

УДК [004.85:004.82].032.26:004.6-022.225

Е.В. Мельникова

## Глубокое машинное обучение в оптимизации научно-исследовательской деятельности\*

*Дана общая характеристика машинного обучения – подобласти искусственного интеллекта. Раскрыто содержание процесса глубокого обучения, представлены его основные черты и особенности как технологии искусственного интеллекта высокого уровня. Проанализированы отличия глубокого обучения от обычного машинного. Рассмотрена архитектура моделей глубокого обучения. Изложены вопросы применения глубокого обучения в искусственных нейронных сетях, отражены основные процессы в работе нейросетей. Отмечена значимость нейросетей глубокого обучения для обработки больших данных. Приведены конкретные примеры применения алгоритмов глубокого обучения в различных областях науки, включая наукометрию, библиометрию, медицину, геосейсмику и др. Подчеркнута важная роль глубокого обучения в оптимизации научно-исследовательской деятельности, повышении результативности научных исследований.*

**Ключевые слова:** искусственный интеллект, глубокое обучение, машинное обучение, искусственные нейронные сети, большие данные, наукометрия, результативность науки, прогнозные модели, научные репозитарии

**DOI:** 10.36535/0548-0019-2023-02-2

### ВВЕДЕНИЕ

На современном этапе общественного развития значительная часть усилий таких научных дисциплин, как информатика, наука о данных<sup>1</sup> [1], компьютерные науки направлена на повышение производительности имеющихся аппаратных и программных средств и на разработку новых, более совершенных инструментов, которые были бы способны выполнять обработку, хранение и анализ огромных объемов/массивов информации. Эти массивы формируются за счет информации, уже накопленной в мире, и новых потоков данных, которые поступают непрерывно и стремительно. К числу средств, которые могут справиться с этими информационными потоками, относится технология больших данных (*Big Data*). Традиционные инструменты, такие как системы управления реляционными базами данных и некоторые другие, с обработкой и управлением большими данными справиться не могут.

---

\* Статья подготовлена в рамках исследования по теме FFFU-2022-0007 Государственного задания ВИНТИ РАН и при поддержке Российского фонда фундаментальных исследований – проект № 20-07-00014.

<sup>1</sup> Наука о данных (*Data Science*) исследует проблемы поиска, хранения, обработки, анализа и представления информации в цифровой форме. Основателем науки считается Петер Науэр.

Эффективным инструментом для реализации технологии *Big Data* стало глубокое машинное обучение (*Deep Learning*), которому в последние годы придается все большее значение. Основная цель применения алгоритмов глубокого обучения – это содействие в получении нового знания, обладающего большим потенциалом полезности, и на этой основе – в оптимизации деятельности в различных сферах жизни общества, включая сферу науки, повышении их результативности. Что касается науки, то в статье рассматривается применение глубокого обучения в таких ее областях, как наукометрия, геосейсмика, медицина и некоторые другие.

### ОБЩАЯ ХАРАКТЕРИСТИКА ТЕХНОЛОГИИ БОЛЬШИХ ДАННЫХ

Появление больших данных связано с экспоненциальным ростом информации в мире на рубеже XX – XXI вв. По оценкам экспертов международной информационно-аналитической корпорации *IDC* и мирового лидера в области хранения данных – компании *Seagate Technology* [2], к концу 2022 г. было сформировано порядка 80 зеттабайтов информации (рисунок). К 2025 г., как ожидается, общий объем данных в мире вырастет до 175 зеттабайтов.



Рис. 1. Рост мирового объема данных в 2010 – 2025 гг. [2]

Концепция *Big Data* включает в себя как сами большие данные<sup>2</sup>, так и особую технологию работы с ними, которая применяется в рамках автоматизированных систем обработки информации с использованием искусственного интеллекта, имеющих специфическую аппаратную платформу с большими вычислительными мощностями и соответствующее программное обеспечение.

