CAN WEB 2.0 REDUCE PLAGIARISM AND CHEATING?

Katerina Zdravkova
University Ss Cyril and Methodius
Institute of Informatics
Skopje, Macedonia

ABSTRACT

Plagiarism has existed many centuries ago. From the time when Web 1.0 turned into a host of billions of different sites, information access became practically unrestricted. As a consequence, plagiarism has become a crucial problem in all spheres of life, including education and research. Evolution of the Web from traditional Web 1.0 towards Web 2.0 stimulated a transformation of plagiarism. Furthermore, digital age introduced new ways of cheating, such as identity theft, concealment, fraud and cloning.

This paper presents the experience with different types of plagiarism and cheat noticed during the realisation of several courses at undergraduate and graduate level at two universities. It continues with simple techniques used and intended to discover plagiarism, student trick to cover it, and student remarks against it.

With the time, it appeared that Web 2.0 gave an opportunity to ghost writers and identity fraudsters to successfully play the role of excellent students instead of themselves. Unlike plagiarism, there are still no tangible evidences proving the suspicion of these two kinds of cheat. We conclude that the obviousness of Web 2.0 resulted in the reduction of individual external and internal plagiarism. Unfortunately, it inaugurated new types of cheating. While we decide which of two evils is worse, we will stick to traditional oral examination as a best way to make students at least learn what they plagiarised or their identity clones produced in their name.

I. INTRODUCTION

Web 1.0 placed huge amounts of information at everyone's disposal. Some of the sources are legitimately accessible, such as scientific libraries: UCSF Library [1], IEEE eLearning library [2]; journals: ACM Journal/Transactions [3] and IEEE Journals ans Magasines [4] and free online encyclopaedias: Encyclopaedia Britannica [5], McGraw-Hill's Access Science Encyclopaedia [6], and last but not the least Wikipedia [7]. Information which is distributed to subscribers can be accessed illegitimately by unauthorised users, using peer-to-peer file sharing networks [8] or cracked publications. It is no longer a problem to find out what happened few minutes ago, what was discovered only few days ago, and what are the intentions or future plans.

In this era of massive collaboration, open sourcing, rapid sharing, social media and multilingual translations, a student has never had an easier task to find all the "ingredients" to prepare an excellent assignment without doing a real research. It is no longer inevitable to put an immense effort into solving a task because it is very probable that it has already been stated, solved and published on the Web. The challenge to "masquerade another person's work as one's own" [9], i.e. to make a mosaic or a jigsaw puzzle of already prepared assignments or to incorporate huge parts of other's results in an own final product is at the same time enormous and consequently even more tempting.

For the students, very upsetting is the fact that there are more and more ways to detect plagiarism. There are thousands of effective plagiarism scanners, checkers, and detection tools [10] capable of comparing texts and source codes with contents available on the Web [11], or with contents locally stored. No student can argue against their results [12].

On the contrary, in many occasions, plagiarism is either not noticed, or it is tolerated by teachers. But, even when it is not sanctioned, copyright violation still exists. Sooner or later, it can be discovered and proved. In such case, the consequences are painful [13] and scandalous [14, 15].

In the recent years, there is a high pressure for different professionals to publish enormously. The catchphrase explaining this pressure is well known to all of us: "publish or perish". For real researchers, publishing is usually not a problem, but the need to produce huge amounts of essays, articles, and publications is becoming more and more demanding. Therefore, new professional writers who are hired to write in the name of other people become very popular. They are called ghost-writers. They exist everywhere, particularly in ICT centres. Sean Platt, stated: "Most online writers fall into the same wretched rhythm, starving for jobs an endless tsunami of overseas writers who gladly accept Rupees on the Dollar." [16].

