STRATEGIES AND RESPONSES TO PLAGIARISM IN SLOVAKIA

Július Kravjar, Juraj Noge

Abstract: The rapid increase in the number of higher education institutions (HEI), students, ICT and internet penetration after 1989 and a low level of copyright and intellectual property rights awareness contributed to the growth of plagiarism at HEIs in Slovakia. In 2008, the Ministry of Education decided to establish a common national repository for bachelor, master, dissertation, rigorous and habilitation theses. The amendment to the HEI Act dated 2009 stipulates that every thesis or dissertation of a HEI operating under the Slovak legal order shall be sent to the central repository and undergo the originality check before it is defended. The originality check of every new thesis is made against the national repository, as well as against other resources, primarily the internet ones. From 2010 to the end of March 2013, more than 235 thousand theses and dissertations have gone through this process. The quality of the solution (organisational and technical) meets the expectations and has a positive impact on the academic community. This paper describes the implementation of the Central Repository of Theses and Dissertations and of the Plagiarism Detection System into practice and the experience from their operation. It also analyses the impacts of the implementation on HEIs, feedback and reactions of the academia and the public, and provides recommendations regarding the prevention and detection of plagiarism. It also addresses the growth in the number of internet texts and texts in media containing the term plagiátorstvo (plagiarism) and the growth of search hits (search frequency) for this term in the period 2002–2012.

Introduction

The paper analyses the facts that contributed to the spread of plagiarism at HEIs in Slovakia and the responses to the spreading of plagiarism. The academic community and Ministry of Education (MinEdu) realised the need to stop the thriving plagiarism. Time was needed for the awareness to grow into the belief that the fight against plagiarism requires systemic action on a national level. Finally, nationwide central repository of theses and dissertations and nationwide anti-plagiarism system were implemented.

In the second part of the paper, in relation to the implementation of both systems mentioned above, we analyse the number of texts containing the word plagiátorstvo (plagiarism) on the internet and in the media for the period 2002–2012 and the behaviour of the community regarding the frequency of the searches for the word plagiátorstvo (plagiarism). The aim was to identify whether the adoption of the systemic measure at the national level in 2010 maximised the number of texts and their increases for the period up to 2010 inclusive. A similar question was raised regarding the frequency of searches for the word plagiátorstvo (plagiarism) in the examined period.¹

¹The paper is the English analogy of the paper of Kravjar (2013b)
Abbreviations

CR: Central Repository of Bachelor’s, Master’s, Dissertation, Rigorous and Habilitation Theses

APS: Anti-plagiarism System (Plagiarism Detection System, System for Originality Check)

HEI: Higher Education Institution

SRC: Slovak Rector’s Conference

MinEdu: Ministry of Education, Science, Research and Sport of the Slovak Republic

SCSTI: Slovak Centre of Scientific and Technical Information (CVTI SR: Centrum vedecko-technických informácií Slovenskej republiky)

Explanation

A “small doctorate” can be received by a person with a master’s degree. It requires that a candidate passes rigorous examination and defends his/her rigorous thesis.

A prerequisite for the granting of the scientific-pedagogical degree “docent” (assistant professor) is habilitation lecture, submission and defence of the habilitation thesis.

Background Information

Low awareness of copyright, intellectual property rights and academic ethics combined with a growing number of internet users, students and higher education institutions has been a breeding ground for the spread of plagiarism in written papers at HEIs. According to Skalka et al. (2009b)

Internet was the likely catalyst for an avalanche of plagiarism; it provided students, often including primary school pupils, with resources for quick, easy and effortless access to information about the assigned topic and later even complete texts.

A comprehensive overview of the state of academic ethics in Slovakia is analysed by R. Králiková (2009) in her publication “Implementation of rules of academic ethics at Slovak HEIs”. She states that:

The issue of academic ethics started to attract attention in Slovakia especially in the past three years. It usually appeared in relation to the issue of plagiarism among students and teachers. However, plagiarism is just one of the infractions against academic ethics and the issue of academic ethics in Slovakia is much broader than just plagiarism.

