



## Role of librarian in anti-plagiarism campaign with special reference to Turnitin and Urkund

Vinod Kumar<sup>1</sup> and Sandeep Kumar Pathak<sup>2</sup>

<sup>1</sup>Sr. Lib. Inf. Asstt., IIT Mandi

Central Library, IIT Mandi, Kamand -175005, Mandi (HP)

Contact No. 9816608706, email: [vinodmihir@gmail.com](mailto:vinodmihir@gmail.com)

<sup>2</sup>Dy. Librarian

Central Library, IISER Bhopal, Bhopal Bypass Rd, Bhauri, Madhya Pradesh 462066

Contact No. 9009093350, email: [skpathak@iiserb.ac.in](mailto:skpathak@iiserb.ac.in)

### Abstract

**Purpose** – The study aims to explore the role of librarian in anti-plagiarism campaign and consequently evaluate two popular anti-plagiarism software based on their features and by analyzing the similarity report generated by both software.

**Design/methodology/approach** – The study uses a very simple technique of comparison of similarity reports generated by Turnitin and Urkund anti plagiarism software. To achieve the desired results, five manuscripts were submitted to both of the anti plagiarism software and similarity reports were generated. These similarity reports were analyzed and comparison was also done based on these generated reports.

**Findings & Conclusions** –After conducting the study one thing is very clear that it is difficult to decide the actual duplication of work. For example at many points, Turnitin was able to find all possible sources while the Urkund was not. Turnitin software does not have any page restriction but Urkund has. Urkund also restricts no. of submission of manuscript per login however Turnitin do not have such restriction. Despite of properly acknowledging the original source, Turnitin sometimes takes text as duplicate same time, it is time consuming process in Turnitin. Comparatively, Urkund takes less time to provide similarity report. Four out of the five sample papers similarity reports have shown different results. Turnitin offers a comprehensive and accurate similarity index.

**Originality/value** – The study is valuable for those interested in getting the plagiarism report for their written matters. It will help them to decide the tool to be acquired for the purposes at one hand and making the decision on the basis of their requirements at the other. Further it also helps to understand what the plagiarism is and what is not.

**Keywords** - Academic institutions; anti Plagiarism software; Turnitin; Urkund

**Paper type** – Technical paper

**Introduction:** The Plagiarism in the era of internet is one such important aspect specifically in scholastic works which remains in prime focus while conducting research and presenting results. Plagiarism is a road block in research and innovation in academic pursuit. It would be pleasant for anyone to find his/her name in a simple internet browsing. But at the same time no one would be happy to be indulged in copyright issue

which is very common in the current information age. It not only harms the image of an individual but the parent institution's also. In worst cases, it can make one face legal actions also. (Kauffman & Young, 2015) indicated that overall 79.5% of the writers are involved in digital plagiarism. However, many studies shows that more than 50% plagiarized materials is the result of ignorance of researchers and inadequate training in preparing reports and citations. With the advent of ICT the material is easy to copy-paste. At the same time it is also true that now it is very easy to find plagiarized works with the availability of many proprietary and open source plagiarism detection systems. Turnitin and Urkund are such tools which helps professionals in dealing with this problem. While doing this, various features of Turnitin and Urkund have been explored.

**Plagiarism:** In simple words plagiarism is presenting other's work as our own work. If we explore for a technical definition of plagiarism we can find numerous definition having common idea of intellectual theft. Plagiarism is, as defined by Cengage Learning's Encyclopedia ([www.encyclopedia.com](http://www.encyclopedia.com)) is "the unauthorized or unacknowledged appropriation of the words, graphic images, or ideas from another person". It Contains One by one we can define the words used in the definition. First is 'unauthorized appropriation of words' which means the original author does not know the usage of his text which he/she has used in his document and the permission has not been sought for using words. 'Unacknowledged appropriation of words' means the credit has not been given to the original author for using his/her words. However, this unauthorized and unacknowledged appropriation is not limited to text only. It includes images also. The unauthorized use of core idea of the text devised by the original author also falls under the category of plagiarism. Many times the idea of the originator has been copied and interpreted in different words other than the original text. It may be present tense to past tense or vice versa, or changing the perspective of the idea. Also there are



many more ways to rephrase the text. Rephrasing the text is another form of plagiarism. Thus this definition extends the area of plagiarism from text to graphic images and concepts also. Here we are focusing only on text based plagiarism. Except this type there are cases of source code plagiarism also which needs more rigorous detection system.

