

**Методы определения и анализа
научно-технологических фронтиров
в деятельности студентов технических вузов России**

© Т.В. Коренькова, В.Н. Ремарчук

МГТУ им. Н.Э. Баумана, Москва, 105005, Россия

Предложена методика определения и классификации научно-технологических фронтиров применительно к деятельности студентов технических вузов России, выраженной в научных публикациях. Исследование проведено на основе программных, статистических, экспертных и «интеллектуальных» методов сбора и анализа информации. Это позволило сформировать список направлений, которые с наибольшей долей вероятности можно считать научно-технологическими фронтирами в деятельности студентов технических вузов России.

Ключевые слова: фронтиры, научно-технологические фронтиры, студенты технических вузов, научные публикации, токены, искусственный интеллект, NLP, машинное обучение, статистический анализ, экспертный анализ, индекс значимости, индекс динамичности, композитный индекс

Научно-технологические фронтиры (от англ. *frontier* — граница, рубеж, пограничье), или фронтиры научно-технологического развития (далее — фронтиры), — это наиболее популярные, но недостаточно хорошо изученные направления исследований, которые обладают большими перспективами и активно развиваются учеными и исследователями [1]. Устоявшегося в науке определения, как и точного теоретического описания, у данного термина до сих пор нет, поэтому многие аспекты вопроса в настоящее время остаются спорными. Например, определение и методы изучения фронтиров для гуманитарных и технических наук могут отличаться.

Впервые этот термин был употреблен в журнале *Nature* в 1934 г., где его описывали как некую «границу», которая «движется в область неопознанного и необъясненного» [2]. В некоторых научных сообществах и СМИ фронтиры рассматриваются как «передний край науки» или как злободневные вопросы.

Для того чтобы в дальнейшей работе обосновать выбор некоторых направлений в качестве фронтиров, необходимо привести их отличия от «мейнстримов» и других схожих понятий. Специфические особенности фронтиров описал в 1962 г. Томас Кун [3]:

- 1) значительные разногласия научного сообщества;
- 2) вопросы, на которые сложно ответить, применяя обычные методологические подходы;
- 3) использование методологий и концепций, нетипичных для соответствующей области;
- 4) неожиданные выводы, которые бросают вызов существующей парадигме;
- 5) исследования с высокой степенью неопределенности относительно вероятности их успеха.

Главный подход, согласно которому определяются фронтиры в уже существующих системах, — извлечение смыслового контента из полнотекстовых документов и метаданных с помощью программных методов и технологий искусственного интеллекта (ИИ). В качестве примера можно привести систему анализа больших данных iFORA, разработанную командой НИУ ВШЭ. К ее возможностям относится анализ различного рода документов, публикаций, рынков, определение трендов и закономерностей [4]. Кроме того, есть такие системы, как ESI, InCites, Elsevier SciVal и ScienceResearch [5].

Один из базовых алгоритмов определения фронтиров, используемый в системе iFORA, следующий. В первую очередь считается индекс значимости, который показывает актуальность тематик, их популярность и обсуждаемость в научных кругах, а также публикационную активность в настоящее время.

Далее определяется индекс динамичности, который отражает темпы прироста популярности тематики и рассчитывается на основе анализа временных рядов по различным количественным и качественным характеристикам. В итоге с помощью индексов значимости и динамичности строится композитный индекс, который является базой для выявления фронтиров [6]. Поскольку в данном вопросе сложно доверять одному лишь компьютерному анализу в связи с тонкими различиями между фронтами и прочими направлениями, после выполнения всех этапов проводится окончательная экспертная оценка, позволяющая понять, что можно считать фронтами с наибольшей долей вероятности. Необходимость экспертного анализа заключается в том, что найденные даже интеллектуальными технологиями направления могут быть не фронтами, а либо очень широкими понятиями (например, Интернет), либо «мейнстримными», популярными технологиями, возникшими в результате какого-то инфоповода. «То неcodифицированное знание, которое находится в голове у эксперта, незаменимо, и определять прорывные направления развития науки и технологий простой экстраполяцией трендов непрофессионально» [7].

Поскольку под фронтирами могут пониматься направления и технологии различного масштаба, степени охвата и популярности, их подразделяют на три уровня [6]:

1) стратегический уровень — фронтиры-концепции, они являются совокупностью конкретных технологий и не могут применяться в прикладных задачах сами по себе. В качестве примера можно привести ИИ — это актуальное и динамично развивающееся научно-технологическое направление, но в задачах напрямую используются именно технологии, входящие в понятие «искусственный интеллект». Фронтиры этого уровня достаточно легко определить с помощью различных компьютерных методов, потому что они часто упоминаются как в научных публикациях, так и в СМИ;

2) фронтиры-нарративы — во-первых, могут подразумеваться те технологии, которые составляют направления первого уровня. Например, компьютерное зрение, чат-боты — это конкретные технологии ИИ, которые непосредственно применяются в прикладных задачах. Во-вторых, это могут быть бизнес-модели, т. е. новые варианты использования уже существующих технологий. Например, экономика совместного потребления или сервисные модели. Эти фронтиры определяются сложнее и упоминаются в публикациях различного рода реже благодаря своей узкой направленности и, скорее, только в контексте фронтиров-концепций;

3) фронтиры слабых сигналов — очень узкоспециализированные фронтиры, которые развиваются в рамках конкретной задачи или направления и поэтому вряд ли знакомы человеку, не работающему в соответствующей сфере. Исходя из этих особенностей, такие фронтиры могут либо угаснуть и не достичь массового потребления, либо стать популярной и незаменимой технологией. В качестве примера можно привести 3D-дисплеи. Такие фронтиры возможно обнаружить только благодаря оценкам экспертов соответствующих областей.

