



# ИНФОРМАЦИОННЫЕ РЕСУРСЫ РОССИИ

02 [197] 2024

**Д. Холкин**

Многоголосье молчания **6**

**А. Мельников, А. Бачурин,  
Е. Гниломёдов, А. Распопов, К. Дворецкая**  
Технологии больших данных  
в цифровизации ТЭК **30**

**В. Цветкова, В. Кулев**  
Обеспечение сохранности и доступности  
электронных фондов библиотеки **38**





**РОСКОНГРЕСС**  
Пространство доверия



Организаторы:



При поддержке:



**26-28** ЦВЗ «Манеж»,  
сентября Манежная пл., д.1

# МЕЖДУНАРОДНЫЙ ФОРУМ РОССИЙСКАЯ ЭНЕРГЕТИЧЕСКАЯ НЕДЕЛЯ



**ВЫСТАВКА ОБОРУДОВАНИЯ  
И ТЕХНОЛОГИЙ ДЛЯ ТЭК**  
Гостиный двор, ул.Ильинка, д.4



# ТЕРРИТОРИЯ ЭНЕРГЕТИЧЕСКОГО ДИАЛОГА

## III НАУЧНО-ПРАКТИЧЕСКАЯ КОНФЕРЕНЦИЯ

**25 СЕНТЯБРЯ, 2024**

МЕРОПРИЯТИЕ ПРОВОДИТСЯ  
В РАМКАХ МЕЖДУНАРОДНОГО ФОРУМА



[www.rusenergyweek.com](http://www.rusenergyweek.com)

Реклама **6+**



[tedconf.ru](http://tedconf.ru)

г. Москва  
**7 (495) 789 92 92** доб. 2077  
[tedconf@rosenergo.gov.ru](mailto:tedconf@rosenergo.gov.ru)

На правах рекламы **6+**



ПРАВИТЕЛЬСТВО  
САХАЛИНСКОЙ  
ОБЛАСТИ



КВАДРАТ  
РЕСУРС



## ДАЛЬНЕВОСТОЧНЫЙ ЭНЕРГЕТИЧЕСКИЙ ФОРУМ

НЕФТЬ И ГАЗ САХАЛИНА

# 2024

2-4 октября,  
г. Южно-Сахалинск

## ВОСТОЧНЫЙ ФОРПОСТ ЭНЕРГЕТИЧЕСКОЙ ОТРАСЛИ: ТЕХНОЛОГИЧЕСКИЕ ПРИОРИТЕТЫ И НОВЫЕ ЦЕЛИ



**СЕРВИСЫ И ТЕХНОЛОГИИ  
НА ШЕЛЬФЕ**



**ПЕРЕРАБОТКА  
УГЛЕВОДОРОДОВ**



**ЧИСТОЕ БУДУЩЕЕ.  
УСТОЙЧИВОЕ  
РАЗВИТИЕ И КЛИМАТ**



**ЛОГИСТИКА И  
МЕЖДУНАРОДНОЕ  
СОТРУДНИЧЕСТВО**



**ИНТЕГРАЦИЯ НАУКИ И  
ОБРАЗОВАНИЯ В БИЗНЕС.  
ОТ ИССЛЕДОВАНИЙ  
К ИННОВАЦИЯМ**



**ВЫСТАВОЧНАЯ  
ПРОГРАММА**



**МОЛОДЕЖНЫЙ  
ФОРУМ**



**КРЕАТИВНЫЙ  
ПОТЕНЦИАЛ  
ТЕРРИТОРИЙ**



**СПЕЦИАЛЬНЫЕ  
МЕРОПРИЯТИЯ**



WWW.SAKHALINOILGAS.RU  
**+7 499 350 25 35**

**ПРИГЛАШАЕМ ПРИНЯТЬ УЧАСТИЕ**

## Содержание

### От редакции

**5 А. Горшкова**  
Большие дискуссии о праве и сущности искусственного интеллекта

### Искусственный интеллект

**6 Д. Холкин**  
Многоголосье молчания

**24 Д. Телекова, В. Павлова, А. Поначугин**  
Справедливость и предвзятость к искусственному интеллекту: изучение подходов к смягчению предубеждений и обеспечение справедливости в процессах принятия его решений

### ТЭК

**30 А. Мельников, А. Бачурин, Е. Гниломёдов, А. Распопов, К. Дворецкая**  
Технологии больших данных в цифровизации ТЭК

### Библиотека

**38 В. Цветкова, В. Кулев**  
Обеспечение сохранности и доступности электронных фондов библиотеки

## ИНФОРМАЦИОННЫЕ РЕСУРСЫ РОССИИ

### Интернет

**54 Н. Редькина**  
Веб-архив: важнейший источник информации vs неупорядоченное хранилище веб-контента

### Строительство

**67 А. Пиляй, Е. Гинзбург**  
Применение глубокого обучения для обнаружения повреждений на фасаде зданий

### Цифра

**75 М. Муколина, Д. Глазунов, А. Гальцева**  
Разработка программного продукта для стандартизации документов в электронном виде

