

НАУЧНАЯ ЭЛЕКТРОННАЯ БИБЛИОТЕКА КАК СРЕДСТВО БОРЬБЫ С ПЛАГИАТОМ

Альгирдас Аушра

Директор

*ОУ «Балтийский общественный центр информационных технологий и просвещения».
Вильнюс (Литва), эл. почта: info@elibrary.lt*

Переход от индустриального общества к информационному обществу характеризуется многими чертами, одной из которых является то, что информация и знание становятся одним из основных факторов развития и структурирующим социальным фактором, под воздействием которого меняются формы экономической деятельности и социальные взаимоотношения. Сохранение, развитие и рациональное использование этого стратегического ресурса будущего имеют огромное значение для любого общества и государства.

Электронные библиотеки оказываются интегрирующим центром происходящих событий, охватывающих вопросы как локального, так и глобального технологического развития.

1. Рамки авторства и плагиат

Как известно, плагиат — умышленное присвоение авторства на чужое произведение науки, литературы или искусства в целом или в частности. В сегодняшнем понимании плагиат определяется как использование чужого текста, опубликованного в бумажном или электронном виде, без полной ссылки на источник или со ссылками, но когда объем и характер заимствований ставят под сомнение самостоятельность работы, произведения или одной из ее основных частей.

Плагиат может осуществляться в двух видах:

- дословное изложение чужого текста,
- парафраза – изложение чужого текста с заменой слов и выражений без изменения содержания заимствованного текста.

Современное понимание плагиата связано с современным пониманием права и морали, однако представление о плагиате менялось на протяжении столетий.

В антике никакой подписи на книге автор не делал. Об авторстве античных сочинений мы знаем только из косвенных ссылок. В письмах и книгах античные

писатели, ссылаясь на некую мысль, для удобства ссылки указывали автора, личность которого, учитывая соседский характер эллинской цивилизации, не была тайной. То есть авторство не было формализованным институтом с набором правил - от места указания автора до точного определения правил приоритета. Таким образом, авторство античности по своему характеру есть авторство косвенное, отсылочное; не сам автор указывает свое имя на книге, а другие, цитируя книгу, упоминают его имя.

Средневековье не знало авторства в современном его понимании. Оно не знало и современной цитаты. Но, если мы обратимся к истории для определения понятия плагиата на всем протяжении греко-латинского времени, то увидим, как медленно, с определенными и, вполне закономерными, перерывами возникает представление о духовной собственности, о духовной личности человека, об индивидуальном творчестве, результат которого не может быть подделан в своих основах и должен быть защищен от подделки в своем внешнем выражении. Мыслительная работа схоластов способствовала развитию души в этом направлении, сама, в свою очередь, являясь результатом определенного изменения человеческой душевности и вызванная к жизни ростом индивидуальности людей.

Современное содержание понятия "цитата" выработано в работах схоластов. Хотя в самой схоластике "точная цитата" еще не совсем развилась, работа схоластов объективно способствовала развитию понятия "цитата" и понятию авторства текста. Таким образом, авторство рождается из цитаты, причем рождение это происходит достаточно медленно, и даже на исходе античности современное состояние феномена авторства достигнуто не было.

Можно сделать заключение, что у древних греков плагиат не считался зазорным. В древнем мире научных знаний и ученых было не так много. Производитель знания высоко ценился. Распространение знаний происходило либо путем личного контакта с их носителем, либо путем чтения книг, тиражируемых вручную и поэтому представляющих собой весьма редкое явление. Естественно, в такой ситуации большим достижением считалось то, что некто сумел освоить какие-то научные представления деятелей прошлого и пересказал их на новый лад. Такой «переносчик» знаний может считаться ученым и быть всеми уважаемым даже при условии наличия у него относительного невысокого интеллектуального потенциала как производителя нового знания. Плагиат в современном понимании расценивался как положительное явление.

Понятие «авторства» появилось в Новое время, с ростом индивидуализма после Ренессанса. Теперь же идет другой процесс, осознанное авторство начинает разлагаться, и появляется коллективное авторство. Появление авторства - хороший маркер происходящих изменений в душах людей, когда становится важным не то (или не только то), что говорится, а кто и когда это сказал первый. Но изменения человеческой души не закончились, и превращения категории авторства по-прежнему маркируют происходящие в глубинах души процессы.

Наиболее заметен и легче всего доказуем плагиат в музыке. Более затруднительно увидеть плагиат в произведениях искусства и литературы, научные труды, т.к. не совершенна и затруднительная система их легализации. Они чаще всего доступны узкому кругу специалистов данной области. Не менее важным препятствием является и узость специализации представителей данной сферы.

