

ВЛИЯНИЕ ТЕХНОЛОГИЙ БОЛЬШИХ ДАННЫХ НА ДОСТИЖЕНИЕ ЦЕЛЕЙ УСТОЙЧИВОГО ЭКОЛОГИЧЕСКОГО РАЗВИТИЯ

THE IMPACT OF BIG DATA TECHNOLOGIES ON ACHIEVING SUSTAINABLE ENVIRONMENTAL DEVELOPMENT GOALS

**S. Afanasyev
T. Voronin
S. Cherevko**

Summary. This article is devoted to the analysis of the impact of big data collection, storage and analysis technologies on the achievement of sustainable environmental development goals. The positive and negative consequences of technology development are considered, special attention is paid to the negative consequences. Conclusions are drawn regarding the directions for the development of data processing centers, the improvement of which will help minimize negative impacts on the environment. The current best practices in the field of organizing energy-efficient approaches to computing, construction and operation of data processing centers are described. The work uses the results of the project “Intelligent analysis of big data in the tasks of protecting ecology and the environment”, supported by the Ministry of Science and Higher Education of the Russian Federation under the Agreement of Lomonosov Moscow State University with the Project Support Fund of the National Technology Initiative dated December 15, 2021 No. 70–2021–00252.

Keywords: computing, ecology, energy efficiency, sustainable development, big data, artificial intelligence.

Афанасьев Сергей Дмитриевич

Кандидат юридических наук, ведущий специалист,
Центр хранения и анализа больших данных; МГУ имени
М.В. Ломоносова, Москва
sergei.afanasev@digital.msu.ru

Воронин Тимофей Валерьевич

Специалист, Центр хранения и анализа больших
данных, МГУ имени М.В. Ломоносова, Москва
voronin@digital.msu.ru

Черевков Станислав Станиславович

Ведущий специалист, Центр хранения и анализа
больших данных, МГУ имени М.В. Ломоносова, Москва
cherevko@digital.msu.ru

Аннотация. Статья посвящена анализу влияния технологий сбора, хранения и анализа больших данных на достижение целей устойчивого экологического развития. Рассмотрены позитивные и негативные последствия развития технологий, особое внимание уделено негативным последствиям. Сформулированы выводы относительно направлений развития центров обработки данных, совершенствование в которых способствует минимизации негативных воздействий на окружающую среду. Описаны существующие на текущий момент лучшие практики в области организации энергоэффективных подходов к вычислениям, строительству и эксплуатации центров обработки данных. В работе использованы результаты проекта «Интеллектуальный анализ больших данных в задачах охраны экологии и окружающей среды», поддерживаемого Министерством науки и высшего образования Российской Федерации по Договору МГУ имени М.В. Ломоносова с Фондом поддержки проектов Национальной технологической инициативы от 15.12.2021 г. № 70–2021–00252.

Ключевые слова: вычисления, экология, энергоэффективность, устойчивое развитие, большие данные, искусственный интеллект.

Введение

Повестка устойчивого развития становится все более актуальной: в данной области реализуются масштабные фундаментальные и прикладные научно-исследовательские работы, разрабатываются и развиваются инновационные проекты. По мере развития технологий сбора, хранения и обработки данных они постепенно внедряются в процесс изучения и содействия достижению целей и реализации задач в области устойчивого развития. Цели устойчивого развития, принятые на Генеральной ассамблее ООН 25 сентября 2015 г. в рамках Резолюции 70/1 «Преобразование нашего мира: Повестка дня в области устойчивого развития на период до 2030 года», охваты-

вают широкий вызовов глобального развития, для преодоления которых необходима консолидация усилий разных стран. Важность участия в достижении целей устойчивого развития в Российской Федерации обозначена на государственном уровне в нескольких документах стратегического планирования и целеполагания. Среди 17 целей устойчивого развития ООН 3 цели непосредственно связаны с экологической повесткой (цели 13–15).

