

## Предложения по созданию технического средства многоязычного анализа источников научных сведений

В.А. Жмудь<sup>1,2</sup>, А.В. Ляпидевский<sup>2</sup>

<sup>1</sup>ФБОУ ВО «НГТУ», <sup>2</sup>ПАО «Новосибирский институт программных систем»

*Аннотация.* Российская наука (и не только) поставлена в условия, когда научные публикации полагаются самым важным ее продуктом. Это приводит к потере добросовестности некоторыми российскими учеными и преподавателями в массовом масштабе (многие сотни фигурантов). Этическими воздействиями данную проблему решить не представляется возможным, и, по-видимому, это и не требуется. Принятые критерии оценивания научных публикаций не по содержанию и новизне, а по месту их опубликования, явно устарели, что неопровержимо доказывается на многих примерах. Следовательно, актуальны иные средства противодействия этому потоку псевдонаучных публикаций. Вместе с тем, актуальна задача оперативного создания обзоров на самые различные темы, делаются эти обзоры, как правило, недостаточно качественно, трудоемкость их написания при ответственном отношении весьма высока, порой получается, что пока обзор будет готов, он уже устареет. Нейронные сети, нейронные компьютеры и средства анализа больших данных предлагают инструментарий для решения этих проблем в совокупности и максимально эффективно. Обосновывается задача создания такой системы и даются основные указания на самые главные технические требования к ней.

*Ключевые слова:* большие данные, нейронные сети, нейронные вычисления, поисковые системы, автоматический перевод, автоматический анализ текста, наука, научные публикации, финансирование, плагиат, гранты, цитирование, индекс Хирша, теория автоматического управления

### I. ВВЕДЕНИЕ

Публикации на тему фальсификаций в науке не безопасны, они чреватые различными оргвыводами в отношении автора. В частности, автору этих строк указывали на его собственную не безупречность, полагая, что это достаточный и окончательный аргумент для отказа разработки этой темы.

Автор считает, что безупречных людей нет, что он может во многих случаях доказать. Если наличие каких-либо вин в прошлом и настоящем является основанием не замечать любые вины всех остальных, все человечество должно в этом случае молчать при виде любых видов нарушений справедливости и нравственности. В этом случае нравственность не совершенствовалась бы, начиная с каменного века. Персональные вины автора перед кем бы то ни было не являются основанием для превращения его в безгласное и послушное орудие чужой воли. Если обвинения против автора могут быть доказаны, автора следует, по-видимому, наказать за это, при условии, что ранее он не был наказан. Если же обвинения не доказаны (не доказуемы), они голословны. Высказывание недоказанных обвинений регулируется ответственностью за моральный ущерб, но реагировать на них не обязательно.

Автор должен признать, что в отношении этики научных публикаций также не безгрешен в тех сферах, которые, по его мнению, не являются серьезными, хотя современным критериям высшей этики подобные действия, безусловно, не соответствуют. В частности, излишнее самоцитирование, на взгляд автора является не вполне этичным, но это не является ни преступлением, ни злоупотреблением полномочий. Опубликование тех же материалов

на нескольких языках – это вообще крайне тонкий вопрос. Были времена, когда никому бы и в голову не пришло, что это не этично. И лишь совсем недавно некоторые журналы стали обращать на это внимание, включили в свои требования и такое требование, чтобы ранее опубликованные материалы на любом языке не присылались на опубликование в этом журнале, поскольку таковая публикация была бы вторична, и приравнена ими к автоплагиату. Далеко не все журналы до настоящего времени этот критерий взяли на вооружение. В частности, еще пару лет назад даже руководство весьма высокой публикационной коллегии отвечало, что нет ничего проблемного в том, что к публикации предлагаются материалы, ранее опубликованные на другом языке. Нынче это уже не так. Поскольку этические нормы изменяются, невозможно соответствовать сегодняшним нормам в своих всех без исключения ранее выпущенных за всю жизнь публикациях. Тем не менее в настоящее время в дискуссиях об этичности подобного опубликования автор данной статьи стоит на позиции, что опубликование на другом языке, по-видимому, действительно, является автоплагиатом, кроме случаев, когда статья попадает в двуязычный журнал, поскольку в этом случае это формально – одна и та же статья, выходящая в едином месте в одно и то же время, автор не пытается делать вид, что это две разные публикации. При этом также следует учитывать и то, что такие две публикации одного материала на двух языках в двух видах одного и того же журнала не являются двумя публикациями, и это принципиально. Следовательно, автор не имеет морального права давать в списке своих трудов две эти публикации как разные, они должны даваться либо в едином пункте, либо должна

даваться только одна из двух ссылок. Этого не делается до сих пор многими авторами при формировании своих списков трудов для различных рейтингов и иных действиях с целью поднятия своего уровня, подтверждения квалификации, получения степеней, должностей, званий и так далее. Отметим, что и база данных РИНЦ подобные публикации зачастую формирует в списке как две различные публикации, что вроде бы освобождает автора от ответственности за такой способ составления списка – если база данных извещает, что это две разные публикации, тогда и автору спорить с этим проблематично. Тем более, что каждая из таких публикаций имеет свой индивидуальный перечень цитирования. Если бы можно было эти цитирования автоматически слить, это было бы правильно, удалять эти цитирования было бы несправедливо по отношению к автору.

Вернемся к проблеме.

## II. СУТЬ ПРОБЛЕМЫ

Когда из средств массовой информации мы узнаем о том, что, оказывается, в среде российских ученых имеется множество плагиатчиков, стоит ли удивляться? Гонка за количеством публикаций со стороны работников науки и высшего образования организована сверху, кто в ней не участвует, тот проигрывает. Следовательно, участвовать вынуждены все.

Если объявляется во многих инстанциях конкретно для работников науки и образования, что они лишь в том случае могут рассчитывать на премии и иные виды финансовых поддержек, включая даже очень и очень большие величины, если они будут иметь достаточное количество публикаций, то стоит ли удивляться, что эти работники будут прилагать все виды усилий для получения как можно большего количества этих самых публикаций? А если они будут видеть, что их коллеги, осуществившие опубликование малозначительных и вовсе незначительных сведений, получили уже достаточное вознаграждение, в этом случае лишь очень ленивые и (или) болезненно-порядочные люди не сделают ничего для того, чтобы увеличить количество своих публикаций.

Надо было бы посмотреть статистику по количеству научных публикаций до того, как руководители научных и образовательных сфер изменили свою политику финансирования, подчинив ее полностью публикационной активности (что является одним из индикаторов планирования развития науки). По указанным данным следует построить некоторые статистические прогнозы количества публикаций на последующие годы. Все, что превышает эти прогнозы – это следствие изменения политики. То есть весь прирост появился ради прироста как такового. Даже до начала указанных призывов словом и рублем, далеко не все публикации осуществлялись ради

того, чтобы сделать полученные сведения достоянием научной общественности. Львиная доля научных публикаций делается, делается и будет делаться ради возвышения их авторов – это путь к диссертациям, путь к славе, деньгам, и не малым. В этом смысле все ученые, по отдельности и вместе взятые, мало чем принципиально отличаются от пресловутой Ольги Бузовой [1], которая из популярности делает деньги. Этим же самым занимаются практически все ученые – достигают популярности, чтобы делать деньги. Или достигают популярности, чтобы добиваться известности, из которой делать деньги, или делать научные степени, из которых опять-таки делать деньги. Или делать остепененных учеников, что также повышает известность и приносит деньги.

Если руководством покупается публикационная активность, значит, ученые будут производить публикационную активность. Если министерством покупается цитируемость, значит, все, кто понимает механизм этой формальной машины, будет работать для ускорения персонального механизма по увеличению персональных показателей цитируемости. Если в другом месте это все продается, значит те, у кого это покупают, будут это покупать сами при условии, что **купить можно дешевле, чем продать**. Если за тысячу евро в одном месте вас включают в авторский коллектив, и за это вы в другом месте получите премию в полторы тысячи, для вас сделка выгодная. Поскольку не все измеряется только наличными деньгами, есть и другие формы платежей, то даже если премия будет меньше стоимости вхождения в публикацию, другие дополнительные преимущества могут покрыть и перекрыть разницу, следовательно, интерес сохранится.

В статьях [2–7] мы обсуждали проблему мошенничества на рынке научных публикаций. Разумеется, некоторые даже должностные лица прочитали эти статьи, возможно, даже сделали для себя некоторые выводы. Во всяком случае та самая организация ООО «Международный издатель», которая в открытую занимается торговлей публикационной активностью, в том числе продает места соавторства в уже готовых написанных статьях, подверглась впоследствии критике в других средствах массовой информации, в статьях на эту же тему и со ссылкой на сайт этой же мошеннической организации [8]. Сказано, в частности: «сайт 123mi.ru утверждает, что предоставил авторские права более чем 10 000 исследователей, продавая слоты на рукописи, написанные другими, которые уже были приняты журналами» [8].

