

cos similarity heatmap - layer 8  
ert-xlarge-v2\_uniform-tokens\_{Qx,Ky}\_trained-params\_dense-off\_24-layers)

