

# Вопросы импортозамещения и сокращения зависимости разработчиков программного обеспечения с элементами искусственного интеллекта в Российской Федерации

Май 2022



Центр компетенций  
НТИ по технологиям  
хранения и анализа  
больших данных МГУ

# Содержание

<b>Содержание</b>	<b>2</b>	Сопутствующее программное обеспечение	18
<b>Введение</b>	<b>3</b>	<b>Оценка вызовов</b>	<b>20</b>
<b>Обзор ключевых технологий ИИ</b>	<b>6</b>	<b>Предложения по снижению риска влияния</b>	<b>23</b>
Общие сведения	6	Меры поддержки развития ИИ в России	23
Машинное обучение (Machine Learning)	9	Предложения по нивелированию риска влияния санкций в сфере ИИ	24
Извлечение текстовых данных (Text Mining)	9	<b>Приложение А Основные компоненты проектов в сфере ИИ и их аналоги</b>	<b>28</b>
Распознавание речи (Speech Recognition)	10		
Обработка и понимание естественного языка (Natural Language Processing and Understanding)	10		
Компьютерное зрение (Computer Vision)	11		
Системы поддержки принятия решений (Intelligent decision support system)	11		
ИИ-роботы и беспилотный транспорт (AI robots and autonomous vehicles)	12		
Автоматическое машинное обучение (AutoML)	13		
Перспективные методы ИИ (Advanced AI Methods)	13		
<b>Обзор ключевых компонентов ИИ</b>	<b>15</b>		
Доступ к лучшим мировым практикам	15		
Инструменты работы с данными	15		
Инструменты работы дата-аналитиков	16		
Библиотеки алгоритмов	17		
Библиотеки данных	17		
Инструменты представления данных	18		

# Введение

Руководство Российской Федерации уже несколько лет уделяет пристальное внимание технологиям искусственного интеллекта (ИИ). Одним из основополагающих документов является Национальная стратегия развития искусственного интеллекта на период до 2030 года<sup>1</sup>, утвержден федеральный проект «Искусственный интеллект». Согласно отчету о состоянии рынка ИИ в 2021 году аналитической платформы Statista, Россия занимала 8 место в мире по количеству инвестиций в исследования и развитие ИИ, однако объем инвестиций был в 10 раз ниже, чем у Китая и США, занимающих лидирующие позиции по объему инвестиций<sup>2</sup>. База данных Claims Direct показывает, что за последние пять лет количество выданных патентов в этой области увеличилось с 10 тыс. до 80 тыс. в 2021 году. Лидером стала IBM – 1 813 заявок на изобретения. В число американских компаний, которые активно разрабатывают научно-исследовательские проекты с применением нейросетей, также входят Google (1 167), Adobe (580), Intel (1 131) и Microsoft (948). В Китае лидируют Baidu (317), Tencent (306), Huawei (272), в Европе - Bosch (590) и Siemens (333)<sup>3</sup>. По данным Института статистических исследований и экономики знаний НИУ ВШЭ, за последние 10 лет 91% российских регионов подали патентные заявки на изобретения в области ИИ. В целом, по данным WIPO, Россия занимает 16-е место в мире по числу подобных изобретений. Таким образом, несмотря на имеющиеся достижения, можно сделать вывод, что Россия не является лидером в развитии технологий ИИ и в настоящее время находится в зависимой позиции от зарубежных разработок.

Вместе с тем положительной характеристикой развития технологий ИИ в России является достаточно глубокое проникновение указанных технологий в повседневную деятельность бизнеса. Как показало<sup>4</sup> совместное исследование компаний PwC и ABBYY «Digital IQ 2020 в России», к 2022 году доля компаний – респондентов исследования, использующих средства искусственного интеллекта в своей повседневной деятельности, оценивалась в 68%, что составляет увеличение на 76% по сравнению с аналогичным показателем в 2020 году:

Рисунок 1. Проникновение инновационных технологий в российские компании<sup>5</sup>



<sup>1</sup> URL: <https://www.garant.ru/products/ipo/prime/doc/72738946/>

<sup>2</sup> URL: <https://rdc.grfc.ru/2021/11/artificial-intelligence-market-analysis/>

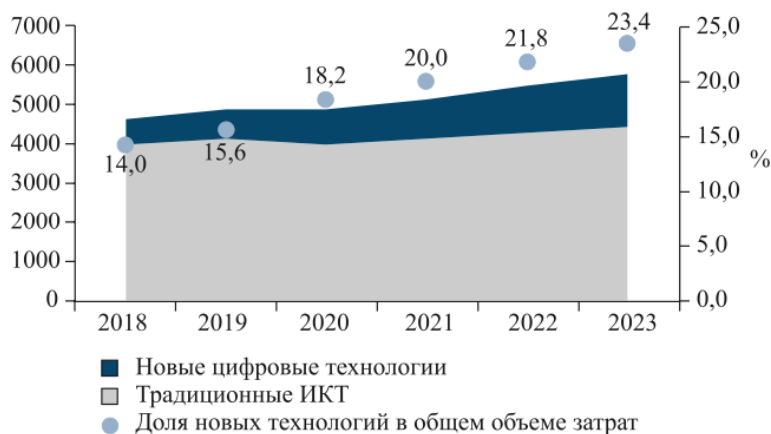
<sup>3</sup> URL: <https://tass.ru/obschestvo/13448243>

<sup>4</sup> URL: <https://www.pwc.ru/publications/digital-iq-2020/pwc-abbyy-digital-iq-2020.pdf>

<sup>5</sup> Совместное исследование PwC и ABBYY «Digital IQ 2020 в России» // URL: <https://www.pwc.ru/publications/digital-iq-2020/pwc-abbyy-digital-iq-2020.pdf>

Расширению практического применения новейших технологий, включая искусственный интеллект, также способствует увеличение финансирования разработки и внедрения новых цифровых технологий, устойчиво наблюдаемое в мировой экономике в последние годы. Как показывает<sup>6</sup> исследование НИУ ВШЭ, доля расходов на новые цифровые технологии в общем объеме затрат на информационно-компьютерные технологии увеличилась с 14% в 2018 году до 20% в 2021 году. К 2023 году ожидается рост данного показателя до 23,4%. Наблюдается также и устойчивый рост общего объема расходов на новые цифровые технологии, включая искусственный интеллект, в денежном выражении.

Рисунок 2. Объем (млрд долларов) и доля расходов на новые цифровые технологии в мире<sup>7</sup>



Вместе с тем в настоящий момент наблюдается ряд существенных рисков, способных замедлить развитие технологий искусственного интеллекта в России. В свете последних геополитических событий в мире в отношении Российской Федерации введены ряд ограничительных мер (санкции), в том числе в сфере ИТ-технологий. В феврале 2022 года санкционное давление резко усилилось со стороны таких стран, как США, Великобритания, Япония, Австралия, Канада, стран Европейского Союза и других государств. В том числе ограничения затронули возможности использования информационных технологий, программного обеспечения, правообладателями которыми являются организации, находящиеся в юрисдикции стран, присоединившихся к санкциям. Многие западные компании объявили о приостановке отгрузки продукции в Россию: Hewlett-Packard Enterprise, Dell, Apple, Philips & AOC Monitors and Displays, Canon, Xerox, Cisco, Ericsson, IBM, Nokia, Motorola Solutions, Siemens, Poly, крупнейшие производители процессоров AMD и Intel приостановили поставки своей продукции, NVIDIA приостановила продажу видеокарт, TSMC приостановила все продажи и т.п. Вендоры приостановили поставку программного обеспечения и предоставления услуг: Zoom Video Communications, Oracle, SAP, Autodesk, Adobe, Microsoft, Apple, EPAM, Atlassian (разработчик продуктов Jira, Confluence, Bitbucket), Acronis. Консалтинговые компании и глобальные интеграторы в сфере ИТ объявили об уходе из России: Accenture, DXC Technology<sup>8</sup>.

В текущей обстановке накладываемые ограничения на импорт информационных технологий в Россию максимизируют актуальность вопросов импортозамещения, в том числе в сфере развития ИИ. Необходимо оценить риски влияния санкционного воздействия на развитие ИИ в России – определить степень риска.

Следует обратить особое внимание, что среди зарубежных разработчиков Open Source ПО формируется тренд на включение в открытый код зловредных программ - специализированных команд, направленных на нанесение намеренного ущерба российским пользователям,

<sup>6</sup> URL: <https://conf.hse.ru/mirror/pubs/share/463148459.pdf>

<sup>7</sup> Доклад НИУ ВШЭ «Цифровая трансформация отраслей: стартовые условия и приоритеты» // URL: <https://conf.hse.ru/mirror/pubs/share/463148459.pdf>

<sup>8</sup> URL: <https://www.crn.ru/news/detail.php?ID=160359>

определяющих их по IP, - так называемое «протестное» ПО. Авторы некоторых библиотек в последних версиях:

- добавили в код команды, которые выводят специфические политизированные сообщения<sup>9</sup>;
- добавили в код команды, которые удаляют<sup>10</sup> или переписывают все файлы на компьютере<sup>11</sup>.

Таким образом, возникает необходимость предварительной верификации (валидации) и/или тестирования Open Source продуктов перед его использованием (применением) и размещением в отечественном репозитории.

Несмотря на активную поддержку развития ИИ в России со стороны государства, Российская Федерация все еще находится в зависимой позиции от разработок в сфере ИИ зарубежных стран. Применяются в большей части продукты и технологии ИИ, созданные разработчиками, находящимися в юрисдикции «западных» стран.

Крупные российские компании, принимающие участие в проектах по разработке и применению продуктов с технологиями ИИ (Сбер, Яндекс и т.п.), исследовательские фонды создают различные альянсы и ассоциации (Альянс в сфере искусственного интеллекта<sup>12</sup>, Ассоциация содействия ИИ в публичном секторе<sup>13</sup> и т.п.), которые, в свою очередь, активно участвуют в развитии и популяризации технологий ИИ. Коммерческими организациями создаются сборники успешных реализованных проектов с применением технологий ИИ (сборники кейсов, опыт применения) (например, библиотека AI Russia Works<sup>14</sup>), тематические базы знаний (например, база знаний ИИ ICT.moscow<sup>15</sup>), тематические порталы (например, портал искусственного интеллекта<sup>16</sup>) и репозитории датасетов (например, наборы данных Центра диагностики и телемедицины г.Москва<sup>17</sup>) и т.д. Однако такой разрозненный подход приводит к «лоскутному» покрытию информационного вакуума, а также к сложности поиска информации в разрозненных источниках. Необходимо создание единого репозитория тематической базы знаний и технологий в сфере ИИ. Со стороны государства в рамках федерального проекта «Искусственный интеллект» запланировано мероприятие по созданию национального портала в сфере искусственного интеллекта на 2022 год<sup>18</sup>.

Кроме того, следует отметить, что в России предпринимаются попытки создания организаций, направленных на объединение разработчиков программного обеспечения с открытым кодом и стимулирование таких разработок<sup>19</sup>, однако идея в настоящее время не реализована.

