

Presented at:

DH2015
Global Digital
Humanities

“Everything on Paper Will Be Used Against Me”: Quantifying Kissinger

*A Computational Analysis of the Digital National Security Archive’s
Henry A. Kissinger Correspondence Collection, 1968-1977*

Micki Kaufman, MA

Doctoral Student, Department of History, CUNY Graduate Center
New York, NY USA


Email: mkaufman@gradcenter.cuny.edu

Twitter: @MickiKaufman

<http://blog.quantifyingkissinger.com>

**THE
GRADUATE
CENTER**
CITY UNIVERSITY
OF NEW YORK

Generously supported by a
CUNY Graduate Center
Provost’s Digital Innovation Grant
2012-2013, 2013-2014, 2014-2015



“Digitisation by itself is not sufficient to break through the fog of stories and the confusion of a society divided by competing mythologies. Cautious and judicious curating of possible data, questions, and subjects is necessary.

We must strive to discern and promote questions that are synthetic and relevant and which break new methodological ground.”

-- Guldi & Armitage,

“The History Manifesto” (2014)



John Ehrlichman, 1982
Former Chief Domestic Advisor

Archives
Egypt

Methodological Ground



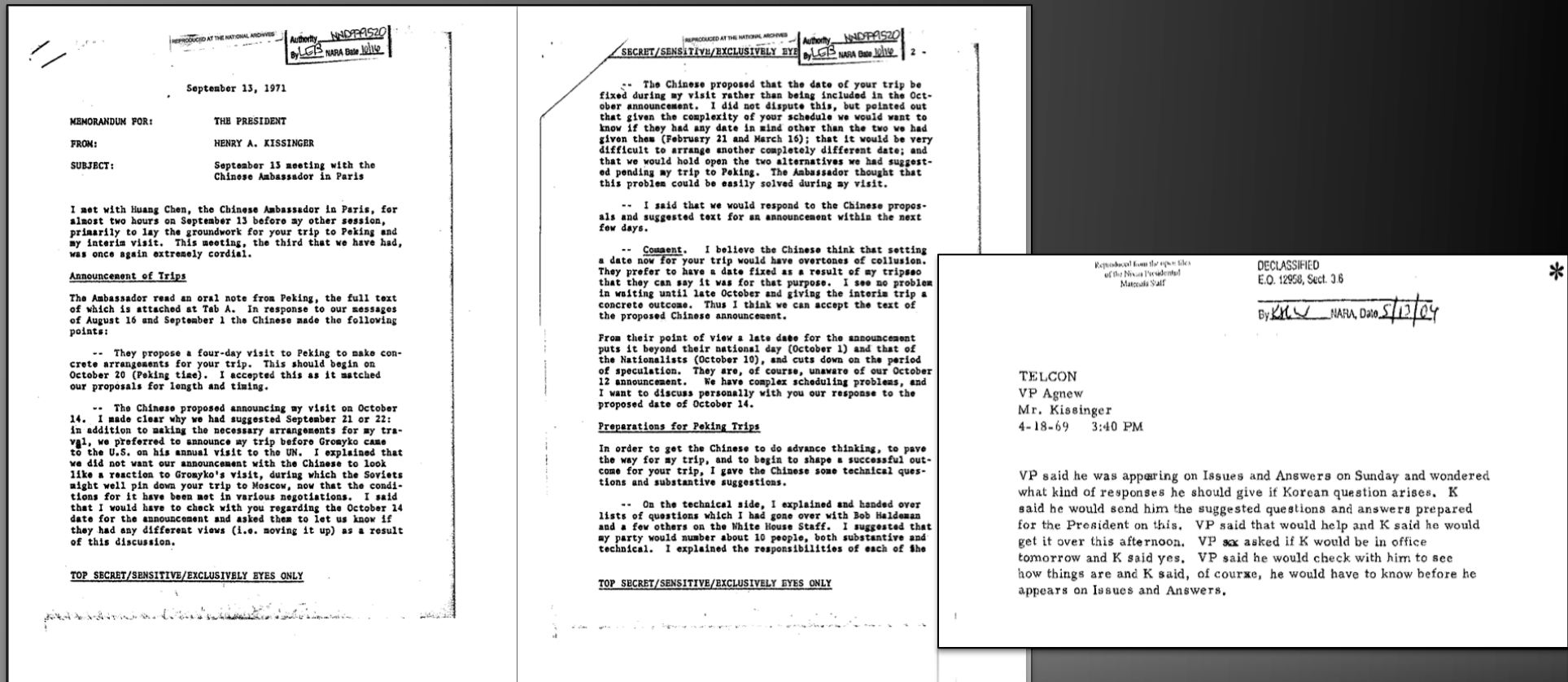
Using a host of interdisciplinary text, data and network analysis techniques* this project attempts to explore the use of quantitative methods to facilitate new questions – and re-examine current understandings – about the people, policies and processes of the Kissinger correspondence by distant and/or networked reading.

* eg. 'word frequency/ngram analysis,' 'word collocation,' 'topic modeling,' 'network mapping,' 'sentiment analysis,' 'data visualization' and others.

About the Sources

A fascinating 'Medium Data' problem: the DNSA (Digital National Security Archive)'s Kissinger Collection contains more than 2 gigabytes of text and pdf files comprising 17,500+ meeting memoranda ('memcons') and teleconference transcripts ('telcons').

These documents, among millions of existing documents, have been curated by the DNSA and detail Henry A. Kissinger's correspondence during his tenure as National Security Adviser (1969-1975) and as US Secretary of State (1973-1977).



We accessed, ingested, processed and analyzed both the text content AND the document metadata.

DIGITAL NATIONAL SECURITY ARCHIVE
The documents that made U.S. policy

HOME | COLLECTIONS GUIDE | INFORMATION RESOURCES | MY ARCHIVE | HELP

QUICK SEARCH:
[Search Box] **Go**
Need [HELP?](#)

SEARCH
» DOCUMENTS
» BIBLIOGRAPHY
» CHRONOLOGY
» GLOSSARY

SEARCH DOCUMENTS
DOCUMENTS: FULL RECORD

Add to Marked List

Citation: [\[Attendance at National Security Council Meeting; Best Available Copy\]](#)
Non-Classified, Memorandum of Telephone Conversation, Annotated Copy, January 21, 1969, 1150 Local Time, 1 pp.

Collection: Kissinger Telephone Conversations

Item Number: KA00001

Origin: United States. Assistant to the President for National Security Affairs

Location of Original: National Archives. Richard Nixon Presidential Library and Museum. Henry A. Kissinger Telephone Conversation Transcripts (Telcons). Chronological File. Box 1. January 21-31, 1969

Individuals/Organizations Named: [Nixon, Richard M.](#) | [Volcker, Paul A.](#) | United States. National Security Council

Participant(s): Kennedy, David M.

Subjects: [Bureaucracy](#) | [Meetings and schedules](#)

Abstract: David Kennedy indicates to Henry Kissinger that he would like to be invited to the next meeting of the National Security Council.

