

# **Advanced Syllabus Analyzer**

## **Applied Project Final Report**

by  
Yabo Dong  
yd1540  
Spring, 2022

A paper submitted in partial fulfillment of the requirements for the degree of  
Master of Science in Management and Systems  
at the  
Division of Programs in Business  
School of Professional Studies  
New York University

---

## **Declaration**

I grant powers of discretion to the Department, SPS, and NYU to allow this paper to be copied in part or in whole without further reference to me. This permission covers only copies made for study purposes or for inclusion in Department, SPS, and NYU research publications, subject to normal conditions of acknowledgement.

## **Acknowledgements**

I sincerely thank Dr. Andres Fortino for their service as my Applied Project Sponsor and New York University School of Professional Studies for providing the opportunity to conduct this Applied Project. I also thank Prof. Israel Moskowitz for his Applied Project course and for his mentoring, which prepared me for my successful project. My thanks also go to all the instructors in the Management and Systems Master's program, from whom I learned a great deal.

## Abstract

The purpose of this project is to develop a module to perform TF-IDF text similarity scoring analysis for New York University School of Professional Studies. The module is programmed in R programming language and hosted on the shinyapp.io.server. This module is featuring keyword extraction, online source scrape, cloud computing, and Shiny-based user experience. This module allows users to upload a text file of the syllabus locally, and then it will conduct text similarity analysis against the syllabus and O\*Net Occupation Database and return 20 most matched occupations that the students are prepared for after taking the courses. The users then can choose up to 5 occupations, and the module will scrape 15 jobs for each occupation from Indeed.com and conduct term frequency analysis for the syllabus and scraped job descriptions.

The underlying algorithms of this module are as following: extract and clean the text, drop all non-alphabetical characters, eliminate multi-space, and lemmatize the all the words; apply TF-IDF analysis and compute the similarity score against the text and the occupation descriptions; scrape job information and conduct term frequency analysis for the text and job descriptions.

With this new tool, instructors at New York University School of Professional Studies would have an easy-to-use R tool to ascertain the matched occupations with the course and adjust the course content or structure accordingly. The module works for both Windows and Macintosh operating system. Samples and comprehensive are provided in the GitHub Repository.

URL of the Advanced Syllabus Analyzer:

[https://nyuprof.shinyapps.io/Advanced\\_Syllabus\\_Analyzer/](https://nyuprof.shinyapps.io/Advanced_Syllabus_Analyzer/)

GitHub Repository:

[https://github.com/yd1540/Advanced\\_Syllabus\\_Analyzer](https://github.com/yd1540/Advanced_Syllabus_Analyzer)

# Table of Contents

<i>Declaration</i> .....	<i>ii</i>
<i>Acknowledgements</i> .....	<i>iii</i>
<i>Abstract</i> .....	<i>iv</i>
<i>Abbreviations or Symbols</i> .....	<i>vii</i>
<i>Introduction</i> .....	<i>ix</i>
Background .....	1
Company Name .....	1
Sponsor Information .....	2
<i>Problem Description/Opportunity</i> .....	3
<i>Importance of the Project</i> .....	4
<i>Approach and Methodology</i> .....	5
Work Breakdown Structure .....	5
Industry Analysis – Education Industry .....	6
Industry Trends.....	6
The customer .....	13
The competitors .....	13
SWOT Analysis .....	11
Current project management methodologies .....	12
Agile transformation.....	12
Project Implementation strategy .....	13
<i>Project Objectives and Metrics</i> .....	15
The Goal of the Project .....	15
Project Deliverables and Metrics .....	15
<i>Risk Analysis</i> .....	17
<i>Issues Encountered</i> .....	18
<i>Project Chronology and Critique</i> .....	19
Critique .....	19
<i>Lessons Learned</i> .....	21
<i>Conclusion and Summary</i> .....	23
<i>Limitations, Recommendations and Scope</i> .....	28
<i>Literature Survey</i> .....	29

1. Introduction .....	29
2. Project .....	30
3. Organization.....	32
4. Customer Collaboration.....	33
5. Working Software .....	34
6. Impediments to Implementing Agile Methodology.....	35
7. Incorporating Agile Values into the Workplace.....	36
8. Conclusion .....	37
<i>Reference.....</i>	<i>51</i>

## **Table of Figures**

Figure 1 Upload the syllabus page.....	23
Figure 2 Text similarity analysis result.....	24
Figure 3 Select occupations page .....	25
Figure 4 Scrape job page .....	26
Figure 5 Term frequency analysis result.....	26

## **Table of Tables**

Table 1 Competitors of NYU SPS.....	9
Table 2 SWOT Analysis of NYU SPS .....	11



## **Abbreviations or Symbols**

NYU: New York University, a private research university located at New York City

SPS: School of Professional University at New York University

MASY: Master of Science in Management and Systems is a master program offered through the Management & Technology department within the Division of Programs in Business at the School of Professional Studies at New York University

R: A programming language and free software environment for statistical computing and graphic

# **Introduction**

## **Background**

NYU SPS appeals to develop a module to perform TF-IDF text similarity analysis and find occupations that match with the syllabus in R program. I will program the required functions in R languages and deliver the interface by shinyapp. With the new module, instructors at NYU SPS would have an easy-to-use R tool to ascertain the matched occupations with the course and adjust the course content or structure accordingly. The module works for both Windows and Macintosh operating system. Users need to upload a text file of the syllabus. After submitting, the module will present 20 occupations that match with the syllabus with a new column indicating the similarity score.

## **Company Name**

New York University (NYU) is a private research university based New York University. The university locates at 7 East 12<sup>th</sup> Street, New York, NY 10003. The New York University School of Professional Studies (also known as SPS), previously known as the New York University School of Continuing Education, is one of the schools and colleges that compose New York University. Founded in 1934, the school offers undergraduate, graduate, and continuing education programs. NYU SPS has departments and divisions in liberal arts, global affairs, publishing, business, English, hospitality, sports business and management, and real estate. As of fall 2020, the school has a total enrollment of approximately

3,634 graduate students, 2,119 undergraduate students, and 11,000 continuing education students.

## **Sponsor Information**

Dr. Andres Fortino is the Clinical Associate Professor and MASY ACP Leader at NYU. Over the past 40 years, he gained a great deal of experience in higher education as a faculty member, as an administrator, as well as an executive in many colleges and universities. He is currently a Clinical Associate Professor of Management and Systems at the NYU School of Professional Studies and has taught in the Management and Systems (MASYS) master's program since 2012. Previously, he held the Campus Provost and Dean of Academic Affairs position at DeVry College of New York. Formerly (2006-10). He was on the faculty and held an Associate Provost's position at the Polytechnic Institute of NY, where he managed two campuses and numerous graduate programs, among other duties. Before that (2004-06), he was Dean of the Marist College School of Management, where he successfully initiated many new graduate programs. And he spent six years before that (1998-06) on the faculty and as Associate Dean at George Mason University School of Management, where he managed their MBA and other master's programs.

## **Problem Description/Opportunity**

NYU SPS wishes to develop a tool to perform TF-IDF text similarity analysis and find out the occupations that students are prepared after taking the courses. Instructors currently do not have an easy-to-use R tool to ascertain the occupations that match with the course. The syllabus is developed with the template provided by NYU and the content the instructor will cover during the course. We wish to develop an easier to use tool in R and make sure the module works for both Windows and Macintosh operating system.

Deliverables of this project include:

1. Become familiar with TF-IDF text analytics in R and how to develop shiny apps.
2. Build functions including text similarity analyzer, job scraper, and text frequency analyzer.
3. Build an interface that integrates the functions above and enable user to upload a text file of the syllabus. The output should include 1) 20 occupations and checkboxes that allow user to select and submit the results. 2) a table of jobs scraped from Indeed.com that can be downloaded. 3) Term frequency of the syllabus and the scraped job description.
4. Build a shiny easy-to-use interface.
5. Test the validity of the functions using client-defined data sets.
6. Deliver clear instructions on how to use the module.
7. Final project files and supporting documentation should be delivered via GitHub repository.

This module offers opportunity for instructors at NYU SPS and students a free and powerful tool to ascertain occupations after taking the courses. Besides, the module also provides function to scrape jobs from Indeed.com, which makes it a useful tool for job hunting.

## **Importance of the Project**

The syllabus is developed with the templates and instructors' perception of how the course will prepare students for job hunting. The importance of this project is to offer instructors a tool that can conduct similarity analysis and deliver the matched occupations, which can help instructors ascertain the occupations, improve the efficiency of creating the syllabus, and help instructors to restructure the course.

Further that, the module greatly reduces the repetitive work of text comparing and similarity scoring. Also, it does not require users to have a knowledge of coding. Users only need to prepare the syllabus in a text file format. Finally, this project is purely non-profit and education oriented.

# **Approach and Methodology**

Work Breakdown Structure, Situational Analysis, SWOT Analysis, Agile Methodology, Project Plan, and Resource Loading are used in the project.

