

The Transformation of the Concept of Terrorism - A Text Analysis from New York Times 1850-1980

Annual Meeting of American Society of Criminology

Yu-Hsuan Liu & Tzu-Ying Lo,
11/17/2021, Chicago



Literature Review

The Definition of Terrorism (Hoffman, 2006)

- Vague idea or impression for most people
- Modern media:
 - Complex and convoluted messages
 - Labelling of a range of violent acts as terrorism

Literature Review

- The Transformation of the Definition of “Terrorism” (Hoffman, 2006)
 - The **end of the 18th** century:
 - positive connotation, virtue, revolutionary, and democracy.
 - In the **1930s**:
 - Revolutionary movement and violence against the government and their leaders.
 - Following **WWII**:
 - Revolutionary, national liberation and self-determination.
 - In the **1960s and 1970s**:
 - Nationalism and separatism outside of colonial oppression.

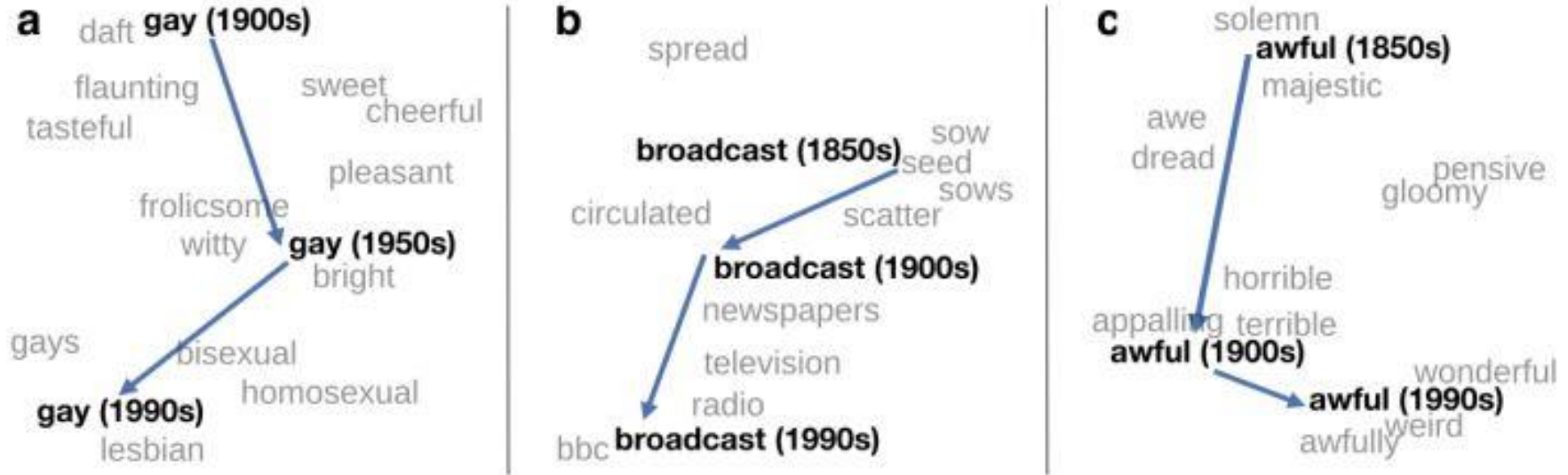
What Do We Know?

- Public's perception of “terrorism” was affected by media (Nellies and Savage, 2012; Mitnik et al. , 2018)
- The conception of terrorism kept changing (Hoffman, 2006)
- The definition of “terrorism” is conceptualized by counting the frequencies of its elements (Schmid et al., 1988) from scholars' definitions.

What Gap Can We Bridge?

- Few studies of media and terrorism use news texts before 1970. (Mitnik et al. , 2018; Jetter, 2019; Chermak and Gruenewald, 2006)
- Few supporting evidence of Hoffman's₍₂₀₀₆₎ argument of the transformation of the definition of terrorism.

