



University of Siena
School of Economics and Management
piazza S. Francesco, 7-8, 53100 Siena Italy

Anti-plagiarism System Guidelines for first and second cycle degree theses in SEM

Claudio Pacati

September 7, 2022 – v. 2.1.1

Contents

Introduction	3
1 Purpose and general functions of the system	3
2 General aspects of the procedure	3
2.1 The assignment of the thesis	3
2.2 Normal mode and expert mode	4
2.3 The report and the synthetic indicator	4
Appendix	5
A Answers to teachers' Frequently Asked Questions	5
B Answers to students' Frequently Asked Questions	6

List of Figures

1 System screen with the deletion button highlighted	8
2 Test on a file containing the first paragraph of the <i>Promessi Sposi</i>	8
3 Test on a file containing the first paragraph of the <i>Promessi Sposi</i> with some modifications	9

Introduction

To meet the request of the administrators of the courses of the DEPS and DISAG Departments, the School of Economics and Management (SEM) has prepared an operational protocol to be used so that all theses of the first and the second cycle degrees are subjected to the screening of the [turnitin®](#): a system set up and managed in collaboration with the Library system coordination division. The system is accessible via the Usiena integra e-learning platform using the UnisiPass institutional account.

This guide aims to illustrate the application procedure. Common points and specific aspects are presented for the case of theses with discussion (four-year degrees theses, second cycle degrees and first cycle degrees theses with the traditional final exam) and without discussion (first cycle degrees theses with the final exam, the so called the experimental degree course exam).

1. Purpose and general functions of the system

The plagiarism¹ phenomenon in the first and second cycle degrees is growing very fast in all Italian and foreign universities, as a direct result of the extreme ease in using the *Google-and paste* “practice”. If this is somehow not stopped, it could seriously undermine the reputation of our courses and damage the external perception of the (until now good) value of the degrees issued from us.² For this reason, **SEM has now decided that all theses, submitted by students, should be subjected to this screening system before their final merit is examined.**

The use of the system aims to operate a “mechanical” checking system in order to identify any **potentially** non-original parts (i.e. potential plagiarisms) of a text. For this purpose, the system has access to a large proprietary database containing public and non-public documents, and compares the content of the text submitted with the content of the sources in the database, searching for matching of paragraphs or parts of paragraphs³. The result of the so-called *Originality Report* produced by the system, contains:

- a synthetic numeric indicator that measures the percentage of the text submitted for which textual matching has been found in a source;
- the list of the sources where matching has been found.

It is then possible to highlight all the matches identified between the text submitted and the sources, and functions are available for a thorough investigation of the matches (exclusion of sources from the analysis, ...). The governing board of SEM has identified a **flat threshold value of the synthetic index, equal to 25%** (see section 2.3). It is worth observing that the use of the system is also an advantage for the graduate: his thesis becomes part of the sources’ database and any future possible plagiarism by others will be recognized by [turnitin®](#) and other anti-plagiarism systems⁴.

2. General aspects of the procedure

2.1 The assignment of the thesis

The uploading of the thesis or its parts is carried out by the graduate, according to the instructions and timing found on the [SEM website](#).

The deadline indicated for the thesis assignment is essential, beyond this date there is no guarantee that the necessary IT operations will be carried out in order to allow the student to upload the thesis on time. Nothing prevents one from presenting, or is even much encouraged in fact, to present the form well before the deadline, i.e. as soon as the

¹The aim of this guide is not defining and explaining what plagiarism is. However, this definition may be reported: “Plagiarism is defined as submitting as one’s own work, irrespective of intent to deceive, that which derives in part or in its entirety from the work of others without due acknowledgement. It is both poor scholarship and a breach of academic integrity” ([University-wide statement on plagiarism, University of Cambridge](#)), as well as the mini-guide on [this page](#), that contains navigable references to guides of other institutions.

It should also be pointed out that plagiarism could have civil law consequences (copyright violation, protected by the Italian Civil code, Book V, Title IX, [art. 2575 et seq.](#), and the Law [22 aprile 1941, n. 633](#)), and penal law consequences (see D. Terracina, *La tutela penale del diritto d’autore e dei diritti connessi*, G. Giappichelli Editore, Torino 2006).

²It should be noted that the loss of reputation concerns not only the entire SEM, but also – and in some cases above all – the supervisor of a thesis that plagiarises (in whole or in part) the work of others.

³The length of the parts depends on the configuration parameters in the system.

