Experiences of University Libraries in Kerala with Anti-Plagiarism Softwares iThenticate and Turnitin

Mini G. Pillai

1. Introduction

Academic writing serves the purpose of communicating thoughts, ideas, and results of research to the public. It is an important part of 'Scientific temper'- a way of life, the famous term coined by Jawaharlal Nehru. Nowadays research scholars and teachers are enthusiastic in publishing the results of their research in conference proceedings, books, and journals. Unleashing one's own work to the public not only gives a person satisfaction and fame but also proves to be academically and professionally beneficial. This has resulted in the urge to publish more. But what is more important is to keep the originality and ethics in the context of the published paper. It is unethical to steal the contents of other's works without giving them proper acknowledgement. Recently this unethical activity called "plagiarism" has grown to an alarming state.

2. Plagiarism

Plagiarism has been defined by different sources. Oxford Dictionary & Thesaurus defines it as "the act of taking the work or idea of someone else and pretend it is one's own".

Dictionary.com web site defines the term as "an act or instance of using or closely imitating the language and thoughts of another author without authorization and the representation of that author's work as one's own, as by not crediting the original author" (Dictionary.com, 2015).

The Internet has made it easier for everyone to plagiarize. Besides the fact that plagiarism tarnishes the image of many reputed institutions, it also violates the copyright of an individual. One out of three editors encounters plagiarism regularly (ithenticate.com, 2015).

2.1. Why Plagiarism?

There are many reasons for a person to indulge in plagiarism. Prabhu Sankar and Ramasesh (2014) have identified some of them as follows:

1. Lack of knowledge-The author may not have deep knowledge on the subject of research.

2. Laziness-The author may not be patient and active enough to work hard.

3. Poor command over language-In this case, the author is forced to go for plagiarism due to the lack of command over English language.

4. Fear of failure/pressure to publish for career advancement-The present grading system in academic institutions puts the writers under pressure to upgrade himself by writing articles or presenting papers

5. Poor research skills-even though the writer may be poor in research skills, the demand for published works on many occasions force him to plagiarize.

6. Desire for instantaneous recognition-Recognition comes where there is sincerity and dedication for a long period. When a person desires for instantaneous recognition, he commits plagiarism.

2.2. Types of Plagiarism

According to El Tahir Ali (2011), depending on the depth of stealing, plagiarism can be classified into five categories:

1. **Copy & Paste Plagiarism**: The act of simply copying the text and pasting without acknowledging the original author.

2. **Word Switch Plagiarism**: Plagiarism by modifying the keywords and phrases of the original text.

3. **Ghost Writer**: The writer writes for someone else.

4. **Self- Plagiarism**: The author republishes from his own earlier published papers

5. **Paraphrasing**: Here the contents of many papers are expressed in one's own words without giving due acknowledgement.

There are two ways of detecting plagiarism; one being the manual method and the other, the use of anti-plagiarism software. The manual method being very tedious and time consuming, the use of anti-plagiarism softwares is gaining more popularity.

2.3. Commercial/Open Source Softwares

Anti-plagiarism softwares are available commercially as well as on open source. Some of them are listed in the table 1.

Table 1:

Available Anti-plagiarism Softwares

Name of Software	Country of Origin	Licensing
1.Turnitin	USA	Commercial
2. PlagScan	Germany	Commercial
3. Viper	England	Open source
4. iThenticate	USA	Commercial
5. Plag Tracker	Ukrain	Commercial
6. Urkund	Sweden	Commercial
7. Source forge	USA	Open source
8. Scanmyessay.com	England	Open source
9. Paper Rater	USA	Open source
10.PlagAware	Germany	Commercial

3. Anti-plagiarism and INFLIBNET Centre

INFLIBNET has been in the forefront for introducing so many revolutionary initiatives in the field of academic libraries. These initiatives include SOUL (Software for University Libraries), Shodhganga - the online digital thesis repository and UGC-INFONET Digital Library Consortium. In the case of anti-plagiarism initiatives too INFLIBNET has played a key role. Though the main motivation behind Shodhganga project was that "fruits of public funded research are to be made "public", it had another aim of preventing or reducing plagiarism. In spite of all these efforts, many universities reported evident cases of plagiarism. So the next step was to introduce anti-plagiarism softwares. In 2014, INFLIBNET allocated One crore rupees for providing access to anti-plagiarism softwares. A trial access of two anti-plagiarism softwares, namely iThenticate and Turnitin was provided to hundred universities in the first phase and ten other universities in the second phase, for one year. Universities in Kerala were also provided access to these softwares. This paper attempts to study the experiences of university libraries in Kerala with these softwares during the trial access period.