В рамках исследования сформированы актуальные предложения по снижению риска влияния возможных ограничений (санкций) экспорта технологий со стороны зарубежных стран на развитие применения продуктов с элементами ИИ в Российской Федерации и мер по импортозамещению.

<sup>9</sup> URL: <https://github.com/medikoo/es5-ext/commit/28de285ed433b45113f01e4ce7c74e9a356b2af2>

<sup>10</sup> URL: <https://github.com/vuejs/vue-cli/issues/7054>

<sup>11</sup> URL: <https://github.com/RIAEvangelist/node-ipc/issues/233>

<sup>12</sup> URL: <https://a-ai.ru/#about>

<sup>13</sup> URL: <https://gos.ai/>

<sup>14</sup> URL: <https://ai-russia.ru/library/>

<sup>15</sup> URL: <https://ict.moscow/projects/ai/>

<sup>16</sup> URL: <http://www.aiportal.ru/>

<sup>17</sup> URL: <https://mosmed.ai/datasets/>

<sup>18</sup> URL: <https://tass.ru/ekonomika/13959441>

<sup>19</sup> URL: [https://www.rbc.ru/technology\\_and\\_media/04/10/2021/615757b29a7947a3df10e734](https://www.rbc.ru/technology_and_media/04/10/2021/615757b29a7947a3df10e734)

# Обзор ключевых технологий ИИ

## Общие сведения

В качестве искусственного интеллекта принято понимать область информационных технологий, которая направлена на разработку программных и программно-аппаратных средств и продуктов, способных выполнять ряд задач (в первую очередь связанных с анализом данных) творческого характера, которые принято связывать с человеческим разумом. Системы искусственного интеллекта способны имитировать свойства человеческого разума или поведения для выполнения задач определенного характера. Важным свойством разработок, основанных на технологиях искусственного интеллекта, является также способность к постепенному самостоятельному обучению или обучению с помощью со стороны экспертов на основе новых данных и предыдущих результатов выполнения задач, на которые ориентирована система.

Системы и решения искусственного интеллекта позволяют эффективно автоматизировать выполнение задач, которые ранее могли осуществляться только при помощи человеческого труда, при этом производительность искусственного интеллекта, способного обрабатывать большие объемы данных (так называемые «большие данные»), во много раз превосходит производительность человеческого труда. Применение искусственного интеллекта позволяет проводить анализ всего объема данных, накопленного предприятием или отраслью, составлять надежные прогнозы и осуществлять автоматизацию сложных задач.

Необходимыми предпосылками для широкого распространения и практического применения систем искусственного интеллекта являются:

- доступность высокопроизводительных и относительно недорогих вычислительных ресурсов;
- доступность больших объемов данных для обучения систем искусственного интеллекта.

Среди форм искусственного интеллекта принято выделять сильный и слабый виды. Сильный искусственный интеллект представляет собой теоретическую систему ИИ, обладающую всеми свойствами человеческого интеллекта и способную решать задачи любого рода без необходимости предварительного обучения или изменения системы. Сильный искусственный интеллект в настоящий момент существует лишь как теоретическая концепция. К слабому искусственному интеллекту относятся системы искусственного интеллекта, воспроизводящие лишь отдельные свойства и стороны человеческого интеллекта и направленные на решение конкретных узкоспециализированных задач. К слабому искусственному интеллекту относятся все современные системы искусственного интеллекта.

В существующих аналитических обзорах, связанных с оценкой емкости мирового и российского рынков решений в области искусственного интеллекта, содержатся<sup>20</sup> достаточно позитивные сценарии роста указанных рынков. Ожидается, что общий объем мирового рынка ИИ в 2024 году составит 137 млрд долларов США, продемонстрировав, таким образом, рост приблизительно в 7 раз по сравнению с 2018 годом и рост более чем в 3 раза по сравнению с 2020 годом. Около половины мирового рынка решений искусственного интеллекта в денежном выражении, как ожидается, будут занимать рекомендательные системы и интеллектуальные системы поддержки принятия решений. Важными сегментами также будут являться системы и технологии

---

<sup>20</sup> URL: [https://www.cnews.ru/news/top/2019-10-28\\_rossijskij\\_iskusstvennyj](https://www.cnews.ru/news/top/2019-10-28_rossijskij_iskusstvennyj)

компьютерного зрения, а также обработки естественного языка (NLP). Динамика ожидаемого роста мирового рынка решений ИИ и отдельных его сегментов, а также доля отдельных сегментов мирового рынка ИИ приведены на графиках ниже.

Рисунок 3. Прогноз объема мирового рынка решений в области искусственного интеллекта, \$ млрд<sup>21</sup>

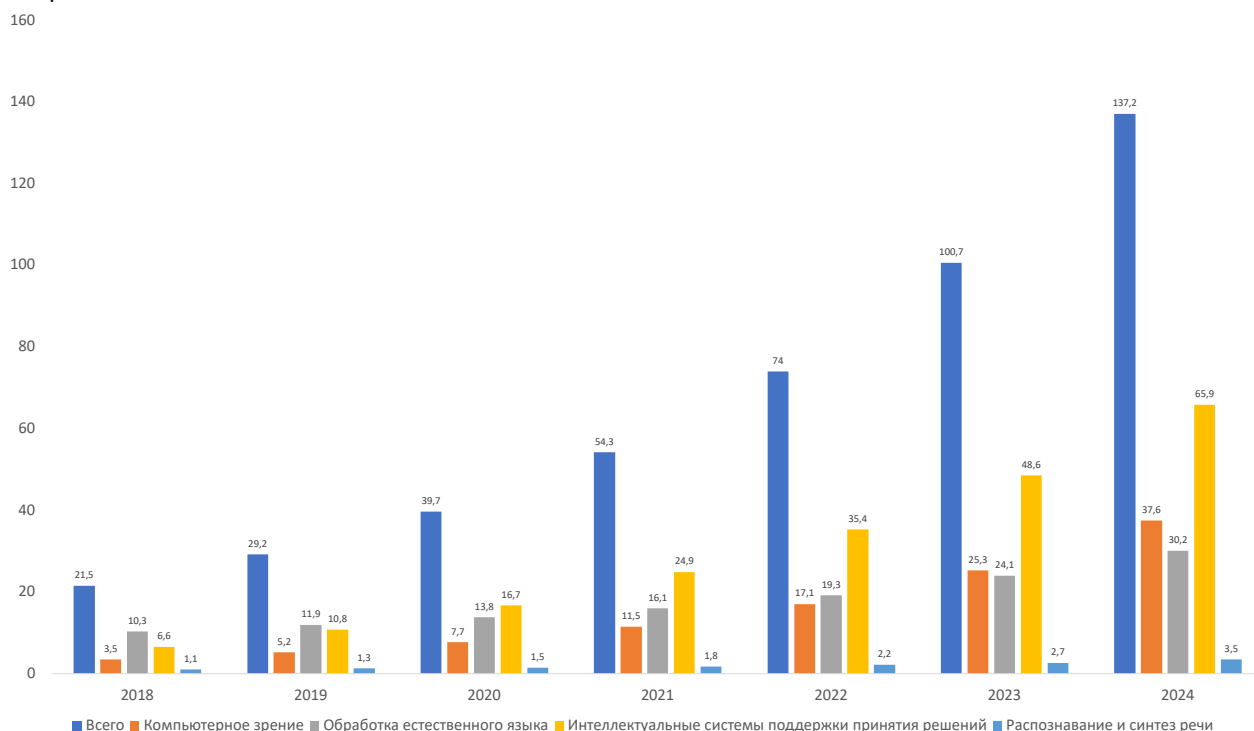


Рисунок 4. Структура мирового рынка решений в области искусственного интеллекта в 2021 году, %<sup>22</sup>



<sup>21</sup> составлено авторами на основе данных портала CNews // URL: [https://www.cnews.ru/news/top/2019-10-28\\_rossijskij\\_iskusstvennyj](https://www.cnews.ru/news/top/2019-10-28_rossijskij_iskusstvennyj)

<sup>22</sup> составлено авторами на основе данных портала CNews // URL: [https://www.cnews.ru/news/top/2019-10-28\\_rossijskij\\_iskusstvennyj](https://www.cnews.ru/news/top/2019-10-28_rossijskij_iskusstvennyj)

В существующей рыночной аналитике также представлен<sup>23</sup> прогноз развития отечественного рынка решений в области искусственного интеллекта. В соответствии с прогнозом, ожидается рост российского рынка технологий искусственного интеллекта опережающими темпами по сравнению с общемировыми. Объем рынка в 2024 году оценивается в 160 млрд рублей, что на порядок превышает объем российского рынка ИИ в 2020 году. Наиболее важной и конкурентоспособной в том числе и на мировом уровне в соответствии с прогнозом является отрасль решений в области компьютерного зрения, которая, как ожидается, составит около 40% всего российского рынка решений ИИ в 2024 году. Динамика ожидаемого роста российского рынка решений ИИ и отдельных его сегментов, а также доля отдельных сегментов российского рынка ИИ приведены на графиках ниже.

Рисунок 5. Прогноз объема российского рынка решений в области искусственного интеллекта, млрд рублей<sup>24</sup>

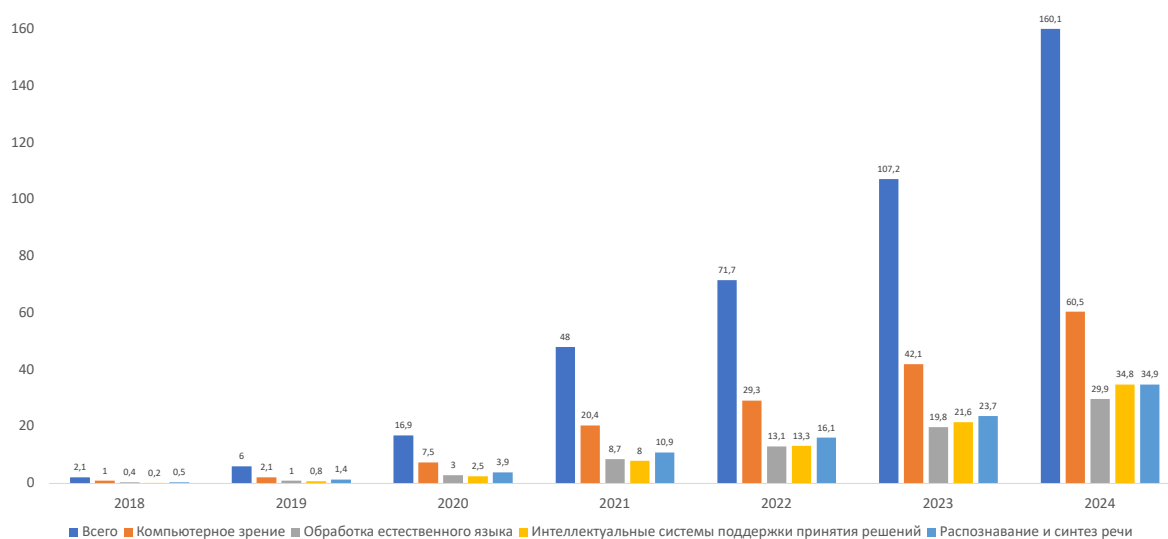
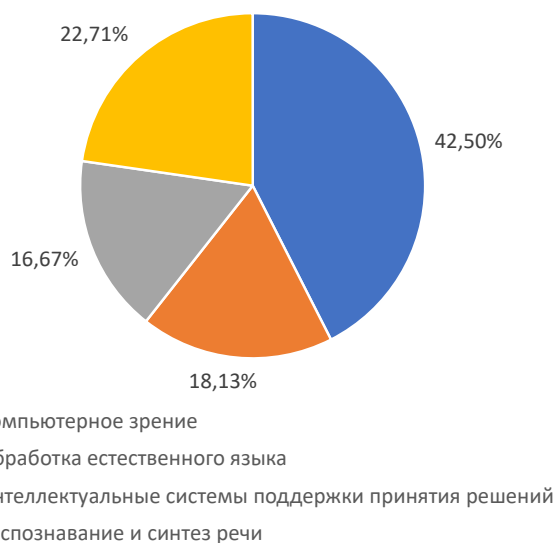