2. Плагиат в Интернете

Плагиат с появлением Интернета превратился в серьезную проблему. Попав в Интернет, знание становится достоянием всех, соблюдать авторское право становится все более трудно и даже невозможно. Авторитет производителя знаний постепенно снижается из-за сложности его идентификации.

В марте 2006 года был обнародован третий ежегодный отчет школы журналистики при Колумбийском университете. Эксперты проанализировали работу американских СМИ в течение одного дня – 11 мая 2005 года. Интернет-поисковик в этот день проиндексировал 14 000 новостных ссылок. Исследователи прочитали все и обнаружили, что они посвящены только 24 новостям. Специалисты Колумбийского университета также тщательно изучили печатные СМИ и вечерние новости эфирных телеканалов. Американские исследователи делают вывод, что новости в американских СМИ 11 мая 2005 года оказались бесконечным повторением двух десятков историй, при этом большинство материалов оказались новостями-однодневками.

Публичная Интернет-библиотека Public.ru сообщает о проведенном подобном исследовании. Чтобы проследить, какие именно публикации получают массовую, так называемую веерную перепечатку в региональной прессе, был проведен анализ контента региональных изданий за последнюю декаду марта 2006 года (за период с 20 по 31 марта). При исследовании отбирались буквальные информационные «дубли», распространяющиеся по всей территории нашей страны в виде печатных материалов. Всего за 11 дней таких материалов было найдено 67. Кроме того, на примере

нескольких материалов было проверено, не встречались ли эти же материалы в местных СМИ раньше, чем рассматриваемый период времени, и оказалось, что такие случаи также имели место. В рейтинге наиболее популярных тем на первом месте оказались "семейные советы" - жизненные истории и письма читателей в редакцию с рассказами о личной жизни. На втором – юмористические сюжеты или анекдоты. Тема здоровья заняла третью строчку. А статьи, посвященные актуальным событиям федерального масштаба, - девятую.

По мере распространения Интернета плагиат становится все большей проблемой не только в вузах, но уже даже в школах. Ситуация усугубляется тем, что школьники и студенты обычно на порядок превосходят своих учителей и преподавателей по уровню компьютерной грамотности. Как известно, учащиеся обычно копируют из онлайн-коллекций рефератов целые страницы текста.

Данные исследований о плагиате в научных кругах пока отсутствуют.

По данным последнего опроса, проведенного профессором Доном Маккейбом (Don McCabe) из Ратгерского университета (Rutgers University), 37% из 30 тыс. студентов в 34 колледжах в США пользуются заветным "copy&paste" для копирования кусков чужих работ, найденных в Сети, в свои собственные. В 1999 году таковых было только 10%. По последней статистике, 38% студентов признаются в копировании чужих рефератов на компьютере. По другим данным, 80% студентов хотя бы однажды обманывали преподавателей.

Университеты разных стран продолжают бороться с воровством рефератов и курсовых из Интернета. Около 20% университетских профессоров в США уже используют в своей деятельности программные продукты, выявляющие факты плагиата. Все большее количество издательств, юридических и других компаний также начинают применять в повседневной деятельности программное обеспечение, позволяющее анализировать огромные объемы текстовой информации и за считанные секунды выделять в них идентичные блоки.

Один из популярных Интернет-сервисов, позволяющих выискивать плагиаторов, создан компанией iParadigms LLC из г. Оклэнд, Калифорния. Его история берет начало в 1996 году, когда разработанная специалистами компании программа начала использоваться преподавателями Калифорнийского университета в г. Беркли для анализа студенческих работ. В отличие от обычных поисковых, Интернет-систем типа Google, системы осуществляют поиск по введенным ключевым словам. Программное обеспечение iParadigms создает цифровой образ всего проверяемого

документа, после чего сравнивает его с материалом, в том числе в Интернете, в специализированных базах данных, в академических ресурсах и в базах данных СМИ. Популярность нового сервиса оказалась столь велика, что его услугами с прошлой осени стал пользоваться даже Совет Безопасности ООН. Сегодня детектор плагиата Turnitin применяется более чем в 2500 высших учебных заведениях США и более чем в 1000 - за рубежом. Коммерческая версия программы iThenticate появилась в январе этого года.

Пользование iParadigm обходится университетам в \$500 ежегодных лицензионных отчислений плюс \$0,6 за каждого студента. Коммерческие пользователи платят \$1000 в год плюс \$10 за каждую страницу проанализированного текста. Для печатных средств массовой информации расценки определяются с учетом количества слов в анализируемом тексте и тиража издания.