## **Work Breakdown Structure**

### 1. Initiation

- 1.1. Evaluation & Recommendations
- 1.2. Develop Project Proposal
- 1.3. Develop Project Charter
- 1.4. Deliver Project Sponsor Agreement

### 2. Planning

- 2.1. Create Preliminary Scope Statement
- 2.2. Develop Communication Plan and Project Plan
- 2.3. Develop Functional Specifications Requirements
- 2.4. Deliver Project Sponsor Agreement Part# 1

### 3. Execution

- 3.1. Design Functions: Text Similarity Analysis Function, Scrape Function,  
Term frequency Analysis
- 3.2. Testing Functions
- 3.3. Design Shiny Interface
- 3.4. Testing and Improve the Module
- 3.5. Upload the Module to Cloud

### 4. Control

- 4.1. Project Management

- 4.2. Risk Management
- 4.3. Project Status Reporting
- 5. Closeout
  - 5.1. Limitations, Recommendations and Scope for Future Work
  - 5.2. Upload Files to Repository
  - 5.3. Deliver Sponsor Agreement Part#2 and Gain Formal Acceptance

## **Industry Analysis – Education Industry**

Based on the Bureau of Labor Statistics description, the education industry consists of institutions that provide education and training in a wide range of areas. The education and training are provided by specialized establishments, such as schools, colleges, universities, and training centers. These establishments may be privately owned and run for-profit and non-profit, or they can be owned and operated by the government. They could also provide food and or accommodation services to the students.

This paper will focus on universities that provide teaching and research facilities, including graduate schools and professional schools. In 2020, the U.S. education industry garnered revenue of about US\$ 1.1 trillion in 2020 and was set to earn revenue of nearly US\$ 2.3 trillion by 2028 (Research, 2022).

## **Industry Trends**

With time, the education business is witnessing even more changes. By harnessing new technology, startups are filling in the holes and flaws in the sector, bringing fresh and inventive ideas to the table.

The emergence of online academics: Because of the real-time and participatory environment, online academics are becoming increasingly popular,

particularly during pandemics. Students can now complete their courses online by utilizing interactive multimedia technologies such as live-streaming and video-conferencing. Over the Internet, students and instructors engage and have discussions in a classroom-like setting. With more and more online programs and distant learning methods springing up, Universities aren't left behind.

AR (augmented reality) and VR (virtual reality) are currently being used to improve learning by presenting virtual 3D models in 2D areas. Instructors, particularly in science and engineering, may now create interactive lectures and models without building physical models. Simulations that would have previously only been done on a computer may now be run in a VR area thanks to the usage of VR (virtual reality). This allows users to experience the same situations as they would in the actual world, which is a significant advancement in educational research and instruction.

## **Porter's Five Forces Analysis**

### **The buyer's power**

Students can be considered the education industry's customers. They require the assistance of committed educators to attain their educational objectives and acquire the information and skills they require. Students need support for technology-based learning methods, improved educational results, and achieving professional goals with each academic year. All other industries can be called education consumers since they rely on trained workers from it.

### **Industry rivalry**



The table below lists the competitors of the New York University School of Professional Studies in the education industry.

<b>Institutions</b>	<b>Description</b>
Columbia University – School of Professional Studies	The school offers seventeen master's degrees, courses for advancement and graduate school preparation, certificate programs, summer courses, auditing, and lifelong learning programs in New York City and abroad.
Boston University -- Metropolitan College	The school offers 78 undergraduate and graduate degree and certificate programs, available part-time and full-time, along with individual courses that are open to the public
CUNY School of Professional Studies	The school provides online programs that meet the needs of adults who are looking to finish a bachelor's degree, progress from an associate degree, earn a master's degree or a certificate in a specialized field, and advance in the workplace or change careers.
Northwestern University – School of Professional Studies	The School of Professional Studies offers highly specialized degree programs at the undergraduate, post-baccalaureate, graduate, post-graduate, and professional development levels.
Northeastern University –	The school offers academic programs grounded in theory and applied in practice, pushing through the

College of Professional Studies	traditional barriers between work and education and providing a rich educational program
---------------------------------	--

*Table 1 Competitors of NYU SPS*

Universities must use several competitive advantages in the education business if they are to continue to grow.

Information technology: The proper application of information technology in its different forms can help educational institutions obtain a competitive advantage in the education market. Delivering the educational resources and courses through technology, instructors may now provide lectures, offer homework, and administer assessments through the Internet. On a global scale, technology also provides the tools and infrastructure for managing and delivering education services. How successfully consumers and institutions embrace, and exploit information technology will determine the education industry's long-term viability.

Governmental and non-governmental funding: Through loans, grants, and scholarships, students get many official and non-governmental funds. On the other hand, education institutions get government funding for initiatives and curriculum improvement. Corporations also provide funding to universities in return for a competent workforce and university-led research initiatives. Accountability for how this money is spent is becoming increasingly important. Using the budget to undertake educational programs that will assist product graduates in getting employable skills can provide educational institutions with a competitive advantage.

**Supplier Power**

The supplier of an education institution could be the professors and infrastructure, both of them will impact the pedagogical choices. For example, there will be an increase in wage costs if a university decides to move the focus to STEM courses to fulfill the increased demand. According to the National Compensation Survey of the Bureau of Labor Statistics, the average post-secondary STEM teaching compensation is \$85,827, while the average humanities teaching salary is \$78,134. Without them, the average salary is \$71,123, or \$14,000 less per post than STEM!

### **The Threat of New Entrants**

As employer-sponsored on-the-job training and pension funds become extinct, certificates, EMBA's, and associate degrees are becoming increasingly popular. Employees increasingly expect to be job changers for the rest of their lives, with the burden of retraining falling on their own shoulders. The market for credentials gained for a job switch or upgrading is heating up as a result. These stackable credentials are an important part of education's 'as-needed' philosophy, as well as a contributor to the concept of the student-as-consumer. Prepare thyself, thyself, thyself.

### **The Threat of Substitutes**

Students are increasingly viewing education through the lens of a consumer. This translates to grade inflation, more special needs accommodations, and a shopper attitude loosening the hold of the 4-year tradition. According to the National Center for Education Statistics' Persistence and Attainment of Postsecondary Degrees study, 'mixed menu' education is on the rise. Gap years, 2-

year programs, online courses, summer courses, and life experience are now interwoven throughout the four-year degree. The rise in traditional tuition, the rise in student debt, the decline in ROI, and the pressures on colleges to adapt to meet an ever-widening range of student needs are all driving these competitive possibilities.

## SWOT Analysis

The following is a SWOT analysis of the New York University School of Professional Studies

<p><b>Strengths</b></p> <ul style="list-style-type: none"> <li>• Great campus location</li> <li>• High education quality and reputation</li> <li>• Enable students with promising career prospects</li> </ul>	<p><b>Weaknesses</b></p> <ul style="list-style-type: none"> <li>• High tuition</li> <li>• Increased living expenses for students</li> <li>• Student preparedness at the entrance</li> </ul>
<p><b>Opportunities</b></p> <ul style="list-style-type: none"> <li>• Digital transformation to online</li> <li>• Going up the collection of programs to face the marketplace</li> <li>• Broader connections with alumni</li> </ul>	<p><b>Threats</b></p> <ul style="list-style-type: none"> <li>• Intense competition from other universities</li> <li>• Less international student enrollment due to pandemic</li> </ul>

Table 2 SWOT Analysis of NYU SPS

## **Current project management methodologies**

Currently, the New York University School of Professional Studies provides templates and instructions on how to write a syllabus. The templates include general course information, course description, learning outcomes, expectations, assessment strategy, and course schedule. While the templates and instructions are exhaustive, it is also cumbersome and makes it obscure to reveal the matched and possible occupations after taking this course. By defining mandatory sections to be included in the syllabus, the school ensures the syllabus can provide detailed information about the procedure itself but ignores how the course can prepare students for career life. The prevalent way of creating a syllabus at New York University is to use the templates and instructions provided by schools.

## **Agile transformation**

It has been established that the education industry is highly competitive. Students require high-level education quality and career preparation, including a clear syllabus and high performance in an agile fashion. New York University School of Professional Studies needs to transform itself to be agile and nimble to stay competitive in the education industry. It needs to adopt a methodology that will allow instructors to ascertain the most likely jobs students are prepared for after taking the courses.

The opportunity that the New York University School of Professional Studies has on hand is to develop and deliver a module that is fully aligned with the business objectives of the instructors and its consumers. By adopting agile methodologies to build the module, NYU SPS has the opportunity to provide its

customers and product management visibility and control to create a market to differentiate from competitors. It also has the chance to get the module in application quicker due to agile methods. This will ensure a high return on investment for the project. By adopting agile methodologies, it can increase the efficiency and productivity of the instructors and improve the quality of the syllabus, which is expected to have a positive user experience with the module.

## **Project Implementation strategy**

"Individuals and Interactions" are valued more than "Processes and Tools," according to the agile manifesto. On the other hand, the correct tools may make a tremendous impact, especially when the goal is to make individuals and their interactions more productive and meaningful. Because the target user of the module will be instructors at NYU SPS, choosing the correct tools to automate agile procedures is just as crucial as deciding to transition to agile development. A list of functional requirements should be carefully established when evaluating agile project management systems to make the tool selection process more manageable.

The goal of agile development and breaking down stories into something that can deliver value in short iterations is to gather immediate feedback from the sponsor. The product managers must be taught to break down significant needs into smaller, more manageable deliverables for this to work.

Module development quality has always been a source of contention, and it has always been acceptable to have "bugs" in a program. For this project, the testing for quality will be conducted at the end of the development phase. In an

agile setup, the tester that mainly focuses on requirements confirmation is in a weekly meeting with the sponsor.

