

**ПРОВЕРКА ПЛАГИАТА В ВЕБ-ПРИЛОЖЕНИЯХ**

Голобурда А.В., Попова Ю.Б.

БНТУ, г. Минск, Беларусь, [goloburda.av@gmail.com](mailto:goloburda.av@gmail.com)БНТУ, г. Минск, Беларусь, [julia\\_porova@mail.ru](mailto:julia_porova@mail.ru)*Введение*

В настоящее время влияние информационных технологий на систему образования весьма велико. Бесспорны позитивные стороны такого внедрения, поскольку происходит повышение доступности информации, а также ускорение и упрощение коммуникаций. Обучающиеся приобретают возможность быстро и практически без затрат пользоваться источниками и нужной информацией, получение которой ранее потребовало бы большое количество времени и сил. Аналогичное утверждение относится и к коммуникации: студент может связаться с преподавателем из дома, из другого города и даже из другой страны. Информационные технологии также позволяют сделать нагляднее подачу материала с помощью различного рода анимаций, схем и прочих средств.

Таким образом, основное преимущество применения информационных технологий в образовании — это ускорение, упрощение и расширение обмена информацией. Однако именно это является одновременно и недостатком, минусом процессов информатизации, потому что такая легкость порождает проблему, которая становится все актуальнее в последнее время — проблему плагиата в образовании [1].

Плагиат превратился в существенную проблему высшего образования большинства стран мира. По некоторым встречающимся оценкам доля самостоятельных работ обучающихся, имеющих очевидные признаки плагиата, достигает 70 и более процентов. Это связано с распространением сети Интернет, возрастающей загруженностью преподавателей, а также неуважением к авторскому праву и невысоким уровнем академической культуры [2].

Поэтому многие образовательные учреждения осознают необходимость проведения политики недопустимости плагиата и предпринимают конкретные меры для его выявления. К более действенным из них следует отнести внедрение соответствующего программного обеспечения.

*1. Обзор систем выявления плагиата*

В настоящее время существует множество on-line сервисов и специализированных программ, как бесплатных, так и коммерческих, предназначенных для выявления плагиата. Рассмотрим некоторые из них:

1) Наиболее известной из таких систем является лицензированная программа Turnitin, разработанная компанией iParadigms и доступная по адресу: <http://www.turnitin.com>. Программа используется более чем в десяти тысячах университетов и колледжей 126 стран мира [3]. Алгоритм работы в ней следующий. Преподаватель создает аккаунт студенческой группы на сайте Turnitin. После этого студенты получают персональные пароли и могут отправлять свои работы в электронном виде. Войдя под собственным паролем, студент видит на странице сроки сдачи работы, результат ее проверки в виде процента уникальности текста, а также помеченные определенным образом заимствования и их оригинальные источники. Преподаватель видит весь список группы с подобными результатами каждого студента. Таким образом, проверка на плагиат практически не требует от преподавателя дополнительных усилий. Каждый текст сравнивается с более чем 135 млн. ранее загруженных студенческих работ, со статьями более 90 тыс. периодических изданий и 13,5 млрд. Интернет-страниц, а также с другими источниками, с которыми у владельцев ресурса заключены соглашения об использовании. Одной из сильных сторон Turnitin является его популярность в США и странах Западной Европы. Туда отправляют большое количество работ: студенты — свои, преподаватели — студенческие, и их база поиска постоянно растет. В результате существует возможность находить плагиат не только из Интернета, но и из

прошлогодних работ. К недостаткам Turnitin можно отнести малую популярность в нашей стране, что делает систему малоэффективной при проверках лабораторных, практических и контрольных работ.

2) В 2008 году появилась российская система «Антиплагиат» от фирмы-разработчика «Форексис», расположенная по адресу <http://www.antiplagiat.ru>. Первым ее стал системно использовать государственный университет «Высшая школа экономики». Основное назначение системы «Антиплагиат» — это проверка текстовых документов на наличие заимствований. Необходимость такой проверки чаще всего возникает при проверке учебных работ обучающихся. Другой, не менее важной областью применения системы является анализ научных работ. «Антиплагиат» использует уникальные алгоритмы, разработанные российскими учеными, что гарантирует быстрый и эффективный поиск заимствованных фрагментов, а также обеспечивает корректную обработку текстов на русском языке [4].

Принцип работы системы «Антиплагиат» следующий: необходимо открыть сайт и загрузить документ, например, реферат. Сервис ищет в сети Интернет похожие тексты, и через непродолжительный промежуток времени выдает результаты: процент уникальности текста и ссылки на настоящего автора.

