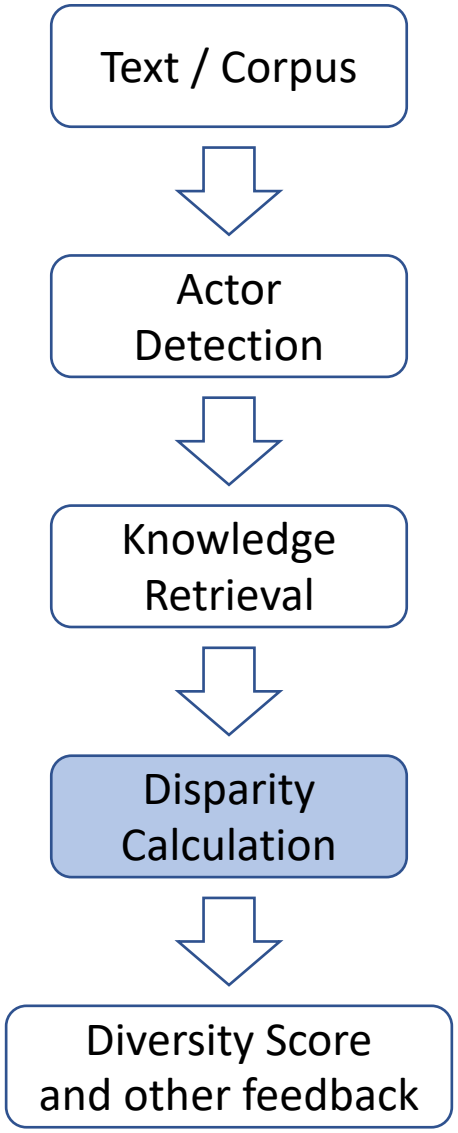
Meanwhile Trump, who spoke hours later at a rival rally for the Republican ORG candidates in Dalton, Georgia GPE continued to deny that he lost the presidential election and to recite debunked claims about election fraud.

"If the liberal Democrats ORG take the Senate ORG and the White House – and they're not taking this White House," Trump said of Democrats ORG "we're going to fight like hell."

If Democrats ORG win both seats – no easy feat – the Senate ORG would be evenly divided, with Kamala Harris PERSON the vice-president-elect, serving as the tie-breaking vote. If Republicans ORG win at least one of the races, Mitch McConnell PERSON will remain the Senate ORG majority leader, making it far more difficult for the president-elect to deliver on top policy priorities such as healthcare, taxation and climate.

Three million Georgia GPE voters cast ballots during the early voting period, which ended on Thursday – a record for runoff elections in the state. Nearly half a billion dollars has been spent on the twin races, as residents are bombarded with political ads and messaging urging them to vote in Tuesday's elections.

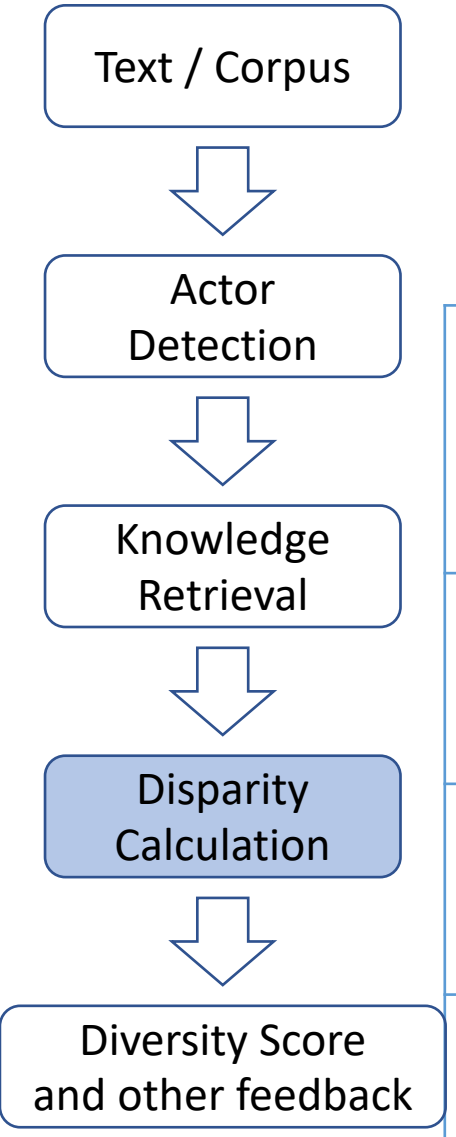
Diversity Searcher Pipeline

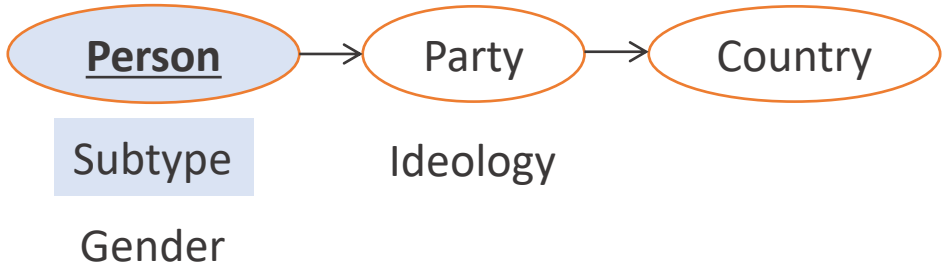


Example:

- Donald Trump
- Hillary Clinton
- Tayyip Erdoğan

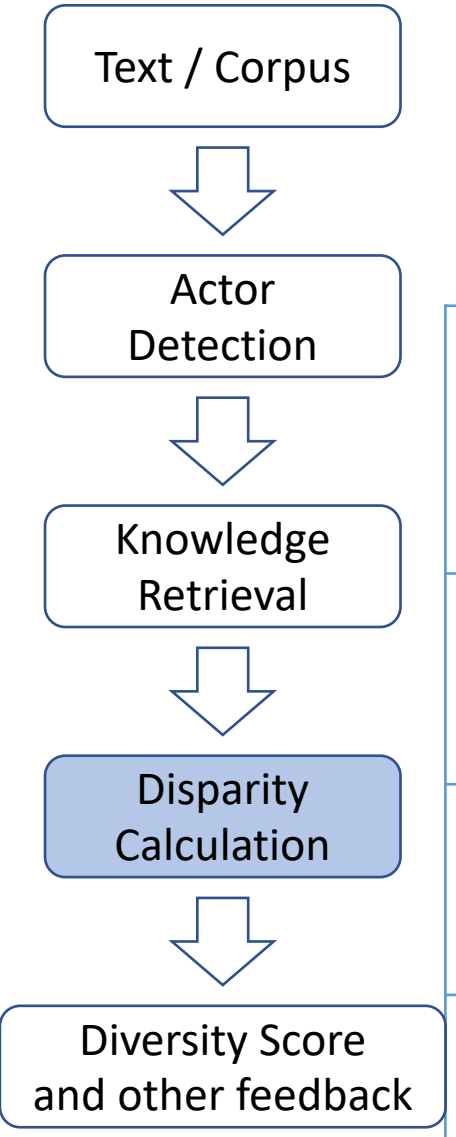
Diversity Searcher Pipeline

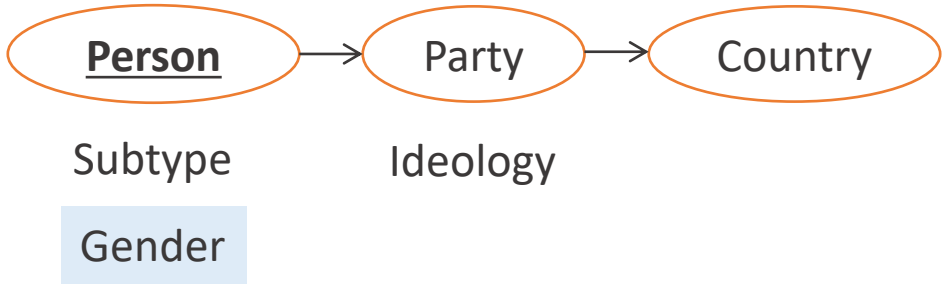




	Type	Subtype	Gender	Party	Party Ideology	Country
Donald Trump	Person	Politician	Male	Republican Party	Conservative	US
Hillary Clinton	Person	Politician	Female	Democratic Party	Liberal	US
Tayyip Erdogan	Person	Politician	Male	AKP	Conservative	Turkey

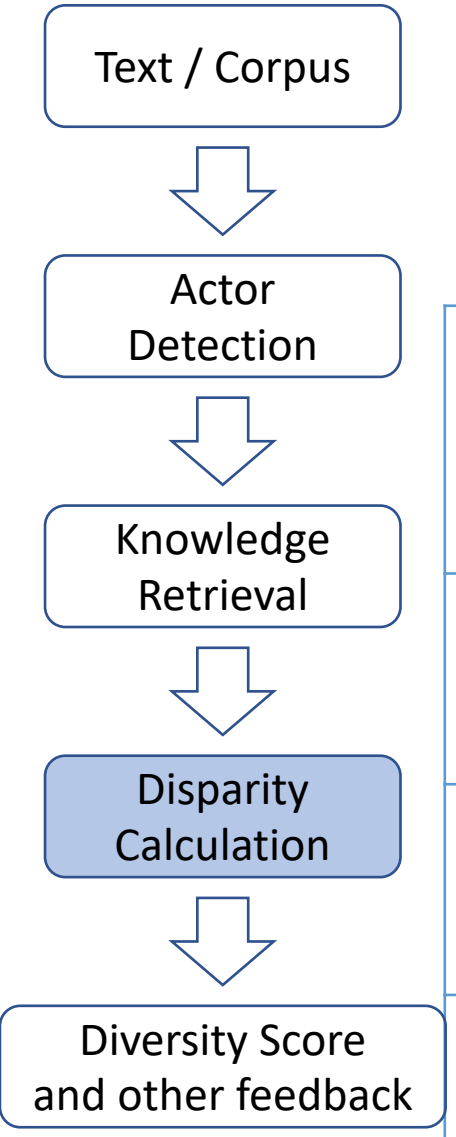
Diversity Searcher Pipeline

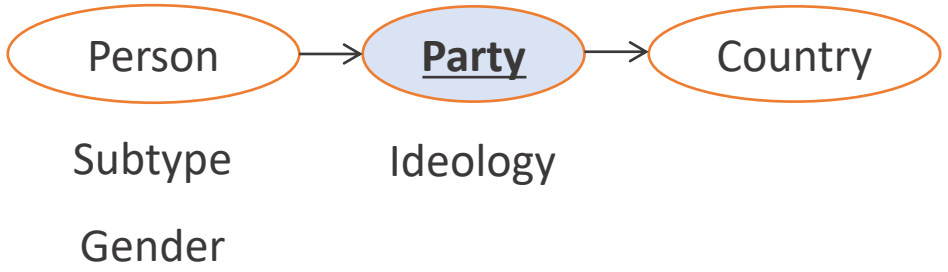




	Type	Subtype	Gender	Party	Party Ideology	Country
Donald Trump	Person	Politician	Male	Republican Party	Conservative	US
Hillary Clinton	Person	Politician	Female	Democratic Party	Liberal	US
Tayyip Erdogan	Person	Politician	Male	AKP	Conservative	Turkey

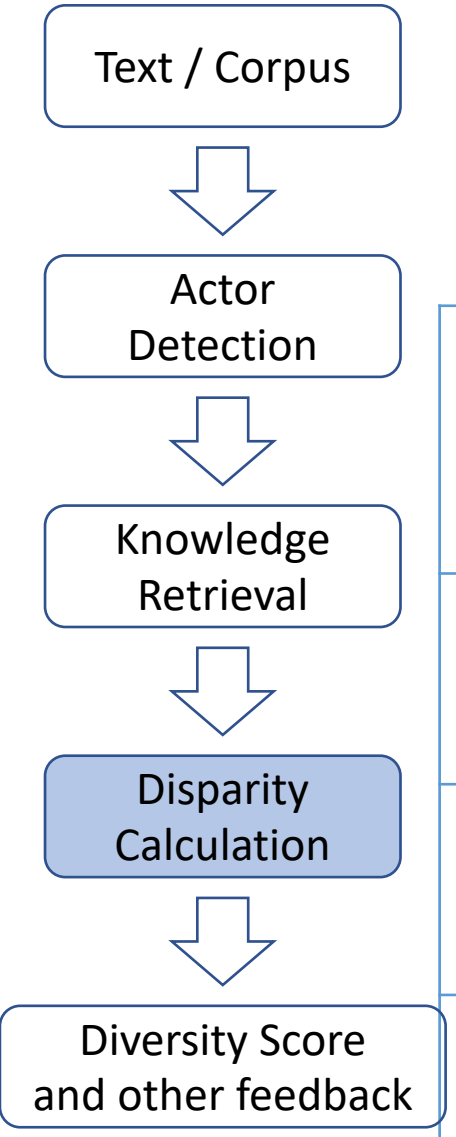
Diversity Searcher Pipeline

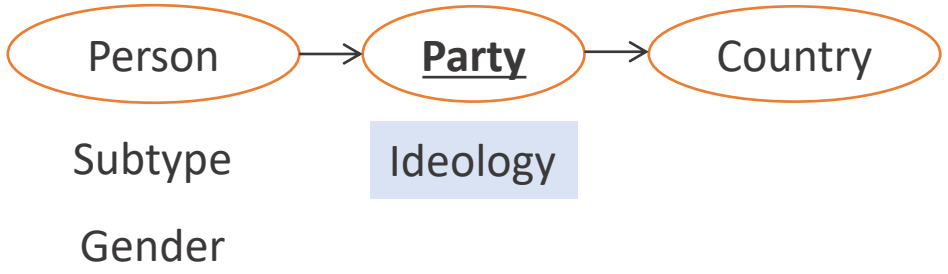




	Type	Subtype	Gender	Party	Party Ideology	Country
Donald Trump	Person	Politician	Male	Republican Party	Conservative	US
Hillary Clinton	Person	Politician	Female	Democratic Party	Liberal	US
Tayyip Erdogan	Person	Politician	Male	AKP	Conservative	Turkey

