

January 2019

# Similarity Index of Doctoral Theses Submitted to Universities in Kerala: An investigation

SURESH KUMAR PK Dr.

*Assistant Librarian, Kerala University Library, pksuresh@yahoo.com*

Follow this and additional works at: <http://digitalcommons.unl.edu/libphilprac>



Part of the [Library and Information Science Commons](#)

---

KUMAR PK, SURESH Dr., "Similarity Index of Doctoral Theses Submitted to Universities in Kerala: An investigation" (2019).  
*Library Philosophy and Practice (e-journal)*. 2130.  
<http://digitalcommons.unl.edu/libphilprac/2130>

# Similarity Index of Doctoral Theses Submitted to Universities in Kerala: An investigation

**Dr.P.K Suresh Kumar**  
Assistant Librarian (UGC)  
Kerala University library  
Thriuvananthapuram-34  
*pksuresh@yahoo.com*

## Abstract

*Plagiarism has grown to such an alarming rate that the academic community finds it difficult to spot an authentic work among the plethora of published literature. Anti-plagiarism softwares have become popular as a solution to overcome this problem. This study aims to describe various aspects of plagiarism, to assess the contribution of PhD theses in ShodhGanga by the universities in Kerala, to identify the subject wise contribution of PhD theses in various universities in Kerala during 2010-17, to find the extend of similarity of content in the PhD theses of the Universities in Kerala and to rank the universities in Kerala with respect to the similarity index. The study is based on 70 PhD theses of three Universities in Kerala viz. University of Kerala, M.G University and University of Calicut selected at random by simple random sample method from the theses available in ShodhGanga. The theses selected were tested by using Urkund, anti-plagiarism software maintained by INFLIBNET. It was found that Science subjects show least similarity index whereas Social Science subjects have the highest similarity rate. Faculty of Arts (26.6%) has a high degree of similarity followed by Humanities (26%) and lowest in Science (16%). Among the Universities University of Calicut occupied the top position with 21 % similarity index in its PhD theses followed by M.G University having 20.4% and the University of Kerala has least with 17.9% of similarity index. The paper points out the importance of user awareness programmes and training programmes on anti-plagiarism for the research guides, research scholars and library staff members.*

**Key Words:** Plagiarism; Similarity index; Shodh Ganga; Universities in Kerala; INFLIBNET; Urkund; Anti-Plagiarism Softwares

## **1. Introduction**

Higher education in the 21<sup>st</sup> century is witnessing a large number of reported cases of plagiarism. It is a contentious issue in universities, it is perceived, by many to be widespread and increasing among the academic community. The inimical act of copious copying of other scholars' intellectual property devoid of proper attribution or what is popularly regarded as 'cut and paste' syndrome is a clear demonstration of the culture of mediocrity. The incidents of plagiarism nowadays seem to be on the increase especially with the advent of Internet which made information more easily available and accessible without any geographical barriers; therefore, researchers can have access to any document in any part of

the world as easily as possible. The prevalence is traceable to the introduction of information and communication technologies (ICTs) in education as well as the plethora of online resources. (Gow, 2013)

Nowadays the academic community especially 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 him satisfaction and fame but also proves to be academically and professionally beneficial. This has resulted in 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. This unethical activity called "plagiarism" has grown to an alarming state recently ( Mini G Pillai ,2015). At present in Indian Universities, there are no accepted anti plagiarism policy to prevent plagiarism. But sincere efforts are being made in this direction by INFLIBNET. Here LIS professionals can play a vital role in creating awareness among the academic community through orientation, workshops etc. It is true, that no one can prevent plagiarism but sincere efforts can be made to reduce plagiarism in the higher education sector. Many cases were reported with respect to the academic honesty in the Indian universities.

The UGC Regulation "Minimum standards & procedure for award of M.Phil / PhD degree", 1<sup>st</sup> June 2009, mandated all universities to check plagiarism using detection tools of all the theses and dissertations which are uploaded into Shodhganga. This would overcome severe problem of duplication of research and poor visibility and hidden factor in research output. So, all universities of India are bound to check the plagiarism before uploading the theses into Shodhganga. The universities are also mandated to set up an Institutional Repository (IR) which would host all theses and dissertation, permitting anyone to access, browse, and view. For the research scholars at the Ph D level it is necessary to have two publications in peer reviewed journals before they are awarded doctoral degree by the university. So, the researchers at the university have no option but to follow good practices of scholarly communication. Many Indian universities have made mandatory for its students and research scholars to check their theses and dissertations using anti plagiarism software before final submission so that they may be able to check originality and quality of their work submitted for acquiring a higher degree. All of IITs have devised a very strict policy against plagiarism.