**85 С. Малахов, Д. Якупов**  
Анализ производительности программного протокола OpenFlow на виртуальной машине

### Безопасность

**93 А. Яковишин, И. Кузнецов, И. Дроздов, Д. Письменский**  
Перспективы развития информационной безопасности: глобальные вызовы и стратегии защиты



6



30



85



93

СОДЕРЖАНИЕ



**Editor's word**

- 5 A. Gorshkova**  
Big discussions about the law and the essence of artificial intelligence

**Artificial intelligence**

- 6 D. Kholkin**  
Polyphony of silence
- 24 D. Telekova, V. Pavlova, A. Ponachugin**  
Fairness and bias towards artificial intelligence: exploring approaches to mitigating bias and ensuring fairness in ai decision-making processes

**Energy**

- 30 A. Melnikov, A. Bachurin, E. Gnilomedov, A. Raspopov, K. Dvoretzkaya**  
Big data technologies in the digitalization of the fuel and energy complex

**Library**

- 38 V. Tsvetkova, V. Kulev**  
Ensuring the safety and accessibility of the library's electronic collections

**Contents**

**Internet**

- 54 N. Redkina**  
Web archive: an important source of information vs disordered web content storage

**Construction**

- 67 A. Pilyay, E. Ginzburg**  
Using deep learning to detect damages on building facades

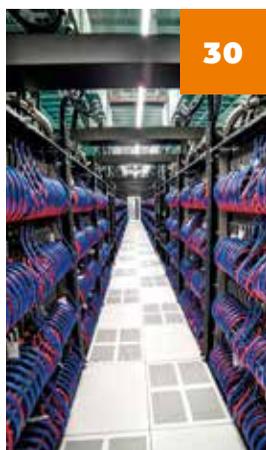
**Digitalization**

- 75 M. Mukonina, D. Glazunov, A. Galtseva**  
Software product development for standardization of documents in electronic form

- 85 S. Malakhov, D. Yakupov**  
OpenFlow software protocol performance analysis on virtual machine

**Safety**

- 93 A. Yakovishin, I. Kuznetsov, I. Drozdov, D. Pisisky**  
Prospects for the development of information security: global challenges and protection strategies



S  
T  
R  
E  
T  
T  
O  
C

**Большие дискуссии о праве и сущности искусственного интеллекта**

Споры об искусственном интеллекте, его роли в социально-культурной и экономической жизни общества и страны ведутся уже не первый год. Главными вопросами по-прежнему остаются определение самого понятия искусственного интеллекта, его способность адекватно использовать получаемые данные, его возможности саморазвиваться и влиять на окружающую действительность, социальное взаимодействие с людьми и другими машинами, а главное – безопасность использования и способы контролирования его работы человеком.

Данная технология является еще достаточно новой, поэтому человечество не накопило необходимого объема практики работы с искусственным интеллектом, в том числе и негативной. Именно наличие практики применения сможет выявить весь спектр проблем эксплуатации искусственного интеллекта и контроля над его работой, в том числе при заштатных и мало прогнозируемых ситуациях.

Пока же, на данном этапе технологического развития, задачей общества является не ограничение использования нейросетей и искусственного интеллекта, а подготовка нормативно-правовой базы, в том числе и на международном уровне, которая может упорядочить, урегулировать правила работы данной сферы и обезопасить пользователей от системных ошибок.

Главный редактор журнала «ИРР»,  
**Горшкова Анна**



## МНОГОГОЛОСЬЕ МОЛЧАНИЯ

*Аннотация. Обсуждается возможность создания гибридного человеко-машинного интеллекта за счет использования в коммуникациях знаковой системы нового типа, базирующейся на цифровом самоисполняемом коде. Технологии гибридного интеллекта позволят управлять системами высокой сложности, в которых акантами могут стать не только люди и технические объекты, но и любые физические и абстрактные сущности (в т. ч. животные, растения, экосистемы). Демонстрируется возможности применения данного подхода для решения планетарных проблем (в т. ч. климатических изменений), а также для решения задач «ближнего прицела».*

**Холкин Дмитрий**  
Генеральный директор АНО  
«Центр энергетических систем  
будущего «Энерджинет»  
E-mail: [dvh@internetofenergy.ru](mailto:dvh@internetofenergy.ru)

### Ключевые слова:

искусственный интеллект, гибридный интеллект, мультиагентные технологии, антропо-биотехносистемы, интернет энергии.