Технология больших данных применяется для решения конкретных прикладных или теоретических задач. Конечная цель – получение новых знаний, новых, нередко неожиданных, смысловых взаимосвязей и закономерностей, обладающих большим потенциалом общественной полезности. Для науки и образования это могут быть знания о конкретном объекте исследования, его характеристиках, свойствах и связях, о появлении новых перспективных научных направлений, о результатах работы ученых, научных организаций, учебных заведений, о развитии науки в стране в целом, о востребованности научных идей, о степени влияния ученых и коллективов в научном сообществе [3]. Большие данные позволяют находить принципиально новые, более рациональные варианты ответов на вопросы, стоящие перед различными научными направлениями.

Для эффективной реализации технологии больших данных необходимо решить целый комплекс сложных задач, включая организацию соответствующих хранилищ для огромных объемов неоднородных данных, имеющих постоянный прирост, – данных из разных источников и разных форматов, структурированных и неструктурированных, качественных и количественных. Объем от 100 терабайтов и более, скорость прироста (более 100 гигабайтов в день) [4, 5] и многообразие видов и форматов данных формируют комплекс основных характеристик, позволяющих отнести данные к категории *Big Data*. Отсюда видно, что понятия «большие данные» и «большой объем данных» не тождественны.

Данные собирают в сетевых хранилищах, которые могут быть как разрозненными, так и объединенными в единую систему. Для исключения возможных потерь и искажения данных их обязательно копируют. В дальнейшем данные обрабатываются алгоритмом, написанным программистами для получения информации в удобном для человека виде.

Для реализации технологии *Big Data* важна организация взаимосвязанных процессов глубинного поиска, обработки и анализа данных, которые (процессы) объединяются в понятие *Data Mining*<sup>3</sup> [6]. Перечисленные задачи работы с большими данными решаются на основе глубокого машинного обучения, которое входит в ядро технологии *Big Data*.

### ОСНОВНЫЕ ЧЕРТЫ И ОСОБЕННОСТИ ГЛУБОКОГО ОБУЧЕНИЯ КАК ПОДОБЛАСТИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Глубокое обучение (*Deep Learning – DL*) – это разновидность обычного машинного обучения (*Machine Learning*) [7], его более высокоразвитая и продвинутая форма. Глубокое обучение, как и машинное обучение, является подобластью искусственного интеллекта.

**Содержание процесса машинного обучения.** Машинное обучение<sup>4</sup> [8] строится на принципах оптимизации исходной модели, которая является математически обобщенным представлением исследуемого массива данных. Суть машинного обучения заключается в «тренировке» модели на первоначаль-

<sup>3</sup> Дословный перевод термина *Data Mining* – «добыча данных». Внутренний смысл термина – обнаружение в сырых данных новой, ранее не известной полезной информации и знаний. Термин введен в 1991 г. Г. Пятецким–Шапиро.

<sup>4</sup> Появление машинного обучения относят к концу 1940-х – началу 1950-х гг. Впервые термин «машинное обучение» был введен в 1952 г. американским ученым-программистом Артуром Самюэлем в рамках написанной им самообучающейся компьютерной программы для игры в шашки. Основные положения были закреплены Самюэлем в 1959 г. в статье “Some Studies in Machine Learning Using the Game of Checkers”.

<sup>2</sup> По мнению специалистов, к категории *Big Data* следует относить потоки данных свыше 100 Гб в день.

ных данных для прогнозирования какого-либо события, явления или результата на основе новых данных.

Процесс машинного обучения состоит из следующих основных стадий: 1) в прогнозную модель<sup>5</sup> передаются входные данные, которые затем могут быть дополнены новыми данными; данные передаются для обучения модели; 2) модель тестируется и развертывается; 3) развернутая модель используется для решения поставленной задачи на основе новых данных путем получения от модели прогнозов по заданному тематическому направлению.