K. Starová (2010) describes the causes supporting plagiarism as follows:

... plagiarism has its roots mainly in the method of teaching; the absence of systematic education (ideally from the primary school level), teaching the students to write argumentative texts, list resources correctly, benefits of quotations (it is much easier to go back to the original source and critically consider the author’s conclusions) and especially ethical principles that would clearly define plagiarism as violation of intellectual property rights and parasitizing on the work of others and not playing down the situation. It is very sad to watch how many teachers who should teach students
Graph 1 and 2. Development of observed indicators over time transformed according to the formula
\[ 100 \times \left( \frac{x_i - x_{\text{median}}}{x_{\text{median}}} \right) \]

ethical principles, explain to them the nature of plagiarism and require them to think
and write originally are themselves guilty of the sin and play down its significance if
caught in the act.

In 1989, there were 13 HEIs in Slovakia and today there are 40; 36 operate under the
Slovak legal order. The development of the number of students, teachers, HEIs (iEDU,
2012) and the number of internet users (www.indexmundi.com, 2012) is shown in Graph
1 and 2. The number of students and teachers apply to HEIs operating under the Slovak
legal order.

The year 2000 was the year of highest annual increase in the number of HEIs (5),
2002 was the year with the highest increase in the number of internet users (nearly 1.5
million, almost 28% of the total population). In 2006, there was the highest annual
increase in the number of students (over 20,000); 2004 and 2005 were also the years
of high annual increases in the number of students (more than 17,000 and more than
18,000 respectively). The highest increase in the number of teachers (774) was recorded
in 2009.

In Slovakia, there were many discussions on the subject of plagiarism, but without
a specific result for a long time. In 2001, the first HEI (Vysoká škola manažmentu—in
Trenčín) in Slovakia started to use an electronic system to detect plagiarism. It was not
only the system to detect plagiarism itself; there were also other related measures: the
rules of academic ethics were adopted, as well as a process defining the procedure for
the investigation of suspected plagiarism (Kročitý, Argaláš, 2012).

In 2006, more than 213,000 students studied at Slovak HEIs for—3.3 times the
number in 1989, there were 32 HEIs—2.4 times the number in 1989; the number
of internet users increased and exceeded 3 million, representing more than 55%
of the total population. However, still only one HEI used an electronic system to
detect plagiarism. In 2008, it was joined by another HEI (University of Economics in
Bratislava) and in 2009 by a third HEI (Comenius University in Bratislava).
Opinions of representatives of HEIs on the use of the system to detect plagiarism varied initially. A significant progress in this area was SRC plenary’s opinion on plagiarism dated 28 September 2006 entitled “Measures to eliminate plagiarism in the processing and presentation of bachelor’s, master’s and dissertation theses” (Slovak Rectors’ Conference, 2006a). The document is important in that:

- Plagiarism is considered a serious problem that must be addressed by informing students how to write papers correctly and how to cite literary sources;
- The best way to reduce plagiarism is prevention;
- In the case of confirmed plagiarism, it requires that the student bear the consequences in accordance with the internal regulations of a HEI.

Another important document approved by SRC on the same day, “The Code of Ethics for HEI Employees” (Slovak Rectors’ Conference, 2006b). However, the Code contains no mention of plagiarism, as if plagiarism never related to teaching and research staff of HEIs.

These two documents did not change the status quo nor did they affect the suppression of plagiarism. A more effective action was needed, with a major impact on combating plagiarism. SRC (2008) adopted a resolution with a potential of a fundamental solution:

The plenary of Slovak Rectors’ Conference revisited the issue of plagiarism. It asked the Ministry of Education of the Slovak Republic to coordinate the relevant activities, especially those relating to the acquisition of the software. It also urges the members of the Slovak Rectors’ Conference to create an electronic archive of theses. It recommends that HEIs modify their regulations to address the issue of plagiarism.

In March 2008, internet media reported (Supuka, 2008) that “the Minister of Education promised to obtain software to control plagiarism for HEIs in Slovakia”. The Ministry of Education decided that a comprehensive solution will be implemented at a national level covering the collection, processing and originality check of specified papers. HEIs in Slovakia will be required to use this comprehensive solution (Skalka et al., 2009a).