На текущий момент технологии больших данных и искусственного интеллекта начинают активно применяться в промышленности, транспорте, в сфере энергетики и ЖКХ, а также в области сельского хозяйства. Данные отрасли являются одними из основных источников

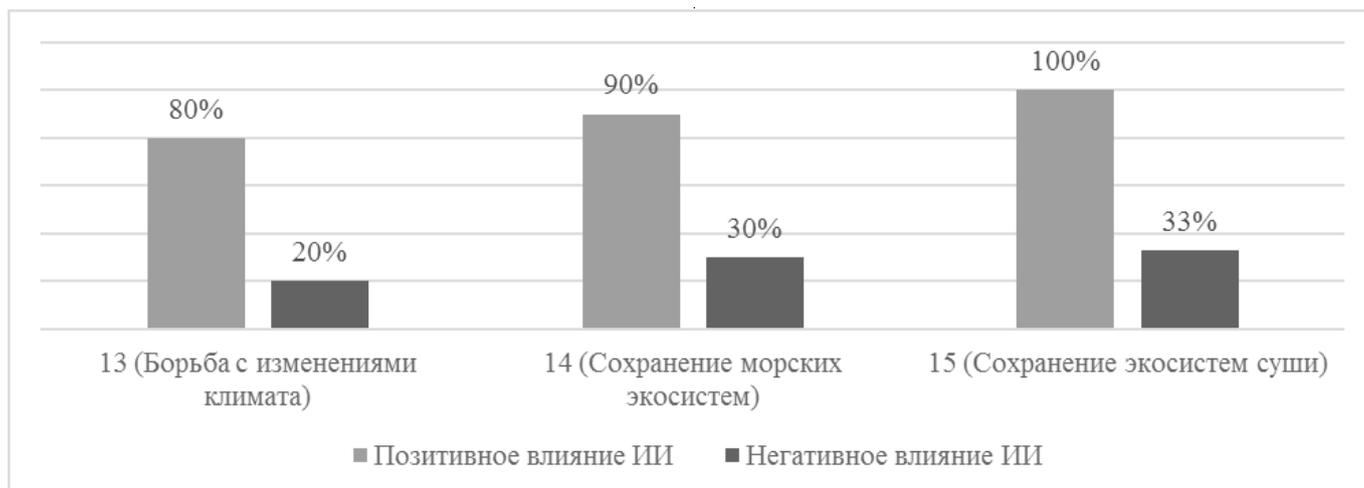


Рис. 1. Доля задач ЦУР ООН, на достижение которых влияет внедрение искусственного интеллекта
 Источник: R. Vinuesa, H. Azizpour, I. Leite, M. Balaam, V. Dignum, S. Domisch, A. Felländer, S.D. Langhans, M. Tegmark, F.F. Nerini. The role of artificial intelligence in achieving the Sustainable Development Goals // Nature Communications volume 11, Article number: 233 (2020) DOI: 10.1038/s41467-019-14108-y // URL: <https://www.nature.com/articles/s41467-019-14108-y> (дата обращения 25.11.2022)

загрязнения окружающей среды, поэтому повышение энергоэффективности процессов в них способно существенно повлиять на достижение ЦУР.

Одновременно с этим, развитие технологий больших данных и ИИ требует увеличения объема вычислительных ресурсов, потребляющих большой объем энергии и, следовательно, являющихся источником загрязнения окружающей среды. Ожидается, что общий спрос на электроэнергию за счет информационных и коммуникационных технологий (ИКТ) потребует до 20% мировой энергии к 2030 году по сравнению с 1–2% на текущий момент [1].

Большие данные и искусственный интеллект неоднозначно влияют на достижение целей устойчивого развития, связанных с экологической повесткой. С одной стороны системы, программные решения и инновации, основанные на технологиях больших данных и ИИ, способны повышать качество мониторинга и прогнозирования процессов в биосистемах и применяться в цифровых решениях, способствующих повышению эффективности расходования ресурсов и минимизации негативного воздействия на окружающую среду, с другой — для их функционирования требуются значительные ресурсы ЦОДов, функционирование которых приводит к пагубному воздействию на среду.

Текущие и перспективные последствия применения больших данных и ИИ в проектах, ориентированных на достижение ЦУР ООН, оцениваются преимущественно позитивно. По данным исследования [2], положи-

тельные эффекты от развития и внедрения ИИ в части достижения ЦУР могут быть получены в первую очередь за счет возможности анализа крупномасштабных взаимосвязанных баз данных для разработки совместных действий, направленных на сохранение окружающей среды. При позитивном влиянии искусственного интеллекта на реализацию 79% всех задач ЦУР ООН и негативном — на реализацию 35% задач, для ЦУР в области характерен большой потенциал эффективного использования ИИ и больших данных: технология позитивно влияет на достижение более чем 80% задач, а по одной из ЦУР внедрение искусственного интеллекта способно позитивно повлиять на реализацию 100% задач (рисунок 1).

