

## От редактора: Вызов Диссернета

Ник Рашби<sup>1,2</sup>

*Научный редактор журнала «Образование и Саморазвитие»*

*<sup>1</sup>Казанский федеральный университет, Казань, Россия*

*<sup>2</sup>Конейшн Текнолоджиз, Великобритания*

*nick.rushby@conation-technologies.co.uk*

ORCID: 0000-0002-2379-1402

### Влияние Диссернета

В марте 2016 года редакторская команда журнала «Образование и Саморазвитие» начала предпринимать целенаправленные действия, чтобы обеспечить вхождение журнала в базы данных Scopus и Web of Science. Эти действия привели к созданию программы развития, которая в том числе направлена на продвижение журнала в качестве международного периодического издания, ориентированного на иностранных исследователей, заинтересованных в российской педагогической науке, а также на российских ученых, ищущих доступное окно в научный мир других стран (Rushby, 2016). В рамках сформировавшейся программы развития журнал перешел к самым высоким публикационным этическим стандартам, а это означало, что мы были готовы (по крайней мере, частично) к экспертизе Диссернета, которая выявила, что некоторые опубликованные в журнале статьи, содержали некорректные заимствования, что в свою очередь вызвало подозрения в редакционной недобросовестности.

В результате, несмотря на то, что выявленные проблемы возникли еще до того, как процесс трансформации журнала был запущен, нам стала очевидной необходимость незамедлительных мер по восстановлению репутации журнала и по исключению возможности возникновения подобных проблем в дальнейшем.

### Ответные действия журнала «Образование и Саморазвитие»

Этическая политика журнала, содержащая предписания для редакционной команды, для рецензентов и для авторов, а также для Казанского федерального университета, который выпускает журнал, были обновлены и представлены на сайте журнала в ноябре 2016 года (См. [www.en.eandsdjournal.org/ethical-policy/](http://www.en.eandsdjournal.org/ethical-policy/)). Эти предписания основаны на руководстве, выпущенном Комитетом по Публикационной Этике (the Committee on Publication Ethics – COPE, 2017). Кроме того, общая этическая политика вместе с конкретными предписаниями для редакционной коллегии, для рецензентов и авторов были разосланы всем этим лицам в качестве прямого руководства.

Далее, журнал принял отставку двух членов редакционной команды, которые имели отношение к нарушению установленных правил. Более серьезным вызовом стала необходимость принять действия по отношению к тем статьям, в которых были обнаружены следы заимствования.

Было решено, что одиннадцать статей (все они были опубликованы в 2015 году) должны быть проверены на предмет некорректных заимствований, а в случае необходимости в их отношении должна быть проведена процедура ретракции. Как далее будет показано, исполнение намеченного – весьма трудоемкая процедура. Очень важно, чтобы случаи возможного плагиата рассматривались с предельным вниманием, а принимаемые решения основывались на фактах, а не предположениях.

### ***Комментарии по поводу программного обеспечения, нацеленного на отслеживание плагиата***

Программы, проверяющие на наличие плагиата, работают по принципу сравнения проверяемой статьи с имеющейся базой данных, которая включает в себя тексты статей, книг, выступлений конференций, диссертаций и прочих материалов, то есть они настроены на поиск каких-либо совпадений между проверяемой статьей и опубликованными ранее трудами. Это означает, что эффективность данных программ очень сильно зависит от полноты и актуальности их баз данных. Объем публикуемых материалов постоянно растет, а в связи с этим базы данных должны все время обновляться. Материал, содержащийся в базе данных, полностью определяет ее функциональность: статьи, которые только что были опубликованы или еще находящиеся в процессе публикации, могут только спустя какое-то время быть добавлены в базу.

Программа iThenticate предоставляет возможность, как включить, так и исключить из проверки библиографию и выделенные кавычками цитаты. Очень вероятно, что статьи на схожую тематику будут ссылаться на одни и те же научные труды. Когда автор делает прямую цитату, он должен заключать ее в кавычки (" " в английском языке и « » в русском) и указывать авторство. Использование коротких цитат не является научным нарушением. Целесообразно не определять как плагиат короткие последовательности слов (обычно до 10 слов), поскольку они могут быть просто встречающимися в письменной речи выражениями. Данные аспекты в значительной мере могут влиять на индекс схожести iThenticate. В одном из примеров (Статья 7) объем совпадений составлял 21%, когда библиография и цитаты были включены в расчет, однако после их исключения объем совпадений сократился до 11%. В данной статье все проведенные сравнения с помощью iThenticate были осуществлены с исключением библиографии и цитат, а также с исключением коротких последовательностей слов (до 10 слов).