Поэтому автор вправе надеяться, что его вклад в разоблачение этих махинаций не прошел напрасно.

И еще интересный отмеченный факт: «Комиссия РАН ... самостоятельно исследовала проблему. ... Сравнивая списки авторов, которые были опубликованы дважды или более, комиссия также выявила очевидные случаи «неясного авторства» - академиков, которые были авторами одной версии документа, но не другой... Летом прошлого года комиссия попросила 541 журнал отозвать в общей сложности 2528 статей. В своем промежуточном отчете комиссия пишет, что 390 журналов уже ответили на запрос, 263 из которых согласились отозвать все подозрительные документы; другие согласились отозвать некоторые из выделенных документов, но не другие, или дали законные причины, по которым документы не следует извлекать».

То есть сегодня реальность несколько изменилась, в чем-то в лучшую сторону, а в чем-то в худшую. А именно: уже очевидно, что факт мошенничества и манипуляций в сфере научных публикаций приобрел невиданный размах. Очевидно, что с ним следует бороться, и с ним уже начинают бороться. Но автору этой статьи очевидно также, что эта борьба не будет легкой.

### III. УРОВЕНЬ ТЕХНИКИ И ПОСТАНОВКА ЗАДАЧИ

Возможно, если бы кто-то из ученых СССР, создающих передовые технологии в интересах государства в целом (Королев, Курчатов, Кошкин, Калашников – это только на букву «К», их легионы), узнал о сегодняшней ситуации, он бы не только удивился, но и возмутился бы этому. Но мы не беремся утверждать за всех абсолютно ученых, работавших в СССР, что все они обладали крайне высокими моральными принципами. Равно я не берусь утверждать, что все ученые, работающие в России сегодня, обладают крайне низкими моральными принципами. Однако, я полагаю, что мало кто будет оспаривать тот факт, что в годы процветания СССР такого безобразия с фабрикацией псевдонаучных публикаций не было, а в наши годы уже невозможно отрицать, что оно имеет место. Следовательно, виновата не человеческая природа конкретных ученых, а виновата в процессе происходящего та нравственно-финансовая среда, которая создана сверху, не будем вдаваться в поиски причины ее создания, ведь не столь важно, допущены эти действия умышленно или неумышленно, сколь важно понять, каковы эти действия, и что необходимо сделать для устранения этого последствия.

А действия были самые простые.

1. Потеря доверия к собственным ученым, к собственным экспертам от науки и техники. Причин для этого я не вижу, кроме как заискивание перед западом тех людей, которые, к счастью, я надеюсь, уже от власти отстранены.

2. Поскольку свято место пусто не бывает, перестав доверять отечественным экспертам, эти лица свой кредит доверия перенесли на экспертов западных, поначалу не понимая ни принципов действия западных экспертиз, ни того факта, что групповые интересы действуют отнюдь не только в местечковых кружках отечественных не самых лучших представителей научных организаций, но это явление гораздо шире, и на западе от него отнюдь не свободны.

3. Постепенно сформировался крайне формальный критерий качества публикаций, который полностью независим от их содержания, а основывается лишь на месте их опубликования.

4. Взят на вооружение дополнительный критерий, состоящий в количестве цитирований автора и в достаточно искусственно придуманном показателе, называемым индексом Хирша [5].

Следовательно, на первом этапе перед российскими учеными (сотрудниками академических институтов и преподавателями университетов, которых в дальнейшем будем иметь в виду в этом списке, но кратко называя их «учеными») были поставлены следующие задачи.

1. Публиковать ежегодно как можно больше статей.

2. Надеяться на то, что эти статьи будут как можно чаще цитироваться другими авторами.

3. Надеяться, что наиболее часто цитируемые статьи будут цитироваться наиболее часто (что необходимо для быстрого роста индекса Хирша).

Кое-кто знаком с утверждением Мичурина: «Мы не можем ждать милостей от природы, взять их у нее – наша задача» [9]. Соответственно, они поняли это так: «Не следует ждать, пока тебя процитируют, обеспечить это – более эффективный путь решения проблемы». Возникли многочисленные клубы взаимного цитирования. Суть их деятельности ясна из названия.

Увеличение количества статей было обеспечено различными путями.

В том числе:

1. Активное опубликование незрелых результатов (наименьшее зло).

2. Опубликовано уже опубликованного на другом языке (до недавнего времени это не считалось неэтичным и сейчас многими редакционными коллегиями неэтичным не считается).

3. Опубликовано ранее изданных результатов (кстати, слово «препринт» означает предварительное опубликование, поэтому повторное опубликование того, что более подробно и в более сыром виде опубликовано в препринте, никогда не считалось зазорным действием).

4. Взаимное включение в соавторство (ты включаешь меня, я включаю тебя, в результате у нас обоих больше публикаций).

5. Публикация якобы научных обзоров (которые на самом деле представляют собой преимущественно куски глав диссертаций, книг или учебников этого автора, слегка дополненных последними сведениями о собственных новых результатах, да и то не всегда).

6. Публикация собственных размышлений «на тему», не подтвержденных никакими научными исследованиями или хотя бы какими-то действиями, напоминающими (хотя бы отдаленно) какие-то научные методы (подчас даже поиск по источникам при этом был крайне скуден).

7. Беззастенчивый плагиат.

8. Все прочие методы, которые мне, вероятно, не известны, но очевидно существуют.

Именно эти методы дали тот небывалый рост активности, который был, по-видимому, отмечен как положительное явление, но таковым вовсе не являлся.

Действительно, попытаться заставить настоящего ученого (именно настоящего) публиковаться в нынешнем году более активно, чем в прошлые годы, равносильно призывам к беременной женщине рожать не через девять месяцев после зачатия, а через все меньшие и меньшие сроки, например, в этом году через семь месяцев, на будущий год через четыре месяца, а в следующий год ежемесячно.

То, что женщина не кошка, понимают почти все. То, что слоненка вынашивают дольше, чем мышонка, понимает гораздо меньшее количество людей. То, что хорошую статью нельзя писать по расписанию, ежеквартально или ежемесячно, боюсь, не понимает никто из тех, кто допущен руководить большой отечественной наукой. Если бы они это понимали, они никогда не сделали бы такой показатель, как «количество статей» хотя бы мало-мальски значимым.

**Теория автоматического управления** утверждает, что управление в целом состоит в следующем.

1. Формируется цель (в данном случае требуемый показатель).

2. Измеряется выход объекта (реальное значение показателя) и сравнивается с этой целью.

3. Разница, называемая ошибкой управления, усиливается и с соответствующим знаком (так, чтобы в целом в петле был отрицательный знак) и с большим коэффициентом подается на вход объекта, то есть на то самое место, в зависимости от воздействия, на которое может изменяться выход объекта.

В переводе на язык публикационной активности это означает следующее.

1. Задается норматив. Например, каждому ученому следует писать не менее трех статей в год.

2. В отношении каждого ученого берется разница того, что требуется и что достигнуто. Например, для ученого, опубликовавшего только две статьи, ошибка будет равна минус единице. Для ученого, опубликовавшего четыре статьи, ошибка будет плюс единица.

3. Вводится регулятор. Например, ПРНД - показатель результативности научной деятельности.

4. Формируется регулятор – вычислитель воздействия на объект. Если ПРНД ниже требуемой планки, на объект вводится отрицательное воздействие, если он выше – вводится положительное воздействие. На языке публикаций это означает, что все ученые, опубликовавшие менее трех статей, наказываются или предупреждаются о будущем наказании, если эта ситуация не будет исправлена. А все, опубликовавшие больше трех статей, поощряются.

После того, как такая система управления сформирована, она работает. Достаточно только изменять «задание», то есть на следующий год можно устанавливать уже не три статьи в год, а четыре, или пять, и так далее.

В результате это сообщество будет полностью управляемым, то есть оно достигнет именно тех показателей, которые сформированы в виде задания. Не следует этому удивляться, поскольку такая система может быть либо работоспособной, либо неустойчивой (то есть развалиться). Пока система не разваливается, она работает. Пока она работает, результат достигается. То есть мы получим именно то, к чему шли, а именно: те, кто способны обеспечивать по три публикации в год, вне зависимости от способа решения этой задачи, эту задачу выполняют. Какой ценой – это другой вопрос. Критерий этики не заложен, следовательно, задача будет выполнена в том числе и за счет несоблюдения этики научных публикаций. Если включить в управление критерий этики, он будет соблюдаться, но только в этом случае. Если задачу нельзя выполнить, не нарушая этических норм, система либо не будет работать, либо развалится (см. выше). Но пока этические нормы в задачу не входят, нарушение их является не препятствием, а средством достижения цели. Те сотрудники, кто это не способны решать поставленную задачу, либо этого делать не хотят (в том числе по этическим соображениям), будут из процесса выталкиваться, так как они будут наказаны. Лица, которые, с одной стороны, не могут решить задачу без нарушения этических норм (например, без повторного опубликования), и которые при этом не готовы поступаться этическими нормами, попадают в класс некомпетентных (*sic!*). Те же, кто способен



перекрывать поставленные цели (пусть даже ценой нарушения этических норм), будут постоянно поощряться, что даст им дальнейшие стимулы для увеличения этого поставленного целевого индикатора, то есть они будут повышать количество статей неуклонно (тем более за счет всех тех способов, которым они вынуждены были обучиться).