It is now needed to be cleared that a material is plagiarized only when if:

- It is copied word by word without giving credit to the original work;
- The text is not copied word by word but rephrased in some other way; and
- The idea is hacked and defined in a language other than the original text.

Similarly a text is not plagiarized if:

- It is copied word by word and a proper credit has been given to the original work; and
- Usage of standard phrases (even if they are found elsewhere also);

The higher % of a plagiarism in a report (generated via a plagiarism detection tool) is not the criteria of a work to be plagiarized. Also a work with low % cannot be said an original or free from all claims of plagiarism. It is only the personal analysis of the report after which the decision regarding the plagiarism is made. The report containing low % might have copied some text word by word from some other source without acknowledging it. It means the paper must not clear from the quality process even it has small portion of plagiarized material. On the other hand a paper with a higher % might have an original work. In many cases it has been found that anti-detection tools misses the acknowledgment given and treats the text as the duplicate. In such cases the report has to pass through a personal examination to decide the actual % of plagiarism.

**Role of librarian:** Faculty members in any of the institutions are always burdened with the teaching and research work. They are always surrounded by students to solve their general or specific problem which demands their much of productive time in routine tasks. In such scenario they do not have much time to educate their scholars about the deepness of plagiarism and its various forms. The increasing duplication of work makes it almost impossible for them to personally check and verify each and every report, rectify it, giving necessary instructions and again getting rectified report and again analyze it. The process of making clarity regarding plagiarism at the initial stage of a work

must be a significant contribution of a librarian in the research work. The librarian's duty is to make users aware of the ill-effects of the plagiarism. Librarian's role is to use anti-plagiarism tools and must have access for the institute to facilitate the process of plagiarism detection. The librarian also must make users aware of these tools. There must be periodic training session for users to train them to make use of tools. It has been found that at the initial stage of their project, users are not very much interested in learning a secondary tool. Their priority is to first come out with some findings and presenting the results is the second stage. Librarian must speak to them and tell them the importance of plagiarism related education in order to avoid any confusion in future. It is also necessary to clear their all confusion they may face at the time of writing their proposals, papers and final reports.

Faculty-librarian bonding is also important in this matter. Faculty is the guide of any of the project running. Without their contribution, the anti-plagiarism campaign cannot be said successful. If faculty will ease their time only then users will be able to attend any training session or any other users awareness program. An academic and research integrity is must in any of the intellectual work and it cannot be achieved without the contribution and support of faculty.

In past many incidents were happened when the duplicate works was identified. The punishment were also imposed on the detection of the plagiarized work. However it is an ethical issue more than the legal. Also an intentional plagiarism makes down the reputation of the person as an individual in the society. An educator or an inventor must not be indulged in the plagiarism issues. He/she must crosscheck all possible faults before the presentation of the work.

There are various plagiarism detection tools through which the originality of any document can be checked. Turnitin and Urkund are two of the most popular proprietary tools. In India, These tools can be said the leaders in the field.

**Turnitin:** Turnitin as its brochures says is a "powerful, trusted online program that helps institutions ensure the highest level of academic integrity." It has wide spread usage among 15,000 institutions. Before extracting 'Similarity Index' and generating 'Originality Report', it claims to have access and scan 165 million journal articles including 'Crossref' database, 700 million student papers/dissertations and 62 billion web pages, It also facilitates instructor automatic peer review and Grading system for students' submitted documents. 'eRater' is the language tool which automatically





checks grammar of the submitted documents. Turnitin supports all major and popular file types including MS-Word, Word XML, OpenOffice formats, PDF, HTML etc. It can be integrated with all popular LMSs. Through its Feedback Studio system Turnitin teaches students writing.

