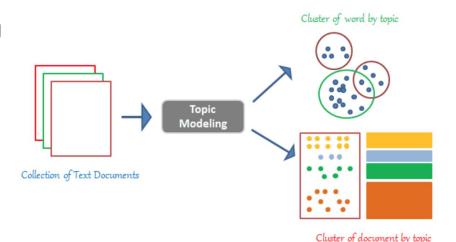
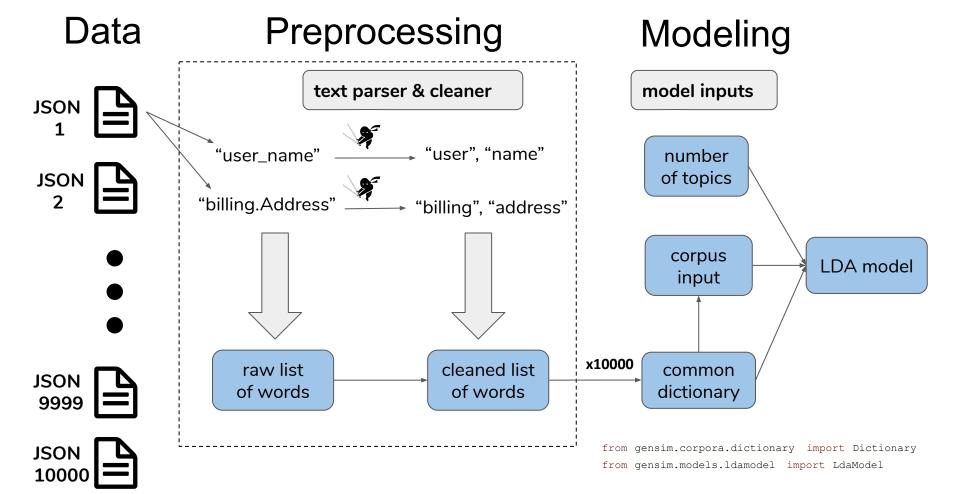
Introducing Topic Modeling with LDA

Topic Modeling: the task of identifying topics that best describe a set of documents

Latent Dirichlet Allocation (a popular example of topic modeling)



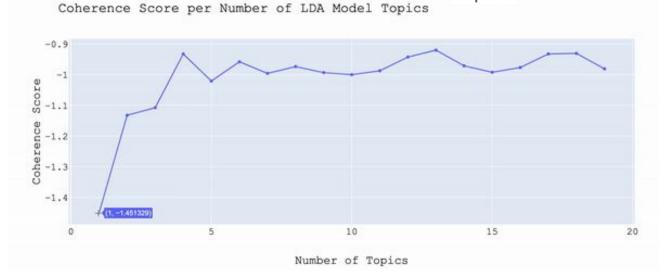
- <u>Latent</u>: hidden topic structures
- <u>Dirichlet</u>: statistical distribution for determining hidden structures
- Allocation: allocation of words per topic



Coherence Model

- Coherence score ⇒ Quality of learned topics
- Intrinsic Measure: UMass

Choosing the number of topics with maximum topic coherence usually offers meaningful and interpretable topics



Word Clouds

