

# Project Proposal

David Orona,  
Nikil Nair

1. What are the names and NetIDs for all team members?
  - David Orona / orona2
  - Nikil Nair / nnair7
2. What is your group name?
  - import team-name
3. Who is the team captain?
  - David Orona / orona2
4. What system have you chosen? Which subtopic(s) under the system?
  - We have chosen the system extension route as the one that we wish to proceed with and we have chosen the MeTA Toolkit as the system that we want to extend/improve upon.
5. Briefly describe any datasets, algorithms or techniques you plan to use:
  - Based off the subtopic, our plan is to utilize and enhance the metapy and MeTA packages. This would include manipulating any functions and methods attached to them.
  - In addition, being that we are bringing the tools to a more modern standard with Python 3.X, it is reasonable to assume we would utilize more contemporary functions of the language, perhaps including NLTK, spaCy, textBlob, gensim, to even something as simple as f-strings for easy usability.
6. If you are adding a function, how will you demonstrate that it works as expected? If you are improving a function, how will you show your implementation actually works better?
  - The functionality that we plan on adding will showcase that the metapy system is operative in settings that it wasn't necessarily configured for, namely new operating systems as well as later versioning. We also plan on implementing cross-functional support with other toolkits and we can showcase that our implementation works by demoing and documenting the results that we obtain. This added functionality, while not an explicit function, is a core aspect of the implementation that showcases the imperative nature of scalability and as a result the persistence of utilization across systems.
7. How will your code communicate with or utilize the system? It is also fine to build your own systems, just please state your plan clearly:

- Our code will attempt to communicate with the MeTA system and extend functionality to available toolkits such as NLTK and gensim. This integration would be a system extension of the current implementation and with the latter, we hope to achieve a wider scope of functionality and usability. We also plan on extending the python versioning as well as the OS support for this package so that its features can incorporate cross-platform utilization.
8. Which programming language do you plan to use?
- We are planning to use Python to fulfill the requirements for this project.
9. Please justify that the workload of your topic is at least  $20 \times N$  hours, N being the total number of students in your team. You may list the main tasks to be completed, and the estimated time cost for each task:
- There is a variety of tasks involved in the scope of this project and as such, we believe that there is an adequate workload of at least  $20 \times N$  hours where N is the number of members within our group. Since our group consists of two members we believe that the average time that it would take to fulfill the requirements that we have set is estimated at around 40 hours. The breakdown of the tasks and the involvement that we estimate completion of are listed below:
    - i. 5 hours: Understand the processes involved in making metapy work as it does, and how we can implement improvements in this regard.
    - ii. 15 hours: Updating the system versioning so it's compatible with the latest systems and OS - we will attempt to add one OS, and then more if time allows.
    - iii. 20 hours: Integrate package functionality with popular toolkit NLTK, gensim, etc. if time allows.
    - iv. Stretch Goal: Time permitting begin modular implementation to add onto extending MeTA toolkit to support text analysis.