The basic strategic goal was defined and it was necessary to develop a strategy to achieve it. One of the major factors that contributed to the success in achieving the strategic objective was finding the support in the Parliament and incorporation in the amendment to the Higher Education Act of the obligations for HEIs to send specified papers in the central repository in order to be checked for originality after registration.

The amendment to the Higher Education Act was approved in 2009 and paved the way for the preparation of a nationwide repository of bachelor, master, dissertation, rigorous and habilitation theses. SCSTI was commissioned by the Ministry of Education to obtain the necessary software by public procurement and it was also made responsible for its operation. The preparation phase was challenging due to a lack of time. The real operation of the central repository and system to detect plagiarism began on 30 April 2010.
Important Moments

The important moments associated with building and operating of both systems are shown below (more details can be found in the paper “Barrier to thriving plagiarism” (Kravjar, 2012)):

- Absence of systemic barrier for plagiarism fight at HEIs;
- Agreement of rectors of HEIs regarding the system for the detection of plagiarism;
- Support for the rectors’ requirement to obtain a system to detect plagiarism from MinEdu;
- Elaboration of the study “Prevencia a odhaľovanie plagiátorstva”² (Skalka, J. et al., 2009c);
- Legislative support from the Parliament—approval of an amendment to the Higher Education Act;
- Appointment of SCSTI as the operator of both systems
- Preparation of methodological materials for HEIs;
- Preparation of the project;
- Procurement of the system, technical, technological and organisational preparation of the systems;
- Effective cooperation in the preparation phase between MinEdu, SCSTI and the supplier of the system SVOP, s. r. o.;
- On 30 April 2010, the routine operation of the CR and APS was started at SCSTI’s premises; all HEIs operating in Slovakia under Slovak legislation are required to send, before the defence, bachelor’s, master’s, dissertation, rigorous and habilitation theses to the CR, where it is archived for 70 years; each work is checked for originality and originality check protocol is delivered to HEIs electronically;
- The new algorithm for the originality check won first place at international competition External Plagiarism Detection Performance at PAN 2011 Lab Uncovering Plagiarism, Authorship, and Social Software Misuse Conference in Amsterdam; it was not only the overall victory, but total victory in all four evaluation indicators;
- Second place at the International Congress ITAPA 2011 (Information Technology and Public Administration) in the category New services;
- Migration of both systems in 2011 to the Data Centre for Research and Development with more efficient, more reliable and more secure ICT infrastructure; the Data Centre is built by SCSTI within the national project “Infrastructure for Research and Development—Data Centre for Research and Development (CVTI SR, 2010)”;
- A change in the system was introduced by another amendment to the Higher Education Act, which imposed on the operator an obligation to make papers stored in the CR available to the public; papers registered in the CR after 31 August 2011 are made available;

²“Prevention and detection of plagiarism”
• In February 2013, we signed up for the competition “European Prize for Innovation in Public Administration” (in the category “Support to Science and Education”) which was announced by the European Commission. In the end of March, we were informed that we had proceeded further and on April 11th we received the information that our application had been shortlisted as one of the 6 most innovative initiatives in our category. The jury will make the final selection on the winning three initiatives in our category based on a hearing and interview.

**Prepared Changes**

CR and APS are live and developing systems responding to everyday operation impulses, as well as to technological trends in the area of plagiarism detection software. We are prepared to develop ANTIPLAG further. Only the most important plans are mentioned:

• Deployment of the winning originality check algorithm from the international competition in Amsterdam ([www.uni-weimar.de](http://www.uni-weimar.de), 2011). This competition involved only English, Spanish and German texts. The algorithm won overall, as well as in each of the four evaluated indicators. One of the indicators was detection of translational plagiarism (cross-languages plagiarism). According to professor Weber-Wulff (Pataki, 2012), who has dealt with plagiarism since 2002, detection of translational plagiarism belongs to greatest weaknesses of anti-plagiarism systems. This claim is based on her experience with testing 48 of such systems from different countries. Detection of translational plagiarism will make the new version of ANTIPLAG a top system in its class.