### **Какое программное обеспечение использовать?**

Журналу «Образование и Саморазвитие», так же как и другим периодическим изданиям, выпускающим материал на двух или более языках, необходимо располагать несколькими программами проверки. Важно, чтобы было единство требований, предъявляемых к материалам как на одном, так и на другом языке.

В виду этого мы рассмотрели следующие программные системы: iThenticate, РУКОНТекст и Антиплагиат.

### ***iThenticate***

Данное программное обеспечение является, пожалуй, самой популярной системой проверки англоязычных материалов. iThenticate утверждает, что они располагают самой большой базой данных, в которую входят:

- 50 миллионов статей, материалы конференций и книг, выпущенных более чем 800 научными, техническими и медицинскими (STM) издательствами,
- 110 миллионов online и offline подписок и научных материалов, предоставляемых тридцатью лидирующими сборщиками информации, базами данных, контент-провайдерами,
- 60 миллиардов текстовых файлов из интернета, накопленных за последние десять лет (10 миллионов новых веб-страниц индексируется ежедневно) (Turnitin, 2016).

iThenticate сравнивает текст рукописей с материалами, которые опубликованы на том же языке. В число представленных языков входят: китайский (упрощен-

ный и традиционный), японский, тайский, корейский, каталонский, хорватский, чешский, датский, финский, французский, немецкий, венгерский, итальянский, норвежский (букмол, нюнорск), польский, португальский, румынский, сербский, словацкий, словенский, испанский, шведский, арабский, греческий, иврит, фарси, русский и турецкий. Принципиальным моментом является то, что сравнение текста происходит в рамках одного языка. Это означает, что некорректные заимствования, которые являются переводом оригинала, не могут быть обнаружены, за исключением тех случаев, когда перевод оригинала также имеется в базе данных. Таким образом, эффективность системы iThenticate определяется количеством материалов на русском языке в их базе данных (Turnitin, 2016).

### ***РУКОНТекст***

Казанский Федеральный Университет располагает лицензионным разрешением на использование системой РУКОНТекст, в основе которой лежит алгоритм, направленный на поиск семантических совпадений в тексте проверяемой рукописи. База данных состоит из 150 миллионов документов и при этом она увеличивается постоянно.

РУКОНТекст определяет процент заимствования по каждому из установленных совпадений, а также считает суммарный процент схожести, который, по сути, является эквивалентом индекса совпадений в системе iThenticate (Rukontekst, 2017).

### ***Антиплагиат***

База данных Антиплагиата включает зарегистрированные online веб-страницы, коллекцию диссертаций из Российской государственной библиотеки, коллекцию Lexpro, содержащую нормативно-правовые документы, а также коллекцию научных статей из электронной научной библиотеки Elibrary.ru.

Данный инструмент позволяет определить объем плагиата. Результатом работы системы становится документ, в котором цветом выделены некорректно заимствованные отрывки, а также приводится список оригинальных источников. Такой отчетный документ дает возможность в отдельности рассмотреть и классифицировать каждую некорректно заимствованную часть (Antiplagiat, 2017).

В момент написания данной статьи, КФУ не располагал независимым доступом к системе Антиплагиат и полностью зависел от предоставляемых Диссернетом отчетов.

### ***Проблема***

Необходимость паритета актуализирует вопрос о сопоставимости результатов проверки русскоязычных статей, полученных путем использования различных программных систем. А именно:

Существует ли корреляция между индексами совпадений, определяемыми системами iThenticate, РУКОНТекст и Антиплагиат для статей, написанных:

- a. На русском языке
- b. На английском языке

Пользователи этих инструментов проверки решают, каким должен быть порог максимально допустимых совпадений. Результаты данной статьи позволят нам обоснованно ответить на следующие вопросы:

- i. Можем ли мы выделить одну систему, которая могла бы эффективно проверять все рукописи вне зависимости от языка?
- ii. Можем ли мы определить порог допустимых совпадений для обеспечения паритета?

