

## Extracting Topics from The News using LDA model for 2023-06-09

[illegible]

Top words

word	frequency of the word
year	93
game	45
season	42
attack	35
office	33
solar	31
go	28
play	27
come	22
injury	21
wednesday	19
look	18
company	17
accord	16
earth	15

two words	frequency
magnet field	13
hall fame	10
duke health	7
heart transplant	6
ariana madix	6
brain dead	5
solar probe	5
bale say	5
pope franci	5
billion year	5

Three word phrases

three words	frequency
fast solar wind	4
donat circulatori death	4
parker solar probe	3
doubl homer run	3
surgeon jacob schroder	3
heart donor brain	3
brain dead research	3
differ approach success	3
success boost number	3
transplant heart donor	3

News source distribution

A pie chart illustrating the distribution of news sources. The chart is divided into 16 segments, each representing a different news source and its corresponding percentage of the total. The segments are color-coded and labeled with their respective percentages. The legend on the right lists the news sources in the same order as the segments in the chart, starting from the top and moving clockwise.

News Source	Percentage
CBS Sports	10%
Reuters	10%
The Washington Post	5%
Buzzfeed	5%
BBC News	5%
YouTube	5%
USA Today	5%
Al Jazeera English	5%
Yahoo Entertainment	5%
CNN	5%
Associated Press	5%
ESPN	5%
The Times of Israel	5%
ABC News	5%
Al Jazeera English	10%
Reuters	10%

Finally, k-means clustering is carried out. The criterion can be selected on the left side, whereby either the 18 food variables, the 5 health variables or all of them in combination may be chosen for the clustering. On the right side, the resulting clusters can then be compared with respect to a chosen variable.