Full Text: [Document image](#)

[Durable URL for this record](#)

MARKED LIST
(0 ITEMS)

[SEARCH HISTORY](#)

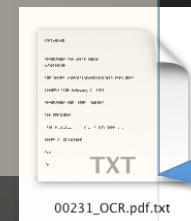
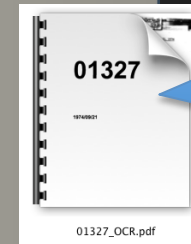
[EMAIL](#)

[PRINT VIEW](#)

[DOWNLOAD](#)

[SAVE](#)

| docid | doctitle | date | classification | category | pages |
|-------|--|-----------|----------------|----------------------------|-------|
| 144 | [U.S. Military Assistance to Israel; Attached to Cover 1 | 4-Jun-70 | Secret | Memorandum of Conversation | 3 |
| 145 | [British Ambassador Freeman and Henry Kissinger Dis | 5-Jun-70 | Secret | Letter | 3 |
| 146 | Persian Gulf | 5-Jun-70 | Secret | Minutes | 7 |
| 147 | [Discussion of Middle East Negotiations] | 6-Jun-70 | Secret | Letter | 3 |
| 148 | Jordan | 6-Jun-70 | Top Secret | Minutes | 13 |
| 149 | Jordan | 6-Jun-70 | Top Secret | Minutes | 7 |
| 150 | My June 10 Meeting with Ambassador Dobrynin | 6-Jun-70 | Top Secret | Memorandum | 13 |
| 151 | Cambodia | 6-Jun-70 | Top Secret | Minutes | 11 |
| 152 | Mediterranean, Greece, Italy: NSSM 90 | 17-Jun-70 | Secret | Memorandum of Conversation | 8 |
| 153 | [Meeting with Pakistani Ambassador Agha Hilaly] | 17-Jun-70 | Top Secret | Memorandum of Conversation | 3 |
| 154 | Middle East | 22-Jun-70 | Secret | Minutes | 14 |
| 155 | [Strategic Arms Limitation Talks Agreement] | 23-Jun-70 | Top Secret | Memorandum of Conversation | 2 |
| 156 | Verification Panel | 24-Jun-70 | Top Secret | Minutes | 12 |
| 157 | Vietnam Cease-Fire and Peace Initiatives | 27-Jun-70 | Top Secret | Minutes | 24 |
| 158 | U.S. Military Assistance to Indonesia | 1-Jul-70 | Top Secret | Memorandum of Conversation | 6 |
| 159 | [Meeting with Anatolii Dobrynin] | 7-Jul-70 | Top Secret | Memorandum of Conversation | 2 |
| 160 | The Middle East | 9-Jul-70 | Top Secret | Minutes | 12 |
| 161 | U.S. Military Assistance to Indonesia | 10-Jul-70 | Top Secret | Memorandum of Conversation | 6 |
| 162 | My Recent Conversations with Ambassador Dobrynin | 16-Jul-70 | Top Secret | Memorandum | 10 |
| 163 | Defense Budget | 17-Jul-70 | Top Secret | Minutes | 14 |
| 164 | Middle East | 20-Jul-70 | Top Secret | Minutes | 5 |
| 165 | My July 20 Meeting with Dobrynin | 21-Jul-70 | Top Secret | Information Memorandum | 7 |
| 166 | NSC Meeting: Vietnam Ceasefire and Possible Diplom | 21-Jul-70 | Top Secret | Memorandum of Conversation | 6 |
| 167 | [Meeting with Indian Ambassador Jha; Attached to Co | 22-Jul-70 | Secret | Memorandum of Conversation | 6 |
| 168 | MBFR Verification Panel | 30-Jul-70 | Secret | Minutes | 11 |
| 169 | Conversation with Vorontsov--Map Room--August 4, | 4-Aug-70 | Classification | Memorandum of Conversation | 2 |
| 170 | Meeting between Israeli Ambassador Rabin and Mr. K | 6-Aug-70 | Top Secret | Memorandum | 8 |
| 171 | Conversation with Vorontsov--August 7, 1970 | 7-Aug-70 | Classification | Memorandum of Conversation | 2 |
| 172 | Middle East | 12-Aug-70 | Secret | Minutes | 9 |
| 173 | [Discussion with Yitzhak Rabin] | 15-Aug-70 | Top Secret | Memorandum of Conversation | 8 |
| 174 | German-Soviet Negotiations | 17-Aug-70 | Secret | Memorandum of Conversation | 5 |
| 175 | My Meeting Today with Vorontsov on Possible Summ | 19-Aug-70 | Top Secret | Action Memorandum | 2 |
| 176 | Conversation with Brandt's Foreign Policy Advisor, Eg | 20-Aug-70 | Secret | Information Memorandum | 6 |
| 177 | U.S. Strategies and Forces for NATO (NSSM 84) [and] | 31-Aug-70 | Top Secret | Minutes | 17 |



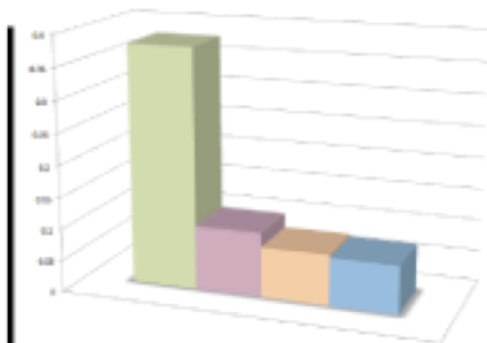
Both contain vital information regarding the correspondence and its context, even for text analysis!

Methodology: 'Topic Modeling'

1. MALLET software analyzes the frequency of all documents' words

responded that more ^{any/just} means for cadres ^{was} necessary. **President**
Nixon mentioned that **Hanoi** has never had ^{real} elections and is thus
employing a double standard. **Thieu** pointed out that 56% of those "elected"
in **North Vietnam** were women. This shows the magnitude of their man-
power problem. He reiterated that there would be elections after
the withdrawal of non-**South Vietnamese forces**. Thieu was prepared
for good international supervision -- even without troops. **President**
Nixon wondered whether the GVN could siphon off the **political forces**
in the center to weaken the Viet Cong. **Thieu** responded that when we
have a common position on our side, we can have a united front. What
made the middle ground in Saigon so uncertain was the fear that the
U.S. would withdraw support. Hence, many politicians were holding
themselves available for a coalition **government** with the NLF.

President Nixon asked why not a united front now; the GVN
is going to win and that is a great asset. Thieu stated very frankly
that there was a ^{of spirit} sagging in Saigon. Many still believe that the Viet



North-South-Thieu-Vietnam
.379

US-Nixon-Relations
.102

Soviet-Dobrynin
.08

Air-Military-Forces
.076

Methodology: 'Topic Modeling,' cont'd.

