

Методический концепт экспертной оценки научно-исследовательских работ студентов в условиях информационного общества

Methodological concept of expert evaluation of research work of students in conditions of information society

Пунчик Вероника Николаевна¹

Punchyk Veranika

Пунчик Николай Николаевич²

Petrov Valerii

1. Кандидат педагогических наук, доцент, доцент кафедры педагогики
Белорусского государственного педагогического университета им. М. Танка
PhD in Pedagogical sciences, Associate Professor, associate professor of the Department of pedagogy of Belarusian State Pedagogical University named after M. Tank
e-mail: zelda@tut.by

2. Старший преподаватель кафедры информационных технологий Минского инновационного университета
Senior lecturer of the Department of information technologies of Minsk Innovation University
e-mail: PuNik@tut.by

Аннотация

Раскрываются основные аспекты проблемы плагиата, имманентные процессу организации современного научного исследования и его экспертизе в условиях информационного общества. Осуществлен краткий обзор программных систем выявления заимствований (программ контроля плагиата), выделены основные этапы их работы и используемые методы, а также определена их роль в процессе экспертизы научных работ. Обозначены направления предотвращения плагиата экспертами и научным сообществом. Подчеркнута роль правил цитирования и ссылок в научных исследованиях. Охарактеризованы понятия «белые» и «серые» источники, заимствования, цитирование, оригинальный текст.

Ключевые слова: плагиат, научное исследование, научная экспертиза, научное знание, информация, информационная среда, программные системы выявления заимствований (системы контроля плагиата), библиографическая ссылка.

Abstract

The article describes the main aspects of the problem of plagiarism inherent in the process of organizing modern scientific research and its expertise in the information society. The authors carried out a brief review of software systems to identify the borrowings (plagiarism control programs), the main stages of their work and the methods used are distinguished, and their role in the review process of scientific papers is identified. The article gives ways of preventing plagiarism by experts and the scientific community. The role of the rules of citation and references in research is emphasized. The concepts of «white» and «grey» sources of borrowing, quoting the original text are characterized.

Keywords: plagiarism, research, scientific activities, scientific knowledge, information, information environment, software system for identifying borrowings (plagiarism control programs), bibliographic reference.

Поступила в редакцию / Received: 16.03.2017

Web: <http://elibrary.miu.by/journals!/item.iot/issue.50/article.5.html>

Введение

В настоящее время учреждения высшего образования применяют разнообразные формы и методы привлечения студентов к проектно-исследовательской деятельности и научному творчеству, способствуя решению задач выявления и развития интеллектуального и творческого потенциала личности обучаемых. Для организации научной деятельности студентов, ориентированной на создание нового, принципиальное значение имеет опора на достоверное знание, представленное в научных источниках. В условиях информационного общества изменился механизм функционирования и распространения научного знания, поскольку информационные технологии в целом способствуют процессу распространения имеющегося знания, а не производству нового. Современные средства массовой информации породили особую среду Интернет-коммуникации, характеристикой которой является растворение границ текста как материального носителя информации. Феномены гипертекста и гиперпространства приводят к так называемой анонимности текста: он «как бы ничей», общий в виртуальном текстовом континууме. В такой ситуации обнаруживается отчетливая тенденция размытости, клиповости воспоминаний об источнике информации, и «значительно снижается тот порог торможения, который заслоняет от манипулирования текстами

и присвоения их фрагментов без указания на источник авторства» [1, с. 32]. Более того, собственно новые идеи, свежие мысли, порожденные творческой личностью, почти немедленно распространяются в информационной среде как анонимные самоочевидности, тем самым создавая благоприятные условия для плагиата, особенно медиально обусловленного технологизацией коммуникативного процесса.

В научном и педагогическом сообществах проблема выявления и оценки плагиата в условиях информационного общества не теряет своей актуальности, о чем свидетельствуют публикации Е.В. Беляевой, А.В. Никитова, А.С. Новикова, В.Е. Чернявской и др. Данная проблема имеет различные аспекты рассмотрения: нормативно-правовой, психологический, социальный, этический и др. В данной статье обратимся к методическому аспекту выявления плагиата в научно-исследовательских работах студентов, его качественной интерпретации в условиях информационного общества и определим ее основной концепт. Глубокое понимание данной проблемы невозможно без обращения к объяснению алгоритмов функционирования систем антиплагиата.