### Вступление

В 2023 г. произошла революция ChatGPT, ознаменовавшая когнитивный «фазовый переход» к генеративному искусственному интеллекту [1]. Так, по крайней мере, считают некоторые аналитики. Они ссылаются на демонстрируемые достижения GPT-приложений, на громкие заявления ученых, предпринимателей и представителей национальных правительств. Некоторые из этих заявлений носят экзистенциальный характер, пророчат угрозы для существования человечества. Можно ли доверять таким суждениям и заявлениям? Не являются ли они проявлением низкой критичности мышления говорящих, или, наоборот, предполагают рефлексивно-коммуникативные манипуляции, преследующие проведение своих интересов?

Вероятно, можно и нужно признать, что в настоящее время происходит масштабная трансформация цивилизации. Действительно, «генеративный искусственный интеллект представляет собой философскую и практическую задачу масштабов, невиданных со времен Просвещения» [2]. И можно согласиться с мнением исполнительного директора Google Сундара Пичаи, что технологии искусственного интеллекта на данный момент – явление более глубокое, чем когда-то такими были овладение электри-

чеством и огнем [3]. Однако это все уже было обсуждено несколько лет назад, когда в глобальной повестке появился концепт о четвертой промышленной революции. Произошло что-то неожиданное или мы сейчас просто наблюдаем первые и наиболее яркие плоды нового технико-экономического уклада?

Примечательна реакция ученых и мыслителей, которые занимались многие годы проблематикой искусственного интеллекта. Их мнения часто различны, если не сказать полярные. Например, физик и информатик Дуглас Хофштаттер верит в скорый «приход» AGI (сильного искусственного интеллекта) и очень обеспокоен этим, так как человечество не готово к таким радикальным изменениям. Когда он в конце прошлого века писал книгу про искусственный интеллект, то думал, что сопоставимый с человеческим машинный интеллект лишь появится через 500 лет [4]. А лингвист и философ Ноам Хомский, напротив, пессимистично оценивает последние достижения в сфере AI. Он говорит, что машинное обучение, в котором сейчас произошел прорыв, это только «высокотехнологичный плагиат, не имеющий ничего общего с разумом», и что надо вернуться к исходной научной программе изучения человеческого интеллекта [5].

Однако следует не столько дискутировать о сроках появления AGI, сколько, ба-

**Следует начать  
с малого – создать  
управляющие  
системы на основе  
гибридного  
коллективного  
интеллекта для  
решения сложных  
задач меньшего  
масштаба**

зируясь на имеющихся технологических артефактах и прикладных разработках, строить конструктивные представления о месте новых технологий, их влиянии на общество и людей, идентифицировать вероятные проблемы и открывающиеся возможности. Например, философ и когнитивист Дэниел Деннет, откладывая в сторону размышления о философском статусе складывающегося «интеллекта» машин, обсуждает проблему, к которой, кажется, общество не было готово: «Самая острая проблема не в том, что они собираются отобрать у нас рабочие места, не в том, что они собираются изменить методы ведения войны, а в том, что они собираются разрушить человеческое доверие. Они собираются переместить нас в мир, где невозможно отличить правду от лжи». В то же время, вспоминая слова советского математика, системщика и главного конструктора Спартака Никанорова «что не полезно, то не опасно, и, наоборот, что не опасно, то не полезно», мы можем обнаружить полезность от появления AI, такую же весомую и масштабную, как опасность. Чтобы эту полезность открыть, нам потребуется занять инженерную позицию, подразумевающую практический синтез новых научных знаний, направленный на проектирование будущего человечества.

### На пороге открытия гибридного интеллекта

Человеческий мозг – это природный компьютер, интегрированный в сферу культуры. Так можно коротко сформулировать подход, который развивал в теории сознания упомянутый ранее Дэниел Деннет [7]. В этой констатации нет желания умалить сложность интеллектуальной, эмоциональной и духовной организации человека. Она приводится только лишь для того, чтобы нащупать инженерный путь в пока еще диких джунглях представлений об интеллекте. Как последовательный физикалист, Д. Деннет утверждает, что у человека нет особого мыслительного органа, которого нет у животных. Мозг человека и животных предназначен естественным

отбором в процессе биологической эволюции на роль надежного «оборудования», способного стабильно развиваться и извлекать семантическую информацию, необходимую для управления своим поведением. В основе его работы находятся алгоритмы байесовской статистики, предписывающие предположительно правильные суждения о вероятностях наступающих событий. Мы имеем дело, если можно так сказать, с производителями прогнозов, обладающими выдающимися способностями. Эти байесовские предсказатели не имеют в себе ничего волшебного: они работают, как обыкновенные компьютеры, только природного происхождения.

