Project Title – I have three working titles. The first title is representative of the 'source' task—neural metaphor detection. The second and third titles are possible extensions of the source task.

1) Neural Metaphor Detection using Bidirectional LSTMs with Attention and Deep Word Embeddings with Named Entity Tagging
2) Extended Metaphor Detection using Bidirectional LSTMs and Deep Word Embeddings
3) Neural Metaphor Detection and its Application to Internet Privacy Policies

Team Members – I will be working alone.

Overview – The first title describes a project that is solely focused on the source task. The title could simply be 'Neural Metaphor Detection'—I'm only vaguely familiar with bidirectional LSTMs and attention mechanisms, and there are other neural network architectures described in the literature that perform the task well [1]. In their description of a bidirectional LSTM model for metaphor detection, Aggarwal and Singh mention that their model has trouble discriminating between different entities and could be improved with POS and/or named entity tagging [2]. In that vein, the second half of the title refers to innovations in training data preprocessing aimed at improving model performance. In reference to 'Deep Word Embeddings', there are many examples in the literature of inputting training data as ELMo, BERT, and GloVe embeddings, and I would like to explore the performance of these different embeddings, and potential combinations of them, during the model training phase.

The second title describes a project where an acceptable, working model for metaphor detection is constructed and applied to another related task: detecting extended metaphors in a document. I haven't thoroughly explored this idea yet, but here's the gist: using a domain-specific corpus, such as a corpus of privacy policies [4], run the model and compile a list of terms used non-literally, keeping track of their position in the document. Using deep word embeddings, compute the similarity of the non-literal words, searching for non-literal usages with similar conceptual mappings, taking into account the context in which the non-literal usages appear. Deep word embeddings may or may not help with this task.

The third title also describes a project that applies a neural metaphor detection model to a specific domain/corpus. This is also not a fully formed idea, just a possible extension if the initial task is completed quickly and I would like to do more work. Reidenberg et al. describe a system of scoring ambiguous language in privacy policies [5]. I think it would be interesting to apply this method in conjunction with metaphor detection to see if there's any overlap between highly ambiguous language usage and non-literal language usage.

Data Sources – There are several datasets commonly used for building neural metaphor detection models, the most robust and nuanced of which seems to be the VU Amsterdam Metaphor Corpus [6]. The VU Amsterdam Corpus contains annotations for three types of metaphor: indirect, direct, and implicit. Training a model to be aware of these different metaphor types could be useful for various model applications. The MOH-X dataset has examples of literal and non-literal language usage in relation to specific verb occurrence [7]. The sentence object is also included with each sample. The TroFi dataset contains literal/non-literal language samples for a small set of verbs [8]. It may be interesting to train a model on the TroFi dataset and see if it scales to non-literal usage associated with verbs outside the TroFi vocabulary.

Project Plan – The problem of neural metaphor detection is well described in the literature, and the first order of business will be to conduct further research into model architecture and input data processing techniques. Once I've drawn up a rough plan for constructing a model architecture, I'll create the model and begin testing it on deep word embeddings and training data from one or more of the datasets mentioned above. A large part of this process will be deepening my knowledge of deep learning and finding ways to improve and innovate on model architectures and data preprocessing techniques described in the literature. If the model building phase proceeds quickly and produces decent results, I will extend the project by applying the model towards solving one the problems described above. As we have not yet covered the deep learning material and my own experience with deep learning is very limited, I cannot provide specific architectural details.

Related Research – A fair amount of research has been produced on creating neural network architectures for metaphor detection over the past several years. Most of the recent literature describes methods involving deep embeddings, bidirectional LSTMs, attention mechanisms, and various other RNN architectures. In my brief search, I was unable to find research that applied neural metaphor detection models to domain-specific tasks. I was also unable to find research on extended metaphor detection. Tong et al. [9] provide a thorough overview of the current state-of-the-art research into neural metaphor detection.

Proposed Evaluation – Standard metrics for model evaluation—precision, recall, F1 scoring, accuracy, and AUROC scoring—will be used to evaluate and compare my model against other models trained on the same datasets. In addition, I will also manually evaluate instances of false positives and false negatives, searching for patterns that may provide insights on how to improve model performance.

Paper Summary – I may use some of the ideas in this paper, but I do not plan on directly extending the metaphor detection model they describe. That is, I do not plan on building an ensemble of RNNs to solve the task. I chose to provide a summary of this paper because it is well-written and provides examples of model architectures and data preprocessing that are unintuitive and potentially useful.

Title: Metaphor Detection using Ensembles of Bidirectional Recurrent Neural Networks
Authors: Jennifer Brooks and Abdou Youssef
Publication: Association for Computational Linguistics
Volume: Proceedings of the Second Workshop on Figurative Language Processing
Date: July, 2020
URL: https://aclanthology.org/2020.figlang-1.33.pdf

Brooks and Youssef describe a metaphor detection model constructed with ensembles of bidirectional recurrent neural networks with the goal of "improv[ing] metaphor detection to facilitate the interpretation and translation of natural language in discourse." For training, they used the VU Vrije University Amsterdam Metaphor Corpus (VUAMC) mentioned above. They evaluated their model using precision, recall, and F1 scoring. Their best model achieved an F1 score of 0.703.

For input, they prepared 11-grams, each from a single sentence, for each word in the training set. They represented each word in the 11-gram as a 1,024-dimensional ELMo embedding concatenated with a 300-dimensional GloVe embedding. They note that mixing ELMo and GloVe embeddings resulted in better performance than using one or the other in isolation. The model output, passed through a Softmax activation, is a two-dimensional vector of probabilities for the center word in the 11-gram representing metaphorical language usage.

The authors tested two different architectures. The first architecture is a many-to-one bidirectional LSTM where the outputs of the "forward and backward LSTM cells in the attention layer are concatenated only at the output for the" target word. This architecture is not fully connected, i.e. each attention cell receives output from a single bidirectional LSTM cell, in an effort to respect the sequential processing order of attention cells. The authors note that they achieved better results without fully connecting the LSTM and attention layers. The second model is a many-to-many bidirectional LSTM with bidirectional attention, with the main difference from the first model being that it updates weights in relation to the model's performance on both the center target word and the surrounding context words (the output is still a two-dimensional vector of probabilities for metaphorical usage of the target word). Their ensemble consisted of "five models per architecture trained independently on all parts of speech"; results improved when logically combining outputs of the ensembles, which were collated into multiple voting pools.

In their conclusion, they mention that metaphorical usage of verbs is over-accounted for in VUAMC dataset, and that future work in metaphor detection should place more emphasis on other parts of speech, in particular nouns.

Informal List of Sources

1) https://aclanthology.org/2020.figlang-1.33.pdf
2) https://arxiv.org/pdf/2009.12565.pdf
3) https://privacypolicies.cs.princeton.edu/
4) https://privacypolicies.cs.princeton.edu/
5) https://www.oag.ca.gov/sites/all/files/agweb/pdfs/privacy/reidenberg-ambiguity.pdf
6) http://www.vismet.org/metcor/documentation/home.html
7) https://github.com/gititkeh/visibility_embeddings/blob/master/datasets/mohx/mohx.csv
8) http://natlang.cs.sfu.ca/software/trofi.html
9) https://aclanthology.org/2021.naacl-main.372.pdf