Существуют высокопроизводительные и мощные системы, которые уже занимаются задачами определения и анализа фронтиров, используя методы сбора и обработки больших данных, технологии ИИ, визуализации и т. д. Поэтому цель и новизна данной работы заключается именно в решении этой задачи в контексте научной деятельности студентов технических вузов России, предложении и пошаговом описании несложной, но эффективной методики, которая опирается на приведенные выше теоретические сведения. При этом в ходе работы не только выявляются конкретные фронтиры, но и пошагово объясняются алгоритм и методы их определения с целью использования полученной методики в решении будущих задач.

Решение задачи определения фронтиров проведено применительно к деятельности студентов технических вузов России, поскольку

сегодня все больший упор делается на развитие талантов среди студентов и на привлечение их к решению серьезных прикладных задач и кейсов. Это выражается в растущем с каждым годом количестве проводимых конкурсов, олимпиад, грантов, программ стажировок и курсов для студентов. Таким образом, студенты намного раньше приобщаются к научным исследованиям, формированию своих научных интересов, реализуя ведущий принцип инженерной школы «образование через науку». Но для того чтобы поддерживать у студентов мотивацию работать и развиваться и намного чаще находить настоящие таланты, необходимо учитывать их интересы, принимать решения о продвижении определенных программ обучения и стратегий развития организаций, привлекающих к своей деятельности обучающихся.

В качестве примера можно привести развитие и актуализацию среди студентов такого направления, как ИИ. В последние несколько лет во Всероссийской студенческой олимпиаде «Я — профессионал» появились направления «искусственный интеллект» и «машинное обучение», а победителям и призерам предлагаются стажировки в компаниях по соответствующим направлениям [8]. Также все чаще появляются новости о планах внедрения программ по обучению основам ИИ в российских вузах, причем не только технических. Этот факт подтверждается рекомендациями и заявлениями Минобрнауки России: «Компетенции ИИ мы будем формировать по всему перечню специальностей и направлений и ставим себе задачу сформировать у каждого выпускника вуза цифровую культуру и цифровые компетенции» [9].

Из приведенных выше примеров можно сделать вывод, что понимание профессиональных интересов студентов и потребностей организаций очень важно для развития молодежного научного сообщества, поэтому необходимо знать, какие фронтиры сейчас активно преобладают в студенческих работах. Для получения более показательных результатов проанализированы работы обучающихся в технических вузах, так как найти научно-технологические фронтиры в них наиболее вероятно.

Предложенная методика определения фронтиров, опирающаяся на уже существующие системы и подходы, состоит из следующих этапов:

- 1) определение объекта анализа;
- 2) сбор и агрегация данных;
- 3) первичная обработка данных программными средствами;
- 4) формирование датасета из признаков и токенов (под токеном в данной статье подразумевается минимальная лингвистическая единица текста, обладающая смыслом), а также назначение весов признаков [10];

- 5) построение индекса значимости по каждому токену [6], равного взвешенной сумме значения частоты токена по каждому при знаку;
- 6) кластеризация полученных данных методами машинного обучения для отбора дальнейших направлений [11];
- 7) построение «облаков слов» для каждого выделенного направления, чтобы определить более узкие тематики на основе их близости к ключевым словам в тексте [12];
- 8) анализ динамики развития отобранных направлений на основе временных рядов публикационной активности;
- 9) построение индекса динамичности [6], равного отношению количества публикаций за последние 3 года к количеству публикаций за предыдущие 3 года;
- 10) построение композитного индекса [6], равного произведению индексов значимости и динамичности;
- 11) проведение более детального экспертного анализа выделенных направлений с целью определения вероятности их причисления к фронтирам и разбиению по уровням [7].

В первую очередь при решении поставленной задачи необходимо было выбрать объекты анализа. В качестве них были взяты научные работы студентов, опубликованные в журналах ведущих технических вузов России.

Следуя методике, описанной выше, сначала потребовалось получить индекс значимости тематик, показывающий их актуальность. Поэтому для анализа были выбраны публикации, выложенные с 2019 по 2021 г. Под этот фильтр подошло 1252 доступные к просмотру научные публикации журналов следующих технических вузов России: МФТИ, МГТУ им. Н.Э. Баумана, СПбПУ, НИЯУ МИФИ, МИРЭА, МАИ, МЭИ, НИУ ИТМО, МГТУ «СТАНКИН». Источником данных статей была выбрана научная электронная библиотека eLIBRARY.RU благодаря своей надежности, доступности и удобству анализа [13]. Основные используемые в этой задаче методы: feature engineering, статистический анализ (анализ временных рядов, кластеризация), NLP (нейролингвистическое программирование), экспертный анализ, визуализация (построение диаграмм, временных рядов).