⁴The producers of anti-plagiarism systems regularly exchange the acquired documents.

supervisor accepts the graduate. In addition, if the supervisor chooses to use the system in expert mode (see section 2.2 below), he/she can only do so after the form has been presented and the student has been enabled by the system.

2.2 Normal mode and expert mode

On the assignment form, the supervisor chooses in which mode he/she will use the system: “normal” or “expert”.

In normal mode, the student will only submit the final and conclusive thesis in the system and only this version will be analysed.

The expert mode allows to use the system in a more gradual way. It is designed for a supervisor who wants to check the thesis during the writing up phase, which could be for example chapter by chapter. The typical procedure should be the following: as soon as a graduate student has written a new part of the thesis and wants to submit this for correction, the student himself needs to upload it onto the system and also inform his/her supervisor. In order to analyse the text the supervisor will retrieve it from the system and then read and correct it. It is then necessary to report to the Student and teaching office that the supervisor will use this mode because these partial sections of the thesis, once loaded, become part of the source database: indeed in the analysis phase of the final version the system needs to be explicitly told not to consider the previous drafts which would otherwise give false assessments. As the supervisor is aware of the history of the revisions of the draft, he/she is the one who has to undertake the task of setting the system in order to ignore the previous drafts. In this mode, the final version of the thesis is loaded by the graduate student in the “final version” slot.

2.3 The report and the synthetic indicator

The system produces a synthetic report, which contains a numeric indicator that quantifies in percentage the matching between the draft and the sources. The system **cannot and should not** provide a final assessment of the presence of actual plagiarisms in the thesis. Final assessment must be carefully considered and given by the evaluator (the supervisor or the commission of the final exam). The system is just a tool, among others, available to the evaluator. Regarding the operation mode of the research software, the system can **only** provide the matching of paragraphs or parts of paragraphs, and therefore it is sometimes possible to fall into **false positives**, such as:

- bibliographical references;
- textual citations between quotes not recognized by the system;
- commonly used phrases (e.g. “without any doubt, we can certainly say that”);
- citations not quoted but clearly expressed in the text (e.g. “Art. 640 identifies the fraud as an act that, using artifices or scams, misleads and procures to oneself or others an unfair profit by harming others.”);
- commonly used mathematical formulas, written with standard symbols (e.g. “the famous Black & Scholes formula states that the price of a European call is $C = SN(d_1) - Ke^{-rT}N(d_2)$ ”).

It is worth noting that there may also be **false negatives** not recognized by the system. The most common case is that of paraphrased text⁵ from a source (quoted or not quoted).

The possible presence of false positives implies that a non-zero value of the synthetic index does not necessarily mean that the draft contains plagiarized parts.⁶

Therefore, on the basis of the experience gained during an experiment, SEM has identified a flat threshold value of the synthetic index equal to 25%: any level at or over this threshold gives a **clue of possible plagiarism**, which the evaluator necessarily has to analyse carefully by checking the matching sections identified by the system.

The identified threshold **should not be interpreted as a “modest amount” of tolerable plagiarism**: it is simply a maximum amount of false positives that are normally found in a text. Furthermore, the possibility of false negatives implies that any level of the indicator within the threshold does not guarantee that there is no plagiarism.

⁵The evaluator (supervisor) with a minimum of experience often recognizes these cases by the twisted phrasing and the use of unusual terms which replace the original text in the paraphrase.

⁶For example, the synthetic index calculated by the system after the analysis of the text of the Italian version of this guide (before inserting the Italian version of this footnote) is 6%; analysing the report, we notice that the system detects only false positives: “Università di Siena”, “School of Economics and Management” (that is present in the heading of each page), “Piazza S. Francesco”, bibliographical references in footnote n. 1, the parts of the phrase “Al momento della presentazione della domanda di laurea” and “l’istituzione che ha rilasciato il titolo di studio”.

Appendix

A. Answers to teachers' Frequently Asked Questions

QUESTION 1. *The system report shows a synthetic index of 35% for my student's thesis, i.e. well beyond the level of alert. Should I reject the thesis?*

ANSWER: No. Such a decision cannot be dictated by the system. You need to check, by using the tools that the system provides you with and decide if there really is plagiarism and consequently then act on the basis of the results of your analysis.

QUESTION 2. *At the graduation session, the thesis of the previous candidate presented a synthetic index of 30%, while the candidate's thesis being exposed now has a value of 10%. Can I ask for 1 point (2 points, ..., n points) in addition for her thesis, given that she presents a smaller amount of plagiarism?*

ANSWER: No. The synthetic index does not measure the amount of plagiarism, but only the percentage of matching found with the system's database sources. A higher value of the index could simply mean a higher number of false positives.