2. MALLET assigns 'topic weights' to each document based on shared patterns of word frequency.

Computer-Generated Topic List

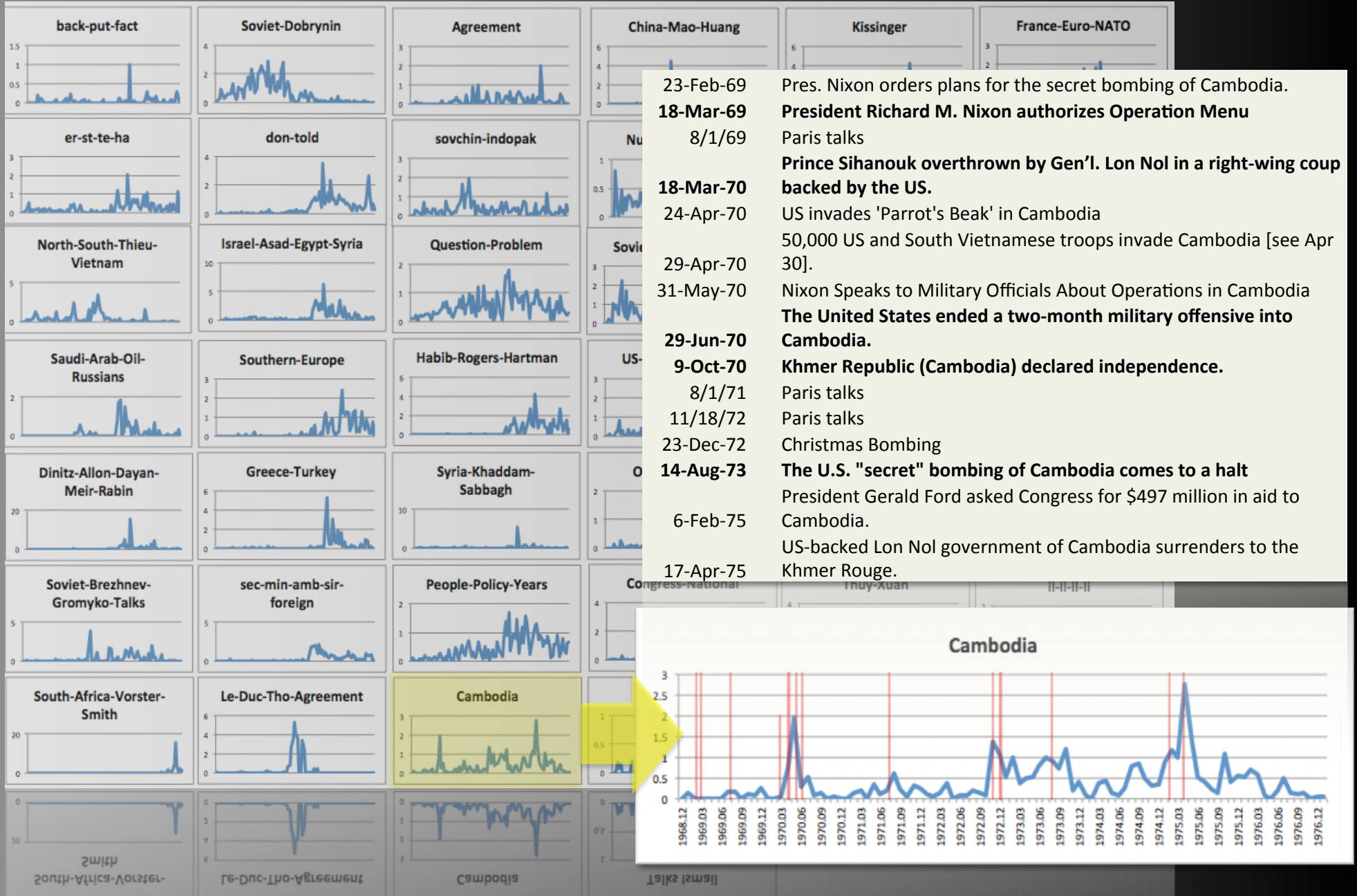
1. back-put-fact
2. Kissinger
3. er-st-te-ha
4. Johnson-Sec-Sens
5. **North-South-Thieu-Vietnam**
6. **SALT-Missile-Talks**
7. Saudi-Arab-Oil-Russians
8. Negotiations
9. Dinitz-Allon-Dayan-Meir-Rabin
10. Laughter
11. **Soviet-Dobrynin**
12. **France-Euro-NATO**
13. don-told
14. **Chou**
15. Israel-Asad-Egypt-Syria
16. air-military-forces
17. **Southern-Europe**
18. Archives
19. **Greece-Turkey**
20. se-ve-tl-al
21. Agreement
22. il-ll-il-li
23. **sovchin-indopak**
24. **Thuy-Xuan**
25. Question-Problem
26. Congress-National
27. **Habib-Rogers-Hartman**
28. People-Policy-Years
29. **Syria-Khaddam-Sabbagh**
30. sec-min-amb-sir-foreign
31. **China-Mao-Huang**
32. **Soviet-Brezhnev-Gromyko-Talks**
33. **Nuclear-Weapons**
34. **South-Africa-Vorster-Smith**
35. **Soviet-Group-Options**
36. **Le-Duc-Tho-Agreement**
37. US-Nixon-Relations
38. **Cambodia**
39. **Oil-Energy-Econ**
40. Talks Ismail

Human-Generated Topic List

- Wars in Indochina: not only **Vietnam**, but also **Laos** and **Cambodia**, and the White House's managerial role in United States military operations in these countries
- U.S. (especially Kissinger's) interactions with the North **Vietnamese**, the **Chinese**, and the **Soviets** in trying to bring conflicts in Indochina to an end
- the final stages of the Indochina wars, including U.S. reactions to **the collapse of the client regime in Saigon** and the rise of the **Cambodian Khmer Rouge**
- **U.S.-Soviet détente** and Kissinger's conduct of **"back channel" diplomacy** with the **Soviet leadership**
- **U.S.-China rapprochement**, including initial White House efforts to communicate with Beijing; Kissinger's "secret trip" to China, July 1971; and subsequent **high-level meetings with Chinese officials**, including visits by presidents Nixon and Ford in 1972 and 1975, respectively
- developments in South Asia, including the 1971 **India-Pakistan** war and the Nixon/Kissinger tilt to Pakistan during that crisis
- the **Middle East**, including U.S. conduct during the 1973 **October War** and Kissinger's role as **shuttle diplomat during 1974-1975**
- negotiations to end minority white rule in **Rhodesia/Zimbabwe**
- **U.S.-European relations**, including policy coordination and consultations on a variety of "hot spots" and "hot issues," including the rise of Euro-Communism
- the 1974 **Cyprus** crisis and U.S. relations with **Greece and Turkey**
- international economic, energy and raw materials policies
- relations between the mass media and policymakers

