

Научные заделы России по сквозным технологиям цифровой экономики



БАКАРОВ Амир Асхатович
Программист, Федеральный исследовательский центр «Информатика и управление» Российской академии наук

Аннотация

В работе предложен подход к анализу научных заделов России по сквозным технологиям цифровой экономики. Подход апробирован на примере сквозных технологий применительно к приоритетным отраслям экономики и предметным областям, которые определены программой «Цифровая экономика Российской Федерации». Выявлено существенное различие между Россией и миром в структуре и количестве научных заделов. Проведен детальный анализ публикационной активности по направлению «Нейротехнологии и искусственный интеллект», сделаны выводы относительно дальнейших направлений исследований.

Ключевые слова:

цифровая экономика, цифровая технология, сквозная технология, научный задел, наукометрический анализ.



ДЕВЯТКИН Дмитрий Алексеевич
Главный специалист, Федеральный исследовательский центр «Информатика и управление» Российской академии наук

1. Введение

В современном мире происходит революционная трансформация традиционных видов экономической деятельности под натиском цифровых технологий. Успешные примеры такой трансформации приводят к экономическому росту, повышению производительности труда, улучшению качества предоставляемых услуг как в отдельных предприятиях, так и в масштабах целых отраслей экономики [1]. Достижение экономических и социальных эффектов невозможно без целенаправленной государственной политики, направленной на поддержку развития и использования цифровых технологий в отраслях экономики [2]. При реализации государственной политики приходится решать задачу выбора приоритетов, когда из множества отраслей экономики и обширного перечня цифровых технологий необходимо выбрать те, которые принесут максимальный эффект.

В 2017 году Правительством Российской Федерации была разработана и утверждена программа «Цифровая экономика Российской Федерации», которая направлена на создание экосистемы цифровой экономики Российской Федерации, в которой данные в цифровой форме являются ключевым фактором производства во всех сферах социально-экономической деятельности [3]. Одно из направлений реализации программы связано с формированием исследовательских компетенций и технологических заделов для



ЕРШОВА Татьяна Викторовна
Кандидат экономических наук, директор Национального центра цифровой экономики, Московский государственный университет имени М.В. Ломоносова



ТИХОМИРОВ Илья Александрович
Кандидат технических наук, доцент, Федеральный исследовательский центр «Информатика и управление» Российской академии наук



ХОХЛОВ Юрий Евгеньевич
Кандидат физико-математических наук, доцент, академик Российской инженерной академии, председатель Совета директоров Института развития информационного общества, заведующий базовой кафедрой цифровой экономики Института развития информационного общества Российского экономического университета им. Г.В. Плеханова

цифровой экономики путем развития так называемых сквозных технологий, которые используются в различных предметных областях и отраслях экономики.

В качестве приоритетных сквозных технологий в программе «Цифровая экономика Российской Федерации» обозначены:

- большие данные;
- нейротехнологии и искусственный интеллект;
- системы распределенного реестра;
- квантовые технологии;
- новые производственные технологии;
- промышленный интернет;
- сенсорика и компоненты робототехники;
- технологии беспроводной связи;
- технологии виртуальной и дополненной реальностей.

Предполагается, что развитие именно этих технологий способно обеспечить отечественным предприятиям высокий рост производительности и конкурентоспособности на глобальном мировом рынке.

В конце 2017 года был утвержден план мероприятий (дорожная карта) по направлению «Формирование исследовательских компетенций и технологических заделов» программы «Цифровая экономика Российской Федерации» [4]. Несколько мероприятий данного плана направлены как на выявление потребностей конкретных отраслей экономики в сквозных технологиях, так и на наличие в стране соответствующих научно-технологических заделов.

Научный задел представляет собой результат фундаментальных научных исследований (новые знания о явлениях, эффектах, законах, закономерностях и т.п.), напрямую не связанный с существующими или перспективными артефактами, техническими средствами и технологиями [5].