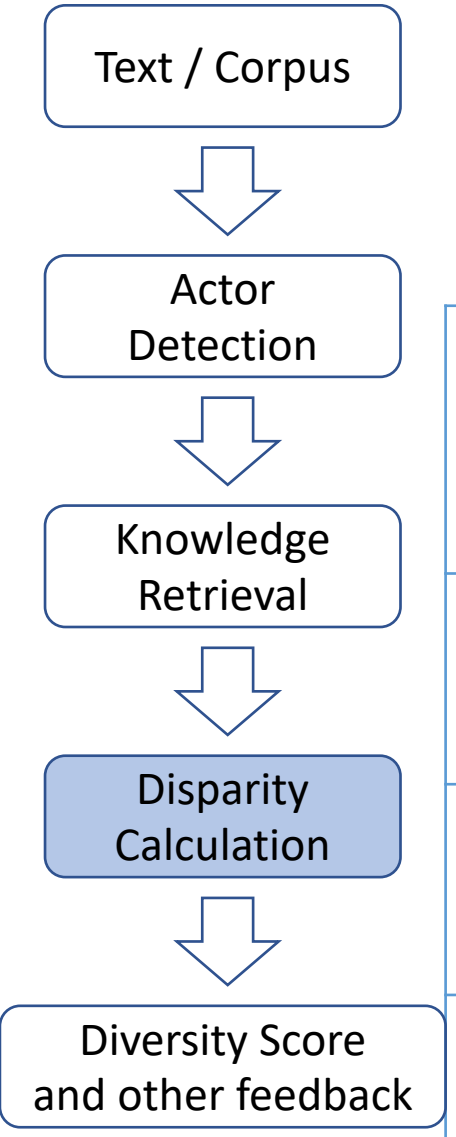
Diversity Searcher Pipeline

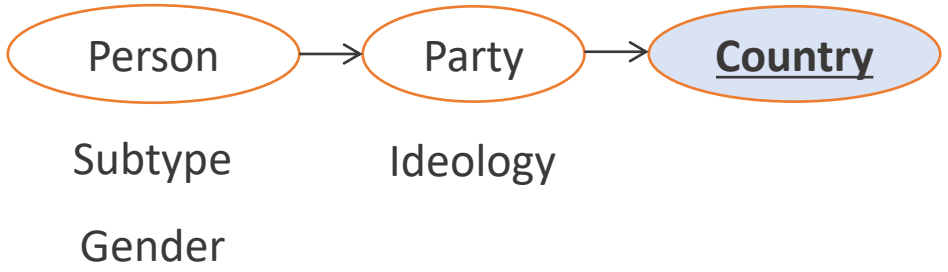




	Type	Subtype	Gender	Party	Party Ideology	Country
Donald Trump	Person	Politician	Male	Republican Party	Conservative	US
Hillary Clinton	Person	Politician	Female	Democratic Party	Liberal	US
Tayyip Erdogan	Person	Politician	Male	AKP	Conservative	Turkey

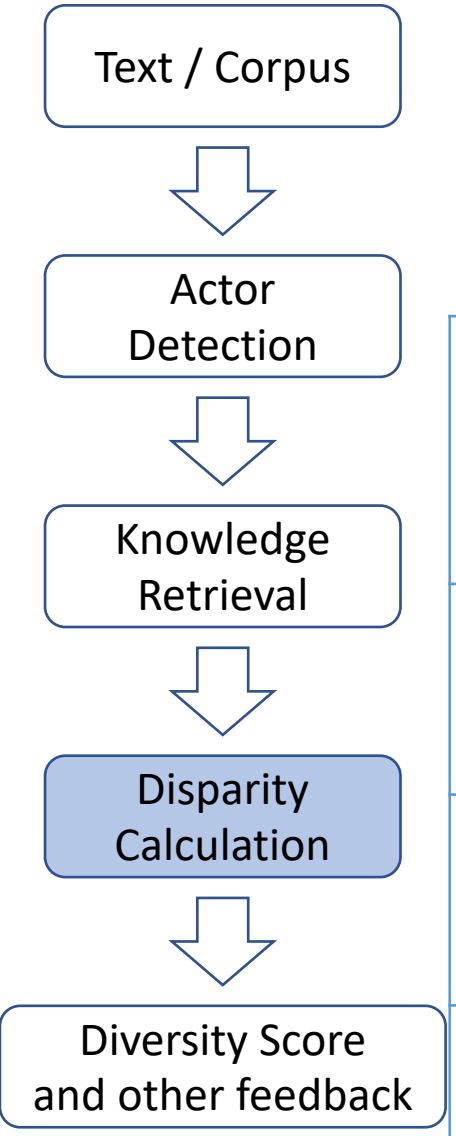
Diversity Searcher Pipeline

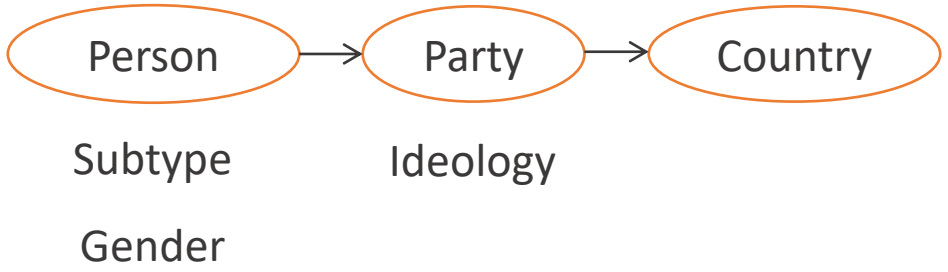




	Type	Subtype	Gender	Party	Party Ideology	Country
Donald Trump	Person	Politician	Male	Republican Party	Conservative	US
Hillary Clinton	Person	Politician	Female	Democratic Party	Liberal	US
Tayyip Erdogan	Person	Politician	Male	AKP	Conservative	Turkey

Diversity Searcher Pipeline





	Type	Subtype	Gender	Party	Party Ideology	Country
Donald Trump	Person	Politician	Male	Republican Party	Conservative	US
Hillary Clinton	Person	Politician	Female	Democratic Party	Liberal	US
Tayyip Erdogan	Person	Politician	Male	AKP	Conservative	Turkey