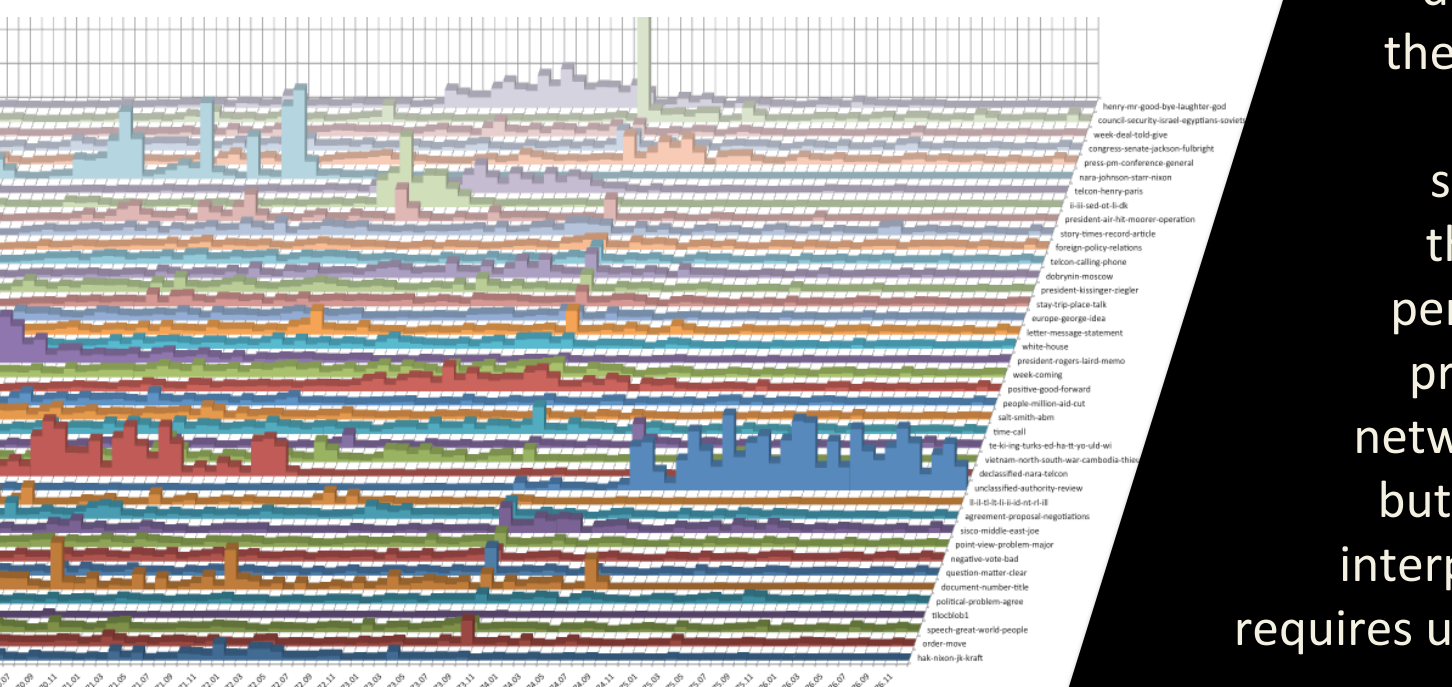
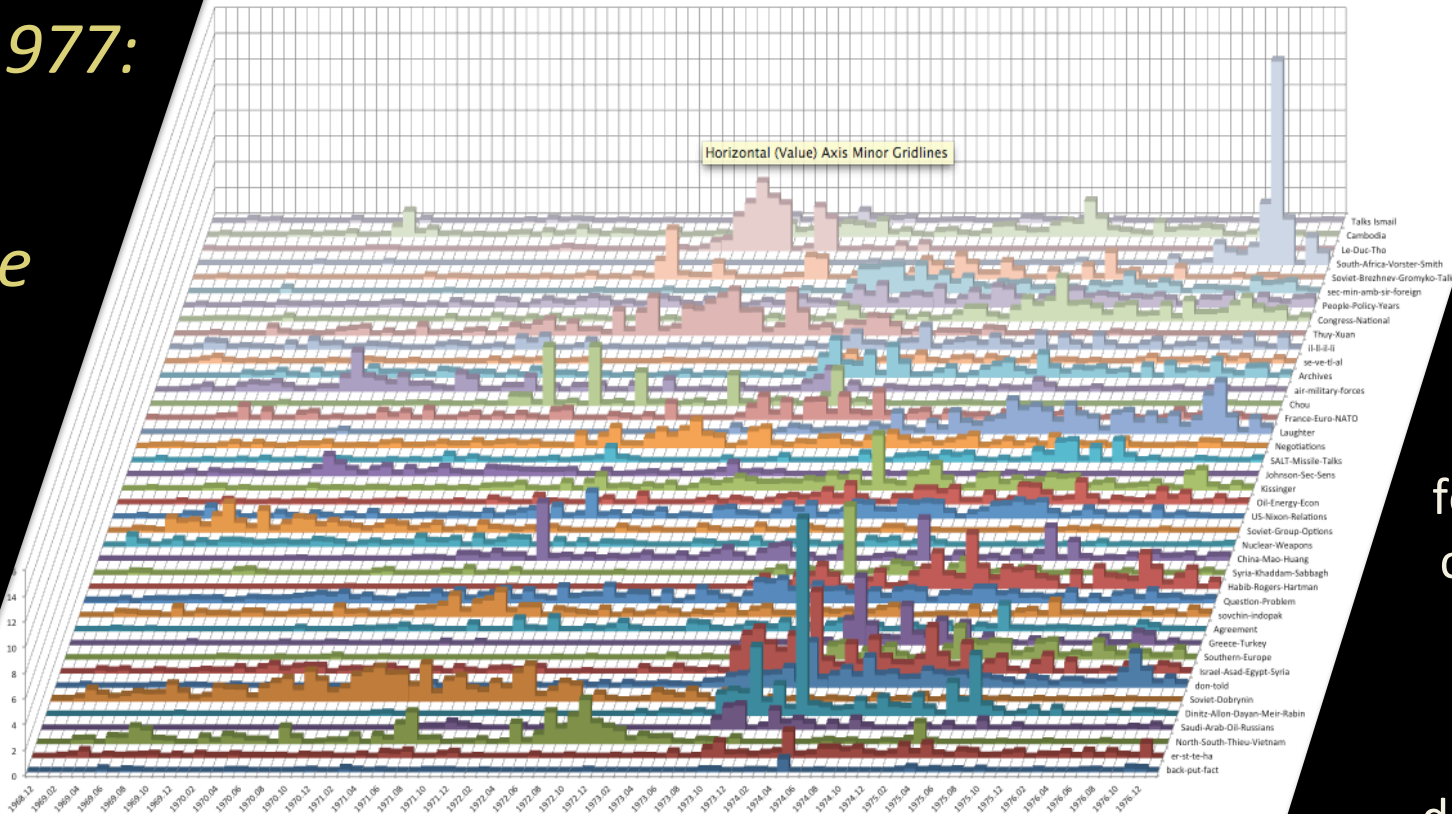
Methodology: 'Topic Modeling,' cont'd.

