



АВДЕЕВА Нина Владимировна -
начальник управления
функционирования и мониторинга
клиентского сервиса
Российской государственной
библиотеки
Адрес: 119019, г. Москва,
ул. Воздвиженка, 3/5
e-mail: Avdeeva@rsl.ru



СУСЬ Ирина Валерьевна -
кандидат технических наук,
доцент, ведущий специалист
сектора исследования документов
отдела перспективного развития
Российской государственной
библиотеки
Адрес: 119019, г. Москва,
ул. Воздвиженка, 3/5

УДК 004.912

РОЛЬ ЭКСПЕРТА В ОЦЕНКЕ КАЧЕСТВА НАУЧНЫХ ДОКУМЕНТОВ С ПОМОЩЬЮ ПРОГРАММНЫХ СИСТЕМ

Научная деятельность ученых подразумевает публичность и преемственность. Все исследования обсуждаются еще на этапе постановки эксперимента и подвергаются анализу в процессе их проведения. Результаты выносятся на суд научного сообщества в виде публикаций в научных отраслевых журналах и квалификационных работ, обсуждаются на ученых советах и конференциях. При этом с самого своего начала научная деятельность человека проходит под руководством более опытного наставника. Становлению молодых ученых помогают руководители научных, дипломных и диссертационных работ, консультантами более зрелых ученых выступают маститые в этой области знаньи коллеги. Основной показатель научной зрелости ученого - способность его к научному творчеству, т.е. выражению результатов своего труда в оригинальных, самостоятельно написанных научных текстах (публикациях, монографиях, отчетах НИР, диссертациях, методиках и т.д.).

Создание оригинального научного текста требует наличия у автора соответствующих навыков и компетенций, а также умения применять их на практике. Их отсутствие компенсируется контрафактом авторского права за деньги или путем подлога.

К сожалению, в последние десятилетия российское академическое сообщество все чаще сталкивается с нарушением этических норм и взаимной требовательности. «Сейчас, к примеру, даже отсутствует серьезное осуждение тех лиц, кто явно за плату заказывает диссертации, а затем, получив ученую степень, даже начинает свысока глядеть на окружающих» [3].

В практике всех зарубежных университетов существует понятие «Кодекса чести», основные положения которого отражены в уставах. «Личное отношение студентов к плагиату, дезинформации, фальсификации и другим формам ака-

демического мошенничества более чем пренебрежительное» [4].

Значимым моментом в борьбе с плагиатом в академической и образовательной среде является открытость результатов исследований и их публичное обсуждение. Этому способствует положение Постановления Правительства Российской Федерации от 24 сентября 2013 г. № 842 «О порядке присуждения ученых степеней» о размещении на официальном сайте организации, на базе которой создан диссертационный совет, в сети Интернет, полного текста докторской диссертации за три месяца до и девять месяцев после защиты и, соответственно, сроки два и семь месяцев определяются для кандидатской диссертации [6].

В последнее десятилетие наука развивалась стремительно. Стало сложно уследить за всеми исследованиями, проводимыми не только в мире, но и в разных научных школах нашей страны. В помощь ученым пришел интернет. Все институты имеют сайты с информацией о направлениях деятельности и полученных достижениях. Научные журналы размещают на интернет-страницах анонсы статей. Образовались интернет-сообщества ученых, на площадках которых они делятся своим опытом и дискутируют. Появились электронные библиотеки, в том числе Электронная библиотека диссертаций Российской государственной библиотеки, с удобными поисковыми системами.

Однако публичность науки позволяет недобросовестным авторам присваивать чужие мысли. Причины плагиата хорошо раскрыты в статье Р.У. Арифудиной с соавторами [2]:

1) технологический аспект - высокий уровень развития современной компьютерной техники и повышение уровня информационно-технологической грамотности научных исследователей;

2) социокультурный аспект -

требование от исследователей все новых и новых достижений в сжатые сроки и в условиях ограниченности ресурсной базы, борьба любой ценой за получение исследовательских грантов, необходимость получения «валового» образовательного продукта с потерей качества научных работ;

3) личностно-психологический аспект - отсутствие моральных устоев человека, недостаточное развитие общеучебных умений.

Кроме общественного порицания и публичности (открытости) науки основным средством борьбы с использованием и присвоением чужих текстов, идей и решений (плагиатом) должны быть неотвратимость разоблачения и прозрачность системы контроля.