What Analysis Can We Perform?



- Hamilton et al. (2016)

Research Questions/Hypothesis

- Can/How can we use Natural Language Processing (NLP) to ascertain how the conception of "Terrorism" has evolved through time?
- Hypothesis: If Hoffman's assertion is correct, the results of NLP analysis of terrorism-related news texts from the years 1851-1900, 1901-1930, 1931-1950, 1951-1960, and 1961-1980 would be dissimilar.

Why New York Times News Texts?

- Consistency
- Provide news resources back to 1850s with news under the category of “terrorism”
- Convenient to be downloaded in a database

NLP Methods

Data
Retrieval

Pre-
processing

Tokenization

Stemming

Vectorization

Train

Test and
Validation

Classification

Data: 1850-1980 NYT News in the Terrorism Category

Terrorism

Latest

Search

61216 results sorted by **oldest**

July 2, 1918

WARNS AGAINST POGROMS.; Bolshevik Commission Says Reactionaries Foment Anti-Semitism.
Bolshevist Govt warns against pogroms by counter-revolutionaries

July 2, 1918

BOURTSEFF BACKS EX-CZAR'S BROTHER; Former Nihilist Says Michael Is Rallying Forces That Can Save Russia. **GEN. ALEXIEFF TO LEAD THEM** Korniloff and Kaledine Co-operating for the Restoration or Real Democracy. **CZECHS IN SIBERIA OUST RED COUNCIL** Dissolve Bolshevik Soviet at Vladivostok--Duke Michael's Manifesto. **BOURTSEFF BACKS EX-CZAR'S BROTHER**
Bolshevist Council at Vladivostok ousted
Copyright, 1918, by COPYRIGHT, 1918, BY THE NEW YORK TIMES COMPANY. SPECIAL CABLE TO THE NEW YORK TIMES.

July 4, 1918

SOVIETS TAKE INDUSTRIES.; Nationalize All Big Enterprises-- Fear Menshevist Terrorism.



SAY HUNGARIAN REDS MASSACRED 3,000

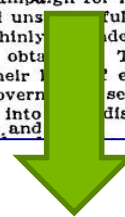
Bloodthirsty Revenge on Apathetic Village People Reported by Anti-Bolshevist Bureau.

COULD NOT GET RECRUITS

Feeble Force of Farmers, with Their Families, Mowed Down with Cannon and Machine Guns.

Copyright, 1918, by Chicago Tribune Co.
VIENNA, June 7, (via Paris, June 15.)
—According to the anti-Bolshevist Press Bureau, the Hungarian Government Bolsheviks have just massacred 3,000 people in the village of Kohlshof, near Oedenburg, in Western Hungary, near the Austrian border.

The statement says that for more than a month the Bolsheviks have been carrying on a campaign for recruits in that neighborhood unsuccessfully. The meetings were thinly attended and no recruits were obtained. To punish the people for their lack of enthusiasm the Budapest Government sent a body of Lenin boys into the district, requisitioning food and