Если требуемые показатели неуклонно повышать, то каждый коллектив также будет неуклонно повышать эти результаты, причем, когда все честные пути будут исчерпаны, то в целом коллектив и каждый сотрудник в частности будут поставлены перед дилеммой – либо признать свою несостоятельность, либо решать задачу не самыми этическими путями. Что и происходит.

Получили следующее: количество публикаций резко растет, а фактической эффективности научной деятельности руководство не увидело.

Следовательно, цели пришлось скорректировать. Но поскольку собственным экспертам мы не доверяем (см. выше), пришлось ввести фактор иностранных баз данных и цитирование из этих баз, то есть пресловутые *Scopus* и *Web of Science*. Если у вас публикации относятся к этим базам данных, мы их будем оценивать высоко, все остальное – малосущественный шум.

Поскольку созданная система автоматического управления работает, результат не заставил себя ждать. Все организации стали отчитываться ростом указанного типа публикаций.

А эффекта от науки не прибавилось.

Следовательно, пришлось еще больше уточнить задачу. Теперь обратились к квартильности журналов. Все журналы по любой тематике можно ранжировать по количеству цитат его статей в году. Этот список далее можно разбить на четыре части, в высшей части те, кто цитируется наиболее часто (первый квартиль, Q1), в самом низу те, кто цитируется наименее часто (четвертый квартиль Q4), в середине два промежуточных квартиля, второй и третий. На этом основании объявлено, что два высших квартиля – это действительно ценные публикации, а два низших квартиля – это ничего не значащие публикации, для которых возник термин «мусорные» [7].

«Наука превращалась в фантазмагорию, опираться на научные публикации стало невозможно. Любая глупость могла быть безо всякого рецензирования напечатана в журнале, называющем себя научным» [10]. В этой статье справедливо отмечается, что настоящие ученые стали заложниками псевдоученых, поскольку нечестными способами в этой сфере можно достичь быстрее и эффективнее гораздо больших формальных показателей, нежели честными, поэтому честные и порядочные ученые

постепенно вытесняются на край науки, на безденежную периферию, тогда как заправляют всем мошенники и манипуляторы в сфере публикационной активности. Столь же справедливо отмечается: «Пока не поменяется ситуация с репутационными механизмами, пока не начнут приниматься вменяемые управленческие решения, вряд ли что-то изменится» [10]. Справедливо и замечание о том, что даже порядочный автор вместо того, чтобы опубликовать одну ценную статью, размазывает ее результаты на множество публикаций, чтобы достичь велемого сверху показателя. Ученые вынуждены ориентироваться не на истинно научные ценности и показатели, а на то, насколько понравится данная статья иностранным рецензентам, они ориентируются на мнения иностранных экспертов вне зависимости от интересов развития отечественной науки, в социальных науках они вынуждены подстраиваться даже под социальные критерии, не являющиеся ни каким боком научными (например, мнение о том, что осознанный выбор пола доказывает более высокий социальный статус общества, является обязательным для специалистов в этой сфере, те, кто его не разделяют, не смогут опубликоваться в журналах, входящих в указанные базы данных).

Вот что сообщил в свое время Ливанов: «И мы выяснили, что некоторые вузы пытаются быстро нарастить объем своих научных публикаций, но публикуя результаты в журналах, которые не являются престижными. По существу, в профессиональном сообществе это называется «мусорными публикациями» [10]. Таким образом, качество публикации оценивается не ее содержанием, а тем, где она опубликована. Даже самую важную и содержательную статью, если она опубликована в журнале, не входящем в эти квартили, и тем более не входящем в эти базы вообще, следует считать научным мусором, но даже самую пустейшую статью, которая опубликована в первых двух квартилях указанных баз данных, следует считать весомым научным результатом данного автора и коллектива, где он работает.

Образно говоря, если вы напишете «Позвони родителям» на воротах собственного гаража, то это – гадость, а надпись даже самого непристойного содержания, помещенная на самом видном месте (например, за деньги) при условии, что это место организовано чисто на коммерческих основах, и так сложилось, что другие люди, не имеющие никакого отношения к этому коммерческому проекту, почему-то для себя решили, что это место само по себе является весьма значительным, является проявлением высокой нравственности.

Цитируемый журнал – важный и научный, не цитируемый – не важный и не научный. Цитируемый – означает читаемый. Больше всего

читают научно-популярные журналы, в которых ничего нового не сообщается.

Что такое цитируемость? Это не более того, как популярность. См. по этому поводу [1]. Популярность (слово от латинского от *populus* – народ) – это не более, чем общедоступность, опубликование на самом посещаемом месте, на месте, предоставленном для всеобщего обозрения. И совершенно не важно, что имеется некий клуб авторитетов, которые по тем или иным принципам решают, кого допустить к этому опубликованию, а кого не допускать. Этот клуб – отнюдь не научное сообщество, они не имеют никакого обязательства ни перед мировой наукой, ни перед наукой Российской Федерации тем более. В подкрепление этой мысли приведу цитату из письма, направленного мне председателем весьма уважаемым и авторитетным человеком, называть которого без его согласия я не могу, но утверждаю, что и для всех членов ВАК его мнение является существенным.

«Что касается рекламы научных публикаций за деньги, то в условиях капиталистической экономики, которая у нас почему-то называется «рыночной», это – обычное предложение услуг, которые являются совершенно законным. Доказать состав преступления, предусмотренный ст. 159 УК РФ могут только сотрудники правоохранительных органов при наличии заявления от потерпевшей стороны, т.е. от конкретного обманутого человека. Значительная часть отечественных научных журналов, и практически все зарубежные издания, это – частные коммерческие проекты. Кстати, базы данных цитирования, включая *Web of science*, *Scopus*, а также РИНЦ (Российский индекс научного цитирования) – это тоже частные компании. С уважением, NN\*».

С другой стороны, действует следующее правило (пункт 5): в перечень ВАК входят «Издания, текущие номера которых или их переводные версии входят хотя бы в одну из международных реферативных баз данных и систем цитирования *Web of Science*, *Scopus*, *PubMed*, *MathSciNet*, *zbMATH*, *Chemical Abstracts*, *Springer* или *GeoRef* считаются включенными в перечень по отраслям науки, соответствующим их профилю. К данным изданиям Требования не предъявляются. (в редакции Приказа Министерства образования и науки Российской Федерации от 12.02.2018 г. N 99)» [12].

Тем самым, мы имеем, с одной стороны, е мнение, с которым приходится согласиться, о том, что **публикация в указанных базах данных вполне законным путем может быть осуществлена просто за деньги, следовательно, без какого-либо научного вклада лица, заплатившего за эту публикацию, при этом совершенно понятно, что если данное лицо не считает себя обманутым, то факта**

**мошенничества здесь нет**, и эта ситуация включает в себя возможность того, что данное лицо не только не является полноценным автором этой работы, но может и вообще ни по каким параметрам не соответствовать понятию «соавтор», то есть оно может вообще не принимать участия ни в сборе материалов, не в подготовке публикации, ни в чем вообще кроме выплаты денег. С другой стороны, ВАК полагает, что опубликование в указанных базах данных, которые являются чисто коммерческими проектами, дает указанному лицу, которое может не быть автором этих публикаций, вполне **законное право на получение научной степени и научного звания** при проведении соответствующих процедур [12].

**Тем самым торговля научными степенями и званиями получается полностью узаконенной.** А поскольку указанные степени и звания позволяют получать государственное финансирование и занимать весьма высокие посты в науке и в руководстве, следовательно, вот вам к вашему вниманию вполне законный коммерческий путь для получения званий, степеней, получения грантов и покупке должностей. Поневоле вспоминается диалог из комедии Бомарше: «Как жаль, что у нас должности продаются! – Да, вы правы! Было бы гораздо лучше, чтобы их давали даром!» [13].

К сожалению, в настоящее время эффективного средства выявления недобросовестных публикаций не существует.

Приведу пример из своей практики. Студент принес дипломный проект. Он был идеален. Поиск с помощью базы данных «Антиплагиат» [14] не выявил никакого плагиата, оригинальность текста составила 100%.

Вместе с тем было очевидно по известному уровню данного студента, что он не мог в указанные сроки подготовить указанный дипломный проект, хотя, строго говоря, утверждать это категорически также было невозможным, поскольку студент брал академический отпуск на год, а за год может случиться всякое, возможно, он прошел курсы сверхэффективного тренинга. Он утверждал, всяком случае, что он нанял репетитора-руководителя и воспользовался возможностью провести все исследования в партнерской организации, что опровергнуть не представляется возможным, да ведь и преподаватель – это не следователь. Студент сказа, нет причины не верить, хотя сомнения были большие и вполне обоснованные.