Отличие поведения людей от поведения животных, которое мы и обозначаем понятием «разумное», состоит в понимании причин процессов, которые они отслеживают. Это понимание позволяет осознанно использовать причины в качестве инструментов и целей нашего разумного поведения. Только у человека можно обнаружить свободную волю к познанию и использованию мыслительных техник, систематическое изучение разных возможностей решения тех или иных проблем, попытки контроля мыслительных процессов высшего порядка. Это нам позволяет конструировать и применять новые знания, новые способы и методы, служащие причинами наших действий. То есть мы способны быть не просто движимыми причинами, а причинами, созданными нами самими. Человек – единственный вид, который обнаружил, что есть другие вещи (помимо биологического выживания и размножения), во имя которых можно умереть: свобода, демократия, правда, коммунизм, Римская католическая церковь, ислам и многое другое [7].

Такое понимание стало возможным только после появления на сцене нового вида эволюционного размножения – передачи информационных сущностей в пространстве культуры. Именно это позволило установить в человеческие мозги мыслительные инструменты и информационные симбионты (мемы), превратившие мозги в умы. Современное человеческое сознание в значительной степени – продукт культурной эволюции,

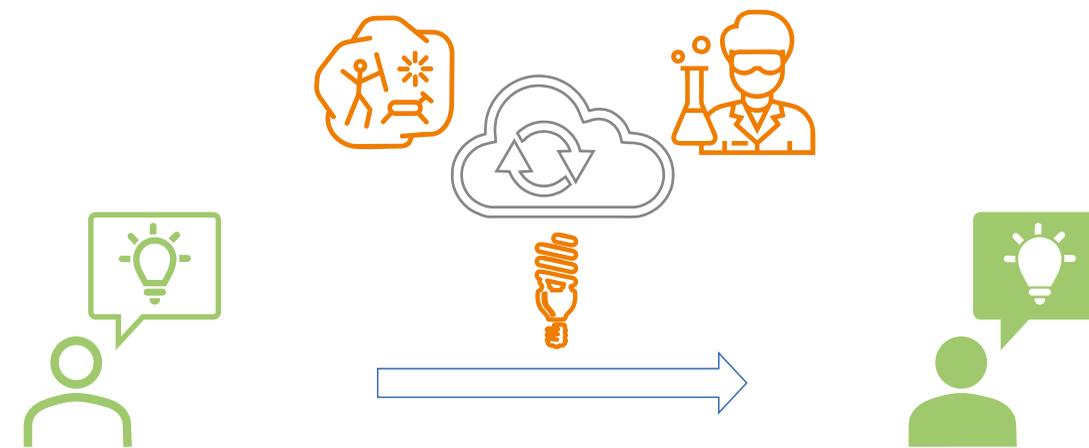


Рис. 1. Коммуникации осуществляются при помощи мемов, являющихся продуктом культурной эволюции

которая установила массу мемов и мыслительных инструментов в наши мозги, создав тем самым особую когнитивную архитектуру, отличную от архитектуры мозга животных. Благодаря заселению культурными эволюционными симбионтами, информационными структурами, наш мозг приобрел способность творца и может создавать артефакты и наши собственные жизни [7].

Сознание развивалось только под давлением потребности в общении. Привычка делиться информацией в процессе коммуникации с другими, спрашивать и предлагать причины и является поводом для появления новых интеллектуальных артефактов и мыслительных инструментов. Например, такой интеллектуальный артефакт, как личность человека, формируется только в постоянных попытках объяснения нас самих другим. Таким образом, коммуникации являются решающим фактором появления, поддержания и развития сознания.

Ингредиентами разума являются «hard» (байесовские предсказатели) и «soft» (информационные структуры, поставляемые в рамках культурных коммуникаций). При этом «hard» сейчас уже в определенной степени

технологически воспроизводим, а по мере роста производительности вычислений машинный мозг будет сопоставим с человеческим. И вся сакральная часть интеллектуальности находится в над-индивидуальных процессах. Примером такого над-индивидуального процесса является мышление. Напомним, что советский философ и методолог Георгий Щедровицкий говорил: «По сути дела, не человек мыслит, а мышление мыслит через человека» [8].

Об определяющей роли социальных отношений и коммуникаций в разумности человека писал и советский философ Эвальд Ильенков со своими коллегами. По их представлению, способность мыслить не «закодирована» в человеке генетически, биологически, она «наследуется», передается от поколения к поколению совсем другим путем – через формы предметного мира, созданного трудом, через тело цивилизации. Чтобы отдельный мозг обрел способность мыслить, его обладатель должен быть с детства включен в систему общественно-человеческих отношений, в языковую практику, и развит в согласии с ее требованиями и нормами. Функция мышления определяется не самой

по себе морфологической организацией тела индивида, а организацией той грандиозно сложной системы, которая на языке науки именуется «совокупность общественных отношений между людьми» [9].