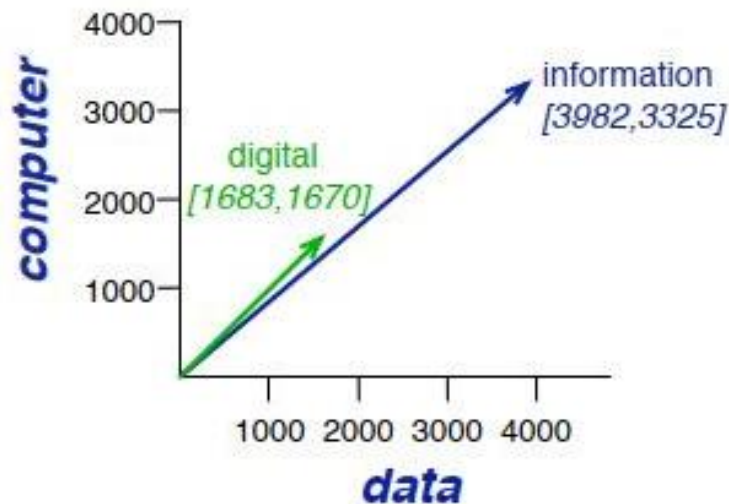
Data



year	month	day	pubdate_y	text
1851	November	24	Nov. 24, 1851	benevolence,â to aid the cause of J y i...
1851	November	27	Nov. 27, 1851	did not} break the tie between e...
1851	December	16	Dec. 16, 1851	ambition 3 he flattered them with the dea...
1851	December	18	Dec. 18, 1851	German Radicalism. New Work Daily Times (18...
1851	December	29	Dec. 29, 1851	is lure in establishing freedom to the peop...
1852	January	10	Jan. 10, 1852	I thâ™ and the duties of a â
1853	May	7	May 7, 1853	a impression. : logical of the equality tha...
1853	June	18	June 18, 1853	urope is foolish enough to take the Course,...
1855	February	1	Feb. 1, 1855	These are my jewels.â a of He A e a A ...
1855	April	26	April 26, 1855	Is your name Lorenzo pg Days?â â
1856	January	25	Jan. 25, 1856	brownâ or mulattoes have led to most of the...
1856	March	17	March 17, 1856	in small parties to eton horses, and not lo...
1856	March	26	March 26, 1856	I blushed when I read it,â and prompted Or...
1856	August	13	Aug. 13, 1856	claims,â erect their log houses, and form ...
1857	January	17	Jan. 17, 1857	Of any person shall the upon himself any of...
1857	January	27	Jan. 27, 1857	Kansas AffairsMessage of Governor Weary. ...
1857	February	27	Feb. 27, 1857	Terroristâ at the not election, all will...

NLP Methods - Vectorization

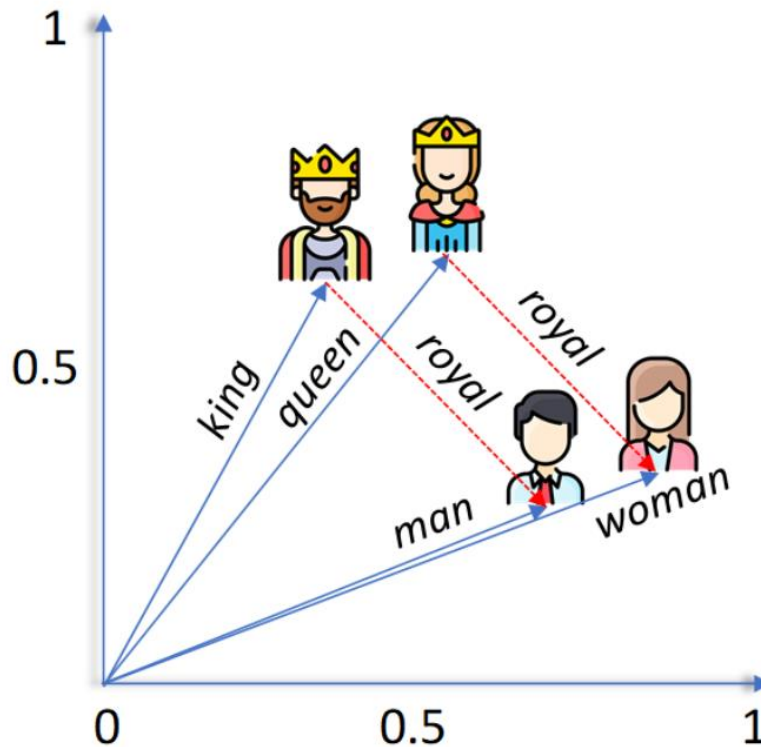
	aardvark	...	computer	data	result	ple	sugar	...
cherry	0	...	2	8	9	442	25	...
strawberry	0	...	0	0	1	60	19	...
digital	0	...	1670	1683	85	5	4	...
information	0	...	3325	3982	378	5	13	...