Главное превосходство системы «Антиплагиат», как считают создатели проекта, — это повышение качества образования и побуждение обучающихся к самостоятельному написанию работ. Для преподавателей этот проект стал хорошим помощником, потому что они легко могут определять уникальность работы обучающегося. Также следует отметить, что алгоритмы программы защищены от манипуляций, которые выполняются обучающимися для корректировки результата и увеличения уникальности работ. А именно, сервис защищен от подмены кириллических букв латинскими, от перестановки абзацев, от перестановки слов и перегруппировки предложений, от замены пробелов точками, точек запятыми и отдельных слов их синонимами. Кроме того, выполняя данные манипуляции с работами, обучающийся достаточно углубленно изучает тему задания, выполняет своего рода самостоятельную работу, а не элементарно списывает чужие мысли из Интернета.

Помимо перечисленных преимуществ системы, необходимо отметить и недостатки. В первую очередь — это сложность проверки огромного объема работ, что практически постоянно приходится делать преподавателям. Помимо всего этого, порой система необъективно оценивает работу. Нередко ответственные студенты, добросовестно выполняющие все задания собственными силами, имеют все шансы получить отрицательную оценку своей работы после проверки системой. Сложность состоит в том, что цитаты или ссылки на источники система часто определяет, как плагиат. Получается, что уникальность текста снижается, хотя на самом деле студент выполнил качественно и грамотно задание самостоятельно.

3) Программа Advego Plagiatius осуществляет онлайн-проверку с использованием поисковых систем. В отличие от подобных систем, Advego Plagiatius не использует Яндекс.XML. Программа выдает процент совпадения текста и выводит найденные источники. К недостаткам данной программы можно отнести то, что в ней отсутствуют преобразование регистра, обработка и замена латинских букв в русских словах на аналогичные буквы русского алфавита для слов на русском языке. Также отсутствует собственная база для поиска. Из-за особенностей работы возникают ситуации, когда результаты проверки отличаются от раза к разу.

Программа Praide Unique Content Analyser II проверяет текст с использованием поисковых систем. Существует возможность выбора используемых поисковых систем, а также добавление новых. Проверка осуществляется шинглами (отдельно взятыми частями текста, по которым и проходит проверка на уникальность), длину которых можно менять. Можно задавать количества слов перекрытия шинглов. В результате выводится детальный отчет по проверке в каждой поисковой системе. В программе отсутствуют подмены букв, обработка стоп-слов, а также поддержка работы с собственной базой.

4) Сервис Unplag также является сервисом проверки на плагиат. Он может выполнять проверку как онлайн, так и сравнивать документ с собственной сохраненной базой документов пользователя. Сервис поддерживает работу с различными типами документов и имеет персональную и корпоративную версии.

5) Система Plagiatinform проверяет документы на наличие совпадений в сети Интернет и в локальной базе. Система может находить плагиат в виде документов, скомпонованных из кусков текста нескольких источников. Проверка может выполняться с помощью быстрого или углублённого поиска. Результаты проверки выдаются в виде отчёта. В качестве недостатка можно указать отсутствие преобразования букв.

6) Сервис [www.miratools.ru](http://www.miratools.ru) выполняет онлайн проверку текста на плагиат. Система использует результаты выдачи поисковых систем. Существует возможность замены латинских букв на русские. Также имеется возможность изменения длины и шага шингла, используемых для проверки. В результате проверки выдаётся процент совпадений и найденные источники. Система не имеет собственной базы, существует ограничение на длину текста в 3000 символов и ограничение на число проверок в течение суток.

7) Сервис [www.istio.com](http://www.istio.com) осуществляет проверку текста на наличие плагиата с использованием поисковых систем Яндекс.XML и Yahoo.com. По результатам проверки выдаётся сообщение о том, является ли текст уникальным или нет, и выдаётся список источников. Сервис дополнительно проверяет орфографию и делает анализ наиболее частотных слов. У системы отсутствует преобразование букв и поиск по собственной базе [5].

Почти во всех рассмотренных выше приложениях по выявлению плагиата существенным недостатком является сложность либо невозможность интеграции с сторонними приложениями.

## *2. Основные требования к системам выявления плагиата*

Автоматизированные системы управления обучением являются основой современного учебного процесса и используются для разработки, управления и распространения учебных онлайн-материалов с обеспечением совместного доступа. Данные системы сейчас используются практически в каждом учебном заведении. Поэтому одним из обязательных условий системы по выявлению плагиата должна быть интеграция с системой управления учебным процессом. Это позволит не только упростить преподавателю процесс проверки до одного клика, но и даст возможность сверить лабораторные, практические и контрольные работы с последними за несколько лет.