После сбора всех статей была проведена предварительная обработка текстов на языке программирования Python с использованием библиотек для анализа и обработки данных: Pandas, NLTK, Regex. Обработка включала в себя удаление лишних символов и стоп-слов. Далее для очищенного текста были построены частотные словари для токенов [14] (слов, биграмм и триграмм), при этом для каждой части значимой статьи был построен свой частотный словарь и даны разные веса важности, распределенные так, чтобы их сумма была равна единице:

- ключевые слова. Это те токены, которые определяют основной смысл статьи и назначаются самими авторами, поэтому им был дан самый высокий вес — 0,28. В ключевых словах с наибольшей долей вероятности можно как раз обнаружить фронтиры первого уровня;

- название публикации. Им был дан меньший вес — 0,25, поскольку названия должны быть более компактными, чем список ключевых слов, поэтому они могут не отображать всей сути публикации;

- ключевые слова публикаций, цитируемые в иных изданиях. Такая метрика была выбрана для того, чтобы получить более полное представление об анализируемых публикациях, заметить, возможно, те направления, которых касаются косвенно или описывают другими словами. Вес данного признака — 0,19;

- аннотация. В аннотациях текст составляют таким образом, чтобы была понятна суть статьи без лишних подробностей. Однако в этом случае в список могут попасть лишние токены, не имеющие значения в рамках решаемой задачи, поэтому им был дан вес 0,17;

- список литературы. Довольно значимая часть для анализа, однако может содержать в себе очень много лишней информации, которую сложно отфильтровать, например, названия издательств и журналов, инициалы авторов, ссылки и прочие элементы. Вес данного признака — 0,08;

- полный текст публикации. В нем полезные с точки зрения решаемой задачи токены могут быть смешаны со всем остальным информационным шумом, который сложно распознать методами NLP, и, соответственно, этому признаку был дан самый минимальный вес — 0,03.

После того как токены получили свое значение частоты по каждому признаку, были добавлены еще две характеристики:

- взвешенная сумма по каждому признаку токена — индекс значимости каждой тематики;

- количество признаков, в котором значение токена было не равно нулю — для того, чтобы понимать, насколько и где тематика обсуждается как самим автором публикации, так и авторами из списка литературы.

Фрагмент таблицы, полученной после проведенных этапов анализа, отсортированной в порядке убывания частоты в тексте, приведен на рис. 1.

Далее были получены кластеры токенов, объединенные на основе схожести значений признаков. Для этого использовалось «обучение без учителя» — способ машинного обучения, в котором модель решает прогнозные задачи без вмешательства со стороны человека, т. е. данные не размечены на группы: выводы о том, являются ли они фронтирами или нет, делает алгоритм [12]. Для решения этой задачи был выбран алгоритм кластеризации данных, основанный на плотно-

сти точек-признаков — DBSCAN. Получившиеся кластеры данных были в основном объединены по признаку «количество ненулевых признаков», так как он является дискретным и имеет небольшую вариативность.

Токен	Частота в тексте	Частота в аннотации	Частота в списке литературы	Частота в ключевых словах	Частота в ключевых словах цитируемых	Частота в названии	Взвешенная сумма (индекс значимости)	Количество ненулевых признаков
момент	1007	23	85	0	0	1	35,14	4
моделирование	1005	39	201	15	11	60	72,47	6
модельной	787	50	149	0	0	8	51,28	4
частоты	743	50	62	0	0	4	35,14	4
оптимизации	742	82	159	0	0	12	50,53	4
визуализации	728	41	40	0	0	1	30,64	4
рассмотрим	728	0	36	0	0	0	24,89	2
установки	690	46	41	0	0	4	31,25	4
плотности	688	26	35	0	0	1	28,28	4
молельной	655	42	68	0	0	9	33,11	4
аппаратов	654	60	227	0	0	26	53,58	4
цифровой	648	53	85	0	0	17	38,17	4
плоскости	640	34	41	0	0	4	27,86	4
available at	640	0	24	0	0	0	19,78	2
аппарата	639	77	121	0	0	24	46,67	4
mathematical	637	2	35	0	0	0	20,54	3
ноды	625	20	43	0	0	2	24,78	4
множество	625	12	76	0	0	0	25,69	3
economic	602	0	37	0	0	1	20,06	3
затрат	601	21	41	0	0	3	24,36	4
the article	600	5	0	0	0	0	17,50	2
она	599	16	32	0	0	0	21,97	3
создание	598	30	88	0	0	2	29,44	4
лоды	598	30	44	0	0	4	28,28	4
результат	598	13	42	0	0	0	22,28	3
производство	597	19	77	4	6	2	28,94	6
особенности	597	88	93	0	0	22	44,50	4
услуг	597	23	47	0	0	2	24,83	4

Рис. 1. Фрагмент таблицы признаков токенов

Для того чтобы получить лучшее представление, данные были также отсортированы сначала по количеству ненулевых признаков, затем по взвешенной сумме и по всем остальным признакам в порядке убывания. В итоге, после ручной проверки токенов, находящихся в «топе» в результате кластеризации и сортировки, для дальнейшего анализа были взяты следующие крупные тематики с соответствующими индексами значимости:

- цифровизация — 102,67;
- беспилотные летающие аппараты (БПЛА) — 97,53;
- моделирование — 94,8;
- композитные материалы — 29,44;
- космический аппарат — 26,55;
- аддитивные технологии — 24;
- ИИ — 21,44;
- система экологического менеджмента — 16,22;
- интернет вещей — 12,14.

Для определения более узких, схожих тематик и «кандидатов» во фронты потребовалось построить «облака слов» на основе частотности токенов, находящихся в одном контексте с выделенными выше направлениями [5]. Была выполнена вторичная обработка данных средствами NLP и машинного обучения. Для всех найденных токенов и каждого направления провели лемматизацию и фильтрацию по частям речи таким образом, чтобы была выше вероятность

попадания более осмысленных токенов в «облако слов». Примеры составленных «облаков слов» из направлений, выделенных выше, приведены на рис. 2.

