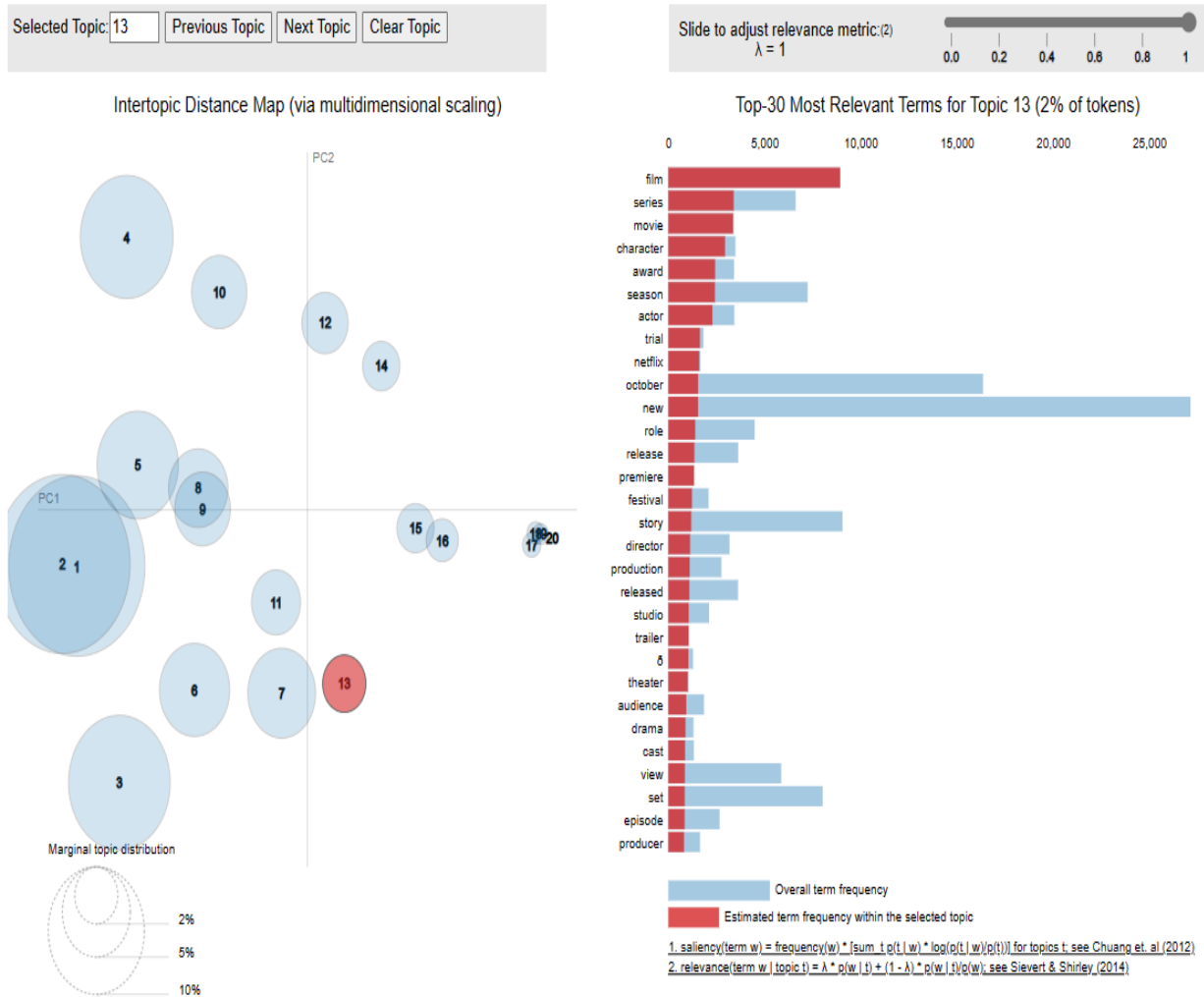


PyLDA Visualization to visualize and interpret the topics extracted from the corpus of news articles.