Рисунок 6. Структура российского рынка решений в области искусственного интеллекта в 2021 году, %<sup>25</sup>



<sup>23</sup> [https://www.cnews.ru/news/top/2019-10-28\\_rossijskij\\_iskusstvennyj](https://www.cnews.ru/news/top/2019-10-28_rossijskij_iskusstvennyj)

<sup>24</sup> составлено авторами на основе данных портала CNews // URL: [https://www.cnews.ru/news/top/2019-10-28\\_rossijskij\\_iskusstvennyj](https://www.cnews.ru/news/top/2019-10-28_rossijskij_iskusstvennyj)

<sup>25</sup> составлено авторами на основе данных портала CNews // URL: [https://www.cnews.ru/news/top/2019-10-28\\_rossijskij\\_iskusstvennyj](https://www.cnews.ru/news/top/2019-10-28_rossijskij_iskusstvennyj)



Важным замечанием к представленным прогнозам является тот факт, что указанные прогнозы не учитывают возможное негативное влияние наложенных в адрес Российской Федерации технологических санкций. Как следствие, эти прогнозы можно полагать лишь наиболее оптимистическим сценарием развития технологий и рынка искусственного интеллекта в России.

Далее представлены основные классы и технологии систем искусственного интеллекта, имеющие в настоящее время широкое применение, а также предложено описание некоторых основных перспективных методов искусственного интеллекта.

## **Машинное обучение (Machine Learning)**

Машинное обучение является областью искусственного интеллекта, которая использует данные и алгоритмы, чтобы воссоздать особенности человеческого обучения, в том числе постепенное повышение качества результата действий в процессе обучения. В рамках машинного обучения разрабатывается специализированная модель, которая обучается принятию решений или прогнозированию на определенном тренировочном наборе данных (датасете). После завершения обучения модель может выполнять аналитические задачи на произвольных данных (имеющих, однако, такие же тип и структуру, как и у тренировочных данных) без необходимости дополнительной настройки или программирования.

Выделяют несколько основных методик создания моделей машинного обучения, к которым относятся:

- обучение с учителем. В рамках данной методики обучение модели проводится с использованием заранее размеченных вручную данных;
- обучение без учителя. С помощью данного вида обучения создаются модели машинного обучения, способные анализировать неразмеченные данные и автоматически осуществлять их кластеризацию и классификацию;
- обучение с частичным привлечением учителя использует небольшой объем размеченной информации для управления процессом классификации и выделения признаков из неразмеченного датасета большего объема;
- обучение с подкреплением использует «поощрение» за правильный выбор при анализе данных и, в результате цепочки таких шагов, достигает необходимой точности анализа.

Подмножеством алгоритмов машинного обучения является так называемое глубокое обучение. Модели глубокого обучения формируются на основе очень больших наборов данных, при этом данные могут быть как структурированными и размеченными, так и иметь неструктурированный характер. На основе этих данных алгоритмы глубокого обучения автоматически выделяют признаки, характерные для тех или иных категорий данных и необходимые для идентификации соответствующих категорий данных. Глубокое обучение не требует предварительной ручной обработки данных, однако для высокой точности полученной модели необходим большой объем данных для обучения.

Машинное обучение является технологической основой для реализации многих систем искусственного интеллекта.

## **Извлечение текстовых данных (Text Mining)**

Под извлечением (парсингом) текстовых данных, как правило, понимают комплекс технологий, обеспечивающих процесс автоматизированного анализа больших объемов текстовой информации

(в том числе распределенной по нескольким документам) в целях оперативного и точного выявления фактов, сущностей и взаимоотношений между ними, которые могли бы остаться незамеченными при ручном анализе этого объема информации. Выявленные сведения конвертируются в структурированную форму, подходящую для дальнейшей описательной или предиктивной аналитики и интеграции с базами данных и системами Business Intelligence. Как правило, алгоритмы извлечения текстовых данных в качестве первого этапа осуществляют структуризацию входящей текстовой информации (наиболее распространенным инструментом для этого являются алгоритмы парсинга), при помощи статистических методов выявляют образы в полученной структурированной информации, после чего проводят оценку и интерпретацию сформированных образов и данных.

Результаты извлечения текстовых данных далее используются для решения задач классификации текстов (отнесения текста к тому или иному виду среди заранее заданного множества), кластеризации документов (выявления групп схожих документов среди заданной выборки документов), выявления идей и сущностей в тексте, анализа тональности и характера текста, автоматического реферирования, выявления именованных сущностей в тексте и моделирования взаимоотношений между этими сущностями.

Технологии извлечения текстовых данных являются тесно связанными с методами обработки естественного языка и машинным обучением.

## **Распознавание речи (Speech Recognition)**

К распознаванию речи следует отнести комплекс методов, позволяющих с достаточно высокой точностью преобразовать устную речь в письменный машиночитаемый вид. Наиболее продвинутые современные системы распознавания речи широко используют методики машинного обучения и оценивают грамматику, структуру, синтаксис и композицию устной речи для ее приведения в машиночитаемый вид, а также обладают способностью к самостоятельному обучению и повышению точности распознавания на основе предыдущего взаимодействия с пользователями. Современные системы распознавания речи также обладают способностью к адаптации под специфику окружения, в котором они будут использоваться, в частности, они могут учитывать отраслевую специфику и распознавать не только слова из базового словарного запаса системы, но и жаргонные термины и названия продуктов компании, а также адаптироваться под особенности речи пользователей и фоновый шум, повышая, таким образом, точность распознавания речи. Современные системы распознавания речи также обладают способностью автоматической идентификации участников в случаях, если распознается диалог двух и более сторон, и присвоения им соответствующих меток в полученном письменном тексте.

Технологической основой современных систем распознавания речи являются, как правило, статистические алгоритмы, в частности, скрытые марковские модели, определяющие неизвестные параметры модели на основе наблюдаемых. Широкое распространение в решении задач распознавания речи также получили искусственные нейронные сети, в том числе использующие алгоритмы глубокого обучения. Также развиваются подходы построения полностью автоматизированных моделей распознавания речи, которые не требуют отдельной разработки и обучения акустической и языковой компонент модели.

## **Обработка и понимание естественного языка (Natural Language Processing and Understanding)**

В качестве технологий обработки и понимания естественного языка понимаются алгоритмы, использующие техники и методики искусственного интеллекта и машинного обучения для анализа произвольного текста. При помощи указанных технологий можно выполнять

идентификацию идей и концепций в тексте в целом, автоматическое формирование его краткого пересказа или определяющих текст ключевых слов, выявление тональности текста, извлечение именованных сущностей, машинный перевод. Современные решения обработки естественного языка могут анализировать и выявлять смысл текста как на основе его семантики, так и синтаксиса. В рамках синтаксического анализа NLP-решение определяет смысл текста на основе грамматических правил языка. При помощи семантического анализа NLP-решение определяет смысл и контекст текста на основе смысла и структуры отдельных предложений.

Наиболее продвинутыми и точными средствами обработки естественного языка в настоящее время являются статистические модели обработки естественного языка, опирающиеся на методики и алгоритмы машинного и глубокого обучения с применением нейронных сетей.

Положительным фактором, способствующим успешному импортозамещению технологий в сфере обработки и понимания естественного языка, является то, что уже в настоящее время передовые российские финансовые и технологические компании разработали и применяют на практике высокоэффективные модели обработки естественного языка, являющиеся конкурентоспособными по функциональным возможностям и качеству выполнения задач в том числе и на мировом уровне. К числу таких моделей можно отнести, в частности, RuGPT-3 Сбербанка и YaLM компании «Яндекс».

## **Компьютерное зрение (Computer Vision)**

Компьютерное зрение является областью искусственного интеллекта, которая разрабатывает алгоритмы, позволяющие компьютерам извлекать информацию из изображений, видеофайлов и другой визуальной информации. Современные системы компьютерного зрения, как правило, используют в качестве технологической основы методики глубокого обучения и сверточные нейронные сети. При помощи методик глубокого обучения на основе большого количества накопленных и подготовленных визуальных данных система компьютерного зрения обучается идентифицировать изображенные или видимые ей объекты.

Системы компьютерного зрения в настоящее время широко используются для решения следующих практических задач:

- классификация изображений;
- выявление объектов на изображении;
- отслеживание объекта в кадре;
- интеллектуальный поиск изображений;
- оптическое распознавание символов (Optical Character Recognition, OCR);
- распознавание лиц;
- восстановление изображений.

Важным свойством систем компьютерного зрения, основанных на нейронных сетях, являются высокие требования к объему информации (датасету), необходимой для их обучения, и к производительности оборудования, на котором исполняется программный код системы компьютерного зрения. В связи с этим значительная доля современных систем компьютерного зрения реализуется в облачном варианте.

## **Системы поддержки принятия решений (Intelligent decision support system)**

Системы поддержки принятия решений являются информационными системами, которые применяют различные технологии искусственного интеллекта для анализа собранной в рамках системы информации и построения рекомендаций или прогноза для пользователей

информационной системы. Для этого механизмы искусственного интеллекта, встроенные в систему поддержки принятия решений, эмулируют поведение человеческих консультантов в процессах сбора и анализа информации, идентификации и диагностики проблем в бизнес-процессах, предлагают возможные пути решения выявленных проблем.

Значительная доля промышленных систем поддержки принятия решений является дальнейшим развитием экспертных систем, которые основаны на подготовленной экспертами в соответствующей предметной области базе знаний и логических правилах, определяющих выбор системы в той или иной ситуации. Такие экспертные системы в некоторых ситуациях были способны быстро и точно рекомендовать оптимальное решение, однако их способности были крайне ограничены в условиях неопределенности или неполной определенности внешних факторов. Переход к системам поддержки принятия решений, использующим инструментарий искусственного интеллекта (в частности, такие системы широко используют методы машинного обучения, интеллектуального анализа текста и данных, нейронные сети), позволяет эффективно использовать их рекомендательные и прогностические способности в том числе и в условиях неопределенности, а также формировать полезные для бизнеса рекомендации на основе больших объемов необработанных данных.

