**Summative Assignment 2016**

**Module: Software Systems and Applications III:** Web Technology (COMP3381)

**Assignment weighting:** 25%

**Lecturer:** Boguslaw Obara

**Deadline:** 14:00, 30/11/16

**Submission:** code and examples should be submitted via DUO

**General Requirements: (10%)**

This exercise asks you to write a Python program to understand the principles of a text summarization concept. You may use another language if you prefer, but the same principles of readable and reusable code applies.

Each function should include:

* A good help text explaining the function usage, found by help function. Try using headers!
* Enough comments to aid the user in understanding the algorithm.

Try to keep all your functions small. They should be re-used by other functions making the code easier to read and keeping it organized.

Students will be marked based on: the quality, readability and usability of their code.

**Tasks:**

1. Implement a text summarization algorithm:
   1. Parse Document (20%).
   2. Build the Count Matrix (10%).
   3. Use Singular Value Decomposition (20%).
   4. Cluster (20%).
2. Present text summarization results in a graphical form, e.g. using matplotlib (10%).
3. Validate the implemented algorithm on two selected text documents (in addition to any documents used in the system development) (10%).

**References:**

1. Dipanjan Das and André F. T. Martins, A Survey on Automatic Text Summarization, 2007
2. V. Dalal, A Survey of Extractive and Abstractive Text Summarization Techniques, 2013
3. Vishal Gupta, A Survey of Text Summarization Extractive Techniques, 2010