

The Final Countdown:

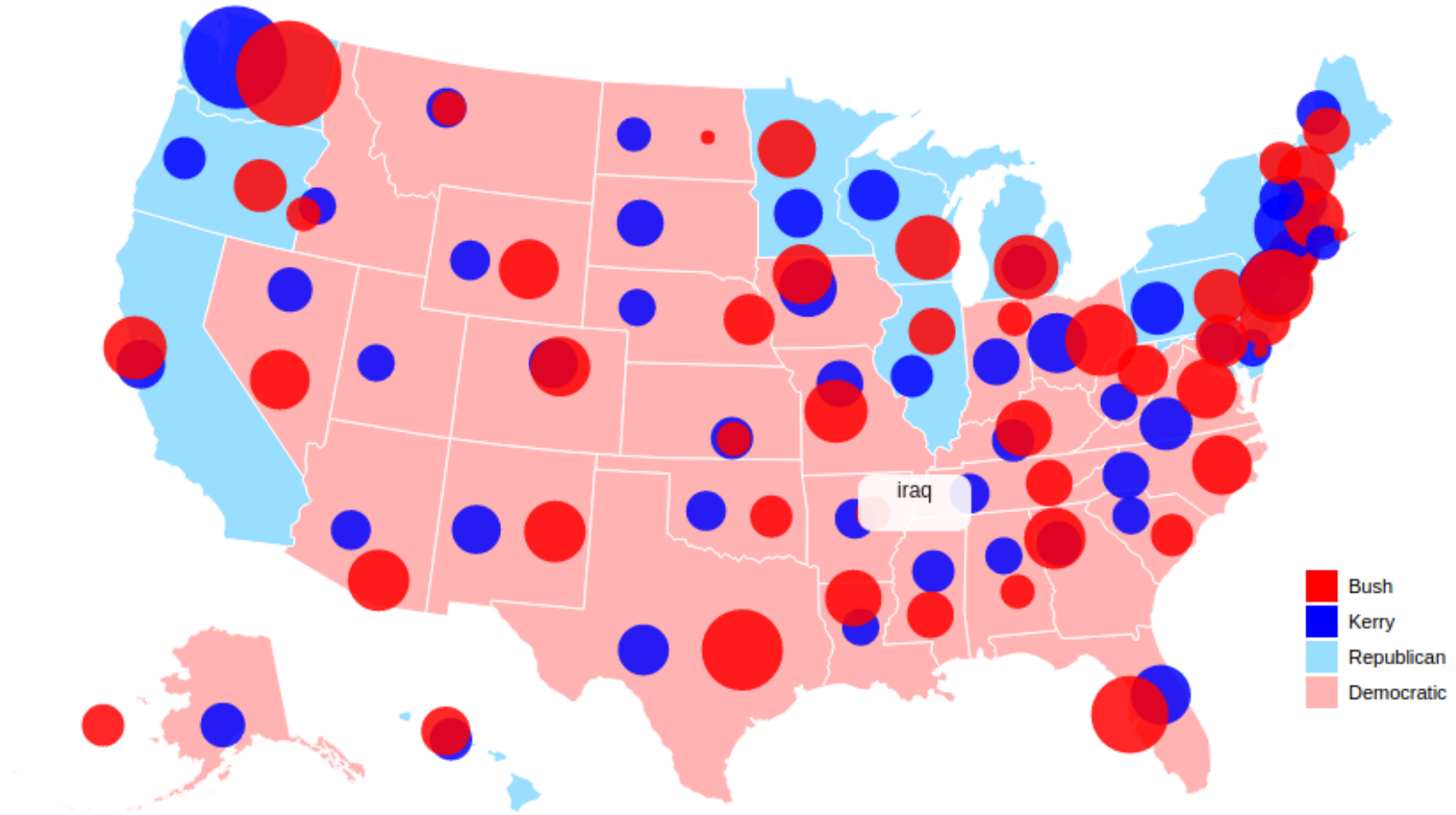
Talking Points by Place and Party in the last days of the 2004 Election

Research Question:

Which places and topics did the two presidential candidates mention most in the days leading up to the 2004 election?

Method:

1. Parse and clean the text from WARC files of party/candidate websites
2. Extracted Named Entities (candidate names, state names)
3. Topic modeling (LDA, TF-IDF)
4. Semantic Web extension - Google Knowledge Graph API and Wikidata



<http://www.cs.odu.edu/~maturban/hackathon/>

```

bash
nc@nchah:warcbase-scripts$ python kg-api.py 'george w. bush'
Displaying results... (limit: 5)

George W. Bush
- entity_types: Thing, Person
- description: 43rd U.S. President
- detailed_description: George Walker Bush is an American politician who served as the 43rd President of the United States from 2001 to 2009 and 46th Governor of Texas from 1995 to 2000.
- identifier: kg:/m/09b6zr
- url: http://www.georgewbush.com/
- resultScore: 2214.081787

George H. W. Bush
- entity_types: Thing, Person
- description: 41st U.S. President
- detailed_description: George Herbert Walker Bush is an American politician who served as the 41st President of the United States from 1989 to 1993 and 42nd Governor of Texas from 1975 to 1979.
- identifier: kg:/m/0341s
- url: N/A
- resultScore: 774.34845

George W. Bush Presidential Center
- entity_types: Place, TouristAttraction, Thing
- description: Building complex
- detailed_description: The George W. Bush Presidential Center, which opened in 2002, is a complex of buildings in Dallas, Texas, that houses the George W. Bush Presidential Library and the George W. Bush Presidential Center.
- identifier: kg:/m/0277mw9
- url: http://www.georgewbushlibrary.com/
- resultScore: 542.113708

Dixie Chicks
- entity_types: MusicGroup, Thing
- description: Musical band
- detailed_description: The Dixie Chicks are an American country music band consisting of three sisters: Laura, white, and Natalie.
- identifier: kg:/m/018ndc
- url: http://www.dixiechicks.com/
- resultScore: 438.937195

W.
- entity_types: Movie, Thing
- description: 2008 film
- detailed_description: W. is a 2008 American biographical drama film loosely based on the life of George W. Bush.
- identifier: kg:/m/03m5y9p

```

The screenshot shows a Google search for "george w bush". The top result is the Wikipedia entry for George W. Bush. Below it are news snippets from Hot Air, RedState, and the Washington Times. To the right is a summary card for George W. Bush, including his birth date, spouse, and presidential term. Below the search results is the Wikidata entry for George W. Bush, which includes a table of labels in various languages and their corresponding descriptions.

| Language | Label | Description | Also known as |
|----------|----------------|---|--|
| English | George W. Bush | 43rd President of the United States | George Bush George Walker Bush Bush, George W. Bush Jr. |
| Korean | 조지 W. 부시 | 미국의 정치인이자 제43대 대통령 | |
| Galician | George W. Bush | No description defined | |
| French | George W. Bush | 43e président des États-Unis d'Amérique | Georges W. Bush George W Bush Georges W Bush George Walker Bush |

<https://github.com/nchah/warcbase-scripts>