

# How would you train the model?

## Probabilistic Latent Semantic Analysis:

$$p(w|d) = \sum_{t \in T} p(w|t) p(t|d) = \sum_{t \in T} \phi_{wt} \theta_{td}$$

## Parameters of the model:

- $\phi_{wt}$  – probability of word  $w$  in topic  $t$
- $\theta_{td}$  – probability of topic  $t$  in document  $d$