# Initial Report

Name: Worawat Kimpara
Student Number: 1744034
Email: KimparaW@Cardiff.ac.uk
Supervisor: Martin Caminada
Date: 3rd February 2020

## Contents

## Project Title

Automatic Generation of Word Documents for the Year in Industry Programme

## Project Description

For its Year in Industry programme, the School of Computer Science and Informatics uses a document called the Skills Framework for the Information Age (SFIA). This is an industry standard that contains a wide variety of IT-related professional skills (such as "Software Development" or "Testing"). For each of these skills, different levels are identified, at which a professional can operate. In general, the higher levels indicate that a person is doing work of greater complexity, with more influence and at a higher level of autonomy.

Currently, the students in the placement program have to go through the SFIA document and select the two skills that they think are most appropriate for them to reach level 3 (BSc students) or level 4 (MSc students) of at least one of the selected skills by the end of their placement. For each skill chosen, the student chooses the appropriate range starting from their current level and increasing up to a level that is the maximum feasible level that the student may reach during their program. Currently, the forms are created manually, and fields are manually populated by copying between the SFIA document and a word document template. As currently, increasing volumes of students are entering the year in industry program, it is becoming increasingly burdensome for staff to manually do the work. Therefore, the main goal of this project will be to, at the very least, allow the user of the service to provide an input (such as “PROG” and TEST”) and then for the service to generate a document and allow the user to download it and modify it with their name, date, employer, etc. The second part involves the process of selecting the relevant skills for the students. Using natural language selection and machine learning, it may be possible to generate recommendations for students on the most relevant skills for them using the input of a short paragraph about them and what they’re looking for in the program. This would greatly reduce the effort needed by the students on selecting their most appropriate skills as the large volume of information within the SFIA document may be overwhelming when they are trying to manually select the most appropriate skills. Using recommendations, students will be able to analyze the recommendations in depth to see if they’re appropriate and avoid the hassle of reading every skill in the document in depth.

The service will be web based and built on Django Web Framework which can be easily called to allow for simplistic integration with current university services. The use of Django is due to its simplistic content management system and its user authentication system which may be useful for administrators to manage resources such as skill descriptions, templates, learning data for the natural language algorithms, etc. An additional advantage to using Django will allow for the use of python libraries such as docx and Natural Language Toolkit which will be essential in the development of this project.

## Project Aims and Objectives

- Extract all skill and level descriptions for SFIA and store them in a more readily accessible JSON format
- Generate documents populated with multiple skills with specified ranges of levels using a python script running locally
- Create a Django web service to house the document generation in a view
- Create documentation to allow for integration with other university services
- Utilize natural language processing and/or machine learning to detect appropriate skills based on an input paragraph
- Integrate the natural language processing algorithm into the web service

## Work Plan

| Week Number | Plan |
|---|---|
| 1 | Extract all skill and level descriptions for SFIA and store them in a more readily accessible format using a script |
| 2-3 | Create a local python script to generate documents populated with multiple skills with specified ranges of levels |
| 4 | Meet with supervisor to present current progress relating to document generation |
| 4-5 | Create a Django web service to house the document generation in a view |
| 6 | Create documentation to allow for integration with other university services |
| 7 | Meet with supervisor to present current progress relating to the web service |
| 7-8 | Create an algorithm utilizing natural language processing and/or machine learning to detect appropriate SFIA skills based on an input paragraph |
| Easter Break | Integrate the natural language processing algorithm into the web service |
| 9 | Meet with supervisor to discus current progress relating to natural language detection |
| 9-10 | Complete any fixes or changes and improve natural language processing algorithm |
| 10-12 | Complete final report and perform fixes and/or changes if necessary |