



Christopher Nolan’s Oppenheimer: A Big Data Analysis

Namkil Kang

Far East University South Korea

| Received: 24.09.2023 | Accepted: 27.09.2023 | Published: 05.10.2023

*Corresponding author: Namkil Kang

Far East University South Korea

Abstract

This paper aims to analyze 29 articles of Google written from 2022 to 2023 regarding Christopher Nolan’s Oppenheimer. A point to note is that the 7-word sentence was the most frequently used, followed by the 6-word sentence, the 5-word sentence, and the 4-word sentence. A further point to note is that in the word cloud, the word Oppenheimer was the most pivotal in 29 articles, followed by the name Nolan, the word film, and the word bomb, in descending order. A major point of this paper is that topic 16 was the most widely used, followed by topic 1, topic 11, and topic 2, in that order. With respect to 18 topics, it is worthwhile noting that the keyword Oppenheimer is linked to ten topics, thus counting as the most pivotal. When it comes to the frequency of pivotal words, the keyword Oppenheimer was the most frequently used, followed the keyword film, the keyword Nolan, and the keyword bomb, in descending order. Finally, this paper argues that block modelling makes networks much more simple and easier by abbreviating them and grouping them.

Keywords: Oppenheimer, Nolan, big data, NetMiner, topic, keyword, block modeling

1. Introduction

The main goal of this paper is to analyze 29 articles of Google written from 2022 to 2023 regarding Christopher Nolan’s Oppenheimer. Our research was carried out by the software package NetMiner. First, we aim to explore the word formation of nouns, their frequency and their proportion. Also, attention is paid to the frequency of 29 documents. Second, we aim at searching into the word cloud which represents 29 articles of Google. The word cloud enables us to find out which words are more central and pivotal. Third, we attempt to inquire into 18 topics that showed up in 29 articles of Google and their keywords. Also, we attempt to look into the frequency of 18 topics and explore which topics are the preferable ones for authors. We also aim at probing into the map of 18 topics and keywords related to them. Fourth, we aim to consider how frequently central keywords turned up in 29 articles of Google. Finally, we aim at investigating the block modelling of 29 articles in which the relevant words form a group. In the block modelling, a group becomes a node by abbreviating networks, which enables us to interpret them easily. Simply put, block modelling makes networks much more simple and easier.

In section 2,1, we aim to explore the word formation of nouns and the frequency of documents. Table 1 shows how many nouns form a sentence:

Table 1 Sentence formation of nouns

Value	Frequency	Proportion	Cumulative Proportion
2.0	23	0.009	0.009
3.0	122	0.047	0.056
4.0	294	0.113	0.169
5.0	322	0.124	0.293
6.0	334	0.129	0.422
7.0	340	0.131	0.553
8.0	286	0.11	0.664
9.0	229	0.088	0.752

2. Results

2.1. Words and documents

10.0	193	0.074	0.826
11.0	120	0.046	0.873
12.0	79	0.03	0.903
13.0	51	0.02	0.923
14.0	41	0.016	0.939
15.0	28	0.011	0.949
16.0	20	0.008	0.957
17.0	15	0.006	0.963
18.0	16	0.006	0.969
19.0	13	0.005	0.974
20.0	7	0.003	0.977
21.0	5	0.002	0.979
22.0	4	0.002	0.98
23.0	4	0.002	0.982
24.0	6	0.002	0.984
25.0	3	0.001	0.985
26.0	4	0.002	0.987
27.0	3	0.001	0.988
28.0	2	0.001	0.989
29.0	5	0.002	0.991
30.0	3	0.001	0.992
31.0	2	0.001	0.993
32.0	5	0.002	0.995
33.0	1	0	0.995
34.0	1	0	0.995
35.0	1	0	0.996
36.0	2	0.001	0.997
37.0	1	0	0.997
38.0	2	0.001	0.998
44.0	1	0	0.998
45.0	1	0	0.998
46.0	1	0	0.999
53.0	1	0	0.999