QUESTION 3. *I found a plagiarism (I know the source) in the thesis I'm reading, but the system does not recognize it as such. Can I reproach the graduate and ask him to remove it?*

ANSWER: Yes, not only you can, but you have to. The system can produce false negatives.

QUESTION 4. *I decided to use the system in "expert" mode and I used the system to examine the various chapters of the thesis that the graduate presented for the correction without finding plagiarism. Now however, with the full thesis inserted, the system displays an indicator of 100%. How can this happen?*

ANSWER: The chapters that the student had loaded gradually were entered onto the database. The system therefore found some matching between the full version and the individual chapters. By using the appropriate features of the system, you will need to eliminate, from the matching research, the previously loaded chapters and repeat the analysis.

QUESTION 5. *One of my graduates who attends a double degree program has delivered the thesis and the report shows a value of 100% for the synthetic indicator. Is this a complete plagiarism?*

ANSWER: Not necessarily. The graduate may have already presented the same thesis (as expected) in the partner university that uploaded it to the system. You have to check using the appropriate features of the system.

QUESTION 6. *In the system report I noticed that one of my second-cycle degree students had copied part of his first-cycle degree thesis which had a similar subject. Is this legitimate?*

ANSWER: No. Although *self-plagiarism* is not a crime, it is ethically wrong and not admissible.

QUESTION 7. *The thesis of my graduate is identical (only the author and the title are different) to a thesis presented and discussed by another person at another university. Is this just plagiarism or is it a more serious act?*

ANSWER: It could be considered as fraud or material falsity.

QUESTION 8. *The student did not submit the thesis assignment form within the deadline and now he cannot upload the thesis in the system because he is not enabled. Can we avoid uploading the thesis in order to allow the student to graduate on time?*

ANSWER: No. The loading of the thesis in order to be analysed by the system is mandatory. The student should submit the thesis assignment form as soon as possible, but there is no guarantee that he/she will be enabled in good time.

QUESTION 9. *I realize if one of my students copies a bit and, in that case, I consider this in my finale grade proposal. Why adopt this system?*

ANSWER: When looking at the amount of books, articles, web sites, degree theses of other universities available for consultation, are you sure you can always notice plagiarism? But even if it were possible, plagiarism, even a "limited" amount of it, is an intellectually incorrect practice, with possible civil and criminal implications. In addition, in the future, the thesis could be read by others (competition commissions, potential employers, colleagues, students, ...) who might recognize plagiarized parts: remember that on the title page there is your name marked as the supervisor

who approved the thesis, as well as of course the name of our university which has issued the degree, also on the basis of that thesis.

QUESTION 10. *After my student has uploaded the final version of the thesis, I think that it is appropriate to make some modifications and then reload the new version of the thesis for a new analysis. The graduate tells me that the system does not allow it, because the old version is already present. How should I behave?*

ANSWER: You can delete the file uploaded by the graduate, releasing the space for a new upload. In order to do this, use the tool located on the system's screen at the bottom right of the line of that student, as shown in figure 1 on page 8.

QUESTION 11. *Is the system able to locate copied parts with some changed words here and there?*

ANSWER: Yes. As an example, in figure 2 on page 8 there is the report produced by the system regarding a page containing the first paragraph of the *Promessi Sposi*. As you can see, the system recognizes the complete plagiarism, indicating one of the sources identified, in this case it is a web page. In figure 3 on page 9, you can see the report produced by a file obtained modifying here and there some words in the text. The system is not fooled, and recognizes the almost complete plagiarism.

B. Answers to students' Frequently Asked Questions

QUESTION 12. *I have finished writing the thesis. Before the official deposit, can I check for plagiarism in my thesis?*

ANSWER: Answer: No. The system is a tool to help the supervisor and the commission to evaluate the thesis. It is not made for making "attempts". In any case, there is no reason why anybody should be worried or in the slightest concerned about plagiarism, as it is statistically impossible to happen unintentionally. Those worried can only be hiding something.

QUESTION 13. *I did not submit the thesis assignment form within the scheduled time limit and now I cannot upload the thesis onto the system because I am not authorized. Can we avoid uploading the thesis? Otherwise I risk not graduating on time!*

ANSWER: Read the answer to question 8 in teachers' Frequently Asked Questions section above.

QUESTION 14. *In my thesis, I occasionally used some very well-written and meaningful phrases, taken from texts that I read and quoted in the bibliography. The system detects them as plagiarism, is that correct?*

ANSWER: It depends. Let's look at an example. Suppose that A. Uthor (2016) writes in a text the phrase "It is absolutely obvious that regarding taste, fragrance and even aesthetics, the plain hamburger is absolutely to be preferred to the cheeseburger."