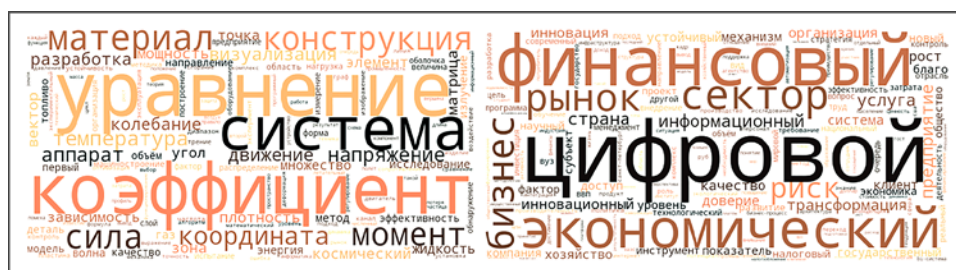


Рис. 2. Примеры «облаков слов» для тематик «моделирование» и «цифровизация»

Построенные для всех направлений «облака слов» помогли пополнить список направлений для последующих этапов анализа. Между некоторыми из них сразу удалось установить отношения «фронтиры-концепции — фронтиры-нарративы», например, в случае с ИИ и одной из его технологий — нейронными сетями. Некоторые понятия связаны между собой схожестью тематик и не обязательно связаны отношением «целое — часть». Окончательный список для последующих этапов анализа представлен в виде диаграммы связей с распределением по уровням (рис. 3).

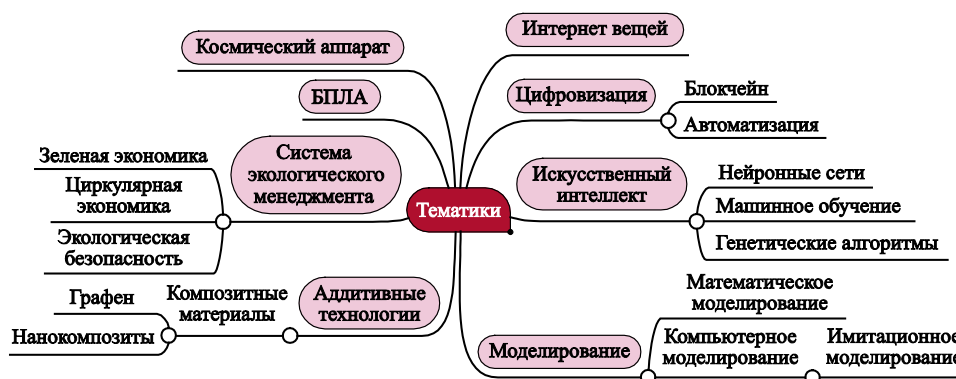


Рис. 3. Диаграмма связей для тематик

Следующим шагом в реализации поставленной задачи было построение индекса динамичности, опирающегося на публикационную активность выделенных выше тематик [10]. Для каждой из них были построены временные ряды частотности научных публикаций с 2010 по 2021 г., в которых эти тематики были отмечены в качестве ключевых

слов как на русском языке, так и в переводе [13]. Гипотеза такова, что фронтами можно считать те направления, публикационная активность которых резко возросла в последние несколько лет. Поэтому для каждой тематики были также рассчитаны индексы динамичности, равные отношению количества публикаций за 2019–2021 гг. к количеству публикаций за 2016–2018 гг. В итоге были рассчитаны композитные индексы, равные произведению индексов значимости и динамичности. Получившиеся графики и композитные индексы приведены на рис. 4.

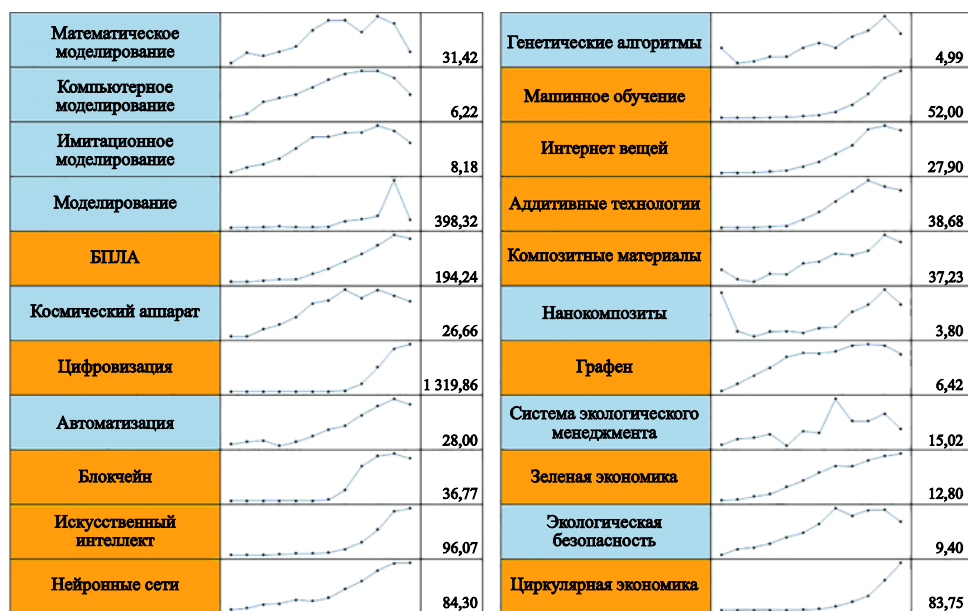


Рис. 4. Временные ряды публикационной активности тематик с композитными индексами