Безнаказанность порождает дальнейшие хищения интеллектуальной собственности. Противостоят подмене новых научных достижений компиляциями из более ранних текстов должен институт экспертов. В настоящее время в России не существует системы экспертной деятельности по авторской ответственности, отсутствует орган, реализующий данную политику, и не разработаны механизмы функционирования данной системы.

На сегодняшний день проверять оригинальность дипломных и диссертационных работ по указанию Министерства образования и науки Российской Федерации обязаны образовательные учреждения и диссертационные советы. Так как тема оценки оригинальности текстов документов новая для отечественной науки, существует ряд нерешенных вопросов: квалификация эксперта, критерии оригинальности текста, база данных, правила проверки и др. А вот задача использования систем компьютерного анализа и сравнения текста решается успешно.

Для оценки оригинальности текстов предоставляемых к защите диссертаций в настоящее время существует достаточно большое количество сервисов и программ, в

том числе на русском языке. Наиболее распространенной в научной среде является система «Антиплагиат», включающая обширную базу данных и точно определяющая текстовые совпадения. Пакет услуг «Антиплагиат.ВУЗ» востребован многими учебными заведениями. «Эвристичность этой системы («Антиплагиат.ВУЗ») состоит в том, что она осуществляет онлайн-поиск по большому количеству документов, хранящихся в коллекциях Российской государственной библиотеки. Важно отметить, что система находит источник заимствования, но не определяет, является ли он первоисточником. Ответ на вопрос, является ли заимствованный фрагмент именно плагиатом, а не законной цитатой, система оставляет на усмотрение экспертов» [5]. Поэтому подготовка специалистов семантического анализа научных текстов с помощью автоматизированных систем выходит на первый план.

Российская государственная библиотека обладает уникальным опытом по использованию специализированного программного обеспечения для проверок научных текстов на наличие некорректных заимствований. С 2009 года штатом экспертов РГБ было выполнено 3 498 таких проверок. Проверка научных документов на предмет заимствований с использованием автоматизированной системы специализированной обработки текстовых документов «Антиплагиат.РГБ» - единственная официальная независимая объективная экспертиза, проводимая квалифицированными специалистами по разработанным методикам.

Автоматизированная система «Антиплагиат» - это инструмент эксперта для обнаружения некорректных заимствований в текстовых документах. Данная система не обнаруживает плагиат, а всего лишь указывает на наличие совпадений. Определить характер совпадений и делать выводы должен эксперт.

В основе проверки лежит постулат, что использование любой информации (текста, количественных данных, иллюстраций) при написании документа (диссертации, монографии, статьи и других научных трудов) возможно лишь при соблюдении правил научного цитирования, важнейшими из которых являются: минимальный объем и необходимость цитаты, достоверность ссылки и ее корректное оформление.

Под корректным цитированием мы понимаем использование текстовой, табличной или иллюстрационной информации, в явном или неявном виде включенной в логическую цепь рассуждений автора документа, снабженной ссылкой на источник заимствования. Ссылка должна содержать всю необходимую информацию об источнике, быть указана в тексте (в виде краткой ссылки) и в списке литературы (в виде библиографического описания).

Правило цитирования текста позволяет воспроизводить небольшие цитаты по вторичному источнику (с обязательной ссылкой на него) при невозможности обращения к первоисточнику.

Заимствование из документов обширных фрагментов текста вместе с имеющимися в нем ссылками на первоисточники не соответствует правилам научного цитирования независимо от наличия ссылки на «вторичный источник». Очевидно, что компиляция многостраничных отрывков чужих текстов, не подвергнутых критическому осмыслению и анализу с учетом задач, поставленных автором, заимствование чужого текста вместе со ссылками, оценочными суждениями и комментариями, нарушает ключевое требование к любой научной работе, в том числе к диссертационной: «диссертация должна быть написана автором самостоятельно... и свидетельствовать о личном вкладе автора диссертации в науку» [6, п.10].

Соотношение объемов ори-

гинального авторского текста и дословно воспроизведенных цитат может быть различным в разных областях наук, поскольку определяется научными традициями и задачами конкретного исследования. Например, в разных диссертациях могут быть воспроизведены дословно выдержки из одного и того же законодательного документа, и такие фрагменты текста являются «корректными цитатами». Однако текст законодательного документа, выданный за авторский текст, то есть использованный в качестве аналитического обзора проблемного поля научного исследования объемом в несколько страниц, даже со ссылками на этот документ, корректным заимствованием не является. «Обзор литературы» также не может состоять преимущественно из дословно воспроизведенных цитат, тем более, если они не взяты в кавычки и не указаны источник и страницы, на которых расположен текст цитаты.

