

# Application of text mining in the biomedical domain

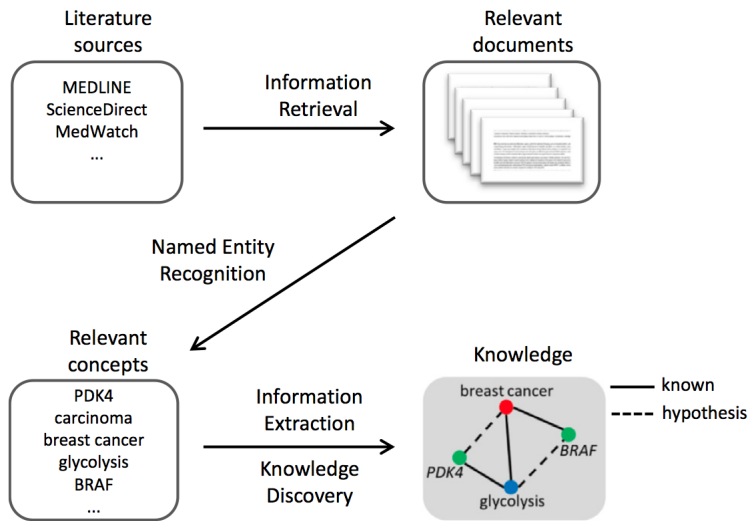
Nikita Kartashov

February 21, 2015

# Motivation for text mining (TM)

- Lots of available data
- Existing techniques for automated text processing
- Possibility of faster knowledge generation

# TM process



- Keywords - just a number of words
- Concepts - extended groups of keywords
- Gene lists - list of genes
- Ontologies - description of the concepts and relationships

# Relations discovery techniques available

- Co-occurrence based - two things occur together, therefore related
- NLP-based - two things mentioned in a special context - therefore related
- Using ABC principle
  - Finding indirect relationships
  - Finding the "middle man" of the possible relationship

# Purposes

- Genome and gene expression annotation
- Drug–target discovery
- Drug repositioning
- Adverse effects discovery
- Electronic health records mining
- Domain specific databases

That's all, folks!