На основе анализа получившихся графиков, композитных индексов и ручной оценки были выбраны тематики, которые изучались далее на предмет их принадлежности к фронтам уже с помощью экспертного анализа (выделены на рис. 4 желтым цветом). При этом категории моделирования и автоматизации не были выбраны для дальнейшего анализа, так как эти понятия очень широкие, а интерес к остальным исключенным направлениям в последние годы начинает падать по сравнению с другими временными промежутками. Для этого были вручную проанализированы новости, программы развития, конкурсы по отобранным тематикам, поисковые запросы и прочие факторы. Можно увидеть корреляцию некоторых выделенных направлений с теми программами развития, которые обозначены в программе поддержки университетов и кадрового обеспечения

приоритетных направлений науки «Приоритет 2030», что свидетельствует о жизнеспособности описанной методики определения фронтиров [15]. В частности, в большей мере это касается информационных технологий (ИТ), цифровизации и «зеленой» экономики. Взаимосвязь с выделенными тематиками можно обнаружить также в списке 16 высокотехнологичных направлений, выделенных Правительством РФ и проанализированных в «Белой книге» технологий [16]. Приведем краткий обзор по каждой из отобранных тематик:

- цифровизация (цифровая трансформация) — это повсеместное внедрение цифровых технологий в разные сферы жизни: промышленность, экономику, образование, культуру, сферы обслуживания и т. п. Связано это явление со стремительным развитием ИТ и информационно-коммуникативных технологий. Важность и необходимость данного направления подтверждает тот факт, что в России цифровизация была выбрана в качестве национальной цели развития до 2030 г. Эта программа уже включает в себя увеличение бюджетных мест и дополнительных программ по ИТ-специальностям, повышение квалификаций преподавателей и руководителей, цифровизацию персональных данных и прочие важные нововведения [17]. Все эти факторы подтверждают, что цифровизацию можно считать фронтиром, предположительно, первого уровня;

- блокчейн — сквозная технология *цифровой экономики*, механизм децентрализованного подтверждения операций, основанный на современных криптографических алгоритмах. Поскольку с развитием цифровизации вопрос безопасности становится все более актуальным, технология блокчейна уже активно внедряется в основы крупных платформ, проектов для обеспечения безопасности функционирования бизнеса, государства и общества [18]. Данная технология уже прошла волну мейнстрима, более прочно закрепилась на рынке, и поэтому ее можно считать фронтиром второго уровня;

- ИИ — комплекс технологий, позволяющих имитировать работу человеческого мозга. Данное направление достигло такого развития, что теперь является федеральным проектом РФ и входит в состав Национальной программы «Цифровая экономика Российской Федерации». Его поддержку подтверждают создание крупных исследовательских центров по ИИ, поддержка ИИ-стартапов, введение дополнительных программ обучения, обеспечение доступа к датасетам [17]. Ежегодно в РФ также организуются международные конференции, проводятся хакатоны, олимпиады и прочие конкурсы по ИИ, что позволяет причислить его к фронтирам первого уровня [19];

- машинное обучение и нейронные сети — одни из самых передовых технологий ИИ, которые применяются в крупнейших российских компаниях и их разработках [20]. На «цикле хайпа» можно увидеть,

что эти технологии уже прошли пик завышенных ожиданий и прогнозируется, что они должны достичь плато эффективности в течение 2–5 лет [21]. Прикладной характер и устойчивость описанных технологий позволяют отнести их к фронтам второго уровня;

- интернет вещей (IoT) — это концепция сети передачи данных между устройствами. В последние несколько лет данная тематика активно развивается в РФ в таких областях, как электроэнергетика, здравоохранение, сельское хозяйство, «умный дом» и т. д. Многочисленные исследования прогнозируют увеличение объема этого рынка в РФ и в дальнейшем будущем [22], поэтому IoT можно считать фронтом первого уровня;

- циркулярная экономика — это экономика, в которой на инновационной основе обеспечивается возобновление и воспроизводство ресурсов, формируются инструменты их циклического вовлечения в экономическую систему. Неразрывно стоит рядом зеленая экономика, принципы которой в последнее время набирают всю большую востребованность на государственном уровне и входят в привычный уклад жизни обычных граждан [23]. Например, в России разработан национальный проект «Экология», в котором делается активный упор на утилизацию отходов [24]. Эти направления можно считать фронтами второго уровня — бизнес-моделью, поскольку здесь речь идет о новом варианте использования уже существующих технологий;

- аддитивные технологии — технологии послойного наращивания и синтеза объектов. Самый яркий пример таких технологий — 3D-печать, которая активно развивается в РФ, а ее продукты поставляются крупным промышленным организациям [25]. Эта тематика является фронтом первого уровня;

- графен — двумерная структура, в которой атомы углерода встроены в вершины правильных шестиугольников. Изучение этой структуры до сих пор продолжается и развивается по всему миру: открываются новые свойства, способы применения, связанные с созданием композитных материалов на основе графена, в том числе в РФ [26]. Можно считать эту тематику фронтом третьего уровня, потому что она очень узкоспециализированная и еще не перешла уровень «слабых сигналов»;

- БПЛА — летательный аппарат без экипажа на борту. В последние годы в РФ активно развивается производство БПЛА разных типов, особенно военных, строительство заводов и предприятий [27], а также разработка программ по этому направлению в авиационных вузах. Данное направление можно считать фронтом либо второго уровня, либо третьего.