Methods

- Word2Vec

Preserves the relationship between words, and deals with addition of new words in the vocabulary

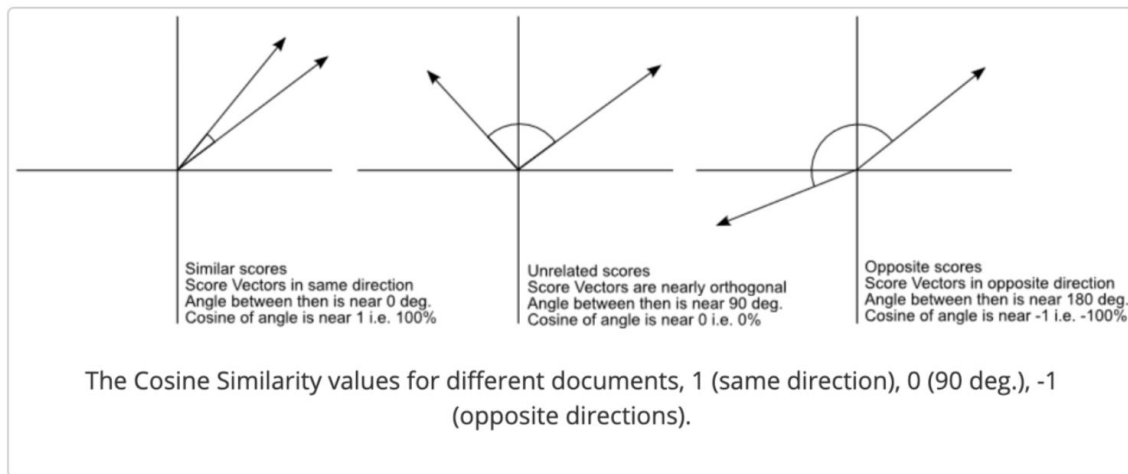


Z_ai (2020), Deep Learning for NLP: Word Embeddings.
<https://towardsdatascience.com/deep-learning-for-nlp-word-embeddings-4f5c90bcdab5>

Methods

- Cosine Similarity

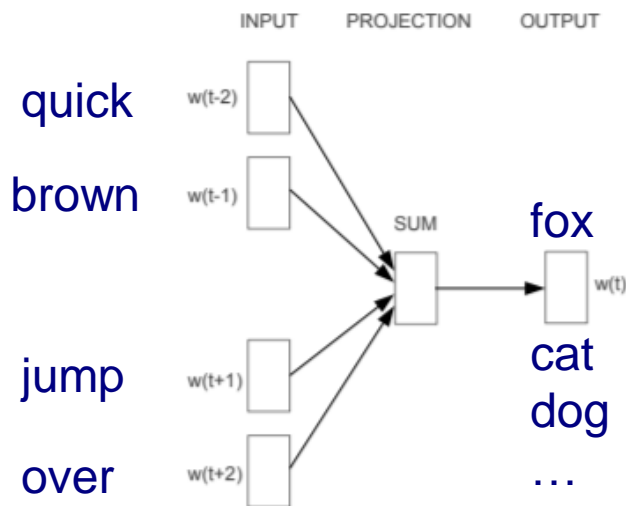
$$\text{Cos}\theta = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|} = \frac{\sum_1^n a_i b_i}{\sqrt{\sum_1^n a_i^2} \sqrt{\sum_1^n b_i^2}}$$