The phenomenon of ghost-writers is directly connected with eLearning, no matter its generation 1.0, or 2.0. Namely, instead of sending the final essay, discussion, blog, or wiki to a student who will afterwards deliver it, new ghost-writers get the identity of students and act instead of them. They are active participants ready to prepare all the assignments credited to the student who hires and pays them. At the moment, it seems that China is the world leader of this infamous tendency [17]. If the plagiarism can be detected [18], the existence of ghost-writers deliberately possessing the identity of a student can be only suspected, but never proved.

This paper continues with the presentation of our experience tackling two types of plagiarising essays, external and internal and the simplest techniques implemented to uncover and prove plagiarism. Both types can be also detected during online submissions within forums, wikis and blogs. Next session is dedicated to teacher's suspicions of identity exchange and student's anonymous assertion proving the suspicion. In the concluding part the comparison of plagiarism and cheat in eLearning 1.0 and eLearning 2.0 is

presented. The paper culminates with currently the only reasonable solution to these student scams.

II. PLAGIARISM AND CHEATING WITHIN ELEARNING 1.0

Before the E-era, there were three basic types for assessing students' knowledge: home work assignments, question papers and oral exams. Sources to prepare home works were limited and usually available to teachers, so whenever the task was original, there was a little likelihood to copy anything without being detected. The list of cheating techniques implemented at written exams was, and still is definite, but far from trivial. However, public oral exam has always been the best way to measure whether student possessed sufficient knowledge to successfully finish the course. But, very few teachers nowadays implement this exhaustive and time consuming assessment.

Since it was initially launched in 2002, Professional Ethics course has been organized as an eLearning course. Students were preparing individual and group essays using materials from the Internet mainly written in English. Plagiarism was noticed, particularly when the writing style was untypical for students. This plagiarism was completely based on external sources, thus we call it an external plagiarism. At that time, final grade was mainly based on oral examinations, polishing the irregularities due to plagiarising external sources.

Many course topics became popular in professional blogs offering almost completely prepared home work essays. As a consequence, external plagiarism has immensely increased. It was exhausting to manually discover whether an essay revealed a research and critical thinking or it was a copy of someone else's study.

There are many plagiarism detecting tools, either commercial [11] or open-source [19]. We implemented the second and discovered several hard cases of literal copying. But, simple Google search also works perfectly. Students usually do not even try to hide it, thinking that mentioning the source is sufficient (Fig. 1.). Detecting external plagiarism is time-consuming for the teacher, but easy to detect.

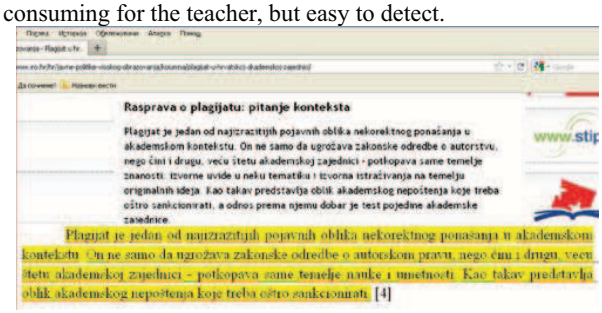


Figure 1. Students think that referencing the source is an amnesty for plagiarism

Meanwhile, Wikipedia with its millions of articles became a very powerful source. Powered by Google translate, it was easier to produce fair essays particularly when the student was willing to edit them manually. However, there were many funny translations such as "чадор мандат", meaning "an umbrella term" confirming the plagiarism due to unedited machine translation (Fig. 2).

We intend to make own plagiarism tool connected with Google and Google translate. It will not prove more than manual check, but the process will be much easier and more efficient.

The most difficult plagiarism to explain to students occurs when they slightly paraphrase many different sources found in their mother tongue and combine them with fair translation of sources into a foreign language, predominantly English. Plagiarism tools are usually not sensitive to this kind of plagiarism, particularly when the translation is fair. But, the problem of stealing other's intellectual property still exists.