**Отличие глубокого обучения от машинного.** Машинное обучение объединяет в себе теорию и практику разработки самообучающихся компьютерных программ и представляет собой обучение компьютера обработке больших массивов данных (не больших данных) на основе методов искусственного интеллекта [9].

Глубокое обучение – это разновидность машинного обучения, которое реализуется на основе искусственных нейронных сетей и является подобластью искусственного интеллекта более высокого уровня. Модели и методы глубокого обучения формируют новую парадигму в области машинного обучения и искусственного интеллекта. В рамках глубокого обучения как более совершенного вида машинного обучения создаются более сложные и более самостоятельные обучающиеся программы [10]. В основе методологии *Deep Learning* лежат математические и вычислительные методы особой сложности, которые делают возможной обработку больших данных. Эти преимущества глубокого обучения приводят к тому, что оно приобретает все большее распространение.

Процесс обучения называется глубоким, так как структура искусственных нейронных сетей имеет определенную (часто большую) глубину, которая формируется за счет многослойности нейросетей. Глубоко обучаться могут только многослойные нейросети. Они состоят из нескольких входных, выходных и скрытых слоев. Входные данные, проходя через слои сети, обрабатываются каждым последующим слоем, приобретая все больше характеристик, способствующих решению поставленной задачи прогнозирования. Благодаря такой структуре нейросети, компьютер/алгоритм может путем собственной обработки данных обучиться, как создавать точный прогноз. В случае же обычного машинного обучения алгоритму необходимо извне сообщить, как выполнять точный прогноз, используя дополнительные сведения (например, путем получения новых данных).

Если отразить процесс более подробно, то при машинном (или контролируемом) обучении компьютер извлекает знания через извне управляемый опыт – внешнего учителя, когда человек дает программному алгоритму компьютера, обрабатывающему данные, некоторое количество конкретных примеров, гипотез для построения прогнозов и затем вручную исправ-

ляет ошибки компьютера в работе с данными; эксперты делают аналитику обработанных данных и принимают решения о дальнейших действиях.

При глубоком обучении у компьютера есть только программное обеспечение и данные, которые ему необходимо обработать без внешнего учителя, установить их свойства и связи. Компьютер на основе многоуровневых вычислений самостоятельно проектирует свои функции, учится решать проблемы, т. е. самостоятельно обучается, делает выводы, строит прогнозы и корректирует свою работу [11]. На основе самообучающихся программ компьютеры по частным случаям учатся выявлять общие смысловые взаимосвязанности и закономерности окружающего мира. В результате человек исключается при анализе данных, прогнозировании и принятии решений о необходимых действиях; компьютер самостоятельно принимает решения, исходя из собственного опыта, а не команд человека.

**Сферы применения глубокого обучения.** Методология *Deep Learning* описывает не один, а целый спектр специальных методов, алгоритмов, которые могут быть использованы для обучения сложных моделей прогнозирования, конструируемых на базе искусственных нейронных сетей – сверточных, импульсных, противоборствующих и др. Это имеет большое практическое значение: нейронные сети глубокого обучения<sup>6</sup> [12] сейчас все более широко применяются в распознавании речи, анализе изображений, нейро-машинном переводе, акустическом моделировании. Возможны приложения во многих других сферах, агрегирующих большие данные. Наиболее широко нейросетевые приложения глубокого обучения используются крупнейшими мировыми IT-корпорациями, включая *Google*, *Microsoft* [10], *Facebook*, *Baidu* и др. Применение глубокого обучения в значительной степени определяется объемом вычислительных мощностей компаний.

Что касается сферы науки, то применение алгоритмов глубокого обучения неуклонно расширяется, охватывая нейроинформатику, физику, вычислительную биологию, медицину, наукометрию, библиометрию [13], фармацевтику, астрономию и другие научные области.

**Архитектура моделей глубокого обучения** конструируется из базовых строительных блоков, которые могут гибко изменяться в зависимости от потребностей конкретного нейросетевого приложения. Различные сетевые архитектуры формируются из разного количества строительных блоков и на основе разных взаимосвязей между ними.