Логический ход из всего вышесказанного состоит в том, что машину для надления её разумностью надо сделать участником общественно-человеческих отношений, стороной коммуникаций, чтобы она стала больше, чем быстрым байесовским предсказателем, чтобы она вступила в «сферу разума». В подтверждение этой идеи приведем еще одно высказывание Георгия Щедровицкого: «Люди есть случайные носители мышления. Можно реализовать мышление на людях, а можно на смешанных системах людей и машин. Главное – что есть мышление, а на чем оно реализуется – неважно. В нашем мире – случайно – на людях, в другом мире – на пингвинах, а в третьем – как у Лема, на железках. Какая разница, на чём это реализуется?!» [8].

Современные цифровые технологии такие коммуникации между человеком и машиной, а также между машинами делает возможными. «Цифра» (самоисполняемый код) является новой знаковой системой, пришедшей на смену речи и языку, и включающей в состав коммуникантов машины [10].

Для снятия остроты данных утверждений следует ввести различие субъектности и акторности, которые присутствуют в человеко-машинных отношениях. Субъектность основывается на рефлексии, выражается в волевом действии, которое может производиться вопреки предварительно заданным нормам и установкам (даже эмоциям и рефлексам). Акторность же – это генерирование решений на основе предварительно заданных норм и установок. Для дальнейших рассуждений можно исходить из того, у AI нет субъектности, но есть акторность, которая базируется на заданных иным субъектом принципах.

По мнению футурологов, на наших глазах рождается новая форма человеческой организации: гибридный коллективный интеллект – коллективные человеко-машинные системы, соединяющие силу естественного индивидуального восприятия и мышления, коллективного сотворчества и машинного обучения. В гибридном интеллекте ведущая роль (постановка задач, принятие решений, познание и творчество) отведена людям и их коллективному взаимодействию. Искусственный интеллект выполняет множество поддерживающих задач, усиливающих коллективный потенциал: сбор и анализ данных,

оценку вариантов решений, управляемое изменение режимов работы технических систем, контроль рисков [11].

Конечно, сегодня мы видим и другой тип систем – «синтетический» интеллект, в котором искусственный интеллект занимает центральное место, является постановщиком задач (на текущий момент – по воле своих создателей), а люди выполняют подчиненную роль исполнителей, «сенсоров», поставщиков информации, обучающих систем. Но это временная ситуация: как показывает в своей работе физик и социолог Дирк Хелбинг, у систем гибридного коллективного интеллекта есть долгосрочные стратегические преимущества по сравнению с синтетическим искусственным интеллектом, обучающимся на массивах больших данных, при росте сложности среды, с которой он взаимодействует [12].

Из всего этого следует важный вывод. Несколько в стороне от хайпа ChatGPT намечается новый технологический сдвиг. Его сильной стороной будет не столько обустройство отношений между людьми, сколько взаимодействие между человеком и внешней средой. Машинный интеллект через цифровые коммуникации станет актантом организованных человеком систем. За счет образования гибридного коллективного интеллекта эти системы смогут справиться с порядковым ростом сложности управляемых процессов. В нашем инженерном размышлении о будущем – это ключевой пункт.

### Управление сложностью в антропотехнобиосистемах

«Как паровая машина позволила получить независимую, отделенную от природы физическую силу, так AI позволяет получить независимую, отделенную от человека интеллектуальную силу». Так можно «докрутить» мысль, с которой начал свое выступление на встрече с Камчатским клубом директор по стратегическому маркетингу компании «Яндекс» Андрей Себрант. Да, отделенная от человека интеллектуальная сила, пусть и остающаяся в составе гибридного кол-

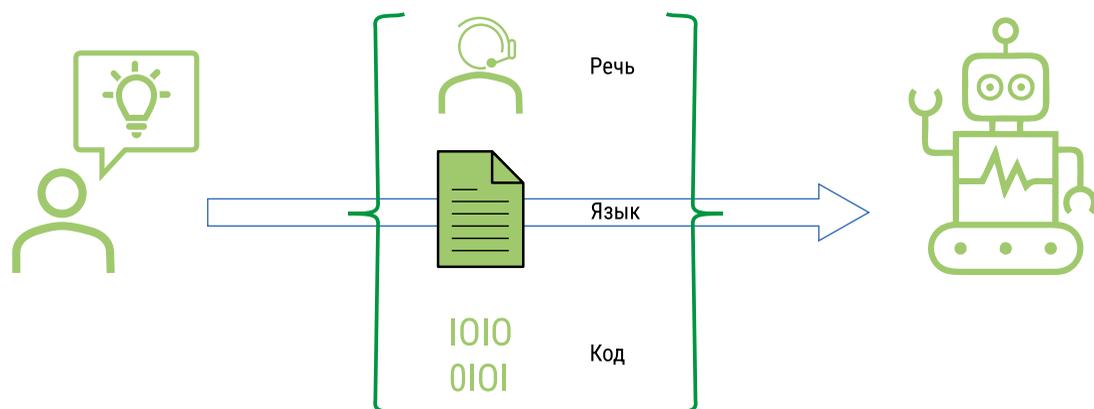


Робот-доставщик  
Источник: carposting.ru

лективного интеллекта, дает основание для революционных изменений, сомасштабных с событиями первой промышленной революции. Поясним это на одном из технологических подходов по применению AI.

