

SemBioNLQA

Module 1

Query Formulation

MetaMap

Entity 1

Entity 2

Entity n

Query

Question Classification

Feature extraction

Patterns

Machine learning system

Question type

Module 4

Answer Processing

Yes/No answer
extraction

SENTIWORDNET

Factoid answer
extraction

UMLS
BioPortal
Term frequency

List answer
extraction

UMLS
BioPortal
Term frequency

Summary answer
extraction

UMLS
Porter stemmer
BM25

Module 2

Document Retrieval



PubMed



MEDLINE database

Relevant documents

Re-rank documents
(UMLS similarity)

Top-ranked documents
 d_1, d_2, \dots, d_k

Module 3

Passage Retrieval

Sentence splitter
(Stanford CoreNLP)

Passages ranking
(BM25)

Top-ranked passages
 p_1, p_2, \dots, p_n

Candidate
answers

Questions
????

User

ANSWERS