Методические аспекты применения систем антиплагиата

Как инструмент борьбы с плагиатом в учебных и научных исследованиях за последнее десятилетие разработан ряд программных продуктов, позволяющих определить заимствования в различных текстах, и в частности в научных работах. Эти программные продукты принято называть системами антиплагиата, или системами обнаружения и контроля плагиата.

Среди программных средств обнаружения плагиата можно назвать:

- система *Turnitin*, созданная американской компанией *iParadigms* – www.turnitin.com;
- программа *CopyCatch*, разработанная британской компанией *CFL Software* – www.cflsoftware.com;
- программа *Плагиат.НЕТ*, созданная российской компанией – www.mywebs.ru/plagiatanet.html;
- программа *Детектор плагиата*, разработанная в Московском университете печати – www.detector-plagiata.ru;
- система «*Антиплагиат*», созданная в 2005 г. российской компанией «*Forecsys*» – www.antiplagiat.ru;
- система *ПлагиатКонтроль*, разработанная в Белорусском государственном университете.

Отечественная система *ПлагиатКонтроль* обладает рядом преимуществ: в первую очередь это – работа в специфической белорусско-русской информационной среде одновременно с двумя языками: русским и белорусским, затем – она имеет настраиваемую структуру, что позволяет эффективно осуществлять поиск заимствований как в интернет-источниках, так и во внутренних базах текстов организаций с их постоянным пополнением.

Детали используемых различными программными продуктами алгоритмов проверки текста на плагиат их разработчиками не афишируются, но их теоретической основой являются методы информационного поиска – научного направления, базирующегося на методах информатики, библиотечного дела, лингвистики, семиотики, имеющего в качестве предметного поля принципы и методы поиска неструктурированной документальной информации. Исследования в области теории информационного поиска находят широкое применение в различных предметных областях, в первую очередь при организации поиска в сети Интернет, и, в частности, при создании систем обнаружения заимствований, с помощью которых должна быть решена нетривиальная задача определения плагиата в текстовых документах.

К типовым процедурам обнаружения заимствований относятся: *канонизация, стемминг и лемматизация, сравнение текстов* [2].

Канонизация (нормализация) текста – исключение некоторых символов и их последовательностей (знаки препинания, цифры, смайлы, скобки, кавычки, последовательности пробелов, управляющие символы и т. д.), а также слов, не несущих смысловую нагрузку (предлоги, союзы, междометия, стоп-слова).

Стемминг и лемматизация, позволяющие привести словоформы и производные формы слова к общей форме, что дает возможность распознавать замаскированный плагиат, например, изменение падежей, превращение одной части речи в другую.

Сравнение текстов. При этом арсенал применяемых методов очень широк. К примеру, метод шинглов и его модификации, методы сигнатур, индексирование, метод контрольных сумм (дактилограмм), метод опорных слов и др. При использовании метода шинглов текст разбивается на цепочки (шинглы) последовательно расположенных слов в нормализованном тексте. Количество слов в цепочке из слов называется «длиной шингла». Чем короче шингл, тем более точным будет анализ уникальности. Шинглы проверяемого текста сравниваются с шинглами текстов, размещенных в указанных источниках. Наличие одинаковых шинглов в текстах определяет степень их схоже-