• The originality check protocol in its paper version will have a more advanced “brother”: a web version with all text matches that ANTIPLAG detected, not only text matches from the first fifteen documents with the highest percentage of over-the-threshold match (which limits the paper version of the protocol).

• Extension of the comparison corpus with papers in the Czech language from internet sources.

• Preparation to extend the comparison corpus with papers in the English language from internet sources.

ANTIPLAG is used by HEIs 24/7 throughout the year. The seasonality of additions in the number of papers in the CRTD is illustrated in Graph 3.

**Reactions to the Implementation of CR and APS into Routine Operation**

According to the statements of HEI officials in the media, the mere publication of the information that operation of such systems is being prepared had a positive preventive effect. Students and teachers realised that their papers will be under permanent supervision of the plagiarism detection system and they began to approach the creation of papers and work with literature more seriously.
Graph 3. Annual cycles of theses and dissertation increases in the CR

Reactions to the implementation of the CR and APS systems in the media agree that they contributed to:

• Increased responsibility of students and teachers;
• Greater student autonomy in the creation of papers;
• Improvement in the quality of papers;
• A higher level of citations.

One quote for all (www.topky.sk, 2010):

Zuzana Pucherová from the Department of Ecology and Environmental Sciences at the Constantine the Philosopher University in Nitra noted that the anti-plagiarism system forced students and teachers to be more responsible. “A teacher has to approach the reading of partial texts with greater precision in order to avoid overlooking something,” she commented. According to her, although a student is the author, the responsibility is borne by teachers/tutors.

The implementation of both systems was a significant step; however, the originality check only covers bachelor’s, master’s, dissertation, rigorous and habilitation theses. There still remains a range of publications produced by HEIs, which are not covered by the originality check. A critique of this situation and plagiarism among teachers was included in E. Mistrík’s (2012) article “Why academics remain silent about plagiarism”. He looks for and finds reasons why academics remain silent. He reveals why some teachers resort to unethical practices and specifies them. He also writes:

However, I feel that the most widespread from various forms of plagiarism is using the papers of their students/postgraduates (“But I taught him”, “But he/she did under my
supervision”), or of their subordinates (“After all, he/she did in my laboratory”, “But he/she is in my department”).

It is a pity that the amendments to the Higher Education Act did not consider feedback. HEIs are not required to disclose information, such as the number of suspected plagiarism cases investigated, how many of them represented a positive finding, and what were the sanctions against the plagiarists. Even with a nationwide system in place, we do not know whether the number of plagiarists is increasing or decreasing. There is a lack of exact data on plagiarism.

The Occurrence of the Term plagiátorstvo (plagiarism) in Texts on the Internet and in the Media

The launch of CR along with APS into live operation in 2010 represents a significant milestone in the fight against plagiarism of theses (bachelor, master, rigorous, dissertation, habilitation) at HEIs in Slovakia. The pre-implementation period and the first months of the operation of these systems was a period when plagiarism often appeared in the traditional media and on the internet.

Our objective was to determine whether this breakthrough event affected the number of texts containing the word plagiátorstvo (plagiarism), as well as their annual increase on the internet and in the media in the period 2002–2010. We assumed 2010 to be the year with both indicators at the maximum.

The base year 2002 was selected based on a recommendation (Ionescu, 2010): Search for a specified time interval: “This option works best for pages published after 2001”. We used Google advanced search with the following parameters:

• Search in Slovak language;
• Search in Slovakia;
• Search only in the text of pages (not in the page title, not in the URL address, not in the links to the page);
• Strictly safe search;
• Last update at any time;
• Any format;
• Rights to use unfiltered by license;
• Unspecified site or domain.

Given that the search for individual years provided either “Number of results” or “Approximate number of results”, we performed the search on a monthly basis in order to reach homogeneity—the search engine always gave “Number of results”. After displaying “Number of results”, we always checked if there is no message after the last result saying: “In order to show you the most relevant results, we have omitted some entries very similar to the N already displayed. If you like, you can repeat the search with the omitted results included.” If there was such a message, the search was repeated and “Number of results” afterwards was considered final.