Figure 2. Google Translate is an ally of plagiarism

For years, number of students became huge, and it was no longer possible to define individual topics. Copying from colleagues from same generation, or from previous generation has soon become obvious too. We call this copying an internal plagiarism. This scam is obvious when the borrower is too naïve or too insolent. Sometimes, document has no editing time, making the teacher more careful. Sometimes it uncovers another student from the same group. Even without these small mistakes, internal plagiarism within same generation is easily detectable, particularly when only one teacher grades the essays. Some students handle a clone or an identical copy of other colleague's essays sometimes forgetting to at least change basic document properties. In many occasions, student doesn't even try to cover it. But it is not immodest to claim that most teachers have impeccable memory.

Automatic tool can easily detect the same problem by storing all the essays and comparing one with others. Whenever internal plagiarism is caught, students insist on "group" effort.

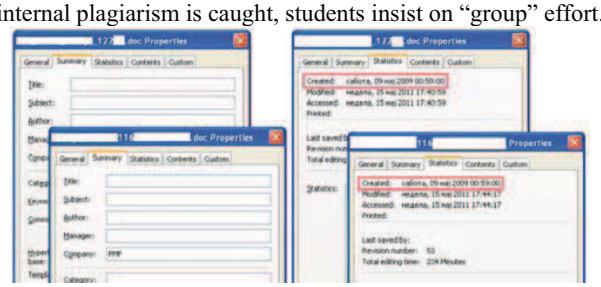


Figure 3. Whenever the time of creation is from earlier period, it is easy to find the earlier author

Concerning repeated assignments from past generations, in many occasions document properties reveal a date in history,

indicating that it was done in the past by another colleague (Fig. 3.).

Interestingly, students prefer those past essays or solutions which are atypical and uncommon, enabling the teacher to immediately recall previous solution and to easily recognize them.

Several years ago, it has been noticed that although different, many essays look alike. Most of them used same broader set of sources, quite regularly distributed between several essays. There were also few irregular words typical for certain dialects, and some philosophical phrases which were not used in ordinary communication. Furthermore, some essays were uploaded from the same IP address. Teacher's suspicion was that all these essays originated from the same author. Few independent teachers from other universities noticed the same. The suspicion concerning this "plagiarism sins" was verified and unfortunately confirmed with an open source plagiarism checker [19] showing substantial similarity between two or several essays. But, it was neither not possible to categorically prove this speculation, nor to find who the real author was. Together with teacher's feeling of ghost-writer presence, "couloirs rumours" proved that students hired other students to prepare their essays for a proper compensation. The amount of students was huge, thus oral examining was replaced by e-testing. Students instantly hacked it, making it obsolete.

Instead of stimulating ethical behaviour, it appeared that the course enabled more plagiarism and cheating. Teacher's grading role turned into searching for fakers and cheaters and sanctioning. This was exactly opposite of the course goal.

As the old proverb says "It is better to prevent than to cure". The first available methods to avoid undesired student activities were eLearning 2.0 facilities of our course management system. They partly reduced individual external and internal plagiarism, but new forms of cheat were born. Next session presents only those tricks or scams we have discovered so far.

III. PLAGIARISM AND CHEATING WITHIN ELEARNING 2.0

In order to reduce undesirable and unethical student behaviour, several Web 2.0 techniques have been implemented in the course. Individual essays were steadily replaced by basic social networking components, predominantly discussion forums, wikis and external blogs. Collaborative creation of final products was also forced. It comprised a creation of either joint essays, or joint wikis. Each joint outcome was coordinated and moderated by students who demonstrated high level of self-consciousness, meaning that their previous essays and individual discussions were faultless.