сти. Метод сигнатур заключается в том, что каждому тексту ставится в соответствие некоторый уникальный набор характеристик, однозначно его характеризующий, своего рода подпись. При этом происходит свертка большого текста на небольшое количество его характерных признаков. Совпадение сигнатур означает о схожесть самих текстов. Применение метода увеличивает скорость обработки. Различные методы индексирования позволяют в ряде случаев обеспечить быстрый и точный поиск информации. Комбинированные схемы использования индексов позволяют выявить полное совпадение фраз, в которых используемые слова встречаются часто, но сама фраза – относительно редко, что служит предпосылкой обнаружения объемного совпадения сравниваемых текстов. Метод k-граммных индексов позволяет найти сознательно сделанные замены символов одного алфавита на символы другого. Скрытое семантическое (или латентно-семантическое) индексирование позволяет создать новое сжатое представление каждого документа в базе данных и на его основе быстро вычислить сходство между документами. Причем в ряде случаев этот метод дает возможность разрешить проблему обнаружения синонимии. В некоторых случаях используются классические методы информационного поиска, основанные на методах взвешивания терминов, такие как: TF (основан на мере отношения числа вхождения некоторого слова к общему количеству слов документа), TF*IDF (в нем используется мера IDF – обратная частота документа; эта мера больше у тех слов, которые часто использовались в одном документе и реже – в другом) и их разнообразные модификации.

В системах антиплагиата введена поддержка так называемых «белых» коллекций, включающих, например, нормативные и правовые документы. Сравнение может проводиться не только с текстами в свободном доступе в Интернет, но и со специальными модулями, например «Коллекция диссертаций РГБ», «Коллекция eLibrary» и др. Поиск по специальным модулям системы контроля плагиата обычно доступен на платной основе.

После проверки система контроля плагиата предоставляет отчет, в котором в наглядной форме представляются его результаты: в проверяемом цветом тексте отмечены оригинальные и неоригинальные фрагменты, указывается процент оригинальности, процент совпадения с каждым из источников заимствования.

Современные версии систем контроля плагиата позволяют обнаружить ухищрения, которые ранее помогали недобросовестным студентам и исследователям скрыть заимствование, например, замена символов одного алфавита аналогичными по написанию символа-

ми другого алфавита; изменение формата текста; применение псевдографики и скрытого текста; перестановка местами фрагментов текста; перестановка слов в предложении; манипуляции со стоп-словами и т. п.

Очень важным аспектом применения систем контроля плагиата выступает понимание педагогическим сообществом того факта, что применение подобных систем не решает проблему плагиата [3], однако может автоматизировать часть работы по качественной оценке работы обучающегося. Сам по себе отчет системы контроля плагиата является формализованным и обязательна его дальнейшая интерпретация: только эксперт (педагог, осуществляющий проверку работы) может дать оценку имеющимся заимствованиям, классифицируя их как корректное цитирование, авторскую интерпретацию, творческое развитие или плагиат [4].

Структура отчета позволяет выделять в проверяемом тексте заимствованные части, выдает все ссылки на Интернет-ресурсы, с которых они были «скачаны». Эксперту необходимо сопоставить объем заимствований с объемом текста, проверить наличие и корректность ссылок, чтобы выяснить, присутствует ли в работе плагиат. Отчет систем контроля плагиата показывает заимствования из чужих текстов, но не идентифицирует их именно как плагиат.

В системах антиплагиата присутствует поддержка так называемых «белых» коллекций, включающих достоверные источники цитирования. К ним в первую очередь относятся: нормативные и правовые документы, монографии, диссертационные исследования и авторефераты, научные статьи, данные энциклопедий, хрестоматий и атласов. Цитаты из таких коллекций и документов выделяются в тексте отчета особым образом, что помогает преподавателю корректно определить характер заимствования: различить цитирование и заимствование. Но при этом итоговый ответ на вопрос, является ли тот или иной фрагмент текста заимствования корректным или плагиатом, может дать только эксперт, владеющий спецификой письменного изложения научного материала в своей предметной области.

Цитирование – дословная выдержка из какого-либо текста «белой» коллекции. Цитирование является корректной формой заимствования, при этом цитируемый текст должен быть оформлен в кавычки («прямая цитата») с указанием автора цитаты, названия произведения и страницы цитирования либо корректной ссылкой на электронный источник также с указанием автора и произведения. Возможна также интерпретация чужих мыслей с указанием прямого авторства (например, «Развивая идеи автора о...», «Соглашаясь с высказыванием автора о...»). При условии оформле-

ния границ цитируемой идеи и библиографической ссылки цитирование не является плагиатом. Если цитата не оформлена соответствующим образом, она является некорректным заимствованием и интерпретируется как плагиат.