Научно-исследовательские и коммерческие проекты, ориентированные на сбор и обработку больших данных о состоянии и изменениях процессов окружающей среды способны оказать существенное влияние на достижение целей устойчивого развития, однако, крайне актуальным вызовом становится разработка энергоэффективных технологических решений и подходов к организации процессов сбора, хранения и анализа больших данных и к обеспечению устойчивого развития центров обработки данных (ЦОД).

Влияние развития технологий сбора, хранения и анализа больших данных на достижение целей устойчивого развития должно оцениваться исходя из чистого влияния технологии, то есть важно учитывать всю совокупность позитивных и негативных последствий развития технологии и ее применения не только в об-

ласти непосредственно реализации задач ЦУР ООН, но и в других направлениях. Например, применение больших данных и ИИ в промышленности для задач оптимизации энергопотребления оборудования должно анализироваться и изучаться как в качестве фактора, позитивно влияющего на достижение целей устойчивого развития за счет повышения энергоэффективности производственных процессов, так и в качестве фактора, негативно влияющего на достижение ЦУР в связи с использованием большого количества вычислительных ресурсов ЦОД для хранения и обработки данных и большого объема потребления ресурсов.

По мере увеличения объемов собираемых и обрабатываемых данных увеличивается потребление мощностей ЦОД. На функционирование центров обработки данных приходится около 1% глобальных выбросов парниковых газов. Каждый год операторам ИТ-инфраструктуры необходимо адаптироваться к увеличению объема обрабатываемых данных, ожидается, что объем хранимых данных в 2025 г. будет в 5,3 раза больше, чем в 2018 г., что, безусловно, приведет к существенному увеличению объемов потребления ресурсов вычислительными системами [3]. Несмотря на увеличение спроса на мощности ЦОДов, одной из актуальных проблем в области обеспечения устойчивого развития в условиях развития технологий сбора, обработки и хранения больших данных, остается низкий уровень утилизации серверов в некоторых ЦОДах. Уровень эффективного использования наиболее мощных серверов с высочайшим уровнем потребления ресурсов даже в случае простоя может составлять менее 50% [4]. Данный вызов особенно характерен для корпоративных ЦОДов и локальной вычислительной инфраструктуры небольших компаний-операторов центров.

Ключевые направления развития ЦОДов определяются показателями устойчивого развития ЦОДов, большая часть из которых определены консорциумом Green Grid. Основными показателями эффективности ЦОД, связанными с уровнем последствий его функционирования для экологии, являются [5]:

1. Уровень эффективности расходования энергии (PUE), измеряемый как отношение потребляемой инфраструктурой мощности к мощности, отдаваемой серверам, дискам и сетевому оборудованию. При PUE, равном 1,0, ИТ-оборудование получает 100% потребляемой мощности в центре обработки данных, однако, среднее значение показателя составляет 1,67 и колеблется от 1,11 до 3,0 в зависимости от ЦОДа.
2. Эффективность использования углерода (CUE) — отношение выбросов CO₂, вызванных общим энергопотреблением центра обработки данных, к энергопотреблению ИТ-оборудования.

3. Недостаточный уровень эффективности использования воды (WUE), рассчитываемый как соотношение между использованием воды в системе центра обработки данных (водные петли, адiabатические башни, увлажнение, производство энергии на основе воды и т.д.) и потреблением энергии ИТ-компонентов. Около половины потребления воды приходится на системы охлаждения [6].

Менее используемыми метриками экологичности ЦОДов являются показатели, связанные с утилизацией электронной продукции, эффективностью управления воздушными потоками в ЦОД и др. Система показателей экологичности ЦОДов не определена окончательно и находится в стадии активной проработки в рамках ряда международных проектов, например, All4Green, CoolEmAll, GreenDataNet, RenewIT, GENiC, GEYSER, Dolfin и DC4Cities.

Эффективность использования ресурсов в ЦОДах определяется преимущественно на этапе проектирования и строительства, при этом для ЦОДов характерен длительный цикл эксплуатации и центры, создаваемые на текущий момент, будут эксплуатироваться в дальнейшем в течение десятков лет.

На текущий момент существует большое количество успешных практик повышения эффективности использования ресурсов в ЦОДах, они связаны с использованием энергии из альтернативных источников, применением новых подходов к проектированию инженерных систем ЦОДов, внедрением интеллектуальных решений по оптимизации потребления ресурсов и др. Соответствующие технологии изучаются, развиваются и внедряются с начала 2000-х годов, однако лишь в последние годы они становятся одним из ключевых направлений развития сферы проектирования и эксплуатации ЦОД [7] и все чаще учитываются в проектах по созданию новых, развитию и масштабированию существующих центров обработки данных.

