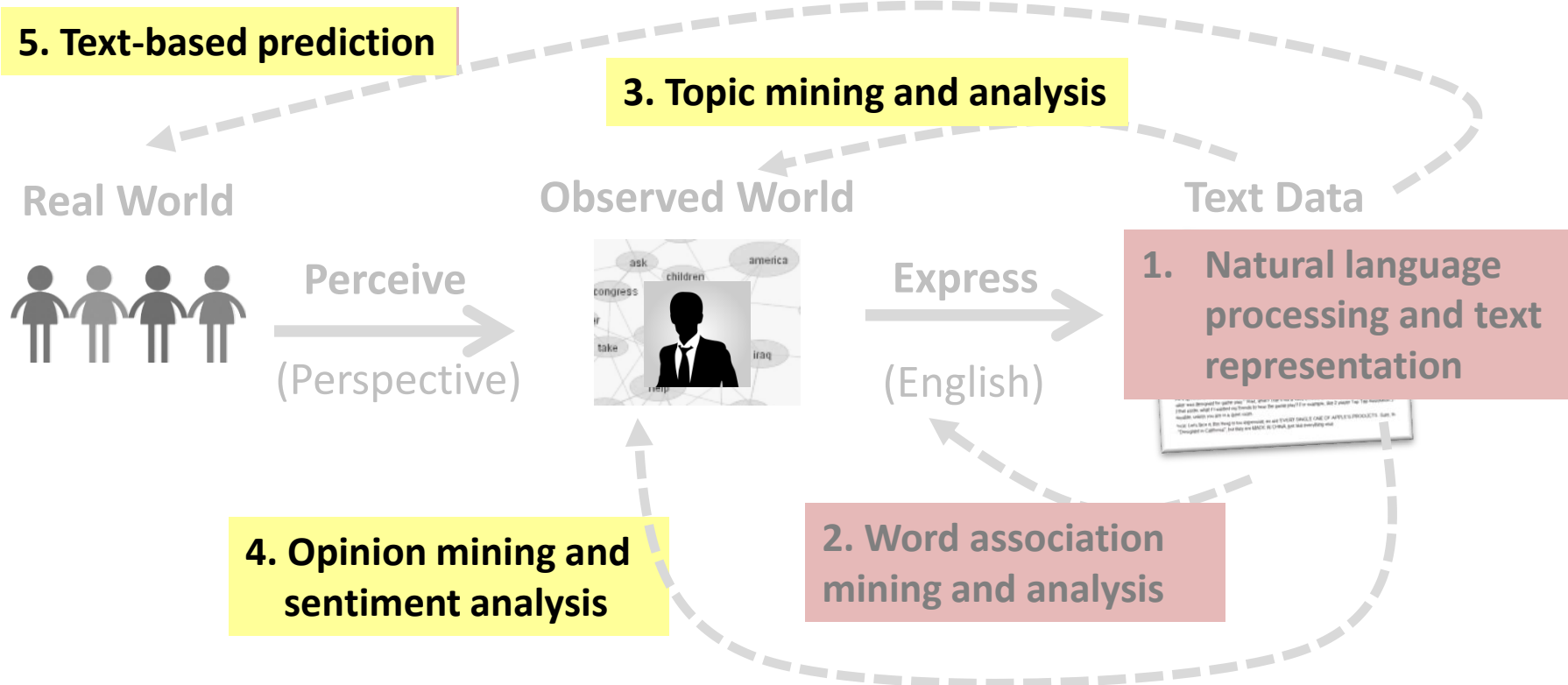# Contextual Text Mining: Motivation

ChengXiang "Cheng" Zhai
Department of Computer Science
University of Illinois at Urbana-Champaign
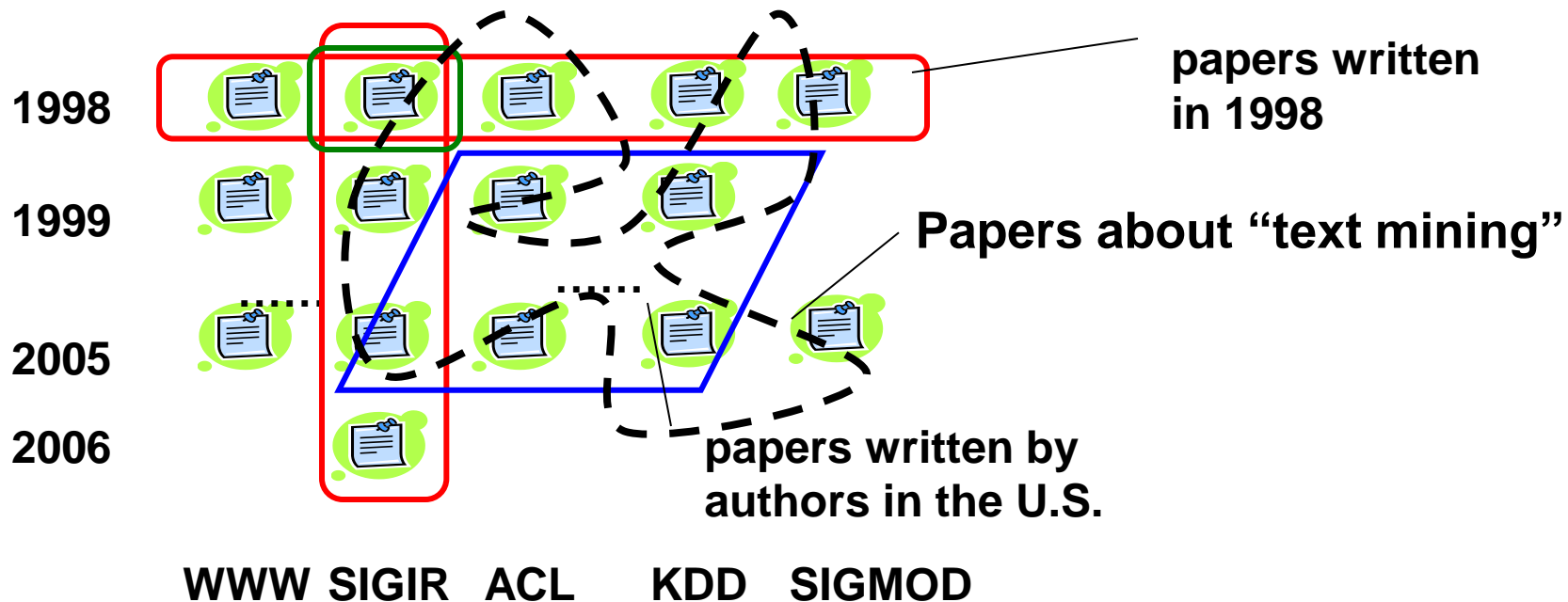
Contextual Text Mining

5. Text-based prediction

3. Topic mining and analysis

Real World

Observed World

Text Data

Perceive

(Perspective)

Express

(English)

1. Natural language processing and text representation

2. Word association mining and analysis

4. Opinion mining and sentiment analysis

Contextual Text Mining: Motivation

- Text often has rich context information
  - Direct context (Meta-Data): time, location, authors, source, …
  - Indirect context (additional data related to meta-data): social network of the author, author's age, other text from the same source, etc.
  - Any related data can be regarded as context
- Context can be used to
  - Partition text data for comparative analysis
  - Provide meaning to the discovered topics

Context = Partitioning of Text

Enables discovery of knowledge associated with different context as needed

4

Many Interesting Questions Require Contextual Text Mining

- What topics have been gaining increasing attention recently in data mining research? (time as context)
- Is there any difference in the responses of people in different regions to the event? (location as context)
- What are the common research interests of two researchers? (authors as context)
- Is there any difference in the research topics published by authors in the USA and those outside? (author's affiliation and location as context)
- Is there any difference in the opinions about a topic expressed on one social network and another? (social network of authors and topic as context)
- Are there topics in news data that are correlated with sudden changes in stock prices? (time series as context)
- What issues "mattered" in the 2012 presidential election? (time series as context)