Аналогичные детекторы плагиата предлагают сегодня на рынке и другие компании. Многие - к примеру, Glatt Plagiarism Services, MyDropBox, а также CFL Software Development, - сообщают о росте популярности своих продуктов не только в университетских стенах, но и за их пределами. Плагиат выискивают компании, выпускающие учебную продукцию, юридические фирмы, Интернет-компании, а также государственные органы. При этом зачастую выявляются весьма любопытные факты.

К примеру, программа iParadigm помогла газете The Hartford Courant подтвердить, что президент Центрального университета штата Коннектикут Ричард Джадд (Richard Judd) в своем обзорном материале использовал готовые предложения не только из газеты The New York Times, как предположил один из читателей, но и еще из трех источников - в общей сложности 11% его текста были "позаимствованы". Предварительная попытка проверить обоснованность подозрений внимательного читателя с помощью обычной поисковой системы успеха не принесла.

В Университете Роберта Гордона из Абердина (Шотландия) создан Интернет-сервис Turnitin. Систему используют так: студента просят сдать реферат не только в бумажном, но и в цифровом виде, после чего файл закачивают на сервер Turnitin (www.turnitin.com). Программа сравнивает текст с образцами из своей базы данных объемом более 25 ТБ и в течение нескольких секунд вычисляет плагиатора. Онлайн-сервис Turnitin содержит в своей базе более 800 млн. веб-страниц и более 4,5 млрд. рефератов и курсовых, а также множество других источников информации. Робот Turnitin бороздит просторы Интернета и ежедневно добавляет в БД около 40 млн. новых страниц, обрабатывая файлы размером до 2 МБ в разных форматах, на что не

способны даже многие специализированные поисковые машины. Преподаватели могут самостоятельно загрузить на сервер работы своих студентов (до 1000 заархивированных файлов за одну загрузку), чтобы получить гарантию, что эти тексты не будут в дальнейшем использованы их однокурсниками.

В прошлом году благодаря Turnitin в Университете Роберта Гордона было поймано такое огромное количество студентов-плагиаторов, что, по словам профессора Джастина Гринвуда (Justin Greenwood), его первой мыслью было вообще отменить рефераты и проверять студентов исключительно в очном тестировании. В конце концов, пришлось перейти к новому методу согласования тем.

Подобная система уже будет работать и в русских вузах. В России будет разработана информационная база по рефератам, представленным в Интернете, которая позволит каждому преподавателю проверить студента на добросовестность. Проверку будет осуществлять специальная компьютерная программа, разработка которой начнется вскоре. Технические детали проекта пока неизвестны, поэтому непонятно каким именно образом будет вестись сверка текста представленной студентом работы с огромным текстовым массивом рефератов, представленным в Сети. Так же неясно, как будет вестись борьба с нерадивыми студентами, заказывающими рефераты и курсовые у профессиональных "писателей" научных работ.

Для идентификации плагиата применяются и другие средства. Одно из них – индекс научного цитирования, который появился в 60-х годах прошлого века. Идею этого статистического метода сформулировал химик Юджин Гарфильд. Он предложил в качестве исходных данных для оценки использовать список научной литературы, который есть в конце любой научной статьи. Если статья полезна и ее материалы на что-то влияют, значит, ее должны часто цитировать. На сегодняшний день это специализированный информационный продукт, в котором собирается и обрабатывается полная библиографическая информация о журнальных статьях, аннотации и статейные списки цитируемой в данных статьях литературы. Такая база позволяет находить как публикации, цитируемые в отдельно взятой статье, так и публикации, цитирующие эту статью. Этот механизм дает возможность интегрировать публикационные и цитатные показатели по всей вертикали социального института науки: от научного сотрудника-автора, структурного подразделения и учреждения, где работает круг авторов, до министерств и ведомств или целых административно-географических регионов. Проект Российского индекса научного цитирования

осуществляет www.elibrary.ru. Подобный проект уже создан в Китае (Китайский индекс научного цитирования включает свыше 1000 национальных журналов). База данных по европейским научным журналам “EuroFactor” все еще находится в начальной стадии.

Существуют и другие методы установления плагиата. Среди них публикация научных и других ресурсов в виде ресурсов открытого доступа в личных сайтах интернета, на страницах высших учебных заведений, в научных электронных библиотеках. Последнее заслуживает более тщательного анализа.

3. Научная электронная библиотека и борьба с плагиатом

Научная электронная библиотека наряду с программами идентификации текста, индексом научного цитирования и др. также является средством в борьбе с плагиатом.