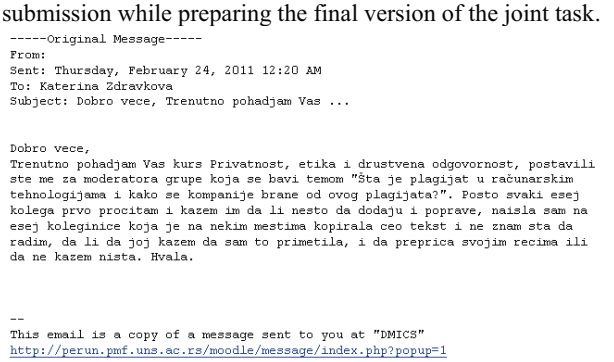
The effort to participate in such an interactive course was highly challenging both for the students and for the teacher. It needed an instant presence and immediate reaction throughout the whole course. For example, students accessed each discussion forum in average more than 40 times, and contributed into it with in average more than 7 posts. Discussion forums intended to support the creation of joint products were ever more visited and the average contribution exceeded 9 posts per student.

In the independent wikis students created hundreds of interlinked pages editing them in average almost 32 times. Discussion supported wikis were edited in average almost 25 times per student. Was there any benefit from all these efforts?

Every individual post was graded soon after it was posted. Since plagiarised posts were awarded zero points, only the most stubborn students tried to repeat this fake in the discussions a number of times. At the end, they realised that the teacher was rigorous and always sanctioned plagiarism, so they finally gave up submitting another's works.

However, there were few occasions of internal plagiarism within different topics of same discussion forum or wiki, but it was rare and it was obvious who the original writer was. All the creation was very overt, and with time, fewer and fewer students committed both, external and internal plagiarism.

Individual essays, posts and wikis leading to joint products were not individually and instantly graded. Therefore, students felt more comfortable and some of them didn't obey basic rules. Luckily, all the moderators noticed copyright theft. Some moderators informed the teacher (Fig. 4.). They were suggested not to undertake any educational action, but to correct the submission if possible, or to ignore that submission while preparing the final version of the joint task.



-----Original Message-----
From:
Sent: Thursday, February 24, 2011 12:20 AM
To: Katerina Zdravkova
Subject: Dobro vece, Trenutno pohadjam Vas ...

Dobro vece,
Trenutno pohadjam Vas kurs Privatnost, etika i drustvena odgovornost, postavili ste me za moderatora grupe koja se bavi temom "Sta je plagijat u računarskim tehnologijama i kako se kompanije brane od ovog plagijata?". Posto svaki esej kolega prvo procitam i kazem im da li nesto da dodaju i poprave, naisla sam na esej koleginice koja je na nekim mestima kopirala ceo tekst i ne znam sta da radim, da li da joj kazem da sam to primetila, i da preprica svojim recima ili da ne kazem nista. Hvala.

--
This email is a copy of a message sent to you at "DMICS"
http://perun.pmf.uns.ac.rs/moodle/message/index.php?popup=1

Figure 4. Conscientious students reacted against plagiarism following the positive recommendation for whistleblowers

Few moderators took the opportunity to correct plagiarised parts not only when they prepared the final joint essay, but also within broader wiki. Some plagiarists reacted against this moderator's interference during tutorials, but when they found out that moderators were advised to do so, they self initiatively deleted or corrected the mistake.

It would be ideal to conclude that the implementation of Web 2.0 significantly reduced plagiarism. This conclusion is partially true and it actually affects those students who would play according to rules agreed at the beginning of the course even if the individual essays were the only way to assess their knowledge and participation in the course.

Student "solidarity", cohesion and mutual connection appeared to be more powerful than we could ever imagine. It seemed that there were few groups of two or several students with interconnected posts or wiki pages. Their writing styles looked alike a lot. Moreover, there was a great correlation between their access times and IP addresses. All these might

indicate mutual face-to-face collaboration, which is stimulating and welcome.

Unfortunately, it might also suggest that some students possessed the passwords of other colleagues and added posts and new pages in their own name and in the name of others.

Anonymous feedback of a disappointed student declaring that: "The forums were not a reliable way to assess student knowledge, because they were abused by some students. To be more precise, they (N.B. students) gave their passwords to acquaintances disabling other students to submit their posts." proved the unpleasant doubt (Fig. 5).