Пришлось действовать неформально. Не могу раскрывать способ, которым я воспользовался, по той причине, что, если мою статью прочитают те, кто торгует такими изделиями, как «диплом под ключ» или «курсовая работа под ключ», а, судя по их умению, я не удивился бы, что они могут сделать и кандидатскую и докторскую диссертацию под ключ, но способ отыскания

первоисточника я нашел, и это мое ноу-хау. Если соответствующие специалисты заинтересуются сотрудничеством со мной, я этот способ им раскрою не публично. Таким образом, было выявлено, что «дипломный проект» представляет собой перевод двух относительно объемных англоязычных статей: одной обзорной, другой оригинальной с описанием методов исследований и результатов. Статьи были написаны за три года до получения студентом задания на дипломную работу и опубликованы на английском языке в трудах одной конференции, состоявшейся в Таиланде.

Все иллюстрации были тщательно отредактированы, в них все англоязычные надписи были заменены на русскоязычные, текст был тщательно переведен, не с помощью автоматического переводчика, а, как минимум, окончательно был отредактирован корректором. Перечень литературы был выправлен по требованиям отечественных ГОСТ. Все было сделано так, что не подкопаешься.

Безусловно, студент был отчислен, но я не берусь утверждать, что он не поступил в другой университет и не защитил там этот диплом, скорее всего, если это так, то он получил оценку «отлично».

На этом основании можно сколько угодно утверждать, что я занимаюсь не своим делом, когда беспокоюсь о плагиате или публикую статьи на эту тему, но я имею все основания утверждать, что именно эта деятельность является одной из основных профессиональных видов деятельности преподавателя, заведующего кафедрой, который обеспечен уровнем современного преподавания в стране, в которой он прожил всю жизнь и которую не собирается покидать даже при том условии, что с процессами в сфере высшего образования не всегда можно согласиться.

Процессы идут, торговля в этой сфере процветает, в Томске на улице Ленина, где располагаются три основных университета – ТУСУР, ТПУ и ТГУ, имеются точки продажи микро наушников телесного цвета и других еще более изощренных технических средств для обмана преподавателя, также в сети интернет вы легко найдете сайты с предложениями за деньги подготовить вам любые квалификационные работы – курсовые проекты, а также дипломный проект по специальности и магистерскую или бакалаврскую диссертацию.

Поскольку средство «Антиплагиат» [14] не действенно, что я утверждаю из собственного опыта, следовательно, нужны другие средства борьбы с этим явлением. Возможно, более эффективные средства поиска существуют, мне они не известны.

Таким образом, для фактически всех преподавателей, руководителей университетов всех уровней, для всех редакционных коллегий, для всех организационных комитетов всех

конференций необходимо средство анализа всех опубликованных материалов на всех языках с целью выявления плагиата. С учетом возможностей слабого редактирования текста на окончательном языке с целью уничтожения полной идентичности (замена синонимов, перестановка мест и иные способы, не все из которых широко известны), требуется не простой формальный способ сопоставления текстов, но более интеллектуальный, который может быть выполнен лишь с применением современных систем искусственного интеллекта, то есть анализ не по отдельным словам и совпадению слов из разных языков, а анализ содержания как такового с сопоставлением именно содержательной части, что в десятки раз сложнее, но при сегодняшнем уровне техники эта задача разрешима.

Действительно, все мы хотим, чтобы новые публикации были действительно новыми, а не перекрашенным старьем. Научным рецензированием должны заниматься ученые. Наряду с тем, что они должны оценивать действительный уровень предлагаемого достижения, они должны еще и проделывать весьма трудоемкую работу по анализу на плагиат. Просто загрузить файл в систему [14] и дождаться отчета – не эффективно. Эта система имеет пути обхода. Кроме того, подобные услуги на высоком качественном уровне не бесплатны, следовательно, проверить качество публикации не представляется возможным. В дополнение к этому отметим, что далеко не все журналы предоставляют открытый доступ к своим публикациям, следовательно, если, например, автор отправит одну и ту же статью в несколько журналов одновременно (что случается), то эти журналы не смогут выявить этого факта. Если же существовала бы система, в которую можно было выкладывать не только статьи в открытый доступ, но также и статьи без открытого доступа какому бы то ни было читателю, но с возможным доступом к информационной системе на предмет совпадения исследуемого текста с этим текстом (частично или полностью), то есть пользователь мог бы лишь узнать, новый это текст или не новый, но не мог бы ознакомиться с тем текстом, с которым шло сравнение, поскольку он пока еще закрыт для доступа, это был бы совершенно новый шаг по выявлению плагиата. К этому следует добавить многоязычность, в идеале, по крайней мере, сравнение по основным европейским языкам, желательно присовокупить также японский, китайский, корейский, дальше по мере развития системы и другие языки, менее популярные в технической литературе, для нас также важен русский язык и языки стран бывшей СССР, а также стремительно развивающихся стран – Монголии, Индии и так далее. Рецензировать статью должен авторитетный ученый, а авторитетным является тот ученый, который активно работает в этой сфере сегодня,

но если он активно работает в науке, то очень жалко его времени, если оно расходуется на простой поиск ответа на вопрос, является ли данный материал оригинальным, или же он является заимствованным.

Экономить труд такого ученого – святая задача специалистов по информационным технологиям.

К этой проблеме примыкает еще ряд задач. Крайне актуально составление достаточно эффективного результата научного поиска по всем источникам (журналам, патентам, трудам конференций и сайтам продаж) для определения уровня техники. Эту задачу приходится решать, как минимум, в трех случаях.

Во-первых, при выборе темы аспирантских или докторантских исследований, при обосновании проекта, подаваемого на конкурс РФФИ, РНФ и т.п., при написании введения кандидатской и докторской диссертации. Убежден, что подавляющее большинство подобных обзоров недостаточно полно, поскольку работа эта чрезвычайно трудоемка, полноценный обзор может быть по трудоемкости больше, чем подготовка полноценной докторской диссертации, хотя формально – это лишь ее часть.

Во-вторых, анализ поданных в те же самые организации проектов, претендующих на гранты. По сути, эксперт должен проделать ту же самую работу, которую проделал автор проекта на грант, но при этом эксперт, вероятнее всего, не является столь большим специалистом, как автор проекта, поскольку автор выбирает себе тему сам, а эксперту эту тему дают на экспертизу.

В-третьих, подобный анализ необходим при фактическом выполнении любой НИР и ОКР, вне зависимости от источника финансирования, даже если это финансируется не по гранту, то есть исполнитель не предпринимал подобного поиска с целью получения финансирования, а задача пришла извне.

Кроме того, имеется задача отыскания наилучшего эксперта в конкретной тематике работ. Эта задача также не столь проста, но она может быть решена почти формально при наличии подобной экспертной системы.

Также подобная система была бы полезна не только в науке, но и в технике и технологиях, можно было бы достаточно эффективно и быстро найти наилучшее техническое решение, лучшего поставщика продукции. Можно было бы размещать заказы и осуществлять закупки не по конкурсам и котировкам, а на основании действительно объективной оценки возможностей всех исполнителей или всех поставщиков.

Кроме того, в этом случае отпала бы какая бы то ни было потребность в так называемых научных обзорах, против которых имеется множество возражений. Главные возражения

состоят в том, что каждый известный автору научный обзор несет на себе отчетливые признаки субъективности, неполноты, вторичности и опоздании во времени на существенный период. В отношении научных обзоров целесообразно было бы предложить дополнительные принципы этики (редакция наша):

1. Неэтичными являются публикации, сделанные только на основании других публикаций без привлечения собственных результатов научных исследований, кроме явно вторичных публикаций, каковыми являются научные обзоры, монографии, учебники и учебные пособия, справочники; в частности, неэтичными являются обзоры, написанные по другим обзорам (обзор обзоров).

2. Неэтичными являются публикации, утверждающие (согласно названию или аннотации) о том, что они предоставляют сведения о современном состоянии и (или) тенденциях развития какого-либо научного направления, в том случае, когда в них представлены преимущественно устаревшие сведения (более 5-летней давности) или свежие сведения ограничены собственными публикациями автора, или научного руководителя автора, или учеников автора, или научной школы, к которой принадлежит данный автор, или организации по месту работы данного автора.

3. Неэтичными являются публикации, утверждающие (согласно названию или аннотации) о том, что они предоставляют сведения об истории развития какого-либо научного направления, в том случае, когда в них представлены в существенной мере (15% и более) сведения из собственных публикаций автора, или научного руководителя автора, или учеников автора, или научной школы, к которой принадлежит данный автор, или организации по месту работы данного автора.