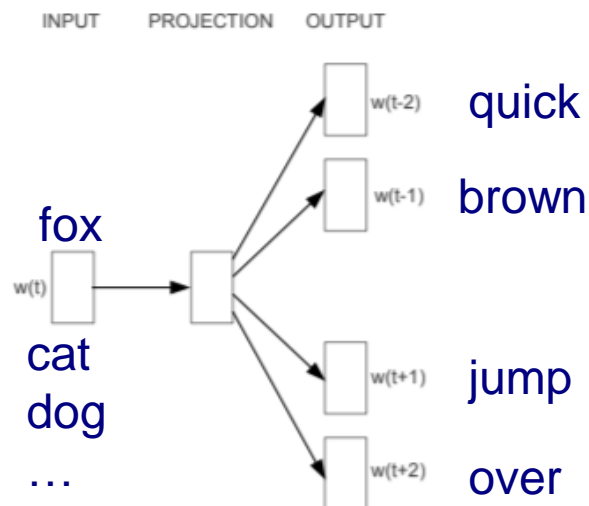
Christian S. Perone (2013), Machine Learning :: Cosine Similarity for Vector Space Models (Part III).
<https://blog.christianperone.com/2013/09/machine-learning-cosine-similarity-for-vector-space-models-part-iii/>

Methods

- Embedding Models (Mikolov et al., 2013)



Continuous Bag-of-Words Model(CBOW):
Predicts the current word based on the context



Skip-gram: Predicts surrounding words given the
current word

Descriptive Statistics

Years	Numbers of Articles	Numbers of Tokens
1851-1900	885	502,212
1901-1930	2,126	664,334
1931-1950	3,665	902,281
1951-1960	1,885	368,161
1961-1980	14,104	3,917,863

token	raw_frequency	probability(%)
peopl	2508	0.577085
govern	2314	0.532446
parti	1740	0.40037
republican	1547	0.355961
south	1445	0.332491
countri	1422	0.327199
war	1408	0.323978
everi	1303	0.299817
nation	1288	0.296366
citi	1265	0.291074
democrat	1211	0.278648
hous	1197	0.275427
power	1197	0.275427
polit	1189	0.273586
order	1181	0.271745
elect	1169	0.268984
public	1137	0.261621
vote	1121	0.25794
senat	1101	0.253338
gener	1099	0.252877
present	1092	0.251267
act	1073	0.246895
give	1046	0.240682
mani	1027	0.23631
come	1011	0.232629

1901-1930



token	raw_frequency	probability(%)
govern	4731	0.818882
war	2943	0.5094
russia	2793	0.483436
peopl	2388	0.413335
histor	2206	0.381833
russian	2158	0.373525
polic	2119	0.366775
gener	2096	0.362794
terrorist	2060	0.356562
german	1984	0.343408
class	1919	0.332157
countri	1785	0.308963
special	1750	0.302905
organ	1640	0.283865
nation	1623	0.280923
street	1621	0.280577
mani	1616	0.279711
polit	1609	0.278499
citi	1603	0.277461
act	1518	0.262748
parti	1506	0.260671
order	1452	0.251325
soviet	1428	0.24717
public	1380	0.238862
present	1356	0.234708

1931-1950



token	raw_frequency	probability(%)
govern	6398	0.818667
terrorist	5694	0.728586
british	4525	0.579004
war	4153	0.531404
histor	3816	0.488283
jewish	3490	0.446569
german	3342	0.427631
polic	3219	0.411893
nation	3026	0.387197
gener	2789	0.356871
polit	2765	0.3538
act	2648	0.33883
special	2644	0.338318
peopl	2556	0.327057
organ	2438	0.311959
arab	2436	0.311703
world	2422	0.309911
parti	2292	0.293277
york	2234	0.285855
presid	2184	0.279458
countri	2157	0.276003
communist	2116	0.270757
japanes	2103	0.269093
class	2091	0.267558
militari	1938	0.24798