Even worse, false students were tireless and intrusive, living less space for real participants. Such students are in fact ghost-writers who employ deliberate identity fraud. We call this attitude an identity exchange.

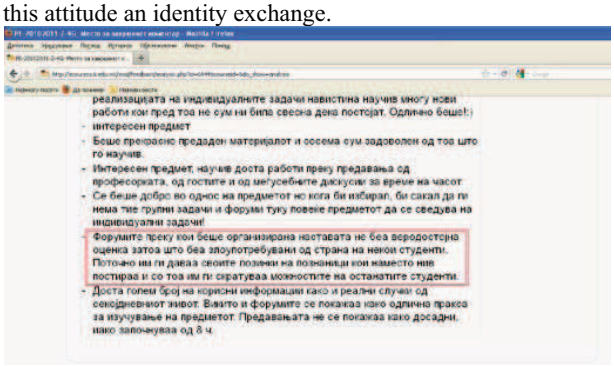


Figure 5. Student reaction confirming identity exchange

Students who were honest and ethical passed the course with high final grades. They were active during the whole course and eager to participate in the feedback before others. Their personal feeling was that they learned a lot and that the course positively influenced their future ethical and professional behaviour. Unfortunately, we didn't record the intermediate results, so the final feedback results (Table 1.) are not as favourable as they were in the beginning. It is important to emphasise that the topics students selected the most were those which were done using Web 2.0 tools. The least selected are exactly those which were delivered as individual essays.

The question dealing with the course itself showed that no student selected the options that the course was obsolete. Only 11.29% liked it partially. Another 14.52% found the course influential and necessary, while the majority of students (74.19%) approved the course and confirmed that it increased the awareness and that every IT professional must attend it as it is.

Table 1. Personal feeling of the most recent students about increased awareness of particular topics in the course

Topic	selected by	percentage
Ethics and ethical codex	31	0.5000
Privacy and its protection	49	0.7903
Information security	48	0.7742
Computer security	16	0.2581
Computer reliability	23	0.3710
Intellectual property	41	0.6613
Information access	25	0.4032
Professional behaviour	33	0.5323

No one complained about sanctioned plagiarism and cheating, which was very good news. Students sanctioned for identity exchange had an opportunity to get the points back during tutorials. When they "orally defended" the essay or posts, i.e. knew what others wrote in their name they got the full points. For those who were punished with zero points for plagiarism traditional oral exams was introduced again.

Very few students benefited from oral presentations and exams, but at the end, they learned at least what they faked.

IV. Conclusion and further ideas how to eliminate student scams

The initial question in this paper was: "Can Web 2.0 reduce plagiarism and cheating?". The answer is absolutely unknown (Table 2.).

Table 2. Comparison between cheats and scams

Cheats and scams	eLearning 1.0	eLearning 2.0
Plagiarism	Detectable (documents)	Detectable (online texts)
Ghost-writers	Suspected	Suspected
Identity exchange	Almost non existent	Year by year more frequent

To be exact, external plagiarism is easily detectable independently on the way of delivery. In eLearning 1.0, students prepare documents, while in eLearning 2.0 they prepare online texts.

In traditional eLearning 1.0, teacher needs to perform a comprehensive search for internal plagiarism. Sometimes naïve or insolent students facilitate the process by leaving document properties unchanged. In eLearning 2.0, the comparison has to be done over online texts.

Existent or new computerized tools can only facilitate the process of chasing plagiarism. In other courses it might not be crucial, but in Professional Ethics, it is an essential obligation.

The presence of ghost-writers remains the same. It is only suspected, but unfortunately, it can't be firmly proved. The only difference nowadays is that contemporary Web2.0 ghost-writers are no longer passive contributors and essay providers. They are omnipresent and they play the role of the students who hired them.

Unfortunately, in earlier periods, deliberate fraudsters were almost non existent. Identity exchange was not necessary because it was not important to submit and share the texts online. Collaboration was totally hidden and only the final product was delivered.