В настоящее время при создании приложения можно использовать различные виды инструментов. Это могут быть как обычные десктопные приложения, так и веб-приложения. Достоинствами браузерных приложений являются:

- отсутствие установки;
- отсутствие особых требований к аппаратной платформе;
- наличие поддержки старых версий программ и обратной совместимости;
- мобильность.

Однако у веб-приложений существуют и слабые стороны:

- зависимость от наличия Интернета;
- хранение информации на удаленном сервере, что может привести к потере или искажению информации.

Разрабатываемый программный продукт должен быть реализован как самостоятельное приложение с простым и понятным интерфейсом. Интеграция со сторонними приложениями будет реализована с помощью веб-службы.

Веб-служба, веб-сервис (англ. web-service) — это сетевая технология, обеспечивающая межпрограммное взаимодействие на основе веб-стандартов. Веб-сервис идентифицируется строкой URL. Веб-сервис имеет программный интерфейс, представленный в машинно-обрабатываемом формате WSDL. Другие системы взаимодействуют с этим веб-сервисом путем обмена сообщениями протокола SOAP. В качестве транспорта для сообщений

используется протокол HTTP. Описание веб-сервисов и их API могут быть найдены средствами UDDI. Концептуальная схема технологии приведена на рисунке 1, а связь между протоколами — на рисунке 2 [6].



Рисунок 1 - Концепция веб-сервиса

- SOAP (Simple Object Access Protocol) — протокол обмена сообщениями между потребителем и поставщиком веб-сервиса;
- WSDL (Web Services Description Language) — язык описания внешних интерфейсов веб-службы;
- UDDI (Universal Discovery, Description and Integration) — универсальный интерфейс распознавания, описания и интеграции, используемый для формирования каталога веб-сервисов и доступа к нему



Рисунок 2 - Протоколы веб-сервисов

Все спецификации, используемые в технологии, основаны на XML и, соответственно, наследуют его преимущества и недостатки.

К преимуществам данного типа взаимодействия можно отнести:

- веб-службы обеспечивают взаимодействие программных систем независимо от платформы. Например, Windows-C#-клиент может коммуницировать с Java-сервером, работающим под Linux;
- веб-службы основаны на базе открытых стандартов и протоколов. Благодаря использованию XML достигается простота разработки и отладки веб-служб;
- использование Интернет-протокола обеспечивает HTTP-взаимодействие программных систем через межсетевой экран. С другой стороны, веб-службы не привязаны намертво к HTTP — могут использоваться и другие протоколы.

Но есть и недостатки:

- меньшая производительность и больший размер сетевого трафика за счёт использования текстовых XML-сообщений. Однако на некоторых веб-серверах возможна настройка сжатия сетевого трафика;

– аспекты безопасности. Ответственные веб-службы должны использовать кодирование, возможно — требовать аутентификации пользователя.

### *Заключение*

Таким образом, проведя анализ существующих систем по выявлению плагиата и выявив все слабые и сильные стороны каждой из них, была разработана система требований к проектируемому программному обеспечению с учетом недостатков существующих аналогов. Главным преимуществом планируемого приложения определения плагиата будет простота интеграции в уже существующие системы управления учебным процессом, что позволит преподавателю одним кликом мыши найти одинаковые работы не только внутри студенческой группы и курса, но и сравнить их с прошлогодними. Это позволит более объективно оценить знания студента.

### Список использованной литературы

1. Петракова, И.Н. Философский анализ проблемы плагиата в образовании: к вопросу о причинах // Известия Тульского государственного университета. Гуманитарные науки, 2014 — 1с.
2. Ахметов, Р.Ш. Противодействие студенческому плагиату как необходимое условие эффективности обучения [Электронный ресурс]. — Режим доступа: <http://pandia.org/text/78/454/19874.php> — Дата доступа 26.10.2016.
3. Turnitin quick facts [Электронный ресурс]. — Режим доступа: [https://turnitin.com/static/resources/documentation/turnitin/sales/Turnitin\\_Quick\\_Facts.pdf](https://turnitin.com/static/resources/documentation/turnitin/sales/Turnitin_Quick_Facts.pdf) — Дата доступа 26.10.2016.
4. Антиплагиат [Электронный ресурс]. — Режим доступа: <http://www.antiplagiat.ru/Page/About> — Дата доступа 26.10.2016.
5. Выявление плагиата [Электронный ресурс]. — 2016 — Режим доступа: [https://ru.wikipedia.org/wiki/Выявление\\_плагиата](https://ru.wikipedia.org/wiki/Выявление_плагиата) — Дата доступа 25.10.2016.
6. Веб-сервисы как средство интеграции приложений в WWW [Электронный ресурс] — 2014 — Режим доступа: <http://www.4stud.info/networking/web-services.html> — Дата доступа: 26.10.2016.