**Методология**

Десять статей (пять написаны на русском языке, пять – на английском), которые были опубликованы журналом «Образование и Саморазвитие», были выбраны случайным образом из недавних номеров. Решение об использовании недавних статей было связано с возможностью быстрого поиска их MS Word файлов, а не pdf, используемых для публикации.

Русскоязычные статьи (озаглавленные «Статьи 1-5») были отредактированы так, чтобы исключить англоязычные компоненты (названия, аннотации, ключевые слова на английском языке). Схемы и таблицы не исключались, но они обычно не проверяются системами проверок.

Англоязычные статьи (озаглавленные «Статьи 6-10») были отредактированы схожим образом (с исключением названий, аннотаций, ключевых слов на русском языке).

Все десять статей затем были проверены с использованием iThenticate и РУКОНТекст. В момент проверки у нас не было доступа к системе Антиплагиат. Результаты обнаруженных совпадений представлены в Таблице 1. При использовании системы iThenticate были исключены из проверки библиографии, цитаты и короткие совпадения (длинною меньше десяти слов).

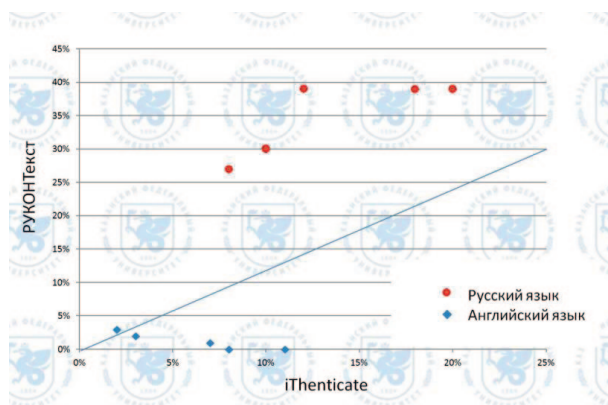
**Результаты**

Полученные нами Индекс Схожести (SI) (рассчитываемый системой iThenticate), а также Процент Заимствований (рассчитываемый системой РУКОНТекст) представлены в Таблице 1.

Как и можно было предположить, для англоязычных статей Индекс Схожести iThenticate был выше, чем Процент Заимствований РУКОНТекст, в то время как для русскоязычных статей, наоборот, Процент Заимствований РУКОНТекст был выше Индекса Схожести iThenticate. С уверенностью можно предположить, что данные результаты отражают содержание баз данных двух систем. База данных, используемая РУКОНТекст, в большей степени содержит русскоязычные материалы, не все из них представлены в базе iThenticate. Данный факт был также подтвержден нашим анализом источников, выявленных каждой из систем.

**Таблица 2. Показатели совпадений для десяти статей, выбранных для проверки**

Номер статьи	Язык	Длина (слова)	Процент совпадений	
			iThenticate	РУКОНТекст
1	Русский	2282	10%	30%
2		2775	12%	39%
3		4236	20%	39%
4		6473	8%	27%
5		6456	18%	39%
6	Английский	4141	1%	1%
7		7404	11%	0%
8		3762	5%	0%
9		3002	2%	3%
10		3566	2%	2%



**График 1.** Показатели схожести РУКОНТекст и iThenticate для десяти статей, выбранных для проверки

Более того, внимательное рассмотрение результатов по каждой статье позволило обнаружить и другие проблемные места в системе РУКОНТекст.

Например, в одной из статей РУКОНТекст выявил один источник заимствования (<http://studopedia.ru/>), которому приписал 38% от всего объема статьи. Обе статьи (проверяемая нами и определенная системой как источник) были опубликованы в 2015 году. Однако, из проверяемой нами статьи видно, что она была написана в начале 2014 года. Studopedia же не указывает конкретной даты публикации и не уточняет авторов. В данном случае совершенно справедливо можно предположить, что Studopedia некорректно заимствовала материал из статьи, опубликованной в журнале «Образование и Саморазвитие», а РУКОНТекст перепутал оригинальный источник и статью с некорректными заимствованиями.

В случае со второй статьей, в которой Диссернет выявил высокий процент плагиата, РУКОНТекст не смог определить источники заимствования и пришел к выводу, что статья не содержит плагиата.