3. Analyze the changing frequency of topics in relation to the timeline, historical record and/or historiography.



1969-1977:

The 'Longue Duree' views



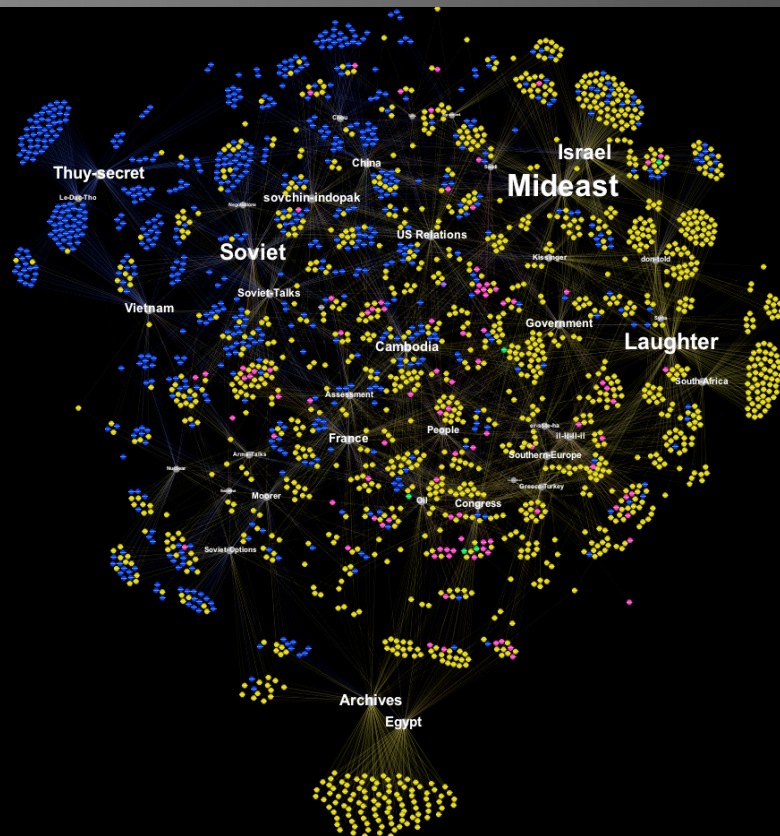
One combined view of all topics across the time span, for both forms of correspondence.

Such a 'high-altitude' distant reading of the material can be informative: it suggests possible themes, subjects, periodizations, and prompts a closer / networked reading – but it also demands interpretive rigor and requires user interactivity.

Interactive and/or Dynamic Visualizations

Work with the dimensions of the visualization to help unpack / understand the data (space, time, metadata).

<http://www.quantifyingkissinger.com>



For example, combining an analysis of the text and metadata in an animated display allows one to ask questions about the shifting relationship between subject matter, status and secrecy.

Methodology: Word Collocation Analysis, Target Word: 'Bombing'

AntConc 3.3.5m (Macintosh OS X) 2012

Corpus Files

- 00001_OCR.pdf.txt
- 00002_OCR.pdf.txt
- 00003_OCR.pdf.txt
- 00004_OCR.pdf.txt
- 00005_OCR.pdf.txt
- 00006_OCR.pdf.txt
- 00007_OCR.pdf.txt
- 00008_OCR.pdf.txt
- 00009_OCR.pdf.txt
- 00010_OCR.pdf.txt
- 00011_OCR.pdf.txt
- 00012_OCR.pdf.txt
- 00013_OCR.pdf.txt
- 00014_OCR.pdf.txt
- 00015_OCR.pdf.txt
- 00016_OCR.pdf.txt
- 00017_OCR.pdf.txt
- 00018_OCR.pdf.txt
- 00019_OCR.pdf.txt
- 00020_OCR.pdf.txt
- 00021_OCR.pdf.txt
- 00022_OCR.pdf.txt
- 00023_OCR.pdf.txt
- 00024_OCR.pdf.txt
- 00025_OCR.pdf.txt
- 00026_OCR.pdf.txt
- 00027_OCR.pdf.txt
- 00028_OCR.pdf.txt
- 00029_OCR.pdf.txt
- 00030_OCR.pdf.txt
- 00031_OCR.pdf.txt
- 00032_OCR.pdf.txt
- 00033_OCR.pdf.txt
- 00034_OCR.pdf.txt
- 00035_OCR.pdf.txt
- 00036_OCR.pdf.txt
- 00037_OCR.pdf.txt
- 00038_OCR.pdf.txt
- 00039_OCR.pdf.txt
- 00040_OCR.pdf.txt
- 00041_OCR.pdf.txt
- 00042_OCR.pdf.txt
- 00043_OCR.pdf.txt
- 00044_OCR.pdf.txt
- 00045_OCR.pdf.txt
- 00046_OCR.pdf.txt

Total No. of Collocate Types: 75 Total No. of Collocate Tokens: 305

| Rank | Freq | Freq(L) | Freq(R) | Stat | Collocate |
|------|------|---------|---------|---------|---------------|
| 1 | 2 | 0 | 2 | 8.54807 | thereafter |
| 2 | 3 | 0 | 3 | 6.74071 | bombing |
| 3 | 2 | 0 | 2 | 5.71848 | circumstances |
| 4 | 2 | 0 | 2 | 5.00017 | reached |
| 5 | 2 | 0 | 2 | 4.73847 | open |
| 6 | 2 | 0 | 2 | 4.46337 | hanoi |
| 7 | 2 | 0 | 2 | 3.88737 | continue |
| 8 | 2 | 0 | 2 | 3.57378 | issues |
| 9 | 2 | 0 | 2 | 3.41699 | own |
| 10 | 2 | 0 | 2 | 3.38163 | down |
| 11 | 2 | 0 | 2 | 3.36772 | questions |
| 12 | 2 | 0 | 2 | 3.20628 | authority |
| 13 | 2 | 0 | 2 | 3.04718 | north |
| 14 | 6 | 0 | 6 | 3.01635 | when |
| 15 | 2 | 0 | 2 | 2.95756 | negotiations |
| 16 | 4 | 0 | 4 | 2.91018 | here |
| 17 | 3 | 0 | 3 | 2.85731 | situation |
| 18 | 3 | 0 | 3 | 2.34910 | exclusively |
| 19 | 4 | 0 | 4 | 2.27640 | u |
| 20 | 2 | 0 | 2 | 2.12381 | states |
| 21 | 2 | 0 | 2 | 2.05204 | point |
| 22 | 5 | 0 | 5 | 1.90383 | dr |
| 23 | 2 | 0 | 2 | 1.87410 | also |
| 24 | 3 | 0 | 3 | 1.84291 | were |