1951-1960



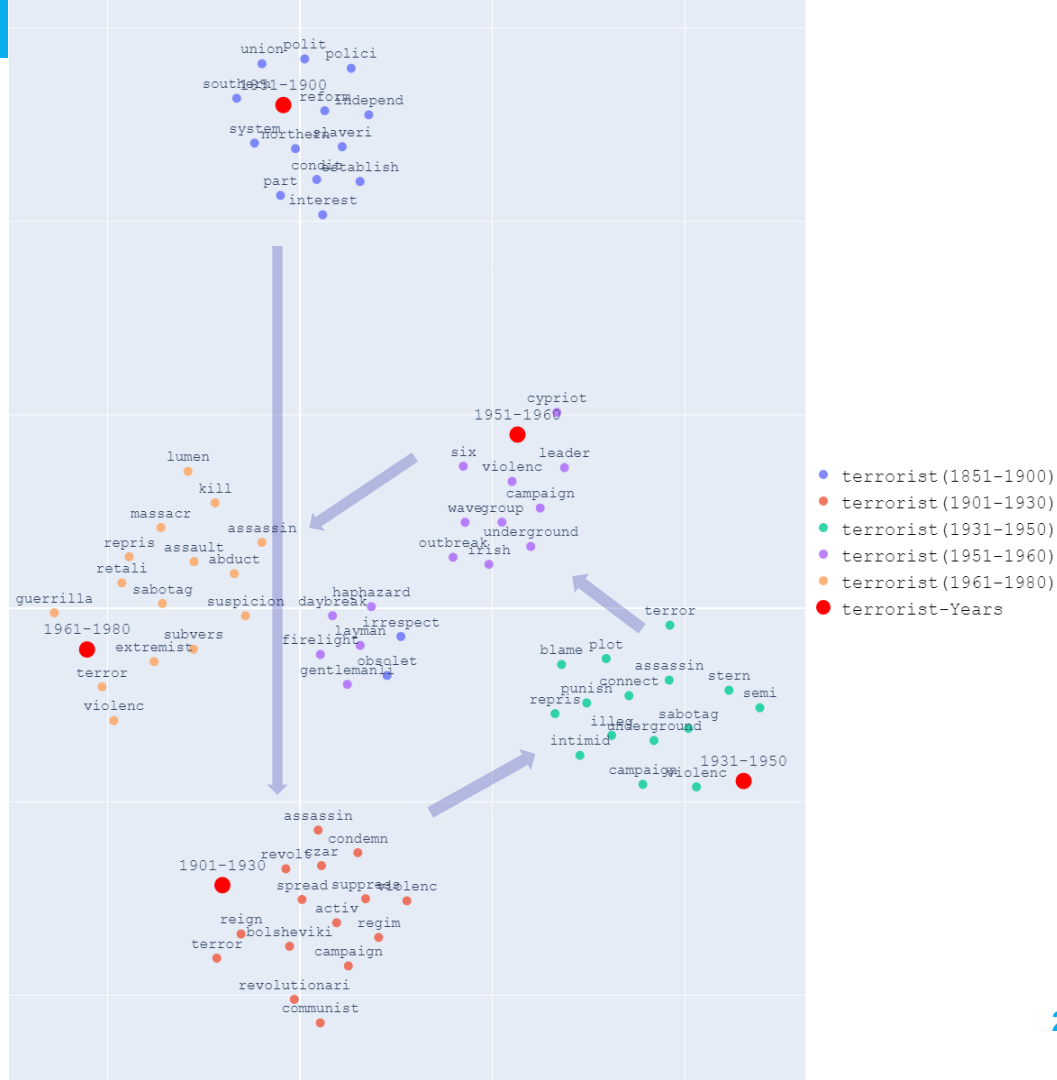
token	raw_frequency	probability(%)
terrorist	3376	1.060761
govern	2632	0.826992
french	2343	0.736186
british	2066	0.649151
special	2000	0.628413
histor	1917	0.602334
communist	1840	0.57814
gener	1430	0.449315
presid	1368	0.429835
polit	1358	0.426692
nation	1342	0.421665
typhu	1331	0.418209
war	1294	0.406583
polic	1200	0.377048
peopl	1034	0.32489
class	1032	0.324261
act	990	0.311064
parti	975	0.306351
soviet	970	0.30478
hit	966	0.303524
militari	955	0.300067
armi	932	0.29284
world	898	0.282157
organ	847	0.266133
union	847	0.266133