Web 2.0 initiated the concept of students of professional ghost-writers who undertake the role of the student. Students don't take care that the authentication systems is fully centralized, so those who do the identity exchange have access to everything they posses within faculty system.

This is a very bad news, but students don't care. They get what they want, because it is completely impossible to prove

the identity exchange. There are always reasonable arguments proving that the fraudsters never existed.

To finally summarise, Web 2.0 managed to reduce individual external and internal plagiarism. Ghost-writers remained present converting into intentional identity fraudsters. The worst scenario is always the most probable, because students always invent many scams teachers can not be even aware of.

For many students, the greatest intellectual challenge is to invent a new way to cheat instead of working honestly.

We have fixed one thing, and broken another. New ghost-writers and identity fraudsters spring like mushrooms after the rain.

By next year we should find a way to reduce the intentional identity fraud. Otherwise, the old good oral examination will become the only way to assess students. It's old fashioned, unpopular, tiring, and time consuming, but it always works.

V. REFERENCES

[1] University of California, San Francisco Library; *Online Journals*, http://www.library.ucsf.edu/collections/journals

[2] *IEEE eLearning Library*, http://www.ieee-elearning.org/

[3] *ACM Journals/Transactions*, http://portal.acm.org/pubs.cfm

[4] *IEEE Journals & Magazines + IEEE Journals, Transactions, and Letters,* http://www.ieee.org/publications_standards/publications/journmag/journals_magazines.html

[5] *Britannica Online Encyclopaedia*, http://www.britannica.com/

[6] *McGraw-Hills' Access Science Encyclopaedia of Science & Technology Online*, http://accessscience.com/

[7] *Wikipedia, the Free Encyclopaedia*, http://www.wikipedia.org

[8] *BitTorrent - Delivering the World's Content*, http://www.bittorrent.com/

[9] Dave, "Plagiarism - Definition, Causes & Consequences: Brief introduction to plagiarism, its causes, consequences & resolutions" Version 21. *Knol. 2009 Jun 8*, http://knol.google.com/k/dave/plagiarism-definition-causes/2pr18mcjtayt9/4

[10] Kakonen, T., Mozgovoy, M. (2010) Hermetic and Web Plagiarism Detections Systems for Student Essays - An Evaluation of the State-of-the-Art, *Journal of Educational Computing Research, Vol. 42, Number 2*/2010, pp. 135 - 159

[11] *Plagiarism Checker | Plagiarism Detection software from iThenticate*, http://www.ithenticate.com/

[12] Tackett. J., Claypool, G. A., Wolf F., Antenucci J. (2010) "The Impact of Plagiarism Detection Software on College Plagiarism", *Journal of Business and Accounting, Volume 3, Number 1*, pp. 69 - 80

[13] Attwood, R. "143 students expelled for plagiarism", *Times Higher Education*,http://www.timeshighereducation.co.uk/

[14] Pidd, H. "German defence minister resigns in PhD plagiarism row", *Guardian.co.uk*, http://www.guardian.co.uk/world/2011/mar/01/german-defence-minister-resigns-plagiarism

[15] "German MEP Silvana Koch-Mehrin quits in plagiarism row", *BBC News*, http://www.bbc.co.uk/news/world-europe-13369608

[16] Platt, S, "How Good Writers Can Make A Great Living", *Ghost-writer Dad*, http://ghost-writerdad.com/

[17] Medeiros, I., "Education in China: Ghostwriting hits frightening levels at universities", *Thoughts on Design, Technology and Culture*, http://designative.info/2010/03/26/education-in-china-ghostwriting-hits-frightening-levels-at-universities/

[18] Shen, Y. Yuan, Z. Liu, L. Dong, H., Research of anti-plagiarism monitoring system model, Wuhan University Journal of Natural Sciences, 2007, Vol 12; number 5, pp. 937-940

[19] Bloomfield, L.A,. "Software to detect plagiarism: WCopyfind", http://plagiarism.phys.virginia.edu/Wsoftware.htm