56.0	1	0	1
72.0	1	0	1
Total	2593	1	

It is important to mention that when the 7-word expression forms a sentence, its frequency is 340 tokens (the highest). Notice that its proportion and cumulative proportion are 0.131 and 0.553, respectively. It is worthwhile noting that the sentence formation of the 7-word expression is followed by that of the 6-word expression. More specifically, when the 6-word expression forms a sentence, its frequency is 334 tokens. We note that its proportion and cumulative proportion are 0.129 and 0.422, respectively. It must be pointed out that the sentence formation of the 6-word expression is followed by that of the 5-word expression. When the latter forms a sentence, its frequency is 322 tokens. Note that they account for 12.4% (the third highest). It is interesting to mention that when the 4-word expression forms a sentence, its frequency is 294 tokens (the fourth highest). It therefore seems clear that the 7-word sentence was the most frequently used, followed by the 6-word sentence, the 5-word sentence, and the 4-word sentence, in that order.

Now attention is paid to the frequency of documents. Table 2 shows the frequency of each article and its proportion:

Table 2 Frequency of each article

Value	Frequency	Proportion	Cumulative Proportion
Document 1	41	0.032	0.032
Document 10	17	0.013	0.045
Document 11	60	0.046	0.091
Document 12	115	0.089	0.18
Document 13	112	0.086	0.266
Document 14	19	0.015	0.281
Document 15	12	0.009	0.29
Document 16	31	0.024	0.314
Document 17	51	0.039	0.354
Document 18	50	0.039	0.392
Document	26	0.02	0.412

19			
Document 2	28	0.022	0.434
Document 20	71	0.055	0.489
Document 21	19	0.015	0.503
Document 22	64	0.049	0.553
Document 23	15	0.012	0.564
Document 24	60	0.046	0.611
Document 25	16	0.012	0.623
Document 26	44	0.034	0.657
Document 27	49	0.038	0.695
Document 28	6	0.005	0.7
Document 29	20	0.015	0.715
Document 3	43	0.033	0.748
Document 4	40	0.031	0.779
Document 5	80	0.062	0.841
Document 6	50	0.039	0.88
Document 7	53	0.041	0.92
Document 8	97	0.075	0.995
Document 9	6	0.005	1
Total	1295	1	

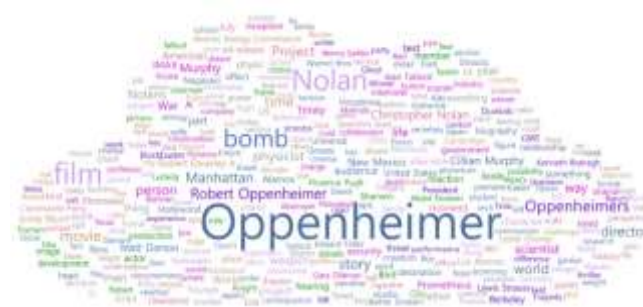
It significant to note that article 12 includes 115 sentences. It obtains the highest frequency (115 sentences) and the highest proportion (8.9%). That is to say, 115 sentences that showed up in article 12 account for 8.9%. It is quite interesting to mention that article 12 is followed by article 13. When it comes to the latter, it includes 112 sentences. Interestingly, its frequency is 112 (112 sentences) and they account for 8.6%. It must be stressed that 71 sentences consist of article 20 (the third highest). 71 sentences that constitute article 20 account for 5.5%. It is worth observing that 64 sentences are made up of article 22. 64 sentences that showed up in article 22 account for 4.9%. It is worth pointing out that 60

sentences constitute article 24. Its frequency is 60 and they account for 4.6%. It therefore seems clear that article 12 has the highest proportion (8.9%), followed by article 13 (8.6%), article 20 (5.5%), article 22 (4.9%), and article 24 (4.6), in that order.

2.2 A word cloud representing 29 articles

In section 2.2, we aim at searching into a word cloud that shows the outline of 29 articles. Figure 1 shows the picture of pivotal words that turned up in 29 articles regarding Christopher Nolan's *Oppenheimer*:

Figure 1 A word cloud



It is important to note that as shown in Figure 1, the word *Oppenheimer* was represented as the biggest. It amounts to saying that it is the most important and pivotal one of all keywords. It would be unfair not to contend that the word *Oppenheimer* is followed by the word *Nolan*. Quite interestingly, the name *Nolan* is the second biggest. We take this as confirming evidence that this name is one of the most pivotal keywords. It is worthwhile noting that the name *Nolan* is followed by the word *film*. This might be due to the fact that *Oppenheimer* is a 2023 epic biographical film, written and directed by Christopher Nolan. It is worth noticing that as exemplified in Figure 1, the word *bomb* is the fourth biggest. This might be due to the fact that Robert Oppenheimer was the American theoretical physicist, called the father of the atomic bomb. It seems thus appropriate to contend that the word *Oppenheimer* was the most pivotal in 29 articles, followed by the name *Nolan*, the word *film*, and the word *bomb*, in descending order.

