

## **Chapter 2**

### **Review of Related Literature and Studies**

This chapter provides an overview of literatures on the previous studies and research conducted by the researcher about Record Management System with Plagiarism.

#### *The Plagiarism.*

A lot of people know what plagiarism is but the knowledge about it was limited and don't know when and how it works. Beeley R. & Paulos L. (2006) stated that plagiarism is about owning the creation or work made by someone else and passing it without giving credit to the owner. The authors also stated that the word plagiarism came from the Latin word *plagiarius* literally mean *kidnapper*. For instance, a writer claims the work of another writer as if the work was made by that writer without giving credit to the owner, it is considered as plagiarized even if it is just a mistake.

Knowing all about plagiarism is very important in working a research paper and at the same time, as a research writer, it is also important that there is system that will check and detect plagiarism locally as how Beeley R. & Paulos L. (2006) define plagiarism. Compile all the research paper and check if there is some plagiarized work with no

credits at all. To avoid some problems about plagiarism, wide knowledge about plagiarism is one of the solution and cite all the sources is the best way.

### *The Laws Against Plagiarism.*

Plagiarism is under Cybercrime Law in the Philippines, specifically the Philippines Cybercrime Prevention Act (RA 10175). According to Department of Justice, plagiarism itself is not a crime but it is illegal if it infringes an author's intellectual property rights. As what DOJ stated, plagiarism now carries a penalty of 3 to 6 years imprisonment and a fine of P50,000 – P150,000 (\$1,200 - \$3,600), if prosecuted under the law.

However, Justice Secretary Lella de Lima point out that these penalties will not be effective to someone copying news items or works created by the government.

### *Plagiarism Detection*

According to Kang, MK. et.al (2008), plagiarism was committed when there are minor alterations and manipulations from the original paper work of others. This includes insertion, deletion or substitution of original used words. With those simple alterations requires intensive string comparisons, it becomes a very huge task.

Each task in plagiarism detection correspond several modules as enumerated by Pai, A. (2013). The first on his modules is to cross-check the title of the submitted documents with in the databases and it will going to identify the author. Second is to calculate the percentage and logarithmic value to number of matching words in the submitted documents. This time in the third module, the system will identify and count the match sentences with in the 2 documents. In the last modules found the most difficult task in plagiarism detection as he declared, it is detecting paraphrasing. Paraphrasing falls in the plagiarism when someone copies the ideas or concept from a registered paper without a citation. In this task, it requires lots computation power, time and resources as he clarified.

### *Plagiarism Algorithms*

There are lots of ways and algorithm proposed for this plagiarism detection. The two most common are the Levenshtein Distance and the Smith Waterman algorithm.

According to Goosken, G. (2008), Levenshtein Distance Algorithm measures the edit distance measure of the 2 string that quantifies the distance of the pronunciations of corresponding words or dialects. It calculates the minimal cost required to change a string or substring to another. In deleting or adding character of original string gets 1 cost, and the insertion will cost 2.

The Smith-Waterman algorithm is a classical method of comparing two strings with a view to identifying highly similar sections within them. It is widely-used in finding good near-matches, or so-called local alignments, within biological sequences Irving, R. (2004).

### *Record Management System*

Johnson J. and McElroy J, Jr. (1995) stated that record management system is required for the organization to composed an information filter for satisfying that data is available to the system for storage are complete and not redundant. Because almost all institution or organizations must need to hold their data. If Record management system applied to our studies we claim that all transactions in a specific respondent can easily access the files. It also secures the data available to the system.

### *Search Engine Optimization*

According to <https://moz.com/beginners-guide-to-seo>, search engine optimization focus on growing visibility of the search engine results.

Search engine optimization encompasses both the technical and creative elements required to improve the performance of the system. In relate to our studies, search engine enhance the performance of the system for

searching the stored files in the database. Also it is need to make the system fast and reliable .