4. Turnitin and iThenticate

Turnitin and iThenicate have been developed by iParadigms, Oakland, California. Turnitin was designed for instructors to check student papers against a database of more than 45 billion web pages, more than 337 million student papers, and more than 130 million articles from academic books and publications. When a paper is submitted, by default it is added to the student database. But the instructor can change the settings. Submissions can be maximum length of 100 pages or 20 MB. (On exceeding the limit, non-textual information can be excluded). It accepts all formats like MS Word, Word XML, WordPerfect, PostScript, PDF, HTML, RTF, HWP, plain text, and MS PowerPoint (Turnitin.com, 2015).

iThenticate is most suited for post graduate students, academic staff, and research scholars to check their own work. The submissions are matched against 42 million scholarly journals, conference proceedings, and books from more than 590 leading scientific, technical and medical publishers, 93 million online and offline subscription content and research titles from 30 leading aggregators, databases and content providers, 52 billion web pages archived nearly a decade and also the ProQuest. The submissions are not added to the database. Limit of the submission length is 200 pages or 40 MB. It also accepts all formats including MS Word, Word XML, WordPerfect, PostScript, PDF, HTML, RTF, HWP, plain text, MS PowerPoint, and open office (ODT). iThenticate supports more than 30 languages. But it does not support checking synonym or sentence structure (ithenticate.com, 2015).

Fig I: Report Generation in iThenticate

The figure shows how a similarity report is generated in iThenticate. The submitted document and the original source are shown side by side. The percentage of similarity is shown at the right top corner. In this case it is 7%.

5. Review of Literature

As anti-plagiarism softwares have been introduced only recently in university environment, there is a scarcity of available literature in this field. Yet, research papers which deal with the concept of plagiarism and features of different anti-plagiarism softwares have been reviewed.

Prabhu Sankar and Ramasesh (2014) identified the types of plagiarism, reasons for committing plagiarism, and the consequences thereafter. They emphasized the importance of addressing plagiarism at school and college levels. The importance of plagiarism tools lies in the fact that they prove to be handy for the researcher and supervisor to avoid errors in the thesis besides giving an opportunity to avoid plagiarism.

The paper "Plagiarism Detection Softwares: misconceptions" (2013) which was published in iThenticate official newsletter lists out some of the common misperceptions about these softwares. They are: a) plagiarism detection software automatically detects plagiarism b) plagiarism detectors are inaccurate. c) Plagiarism detectors are easy to deceive d) plagiarism detectors are only useful for uncovering unethical work and e) plagiarism detection softwares are time consuming.

In response to these wrong notions, the authors reported that there is a human element in plagiarism detection, plagiarism detection softwares overcome all cheating methods devised so far, they are useful for detecting accidental duplication and common mistakes in writing and even though they are slightly time consuming, pre-checking gives the authors added confidence.

Bella (2014) also described the different types of plagiarism. According to him, self plagiarism is not as grave as other types of plagiarism. He recommended that the journals should develop a stringent policy towards this and plagiarism when noticed should certainly be reported.

Saini and Prakash (2012) , highlight the popularity of iThenticate and Turnitin as anti-plagiarism softwares in academic environment, based on their features.

El Tahir Ali, Hussam, Abdulla and Snasel (2011) pointed out that plagiarism detection is one of the most important issues of journals, research centres, and conferences. They had also compared five plagiarism detection softwares based on their features and performance. The comparative study reveals that PlagAware is the best one and iThenticate stands next to it.

6. Objectives of the Study

1. To know how many universities in Kerala have got trial access of the anti-plagiarism softwares.

2. To identify the different categories of users and their extent of usage.

3. To find the maximum and minimum percentages of similarity recorded.

4. To find out which subjects have the maximum and minimum similarity.

5. To identify whether any orientation programmes were conducted in these universities.

6. To know whether the person(s) performing the plagiarism detection has acquired training in using the software.

7. Scope and Limitations

The study was conducted in university libraries in Kerala. There are twelve universities in Kerala. Out of them five universities are recently established focusing on specific subjects in which no substantial amount of research has been done. Central University of Kerala has not been given access to these softwares. So they are not included in the study. Therefore the sample consists of the six university libraries, namely the libraries of Kerala University, Thiruvananthapuram, Mahatma Gandhi University, Kottayam, Cochin University of Science and Technology, Ernakulam, Sree Sankaracharya University of Sanskrit, Kaladi, University of Calicut, Kozhikode, and Kannur University, Kannur. Due to time constraints, only the research scholars of Mahatma Gandhi University, Kottayam were interviewed to draw their opinions.

8. Methodology

Survey method was adopted using questionnaires to collect data from university libraries. Interview method was also adopted to collect data on the apprehensions and opinions of the research scholars of Mahatma Gandhi University, Kottayam.