## **ИИ-роботы и беспилотный транспорт (AI robots and autonomous vehicles)**

Искусственный интеллект является мощным инструментом автоматизации задач в промышленности. Применение технологии ИИ помогает увеличить гибкость фабричных роботов, расширить возможности их применения, оптимизировать выполняемые роботами операции за счет способности методов искусственного интеллекта к самостоятельному обучению. Как следствие, применение ИИ-роботов позволяет сделать производство эффективнее, сократить количество брака. В мировой практике достаточно широко распространено использование искусственного интеллекта в рамках роботизированных конвейерных операций. В сочетании с методами компьютерного зрения ИИ способен выявлять брак в продукции или, в случае производства сложных устройств, выявлять некорректные операции, предотвращая, таким образом, сборку некондиционного устройства или его повреждение. ИИ-роботы также зачастую используются для упаковки продукции.

Важным смежным направлением разработки специализированных средств искусственного интеллекта является развитие беспилотного транспорта. Конечной целью развития беспилотного транспорта является создание наземных транспортных средств, которые способны получать и оценивать информацию об окружающей обстановке и безопасно двигаться к назначенной цели без необходимости управления ими водителем. Для анализа окружающей обстановки беспилотный транспорт использует комплекс технических средств, включающий камеры, радары, лидары, сонары, геопозиционные системы и инерционные средства измерения. На основании входящей информации от указанных технических средств специализированная система искусственного интеллекта определяет маршрут транспортного средства, идентифицирует препятствия и релевантные дорожные знаки и подает транспортному средству управляющие команды.

На текущий момент наиболее продвинутые системы автопилотирования, допущенные к использованию на дорогах общего пользования, не имеют возможности полностью автоматического управления автомобилем в течение всей поездки, а могут использоваться лишь в некоторых ситуациях (как правило, достаточно простых с точки зрения требований к анализу окружающей обстановки, к примеру, движению по автомагистрали) и требуют сохранения внимания водителя во время поездки в случае, если понадобится его вмешательство. Дальнейшее развитие систем автопилотирования связано с возрастающей сложностью алгоритмов оценки и анализа окружающей обстановки, которым требуется корректно идентифицировать большое количество разнообразных объектов. Кроме того, критически

важным для создания беспилотного транспорта является доступ к упомянутым выше техническим средствам, необходимым для получения сведений об окружающей обстановке.

## Автоматическое машинное обучение (AutoML)

Автоматическое машинное обучение является совокупностью методов автоматизации различных аспектов применения машинного обучения для решения проблем обработки информации. AutoML возникло как ответ на вызов, связанный с непрерывно возрастающей технологической сложностью применения современных моделей машинного обучения и имеет инструменты, позволяющие автоматизировать каждый из этапов создания модели машинного обучения.

Автоматическое машинное обучение в настоящее время, как правило, используется на этапе прототипирования модели для быстрой проверки гипотез и исследования данных. После получения удовлетворительных результатов предложенная инструментами автоматического машинного обучения модель дорабатывается инженерами машинного обучения в целях получения более точных результатов работы модели и увеличения производительности ее работы.

Автоматическое машинное обучение может применяться на всех этапах создания модели ML, а именно:

- сбор и подготовка данных, включая выявление типа данных и определение задачи модели;
- конструирование признаков модели;
- обнаружение и обработка отсутствующих или асимметричных данных;
- выбор типа модели;
- создание при необходимости ансамбля методов обучения модели;
- оптимизация параметров алгоритма обучения модели;
- выбор процессов работы модели с учетом ограничений по времени и машинным ресурсам;
- оценка качества и валидация модели;
- анализ полученных результатов;
- создание пользовательского интерфейса и средств визуализации.

## Перспективные методы ИИ (Advanced AI Methods)

К перспективным методам искусственного интеллекта относятся разработки, направленные на создание принципиально новых систем и методов применения искусственного интеллекта, а также готовой технической продукции, использующей на практике такие методы. К одним из наиболее важных направлений развития перспективных методов ИИ относится научно-исследовательская проработка вопросов создания сильного искусственного интеллекта, обладающего всеми свойствами человеческого интеллекта и способного решать любые задачи без предварительного обучения или подготовки. К прочим перспективным методам искусственного интеллекта можно отнести:

- разработку алгоритмов квантового машинного обучения и проработку вопросов использования квантовых вычислений для решения задач искусственного интеллекта;
- использование средств искусственного интеллекта для автономного сквозного
- создание интерпретируемых (объяснимых) моделей искусственного интеллекта;
- создание методов оценки предвзятости систем искусственного интеллекта;
- создание систем искусственного интеллекта, осуществляющих

- проектирования сложных объектов, в том числе программно-аппаратной части технических устройств;
- создание гибридных моделей искусственного интеллекта, комбинирующих методы искусственного интеллекта с классическими алгоритмами, для использования в плохо формализуемых прикладных областях;
- централизованное и децентрализованное управление группами однородных и неоднородных объектов;
- создание систем искусственного интеллекта, направленных на автоматизированное обогащение и улучшение качества больших данных.

# Обзор ключевых компонентов ИИ

В настоящем разделе представлен общий обзор ключевых технических компонентов и источников знаний, которые необходимы для успешной реализации проектов разработки и внедрения систем искусственного интеллекта.

## Доступ к лучшим мировым практикам

Одним из основных аспектов успешной реализации проектов в области искусственного интеллекта является доступ к накопленным мировым сообществом знаниям в области ИИ. Критически важное значение для развития отечественной отрасли ИИ имеет наличие доступа к ресурсам, публикующим международные научные статьи в области искусственного интеллекта, и ресурсам, осуществляющим поиск международных научных статей. Среди таких ресурсов можно выделить, в частности: Google Scholar; ArXiv.org; Web of Science; ScienceDirect; Sci-Hub; ResearchGate; IEEE Xplore.

Кроме того, достаточно значимую практическую роль для профессионалов в сфере искусственного интеллекта имеют веб-ресурсы, на которых профессиональное сообщество имеет возможность обсуждения вопросов, возникающих в ходе реализации проектов в сфере искусственного интеллекта в частности и реализации проектов в сфере информационных технологий в целом. Одним из наиболее важных подобных ресурсов является веб-сайт Stack Overflow, предоставляющий наиболее крупную в мировом масштабе площадку для обсуждения пользователями вопросов, возникающих на различных этапах реализации проектов в сфере информационных технологий.

Необходимо отметить, что представленные ресурсы в подавляющем большинстве размещаются на площадках недружественных зарубежных юрисдикций.

## Инструменты работы с данными

Важнейшей составляющей разработки точных и эффективных систем искусственного интеллекта является доступ к большим объемам обучающих данных для системы ИИ, а также сбора, обработки и хранения входящих данных для системы искусственного интеллекта и результатов работы этой системы. В связи с этим важным фактором является наличие современных программных и аппаратных средств хранения и обработки данных.

Основным инструментом работы с данными являются хранилища данных. Хранилища данных могут быть реализованы в локальном или сетевом (облачном) виде. Для реализации хранилища данных в локальном виде необходимым условием является наличие специализированного оборудования для хранения данных – запоминающих устройств, реализованных в виде жестких дисков (HDD) или твердотельных накопителей (SSD). И жесткие диски, и твердотельные накопители в настоящий момент производятся исключительно за рубежом.

Для обеспечения работы программной составляющей хранилища данных необходимо специализированное программное обеспечение – системы управления базами данных (СУБД). СУБД можно разделить на две основные категории: реляционные СУБД, управляющие структурированными данными в табличном представлении, и NoSQL – СУБД, способные

управлять неструктурированными данными. Как реляционные, так и NoSQL – СУБД представлены и коммерческими проприетарными, и открытыми разработками. К наиболее популярным открытым реляционным СУБД можно отнести продукты MySQL, PostgreSQL и MariaDB. Примерами популярных NoSQL – СУБД являются Neo4j, MongoDB и Redis. Большинство представленных продуктов доступно в режиме свободного распространения, но при этом подавляющее большинство команд разработчиков размещаются в недружественных зарубежных юрисдикциях (за исключением распределенной команды разработчиков PostgreSQL). Существуют отечественные аналоги реляционных СУБД: Линтер (Россия), Ред База Данных (Россия).

Важным аспектом обработки данных при реализации проектов ИИ также является сбор и обработка данных, полученных из разных источников, и их хранение в едином хранилище данных (озере данных). Существует два основных подхода к реализации этого процесса. В рамках процедур ETL (Extract, Transform, Load; извлечение, преобразование, загрузка) первым этапом является получение исходных данных в различных форматах, после чего происходит обработка этих данных и их приведение к единому формату, поддерживаемому хранилищем данных. После завершения обработки данные загружаются в хранилище данных и могут быть использованы системой ИИ.

В рамках схемы ELT (Extract, Load, Transform; извлечение, загрузка, преобразование) входящие данные сразу загружаются в хранилище данных в том виде, в котором они были получены, без проведения обработки. Приведение данных к виду, в котором они требуются системе, обслуживаемой хранилищем данных, происходит в момент запроса данных системой.

## Инструменты работы дата-аналитиков

Существенным аспектом успешной реализации проектов, связанных с искусственным интеллектом, является предварительный анализ и подготовка данных для проекта. Для этих целей дата-аналитики используют комплекс программного обеспечения, в том числе языки программирования, библиотеки для них, фреймворки работы с большими данными и др. Для успешной реализации проектов, связанных с искусственным интеллектом, необходимо обеспечить доступ дата-аналитиков к таким инструментам работы или их аналогам. Среди инструментов работы дата-аналитиков можно выделить:

- статистические и математические пакеты программного обеспечения, например, SAS, SPSS, Matlab. Все представленные продукты распространяются на условиях коммерческого программного обеспечения, команды разработчиков размещаются в недружественных зарубежных юрисдикциях. Существуют отечественные аналоги: Polymatica (Россия), Статэк (Россия);
- фреймворки работы с большими данными, например, Apache Hadoop. Все представленные продукты распространяются на условиях свободной, открытой лицензии, команды разработчиков размещаются в недружественных зарубежных юрисдикциях. Отечественных аналогов или продуктов разработчиков команд дружественных юрисдикций не выявлено;
- специализированные системы управления базами данных, например, Apache Spark, MongoDB, SAP Hana. Подавляющее большинство представленные продукты распространяются на условиях свободной, открытой лицензии, команды разработчиков размещаются в недружественных зарубежных юрисдикциях. Отечественных аналогов или продуктов разработчиков команд дружественных юрисдикций не выявлено;
- языки программирования, например, Python, R. Интерпретаторы распространяются с открытым исходным кодом;



- специализированные библиотеки для языков программирования, например, TensorFlow, PyTorch, scikit-learn, matplotlib, NLTK. Подавляющее большинство представленные библиотек размещаются в недружественных зарубежных юрисдикциях;
- инструменты визуализации данных, например, Tableau, Qlik, PowerBI. Все представленные продукты распространяются на условиях коммерческого программного обеспечения, команды разработчиков размещаются в недружественных зарубежных юрисдикциях. Существуют отечественные аналоги: Visiology (Россия), N3.Аналитика (Россия), Visary BI (Россия), Форсайт Аналитическая Платформа (Россия), Интеград Аналитика (Россия), Криста BI (Россия);
- инструменты подготовки данных, например, BigML, Trifacta, Knime. Все представленные продукты распространяются на условиях свободной, открытой лицензии, команды разработчиков размещаются в недружественных зарубежных юрисдикциях. Отечественных аналогов или продуктов разработчиков команд дружественных юрисдикций не выявлено;
- специализированные инструменты анализа данных, например, интерактивный блокнот Jupyter, распространяется на условиях свободной, открытой лицензии, команда разработчиков размещается в недружественных зарубежных юрисдикциях.

