

cos similarity heatmap - layer 23  
bert-xlarge-v2\_uniform-tokens\_{Qx,Ky}\_trained-params\_dense-on\_24-layers)