# **Project Objectives and Metrics**

## **The Goal of the Project**

The goal of the project is to create a module to perform TF-IDF text similarity analysis against syllabus and O\*Net occupation database within an easy-to-use Shiny interface.

Scenarios that instructors can upload a text file of the syllabus, the module will then analyze 20 occupations that match with the syllabus most, the user can select up to 5 occupations, and the module will scrape 15 jobs for each occupation from Indeed.com and conduct term frequency analysis against syllabus and scraped job descriptions.

As the person conducting the project, I will account for the project throughout its lifecycle. I will be the project manager that requires me to initiate the project. I will plan, execute, and monitor the project that are similar to what I learned and did in the Project Management course. Parse and Visualize Data. I will use the scrape and visualize data skills that I learned from this course. During this project, I will analyze the syllabus and data parsed from Indeed.com and predict which industries and technologies best match this syllabus, which is similar to what I learned from Data Analytics and Text Analytics courses.

## **Project Deliverables and Metrics**

- Objective1 – Have weekly meetings with the clients and specify detailed requirements, including features and functions of the module by the end of the course.



Measurement: Deliver the module and required documents and discuss any changes with the client.

- Objective2 – Build an Advanced Syllabus Analyzer based on Shiny interface.

Measurement: Deliver in R language by the end of the project.

- Objective3 – Provide clients and users with explicit instruction for using the module by the end of the course.

Measurement: Deliver the module and user manual by the end of the course and save the project code and supporting documents in GitHub.

- Objective4 – Use data sets specified by the client to test the validity of the algorithm and module.

Measurement: Deliver the final module by the end of the course so that the module's functionality can be demonstrated in the presentation.

## Risk Analysis

1. R updates to a new version and some functions or packages are changed. The impact area will be in schedule. If the risk is triggered, I need to spend time to be familiar with the new version and learn the syntax of new functions or packages. The severity of this risk is low.
2. Indeed.com does not allow users to parse job data from its website. The impact area will be in quality. If the risk is triggered, it might cause no source to provide job lists in the second phase of the project. The severity is medium. The way to mitigate this risk can be using other job boarding websites, such as Handshake, LinkedIn.
3. The project process stagnates at the coding phase. The impact area will be in schedule. If the risk is triggered, the project may not complete on the required date. The severity is high. To avoid this risk, I will have weekly meetings with the sponsor, report the current progress, and discuss the issues during the coding phase, and look for training courses from the Internet and get familiar with R and text mining techniques.
4. The final delivery does not meet with the expectation. The impact area is in quality. If the risk is triggered, extra time or resources will be needed to improve the performance of the project. The severity level is medium. To avoid this risk, I will have meetings with sponsor and update him about the progress and current performance and reach an agreement with the sponsor on the final expectations.

## **Issues Encountered**

The main issue I faced was that I did not have an experience with R and shiny app. My experience focuses on using Python to conduct data analysis and use HTML5 and CSS to build interface. However, the syntax and logic of R and Shiny are different. For instance, code in Shiny would be rendered only when users open the specific tab, and a reactive function is required to call a function. To solve this issue, I educated myself with Datacamp that provides a R and Shiny training, searched several articles and websites that provide threads for creating the required functions, and used Stack Overflow to optimize most of the problems. Then I delivered the module via cloud before the required date.

The other issues were about improving the format, style, and the performance of the interface. Examples include how to limit the user to select up to 5 occupations, how to create a popup window, and how to present a process bar when the function is running, etc. I used Stack Overflow to gain insights and possible solutions to improve the performance and delivered an easy-to-use shiny interface. However, there is still a huge space to improve on.

## **Project Chronology and Critique**

Jan 10, 2022: Delivered project proposals

Jan 28, 2022: Delivered the project specifications requirements

Feb 09, 2022: Delivered the project sponsor agreement and pre-project sponsor acceptance

Feb 15, 2022: Delivered the project charter

Feb 16, 2022: Created function: TF-IDF text similarity analyzer

Feb 22, 2022: Delivered communication plan

Feb 24, 2022: Created function: Indeed.com job scraper

Mar 01, 2022: Delivered Risk management plan

Mar 03, 2022: Delivered WBS

Mar 05, 2022: Created function: Term frequency analyzer

Apr 10, 2022: Created the shiny interface

Apr 12, 2022: Delivered situational analysis

Apr 19, 2022: Delivered literature survey research

Apr 24, 2022: Tested and published the module

Apr 29, 2022: Presentation for sponsor

May 03, 2022: Project ends

### **Critique**

The room for improvement in this project is the responsiveness of the front-end, the running time of the program, and the interface.

First, for the responsiveness of the front-end, the current module requires the user to manually clean the syllabus, only keep the parts that can contribute to

the similarity analyzer and convert the syllabus into a text file. The program should accept more types of files and automatically conduct text mining. In this way, the users only need to upload the file to the module. Also, after submitting the selected occupations, the user will first go the job tab and then manually click the term frequency tab to view more results. These actions should be activated when users click submit button.

Second, the running time of the module is a subtle area that could be improved. Certain functions such as term frequency analysis and job scraping can be streamlined and improved, and how the functions are triggered within shiny interface can also be improved.

Third, the interface can be more convenient. Certain functions such as data table filter that allows user to filter the scraped jobs can be created and improve the usability. The third tab "Term frequency Tab" can be dynamic and let user select the frequency boundary. There might be space for improving the overall performance of the interface.

Overall, the project meets the requirements of the sponsor and is delivered within the time limit, but there is space for continuous improvement. For the details of the limitations, please refer to *the Limitations, Recommendations and Scope for Future Work*.

## **Lessons Learned**

I learned how to manage a project and develop software during the entire project implementation. In addition, I learned about software development life cycles, project management, and risk management. After completing this job, I am more confident in taking on more complex software development assignments.

First, my R programming and text mining skills were improved during the process. I learned how to extract text content from a file and transform into a vector, drop all non-alphabetical characters and multi-space, then use TF-IDF to compute similarity score and present the result into a dynamic data table. This tool also embeds term frequency analyzer and presents the result by descending order. By creating the job scraper, my HTML was honed by looking into different HTML elements and attributes. The module will be a convenient tool for instructors to ascertain the occupations matched with the course, and help students get prepared for the job launching.

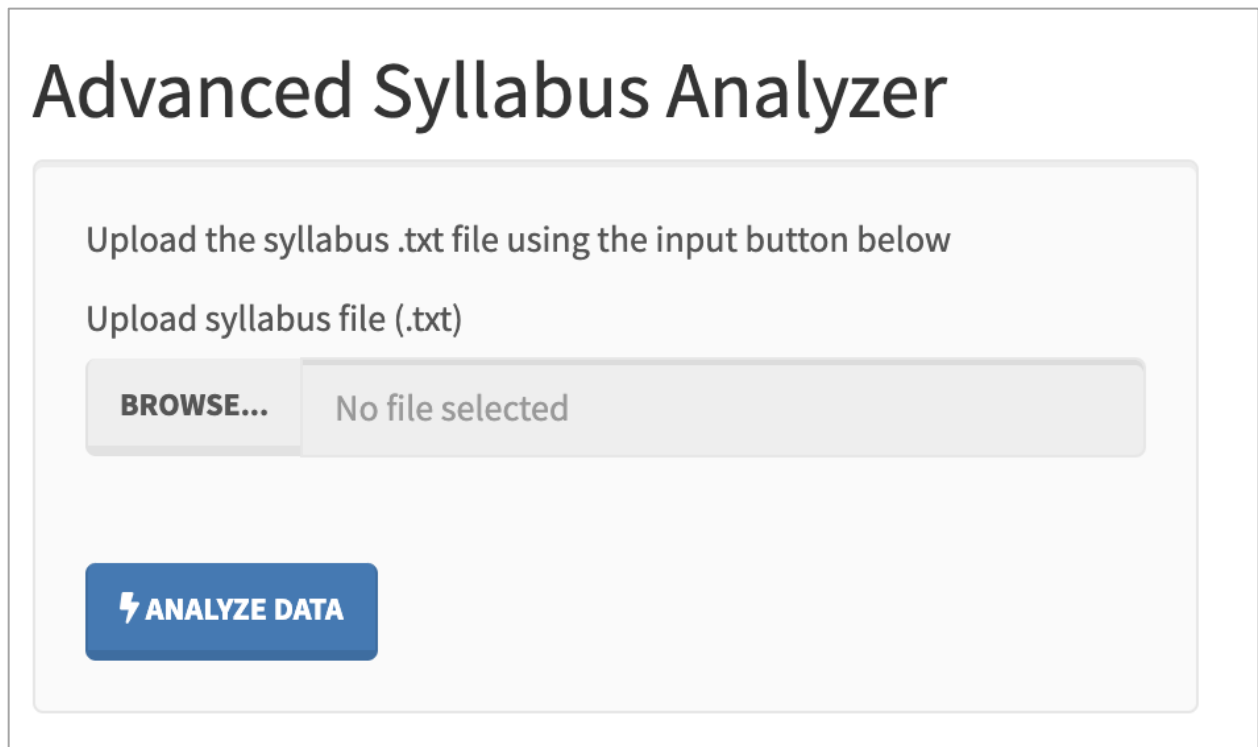
Second, as the project manager of this project, I have learned skills of how to create project proposal and project charter, and then design and implement the project. Also, I applied agile methodology and had weekly meetings with the sponsor to update him about the current process and receive suggestions for improvement and the next steps. Last but not least, I obtained knowledge about how to create documentation. My understanding of project management has been vastly expanded by the reasoning process. I improved my critical thinking skills by coming up with a contingency plan that included all possible risks.