В 1993 году в Литве было основано Объединение научных библиотек Литвы (ОНБЛ), целью которого является помощь в осуществлении проектировании, внедрении, эксплуатации и финансировании проектов интегральной информационной системы библиотек Литвы (ИИСБЛ). В соответствии с решением Союза научных библиотек Литвы, в 1995 году Национальная библиотека Литвы им. Мартинаса Мажвидаса объявила конкурс на подготовку проекта ИИСБЛ. Победитель конкурса Каунасский технологический университет подготовил проект ИИСБЛ, который в 1996 году был внесен в Государственную программу по развитию связей и информатики. На 1997-1999 год проект получил финансирование государства и фонда Сороса.

С целью лучшего внедрения ИИСБЛ в вузах библиотек Литвы, в 1997 году была основана Ассоциация научных библиотек Литвы (АНБЛ). Она дала поручение Каунасскому технологическому университету на основе проекта ИИСБЛ создать проект компьютеризации всех научных библиотек Литвы «Сеть научных библиотек Литвы» (СНБЛ). Основная цель данного проекта – обеспечение научных библиотек вузов Литвы современной, признанной в мире, научной системой. Данный проект получил поддержку конфедерации ректоров Литвы, ему было назначено дополнительное финансирование. С 1998 года по поручению конфедерации ректоров Литвы, Каунасский технологический университет разрабатывает проект СНБЛ. В 2004 году в СНБЛ входило 16 университетских библиотек Литвы, 20 библиотек неуниверситетских высших учебных заведений (коллегий), библиотека Академии Наук Литвы и 39 библиотек государственных научных институтов.

Публикация научных работ в электронной библиотеке – это одна из основных возможностей доказать автору свои права на авторство работы. Не секрет, что среди ученых наиболее часто воровство идей происходит на научных конференциях, семинарах. Ученый, не успевший оформить свою идею в виде публикации, на следующей конференции или до нее может увидеть нечто очень схожее с его суждениями или тождественное с его идеями.

Не менее актуально это и для студентов ВУЗов, защищающих свои работы на бакалавра или магистра, Присвоение работы студента его научным руководителем также не столь редкое явление. Об этом все чаще задумываются студенты и торопятся опубликовать свои работы в изданиях или нашей научной электронной библиотеке. Инициатива Сельскохозяйственного университета Литвы публиковать работы студентов даже до их защиты может послужить хорошим примером и для других вузов страны.

Многие научные журналы стараются сделать свои материалы общедоступными, невзирая на то, что это, казалось бы, противоречит их финансовым интересам. Однако серьезных исследований, посвященных влиянию открытого сетевого доступа на цитируемость публикаций, до сих пор было сравнительно немного.

Канадский исследователь Гюнтер Айзенбах из университета Торонто установил, что научные статьи, находящиеся в открытом сетевом доступе с момента публикации, цитируются гораздо активнее даже в том случае, если журнал, в котором они опубликованы, есть в большинстве научных библиотек. Резко повышенный уровень цитируемости сохраняется даже после того, как в открытый доступ по истечении определенного срока попадают и все остальные статьи из данного журнала. Гюнтер Айзенбах избрал для анализа один из самых серьезных и уважаемых биологических журналов — Proceedings of the National Academy of Sciences of the USA (PNAS). В этом журнале с недавних пор часть статей находится в открытом доступе с момента публикации (для этого авторы должны заплатить журналу \$1000), а спустя 6 месяцев в открытый доступ попадают все остальные статьи. Айзенбах проанализировал цитируемость 1492 статей, опубликованных в PNAS с 8 июня по 20 декабря 2004 года, из которых 212 (14.2%) находились в открытом доступе. Выяснилось, что статьи, находящиеся в открытом доступе с момента публикации, вносят гораздо больший вклад в науку, поскольку их начинают цитировать раньше и в дальнейшем цитируют намного активнее, чем статьи, доступные лишь по платной подписке или в библиотеке и становящиеся открытыми лишь спустя 6 месяцев. Было

подсчитано, что вероятность цитирования "открытых" статей в течение первых 10 месяцев после публикации в 2.1 раза выше, чем у платных, а в последующие 6 месяцев – в 2.9 раза выше.

В свете этих данных крайне недальновидной выглядит информационная политика некоторых издательств, контролирующих выпуск и распространение научных изданий. Эти издательства за редчайшими исключениями категорически запрещают вывешивать в сети тексты публикуемых статей под тем предлогом, что от этого якобы могут упасть доходы от подписки. При крошечных тиражах некоторых изданий едва ли погоня за ничтожной прибылью может оправдать колоссальные потери, которая наука, несомненно, несет из-за такой политики.

