

# Additive Regularization for Topic Models

How to combine all those extensions in one model?

**PLSA:**  $\mathcal{L} = \sum_{d \in D} \sum_{w \in W} n_{dw} \log \sum_{t \in T} \phi_{wt} \theta_{td} \rightarrow \max_{\Phi, \Theta}$

**ARTM:**  $\mathcal{L} + \sum_{i=1}^n \tau_i R_i(\Phi, \theta) \rightarrow \max_{\Phi, \Theta}$

Example of a regularizer – diversity of topics:

$$R_i(\Phi) = - \sum_{t \neq s} \sum_w \phi_{wt} \phi_{ws}$$

K. Vorontsov, A. Potapenko Additive Regularization of Topic Models, 2015.