Understanding European Institutional Policy Discourse on the Council of Europe's Convention on Preventing and Combating Violence Against Women through Automated Content Analysis

Karen Vogel, PhD
Hamline University

Bridget McInnes, PhD
Minnesota Supercomputing Institute
Securboration

Context of Research

- Context of Research
  - Academic Literature on Violence Against Women, Domestic Violence, Gender Policy Reform in European Institutions and Discourse Policy

- Focus
  - Convention on Preventing and Combating Violence Against Women

- Discourse and Policy in:
  - European Council of the European Union
  - Council of Europe

Research Questions

- To what extent is discourse (words) about preventing and combating violence against women and domestic violence appear in the speeches of key political figures of the European Council of the EU and the Committee of Ministers of the Council of Europe?
  - to what extent does the content of speeches and minutes show the privileging of gender policy debate and human rights?

- To what extent has reference to the Convention been prevalent in the dialogues of key European political leaders after the opening for signature and ratification of the Convention?
  - Do we see a similarity between the Convention in dialogues in both the EU and the Council of Europe which may signal continued interest in the Convention?

- How effective is automated content analysis in helping us answer these questions?

Combating Violence against Women and Domestic Violence

- Evolution of Discourse
  - European Union and Council of Europe

- Policy Reform Initiatives and Capacity Building

- Council of Europe work on binding agreement

- Convention as test case for understanding durable dialogue
-

Data

- Council of Europe

  o Speeches from the Committee of Ministers

  o Download link: www.coe.int/t/cm/WCD/fulltextSearch_en.asp#

  o Queried
    - Term: speeches
    - Date Range: 1 Jan 2008 -15 Jan 2013

  o Results:
    - Number of documents: 251
    - Date Range: 8 January 2009 – 15 January 2013

Data (cont.)

- European Council Data

  - Speeches from European Council Member

  - Download link:
    www.consillium.europea.eu/uedocs/cms_data/docs/pressdata/en/ec/

  - Filtered:
    - Removed:
      - Council of European Union documents
      - European Parliament documents
      - Non-English documents
    - Date Range: 1 Jan 2009 -15 Jan 2013

  - Results:
    - Number of documents: 653
    - Date Range: 6 October 2009 – 8 May 2012

Data chart: Number of documents in the datasets broken down by year

Year	Council of Europe	European Council
2009	67	13
2010	65	258
2011	56	267
2012	59	106
2013	4	0
Total	251	653

Method: Ngram Modeling

- Ngrams: sequence of N words in a sentence

Among the international human rights treaties

1grams	2grams	3grams
among	international human	international human rights
international	human rights	human rights treaties
human	rights treaties	
rights		
treaties		

Method: Cosine Document Similarity

Speeches consisted of:

Conformity with the international declaration of human rights

Feature	Convention Vector	Speech Vector
among	12	0
international	19	10
human	23	18
rights	23	18
treaties	8	0
international human	13.4	0
human rights	15.6	22.3
rights treaties	13.4	0
international human rights	23.4	0
human rights treaties	25.2	0

Convention Vector

Speech #1 Vector

Speech #2 Vector

θ_1

θ_2
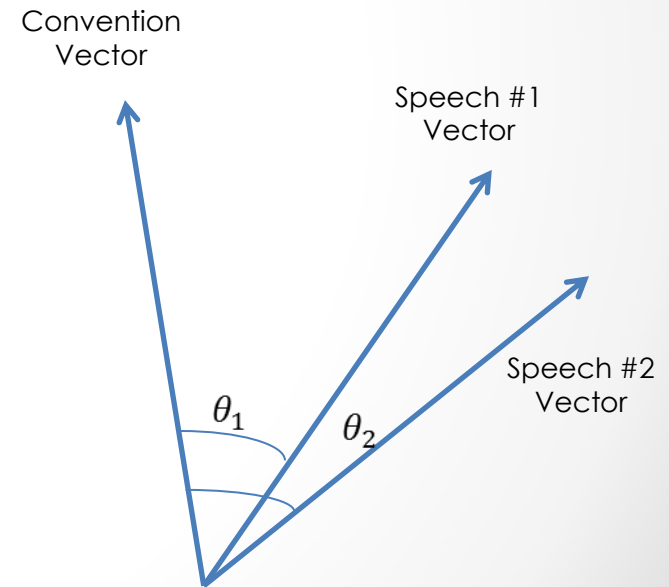
$$similarity(A, B) = \cos \frac{A \cdot B}{\|A\| \|B\|}$$