Last, my communication and collaboration skills were improved during the project. I worked with fellow classmates and discussed the possible solutions at the

stage of proposing the project. During the weekly meetings with sponsor, I presented the progress and the next step to keep him updated. I believe after this project, I will be confident to take complex technical development projects.

## Conclusion and Summary

The project is delivered on time with required functions and quality. It allows users to upload a text file of the syllabus and analyzer 20 occupations that matched with the course most, the user then need to select up to 5 occupations. After clicking the submit button, the module will automatically scrape 15 jobs for each occupation from Indeed.com and conduct term frequency analysis against the scraped job description.



**Advanced Syllabus Analyzer**

Upload the syllabus .txt file using the input button below

Upload syllabus file (.txt)

**BROWSE...** No file selected

**⚡ ANALYZE DATA**

*Figure 1 Upload the syllabus page*



Matched\_Occupation Jobs\_from\_Indeed Text\_Frequency\_Analysis

Show 10 entries Search:

	Title	similarity_score
1	Nurse Practitioners	0.15125
2	Interpreters and Translators	0.14818
3	Prepress Technicians and Workers	0.10519
4	Computer and Information Systems Managers	0.10177
5	Art Therapists	0.09155
6	Team Assemblers	0.0798
7	Adult Basic and Secondary Education and Literacy Teachers and Instructors	0.07154
8	Chemists	0.06661
9	Information Security Analysts	0.06576
10	Business Intelligence Analysts	0.0644

Showing 1 to 10 of 20 entries Previous 1 2 Next

≡ SELECT OCCUPATIONS ▾

Figure 2 Text similarity analysis result

≡ SELECT OCCUPATIONS ▾

## Select up to 5 occupations:

- Nurse Practitioners
- Interpreters and Translators
- Prepress Technicians and Workers
- Computer and Information Systems Managers
- Art Therapists
- Team Assemblers
- Adult Basic and Secondary Education and Literacy Teachers and Instructors
- Chemists
- Information Security Analysts
- Business Intelligence Analysts
- Clinical Data Managers
- Social Science Research Assistants
- Occupational Health and Safety Technicians
- Credit Analysts
- Survey Researchers
- Cartographers and Photogrammetrists
- Data Entry Keyers
- Logisticians
- Mining Machine Operators All Other
- Dispatchers Except Police Fire and Ambulance

⚡ SUBMIT

Figure 3 Select occupations page

**Advanced Syllabus Analyzer**

Upload the syllabus .txt file using the input button below  
 Upload syllabus file (.txt)

BROWSE... Fortino\_MASY1-GC5000-101\_Special Top  
 Upload complete

ANALYZE DATA

Matched\_Occupation Jobs\_from\_Indeed Term\_Frequency\_Analysis

DOWNLOAD

You are searching for occupations: Information Security Analysts, Business Intelligence Analysts, Clinical Data Managers

Please use the search box to filter the data. Sometimes there are no jobs found for certain occupations

Show 10 entries Search:

job_title	company_name	job_location	job_occupation	Description	link
1 Information Security Compliance Analyst	Interpublic Group	Remote in Jersey City, NJ	Information Security Analysts	Coordinate and moderate various training sessions on IT SOX controls and information security policies. The Associate will be responsible for assisting with...	https://www.in

Figure 4 Scrape job page

**Term frequency of Syllabus (freq >= 5)**

keyword	ngram	freq
data	1	23
assignment	1	16
mining	1	11
course	1	11
exam	1	10
class	1	9
text	1	9
lab	1	9
text data	2	8
final exam	2	8

**Term frequency of Job Descriptions (freq >= 5)**

keyword	ngram	freq
security	1	19
data	1	18
experience	1	16
business	1	12
information	1	9
information security	2	8
need	1	7
clinical	1	7
management	1	6
issue	1	5

Figure 5 Term frequency analysis result

The underlying algorithms of the project are as follows: extract the text content from the source file, cleanse the text and drop all non-alphabetical characters and multi-space, and lemmatize all the words; apply TF-IDF text similarity scoring analysis. After user selects the occupations, use XML2 to open the

URL and scrape the Indeed.com website, and use term frequency analyzer against the syllabus and job descriptions.

## **Limitations, Recommendations and Scope**

Even the project was delivered on time and was able to perform the functions required by the sponsor. There are some limitations within this project and some limitations may be improved in the future project at NYU SPS.

First, the module only accepts UTF-8 text file. The user needs to clean the syllabus and only keep parts that will have contribution to the analysis, and then manually convert the syllabus into a UTF-8 text file. Other languages and non-alphabetical would not be analyzed. Future work would enable the function that works with other types of files, such as PDF, Word doc, etc. and would automatically extract the main content from the syllabus, which allows to user to only upload a syllabus.

Second, the current text similarity analyzer is based on TF-IDF, which may not be the most efficient one. Future work could try other ways to conduct text similarity analysis and select the most efficient one to implement in this module, which could be shortening the overall running time.

Third, the current front-end provides limited responsiveness and consistency. For future work, after user clicking the submit button, the module should automatically call the job scrape function and term frequency analysis function. Meanwhile, filters could be embedded within the data table output and allow user to filter the output, which would improve the usability and convenience of the module.

# Literature Survey

## 1. Introduction

Project management theory and practice have supplied a great range of practical mythologies for project management efficiency. A project management methodology is a collection of methods, strategies, procedures, best practices, and other tools used to manage a project. It is frequently based on a specific project management strategy, which establishes rules and criteria that outline how a project should be handled (Spundak, 2014). One of the best-known project management process methodologies is Project Management Institute (PMI). It is a process methodology that proposes managing a project with ten knowledge areas, including project integration, scope, time, cost, quality, human resource, communication, negotiation in procurement, risk management, and stakeholder management (PMI, 2013). However, the major problem is that PMI is usually proposed regardless of the type of project. However, the architecture of specific categories of projects differs significantly; as a result, the management strategies for these projects must also change (C. & B., 2012). Agile methodologies were created in response to the apparent variations between different types of projects, particularly the peculiarities of IT projects. Agile methodologies are based on the following principles: flexible work, continuous reconsideration of the completed section or task and introduced alterations, the active role of the client, a specific manager of project team organization, continuous communication among all participants, regular meetings of the project team (Rasnacis & Berzise, 2017). However, educational institutions such as universities may find it challenging to

deploy agile because of its high emphasis on client connection, self-organizing teams, verbal communication over written documentation, prototyping, and requirements flexibility.

This study will concentrate on a few key topics that will aid in the adoption of agile approaches. First, it will give the agile manifesto. Second, the challenges to applying agile methods will be identified. Finally, it will pinpoint best practices and areas where agile may be fine-tuned for adoption.

## **2. Project**

When evaluating the mechanics of the project, the project grouping indicates the importance of essential adoption variables. The duration of the project, rising ambiguity and changing needs within the project, and the project's criticality are all factors that might influence whether agile methodologies are acceptable. Today's project manager must produce actual outcomes in shorter periods while being continually inundated with countless modifications and risk-laden decisions (Highsmith, Opening Statement, 2001). Acceptance of change, frequently manifested as shifting needs, is a key value of agile development. The agile method becomes more appropriate as the project's uncertainty grows. One of the most significant causes of risk in a project is the requirements change. This, however, is in direct opposition to the agile approach, as stories (requirements definitions in the agile technique) are created throughout the project (Kirkpatrick, Walker, & Firth, 1992). Businesses could be divided into two management styles, each best suited to their working environment (Burns & Stalker, 1961). Organic systems are suited

for settings of change and flux, whereas mechanized systems are appropriate for stable surroundings. Even though this classification was introduced years before agile approaches, projects adopting agile methodologies demonstrate a strong link with the concept of organic systems. Finally, agile approaches like ASD are not appropriate for critical systems like air traffic control software. However, this is not a massive worry because most systems do not have this level of criticality (Emery, 2001).

Another core value of agile development is adaptability. Agile development is based on a change model that differs from traditional development approaches, assuming gradual change and continuous improvement (Kelly, 2008). In a firm that uses agile approaches, incremental change means that requirements will alter because of regular input and will thus adapt as the project advances. As a result of this evolution, ongoing improvement based on user feedback is possible instead of traditional models where improvement occurs after the initial development cycle. Agile development teams are anticipated to be capable of comprehending change and having the power to adjust as needed throughout the development lifecycle. According to a review study that presents a conceptual framework that highlights the links between the significant predictor variables and agile software development performance (Misra, Kumar, & Kumar, 2008). According to the author's empirical findings, agile approaches thrive in high-change contexts. The advantage of high change behavior is that it eliminates the element of surprise when the end-user receives the final product since it has been continually modified to match the end users' needs.



