

«АНТИПЛАГИАТ.РГБ»: найти и обезвредить

Авторы



**Нина Владимировна
АВДЕЕВА**
начальник управления
функционирования
и мониторинга
клиентского
сервиса РГБ



**Ольга Владимировна
НИКУЛИНА**
начальник отдела
поддержки доступа
к электронным
ресурсам РГБ



**Андрей
Владимирович
САЗАНОВ**
эксперт РГБ, доктор
исторических наук,
профессор МГУ
им. М.В. Ломоносова,
Московского
института лингвистики,
эксперт и научный
консультант Первого
канала Российского
телевидения

По данным на 1 сентября 2012 г., в Электронной библиотеке диссертаций Российской государственной библиотеки (ЭБД РГБ) содержится 725 810 полных текстов диссертаций и авторефератов, что делает коллекцию уникальным по полноте и значимости ресурсом для науки и образования.

Российская государственная библиотека стремится повысить качество российского образования и науки, показывая авторам необходимость самостоятельного написания научных работ, исключая создание текстов путём компиляции найденных в различных источниках материалов по заданной тематике.

Этой цели служит проведение проверки текстов на предмет наличия некорректных заимствований (плагиата) по ЭБД РГБ с использованием системы «Антиплагиат.РГБ» и выдачей официального заключения эксперта.

Для обеспечения широкой доступности и сохранности диссертационного фонда в 2003 г. руководством РГБ было принято решение о создании на основе современных информационных технологий ЭБД РГБ. Формирование Электронной библиотеки диссертаций РГБ проводилось в три этапа. На первом был сформирован стартовый пакет примерно из 28 тыс. полных текстов диссертаций по наиболее востребованным наукам: экономике, юриспруденции, педагогике, психологии и философии. Второй этап начался в 2004 г., когда ЭБД РГБ стала пополняться диссертациями по всем специальностям, исключая медицину и фармацию, что, в общей сложности, составляло около 30 тыс. единиц в год, в том числе 20 тыс. кандидатских и 10 тыс. докторских диссертаций. Дополнительно в 2006 г. в рамках проекта ретроконверсии были оцифрованы все диссертации за 1985 г. Наконец, на третьем этапе, с 2007 г. в ЭБД РГБ поступают диссертации по всем дисциплинам, включая работы по медицине и фармации [4]. В результате на 1 сентября 2012 г. в ЭБД РГБ содержится 357 221 диссертация и 368 589 авторефератов.

Одной из наиболее востребованных услуг, направленных на получение каждым читателем точной и актуальной научной информации, является проведение проверки текстов на

предмет наличия некорректных заимствований (плагиата) по ЭБД РГБ с использованием системы «Антиплагиат.РГБ» и выдачей официального заключения эксперта [5].

Хотя точного определения плагиата в современном российском законодательстве нет, под ним обычно понимают трактовку, изложенную в Информационном письме Генеральной прокуратуры РФ от 30.03.2001 № 36-15-01 «О практике применения законодательства по защите интеллектуальной собственности, состоянии прокурорского надзора и мерах по усилению борьбы с пиратством в аудиовизуальной сфере» и п. 3 Постановления Пленума Верховного Суда РФ от 26.04.2007 № 14.

Согласно первому документу, «присвоение авторства (плагиат) **выражается** в том, что лицо провозглашает себя автором чужого произведения, выпуская его полностью или частично под своим именем (псевдонимом). Присвоением авторства признаётся выпуск произведения, созданного совместно с другими авторами, без указания фамилий соавторов, использование в своих трудах чужого произведения без ссылки на автора».

Во втором документе «присвоение авторства (ч. 1 ст. 146 УК РФ) **может состоять** в объявлении себя автором чужого произведения, выпуске чужого произведения (в полном объ-

ёме или частично) под своим именем, издании под своим именем произведения, созданного в соавторстве с другими лицами без указания их имён» [3].

Не подлежит сомнению, что соблюдение авторских прав и корректное использование материалов из различных источников являются важными условиями работы любого серьёзного и грамотного специалиста. Вместе с тем современные информационно-коммуникационные технологии значительно облегчают как чтение, так и заимствование фрагментов чужих работ, чем активно пользуются недобросовестные авторы. Широкое распространение плагиата в учебных заведениях и, отчасти, научных учреждениях неоспоримо.

