

Project 2: Implement PageRank

In the data files under /gpfs/projects/AMS598/projects2025_data/project2_data/, each row has two webpage numbers, indicating that the first webpage contains a link pointing to the second webpage. For example, (149, 35725) means webpage 149 has a link to webpage 35725. **Use the taxation method (beta = 0.9) to calculate and report the top 10 webpages with the largest pagerank values.**

Requirements:

1. Use mpi4py on SeaWulf for the computation.
2. For the pagerank algorithm, you only need to run four (4) map/reduce iterations.
3. Use 5 processes; however, your program should be able to run with any arbitrary number of processes.
4. Use your own directories under /gpfs/projects/AMS598/class2025/LastName_FirstName/ to save your intermediate files so we can run your script.

Submission:

1. Write a report about the analysis and submit a PDF file about the analysis to BrightSpace.
2. Submit one script file and one slurm file to BrightSpace.
3. Upload the same script file and the same slurm file to the following directory so your analysis and results can be repeated:
/gpfs/projects/AMS598/projects2025_submission/project2_submission/
LastName_Firstname /