Заимствование – фрагмент или дословная выдержка из какого-либо текста без оформления цитаты и библиографической ссылки. Заимствование в таком случае интерпретируется как плагиат. Заимствование допустимо в отношении таких атрибутов, как формулы, титульные листы, шапки таблиц, стандартные и вводные словосочетания и т. д., объемы, уместность и корректность использования которых в работах может оценить только эксперт.

Оригинальный текст – текст, с которым система контроля плагиата не нашла совпадений. Важно отметить, что система контроля плагиата может обнаружить совпадение только с теми источниками, которые

представлены в электронном виде и доступны ей для сравнения. Однако существуют источники, которые представлены только в бумажном виде и не представлены в Интернет (например, архивные материалы, раритетные издания, печатная продукция, не подвергнутая оцифровке). В таком случае система контроля плагиата расценит их как оригинальные, хотя таковыми они являться не будут – и тогда только эксперт сможет оценить качество работы.

В России практически все вузы нормативно обязывают проверять учебные и научные работы на наличие плагиата с помощью различных программных средств, и зачастую работы, нарушающие принятые нормативы, не допускаются к дальнейшему рассмотрению. Подобная тенденция наблюдается и в Республике Беларусь. Некоторые усредненные требования к уникальности учебных и научных работ представлены в таблице 1 [5].

Таблица 1 – Нормативы уникальности текста учебных и научных работ в процентах

Дисциплины	Контрольная работа	Реферат	Отчет по практике	Курсовая работа	Дипломная работа	Эссе
Социально-гуманитарные	30 и выше	40 и выше	50 и выше	50 и выше	60 и выше	80 и выше
Естественнонаучные	40 и выше	40 и выше	50 и выше	60 и выше	70 и выше	70 и выше
Технические	-	30 и выше	50 и выше	60 и выше	70 и выше	70 и выше

В связи с официальным введением систем контроля плагиата появились ее сторонники и противники. Если учитывать доводы обеих сторон и принимать во внимание мнение противников программы, общий анализ ситуации, тем не менее, позволяет сделать вывод, что к сторонникам относятся преимущественно те научные работники и ученые, которые не используют напрямую ресурсы Интернет в качестве источника своих работ, а к противникам – преимущественно те, кто использует. При этом на проверку часто оказывается, что те источники, которые были использованы как печатные, представлены в Интернет, и их применение интерпретируется программой как заимствование. В целом это только подтверждает тот факт, что разделение источников на электронные и бумажные в условиях информационного общества является относительным, а существенным, по нашему мнению, должен выступать признак, является ли источник «белым» для данного вида исследования.

Заключение

Анализ состояния проблемы в образовательной практике, руководство научными и учебными исследова-

ниями различного уровня, аппроксимация этических норм науки к условиям постмодернизма и информационного общества, а также собственный опыт применения систем контроля плагиата позволили авторам данной статьи сформулировать следующие рекомендации преподавателям, выступающим в роли экспертов, которые определяют границы применения компьютерных систем в оценке качества научно-исследовательских работ. Эти рекомендации выступают методическим концептом экспертной оценки научно-исследовательских работ студентов в условиях информационного общества и отражают основные пункты дорожной карты экспертизы, на которые должен ориентироваться эксперт при использовании систем антиплагиата.

- К «белым» источникам, помимо прямых и косвенных цитат со ссылками, эксперту целесообразно также отнести: интерпретацию общеизвестных фактов, самоцитирование (не более 25 % от работы).
- К «серым» (недостоверным) источникам, использование которых не допускается в исследовательской работе, отнести: использование данных из Интернет-баз рефератов, курсовых, дипломов

и других банков готовых работ, т. е. плагиата чужих выводов, аргументов, мыслей и идей. Недопустимо заимствование в области выводов, итогов, результатов, полного перечня использованных источников. Статистическое совпадение с «серыми» источниками должно составлять не более 1–2 %. Важно понимать, что система контроля плагиата находит источник заимствования текста, но не может определить, является ли он первоисточником.