4. Неэтичными являются ссылки на публикации, которых авторы данной публикации не читали и не видели (заимствование ссылок из чужих публикаций), в подобных случаях необходимо ссылаться именно на те источники, из которых взяты цитируемые или пересказываемые сведения (результаты, положения).

Примените эти три принципа к известным вам обзорам и делайте свои выводы.

#### IV. ВЫДВИГАЕМЫЕ ТЕХНИЧЕСКИЕ ТРЕБОВАНИЯ К ПОСТАНОВКЕ ЗАДАЧИ

Разработанное программно-аппаратное средство должно выявлять все виды и формы заимствования, анализируя тексты представленных научно-технических материалов, вне зависимости от текстового или иного редактора, форматирования, вне зависимости от языка (в пределах согласованного минимального



перечня языков и использованием всех существующих сетевых средств автоматического перевода или распознавания текстов с учетом их прогресса).

Результат анализа должен четко отличать корректное цитирование с указанием на источник заимствования (маркируется одним способом) от необоснованного заимствования без ссылки на источник (плагиат, автоплагиат).

Факультативно на будущее можно потребовать следующие дополнительные опции, которые не так легко сделать в настоящее время, но недалеко время, когда это станет возможным.

Крайне желателен анализ представленных графиков, таблиц, рисунков, с учетом преобразования изображений в различные виды представления изображений (\*.jpeg, \*.bmp и т.п.) и с учетом вносимых фрагментарно исправлений в эти изображения (обрезка, изменение яркости, контрастности, изменение некоторой части изображений с целью изменения текстовых элементов).

Анализ должен включать (при необходимости) опциональную возможность поиска по базе патентов с целью выявления ключевых признаков отличия и формирования по ним предельно формального перечня, указывающего на совпадения отличительных признаков в случае выявления такового, при этом должны учитываться возможные синонимы, например, термин «усилитель» идентичен термину «средство усиления сигнала», «преобразователь с коэффициентом», «линейное усиливающее звено» и некоторым другим.

Основное требование к системе: ее возможности по выявлению автоматически сформированных текстов должны превышать возрастающие возможности по автоматическому созданию таких текстов, следовательно, система должна позволять наращивание опций, усовершенствование по мере использования. Поэтому такая система должна быть обязательно самообучающейся, то есть быть способной самостоятельно или с подсказками выявлять собственные недостатки и устранять их путем усложнения алгоритмов или накопления путей решения аналогичных задач.

## V. ОПАСНОСТИ НЕЦЕЛЕВОГО ИСПОЛЬЗОВАНИЯ

Следует заблаговременно исключить возможность доступа к такой системе со стороны недобросовестных исполнителей, поскольку если система обладает возможностями идентификации плагиата, существует большая вероятность что она в другом режиме или с небольшими доработками может быть преобразована в средство создания псевдо новых текстов, лишенных формальных признаков плагиата вопреки фактическому способу создания этих текстов методом плагиата с

последующей косметической доработкой. То есть средство выявления мошенничества не должно становиться средством мошенничества.

## VI. ТЕХНИКО-ЭКОНОМИЧЕСКОЕ ОБОСНОВАНИЕ

Экономия от использования такого технического средства представляется значительной.

Во-первых, очистка потока публикационной активности от мусора, не содержащего ничего нового, упростит ознакомление научной общественности с действительно новыми публикациями, достойными ознакомления. Это сэкономит много времени и повысит эффективность труда ученых, который должен, вообще говоря, очень хорошо оплачиваться, при условии, что он будет действительно эффективным, а не направленным на создание указанного потока мусорных публикаций.

Во-вторых, можно будет обоснованно лишить финансирования за создание мусорных публикаций, причем, терминология сменится, полезными или мусорными будут считаться публикации не вследствие места их опубликования, вследствие того, содержат или не содержат они новые результаты, суть, обоснованные поправки к теориям и т.п.

В-третьих, может отпасть необходимость какого-либо научного рецензирования, коль скоро искусственный интеллект может это сделать более качественно, чем случайно найденный специалист, который согласился рецензировать статью, которая в целом не представляет для него лично никакого интереса. Действительно, если рецензент работает в точности в той же сфере, тогда статья может задевать его личные интересы, поскольку, вероятно, обгоняет хотя бы в чем-то результаты данного рецензента. Поэтому может оказаться, что рецензент, наиболее компетентный в сфере данной статьи, является далеко не самым объективным, его **компетентное мнение может оказаться недостаточно объективным**. С другой стороны, специалист, работающий в другой сфере, не заинтересован к положительном или отрицательном решении, в этом смысле он больше шансов имеет быть полностью объективным, но он в этом случае не имеет интереса в чтении этой статьи, следовательно, он не захочет посвятить достаточно много времени ее изучению, а для ее объективной оценки ему еще потребовалось бы ознакомиться с самыми новейшими публикациями в этой сфере, поэтому его **непредвзятое мнение не может быть в достаточной степени компетентным**.

Из этого следует, что институт научного рецензирования фактически устарел, от него мы пока не можем отказаться лишь вследствие отсутствия эффективного средства компьютерного интеллекта, средства анализа и

сопоставления, извлечения сути и сравнение этой сути из извлеченной сути других аналогичных публикаций, которое только и могло бы быть одновременно и объективным и компетентным.

Таким образом, это позволило бы отказаться от показавшей себя порочной (и при этом явно коммерческой) системы ценностей, ориентированной на вхождение или не вхождение журнала в ту или иную базу данных, которая, как мы видим, имеет полное право обеспечивать это вхождение всего лишь за деньги. Вместо этого возникла бы иная система оценки новизны и полезности, несколько идентичной системе деклараций отличий в патентном деле, но более широкой и понимаемой во всем научном мире одинаково.

Например, если речь идет о цифроаналоговых преобразователях, задача решается предельно просто – быстродействие и разрядность определяют их почти полностью.

Если речь идет о создаваемых стандартах частоты, то их все можно ранжировать по характеристикам: значение частоты, воспроизводимость, стабильность, функции Аллана в зависимости от времени, а второстепенные признаки – возможность транспортировки, стоимость, сложность, длительность работы, сложность обслуживания и так далее [15].

Если речь идет о синтезе систем для автоматического управления, то есть фактически о регуляторах [16], также следует сравнить их по возможностям, если какие-то методы предлагают худшие возможности для идентичных объектов, следовательно, эти методы не интересны, даже если статья новая. Она имеет право на опубликование, но не представляет существенного интереса для специалистов, даже при условии опубликования в высокорейтинговом журнале. Если же результаты действительно лучшие из возможных [17], то она представляет интерес для специалистов даже и в том случае, если опубликована далеко не в самом высокорейтинговом журнале.

Цитирование как показатель ценности статьи – это критерий, который сформировался при отсутствии лучших критериев, сегодня мы уже видим, что это далеко не лучший критерий, поэтому экспертная система на основе искусственного интеллекта могла бы лучше справиться с задачей оценки полезности публикации. Разумеется, переход не может быть мгновенным, этого никто и не ждет. Но уход от цитирования как средства научной оценки и экспертизы почти во всех случаях давно назрел, эту задачу следует решать, или отсутствие решения доведет ситуацию до состояния фарса.

Сегодня стоимость высококачественного обзора по тематике колеблется от 40 тысяч рублей (такие расценки предлагают некоторые

журналы авторам с высоким индексом цитирования, но в этом случае само содержание не интересует редакционную коллегию, надо признаться, что имеются все признаки, указывающие на то, что единственная цель таких обзоров – это повышение рейтинга самих этих журналов путем повышения среднего показателя индекса Хирша авторов данного журнала) до нескольких миллионов рублей (в этом случае, вероятно, необходимость таких обзоров также не вполне достаточно обоснована, просто это эффективный способ забить средства заработной платы исполнителей и зависить стоимость выполнения проектов). Как бы мы ни относились к этим работам, за них выплачиваются реальные деньги из государственного бюджета. В случае создания такой системы все эти обзоры утратят актуальность коренным образом. Государство будет платить не за труд ученых и не очень ученых читателей чужих статей, а только за компьютерное время (амортизацию техники, сопровождающие расходы и электроэнергию), а также за доступ к источникам, представляющим интерес. В настоящее время источники, не предоставляющие открытого доступа, попросту не попадают в эти заказные обзоры, что также характеризует научную ценность получаемых обзоров не излишне высоко.

## VII. ПРИМЕР «КАК ЕСТЬ» И «КАК НАДО»

Допустим, нас интересует тема ПИД-регуляторов. Поиск на русском языке не представляет существенного интереса, поскольку научно-техническая сфера в этой области русскоязычными авторами также покрывается в англоязычных публикациях. Исключением может являться наше персональное знакомство с перечнем журналов или авторов, представляющих наш отдельный интерес, поиск по этим признакам очевиден. Поэтому ищем в интернет-пространстве на английском языке.