### **3. Organization**

The organization grouping highlights the organizational environment, including organizational and reporting structure, procedure, documentation requirements, and workplace arrangement (McAvoy & Sammon, 2005). Instead of the typical command and control system, the interaction between managers and developers is one of cooperation. While discussing employee loyalty, Reichheld raises the critical point that businesses must be adaptive in today's dynamic work environment (Reichheld, 2001). To ensure this flexibility, managers must reduce decision-making procedures and implement an organic management style.

Greening also highlights the need for a senior engineer in XP teams, ensuring that expertise is shared across the group (Greening, 2001). Rainwater also characterizes extreme programming management as predominantly carried out by the coach, who has final accountability for the project (Rainwater, 2002). Process, or the lack thereof, clearly plays a big part in agile methodologies (Highsmith, 2001). "A decent method will not save the project from failure if the team does not have strong players," says the agile approach, emphasizing the significance of persons over processes (Martin, 2003). As a result, the lack of strict protocols is welcomed. "Fast firms have a clear sense of purpose, backed up by straightforward principles. They decide what is important and dismiss the rest" (Gandossy, 2003). For example, it discusses the challenges of implementing Windows XP in a corporation with a structured software development process (Greening, 2001).

## 4. Customer Collaboration

The customer grouping identifies the project's customer or customers and the client's location and engagement as essential adoption elements to consider when adopting an agile approach. The customer grouping identifies the project's customer and the client's location and engagement as essential adoption elements to consider when adopting an agile approach (Beck & Fowler, 2001). Customer surveys were utilized as a replacement, and efforts were made to have client representatives on-site as often as feasible, according to Reifer, who demonstrated that the principle of customer engagement was not mirrored in genuine agile projects (Reifer, 2002). Furthermore, although allowing diverse definitions of client, or customer stand-in, Beck and Fowler asserted that a customer must be an integrated part of the team. The project will fail if this integration into the team is not achieved (Beck & Fowler, 2001).

Customer participation is the most potent agile value that connects the development team with the customer (Hanssen & Fægri, 2006). They conducted a longitudinal case study on a software firm that used an agile development approach and presented its results in a software product set, including requirements, advantages, costs, and dangers (Hanssen & Fægri, 2006). The tight customer collaboration had a powerful motivational influence on the company's developers. They also mentioned that the engineers' trust has grown because of consistently exceeding client expectations (Hanssen & Fægri, 2006). They found that real consumer involvement is critical for product development innovation (Hanssen & Fægri, 2006). Another study looked at the association between customer

cooperation and software project overruns and found that projects with effective customer collaboration, supported by frequent communication, had more minor effort overruns (Molokken-Ostvold & Furulund, 2007).

## **5. Working Software**

The emphasis on working software above detailed documentation is also the leading value agile practitioners advocate. The need to make the code basic, unambiguous, and self-documenting to reduce the amount of documentation was required (Abrahamsson, Salo, Ronkainen, & Warsta, 2002). Self-documenting code aims to allow a reader with no prior knowledge of the system to comprehend the source code, its design, and function. This is accomplished by following a set of naming standards for software programs, functions, and variables (Schach, 2002). Documentation may become cumbersome in a standard waterfall process since every conceivable scenario must be recorded. A user manual, for example, would have to guide users through every single problem or exception that they may experience (Boehm & Turner, 2003). Holcombe highlights the need to write simple code that is clear and provides enough detail for the target reader to easily edit, update, or extend it further (Holcombe, 2008). The book *Agile Project Management: Creating Innovative Products*, citing the limited value of thorough documentation, claims that documentation can only supply 15 to 25% of the understanding (Highsmith, 2004). Documentation, according to Highsmith, provides merely content and no context. The most important lesson from these readings is

that agile software development should not treat code and documentation as two distinct entities and that code should be self-documenting (Highsmith, 2004).

## **6. Impediments to Implementing Agile Methodology**

According to two agile characteristics, interruptions are likely to occur in agile software development teams. First, significant client participation and rapid change reactions stymie agile software development teams during planning and iterations (Moe, Dingsøy, & Dybå, A teamwork model for understanding an agile team: A case study of a Scrum project., 2010). Second, agile methodologies emphasize individuals and direct relationships, which necessitates coordination and the involvement of other team members, resulting in process disruptions by others (Matook & Vidgen, 2014).

The high level of uncertainty in agile software development projects causes delays in obtaining and validating requirements and making development decisions (Drury, Conboy, & Power, 2012). Customers who seek changes or additional requirements cause much disruption for agile developers (Tanner & Mackinnon, 2015). Furthermore, new information interrupts agile software development teams, causing them to alter their plans drastically. Daily standups, planning sessions, and burndown charts are among the many agile software development teams (Moe, Aurum, & Dybå, 2012).

When team members seek assistance or bottlenecks emerge, the tight cooperation in agile software development teams produces disruptions. While knowledge sharing and a typical mentality enhance agile software development

teams, the quantity of direct contact is considerable and causes disruptions (Przybilla, Wiesche, & Krcmar, 2018). Agile software development teams use informal control techniques, such as clan control, to build group norms and shared values (Wiedemann & Wiesche, 2018).

Daily standups and pair programming add to the number of interruptions. Daily standup meetings provide an overview of project progress and resolve issues, although developers have complained that they are wasteful and overly focused on status reporting (Stray, Sjøberg, & Dybå, 2016). During pair programming, peers interrupt developers during cooperative activities, prompting them to fix code they have written (Balijepally, Mahapatra, Nerur, & Price, 2009). In addition to these minor disruptions, today's work patterns in software development teams include developers working in numerous teams at the same time. This not only disrupts the developer, whose jobs essentially mix throughout the day, but it also disrupts the software development team, as an absent team member might create delays in decision-making, quality assurance, or other route dependencies.

## **7. Incorporating Agile Values into the Workplace**

Large businesses tend to create deep hierarchical structures, resulting in inflexible environments. Though agile approaches do not mandate a specific organizational structure, they need more frequent and open communication among stakeholders. Due to the fewer levels of administration, a lean and flat organization will be more responsive (Coplien, 2006). A flat organization with lighter processes and fewer bureaucracy is required for firms to implement agile development

successfully. A flat development structure permits engineers to make technical decisions while encouraging decentralized development-oriented decision-making. Empowerment should be managed elegantly through the consistent application of standard interfaces and design principles.

Technical leadership skills are another part of an agile software company. According to an experiment in Norway, software experts in technical leadership roles offered better software project time estimates. They were more qualified to make rapid technical judgments than non-technical leaders (Moløkken & Jørgensen, 2005). Non-technical leaders' performance is also attributed to their lack of technical detail understanding, according to the study. One of the most important findings of the study is that human resource management varies in agile software development, necessitating leadership with technical expertise and experience (Moløkken & Jørgensen, 2005).

## **8. Conclusion**

Large software firms may use agile software development to deliver high-quality software cost-effectively, provided they modify a few things about their company. The first step toward agility is to change the culture of the organization. Agile development approaches will be more successful in leaner and less bureaucratic firms. It is also critical to have the correct amount of technical leadership at the management echelons. A manager must have sufficient technical expertise to make sound judgments. Relevant stakeholders should be taught about agile approaches, and management should be aligned entirely with the endeavor to implement agile methodologies. There may be attempts to weaken the essential

ideals of agile approaches during the adoption process. It is critical to keep track and guarantee that key principles are executed consistently. The organization's agile community should oversee adoption activities and assure compliance. A competent product management team should serve as a link between the development team and the client. This group should make sure that consumer input is recorded, that the project's scope is determined, and that priorities are revised as the project continues. Finally, several software best practices might aid in the implementation of agile. Design patterns aid in identifying common business patterns and the development of software frameworks to abstract such patterns.

# Appendices

## Appendix#1: Sponsor Project Agreement

**New York University  
MS in Management and Systems  
Applied Project  
Project Sponsor Agreement**

### 1. Goals of the Program

#### For Participating Organizations

- Receive help from highly trained NYU graduate student.
- Receive assistance at no cost.

#### For NYU Graduate Students

- Manage and implement a meaningful project aligned with their professional and educational goals
- Hands-on experience interacting with a start-up or operational small business or organization
- Earn credit toward completion of graduate degree by conducting an unpaid Applied Project under the mentorship of an NYU-SPS professor.

### 2. Project Sponsor and Student Responsibilities

- Student prepares project planning documents
- Sponsor reviews and approves student's project plan
- Student submits project plan to faculty supervisors for approval
- Student conducts project according to plan
- At predetermined milestones sponsor reviews and approves status reports submitted by student
- Status reports reviewed and evaluated by faculty supervisors to assure student effort and project meet course requirements
- Project sponsor and student participate in periodic project reviews with NYU
- At project completion project sponsor completes evaluation forms
- Student prepares final report

### 3. Project Selection Process

- Project Evaluation Committee reviews proposed projects
- Projects are:
  - Relevant to MS degree course content
  - Significant to the participating organization
  - Substantial in terms of duration and scope
  - Challenging to the student
  - Capable of being measured against predetermined goals

### 4. The MS in Management and Systems

#### Concentrations in:

- Strategy and Leadership
- Systems Management
- Database Technologies
- Enterprise Risk Management

#### Students Study Courses in:

- Business Management
- Marketing
- Information Technology
- Database Development
- Financial Management
- Project Management
- Et al.



**Typical Participating Student Profile**

- Students selected to participate in this program meet stringent criteria
- Have completed all coursework
- High achievers with high level GPAs and strong academic credentials
- Many international students: some students have prior business experience.
- Students are highly motivated for success.