In Indian Universities INFLIBNET has played a key role in reducing or preventing plagiarism. The issue of academic integrity or dishonesty is so crucial in the Universities. In

2014 INFLIBNET has allocated one crore rupees for providing access to anti-plagiarism softwares and a trial access of two anti-plagiarism softwares, namely *iThenticate* and *Turnitin*. It 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. But later these softwares were found to be biased and not reliable in Indian universities. Later INFLIBNET Centre has selected 'Urkund' Anti-Plagiarism software from M/s. eGalactic, Pune ([www.urbund.com](http://www.urbund.com)) through the global tender and access is recommended on 1st August, 2015 onwards. Universities who had signed MoU and eligible for funding from UGC will be getting the software free of cost from INFLIBNET Centre. The accounts were created on 1st August 2015. University Coordinators who were recommended for managing Anti plagiarism software by the University will be responsible for providing access to the research guides of the concerned universities to provide links to upload the file for plagiarism checking. In order to test the document, user can email to the analysis address contained in their registration email or upload through the user account. Results once generated will be mailed to the mail address. Once successful, they can enter into their Dashboard for managing the submissions. But at present these kinds of software is for the theses prepared in English language and are not applicable for theses written in Marathi, Sanskrit, Hindi, Urdu, Malayalam and all other regional languages.

Different universities setup certain percentage of plagiarism. i.e. percentage of similarity. Because of this practice quality and standard of theses report has increased up to certain extent. Now a days the academic community is more concerned about the degree of similarity in their research output. Hence this study intends to make awareness about various aspects of plagiarism and to test the degree of similarity in the PhD theses of the Universities in Kerala.

### **3. Objectives of the study**

The major objectives of the study are:

- i. To describe various aspects of plagiarism
- ii. To assess the contribution of PhD theses in ShodhGanga by the universities in Kerala.
- iii. To identify the subject wise contribution of PhD theses in various universities in Kerala during 2010-17

- iv. To find the extent of similarity of content in the PhD theses of the Universities in Kerala.
- v. To rank the universities in Kerala under study with respect to the similarity index.

#### **4. Data and Methodology**

There are state, central, deemed and private universities in Kerala. Even though Kerala has 14 universities, the present study is focused on only three universities viz. University of Kerala, MG University and University of Calicut. The scope of the present study is limited to the PhD theses available in ShodhGanga during 2010-2017 from these universities. The data for the study are 70 PhD theses selected at random by simple random sample method from a population of 564 theses of three major universities in Kerala viz. University of Kerala, MG university and University of Calicut available in ShodhGanga during 2010-17. The theses selected are downloaded from ShodhGanga in pdf format, make it a single file and checked the percentage of similarity using Urkund, an anti-plagiarism software managed by INFLIBENT. The data stored on the home page of Urkund after login. The reports received through e-mail are collected, analyzed by excluding other works of the same author. The similarity indices thus obtained are tabulated, interpreted and results are drawn. The similarity index obtained through the Urkund Software is not equivalent to the percentage of plagiarism since the contents of high profile theses may be used by other scholars for their work cannot be detected as plagiarism through urkund software. This is one of the major constraints of the software used for this study.

#### **5. Review of Literature**

In a study Cleary and Sayers (2017) opined that it is the responsibility of an academic faculty to provide support to students and researchers who are unfamiliar with the honesty in academic writing. Plagiarism software detection shall insist before the submission of papers for publication. Prabhu Sankar and Ramasesh (2014) identify different types of plagiarism, reasons for committing plagiarism and the consequences thereafter. They emphasize 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 theses besides giving an opportunity to avoid plagiarism. Bella (2014) also describes the different types of plagiarism. According to him, self plagiarism is not as grave as other types of plagiarism. He recommends that the journals should develop a stringent policy towards this and plagiarism when noticed should certainly be reported.