При проверке документа оценки корректности совпадения эксперт устанавливает в каждом конкретном случае в зависимости от вида (текст, рисунок, таблица, схема), сущности, важности, контекста совпадающих блоков, аутентификации цитат, наличия и достоверности ссылок, срока защиты работы-экспортера, связи авторов источников с автором проверяемой работы и т.д.

К корректным совпадениям могут быть отнесены: блоки, содержащие названия конференций, нормативно-правовых актов, ГОСТов; небольшого объема отрывки, содержащие терминологию, стандартные фразы и устойчивые выражения, принятые в научной области проверяемого документа; случайные и неинформативные совпадения; основные сведения о подготовленной к защите диссертации; библиографические описания первоисточников в списке литературы. Корректными заимствованиями являются цитаты -

дословно или косвенно воспроизведенные выдержки из какого-либо источника со ссылками на источники заимствования. При этом важно, чтобы чужой текст, использованный в тексте как цитата, можно было идентифицировать (с помощью кавычек, ссылок, упоминания автора высказывания вблизи заимствованного фрагмента).

Некорректные заимствования - это не имеющие ссылок фрагменты проверенного экспертом документа, дословно совпадающие с отрывками доступных эксперту документов из коллекции или же фрагменты, снабженные неправильными ссылками. Некорректными заимствованиями являются перефразированные относительно первоисточника или вторичного источника отрывки текста, не сопровождаемые правильными ссылками.

В случае обнаружения некорректных заимствований эксперт определяет их характер и значимость для того раздела работы, в котором они выявлены, а также для всего исследования в целом.

Оценку оригинальности работы эксперт осуществляет по значимости для проверяемого документа некорректных заимствований: он определяет их характер, местоположение в документе, объем заимствования. Например, заимствования а priori не может быть в некоторых разделах квалификационного документа - формулировке новизны, тезисах на защиту, гипотезе, выводах и т.д.

«Любая автоматизированная система проверки текстовых документов является лишь инструментом, посредником между экспертом и автором, а сама проверка - трудоемким интеллектуальным процессом с индивидуальным подходом в каждом случае» [7].

Исходя из вышесказанного, становится понятно, как велика роль эксперта при оценке оригинальности научных текстовых документов. Нельзя делать вывод о работе только по проценту оригинальности, выданному автоматизирован-

ной системой, т.к. она не определяет легитимность совпадений, не учитывает специфики текста. Также не представляется возможным выработать единый критерий оценки результатов проверки текстов - в рамках каждой специальности и категории работ должны быть свои правила - и только опытный специалист может принять верное решение. Ключевым моментом в принятии решения является факт присвоения авторства и умышленность свершаемых действий.

Для независимой оценки научных документов необходимо привлекать экспертов, обладающих не только знаниями в определенной отрасли, но и владеющих методологией и навыками автоматизированной проверки текстов, способных к латентно-семантическому анализу.

Эксперты становятся обязательным аналитическим звеном цепочки создания научных документов. Поэтому необходимо расширять круг анализируемых документов - не только дипломы и диссертации, но и статьи, монографии, отчеты НИР и др. Это позволит культивировать правила научного цитирования и на ранних этапах «очищать» научные круги от лжеученых.

В последнее время эксперты ассоциируются только с проверкой текстовых документов на плагиат, но этим их деятельность не ограничивается. В первую очередь эти специалисты обучают авторов правилам научной этики и оформления научных трудов. Эксперты, оценивающие научные работы, выполняемые по государственному заказу, могут проанализировать эффективность использования финансов, снизить дублирование работ.

Важная роль экспертов состоит в совершенствовании и обучении появившихся недавно интеллектуальных автоматизированных систем, верификации результатов, выдаваемых этими системами. В качестве примера можно привести разрабатываемую в рамках государственного задания

Министерства образования и науки Российской Федерации «Информационную систему формирования и структуризации единой цифровой коллекции первоисточников научных работ удостоверенного качества с обеспечением планового пополнения современными научными произведениями и трудами для создания условий публичного доступа к коллекции», названную «Научный архив».