В последние годы российскими учеными было проведено несколько исследований, посвященных методике анализа научно-технологических и научно-технических заделов, прежде всего в оборонной промышленности (см. [6], [7]). Параллельно проводился анализ научных и научно-технологических заделов, в первую очередь с точки зрения публикационной и патентной активности. В 2016 году были разработаны подходы и инструменты анализа научно-технологических заделов предприятий [8] путем анализа как публикационного, так и патентного ландшафта [9–11]. Отметим также исследование центров компетенции Российской Федерации по направлению «искусственный интеллект», которое позволило выявить ведущие научные организации и исследователей, работающих в этой предметной области [12].

В то же время общий анализ заделов по сквозным технологиям в разрезе отраслей экономики в стране и в сравнении с мировыми трендами еще не проводился.

Такой анализ позволил бы основным заинтересованным сторонам (власти, бизнесу, научно-образовательному сообществу) понять, каким сквозным технологиям стоит уделить особое внимание, а какие являются узкоспециализированными. Это также позволит директивным органам определять направление вложений ресурсов на развитие сквозных технологий и их использование в конкретных отраслях экономики, что будет способствовать росту конкурентоспособности российских предприятий на высокотехнологичных рынках.

Целью настоящей работы является разработка методологии и проведение сравнительного анализа научных заделов Российской Федерации по сквозным технологиям в разрезе отраслей экономики. Апробация методологии проведена на примере мировых и отечественных заделов по развитию и использованию сквозных технологий. Подход, предложенный в статье, может быть в дальнейшем распространен на другие технологии, отрасли экономики или предметные области.

2. Методология анализа научных заделов по сквозным технологиям цифровой экономики

В настоящем исследовании будем рассматривать научные заделы с точки зрения объема и динамики публикационной активности, отраженной в зарубежных наукометрических базах данных. Считается, что именно эта информация отражает реальное положение дел в России и в мире в части научных заделов [8].

2.1. Построение ключевой лексики

В предлагаемой методологии необходимо сформировать списки ключевой лексики, характеризующие сквозные технологии. Списки ключевой лексики формировались в автоматизированном режиме с последующей экспертной обработкой специалистами по сквозным технологиям. Для извлечения ключевой лексики были использованы технологии анализа больших данных и автоматической обработки текстов. На первом шаге в коллекциях российских и зарубежных научных публикаций проводился поиск по запросам на основе ключевых слов, характеризующих названия сквозных технологий. Далее при помощи алгоритма поиска похожих документов [13] перечень найденных документов расширялся. На втором шаге из текстов документов поисковой выдачи, полученной на первом шаге, автоматически строилась ключевая лексика [14]. Итоговый список ключевой лексики верифицировался экспертами.

Приведем в качестве примера списки ключевой лексики, построенные по описанной выше схеме для сквозной технологии «Нейротехнологии и искусственный интеллект» (Таблица 1).

Таб. 1. Ключевая лексика на английском языке для сквозной технологии «Нейротехнологии и искусственный интеллект».

Ключевая лексика на английском языке
neurotechnologies, artificial intelligence, models of cognitive processes, knowledge representation, representation of knowledge gaps, statistical learning, parametric learning, inductive inference, learning by examples, training by analogy, learning as discovery, expert system, consulting expert systems, diagnostic expert systems, pattern recognition, speech recognition, image processing, heuristic methods, discourse models, models of cognitive psychology, turing test, intelligent system, applied systems of artificial intelligence, natural language processing, computer vision, knowledge manipulation languages, logic in artificial intelligence, models of information perception, models of reasoning, intelligent robotic systems, intellectual knowledge bases, software implementation of intelligent systems, hardware implementation of intelligent systems, machine learning, machine reasoning, reasoning system, semantic network, knowledge network, frame, frame system, representation of ambiguity, representation of uncertainty, neural networks, deep blue, neuroscience, neurolinguistics, neuroeconomics, neuromarketing, neuroengineering, cognitive science, decision making, knowledge engineering, knowledge bases, ontology engineering, support vector machine, random forest, naive bayes, conditional random fields, gradient boosting, adaptive boosting, AdaBoost, self-organizing maps, decision tree, cross-validation, training set, predictive model, reinforcement learning, supervised learning, unsupervised learning, fuzzy logic, topic modeling, word2vec, doc2vec, tensorflow, keras, vector space models of text, network input, network architecture, deep network, deep algorithm, deep learning, layer of neurons, discrete network, connection weight, convolutional layer, convnet, activation function, neural network model, interneuronal communication, belief network, backpropagation, synaptic coefficients, artificial network, network learning, deep learning, autoencoder, perceptron, multilayer perceptron, fully connected network, multilayer network, recurrent network, recurrent layer, Boltzmann machine, unsupervised representation, approximate inference, opencl, variational inference, component perception, propagation target, graph computation, transfer function, embedding layer

