

cos similarity heatmap - layer 8
bert-large-uncased_wikitext_{x,y}_trained-params_dense-on_24-layers)

