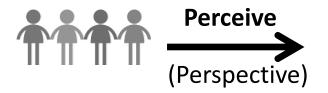# Key High-Level Take-Away Messages

13. Joint mining of text and non-text
14. Contextual PLSA
15. NetPLSA
16. Causal topic mining

6. Probabilistic Topic Model (PLSA, LDA)
7. Generative model; ML estimate; EM
8. Text clustering: model vs. similarity-based
9. Text categorization: generative vs. discriminative
10. Evaluation of clustering and categorization

Real World

Observe

Perceive

(Perspective)

1. NLP ➜ Text representation ➜ Knowledge discovery
2. Robust TM = Word-based rep + Statistical analysis

(English)

representation

11. Sentiment classification: ordinal regression
12. Latent Aspect Rating Analysis

3. Paradigmatic and syntagmatic relations
4. Text similarity: Vector space, BM25
5. Co-occurrence analysis: Entropy, MI

3