**5. Sponsor and Project Information**

Type of Organization	<input type="checkbox"/> For Profit <input checked="" type="checkbox"/> Not for Profit				
Name of Organization	NYU School of Professional Studies				
Address	7 East 12 <sup>th</sup> Street				
City	New York	State	New York	Zip	10003
	City				
Project Sponsor	First Name	Andres	Last Name	Fortino	
Title	Clinical Associate Professor and MASY ACP Leader				
Phone	845-242-7614				
Email	agf249@nyu.edu				
Web Site	<a href="https://www.sps.nyu.edu/">https://www.sps.nyu.edu/</a>				
Type of Business	Education				

Student Name	Yabo Dong
Project Title	Advanced Syllabus Analyzer

Description of Project	
The project is about developing an Advanced Syllabus Module using the knowledge the student received from the MASY program at NYU SPS. This project involves building a module that meets the specific with R requirements and makes sure it will operate appropriately for the clients. The project's final deliverables must meet the client's needs, which include ease of use as a career tool and service for research. Additionally, the final deliverables will consist of the user manual for the module.	
Estimated Hours of Student Participation (min. 250 hours)	300 hours

Anticipated Results
The project's goal is to build a module in R with a Shiny interface that accepts a syllabus, scores it against job descriptions scraped from Indeed.com, and provides top industries and occupations that match the syllabus. The explicit instructions for using the module will also be provided, and the project code and supporting documents will be saved in GitHub.

Knowledge and expertise student will need to be able to complete the project

- Knowledge of project management
- Knowledge of using R to build TF-IDF text analysis model
- Knowledge of using Shiny app to build an interface
- Knowledge of how to parse data from internet and visualize data
- Knowledge of data analysis and prediction

The sponsor is asked to communicate with and interact with the student regularly as the student conducts and completes the project. This typically takes the form of a weekly status update between the student and sponsor as well as project status reports at predetermined milestones agreed between sponsor and student. In addition, the project sponsor may be asked to be available for approximately one (1) virtual meeting with the NYU Professor to review progress, address questions and concerns. Such an optional meeting may be requested by the sponsor or the Professor.

Sponsor agrees to be available to student for status update meetings

- Yes  
 No

Describe the form and frequency of supervision of the student by the Project Sponsor.

## 6. Sponsor Agreement

Students are interns, not professional consultants. NYU is not responsible for the outcomes of projects undertaken by students. Work is on a best-efforts basis; no guarantees or warranties are expressed or implied. Organization is responsible for evaluating work presented, determining its value and whether to use it or not. Some projects may require on-going management or even re-work by the Organization after the student completes their Applied Project.

Please note that in order to post an unpaid position, the internship must encompass all 6 components below:

1. The internship, even though it includes actual operation of the facilities of the employer, is similar to training which would be given in an educational environment;
2. The internship experience is for the benefit of the intern;

3. The intern does not displace regular employees, but works under close supervision of existing staff;
4. The employer that provides the training derives no immediate advantage from the activities of the intern; and on occasion its operations may actually be impeded;
5. The intern is not necessarily entitled to a job at the conclusion of the internship; and
6. The employer and the intern understand that the intern is not entitled to wages for the time spent in the internship.

I have read and agree with the above Terms and Conditions for employers.

Please complete and sign this form in the space provided below and return to the course professor via the student who will upload the document to the course assignment area. For any questions, please email the professor: Prof. Israel Moskowitz [im36@nyu.edu](mailto:im36@nyu.edu).

**I agree to all of the above**

Participating Organization NYU School of Professional Studies Date 2/6/22

By (signature): *Andres Fortino*  
Project Sponsor

Printed Name: Dr. Andres Fortino  
Please print in English

Title: Clinical Associate Professor of Management and Systems

**7. Student Agreement**

Students who are planning to conduct an unpaid Applied Project must read and agree to the "Important Considerations Before Accepting a Job or Internship" contained on the following web page(s): <http://www.nyu.edu/life/resources-and-services/career-development/find-a-job-or-internship/important-considerations-before-accepting-a-job-or-internship.html>.

**Students do not register their Applied Project with the Wasserman Center.**

I agree to the all of the above

Student Name: Yabo Dong Date 02/01/2022  
Please print in English

Signature: *Yabo Dong*  
Please do not type signature. Use only handwritten or e-signature.

# Appendix#2: Sponsor Project Acceptance

## Sponsor's Project Acceptance Document – PART 1

*This document is the means by which your project sponsor formally agrees that your project has been satisfactorily completed and that it meets the project goal and objectives that were set at the onset of the project. It is therefore important that you describe the goal, objectives, and related metrics in the appropriate section below. The "PLAN" section is to be prepared at the beginning of the project and the "RESULTS" and "ACCEPTANCE" sections after your project has been completed. Your sponsor should provide input and sign where indicated. The signed document will also be a required section in your Project Final Report. This document is a template whose sections may be expanded as necessary.*

**Project Title:** Advanced Syllabus Analyzer

**Student Name:** Yabo Dong

**Sponsoring Organization:** NYU School of Professional Studies

**Project Sponsor Name and Title:**  
Andres Fortino Clinical Associate Professor and MASY ACP Leader

**Project Sponsor Contact Information (email and phone):** agf249@nyu.edu  
845-242-7614

### PROJECT PLAN

At project start, show the project goal; the project objectives and related metrics to be used to show successful project completion. Sponsor should sign to indicate agreement.

**Project Goal** \_\_\_\_\_  
The project's goal is to build a module in R with a Shiny interface that accepts a syllabus, scores it against job descriptions scraped from Indeed.com, and provides top industries and occupations that match the syllabus.

**Objective #1** \_\_\_\_\_  
Have weekly meetings with the clients and specify detailed requirements, including features and functions of the module by the end of the course.

**Objective #2** \_\_\_\_\_  
Build an Advanced Syllabus Analyzer based on Shiny interface.

**Objective #3** \_\_\_\_\_  
Provide clients and users with explicit instruction for using the module by the end of the course.

**Objective #4** \_\_\_\_\_  
Use data sets specified by the client to test the validity of the algorithm and module.

**I agree with the above planned project goal, project objectives, and related metrics.**

*Andres Fortino*

2/6/22

**Project Sponsor Signature**

**Date:**

**Printed Name:** Dr. Andres Fortino

Please print in English

## Sponsor's Project Acceptance Document Part 2

### PROJECT RESULTS

Student includes below in "Project Results" the goal and ALL of the Objectives in your project, which you listed in "Project Plan." You may add additional Objectives, as needed.

Planned Start Date: 01/25/2022  
Actual Start Date: 01/25/2022

Planned End Date: 05/03/2022  
Actual End Date: 05/03/2022

If actuals differ from planned dates, the revised dates (Actual) are accepted by the sponsor if initialed here: **Sponsor Initials** *AGF*

#### Project Goal

Was the project goal achieved as planned?  Yes  No, Reason missed: \_\_\_\_\_

If NO, please explain why this is an acceptable deviation. \_\_\_\_\_ **Sponsor Initials** *AGF*

#### Project Objective #1: <as shown above in Plan section>

Did the student's project meet this objective with associated measures and metrics as established at project inception? **Objective#1**  has or  has not been met. **Sponsor Initials** *AGF*  
If not met please explain why this is or is not an acceptable deviation.

#### Project Objective #2: <as shown above in Plan section>

Did the student's project meet this objective with associated measures and metrics as established at project inception? **Objective#2**  has or  has not been met. **Sponsor Initials** *AGF*  
If not met please explain why this is or is not an acceptable deviation.

#### Project Objective #3: <as shown above in Plan section>

Did the student's project meet this objective with associated measures and metrics as established at project inception? **Objective#3**  has or  has not been met. **Sponsor Initials** *AGF*  
If not met please explain why this is or is not an acceptable deviation.

### PROJECT ACCEPTANCE

#### Sponsor's Evaluation of student's performance:

A  A-  B+  B  B-  C+  C  C-  F

**Sponsor's Comments:** Yabo did a great job with a difficult project and should be commended for his willingness to coach others

Project was completed satisfactorily and is hereby accepted

Project was completed satisfactorily but did not meet all objectives, as shown above.  
The Project is, nevertheless, accepted.

*Andres Fortino*

04/29/22

**Project Sponsor Signature**

**Date:**

*Yabo Douz*

04/29/22

**Student Signature**

**Date:**

**Please sign in English. Sign by hand or e-signature. Do not type or use handwriting font.**

0ae9c3a13078d222cb7a15195723b2fe

## Appendix#3: Project Charter

# Advanced Syllabus Analyzer Project Charter

**Project Manager:** Yabo Dong

**Sponsor:** Andres Fortino

**Prepared by:** Yabo Dong

**Name and Location of Client Organization:**

NYU School of Professional Studies, 7 East 12<sup>th</sup> Street

**Revision History**

Revision date	Revised by	Approved by	Description of change

**Project Goal**

The project aims to build a module in R with a Shiny interface that accepts a syllabus, scores it against job descriptions scraped from Indeed.com and provides top industries and occupations that match the syllabus.

**Problem/Opportunity Definition**

Professors at NYU School of Professional Studies plan to specify what occupations and jobs match the courses they teach, which intend to improve the course structure and help students understand how the courses fit their career goals.