9. Results and Discussions

Questionnaires were sent to eight university libraries. But only five out of them were duly filled and returned. Central University of Kerala and Kerala Agricultural Universities responded that they have not got access to these softwares.

9.1. Access to Software

All the five universities which responded to the study have got access to either of the software whereas Sree Sankaracharya University of Sanskrit and University of Calicut have access to both the softwares and access was enabled to all these universities before August 2014. Cochin University of Science and Technology has got access (www.inflibnet.ac.in, 2015) But they did not respond to the questionnaire.

9.2. Person in Charge of the Software

The Deputy Librarian is in charge of anti-plagiarism software in university libraries.

9.3. Number of Persons with ID and Password

Three university libraries, namely University of Kerala, University of Calicut and Kannur University have authorized two staff members to use the software with the login ID and password whereas in Mahatma Gandhi University, three persons have been endorsed with login ID and password and in Sanskrit University, the Deputy Librarian is the sole authority to use the software. Research scholars have to check for plagiarism before the submission of their thesis. Provision of user ID and password to a limited number of persons will surely constrain the chances of misuse.

9.4. Awareness Programmes to Users

Only University of Calicut and University of Kerala have conducted user awareness programmes on anti-plagiarism softwares. As this is a new initiative to the universities in Kerala, research scholars, and research guides have many doubts and

anxieties. The research guides may have their suggestions on modifying search and the items to be excluded while checking a Ph. D. thesis. These suggestions will help the authorities in framing the policies for use of these softwares.

9.5. Availability of Trained Personnel

The respondents were asked whether the person in charge of the software has got training. All of them have got training from the INFLIBNET centre. Turnitin and iThenticate require a human element in effectively checking plagiarism such as excluding certain chapters and using modifying devices such as filters. Therefore the person should be trained both in the working of the software and research methodology.

9.6. Number of Documents Checked

In three university libraries, the number of documents which were submitted for anti plagiarism detection is below 50 during the trial access period. In Mahatma Gandhi University it is more than 50 and in University of Calicut it is more than 100. In University of Calicut, user awareness programmes were conducted and the users are well aware of the anti-plagiarism softwares.

9.7. Highest and Lowest Percentage of Similarity

Universities recorded highest percentage of similarity ranging from 40% (Sree Sankaracharya University) to 90% (Mahatma Gandhi University) and lowest percentage ranging from 12% (University of Kerala) to 0% (University of Calicut). Increased awareness of plagiarism among the users in University of Calicut is clearly indicated by the low percentage.

9.8. Subjects having Highest and Lowest Similarity

In all the five universities, Social Science subjects show the highest similarity and Science subjects show the lowest similarity.

9.9. Ranking of Category of Users

Universities were asked to rank the users on the basis of their usage. Table 2 shows the ranks obtained by the category of users. Here it seems that in all universities, library staff comes in the first position and teachers in the last position. Teachers and research guides should be made aware of these softwares as they may become members of the body for framing anti-plagiarism policies.

Table 2:
Ranking of Users based on Usage

Category of Users	Rank
Library staff	1
Research scholars	2
M Phil/P G students	3
Teachers/Research Guides	4

9.10. Ranking of Category of Documents

The Documents submitted for plagiarism detection were also ranked. Ph. D. theses come in the first position followed by research papers, assignments, and dissertations. No book was submitted in any of the universities. Table 3 shows the ranking of documents.

Table 3:
Ranking of Category of Documents

Category of documents	Rank
PhD theses	1
Research papers	2
Assignments	3
M.Phil/P G Dissertations	4

9.11. Exclusion of Chapters

Among the five universities studied, three universities, namely University of Kerala, Mahatma Gandhi University and Kannur University submit the whole thesis for plagiarism detection while Sree Sankaracharya University of Sanskrit and University of Calicut exclude bibliography and appendices from being checked. In a thesis, there are other chapters which can be excluded while submitting for plagiarism detection such as Review of literature, Preliminary pages like certificate, title page and index. The INFLIBNET Centre has directed each university to frame its own policies on plagiarism detection. It is hoped that the authorities may take appropriate decision on such matters.

9.12. Features Liked most and Disliked most

Both Turnitin and iThenticate have many useful features for plagiarism detection, of which, all universities appreciate the

features of submission process, originality check, speed of report generation. They dislike the features for modification (they are inadequate) and the submission process in Turnitin is also disliked by Kannur University.

9.13. Other Anti-Plagiarism Softwares Used

None of the five universities except Sree Sankaracharya University of Sanskrit use any other software for plagiarism detection. Sree Sankaracharya University of Sanskrit uses Viper in addition to these two softwares. Plagiarism is a hot topic in the academic community. But anti-plagiarism softwares are yet to gain momentum even in universities. At this juncture, the intervention of INFLIBNET by providing two anti-plagiarism softwares is appropriate and timely.