Diversity Searcher Pipeline

Text / Corpus

Actor Detection

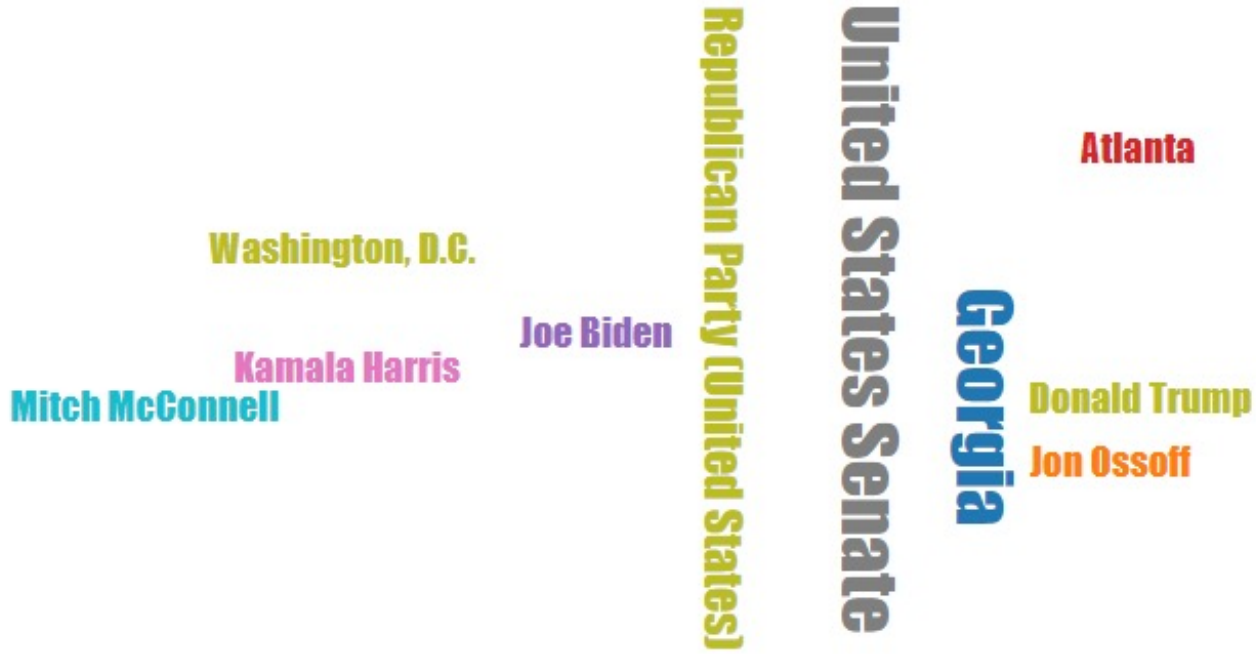
Knowledge Retrieval

Disparity Calculation

Diversity Score and other feedback

Diversity Searcher

Wordcloud



Stirling Diversity: 0.39

Diversity Searcher Pipeline

Text / Corpus

Actor Detection

Knowledge Retrieval

Disparity Calculation

Diversity Score and other feedback

Diversity Searcher

$$\Delta = \sum_{i \neq j} (d_{ij})^{\alpha} (p_i p_j)^{\beta}$$

Stirling Diversity: 0.39

Stirling Diversity: 0.39

Diversity Searcher Pipeline

Text / Corpus

Actor Detection

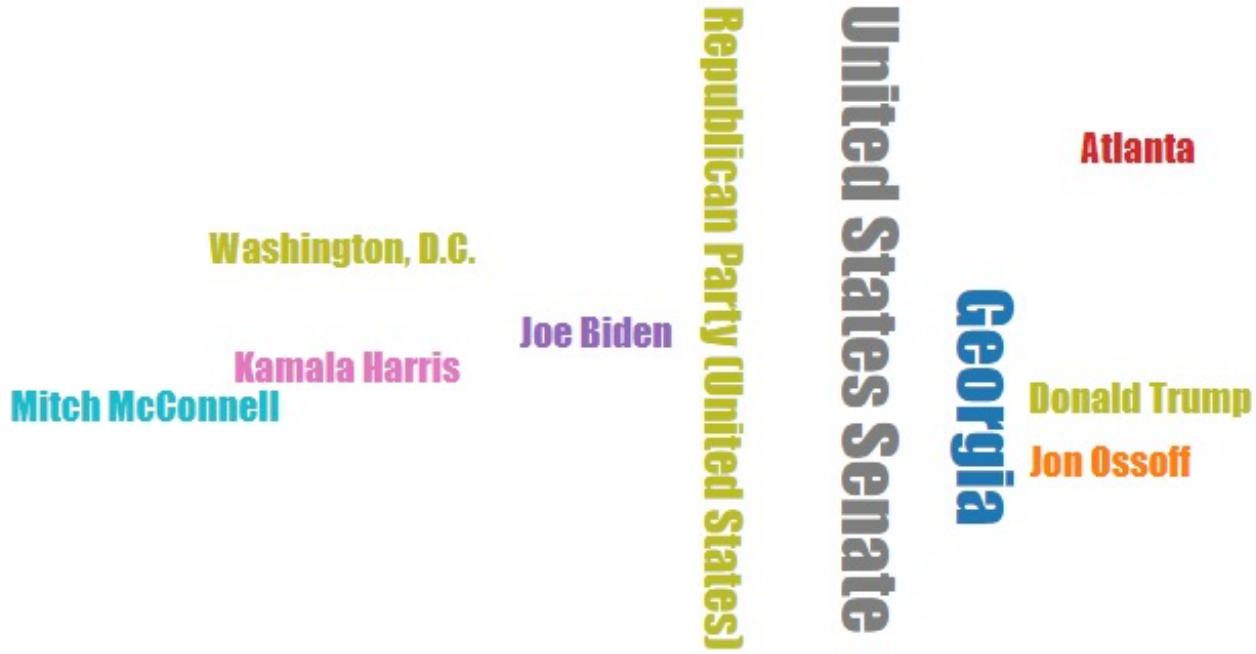
Knowledge Retrieval

Disparity Calculation

Diversity Score and other feedback

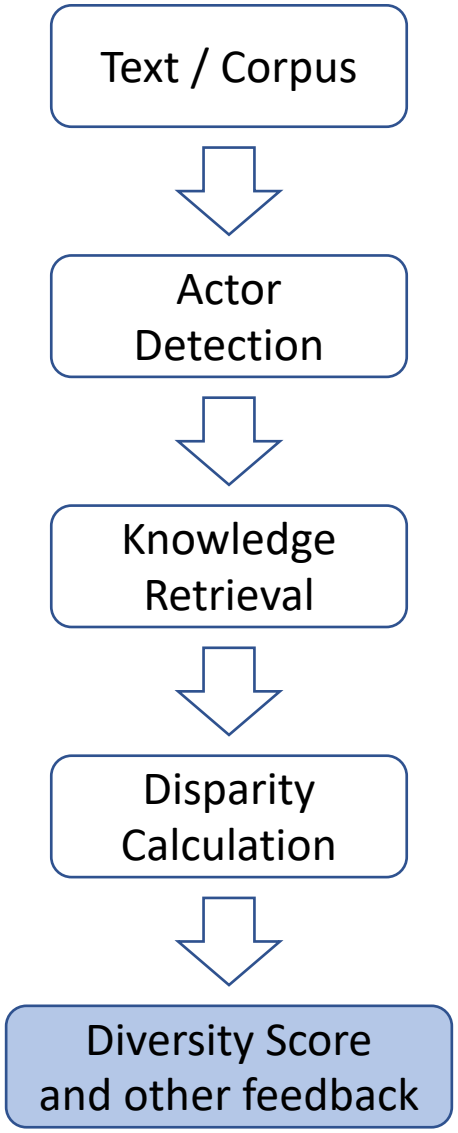
Diversity Searcher

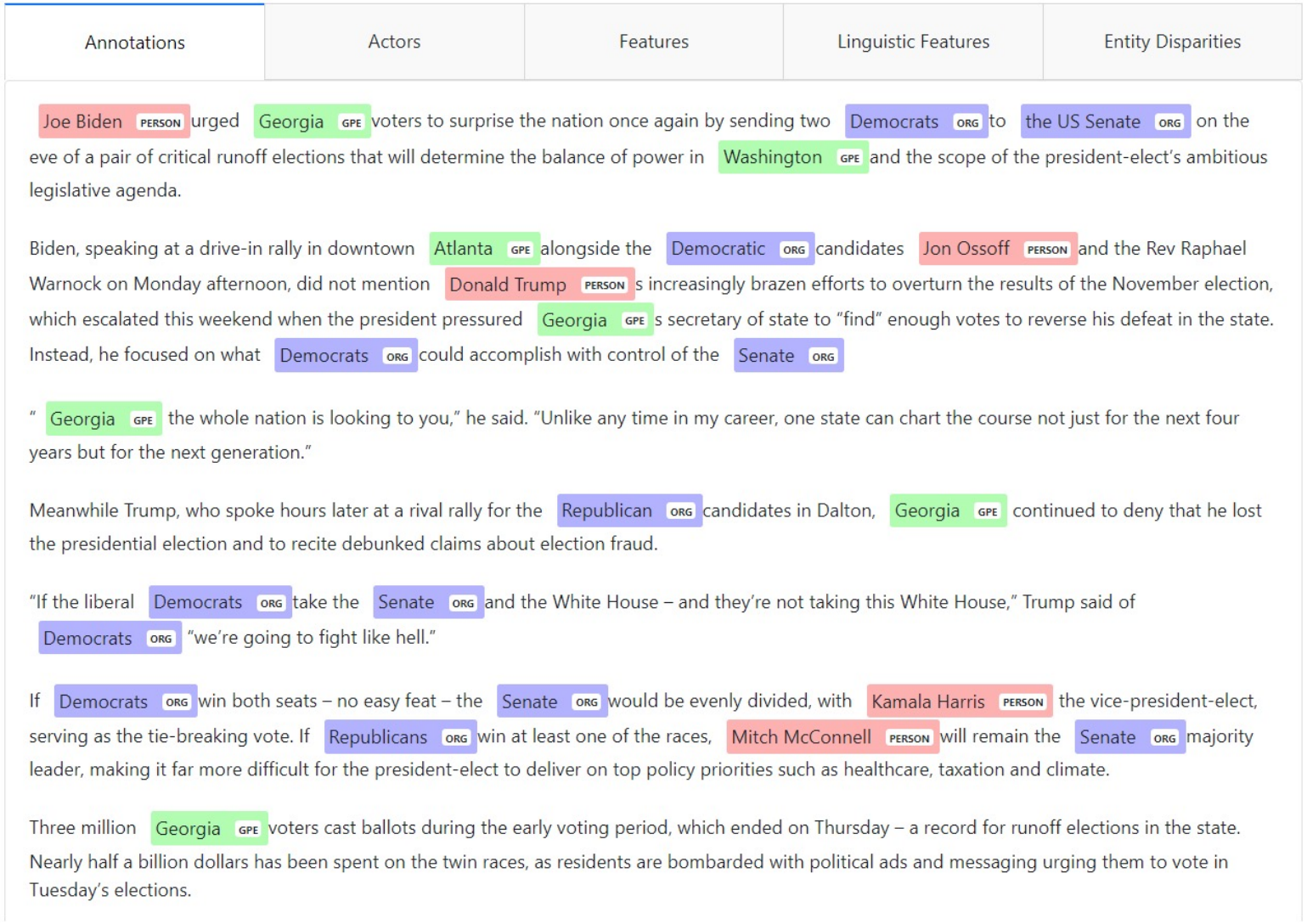
Wordcloud



Stirling Diversity: 0.39

Diversity Searcher Pipeline





Diversity Searcher Pipeline

Joe Biden PERSON urged Georgia GPE voters to surprise the nation once again by sending two Democrats ORG to the US S
eve of a pair of critical runoff elections that will determine the balance of power in Washington GPE and the scope of the presi
legislative agenda.

Biden, speaking at a drive-in rally in downtown Atlanta GPE alongside the Democratic ORG candidates Jon Ossoff PERSON a
Warnock on Monday afternoon, did not mention Donald Trump PERSON s increasingly brazen efforts to overturn the results of t
which escalated this weekend when the president pressured Georgia GPE s secretary of state to "find" enough votes to reverse
Instead, he focused on what Democrats ORG could accomplish with control of the Senate ORG

" Georgia GPE the whole nation is looking to you," he said. "Unlike any time in my career, one state can chart the course not jus
years but for the next generation."

Meanwhile Trump, who spoke hours later at a rival rally for the Republican ORG candidates in Dalton, Georgia GPE continue
the presidential election and to recite debunked claims about election fraud.

Diversity Score
and other feedback

leader, making it far more difficult for the president-elect to deliver on top policy priorities such as healthcare, taxation and climate.

Three million Georgia GPE voters cast ballots during the early voting period, which ended on Thursday – a record for runoff elections in the state.
Nearly half a billion dollars has been spent on the twin races, as residents are bombarded with political ads and messaging urging them to vote in
Tuesday's elections.