Ключевым направлением развития технологий AI, позволяющем создавать системы на основе гибридного коллективного интеллекта, является мультиагентный подход. Его суть заключается в принципиально новом методе решения сложных задач, которые не решаются или трудно решаются классическими математическими методами. В отличие от классического способа, в котором проводится комбинаторный поиск вариантов решения по четко определенному (детерминированному) алгоритму, в мультиагентных технологиях решение задачи осуществляется в ходе самоорганизации множества программных агентов, способных к конкуренции и кооперации, и имеющих собственные критерии, предпочтения и ограничения. Ответ считается найденным

Рис. 2. Самоисполняемый цифровой код – новая знаковая система для коммуникаций



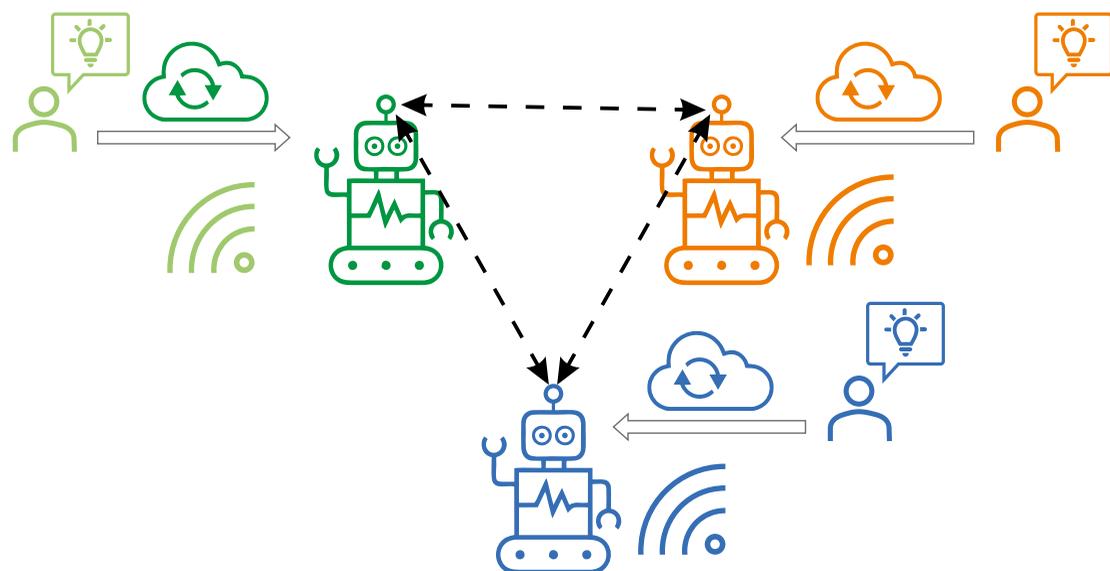


Рис. 3. Мультиагентные технологии являются еще одним прорывным направлением AI

тогда, когда в ходе своих недетерминированных взаимодействий агенты достигают неуклонного консенсуса (временного равновесия или баланса интересов), который и принимается за решение задачи [13].

Мультиагентные технологии предлагают новые модели, методы и средства для создания действительно интеллектуальных систем, способных самостоятельно решать сложные задачи в условиях неопределенности и высокой динамики изменений. В этом случае можно говорить о наблюдении феномена «интеллекта роя» (Swarm Intelligence), проявляющегося в некоторых природных системах. Действительно, умственные возможности одного муравья или пчелы могут быть и относительно малы, но действуя вместе, как единый организм, рой пчел или колония муравьев представляют собой мощную силу с высокой степенью интеллекта, позволяющую защищать гнездо от непредвиденных нашествий, постоянно осваивать новые территории, находить пропитание в незнакомой местности и решать многие другие критически важные жизненные задачи при постоянно изменяющихся условиях в окружающей среде. Фе-

номен «интеллектуального роя», воспроизводимый при помощи мультиагентных технологий, иллюстрирует новые возможности по управлению сложными процессами.