1961-1980



token	raw_frequency	probability(%)
terrorist	22699	0.67199
govern	20489	0.606565
israel	18132	0.536787
presid	15188	0.449632
histor	14855	0.439773
special	13978	0.41381
polit	13763	0.407445
age	13227	0.391577
peopl	12889	0.381571
polic	12472	0.369226
nation	11955	0.353921
war	11691	0.346105
south	10166	0.300958
militari	10051	0.297554
world	9404	0.2784
citi	9314	0.275735
gener	9189	0.272035
act	8564	0.253532
armi	8142	0.241039
organ	8045	0.238167
countri	7909	0.234141
west	7767	0.229937
mani	7538	0.223158
come	7510	0.222329
intern	7461	0.220878

Similar Words



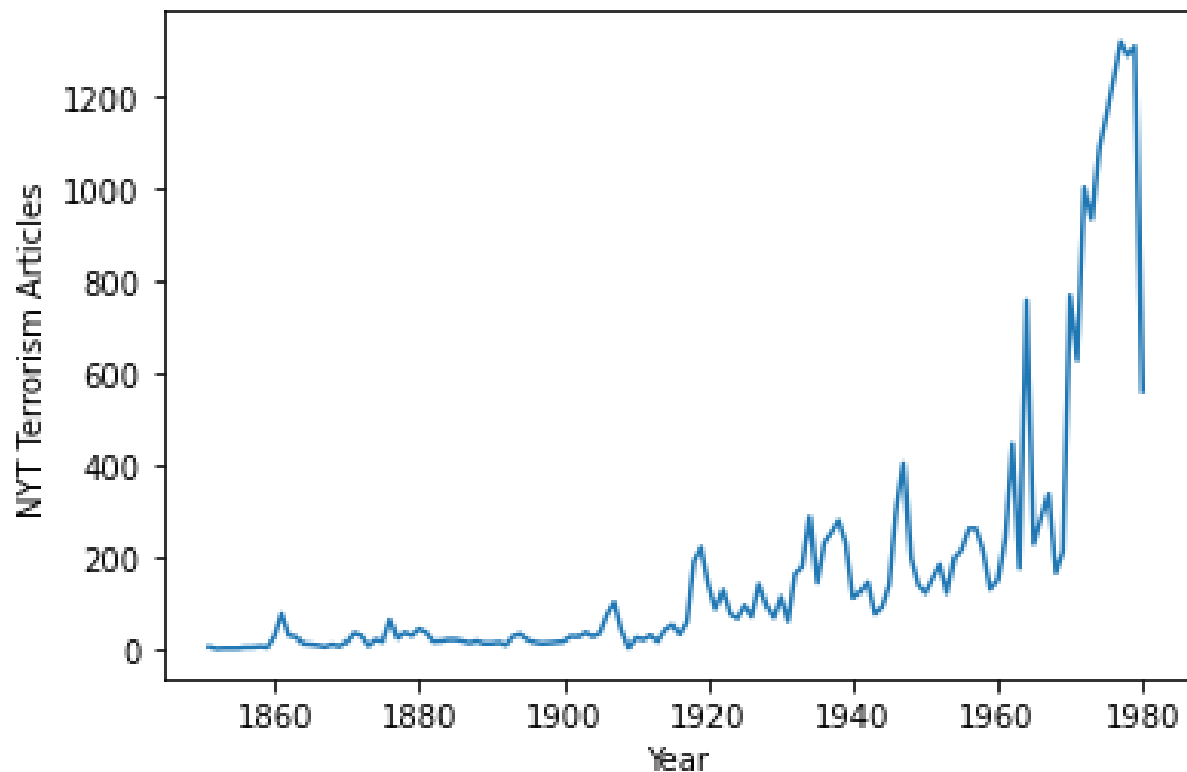
Conclusion

- Provided evidence supporting Hoffman's (2006) claim– the transformation of the concept of terrorism
- The concept of terrorism evolves over time.