- Необходимо выявлять некорректное копирование цитат авторов не из первоисточников: «Как отмечает автор А, автор Б писал...», – в подобном случае целесообразно изучить первоисточник Б и приводить цитату из него. В отчетах системы контроля плагиата эксперт может определить это, если как заимствование выделена не только цитата, а все предложение.
- Естественным является факт того, что даже если работа написана с четким указанием авторства, но состоит сплошь из чужих цитат, даже с прямым указанием их авторства, т. е. является компилированием, отнести ее к исследованиям высокого качества нельзя.
- В исследовательской работе не должно содержаться более 20 % заимствования (цитирования) из одного источника (в том числе «белого»). Присутствие большего процента заимствования снижает научную ценность исследования.

В целом считаем, что жесткие нормативные требования к учебным работам без интерпретации экспертом не являются обоснованными. Безусловно, недопустимо, чтобы текст курсовой работы имел многократные совпадения с текстом других курсовых работ, но если при заданных структуре и списке авторитетных источников уникальность текста составляет, например, 40 %, то нецелесообразно без анализа работы экспертом исключать работу из рассмотрения.

Таким образом, можно заключить, что в современном информационном обществе проверка научно-исследовательских работ студентов на плагиат является неотъемлемым этапом экспертизы. Однако следует

учитывать, что системы обнаружения заимствований, как и все применяемые информационные технологии в целом, являются вспомогательным инструментом интеллектуальной деятельности и должны дополнять экспертную оценку, а не заменять ее, и тем более не становиться инструментом манипулирования.

Литература / References

1. Чернявская, В.Е. Плагиат как социокультурный феномен / В.Е. Чернявская // Известия Санкт-Петербургского университета экономики и финансов. – 2011. – № 3. – С. 26–32.
Chernyavskaya, V.Ye. Plagiat kak sotsiokul'turnyy fenomen / V.Ye. Chernyavskaya // Izvestiya Sankt-Peterburgskogo universiteta ekonomiki i finansov. – 2011. – No. 3. – P. 26–32.
2. Пунчик, В.Н. Поликонтекстный анализ феномена «плагиат» в информационном обществе / В.Н. Пунчик, З.В. Пунчик // Социология. – 2016. – № 1. – С. 83–92.
Punchik, V.N. Polikontekstnyy analiz fenomena «plagiat» v informatsionnom obshchestve / V.N. Punchik, Z.V. Punchik // Sotsiologiya. – 2016. – No. 1. – P. 83–92.
3. Берков, В.Ф. Признание истин в науке / В.Ф. Берков // Проблемы управления. – 2007. – № 2. – С. 157–161.
Berkov, V.F. Priznaniye istin v nauke / V.F. Berkov // Problemy upravleniya. – 2007. – No. 2. – P. 157–161.
4. Никитов, А.В. Плагиат в работах студентов и аспирантов: проблема и методы противодействия / А.В. Никитов, О.А. Орчаков, Ю.В. Чехович // Университетское управление: практика и анализ. – 2012. – № 5. – С. 61–68.
Nikitov, A.V. Plagiat v rabotakh studentov i aspirantov: problema i metody protivodeystviya / A.V. Nikitov, O.A. Orchakov, Yu.V. Chekhovich // Universitetskoye upravleniye: praktika i analiz. – 2012. – No. 5. – P. 61–68.
5. Пунчик, В.Н. Методические аспекты применения системы антиплагиата в оценке качества учебных и научных работ / В.Н. Пунчик, Н.Н. Пунчик // Управление в социальных и экономических системах: материалы XXIV Международ. науч.-практ. конф., Минск, 14 мая 2015 г./ редкол.: Н.В. Суша (предс.) и др.; Минский ун-т управления. – Минск: МУУ, 2015. – С. 174–176.
Punchik, V.N. Metodicheskiye aspekty primeneniya sistemy antiplagiata v otsenke kachestva uchebnykh i nauchnykh rabot / V.N. Punchik, N.N. Punchik // Upravleniye v sotsial'nykh i ekonomicheskikh sistemakh: materialy XXIV Mezhdunar. nauch.-prakt. konf., Minsk, 14 maya 2015 g. / redkol.: N.V. Susha (preds.) i dr.; Minskiy un-t upravleniya. – Minsk: MUU, 2015. – P. 174–176.