В результате проведенного теоретического и практического исследования была предложена и описана авторская методика опреде-

ления фронтиров в контексте деятельности студентов технических вузов России. Для каждой значимой тематики, отобранной с точки зрения статистических, интеллектуальных и экспертных методов анализа, был построен композитный индекс и составлен и проанализирован окончательный список выявленных фронтиров, распределенных по уровням. Преобладающее число фронтиров было найдено в следующих тематиках: ИТ-направления, экологическая сфера и аддитивные технологии.

Релевантность полученных фронтиров подтверждает факт корреляции с направлениями, обозначенными в программе «Приоритет 2030» и в «Белой книге технологий».

Итоги проделанной работы можно использовать в виде источника направлений, которые следует развивать в деятельности студентов, а также применять описанную методику для определения фронтиров в дальнейшем.

ЛИТЕРАТУРА

- [1] Фронтиры науки. *Цифровой прорыв*. URL: <https://leadersofdigital.ru/event/63008/case/704366> (дата обращения 25.01.2022).
- [2] The frontiers of science. *Nature*, 1934, no. 134, pp. 670, 671.
- [3] Кун Т. *Структура научных революций*. С вводной статьей и дополнениями, Москва, Прогресс, 1977, 300 с.
- [4] iFORA: трехмерный взгляд на растущие области науки и технологий. *Институт статистических исследований и экономики знаний*. URL: <https://issek.hse.ru/news/254274661.html> (дата обращения 25.01.2022).
- [5] Девяткин Д.А. Метод выделения направлений научных исследований на основе анализа больших массивов научных публикаций. *Выявление приоритетных научных направлений: междисциплинарный подход*. Москва, ИМЭМО, 2016, с. 110–113.
- [6] Семинар «Фронтир научно-технологического развития: определение и поиск». *YouTube*. URL: <http://surl.li/bgcjw> (дата обращения 25.01.2022).
- [7] Быкова Н. Научную экспертизу доведут до автоматизма. *Российская академия наук*. URL: <http://www.ras.ru/news/shownews.aspx?id=6e44dda1-d970-4c43-b6e9-71bc4df7f477> (дата обращения 25.01.2022).
- [8] Олимпиада «Я — профессионал». URL: <https://yandex.ru/profi/> (дата обращения 26.01.2022).
- [9] Минобрнауки внедрит в программы вузов образовательный модуль в области ИИ. *TACC*. URL: <https://tass.ru/obschestvo/12796581> (дата обращения 25.01.2022).
- [10] Как мы фронтиры науки с помощью тематического моделирования искали. *Хабр*. URL: <https://habr.com/ru/company/sberbank/blog/575726/> (дата обращения 26.01.2022).
- [11] Храмоин И.В. Анализ динамики приоритетных направлений исследований методом кластеризации научных коллективов в предметнофреймовом пространстве. *Выявление приоритетных научных направлений: междисциплинарный подход*. Москва, ИМЭМО, 2016, с. 114–119.
- [12] Смирнов И.В. Анализ социально-экономических данных. *Выявление приоритетных научных направлений: междисциплинарный подход*. Москва, ИМЭМО, 2016, с. 98–102.

- [13] Научная электронная библиотека eLIBRARY.RU. URL: <https://www.elibrary.ru/> (дата обращения 26.01.2022).
- [14] Швец А.В. Метод выявления научных коллективов на основе анализа больших массивов научно-технической информации. *Выявление приоритетных научных направлений: междисциплинарный подход*. Москва, ИМЭМО, 2016, с. 103–109.
- [15] *Приоритет 2030*. URL: <https://priority2030.ru/> (дата обращения 28.01.2022).
- [16] Соколова М.Ю., Эйделькин Л.Д., ред. *Развитие отдельных высокотехнологических направлений. Белая книга*. Москва, Национальный исследовательский университет «Высшая школа экономики», 2022, 188 с.
- [17] *Цифровая экономика РФ*. URL: <https://digital.gov.ru/ru/activity/directions/858/> (дата обращения 27.01.2022).
- [18] «Признание блокчейна уже наступило». Как технология развивается в России. *РБК*. URL: <https://www.rbc.ru/crypto/news/5feadd249a7947b4da846630> (дата обращения 27.01.2022).
- [19] Эволюция мозга: куда ведет искусственный интеллект. Как изменится мир в ближайшем будущем. *AI Journey*. URL: <https://ai-journey.ru/news/21> (дата обращения 27.01.2022).
- [20] Сбер создал первую мультимодальную нейросеть ruDALL-E, которая генерирует картинки по описанию на русском языке. *СберБанк*. URL: https://www.sberbank.ru/ru/press_center/all/article?newsID=3e84a9bc-973e-4d33-9de5-1c56851e90d&blockID=1303®ionID=77&lang=ru&type=NEWS (дата обращения 28.01.2022).
- [21] The 4 Trends That Prevail on the Gartner Hype Cycle for AI. *Gartner*. URL: <https://www.gartner.com/en/articles/the-4-trends-that-prevail-on-the-gartner-hype-cycle-for-ai-2021> (дата обращения 28.01.2022).
- [22] «Интернет вещей» (IoT) в России: технология будущего, доступная уже сейчас. *PwC*. URL: <https://www.pwc.ru/ru/publications/IoT.html> (дата обращения 28.01.2022).
- [23] Россия взяла курс на развитие «зеленой» экономики. *Российская газета*. URL: <https://rg.ru/2020/07/21/rossiia-vziala-kurs-na-razvitie-zelenoj-ekonomiki.html> (дата обращения 28.01.2022).
- [24] Паспорт Национального проекта «Экология». *Минприроды России*. URL: https://www.mnr.gov.ru/activity/directions/natsionalnyy_proekt_ekologiya/ (дата обращения 28.01.2022).
- [25] В России разработали революционную технологию 3D-печати. *Lenta.ru*. URL: <https://lenta.ru/articles/2019/05/25/aniz/> (дата обращения 28.01.2022).
- [26] Россия запустила крупнейшее в мире производство графеновых нанотрубок. *РИА новости*. URL: <https://ria.ru/20200211/1564535545.html> (дата обращения 28.01.2022).
- [27] В России построили первый завод по производству ударных беспилотников. *РИА новости*. URL: <https://ria.ru/20211222/zavod-1764908105.html> (дата обращения 28.01.2022).