Применение аналогичного подхода к построению списков ключевой лексики для отраслей экономики оказалось нерезультативным. В качестве отправной точки использовалась Статистическая классификация видов экономической деятельности в Европейском экономическом сообществе (редакция 2) NACE Rev 2 [15]. Выяснилось, что для основных отраслей экономики возможные списки ключевой лексики оказываются нечитабельными. Например, для отрасли «промышленность» ключевая лексика, содержащая все типы обрабатывающей промышленности, занимает более 50 страниц текста. Поэтому в дальнейшем было решено ограничиться построением ключевой лексики только для сквозных технологий, а для отраслей экономики использовать коды классификаторов, о чем пойдет речь в следующем разделе.

2.2. Выбор отраслей экономики и источников данных для наукометрического анализа

Для проведения наукометрического анализа научных заделов необходимо определить предметную область, в которой могут найти применение сквозные технологии.

С этой целью в рамках данного исследования были выбраны отрасли экономики и предметные области, определенные как приоритетные в направлении «Формирование исследовательских компетенций и технологических заделов» [4] программы «Цифровая экономика Российской Федерации»:

- промышленность;

- энергетика;
- строительство;
- сельское хозяйство;
- транспорт и логистика;
- образование;
- здравоохранение;
- государственное управление.

В качестве источника данных для наукометрического анализа была выбрана база данных Web of Science [16], для которой можно установить релевантное соответствие между ее собственными кодами классификатора (категориями) и представленными выше отраслями экономики и предметными областями.

Классификатор Web of Science Categories состоит из 250 предметных категорий, которые характеризуют релевантные журналы и издания. Для дальнейшего анализа было установлено соответствие между Web of Science Categories и анализируемыми отраслями экономики (Таблица 2).

Таб. 2. Соответствие между анализируемыми отраслями экономики и кодами классификатора Web of Science Categories

Отрасль экономики	Код Web of Science Categories
Промышленность	Engineering, Manufacturing; Engineering, Industrial; Engineering, Multidisciplinary
Энергетика	Energy & Fuels
Строительство	Construction & Building Technology
Транспорт и логистика	Transportation
Сельское хозяйство	Agriculture, Dairy & Animal Science; Agriculture, Multidisciplinary
Образование	Education & Educational Research; Education, Scientific Disciplines; Education, Special
Здравоохранение	Public, Environmental & Occupational Health; Health Policy & Services; Health Care Sciences & Services
Государственное управление	Public Administration

2.3. Выбор параметров для анализа

В качестве выборки для наукометрического анализа была использована база Web of Science Core Collection. Эта коллекция содержит записи статей только из наиболее значимых и влиятельных научных изданий в мире. При этом из анализируемой выборки были исключены публикации, находящиеся в индексе Emerging Sources Citation Index (ESCI). Дело в том, что ESCI — отдельный индекс, появившийся в базе

Web of Science в 2015 году. Этот индекс содержит публикации журналов из быстрорастущих научных областей. Основная цель введения ESCI состояла в расширении возможностей базы данных Web of Science за счет регионов мира, быстро растущих в научном отношении, и новых научных областей, в которых журналы «отставали» от изданий из более традиционных областей. Однако не всегда статьи из таких журналов являются достаточно качественными и не всегда отражают интерес ведущих научных центров по отдельным направлениям науки и технологий. Появление этой коллекции в составе Web of Science также привело к появлению большого количества публикаций, начиная с 2015 года, что вносит искажения в результаты анализа за период, включающий предшествующие и последующие годы. Этим объясняется исключение коллекции Emerging Sources Citation Index из выборки для данного исследования.