Развитие информационных технологий, возникновение многообразных коллекций электронных документов сделали актуальным вопрос о необходимости быстрой оценки качества трудов и формирования «чистых» электронных фондов. Помочь специалистам в экспертизе неограниченного объема постоянно создаваемых и публикуемых научных трудов призвана интеллектуальная система «Научный архив».

«Документы, успешно прошедшие стилистический анализ, система проверяет на оригинальность, то есть на отсутствие некорректных заимствований из более ранних по времени создания

источников доступных коллекций. Путем сопоставления с ними она также определяет объем текста, не совпадающего ни с какими другими документами, то есть уникального. Кроме того, система подсчитывает долю так называемого «цитатного» текста (использованных в данном документе цитат из более ранних публикаций). Обнаружение случаев, когда большую часть научной работы составляют корректно оформленные цитаты из чужих научных трудов, поможет отделить компилятивные документы от остальных текстов. Наконец, система может определить и объем таких фрагментов проверяемого текста, которые сами оказались востребованными в более поздних научных работах. При этом под «востребованностью» понимается не только корректное использование в виде цитат или пересказа со ссылкой на проверяемый документ, но и некорректное, то есть без всяких обозначений происхождения заимствованных фрагментов.

По результатам автоматизированной проверки документов в соответствии с указанными критери-

ями (оригинальности, уникальности, цитатности и востребованности) система ранжирует и сортирует документы, создавая коллекцию документов, представляющих собой уникальные авторские тексты, отвечающие всем требованиям научного стиля и содержащие новые ценные знания» [1].

Однако при разработке и апробации интеллектуальных систем теоретически возможно наделение научных документов ошибочными рангами. Поэтому эксперты РГБ, обладающие навыком аналогичных проверок и привлеченные к проекту, проводили обучение системы, обнаруживали погрешности в ее работе и осуществляли корректировку результатов ранжирования.

Проведенный анализ значимости экспертной проверки свидетельствует о необходимости создания института независимой экспертизы качества научных документов, унификации основных принципов его деятельности и широкого вовлечения в оценку научных трудов, начиная с дипломных работ и статей в журналы, заканчивая диссертациями и монографиями.

Список литературы:

1. Авдеева Н.В., Ледовская В.М., Лобанова Г.А. Методика подробной экспертной проверки научных трудов низкого ранга как способ верификации результатов их автоматизированного ранжирования по качеству // Информационные ресурсы России. - 2016. - № 4. - С. 2-5

2. Ариффулина Р.У., Карпова Ю.Н., Шесслер О.В., Суханова М.Э., Яременко И.Ю. К проблеме плагиата в сфере науки и образования // Современные наукоемкие технологии. - 2016. - № 2. - С. 55-59.

3. Бабаев Б.Д., Чекмарев В.В. Об институте экспертов ВАК и другом: наши соображения // Экономика образования. - 2011. - № 3. - С. 158-161.

4. Герцен С.М. Меры предотвращения академического мошенничества (из опыта зарубежных вузов) // Интернет-журнал Науковедение. - 2013. - № 4 (17) - С. 28 [Электронный ресурс] / Научная библиотека КиберЛенинка. - Режим доступа: <http://cyberleninka.ru/article/n/меры-predotvrascheniya-akademicheskogo-moshennichesk-tva-iz-opyta-zarubezhnyh-vuzov#ixzz4ALXpD6IK>.

5. Глебов А.А. Анализ теоретических источников в диссертационном тексте // Известия ВГПУ. - 2014. - № 4 (89). - С. 11-14 [Электронный ресурс] / Научная библиотека КиберЛенинка. - Режим доступа: <http://cyberleninka.ru/article/n/analiz-teoreticheskikh-istochnikov-v-dissertatsionnom-tekste#ixzz4ALXQH3eH>.

6. Положение о присуждении ученых степеней. Утверждено Постановлением Правительства Российской Федерации от 24 сентября 2013 г. № 842.

7. Скаковская Л.Н. Использование информационных систем при оценке качества квалификационных работ / Скаковская Л.Н., Медведева О.Н., Мидоренко Д.А., Потехина А.В., Жданова О.В., Супонев Н.П. // Высшее образование в России. - 2015. - № 5. - С. 110-114 [Электронный ресурс] / Научная библиотека КиберЛенинка. - Режим доступа: <http://cyberleninka.ru/article/n/ispolzovanie-informatsionnyh-sistem-pri-otsenke-kvalifikatsionnyh-robot#ixzz4AKra6IHq>.