Method: Ranking

Narrow view content within a specific speech

| Speech Ranking based on Cosine	Cosine
Speech #1	0.75
Speech #2	0.63
Speech #3	0.62
Speech #4	0.60
Speech #5	0.54
Speech #6	0.52
Speech #7	0.39
Speech #8	0.28
Speech #9	0.25
etc.	…

Broad view of content over time period

Bucket Ranking based on date range	Cosine
Oct – Dec 2009	0.32
Jan – Mar 2010	0.53
Apr – Jun 2010	0.67
Jul – Sep 2010	0.60
Oct – Dec 2010	0.75
Jan – Mar 2011	0.52
Apr – Jun 2011	0.85
Jul – Sep 2011	0.28
Oct – Dec 2011	0.36
Jan – Mar 2012	0.58
Apr – Jun 2012	0.26
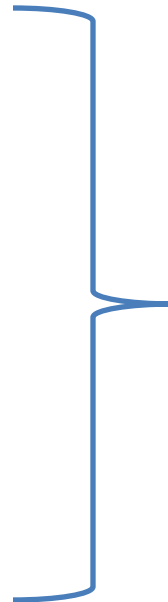
Similarity Results
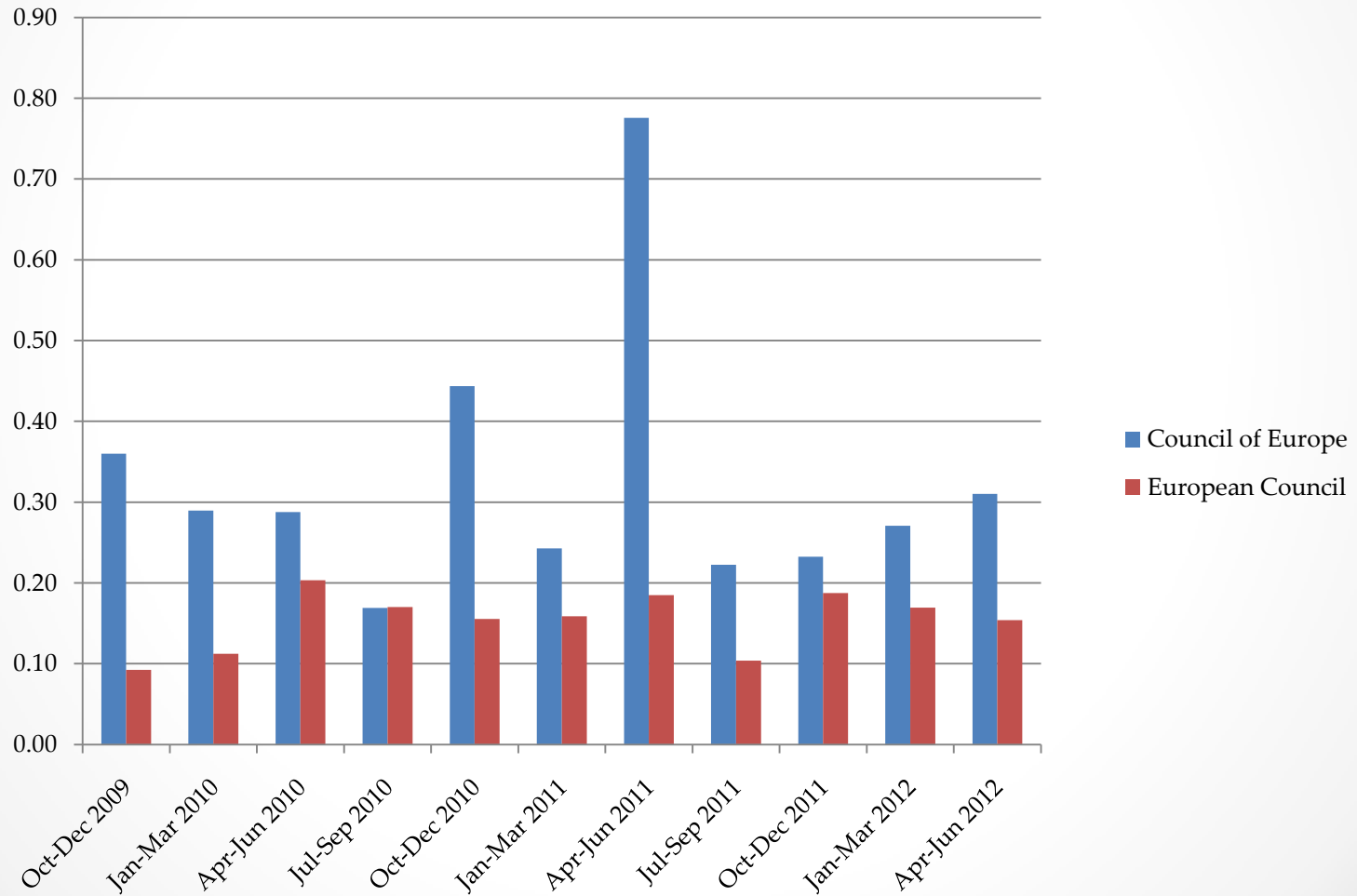
Feature Results