Topic number 13 selected:

Out[71]:



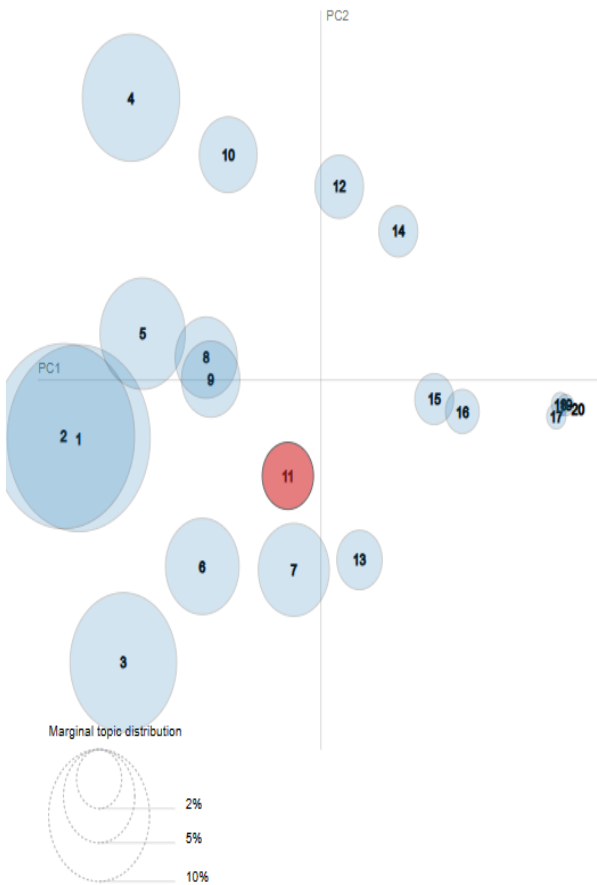
Topic number 11 selected:

Out[71]:

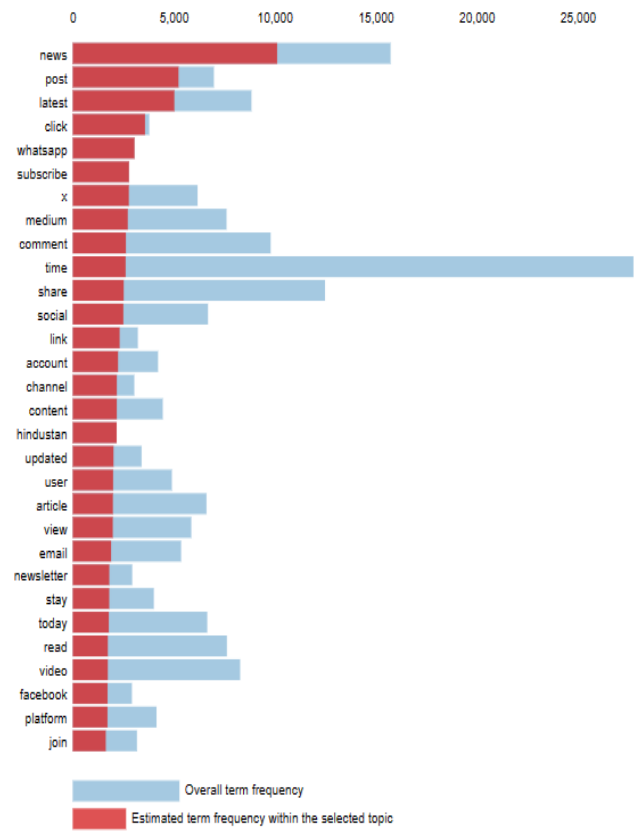
Selected Topic:

Slide to adjust relevance metric:(2)
 $\lambda = 1$

Intertopic Distance Map (via multidimensional scaling)