Схожая проблема была обнаружена при проверке одного из текстов с помощью системы iThenticate – дата была указана лишь для одного текста из двух сравниваемых, а iThenticate ошибочно установил порядок выхода публикаций.

В третьем примере РУКОНТекст определил несколько источников заимствования и выдал процент заимствования для каждого из них. Суммарный процент всех заимствований получился 292.6%! На основании этого уместно полагать, что вероятность внутренних совпадений материала из разных источников, используемых системами для сравнения, достаточно высока. Другими словами, нельзя исключить, что небольшой отрывок из статьи может быть заимствован рядом других авторов, в том числе и некорректным способом, что впоследствии может привести к очень высокому проценту схожести оригинальной статьи.

Для защиты данной гипотезы рассмотрим еще одну из тестируемых нами русскоязычных статей. В ее случае РУКОНТекст выявил 302 источника, что в сумме дало выше 100% схожести. Два из значительных заимствований были из одного источника. В данном случае нам не совсем ясно, почему РУКОНТекст привел их в отчетном списке как два различных источника.

Когда эта же статья была проверена с помощью iThenticate, было выявлено 3% схожести и 2 два источника заимствования. При этом в базе данных системы каждый из этих источников был скопирован до 17 раз. iThenticate признает наличие

данной проблемы и работает над ее устранением, но в случае с РУКОНТекст ситуация не ясна. Возможно, что они не знают о существовании данной проблемы.

В свете изложенных результатов, мы не можем быть уверены в надежности показателей, определяемых системой РУКОНТекст. Обе системы (iThenticate и Антиплагиат) дают возможность визуального построчного сравнения проверяемого текста и установленных источников, а также позволяют подтвердить (или опровергнуть) наличие предполагаемого плагиата.

Несмотря на то, что наше исследование охватило лишь малое количество статей, которое не может считаться статистически значимым, мы все же можем сделать ряд выводов:

а. iThenticate может эффективно использоваться для проверки англоязычных материалов, но не для русскоязычных.

б. Очень важно проводить визуальное сравнение материалов, чтобы уменьшить вероятность технической ошибки системы.

### Статьи Диссернета

Диссернет обнаружил одиннадцать статей с некорректными заимствованиями. Вместе с тем мы решили проверить все англоязычные статьи, опубликованные в журнале «Образование и Саморазвитие» начиная с 2014 года, и в данный момент они все очень внимательно рассматриваются.

Комитет Публикационной Этики (The Committee on Publication Ethics, COPE), который является авторитетной организацией с международным признанием, сформулировал набор предписаний и руководств для редакторов, решающих проблемы, связанные с научными нарушениями (плагиат, избыточные публикации) (COPE, 2017). Журнал «Образование и Саморазвитие» в полной мере следует всем указанным COPE предписаниям по работе с плагиатом (см. Схема 2).

Очень важно понимать, что системы проверки на плагиат не являются совершенными. Они могут выявить подозрительные статьи, однако после того как возникает подозрение, человек должен перепроверить и рассмотреть все предполагаемые источники заимствования. Таким образом, процедура проверки становится длиннее, но она необходима для предотвращения несправедливых обвинений.

Наше расследование подтвердило, что восемь статей действительно содержали плагиат. Мы связались со всеми авторами и задали им вопрос, могут ли они предоставить какие-то объяснения.

Хотя некоторые авторы согласились с решением о ретракции, несколько авторов попыталось оспорить наше решение, ссылаясь на то, что в России повторные публикации одних и тех материалов не могут рассматриваться как плагиат, а поэтому они должны быть допустимы. Данная культурная особенность, касающаяся публикационной этики, подробно рассматривается Рашби в статье «Публикационная этика – моральные принципы и культурный диссонанс» (Rushby, Publication ethics – moral principles and cultural dissonance, 2017). Признавая, что взгляды на публикационную этику варьируются в различных странах, Рашби приходит к выводу, что «научная публикационная активность все чаще выходит за рамки одной страны – происходит глобальная интернационализация науки, но единая публикационная система может быть создана только в том случае, если авторы и журналы придут к общему пониманию того, что же все-таки приемлемо в научных публикациях» (Rushby, 2017).