Статья поступила в редакцию 20.02.2022

Ссылку на эту статью просим оформлять следующим образом:

Коренькова Т.В., Ремарчук В.Н. Методы определения и анализа научно-технологических фронтов в деятельности студентов технических вузов России. *Гуманитарный вестник*, 2022, вып. 1. <http://dx.doi.org/10.18698/2306-8477-2022-1-765>

Коренькова Татьяна Вячеславовна — студентка кафедры «Информационная аналитика и политические технологии» МГТУ им. Н.Э. Баумана.
e-mail: korenkovatv@student.bmstu.ru

Ремарчук Валерий Николаевич — д-р филос. наук, профессор, декан факультета «Социальные и гуманитарные науки», заведующий кафедрой «Информационная аналитика и политические технологии» МГТУ им. Н.Э. Баумана.
e-mail: dekan.fsgn@bmstu.ru

Methods for determining and analyzing scientific and technological frontiers in the activities of students of technical universities in Russia

© T.V. Korenkova, V.N. Remarchuk

Bauman Moscow State Technical University, Moscow, 105005, Russia

The paper introduces a methodology for determining and classifying scientific and technological frontiers related to the activities of students of technical universities in Russia, expressed in scientific publications. The study is based on software, statistical, expert and AI methods of collecting and analyzing information. This made it possible to form a list of areas that most likely can be considered scientific and technological frontiers in the activities of students of technical universities in Russia.

Keywords: *frontiers, scientific and technological frontiers, students of technical universities, scientific publications, tokens, artificial intelligence, natural language processing, machine learning, statistical analysis, expert analysis, significance index, dynamic index, composite index*

REFERENCES

- [1] Frontiry nauki [Frontiers of science]. *Tsifrovoy proryv* [Digital breakthrough]. Available at: <https://leadersofdigital.ru/event/63008/case/704366> (accessed January 25, 2022).
- [2] The frontiers of science. *Nature*, 1934, pp. 670–671.
- [3] Kuhn T.S. *The Structure of Scientific Revolutions*. University of Chicago Press; 4th ed., 2012, 264 p. [In Russ.: Kuhn T.S. *Struktura nauchnykh revolyutsiy. S vvodnoy statey i dopolneniyami*. Moscow, Progress Publ., 1977, 300 p.].
- [4] iFORA: trekhmerny vzglyad na rastuschie oblasti nauki i tekhnologii [iFORA: a 3D view of growing fields of science and technology]. *Institut statisticheskikh issledovaniy i ekonomiki znaniy* [Institute for Statistical Studies and Economics of Knowledge]. Available at: <https://issek.hse.ru/news/254274661.html> (accessed January 25, 2022).
- [5] Devyatkin D.A. Metod vydeleniya napravleniy nauchnykh issledovaniy na osnove analiza bolshikh massivov nauchnykh publikatsiy [The method of identifying areas of scientific research based on the analysis of large arrays of scientific publications]. *Vyyavlenie prioritetykh nauchnykh napravleniy: mezhdistsiplinarny podkhod* [Identification of priority research areas: an interdisciplinary approach]. Moscow, IMEMO Publ., 2016, pp. 110–113.
- [6] Seminar «Frontir nauchno-tekhnologicheskogo razvitiya: opredelenie i poisk» [Seminar “Frontier of scientific and technological development: definition and search”]. *YouTube*. Available at: <http://surl.li/bgcjw> (accessed January 25, 2022).
- [7] Bykova N. Nauchnuyu ekspertizu dovedut do avtomatizma [Scientific expertise will be brought to automatism]. *Rossiyskaya akademiya nauk* [Russian Academy of Sciences]. Available at: <http://www.ras.ru/news/shownews.aspx?id=6e44dda1-d970-4c43-b6e9-71bc4df7f477> (accessed January 26, 2022).
- [8] *Olimpiada «Ya-Professional»* [Olympiad “I am a professional”]. Available at: <https://yandex.ru/profi/> (accessed January 26, 2022).
- [9] Minobrnauki vnedrit v programmy vuzov obrazovatelny modul v oblasti II [The Ministry of Education and Science will introduce an educational module in the field of AI into the programs of universities]. *TASS* [TASS. Russian news