The 2002–2006 period may be characterised as a period of very low occurrence of information about plagiarism on the internet with given search parameters. This
corresponds with the finding of R. Králiková (2009). However, the internet penetration amounted to 40% and 56% in 2002 and 2006 respectively. A sharp increase in the number of internet texts containing the word *plagiátorstvo* (plagiarism) occurred after 2006 (Graphs 4 and 5).

The data relating to the media monitoring was provided by Storin (Bartko, 2012, 2013). The occurrence of the variants of the word *plagiátorstvo* in the media was monitored (radio, television, press, internet media). The monitoring of radio and press covers the 2002–2012 period, the television monitoring started in 2004 and internet media monitoring started in 2007. Like with the internet search, a very slow growth in the number of texts on plagiarism is characteristic for the initial years 2002–2007 (Graphs 6 and 7).

We assumed that the frequency of texts with the word *plagiátorstvo* on the internet, in media monitoring and the frequency of searches for this word by internet users for the period until 2010 (including their increases) will culminate in 2010. This assumption
was confirmed only partially. The results overview is shown in Table 1; the columns “stagnation” and “moderate growth” do not apply to the last three columns of the table.

Regarding the texts containing the word *plagiátorstvo* on the internet and in the media, only in one case for the 2002–2010 period, the maximum is not achieved in 2010—in the case of press monitoring (maximum in 2009). Regarding the increases, there are four exceptions: overall monitoring, radio and press (maximum in 2009), television monitoring (maximum in 2008). The number of internet texts containing the word *plagiátorstvo* shows a rising trend after 2010, contrary to the media.

The maximum of internet texts with the word *plagiátorstvo* was achieved in December 2010, the value over 50% of the maximum was achieved in November 2009. Values over 40% of the maximum were achieved in October, November 2009 and in March, October, November and December 2010. The maximum increase in the number of texts was not concentrated in the first month of using CRTD and APS, but in the period seven months later.

**The Number of Search Hits for the Word *plagiátorstvo* on the Internet**

On one hand, interest in plagiarism leads to an increasing number of texts about plagiarism on the internet; on the other hand, there is the activity of internet users searching for the word *plagiátorstvo* (*plagiarism*). Google Trends (www.google.com/trends/) provides the ability to quantify the frequency of searched words entered for a period since 2004 per months in percentage compared to the month in which the maximum number of searches was recorded. It should be noted that the zero value does not mean any query, but that the number did not exceed the threshold value. We assumed that the interest in the search for the word *plagiátorstvo* would peak in 2010.

When entering the word *plagiátorstvo* in Google Trends and defining Slovakia, we got a graph with only one over-the-threshold value (June 2011). Given that internet users largely ignore diacritics when entering search queries, we also looked at the trends in searching for the word *plagiatorstvo*. The maximum number of queries (Graph 8) was recorded in May 2010; over-the-threshold values are shown in January to April 2010. An increased interest for the search grows every year in the period of submission of papers.

Google Trends allows identifying not only relative numbers of searches for one word, but also for multiple words. When identifying the number of queries for the words *plagiatorstvo* and *plagiátorstvo* (Graph 9), we get a higher number of over-the-threshold values (a sum of two below-the-threshold values exceeding the threshold value). Results for *plagiátorstvo* and *plagiatorstvo* (Graph 10) are not identical with the results for *plagiatorstvo* and *plagiátorstvo*; however, the maximums are identified in the same months.

Despite the “non-commutativity” in identifying the number of searches, the trends are well-identified. Identifying the number of searches are not invariant in time; a shift in identification by 24 hours identifies small deviations compared with the previous day.

Results regarding the number of searches for the word *plagiátorstvo* document that the use of Slovak language without diacritics when entering search queries dominates the grammatically correct Slovak language.
Figure 11.2. Graph 8. Relative number of searches for the word \textit{plagiatorstvo}

Figure 11.2. Graph 9. Relative number of searches for the word \textit{plagiatorstvo} and \textit{plagiátorstvo}

The first month of the use of CRTD and APS was May 2010 and May showed the maximum number of searches. The local maximums of the number of searches have a year’s periodicity and they belong to the main period of submission of papers (March to May).
Table 1 summarises results regarding the maximums of texts containing the word *plagiátorstvo* on the internet, in the media and related to the maximum interest in searching for the word *plagiátorstvo* or *plagiautorstvo*.