2.3 Topics and keywords

Section 2.3. is devoted to searching into 18 topics and their keywords. Table 3 shows 18 topics that showed up in 29 articles and their keywords:

Table 3 Topics and keywords

	1st Keyword	2nd Keyword	3rd Keyword	4th Keyword	5th Keyword
Topic-1	Bomb	test	Trinity	weapon	New Mexico
Topic-2	Project	Manhattan	War	World	II
Topic-3	Oppenheimer	project	Nolan	director	Cillian Murphy
Topic-4	Oppenheimer	movie	work	Nolan	Los
Topic-	Nolan	Oppenh	film	story	Murphy

5		eimer			
Topic-6	Person	security	Nolans	audience	film
Topic-7	Oppenheimer	time	Nolan	Nolans	Murphy
Topic-8	Way	Oppenheimer	world	something	Oppenheimers
Topic-9	Nolan	person	part	director	Murphy
Topic-10	Film	Oppenheimer	scientist	Nolan	Christopher Nolan
Topic-11	Oppenheimers	Oppenheimer	hearing	director	Nolans
Topic-12	Movie	year	story	bomb	Oppenheimer
Topic-13	World	Oppenheimer	scientist	man	bomb
Topic-14	Oppenheimer	life	member	story	war
Topic-15	Robert Oppenheimer	Prometheus	American	Bird	father
Topic-16	Film	time	director	IMAX	team
Topic-17	Bomb	physicist	life	Oppenheimer	cast
Topic-18	Oppenheimer	film	thing	IMAX	Cillian Murphy

Topic-3	33
Topic-4	87
Topic-5	52
Topic-6	56
Topic-7	76
Topic-8	67
Topic-9	86
Topic-10	41
Topic-11	92
Topic-12	70
Topic-13	77
Topic-14	54
Topic-15	71
Topic-16	125
Topic-17	65
Topic-18	49

Noteworthy is that the keywords *Bomb*, *test*, *Trinity*, *weapon*, and *New Mexico* consist of topic 1. Quite interestingly, the word *bomb* is the 1st keyword, which may have happened since Robert Oppenheimer, the theoretical physicist who led the Manhattan Project Laboratory developed the atomic bomb. It is worth pointing out that topic 9 includes the keywords *Nolan*, *person*, *part*, *director*, and *Murphy*. As illustrated in Table 3, the name *Nolan* is the 1st keyword, which may have taken place since Oppenheimer was an epic biographical film written and directed by Christopher Nolan. It is interesting to note that the keywords *World*, *Oppenheimer*, *scientist*, *man*, and *bomb* constitute topic 13. More interestingly, the word *world* is the 1st keyword, which may have taken place since the film *Oppenheimer* focused on Oppenheimer's regret over his role in developing the atomic bomb. Now we turn our attention to Table 4:

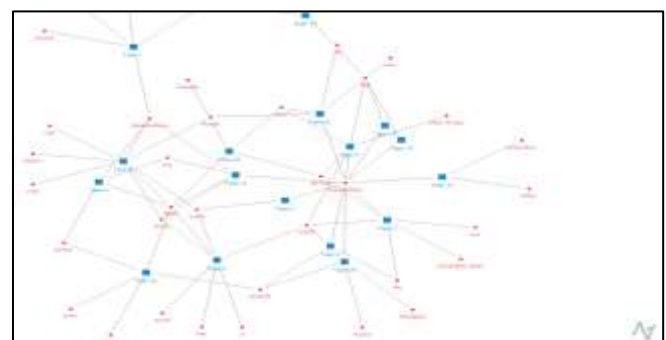
Table 4 Frequency of each topic

	Frequency of each topic
Topic-1	105
Topic-2	89

It is worth mentioning that topic 16 has the highest frequency (125 tokens). More specifically, it turned up 125 times in 29 articles of Google. The keywords *Film*, *time*, *director*, *IMAX*, and *team* constitute topic 16. It is significant that topic 1 showed up 105 times in 29 articles. Topic 1 includes the keywords *Bomb*, *test*, *Trinity*, *weapon*, and *New Mexico*. Note that the frequency of these five keywords is 105 tokens, which are deemed to be the widely used ones in 29 articles of Google. What is interesting is that topic 11 occurred 92 times in 29 articles (the third highest). The keywords *Oppenheimers*, *Oppenheimer*, *hearing*, *director*, and *Nolans* are made up of topic 11. It is appropriate to mention that topic 2 appeared 89 times in 29 articles (the fourth highest). The keywords *Project*, *Manhattan*, *War*, *World*, and *II* consist of topic 2. It is therefore evident that topic 16 was the most widely used, followed by topic 1, topic 11, and topic 2, in that order.

Now attention is paid to Figure 2:

Figure 2 18 topics and their keywords



It is worth mentioning that topic 4, topic 6, topic 9, topic 10, topic 11, and topic 13 have a commonality. That is to say, these six

topics have the keyword *Nolan* in common. This might be due to the fact that the film *Oppenheimer* is a 2023 biographical film written and directed by Christopher Nolan. Particularly noteworthy is that topic 4, topic 9, and topic 13 have one thing in common. These three topics have the keyword *movie* in common. It should be noted that *Oppenheimer* is linked to topic 1, topic 2, topic 3, topic 5, topic 12, topic 13, topic 14, topic 16, topic 17, and topic 18, hence indicating that these ten topics have the keyword *Oppenheimer* in common. This may have taken place since the film *Oppenheimer* was about Oppenheimer's development of the atomic bomb and his regret over his role. Finally, notice that the keyword *world* is linked to topic 3, topic 4, and topic 9. This keyword may have been widely used since it came from world war two. We thus conclude that the keyword *Oppenheimer* is linked to ten topics, thus counting as the most pivotal.

2.4 Frequency of nouns

This section focuses on searching into the frequency of 33 nouns that showed up in 29 articles of Google. Table 5 shows the use of 33 nouns that turned up in 29 articles:

Table 5 Frequency of 33 nouns

Number	Words	Frequency
1	Oppenheimer	338
2	film	174
3	Nolan	173
4	bomb	132
5	movie	66
6	time	59
7	story	58
8	Project	58
9	world	56
10	Robert Oppenheimer	55
11	Manhattan	55
12	director	54
13	person	50
14	physicist	49
15	scientist	44
16	way	43
17	life	43
18	Nolans	43
19	man	39

20	Murphy	38
21	year	37
22	Christopher Nolan	37
23	weapon	35
24	Cillian Murphy	33
25	work	32
26	audience	31
27	war	29
28	Project	29
29	thing	28
30	father	28
31	New Mexico	27
32	American	27
33	moment	25

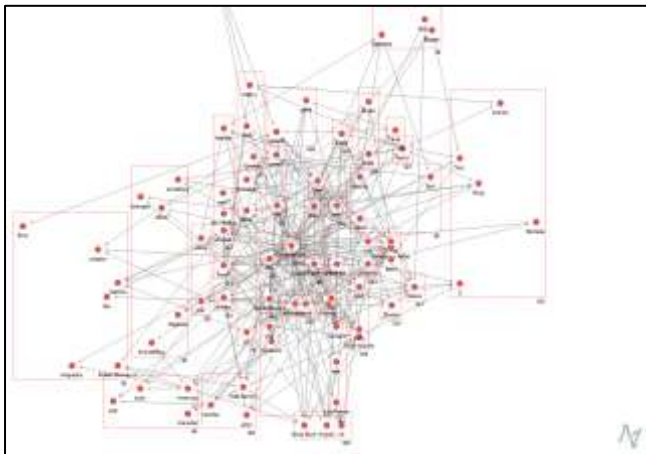
It is important to mention that *Oppenheimer* turned up 338 times in 29 articles of Google. Simply put, this keyword obtains the highest frequency (338 tokens) in 29 articles. This may have happened since the film was about Oppenheimer. It must be emphasized that the keyword *Oppenheimer* is followed by the keyword *film*. More specifically, the latter showed up 173 times in 29 articles. What is interesting is that the keyword *film* is followed by the name *Nolan*. To be more specific, the keyword *Nolan* occurred 173 times in 29 articles. This might be due to the fact that Oppenheimer is a film written and directed by Nolan. It is significant to note that the keyword *bomb* ranks fourth in 29 articles. More specifically, this keyword appeared 132 times in 29 articles of Google. It is very interesting that the keyword *father* showed up 28 times in 29 articles of Google. This might have taken place since Oppenheimer was called the father of the atomic bomb. It therefore seems

