During my PhD course semesters in software engineering department, by taking some courses such as Semantic Web, Metadata managements and Big Data, parallel to my studies in Machine Learning and Data Science, I found the information exaction is one of the interesting areas to me to research. After my course semesters, we got a project proposal to populate a domain specific knowledge base. By ending beginning studies, Triple Extraction (relation extraction) was the task that I felt close to myself then by encouragement of our supervisor I started to deep research in this area as a part of project and my PhD thesis which entitled “Development a Tool and Method for Domain Specific Relation Extraction”.

Joining to a research group that was working on similarity (semantic and network) based prediction techniques for drug-drug interactions and biomedical data, gave me the opportunity that to make a good experience in data analysis and using machine learning techniques on linked open data. In that research we were leveraging semantic and topological representation of linked open data of drug and deses data sets such as Drugbank and Pubmed.

In addition to experiences in biomedical data analysis, for my thesis studies and our knowledge extraction project, I tried to analyze some relation extraction tools and algorithms such as Reverb, DIPRE, KnowLife and etc. Implementing simple examples to shallow usage of NLP tools such as Stanford Core NLP and NLTK got useful. Also, as my first experience to completely usage of a triple extraction tool, I tried to modify BRED project (Word Embedded Version of Snowball, Implemented by python 2) to be runnable on python 3.5 by Anaconda Interpreter

During our work on our project, we studied creation and population models of some well-known knowledge bases such as DBpedia, Yago, KnowledgeVault. We found that DBpedia is the well-known and skeleton of LOD cloud, so we studied DBpedia construction and refinement process. Finally, to set a simple and porotype version of our extraction pipeline for our domain specific knowledge base generation project and my PhD thesis, we used DBpedia to generate our evaluation dataset for domain specific tasks.

At the end, due to my research and implementation background in related area and my interesting, I am confident that if I take this position I will make an effective contribution to the project and also, it would be good experience to me to improve my information in theoretical and implementation perspectives.