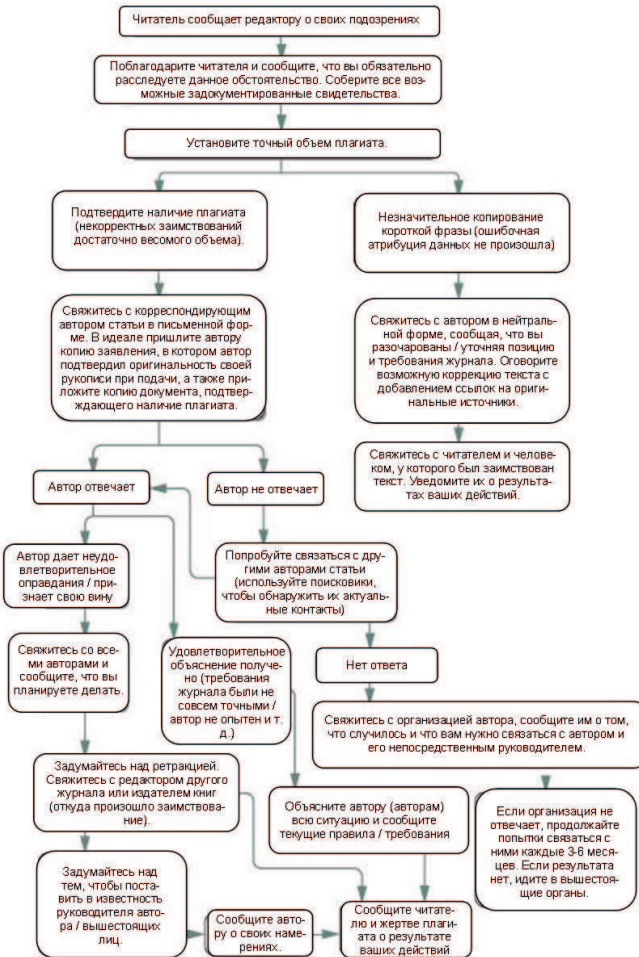
В семи случаях, обнаруженные факты указывали на то, что статьи являются избыточными (содержат самоплагиат). Выяснилось, что статьи были разосланы и опубликованы в нескольких источниках практически одновременно. В данный





## Что делать, если вы заподозрили плагиат?

### (b) Предположительный плагиат в опубликованных материалах



**Обратите внимание!**  
Инструкции для автора должны содержать точное определение плагиата, а также полную политику журнала в отношении плагиата.

Developed for COPE by Liz Wager of Sideview ([www.lizwager.com](http://www.lizwager.com)) © 2013 Committee on Publication Ethics First published 2006  
A non-exclusive licence to reproduce these flowcharts may be applied for by writing to: [cope\\_administrator@publicationethics.org](mailto:cope_administrator@publicationethics.org)

publicationethics.org

Схема 2. Указания COPE «Как действовать, когда появились подозрения в плагиате»  
(Используется в соответствии с лицензией Creative Commons AttributionNonCommercial-NoDerivs, <http://publicationethics.org/>)

момент журнал «Образование и Саморазвитие» (<http://en.eandsjournal.org/for-contributors/>) требует от авторов подтверждения того, что они подают в полной мере оригинальные материалы, которые нигде раньше не публиковались и в данный момент не рассматриваются другими журналами. Данный факт авторы подтверждают в своих сопроводительных письмах к главному редактору, когда они отправляют свои статьи. Раньше такого подтверждения журнал не требовал, и мы точно не знаем, предъявляют ли схожее требование другие журналы, которые были определены как источники заимствования.

Вследствие этого было решено выпустить по каждой из семи статей «постановление о признании избыточной публикации». Эти постановления привлекают внимание к тому факту, что схожая версия данной статьи была опубликована где-то еще.

Журнал «Образование и Саморазвитие» будет принимать такие же действия по отношению к любым другим напечатанным статьям, в которых будут обнаружены доказуемые признаки плагиата или самоплагиата (избыточности).

### Выводы

Из одиннадцати статей, выделенных Диссернетом, мы произвели процедуру ретракции для четырех из них, в которых системы проверки обнаружили недопустимый объем некорректных заимствований. В оставшихся семи статьях места, выделенные системами проверки как заимствования, являются результатом одно-временной подачи статей в те журналы, где не требуется подтверждение того, что статья ранее не печаталась и не рассматривается в данный момент для печати в других журналах. В виду этого в их адрес мы выпустили постановления о признании избыточной публикации.

