

## Business report extraction



Define HPC Keywords



EDGAR API



Business section of  
filtered 10-K report



Filter business  
report contain HPC  
keywords



## Business model canvas extraction



Extract business model canvas in dictionary  
format and identify if healthcare related



## Visualizations



Trend of business model  
transition



## topic modeling for each canvas component



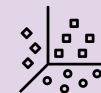
Merge topic & topic words  
regeneration



Define optimal topic  
number using dendrogram



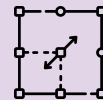
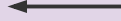
## Top2Vec



Clustering & topic  
generation



lower dimensional  
embedding



Text embedding