В настоящее время услуга РГБ по выявлению плагиата доступна для юридических и физических лиц. Юридические лица предоставляют тексты на проверку на основании заключённого с РГБ договора (проект договора на проведение определённого количества проверок и подробная информация об услуге представлены на сайте <http://diss.rsl.ru> в разделе «Антиплагиат»). Физические лица, оформив заказ в интернет-магазине РГБ на сайте <http://store.rsl.ru>, могут прислать на проверку только произведения или статьи, авторами которых они являются.

Система «Антиплагиат.РГБ», с помощью которой производится проверка присланных текстов, — это комплекс программных средств для быстрого поиска некорректных заимствований в электронных документах по коллекциям источников большого объёма, которой и является ЭБД РГБ. Функциональное ядро сервиса использует уникальные алгоритмы, которые позволяют быстро и эффективно проводить поиск заимствованных фрагментов и гарантируют корректную обработку русскоязычных текстов.

Система имеет надёжную защиту от простых средств «обхода» проверки и распознаёт:

- замены русских букв латинскими;
- изменения формы слов;
- замены слов синонимами;
- изменения порядка слов;
- перестановки страниц, абзацев, предложений.

Система «Антиплагиат.РГБ» используется только для поддержки экспертизы текстовых документов на наличие заимствований и является надёжной технической составляющей процесса проверки. Основную же аналитическую деятельность выполняет эксперт — специалист, обладающий учёной степенью кандидата или доктора наук и имеющий опыт редактирования и рецензирования научных статей, трудов и отчётов. Эксперт не только подготавливает текстовый документ для его программной проверки, но и обрабатывает отчёт, сформированный системой «Антиплагиат.РГБ». Необходимо подчеркнуть, что на эксперта возлагается самая главная задача: провести смысловой анализ выданных системой результатов и сформировать экспертное заключение. Таким образом, удалось создать уникальное сочетание работы человека и машины, позволяющее максимально полно проверить и проанализировать любой текстовый документ. Результатом этой работы является обоснованное экспертное заключение.

Сам процесс работы с документом проходит следующие стадии.

Работа начинается сразу же после получения документа от заказчика, когда эксперт проверяет соответствие формата присланного текста тем, которые поддерживает система «Антиплагиат.РГБ» (форматы pdf, doc или txt), и проводит загрузку документа в систему. В результате выявляется первоначальный процент совпадений, который в дальнейшем указывается в экспертном заключении.

Далее эксперт производит первичную обработку документа. Она заключается в удалении всех малоинформативных частей, автопереносов, а также фрагментов, которые, не являясь заимствованиями по сути, опознаются системой в качестве таковых. К ним относятся: титульный лист, содержание, список литературы, графические объекты (рисунки, схемы), формулы, таблицы. Полученный таким образом сокращённый вариант работы загружается в систему «Антиплагиат.РГБ». Он и является основным для дальнейшей обработки текста и анализа выявленных системой совпадений.

При этом система «Антиплагиат.РГБ» обеспечивает:

- формирование отчёта о фрагментах заимствования с указанием источников из ЭБД РГБ;
- указание объёма заимствования по каждому из выявленных источников из ЭБД РГБ;
- выделение общего процента заимствования в проверяемом документе.

Дополнительно для более детального изучения найденных заимствований эксперт РГБ имеет возможность просмотра полных текстов диссертаций и авторефератов из ЭБД РГБ, на которые ссылается система.

Эксперт из предоставленного системой «Антиплагиат.РГБ» отчёта вручную исключает названия глав и параграфов, наименования ГОСТов и нормативных документов, публикаций, физических и юридических лиц и т.д., проверяет правильность оформления цитат, удаляет сноски.

Основной обязанностью эксперта является оценка характера и значимости обнаруженных системой заимствований. При этом определяются степень дословности и объём совпадений, анализируется их последовательность (логическая последовательность или разбросанный порядок в тексте работы), изучается список литературы на наличие в нём обнаруженных источников цитирования.

В результате экспертной проверки в отчёте остаются только те заимствования, для которых доказана их некорректность. При необходимости эксперт проводит более подробный анализ текста и также по исходному варианту работы проверяет таблицы, изображения, графики и формулы, анализирует сформированный список литературы. Последнее представляется важным, поскольку любая библиография диссертации, оформленная в виде списка использованной литературы, является составной и, соответственно, уникальной частью диссертационного исследования. Однако встречаются случаи, когда значительная часть списка литературы механически заимствована диссертантом из другого диссертационного исследования напрямую.

Далее эксперт оценивает связь обнаруженных некорректных заимствований с целью, задачами работы, выносимыми на защиту положениями и составляет заключение об оригинальности работы.

