Code:	COMP47600
Title of Course:	Using Text Analytics to Discover Meaning
Proposer:	Mark Keane

1) General Information

Subject Area:	Computer Science
Module ID:	COMP47600
Semester:	Semester one
Module Co-ordinator:	Mark Keane
College:	Science
Level:	Masters
Module Type:	Postgraduate Module
Capacity:	100
Short Title:	Text Analytics
Long Title:	Using Text Analytics to Discover Meaning
School:	Computer Science
Credits:	10
Passing Grade:	40%

2) Descriptions/Outcomes

Module Description
This course aims is to cover how text analytics is currently used to find important regularities and discover meaning in big data. As such, the course will cover the fundamental techniques and some sample application areas where text analytics is deployed. Initially, the course will cover how raw textual data is pre-processed, the natural language techniques (NLP) used to prepare data for subsequent analysis and the paradigms used for system evaluation. The key techniques used in text analytics will be reviewed; including techniques for computing similarity, classification and clustering of texts, sentiment analysis, and discovering temporal regularities. Classic examples of text analytics from social media, predictive analytics and news media will be discussed as examples of the application of these techniques.

Learning Outcomes
At the end of the course students should have a thorough knowledge of the main techniques used in text analytics, some familiarity with the software used to implement these techniques and a knowledge of some of the main application areas. Students should have developed a knowledge of the main application areas in which these techniques prove useful and know how to evaluate new text-analytics systems.

3) Workload FTE

Lectures:	24
Practicals:	12
Specified Learning Activity:	144
	180

4) Prior Learning

None required.

5) Assessment

Examination	60%
Coursework	40%

6) Course Outline*

Lecture 1: Introduction and Course Overview

Lecture 2: Preparing Data

Lecture 3: Using Simple Frequencies

Lecture 4: Beyond Frequencies

Lecture 5: Similarity

Lecture 6: Classifiers

Lecture 7: Clustering

Lecture 8: Evaluation

Lecture 9: Sentiment Identification

Lecture 10: Sentiment Use

Lecture 9: Text Analytics Application (tbd)

Lecture 12: Words in Time

Lecture 12: Exams

* As this is a news course, the lecture names and contents may change.