**Summary**

The implementation of the strategic objective to deploy national CR and APS was a fundamental systemic step and brought positive effects in the fight against the spreading of plagiarism. The operation of both systems is a preventive tool; it contributes to improvements in education, to a better understanding and application of academic ethics, copyright and intellectual property rights, and to improvements in students’ work with literature. The papers registered in CRTD from 1 September 2011 are available to the public (www.crzp.sk/crzpopacxe). The public is able to verify suspected plagiarism. We can conclude that theses and dissertation are under three types of control: the tutor’s control, the examination committee’s control and the public control.

However, we do not have exact plagiarism data due to non-existent feedback from HEIs regarding the results of investigations into suspected plagiarism.

The originality check does not cover all HEI publications; there is still room to implement the originality check for other types of HEI publications.

The use of plagiarism detection systems has had longer tradition in many countries than in Slovakia. The implementation of CR and APS and their mandatory use nationwide (at HEIs operating under Slovak legislation) is a unique solution in Europe and very likely in the world.
Table 1
*A brief overview of the results on the internet, media monitoring and the number of searches*

<table>
<thead>
<tr>
<th>The results overview</th>
<th>Number of texts</th>
<th>Relative search number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stagnation</td>
<td>Moderate growth</td>
</tr>
<tr>
<td><strong>Media monitoring total</strong></td>
<td><strong>—year base</strong></td>
<td>2004–2007</td>
</tr>
<tr>
<td><strong>Media monitoring: i-media</strong></td>
<td><strong>—annual base</strong></td>
<td>—</td>
</tr>
</tbody>
</table>

Education and prevention play an important role in combating plagiarism. HEIs are aware of this and improve the moral and knowledge level of students in this regard. However, there is one major drawback: our Higher Education Act does not recognise a withdrawal of a university degree. If moral principles are violated in the preparation of the thesis and it remains undetected until the award of a degree, the degree remains intact in the hands of the owner (Húska, 2012). Therefore, we have seen no withdrawal of degrees yet, for example, in Germany (case Guttenberg, Schavan) and in Hungary (case Schmitt). Hopefully a change is in sight—in August 2012. the newspapers informed (TASR.SK, 2012) that a fraudulently obtained degree will be qualified as a criminal offense. The issue should be subject to the amendment of the Criminal Code.

The number of text containing the word *plagiátorstvo* was very low on the Slovak internet and in the media until 2006 (internet search with the parameters specified above) or until 2007 (media) which corresponded with the number of searches for this word. This was a reflection of the situation of the whole society—the issue of plagiarism, copyright, intellectual property rights was not given sufficient attention by the society. Significant changes started in 2007/2008, when the number of texts on
plagiarism significantly increased, and the issue of plagiarism became a public matter. The term plagiarism is no longer a Cinderella of Slovak internet and media. We are more open to the issue of plagiarism, its presence is openly admitted and there is much more written and spoken word on the topic. The interest in the search for the word plagiátorstvo or plagiatorstvo culminates in spring every year which is the main period for submission of theses and dissertations (Graphs 3, 8, 9 and 10). A large contribution to that was the effort that led to the nationwide implementation of CR and APS.

References


Authors

Július Kravjar, julius.kravjar@cvtsr.sk,
Juraj Noge, juraj.noge@cvtsr.sk,
Slovak Centre of Scientific and Technical Information, Slovakia

Copyright © 2013 The authors listed on page 201 grant to the IPPHEAE 2013 Conference organisers and educational non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full and this copyright statement is reproduced. The authors also grant a non-exclusive licence to Mendel University in Brno, Czech Republic, to publish this document in full on the World Wide Web (prime sites and mirrors) on flash memory drive and in printed form within the IPPHEAE 2013 conference proceedings. Any other usage is prohibited without the express permission of the authors.