Gopalakrishnan (2013) in a paper lists out some of the common misperceptions about anti-plagiarism softwares such as plagiarism detection software automatically detects plagiarism, are inaccurate, easy to deceive, only useful for uncovering unethical work and are time consuming. In response to these mis-notions, the author 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.

The studies reviewed above shows that most of the studies are related to the anti-plagiarism software. But no study has been conducted so far to assess the extent of plagiarism in PhD theses. Hence the present study is more significant.

## 6 Analysis and Discussions

The files downloaded from Shodh Ganga were uploaded in Urkund and the reports received are collected, analyzed, tabulated and interpreted.

### 6.1 Contribution of PhD theses in ShodhGanga by the Universities in Kerala

The share of PhD theses in the repository of Indian theses and dissertations of INFLIBNET by the universities in Kerala is given in Table 1.

**Table 1: Contribution in Shodh Ganga by the universities in Kerala**

Sl.No.	Name of the University	Number of theses	Percentage
1	Mahatma Gandhi University	2275	1.52
2	CUSAT	1897	1.26
3	Kannur University	313	0.21
4	Kerala University	2178	1.45
5	Calicut University	890	0.59
6	Sree Sankaracharya University of Sanskrit	445	0.30
	Other Universities in India (274)	1,41,984	94.67
	<b>Total</b>	<b>1,49,982</b>	<b>100</b>

Nearly 1.5 lakhs theses in different subjects are available in ShodhGanga for full text access and downloading. Table 1 clearly indicates that only 5.33% of the PhD theses are contributed by the Universities in Kerala in the institutional repository of PhD theses. There

were 280 Indian Universities added PhD theses to ShodhGanga and only six universities in Kerala were signed MOU with INFLIBNET.

## 6.2 Sampling Distribution of PhD theses

Number of theses available in ShodhGanga during 2010-2017 by the Universities in Kerala under study is given in Table 2.

**Table 2: Sampling distribution of PhD theses**

Sl.No.	Universities Under Study.	Total number of theses	Sample selected	Percentage
1	University of Kerala	163	20	12.27
2	MG university	336	40	12.0
3	Calicut University	65	10	15.38
	Total	564	70	13.22

Table 2 shows that 70 sample theses are selected for the study from a population of 564. Approximately 12 percent of the theses are selected from each university under study.

## 6.3 Subject-wise distribution of theses

The subject wise distribution of PhD theses under study is provided in Table 3.

**Table 3: Subject distribution of PhD theses**

Sl.No	University of Kerala	MG University	University of Calicut	Total	Percentage	Faculty
Mathematics	1	3	1	5	7.1	Science-26 (37.14%)
Physics	2	3	0	5	7.1	
Chemistry	2	3	1	6	8.6	
Botany	1	2	1	4	5.7	
Zoology	2	3	1	6	8.6	
Political	2	4	0	6	8.6	
History	2	4	1	7	10.0	Arts-36 (51.42%)
Economics	2	4	1	7	10.0	
Psychology	2	2	0	4	5.7	
Education	2	3	1	6	8.6	
Sociology	2	3	1	6	8.6	
English	2	4	2	8	14.3	Humanities-8 (11.43%)

<b>Total</b>	<b>22</b> <b>(31.42%)</b>	<b>38</b> <b>(54.29%)</b>	<b>10</b> <b>(14.29%)</b>	<b>70</b> <b>(100.0)</b>	<b>100.0</b>	
--------------	------------------------------	------------------------------	------------------------------	-----------------------------	--------------	--

It is observed from Table 3 that majority of theses (51.42%) are under the Arts faculty followed by 37.14% in Science and 11.43% in the faculty of Humanities. Among the universities maximum number of theses are from MG university (54.29%), 31.42% from Kerala University and rest 14.29% from the University of Calicut. This dispersion is because of the difference in number of theses available in ShodhGanga by the concerned universities during 2010-2017. Since the lack of reliable results in the similarity checking of language other than English resulted in the omission of such language theses.