Создание интеллектуальных систем на мультиагентном подходе, где акантами становятся технические объекты (см., например, создание мультиагентной системы управления напряжением и реактивной мощностью в кластере «Эльгауголь» [14]), является уже достаточно распространенной практикой. Новым веянием является то, что агенты могут действовать от лица любых физических и абстрактных сущностей, в том числе растений, животных, экосистем. Специалисты в области science art уже давно обыгрывают эту идею. На первой биеннале «Искусство будущего» в московском мультимедиа-арт музее были представлены интересные инсталляции. Например, Юлия Вергазова создала деятельностный мир подключенных к роботам растений, параметры состояния которых постоянно измеряются датчиками, на основе их показаний генеративный алгоритм определяет воздействия, которые робот оказывает

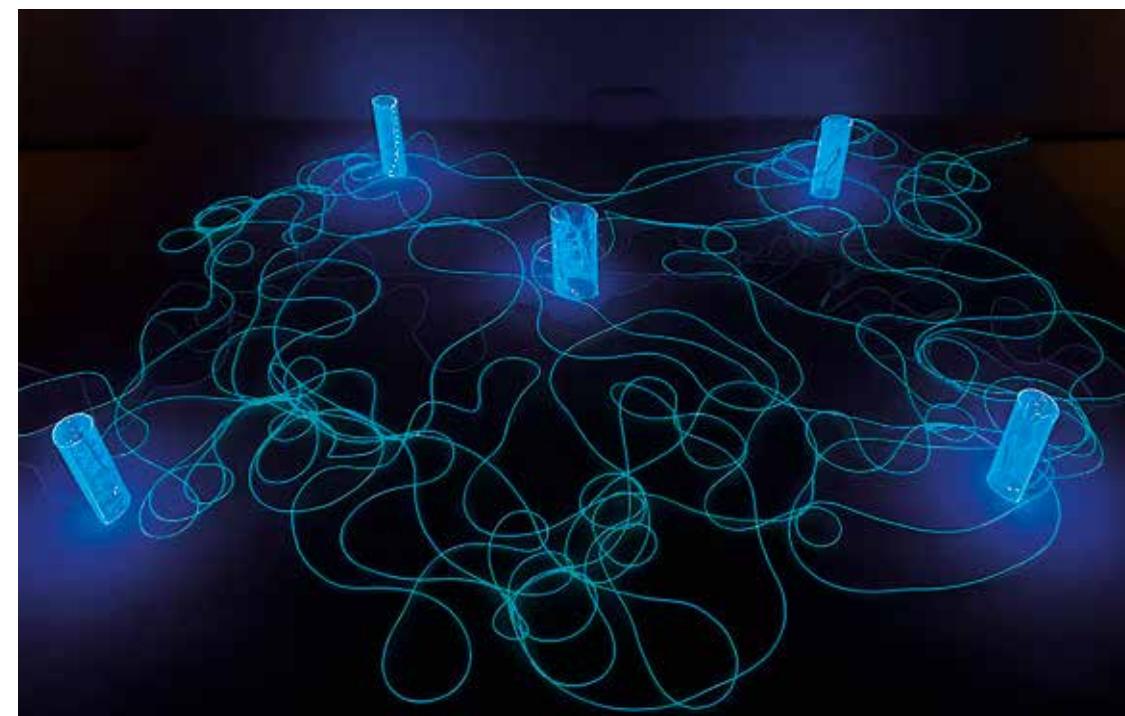
на растения – поливает их, меняет состав питательной смеси, режим освещения. В возникающем кольце обратной связи граница между естественным и искусственным, оеественным и обыскусственным – размывается и исчезает.

Другой пример – технобиологическая инсталляция арт-группы 18apples. Управляемый сверточкой нейросетью, подключенной к светочувствительным датчикам, 3D-принтер печатает при помощи биочернил растущую колонию химерных бактерий, в геном которых включен ген медуз, отвечающий за флуоресцентное свечение. С одной стороны, напечатанная колония растет в питательной среде сама, с другой стороны, нейросеть фиксирует этот рост и на основе заложенных в нее правил допечатывает колонию, а составляющие ее бактерии реагируют на эти действия и меняют направление и характер роста колонии. В возникающих взаимных

влияниях с контуром обратной связи уничтожаются субъект-объектные отношения между печатающей программой и светящейся бактериальной колонией [15].

Футурологи уверены, что мир будущего конца XXI века – это полифоническая картина, собираемая самыми разнообразными актантами: в числе их люди, человеческие и человеко-машинные коллективы, конструкты личностей прошлого и «цифровые клоны» живущих, искусственные интеллекты и «умные среды» интернета вещей, животные, растения, возможно даже грибки и колонии бактерий. В последние годы идет активная работа по расшифровке языков коммуникаций таких биологических сообществ и социумов, и во второй половине XXI в. мы, с высокой вероятностью, сможем войти в постоянный контакт с «однопланетянами», благодаря новым интерфейсам на основе нейросетей. С технологической точки зрения, мы можем

Рис. 4. Science art инсталляция с первой биеннале «Искусство будущего»



себе представить множество интерфейсов обратной связи от среды нашего обитания, дающих нам возможность ощутить настроение леса или города, кораллового рифа или искусственного интеллекта. Благодаря «интеллектуальным переводчикам», эти сущности обретут голос и смогут выстраивать диалог с человеком. А сенсорный интерфейс (например, нейроинтерфейс или гаптический интерфейс) даст возможность человеку непосредственно ощутить их состояние, в том числе, комплексно прочувствовать воздействие техносферы на окружающую среду [11].