The author has conducted interviews with some of the research scholars of Mahatma Gandhi University to draw their opinions. Most of them are having some apprehensions on the introduction of these softwares. But librarians can alleviate their anxieties by providing them with awareness programmes on the ways to avoid plagiarism and on the working of these softwares. The following are some of the questions put forward by them.

■ What percentage of copied content is acceptable with citation in a thesis/ dissertation?

Anything copied as such from a work should be put in quotes and due acknowledgement should be given. While checking plagiarism, matter in quotes can be excluded using filters.

■ What is the cut off percentage of content permissible as plagiarism in a thesis/ dissertation?

There is no cut off percentage. As per international standards, up to seven continuous words will not be treated as plagiarism. According to the World Association of Medical Editors (WAME), plagiarism is repetition of 6 consecutive words, or overlapping of 7–11 words in a set of 30 words.

■ How safe is to use iThenticate? Will the material uploaded for plagiarism checking be misused?

Unlike Turnitin, where all submissions are by default added to its database, iThenticate do not include the submissions to its database. Each university entrust a responsible person with user

ID and password to the plagiarism detection. So there is less chance of misuse.

- What are generally omitted from the general idea of plagiarism by iThenticate?

  It depends on the policy formulated by each university. Quotes in quotation marks, Phrases, Small matches can be excluded. Also preliminary pages such as title page, certificate, acknowledgement and certain chapters like bibliography, review of literature, appendices can be excluded

- What is the validity of the report generated by iThenticate regarding plagiarism?

  iThenticate is an internationally accepted anti-plagiarism software and it is provided to Indian Universities by the University Grants Commission

- What are the databases iThenticate cross check against for detecting plagiarism?

  See second paragraph 4

- What are the similar softwares like iThenticate used to detect plagiarism in research?

  See Table 1

10. Suggestions

1. Though anti-plagiarism softwares were developed as early as in 1993, it is a novel initiative as far as researchers in Kerala are concerned. As a result, they are anxious and keen in getting ample information on the working of such softwares. Therefore, orientation programmes should be conducted for the research guides and students.

2. A certificate regarding the originality of documents will become mandatory in universities within no time. The researches should be made aware of plagiarism and the anti-plagiarism softwares. Therefore these should be included in the curriculum for coursework.

3. Most of the anti-plagiarism softwares need a human element in checking plagiarism, that is, or refining the search and avoiding certain chapters or parts of the document being checked. Therefore library staff members should be provided

training in both the working of the software and the process of search.

11. Conclusion

The famous poet Javed Akthar says, "True originality is next to impossible. One is impacted by certain authors, certain ideas and it seeps into his/her work. The newness comes in the way it is presented or in the way it puts together old elements. Occasionally this might lead to an insight or a form that is unexpected" (Raman, 2015). If an author is greatly influenced by one or more scholars, he should evidently acknowledge them. This in fact, improves the quality. What people look for in a research work is not necessarily originality, but honesty. It is optimistically expected that the introduction of anti-plagiarism softwares in universities may lead to originality in academic writing and thereby upholding honesty to oneself.

Acknowledgement

References:

Bella, H. P. (2014). Plagiarism. *Saudi J Med Med Sci,* 2(127).

Bhaktiyari, K. et al. (2014). Ethical and unethical methods of plagiarism prevention in academic writing. *International Education Studies,* 7 (7), 52-58.

El Tahir Ali, D. A., Hussam, M., & Snasel, V. (2011). Overview of and comparison of plagiarism detection tools. In *Proceedings of the Dateso 2011 Annual International Workshop on DAtabases, TExts, Specifications and Objects* 161-172

Plagiarism detection software: misconceptions (2013). *iThenticate Newsletter,* Ed. 16.

Prabhu Sankar, M., & Ramasesh, C. P. (2014). Anti-plagiarism software-a tool to ensure quality. In National Conference on Librarians and Librarianship in Transition: challenges and opportunities, August 22-23, 2014

Raman, U. (2015, February 9). Something new, something fresh. *The Hindu*, p. 4.

Saini, D. K., & Prakash, L. S. (2012). Plagiarism detection in web-based learning management systems and intellectual property rights in the academic environment. *International Journal of Computer Applications*, 57 (14), 6-11.

www.Dictionary.com (Accessed on 05-02-2015)

www.inflibnet.ac.in (Accessed on 05-02-2015)

www.ithenticate.com (Accessed on 10-02-2015)

www.Turnitin.com (Accessed on 10-02-2015)

+ + +