## Библиотеки алгоритмов

Для эффективной реализации проектов в области искусственного интеллекта необходимо наличие доступа специалистов в сфере ИИ к специализированным библиотекам алгоритмов для языков программирования, упрощающих работу с большими данными, необходимыми для обучения моделей машинного обучения, реализующих работу моделей машинного обучения и других подмножеств ИИ, визуализирующих данные и результаты работы алгоритмов ИИ. К числу наиболее важных и распространенных современных библиотек алгоритмов искусственного интеллекта можно отнести: TensorFlow; Microsoft CNTK; Theano; Caffe; Keras; Torch; Spark Mlib; Sci-kit Learn; MIPack. Необходимо отметить, что представленные ресурсы в подавляющем большинстве размещаются на площадках недружественных зарубежных юрисдикций.

## Библиотеки данных

Неотъемлемой частью информации, необходимой для успешной разработки и реализации проектов в сфере искусственного интеллекта, являются наборы данных (датасетов), на основе которых осуществляется обучение модели машинного обучения. Датасеты, как правило, имеют значительный объем, а информация в них должна быть соответствующим образом подготовлена и размечена для эффективного использования в процессах обучения алгоритмов искусственного интеллекта.

Как правило, датасеты поставляются совместно с библиотеками алгоритмов машинного обучения. Вместе с тем существует ряд ресурсов, предназначенных для поиска и/или публикации специализированных датасетов, обеспечение доступа к которым является важной задачей для дальнейшей успешной реализации проектов в области ИИ. К таким ресурсам можно отнести: Google Dataset Search Engine, Kaggle, Earth Data, Amazon Datasets, Azure Datasets, Data World, Lionbridge AI, UCI Machine Learning Repository. Представленные ресурсы в подавляющем большинстве размещаются на площадках недружественных зарубежных юрисдикций. Существуют аналогичные отечественные ресурсы: Инфраструктура научно-исследовательских данных (ИНИД) Центра перспективных управленческих решений (ЦПУР) (Россия); Университетский консорциум исследователей больших данных (Россия); Яндекс.Толока – открытые датасеты (Россия); Хаб открытых данных АНО «Информационная культура» (Россия);

Отраслевой реестр датасетов АО «Концерн Росатом» (Россия); Каталог каталогов данных datacatalogs.ru АНО «Инфокультура» (Россия).

## Инструменты представления данных

Немаловажным программным компонентом, обеспечивающим эффективную реализацию проектов искусственного интеллекта, а также являющимся ценным средством доведения результатов работы систем искусственного интеллекта до широкой аудитории, являются инструменты графического представления (визуализации) данных. Эти инструменты могут представлять собой как библиотеки для языков программирования, так и самостоятельные программные продукты. В число наиболее популярных современных инструментов представления данных в области искусственного интеллекта входят: Matplotlib, Seaborn, Microsoft Power BI, Google Charts, Tableau, Qlik Sense.

Подавляющее большинство представленных продуктов распространяются на условиях коммерческого программного обеспечения, команды разработчиков размещаются в недружественных зарубежных юрисдикциях. Существуют отечественные аналоги: Visiology (Россия), N3.Аналитика (Россия), Visary BI (Россия), Форсайт Аналитическая Платформа (Россия), Интеград Аналитика (Россия), Криста BI (Россия).

## Сопутствующее программное обеспечение

Как и при разработке программного обеспечения других типов, достаточно важное значение при реализации проектов искусственного интеллекта имеет сопутствующее вспомогательное программное обеспечение, упрощающее и делающее более эффективными процессы разработки. Среди сопутствующего программного обеспечения необходимо отдельно выделить среды разработки и средства реализации репозитория программного обеспечения.

Интегрированная среда разработки представляет собой единый комплекс программных средств, используемый для разработки ПО. В ее состав, как правило, входят редактор исходного кода программы, компилятор (для компилируемых языков программирования) или интерпретатор (для интерпретируемых языков программирования), средства автоматизации сборки и отладчик программы. Современные среды программирования также зачастую поддерживают интеграцию со средствами управления версиями ПО и репозиториями ПО. Большинство современных интегрированных сред разработки поддерживают несколько языков программирования. Наиболее популярным в настоящее время языком программирования для нужд анализа данных и конструирования систем искусственного интеллекта является Python. Таким образом, необходимо обеспечить доступ отечественного сообщества в сфере искусственного интеллекта к интегрированным средам разработки (или в краткий срок создать их аналоги), поддерживающим язык программирования Python. В их число входят: NetBeans, PyCharm, Geany, IDLE (является частью стандартного комплекта поставки языка программирования Python), Eclipse, Microsoft Visual Studio. Подавляющее большинство представленных продуктов, за исключением Microsoft Visual Studio, распространяются на условиях свободной, открытой лицензии, команды разработчиков размещаются в недружественных зарубежных юрисдикциях. Отечественных аналогов или продуктов разработчиков команд дружественных юрисдикций не выявлено.

Важным аспектом дальнейшего развития ИИ-сообщества также является сохранение доступа к публичным репозиториям, в которых размещены открытые проекты, связанные с искусственным интеллектом, и их исходный код. Такие репозитории, как правило, поддерживают одну или несколько систем управления версиями ПО (CVS, Git, Mercurial, SVN), а также имеют функционал, связанный с отслеживанием ошибок и публикацией документации программного обеспечения. Использование репозитория при разработке ПО, в том числе проектов в области искусственного интеллекта, увеличивает надежность хранения разработанного программного

кода, удобство разработки, облегчает процессы разработки ПО с использованием распределенных команд, а также упрощает управление версиями разработанного программного обеспечения. В число основных публичных репозиторий ПО входят: Azure DevOps Services, Bitbucket, GitHub, GitLab, SourceForge. Подавляющее большинство представленных продуктов распространяются на условиях коммерческого программного обеспечения, команды разработчиков размещаются в недружественных зарубежных юрисдикциях. Существуют отечественные аналоги: GitFlic (Россия).

Таблица 1. Основные компоненты проектов в сфере ИИ и их аналоги

# Оценка вызовов

В интересах оценки рисков санкционного влияния на развитие ИИ в Российской Федерации представляется необходимым исследовать условия предоставления лицензий к ключевым компонентам проектов в сфере искусственного интеллекта, определить юрисдикцию разработчиков компонентов, выявить аналоги отечественного происхождения и аналоги, разработанные в юрисдикциях, не являющихся недружественными. По результатам такого анализа (Приложение А) не выявлены продукты отечественных разработок и источников знаний или разработчиков «дружественных» стран, в следующих категориях: системы управления нереляционными базами данных, пакеты прикладных программ для математических вычислений, фреймворки хранения и обработки больших данных, специализированные библиотеки алгоритмов, средства подготовки данных, интегрированные среды разработки.

Со времен существования Советского Союза и по настоящее время на Россию «западными» странами формируются и накладываются ограничения в разных сферах взаимодействия для сдерживания технологического развития в нашей стране. В свете развития последних геополитических событий со стороны мирового сообщества на Российскую Федерацию и ее резидентов наложили беспрецедентное количество ограничений, в том числе в сфере ИТ технологий. Накладываемые ограничения на импорт информационных технологий в Россию поднимают вопросы импортозамещения в наиболее актуальную повестку, в том числе в сфере развития ИИ. Необходимо оценить риски влияния санкционного воздействия на развитие ИИ в России – определить степень риска.

Для данных целей в рамках исследования была разработана модель оценки риска влияния санкционного воздействия на развитие ИИ в России. Модель оценки риска влияния санкционного воздействия на развитие ИИ в России сформирована на основе методов экспертных и рейтинговых оценок. Для формирования модели оценки риска выделены критерии (зависимости) санкционного воздействия на развитие ИИ в России, в наибольшей степени влияющие на возможность введения ограничений и запретов: юрисдикция, в которых находится разработчик программного обеспечения; типы лицензий программного обеспечения; тип программного обеспечения. В качестве дополнительной оценки влияния санкционного воздействия предложено использовать зависимость российских разработок от применения зарубежных технологий – объем (или долю) применения тех или иных разработок в сфере ИИ.

Разработана шкала ранжирования критериев для оценки риска, в зависимости от совокупности критериев используемому программному обеспечению предлагается установить градацию ранжирования критериев, представленную в таблице 1.

Таблица 1. Модель ранжирования совокупности критериев оценки риска вероятности применения ограничительных мер (санкций) на программное обеспечение в сфере ИИ.

ТИПЫ ПРОГРАММНОГО ПРОДУКТА	ТИПЫ ЛИЦЕНЗИЙ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ					ЮРИСДИКЦИЯ РАЗРАБОТЧИКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ
	Коммерческое ПО	Условно бесплатное ПО	Свободно распространяемое ПО	Open source ПО	Public domain ПО	
Оконечный (готовый) продукт	10	8	6	4	4	Страны группы высокого риска
	8	6	4	2	2	Страны группы среднего риска
	5	3	2	1	1	Страны группы низкого риска
Сопутствующий продукт	9	7	5	3	3	Страны группы высокого риска
	6	4	2	1	1	Страны группы среднего риска
	3	2	1	1	1	Страны группы низкого риска
Элементы ИИ требующие доработки (настройки)	7	5	3	1	1	Страны группы высокого риска
	3	2	1	1	1	Страны группы среднего риска
	3	2	1	1	1	Страны группы низкого риска

В качестве весовых коэффициентов для оценки риска влияния применения ограничительных мер (санкций) предлагается использовать значение доли использования (применения) тех или иных продуктов, подпадающих под критерии.

Оценку риска влияния применения ограничительных мер (санкций) на рынок используемого программного обеспечения с элементами ИИ предлагается рассчитывать по формуле:

$$K_{\text{риск}} = \sum_{i=1}^n D_i * P_i,$$

где  $K_{\text{риск}}$  – коэффициент риска,

$D_i$  – доля программного обеспечения соответствующей совокупности критериев оценки, используемого на российском рынке;

$P_i$  – коэффициент ранжирования соответствующей совокупности критериев оценки риска влияния ограничительных мер (санкций) на рынок используемого программного обеспечения с элементами ИИ;

$n$  – количество рассчитанных долей программного обеспечения соответствующей совокупности критериев оценки, используемого на российском рынке;

$i$  – порядковый номер рассчитанной доли программного обеспечения соответствующей совокупности критериев оценки, используемого на российском рынке.