## ГЛУБОКОЕ ОБУЧЕНИЕ В ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЯХ

Искусственная нейронная сеть (ИНС)<sup>7</sup> [14] представляет собой систему соединенных процессоров (искусственных нейронов), которая отдаленно ими-

<sup>5</sup> Прогнозные модели создаются для обработки и интерпретации заданных массивов данных, их последующего анализа и выдачи результатов путем прогнозирования (составления прогнозов). Данные могут относиться к конкретным областям научных исследований, отраслям экономики, сферам бизнеса.

<sup>6</sup> Глубокое обучение было впервые применено к искусственным нейросетям в 2007 г. канадским ученым Дж. Хинтоном, который разработал алгоритм *Deep Learning* для многослойных нейросетей.

<sup>7</sup> Понятие искусственных нейронных сетей ввели американские ученые У. Маккалок и У. Питтс в начале 1940-х гг.

тирует сеть нервных клеток живого организма, т. е. биологическую нервную систему. Искусственные нейроны, как и биологические, принимают сигналы от одних связанных с ними процессоров и передают их другим через специальные элементы – синапсы, которые ослабляют или усиливают сигналы, поступающие к нейрону. Эксперт, задавая параметры синапсов, может получать на выходе правильные результаты преобразования входной информации. Каждый нейрон может иметь целое множество синапсов. Нейроны вместе со связанными с ними синапсами формируют нейроузлы. Нейросеть в целом состоит из *n*-го количества слоев нейроузлов.

Благодаря своей комплексной структуре искусственная нейросеть способна выполнять сложные задачи, с которыми не справляются традиционные средства. Что касается использования в ИНС алгоритмов глубокого обучения, то, как было отмечено, благодаря структуре ИНС глубокое обучение четко справляется с обработкой больших данных, поиском закономерностей в неструктурированных данных различных форматов, таких как изображения, звук, видео, текст [10]. С учетом больших возможностей глубоко-обученных нейросетей сфера их практического применения расширяется.

Такие сети используются в распознавании образов, распознавании речи, прогнозировании, обнаружении объектов (классификации и локализации изображений), нейросетевом сжатии данных и решении других практических задач в таких областях, как электроэнергетика, транспорт, здравоохранение, розничная торговля, банковско-финансовая сфера и др. Важным условием эффективной работы нейросетей глубокого обучения является их использование на вычислительных ресурсах большой мощности, способных обрабатывать *Big Data*.

## ПРИМЕНЕНИЕ ГЛУБОКОГО ОБУЧЕНИЯ В СФЕРЕ НАУКИ

Вместе со взрывным ростом общего объема данных в мире, динамично развивающаяся наука и научные исследования, как было отмечено, приводят к существенному увеличению объема *научных данных*, включая их наиболее значимую часть – данные о результатах исследований.

**В научных информационных ресурсах.** Большие научные данные концентрируются в специально создаваемых для этого хранилищах – базах данных крупных научных депозитариев, сетевых репозитариев, систем индексации и цитирования научной информации, научных банках данных [15]. В технологическое ядро таких баз и банков данных входит комплекс процедур глубокого обучения, которые, наряду с другими факторами, обеспечивают эффективность работы с большими научными данными.

Одним из видов хранилищ результатов научных исследований, работающих на основе алгоритмов глубокого обучения и искусственных нейросетей, являются крупные научные репозитории. В качестве примера можно привести репозитории, создаваемые в различных странах мира в рамках концепции открытой науки и работающие по принципу *FAIR* –

«справедливого доступа к данным» [16, 17]. Преимущество таких репозитариев заключается в том, что, помимо предоставления пользователям доступа к научным данным, они способны сохранять в рамках единой платформы и главное – обрабатывать на высоком уровне обобщения крупные и сверхкрупные массивы результатов научных исследований, проводимых учеными со всего мира. После обработки этих массивов определенным набором технологических средств на выходе могут получаться поистине бесценные научные результаты, обнаруживаться неожиданные научные взаимосвязи и закономерности, значительно сокращающие путь к научным открытиям. Перечень таких репозитариев представлен, например, на сайте Швейцарского национального научного фонда *SNF* [18], работающего в рамках концепции *FAIR*. Цель создания таких фондов и репозитариев в различных странах мира, включая США и страны Европы, заключается, помимо прочего, в организации системы управления научными данными в общемировом масштабе.