- agency]. Available at: <https://tass.ru/obschestvo/12796581> (accessed January 27, 2022).
- [10] Kak my frontiry nauki s pomoschy tematicheskogo modelirovaniya iskali [How we searched the frontiers of science using topic modeling]. *Habr*. Available at: <https://habr.com/ru/company/sberbank/blog/575726/> (accessed January 27, 2022).
- [11] Khramoin I.V. Analiz dinamiki prioritnykh napravleniy issledovaniy metodom klasterizatsii nauchnykh kollektivov v predmetnofreymovom prostranstve [Analysis of the dynamics of priority areas of research by the method of clustering research teams in the subject-frame space]. *Vyyavlenie prioritnykh nauchnykh napravleniy: mezhdistsiplinary podkhod* [Identification of priority research areas: an interdisciplinary approach]. Moscow, IMEMO Publ., 2016, pp. 114–119.
- [12] Smirnov I.V. Analiz sotsialno-ekonomicheskikh dannykh [Analysis of socio-economic data]. *Vyyavlenie prioritnykh nauchnykh napravleniy: mezhdistsiplinary podkhod* [Identification of priority research areas: an interdisciplinary approach]. Moscow, IMEMO Publ., 2016, pp. 98-102.
- [13] *Nauchnaya elektronnyaya biblioteka eLIBRARY.RU* [Scientific electronic library eLIBRARY.RU]. Available at: <https://www.elibrary.ru/> (accessed January 29, 2022).
- [14] Shvets A.V. Metod vyyavleniya nauchnykh kollektivov na osnove analiza bolshikh massivov nauchno-tehnicheskoy informatsii [Method for identifying research teams based on the analysis of large arrays of scientific and technical information]. *Vyyavlenie prioritnykh nauchnykh napravleniy: mezhdistsiplinary podkhod* [Identification of priority research areas: an interdisciplinary approach]. Moscow, IMEMO Publ., 2016, pp. 103-109.
- [15] *Prioritet 2030* [Priority 2030]. Available at: <https://priority2030.ru/> (accessed January 29, 2022).
- [16] Sokolova M.Yu., Eydelkind L.D., ed. *Razvitie otdelnykh vysokotekhnologichnykh napravleniy. Belaya kniga* [Development of individual high-tech areas. The white book]. Moscow, HSE Publ., 2022, 188 p.
- [17] *Tsifrovaya ekonomika RF* [Digital economy of the Russian Federation]. Available at: <https://digital.gov.ru/ru/activity/directions/858/> (accessed January 30, 2022).
- [18] Priznanie blokcheyna uzhe nastupilo. Kak tekhnologiya razvivaetsya v Rossii [Blockchain recognition has already arrived. How technology is developing in Russia]. *RBK* [RBC]. Available at: <https://www.rbc.ru/crypto/news/5feadd249a7947b4da846630> (accessed January 30, 2022).
- [19] Evolyutsiya mozga: kuda vedet iskusstvenny intellekt. Kak izmenitsya mir v blizhayshem buduschem [Brain evolution: where artificial intelligence leads. How the world will change in the near future]. *AI Journey*. Available at: <https://ai-journey.ru/news/21> (accessed January 30, 2022).
- [20] Sber sozdal pervuyu multimodalnuyu neyroset ruDALL-E, kotoraya generiruet kartinki po opisaniyu na russkom yazyke [Sber created the first multimodal neural network ruDALL-E, which generates pictures according to the description in Russian]. *Sberbank*. Available at: https://www.sberbank.ru/ru/press_center/all/article?newsID=3e84a9bc-973e-4d33-9de5-91c56851e90d&blockID=1303@ionID=77&lang=ru&type=NEWS (accessed January 31, 2022).
- [21] The 4 Trends That Prevail on the Gartner Hype Cycle for AI. *Gartner*. Available at: <https://www.gartner.com/en/articles/the-4-trends-that-prevail-on-the-gartner-hype-cycle-for-ai-2021> (accessed January 31, 2022).

- [22] Internet veschey (IoT) v Rossii: tekhnologiya buduschego, dostupnaya uzhe seychas [Internet of Things (IoT) in Russia: future technology available now]. *PwC*. Available at: <https://www.pwc.ru/ru/publications/IoT.html> (accessed January 31, 2022).
- [23] Rossiya vzyala kurs na razvitie «zelenoy» ekonomiki [Russia has set a course for the development of a “green” economy]. *Rossiyskaya gazeta* [RG.ru]. Available at: <https://rg.ru/2020/07/21/rossiia-vziala-kurs-na-razvitie-zelenoj-ekonomiki.html> (accessed January 31, 2022).
- [24] Pasport Natsionalnogo proekta «Ekologiya» [Passport of the National project “Ecology”]. *Minprirody Rossii* [Ministry of Natural Resources and Environment of the Russian Federation]. Available at: https://www.mnr.gov.ru/activity/directions/natsionalnyy_proekt_ekologiya/ (accessed January 31, 2022).
- [25] V Rossii razrabotali revolyutsionnyu tekhnologiyu 3D-pechati [Revolutionary 3D printing technology was developed in Russia]. *Lenta.ru*. Available at: <https://lenta.ru/articles/2019/05/25/aniz/> (accessed January 31, 2022).
- [26] Rossiya zapustila krupneyshee v mire proizvodstvo grafenovykh nanotrubok [Russia launched the world's largest production of graphene nanotubes]. *RIA novosti* [RIA news]. Available at: <https://ria.ru/20200211/1564535545.html> (accessed January 31, 2022).
- [27] V Rossii postroili pervy zavod po proizvodstvu udarnykh bespilotnikov [The first plant for the production of strike drones was built in Russia]. *RIA novosti* [RIA news]. Available at: <https://ria.ru/20211222/zavod-1764908105.html> (accessed January 31, 2022).

Korenkova T.V., student, Department of Information Analytics and Political Technologies, Bauman Moscow State Technical University.
e-mail: korenkovatv@student.bmstu.ru

Remarchuk V.N., Dr. Sc. (Philos.), Professor, Dean of the Faculty of Social and Humanitarian Sciences, Head of the Department of Information Analytics and Political Technologies, Bauman Moscow State Technical University. e-mail: dekan.fsgn@bmstu.ru