Methods (continued)

- Concept ontology contained entities representing five quality metrics from SNOMED terminology (Diabetes, Hypertension, Cardiac Surgery, Transient Ischemic Attack, CNS tumor).
- All ontologies and patient notes (RDFs) were imported into AllegroGraph® triple store as classes and instances respectively (Figure 1).
- SPARQL information retrieval protocol was used for reporting extracted concepts under four settings:
  - Baseline NLP output
  - Inclusion of a concept ontology
  - Exclusion of negated concepts
  - Inclusion of a section header ontology
- Existing manual abstraction data from surgical clinical reviewers, on the same set of patients and documents, was considered as the gold standard.
- Statistical agreement tests (precision, recall, F Measure) were applied for comparing the results of manual abstraction and the proposed ontology-based approach.

Results

- The overall (micro-average) F-Measure increased from 68% to 82% (Chart 1).
- F-Measure increase was highest for CNS tumors (63%), Cardiac Surgery (57%), and TIA (32%).
- F-Measure increase was lowest for Hypertension (9%) and Diabetes (1%).
- We also compared the isolated effect of each ontology layer to the base NLP output (Chart 2).
- The overall (micro-average) effect of section ontology was 11% and 5%
  - (F-Measure) higher compared to concept ontology and negation context respectively

Summaries of Conclusions

- The application of ontology-based approach for natural language processing in our study has provided complementary mechanisms for increasing the performance of NLP systems.
- The pivot point for extracting more meaningful quality metrics from clinical narratives is the abstraction of contextual semantics hidden in the notes.
- We have defined some of these semantics and quantified them in multiple complementary layers.
- Rigorous evaluations are still necessary to
  - Create evaluation guidelines for assessment of performance of ontology-based information extraction systems
  - Provide a consistent baseline for the purpose of comparing alternative approaches

Reference


Acknowledgement

This project was supported by Grant No. 10510592 for Patient-Centered Cognitive Support under the Strategic Health IT Advanced Research Projects (SHARP) from the Office of the National Coordinator for Health Information Technology.

Please contact the first author via email: sina.madani@uth.tmc.edu