Анализ различных временных интервалов показал, что наибольшая публикационная активность по сквозным технологиям приходится на период с 2008 по 2017 годы (последние 10 лет), за эти годы Web of Science Core Collection (без учета Emerging Sources Citation Index) содержит более 23 миллионов записей. Следует отметить, что публикационная активность для нарождающихся цифровых технологий «третьей волны» [17] или новых применений цифровых технологий предыдущих поколений не всегда имеет след в 10 лет и более. Например, публикации по сквозным технологиям «систем распределенного реестра» начали появляться только в последние годы в связи с возросшим интересом к этой технологии.

3. Анализ научных заделов Российской Федерации по сквозным технологиям

3.1. Публикационная активность

Проведенный на основе предложенной в разделе 2 методологии анализ позволил построить матрицу публикационной активности научного сообщества по сквозным технологиям в разрезе отраслей экономики (Таблица 3).

Первая цифра в ячейках таблицы указывает количество публикаций в мире, вторая цифра — количество публикаций с российскими аффилиациями. Уровень интереса научного сообщества к развитию и использованию сквозной цифровой технологии в отрасли экономики представлен в виде цветового индикатора. Зеленый цвет характеризует высокий уровень научных заделов (от 1000 публикаций в мире), желтый — средний (от 100 до 1000 публикаций в мире), серый — отсутствие заделов (меньше 100 публикаций в мире). Таким образом, таблица представляет собой результат исследования интереса научного сообщества к 9 сквозным технологиям в разрезе 8 отраслей различных отраслей экономики.

Был также проведен анализ динамики публикационной активности, который выявил положительный тренд: для большинства сквозных технологий наблюдается существенный рост. При этом для технологий «систем распределенного реестра» или «квантовых технологий» первые публикации в разрезе конкретных отраслей экономики появились только в 2017 году, что демонстрирует только зарождающийся интерес к данным технологиям.

Таб. 3. Матрица публикационной активности научного сообщества по сквозным технологиям в разрезе отраслей экономики (данные Web of Science с 2008 по 2017 годы).

Отрасль экономики / сквозная технология	Матрица публикационной активности научного сообщества по сквозным технологиям в разрезе отраслей экономики (данные Web of Science с 2008 по 2017 годы)									
	больше данные	нейротехнологии и искусственный интеллект	системы распределенного реестра	квантовые технологии	новые производственные технологии	промышленный интернет	компоненты робототехники и сенсорика	технологии беспроводной связи	технологии виртуальной и доп. реальности	
промышленность	5 610 33	48 177 463	83 0	217 5	49 866 753	5 489 38	26 797 402	4 197 35	5 223 95	
энергетика	2 153 10	19 896 115	29 0	30 2	13 126 272	1 291 10	4 625 41	1 664 4	1 887 65	
строительство	403 0	8 717 47	5 0	1 0	4 829 69	348 2	1 835 11	302 1	443 2	
транспорт и логистика	761 2	7 653 24	19 0	12 0	392 10	1 049 2	3 881 7	4 669 14	328 1	
сельское хозяйство	214 0	15 392 34	1 0	3 0	484 2	281 0	1 611 1	421 0	399 0	
образование	1 975 16	12 597 116	31 0	55 1	747 15	761 0	761 0	634 0	3 046 27	
здравоохранение	1 466 3	24 738 45	40 1	0 0	752 2	395 2	1 428 5	581 3	1 311 5	
государственное управление	158 4	1 943 9	5 0	0 0	11 0	19 0	14 0	25 0	25 0	

Наибольший интерес у научного сообщества вызывают сквозные цифровые технологии направления «нейротехнологии и искусственный интеллект», демонстрирующие большие научные заделы и положительную динамику во всех исследуемых отраслях экономики и предметных областях. Наивысший интерес отмечен в промышленности, при этом следует отметить также высокую публикационную активность практически по всему спектру анализируемых сквозных технологий.

