cos similarity heatmap – layer 7
(...ert-xlarge-v2_uniform-tokens_(Qx,Ky)_trained-params_dense-off_24-layers)