Экспертное заключение содержит информацию о присланной работе (ФИО автора, название и др.), данные об актуальности ресурса, по которому производится проверка, процент оригинальности текстового документа в результате первичной проверки, перечень источников цитирования с указанием процента совпадения по каждому из них, а главное, подробное описание характера заимствований с указанием разделов документа, где они встречаются.

Заключение об оригинальности в данном случае является итогом проведённой независимой экспертизы, предоставляющим объективные данные. Заключение обладает юридической силой, т.к. подписывается не только экспертом, но и руководителем РГБ, и заверяется гербовой печатью.

При этом необходимо подчеркнуть, что окончательное решение о значимости выявленных некорректных заимствований и оригинальности работы остаётся за диссертационным советом, куда поступает работа, и Высшей аттестационной комиссией Министерства образования и науки Российской Федерации.

По состоянию на 1 сентября 2012 г. проверен 251 текстовый документ, из них 188 могут считаться оригинальными по отношению к коллекции ЭБД РГБ, 40 содержат несущественные заимствования, а 23 документа содержат существенные заимствования из ЭБД РГБ, которые могут быть расценены как плагиат (см. диаграмму).

Общий анализ работ, присылаемых на проверку в РГБ, показал, что зачастую авторы не умеют корректно приводить выдержки из проанализированных источников и делать на них библиографические ссылки, злоупотребляют цитированием, не формируя собственного мнения по рассматриваемым вопросам. Большинство авторов, в чьих работах был обнаружен плагиат, не перерабатывали чужой текст, ограничиваясь смысловым подбором фрагментов, соединяя их друг с другом вводными словами («мы показали, что...» и т.п.).

Большую часть присланных на проверку документов составляют диссертации, но также присылают и НИОКР. Так, к 1 сентября 2012 г. на проверку предоставили:

- диссертации по экономическим наукам (89 документов);
- диссертации по юридическим наукам (35 документов);
- отчёты о НИОКР (30 документов);
- диссертации по биологическим наукам (26 документов);
- диссертации по техническим и педагогическим наукам (по 20 документов).

Помимо системы «Антиплагиат.РГБ» существует около десятка программ, позволяющих выявить плагиат в текстовых документах. И если система «Антиплагиат.РГБ» проводит проверку по ЭБД РГБ,



Диаграмма. Результаты проверки текстовых документов на предмет наличия некорректных заимствований (плагиата)

то другие программы в качестве источников для сравнения в основном используют тексты из сети Интернет.

Однако диссертации из ЭБД РГБ обладают преимуществом в достоверности, поскольку в случае с интернет-ресурсами зачастую трудно определить первоначальный источник заимствования. К этому добавим анонимность многих статей, выложенных на различных сайтах.

В любом случае необходима комплексная проверка, включающая как ЭБД РГБ, так и другие авторизованные источники и интернет-ресурсы. В этой связи дальнейшие планы развития услуги РГБ подразумевают расширение базы источников, на основе которой производится проверка текстов, и включение в неё не только диссертаций, но и электронных библиотек, а также интернет-источников. Такая комплексная проверка текстов на предмет заимствований по ЭБД РГБ и другим источникам даст гораздо более достоверные цифры, что может повлиять на общую экспертную оценку представленных работ. ■

Использованная литература:

1. Авдеева Н.В. Электронная библиотека диссертаций Российской государственной библиотеки: история создания и перспективы развития // Информационные ресурсы России. – 2009. – № 5 – С. 17–21.
2. Авдеева Н.В., Ботов П.Ю., Букаев А.С., Вислый А.И., Груздев И.А., Житлухин Д.А., Романов М.Ю., Чехович Ю.В. Внедрение системы «Антиплагиат» в Российской государственной библиотеке // Материалы конференции «Интеллектуализация обработки информации» – октябрь, 2010. – С. 499–503.
3. Фёдорова О.А. Понятие «плагиат» в уголовном законодательстве // Применение уголовного и уголовно-процессуального законодательства в правоохранительной деятельности, судебном производстве и прокурорском надзоре: научно-практический журнал. Москва: ИНГН, 1 (3) – январь – апрель, 2012. – С. 5–7.
4. Avdeeva N. Innovative services for libraries through the virtual reading rooms of the digital dissertation library, Russian State Library // IFLA Journal. – 2010. – Vol. 36, Issue No. 2. P. 138–144.
5. Сайт проекта «Электронная библиотека диссертаций Российской государственной библиотеки» <http://diss.rsl.ru>.