- If you wrote this sentence without citing, i.e. implicitly attributing it to yourself, it is plagiarism and the system has correctly identified it. It is irrelevant that Uthor's text appears in the bibliography.
- If, however, in your writing you put something like "As Uthor (2016) correctly points out, it is absolutely evident that [...]", we are in a particular case: you mention Uthor, but it is not completely clear if you are reporting textually Uthor's observation, and not your elaboration that agrees with that of Uthor. For this reason, it is better to write the sentence cited in quotation marks.
- If the sentence is cited in quotation marks, the system has identified a false positive and your supervisor will be certainly able to recognize it as such.

QUESTION 15. *So, for being cautious, should I write in quotation marks and indicate the source of everything that I have found in books, articles, web sites?*

ANSWER: Yes, but not for caution, but for intellectual honesty. If, however, in doing so, your thesis becomes a list of quoted phrases, the commission will legitimately wonder what your original contribution is.

QUESTION 16. *I have taken some parts of my thesis from some texts and web sites without modifying them: the system has "caught" me. To get a "clean" report, do I just need to change some of the words of the incriminated phrases?*

ANSWER: First of all, it should be underlined that the purpose is not to "have a clean report", but to submit a thesis without contents of others claimed to be yours. The *copy-paste-and-modify-some-words* is as wrong as the *copy-and-paste-without-modifications*, regardless of whether the system will "find out". Read also the contents of note 5 on page 4 and the answer to question 11 in the section dedicated to teachers' questions.

Titolo	Data di inizio	Data di scadenza	Data di pubblicazione	Commenti disponibili	Esporta
Laureando ██████████ (Parte3 )	██████████ 2017 - 20:25 	██████████ 2017 - 20:25 	██████████ 2017 - 20:25 	100 	   
Compiti Peermark (0)					
Show <input type="text" value="10"/> Entries	Cerca <input type="text"/>	 Aggiorna consegne degli elaborati	 Notify Non-Submitters	 Posta in arrivo Turnitin (0)	
					Precedente <input type="text" value="1"/> <input type="text" value="2"/> Successivo

SHOWING 1 TO 10 OF 11 ENTRIES.

<input type="checkbox"/>	Studiante	Titolo della consegna dell'elaborato	N°elaborato Turnitin	Submitted	Similarity	Voto	Voto complessivo	
<input type="checkbox"/>	██████████	finale	76 ██████████ 37	██████████ /2017 12:37	██████████	N/A --	.	

Figure 1: System screen with the deletion button highlighted

1 Quel ramo del lago di Como, che volge a mezzogiorno, tra due catene non interrotte di monti, tutto a seni e a golfi, a seconda dello sporgere e del rientrare di quelli, vien, quasi a un tratto, a restringersi, e a prender corso e figura di fiume, tra un promontorio a destra, e un'ampia costiera dall'altra parte; e il ponte, che ivi congiunge le due rive, par che renda ancor più sensibile all'occhio questa trasformazione, e segni il punto in cui il lago cessa, e l'Adda ricomincia, per ripigliar poi nome di lago dove le rive, allontanandosi di nuovo, lascian l'acqua distendersi e rallentarsi in nuovi golfi e in nuovi seni. La costiera, formata dal deposito di tre grossi torrenti, scende appoggiata a due monti contigui, l'uno detto di san Martino, l'altro, con voce lombarda, il Resegone, dai molti suoi cocuzzoli in fila, che in vero lo fanno somigliare a una sega: talché non è chi, al primo vederlo, purché sia di fronte, come per esempio di su le mura di Milano che guardano a settentrione, non lo discerna tosto, a un tal contrassegno, in quella lunga e vasta giogaia, dagli altri monti di nome più oscuro e di forma più comune. Per un buon pezzo, la costa sale con un pendii lento e continuo; poi si rompe in poggi e in valloncelli, in erte e in is pianate, secondo l'ossatura de' due monti, e il lavoro dell'acque. Il lembo estremo, tagliato dalle foci de' torrenti, è quasi tutto ghiaia e ciottoloni; il resto, campi e vigne, sparse di terre, di ville, di casali; in qualche parte boschi, che si prolungano su per la montagna.