Развитие технологий сбора, хранения и анализа больших данных требует увеличения объема эксплуатируемых вычислительных ресурсов, которые могут быть развернуты на мощностях сторонней организации — оператора ЦОДа, облачного провайдера или использоваться на локальной инфраструктуре организации. В первом случае инфраструктура предоставляется в формате облачных технологий и процессы проектирования, строительства и администрирования ЦОДа реализуются провайдером облачных услуг. Во втором случае аналогичные работы выполняются непосредственно организацией, которая в дальнейшем использует инфраструктуру собственных сер-

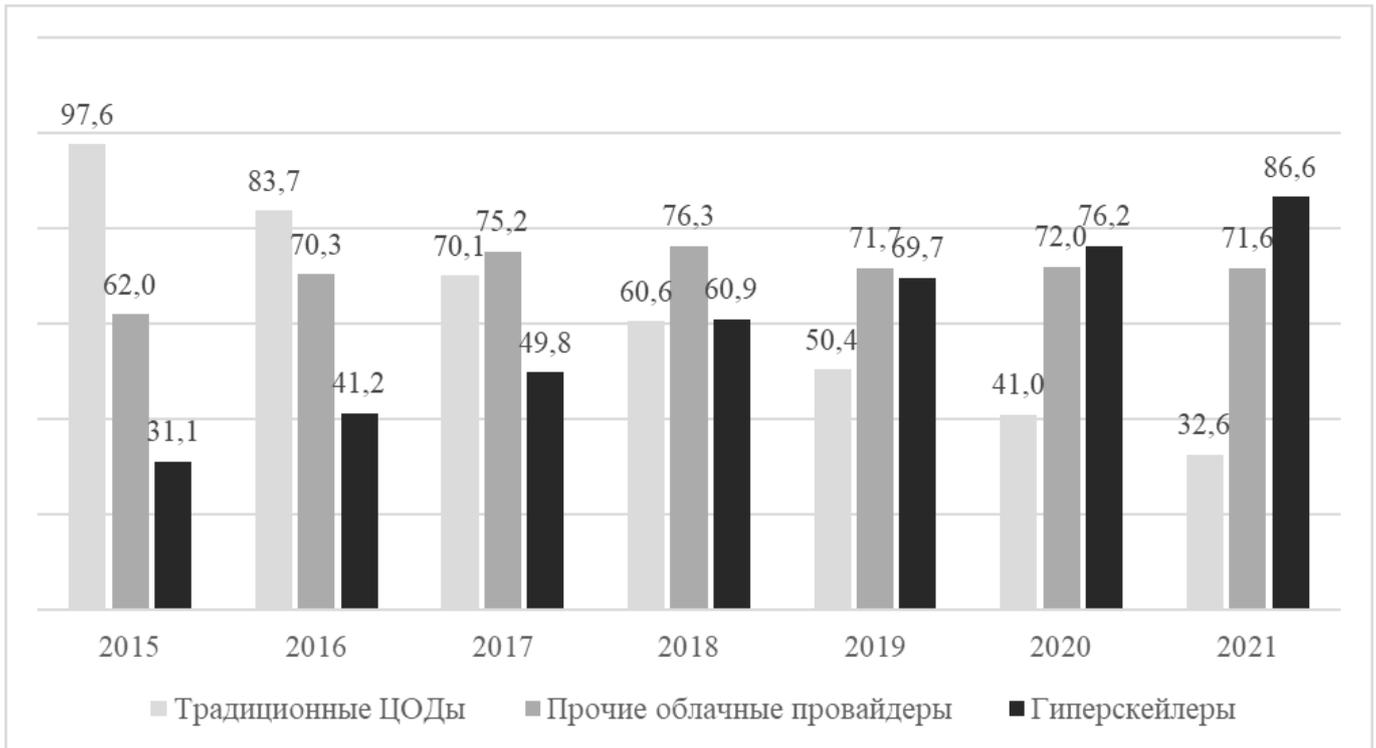


Рис. 2. Объем потребления энергии в ЦОДах, ТВт·ч

Источник: Официальный сайт Statista «Energy demand in data centers worldwide from 2015 to 2021, by type(in terawatt hours)» // URL: <https://www.statista.com/statistics/186992/global-derived-electricity-consumption-in-data-centers-and-telecoms/> (дата обращения 25.11.2022)

веров для размещения рабочих нагрузок, хранения и анализа данных.

Таким образом, обеспечение устойчивого развития технологий сбора, хранения и анализа больших данных и минимизация негативных последствий от развития технологий больших данных и искусственного интеллекта на окружающую среду требуют реализации мероприятий, способствующих проектированию и эксплуатации ЦОДов с высоким уровнем энергоэффективности на уровне крупнейших корпораций и организаций, использующих локальную ИТ-инфраструктуру, а также на уровне операторов ЦОД и облачных провайдеров, предоставляющих вычислительную инфраструктуру в формате as-a-service.

За период с 2015 до 2021 существенно возрастает объем энергопотребления гиперскейлеров [8] — ведущих мировых облачных провайдеров, к которым относятся Amazon Web Services, Microsoft Azure и Google Cloud Platform, при этом значительно уменьшается объем потребления энергии операторами ЦОД. Динамичное развитие облачных технологий и облачных провайдеров определяет особую роль позиции и инвестиций данных компаний в строительство и эксплу-

атацию эффективных по параметрам потребления ресурсов ЦОДов, а также уровня их участия в процессах формирования стандартов и требований к ЦОДам, что продемонстрировано на рисунке 2.

Гиперскейлеры и крупнейшие национальные облачные провайдеры акцентируют приоритет на устойчивое развитие при описании показателей функционирования их инфраструктуры, а также при планировании и запуске новых ЦОДов или масштабирования существующих центров. Amazon Web Services — ведущий глобальный облачный провайдер, активно инвестирует в технологии устойчивого развития и разрабатывает собственные инфраструктурные решения, позволяющие повысить ресурсоэффективность серверов. Согласно исследованиям 451 Research, инфраструктура AWS в 3,6 раз более энергоэффективна, чем корпоративный ЦОД в США, и до 5 раз более энергоэффективна, чем ЦОД в Европе. На текущий момент в ЦОДах провайдера используются особые процессоры, применяются высокоэффективные системы охлаждения, алгоритмы прогнозирования и отслеживания производительности и другие решения, способствующие оптимизации потребления ресурсов. К 2025 году предполагается переход ЦОДов провайдера на возобновляемую энер-

гию, что позволит снизить углеродный след рабочей нагрузки пользователей облачных сервисов более чем на 90%.

Крупнейшие компании, имеющие стратегии в области ESG, активно инвестируют и внедряют инновационные подходы к организации процессов сбора, хранения и анализа данных, направленные на снижение энергопотребления и минимизацию вреда окружающей среде от функционирования корпоративных ЦОДов. Facebook [9] использует энергоэффективные ЦОДы, сертифицированные LEED Gold, потребление воды в данных ЦОДах на 80% меньше, чем в стандартных ЦОДах. В ЦОДах компании используются особые подходы к обеспечению и регулированию влажности воздуха и организации систем охлаждения. Компания является членом Совета по экологическому строительству США и активно участвует в Консультативном комитете LEED в рамках формирования системы экологического рейтинга для центров обработки данных. В Google применяют технологии, позволяющие снизить энергопотребление ЦОДов в 2 раза. Целью компании к 2030 году является переход на безуглеродную энергию в режиме 24/7 во всех центрах обработки данных. Для оптимизации загрузки и повышения энергоэффективности ЦОДов используют машинное обучение [10].

Роль облачных провайдеров, на ЦОДы которых приходится все больший объем потребления ресурсов, в обеспечении устойчивого развития подходов к работе с данными, чрезвычайно высока. Для облачных провайдеров доля издержек, связанных с поддержанием функционирования ЦОДов, является одной из наибольших в общем объеме затрат, поэтому инвестиции в устойчивое развитие для них необходимы не только в связи с общим приоритетом на сбережение природных ресурсов, но и экономически оправданы в долгосрочной перспективе. Одновременно с этим, крупнейшие корпорации с высоким уровнем цифровизации также активно развивают ИТ-инфраструктуру и могут потреблять ее в объеме, аналогичном или превышающим объемы потребления инфраструктуры ЦОДов облачных провайдеров. Организации, создающие и развивающие собственные ЦОДы, также должны инвестировать в их экологичность, внедряя лучшие практики в данной области. На текущий момент многие компании, в особенности — международные, упо-

минают экологичность вычислительной инфраструктуры среди приоритетов корпоративного устойчивого развития.

