Problems with "Term as Topic"

- Lack of expressive power   **➔ Topic = {Multiple Words}**
  - Can only represent simple/general topics
  - Can't represent complicated topics
- Incompleteness in vocabulary coverage   **+ weights on words**
  - Can't capture variations of vocabulary (e.g., related words)
- Word sense ambiguity   **➔ Split an ambiguous word**
  - A topical term or related term can be ambiguous (e.g., basketball star vs. star in the sky)

A probabilistic topic model can do all these!