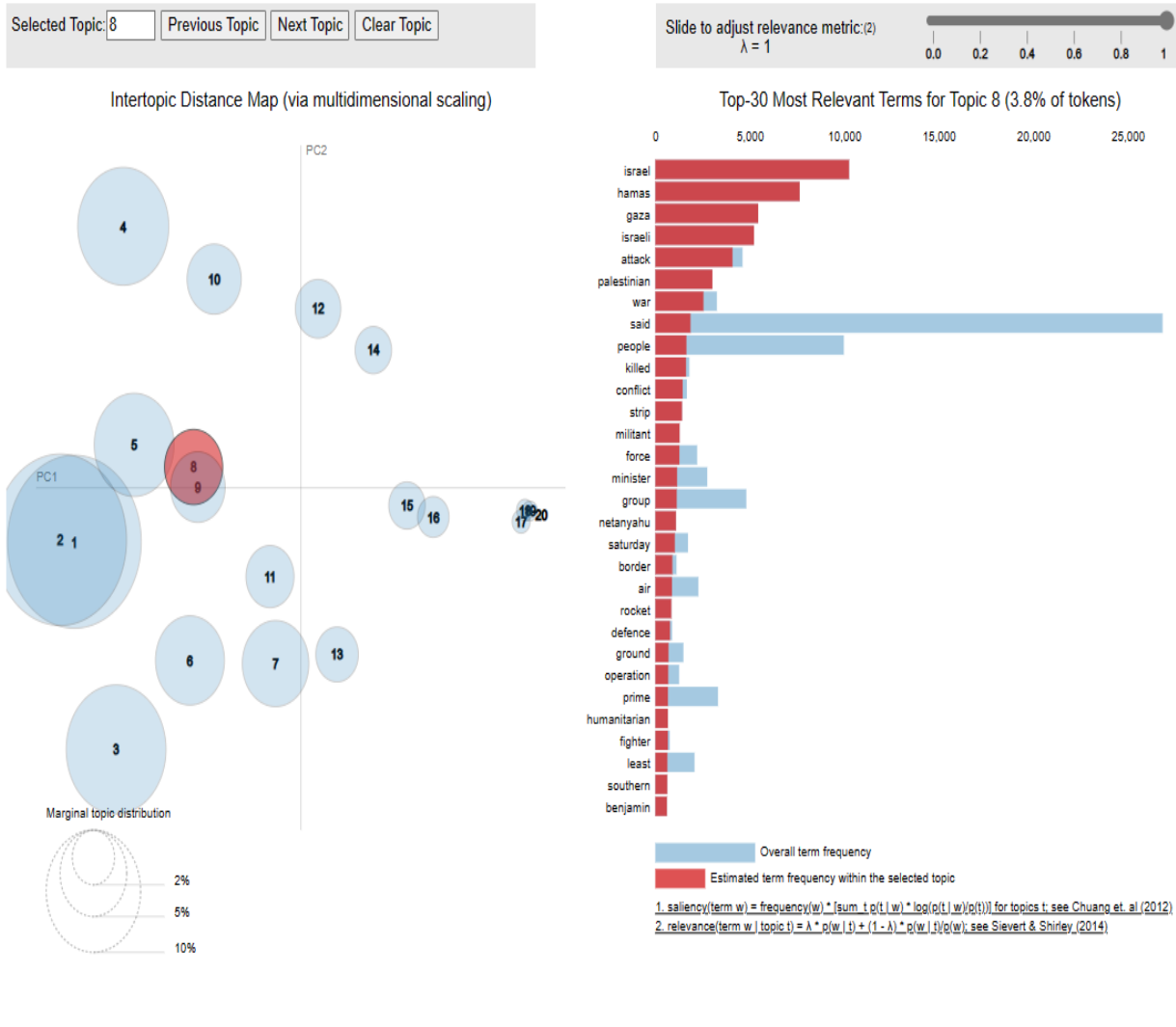
Top-30 Most Relevant Terms for Topic 11 (2.6% of tokens)



1. $saliency(\text{term } w) = \text{frequency}(w) * [\sum_t p(t|w) * \log(p(t|w)/p(t))]$ for topics t ; see Chuang et. al (2012)
2. $relevance(\text{term } w | \text{topic } t) = \lambda * p(w|t) + (1 - \lambda) * p(w|t)/p(w)$; see Sievert & Shirley (2014)

Topic number 8 selected:

Out[71]:



Topic Number selected 1:

Out[71]:

