

**СОДЕРЖАНИЕ**

<b>Цукерман Г.</b> Пришло ли время для количественных исследований недобросовестного поведения в науке?	3
<b>Ганади Нежад Ф., Осарех Ф., Гане М.Р.</b> Предсказание предметных направлений международных исследований в области библиотечного дела и информатики до 2030 года с использованием метода глубокого обучения	16
<b>МАТЕРИАЛЫ МКСНТИ</b>	
<b>Абдувалиев А.А.</b> Инновационное развитие Узбекистана: анализ сильных и слабых позиций в рейтинге Глобального инновационного индекса	32

# **РЕДАКЦИОННАЯ КОЛЛЕГИЯ**

## **ГЛАВНЫЙ РЕДАКТОР**

**д.филол.н. ГИЛЯРЕВСКИЙ Р.С.**

## **ЧЛЕНЫ РЕДКОЛЛЕГИИ:**

**к.т.н. БЫКОВ В.А. (РОССИЯ),**

**к.физ.-мат.н. ВАРДАНЯН Г. Г. (АРМЕНИЯ),**

**д.т.н., проф. ВОЙТОВ И. В. (БЕЛАРУСЬ),**

**МИРАЛИЕВ К. Х. (ТАДЖИКИСТАН)**

## **РЕДАКТОРЫ:**

**ОВЧЕНКОВА Е.А., САРКИСЯН Д.Б.**

## Пришло ли время для количественных исследований недобросовестного поведения в науке?\*

**ГАРРИЕТ ЦУКЕРМАН  
(HARRIET ZUCKERMAN)**

Колумбийский университет,  
шт. Нью-Йорк, г. Нью-Йорк, США

*Недобросовестное поведение в науке – современная и содержательно важная проблема в социальном изучении науки. Но при отсутствии полных и надежных данных, необходимых для анализа, на пути его количественного исследования стоят серьезные препятствия. Доступные базы данных, включая государственные, не совершенны, а сбор новых данных сопряжен со своими проблемами. Во-первых, мало что известно о необъективности официальных правительственных отчетов. Во-вторых, официальные отчеты исключают случаи злоупотреблений кроме фабрикаций, фальсификации и плагиата свидетельств (fabrication, falsification, and plagiarism of evidence — FFP). В-третьих, несмотря на целесообразность использования официальных данных, имеющаяся официальная информация является скудной; она мало что говорит о нарушителях и не позволяет определить потенциальные причины их действий и среду, в которой происходили нарушения. В-четвертых, он (анализ) также не выдерживает испытания на возможность оценки популяций, подверженных риску, что делает невозможным подсчет случаев «заболеваемости». Здоровая доля скептицизма необходима при оценке, как результатов текущих количественных исследований, так и предложений по исправлению ситуации.*

### ВВЕДЕНИЕ

Недобросовестное поведение в науке вызывает значительный интерес у экономистов, политологов, политиков, социологов, исследователей в области науки\*\*,

и, конечно, ученых всех мастей. Оно является также процветающей областью исследований согласно БД *Google Scholar*, грубому, но готовому показателю исследовательской активности. В 2019 г. в его базе

\* Перевод Zuckerman H. Is “the time ripe” for quantitative research on misconduct in science? // *Quantitative Science Studies*. — 2020. Vol. 1, No. 3 (summer 2020). — P. 945-958. — [https://doi.org/10.1162/qss\\_a\\_00065](https://doi.org/10.1162/qss_a_00065). — [http://direct.mit.edu/qss/article-pdf/1/3/945/1869848/qss\\_a\\_00065/pdf](http://direct.mit.edu/qss/article-pdf/1/3/945/1869848/qss_a_00065/pdf)

\*\*В целом эмпирические исследования в науковедении склоняются в сторону качественных опросов и небольших тематических исследований. По крайней мере, так кажется,

судя по недавним исследованиям, опубликованным в этой области, и содержанию таких журналов, как *Social Studies of Science* и *Science, Technology and Human Values*. Эта тенденция согласуется с мнением некоторых о том, что количественные данные и сопутствующий им анализ неизбежно вводят в заблуждение при попытке исследовать то, чем занимаются ученые, и, действительно, многие другие социальные явления.

данных содержится 1 310 статей, книг, презентаций на конференциях и других материалов по тем или иным аспектам «научной недобросовестности». Действительно, в *Google Scholar* резко возросло количество статей о «научной недобросовестности»: в 1981 г. таких статей было менее 10, а в 2011 г. их количество подскочило до 725 [1, рис. 4.1., с. 96]\*. Абсолютное число записей и крутизна связанной с ними кривой роста удивили обозревателя, который писал на эту тему в 1977 и 1984 гг. [2,3], когда эта тема была весьма далека от «растущей индустрии». Возникает вопрос, сохранился ли явно растущий интерес к недобросовестному поведению, о котором сигнализировала БД *Google Scholar* в первые 30 лет, или он угас, как это часто бывает в эмпирических исследованиях этой тематики. Рисунок сравнивает темпы роста упоминаний в *Google Scholar* за 2006-2011 гг. и за 2012-2019 гг. и дает ответ на этот вопрос.

Интерес к недобросовестному поведению в науке не только не ослабевает, но и, как отражено на рисунке, возрастает. Рост абсолютных цифр по-прежнему поражает, а ежегодные темпы роста в

последующие годы стали даже быстрее, чем раньше. С 2012 по 2019 гг. они выросли в 5,89 раза, а с 2006 по 2011 гг. в 1,61 раза. Не придавая большого значения приведенным здесь точным цифрам, можно сделать вывод, что феномен недобросовестного поведения явно привлек к себе большое внимание ученых.

Однако остается неизвестным, что покажет контент-анализ этих статей: какого рода научную сферу они представляют, какие темы они затрагивают, в какой степени эти статьи отражают эмпирические исследования, и каковы методы и данные, используемые ими, а тем более к каким дисциплинам относятся их авторы; и если это журнальные статьи, то как они распределены по журналам, и какая их доля соответствует тому количественному анализу, который стремится публиковать журнал *Quantitative Science Studies*. Указывает ли вся эта активность на то, что существует или готовится к созданию солидный массив количественных исследований недобросовестного поведения? Я признаюсь, что не проводила существенной работы, которая требует ответа на эти вопросы.

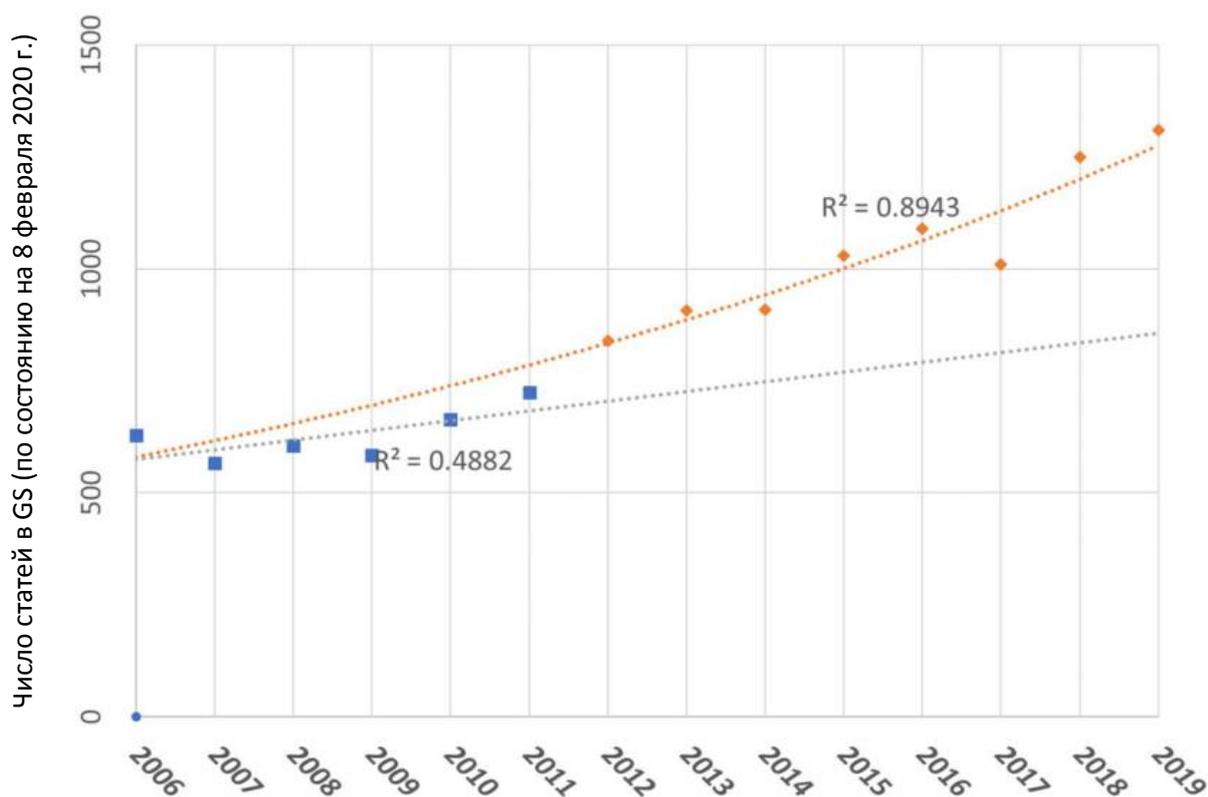


Рис. Отображение научного недобросовестного поведения в Google Scholar за 2006-2020 гг. Источник: Рисунок основан на работе авторов [1, рис. 4.1., с. 96], но был изменен и переработан для этой статьи Лутом Лейдесдорффом 8 февраля 2020 г.

\* *Google Scholar* был запущен в 2004 г., но в нем перечислены публикации, относящиеся к данному исследованию, начиная с 1980 г. Охват исследовательской литературы включает все виды публикаций, но ограничивается теми, что на английском языке. Приведенное здесь число немного отличается от того, что показано на графике Бен-Иегуды и Оливера-Лумермана [1, рис. 4.1., с. 96].

Когда появился шанс вернуться к теме недобросовестного научного поведения, особенно его распространенности, обзор имеющейся сегодня литературы был явно востребован. Чем больше читаю, тем больше убеждаюсь, что недобросовестное поведение стало более обсуждаемым, чем количественно изученным, а когда проводились количественные исследования, они наталкивались на существенные методологические препятствия. Они происходили в большинстве из-за неадекватности доступных данных, а также из-за фундаментальных трудностей сбора новых данных. Тогда в качестве замены большой ортодоксальной исследовательской работы, которую я первоначально стремилась написать, в оставшейся части этих комментариев я рассматриваю важные проблемы, возникающие в процессе проведения количественного исследования недобросовестного поведения в науке, и меры, которые можно предпринять для улучшения нынешней ситуации. Как станет очевидным позже, мои комментарии опираются на эволюцию темы недобросовестного научного поведения в США гораздо чаще, чем хотелось бы\*. Это, вероятно, просто результат значительно большего числа доступных в США данных и большего числа проведенных исследований этой сферы в США. И вовсе не означает, что недобросовестное научное поведение в США является более частым, чем где-либо еще.

Мои комментарии означают, что чаще всего недостатки доступных данных заключаются в изменении определения недобросовестного поведения под воздействием официального контроля. Связанные с этим сложности входят в потенциал исследований. Это введение технологий по обнаружению недобросовестного поведения, последствие внимания со стороны прессы на недобросовестное поведение в рамках общественного доверия к науке и повторные наблюдения из-за потребности выяснить причины недобросовестного научного поведения.

## НЕДОСТАТКИ ДОСТУПНЫХ ДАННЫХ

Поскольку моя статья [2] опубликована более 40 лет назад, странно, что ничего подобного эпидемиологии недобросовестного поведения не существует в науке, области, в которой данные считаются основополагающими, бережное хранение доку-

---

\* Научное недобросовестное поведение и интерес к нему не являются опосредованными США феноменами. Агентства, нацеленные на контроль и сдерживание научного недобросовестного поведения учреждались в Канаде, Китае, Дании, Израиле, Норвегии и Соединенном королевстве, в числе других государств. Они существенно различаются в целях и видах деятельности [4]. Сравнительный межнациональный анализ недобросовестного поведения и реакций на него поэтому представляется особенно проблематичным.

ментов является обязательным, а статистическая сложность признанной. Кажется вероятным, что отсутствие хорошего свидетельства происходит из длительно существующего убеждения у ученых в том, что недобросовестное поведение чрезвычайно редко встречается и что это не является проблемой для продолжающейся научной практики, и таким образом не нуждается в дальнейшем внимании.

В той статье я также отмечала, что анализ недобросовестного поведения в науке представляет те же самые проблемы, с которыми ученые по отклоняющемуся от нормы или криминальному поведению сталкиваются десятилетиями. Например, официальные данные по различным случаям весьма вероятно являются противоречивыми, поскольку противоречие существует в методах обнаружения отклонений от нормы и, таким образом, в отчетах, основанных на них. В какой степени существует, следовательно, недобросовестное поведение в науке? В то время были проведены отдельные исследования с использованием официальных данных, взятых из баз данных США, они являются ограниченными и в любом случае нет способа узнать о противоречиях, наследуемых ими. Мы лишь знаем, что официальные данные, т.е. данные, собранные правительственными агентствами, виновны в плохом контроле за большинством последующих примеров недобросовестного поведения и ограничены только актами «доказанных» нарушений. Нет никакой информации о вынесенных обвинениях, а только о признанных несущественными или о характеристике жалоб на то, при каких обстоятельствах они последуют; нет аналогов официально оговоренных наказаний за отдельные действия; и что очень важно, нет ни одной попытки идентификации сообществ, находящихся в зоне риска вовлеченности в недобросовестное поведение. Без попыток идентифицировать процент привлеченных к ответственности нарушителей не представляется возможным подсчитать их число. Проблемы множатся, если заинтересоваться определением того, встречаются ли изменения в сфере распространения отдельных актов недобросовестного поведения или в других его проявлениях. При отсутствии подобной информации оценка эффективности действий по сокращению недобросовестного поведения не представляется возможной.

Список недостатков также включает довольно сжатое определение недобросовестного поведения, которое сейчас включает фабрикацию, фальсификацию и плагиат свидетельства (или часто используемый акроним *FFP*)\*. Эти три понятия являются

---

\* Определения, используемые Национальным институтом здравоохранения (NIH) и Национальным научным фондом (NSF) отличаются друг от друга, но не существенно. Согласно последнему: Недобросовестное поведение является «формой существенного отклонения от

несомненно важными нарушениями стандартов, которые по убеждению ученых требуются научным вкладом, чтобы быть доверительными. Но являются ли они действиями, заслуживающими анализа? – Нет. (Скоро я вернусь к этому вопросу).

Другие источники информации также имеют проблемы. Самоотчеты нарушений считаются лучшими показателями, чем официальные отчеты в литературе по уголовному праву. Но они также подозрительны по ряду причин. Насколько прочным является свидетельство, собранное в опросах тех, кто сообщает о своей вовлеченности в недобросовестное научное поведение? Межсубъективность – всегда проблема в изучении опроса, и это тоже подпадает под наш случай. Какие действия респонденты вспоминают при упоминании, что они сфабриковали или сфальсифицировали данные? Было бы прозрачным узнать в мельчайших деталях, что считается фальсификацией самими респондентами опроса. Например, где необходимо провести линию между сообщением всех свидетельств, приводимых исследователями, и только теми результатами, которые они считают надежными? Тщательное исследование исторических свидетельств признанной фальсификации показывает (см. [5]), что эта линия размыта, и что исследователи обоснованно понимают, какие наблюдения вероятнее всего будут правильными, а какие – нет.

Фабрикация свидетельства является более простой, но вопросы в опросах, например, должны изучать то, насколько важной была признанная фабрикация относительно сообщаемых научных результатов, частоту встречаемости фабрикации, знают ли о ней другие лица, и были ли они когда-нибудь выявлены\*\*. На фоне почти отсутствия информации по интерпретациям респондентами изучаемых феноменов в подобного рода опросах, можно в лучшем случае догадываться о противоречиях в самоотчетах, в том числе недооценивают ли или переоценивают они нарушения, которые они признаны измерять (хотя многие верят, что это так на самом деле).

Аналогично, надежность подсчетов свидетельства из первых рук, сообщаемых в опросах или публично доступных через отчеты осведомителей, также является не самоочевидной. Несмотря на официальную защиту, предоставленную осведомителям, вероятнее всего, что они недооценивают

случаи недобросовестного поведения. Например, утверждается, что осведомление является превосходством, которого могут достичь лишь несколько работающих ученых, и, кажется очевидным, те, кто молод и те, которые занимают невысокие позиции. Хотя старшие научные сотрудники могут быть менее уязвимыми, чем их коллеги рангом ниже, и хотя это непосредственно в их интересах сообщать о нарушениях, которые они считают правдивыми, в краткосрочной перспективе они могут быть также сопротивляющимися, чтобы сообщить о инцидентах недобросовестного поведения, свидетелями которых они были и которые совершались на их глазах, или потому что им не хотелось причинять вред карьерам коллег.

Более того, некоторые акты недобросовестного поведения более прозрачны, чем другие для «невооруженного глаза». Как часто совершаются пойманные на месте проступки, такие как нарушение этики исследования Уильяма Саммерлина несколько десятилетий назад? Оно произошло, когда скептик-технолог обнаружил, что черное пятно у белой мыши можно легко удалить с помощью алкоголя. Это черное пятно было важным по утверждению исследователя, поскольку мышь с пятном была свидетельством успешно проведенной трансплантации кожи на генетически неродственных животных [6].

Более вероятно, что недобросовестное поведение выявляется специальной проверкой записей исследований, хотя даже эти данные могут быть и бывали искусно подделаны злоумышленниками или потому, что само по себе свидетельство кажется им не располагающим к доверию.

Сегодня используются и другие средства обнаружения, включая программное обеспечение, признанное для охвата плагиата текста и изображений, а также фабрикации и «исправления» изображений, взятых из чужих публикаций, чтобы они казались оригинальным вкладом в исследование. Проверка статистики тоже используется для измерения того, являются ли сообщаемые данные настолько за гранью возможного, что они вероятнее всего являются мошенническими. Однако здесь, опираясь на исследование историков науки, можно привести пример классического утверждения, что статистические наблюдения Грегором Менделем генотипов душистого горошка были также «слишком хорошими, чтобы быть правдой». Это мнение было признано необоснованным, и более того, Рональда Фишера, часто признаваемый ответственным за обвинение Г. Менделя, не был его обвинителем, а скорее лишь еще одним статистиком, примкнувшим к обвинению [7]. Вкратце, новые ситуации могут вносить противоречия в исследование недобросовестного поведения за счет роста шансов обнаружения недобросовестного поведения в тех их типах, которое они были предназначены открывать, но в то же время оставляя нетронутыми дру-

---

принятых практик научного сообщества в целях поддержания добросовестности научного документа. Подобное недобросовестное поведение должно быть совершено на международном, осознанном, или безответственном пренебрежении уровне по отношению к принятым практикам. Утверждение должно быть доказано перевесом доказательств». Позже в этих комментариях мне нечего будет добавить к этому изложению ([https://grants.nih.gov/policy/research\\_integrity/overview.htm](https://grants.nih.gov/policy/research_integrity/overview.htm)).

\*\*см. [https://grants.nih.gov/policy/research\\_integrity/overview.htm](https://grants.nih.gov/policy/research_integrity/overview.htm)

гие формы отклонения от нормы. Следует также учитывать вероятность того, что все подобные технологии могут быть подвержены манипуляции со стороны изощренных злоумышленников, и опередить их будет нелегко.

Доступные данные по отзывам статей также представляют проблемы для исследования недобросовестного поведения. Что дает нам лавина статей, приведенных в *RetractionWatch* – блоге и базе данных, учрежденных в 2002 г., как средство, позволяющее сделать опубликованную науку «более прозрачной»? Путем сбора отзыва статей, опубликованных не только в совокупности журналов, но и по дисциплинам, и во времени, создатели хотели сделать более видимыми отозванные статьи и посредством этого сократить их использование другими, которые не осознают, что они могут использовать фальшивые или ошибочные научные статьи в своем собственном исследовании. Как выразил позже это один из ее основателей, они стремились выявить «лженауку» для неосведомленных ученых и также стать «окном по самокоррекции характера науки, ... что может обеспечить внимание к случаям научного обмана» [8].

По состоянию на январь 2020 г. эта БД привела список максимального числа 21 792 статей, которые почти все были отклонены авторами или издательствами. Это такое большое число, что ученые, ищущие надежные базы данных для использования в исследованиях по недобросовестному поведению, применяют именно ее. Однако есть также причины для скептицизма относительно надежности этих данных и исследований, их использующих [9].

Как знают те, кто анализировал опубликованные списки отозваний, подробности отозванных статей и причины, учитывающие их отклонение, весьма разнообразны. Некоторые из них более ценны, чем другие, но их большинство достаточно редко. Более того, причины, учитывающие отозвания, часто неоднозначны и, таким образом, трудны для интерпретации и классификации. Другие достаточно четкие, чтобы различать между теми, которые сообщают о недобросовестном поведении, и теми, которые оспаривают научные практики, и теми, кто фиксирует ошибки разного рода. (Я еще вернусь к ошибкам в тексте). До сих пор считается, что иные статьи отзываются из-за того, что сообщаемое свидетельство было представлено менее убедительным, чем более поздние результаты, собранные после представления статьи. Отозванные статьи содержат данные по анализу воздействия отклонений на цитирование до и после того, как произошло отклонение [10]. По воздействию отозваний на цитирование статей их соавторов, можно судить о наличии недобросовестного поведения, довольно часто явно не приводимого.

Другое направление исследования по недобросовестному поведению основывается на опросах

точек зрения ученых об относительной серьезности «спорных научных практик». Эти исследования расширяют перспективы отклонения от нормы в науке и вариации в ощущаемой тяжести отдельных актов. По сообщениям в этих статьях спорные научные практики многочисленны. В числе прочих они включают отсутствие у коллег права на авторство, нарушения правил защиты людей и животных, неудачу в поисках источников финансирования, которая может прекратить исследование, а также самоплагиат или «переработка текста», следы перепечатки чей-то собственной первой работы [11]. Некоторые из них явно нарушают правила финансирующих агентств или журналов, стремящихся не поощрять подобные практики в статьях, которые они публикуют. Другие нарушения, такие как «самоплагиат» (вероятный оксюморон) кажутся менее вредными, чем *FFP*, но многие журналы известны отказом публиковать статьи, содержащие ранее изданные тексты авторов, если дано недостаточное обоснование этого и того, что представленная рукопись содержит материал, заслуживающий публикации. Несмотря на существование единственного устава Национального института здравоохранения\*, в котором самоплагиат не считается научным недобросовестным поведением, он неприемлем для некоторых ученых, имеющих значительный послужной список исследований и многочисленные публикации, поскольку они подверглись публичной критике за воспроизведение своих более ранних работ. Вопрос этот остается буквально «спорной научной практикой».

В принципе данные, собранные высококачественными опросами, связанные со спорной научной практикой, дополняют наше понимание недобросовестного поведения и практики науки в целом. Могу добавить, что дальнейшее изучение причин убеждения ученых в том, что некоторые практики неприемлемы, тогда как другие – тривиальны, разумно, и что удвоение свидетельств о случаях спорной научной практики будет даже более значимым. А усиление акцента на теме, большее внимание на последствия спорной научной практики сделает более ясной надежность результатов и злоумышленников и их жертв для тех, кто ее отслеживает. Рассмотрение последствий недобросовестного поведения иногда, но не часто, фигурирует в обсуждениях затрат на них. Цена, которую такое недобросовестное поведение накладывает на состояние науки, – ключевой вопрос в эпидемиологии недобросовестного поведения, поскольку эпидемиология болезни – основа состояния общества.

Наконец, новая линия исследования недобросовестного поведения сфокусирована на влиянии

\* [https://grants.nih.gov/policy/research\\_integrity/overview.htm](https://grants.nih.gov/policy/research_integrity/overview.htm); и его эквивалент Национального научного фонда, <https://www.nsf.gov/oig/regulations/>.

организационных контекстов, в которых оно встречается, и провоцирует ряд вопросов, полностью соответствующих созданию эпидемиологии недобросовестного поведения в науке. Многообещающие исследования сфокусированы, например, на влиянии переменных, таких как степень социальной интеграции в организациях и готовность тех, кто работает в них, обходить правила и сообщать о недобросовестном поведении, когда, они верят, что оно произойдет [1, гл. 3]. Другим примером, является предположение, что контроль над научными практиками может быть менее последовательным, чем может показаться, когда организации разрешают «несоблюдение требований, считающихся непризнанными и/или неизменным», допускающими гибкость в характере контроля научной практики [12].

В любом случае при решении ученым каждой проблемы большое значение имеет выбор подходящего массива данных. Я не предполагаю, что количественное исследование по важным аспектам недобросовестного поведения невозможно. Я полагаю, что современные доступные базы данных, созданные для этих целей, плохо подходят для исследования недобросовестного поведения, и что много работы потребуется, чтобы усилить подручные данные, и еще больше потребуется для сбора новых и лучших данных для исследования. Начиная отсюда, двигаться будет нелегко.

## **НЕДОБРОСОВЕСТНОЕ ПОВЕДЕНИЕ: КАК ЕГО ОПРЕДЕЛЕНИЕ ОБЛЕГЧАЕТ И УСЛОЖНЯЕТ ИССЛЕДОВАНИЕ**

История институционализации правительственного контроля за исследованиями показывает, что достижение согласия между заинтересованными сторонами о недобросовестном научном поведении давалось не легко. Оно началось со слушаний Конгресса основной комиссией в 1981 г., запросившей публичного обнародования дел о недобросовестном научном поведении в четырех основных научных центрах в 1980 г. Выборочно было обнародовано 12 дел недобросовестного научного поведения в США в 1974 – 1981 гг. Внимание Конгресса к недобросовестному научному поведению было приковано на протяжении 1980-х гг. из-за дополнительных обвинений в недобросовестном научном поведении и отчетов о том, что национальный институт здравоохранения, университеты и другие научные организации неадекватно отвечали на эти обвинения\*. Этот отчет лишь намекает на противоречивый тон и содержание этих дискуссий. Обвинения были построены на растрате денежных средств налогоплательщиков на проведение исследований и на безответственности грантополучателей относительно контроля за тем, как было про-

ведено исследование и как были распределены данные целевые денежные средства. Эти обвинения попали в прессу и стимулировали дальнейшие противоречия.

Стало ясно, что консенсус должен был быть достигнут по тому, как недобросовестное научное поведение следует определять, по тому, какой феномен оно описывает и что остается за его рамками, а также необходимо понять, как и кто его должен изучать, как это надо рассматривать в суде, и какие наказания следует возложить на тех, кто виновен в этом преступлении. В *НИН*, основном спонсоре биомедицинских наук в США, в марте 1989 г. были учреждены Офис научной добросовестности и его организация-близнец, Офис по контролю за научной добросовестностью, с целью реализации этих задач. Первый был расположен в самом *НИН*\*\* , а второй – в офисе помощника секретаря по здравоохранению\*\*\*. То, что было создано два офиса, а не один, было ранним показателем организационной сложности, содержащейся в контроле за исследованиями. Национальный научный фонд, также основной спонсор исследований в США, позднее распределил свои собственные роли, связанные с недобросовестным научным поведением и то, как его следует обрабатывать. Он расположил эти виды деятельности в офисе генерального инспектора.

Потребовались годы переговоров, например, для замены общего термина мошенничество научным недобросовестным поведением, а научное недобросовестное поведение\*\*\*\* – недобросовестное поведение в ходе исследований, которое затем было подчинено проблеме поддержания «добросовестности в ходе исследований». Таким образом, произошел процесс нормализации, в котором недобросовестное поведение стало тем, чем оно не было, и в соответствии со способом, каким неподходящий термин научного недобросовестного поведения (которое после всего стало вовсе не научным) был не принят во внимание.

Это официальное определение, ограниченное фабрикацией, фальсификацией и плагиатом, было окончательно согласовано *НИН*, *NSF*, и федеральным правительственным отделением здравоохранения и обслуживания людей и др. в 1987-1989 гг. Потребовалось еще три года для национальной

---

\*\* Об истории регулирования *НИН* см.

<https://ori.hhs.gov/historical-background> [13].

\*\*\* <https://ori.hhs.gov/historical-background>. См. также [https://www.nsf.gov/oig/\\_pdf/presentations/session.pdf](https://www.nsf.gov/oig/_pdf/presentations/session.pdf).

\*\*\*\* Данный термин сам по себе эволюционировал от обмана и неправильного представления результатов до научного недобросовестного поведения и сегодня до целостности исследований, сигнализируя об усилиях классифицировать подобный феномен, включающий и усилия по формированию его публичного имиджа.

---

\* <https://ori.hhs.gov/historical-background>.

академии наук, чтобы принять его в 1992 г. Дальнейшие рассмотрения состоялись в 1996 и 1999 гг. в связи с продолжающимися и острыми спорами по поводу терминов, которые определение недобросовестного поведения должно содержать и как контроль должен осуществляться и кем [14].

Организационная сложность была помножена на требование всем получателям финансирования, включая колледжи и университеты, научные лаборатории, больницы и иные организации получать федеральные фонды, либо в США, либо за границей в целях соблюдения официальных правил и участия в научной политике.

Главными причинами, учитывая сведение недобросовестного поведения только к трем действиям (*FFP*), были такие, что первые определения «охватывали слишком многое», были неоднозначными и не готовыми к применению в запросах относительно недобросовестного поведения. Они слегка похожи с закулисными политическими спорами о том, что официальные утверждения не начинают учитывать нюансы сообщений, которые они стремятся обобщить.

После учреждения критериев *FFP* поднялось движение за унифицированное определение, которое можно было применять во всех организациях и во всех областях науки. (Федеральное правительство США приводит список по крайней мере из 35 отдельных организаций и дополнительных обособленных подведомственных организациях, занимающихся поддержкой научного исследования в той или иной форме). Снова, множественные обмены между релевантными деятелями завершились фокусом на стандарте *FFP*, причиняющим наибольший вред всем, имеющим отношение к публичному документированию науки.

Это оставляет ряд практик, некоторые из которых считаются неприемлемыми, без списка действий, наказываемых федеральным правительством, – решение, ориентированное быть административно практически (см. раннюю дискуссию о спорных научных практиках). Даже сейчас кажется, что унификация определения недобросовестного поведения в ходе исследования была не вполне достигнута, например, по тому, нарушают ли основной стандарт поведения или нет рецензенты, которые недобросовестно используют в своем собственном исследовании любые идеи, методы или результаты, полученные в ходе рецензирования, без указания их источников? *NSF* явно препятствует такому недобросовестному использованию, но не все другие организации так делают. Что касается сегодняшнего дня, то *FFP* стандарт, кажется, сохранил некоторую поддержку в научном сообществе, но присутствие правительства в процессе контроля продолжает раздражать тех, кто все еще верит в давно поощряемое в науке обязательство «самоконтроля».

Сужение определения и противоречия вокруг него тесно связаны с качественными исследованиями недобросовестного поведения. Их фиксация ограничена масштабом официально маркированного недобросовестного поведения, и таким образом оказало влияние на создание данных, собранных многими группами по контролю, более унифицированными, чем в противном случае они могли бы быть; полезнее было бы проводить эмпирические исследования на основе официальных документов. В то же время объяснение того, какими могли бы быть эти три формы недобросовестного поведения (*FFP*), сложнее, чем кажется. Рассмотрим нынешнее определение недобросовестного поведения, распространенное *NSF*: недобросовестным поведением является «форма существенного отклонения от принятых практик научного сообщества поддержания добросовестности научного документа». Недобросовестное поведение должно быть обязательно намеренным или осознанным или заключаться в безответственном пренебрежении к принятым практикам. Акцент на намерении, осознанности или безответственном пренебрежении должен ослабить предположение о самообмане как источнике виновности при недобросовестном поведении. Самообман имеет иную историю в науке, включая неспособность ученых подчинять свои собственные идеи и исследования скептицизму так же строго, как они обязаны подчинять ему работу других. Исключение самообмана из предположений о недобросовестном поведении не является равнозначным освобождению нарушителей уголовного права от ответственности за их действия, если они неспособны оценить серьезность того, что они делают, и вводит меру пощады в оценку недобросовестного поведения. Однако, намеренный или нет вред, причиненный исследовательскому документу (и его влиянию на доверие ученых) является одинаковым, независимо от того, имеет ли злоумышленник намерение или нет. Мне непонятно, почему так много внимания уделяется намерению обмануть в оценке недобросовестного поведения. Возможно потому, что есть прецеденты в нарушениях правил или они допускают некоторую меру терпимости, которую исследователи считают необходимой.

Акцент на сохранность и целостность записей исследований в официальных заявлениях приравнивает эти качества научного документа\* к благополучию самой науки. Поддержке доверия среди ученых и необходимости для них полагаться на работу другого ученого, опубликованную или нет, как ни странно, не уделяется должного внимания. Доверие среди ученых является важной основой скорее для прагматического, чем убедительного обоснования.

\* [https://grants.nih.gov/policy/research\\_integrity/overview.htm](https://grants.nih.gov/policy/research_integrity/overview.htm).

В качественных исследованиях недобросовестного поведения ясность относительно значения фабрикации, фальсификации и плагиата – абсолютное требование. Вопросы также возникают относительно того, как много фабрикация или фальсификация данных может быть использовано, чтобы определить их как недобросовестное поведение? Аналогично должны изучаться вопросы о степени, в которой работа других является недобросовестно использованной. Как много текста из чужой работы может быть «заимствовано» без указания ее атрибутов, чтобы определить это как плагиат? Если это случилось, плагиат является одним из трех действий, считающихся наказуемым нарушением, и юридический закон действует только тогда, когда можно считать, что процитировано слишком много текста, собственного и чужого без разрешения. Аналогично, существует значимый юридический документ, касающийся воспроизводства рисунков без авторизации. Закон, касающийся воспроизводства как текста, так и изображений, специально защищает интеллектуальную собственность и, таким образом, финансовые интересы авторов и художников. В науке, тем не менее, кража текста и рисунков также может нанести финансовый урон, такой же важный или даже больший, каким являются потери коллегами престижа и репутационных наград, с ним связанных.

Трудной является оценка потерь авторов и художников, чьи работы подверглись плагиату, и затем компенсировать их им, так же трудно оценить потери ученых из-за кражи их интеллектуальной собственности, и нелегко компенсировать символические награды признания, которое они бы получили, наравне с финансовой компенсацией, часто с ней связанной. Вкратце, научное недобросовестное поведение – трудный феномен для изучения, если предполагается качественное исследование\*.

## **АКТУАЛЬНЫЙ И ПРЕДСТОЯЩИЙ ФОКУС ВНИМАНИЯ В ИССЛЕДОВАНИИ: КРИЗИС РЕПЛИКАЦИИ, РОСТ ОТЗЫВА СТАТЕЙ И ТЕХНОЛОГИИ ОБНАРУЖЕНИЯ FFP**

### **Кризис репликации**

Предоставляет ли кризис репликации новые возможности для качественного исследования по недобросовестному поведению? В США средства массовой информации уделяют большое внимание требованиям к опубликованным результатам исследований, особенно, но не только в экономике, медицине и социальной психологии, чтобы они не терпели неудачу в попытках их репликации. Эти неудачи возникают не из-за внутренних трудно-

стей, складывающихся из-за того, что трудно их реплицировать, как например, эксперименты сфокусированные на феноменах, которые при желании нельзя повторить, такие как редкие астрономические явления. Напротив, кризис репликации возникает, как утверждает, благодаря ученым, применяющим часто используемые, но ошибочные исследовательские проекты и аналитические методы. Здесь фокус лежит не только на недобросовестном поведении отдельных исследователей, но рождается из распространенного принятия научных практик, вносящих противоречия в результаты исследований.

Некоторые высказывают крайние мнения «о росте обеспокоенности в том, что в современных исследованиях ложные результаты могут быть большинством или даже подавляющим большинством опубликованных научных утверждений» [15]. Критикуемые практики включают использование очень небольших выборок, анализ слабых эффектов, почти полное отсутствие усилий со стороны исследователей реплицировать свои собственные результаты и использование недостаточно требовательных статистических погрешностей, самое известное из которых – классическое измерение  $p \leq .05$  [16].

Кризис репликации по понятным причинам привлекал значительный интерес тех, кто проводит исследование самих исследований. Действительно, некоторые из самых активных исследователей по кризису репликации характеризуют свою работу как закладывающую основы новой дисциплины, «метанауки о науке», направленной на улучшение научных методов и практик [17].

Кризис репликации также создал предложения по восстановлению, например, предложения заменить общий статистический стандарт значимости  $p \leq 5\%$  с помощью одного только  $p \leq 1\%$  [18]. Повышение ставки важности таким способом значительно сократит шансы исследователя не получить никаких «значимых» результатов и таким образом грозит отсутствием опубликованных результатов.

Другие предложения включают рост размеров выборок, использование значительно большей рандомизации, найм многих команд для независимой работы над одними и теми же вопросами, движение к практике, называемой «предрегистрацией» научных планов. В предрегистрационных протоколах от авторов, желающих опубликоваться, требуется представить их научные предложения (а не их готовые статьи) в журналы, согласные их рассмотреть. Предложения излагают заранее используемые методы для сбора и анализа данных, а также их ожидаемые результаты. Перед журналами тогда стоит задача определить, удовлетворяют ли установленным у них стандартам полученные ими предложения, и если планы удовлетворяют им, и они осуществимы, то публикация результатов

\* См. [1, с. 19-62] для анализа 748 судебных дел, взятых из БД NIH, NSF, данных из Дании, книг и интернета.

предрегистраемого исследования обеспечивается. Она ставит целью препятствовать авторам в отборе и сообщении результатов, появившихся «случайно», несмотря на их удовлетворение условных мер «значимости». (термин, сейчас используемый для комплектации и выбора согласно значимости – это *p-hacking*).

Предрегистрация имеет как критиков, так и сторонников. Некоторые ученые указывают на неизбежную мультипликацию статей, содержащих ожидаемые и вероятно скучные результаты. Другие предполагают, что предрегистрация может быть полезна, если она соотносится с исследованием, которое достигло стадии проверки гипотезы, но не подходит для исследования, нацеленного на создание гипотез. Другие ученые все еще выражают обеспокоенность относительно того, что она сократит мотивы ученых принимать всерьез неожиданные результаты и гораздо реже рассматривать то, как они могли возникнуть. Большие возможности серендипности в исследовании (когда она случается) не следует забывать при стремлении применить один ко всем набор правил [19]. Предрегистрация может быстро увеличить быстроту репликации и определить тех, кто готов сообщать результаты, появляющиеся случайно, но это затратное определение по отношению к слабой научной практике, которая может быть не поддержана иным способом.

### **Отзыв статей как потенциальные данные для качественного анализа недобросовестного поведения**

Абсолютное число отозванных публикаций не только огромно, как отмечалось ранее, но увеличивается ежегодно. Тем не менее, если учесть темпы роста числа опубликованных статей, рост числа отзывов фактически выровнялся с 2012 г. То же самое свойственно росту недобросовестного поведения и иных феноменов, о которых отзывы статей могут сигнализировать. Однако сырые данные числа отзывов, публикуемые сейчас в базе данных по ретракции, являются очень большими, и по одной этой причине они, кажется, привлекают ученых, ищущих качественную информацию о недобросовестном поведении. Но насколько полезной она является для таких исследований?

Ретракции предназначены выделять опубликованные статьи, содержащие свидетельство об исследовании, которое ненадежно, либо ошибочно, либо обманно. Ошибки встречаются в научных проектах, даже когда удовлетворены оговоренные стандарты практики. Ученые знают это. Они могут также вытекать из некомпетентной или некачественной научной практики (включая, например, загрязнение выборок). В ранней статье по недобросовестному поведению [2] я проводила различие между «репутационными» ошибками (те, что воз-

никают несмотря на большие усилия ученых избежать их) и «нерепутационными» ошибками (возникают из-за несоответствия принятой научной практики). Если учитывается намерение в оценке случаев недобросовестного поведения, то авторы, допускающие репутационные ошибки, явно невиновны. Но те, кто игнорирует современные стандарты практики, намеренно или нет, виновны. Оба класса ошибок недооценивают надежность научного свидетельства, вне зависимости от стоящего за ним намерения. И оба класса ошибок тратят впустую время младших научных сотрудников, которые их учитывают и еще больше тех, кто следует им в исследовании.

Таким образом, если ретракции были признаны полезным свидетельством для исследований о недобросовестном поведении, они должны быть достаточно детализированы, чтобы позволить ученым классифицировать ретракции в три группы: те, кто включает недобросовестное поведение, соответствующее стандарту *FFP*, те, кто включает репутационные ошибки, и те, кто включает не репутационные ошибки, близкие к определению *NSF* о виновности из-за «безответственного пренебрежения» к текущим стандартам. Тексты большинства ретракций не позволяют совершать такие различия. Более того, отдельные журналы публикуют сведения о ретракциях, а другие – нет. Области, в которых публикуются ретракции, также разнообразны, как и взгляды редакторов на необходимость публиковать ретракции. Раз так, то в их сегодняшнем состоянии данные ретракции – проблемные показатели недобросовестного поведения\*.

Также утверждается, что рост числа ретракций является свидетельством того, что рецензирование, являясь по существу барьером для недобросовестного поведения, терпит неудачу в его предотвращении. Выявляет ли рецензирование обманное представление исследования от достоверного? Вероятнее всего нет, так как часто разрешает публикацию обманных статей. Но может ли рецензирование обслуживать эту цель? Ответ на этот вопрос далек от ясности. Оно может служить в качестве определителя обманных статей, поскольку авторы не выбирают строго рецензируемых журналов, чтобы избежать подобного рассмотрения. С другой стороны, рецензенты обычно не имеют доступа к сырым данным, на которых основано исследование, и, следовательно, не способны проверить сами данные, на которых основано исследование и сделаны выводы. Представляется возможным спросить редакторов о том, как много представленных в их журнал статей было отклонено, поскольку они

\* См. [20, с. 817-823] для анализа противоречивых результатов в публикациях, утверждающих оценку различного рода недобросовестного поведения с использованием данных ретракции.

или рецензенты посчитали их обманными. Существует ли подобное свидетельство, широко не известно. В целом, рецензирование стремится отделять те статьи, которые удовлетворяют стандартам научной практики и предоставляют надежные результаты, от тех, которые им не удовлетворяют. И действительно так функционирует рецензирование в науке.

Исследование, в которое я была вовлечена годы назад, имело две эмпирические цели: первая – сравнить быстроту ретракций журналов в естественных, общественных и гуманитарных науках, и вторую – опробовать способы, которыми рецензенты обрабатывают представления, полученные от ученых разного профиля. Нам удалось узнать: быстрота принятия решений гораздо больше в естественных науках относительно других дисциплин на десятки лет [21]. Мы предположили, что относительно более короткие сроки принятия решений в естественных науках зависят отчасти от преобладания практики рецензирования до представления их коллегам и отчасти в результате преобладания более унифицированных стандартов в науках, в которых публикуется принятая статья. Более того, что касается особой важности этого обсуждения, когда речь идет о высоких скоростях принятия решений, редакторы журналов, изучаемых нами подробно, отвечали, что они верят, что рецензенты отделяют обоснованные представления от остальных и что непосредственно достаточно хорошее исследование будет преобладать и менее удовлетворяющее исследование «потонет» в огромной массе опубликованных статей, используемых реже, если они вообще не утонут. В результате, высокая скорость принятия статей к публикации, благоприятствующая им, позволит сообществу оценить то, что чрезвычайно полезно, а что нет, и что будет эффективной стратегией издания.

Заслуживает внимания, что ни одна из огромного числа рецензируемых статей, сформировавших основу исследования, не была определена как обманная. Это не говорит о том, что оценки не были сделаны относительно доверия к представлению статей, а просто те, что потерпели неудачу в удовлетворении теста, были отклонены. Значимым было то, что определенное число статей имело пометку «CP», что на жаргоне издателей в свое время подразумевало «псих». Авторы этих отклоненных статей ими и были. (Некоторые из этих представлений, написанные карандашом на разлинованной бумаге, являются показателем того, что наше исследование проводилось давно). Редакторами не было сделано ни одного усилия, чтобы определить, были ли эти статьи обманными, или просто целиком неверными; они просто не были опубликованы.

До тех пор, пока многое не узнается о «демографии» журналов, которые печатают ретракции, включая дисциплины, ими представленные, о причинах, цитирующих ретракции, ими опубликован-

ные, и о тех, кто принимает решение об отклонении, исследования о недобросовестном поведении, основанные на базах данных по ретракции, должны рассматриваться как носящие более предполагающий характер, чем как надежные.

### **Технологии обнаружения FFP**

Использование новых технологий, сегодня доступных для обнаружения недобросовестного поведения, еще не изучалось (по крайней мере, по моим сведениям); также не оценивалась их эффективность. В числе тех, которые часто упоминаются, поиск плагиата, статистические методы идентификации данных «слишком хороших, чтобы быть правдой», и программное обеспечение, идентифицирующее фотографии и графические изображения, которые были подвергнуты манипуляции или редактированию. Поиск плагиата, говорят, используется инструкторами колледжей, чтобы проверить оригинальность статей, представленных студентами. Он также применяется отдельными журналами для идентификации представлений, содержащих плагиат материала. Статистические методы анализа применяются для данных, чтобы идентифицировать те, которые слишком хороши, чтобы быть правдой: которые вероятно (но не доказано) должны содержаться в ряде дополнений или отклонений. Сегодня возможно использовать программное обеспечение, чтобы выявлять, «подделывались ли» изображения или графики. Информация о том, как получить доступ к этим программам, доступна на сетевом сайте *NIH*, который подробно детализирует свои ресурсы для защиты добросовестности исследований. Определение того, как часто эти технологии используются и кем и при каких обстоятельствах, а также степень их эффективности в идентификации научного недобросовестного поведения, по-видимому, обеспечит некоторое свидетельство относительно случаев недобросовестного поведения, которые они стремятся выявить. Но, безусловно, нет способа узнать, сокращает ли их нынешнее использование уровень недобросовестного поведения относительно недобросовестного поведения, случившегося в прошлом, или выявляет ли их использование недобросовестное поведение, которое иначе бы произошло. Нам неизвестно сегодня, насколько обоснованными они оказываются и как часто их использование приводит к отклонению статей для публикации, к ретракции плохих статей или к идентификации иначе ненадежного исследования. Лучшее знание об эффективности этих изобретений было бы полезным.

### **ПОСЛЕДСТВИЯ НЕПРАВОМЕРНОГО ПОВЕДЕНИЯ В НАУКЕ ДЛЯ ДОВЕРИЯ К НАУКЕ**

Продолжение внимания медиа к недобросовестному научному поведению, к кризису репликации и к сегодняшней частоте ретракций публика-

ций кажется обоснованной причиной для публичного скептицизма о кредитоспособности науки. Отчеты медиа уверенно ведут к росту беспокойства среди ученых о готовности общества поддержать научные и технические исследования (по крайней мере, так говорят ученые). Подручные сведения, основанные на многолетних обзорах общественного доверия к науке в США, тем не менее, предполагают иное. Повторные опросы о доверии общества к науке за последние три десятилетия или больше, не показывают основных изменений в измерениях доверия. Существенная доля граждан США говорит, что они имеют средний или высокий уровень доверия как к науке, так и к лидерам научного сообщества. Существенная часть людей верит, что ученые «работают на благо гуманности» [22]. Я не буду комментировать вывод о том, что наука в США оценивается вторым номером по доверию в обществе среди всех занятых групп только у военных, и выше стоит у физиков, клерков, судей, политиков и юристов. (Важность этого вывода предоставляется на суд читателей). Тем не менее, было бы неправильно обойти стороной и не отметить существование сильных антинаучных настроений среди некоторых секторов популяции США. «Антивакцинисты», которые считают вакцину от болезней вредной, многочисленны, особенно в определенных религиозных группах и в отдельных регионах. Это они отрицают науку об изменении климата как не имеющую под собой научной основы. Те, кто не согласен с ними, придерживающиеся экзотических взглядов на диету и физические упражнения, также многочисленны, начиная от розничных покупок потребляемых ими продуктов, и веру в астрологию, долгое время были основной популяцией в США. Какими бы ни были источники этих убеждений, недобросовестное научное поведение, кажется, не усиливает их и широко не распространяется. Несмотря на обоснованное внимание в прессе к недобросовестному поведению в науке, публичная поддержка науки кажется не подверженной этому воздействию.

### **КАКИЕ УРОКИ МОЖНО ИЗВЛЕЧЬ ИЗ НЕДОБРОСОВЕСТНОГО ПОВЕДЕНИЯ И ЗАЩИТЫ КАЧЕСТВЕННОГО ИССЛЕДОВАНИЯ ЭТОГО ФЕНОМЕНА?**

1. Размышление о недобросовестном поведении с точки зрения эпидемиологии, т.е. фокусировании на вопросах об его проявлениях и распределении, эффективности средств контроля или их восстановления, и бережный анализ характеристик «популяций под риском» в приведенных примерах поможет сформулировать вопросы, заслуживающие ответов, идентифицировать те, которые можно изучать, и определить, что нужно сделать, чтобы сделать исследование возможным. Адаптация этой перспективы сможет управлять решениями,

которые исследователи принимают при использовании доступных баз данных или сборе новых свидетельств для исследования. Это было наблюдение, сделанное мной в 1977 г. Это остается скорее вдохновением, чем реальностью.

2. В использовании подручных данных, а также сборе новых обоснованных и надежных данных по научному «недобросовестному поведению» полно подводных камней, таких как встречаются у криминалистов при изучении незаконного поведения. Публичная документация по недобросовестному поведению противоречива. Нет ни одного доступного свидетельства по обвиняемым в нем, и тем, кто оправдан или просто признан невиновным. Самоотчеты являются спорными по разным причинам, так же, как и отчеты очевидцев. Таким образом, усилия для оценки причин недобросовестного поведения, его распределения, и имеющимся его проявлениям также, весьма вероятно, являются неоднозначными. Тем не менее, исследование криминального поведения имеет гораздо более длинную историю, чем недобросовестного поведения. Бережное исследование уже опробованных методов и встречающихся препятствий может принести пользу исследованию недобросовестного поведения.

3. Разногласия относительно того, лучше ли описывается недобросовестное поведение как «вершина айсберга» (очень часто) или как результат случайных «плохих яблок», загрязняющих научный баррель (очень редко), нельзя разрешить в свете доступной информации. Причины в защиту той или иной позиции говорят меньше о недобросовестном поведении и его случаях, чем о тех, кто их поддерживает.

4. Несмотря на основные проблемы с данными по недобросовестному поведению, возможности качественного анализа существуют и должны изучаться. Приведу один пример. Официальные данные по характеристикам ученых, признанных виновными в недобросовестном поведении, имеют свои недостатки (они, кроме всего, являются только теми, которые были найдены в некоторой неизвестной большой группе злоумышленников). Все-таки возможно идентифицировать характеристики тех, кто вовлечен в фабрикацию, фальсификацию или плагиат (в один из них, или два, или все три), узнать, посредством каких способов их недобросовестное поведение было открыто, в какого рода организациях недобросовестное поведение встречается, какие исследовательские процедуры работают и какова тяжесть наложенных наказаний. Знание об этих случаях должно открыть, какая же система контроля работает, если это не эпидемиология широко воспринимаемого недобросовестного поведения. Определение характера вопросов, на которые могут ответить доступные данные, является частью проблемы.

5. Дальнейшее внимание к последствиям, которые влечет недобросовестное поведение для документации науки, следует поощрять. Это открывает ряд вопросов об эффектах недобросовестного поведения, которые едва ощутимы. Линия интереса должна включать дальнейшее исследование быстроты цитирования и сопротивления ему на публикации, признанные нарушающими стандарт *FFP*. Есть ли и какая разница в том, имеет ли подобное исследование влияние на последующие исследования. Юджин Гарфилд [23, 10] начал проводить такого рода исследования десятилетия назад, и его выводы остаются актуальными [24]. Другие исследователи начали изучать влияние отозвания статей (независимо от причин ретракции) на репутацию первых авторов и их соавторов. Есть указание на то, что цитирования, принадлежащие первому автору ретракционных статей, падают после объявления ретракции, и есть также свидетельство, что их соавторы страдают от своего рода коллатерального вреда [25]. Эти выводы являются высоковероятными, но нуждаются в большем подтверждении. Помимо учета эффектов цитирования сложная работа требуется по последствиям недобросовестного поведения и обвинениям в нем для карьеры ученых, при опоре как на более длинный, так и на краткий срок, для институтов, в которых они работают, для их коллег и для быстроты возникновения идей, которые они стремятся воплотить.

Читатели могут посчитать эти комментарии о научном недобросовестном поведении в защиту проведения качественного исследования по нему информативным и даже возможно наводящим на размышление.

Журнал *Quantitative Science Studies* будет большим фондом по вопросам общественного исследования науки для тех из нас, кто полагает, что наука является чрезвычайно важной частью общественной жизни.

**Финансирование.** Данное исследование не получило никакого финансирования.

## ЛИТЕРАТУРА

1. Ben-Yebuda N., Oliver-Lumerman A. Fraud and misconduct in research: Detection, investigation, and organizational response. — Ann Arbor, MI: University of Michigan Press, 2017.
2. Zuckerman H. Deviant behavior and social control in science/ E. Savarin (Ed.), Deviance and social change. — Beverly Hills, CA: Sage, 1977. — P. 87–138.
3. Zuckerman H. Norms and deviant behavior in science // Science, Technology, & Human Values. — 1984. — Vol. 9, No. 1. — P. 7–13.

4. Steneck N. H., Mayer T. A global guide to the responsible conduct of research. — London: Imperial College Press, 2017.

5. Holton G. “We do our damndest”: The evolution of trust in scientific findings / D. H. Guston and K. Keniston (Eds.) The fragile contract: University science and the federal government (pp. 59–81). — Cambridge, MA: MIT Press, 1994.

6. Hixon J. R. The patchwork mouse. — Norwall, MA: Anchor Press, 1976.

7. Radick G. Beyond the Mendel-Fisher controversy // Science. — 2015. — Vol. 350, No. 6257. — P. 159–160. — <https://doi.org/10.1126/science.aab3846>

8. Oransky I., Marcus A. Why write a blog about retractions. — Retraction Watch, 2010. — <http://retractionwatch.com/2010/08/03/why-write-a-blog-about-retractions/> Retrieved February 27, 2020.

9. Brainard J., You J. What a massive database of retracted papers reveals about science publishing’s “death penalty.” // Science. — 2018. — Vol. 25, No. 1. — P. 1–5.

10. Garfield E. What do we know about fraud and other forms of intellectual dishonesty in science? Part 1: The spectrum of deviant behavior in science // Current contents. Clinical Medicine — 1987. — N. 15(14). — P. 3.

11. Horbach S. P. J. M., Halfman W. The extent and causes of academic text recycling or “self-plagiarism.” // Research Policy. — 2019. — Vol. 48, No. 2. — P. 492–502. <https://doi.org/10.1016/j.respol.2017.09.004>

12. Leabey E., Montgomery K. The meaning of regulation in a changing academic profession / J. C. Hermonowicz (Ed.), The American academic profession: Changing forms and functions (pp. 295–311). — Baltimore, MD: Johns Hopkins Press, 2011.

13. Steneck N. H. Research universities and scientific misconduct: History, policies, and the future // Journal of Higher Education. — 1994. — Vol. 65, No. 3. — P. 310–330.

14. Wadman M. Hostile reception to US misconduct report // Nature. — 1996. — Vol. 381. — P. 639.

15. Ioannidis J. P. Why most published research findings are false // PLOS Medicine. — 2005. — Vol. 2, No. 8, e124.

16. Fanelli D., Costas R., Ioannidis J. P. Meta-assessment of bias in science // Proceedings of the National Academy of Sciences. — 2017. — Vol. 114, No. 14. — P. 3714–3719.

17. Ioannidis J. P., Fanelli D., Dunne D. D., Goodman S. N. Meta-research: Evaluation and improvement of research methods and practices // PLOS Biology. — 2015. — Vol. 13, No. 10.

18. Fanelli D. Is science really facing a reproducibility crisis, and do we need it to? // PNAS. — 2018. — Vol. 115, No. 11. — P. 2628–2631. — [www.pnas.org/cgi/doi/10.1073/pnas.1708272114](http://www.pnas.org/cgi/doi/10.1073/pnas.1708272114)

19. Shiffrin R. M., Börner, K., Stigler, S. M. Scientific progress despite irreproducibility: A seeming paradox //

Proceedings of the National Academy of Sciences. — 2018. — Vol. 115, No. 11. — P. 2632–2639.

20. *Hesselmann F., Graf V., Schmidt M., Reinhart M.* The visibility of scientific misconduct: A review of the literature on retracted journal articles // *Current Sociology*. — 2017. — Vol. 65, No. 6. — P. 814–845. — <https://doi.org/10.1177/0011392116663807>

21. *Zuckerman H., & Merton R. K.* Patterns of evaluation in science: Institutionalization, structure and functions of the referee system // *Minerva*. — 1971. — Vol. 9, No. 1. — P. 66–100.

22. *National Science Board.* Science and engineering indicators 2018. — Washington, DC: National Science Foundation, 2018.

23. *Garfield E.* What do we know about fraud and other forms of intellectual dishonesty in science? Part 2: Why does fraud happen and what are its effects? // *Current Contents. Clinical Medicine*. — 1987. — N. 15(15). — P. 3–10.

24. *Zuckerman H.* The sociology of science and the Garfield effect: Happy accidents, unanticipated developments and unexploited potentials // *Frontiers in Research Metrics and Analytics*. — 2018. — Vol. 3. — Article 20. — <https://doi.org/10.3389/frma.2018.00020>

25. *Azoulay P., Bonatti A., Krieger J. L.* The career effects of scandal: Evidence from scientific retractions // *Research Policy*. — 2017. — Vol. 46, No. 9. — P. 1552–1569.

## Предсказание предметных направлений международных исследований в области библиотечного дела и информатики до 2030 года с использованием метода глубокого обучения\*

**Фарзана ГНАДИ НЕЖАД  
(Farzaneh GHANADI  
NEZHAD)**

Университет Шахида Чамрана в Ахвазе,  
г. Ахваз, Иран

**Фарида ОСАРЕХ  
(Farideh OSAREH)**

Университет Шахида Чамрана в Ахвазе,  
г. Ахваз, Иран

**Мохаммед Реза ГАНЕ  
(Mohammad Reza GHANE)**

Отдел оценки и развития сбора данных,  
Региональный информационный Центр  
науки и технологий (RICeST), г. Шираз, Иран

*Данное исследование стремится предсказать предметные направления исследований в области библиотечного дела и информатики до 2030 года на основе моделирования предыдущих тем исследований в данной области, которое было сделано с использованием интеллектуального анализа текста и подхода глубокого обучения. Чтобы построить модель предыдущих исследований и предсказать возможные темы были использованы алгоритмы глубокой нейронной сети после предварительной обработки и тематической классификации исследований. В исследовательский массив были включены 90311 журнальных статей в библиотечных и информационных изданиях, проиндексированных на веб-сайте Web of Science за период с 1945-2020 гг. Все исследовательские процессы были реализованы на языке программирования Python. Результаты показали, что наибольшее количество исследований в будущем будет связано с Интернетом и веб-исследованиями, и темпы роста этих тем со временем будут выше. Однако темы, связанные с библиотеками и их рабочими процессами, а также с другими традиционными дисциплинами, такими как теоретические основы, будут иметь более низкие темпы роста в изучении библиотековедения и информатики. Отсюда следует, что знание важных предстоящих проблем, помогающее планировать будущие исследования, может выявить пробелы в исследованиях и инвестиционные возможности в секторе исследований и разработок, тем самым поддерживая исследователей, университеты и соответствующие исследовательские институты в разумном выборе проектов.*

\*Перевод Ghanadinezhad F., Osareh F., Ghane M. R. Forecasting the subject trend of international Library and Information Science research by 2030 using the Deep Learning approach//International Journal of Information Science and Management. — 2022. — Vol. 20, No. 1. — P/ 471-492. — <https://ijism.ricest.ac.ir/index.php/ijism/article/view/2361/559>

## ВВЕДЕНИЕ

В эпоху стремительного развития, особенно цифровых технологий, в библиотечном деле и информатике чувствуется целесообразность в систематических исследованиях и размышлениях о будущем с тем, чтобы адаптироваться к современным потребностям. Планирование и подготовка к решению будущих проблем и поддержанию статуса и выживания в условиях конкуренции требуют таких системных исследований, чтобы в этой области можно было отслеживать и следить за развитием событий и прогнозировать будущие изменения. В противном случае область останется вне конкуренции и потеряет свои позиции в будущем. Как указывает Годе [1], с помощью будущих исследований, в дополнение к анализу проведенных тенденций, можно придумать и оценить возможное, вероятное и желательное будущее. Обзор исследований показывает, что исследователи в различных областях осознали важность и сложность проблемы перспективных исследований и прогнозирования предметных направлений научных работ и всегда пытались спрогнозировать будущие тенденции публикации с разных точек зрения, используя множество методов. Выявление возникающих проблем является одной из наиболее важных целей, используемых для прогнозирования предметных направлений научных работ. Анализ возникающих проблем – это условие будущих исследований, стратегического управления и разработки политики, описывающее процесс выявления тем, которые не были важными или влиятельными в прошлом, но могут оказать влияние в дальнейшем [2].

Для определения новых тем и прогнозирования перспективных направлений исследований в подобных работах использовались различные методы. Одним из распространенных подходов в этой области является использование мнений и опыта экспертов [3-6]. Однако следует отметить, что эксперты могут предвзято относиться к прогнозированию будущего или что те, кто участвует в опросе, могут обладать недостаточным опытом. Также может быть сложно рассчитать уровень достоверности прогнозов. С другой стороны, из-за постоянного увеличения научных разработок и их быстрого развития профессионалам трудно поспевать за быстрыми изменениями в исследованиях [4; 7-8]. Эти изменения делают использование подходов, основанных на экспертных заключениях, менее практичным. В настоящее время больше внимания уделяется автоматизированному и масштабируемому методам расчета передовых направлений научного производства. Также рассматривалось прогнозирование предметной направленности исследования полностью автоматическим способом. Де Солла Прайс [9] впервые ввел концепцию активных областей исследований как растущих сетей, чтобы показать будущее направление тем. Позже анализ

совместного цитирования был предложен Смоллом (1973, цитируется в [10]) для определения новых тем в дальнейших исследованиях. В 1974-2015 гг. для выявления новых предметных областей в исследованиях в основном использовался анализ сетей цитирования через анализ совместного цитирования, библиографических пар и прямого цитирования [10, 11-12]. Можно отметить, что в данный период в этой области преобладают подходы, основанные на цитировании [13]. Во время и после этого периода в некоторых исследованиях использовались подходы, основанные на лексике, такие как анализ ключевых слов и анализ сопутствующих слов, для изучения будущего направления исследований [14, 15, 12, 16-19].

Следует отметить, что использование таких методов количественного анализа, как библиометрия, наукометрия и информатика, для определения направлений будущих исследований имеет некоторые ограничения, т.е. временные задержки, возникающие в результате использования методов, основанных на цитировании (поскольку требуется время для цитирования исследований). Это приводит к тому, что библиометрический анализ и анализ совместного цитирования основываются на более старых источниках, и их результаты не могут показать последние тенденции в тематике и будущее направление исследований. Более того, большее количество цитирований может не обязательно указывать на более высокое качество исследовательских работ и тем. Это происходит потому, что люди могут больше цитировать свои исследования, своих коллег или журналы, которые они выбирают для публикации своих результатов. Такие проблемы и растущий объем текстовых ресурсов привели к ограничениям в применении подходов, основанных на анализе лексики, таких как гомологичный метод, для будущих исследований научных продуктов. Эти проблемы заставили исследователей задуматься об альтернативных или дополнительных решениях и методах для определения новых областей исследований и перспектив будущих исследований [8].

С 2015 г. произошли изменения в определении новых тем исследований и подходов к прогнозированию направлений научного производства. В связи с этим авторы работы [20] предложили основанную на ключевых словах модель для анализа данных патентных баз данных, чтобы отслеживать новые технологии с помощью связей между ключевыми словами и методом документа с обратной частотой терминов (TF-IDF). В другом исследовании [21] были определены и представлены показатели, позволяющие определять актуальные темы для того, чтобы выявлять новые области исследований и улучшать наукометрические показатели. В этот период, наряду с развитием крупномасштабных технологий обработки текстов, постепенно использовались методы машинного обуче-

ния для выявления новых областей и прогнозирования исследовательского процесса из-за их высокой эффективности при анализе больших объемов данных [22]. Использование онтологий и отношений ключевых слов [4; 23], а также интеллектуальный анализ данных, интеллектуальный анализ текста и алгоритмы машинного обучения [3, 24, 25, 23, 26-28, 7, 6, 29-30] являются другими стратегиями, которые могут помочь определить новые темы и приоритеты будущих исследований. Как отмечают авторы [31], в прогнозировании науки и технологий в настоящее время предпринимается больше усилий для использования не традиционных наукометрических методов, а подходов машинного обучения, глубокого обучения и методов моделирования по поиску закономерностей и точных взаимосвязей между данными и правилами вывода, которые позволяют определить будущие тенденции.

Изучение будущего библиотечного дела и информатики проводится в разных измерениях. Одним из важных направлений является изучение потребностей в исследованиях и будущих областях исследований. Из-за междисциплинарного характера библиотечного дела и информатики особое значение имеет внимание к будущим предметным областям в этой области. В основных исследованиях, касающихся предметных тенденций в библиотечно-информационных исследованиях, анализируются статьи, опубликованные в журналах в разные периоды [32-34]. Они касались тематического анализа диссертаций [35, 36], докладов конференций [37], структуры и картографирования областей [38-40], контент-анализа коллекции текстов в этой области, включая статьи, диссертации и т.д. [41], а также контент-анализа содержания других текстов в данной области, включая приоритеты исследований [42]. В упомянутых исследованиях преимущественно анализировались прошлые и нынешние тенденции научных результатов в этой области с использованием контент-анализа и наукометрических подходов. Внимание к предметным направлениям будущих исследований в этой области было несколько упущено.

Настоящее исследование направлено на предсказание будущих исследований в области библиотечного дела и информатики на основе более ранних исследований в этой области. Представление о будущих предметных направлениях в этой области поможет исследователям определить наиболее важные проблемы и потребности и принять соответствующие меры для выбора темы исследований. Однако предположим, что будущие потребности не будут удовлетворены. В этом случае значительная часть исследовательских ресурсов и кредитов будет потрачена на темы, которые не имеют необходимого приоритета и важности. Мы не можем ожидать, что будем использовать их результаты для решения проблем и развития этой области.

Проведенные исследования в области библиотечного дела и информатики определяли будущее направление исследований и выявляли новые темы. В этих исследованиях использовались такие подходы, как наукометрический, так и совместной встречаемости слов. Исследования показали, что такие темы, как цифровые библиотеки, поиск информации, управление знаниями, информационно-коммуникационные технологии, Интернет, веб-технологии, семантическая сеть, социальные сети, поисковые системы, метаданные, онтология, анализ эмоций, обработка естественного языка, машинное обучение, интеллектуальный анализ данных, интеллектуальный анализ текста и веб-анализ, являются новыми направлениями в этой области [43-49]. Учитывая важность проведения подобных исследований для будущей политики и планирования, мы стремились использовать подход глубокого обучения для прогнозирования тенденций международных исследовательских тем в области библиотечного дела и информатики, чтобы сделать шаг к определению потребностей будущих исследований.

## МЕТОДОЛОГИЯ ИССЛЕДОВАНИЯ

Настоящее исследование было проведено с использованием метода интеллектуального анализа текста и глубокого обучения. После предварительной обработки и тематической классификации исследований были использованы алгоритмы глубокой нейронной сети для моделирования предметных направлений уже проведенных исследований и прогнозирования будущих тенденций на основе более ранних моделей данных. В исследовании были включены статьи из всех библиотечных и информационных научных журналов, которые были проиндексированы в Web of Science (WoS) с 1945-2020 гг. и перечислены под заголовком "Информатика и библиотечное дело" в Journal Citation Reports (JCR). Таким образом, было извлечено 89 публикаций, содержащих 90311 статей. В поле название публикации (SO) были получены все статьи в разделе Расширенного поиска. Записи, полученные в разделе "Тип документа", были ограничены статьями, английским языком, а в разделе "Тематическая категория" - информатикой и библиотечным делом. Наиболее важными исполнительными этапами настоящего исследования были следующие:

### Выбор предметной области для анализа

Для проведения настоящего анализа были использованы названия статей. Во-первых, библиографическая информация о статьях, включая название и год публикации, была извлечена из базы данных WoS. Затем была создана база данных на основе информации, хранящейся в программном обеспечении Excel 2016.

## **Выбор тематического плана классификации научных продуктов**

Классификации, представленные в специализированных базах данных, часто использовались в предыдущих исследованиях для систематизации исследований в этой области. Однако следует отметить, что каждая из этих классификационных схем сама по себе не может охватить различные аспекты библиотечного дела и информатики, а также новые и междисциплинарные предметные области. Более того, сферы исследований в этой области находятся под влиянием различных социальных, культурных и технологических изменений в сообществе и внутри отрасли. Эти изменения указывают на необходимость пересмотра существующих классификационных схем. Необходимо было создать систему, которая охватывала бы все предметные сферы исследований в этой области, включая традиционные и современные темы, другими словами, все предметные области за последние 76 лет и предсказание будущих тем в этой области. Поэтому, изучив и объединив существующие схемы классификации и проконсультировавшись с экспертами в этой области, были подготовлены списки, составленные исследователями в 31 предметной области, которые стали основой анализа.

## **Подготовка и предварительная обработка данных**

Чтобы преобразовать названия статей и полученные из них слова и фразы в стандартный и структурированный формат были выполнены операции предварительной обработки данных. Эти операции включают удаление стоп-слов, стемминг (в структурированном формате) и лемматизацию, удаление повторяющихся данных, выбросов и пропущенных значений, а также извлечение параллельных слов и словарного запаса с использованием методов n-граммы (`ngram_range=[1, 3]`). Наконец, метод TF-IDF был использован для выбора ценных функций и ключевых слов, и документы были отображены в виде модели векторного пространства (VSM), известной как Мешок слов (BOW – Bag of Words), и была сформирована матрица термин-документ. Мы нормализовали данные с помощью алгоритма Min-Max Scaler. Размеры матрицы были уменьшены с помощью алгоритма PCA, документы на английском языке отображались в виде матрицы терминов в размерах 8438\*77161. Случаи употребления каждого термина в документе были узнаваемы по названию статей. Более того, вес этого слова был очевиден.

## **Предметная классификация исследований**

В данном исследовании для классификации исследований в анализируемой области использовался метод автоматической классификации текста, одно из приложений интеллектуального анализа

текста. Результаты показали, что среди моделей, используемых для классификации, модель, полученная с помощью алгоритма логистической регрессии для классификации исследований, имела четкость, точность, полноту и оценку F1 выше 90% и, как следствие, имела лучшую производительность при классификации документов.

## **Использование глубоких нейронных сетей для прогнозирования будущего**

В настоящей работе для прогнозирования предметной направленности исследований использовались глубокие нейронные сети. Сетевая архитектура, используемая в этом исследовании, представляет собой долговременную кратковременную память (LSTM – Long short-term memory). LSTM – это рекуррентная нейронная сеть, предназначенная для хранения, обработки и доступа к информации лучше, чем обычные нейронные сети. Эти сети имеют обратный цикл, который позволяет информации, которую они уже получили, не теряться и оставаться в сети. Однако обычные нейронные сети, используемые специалистами по машинному обучению, не ведут себя подобным образом, что является существенным недостатком для этих сетей [50].

Обратим внимание, что повторяющиеся сети не способны к долгосрочному обучению зависимости и не обладают способностью вспоминать и использовать информацию, которую они узнали достаточно давно. Другими словами, они не могут использовать информацию из прошлого [51, 52]. В связи с этим сети LSTM были разработаны для решения проблемы долгосрочной зависимости.

## **Разработка и внедрение модели глубокой сети**

Перед реализацией модели нейронной сети данные должны быть преобразованы в столбцы, приемлемые для входных данных модели нейронной сети. Эта операция была выполнена с помощью функции придания новой формы (`reshape`) в библиотеке NumPy. Затем входные данные были нормализованы с помощью функции масштабирования Min-Max. Наиболее важной частью разработки модели глубокой сети является определение ее уровней. Для размещения глубоких сетевых уровней была использована последовательная модель, которая является самым простым и базовым типом. Слои укладываются один за другим.

Кроме того, они определяются как последовательный и линейный набор слоев. Каждый слой связан только со следующим слоем. Фактически, в этой модели различные слои добавляются последовательно. Архитектура разработанной модели для глубокой сети в этом исследовании включала плотный входной слой (LSTM), два скрытых плотных слоя, два отсеиваемых слоя и плотный выходной слой, функции которого будут описаны ниже.

**Входной уровень:** Основной уровень и входные данные глубокой сети, которая получает данные, относятся к типу LSTM. Затем входные данные сети были преобразованы в столбцы. Информация в каждом столбце относится к 31 предметному вопросу, который включает 76 данных, относящихся к периоду с 1945 по 2020 гг. для каждого предмета. Параметры задачи включают в себя X1, X2, X3, ..., X76, каждый из которых представляет данные за каждый год. Например, X1 - это данные за 1945 г., а X76 - это данные за 2020 г. Чтобы определить каждый слой, мы должны сначала определить количество нейронов в этом слое. Количество этих нейронов – это параметр, полученный методом проб и ошибок. Наконец, было идентифицировано шесть нейронов, подходящих для этого слоя. Во входном слое использовалась гиперболическая касательная функция (Tanh). Причиной выбора функции Tanh было то, что при разработке модели она была установлена по умолчанию самим программным обеспечением и имела незначительную ошибку при тестировании различных функций.

**Скрытый плотный слой:** Плотный слой с 300 нейронами был добавлен в качестве скрытого слоя для улучшения работы нейронной сети. В этом слое использовалась выпрямленная линейная функция активации (ReLU), наиболее распространенная и часто используемая функция в скрытом слое.

**Уровень отсева:** Этот уровень повышает производительность сети в глубоких сетях. Уровень отсева определяет, сколько процентов сетевых нейронов случайно отключается во время процесса обучения или имеет нулевой ввод, чтобы уменьшить повторяющееся обучение между нейронами, ускорить обучение сети и предотвратить перенапряжение. В настоящей модели нейронной сети этот слой определил, что 0,3 нейрона были исключены в процессе обучения.

**Скрытый плотный слой:** В сеть, разработанную для настоящего исследования, был добавлен еще один скрытый плотный слой со 100 нейронами, и в этом слое использовался ReLU.

**Выпадающий слой:** Выпадающий слой со значением 0,1 также был помещен между скрытым слоем и выходным слоем. Это означает, что 0,1 нейрона было оставлено без внимания для процесса обучения модели.

**Выходной слой:** В выходной секции был помещен плотный слой с нейроном.

**Компиляция модели:** После предварительной обработки и нормализации данных и определения архитектуры модели построенная модель была скомпилирована с использованием команды *модель.компиляция (model.compile)*. В этой секции функция потерь использовалась для отображения производительности сети при обучении данным. На этом этапе для улучшения разработанного процесса обучения сети был использован алгоритм оптимизации ADAM, который является одним из самых

популярных алгоритмов оптимизации и обладает высокой производительностью в системах с низким объемом памяти.

**Соответствие модели:** Наконец, модель была оценена с помощью команды *model.fit*, и модель была выполнена на основе данных. Этот шаг определяет, сколько раз необходимо применить обучение к данным; каждый из них называется эпохой. Следует отметить, что модель будет иметь некоторые ошибки каждый раз, когда процесс обучения повторяется, но через некоторое время это количество ошибок становится равномерным. На данный момент модель, как говорят, сошлась вокруг этого уровня повторения процесса обучения. В настоящем исследовании частота повторения сетевого обучения была основана на этом, и модель сходилась около 100 эпох. Другим параметром, который указывается на этом шаге, является размер пакета. Функция этого параметра такова, что модель вычисляет среднее значение ошибок после просмотра некоторых данных. Затем оптимизирует и обновляет сеть на основе этого среднего значения. Размер партии в настоящей исследовательской модели был установлен равным 30.

Данные исследования были разделены на две категории: образовательные и экспериментальные для измерения эффективности разработанной модели. Следовательно, 80% документов (с 1945 по 2005 гг.) были использованы для обучения модели, а 20% (с 2006 по 2020 гг.) были исключены для тестирования модели. Затем реализованная модель была применена к образовательным данным. Скорость обучения модели и ее производительность были оценены для прогнозирования тестовых данных. Чтобы оценить частоту ошибок, мы сравнили разницу между фактическими значениями и значениями обучения. Кроме того, сравнивалась разница между фактическими значениями и значениями тестируемых данных. Также была рассчитана частота ошибок при изучении модели и ее прогнозировании. В связи с этим был использован метод среднеквадратичной ошибки (MSE).

Языком программирования, использованным в этом исследовании, был Python. Программные пакеты Pandas, Numpy, Hazm и Nltk использовались для ввода данных, обработки, формирования матриц, словообразования, классификации и тестирования, а пакеты Numpy, Pandas, Tensor Flow и Keras – для построения (реализации) глубокой нейронной сети и прогнозирования будущего.

## РЕЗУЛЬТАТЫ

Были изучены предметные направления исследований в области библиотечного дела и информатики за последние 76 лет. В табл. 1 представлено распределение исследований, посвященных каждой из предметных областей данной области, за предыдущие годы.

**Тематические направления исследований в области библиотечного дела и информатики  
на веб-сайте Web of Science в 1945-2020 гг.**

Области исследований	Частота	Процент
Информационно-коммуникационные технологии	7075	9,17
Интернет и веб-исследования	6535	8,47
Библиотеки, архивы и информационные центры	4314	5,59
Исследования, обучение и публикации	3907	5,06
Сохранение данных и извлечение информации	3092	4,01
Наукометрия и информатика	2944	3,82
Информационные системы	2889	3,74
Информационные и основанные на знаниях услуги	2883	3,74
Базы данных	2759	3,58
Преподавание и обучение	2755	3,57
Управление информационными ресурсами	2617	3,39
Организация информации и знаний	2561	3,32
Исследования пользователей	2427	3,15
Научно-информационная коммуникация	2370	3,07
Управление знаниями	2327	3,02
Информационное поведение	2277	2,95
Профессия, занятость и предпринимательство	2239	2,90
Библиотекари и информационные специалисты	1940	2,51
Электронные, цифровые и виртуальные библиотеки	1934	2,51
Поисковые системы	1902	2,46
Экономика и информационный маркетинг	1832	2,37
Этика и информационная безопасность	1832	2,37
Обработка информации и управление	1668	2,16
Информационная и цифровая грамотность	1530	1,98
Интеллектуальный анализ данных и обнаружение знаний	1480	1,92
Искусственный интеллект	1426	1,85
Общество, основанное на информации и знаниях	1243	1,61
Будущие исследования в области библиотечного дела и информатики	1221	1,58
Теоретические основы библиотечного дела и информатики	1171	1,52
Семантическая сеть и онтология	1071	1,39
Организационное управление	940	1,22

Как показано в табл. 1, на международном уровне большинство статей касалось информационных и коммуникационных технологий (9,17%), Интернета и веб-исследований (8,47%), библиотек, архивов и информационных центров (5,59%), исследований, обучения и публикаций (5,06%), а также сохранения данных, и извлечение информации (4,01%). Напротив, наименьшее количество статей приходится на области организационного управления (1,22%), семантической сети и онтологии (1,39%), теоретических основ (1,52%) и будущих исследований (1,58%).

Для того, чтобы спрогнозировать предметные направления исследований в области библиотечного дела и информатики, в первую очередь, была протестирована эффективность разработанной модели глубокой сети. Как упоминалось выше в настоящем исследовании, функция потерь использовалась для отображения производительности

сети при обучении данных. Результаты показаны на рис. 1 для данных обучения, относящихся к первому вопросу (теоретические основы).

На рис. 1 показано, что при 100-кратном выполнении процесса обучения модели частота ошибок снижается, а производительность сети повышается. Производительность модели по другим темам была такой же, но мы не привели их объем в этом исследовании из-за их большого объема. После выполнения реализованной модели на образовательных данных была оценена скорость обучения модели и ее производительность для прогнозирования тестовых данных. Чтобы оценить частоту ошибок, сравнивали разницу между фактическими значениями и значениями обучения и разницу между фактическими значениями и значениями тестируемых данных, а также рассчитывали частоту ошибок при обучении модели и ее прогнозировании с использованием метода MSE. Ре-

зультаты сравнения фактических значений теоретической базы с значениями обучения и значениями, прогнозируемыми моделью, показаны на рис. 2.

На рис. 2 показано, что данные, подготовленные и прогнозируемые моделью, близки к фактическим данным с точки зрения теоретических ос-

нов, а это означает, что разработанная модель хорошо обучена и имеет хорошую производительность при оценке данных. Далее в табл. 2 перечислены данные, полученные в результате вычисления критерия MSE, т.е. оценка ошибки в обучающей и тестовой части модели, как показано в табл. 2.

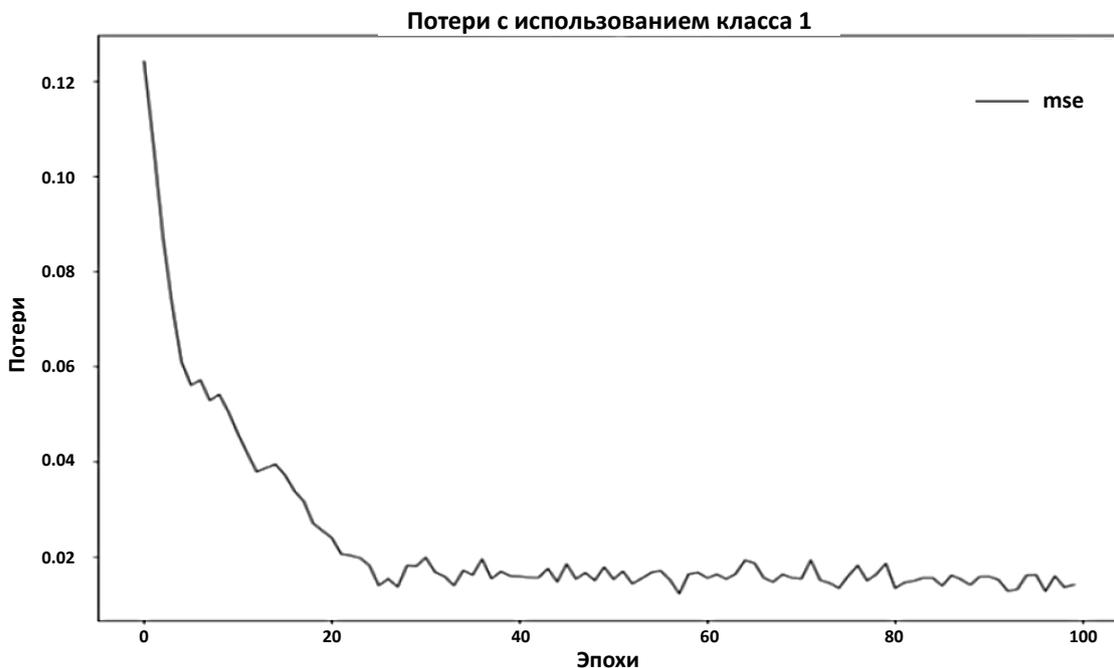


Рис. 1. Результаты частоты ошибок в процессе обучения модели глубокой сети для теоретических основ



Рис. 2. Сравнение фактических значений на теоретической основе с значениями обучения и значениями, прогнозируемыми моделью глубокой сети

**Вычисление среднеквадратичной ошибки при обучении и тестировании модели прогнозирования на международном уровне**

Строка	Области исследований	Среднеквадратичная ошибка	
		Среднее значение обучения	Среднее значение теста
1	Теоретические основы библиотечного дела и информатики	0,77	0,28
2	Библиотеки, архивы и информационные центры	0,49	0,03
3	Организационное управление	0,02	0,57
4	Электронные, цифровые и виртуальные библиотеки	0,03	0,14
5	Профессия, занятость и предпринимательство	0,04	0,72
6	Преподавание и обучение	0,12	2,46
7	Исследования, обучение и публикации	0,22	2,47
8	Информационно-коммуникационные технологии	0,03	3,45
9	Интернет и веб-исследования	0,03	2,43
10	Обработка и управление информацией	0,05	1,88
11	Управление знаниями	0,02	0,74
12	Информационная и цифровая грамотность	0,06	0,09
13	Будущие исследования в области библиотечного дела и информатики	0,11	2,62
14	Управление информационными ресурсами	1,42	1,07
15	Организация информации и знаний	0,57	0,47
16	Сохранение данных и извлечение информации	1,27	2,46
17	Общество, основанное на информации и знаниях	0,15	1,49
18	Семантическая сеть и онтология	0,01	0,90
19	Экономика и информационный маркетинг	0,04	0,60
20	Информационное поведение	0,03	0,26
21	Искусственный интеллект	0,02	1,25
22	Интеллектуальный анализ данных и обнаружение знаний	0,03	2,62
23	Наукометрия и информатика	0,01	2,32
24	Информационные системы	0,62	0,54
25	Поисковые системы	0,009	1,45
26	Исследования пользователей	0,12	1,57
27	Библиотекари и информационные специалисты	0,10	1,18
28	Информационные и основанные на знаниях услуги	0,26	0,63
29	Научные и информационные коммуникации	0,16	1,95
30	Базы данных	0,02	1,03
31	Этика и информационная безопасность	0,21	1,78

Табл. 2 показывает, что ошибка, полученная при обучении и тестировании модели по всем темам, была очень мала. Относительно высокий уровень ошибок в теме информационно-коммуникационных технологий, среди других тем, был обусловлен более нерегулярной тенденцией и внезапным скачком в этих вопросах в последние годы.

После обеспечения точности прогнозирования дизайна глубокой сети модель была переобучена, и было выполнено прогнозирование тем исследований на следующие десять лет. Реализация сети и прогнозирование будущего были выполнены с использованием программы Python, разработанной для этой цели. Согласно столбцам данных, 76 данных с 1945 по 2020 гг. были использованы для формирования сети и прогнозирования будущего. Таким образом, данные за последние 76 лет пере-

даются в сеть, и сеть соответствующим образом обучается. На основе прошлых тенденций затем извлекаются критерии и правила данных и прогнозируются на следующее десятилетие. Другими словами, глубокая сеть действует как система, которая использует значение входных данных для вычисления выходных данных. Каждое из соединений между данными в разных слоях имеет свой вес. Сеть изучает закономерность между входными и выходными переменными, корректируя эти веса на этапе обучения. Конечным результатом модели глубокой сети является прогнозирование количества исследований в каждой из 31 области исследований в сфере библиотечного дела и информатики на следующее десятилетие (т.е. с 2021 по 2030 гг., что показано в табл. 3 с данными за последние шесть лет).

**Предметные направления научных разработок в области библиотечного дела и информатики на международном уровне за последние 6 лет и прогнозирование этой тенденции на ближайшие 10 лет**

Область исследований	Год публикации	Количество научных заглавий															
		2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030
Теоретические основы библиотечного дела и информатики		54	46	59	46	50	56	61	59	67	66	75	74	84	83	94	5
Библиотеки, архивы и информационные центры		62	68	58	56	51	50	47	52	49	48	45	42	38	36	33	29
Организационное управление		34	31	46	41	48	34	37	27	34	31	29	38	38	31	36	37
Электронные, цифровые и виртуальные библиотеки		103	105	108	119	128	132	132	138	141	140	143	148	150	154	158	161
Профессии, занятость и предпринимательство		94	95	98	113	133	129	130	137	140	144	147	150	153	156	159	162
Преподавание и обучение		119	120	121	126	130	138	141	143	146	146	147	148	149	152	157	158
Исследования, обучение и публикации		175	188	205	208	212	223	226	228	231	234	236	237	238	240	241	241
Информационно-коммуникационные технологии		351	360	361	412	418	433	445	454	461	468	474	479	484	488	491	494
Интернет и веб-исследования		449	461	486	510	517	530	549	563	573	578	583	586	587	895	915	597
Обработка и управление информацией		77	91	102	113	129	130	133	131	133	135	136	138	139	143	141	142
Управление знаниями		157	144	149	153	164	166	169	171	174	173	175	177	179	182	185	189
Информационная и цифровая грамотность		94	97	99	100	114	109	114	116	116	117	119	120	122	123	123	125
Будущие исследования в области библиотечного дела и информатики		71	71	80	85	91	104	109	114	116	119	121	121	123	124	126	127
Управление информационными ресурсами		40	39	37	37	36	33	32	30	29	27	26	25	23	22	20	19
Организация информации и знаний		49	47	43	41	41	40	39	38	37	36	38	36	35	33	31	30
Сохранение данных и извлечение информации		85	85	92	94	108	119	123	129	129	139	144	149	156	158	166	170
Общество, основанное на информации и знания		54	46	59	46	50	56	61	59	67	66	75	74	84	83	94	95
Семантическая сеть и онтология		67	77	82	96	101	117	128	142	147	153	158	163	168	172	177	178
Экономика и информационный маркетинг		113	126	139	147	156	169	175	179	183	186	189	192	194	196	198	199
Информационное поведение		142	149	163	170	186	188	194	199	207	213	218	222	224	227	229	233
Искусственный интеллект		64	78	83	94	102	115	118	129	135	139	141	152	160	169	179	181
Интеллектуальный анализ данных и обнаружение знаний		92	101	106	120	144	152	158	160	167	172	178	182	186	192	195	198
Наукометрия и информатика		182	178	152	139	119	117	115	114	111	115	109	108	105	103	100	98
Информационные системы		81	81	84	81	88	78	74	73	75	75	74	76	77	76	76	78
Поисковые системы		121	127	141	152	154	160	175	182	188	189	192	195	197	199	202	207
Исследования пользователей		104	104	117	118	136	138	142	147	148	152	155	159	161	165	167	170
Библиотекари и информационные специалисты		47	50	62	69	78	78	90	93	94	97	100	108	114	117	124	126
Информационные и основанные на знаниях услуги		91	90	89	82	89	90	87	90	91	94	93	94	95	91	95	96
Научные и информационные коммуникации		83	117	112	111	103	113	109	111	115	114	119	125	121	125	126	128
База данных		113	120	129	139	153	154	157	158	161	165	167	176	177	180	184	186
Этика и информационная безопасность		76	68	76	77	93	89	88	93	95	96	97	98	99	100	101	102

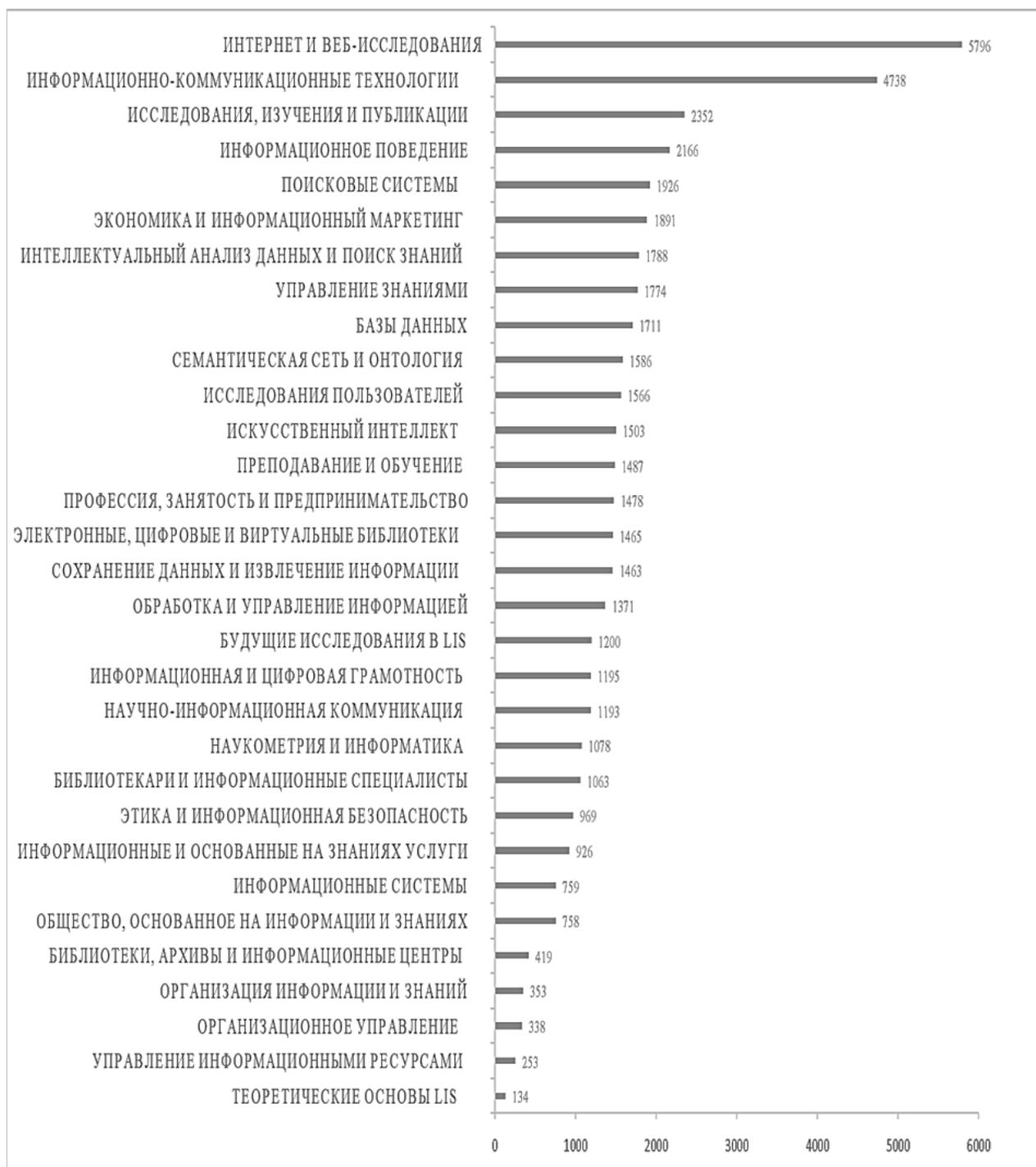


Рис. 3. Прогнозирование количества исследований, посвященных каждой из областей библиотечного дела и информатики на международном уровне на ближайшие 10 лет в порядке приоритетности

Согласно результатам, полученным из табл. 3, темы, которые имели тенденцию к росту в последние годы, будут продолжать расти в последующие периоды. Темы, склонные к снижению в последующие годы, также будут сокращаться. Таким образом, области теоретических основ; библиотеки, архивы и информационные центры; управление информационными ресурсами; организация информации и знаний; и наукометрия и информати-

ка, которые сокращались в последние годы, как ожидается, продолжат сокращаться в следующем десятилетии. Другие темы, которые столкнулись с тенденцией к росту в последние годы, продолжат расти в следующем периоде с небольшими колебаниями. На рис. 3 показан прогноз количества исследований, посвященных библиотечному делу и информатике, на международном уровне на ближайшие десять лет, соответственно.

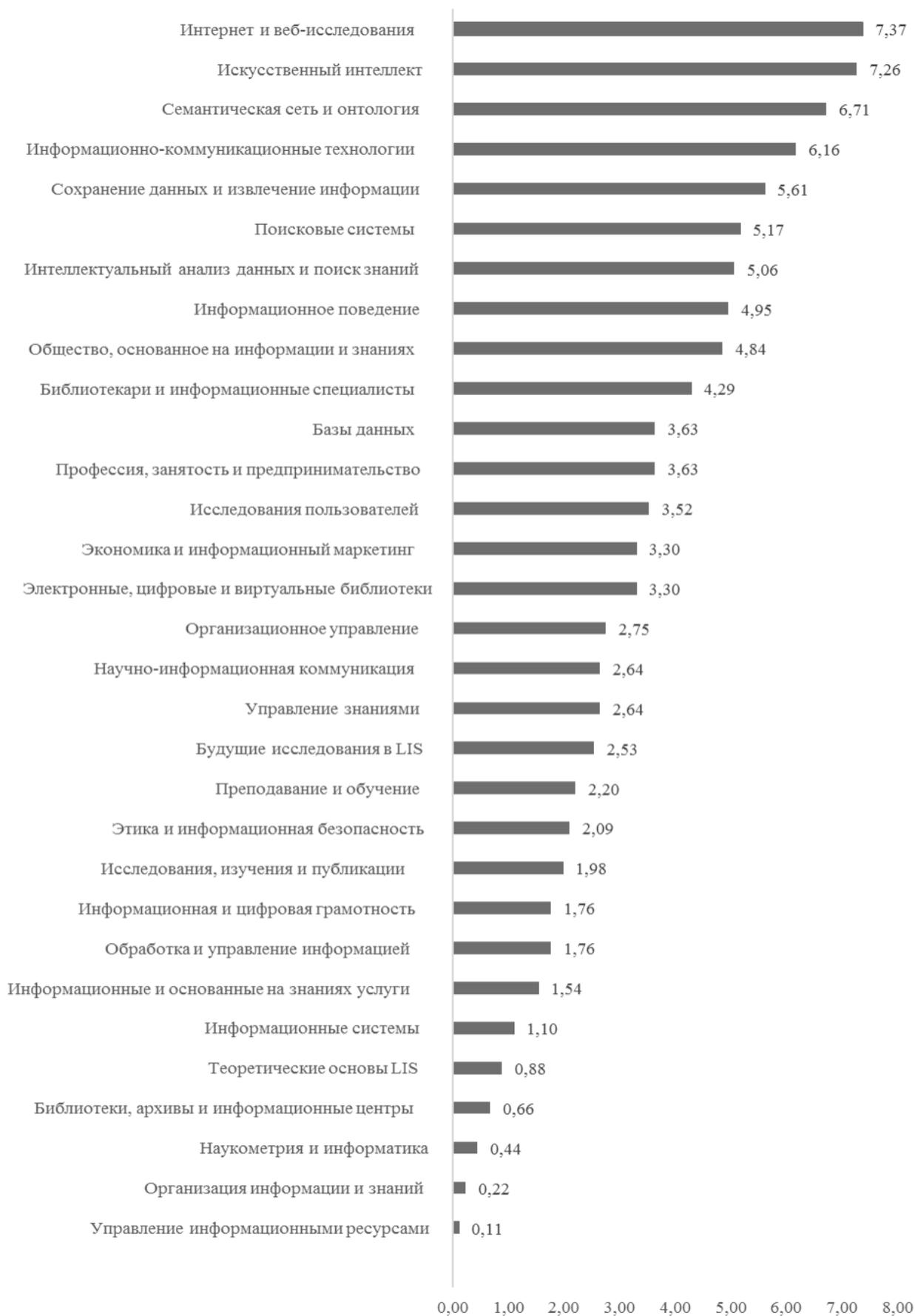


Рис. 4. Прогнозирование темпов роста исследований, посвященных каждой из областей библиотечного дела и информатики, на международном уровне на ближайшие 10 лет в порядке приоритетности

Как показано на рис. 3, в будущем наиболее значительное количество исследований на международном уровне будет проводиться в области Интернета и веб-исследований, информационно-коммуникационных технологий, исследований, обучения и публикации, а также информационного поведения. Такие области, как теоретические основы, управление информационными ресурсами, организационное управление и организация информации и знаний, будут иметь наименьшее количество исследований. Стоит рассмотреть возможность сравнения прошлых предметных направлений исследований в этой области с их будущей тенденцией. В прошлом библиотеки, архивы и информационные центры были одними из областей с наибольшим объемом исследований, в то время как результаты показывают, что в будущем исследований будет меньше. Области семантической сети, онтологии и будущих исследований, которым в прошлом уделялось меньше внимания со стороны исследователей, в перспективе будут иметь больше поклонников среди исследователей. Другой проблемой является прогнозирование процентного роста тем исследований в следующем десятилетии, как показано на рис. 4.

Как видно, тенденция роста исследований во многих областях библиотечного дела и информатики усиливается на международном уровне в следующем десятилетии. Однако значительный рост наблюдается в исследованиях в области Интернета и веб-исследований, искусственного интеллекта, семантической сети и онтологии, информационно-коммуникационных технологий, а также сохранения данных и поиска информации. Согласно прогнозируемым результатам, темы управления информационными ресурсами, организации информации и знаний, наукометрии и информатики, библиотек, архивов и информационных центров, а также теоретических основ будут испытывать гораздо более медленную тенденцию роста.

## ОБСУЖДЕНИЕ

Продолжение всесторонних разработок в будущем и их влияние на научную область выявляют целесообразность систематических исследований будущих знаний. Одним из наиболее важных и влиятельных аспектов библиотечного дела и информатики, на которые повлияют будущие разработки, является изучение тем исследований. Принимая это во внимание, предметное направление будущих исследований в настоящей работе было спрогнозировано на международном уровне. Результаты прогнозирования тем этой области на данном уровне показали, что наибольшее количество исследований в анализируемой области в будущем будет проводиться в области Интернета и веб-исследований, информационно-коммуникационных технологий, исследований, обучения и публикаций,

а также информационного поведения. Рассматривая результаты процентного роста исследовательских тем на международном уровне, мы обнаружили, что тенденция роста исследований во многих предметных областях этой области усиливается в следующем десятилетии. Тем не менее, более быстрый рост Интернет и веб-исследований, искусственного интеллекта, семантической сети и онтологии, информационно-коммуникационных технологий, а также сохранения данных и извлечения информации показывает, что эти результаты могут быть связаны с адаптацией этих областей к будущим разработкам и, следовательно, с острой потребностью библиотек и информатики в использовании выводов из этих областей. Эти выводы также могут быть связаны с последствиями достижений и новых технологий Четвертой промышленной революции, таких как искусственный интеллект, робототехника, блокчейн и Интернет, которые быстро формируются. Согласно автору [53] в книге "Четвертая промышленная революция", эксперты прогнозируют, что развитие, вызванное этой революцией, достигнет своего пика к 2030 году.

Развитие новых информационно-коммуникационных технологий и стремительный рост объема информации создали проблемы в области эффективного поиска и извлечения больших объемов информации, а также анализа и интерпретации полученной информации. Поэтому эксперты в области поиска информации провели исследования, чтобы предложить методы и подходы для решения этих проблем. Следовательно, наиболее важные области исследований в библиотечном деле и информатике включают такие темы, как "интеллектуальные агенты", "алгоритмы ранжирования", "кластерный анализ", "интеллектуальный анализ данных и веб-анализ", "алгоритмы graph web и semantic web" и "Интернет-исследования Веб" [54]. С 1990-х годов, из-за междисциплинарного характера библиотечного дела и информатики, предметы, связанные с технологиями и Интернетом, получили больший рост. Примечательно, что корни этого явления уходят в развитие информационно-коммуникационных технологий и появление Интернета.

Наряду со взрывным ростом информации и роли информации и знаний в экономике, новая эра повлияла на предметные направления. Похоже, что в будущем область библиотечного дела и информатики будет иметь прочные связи с другими областями на международном уровне и будет междисциплинарной, особенно в отношении таких дисциплин, как информационные технологии, информатика, искусственный интеллект, управление и коммуникация [55-59; 40]. Таким образом, изменчивый характер этих областей и быстрое развитие инструментов и технологий приводят к постоянным изменениям фокуса исследований (Saracevic, 1999, цитируется в [40]).

В подтверждение выводов настоящего исследования мы можем сослаться на результаты некоторых предыдущих исследований. Авторы работы [46] пришли к выводу, что, с точки зрения экспертов в области библиотек и информатики, темы, связанные с поиском информации, будут одними из главных исследовательских приоритетов в этой области в научных социальных сетях в ближайшие десять лет. Более того, вопросы, связанные с разработкой политики в области науки и техники, подготовкой к информационной работе и моделями поведения в поиске информации, будут приоритетными для будущих исследований в этой области. По словам автора [47], эксперты в этой области в последние годы использовали возможности информационно-коммуникационных технологий и с помощью экспертов в других областях пытались устранить барьеры для поиска информации. С помощью этого описания можно утверждать, что поиск информации будет одной из наиболее важных исследовательских проблем в библиотечном деле и информатике на сегодняшний день и в будущем, что помогает описать проблемы, с которыми сталкиваются пользователи, и найти их решения. Результаты исследования [60] показали, что с точки зрения экспертов в области библиотек и информатики наиболее важными темами для будущих исследований в этой области будут прикладные темы, темы с опорой на будущее в данной области, и темы, связанные с информационно-коммуникационными технологиями. Согласно оценкам Саммерса и др. [43], намеревающихся наметить будущее информатики на ближайшие десять лет, основными темами в этой области в 2010 году были цифровые библиотеки, поиск информации и управление знаниями. Кроме того, исследование, анализирующее содержание призыва провести семинары, конференции и симпозиумы по библиотечным и информационным наукам в Индии и по всему миру с 2012 по 2014 год, показало, что прикладные темы, основанные на информационно-коммуникационных технологиях, Интернете, веб-технологиях и семантической сети, исследованиях пользователей, цифровых технологиях и социальных медиа представляют перспективу исследований в этой области в сценарии будущего Индии [44]. Кроме того, автор [47] пришел к выводу, что анализ эмоций, обработка естественного языка, машинное обучение, интеллектуальный анализ данных, интеллектуальный анализ текста, веб-анализ, социальные сети и метаданные относятся к числу новых тем в библиотечном деле, и ожидается, что в ближайшие годы информатика будет развиваться быстрее. Результаты исследований, в которых изучалась тематическая тенденция исследований в области библиотечного дела и информатики в последние десятилетия, показывают, что поиск информации был одним из наиболее важных тематических кластеров в этой области во все

периоды. Тем не менее, исследования в последние годы переключились на машинное обучение, извлечение текста, тематическое моделирование, анализ эмоций и обработку текста на основе моделей или алгоритмов [40; 49].

Результаты настоящего исследования показали, что будущие исследователи будут менее востребованы теоретическими основами, библиотеками, архивами и информационными центрами, управлением информационными ресурсами, организационным управлением и организацией информации и знаний. Кроме того, темпы роста этих тем в будущем будут ниже. Тенденция к снижению в этих областях может быть связана с тем, что эти области исследований эволюционировали. Другими словами, эти предметные области по-прежнему актуальны в данной сфере, только под влиянием изменений в обществе и влияния на эту область они были предложены с новым названием или концепцией. Например, изучение библиотек привело к изучению электронных, цифровых, виртуальных и интеллектуальных библиотек. Исследования по приобретению или управлению информационными ресурсами чаще называют информационным управлением. Традиционные темы организации информации, такие как каталогизация и классификация, уступили место организации киберпространства, метаданным, автоматической классификации и автоматической индексации. В этой связи авторы работы [55] признали, что в прошлом традиционные инструменты, такие как тезаурусы, тематические рубрики и классификации, использовались в качестве инструментов для организации знаний в информационных центрах. Однако сегодня Интернет, семантическая паутина, онтологии, метаданные, системы организации знаний, электронные тезаурусы, электронные базы данных и мультимедийные инструменты произвели революцию в организации информации.

Отличительной особенностью результатов настоящего исследования является то, что исследования в области наукометрии и информатики, которые в последние годы значительно выросли, в будущем будут изучаться меньше и иметь тенденцию к более медленному росту. Оправдывая тенденцию к снижению этих исследований, можно отметить, что эволюция носителей информации и увеличение объема текстовых, аудио-, видео- и мультимедийных ресурсов, а также проблемы, связанные с их анализом, привели к ограничениям в применении традиционных наукометрических подходов, таких как метод совместного слова [8]. Наряду с ростом объема информации в различных форматах и средствах массовой информации и развитием крупномасштабных технологий обработки текстов, подходов искусственного интеллекта, машинного обучения, глубокого обучения, а также методов интеллектуального анализа данных и анализа текста из-за их высокой мощности в об-

работке и анализе больших данных [22], методы моделирования [31], использование онтологии и связей ключевых слов [4; 23] и семантический анализ слов [6] заменили наукометрические методы. Очевидно, что в последние годы и даже в будущем количество исследований, проводимых с использованием традиционных наукометрических подходов, будет сокращаться.

## ЗАКЛЮЧЕНИЕ

Быстрое развитие различных аспектов жизни в результате развития технологий и информационно-коммуникационных сетей, Интернета и социальных сетей, а также продолжение этих разработок в будущем и их влияние на область библиотечного дела и информатики показывает необходимость будущих исследований в этой области в различных измерениях, включая анализ будущих тем исследований. Из-за быстрого развития исследовательских тем в области библиотечного дела и информатики, и проблем, связанных с прогнозированием этих тем в будущем, если мы поймем истинную природу и миссию библиотечного дела и информатики, мы можем ожидать более эффективного понимания будущего направления исследований в этой области. Осведомленность о важнейших будущих проблемах, помогающая планировать будущие исследования, может выявить пробелы в исследованиях и инвестиционные возможности в области НИОКР, тем самым поддерживая исследовательские центры разумно выбирать проекты. При таком подходе можно избежать повторения исследовательских тем с разными названиями и сосредоточиться на новых и практических темах с высоким потенциалом для развития сообщества. Предположим, что лица, занимающиеся планированием и принятием решений, особенно в сфере образования и исследований, игнорируют фундаментальные вопросы и конечную миссию этой области в целом. В этом случае в будущем эта область может потерять свое место в научном сообществе.

## ЛИТЕРАТУРА

1. *Godet M.* Creating futures: Scenario planning as a strategic management tool. — Washington, DC: *Economica. Economica Brookings diffusion*, 2006.
2. *Tavakoli Gh., Babakirad A., Ghorooneh H.* Development and designing of the method for emerging issues analysis // *Journal of Security Research*. — 2016. — Vol. 15, No. 53. — P.151-181.
3. *Singh N., Hu C., Roehl W. S.* Text mining a decade of progress in hospitality human resource management research: Identifying emerging thematic development // *International Journal of Hospitality Management*. — 2007. — Vol. 26, No. 1. — P. 131-147. — <http://dx.doi.org/10.1016/j.ijhm.2005.10.002>

4. *Salatino A.* Early detection and forecasting of research trends // 14th International Semantic Web Conference, Bethlehem (PA), USA. — 2015.

5. *Zhang Y., Zhang G., Chen H., Porter A. L., Zhu D., Liu J.* Topic analysis and forecasting for science, technology and innovation: Methodology with a case study focusing on big data research // *Technological Forecasting & Social Change*. — 2016. — Vol. 105. — P. 179-191. — <http://dx.doi.org/10.1016/j.techfore.2016.01.015>

6. *Li X., Xiea Q., Daimb T., Huang L.* Forecasting technology trends using text mining of the gaps between science and technology: The case of perovskite solar cell technology // *Technological Forecasting & Social Change*. — 2019. — Vol. 146. — P. 432-449. — <https://doi.org/10.1016/j.techfore.2019.01.012>

7. *Chen C., Wang Z., Li W., Sun X.* Modeling scientific influence for research trending topic prediction // *Proceedings of the AAAI Conference on Artificial Intelligence*. — 2018, April. — Vol. 32, No. 1.

8. *Abubaya T., Nigatie Y., Kovalchuk S.* Towards predicting trend of scientific research topics using topic modeling // *Procedia Computer Science*. — 2018. — Vol. 136. — P. 304-310. — <https://doi.org/10.1016/j.procs.2018.08.284>

9. *de Solla Price D.* Networks of scientific papers // *Science*. — 1965. — Vol. 149, No. 3683. — P. 510-515.

10. *Small H.* Tracking and predicting growth areas in science // *Scientometrics*. — 2006. — Vol. 68, No. 3. — P. 595-610. — <https://doi.org/10.1007/s11192-006-0132-y>

11. *He Q., Chen B., Pei J., Qiu B., Mitra P., Giles L.* Detecting topic evolution in scientific literature: how can citations help? // *Proceedings of the 18th ACM conference on Information and knowledge management*. — 2009, November. — pp. 957-966.

12. *Small H., Boyack K. W., Klavans R.* Identifying emerging topics in science and technology // *Research Policy*. — 2014. — Vol. 43, No. 8. — P. 1450-1467. — <https://doi.org/10.1016/j.respol.2014.02.005>

13. *Xu S., Hao L., An X., Pang H., Li T.* Review on emerging research topics with key-route main path analysis // *Scientometrics*. — 2020. — Vol. 122. — P. 607-624.

14. *Chen C.* CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature // *Journal of the American Society for Information Science and Technology*. — 2006. — Vol. 57, No. 3. — P. 359-377. — <https://doi.org/10.1002/asi.20317>

15. *Guo H., Weingart S., Borner K.* Mixed-indicators model for identifying emerging research areas // *Scientometrics*. — 2011. — Vol. 89, No. 1. — P. 421-435. — <https://doi.org/10.1007/s11192-011-0433-7>

16. *Wu Z., Ren Y.* A bibliometric review of past trends and future prospects in urban heat island research from 1990 to 2017 // *Environmental Reviews*. — 2018. — Vol. 27, No. 2. — P. 241-251. — <https://doi.org/10.1139/er-2018-0029>

17. *Hoz-Correa A., Munoz-Leiva F., Bakucz M.* Past themes and future trends in medical tourism research: A co-word analysis // *Tourism Management*. — 2018. — Vol. 65. — P. 200-211. — <https://doi.org/10.1016/j.tourman.2017.10.001>
18. *Lara-Rodríguez J. S., Rojas-Contreras C., Oliva E. J. D.* Discovering emerging research topics for brand personality: A bibliometric analysis // *Australasian Marketing Journal (AMJ)*. — 2019. — Vol. 27, No. 4. — P. 261-272. — <https://doi.org/10.1016/j.ausmj.2019.06.002>
19. *Pestana M. H., Sánchez A. V., Moutinho L.* The network science approach in determining the intellectual structure, emerging trends and future research opportunities – An application to senior tourism research // *Tourism Management Perspectives*. — 2019. — Vol. 31. — P. 370-382. — <https://doi.org/10.1016/j.tmp.2019.07.006>
20. *Joung J., Kim K.* Monitoring emerging technologies for technology planning using technical keyword based analysis from patent data // *Technological Forecasting and Social Change*. — 2017. — Vol. 114. — P. 281–292. — <https://doi.org/10.1016/j.techfore.2016.08.020>
21. *Porter A. L., Garner J., Carley S. F., Newman N. C.* Emergence scoring to identify frontier R&D topics and key players // *Technological Forecasting and Social Change*. — 2018. — Vol. 146. — P. 628–643. — <https://doi.org/10.1016/j.techfore.2018.04.016>
22. *Lee C., Kwon O., Kim M., Kwon D.* Early identification of emerging technologies: A machine learning approach using multiple patent indicators // *Technological Forecasting and Social Change*. — 2018. — Vol. 127. — P. 291–303. — <https://doi.org/10.1016/j.techfore.2017.10.002>
23. *Balili C., Segev A., Lee U.* Tracking and predicting the evolution of research topics in scientific literature // 2017 IEEE international conference on big data (big data), (pp. 1694-1697). — IEEE, 2017, December.
24. *Wu F. S., Shin C. C., Lee P. C., Su H. N.* Integrated methodologies for mapping and forecasting science and technology trends: A case of etching technology // *PICMET 2010 Technology Management For Global Economic Growth*, (pp. 1-23). — IEEE, (2010, July).
25. *Budi A., Rizal F. A., Widodo A.* Prediction of research topics on science & technology (S&T) using ensemble forecasting // *International Journal of Software Engineering and Its Applications*. — 2013. — Vol. 7, No. 5. — P. 253-268. — <https://doi.org/10.14257/ijseia.2013.7.5.23>
26. *Bildosola I., González P., Moral P.* An approach for modelling and forecasting research activity related to an emerging technology // *Scientometrics*. — 2017. — Vol. 112, No. 1. — P. 557–572. — <https://doi.org/10.1007/s11192-017-2381-3>
27. *Bai X.* Predicting the number of publications for scholarly networks // *IEEE Access*. — 2018. — Vol. 6. — P. 11842–11848. — <https://doi.org/10.1109/ACCESS.2018.2812804>
28. *Sobrabi B., Khabilijafarabad A.* Systematic method for finding emergence research areas as data quality // *Technological Forecasting and Social Change*. — 2018. — Vol. 137. — P. 280-287. — <https://doi.org/10.1016/j.techfore.2018.08.003>
29. *Choi Y., Park S., Lee S.* Identifying emerging technologies to envision a future innovation ecosystem: A machine learning approach to patent data // *Scientometrics*. — 2021. — Vol. 126, No. 7. — P. 5431–5476. — <https://doi.org/10.1007/s11192-021-04001-1>
30. *Cho J. H., Lee J., Sohn S. Y.* Predicting future technological convergence patterns based on machine learning using link prediction // *Scientometrics*. — 2021. — Vol. 126, No. 7. — P. 5413–5429. — <https://doi.org/10.1007/s11192-021-03999-8>
31. *Santo M. M., Coelho Santos, D. M., Filbo, L. F.* Text mining as a valuable tool in foresight exercises: A study on nanotechnology // *Technological Forecasting & Social Change*. — 2006. — Vol. 73, No. 8. — P. 1013–1027. — <http://dx.doi.org/10.1016%2Fj.techfore.2006.05.020>
32. *Tuomaala O., Järvelin K., Vakkari P.* Evolution of library and information science, 1965–2005: Content analysis of journal articles // *Journal of the Association for Information Science and Technology*. — 2014. — Vol. 65, No. 7. — P. 1446-1462. — <https://doi.org/10.1002/asi.23034>
33. *Luo L., McKinney M.* JAL in the past decade: A comprehensive analysis of academic library research // *The Journal of Academic Librarianship*. — 2015. — Vol. 41, No. 2. — P. 123-129. — <https://doi.org/10.1016/j.acalib.2015.01.003>
34. *Liu G., Yang L.* Popular research topics in the recent journal publications of library and information science // *The Journal of Academic Librarianship*. — 2019. — Vol. 45, No. 3. — P. 278-287. — <https://doi.org/10.1016/j.acalib.2019.04.001>
35. *Zong Q. J., Shen H. Z., Yuan Q. J., Hu X. W., Hou Z. P., Deng S. G.* Doctoral dissertations of Library and Information Science in China: A co-word analysis // *Scientometrics*. — 2013. — Vol. 94, No. 2. — P. 781-799. — <https://doi.org/10.1007/s11192-012-0799-1>
36. *Anna NE. V., Mannan E. F., Srirahayu D. P., Mutia F.* Library and Information (LIS) research topics in indonesia from 2006 to 2017 // *Library Philosophy and Practice (e-journal)*. — 2018. — <https://digitalcommons.unl.edu/libphilprac/1773>
37. *Garner J., Davidson K., Williams V. K.* Identifying serials trends through twenty years of NASIG conference proceedings: A content analysis // *Serials Review*. — 2008. — Vol. 34, No. 2. — P. 88-103. — <https://doi.org/10.1080/00987913.2008.10765160>
38. *Zins Cb.* Knowledge map of information science // *Journal of the American Society for Information Science and Technology*. — 2007. — Vol. 58, No. 4. — P. 526-535. — <https://doi.org/10.1002/asi.20505>
39. *Figuerola C. G., Marco F. J. G., Pinto M.* Mapping the evolution of library and information science (1978–2014) using topic modeling on LISA // *Scien-*

tometrics. — 2017. — Vol. 112, No. 3. — P. 1507-1535. — <https://doi.org/10.1007/s11192-017-2432-9>

40. *Han X.* Evolution of research topics in LIS between 1996 and 2019: an analysis based on latent Dirichlet allocation topic model // *Scientometrics*. — 2020. — Vol. 25, No. 3. — P. 2561–2595. — <https://doi.org/10.1007/s11192-020-03721-0>

41. *Xiao X., Zhang F., Li J.* Library and Information Science research in China: A survey based analysis of 10 LIS educational institutes // *The Journal of Academic Librarianship*. — 2017. — Vol. 41, No. 3. — P. 330–340. — <https://doi.org/10.1016/j.acalib.2015.02.012>

42. *Sarah M., Pete D.* Striking a balance: Priorities for research in LIS // *Library Review*. — 2004. — Vol. 53, No. 3. — P. 167-176. — <https://doi.org/10.1108/00242530410526574>

43. *Summers R., Oppenheim C. H., Meadows J., McKnight C., Kinnell M.* Information Science in 2010: A Loughborough University view // *Journal of the American Society for Information Science*. — 1999. — Vol. 50, No. 12. — P. 1153-1162. — [https://doi.org/10.1002/\(SICI\)1097-4571\(1999\)50:12%3C1153::AID-ASI20%3E3.0.CO;2-H](https://doi.org/10.1002/(SICI)1097-4571(1999)50:12%3C1153::AID-ASI20%3E3.0.CO;2-H)

44. *Kumar Sinha M.* Scenario of changing trends in library and information science education and research an analytical study of brochures of seminars, conferences and workshops organized during January 2012-July 2014 // *Journal of Humanities and Social Science*. — 2016. — Vol. 21, No. 7. — P. 20-43.

