

cos similarity heatmap - layer 3
bert-xlarge-v2_uniform-tokens_{Qx,Ky}_trained-params_dense-on_24-layers)