Для начала осуществим патентный поиск. Зададим на сайте Google.com задачу поиска в виде запроса «PID+Patent». Первая же ссылка приводит нас на страницу патента по данной теме, хотя и относительно старого [18]. Мы можем принять этот патент к сведению, то для полноценного научного обзора эта публикация, относящаяся к 1988 году, не подходит. Однако, на ней имеется, как минимум, **три пути дальнейшего поиска, причем каждый путь ветвится.**

Во-первых, внизу дана кликабельная ссылка, связанная с классом данного патента. Она приводит нас на сайт [19]. На нем представлены 95 патентов. Мы можем осуществить выбор только публикаций за последние 5 лет, то есть, начиная с 2015 года, остальные сочтем не современными. К сожалению, самый новый патент в этом перечне относится к 2011 году, но это уже значительный прогресс в нашем

движении [20]. Далее эту ссылку можно раскручивать аналогичным путем.

Во-вторых, мы имеем кликабельные ссылки авторов этого патента и можем посмотреть, что они написали или запатентовали.

В-третьих, мы имеем внизу два перечня, первый – это патенты, на которые данный патент ссылается, они нам не интересны, так как они еще более старые, а второй список – это «похожие документы». В этом перечне мы можем обратить внимание либо на самый свежий патент или несколько самых свежих патентов, либо на наиболее подходящий по тематике, либо принять во внимание оба эти признака. В этом списке мы находим два патента уже за 2013 год и один патент за 2015 год, что и требовалось отыскать [21]. Это китайский патент, имя автора идентифицировать затруднительно, но далее мы можем повторить те же самые процедуры уже с этой страницей. Таким образом, если мы не находим достаточного для нас количества патентов (то есть найденные не представляют интереса), мы можем расширить интервал поиска, взять вместо последних пяти лет десять, или посмотреть близкие классы по идентификационным номерам.

Далее обратим внимание на другие найденные поисковиком ссылки. Третья и четвертая по счету оказываются более продуктивными, поскольку ведут нас к более интересным страницам. В частности, по ссылке [22] мы находим один патент за 2017 год, шесть патентов за 2016 год, и один патент за 2015 год, остальные патенты относятся к более ранним годам, и мы ими можем не интересоваться. Имеются ссылки на другие страницы этого портала по этому классу, но на них представлены более старые патенты, мы можем ими не интересоваться. Ссылка [23] приводит нас к патенту, который со всей очевидностью отвечает запросу, но, во-первых, не столь нов (2010 год), во-вторых, оригинальный текст написан на китайском языке, но на этой странице сайта весь основной текст дан на английском. В перечне аналогичных патентов находим патент [24], который весьма интересен. Поскольку относится к 2016 году, то ест отвечает временным рамкам поиска, во-вторых, интересен по тематике. Все же при ближайшем рассмотрении мы выясняем, что этот патент не относится к разделу ПИД-регуляторы, это адаптивный стабилизатор напряжения, где техника регулирования, конечно, используется, но существо изобретения не представляет существенный интерес. Таким образом, мы вручную проделали некоторую работу лишь для того, чтобы убедиться, что новых технических решений в области ПИД-регуляторов за последние пять лет не так много, как хотелось бы, решений, представляющих существенный интерес вовсе не найдено.

Это представляет собой результат использований только самых очевидных

возможностей поиска, но и они заняли некоторое время, и можно утверждать, что при надлежащем программном средстве все эти результаты могли быть получены быстрее и эффективнее. Алгоритм действия такого программного средства вчерне ясен из описания действий вручную.

Поскольку патентных баз данных много, то имеется широкий фронт работы для автоматизации.

Далее предпримем поиск журнальных публикаций и публикаций в высокорейтинговых трудах конференций.

Например, чтобы исключить различные «инструкции для чайников», многочисленные посты о том, как настроить регулятор, от людей, не являющихся слишком большими специалистами в этой сфере, можно задать поиск в следующем виде: «PID+IEEE». Действительно, сообщество IEEE, как правило, публикует лишь высоконаучные и оригинальные материалы в сфере научной и инженерной мысли, следовательно, такой запрос отмечает много хлама. На первой же странице результата поиска находим два относительно новых источника (2016 год) [25], [26]. На следующей странице также находим несколько свежих ссылок по заданной теме [27], [28], [29], [30], [31]. Получение прав доступа, скачивание этих файлов, автоматизированный анализ содержания – все эти операции могут быть также автоматизированы. Безусловно, существуют и другие журналы, и конференции, поиск их может быть эффективно осуществлен по ссылкам на журналы, по авторам, по поиску в интернете с учетом других ключевых слов поиска, которые бы не ограничивали поиск публикациями сообщества IEEE, но столь же эффективно отсекали несущественные источники. Все это можно автоматизировать, и результат был бы востребован всеми учеными всего мира, всеми экспертами, всеми рецензентами статей, организационных комитетов конференций, экспертами от фондов всех видов, как научных, так и инновационных.

Также эта система смогла бы эффективно находить конкурсы на поддержание научных исследований или на заключение контрактов по конкретным тематикам, наиболее близким и полностью отвечающим параметрам запроса. Еще недавно то, что можно осуществить с помощью элементарных поисковых систем, казалось бы, невозможным, сегодня это привычная данность. Еще недавно люди покупали энциклопедии, справочники, словари, сегодня это бессмысленные траты, поиск в бумажных справочниках и словарях намного более трудоемкий, чем поиск в соответствующих, традиционными поисковыми движками или на тематических сайтах.

Вероятно, сегодня предложение о создании системы, функционально превышающего

систему [14] на порядки, кажется неосуществимым, но не всем и не во всем. Завтра или послезавтра это будет реальностью.

## ЗАКЛЮЧЕНИЕ

Поставлена актуальная задача создания системы автоматического анализа научных и патентных источников методами искусственного интеллекта, нейронных сетей, нейронных вычислений, больших данных. Показана актуальность, техническая и экономическая обоснованность этой задачи, сформулированы основные технические требования.