Search Term Words Case Regex Window Span Same

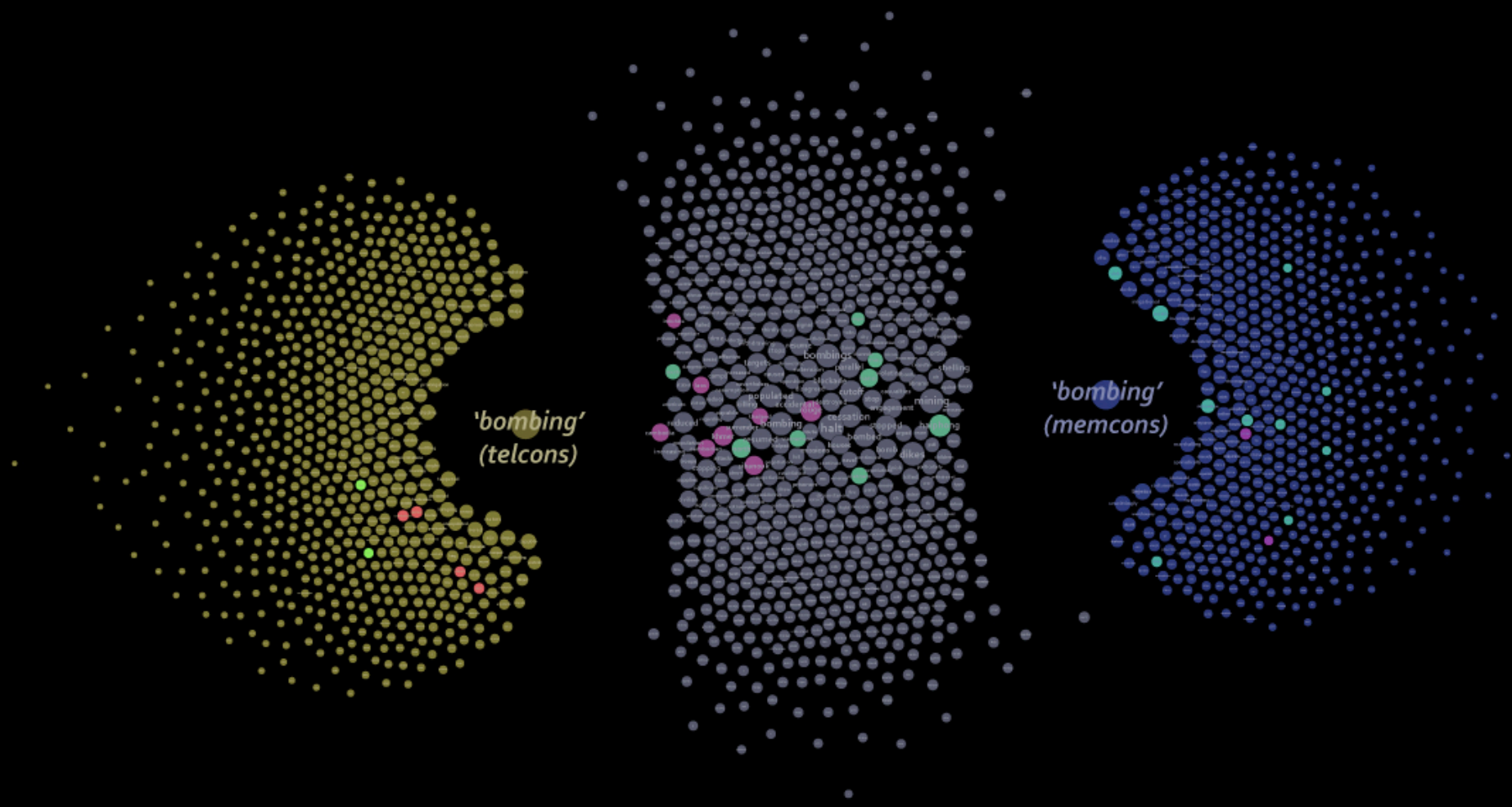
bombing Advanced From... 10R To... 10R

Start Stop Sort Min. Collocate Frequency

Sort by Invert Order 2

Sort by Stat Clone Results

Methodology: Word Collocation Analysis, Target Word: 'Bombing'



The *red dots* represent words related to countries in Indochina other than Vietnam (eg Cambodia, Laos, Thailand, etc.), and the *green dots* represent words related to Vietnam.

Word Collocation Frequency Data

Target Word: "Bombing"

Word Collocation Data

Memcons

of Collocate Types: 1300
of Collocate Tokens: 18387

Telcons

of Collocate Types: 1321
of Collocate Tokens: 20830

Min MI-Score

KT  **2.32658**
KA  **2.36037**

Max MI-Score

KT  **15.65659**
KA  **15.08382**

















by Micki Kaufman

<http://www.quantifyingkissinger.com>

Source: DNSA

TELCONS

| | Documents Containing | Frequency of Occurrence | Number of 20L-20R Collocates (Min 2) | MI-Score Uniqueness of Collocation |
|----------------|----------------------|-------------------------|--------------------------------------|------------------------------------|
| <i>bombing</i> | 317 | 542 | 682 | 11.54599 |

| | | | | | |
|-------------------|------|------|----|---------|-----------|
| <i>le</i> | | | 2 | 5.04971 | 10.09942 |
| <i>duc</i> | | | 0 | 0 | 0 |
| <i>haiphong</i> | | | 7 | 9.96636 | 69.76452 |
| <i>thuy</i> | | | 0 | 0 | 0 |
| <i>hanoi</i> | | | 31 | 8.79663 | 272.69553 |
| <i>ho</i> | | | 5 | 7.91992 | 39.5996 |
| <i>nam</i> | | | 0 | 0 | 0 |
| <i>xuan</i> | | | 0 | 0 | 0 |
| <i>quang</i> | | | 0 | 0 | 0 |
| <i>saigon</i> | | | 3 | 5.94683 | 17.84049 |
| <i>south</i> | | | 25 | 8.03027 | 200.75675 |
| <i>thieu</i> | | | 6 | 6.63674 | 39.82044 |
| <i>tho</i> | | | 0 | 0 | 0 |
| <i>vietnam</i> | 1494 | 1834 | 35 | 7.78736 | 272.5576 |
| <i>vietnamese</i> | | | 30 | 8.35454 | 250.6362 |
| | | | 0 | 0 | 0 |

averages

9 **7.6**

| | | | | | |
|-------------------|-----|-----|----|----------|-----------|
| <i>cambodia</i> | 623 | 931 | 65 | 9.65859 | 627.80835 |
| <i>thai</i> | | | 0 | 0 | 0 |
| <i>laotian</i> | | | 0 | 0 | 0 |
| <i>phnom</i> | | | 4 | 9.27004 | 37.08016 |
| <i>penh</i> | | | 4 | 8.91389 | 35.65556 |
| <i>cambodians</i> | | | 4 | 9.59197 | 38.36788 |
| <i>rouge</i> | | | 2 | 10.49886 | 20.99772 |
| <i>khmer</i> | | | 2 | 10.41139 | 20.82278 |
| <i>thailand</i> | | | 3 | 8.23833 | 24.71499 |
| <i>cambodian</i> | | | 12 | 9.57602 | 114.91224 |
| <i>sihanouk</i> | | | 22 | 10.49068 | 230.79496 |
| <i>pathet</i> | | | 0 | 0 | 0 |
| <i>lao</i> | | | 0 | 0 | 0 |
| <i>laos</i> | | | 23 | 8.81785 | 202.81055 |
| <i>indochina</i> | | | 2 | 7.77094 | 15.54188 |
| <i>phauma</i> | | | 0 | 0 | 0 |

averages

8.9375 **8.6**

MEMCONS

| | Documents Containing | Frequency of Occurrence | Number of 20L-20R Collocates (Min 2) | MI-Score Uniqueness of Collocation |
|--|----------------------|-------------------------|--------------------------------------|------------------------------------|
| | 171 | 483 | 601 | 12.03839 |

| | | | | | |
|--|-----|------|-----|----------|-----------|
| | | | 51 | 6.84867 | 349.28217 |
| | | | 55 | 7.16912 | 394.3016 |
| | | | 10 | 12.14563 | 121.4563 |
| | | | 6 | 5.82291 | 34.93746 |
| | | | 51 | 9.1358 | 465.9258 |
| | | | 0 | 0 | 0 |
| | | | 7 | 7.21365 | 50.49555 |
| | | | 6 | 5.90421 | 35.42526 |
| | | | 6 | 10.01273 | 60.07638 |
| | | | 10 | 6.38723 | 63.8723 |
| | | | 29 | 6.05213 | 175.51177 |
| | | | 19 | 7.64625 | 145.27875 |
| | | | 51 | 6.91691 | 352.76241 |
| | 791 | 7075 | 130 | 8.30548 | 1079.7124 |
| | | | 44 | 7.44658 | 327.64952 |
| | | | 0 | 0 | 0 |

