

NLP with GKS

Project Statement: This research project investigates the use of natural language processing (NLP) to mine formal specifications to allow the security policies to be reasoned about automatically.

Milestones Reached (Oct 6, 2020)

- *Research:* Reading and discussion of the ARSENAL paper and discussion of the paper with our academic advisor, Dr. Nan Niu
- *Project Planning:* Technical feasibility of using an automated NLP from a research paper, every 2 weeks there are advisor meetings, and bi-weekly team meets with daily reports of progress and planning
- *Creation of User Diagrams:* These showed low, medium, and high level demonstrations of how we will go about attempting to implement our project
- *Software Tool Research:* Finding which NLP tools are available and applicable to our project (Stanford Parser, Stanford Core, NLTK, etc.)
- *Testing:* Checking for compatibility of the software to our requirements

Future Milestones (Oct 6, 2020 - Dec 8, 2020):

- Creation of at least 5 IR (intermediate representation) tables with their respective logic formulas
- Streamlined way to take natural language documents or phrases and be able to extract what is considered to be a requirement.
- Ability to create said IR tables in a more automated way, so to save time on any manual work to be done
- Find the NLP tool we will stick to throughout the entire project. Once we find which tool is most “friendly” to our application, we will most likely move forward in our project with said tool
- Updates to User Diagrams as abstraction levels increase or decrease as to maintain proper documentation so that everything researched is archived
- Consolidate our project planning phase and commence project implementation by end of semester

Timeline

Number	Task and Milestones	Start Date	End Date
1	Develop a way to create an Intermediate Representation (IR).	Sept 15, 2020	Oct 12, 2020
2	Research relevant questions to the field of Natural	Aug 21, 2020	Dec 8, 2020

	Language Processing (NLP).		
3	Design the non-automated programming aspect of the processor.	Oct 12, 2020	Oct 29, 2020
4	Update documentation related to the project.	Sept 30, 2020	Dec 8, 2020
5	Specify which NLP processing tool will be used.	Sept 30, 2020	Oct 12, 2020
6	Validate code and tools.	Oct 12, 2020	Nov 6, 2020
7	Verify that the layout and code of the NLP tool have met requirements.	Nov 6, 2020	Dec 8, 2020
8	Perform a test run of the semantic parser. Adjust accordingly.	Nov 6, 2020	Nov 22, 2020
9	Upload content to the GitHub repository as work is completed.	Aug 21, 2020	Dec 8, 2020
10	Maintain constant communication with the faculty advisor.	Aug 21, 2020	Dec 8, 2020
11	Research various parsers that could be of use for this project.	Sept 30, 2020	Oct 12, 2020

Effort Matrix (names in **bold** denote Task Leader)

Number	Task and Milestones	Team Member	Est. Hours
1	Develop a way to create an Intermediate Representation (IR).	Sheffi Kevin Garret	8 9 8
2	Research relevant questions to the field of Natural Language Processing (NLP).	Sheffi Kevin Garret	10 10 10
3	Update documentation related to the project.	Sheffi Kevin Garret	5 5 5

4	Specify which NLP processing tool will be used.	Sheffi Garret	3 2
5	Validate code and tools.	Sheffi Kevin	4 5
6	Verify that the layout and code of the NLP tool have met requirements.	Kevin Garret	4 5
7	Perform a test run of the parser. Adjust accordingly.	Sheffi Kevin Garret	2 2 2
8	Upload content to the GitHub repository as work is completed.	Sheffi Garret Kevin	5 5 6
9	Maintain constant communication with faculty advisors.	Sheffi Kevin Garret	20+ 20+ 20+
10	Research various parsers that could be of use for this project.	Sheffi Kevin Garret	4 4 4