В заключение, необходимо отметить, что технологии сбора, хранения и анализа больших данных, а также алгоритмы искусственного интеллекта и их влияние на достижение целей устойчивого развития должно оцениваться комплексно и учитывать не только преимущества технологий и возможности их использования в областях, связанных с потреблением ресурсов, но и отрицательные последствия их развития, которые являются косвенными и связаны с увеличением объема потребления вычислительных ресурсов и мощностей ЦОДов.

Развитие технологий сбора, хранения и анализа больших данных приводит к существенному увеличению потребления вычислительных ресурсов. На текущий момент показатели устойчивого развития ЦОДов измеряются несистемно и несмотря на то, что на глобальном уровне и в ряде стран большинство крупнейших корпораций замеряют углеродный след и учитывают в этом процессе показатели эффективности вычислительной инфраструктуры, в Российской Федерации данная практика не развита и не является общепринятой. В ближайшие годы важные вызовы связаны с разработкой системы показателей измерения экологичности вычислительной инфраструктуры. Данный процесс должен производиться с вовлечением российских организаций — крупнейших ИТ-компаний, крупнейших промышленных организаций, имеющих и развивающих собственную инфраструктуру и активно инвестирующих в цифровизацию и процессы сбора и анализа больших данных.

Благодарности. Исследование выполнено при финансовой поддержке в рамках реализации программы Центров компетенций Национальной технологической инициативы на базе Московского государственного университета имени М.В. Ломоносова (договор о предоставлении средств юридическому лицу, индивидуальному предпринимателю на безвозмездной и безвозвратной основе в форме гранта, источником финансового обеспечения которых полностью или частично является субсидия, предоставленная из федерального бюджета № 70–2021–00252 от 15.12.2021).

ЛИТЕРАТУРА

1. Jones N. How to stop data centres from gobbling up the world's electricity // Nature. 2018. Vol. 561. No. 7722. PP. 163–167.
2. R. Vinuesa, H. Azizpour, I. Leite, M. Balaam, V. Dignum, S. Domisch et al. The role of artificial intelligence in achieving the Sustainable Development Goals // Nature Communications. 2020. Vol. 11. No. 233. [Электронный ресурс] URL: <https://www.nature.com/articles/s41467-019-14108-y> (дата обращения: 25.11.2022)

3. Replace your electricity consumption with 24/7 carbon-free energy generation, generating revenue and achieving high resiliency for your data center // Официальный сайт «Engie». [Электронный ресурс] URL: [https://www.engie.com/en/campaign/green-data-centers#:~: text=Being%20particularly%20energy-intensive%2C%20the,greenhouse%20gas%20emissions%20\(2\)](https://www.engie.com/en/campaign/green-data-centers#:~:text=Being%20particularly%20energy-intensive%2C%20the,greenhouse%20gas%20emissions%20(2)) (дата обращения: 25.11.2022)
4. The Real Amount of Energy A Data Center Uses // официальный сайт «ACKR». [Электронный ресурс] URL: <https://www.akcp.com/blog/the-real-amount-of-energy-a-data-center-use/> (дата обращения: 25.11.2022)
5. Reddy V.D. et al. Metrics for sustainable data centers // IEEE Transactions on Sustainable Computing. 2017. Vol. 2. No. 3. PP. 290–303.
6. Mytton D. Data centre water consumption // npj Clean Water. 2021. V. 4. No. 1. PP. 1–6.
7. Buyya R., Beloglazov A., Abawajy J. Energy-efficient management of data center resources for cloud computing: A vision, architectural elements, and open challenges // arXiv. 2010.
8. Energy demand in data centers worldwide from 2015 to 2021, by type // Statista. [Электронный ресурс] URL: <https://www.statista.com/statistics/186992/global-derived-electricity-consumption-in-data-centers-and-telecoms/> (дата обращения: 25.11.2022).
9. Data centers // Meta Sustainability. [Электронный ресурс] URL: <https://sustainability.fb.com/data-centers/> (дата обращения: 25.11.2022).
10. Machine learning finds new ways for our data centers to save energy // Google Sustainability. [Электронный ресурс] URL: <https://sustainability.google/progress/projects/machine-learning/> (дата обращения: 25.11.2022).

© Афанасьев Сергей Дмитриевич (sergei.afanasev@digital.msu.ru),

Воронин Тимофей Валерьевич (voronin@digital.msu.ru), Черевков Станислав Станиславович (cherevkvss@digital.msu.ru).

Журнал «Современная наука: актуальные проблемы теории и практики»



Московский государственный университет имени М.В. Ломоносова