Diversity Searcher Pipeline

Text / Corpus

↓

Actor Detection

↓

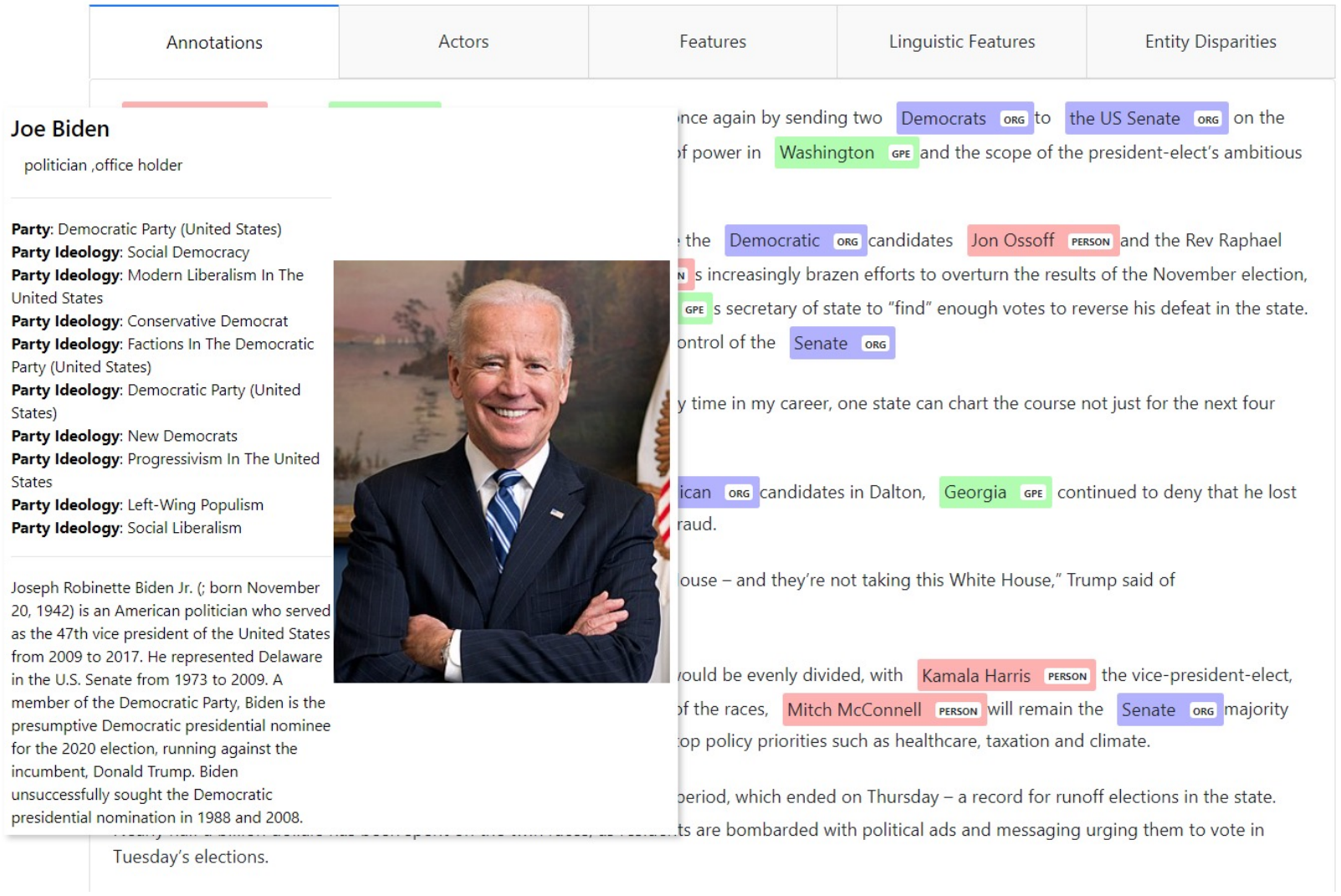
Knowledge Retrieval

↓

Disparity Calculation

↓

Diversity Score and other feedback



Diversity Searcher Pipeline

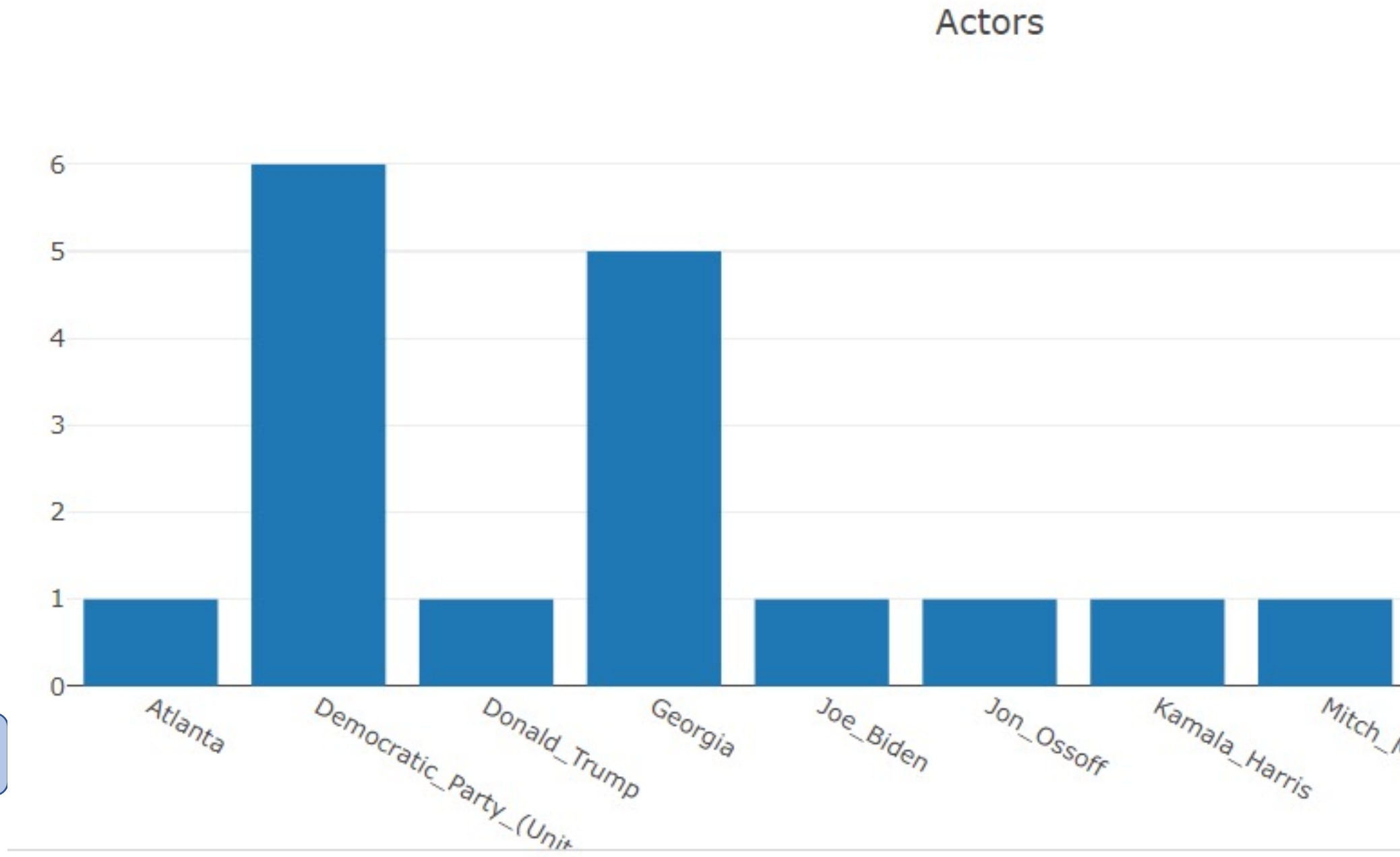
Text / Corpus

Actor Detection

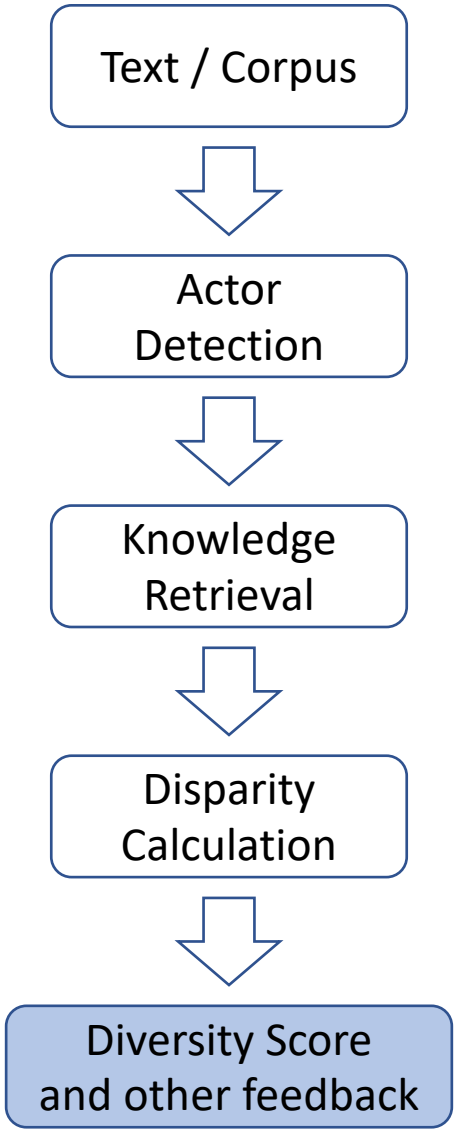
Knowledge Retrieval

Disparity Calculation