Turnitin considers that before generating similarity report, there must be a class (of students) and the submission will only be against an assignment (for the class). In Turnitin, the same process is followed before submitting document for generating originality report. First is creating a 'Class'; and second is adding an 'Assignment' to that class. Multiple assignments can be added in a single class. There is flexibility to specify different parameters for different assignments, which include bibliography/quoted material exclusion, exclusion threshold by word count, multiple submissions by single student, submission in permanent repository of Turnitin. There is also option to choose whether you want to get it checked from institution/student submissions or not to facilitate submission of multiple drafts.

**Urkund:** As per the Urkund document 'Urkund is a fully-automatic machine learning text-recognition system made for detecting, preventing and handling plagiarism.' In this way Urkund is considered as a plagiarism detection system through automatic text-recognition. It is also called a "text-matching tool for monitoring plagiarism." Urkund has more than 5000 institutions as its customer base. In India, MHRD has entrusted UGC the responsibilities to distribute the license of Urkund to Government funded universities and centrally funded technical institutions (CFTIs) in India. UGC has also been given the responsibilities for the training and service supports for the operation of Urkund. Urkund system supports more than 40 LMSs which can be integrated with Urkund. Urkund crosschecks the text through all available sources regardless of the language of the source and covers almost all major languages. Like Turnitin, for text-matching, it can scan through journal articles, published books, web pages and students' submissions. Urkund claims to have reached through all the web pages online, including password protected web pages also. It ensures full data security and safety, as it sells user license to only organizations.

Urkund identifies each submission in the system and does not include the previously submitted document for analysis as original source if it is submitted through a single email. It means it allows corrected submissions multiple times. If the same document will be submitted for checking

plagiarism from a different email address then the previously submitted document will be treated as original source in the analysis report. In the report, it not only shows text-matching results but also suggest correct methods of academic writings.

**Authority & Ownership:** Urkund is a Sweden based company and was started by a group of teachers and students together. Now it is owned by a group of professionals from different field which takes care of its different activities from business development to accounting and technical support.

Turnitin was started as a peer review application by students from four universities of USA in 1998. It was commercially introduced in 2000. Turnitin's parent company iParadigm LLC also runs an informational website plagiarism.org. Turnitin 2 was released in September 2010.

**Support/Documentation:** Turnitin has a very detailed documentation and a very well network of support. Under three different category of Turnitin users i.e. Administrator, Instructor and Student, the usage guide of Turnitin has been provided in detail. Besides, it also provides support to integrate various LMS with Turnitin. Also there is a discussion board to share your own ideas and enquire about others'. Very frequently it sends updates as well as scheduled maintenances emails and alerts. On a single click you can switch to any of twenty popular languages. The best part is to find all documentation in the chosen language.

Urkund has a very illustrative user guide. It provides the detailed presentation of analyzing the generated report. However, the submission part is very easy and does not require much support. The documentation also includes four sections of four different LMs integration with Urkund.