Анализ также выявил отдельные пересечения сквозных технологий и отраслей, которые показывают высокие перспективы относительно других возможных пересечений: так, «технологии беспроводной связи» среди всех исследуемых отраслей показывают наибольший публикационный интерес по отрасли «транспорт и логистика», что говорит о повышенном интересе применения данной технологии.

Сравнение публикационной активности российских исследователей показывает, что российские научные заделы, прежде всего, значительно уступают мировым в количественном отношении (см. таблицу 3). Вместе с тем, можно также сделать вывод и о разнице в структуре заделов: например, «технологии беспроводной связи» в России вызывают наибольший интерес не в отрасли «транспорта и логистика», а в отрасли «промышленность».

4. Анализ публикационной активности по направлению «нейротехнологии и искусственный интеллект»

В данном разделе представлен детальный анализ ситуации в России на примере одной направления — «нейротехнологии и искусственный интеллект», что обусловлено наибольшим интересом научного сообщества к использованию данных сквозных технологий практически по всему спектру отраслей экономики (см. таблицу 3).

4.1. Публикационная активность

Результаты анализа показали, что основная доля научных заделов в России по выбранной сквозной технологии приходится на «промышленность», а наименьший интерес проявили исследователи к предметной области «государственное управление» (Рисунок 4).

Результаты анализа мировой публикационной активности показывают, что количество заделов в России по данному направлению существенно ниже мирового уровня (Рисунок 5). Общее количество публикаций в Российской Федерации близко к 10 тысячам, в мире — близко к одному миллиону.

При этом в разрезе отраслей экономики порядок ранжирования отраслей по количеству публикаций изменился: несмотря на то, что наибольший интерес для научного сообщества все так же представляет «промышленность», а наименьший — «государственное управление», в мире на втором месте находится «здравоохранение», а в РФ — «образование».

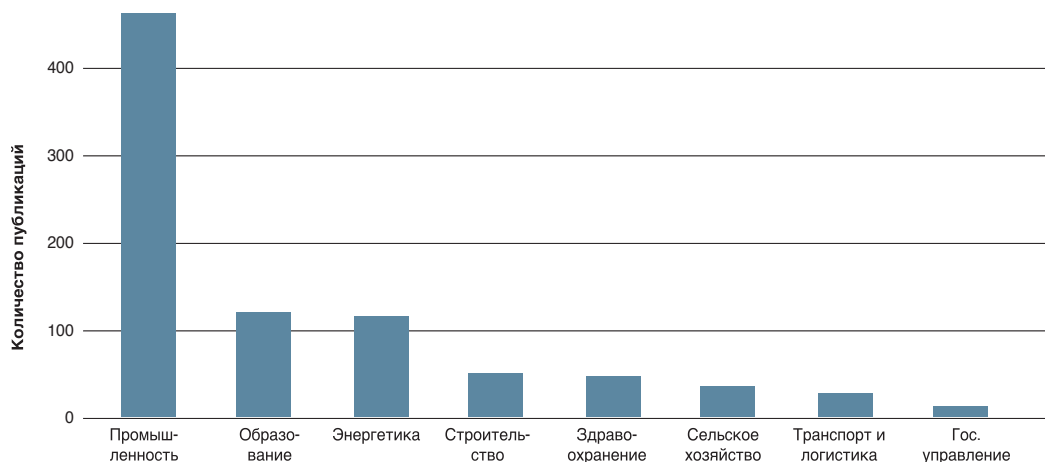


Рис 4. Количество публикаций в РФ по направлению «нейротехнологии и искусственный интеллект» в разрезе отраслей экономики

Если рассматривать мировую публикационную активность по направлению «нейротехнологии и искусственный интеллект» в целом, то наибольший вклад вносит США с 23% от мирового количества публикаций, на втором месте находится Китайская Народная Республика с 18%. Россия занимает 25-е место с менее чем 1% публикаций (отставая от таких стран, как Иран, Малайзия и Греция).