При таком расчёте коэффициент риска будет принимать значение от 1 до 10. Предлагается определить следующую градацию (диапазоны) оценки риска влияния ограничительных мер (санкций) на рынок используемого программного обеспечения с элементами ИИ в зависимости от коэффициента риска:

$K_{\text{риск}} \geq 6$  – высокий риск влияния ограничительных мер (санкций);

$3 \leq K_{\text{риск}} < 6$  – средний риск влияния ограничительных мер (санкций);

$K_{\text{риск}} < 3$  – низкий риск влияния ограничительных мер (санкций).

Для оценки векторов развития политики импортозамещения в сфере программного обеспечения с элементами ИИ в Российской Федерации необходимо определить аспекты (критерии) влияния ограничительных мер (санкций), с которыми связаны наиболее высокие риски.

В рамках исследования проведены расчеты для оценки риска по трем различным базовым сценариям: оптимистичный, нейтральный, пессимистичный. В рамках сценарного анализа использовалось допущение, что наибольшему риску подвержен российский рынок программного обеспечения с элементами ИИ, имеющими зависимость от продуктов, разрабатываемых организациями, расположенными в юрисдикции стран, входящих в группу высокого риска, с примерно равным распределением по типу ПО. В оптимистичном сценарии использования в разработках преимущественно программного обеспечения с элементами ИИ, предоставляемых по типу лицензирования Open source, разработчиков, находящихся в юрисдикции «западных» стран, коэффициент риска воздействия (5,85) соответствует значениям среднего риска влияния ограничительных мер (санкций) ( $3 \leq K_{\text{риск}} < 6$ ). В нейтральном и пессимистичном сценарии коэффициент риска воздействия (5,67 и 7,65 соответственно) соответствует значениям высокого риска влияния ограничительных мер (санкций) ( $K_{\text{риск}} \geq 6$ ).

Таким образом, требуется принятие комплексных мер как оперативного характера, так и последующих среднесрочных мер снижения зависимости российских разработок в сфере ИИ от влияния применения ограничительных мер (санкций) и/или снижения такого влияния.

# Предложения по снижению риска влияния

## Меры поддержки развития ИИ в России

В настоящее время в Российской Федерации со стороны государства активно проводятся мероприятия, направленные на стимулирование и развитие применения разработок в сфере ИТ, в том числе в части развития ИИ. Принят ряд основополагающих документов: Национальная стратегия развития искусственного интеллекта на период до 2030 года; утвержден федеральный проект «Искусственный интеллект»; утверждены и определены директивы<sup>26</sup>, направленные на использование российского ПО в государственных организациях и компаниях, принят существенный пакет мер поддержки ИТ-отрасли; выделяются гранты и субсидии на цифровую трансформацию, внедрение отечественных цифровых решений, на возмещение части затрат на разработку цифровых платформ и программных продуктов; принимаются нормы о запрете использования иностранного программного обеспечения государственными органами и корпорациями.

В целях развития технологий в сфере ИИ в рамках мероприятий федерального проекта «Искусственный интеллект» выделены гранты и субсидии, направленные на:

1. Поддержку исследовательских центров в сфере искусственного интеллекта, в том числе в области «сильного» искусственного интеллекта, систем доверенного искусственного интеллекта и этических аспектов применения искусственного интеллекта. Получателями поддержки в 2021 году были определены 6 организаций<sup>27</sup>.
2. Поддержку пилотных проектов апробации технологий искусственного интеллекта в приоритетных отраслях<sup>28</sup>. Победителями конкурса в 2021 году признаны 9 организаций на общую сумму более 650 млн рублей<sup>29</sup>. Открыт прием заявок на конкурс в текущем году<sup>30</sup>.
3. Предоставление субсидий из федерального бюджета российским организациям на финансовое обеспечение части затрат на разработку конкурентоспособных нишевых аппаратно-программных комплексов для целей искусственного интеллекта<sup>31</sup>.
4. Поддержку малых предприятий по разработке, применению и коммерциализации продуктов, сервисов и (или) решений с использованием технологий искусственного интеллекта, разработчиков открытых библиотек в сфере искусственного интеллекта, акселерации проектов с применением искусственного интеллекта. Победителями конкурса в 2021 году признаны 18 проектов на общую сумму более 330 млн рублей<sup>32</sup>. Открыт прием заявок на конкурс в текущем году<sup>33</sup>.
5. На разработку программ бакалавриата и программ магистратуры по профилю «искусственный интеллект», а также на повышение квалификации педагогических работников образовательных организаций высшего образования в сфере искусственного интеллекта<sup>34</sup>.

<sup>26</sup> URL: <https://digital.gov.ru/ru/documents/8013/>

<sup>27</sup> URL: [https://ac.gov.ru/uploads/Projects/AI\\_support/%E2%84%961.pdf](https://ac.gov.ru/uploads/Projects/AI_support/%E2%84%961.pdf)

<sup>28</sup> URL: <https://sk.ru/news/fond-skolkovo-stal-operatorom-gosudarstvennoj-podderzhki-aprobacii-rossijskih-tehnologij-iskusstvennogo-intellekta/>

<sup>29</sup> URL: <https://dtech.sk.ru/files/78/promezhutochnye-itog0.pdf>

<sup>30</sup> URL: <https://dtech.sk.ru/iskusstvennyj-intelekt/>

<sup>31</sup> URL: <https://gisp.gov.ru/support-measures/list/12447441/>

<sup>32</sup> URL: <https://fasie.ru/competitions/razvitie-ai-2-results/>

<sup>33</sup> URL: <https://fasie.ru/press/fund/kod-ai-3/>

<sup>34</sup> URL: [https://minobrnauki.gov.ru/documents/?ELEMENT\\_ID=38934](https://minobrnauki.gov.ru/documents/?ELEMENT_ID=38934)

Кроме этого, проводятся тематические хакатоны<sup>35</sup>, стимулируется развитие соответствующих практико-ориентированных образовательных программ по искусственному интеллекту<sup>36</sup>, увеличивается количество бюджетных мест в российских вузах по специальностям, связанным с разработкой и развитием технологий ИИ<sup>37</sup>, ведутся работы в области стандартизации, создан и активно работает технический комитет по стандартизации (ТК 164) «Искусственный интеллект»<sup>38</sup> и его рабочие группы, разработан кодекс этики в сфере ИИ, к которому присоединились более 40 крупных организаций и ИТ-компаний<sup>39</sup>.

## Предложения по нивелированию риска влияния санкций в сфере ИИ

В качестве итога данного исследования проблематики снижения зависимости риска влияния применения ограничительных мер (санкций) на развитие рынка продуктов и разработок в сфере ИИ в России, предлагается проработать следующие направления (предложения) в краткосрочной (мероприятия, требующие оперативного исполнения), среднесрочной и долгосрочной перспективе:

1. В краткосрочной перспективе предлагается:

- а) в связи с возрастающими рисками ограничения доступа к различным тематическим репозиториям Open Source программного обеспечения, библиотекам, базам знаний и т.д., в краткосрочной перспективе необходимо оперативно организовать мероприятия по созданию российской платформы Open Source продуктов (репозитория в рунете) и наполнением такого репозитория доступными Open Source продуктами и программным обеспечением, переносимых (копируемых) из публичных репозиториях иностранной юрисдикции, в первую очередь стран объявивших санкции в отношении России, - Azure DevOps Services, Bitbucket, GitHub, GitLab, SourceForge. По информации CNews Сбербанк уже занимается созданием репозитория открытого ПО в рамках работы над «Гостехом», имеет контракт с Github<sup>40</sup>. При этом необходимо разработать и реализовать механизм валидации (тестирования) Open Source продуктов, переносимых из публичных репозиториях иностранной юрисдикции, на предмет включения в них зловредного кода, так как возникают тренды создания так называемого «протестного ПО»;
- б) в краткосрочной перспективе предлагается также разработать публичный реестр технологий, необходимых для разработки в области ИИ, и осуществлять мониторинг рисков и процесса импортозамещения в данной области. На данный момент разработчики ИИ используют устоявшийся стек иностранных технологий в процессе разработки и не имеют мотивации для перехода на отечественные технологии. В то же время отечественные технологические компании разрозненно работают над созданием различных отечественных аналогов решений в области ИИ и смежных областях, однако их решения часто остаются незамеченными и нишевыми. Для решения данных проблем предлагается сформировать открытый реестр импортозамещения технологий в области разработки ИИ с возможностью размещения публичных отзывов и оценок пользователями. В рамках данного реестра предлагается вести учет отечественных разработок в сравнении с их иностранными исторически популярными аналогами. Данный реестр также должен

<sup>35</sup> URL: <https://rsv.ru/competitions/contests/1/180/>

<sup>36</sup> URL: <https://ai.2035.university/fp>

<sup>37</sup> URL: <https://tass.ru/obschestvo/13018437>

<sup>38</sup> URL: <https://www.tc164.ru/>

<sup>39</sup> URL: <https://a-ai.ru/code-of-ethics/#footer>

<sup>40</sup> URL: [https://www.cnews.ru/news/top/2021-09-15\\_natsionalnyj\\_repozitorij](https://www.cnews.ru/news/top/2021-09-15_natsionalnyj_repozitorij)



содержать ссылки на пробный доступ к отечественным решениям или на демонстрацию решения.

Такой реестр позволит пользователям в лице разработчиков ИИ:

- следить за актуальными аналогами иностранных решений и оперативно находить им замену среди отечественных разработок;
- быстро знакомиться с отечественными решениями через демо-доступ или демонстрацию;
- оценивать отечественные решения и оставлять отзывы к ним, что позволит повысить доверие к отечественным решениям среди пользователей;

Компании – разработчики решений получат возможность:

- размещать свои решения в реестре для их популяризации и привлечения новых клиентов;
- формировать доверие к своим продуктам и в то же время получать честный фидбек и дорабатывать решения благодаря системе отзывов.

Рынок решений в области ИИ и анализа данных получит стимулирование внутренней конкуренции, которая будет стимулировать развитие отечественных решений.

Контрольно-надзорные органы смогут осуществлять мониторинг процесса импортозамещения в сфере ИИ в режиме реального времени, а также осуществлять продвижение отдельных решений, использование которых стимулируется государством по тем или иным причинам.

Важно отметить, что система отзывов должна быть сформирована с реализацией принципов защиты от «фейковых» отзывов. Для этого предлагается проводить верификацию пользователей решений, чтобы оставить отзыв в реестре.

Реестр предлагается реализовывать в некоммерческом формате на базе специально созданной или действующей АНО, которая будет являться оператором реестра, с государственным финансированием создания и поддержки функционирования реестра.

2. В среднесрочной перспективе предлагается:

- а) В первую очередь активизировать работы и ускорить создание национального портала в сфере искусственного интеллекта с формированием подразделов: репозитория Open Source программного обеспечения, используемого при разработках продуктов с применением технологий ИИ, в том числе различного сопутствующего ПО; библиотек алгоритмов, датасетов; базы знаний (научных статей, докладов, обмен мнениями участников сообщества и т.д.); опыта применения (текущих практик); информации о мерах поддержки развития сферы ИИ (гранты, субсидии, льготы и т.д.).

Необходима также проработка вопроса создания отечественных веб-ресурсов, являющихся зеркалами международных научных веб-ресурсов.