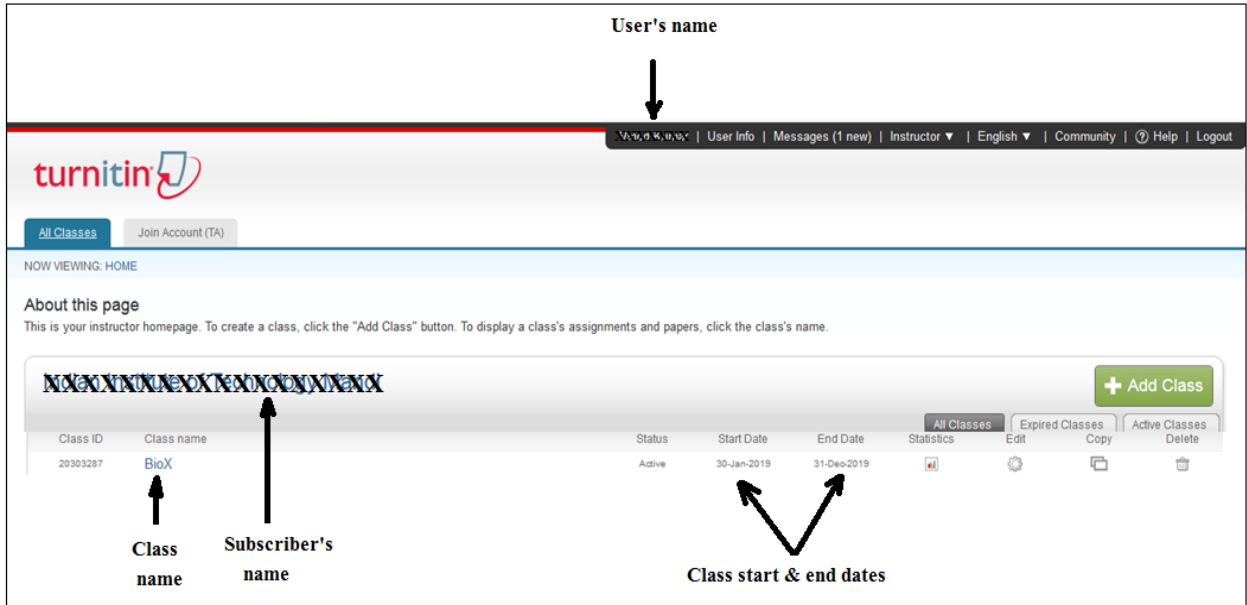
**Data protection and security:** Displayed on the webpage of Turnitin about data privacy is that "Turnitin has procedural safeguards in place to protect personal data, including the use of SSL encryption, redundant servers, and SOC2 certified data centers." However, it also has a detailed privacy policy, which it calls 'Privacy Pledge' is not to use personal data of users for the purpose other than intended.

Urkund also has defined its privacy policy under "Personal Data Processing policy". Urkund commits to not use personal data and submissions beyond the boundary of the intended purposes. It maintains the privacy of personal data as per the regulation mentioned in the European Union's General Data Protection Regulation.



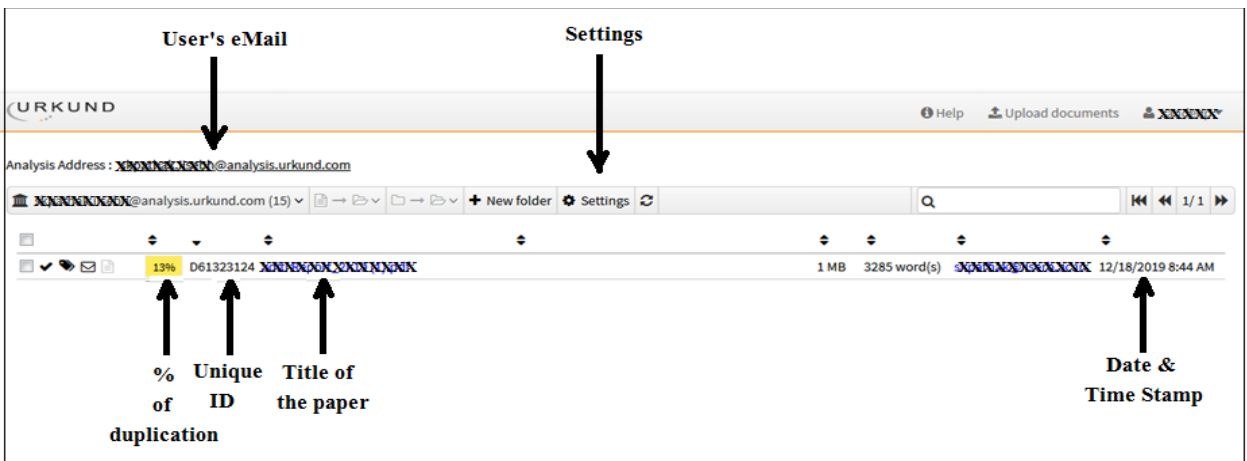
**Interface:** When a user login into his/her account in Turnitin, it shows a clear page except some options on the top which looks like a dashboard of a system. The message is displayed below the dashboard to create a class in requirement to start

using Turnitin. The first time user after login into the account will find the following view of Turitin homepage. The generated report is displayed in the new window.



