School of Information Technology

Department of Computer Science

Mr R Domingues

St no 93398761

Email: riaal.domingues@gmail.com

14 April 2021

MOTIVATION FOR REGISTRATION YEAR THREE FOR MIT BIG DATA SCIENCE AND WORKPLAN

# BACKGROUND

Mr R. Domingues enrolled for MIT Big Data Science in 2019. He completed all the prescribed and elective modules for the first year of enrolment in 2019. In 2020 he enrolled for the Mini-Thesis(MIT 807) and Big Data Science Project (MIT 808). He completed the Big Data Science Project in 2020, and already performed some of the experiments for the Mini-Thesis.

This letter serves as a motivation for registration for a third year of registration for Mr Domingues for the degree MIT Big Data Science to finalise the Mini-Thesis, as well as provide a work plan for finalisation.

# Motivation

The mini-thesis problem statement is already defined. The mini-thesis is investigating the application of Natural Language Processing applied to email bodies to aid forensic investigators sift through vast amounts of emails for potential evidence.

A supervisor has already been assigned and interaction took place to date. Prof. H. Venter is the appointed supervisor, and the topic is within his field of research. Interaction has taken place between the supervisor and Mr. Domingues in 2020 and 2021.

A dataset was obtained and some initial experiments on the data was already performed. The Enron dataset which is often used in research was obtained. Exploratory data analysis has taken place and initial experiments were already performed. Based on the initial experiments the problem statement can now be refined and more detailed experiments conducted.

The thesis structure has already been defined, and work has already started on Chapter 1. The first three chapters will now be written to refine the problem statement, position the work within the context of current literature and define refined experiments to be performed.

Mr. Domingues is a full time employee at Armscor. The MIT Big Data Sciece is a minimum two year degree when studying full time. For full time employees the Master’s study has to take place after hours due to operational matters taking precedence at work. The second year of MIT Big Data Science is therefore difficult to complete within one year on a part time basis, compounded by the challenges experienced in 2020 working from home due to the Covid pandemic.

In summary, Mr Domingues has completed all the coursework, and has made some progress on his mini-thesis within the minimum prescribed period of two years for the degree MIT Big Data Science. He is a full time employee and all studies have to take place after hours. Current progress on the mini-thesis implies that there are no major risks in finalising the mini-thesis in a third year of registration for MIT Big Data Science.

# Work Plan

The proposed work plan to complete the mini-thesis and consequently the degree MIT Big Data Science is:

|  |  |  |
| --- | --- | --- |
| **Ser.** | **Item** | **Target Date** |
|  | Chapter 1: Introduction | End April 2021 |
|  | Chapter 2: Literature study | End May 2021 |
|  | Chapter 3: Experimental design | Mid June 2021 |
|  | Extended experiments | Mid July 2021 |
|  | Chapter 4: Main results | End July 2021 |
|  | Chapter 5 Conclusion | Mid August 2021 |
|  | Revise and proof read | End August 2021 |

R. Domingues (st no 93398761)



Prof H. Venter (Supervisor)