Interconnected Knowledge in Documents

Dr. Siegfried Handschuh
Email: Siegfried.Handschuh@deri.org

Vit Novacek, Tudor Groza
What we did

- CORAAL
  - developed for intelligent document search,
  - currently deployed in one specific domain (adaptable to other domains)

- Problem
  - isolated pieces of knowledge in thousands of documents
  - \( \rightarrow \) Data silos & Information Overload

- Approach
  - We automatically extracted and inferred concepts from more than 10,000 publications,
  - linked them together by various types of relations
Extraction of Interlinked Knowledge

Knowledge query builder

NOT Subject: acute granulocytic leukemia
Relation: is a
Object: T-cell leukemia

Query:

Submit
150

Statement

sorted by: rank; then by... • grouped as sorted

**acute granulocytic leukemia** NOT **T-cell leukemia**

Sources:
- Coding sequence and intron\xe2\x80\x93exon junctions of the c-myb gene are intact in the chron...

Certainty: 0.6640

Contexts: oncology, genetics, pharmacology, biochemistry, biology, cell_research, and clinical_medicine

Inferred: false