Turnitin Dashboard

Urkund has two separate interfaces. These are for submission and display of generated report separately.



Urkund Dashboard (Web Inbox page)

The first page is the homepage of 'Web Inbox page'. Web inbox page is a plain simple page with no graphics and with very less options. By starting using Urkund one may need to have basic knowledge of using the interface. The first view of Urkund Web Inbox page is displayed above. When the user clicks on the percentage of the generated report, it opens a new page with very colourful

interface which is completely different from the web inbox page. The first time user is amazed to see this analysis page which is full of graphics and very interactive. There is also option to switch to the older version of the analysis page.

**Subscription policy:** Turnitin is a proprietary system and access to it is only on subscription basis. The subscriber institution has to specify the







no. of 'licenses' at the time of subscription. Licenses are referred to users who are going to use Turnitin and will have an account under the institutional subscription. The subscription cost increases with the increasing of licenses.

Similarly, Urkund can only be accessed under subscription. Urkund has a clear policy to not provide subscription to an individual to maintain integrity and privacy. In India, Urkund has been centrally funded and the distribution has also been centralized through UGC and INFLIBNET.

**Results & Reports:** Both the systems takes not more than 5 minutes in generating final report. However, the analysis of reports in both of the system many times a time taking process and to come up with a super fine result the user has no other way. Turnitin excludes quoted materials but with some exceptions. Urkund include quoted text as similar text. However we can manually exclude that quoted material if we found that it is properly acknowledged by the author. There is an excellent provision to manually remove a source from the significant part if you find it mistakenly included. This provision is in Turnitin also. However in both systems you cannot decide whether the text taken has been properly acknowledged by the author or not until unless you go to the reference part of the paper.

In other words, you cannot analyze a report without downloading it. In order to properly analyze the report, its best way to first download the report, analyze it and then again download the report after removing irrelevant matching sources. Reports from both of the systems do not include the list of sources which are excluded. Turnitin report includes the submission details like time and date etc. In Urkund, this data can be taken separately.

The generated reports in both of the systems provide you possible sources of original materials in the order of relevancy. Also there is provision to choose alternate source over the suggested sources.

**References:**

1. Kauffman, Yashu & Young, Michael. (2015). Digital Plagiarism: An Experimental Study of the Effect of Instructional Goals and Copy-and-Paste Affordance, *Computers & Education*. 83(4), 44-56
2. Plagiarism (2020) Retrieved from <https://www.encyclopedia.com>
3. Turnitin (2020) Retrieved from <https://www.turnitin.com/>
4. Urkund (2020) Retrieved from <https://www.orkund.com/>

In Turnitin, the report page shows the list of original sources but when you go to exclude sources from the list, it takes a shape of long list of sources and you have to exclude one by one all sources to carefully exclude all sources for the duplicate text. Also there is a difference between checked reports from both of the systems.

**Conclusion:** Turnitin and Urkund both the similarity checking software are good. Both the software are having their advantage and disadvantage. Both the software have option for administrator who can manage login for institutional members. Turnitin has less restriction and more user friendly. Turnitin report and analysis is more descriptive and detailed. Urkund software also communicates through email about receipt of the uploaded manuscript and in separate email, also inform about percentage of similarity report.

While generating and analyzing reports, one thing is very clear that there is no such system exists so far which can give assurance of the originality of the contents. We are in the middle of the age of digitization, which we can say the transformation age. The digitization process was started 2-3 decades ago and it sped up only in the current century. At the present pace of digitization, it may still take many more decades in full digitization and specifically in developing countries. This will always remain the limitation of any new system until unless the complete digitization of work is done. Also a full contract between the system proprietors and content providers is the backbone of the success of anti-plagiarism campaign.

Faculty/scholars generally have no time to learn tips and tricks of such system which does not fall under their priority. Learn to use a secondary tool has never been in their list of preferences. It is the duty of Librarians to spread more awareness about subscribed similarity software in campus and also to tell campus academic community about using the subscribed similarity report software.