ORIGINALITY REPORT

% 100	% 100	% 96	% 0
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.news-eventicomo.it	% 100
	Internet Source	

EXCLUDE QUOTES ON EXCLUDE MATCHES OFF

EXCLUDE BIBLIOGRAPHY ON

Figure 2: Test on a file containing the first paragraph of the *Promessi Sposi*

1
 3
 2
 1
 1
 1

Quel ramo del lago di Como, che volge a sud, tra due catene montuose contigue, tutto a seni e a golfi, a seconda dello sporgere e del rientrare di quelli, viene, quasi a un tratto, a restringersi, e diventa un fiume, tra un promontorio a destra, e un'ampia costiera dall'altra parte. Il ponte, che congiunge le due rive, sembra rendere ancor più evidente all'occhio questa trasformazione, e segna il punto in cui il lago finisce, e l'Adda ricomincia, per riprendere poi nome di lago dove le rive, allontanandosi di nuovo, lasciano l'acqua distendersi e rallentare in nuovi golfi e in nuovi seni. La costiera, formata dal deposito di tre grossi torrenti, scende appoggiata a 1 e monti vicini, l'uno chiamato di san Martino, l'altro, alla lombarda, il Resegone, dalle molte sue cime in fila, che lo fanno som 1 are a una sega: talché non è chi, vedendolo la prima volta, purché sia di fronte, come per esempio dalle mura di Milano che guardano a nord, non lo riconosca, a un tal contrassegno, in quella lunga e vasta giogaia, dagli altri monti di nome più oscuro e di forma più comune. Per un buon pezzo, la costa sale con un pendio lento e continuo; poi si divide in poggi e in valloncelli, in erte e in spianate, secondo l'ossatura dei due monti, e il lavoro delle acque. Il lembo estremo, tagliato dalle foci dei torrenti, è quasi tutto ghiaia e sassi; il resto, campi e vigne, sparse di terre, di ville, di casali; in qualche parte boschi, che si prolungano su per la montagna.

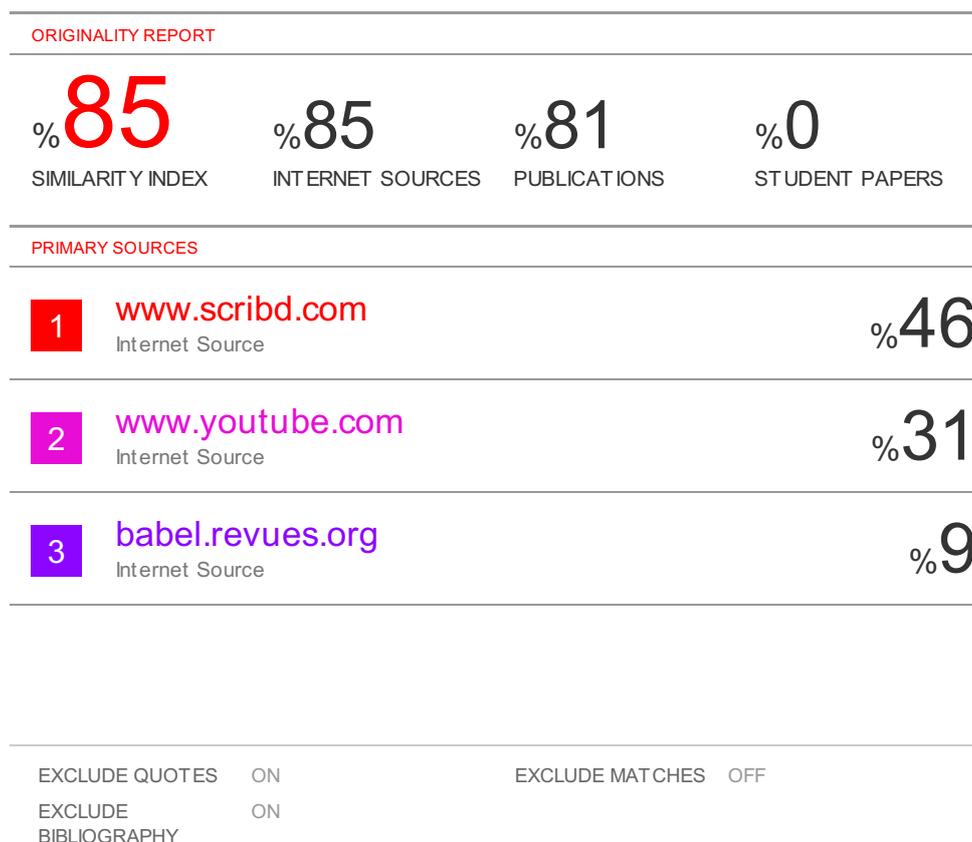


Figure 3: Test on a file containing the first paragraph of the *Promessi Sposi* with some modifications