**Proposed Project Description**

The project is about developing an Advanced Syllabus Module using the knowledge I received from the MASY program at NYU SPS. This project involves building a module that receives a syllabus text file and generates up to 20 occupations options. After the user selects several occupations, the module will parse 20 jobs for each occupation from Indeed.com.

**Project Sponsor**

- Dr. Andres Fortino
- Clinical Associate Professor and MASY ACP Leader
- Role on the project

**Objectives:**

- Have weekly meetings with the clients and specify detailed requirements, including features and functions of the module by the end of the course.

- Build an Advanced Syllabus Analyzer based on Shiny interface. Measurement: Deliver in R language by the end of the project.
- Provide clients and users with explicit instruction for using the module by the end of the course.
- Use data sets specified by the client to test the validity of the algorithm and module.

**Budget objectives:**

	<b>Planned</b>	<b>Actual</b>
Salaries	N/A	N/A
Documentation	N/A	N/A
Construction	N/A	N/A
Mover	N/A	N/A
<b>Total</b>	<b>\$ N/A</b>	<b>N/A</b>

**Project Selection & Ranking Criteria**

**Project benefit category:**

- Compliance/Regulatory  Efficiency/Cost reduction
- Revenue increase

**Portfolio fit and interdependencies**

N/A

**Project urgency**

Medium

**Cost/Benefit Analysis**

**Tangible Benefits**

Benefit: N/A

Value & Probability: N/A

Assumptions Driving Value: N/A

**Intangible Benefits**

Benefit: Improve the efficiency of designing the course structure and help professors and students to clarify the career path that the courses match.

**Cost Categories**

Internal Labor hours: 300 hours

External costs:

Equipment, hardware, or software: R studio, Shiny app

List other costs such as travel & training:

N/A



## Financial Return

This project does not have a financial budget.

## Other Business Benefits

The module can be used on both Windows-based PCs and Apple Macs, thereby eliminating the difficulty of locating an appropriate platform. Professors who use the module for Syllabus Analysis will be benefited from it as a career tool. In addition, this module can be used by those conducting research. Lastly, the project is cost-effective and does not incur additional expenses.

## Assumption

1. The person who conducts the project will have the skills required.
2. Transforming the syllabus from word or pdf documents to plain text files is not included in this project.
3. The project team has access to the O\*Net Occupation database and to parse data from Indeed.com.

## Scope

### ■ Quality

- Build a syllabus analysis module using R, and the module can work on both Windows and macOS.
- The user will only use a text file of the syllabus as the input.

### ■ Time

- Define the proposal and specification requirements within two weeks.
- Code and build the module within five weeks.
- Deliver all the codes and user instructions by May 3, 2022.

### ■ Resource Allocation

Utilize the resources necessary to

- Analyze the syllabus to select the twenty most matched occupations.
- Use the Shiny app to build an interface for the module.
- Parse job details from Indeed.com.

## Constraints

1. The project must be completed by May 3, 2022.
2. No additional documents are needed from the users.
3. No changes in the content of the syllabus.

## Risks and Mitigation Strategies

1. Indeed.com does not enable users to parse job details.

2. R is updated to a new version that requires additional time and resources to be familiar with.

### **Communications Plan**

1. Frequency: Weekly meeting
2. Method: Zoom online meeting
3. Content: update the progress to the client and discuss any changes or improvements.

### **Schedule Overview**

**Project Start Date:** Jan 25, 2022

**Estimated Project Completion Date:** May 03, 2022

#### **Major Milestones**

1. Define project proposal – Jan 25, 2022
2. Define specification requirements – Jan 28, 2022
3. Build a basic module – Mar 04, 2022
4. Build an interface for the module – Mar 31, 2022
5. Test and improve the module performance – Apr 15, 2022
6. Generate user manual and instructions – Apr 22, 2022
7. Upload all the files to GitHub – Apr 25, 2022

#### **Impact of Late Delivery**

1. More time will be needed to finish the project.
2. Staff may spend more time classifying and preparing students for the jobs that match the courses.

### **Facilities, Software, Hardware, and Other Resources**

1. R
2. O\*Net Occupation database
3. Shiny app interface
4. Indeed.com

### **Project Evaluation**

1. Project schedule:  
Jan 25, 2022 – May 03, 2022
2. Project weekly status report and dashboard  
Have weekly meetings with sponsors to update the project progress, receive feedback, and discuss the next steps.
3. Project communication plan, issues log, risk register
4. Project monthly status report

## Appendix#4: WBS

Task Name	Duration	Start	Finish
<b>Plan for the project</b>	3 days	Thu 1/6/22	Sun 1/9/22
Project Specification Requirements	5 days	Mon 1/24/22	Fri 1/28/22
Project Sponsor Agreement	1 day	Tue 2/8/22	Tue 2/8/22
Part one of project sponsor acceptance	1 day	Tue 2/8/22	Tue 2/8/22
Project charter	2 days	Wed 2/9/22	Thu 2/10/22
Communication plan	1 day	Fri 2/11/22	Fri 2/11/22
<b>Create module functions</b>	16 days	Sun 2/20/22	Fri 3/11/22
Create TD-IDF text mining function	6 days	Sun 2/20/22	Fri 2/25/22
Project status report	1 day	Thu 2/24/22	Thu 2/24/22
Risk Management plan	1 day	Sat 2/26/22	Sat 2/26/22
WBS work flow	1 day	Mon 2/28/22	Mon 2/28/22
Create data scrape function	5 days	Mon 2/28/22	Fri 3/4/22
Create keywords analysis function	5 days	Mon 3/7/22	Fri 3/11/22
<b>CreateShinyapp interface</b>	11 days	Mon 3/14/22	Mon 3/28/22
<b>Test module performance</b>	9 days	Tue 3/29/22	Fri 4/8/22
Final Report	27 days	Thu 3/31/22	Fri 5/6/22
Status report part 2	3 days	Fri 4/1/22	Tue 4/5/22
Situational Analysis	4 days	Thu 4/7/22	Tue 4/12/22
<b>Create user manual</b>	3 days	Mon 4/11/22	Wed 4/13/22
Literature Survey	5 days	Wed 4/13/22	Tue 4/19/22
<b>Upload all the files to GitHub</b>	1 day	Thu 4/14/22	Thu 4/14/22

## Appendix#5: Reference

- Abrahamsson, P., Salo, O., Ronkainen, J., & Warsta, J. (2002). *Agile software development methods*. Relatorio Tecnico.
- Balijepally, V., Mahapatra, R., Nerur, S., & Price, K. H. (2009). Are two heads better than one for software development? The productivity paradox of pair programming. *MIS Quarterly*, 91-118.
- Beck, K., & Fowler, M. (2001). *Planning extreme programming*. Addison Wesley.
- Boehm, B., & Turner, R. (2003). *People factorrs in software management: Lessons from comparing agile and plan-driven methods*. CrossTalk.
- Burns, T., & Stalker, G. (1961). *The Management of Innovation*. London: Tavistock Publications.
- C., B., & B., H. (2012). An Empirical Identification of Project Management Toolsets and a Comparison among Project Types. *Project Management Journal*.
- Coplien, J. (2006). Organizational patterns. *Enterprise Information Systems VI*, 43-52.
- Drury, M., Conboy, K., & Power, K. (2012). Obstacles to decision making in Agile software development teams. *Journal of Systems and Software*, 1239-1254.
- Emery, J. (2001). Adaptive software development: An expereince report. *Beyond chaos: The expert edge in managing software development*, 273-280.
- Gandossy, R. (2003). The need for speed. *Journey of Business Strategy*, 29-33.
- Greening, J. (2001). Launching extreme programming at a process-intensive company. *IEEE Software*, 27-33.

- Hanssen, G. K., & Fægri, T. E. (2006). Agile customer engagement: A longitudinal qualitative case study. *Proceedings of the 2006 ACM/IEEE International Symposium on Empirical Software Engineering*, 164-173.
- Highsmith, J. (2001). Opening Statement. *Cutter IT Journal*, 2-4.
- Highsmith, J. (2004). *Agile project management: Creating innovative products*. Redwood City: Addison Wesley Longman Publishing Co.
- Holcombe, W. M. (2008). *Running an agile software development project*. Hoboken: Wiley.
- Kelly, A. (2008). *Changing software development: Learning to become agile*. Chichester: GBR: John Wiley & Sons, Ltd.
- Kirkpatrick, R., Walker, J., & Firth, R. (1992). Software development risk management: an SEI appraisal. *SEI Technical Review*. Software Engineering Institute.
- Martin, R. (2003). *Agile software development. Principles, patterns, and practices*. Prentice Hall.
- Matook, S., & Vidgen, R. (2014). Harmonizing critical success factors in agile ISD projects. *Proceedings of the 20th Americas Conference on Information Systems*. Savannah, Georgia.
- McAvoy, J., & Sammon, D. (2005). Agile Methodology Adoption Decisions: An Innovative Approach to Teaching and Learning. *Journal of Information Systems Education*, 409-420.
- Misra, S., Kumar, V., & Kumar, U. (2008). *Success factors of agile software development*.