Еще один пример научных информационных ресурсов, агрегирующих и обрабатывающих огромные массивы данных о результатах научных исследований, – это цифровая репозитарная библиотека ХатиТраст (США) [19]. Она создана на основе партнерства университетов и научно-исследовательских институтов США в области совместного формирования, постоянного хранения, обработки данных и использования коллекции цифровых материалов научной направленности, изданных в США и других странах мира. В партнерство ХатиТраст, помимо американских, входят университеты, институты, библиотеки, архивы и библиотечные консорциумы Австралии, Канады, Новой Зеландии и некоторых стран Европы. Партнеры используют общую методику сбора данных, единый формат их хранения, а также организуют единообразную обработку данных с применением алгоритмов глубокого обучения. В итоге обеспечивается взаимный открытый доступ через Интернет к научным коллекциям членов партнерства и значительно расширяются возможности выявления скрытых закономерностей и новых научных знаний в огромном агрегированном партнерском массиве научных данных.

**В наукометрии и библиометрии.** Практическим примером применения алгоритмов глубокого обучения может служить Методика наукометрического исследования с использованием технологии обработки больших библиометрических данных, разработанная индийскими учеными в 2021 г. [20] и предназначенная для систем индексации и цитирования, крупных баз данных библиотек, архивов, научных репозитариев, других институций, работающих с информационными ресурсами. При разработке этой Методики был изучен огромный объем библиометрических данных за последнее десятилетие, выгруженных из мировой системы индексации и цитирования *Scopus*. Алгоритмы глубокого обучения позволили эффективно обработать входные данные, представленные в формате изображений. Для выявления скрытой полезной информации из загруженного набора данных авторы Методики Keshav Singh и Sandeep Kumar проанализировали плотность библиометрических се-

тей и интенсивность связей между публикациями, число цитирований, перекрестных цитирований и самоцитирований. На этой основе были выявлены значимые наукометрические закономерности. В итоге глубокое обучение позволило увеличить точность наукометрических оценок и в целом повысило качество научного исследования.

**В медицине.** Глубокое обучение в связке с большими данными в медицинской сфере начинает использоваться для диагностики заболеваний и разработки методик лечения различных болезней [21]. На основе анализа большого объема данных о пациентах могут быть получены прогнозы о том, чем конкретный пациент может заболеть в перспективе, к каким заболеваниям предрасположен его организм. Прогнозы, выдаваемые глубоко обученными нейросетевыми моделями, позволят своевременно принимать профилактические меры.

**В геосейсмике.** Активное применение технологической связки «большие данные – глубокое обучение – нейросети» позволяет эффективно обрабатывать и интерпретировать геофизические данные [22], анализировать состояние почв, определять наличие подземных пустот, температуру пород, обнаруживать важные геосейсмические закономерности и на этой основе прогнозировать наличие и координатное расположение перспективных месторождений, сравнивая новые участки с ранее открытыми залежами, т. е. используя последние как примеры для глубокого обучения разработанных нейросетевых моделей.

**В логистическом моделировании.** Большие данные в логистике формируются из данных о пробках на дорогах, состоянии парка машин, загрузки складов, расположении автозаправочных станций. Применение больших данных в связке с алгоритмами глубокого обучения помогает прогнозировать загрузку складов и рационально планировать грузоперевозки, оптимизировать маршруты, сокращать время доставки и расхода топлива, исключать случаи простоя транспорта. В результате эффективность логистической деятельности значительно возрастает.