Diversity Score and other feedback

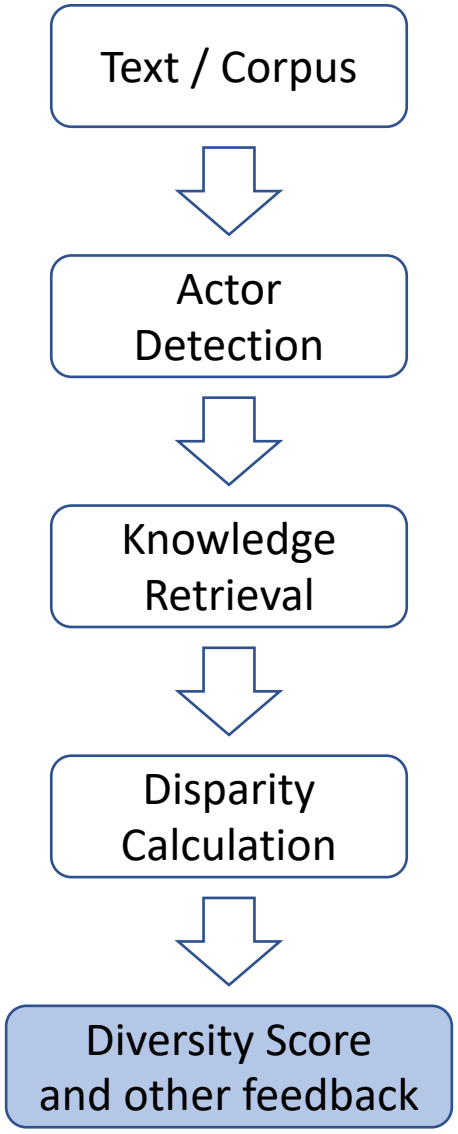


Diversity Searcher Pipeline

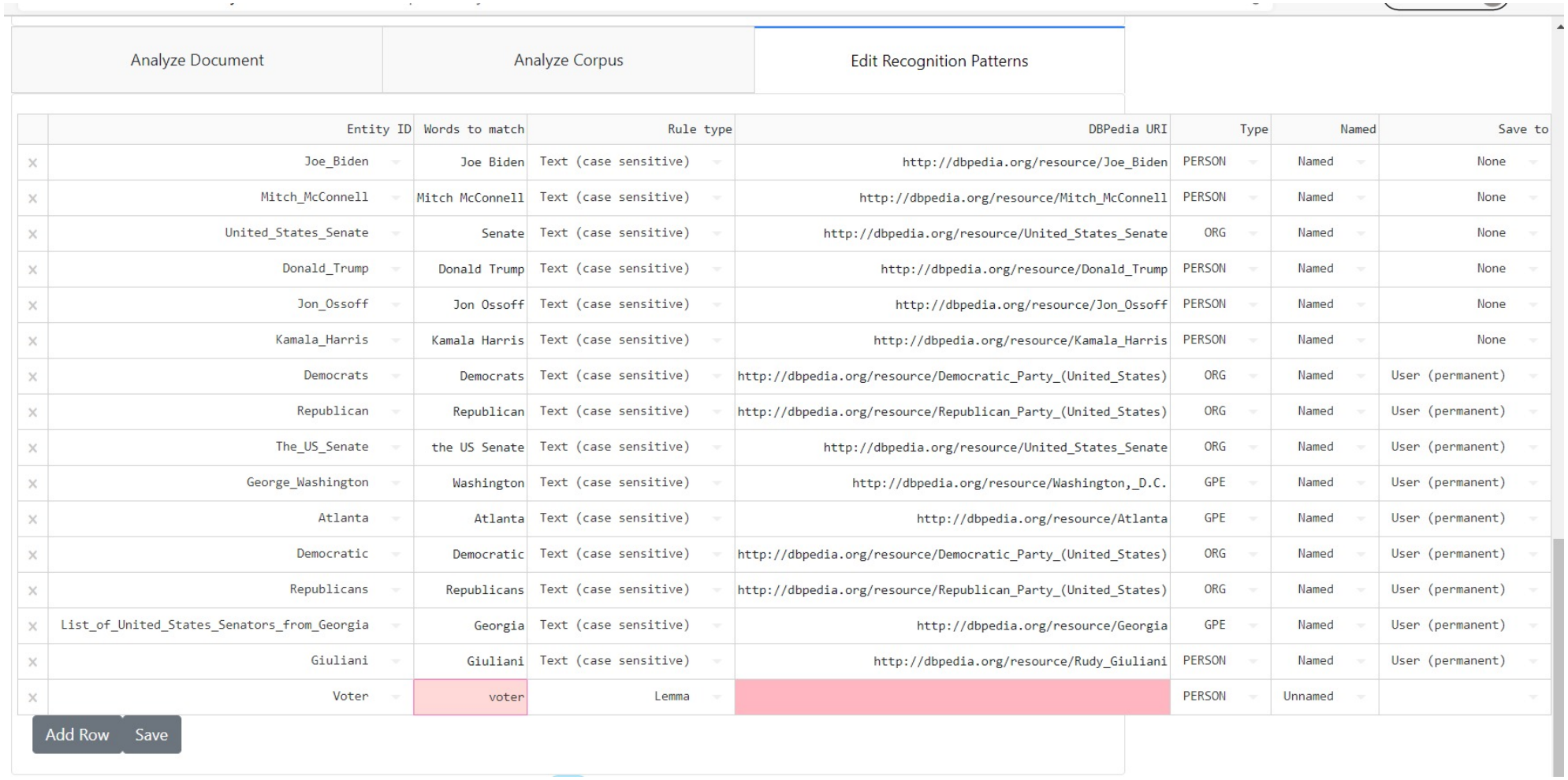


index	total
Atlanta	1
Democratic_Party_(United_States)	6
Donald_Trump	1
Georgia	5
Joe_Biden	1
Jon_Ossoff	1
Kamala_Harris	1
Mitch_McConnell	1
Republican_Party_(United_States)	2
United_States_Senate	5
Washington,_D.C.	1