#### 6.4 Percentage of similarity in the PhD theses

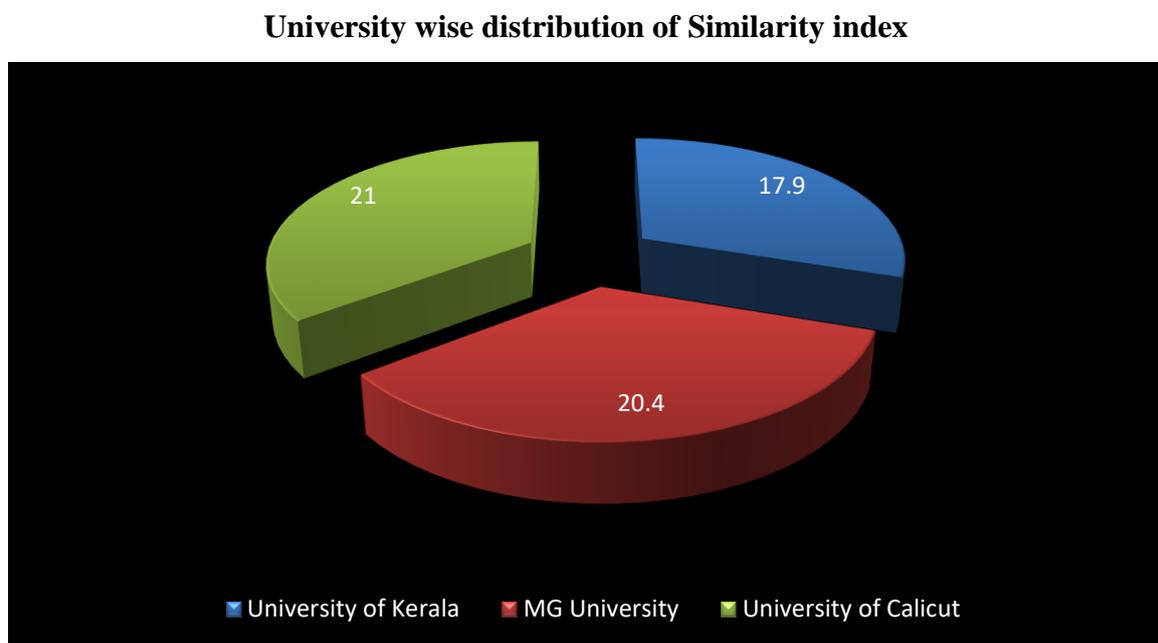
Percentage of similarity in the PhD theses of the Universities in Kerala is given in Table 4. Table 4 shows that the percentage of similarity is high in Sociology and lowest in Botany theses. Faculty of Arts (26.6%) has a high degree of similarity followed by Humanities (26%) and lowest in Science (16%) Among the Universities University of Calicut occupied the top position with 21 % followed by M.G University having 20.4% and the University of Kerala at the bottom level with 17.9% of similarity.

**Table 4: Similarity index in the PhD theses of the Universities in Kerala**

<b>Sl.No</b>	<b>University of Kerala</b>	<b>MG University</b>	<b>University of Calicut</b>	<b>Total</b>	<b>Average</b>	<b>Faculty</b>
Mathematics	14.0	18.0	12.0	44.0	14.7	16.0
Physics	16.0	15.0	NA	31.0	15.5	
Chemistry	18.0	19.0	22.0	59.0	19.7	
Botany	9.0	14.0	18.0	41.0	13.7	
Zoology	12.0	19.0	18.0	49.0	16.3	
Political science	27.0	32.0	NA	59.0	29.5	26.6
History	23.0	28.0	34.0	85.0	28.3	
Economics	20.0	23.0	27.0	70.0	23.3	
Psychology	26.0	22.0	NA	48.0	24.0	
Education	22.0	25.0	26.0	73.0	24.3	
Sociology	28.0	30.0	32.0	90.0	30.0	26.6
English	22.0	28.0	28.0	78.0	26.0	26.0
<b>Average</b>	<b>17.9</b>	<b>20.4</b>	<b>21.0</b>		<b>19.9</b>	

Among Science subjects percentage of similarity of content is high in Chemistry (19.7%) and least in Botany (13.7%). In Arts Sociology (30%) is on top followed by Political Science (29.5%) and Economics at the bottom level with 23.3%. English literature theses have a similarity rate of 26%.

University wise analysis shows that the similarity index is high in the theses of the University of Calicut (21%) followed by MG University (20.4%) and the PhD theses of the University of Kerala have a least amount of similarity (17.9%) . The details are given in Fig. 1.



