

Heatmaps of cos similarity  $\langle \frac{x_i}{\|x_i\|}, \frac{x_j}{\|x_j\|} \rangle$   
(wikitext\_bert-large-uncased\_params-trained\_dense-off\_layers-24, sample 6)

