|  |  |
| --- | --- |
| Project Status Report |  |

# Project Name: Science & Technology Search Engine

February 14, 2017 Team Members: Liyi Li, Yunyan Wu, Jie He, Jinjin Xiao, Yi Luo

|  |  |
| --- | --- |
| The project during the week of 2/8/2017 – 2/14/2017, is due to the following: | 1, Seek for other potential methods to change the way to get all the domain URLs  2, Estimate and make budget plans for different platforms which may use in future  3, Focus on studying the tutorials recommended by Common Search to build the data pipeline and do research on plugins |
| Issues: | 1, Have to use cloud service to handle the large amounts of data  2, Cost of using cloud service is higher than expectation  3, The pipeline tutorials only support Linux and Mac OS X at the moment |
| Milestones accomplished the week of 2/8/2017 - 2/14/2017: | 1, Successfully configured Docker and Git environments following the tutorial and achieve the first 3 steps  2, Explored classification with plugins on URLs  3, Got enough accounts with credits to try AWS EC2 and S3 |
| Milestones planned this week, but not achieved with variance: | 1, The detailed budget plan still need to get more information based on the status of trying AWS  2, Few other potential methods without cost are found  3, Failed to open Spark Web UI |
| Milestones planned for next week: | 1, Build basic pipeline at least on local by using Docker and Spark  2, Continue to try on AWS S3 and EC2  3, Apply plugins on classification |
| Areas/questions for discussion: | How to use cloud service to the pipeline to get the useful URLs  How to configure Docker and connect to other platforms  How to connect to Spark Web UI |
| Last week’s issues forwarded to this week: |  |