**Fig. 1**

## 7. Major Findings

- i. There were 280 Indian Universities added PhD theses to ShodhGanga and only six universities in Kerala were signed MOU with INFLIBNET for creating digital repository. Only 5.33% of the PhD theses in ShodhGanga were contributed by the Universities in Kerala.
- ii. Seventy PhD theses were selected for the study from a population of 564 with an approximate of 12 percent of the theses from each university under study. Majority of theses (51.42%) are under the Arts faculty followed by 37.14% in Science and 11.43% in the faculty of Humanities.

- iii. Among the universities maximum number of theses are from MG university (54.29%), 31.42% from Kerala University and rest 14.29% from the University of Calicut.
- iv. Percentage of similarity is high in Sociology and lowest in Botany theses. Faculty wise analysis shows that Arts (26.6%) has a high degree of similarity followed by Humanities (26%) and lowest in Science (16%)
- v. Among Science subjects percentage of similarity is high in Chemistry (19.7%) and least in Botany (13.7%). In Arts faculty Sociology (30%) is on top followed by Political Science (29.5%) and Economics at the bottom with 23.3%. English literature theses have a similarity rate of 26%.
- vi. University wise analysis shows that the similarity index is high in the theses of the University of Calicut (21%) followed by MG University (20.4%) and the University of Kerala (17.9%) has a least percentage of similarity index.

## **8. Suggestions**

- i. Though anti-plagiarism softwares were developed as early as in 1993, it is a novel initiative of INFLIBNET as far as researchers in Kerala are concerned. The main reason for plagiarism by the academic community is that they do not understand what constitutes plagiarism. As a result, they are anxious and keen in getting ample information on the working of Urkund. Therefore, orientation programmes should be conducted for the research guides and students.
- ii. 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 of PhD.
- iii. The anti-plagiarism software, Urkund, need a human element in checking plagiarism, i.e. for 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.

## **9. Conclusion**

PhD theses of the universities in Kerala have an average similarity index of less than 20 percentage. Subject wise analysis shows that the subjects coming under the faculty of Arts and Humanities have a high rate of similarity index compared to Science. There should be

proper guidelines and plagiarism policy in the research works carried out by the academic community to ensure quality because what people look for in a research work is not necessarily originality, but honesty. It is optimistically expected that the introduction of anti-plagiarism software in universities may lead to originality in academic writing and thereby upholding honesty to oneself.

## References

1. Bella, H P (2014). Plagiarism. *Saudi J Med Med Sci* 2, no.127 .
2. Gopalakrishnan, Jessica (2013). Plagiarism Detection Software Misconceptions. *Ithenticate Newsletter* 16 (July 2013). Retrieved from, [http://www.ithenticate.com/plagiarism-detection-blog/bid/95407/Plagiarism-Detection - Software-Misconceptions-July-Newsletter#.WXNoM4SGPIU](http://www.ithenticate.com/plagiarism-detection-blog/bid/95407/Plagiarism-Detection-Software-Misconceptions-July-Newsletter#.WXNoM4SGPIU)
3. Gow, S. (2013). A Cultural Bridge for Academic Concept of Plagiarism. In *Plagiarism Across Europe And Beyond*, Brno, Czech Republic, June 12-13. Retrieved from <https://plagiarism.pefka.mendelu.cz/files/proceedings.pdf>.
4. Michelle Cleary and Jan Sayers ( 2017). Academic Integrity and Plagiarism. *Nurse Author & Editor*,27(.2: 4) Retrieved from <http://naepub.com/ethical-issues/2017-27-2-4/>
5. Mini G Pillai (2015). Experiences of University Libraries in Kerala with Anti-Plagiarism softwares iThenticate and Turnitin. In G. Devarajan ; K.C Abdul Majeed; Dinesh Koovakkai; P.K Suresh Kumar & S.Beena (Eds.) *Blended Libraries and Information Centres: a blue print for the development of information profession in India* (pp.279-291). Trivandrum: Southern Bookstar.
6. Prabhu Sankar, M and Ramasesh, C P (2014). Anti-plagiarism software-a tool to ensure quality, presented in the National Conference on Librarians and Librarianship in Transition: challenges and opportunities (LIBTRANS), August 22-23,2014.
7. Urkund ( 2016 ). Retrieved from <http://shodhganga.inflibnet.ac.in/newmoredetailsfaq-urkund.html>.