reasonable to contend that the keyword *Oppenheimer* was the most frequently used, followed the keyword *film*, the keyword *Nolan*, and the keyword *bomb*, in descending order. We thus conclude that the keyword *Oppenheimer* was the most occurred one in 29 articles of Google.

2.5 Block modelling

This section is focused on searching into the block modelling of 29 articles of Google. In this block modelling, the relevant words of the film *Oppenheimer* occur and something similar belongs to a block. By abbreviating networks, a group can be represented as a node. Most importantly, block modelling makes networks easier. Take a look at block modelling, the map of 29 articles:

Figure 3 Block modeling



It is important to mention that the keyword *Oppenheimer* is located in the central place of the map. This in turn indicates that this keyword is the most pivotal in 29 articles of Google. Perhaps it is worthwhile noting that the keywords *weapon*, *person*, and *life* belong to a block (group 2), as illustrated in Figure 3. These three keywords form a group by having something similar. Quite interestingly, the keywords *hearing*, *part*, and *New Mexico* form a group by having something similar in common. It must be noted that the keywords *something*, *hydrogen*, *effect*, and *Nagasaki* have one thing in common. That is to say, these four keywords are closely related to the atomic bomb. Finally, it is quite interesting to mention that the keywords *cast*, *character*, *American*, and *book* form a cohesive group by having the characters of the film *Oppenheimer* in common. We thus conclude that block modelling makes networks easier by abbreviating them and grouping them. For the map of big data and synonyms, see Kang (2023a, 2023b, 2023c, 2023d, 2023e, 2023f).

3. Conclusion

To sum up, we have analyzed 29 articles of Google written from 2022 to 2023 regarding Christopher Nolan's *Oppenheimer*. In section 2.1, we have argued that the 7-word sentence was the most frequently used, followed by the 6-word sentence, the 5-word sentence, and the 4-word sentence, in that order. In section 2.2, we have further argued that the word *Oppenheimer* was the most pivotal in 29 articles, followed by the name *Nolan*, the word *film*, and the word *bomb*, in descending order. In section 2.3, we have contended that topic 16 was the most widely used, followed by topic 1, topic 11, and topic 2, in that order. We have also maintained that the keyword *Oppenheimer* is linked to ten topics, thus counting as the most pivotal. In section 2.4, we have shown that the keyword *Oppenheimer* was the most frequently used,

followed the keyword *film*, the keyword *Nolan*, and the keyword *bomb*, in descending order. In section 2.5, we have also shown that block modelling makes networks easier by abbreviating them and grouping them.

References

1. Kang, N. (2023a). K-Pop in BBC News: A Big Data Analysis. *Advances in Social Sciences Research Journal*, 10(2), 156-169.
2. Kang, N. (2023b). K-Dramas in Google: A NetMiner Analysis. *Transaction on Engineering and Computing Sciences*, 11(1), 193-216.
3. Kang, N. (2023c). A Comparative Analysis of Tolerate and Put up with in the COCA. *Semiconductor and optoelectronics* 42(1): 1468-1476.
4. Kang, N. (2023d). Sure of and Sure about in Corpora and ChatGPT. *Journal of Harbin Engineering University* 44(7): 1347-1351.
5. Kang, N. (2023e). Turn out adj and Turn out to be adj in the Now Corpus and ChatGPT. *Journal of Harbin Engineering University* 44(8): 825-831.
6. Kang, N. (2023f). Care for and Like in Corpora and ChatGPT. *Semiconductor and optoelectronics* 42(2): 188-198.