- Moe, N. B., Aurum, A., & Dybå, T. (2012). Challenges of shared decision-making: A multiple case study of agile software development. *Information and Software Technology*, 853-865.
- Moe, N. B., Dingsøy, T., & Dybå, T. (2010). A teamwork model for understanding an agile team: A case study of a Scrum project. *Information and Software Technology*, 480-491.
- Moløkken, K., & Jørgensen, M. (2005). Expert estimation of web-development projects: Are software professionals in technical roles more optimistic than those in non-technical roles? *Empirical Software Engineering*, 7-30.
- Molokken-Ostfold, K., & Furulund, K. (2007). The relationship between customer collaboration and software project overruns. *AGILE 2007*, 72-83.
- PMI. (2013). *A Guide to the Project Management Body of Knowledge (PMBOK Guide)*. Project Management Institute.
- Przybilla, L., Wiesche, M., & Krcmar, H. (2018). The influence of agile practice on performance in software engineering teams: A subgroup perspective. *Proceedings of the 2018 ACM SIGMIS Conference on Computers and People Research*. Buffalo.
- Rainwater, J. (2002). *Herding cats: A primer for programmers who lead programmers*. SpringerVerlag.
- Rasnacis, A., & Berzise, S. (2017). Method for Adaptation and Implementation of Agile Project Management Methodology. *Procedia Computer Science*.
- Reichheld, F. (2001). Lead for Loyalty. *Harvard Business Review*, 76-84.
- Reifer, D. (2002). How to get the most out of Extreme Programming/ Agile methods. *XP/Agile Universe Chicago 2002*, 185-196.

- Schach, S. R. (2002). *Object-oriented and classical software engineering*. New York: McGraw Hill.
- Spundak, M. (2014). *Mixed agile/traditional project management methodology - reality or illusion?* *Procedia - Social and Behavioral Sciences* 119.939 - 948.
- Stray, V., Sjøberg, D., & Dybå, T. (2016). The daily stand-up meeting: A grounded theory study. *Journal of Systems and Software*, 101-124.
- Tanner, M., & Mackinnon, A. (2015). Sources of interruptions experienced during a scrum sprint. *Electronic Journal of Information Systems Evaluation*, 3-18.
- Wiedemann, A., & Wiesche, M. (2018). How to implement clan control in DevOps teams. *Proceedings of the 24th Americas Conference on Information Systems*. New Orleans.

## Appendix#6: Communication Plan

- **Project Introduction Meeting**

**Purpose:** The sponsor specified the requirements, expectation, and metrics about the project.

**Schedule:** Jan 07, 2022 5:00 – 5:30PM

**Place:** Zoom

**Contact:** Dr. Andres Fortino; Yabo Dong

**Location of Agendas and Minutes:** Zoom, 30mins

- **Specification Requirement meeting**

**Purpose:** The sponsor requires a specification requirement and use case of this project.

**Schedule:** Jan 28, 2022 5:00 – 5:30PM

**Place:** Zoom

**Contact:** Dr. Andres Fortino; Yabo Dong

**Location of Agendas and Minutes:** Zoom, 30mins

- **Weekly meeting**

**Purpose:** Report the progress, receive feedback from sponsor, and discuss updates

**Schedule:** Every Friday 5:00 – 5:30 PM

**Place:** Zoom

**Contact:** Dr. Andres Fortino; Yabo Dong

**Location of Agendas and Minutes:** Zoom, 30mins



- **Final Presentation**

**Purpose:** Present the final project and show how the module works.

**Schedule:** Apr 29, 2022 5:30 PM

**Place:** Zoom

**Contact:** Dr. Andres Fortino; Yabo Dong

**Location of Agendas and Minutes:** Zoom, 30mins

# Appendix#7: Status Report#1

## Advanced Syllabus Analyzer

Status Report – Feb, 2022

**To:** Andres Fortino  
Prof Israel Moskowitz

**cc:** N/A

**From:** Yabo Dong

**Date:** Feb 26, 2022

YOUR ANTICIPATED COMPLETION DATE: May 03, 2022

COMPLETION SEMESTER: Fall 2022

Project Status Areas:	Execution Week <06>		
	Green	Yellow	Red
1. Overall Project Status	Yellow		
2. Project Schedule	Green		
3. Project Deliverables	Green		
4. Issues	Yellow		
5. Project Risks	Yellow		
6. Resources & Collaboration	Green		
7. Budget	Green		
8. Change Status	Green		

1 – Overall Project Status
<b>Status – Overall</b>
<ul style="list-style-type: none"> <li>The current progress of the project is at the coding stage.</li> <li>There is unfamiliarity of R syntax.</li> </ul>

2 – Project Schedule	
Tasks that are not on schedule per work plan	Impact
<ol style="list-style-type: none"> <li>Complete function creating by the end of March.</li> <li>Complete interface creating by the middle of April</li> </ol>	<ol style="list-style-type: none"> <li>Communicate with sponsor to receive feedbacks of the module</li> <li>Test and improve the performance of the module.</li> </ol>

3 – Project Deliverables
<p><b>COMPLETED DELIVERABLES:</b> Project proposal; Project charter; Sponsor Agreement; Project Acceptance; Communication Plan; Specification Requirements.</p> <p><b>UPCOMING DELIVERABLES:</b> R code of the module; WBS; Risk management plan.</p>

4 – Issues
<ol style="list-style-type: none"> <li>Unfamiliarity of R syntax and text mining packages.</li> <li>Unfamiliarity of Shiny app.</li> </ol>

5 – Project Risks	
Potential Risks	Possible Mitigation
<ol style="list-style-type: none"> <li>The schedule may be delayed because of the unfamiliarity of R</li> </ol>	<ol style="list-style-type: none"> <li>Search solutions from Internet, consult sponsor and professor for suggestions, and find courses to learn about R.</li> <li>Have weekly meetings with the sponsor to report the progress and improve or change the function requirements based on feedbacks.</li> </ol>

**6– Resources and Collaboration**

- Resources such as O\*Net Occupation Database are available

**7– Budget**

- No monetary budget are included in this project.

**8– Change Status**

Scope Changes	Status (Requested   Approved   Completed)
N/A	N/A

**Comments/Actions**

--

**8 – Sponsor Signoff**

Sponsor indicates agreement with the above status report.	
Approved	

# Appendix#8: Status Report#2

## Advanced Syllabus Analyzer

### Status Report –Month <Apr, 2022>

**To:** Dr. Andres Fortino  
Prof Israel Moskowitz

**From:** Yabo Dong

**Date:** 04/01/2022

YOUR ANTICIPATED COMPLETION DATE: 05/03/2022

COMPLETION SEMESTER: Spring, 2022

Project Status Areas:	Execution Week <11>		
	Green	Yellow	Red
1. Overall Project Status	Green		
2. Project Schedule	Yellow		
3. Project Deliverables	Green		
4. Issues	Yellow		
5. Project Risks	Green		
6. Resources & Collaboration	Green		
7. Budget	Green		
8. Change Status	Yellow		

1 – Overall Project Status
<b>Status – Overall</b>
<ul style="list-style-type: none"> <li>• The current progress of the project is at the stage of building interface.</li> <li>• There is unfamiliarity with the Shiny interface.</li> <li>• Need to discuss with sponsor about the change on the scraped job number for each occupation.</li> </ul>

2 – Project Schedule	
Tasks that are not on schedule per work plan	Impact
3. Create the interface by 04/15/2022 4. Test and validate the function by 04/20/2022	3. Communicate with the sponsor and improve the function based on the feedbacks. 4. Test and improve the function performance.

3 – Project Deliverables
<p><b>COMPLETED DELIVERABLES:</b>            Project proposal; Project charter; Sponsor Agreement; Project Acceptance; Communication Plan; Specification Requirements; Text Similarity Comparison Function, Job Scrape Function, Work Frequency Analysis Function.</p> <p><b>UPCOMING DELIVERABLES:</b>            Shiny Interface, Literature Survey, and Situational Analysis</p>

4 – Issues
1. Unfamiliarity of creating the Shiny interface.

5 – Project Risks	
Potential Risks	Possible Mitigation
1. The schedule may be delayed because of the unfamiliarity of Shiny interface.	1. Search shiny tutorial from the Internet to learn more about the syntax and logic. 2. Meet with sponsor to discuss the current progress. Currently, the project is under control and can be finished before the required date.

### 6– Resources and Collaboration

- Resources such as O\*Net Occupation Database are accessible.

### 7– Budget

- No monetary budget is included in this project

### 8– Change Status

Scope Changes	Status (Requested   Approved   Completed)
1. change the scraped job number for each occupation from 20 to 15	Approved.

### Comments/Actions

N/A

### 8 – Sponsor Signoff

Sponsor indicates agreement with the above status report.	
Approved	

## Assessment Guidelines

The assessment is designated by one of the three “Traffic Light” colors utilizing the following guidelines:

**Each project should establish the appropriate project slippage metrics for yellow vs red status**

Executive Summary:	Assessment		
	Green	Yellow	Red
Overall Project and Most status areas	No major issues, minimal risk to project, on target with expected outcomes, project on schedule, everyone satisfied with progress.	Some major issues, moderate risk to project, must monitor closely, some internal or/and external dissatisfaction with progress. Project plan slipping by 2+ days.	Significant issues, serious risks to project, significant intervention must occur to achieve success, potential for stoppage of project activity. Project slipping by 5+ days, and resources uncommitted to meet deliverables.