Среди организаций наибольший научный вклад в развитие направления «нейротехнологии и искусственный интеллект» вносят University of California (США, 2%), Китайская академия наук (КНР, 1,5%) и Centre national de la recherche scientifique (Франция, 1,3%). Российская академия наук находится на 65-м месте в списке организаций с 0,3% от мирового потока и 35% от всех российских публикаций. В число наиболее публикующихся российских университетов входят МГУ имени М. В. Ломоносова (10% российских публикаций), СПбГУ (4%), НИУ ИТМО (4%) и НИУ ВШЭ (2%).

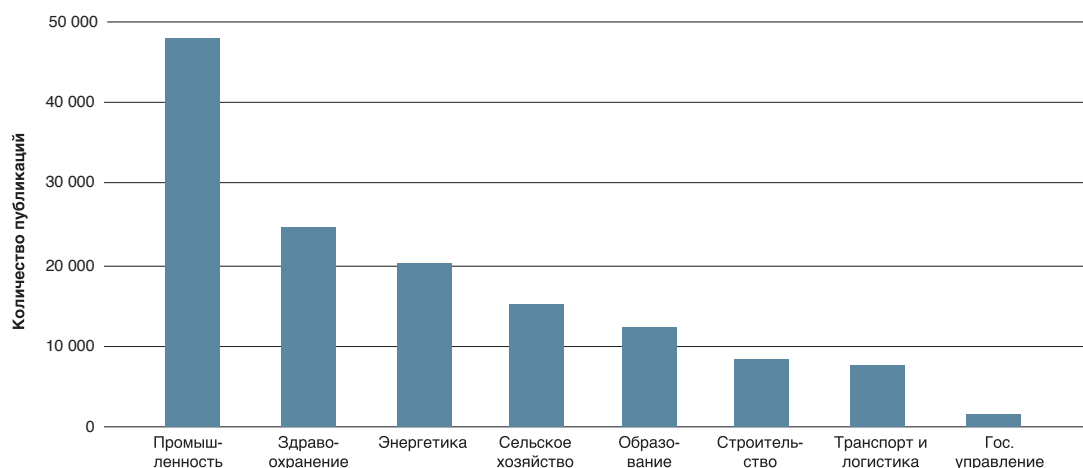


Рис 5. Количество публикаций в мире по направлению «нейротехнологии и искусственный интеллект» в разрезе отраслей экономики

4.3. Результаты и выводы

Проведенный анализ публикационной активности по направлению «нейротехнологии и искусственный интеллект» продемонстрировал наличие некоторых заделов у Российской Федерации и большой количественный разрыв по сравнению с мировыми показателями: Россия существенно отстает от стран-лидеров по количеству публикаций. При этом мировая структура научных публикаций по направлению «нейротехнологии и искусственный интеллект» внутри рассмотренных отраслей экономики также отличается от российской.

Заключение

В статье предложен подход к анализу научных заделов России по сквозным технологиям в разрезе секторов экономики или предметных областей. Предложенный подход также позволяет рассмотреть в подробностях публикационный ландшафт в России и мире. Подход апробирован на примере сквозных технологий применительно к приоритетным отраслям экономики и предметным областям, которые определены программой «Цифровая экономика Российской Федерации».

В результате анализа выявлено существенное различие между Россией и миром в структуре и количестве научных заделов. Определены отрасли экономики и предметные области для применения сквозных технологий в мире и в России, которые вызывают наибольший интерес научного сообщества. Более детальный анализ публикационной активности проведен по направлению «нейротехнологии и искусственный интеллект», в результате чего выявлены научные заделы отдельных организаций и стран.

Отдельно следует подчеркнуть, что наиболее эффективным способом анализа научных заделов является совмещение экспертной оценки и информационно-аналитических инструментов, предложенных в данной работе. При этом может потребоваться дальнейшая детализация и уточнение методологии для различных предметных областей и сквозных технологий.

Полученные результаты можно использовать как ориентир для следующего, более глубокого анализа сквозных технологий, а также как стартовую точку для оценки потребностей отраслей экономики России в научно-технологических заделах. В дальнейшем планируется в дополнение к предложенному подходу разработать и апробировать методику построения патентного ландшафта по сквозным технологиям, что представляется необходимым для оценки технологических заделов в разрезе секторов экономики или предметных областей.

