Overview

- What is text clustering?

- Why text clustering?

- How to do text clustering?

  - **Generative probabilistic models**

  - Similarity-based approaches

- How to evaluate clustering results?