Общественное учреждение Балтийский общественный центр информационных технологий и просвещения, осуществляющий проект Научной электронной библиотеки eLIBRARY.LT, стремится накопить национальные ресурсы, ознакомить общественность с новейшими научными работами в широком смысле, знакомить с достижениями передовой науки, расширять интеллектуальные направления развития человечества. Постоянно внедряя различные проекты, пропагандируются и представляются научные знания, создаются научные традиции.

В электронной библиотеке eLIBRARY.LT размещены ресурсы открытого доступа, предоставляющие научную, учебную и познавательную информацию в следующих направлениях:

- Базы данных (информация, документы);
- Наука Литвы (журналы, издания);
- Литовские издания (книги, аудио, видео);
- Литовские (интеллектуальные) СМИ;
- Зарубежные (научные) издания;
- Материалы (доклады) конференций.

В литовской базе данных собрана обширная информация и документы важнейших госучреждений Литвы, Парламента (Сейма) и Правительства ЛР. Также информация и важнейшие документы организаций ЕС, другая познавательная информация. В разделе Наука Литвы – научные издания Академии Наук Литвы, научные труды государственных учреждений Литвы и научные труды членов Парламента (Сейма) ЛР, издания университетских и неуниверситетских ВУЗов Литвы. В разделе Литовские издания размещаются монографии, научно-популярные и

просветительские издания, книги практического направления, другие сочинения авторов Литвы. В разделе Литовские (интеллектуальные) СМИ размещены издания государственных и общественных учреждений, периодические издания, выпускаемы за рубежом и др. В разделе Зарубежные (научные) издания размещены электронные ресурсы научных, информационных, исторических и высоких технологий. В разделе Конференции накапливаются материалы различных научных, научно-практических конференций и семинаров.

ОУ Балтийский общественный центр информационных технологий и просвещения стремится оказать помощь в создании благоприятных условий для эффективного применения новейших технологий и пользования ими в нашей стране. Организация интересуется нововведениями и всесторонне поощряет их внедрение. Основная цель деятельности организации – поощрение роста информационного развития общества, внедрение новых технологий в области науки и просвещения.

Комитет по развитию информационного общества при Правительстве Литвы присвоил проекту - научная электронная библиотека eLIBRARY.LT статус – лучший проект 2005 года в категории применения ИТ в области просвещения и обучения взрослого населения.

Осуществляя свои функции, научная электронная библиотека eLIBRARY.LT становится виртуальным издательством и просветительским учреждением. Благодаря научной электронной библиотеке для литовского академического общества открывается прекрасная возможность узнать о новейших научных достижениях, идеях, проектах, делах и представить себя мировой общественности, стремящейся к знаниям.

СПИСОК ЛИТЕРАТУРЫ

1. Жданова Р. Понятие плагиата // <http://www.iatp.kz/?id=news&sendnews=1868&lang=2>
2. Arms W. Digital Library. The MIT Press. 2000
3. Arms W. Automated Digital Libraries. D-Lib Magazine, 2000, v.6, #7-8 // <http://www.dlib.org/dlib/july00/arms/07arms.html>
4. Biliæ-Zulle L., Frkoviæ V., Turk T., Azman J., Petroveèki M. Prevalence of Plagiarism among Medical Students, Croat Med, J. – 2005 – 46 (1) – P. 126-131.
5. Bloomfield L. The importance of writing. Originally published on the Commentary Page of the Philadelphia Inquirer on Sunday / Ed J. Timpane. – April 4, 2004.
6. Crane G. Commercial Digital Libraries and the Academic Community. D-Lib magazine, 2001, v.7, #1 // <http://www.dlib.org/dlib/january01/crane/01crane.html>
7. *Gunther Eysenbach. Citation Advantage of Open Access Articles // PLoS Biology. May 2006. V. 4. Issue 5.*
8. Ershova T.V., Hohlov Yu.E. Migrating from the Library of Today to the Library of Tomorrow: Re- or E-volution? // Libraries in the Information Society, IFLA Publications 102. – Munchen: Saur, 2002. - P. 74-81.
9. Sixth Thematic Workshop on Digital Library Architectures // <http://www.delos.info/newsletter/pdf/delos-newsletter-issue2.pdf>
10. The Grid: Blueprint for a New Computing Infrastructure. Ed. by I.Foster and C.Kesselman. Morgan Kaufmann Pub., San Francisco, CA. 1999.

15-20 мая 2006 г.

Будва, Черногория