29.6875 **7.6**

| | | | | | |
|--|-----|------|----|----------|-----------|
| | 373 | 2147 | 23 | 7.52708 | 173.12284 |
| | | | 3 | 7.33918 | 22.01754 |
| | | | 0 | 0 | 0 |
| | | | 0 | 0 | 0 |
| | | | 0 | 0 | 0 |
| | | | 0 | 0 | 0 |
| | | | 4 | 10.23874 | 40.95496 |
| | | | 4 | 8.89172 | 35.56688 |
| | | | 4 | 7.59995 | 30.3998 |
| | | | 4 | 7.35395 | 29.4158 |
| | | | 10 | 8.33016 | 83.3016 |
| | | | 5 | 9.7497 | 48.7485 |
| | | | 7 | 8.60286 | 60.22002 |
| | | | 18 | 7.37119 | 132.68142 |
| | | | 3 | 5.58245 | 16.74735 |
| | | | 2 | 9.71408 | 19.42816 |

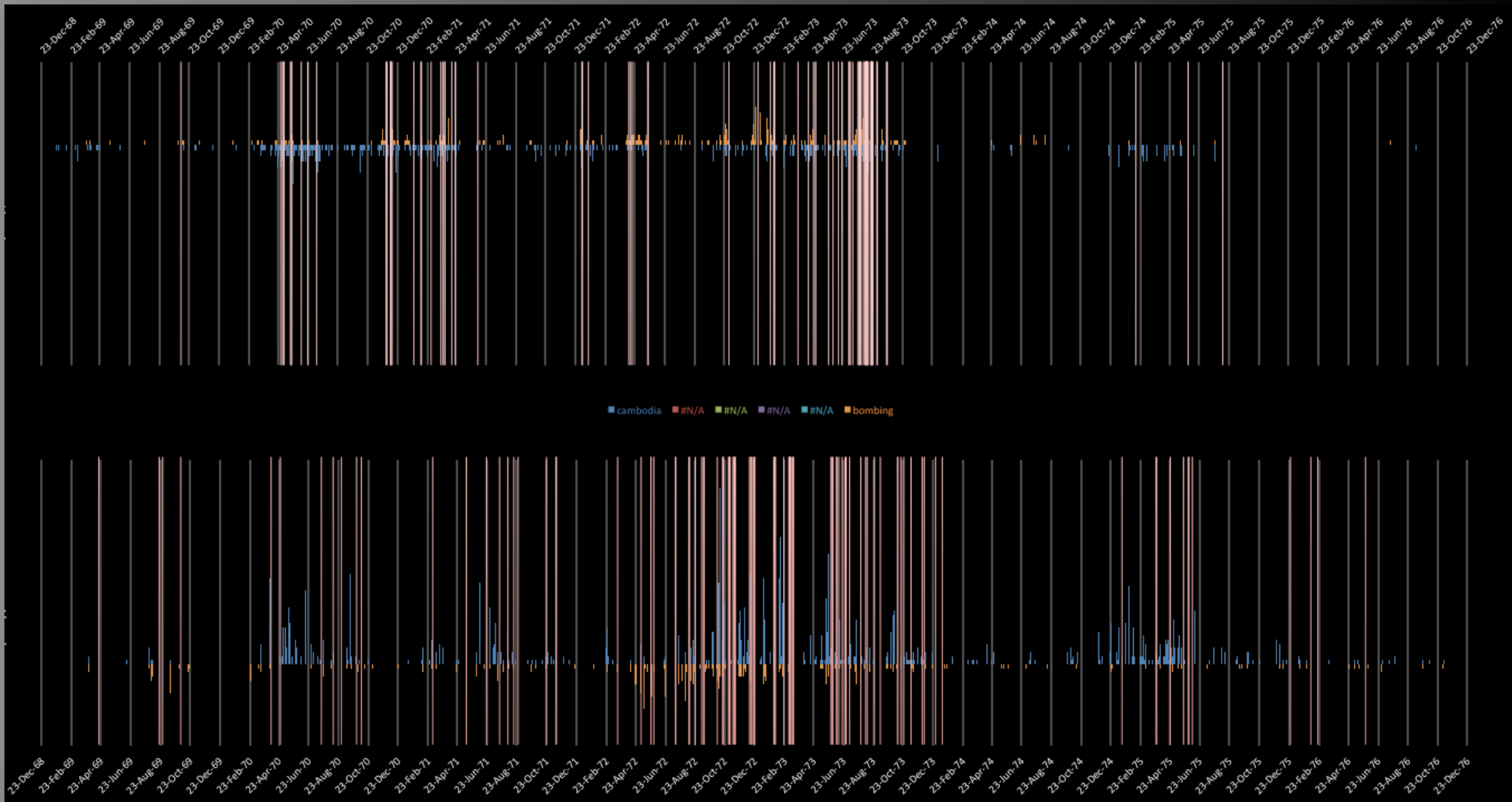
5.4375 **8.2**

Methodology: Word Collocation Frequency, Target Pair: 'Bombing/Cambodia'



