



NLP Marketing Summary

Name: Natural Language Processing (NLP)

Elevator Pitch: Software, dictionaries & know how for computers to understand the meaning of naturally occurring speech and text

Faculty: Dr. Martha Palmer, Department of Linguistics & Computer Science, Dr. Jim Martin, Department of Computer Science, Dr. Wayne Ward, Computational Language & Education Research

Next Steps: Participate in Boulder Innovation Center Facilitated Meeting Exploring the Different Applications for the technology

To Indicate an Interest or For More Information: Eric Gricus, Eric@BoulderInnovationCenter.com, 303.444.2111

Background

Natural Language processing is becoming ubiquitous in everyday life. Applications once thought to be farfetched are now within reach as processing power, machine learning algorithms and the internet extend to all facets of everyday life. All of these applications start with a computer's ability to understand the meaning behind the words.

The Center for Computational Language & Education Research (CLEAR) at the University of Colorado is focused on advancing Human Language Technology and applying it to practical applications. The CLEAR Center goal is the translation of basic research in the area of computational semantics (i.e., getting computers to grasp the meaning of naturally occurring speech and text) into practical systems that can, for example, answer questions given access to texts containing likely answers. In these projects, the faculty apply supervised machine learning techniques to process text to identify entities, events, and the relations between them.

One of the initial research areas has been in the medical field. The faculty researchers are collaborating with Harvard Children's Hospital and the Mayo Clinic to develop systems for the semantic annotation of medical texts, both web-based resources such as Medpedia and patient clinical reports dictated by physicians. The semantic annotations are used to support two applications, medical domain question answering and constructing event timelines from medical reports.

The faculty research team is open to exploring additional applications for their technology. Here are the specific assets related to their research;

- Processes for developing high quality annotated data at a low cost
- Expertise in applying state-of-the-art machine learning techniques to train high performance automatic systems on the data

- Experience with integrating the individual components into end-to-end systems for particular applications, i.e., domain question answering, constructing event timelines from medical reports
- Pre-existing domain specific annotated data for medical domain

The next step for this project is to convene a group of software/IT experts who can assess the assets developed by CLEAR to identify logical commercial targets for the technology.

To Indicate an Interest or For More Information: Eric Gricus, Eric@BoulderInnovationCenter.com,
303.444.2111