Number of Features found in the Documents broken down by Date Range

Date range	European Council	Council of Europe
Oct-Dec 09	11	116
Jan-Mar 10	6	101
Apr-Jun 10	23	120
Jul-Sep 10	7	122
Oct-Dec 10	21	85
Jan-Mar 11	20	62
Apr-Jun 11	26	9
Jul-Sep 11	10	101
Oct-Dec 11	30	140
Jan-Mar 12	23	123
Apr-Jun 12	14	96

Method: Topic Ranking

- Goal: Identify the statistically relevant Ngrams in the documents over specified date ranges

- Method: Log Likelihood Ratio(G^2)

- Idea: is those N-grams that occur often and are statistically relevant with respect to each other describe the context of the documetnts

Topic Results

Date Range	European Council	Council of Europe
Oct-Dec 09	foreign affairs	**human rights**
	climate change	national minorities
Jan-Mar 10	climate change	**human rights**
	broadcast quality	**sexual orientation**
Apr-Jun 10	task force	**national minorities**
	climate change	persons belonging
Jul-Sep 10	task force	parliamentary assembly
	economic governance	higher education
Oct-Dec 10	task force	**human rights**
	economic governance	respondent state
Jan-Mar 11	southern neighbourhood	**human rights**
	financial stability	framework convention
Apr-Jun 11	middle east	minority languages
	southern neighbourhood	periodical report
Jul-Sep 11	plenipotentiary head	venice commission
	polish presidency	**human rights**
Oct-Dec 11	financial stability	minority languages
	fiscal discipline	**human rights**
Jan-Mar 12	single market	**human rights**
	financial stability	periodical report
Apr-Jun 12	single market	periodical report
	human rights	romanian authorities

Analysis

- Results show:
  - Debate present in both institutions over time

- Analysis suggests:
  - that policy consensus on addressing gender violence as a human rights issue has been emerging over the last five years
  - Slowing of discourse momentum after May 2011

Limitations

- The ngrams used in this work consisted of contiguous sequences of words

woman's equality and *equality of women*
are considered two different ngrams

- In the future, we would like to expand this to allow for variations in linguistic phenomena such as
  - *windowing*
  - stemming
  - Remove word order (co-occurrences versus ordered pairs)

Future Directions

- Look more closely at the French and English speaking countries and analyze the dialog of the discussion on whether to ratify the convention (or not)

- This analysis would allow us to look at the connection between domestic political discourse and the European institutional political discourse and to see the extent in which the two discourses (domestic and institutional) are the same/different.

Conclusions

Multiple approaches can increase our understanding of the synergy of European institutions and the emergence of policies on violence against women

Automated text analysis is just one possible tool in that mix which may help us compare the frequency of policy ideas and political discourse in both the European Union and the Council of Europe over time.

Additional Slides

Log Likelihood

	word2	- word2	Totals
word1	n_{11}	n_{12}	n_{1p}
- word1	n_{21}	n_{22}	n_{2p}
Totals	n_{p1}	n_{p2}	n_{pp}

$$G^2 = 2 * \sum_i^j n_{i,j} * \log(\frac{n_{i,j}}{m_{i,j}})$$

$$m_{i,j} = \frac{n_{i,p} + n_{p,i}}{n_{p,p}}$$