## ЛИТЕРАТУРА

- [1] Эксперты раскрыли популярность Ольги Бузовой. Простая как стул. <https://www.nsk.kp.ru/daily/27028.3/4090731/>
- [2] Жмудь В. А. Мошенничество на рынке псевдонаучных публикаций. Автоматика и программная инженерия. 2017. № 1 (19). С. 131–148. <http://www.jurnal.nips.ru/sites/default/files/%D0%90%D0%B8%D0%9F%D0%98-1-2017-14.pdf>
- [3] Жмудь В. А. Сомнительные услуги на рынке псевдонаучных публикаций. Автоматика и программная инженерия. 2017. № 1 (19). С. 110–130. <http://www.jurnal.nips.ru/sites/default/files/%D0%90%D0%B8%D0%9F%D0%98-1-2017-13.pdf>
- [4] Жмудь В. А. Что делать, чтобы коммерческая «помощь отечественной науке» не мешала ее развитию. Автоматика и программная инженерия. 2016. № 4 (18). С. 149–162. [http://www.jurnal.nips.ru/sites/default/files/%D0%90%D0%B8%D0%9F%D0%98-4-2016-18\\_0.pdf](http://www.jurnal.nips.ru/sites/default/files/%D0%90%D0%B8%D0%9F%D0%98-4-2016-18_0.pdf)
- [5] Жмудь В. А. Цели и результаты соревнований по импакт-факторам и h-индексу. Автоматика и программная инженерия. 2016. № 4 (18). С. 104–127. <http://www.jurnal.nips.ru/sites/default/files/%D0%90%D0%B8%D0%9F%D0%98-4-2016-16.pdf>
- [6] В.А. Жмудь. Вскрытие схем нецелевого извлечения крупных сумм из государственного бюджета в сферу лженауки. Автоматика и программная инженерия. 2019. № 2 (28). С. 113–132. [http://www.jurnal.nips.ru/sites/default/files/AaSI-2-2019-13\\_0.pdf](http://www.jurnal.nips.ru/sites/default/files/AaSI-2-2019-13_0.pdf)
- [7] В.А. Жмудь. Мусорные публикации-клоны и их вредное влияние на науку и образование. Автоматика и программная инженерия. 2019. № 2 (28). С. 133–149. [http://www.jurnal.nips.ru/sites/default/files/AaSI-2-2019-14\\_0.pdf](http://www.jurnal.nips.ru/sites/default/files/AaSI-2-2019-14_0.pdf)
- [8] Russian journals retract more than 800 papers after 'bombshell' investigation. <https://www.sciencemag.org/news/2020/01/russian-journals-retract-more-800-papers-after-bombshell-investigation>
- [9] [https://dic.academic.ru/dic.nsf/dic\\_wingwords/1576/%D0%9C%D1%8B](https://dic.academic.ru/dic.nsf/dic_wingwords/1576/%D0%9C%D1%8B)
- [10] А. Кулешова. Мусорная наука: кто, как и почему публикует в России статьи. <https://newizv.ru/article/general/19-07-2019/musornaya-nauka-kto-kak-i-pochem-publikuet-v-rossii-nauchnye-statii>
- [11] Минобрнауки сократит финансирование вузам за "мусорные публикации". [https://akvobr.ru/minobrnauki\\_sokratit\\_finansirovanie\\_vuzam\\_za\\_musornye\\_publicatsii/](https://akvobr.ru/minobrnauki_sokratit_finansirovanie_vuzam_za_musornye_publicatsii/)
- [12] Министерство образования и науки РФ. Приказ №1586 от 12 декабря 2016 г. Об утверждении правил формирования перечня рецензируемых научных... <http://pravo.gov.ru/proxy/ips/?docbody=&nd=102432726>
- [13] Цитаты из Бомарше. <https://zen.yandex.ru/media/id/5d078e0e24ccb50dc611de65/on-slishkom-byi-smeshon-5d359942d11ba200ae4b4def>
- [14] <https://www.antiplagiat.ru/>
- [15] Absolute Frequency Measurements In Precision Laser Spectroscopy of Muonium. Bagayev S.N., Belkin A.M., Dychkov A.S., Farnosov S.A., Fateev N.V., Kolker D.B., Matyugin Yu.A., Okhapkin M.V., Pivtsov V.S., Zakharyash V.F., Zhmud V.A. В сборнике: Proceedings of SPIE - The International Society for Optical Engineering Proceedings of the 1998 Nonlinear Optical Phenomena and Coherent Optics in Information Technologies, ICONO-98. sponsors: Ministry Sci./Techn. Russian Fed., Russian Foundation for Basic Res., Ministry Ed. of the Russian Fed., Basic Optics/Spectrosc. Ed./Res.. Moscow, RUS, 1999. С. 310-318.
- [16] Динамика мехатронных систем. Жмудь В.А., Французова Г.А., Востриков А.С. Учебное пособие / Новосибирск, 2014.
- [17] Zhmud V., Vostrikov A., Semibalumut V. Feedback Systems With Pseudo Local Loops. В сборнике: Testing and Measurement: Techniques and Applications - Proceedings of the 2015 International Conference on Testing and Measurement: Techniques and Applications, TMTA 2015 2015. С. 411-417.
- [18] Tadayoshi Saito, Junzo Kawakami, Susumu Takahashi, Testuo Suehiro, Hiroshi Matsumoto, Kouji Tachibana. Pid controller system. US Patent. US4903192A. <https://patents.google.com/patent/US4903192A/en>
- [19] <https://patents.google.com/?q=G05B13%2F0275>
- [20] Lotfi A. Zadeh. US Patent. US8949170B2. System and method for analyzing ambiguities in language for natural language processing. <https://patents.google.com/patent/US8949170B2/en?q=G05B13%2F0275&num=100>
- [21] Intelligent control method for pulsation vacuum sterilizer based on fuzzy control. China Patent. CN104850010. <https://patents.google.com/patent/CN104850010A/en>
- [22] JUSTIA patents. Proportional-integral-derivative (p-i-d) Patents (Class 700/42) <https://patents.justia.com/patents-by-us-classification/700/42>
- [23] PID regulator circuit. China Patent. CN201732261U. <https://patents.google.com/patent/CN201732261U/en>
- [24] A kind of self-adaptive low-voltage difference linear constant voltage regulator of wide withstand voltage scope and chip thereof. China Patent. CN104808734B. <https://patents.google.com/patent/CN104808734B/en>
- [25] B. Doicin, M. Popescu and C. Patrascioiu, "PID Controller optimal tuning," 2016 8th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Ploiesti, 2016, pp. 1-4. doi: 10.1109/ECAI.2016.7861175. URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&number=7861175&isnumber=7861062>
- [26] L. Xu and F. Ding, "Design of the PID controller for industrial processes based on the stability margin," 2016 Chinese Control and Decision Conference (CCDC), Yinchuan, 2016, pp. 3300-3304. doi: 10.1109/CCDC.2016.7531552. URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&number=7531552&isnumber=7530943>
- [27] S. Tjokro and S. L. Shah, "Adaptive PID Control," 1985 American Control Conference, Boston, MA, USA, 1985, pp. 1528-1534. doi: 10.23919/ACC.1985.4788858. URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&number=4788858&isnumber=4788561>
- [28] I. D. Diaz-Rodríguez, V. A. Oliveira and S. P. Bhattacharyya, "Modern design of classical controllers: Digital PID controllers," 2015 IEEE 24th International Symposium on Industrial Electronics (ISIE), Buzios, 2015, pp. 1010-1015. doi: 10.1109/ISIE.2015.7281610. URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&number=7281610&isnumber=7281431>
- [29] U. Mehta and S. Majhi, "On-line identification and control methods for PID controllers," 2010 11th International Conference on Control Automation Robotics & Vision, Singapore, 2010, pp. 1507-1511. doi: 10.1109/ICARCV.2010.5707428. URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&number=5707428&isnumber=5707203>
- [30] L. Xu and F. Ding, "Design of the PID controller for industrial processes based on the stability margin," 2016 Chinese Control and Decision Conference (CCDC), Yinchuan, 2016, pp. 3300-3304. doi: 10.1109/CCDC.2016.7531552. URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&number=7531552&isnumber=7530943>



[org/stamp/stamp.jsp?tp=&amumber=7531552&isnumber=7530943](http://www.jurnal.nips.ru/sites/default/files/%D0%90%D0%B8%D0%9F%D0%98-4-2016-16.pdf)

- [31] Y. Ashida, K. Hayashi, S. Wakitani and T. Yamamoto, "A novel approach in designing PID controllers using closed-loop data," *2016 American Control Conference (ACC)*, Boston, MA, 2016, pp. 5308-5313. doi: 10.1109/ACC.2016.7526501. URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&amumber=7526501&isnumber=7524873>



**Вадим Аркадьевич Жмуд** - заведующий кафедрой Автоматики НГТУ, профессор, доктор технических наук, заместитель генерального директора Новосибирского института программных систем по науке (ПАО «НИПС»)

**E-mail:** [oao\\_nips@bk.ru](mailto:oao_nips@bk.ru)

630073, Новосибирск,  
Просп. К. Маркса, д. 20  
630090, Новосибирск, просп.  
Академика Лаврентьева, д 6/3



**Александр Ляпидевский** - Кандидат экономических наук, директор Новосибирского института программных систем, автор около 100 научных статей. Область научных интересов и компетенций - программные системы и инструменты, инновационные технологии.

**E-mail:** [nips@nips.ru](mailto:nips@nips.ru)

Россия, Новосибирск, 630090,  
просп. Ак. Лаврентьева 6/1. НИПС.

Статья получена 15.12.2019.

## Proposals for Creating a Technical Tool for Multilingual Analysis of Sources of Scientific Information

V.A. Zhmud<sup>1,2</sup>, A.V. Liapidevskiy<sup>2</sup>

Novosibirsk State Technical University, Novosibirsk Institute of Program Systems, Novosibirsk, Russia

*Abstract.* Russian science (and not only) is placed in conditions when scientific publications are considered its most important product. This leads to a loss of conscientiousness by some Russian scientists and teachers on a massive scale (many hundreds of people involved). It is not possible to solve this problem with ethical influences, and, apparently, this is not required. The accepted criteria for evaluating scientific publications not by their content and novelty, but by the place of their publication, are clearly outdated, which is irrefutably proved by many examples. Consequently, other means of countering this stream of pseudo-scientific publications are relevant. At the same time, the task of quickly creating reviews on a wide variety of topics is relevant, these reviews are made, as a rule, not of high enough quality, the laboriousness of writing them with responsible reference is very high, sometimes it turns out that until the review is ready, it will become obsolete. Neural networks, neural computers and big data analysis tools offer tools to solve these problems together and as efficiently as possible. The task of creating such a system is substantiated and basic instructions are given on the most important technical requirements for it.