Для успешного развития технологий искусственного интеллекта представляется необходимой организация площадки для обсуждения пользователями вопросов, возникающих на различных этапах реализации проектов в сфере информационных технологий, ориентированной на российских программистов, аналитиков данных, экспертов в области ИИ.

- б) Активизировать работы по созданию организации, направленной на объединение разработчиков программного обеспечения с открытым кодом и стимулирование таких разработок - Russian Open Source Foundation.
- в) Продолжить предоставление мер поддержки стимулирования разработки ПО с применением технологий ИИ и обучение специалистов в данной сфере.
- г) Проводить регулярный анализ ключевых компонентов, применяемых при разработке продуктов с технологиями ИИ, в целях выявления имеющихся российских разработок, эквивалентных

иностранному программному обеспечению, определения ниш (сфер и направлений) в которых отсутствуют российские разработки («белые пятна»), организовать ведение соответствующего реестра ПО с его постоянной актуализацией и обеспечить первоочередное стимулирование мерами государственной поддержки (выделение грантов, субсидии, льгот и т.п.) на разработку Open Source программного обеспечения в этих направлениях – закрытие таких «белых пятен».

- д) Разработать дополнительную меру поддержки развития и применения разработок с технологиями ИИ – обеспечить со стороны государства выкуп коммерческого ПО российских разработчиков, применяемого в качестве ключевых компонентов при разработке продуктов с технологиями ИИ, и перевод его в Open Source с последующей публикацией в соответствующем репозитории и предоставлением доступа для использования российскими разработчиками.
- е) Актуализировать нормативные правовые акты в целях стимулирования обеспечения доступа и публикации государственных данных (датасетов), обрабатываемых в органах государственной власти в рамках оказания государственных услуг и/или выполнения государственных функций, в том числе в библиотеке (репозитории) датасетов (данных) национального портала в сфере искусственного интеллекта;
- ж) Предлагается разработать платформу для проведения соревнований в области ИИ. В рамках данного проекта предлагается создать платформу (включая встроенную среду разработки) для:
  - проведения некоммерческих соревнований в области машинного обучения и анализа данных;
  - рейтингования участников по результатам проведения соревнований с целью выявления наиболее квалифицированных кадров;
  - обучения специалистов в соответствующих областях;
  - публикации пользовательского кода для обмена опытом разработчиков;
  - публикации датасетов для машинного обучения.

Такой подход отлично себя зарекомендовал в реализации иностранного портала [kaggle.com](https://www.kaggle.com), который сформировал сообщество более 9 миллионов разработчиков в области ИИ и анализа данных, а также сформировал базу более 50 тыс. датасетов и 400 тыс. шаблонов кода (в формате Jupyter Notebooks). Высокий рейтинг на платформе [Kaggle.com](https://www.kaggle.com) также является общепризнанным показателем компетентности специалиста в области ИИ и анализа данных, на который обращают внимание при отборе кандидатов на вакансии в области ИИ и анализа данных.

Предлагается также дополнить платформу специальной площадкой для проведения хакатонов и соревнований в области DS для коммерческих заказчиков. На данный момент ежегодно проводятся десятки хакатонов, в том числе более 100 хакатонов запланированы до 2024 года в рамках федерального проекта «Искусственный интеллект», однако организаторы традиционно используют разные площадки и решения для организации хакатонов. Проведение соревнований на разных площадках ведет к:

- дополнительным издержкам на адаптацию участников к формату мероприятия, отвлекая от непосредственного решения задач;
- формированию искусственных барьеров для обучения сообщества на лучших практиках, т.к. решения, предложенные участниками, остаются непубличными или теряются по мере прохождения мероприятий;
- не позволяет централизованно учитывать успехи специалистов в области ИИ и анализа данных.

Проведение хакатонов и соревнований с возможностью гибкой настройки, однако в рамках единой платформы, позволит преодолеть перечисленные выше недостатки и барьеры.

В дальнейшем, по мере развития отечественных аналогов ПО для разработки в области ИИ, данная платформа может также использоваться для органичного продвижения отечественных технологий в сообщество разработчиков через проведение соревнований с использованием таких технологий.

Платформу предлагается формировать на базе независимой коммерческой организации для предоставления максимальной гибкости развития платформы и исключения излишней бюрократизации данного проекта. В то же время предлагается организовать государственную поддержку создания и развития проекта, в части финансирования и популяризации решения.

3. В долгосрочной перспективе предлагается:

- а) Развивать международное взаимодействие со странами, активно развивающими технологии в сфере ИИ и не применяющими политику односторонних санкций – Китай, Индия, Иран, ОАЭ, Казахстан, по обмену Open Source программным обеспечением и публикацией в соответствующем репозитории национального портала в сфере искусственного интеллекта.
- б) Организовать международное взаимодействие со странами, активно развивающими технологии в сфере ИИ и не применяющими политику односторонних санкций, в сфере научных публикаций (взаимопризнаваемые публикации научных статей, докладов, проведение конференций и т.п.), опыта разработки, внедрения и применения технологий ИИ – пополнение базы знаний национального портала в сфере искусственного интеллекта.
- в) Экспорт российских технологий иностранным партнерам на коммерческих условиях со стороны государства и/или Russian Open Source Foundation с направлением получаемых доходов (выручки) на финансирование мер государственной поддержки (возмещение расходов государства на развитие ИИ) – создание международных тематических групп по разработке и внедрению технологий ИИ, научных обществ; узкопрофильных тематических рабочих групп, в том числе по вопросам стандартизации, и т.д.

# Приложение А

## Основные компоненты проектов в сфере ИИ и их аналоги

Название продукта	Функционал	Тип лицензии	Происхождение	Аналоги (отечественные, дружественных юрисдикций)
<b>Системы управления базами данных</b>				
MySQL	Система управления реляционными базами данных	GPLv2 (свободн*ая) Проприетарная (для коммерческой версии с расширенным функционалом)	США, Oracle Corporation. Oracle приостановили работу на российском рынке	Линтер (Россия) Ред База Данных (Россия) SoQol (Россия) Arenadata DB (Россия) Postgres Pro (Россия)
PostgreSQL	Система управления реляционными базами данных	PostgreSQL License (свободная, открытая)	PostgreSQL Global Development Group, распределенная команда	Линтер (Россия) Ред База Данных (Россия) SoQol (Россия) Arenadata DB (Россия) Postgres Pro (Россия)
MariaDB	Система управления реляционными базами данных	GPLv2 (свободная)	MariaDB Corporation, Финляндия	Линтер (Россия) Ред База Данных (Россия)*8 SoQol (Россия) Arenadata DB (Россия) Postgres Pro (Россия)
Neo4j	Система управления нереляционными графовыми базами данных	GPLv3 (свободная), APGLv3 (свободная) для базовой версии. Расширения имеют закрытый код и распространяются по коммерческой лицензии	Neo4j Inc., США	NitrosBase (Россия)
MongoDB	Документоориентированная система управления нереляционными базами данных	Server Side Public Licence (свободная)	MongoDB Inc., США. Приостановили работу на российском рынке	ClickHouse (Россия) Tarantool (Россия) NitrosBase (Россия)

Название продукта	Функционал	Тип лицензии	Происхождение	Аналоги (отечественные, дружественных юрисдикций)
Redis	Система управления нереляционными базами данных	BSD (свободная)	Redis Inc., США	ClickHouse (Россия) Tarantool (Россия) NitrosBase (Россия)
<b>Статистическое и математическое программное обеспечение</b>				
SAS	Пакет статистического программного обеспечения	Проприетарная коммерческая	SAS Institute, США	Polymatica (Россия) Статэкс (Россия)
SPSS	Пакет статистического программного обеспечения	Проприетарная коммерческая	IBM, США. Приостановили работу на российском рынке	Polymatica (Россия) Статэкс (Россия)
Matlab	Пакет прикладных программ для математических вычислений	Проприетарная коммерческая	MathWorks, США. Приостановили работу на российском рынке	SimInTech (Россия. Продукт ориентирован на моделирование технических систем) nanoCAD BIM (Россия. Продукт ориентирован на моделирование технических систем) FlowVision (Россия. Продукт ориентирован на моделирование аэро- и гидродинамики)
<b>Фреймворки хранения и обработки больших данных</b>				
Apache Hadoop	Фреймворк программного обеспечения для хранения и обработки больших данных	Apache License 2.0 (свободная, открытая)	Apache Software Foundation, США	Не выявлены*
Apache Spark	Фреймворк распределенной обработки больших данных	Apache License 2.0 (свободная, открытая)	Apache Software Foundation, США	Не выявлены*

Название продукта	Функционал	Тип лицензии	Происхождение	Аналоги (отечественные, дружественных юрисдикций)
-------------------	------------	--------------	---------------	---

\* на Российском рынке существуют аналоги от Аренадаты<sup>41</sup> и VK Cloud Soluitions<sup>42</sup>, но в основе этих платформ также используется Apache Hadoop и Apache Spark

#### Специализированные библиотеки алгоритмов

TensorFlow	Программная библиотека для конструирования и тренировки искусственных нейронных сетей	Apache License 2.0 (свободная, открытая)	Google, США. Приостановили работу некоторых сервисов в РФ	Не выявлены
PyTorch	Программная библиотека для машинного обучения. Адаптация библиотеки Torch для языка программирования Python	BSD (открытая, свободная)	Facebook AI Research Lab. Входит в корпорацию Meta, некоторые сервисы которой заблокированы на территории РФ	Не выявлены
Scikit-learn	Программная библиотека для машинного обучения для языка программирования Python	BSD (открытая, свободная)	Распределенная команда представителей сообщества	Не выявлены
matplotlib	Библиотека средств визуализации данных для языка программирования Python	PSF-based (открытая, свободная)	Распределенная команда представителей сообщества	Не выявлены
Seaborn	Библиотека средств визуализации данных для языка программирования Python	BSD (открытая, свободная)	Распределенная команда представителей сообщества	Не выявлены

<sup>41</sup> URL: <https://arenadata.tech/products/arenadata-hadoop/>

<sup>42</sup> URL: <https://mcs.mail.ru/bigdata/>

<b>Название продукта</b>	<b>Функционал</b>	<b>Тип лицензии</b>	<b>Происхождение</b>	<b>Аналоги (отечественные, дружественных юрисдикций)</b>
NLTK	Библиотека средств обработки естественного языка для языка программирования Python	Apache License 2.0 (свободная, открытая)	Распределенная команда представителей сообщества	Не выявлены
CNTK	Библиотека средств обработки естественного языка для языков программирования Python, C#, C++	MIT (свободная, открытая)	Microsoft, США. Приостановили работу на российском рынке	Не выявлены
Theano	Библиотека средств оптимизации вычислений для языка программирования Python	BSD (открытая, свободная)	Университет Монреаля, Канада	Не выявлены
Caffe	Библиотека средств глубокого обучения для языка программирования Python	BSD (открытая, свободная)	Калифорнийский университет в Беркли, США	Не выявлены
Keras	Библиотека средств глубокого обучения для языка программирования Python	Apache License 2.0 (свободная, открытая)	Распределенная команда представителей сообщества	Не выявлены
Spark Mlib	Библиотека средств машинного обучения для языков программирования Python и R	Apache License 2.0 (свободная, открытая)	Apache Software Foundation, США	Не выявлены
MLPack	Библиотека средств глубокого обучения для языка	BSD (открытая, свободная)	Распределенная команда	Не выявлены