Объявления с данными постановлениями появятся в журнале после данной редакторской заметки.

Мы продолжаем работать над проверкой всех статей, выпущенных журналом за последние три года (примерно 400 статей). На данный момент 18 англоязычных статей проверено, процедура ретракции была проведена для одной из них (в виду обнаруженного плагиата).

Кроме того, мы внимательно проверили все англоязычные статьи, которые в данный момент находятся на рассмотрении для публикации. В виду обнаруженного плагиата одну статью мы отклонили, а одну отправили на доработку, чтобы автор мог полностью исключить некорректные заимствования.

Все процедуры были проведены в соответствии с предписаниями комитета CORE (2017).

Менее 30%	275	72.94%
30-39%	60	15.92%
40-49%	18	4.77%
Более 50%	24	6.37%
Total	377	

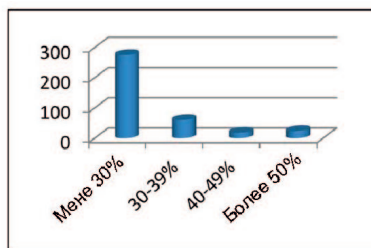


Схема 3. Индекс схожести (SI) рукописей, поданных в ВJET (за 2013 год)



Несмотря на то, что процент плагиата в материалах, присылаемых в журнал «Образование и Саморазвитие», может на первый взгляд показаться тревожаще высоким, он на самом деле сопоставим с опытом других научных журналов, которые используют системы плагиат-проверки. Например, уважаемый журнал *British Journal of Educational Technology* в течение 2013 года с помощью системы iThenticate обнаружил, что 11% всех присланных к ним рукописей получили индекс схожести свыше 40%!

### **Подумайте о редакторе**

Решение проблем, возникших в результате некорректного научного поведения, не является приятным занятием. Оно отнимает достаточно много времени (в среднем уходит 3-4 часа на каждую статью). Кроме того, это очень угнетает. Думаю, немногие из редакторов наслаждаются своей обязанностью сообщать авторам о том, что их статья будет подвергнута ретракции. Несмотря на то, что мы вряд ли будем сочувствовать серийным плагиатчикам, большинство случаев некорректного заимствования возникают в результате того, что авторы не опытные, не знают правил публикационной этики, иногда они находятся под давлением быстро опубликовать статью или просто сбиты с толку журналами, которые очень медленно принимают решения по принятию статей. Научные руководители должны делать больше для того, чтобы их студенты научились проверять свои работы до отправки статьи в журнал.

### **Признательность**

Автор хотел бы выразить благодарность в адрес Лейсан Каюмовой за ее помощь в проверке статей с помощью системы РУКОНТекст, в адрес Ольги Кирилловой за ее помощь и поддержку, и Дарье Ханоланен за перевод статьи на русский язык.

### **Список литературы**

- AntiPlagiat* (2017) *About the system*. Accessed March 1, 2017 at: <https://www.antiplagiat.ru/Page/About>
- Apollodorus (1921). *Bibliotheca (The Library) (in Greek)*. With an English Translation by Sir James George Frazer, in 2 Volumes. Cambridge, MA; London: Harvard University Press; William Heinemann Ltd.
- COPE (2017) *Guidelines*. Committee on Publication Ethics. Accessed February 28, 2017 at: <http://publicationethics.org/resources/guidelines>.
- Rukontekst (2017) *Search for plagiarism and document analysis Rukontekst*. Accessed March 1, 2017 at: <https://text.rucont.ru/>
- Rushby, N.J. (2016) *Transforming a journal from an institutional focus to an international publication: a work in progress!* Paper at the Project 5-100 Seminar, Moscow, 11th November 2016.
- Rushby, N.J. (2017) *Publication ethics – moral principles and cultural dissonance*. *EON, Editorial Office News*. (June) pp4-9. <http://dx.doi.org/10.18243/eon/2017.10.5.3>
- Turnitin (2016) *Largest scholarly comparison database*. Oakland, CA, USA: Turnitin LLP. Accessed February 14, 2017: <http://www.ithenticate.com/content>