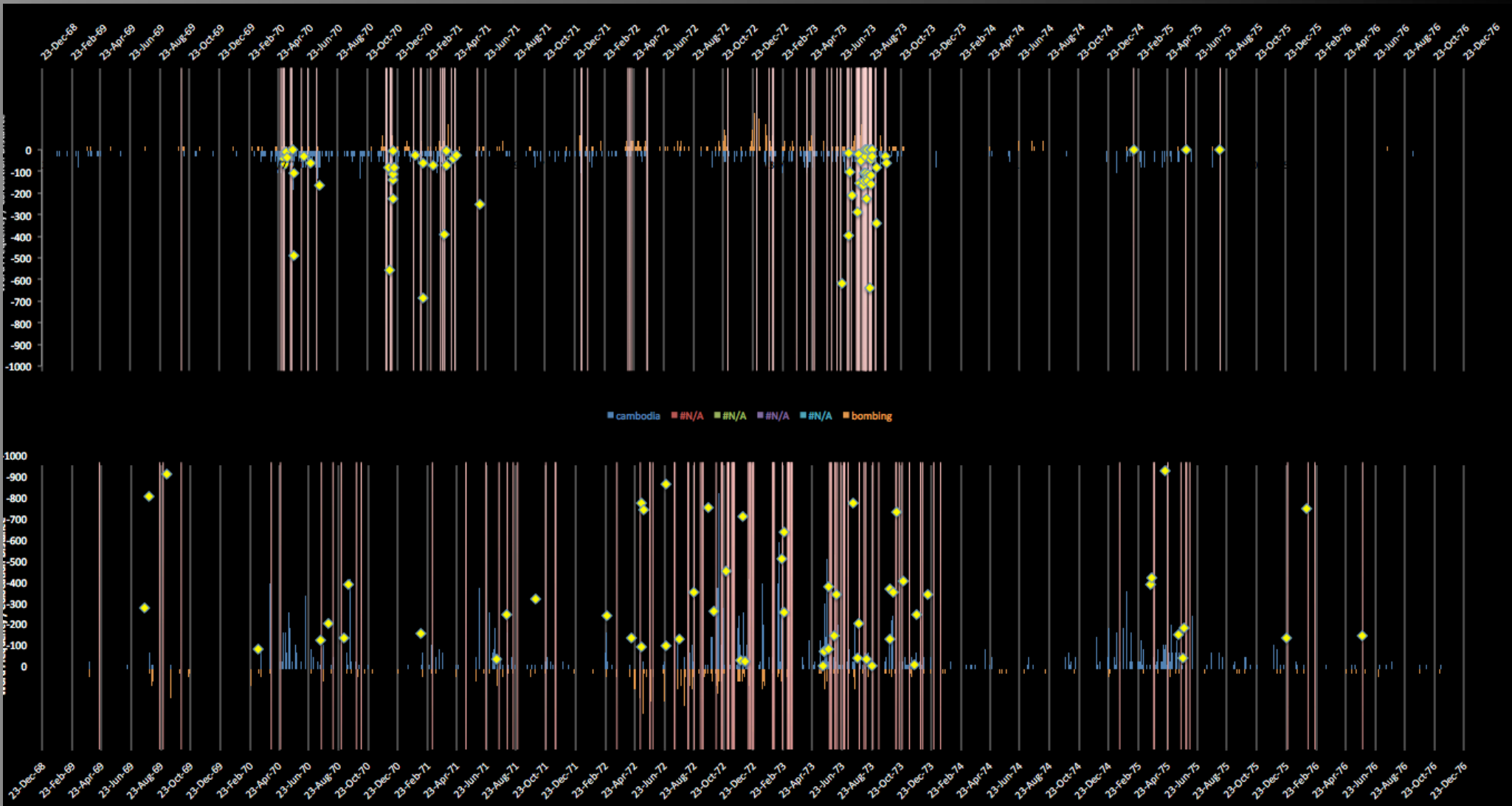
*Frequency over time of the words **cambodia** and **bombing** in the telcons (top) and memcons (bottom)*

Methodology: Word Collocation Frequency, Target Pair: 'Bombing/Cambodia'

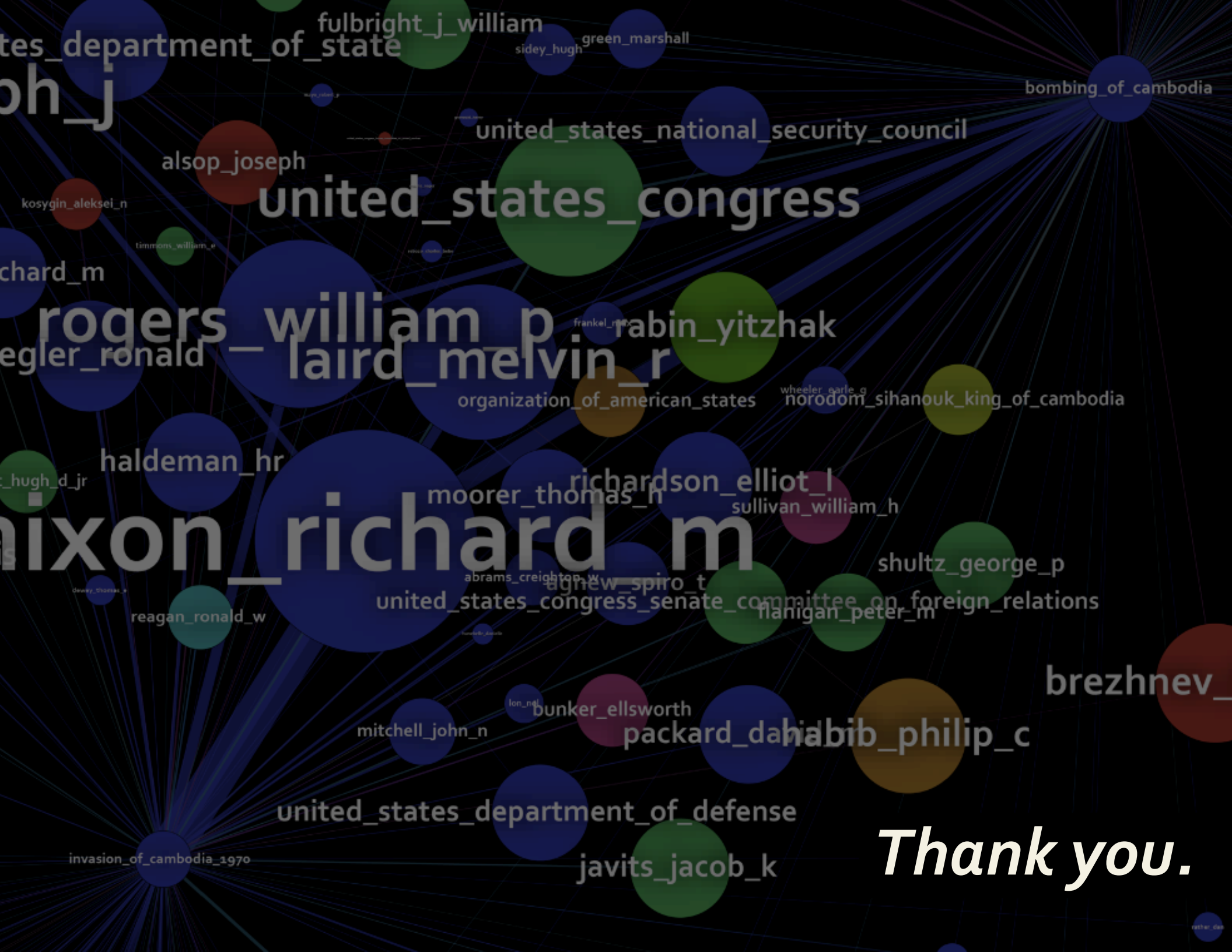


The vertical lines represent instances in the archive where both words occurred in the same document, regardless of collocation distance

Methodology: Word Collocation Frequency, Target Pair: 'Bombing/Cambodia'



The diamonds indicate the specific collocation distance between the words for each document on the timeline 0-1000(!)



tes_department_of_state
fulbright_j_william
sidey_hugh
green_marshall
bombing_of_cambodia
united_states_national_security_council
alsop_joseph
united_states_congress
kosygin_aleksei_n
timmons_william_e
chard_m
rogers_william_p
frankel_max
rabin_yitzhak
egler_ronald
laird_melvin_r
organization_of_american_states
wheeler_earle_g
norodom_sihanouk_king_of_cambodia
haldeman_hr
moorer_thomas_h
richardson_elliott_l
sullivan_william_h
hugh_d_jr
nixon_richard_m
agnew_spiro_t
shultz_george_p
united_states_congress_senate_committee_on_foreign_relations
flanigan_peter_m
reagan_ronald_w
abrams_creighton_w
brezhnev
lon_nel
bunker_ellsworth
packard_daniel
habib_philip_c
mitchell_john_n
united_states_department_of_defense
javits_jacob_k
invasion_of_cambodia_1970

Thank you.