Diversity Searcher Pipeline



Customization of Recognition



Analyze Document | Analyze Corpus | Edit Recognition Patterns

	Entity ID	Words to match	Rule type	DBPedia URI	Type	Named	Save to
×	Joe_Biden	Joe Biden	Text (case sensitive)	http://dbpedia.org/resource/Joe_Biden	PERSON	Named	None
×	Mitch_McConnell	Mitch McConnell	Text (case sensitive)	http://dbpedia.org/resource/Mitch_McConnell	PERSON	Named	None
×	United_States_Senate	Senate	Text (case sensitive)	http://dbpedia.org/resource/United_States_Senate	ORG	Named	None
×	Donald_Trump	Donald Trump	Text (case sensitive)	http://dbpedia.org/resource/Donald_Trump	PERSON	Named	None
×	Jon_Ossoff	Jon Ossoff	Text (case sensitive)	http://dbpedia.org/resource/Jon_Ossoff	PERSON	Named	None
×	Kamala_Harris	Kamala Harris	Text (case sensitive)	http://dbpedia.org/resource/Kamala_Harris	PERSON	Named	None
×	Democrats	Democrats	Text (case sensitive)	http://dbpedia.org/resource/Democratic_Party_(United_States)	ORG	Named	User (permanent)
×	Republican	Republican	Text (case sensitive)	http://dbpedia.org/resource/Republican_Party_(United_States)	ORG	Named	User (permanent)
×	The_US_Senate	the US Senate	Text (case sensitive)	http://dbpedia.org/resource/United_States_Senate	ORG	Named	User (permanent)
×	George_Washington	Washington	Text (case sensitive)	http://dbpedia.org/resource/Washington,_D.C.	GPE	Named	User (permanent)
×	Atlanta	Atlanta	Text (case sensitive)	http://dbpedia.org/resource/Atlanta	GPE	Named	User (permanent)
×	Democratic	Democratic	Text (case sensitive)	http://dbpedia.org/resource/Democratic_Party_(United_States)	ORG	Named	User (permanent)
×	Republicans	Republicans	Text (case sensitive)	http://dbpedia.org/resource/Republican_Party_(United_States)	ORG	Named	User (permanent)
×	List_of_United_States_Senators_from_Georgia	Georgia	Text (case sensitive)	http://dbpedia.org/resource/Georgia	GPE	Named	User (permanent)
×	Giuliani	Giuliani	Text (case sensitive)	http://dbpedia.org/resource/Rudy_Giuliani	PERSON	Named	User (permanent)
×	Voter	voter	Lemma		PERSON	Unnamed	

Add Row | Save

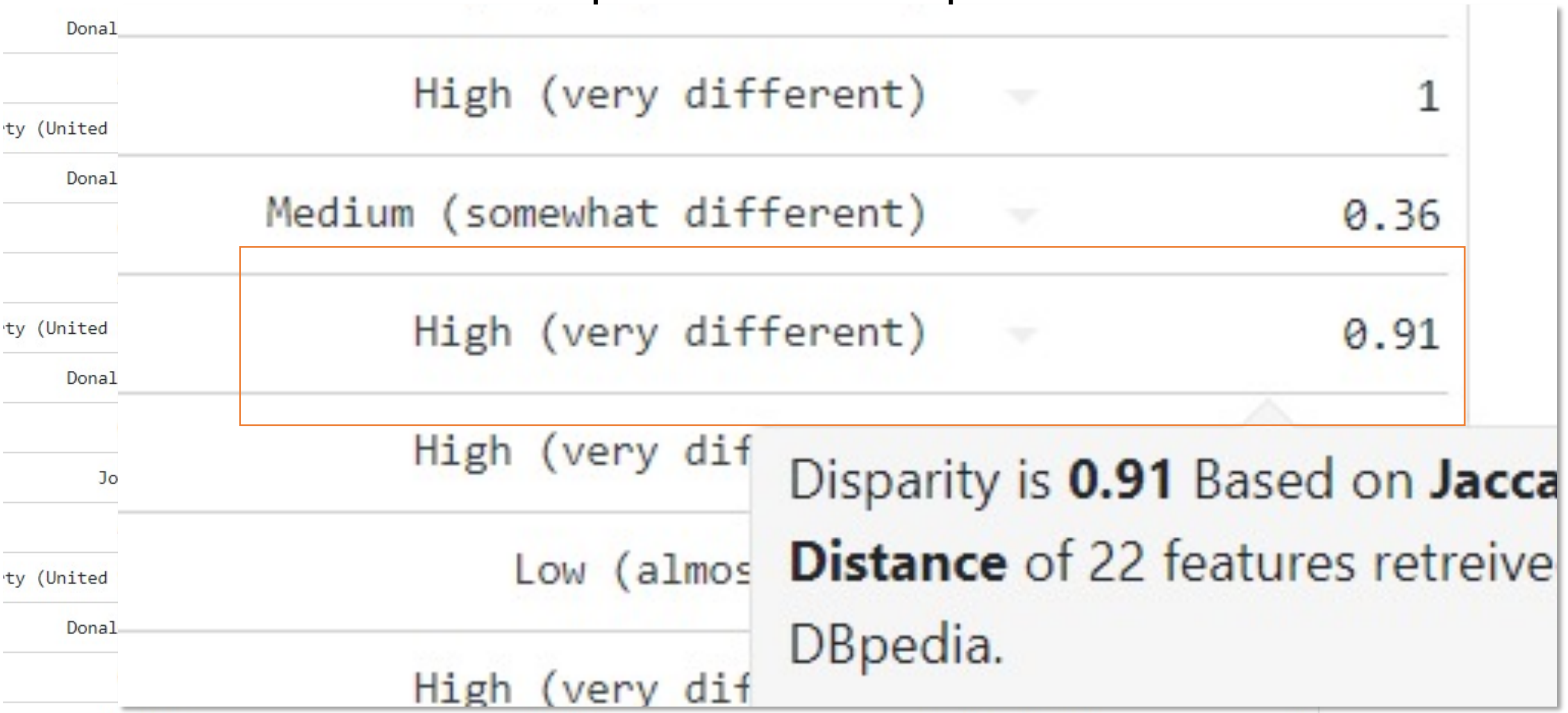
Customization of Recognition

Democrats	Text (case sensitive)	http://dbpedia.org/resource/Democratic_Party_(United_States)	ORG	Named
Republican	Text (case sensitive)	http://dbpedia.org/resource/Republican_Party_(United_States)	ORG	Named
e US Senate	Text (case sensitive)	http://dbpedia.org/resource/United_States_Senate	ORG	Named
Washington	Text (case sensitive)	http://dbpedia.org/resource/Washington,_D.C.	GPE	Named
Atlanta	Text (case sensitive)	http://dbpedia.org/resource/Atlanta	GPE	Named
Democratic	Text (case sensitive)	http://dbpedia.org/resource/Democratic_Party_(United_States)	ORG	Named
Republicans	Text (case sensitive)	http://dbpedia.org/resource/Republican_Party_(United_States)	ORG	Named
Georgia	Text (case sensitive)	http://dbpedia.org/resource/Georgia	GPE	Named
Giuliani	Text (case sensitive)	http://dbpedia.org/resource/Rudy_Giuliani	PERSON	Named
voter	Lemma		PERSON	Unnamed

Customization of pairwise disparities

Donald Trump	Georgia	High (very different)	1
Atlanta	Joe Biden	High (very different)	1
ty (United States)	Joe Biden	Medium (somewhat different)	0.4
Donald Trump	Joe Biden	High (very different)	0.91
Georgia	Joe Biden	High (very different)	1
Atlanta	Jon Ossoff	High (very different)	1
ty (United States)	Jon Ossoff	Medium (somewhat different)	0.36
Donald Trump	Jon Ossoff	High (very different)	0.91
Georgia	Jon Ossoff	High (very dif	
Joe Biden	Jon Ossoff	Low (almos	
Atlanta	Kamala Harris	High (very dif	
ty (United States)	Kamala Harris	Medium (somewhat dif	
Donald Trump	Kamala Harris	High (very dif	
Georgia	Kamala Harris	High (very different)	1
Joe Biden	Kamala Harris	Low (almost same)	0

Disparity is **0.91** Based on **Jaccard Distance** of 22 features retreived fr DBpedia.

Common features	Non-common features	All featu
2	20	22

Customization of pairwise disparities