*Keywords:* big data, neural networks, neural computing, search engines, automatic translation, automatic text analysis, science, scientific publications, financing, plagiarism, grants, citation, Hirsch index, theory of automatic control

### REFERENCES

- [1] Eksperty raskryli populyarnost' Ol'gi Buzovoy. Prostaya kak stul. <https://www.nsk.kp.ru/daily/27028.3/4090731/>
- [2] Zhmud V. A. Moshennichestvo na rynke psevdonauchnykh publikatsiy. *Avtomatika i programmaya inzheneriya*. 2017. № 1 (19). S. 131–148. <http://www.jurnal.nips.ru/sites/default/files/%D0%90%D0%B8%D0%9F%D0%98-1-2017-14.pdf>
- [3] Zhmud V. A. Somnitel'nyye uslugi na rynke psevdonauchnykh publikatsiy. *Avtomatika i programmaya inzheneriya*. 2017. № 1 (19). S. 110–130. <http://www.jurnal.nips.ru/sites/default/files/%D0%90%D0%B8%D0%9F%D0%98-1-2017-13.pdf>
- [4] Zhmud V. A. Chto delat', chtoby kommercheskaya «pomoshch' otechestvennoy nauke» ne meshala yeye razvitiyu. *Avtomatika i programmaya inzheneriya*. 2016. № 4 (18). S. 149–162. [http://www.jurnal.nips.ru/sites/default/files/%D0%90%D0%B8%D0%9F%D0%98-4-2016-18\\_0.pdf](http://www.jurnal.nips.ru/sites/default/files/%D0%90%D0%B8%D0%9F%D0%98-4-2016-18_0.pdf)
- [5] Zhmud V. A. Tseli i rezul'taty sorevnovaniya po impakt-faktoram i h-indeksu. *Avtomatika i programmaya inzheneriya*. 2016. № 4 (18). S. 104–127. <http://www.jurnal.nips.ru/sites/default/files/%D0%90%D0%B8%D0%9F%D0%98-4-2016-16.pdf>
- [6] V.A. Zhmud. Vskrytiye skhem netselevogo izvlecheniya krupnykh summ iz gosudarstvennogo byudzheta v sferu lzhenauki. *Avtomatika i programmaya inzheneriya*. 2019. № 2 (28). S. 113–132. [http://www.jurnal.nips.ru/sites/default/files/AaSI-2-2019-13\\_0.pdf](http://www.jurnal.nips.ru/sites/default/files/AaSI-2-2019-13_0.pdf)
- [7] V.A. Zhmud. Musornyye publikatsii-klony i ikh vrednoye vliyaniye na nauku i obrazovaniye. *Avtomatika i programmaya inzheneriya*. 2019. № 2 (28). S. 133–149. [http://www.jurnal.nips.ru/sites/default/files/AaSI-2-2019-14\\_0.pdf](http://www.jurnal.nips.ru/sites/default/files/AaSI-2-2019-14_0.pdf)
- [8] Russian journals retract more than 800 papers after 'bombshell' investigation. <https://www.sciencemag.org/news/2020/01/russian-journals-retract-more-800-papers-after-bombshell-investigation>
- [9] [https://dic.academic.ru/dic.nsf/dic\\_wingwords/1576/%D0%9C%D1%8B](https://dic.academic.ru/dic.nsf/dic_wingwords/1576/%D0%9C%D1%8B)
- [10] A. Kuleshova. Musornaya nauka: kto, kak i pochem publikuyet v Rossii stat'i. <https://newizv.ru/article/general/19-07-2019/musornaya-nauka-kto-kak-i-pochem-publikuet-v-rossii-nauchnye-stati>

- [11] Minobrnauki sokratit finansirovaniye vuzam za "musornyye publikatsii". [https://akvobr.ru/minobrnauki\\_sokratit\\_finansirovanie\\_vuza\\_m.html](https://akvobr.ru/minobrnauki_sokratit_finansirovanie_vuza_m.html)
- [12] Ministerstvo obrazovaniya i nauki RF. Prikaz №1586 ot 12 dekabrya 2016 g. Ob utverzhenii pravil formirovaniya perechnya resenziruyemykh nauchnykh... <http://pravo.gov.ru/proxy/ips/?docbody=&nd=102432726>
- [13] Tsitaty iz Bomarshe. <https://zen.yandex.ru/media/id/5d078e0e24ccb50dc611de65/on-slishkom-byt-smeshon-5d359942d11ba200ae4b4def>
- [14] <https://www.antiplagiat.ru/>
- [15] Absolute Frequency Measurements In Precision Laser Spectroscopy of Muonium. Bagayev S.N., Belkin A.M., Dyckov A.S., Farnosov S.A., Fateev N.V., Kolker D.B., Matyugin Yu.A., Okhapkin M.V., Pivtsov V.S., Zakharyash V.F., Zhmud V.A. В сборнике: Proceedings of SPIE - The International Society for Optical Engineering Proceedings of the 1998 Nonlinear Optical Phenomena and Coherent Optics in Information Technologies, ICONO-98. sponsors: Ministry Sci./Techn. Russian Fed., Russian Foundation for Basic Res., Ministry Ed. of the Russian Fed., Basic Optics/Spectrosc. Ed./Res.. Moscow, RUS, 1999. C. 310-318.
- [16] Dinamika mekhatronnykh sistem. Zhmud V.A., Frantsuzova G.A., Vostrikov A.S. Uchebnoye posobiye / Novosibirsk, 2014.
- [17] Zhmud V., Vostrikov A., Semibalamut V. Feedback Systems With Pseudo Local Loops. В сборнике: Testing and Measurement: Techniques and Applications - Proceedings of the 2015 International Conference on Testing and Measurement: Techniques and Applications, TMTA 2015 2015. C. 411-417.
- [18] Tadayoshi Saito, Junzo Kawakami, Susumu Takahashi, Testuo Suehiro, Hiroshi Matsumoto, Kouji Tachibana. Pid controller system. US Patent. US4903192A. <https://patents.google.com/patent/US4903192A/en>
- [19] <https://patents.google.com/?q=G05B13%2F0275>
- [20] Lotfi A. Zadeh. US Patent. US8949170B2. System and method for analyzing ambiguities in language for natural language processing. <https://patents.google.com/patent/US8949170B2/en?q=G05B13%2F0275&num=100>
- [21] Intelligent control method for pulsation vacuum sterilizer based on fuzzy control. China Patent. CN104850010. <https://patents.google.com/patent/CN104850010A/en>
- [22] JUSTIA patents. Proportional-integral-derivative (p-i-d) Patents (Class 700/42) <https://patents.justia.com/patents-by-us-classification/700/42>
- [23] PID regulator circuit. China Patent. CN201732261U. <https://patents.google.com/patent/CN201732261U/en>
- [24] A kind of self-adaptive low-voltage difference linear constant voltage regulator of wide withstand voltage scope and chip thereof. China Patent. CN104808734B. <https://patents.google.com/patent/CN104808734B/en>
- [25] B. Doicin, M. Popescu and C. Patrascioiu, "PID Controller optimal tuning," 2016 8th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Ploiesti, 2016, pp. 1-4. doi: 10.1109/ECAI.2016.7861175. URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7861175&isnumber=7861062>
- [26] L. Xu and F. Ding, "Design of the PID controller for industrial processes based on the stability margin," 2016 Chinese Control and Decision Conference (CCDC), Yinchuan, 2016, pp. 3300-3304. doi: 10.1109/CCDC.2016.7531552. URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7531552&isnumber=7530943>
- [27] S. Tjokro and S. L. Shah, "Adaptive PID Control," 1985 American Control Conference, Boston, MA, USA, 1985, pp. 1528-1534. doi: 10.23919/ACC.1985.4788858. URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4788858&isnumber=4788561>
- [28] I. D. Díaz-Rodríguez, V. A. Oliveira and S. P. Bhattacharyya, "Modern design of classical controllers: Digital PID controllers," 2015 IEEE 24th International Symposium on Industrial Electronics (ISIE), Buzios, 2015, pp. 1010-1015. doi: 10.1109/ISIE.2015.7281610. URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7281610&isnumber=7281431>
- [29] U. Mehta and S. Majhi, "On-line identification and control methods for PID controllers," 2010 11th International Conference on Control Automation Robotics & Vision, Singapore, 2010, pp. 1507-1511. doi: 10.1109/ICARCV.2010.5707428. URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5707428&isnumber=5707203>
- [30] L. Xu and F. Ding, "Design of the PID controller for industrial processes based on the stability margin," 2016 Chinese Control and Decision Conference (CCDC), Yinchuan, 2016, pp. 3300-3304. doi: 10.1109/CCDC.2016.7531552. URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7531552&isnumber=7530943>
- [31] Y. Ashida, K. Hayashi, S. Wakitani and T. Yamamoto, "A novel approach in designing PID controllers using closed-loop data," 2016 American Control Conference (ACC), Boston, MA, 2016, pp. 5308-5313. doi: 10.1109/ACC.2016.7526501. URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7526501&isnumber=7524873>



**Vadim Zhmud** – Head of the Department of Automation in NSTU, Professor, Doctor of Technical Sciences. Vice-director of Novosibirsk Institute of Program Systems

E-mail: [oaonips@bk.ru](mailto:oaonips@bk.ru)  
630073, Novosibirsk,  
str. Prosp. K. Marksa, h. 20  
630090, Novosibirsk,  
Pr. Ak.Lavrentievs, 6/1



**Alexander V. Liapidevskiy**, PhD in Economics, director of the Novosibirsk Institute of Program (Software) Systems, the author of about 100 scientific articles. Area of scientific interests and competences - software systems and tools, innovative technologies.  
E-mail: [nips@nips.ru](mailto:nips@nips.ru)  
Russia, Novosibirsk, 630090, prosp. Ak. Lavrentieva 6/1. NIPS.

The paper has been received on 15/12/2019.