45. *Sobeili F., Khasseh A., koranian P.* Thematic trends of concepts in Knowledge and Information Science based on co-word analysis in Iran // *National Studies on Librarianship and Information Organization (NASTINFO)*. — 2018. — Vol. 29, No. 2. — P. 171-190. (перс. яз)

46. *Mansourkiaie R., Babalbavaegi F., Nooshinfard F., Sobeili F.* Study of the future of the dissemination of scientific productions of Knowledge and Information Science in Social Networks from the Iranian experts' viewpoint // *Library and Information Sciences*. — 2019. — Vol. 22, No 3. — P. 136-163. (перс. яз.)

47. *Abdollahzadeh P.* Mapping research topics of Library and Information Sciences based on co-word analysis. Master Thesis. — Tabriz University of Medical Science, 2019.

48. *Baghmohammad M., Mansouri A., Cheashmehsobrabi M.* Identification of topic development process of knowledge and information science field based on the topic modeling (LDA) // *Iranian Journal of Information processing and Management*. — 2020. — Vol. 36, No. 2. — P. 297-328.

49. *Taysem Z.* Forecasting the future of library and information science and its sub-fields // *Scientomet-*

*rics*. — 2021. — Vol. 126. — P. 1527–1551. — <https://doi.org/10.1007/s11192-020-03800-2>

50. *Olab C.* Understanding LSTM networks. 2015. — <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>

51. *Hochreiter J.* Investigations on dynamic neural networks. Diploma thesis. — Technical University of Munich, Germany, 1991.

52. *Bengio Y., Simard P., Frasconi P.* Learning long-term dependencies with gradient descent is difficult // *IEEE Transactions on Neural Networks*. — 1994. — Vol. 5, No. 2. — P. 157-166. — <https://doi.org/10.1109/72.279181>

53. *Schwab K.* The Fourth Industrial Revolution. — Switzerland: World Economic Forum, 2017.

54. *Carlson C. N.* Information overload, retrieval strategies and internet user empowerment. 2003. — [http://eprints.rclis.org/5432/1/Information\\_Overload.pdf](http://eprints.rclis.org/5432/1/Information_Overload.pdf)

55. *Mostafavi I., Osareh F., Tavakolizadeh-Ravari M.* Identifying content structure of knowledge and information science (KIS) studies based on co-word analysis of articles in Web of Science (WoS) database (2009-2013) // *Iranian Journal of Information processing and Management*. — 2018. — Vol. 33, No. 3. — P. 1271-1300. (перс. яз.)

56. *Tang R.* Evolution of the interdisciplinary characteristics of information and library science // *Proceedings of the American Society for Information Science and Technology*. — 2004. — Vol. 41, No. 1. — P. 54-63. — <https://doi.org/10.1002/meet.1450410107>

57. *Chang Y. W., Huang M. H.* A study of the evolution of interdisciplinarity in Library and Information Science: Using three bibliometric methods // *Journal of the American Society For Information Science*. — 2011. — Vol. 63, No. 1. — P. 22-33. — <https://doi.org/10.1002/asi.21649>

58. *Hessey R., Willett P.* Quantifying the value of knowledge exports from librarianship and information science research // *Journal of information science*. — 2012. — Vol. 39, No. 1. — P. 141-150. — <https://doi.org/10.1177%2F0165551512442476>

59. *Dora M., Kumar H. A.* National and international trends in library and information science research: A comparative review of the literature // *IFLA Journal*. — 2019. — Vol. 46, No. 3. — P. 234-249. — <https://doi.org/10.1177%2F0340035219886610>

60. *Ghanadinezhad F., Heidari G.* Identifying and analyzing research priorities in each of the research axes of knowledge and information science from the viewpoint of the faculty and doctoral students of this field // *Iranian Journal of Information processing and Management*. — 2018. — Vol. 34, No. 1. — P. 57-88. (перс. яз.)

## Инновационное развитие Узбекистана: анализ сильных и слабых позиций в рейтинге глобального инновационного индекса

**А. А. АБДУВАЛИЕВ**

Центр научно-технической информации  
при Министерстве инновационного  
развития Республики Узбекистан,  
г. Ташкент, Республика Узбекистан

*С целью выявления конкурентных преимуществ инновационного развития Узбекистана по вхождению в Top-50 стран рейтинга ГИИ к 2030 г. проанализированы сильные и слабые позиции Республики в данном рейтинге. Показано увеличение количества стран с переходной экономикой в группе рейтинга Top-50 стран. Для определения возможности вхождения Узбекистана в Top-50 стран ГИИ к 2030 г. проведен сопоставительный анализ сильных/слабых индексов/субиндексов Узбекистана за 2021 г. со странами с переходной экономикой из группы Top-50.*

### ВВЕДЕНИЕ

За последние годы структурно изменилась научная и инновационная деятельность Узбекистана, в основу которой заложены такие системные меры по созданию благоприятных условий для ее развития, как: стабильное увеличение государственных затрат на науку и инновации; укрепление нормативно-правовой базы; формирование кадрового потенциала; активная интеграция образования, науки и производства; совершенствование механизмов трансферта технологии и коммерциализация научных разработок.

Одним из ключевых показателей результативности реформ, проводимых в построении инновационной экономики Узбекистана, является улучшение позиции страны в рейтинге Глобального инновационного индекса (ГИИ).

Учитывая необходимость реформ в области инновационного развития и актуальность их

ускорения, Президент Республики Узбекистан Ш.М. Мирзиёев в Стратегии действий по пяти приоритетным направлениям развития Узбекистана в 2017-2021 гг. поставил задачу по вхождению Республики к 2030 г. в число 50-ти передовых стран рейтинга ГИИ [1].

### МАТЕРИАЛЫ, МЕТОДЫ И РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Публикуемый ежегодно рейтинг Глобального инновационного индекса более 130 стран и территорий в области инновационной деятельности позволяет оценивать результативность проводимых реформ в этой области. Анализ группы стран ежегодного рейтинга ГИИ свидетельствует о наличии таких критериев оценок, как «сильный» и «слабый» индекс/субиндекс, присвоенных экспертами ГИИ по результативности развития инновационной экономики каждой страны.

Детальный анализ сильных и слабых индексов/субиндексов рейтинга ГИИ страны позволяет обосновать как конкурентоспособные направления, так и отстающие позиции формирующейся инновационной экономики страны. Результаты такого анализа полезны для разработки прогнозных показателей инновационного развития страны.

Количество стран в рейтинге ежегодно меняется, и если в 2012 г. участвовала 141 страна, то в 2021 г. их было уже 132. В разные годы этот показатель постоянно изменялся, что свидетельствует о возросшей конкуренции среди стран-участников. Так, по заключению экспертов ГИИ, страны со средним уровнем дохода на душу населения, такие как Китай, Турция, Вьетнам, Индия и Филиппины изменили глобальный инновационный ландшафт [2, р. 21]. При этом развивающимся странам намного труднее достигнуть положительных показателей в рейтинге по сравнению с развитыми странами.

За девятилетний период участия Узбекистана в рейтинге ГИИ его рейтинг повысился на 41 позицию и в 2021 г. страна заняла 86 место, а не 127-е место в 2012 г. (рис. 1).

Решая задачу вхождения Узбекистана в Топ-50 стран рейтинга ГИИ к 2030 г., важно определить наличие конкуренции между странами-участниками

Топ-50. Это можно сделать на основе сопоставительного анализа за период 2012 г. и 2021 г. (табл. 1 и 2).

Результат группирования и обобщения сильных и слабых индексов/субиндексов 50-ти передовых стран рейтинга по итогам 2012 г. показывает, что во всех странах выявлены как сильные, так и слабые показатели. Средний показатель сильных индексов составил 13,98 ед., а средний показатель слабых индексов – 10,96 ед. Наибольшее количество сильных индексов/субиндексов достигли Дания и Оман (19 ед.), минимальное количество 10 ед. – Франция, Япония, Венгрия и Польша.

По сравнению с 2012 г. результаты 2021 г. по общему количеству сильных/слабых индексов/субиндексов характеризуются количественным ростом сильных показателей (табл. 2).

Данные табл. 2 свидетельствуют о том, что наибольшее количество сильных индексов – 23 ед. у Швейцарии, далее следуют США и Сингапур – 22 ед., наименьшего результата в сильном индексе достигла Украина – 10 ед. Наибольшее количество слабых индексов/субиндексов – 13 ед. отмечено у Ирландии и наименьший слабый показатель 7 ед. у Новой Зеландии. В целом, средний показатель сильных индексов/субиндексов составил 14,14 ед. и средний показатель слабых – 10,7 ед.

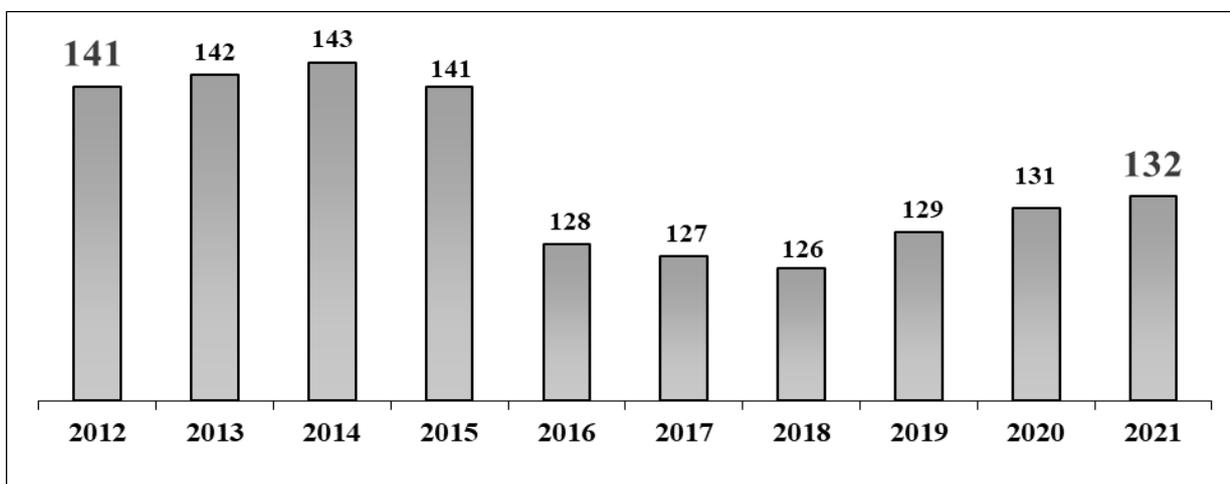


Рис. 1. Динамика рейтинга Узбекистана в ГИИ (2012-2021 гг.)  
[3, р. xix; 4, р. xxi; 5, р. xxv; 6, р. xxi; 7, р. xxxiii; 2, р. 4]

Таблица 1

**Общее количество сильных/слабых индексов/субиндексов топ-50 стран рейтинга ГИИ в 2012 г. [3, р.178-318]**

А*	В	С	А	В	С
1	Швейцария	15/11	26	Словения	14/11
2	Швеция	14/10	27	Республика Чехия	15/11
3	Сингапур	17/10	28	Кипр	14/10

<b>А*</b>	<b>В</b>	<b>С</b>	<b>А</b>	<b>В</b>	<b>С</b>
4	Финляндия	11/10	29	Испания	14/12
5	Великобритания	17/10	30	Латвия	14/10
6	Нидерланды	16/11	31	Венгрия	10/11
7	Дания	19/10	32	Малайзия	12/10
8	Гонконг	16/11	33	Катар	16/13
9	Ирландия	17/11	34	Китай	14/12
10	США	16/10	35	Португалия	13/11
11	Люксембург	15/8	36	Италия	13/11
12	Канада	14/11	37	ОАЭ	14/12
13	Новая Зеландия	16/9	38	Литва	12/11
14	Норвегия	14/10	39	Чили	13/11
15	Германия	13/10	40	Словакия	16/11
16	Мальта	16/12	41	Бахрейн	14/10
17	Израиль	14/10	42	Хорватия	11/13
18	Исландия	12/11	43	Болгария	12/12
19	Эстония	14/11	44	Польша	10/13
20	Бельгия	14/11	45	Монтенегро	15/13
21	Южная Корея	13/12	46	Сербия	12/11
22	Австрия	15/10	47	Оман	19/12
23	Австралия	14/11	48	Саудовская Аравия	15/12
24	Франция	10/10	49	Мавритания	12/13
25	Япония	10/12	50	Молдова	13/10

\*Примечание: **А** – Рейтинг страны в ГИИ; **В** – 50-передовых стран рейтинга ГИИ в 2021 г.;  
**С**– Общее количество сильных/слабых позиций.

Таблица 2

**Общее количество сильных/слабых индексов/субиндексов топ-50 стран рейтинга ГИИ в 2021 г. [2, р.84-173]**

<b>А*</b>	<b>В</b>	<b>С</b>	<b>А</b>	<b>В</b>	<b>С</b>
1	Швейцария	23/10	26	Новая Зеландия	15/7
2	Швеция	16/10	27	Мальта	14/12
3	США	22/10	28	Кипр	12/10
4	Великобритания	17/10	29	Италия	14/12
5	Южная Корея	18/10	30	Испания	13/11
6	Нидерланды	17/10	31	Португалия	14/11
7	Финляндия	18/10	32	Словения	11/11
8	Сингапур	22/11	33	ОАЭ	14/11
9	Дания	13/10	34	Венгрия	13/12
10	Германия	15/11	35	Болгария	13/12
11	Франция	12/10	36	Малайзия	13/10
12	Китай	17/11	37	Словакия	13/11
13	Япония	18/11	38	Латвия	11/10
14	Гонконг	16/11	39	Литва	12/9
15	Израиль	14/12	40	Польша	12/10
16	Канада	17/10	41	Турция	12/12
17	Исландия	11/11	42	Хорватия	11/9
18	Австрия	13/11	43	Таиланд	12/11
19	Ирландия	15/13	44	Вьетнам	12/11
20	Норвегия	15/10	45	Индия	12/11
21	Эстония	12/10	46	Россия	12/12
22	Бельгия	13/10	47	Греция	12/11
23	Люксембург	12/11	48	Румыния	14/12
24	Чехия	15/11	49	Украина	10/12
25	Австралия	14/11	50	Монтенегро	11/10

\*Примечание: **А** – Рейтинг страны в ГИИ; **В** – 50-передовых стран рейтинга ГИИ в 2021 г.;  
**С**– общее количество сильных/слабых позиций.

Сопоставление показателей табл. 1 и табл. 2 позволяет отметить, что в группе 50-ти передовых стран рейтинга ГИИ в 2021 г. по сравнению с 2012 г. отмечены следующие изменения, которые подтверждают усиление конкурентной борьбы в данной группе стран:

1) изменился состав стран группы Топ-50. Так, в 2021 г. из списка выбыли такие страны как, Катар, Чили, Бахрейн, Сербия, Оман, Саудовская Аравия, Мавритания и Молдова, но вошли: Турция, Таиланд, Вьетнам, Индия, Россия, Греция, Румыния и Украина.

Эти изменения позволяют заключить, что конкурентная борьба, по результатам которой выбыли действующие и вошли новые страны, затронула в основном страны с переходной экономикой. В группе Топ-50-ти стран увеличилось число стран с уровнем дохода ниже среднего на душу населения и уменьшилось количество стран с высоким уровнем дохода. Так, если в 2012 г. в группу Топ-50 стран с уровнем дохода ниже среднего на душу населения входила только Молдова (50) [3, р. xviii–xix], то в 2021 г. вошли уже три страны: Вьетнам (44), Индия (45) и Украина (49) [10, р.4];

2) увеличилось количество сильных индексов/субиндексов стран группы Топ-50. Если в 2012 г. наиболее высокий показатель составил 19 ед., то в 2021 г. – 23 ед. (Швейцария). Тем не менее, несмотря на количественные изменения как сильных, так и слабых индексов/субиндексов, в средних значениях этих показателей существенных изменений не наблюдается. Так если в 2012 г. сильные индексы/субиндексы составили 13,98 ед., то в 2021 г. – 14,14 ед.; слабые индексы/субиндексы уменьшились незначительно 10,96 (в 2012 г.) и 10,7 (в 2021 г.), соответственно.

3) на основе вышеизложенного можно заключить, что наибольшая сумма сильных индексов не является основным фактором высокого рейтинга в ГИИ. Так, в рейтинге 2012 г. из Топ-50 стран самое большое количество, 19 сильных индексов набрано в Дании (7 место в рейтинге) и в Омане (47 место). Количественный показатель слабых индексов/субиндексов не влияет негативно на понижение рейтинга. Например, Швейцария, имея 11 слабых показателей, занимает в 2012 г. 1-е место, а Швеция, Сингапур, Финляндия и Великобритания, имея по 10 слабых показателей, вошли в пятерку ведущих 50-ти стран в рейтинге ГИИ.

С целью реализации задачи по вхождению Узбекистана в Топ-50 стран ГИИ к 2030 г., можно сопоставить показатели Вьетнама, Индии, Украины и Узбекистана со странами из группы Топ-50 за 2021 г. (табл. 3).

Приведенный сопоставительный анализ свидетельствует, что количество сильных индексов/субиндексов Узбекистана (12 ед.) идентично с индексами Индии (12) и Вьетнама (12), а в Украине чуть меньше (10 ед.). Совпадение наблюдается и по са-

мым показателям, в числе которых такие индексы, как соотношение учителей и учеников, выпускники факультетов естественных наук и инженерии, интенсивность местной конкуренции, темп роста ВВП на одного работника по ППС и др. Такое сопоставление послужит ориентиром для Узбекистана при формировании конкурентоспособной инновационной политики.

Следует отметить, что достигнутые сильные индексы/субиндексы Узбекистана являются результатом комплексной социально-ориентированной экономической политики страны по формированию инновационной экономики.

Помимо анализа сильных показателей ГИИ, необходимо сравнить и слабые индексы/субиндексы (табл. 4).

Данные табл. 4 подтверждают совпадение слабых индексов/субиндексов (таких, как доля иностранных граждан в высшем образовании, глобальные компании в области исследований и разработок и др.) и их количества, например у Вьетнама и Индии по 11, у Украины 12 и о наименьшем количестве слабых показателей у Узбекистана – 10.

Суммарные показатели табл. 3 и табл. 4 отражают, что в количественном выражении как сильные, так и слабые индексы/субиндексы в анализируемых странах почти одинаковы, но имеются случаи их совпадения. Тем не менее, есть отдельные сильные и слабые индексы у Узбекистана, динамику которых целесообразно рассмотреть в период 2012–2021 гг., что позволит выделить из них конкурентоспособные и выявить продолжительность слабых индексов/субиндексов.

Динамика сильных и слабых индексов/субиндексов Узбекистана в период 2012–2021 гг. (рис. 2) позволяет отметить стабильность количественного показателя сильных индексов/субиндексов (12 ед.), но не стабильность (от 11 до 19 ед.) слабых показателей. Поэтому важно определить разовость или стабильность (продолжительность) достигнутых как сильных, так и слабых индексов/субиндексов Узбекистана в период 2012–2021 гг. (табл. 5)

Данные табл. 5 свидетельствуют о наличии конкурентных преимуществ в представленных направлениях. Соответственно, в дальнейшей политике инновационного развития Узбекистана необходимо учитывать меры сохранения достигнутых конкурентных преимуществ, так как некоторыми индексами/субиндексами за последнее время были утрачены конкурентные преимущества, например, в таких направлениях, как «Легкость разрешения банкротства», «Образование», «Кредитный портфель микрофинансовых организаций», «Торговля, конкуренция и масштаб рынка» и др.

Анализ слабых индексов/субиндексов Узбекистана в рейтинге ГИИ позволяет определить наличие продолжительных проблем в определенных направлениях (табл. 6).

## Сильные индексы/субиндексы Вьетнама, Индии, Украины и Узбекистана в 2021 г. [2]

№ в рейтинге	Индексы/субиндексы	Вьетнам [р. 170]	Индия [р. 93]	Украина [р.163]	Узбекистан [р.169]
1.3.1	Легкость открытия бизнеса				
2.1.1	Расходы на образование (% ВВП)				
2.1.2	Государственное финансирование на учащегося, в среднем, % ВВП на душу населения				
2.1.5	Соотношение учеников и учителей, в среднем образовании				
2.2.1	Прием в высшие образовательные учреждения (% от общего количества поступающих)				
2.2.2	Выпускники факультетов естественных наук и инженерии, %				
2.3.3	Глобальные компании в области исследований и разработок, средние расходы млн долл. США				
2.3.4	Средняя оценка топ 3 университетов по рейтингу QS				
3.1.3	Услуги электронного правительства				
<b>3.2</b>	<b>Общая инфраструктура</b>				
3.2.3	Валовое накопление капитала, % ВВП				
4.1	Кредит				
4.1.2	Внутренний кредит частному сектору, % ВВП				
4.1.3	Кредитный портфель микрофинансовых организаций, % ВВП				
4.2.1	Простота защиты миноритарных инвесторов				
<b>4.3</b>	<b>Торговля, конкуренция и масштаб рынка</b>				
4.3.2	Интенсивность местной конкуренции				
4.3.3	Объем внутреннего рынка, млрд долл. США по ППС				
5.1.4	Валовые расходы на исследования и разработки, профинансированные бизнес-предприятиями, %				
5.1.5	Работающие женщины с учеными степенями, %				
5.3.2	Импорт высоких технологий, % от общего объема торговли				
5.3.4	Чистый приток ПИИ, % ВВП				
6.1.3	Полезные модели по происхождению / млрд ВВП в долл. США по ППС				
6.1.5	Цитируемость документов (индекс Хирша – H-index)				
<b>6.2</b>	<b>Влияние знаний</b>				
6.2.1	Темп роста ВВП на одного работника по ППС, %				
6.2.3	Расходы на компьютерное программное обеспечение (% от ВВП)				
<b>6.3</b>	<b>Распространение знаний</b>				
6.3.3	Экспорт услуг ИКТ, % от общего объема торговли				
7.1.1	Товарные знаки по происхождению / млрд долл. США по ППС ВВП				
7.1.3	Промышленные образцы по происхождению / млрд ВВП в долл. США по ППС				
7.2.1	Экспорт культурных и творческих услуг, % в общем объеме торговли				
7.2.5	Экспорт творческих товаров, % в общем объеме торговли				
7.3.4	Создание мобильных приложений / млрд ВВП в долл. США по ППС				

**Слабые индексы/субиндексы Вьетнама, Индии, Украины  
и Узбекистана в 2021 г. [2]**

№ в рейтинге	Индексы/субиндексы	Вьетнам [р.170]	Индия [р.93]	Украина [р. 163]	Узбекистан [р.169]
1.1.1	Политическая стабильность				
1.2.1	Качество регуляторов				
1.3.2	Легкость разрешения банкротства				
2.1.3	Продолжительность образования				
2.1.5	Соотношение учеников и учителей, среднее образование				
2.2.3	Доля иностранных граждан в высших образовательных учреждениях (%)				
2.3.3	Глобальные компании в области исследований и разработок, средние расходы млн долл. США				
2.3.4	Средняя оценка топ 3 университетов по рейтингу QS				
3.1.1	Доступ к ИКТ				
3.1.2	Охват населения ИКТ				
<b>3.2</b>	<b>Общая инфраструктура</b>				
3.2.3	Валовое накопление капитала, % ВВП				
3.3.1	ВВП на единицу потребления энергоресурсов				
3.3.2	Экологические показатели				
4.1.3	Кредитный портфель микрофинансовых организаций, % ВВП				
<b>4.2</b>	<b>Инвестиции</b>				
4.2.2	Рыночная капитализация, % ВВП				
4.2.4	Принятие венчурного капитала / млрд ВВП в долл. США по ППС				
5.1.1	Занятость в наукоемких отраслях,%				
5.1.5	Работающие женщины с учеными степенями,%				
5.2.3	Валовые расходы на исследования и разработки, финансируемые иностранной стороной,% ВВП				
5.2.4	Совместное предприятие/стратегические альянсы / млрд ВВП в долл. США по ППС				
5.3.3	Импорт услуг ИКТ, % в общем объеме торговли				
6.1.2	Международные заявки по международному сотрудничеству в области патентной кооперации / млрд долл. США по ППС ВВП				
6.1.4	Научно-технические публикации / млрд ВВП в долл. США по ППС				
6.2.2	Количество новых предприятий на душу населения (15 – 64 лет)				
6.3.1	Поступления от интеллектуальной собственности, % в общем объеме торговли				
6.3.4	Чистый отток ПИИ, % ВВП				
7.2.1	Экспорт культурных и творческих услуг, % в общем объеме торговли				
7.2.2	Созданные национальные художественные фильмы				
7.2.3	Рынок развлечений и медиа				
7.2.4	Полиграфия и другие носители, % производство				
7.3.1	Общие домены верхнего уровня (TLD) / населен.				
7.3.3	Ежегодные изменения в Википедии				
7.3.4	Создание мобильных приложений / млрд ВВП в долл. США по ППС				



Рис. 2. Динамика общего количества сильных/слабых индексов/субиндексов Узбекистана в рейтинге ГИИ (2012-2021 гг.) [3, р. xix; 4, р. xxi; 5, р. xxv; 6, р. xxi; 7, р. xxxiii; 2, р. 4]

Таблица 5

**Сильные индексы и субиндексы Узбекистана в рейтинге ГИИ (2012-2021 гг.)**  
[3, р. xix; 4, р. xxi; 5, р. xxv; 6, р. xxi; 7, р. xxxiii; 2, р. 4]

№ в индексе	Индексы/субиндексы	Годы					
		2012	2013	2014	2015	2020	2021
1.3.1	Легкость открытия бизнеса						
1.3.2	Легкость разрешения банкротства						
<b>2</b>	<b>Человеческий капитал и исследования</b>						
<b>2.1</b>	<b>Образование</b>						
2.1.1	Расходы на образование, % ВВП						
2.1.5	Соотношение учеников и учителей, среднее образование						
2.2.2	Выпускники факультетов естественных наук и инженерии, %						
3.1.3	Услуги электронного правительства						
3.1.4	Индекс электронного участия						
<b>3.2</b>	<b>Общая инфраструктура</b>						
3.2.3	Валовое накопление капитала, % ВВП						
4.1.3	Кредитный портфель микрофинансовых организаций, % ВВП						
4.2.1	Простота защиты миноритарных инвесторов						
4.2.3	Сделки по венчурному капиталу / ВВП в млрд долл. США по ППС						
<b>4.3</b>	<b>Торговля, конкуренция и масштаб рынка</b>						
4.3.2	Интенсивность местной конкуренции						
5.2.4	Совместное предприятие/стратегические альянсы / ВВП в млрд долл. США по ППС						
<b>5.3</b>	<b>Поглощение знаний</b>						
<b>6</b>	<b>Знания и технологии</b>						
6.1.1	Патентные заявки по происхождению / млрд долл. США по ППС ВВП						
6.1.3	Полезные модели по происхождению / ВВП в млрд долл. США по ППС						
<b>6.2</b>	<b>Влияние знаний</b>						
6.2.1	Темп роста ВВП на одного работника по ППС, %						
7.1.1	Товарные знаки по происхождению / млрд долл. США по ППС ВВП						
7.2.1	Экспорт культурных и творческих услуг, % в общем объеме торговли						
7.2.2	Созданные национальные художественные фильмы						

**Слабые индексы и субиндексы Узбекистана в рейтинге ГИИ (2012-2021 гг.) [3, р.хix; 4, р.хxi; 5, р.ххv; 6, р.хxi; 7, р.хххiii; 2, р.4]**

№ в индексе	Индексы/субиндексы	Годы					
		2012	2013	2014	2015	2020	2021
1.1.3	Свобода средств массовой информации						
1.2.1	Качество регуляторов						
1.2.2	Верховенство закона						
2.2.3	Доля иностранных граждан в высших образовательных учреждениях, %						
2.3.3	Глобальные компании в области исследований и разработок, средние расходы млн долл. США						
2.3.4	Средняя оценка топ 3 университетов по рейтингу QS						
3.3.1	ВВП на единицу потребления энергоресурсов						
3.3.2	Экологические показатели						
4.1.3	Кредитный портфель микрофинансовых организаций, % ВВП						
4.2.2	Рыночная капитализация, % ВВП						
<b>5</b>	<b>Развитость бизнеса</b>						
<b>5.1</b>	<b>Работники умственного труда</b>						
5.1.2	Фирмы, предлагающие официальное обучение, %						
<b>5.2</b>	<b>Инновационные связи</b>						
5.2.3	Валовые расходы на исследования и разработки, финансируемые иностранной стороной, % ВВП						
6.1.2	Международные заявки по международному сотрудничеству в области патентной кооперации / млрд долл. США по ППС ВВП						
<b>7</b>	<b>Креативность</b>						
7.1	Нематериальные активы						
7.2.3	Рынок развлечений и медиа						
7.3.1	Общие домены верхнего уровня (TLD)						
7.3.4	Создание мобильных приложений / млрд ВВП в долл. США по ППС						

Из продолжительных слабых индексов/субиндексов рейтинга ГИИ Узбекистана выделяются: «качество регуляторов», «верховенство закона», «средняя оценка топ 3 университетов по рейтингу QS», «креативность», «общие домены верхнего уровня (TLD)».

В целом, анализ слабых индексов/субиндексов Узбекистана в рейтинге ГИИ позволяет сделать следующие выводы:

- во-первых, особое внимание необходимо уделить таким показателям, которые из сильных переходят в слабые, например, «кредитный портфель микрофинансовых организаций»;

- во-вторых, требуют внимания только формирующиеся и переходящие на стабильный уровень слабые показатели, например, «глобальные компании, в области исследований и разработок», «кредитный портфель микрофинансовых организаций», «валовые расходы на исследования и разработки, финансируемые иностранной стороной», «создание мобильных приложений» и др.

## **ЗАКЛЮЧЕНИЕ**

Для выявления возможности реализации задачи, поставленной Президентом Республики Узбекистан Ш.М. Мирзиёевым по вхождению страны в Топ-50 стран рейтинга Глобального инновационного индекса (ГИИ) к 2030 г., был проведен анализ сильных и слабых позиции Республики в этом рейтинге. Данный анализ основывался на обобщении сильных и слабых позиций стран Топ-50 рейтинга ГИИ и показателей Узбекистана в сравнении с развивающимися странам из данной группы за 2012-2021 гг.

По результатам анализа были сделаны следующие выводы и предложения.

**1. Установлено, что в мировом пространстве идет конкурентная борьба за высокие позиции как в целом в рейтинге ГИИ, так и в группе стран, входящих в Топ-50.** Это подтверждается ежегодным изменением общего количества стран-участников (тенденция уменьшения) в рейтинге, а

также, ежегодным изменением состава стран в группе Топ-50.

**2. В группе рейтинга Топ-50 стран увеличивается количество стран с переходной экономикой с уровнем дохода ниже среднего на душу населения.** Анализ показал, если в 2012 г. в этой группе была только Молдова, то в 2021 г. в список вошли три страны с переходной экономикой – Вьетнам, Индия и Украина.

**3. Выявлена тенденция увеличения количества сильных индексов/субиндексов в странах группы Топ-50.** Если в 2012 г. наиболее высокий показатель составил 19 ед., то в 2021 г. 23 ед. (Швейцария). Тем не менее, наибольшая сумма сильных индексов не является основным фактором высокого рейтинга в ГИИ. Так, в рейтинге 2012 г. из группы Топ-50 стран самое большое количество (19) сильных индексов набрано Данией (7 место в рейтинге) и в Омане (47 место).

**4. Количественный показатель слабых индексов/субиндексов не влияет негативно на понижение рейтинга.** Например, Швейцария, имея 11 слабых показателей, не утратила 1-е место, а Швеция, Сингапур, Финляндия и Великобритания, имея по 10 слабых показателей, вошли в пятерку Топ-50 рейтинга ГИИ.

**5. С целью определения возможности вхождения Узбекистана в Топ-50 стран ГИИ к 2030 г. проведено сопоставление сильных/слабых индексов/субиндексов Узбекистана за 2021 г. с индексами таких стран, как Вьетнам, Индия и Украина.** Результаты анализа показывают, что в количественном выражении как сильные, так и слабые индексы/субиндексы в анализируемых странах почти одинаковы, в то же время наблюдаются случаи их идентичности.

Тем не менее, для обоснования конкурентоспособности сильных и выявления глубины слабых индексов/субиндексов Узбекистана суммированы показатели 2012-2021 гг. Стабильность количественных показателей и продолжительность неко-

торых сильных индексов/субиндексов свидетельствуют о сформированных конкурентных преимуществах в соответствующих направлениях.

Анализ слабых индексов/субиндексов ГИИ Узбекистана позволил определить продолжительность и нерешенность проблем в некоторых направлениях.

Вкратце, для достижения поставленной цели по вхождению Узбекистана к 2030 г. в Топ-50 стран рейтинга ГИИ политика инновационного развития страны должна ориентироваться на сохранение достигнутых конкурентных преимуществ. Наш анализ показывает, что некоторыми индексами/субиндексами за последние периоды были утрачены сильные позиции. В то же время принятие незамедлительных мер по устранению продолжительных слабых индексов/субиндексов позволит обеспечить недопущение перехода из сильных позиций в ряды слабых позиций индексов/субиндексов («кредитный портфель микрофинансовых организаций»).

## ЛИТЕРАТУРА

1. Указ Президента Республики Узбекистан «Об утверждении Стратегии инновационного развития Республики Узбекистан на 2019-2021 год» // УП-5544 от 21.09.2018 г. — <https://lex.uz/docs/3913186>

2. *Global Innovation Index* Tracking Innovation through the COVID-19 Crisis. — 2021. — 14 th Edition. — P.226.

3. *Global Innovation Index*. Stronger Innovation Linkages for Global Growth. — 2012. — P.464.

4. *Global Innovation Index*. The Local Dynamics of Innovation. — 2013. — P.417.

5. *Global Innovation Index*. The Human Factor in Innovation. — 2014. — P.429.

6. *Global Innovation Index*. Effective Innovation Policies for Development. — 2015. — P.453.

7. *Global Innovation Index*. Who Will Finance Innovation? — 2020. — P.448.

Приглашаем российских и зарубежных авторов к сотрудничеству  
в журнале «Международный форум по информации».  
Оригинальные статьи и другие материалы (рецензии, письма)  
можно присылать на русском или английском языке  
по почтовому адресу, указанному в «Памятке для авторов»  
или по электронной почте: [mfi@viniti.ru](mailto:mfi@viniti.ru).

Ответственный за выпуск *Е.А. Овченкова*

Компьютерная верстка *М. А. Филимонова*

ИД № 04689 от 28.04.2001 г.

Подписано в печать 05.09.2022 г.

Бумага офсетная. Формат 60x84 1/8. Гарн. литер. Печать цифровая

Усл. печ. л. 5,50 Уч.-изд. л. 5,91 Тираж 33 экз.

Адрес редакции: 125190, Россия, г. Москва, ул. Усиевича, д. 20.

Тел. (499) 155-44-95

Отпечатано в ООО «Типография Форпринт».

г. Москва, ул. Маршала Бирюзова, д. 1, корп. 11. Тел. +7 (495) 211-38-33



Федеральное государственное бюджетное учреждение науки  
Всероссийский институт научной и технической информации  
Российской академии наук  
(ВИНИТИ РАН)

## **ИНФОРМАЦИОННОЕ ПИСЬМО**

**о проведении 10-й научной конференции с международным участием**

**«НТИ-2022. Научная информация в современном мире:**

**глобальные вызовы и национальные приоритеты»,**

**посвященной 70-летию ВИНИТИ РАН**

**(25 – 26 октября 2022 г.)**

Всероссийский институт научной и технической информации Российской академии наук (ВИНИТИ РАН), один из крупнейших мировых информационных центров, приглашает принять участие в работе **10-й научной конференции с международным участием «НТИ-2022. Научная информация в современном мире: глобальные вызовы и национальные приоритеты».**

Основная цель конференции – обсуждение актуальных проблем информационного обеспечения науки, аналитической и инновационной деятельности, разработки и внедрения методов интеллектуальной обработки информации, а также обмен опытом между профильными научными, производственными и учебными организациями.

Для участия в конференции приглашаются ученые и специалисты РАН, вузовской и отраслевой науки в области информационных технологий, работники информационных центров и библиотек, служб распространения информационных продуктов и услуг, представители издательств.

На «НТИ-2022» предполагается обсуждение следующих вопросов:

- информационное обеспечение научных исследований в условиях санкционного давления;
- роль реферативных служб в современном мире;
- информационно-библиотечная деятельность и обслуживание;
- развитие классификационных систем;
- проблемы науковедения и наукометрии;
- интеллектуальные технологии обработки информации;
- международное сотрудничество в сфере НТИ;
- методы и технологии популяризации научных знаний.

На конференции планируется проведение пленарного заседания, круглых столов и тематических секций, в том числе отдельной секции молодых ученых, интересующихся проблемами научно-информационного обеспечения научных исследований.

**Рабочие языки конференции:** русский и английский.

**Формат участия:** очно, заочно.

**Участие в конференции бесплатное.**

По итогам конференции будет подготовлен сборник докладов (РИНЦ), лучшие из которых, отобранные экспертами Программного комитета, предполагается опубликовать в журналах «Научно-техническая информация. Серия 1. Организация и методика информационной работы» (РИНЦ, переводная версия издания включена в международные индексы цитирования Scopus, Web of Science) и «Научно-техническая информация. Серия 2. Информационные процессы и системы» (входит в Russian Science Citation Index – RSCI, переводная версия издания включена в международный индекс цитирования Web of Science).

Требования по оформлению представлены на странице конференции <http://www.viniti.ru/news/349-10-ya-nauchnaya-konferentsiya-posvyashchennaya-70-letiyu-viniti-ran>

Прием докладов – до 20 сентября 2022 г.

Публикация программы на сайте конференции – после 18 октября 2022 г.

**Конференция будет проходить по адресу:** г. Москва, ул. Усиевича, д. 20

**Дополнительная информация:**

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

**Контактное лицо:**

Секретарь конференции

Камнева Ирина Евгеньевна

Член оргкомитета конференции

Самсонова Анастасия Станиславовна

**Контактные телефоны:**

+7 (499) 155-45-12, +7 (499) 155-43-33

# **УВАЖАЕМЫЕ КОЛЛЕГИ!**

## **ВИНИТИ РАН предлагает Вашему вниманию Реферативный Журнал в электронной форме**

РЖ в электронной форме (ЭлРЖ) выпускается по всем разделам естественных, технических и точных наук.

Каждый номер ЭлРЖ является полным аналогом печатного номера РЖ по составу описаний документов, их оформлению и расположению. Он сопровождается оглавлением, указателями.

ЭлРЖ представляет собой информационную систему, снабженную поисковым аппаратом и позволяющую пользователю на персональном компьютере:

- читать номер РЖ, последовательно листая рефераты;
- просматривать рефераты отдельных разделов по оглавлению;
- обращаться к рефератам по указателям авторов, источников, ключевых слов;
- проводить поиск документов по словам и словосочетаниям;
- выводить текст описаний документов во внешний файл.

ЭлРЖ могут быть:

- записаны на DVD-ROM;
- передаваться через FTP-сервер (клиенту предоставляется логин и пароль с доступом к FTP-серверу ВИНТИ, с которого он скачивает заказанные журналы).

Электронные реферативные журналы можно заказать за текущий год с любого номера, а также за предыдущие годы.

**Подробную информацию Вы можете получить:**

**Адрес:** 125190, Россия, Москва, ул. Усиевича, 20, ВИНТИ РАН

**Телефон:** 8 499-152-62-11

**E-mail:** feo@viniti.ru