LucidWorks: Vectorize Workflow Module

By David Kniphuisen and Alan Tran
Module Overview

Completion of this module will enable users to:
• Generate a vectorize workflow in LucidWorks.
• Send queries to LucidWorks for progress of the workflow.
• Formulate the data into vectors for machine learning.

Time Required:
• 3 hours out of class to read and learn about the module.
• 1 hour in class to complete the exercises.
Other Modules/Prerequisites

Other modules:
• Follows after the LucidWorks: Overview module.
• Can be done in parallel with the Clustering and Classification modules.

Prerequisites:
• Familiarity with UNIX environment
• Overview of LucidWorks Big Data Software
• Basic cURL commands
Module Content

Body of Knowledge
- Background regarding LucidWorks' workflow
- Vector Space Model

Using LucidWorks
- Step-by-Step guide
- Parameters
Conclusion: Goals

Goals

• Understand the LucidWorks structure
• Perform basic cURL queries
• Create, check, verify the Vectorize workflow