Применение алгоритмов глубокого обучения для обработки больших массивов данных в вышеприведенных и многих других областях науки позволяет значительно оптимизировать научно-исследовательскую деятельность: более эффективно анализировать имеющиеся научные данные, получать неожиданные новые закономерности и знания, определять перспективные темы научных исследований, выявлять новые тенденции в рамках конкретного научного направления, обозначать общественно востребованные направления будущих исследований. Глубокое обучение приводит к повышению результативности науки. Вскрываемые полезные взаимосвязи и закономерности в развитии науки в целом могут помочь государственным органам в упреждающем порядке организовывать финансовую и иную административную поддержку растущим научным направлениям в области как прикладных, так и фундаментальных исследований.

### Перспективы применения глубокого обучения

На современном этапе с технологической связкой «большие данные – глубокое обучение – нейросети», применяемой в широком спектре научных областей и

отраслей экономики, работают преимущественно крупные структуры, институты, компании, предприятия. В перспективе, как ожидается, средние и даже мелкие структуры будут все более активно применять эти технологические инструменты [2]. Для работы с ними не требуется больших вложений в создание собственной аппаратно-программной инфраструктуры и ее обслуживание, так как можно использовать облачные технологии, применяя готовые облачные решения соответствующих профильных компаний.

### ЗАКЛЮЧЕНИЕ

Таким образом, можно заключить, что глубокое обучение как подобласть искусственного интеллекта является важным инструментом, обеспечивающим успешное применение технологии больших данных в различных областях жизни общества, включая сферу науки. Анализ больших данных, реализуемый на основе связки глубокого обучения и искусственных нейросетей, позволяет увидеть скрытые закономерности, которые не может выявить естественный интеллект, и получить новые общественно полезные знания. Глубокое обучение раскрывает беспрецедентные возможности оптимизации научных исследований, государственного управления, производства, медицины, логистики, безопасности, финансов, телекоммуникаций и других общественных сфер.

Для науки глубокое обучение выступает как средство, позволяющее научным организациям и ученым значительно повысить эффективность в проведении глубинного поиска и анализа научных данных, определении основных тенденций современной науки, выявлении новых прорывных научных направлений и точек роста перспективного научного знания.

### СПИСОК ЛИТЕРАТУРЫ

1. Naur P. Concise Survey of Computer Methods. – Lund, Sweden, 1974. – 397 p.
2. Доклад “DataAge 2025” информационно-аналитической корпорации IDC – International Data Corporation (США) / Корпоративный сайт [www.idc.com](http://www.idc.com) (дата обращения 5.11.2022).
3. Мельникова Е.В. Особенности наполнения научных баз данных для эффективного применения технологии Big Data // Информационные ресурсы России / Раздел «Цифровые технологии». – 2021. – № 4(182). – С. 6-11.
4. Денисова О.Ю., Мухутдинов Э.А. Большие данные – это не только размер данных // Вестник технологического университета (г. Казань). – 2015. – Т. 18, № 4. – С. 226-230.
5. Nor Asiakin et al. Exploring big data traits and data quality dimensions for big data analytics application // Springer Science and Business Media // Journal of Big Data. – 2021. – Vol. 8. – P. 1-15.
6. Piattetsky-Shapiro G. Knowledge Discovery in Real Databases (Knowledge Discovery in Data (KDD)). A Workshop Report // AI Magazine. – 1991, January. – Vol. 11, № 5.
7. McCarthy J. Arthur Samuel: Pioneer in Machine Learning. – URL: <https://infolab.stanford.edu> (дата обращения 22.10.2022).