Название продукта	Функционал	Тип лицензии	Происхождение	Аналоги (отечественные, дружественных юрисдикций)
	программирования C++		представителей сообщества	
<b>Средства визуализации данных</b>				
Tableau	Программное обеспечение для интерактивной визуализации данных и бизнес-аналитики	Проприетарная коммерческая	Salesforce, США. Некоторые продукты компании стали недоступны для пользователей из РФ	Visiology (Россия) N3.Аналитика (Россия) Visary BI (Россия) Форсайт Аналитическая Платформа (Россия) Интеград Аналитика (Россия) Криста BI (Россия) Digital Q.Reporting (Россия) FlyBI (Россия) PolyAnalyst (Россия)
Qlik Sense	Облачное программное обеспечение для интеграции, аналитики и визуализации данных	Проприетарная коммерческая	QlikTech International AB, Швеция	Visiology (Россия) N3.Аналитика (Россия) Visary BI (Россия) Форсайт Аналитическая Платформа (Россия) Интеград Аналитика (Россия) Криста BI (Россия) Digital Q.Reporting (Россия) FlyBI (Россия) PolyAnalyst (Россия)
PowerBI	Программное обеспечение для бизнес-анализа и визуализации данных	Проприетарная коммерческая	Microsoft, США. Приостановили работу на российском рынке	Visiology (Россия) N3.Аналитика (Россия) Visary BI (Россия) Форсайт Аналитическая Платформа (Россия) Интеград Аналитика (Россия) Криста BI (Россия) Digital Q.Reporting (Россия) FlyBI (Россия) PolyAnalyst (Россия)
Google Charts	Веб-сервис визуализации данных	Проприетарная коммерческая	Google, США. Приостановили работу некоторых сервисов в РФ	Visiology (Россия) N3.Аналитика (Россия) Visary BI (Россия) Форсайт Аналитическая Платформа (Россия) Интеград Аналитика (Россия)



Название продукта	Функционал	Тип лицензии	Происхождение	Аналоги (отечественные, дружественных юрисдикций)
				Криста BI (Россия) Digital Q.Reporting (Россия) FlyBI (Россия) PolyAnalyst (Россия)
<b>Средства подготовки данных</b>				
BigML	Комплекс инструментов подготовки данных для машинного обучения, обучения, развертывания и автоматизации моделей	Проприетарная коммерческая	BigML Inc., США	PolyAnalyst (Россия)
Trifacta Wrangler	Комплекс инструментов подготовки данных для машинного обучения и аналитики	Проприетарная коммерческая	Trifacta, США	PolyAnalyst (Россия)
Knime	Комплекс инструментов подготовки и интеграции данных для машинного обучения	GNU (открытая, свободная)	KNIME AG, Швейцария	PolyAnalyst (Россия)
<b>Средства поиска и публикации наборов больших данных</b>				
Google Dataset Search	Поисковая система, ориентированная на поиск открытых больших данных	Бесплатное использование. Формальная лицензия не опубликована	Google, США. Приостановили работу некоторых сервисов в РФ	Инфраструктура научно-исследовательских данных (ИНИД) Центра перспективных управленческих решений (ЦПУР) (Россия) Университетский консорциум исследователей больших данных (Россия) Яндекс.Толока – открытые датасеты (Россия) Хаб открытых данных АНО «Информационная

Название продукта	Функционал	Тип лицензии	Происхождение	Аналоги (отечественные, дружественных юрисдикций)
Kaggle	Система поиска и публикации датасетов, онлайн-сообщество специалистов в области машинного обучения и data science	Проприетарная лицензия, бесплатное использование	Google, США. Приостановили работу некоторых сервисов в РФ	культура» (Россия) Отраслевой реестр датасетов АО «Концерн Росатом» (Россия) Каталог каталогов данных datacatalogs.ru АНО «Инфокультура» (Россия)
Hugging Face	Система поиска и публикации датасетов и моделей машинного обучения, документации на ИИ-решения	Проприетарная лицензия, бесплатное использование с ограниченным функционалом. Платный доступ к расширенному функционалу	Hugging Face Inc., США	Инфраструктура научно-исследовательских данных (ИНИД) Центра перспективных управленческих решений (ЦПУР) (Россия) Университетский консорциум исследователей больших данных (Россия) Яндекс.Толока – открытые датасеты (Россия) Хаб открытых данных АНО «Информационная культура» (Россия) Отраслевой реестр датасетов АО «Концерн Росатом» (Россия) Каталог каталогов данных datacatalogs.ru АНО «Инфокультура» (Россия)

Название продукта	Функционал	Тип лицензии	Происхождение	Аналоги (отечественные, дружественных юрисдикций)
				datacatalogs.ru АНО «Инфокультура» (Россия)
Amazon Datasets	Реестр и система поиска датасетов больших данных	Проприетарная лицензия, бесплатное использование	Amazon, США. Приостановили работу на российском рынке	Инфраструктура научно-исследовательских данных (ИНИД) Центра перспективных управленческих решений (ЦПУР) (Россия) Университетский консорциум исследователей больших данных (Россия) Яндекс.Толока – открытые датасеты (Россия) Хаб открытых данных АНО «Информационная культура» (Россия) Отраслевой реестр датасетов АО «Концерн Росатом» (Россия) Каталог каталогов данных datacatalogs.ru АНО «Инфокультура» (Россия)
Azure Datasets	Реестр и система поиска датасетов больших данных	Проприетарная лицензия, бесплатное использование	Microsoft, США. Приостановили работу на российском рынке	Инфраструктура научно-исследовательских данных (ИНИД) Центра перспективных управленческих решений (ЦПУР) (Россия) Университетский консорциум исследователей больших данных (Россия) Яндекс.Толока – открытые датасеты (Россия) Хаб открытых данных АНО «Информационная культура» (Россия) Отраслевой реестр датасетов АО «Концерн Росатом» (Россия) Каталог каталогов данных datacatalogs.ru АНО «Инфокультура» (Россия)

### Интегрированные среды разработки

Название продукта	Функционал	Тип лицензии	Происхождение	Аналоги (отечественные, дружественных юрисдикций)
NetBeans	Интегрированная среда разработки для языков программирования Java, JavaScript, Python, PHP, C, C++ и других	Apache License 2.0 (свободная, открытая)	Apache Software Foundation, США	<p>Не выявлены на настоящий момент.</p> <p>В марте текущего года ФСТЭК был проведен открытый конкурс на выполнение работ по разработке унифицированной среды безопасной разработки отечественного программного обеспечения для языков программирования Java, C, C++. По результатам конкурса исполнителем работ был определен Институт системного программирования им. В.П. Иванникова РАН. Ввод унифицированной среды в промышленную эксплуатацию запланирован на 2024 год</p>
PyCharm	Интегрированная среда разработки для языка программирования Python	<p>Apache License 2.0 (свободная, открытая)</p> <p>Проприетарная коммерческая лицензия для профессиональной версии с расширенным функционалом</p>	JetBrains, Чехия. Приостановили работу на российском рынке	Не выявлены
Eclipse	Интегрированная среда разработки для языка программирования Java. При помощи модулей расширения поддерживает ряд других языков программирования	Eclipse Public License (свободная, открытая)	Eclipse Foundation AISBL, Канада	<p>Не выявлены на настоящий момент.</p> <p>В марте текущего года ФСТЭК был проведен открытый конкурс на выполнение работ по разработке унифицированной среды безопасной разработки отечественного программного обеспечения для языков программирования Java, C, C++. По результатам</p>

Название продукта	Функционал	Тип лицензии	Происхождение	Аналоги (отечественные, дружественных юрисдикций)
				конкурса исполнителем работ был определен Институт системного программирования им. В.П. Иванникова РАН. Ввод унифицированной среды в промышленную эксплуатацию запланирован на 2024 год
Microsoft Visual Studio	Интегрированная среда разработки для языков программирования C, C++, C#, F#, JavaScript. При помощи модулей расширения поддерживает ряд других языков программирования	Проприетарная коммерческая	Microsoft, США. Приостановили работу на российском рынке	Не выявлены на настоящий момент. В марте текущего года ФСТЭК был проведен открытый конкурс на выполнение работ по разработке унифицированной среды безопасной разработки отечественного программного обеспечения для языков программирования Java, C, C++. По результатам конкурса исполнителем работ был определен Институт системного программирования им. В.П. Иванникова РАН. Ввод унифицированной среды в промышленную эксплуатацию запланирован на 2024 год

#### Средства совместной разработки и хранения программного кода

Azure DevOps Services	Набор облачных инструментов совместной работы над программным кодом, включая возможность развертывания программных репозиторийев	Проприетарная коммерческая	Microsoft, США. Приостановили работу на российском рынке	GitFlic (Россия)
-----------------------	--	----------------------------	--	------------------

Название продукта	Функционал	Тип лицензии	Происхождение	Аналоги (отечественные, дружественных юрисдикций)
Bitbucket	Сервис хостинга и совместной разработки программного кода и ИТ-проектов	Проприетарная коммерческая	Atlassian, Австралия. Приостановили работу на российском рынке	GitFlic (Россия)
GitHub	Сервис хостинга и совместной разработки программного кода и ИТ-проектов	Проприетарная коммерческая	Microsoft, США. Приостановили работу на российском рынке, однако сервис GitHub продолжает работу	GitFlic (Россия)
GitLab	Программное обеспечение и веб-сервис для управления репозиториями программного обеспечения и совместной разработки	MIT (открытая, свободная)  Проприетарная коммерческая для профессиональной версии с расширенным функционалом	GitLab Inc., США. Приостановили работу на российском рынке	GitFlic (Россия)
SourceForge	Веб-сервис управления разработкой программного обеспечения и репозиториями исходного кода. Сервис работает только с открытым программным обеспечением	Проприетарная коммерческая	Slashdot Media, США	GitFlic (Россия)



© 2022 Фонд «Центр стратегических разработок» (ЦСР).  
Все права защищены.  
При использовании информации из документа ссылка на ЦСР  
обязательна.

Москва, 125009, Газетный пер., 3–5 стр. 1, 3 этаж  
Тел.: +7 (495) 725-78-06  
Факс: +7 (495) 725-78-14  
E-mail: [info@csr.ru](mailto:info@csr.ru)  
[csr.ru](http://csr.ru)



© 2022 Федеральное государственное бюджетное  
образовательное учреждение высшего образования  
«Московский государственный университет  
имени М.В. Ломоносова». Центр компетенций НТИ по  
технологиям хранения и анализа больших данных (МГУ). Все  
права защищены.  
При использовании информации из документа ссылка на МГУ  
обязательна.  
Москва, 119192, Ломоносовский пр-т., 27 корп. 1, офисы  
E801-804, A818  
Тел.: +7 (903) 186-47-28  
E-mail: [alexander.biryukov@digital.msu.ru](mailto:alexander.biryukov@digital.msu.ru)  
[bigdata.msu.ru](http://bigdata.msu.ru)