Limitation

- Bias in the algorithm
 - When it comes to earlier eras, news articles are far more scarce than they were in the 1960s and 1980s.
 - Some words lose their meaning when they are stemmed.
- Some news are under the wrong categories of NYT news articles.

Limitation



Using the Pre-Trained Data from Google News

Top 15 Similar Words of “terrorist”

terror
terrorists
terrorism
al_Qaeda
Terrorist
Al_Qaeda
extremist
al_Qa'ida
jihadist
al_Qaida
militant
jihadi
Islamic_extremist
al_Qa'eda
Al_Qa'ida

Data resource:
Google Code Archive.
(2013). Word2Vec.
<https://code.google.com/archive/p/word2vec/>

- 言者所以在意，得意而忘言 《莊子·外物》
- Words are for meaning; Once you get the meaning, you can forget the words. (Zhuang Zi)

Appendix -- 22 elements of the definition of

1

Element	Schmid & Jongman Survey (1988) frequency (%)
1. Violence, Force	83.5
2. Political	65
3. Fear, Terror emphasized	51
4. Threat	47
5. Psychological effects and (anticipated) reactions	41.5
6. Victim-Target differentiation	37.5
7. Purposive, Planned, Systematic, Organized action	32
8. Method of combat, strategy, tactic	30.5
9. Extranormality, in breach of accepted rules, without humanitarian constraints	30
10. Coercion, extortion, induction of compliance	28
11. Publicity aspect	21.5
12. Arbitrariness, impersonal, random character, indiscrimination	21
13. Civilians, noncombatants, neutrals, outsiders as victims	17.5
14. Intimidation	17
15. Innocence of victims emphasized	15.5
16. Group, movement, organization as perpetrator	14
17. Symbolic aspect, demonstration to others	13.5
18. Incalculability, unpredictability, unexpectedness of occurrence of violence	9
19. Clandestine, covert nature	9
20. Repetitiveness, serial or campaign character of violence	7
21. Criminal	6
22. Demands made on third parties	4

Resource:

Weinberg et al. (2004)

References

- Bruce Hoffman. (2006). *Inside Terrorism*, Revised and Expanded Edition. Columbia University Press.
- Schmid, A. P., Jongman, A. J., & Stohl, M. (1988). *Political terrorism : a new guide to actors, authors, concepts, data bases, theories, and literature* (Rev., expanded and updated ed.). North-Holland Pub. Co.
- Weinberg, L., Pedahzur, A., & Hirsch-Hoefler, S. (2004). The challenges of conceptualizing terrorism. *Terrorism and Policial Violence*, 16(4), 777-794.
- Mitnik, Z. S., Freilich, J. D., & Chermak, S. M. (2018). Post-9/11 Coverage of Terrorism in the New York Times. *Justice Quarterly*, 1–25.
- Jetter, M. (2019). More Bang for the Buck: Media Coverage of Suicide Attacks. *Terrorism and Political Violence*, 31(4), 779–799.
- Nellis, A. M., & Savage, J. (2012). Does Watching the News Affect Fear of Terrorism? The Importance of Media Exposure on Terrorism Fear. *Crime & Delinquency*, 58(5), 748–768.
- Chermak, S. M., & Gruenewald, J. (2006). The Media's Coverage of Domestic Terrorism. *Justice Quarterly*, 23(4), 428–461.
- Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016). Diachronic word embeddings reveal statistical laws of semantic change. arXiv preprint arXiv:1605.09096.
- Google Code Archive. (2013). Word2Vec. Retrieved from <https://code.google.com/archive/p/word2vec/>
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781.