Благодарности

Данное исследование выполнено при частичной поддержке РФФИ (грант № 18-29-03086).

Авторы признательны А. Г. Макушкину и А. Н. Райкову за многочисленные полезные обсуждения подходов к оценке научно-технологических заделов.

ЛИТЕРАТУРА

1. **World Development Report 2016: Digital Dividends.** Washington, DC: World Bank. — 330 p. doi:10.1596/978-1-4648-0671-1. // URL: <http://www.worldbank.org/en/publication/wdr2016/> (дата обращения: 01.09.2018).
2. **Конкуренция в цифровую эпоху: стратегические вызовы для Российской Федерации.** Доклад о цифровой экономике в России. 2018. Вашингтон, округ Колумбия: Всемирный банк. — 143 с.
3. **Программа «Цифровая экономика Российской Федерации»:** распоряжение Правительства Российской Федерации от 28 июля 2017 г. № 1632 // URL: <http://government.ru/docs/28653/> (дата обращения: 01.09.2018)
4. **О «дорожных картах» по направлениям программы «Цифровая экономика Российской Федерации»** // [Электронный ресурс]. — Режим доступа: <http://government.ru/orders/selection/401/30895/> (дата обращения: 01.09.2018)
5. **Трансфер технологий: Общие положения.** ГОСТ Р 57194.1-2016. — 10 с.
6. КОРОЩУПОВ В. О. **Научно-технологический задел, технологический брокер и потенциал конверсии** // Пути к миру и безопасности. — 2016. — № 2(51). — С. 128-140. doi:10.20542/2307-1494-2016-2-128-140
7. КРАВЧЕНКО А. Ю., ЛЯСКОВСКИЙ В. Л., АРТЕМЕНКО В. Б. **Методика автоматизированной оценки готовности научно-технического задела при планировании работ по созданию сложных технических систем** // Интернет-журнал «Науковедение», 2015, Том 7, № 4. doi:10.15862/137TVN415. [Электронный ресурс]. — Режим доступа: <http://naukovedenie.ru/PDF/137TVN415/> (дата обращения: 01.09.2018).
8. ТИХОМИРОВ И. А., ТОГАНОВА Н. В., АНАНЬЕВА М. И. **Инструменты анализа научно-технологических заделов России** // Труды Института системного анализа Российской академии наук. — 2016. — Т. 66. — № 3. — С. 98-104.
9. КУРАКОВА Н. Г., ЦВЕТКОВА Л. А., ЗИНОВ В. Г. **Патентный ландшафт РФ, созданный резидентами страны: анализ выявленных проблем** // Экономика науки, 2016. — № 1. — С. 64-79.
10. ПОПОВ Н. **Составление и анализ патентных ландшафтов** // Патенты и лицензии. Интеллектуальные права. — 2016. — № 12. — С. 39-46.
11. ПАРФЕНОВА С. Л. И ДР. **Публикационный ландшафт российской науки** // Наука. Инновации. Образование. — 2017. — № 1 (23).
12. ДЕВЯТКИН Д. А., СУВОРОВ Р. Е., ТИХОМИРОВ И. А. **О методике выявления центров компетенции на примере предметной области «искусственный интеллект»** // Информация в современном мире. — 2017. — С. 107 (2)-107 (2).
13. SCHÜTZE H., MANNING C. D., RAGHAVAN P. **Introduction to information retrieval.** — Cambridge University Press, 2008. — Т. 39.
14. ДЮМИН Н. Ю. **Модель автоматизированного построения двуязычных лексиконов** // Вестник Забайкальского государственного университета. — 2012. — № 8.
15. **NACE Rev. 2 — Statistical classification of economic activities in the European Community.** Luxembourg: Office for Official Publications of the European Communities. — 2008. — 363 pp.
16. **Web of Science** // [Электронный ресурс]. — Режим доступа: <http://www.webofknowledge.com> (дата обращения: 01.09.2018).
17. KATZ, RAUL. **Social and economic impact of digital transformation on the economy.** Geneva: International Telecommunications Union. — July 2017. — 41 p.