7.3.3 Semantic Search

With an ever increasing volume of medical literature and clinical reports, it is becoming increasingly important to be able to meaningfully search this information. A major application for Natural Language Processing technologies is to index collections of free text transcripts or documents such that topic specific searches may be run on them. The challenge is to move beyond the limitations of plain keyword searching strategies towards more advanced search techniques, which return ranked matches with high sensitivity and specificity. Clinical searches may be performed over documents within an electronic library, within medical records, or on the internet. Examples of searches include:

- "Show me articles on this website concerned with inflammatory bowel disease"
- "Does this patient have transcripts in their record suggesting a heart rhythm disturbance?"

SNOMED CT was used in techniques developed by Koopman to improve search performance by addressing vocabulary mismatch (using synonyms, e.g. hypertension vs high blood pressure), granularity mismatch (using hierarchical relationships, e.g. antipsychotic vs haloperidol), conceptual implication (using defining relationships, e.g. from renal cyst infer kidney) and inferences of similarity (e.g. using subset membership, e.g. comorbidities anxiety and depression). Koopman also assigned a measure of similarity to each SNOMED CT relationship type, and use this weighting to determine the relevance of each document.

Some commercial tools also provide semantic search, including Cerner's semantic search tool.