**Introduction**

Named entity recognition is the process of finding and determining the types of named entities in a text for example people, locations and organizations and it is an important step in computers understanding written text. When named entity recognition works the computer can determine that the White House is the official residence of the president of the United States instead of a just a regular house that has been painted white, an important distinction to make if one wants to understand a text.

The term named entity recognition comes from the Sixth Message Understanding Conference when researchers were determining the sub tasks involved in information extraction from unstructured text (R. Grishman & Sundheim 1996). The task was separated out from general information extraction due to its difficulty and importance. In order that the task be completed two different approaches have been created, the traditional rule based approach, meaning that the type of an entity is determined based off of a set of hardcoded rules and the classification approach that is based off of machine learning after training data has been used to teach the machine. However, the results discussed in this paper have been found by using a combination of both approaches.

**Literature Survey**

There have been many papers written on the topic of named entity recognition since the term was invented but in this section three will be surveyed for their contributions to the field. The first paper to be examined is a Nadeau and Sekine (2007), which is an overview of the research done in the field. This is an important paper to survey due to its examination of the past research done in the field of named entity recognition over a fifteen-year period. It explores how the field has changed from one based solely off of rule based systems to the increasing use of machine learning today showing how our research which uses a combination of the two is bringing past and current methods together to increase accuracy.

The second paper to be examined is Exploring Entity Recognition and Disambiguation for Cultural Heritage Collections (Hooland, S. Van, M. De Wilde, R. Verborgh, T. Steiner, and R. Van De Walle 2013), which explores the use of, named entity recognitions and term extraction in the field of digital humanities. The end result of the paper was that both forms of recognitions are useful and important however, one must be careful when using it to get meaningful results. Similar to our paper Hooland and his co-authors used a combination approach focusing on both qualitative and quantitative methods. However, this paper focuses on a much more specific type of text while ours is more general in its uses.

The last paper is on using multiple different named entity recognition systems together to create a better system (Illig, Jens, Charles La Rosse, Iliana Simova, Gu ̈nter Neumann, and Bogdan Sacaleanu 2011). The systems work together through a voting system that tries to cover the individual weaknesses of the systems and therefore have fewer errors. This paper also relates to our combination approach since like them we are using multiple approaches to try and cover for individual weaknesses. In conclusion this paper is built upon the work of the many researchers that have come before it.

**References**

Hooland, S. Van, M. De Wilde, R. Verborgh, T. Steiner, and R. Van De Walle. "Exploring Entity Recognition and Disambiguation for Cultural Heritage Collections." *Digital Scholarship in the Humanities* 30.2 (2013): 262-79. Web. 15 Dec. 2016.

Illig, Jens, Charles La Rosse, Iliana Simova, Gu ̈nter Neumann, and Bogdan Sacaleanu. "A Platform for Named Entity Processing." (2011): n. pag. Print.

Nadeau, David, and Satoshi Sekine. "A Survey of Named Entity Recognition and Classification." *David Nadeau and Satoshi Sekine [li.30.1.03nad]*. National Research Council Canada, 10 Aug. 2007. Web. 15 Dec. 2016.

Grishman, Ralph; Sundheim, B. 1996. Message Understanding Conference - 6: A Brief History. In *Proc. International Conference on Computational Linguistics*.