8. Samuel Arthur. Some Studies in Machine Learning Using the Game of Checkers // IBM Journal. – 1959. – № 3(3). – P. 210-229.
9. Elshawi R., Sakr S., Talia D. Big Data Systems Meet Machine Learning Challenges: Towards Big Data Science as a Service // Big Data Research. – 2018. – Vol. 14, December. – P. 1-11.
10. Microsoft Corporation. Deep Learning vs. Machine Learning. – 2022. – April 11. – URL: <https://learn.microsoft.com/ru-ru/azure/machine-learning/concept-deep-learning-vs-machine-learning> (дата обращения 25.10.2022).
11. LeCun Y., Bengio Y., Hinton G.E. Deep Learning // Nature. – 2015. – Vol. 521, May. – P. 436-444.
12. Hinton G. E. Learning Multiple Layers of Representation // Trends in Cognitive Sciences. – 2007. – Vol. 11. – P. 428-434.
13. Мельникова Е.В. Технология Big Data в наборе методов и средств научного исследования в современной наукометрии // Научно-техническая информация. Сер. 1. – 2022. – № 5. – С. 1-7; Melnikova E.V. Big Data Technology in the Set of Methods and Means of Scientific Research in Modern Scientometrics // Scientific and Technical Information Processing. – Springer. – 2022. – Vol. 49, № 2. – P. 102-107.
14. McCulloch Warren, Pitts Walter. A logical calculus of the ideas immanent in nervous activity // Bulletin of Mathematical Biophysics. – 1943. – № 5. – P. 115-133.
15. Гиляревский Р.С., Мельникова Е.В. Особенности доступа к данным в информационной инфраструктуре науки // Научно-техническая информация. Сер. 1. – 2021. – № 3. – С. 8-15; Gilyarevski R.S., Melnikova E.V. Peculiarities of Data Access within the Information Infrastructure of Modern Science // Scientific and Technical Information Processing. – Springer. – 2021. – Vol. 48, № 1. – P. 216-222.
16. Ulrich R., Pampel H., Kindling M. Advancing Services for Open Science // E-Science-Tage 2019. Data to Knowledge / eds. V. Heuveline, F. Gebhart, N. Mohammadianbisheh. – Heidelberg: heiBOOKS, 2020. – P. 194-195.
17. Von der Heyde M. Open Research Data: Landscape and cost analysis of data repositories currently used by the Swiss research community, and requirements for the future // Swiss National Science Foundation. – 2019. – С. 67. – URL: <http://doi.org/10.5281/zenodo.2643460> (дата обращения 12.11.2022).
18. Swiss National Science Foundation– 2022 (Швейцарский национальный научный фонд). – URL: [https://snf.ch/SiteCollectionDocuments/DMP\\_content\\_mySNF-form\\_fr.pdf](https://snf.ch/SiteCollectionDocuments/DMP_content_mySNF-form_fr.pdf) (дата обращения 04.11.2022).
19. Hathitrust digital repository library (Американская Цифровая репозитарная библиотека ХатиТраст). – URL: [https://hathitrust.org/features\\_benefits](https://hathitrust.org/features_benefits) (дата обращения 18.10.2022).
20. Keshav Singh R., Sandeep Kumar S. Emerging trends and global scope of big data analytics: a scientometric analysis // Quality & Quantity. – 2021. – Vol. 55, № 2. – P. 1-26.
21. Комков А.А., Мазаев В.П. и др. Основные направления развития искусственного интеллекта в медицине // Научное обозрение / Раздел «Медицинские науки». – 2020. – №5. – С. 33-40.
22. Краснов Ф.В., Буторин А.В., Ситников А.Н. Автоматизированное обнаружение геологических объектов в изображениях сейсмического поля с применением нейронных сетей глубокого обучения // Бизнес-информатика / Раздел «Анализ данных и интеллектуальные системы». – 2018. – № 2(44). – С. 7-15.

*Материал поступил в редакцию 16.11.22.*

#### **Сведения об авторе**

**МЕЛЬНИКОВА Елена Владимировна** – кандидат технических наук, старший научный сотрудник Отделения теоретических и прикладных проблем информатики ВИНТИ РАН, Москва  
e-mail: [verden.mel@yandex.ru](mailto:verden.mel@yandex.ru)