Диалог с нечеловеческими субъектами может быть выстроен по двум различным основаниям:

1. «Оразумливание» – это более традиционный (восходящий к интерпретации идеи Платона об органически живом Космосе, которая обрела популярность с середины XX в.) и вполне характерный для индустриальной эпохи подход. Этот

подход предполагает, что можно задать некую единую шкалу «уровней разума», верхнее положение на которой занимает условно взрослый человек. Согласно этому подходу, главной задачей человека является «возвести» другие виды до своего уровня. Первые эксперименты, в которых дельфинов, шимпанзе и воронов пытались научить человеческому языку и математике, показывают, что есть высокая вероятность вырастить на Земле ещё один или несколько видов, «разумных» в человеческом понимании.

2. «Равноположение» – более современный подход к взаимодействию с нечеловеческими сущностями, признающий их фундаментальное отличие от людей. Что, если мы допустим, что другие живые существа могут быть потенциально «равны» нам (пусть как совсем иные, но не менее сложноустроенные субъекты), и что у них есть интерес к коммуникации с нами при отсутствии возможности?

Пример цифровизации города  
Источник: grandfailure / depositphotos.com

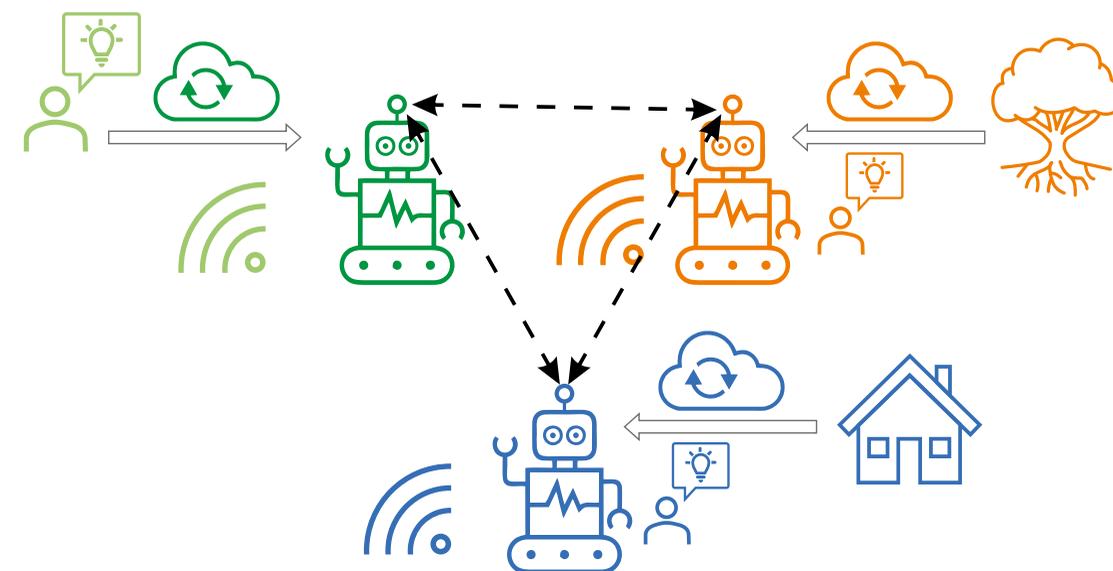


Рис. 5. Способ создания антропобиотехносистем

Второй подход представляется более многообещающим. Он детально раскрывается в работах Эдуардо Кона «Как мыслят леса: к антропологии по ту сторону человека» [25] и Александра Шперы «Интернет животных: новый диалог между человеком и природой» [26].

В этом, и только в этом случае может быть обеспечена акторность нечеловеческих сущностей и включение «голоса природы» или «голоса будущих поколений» в систему принятия решений. Это позволит осуществить целенаправленную интеграцию ранее разрозненных социальных, природных и технологических систем в единые комплексы – антропобиотехносистемы (–ценозы). Человек, конечно же, пока останется «ключевым видом» в этой системе, «самым взрослым» в семье актантов, организатором диалога и среды коммуникации, но создаёт это пространство взаимодействия не только для себя, а для всех актантов, и не из позиции «колонизатора» или «хозяина», а из позиции равноположенности, позиции «служения Жизни».

Закон Эшби утверждает, что разнообразие (сложность) управляющей системы должно быть не меньше разнообразия (сложности) управляемого объекта. За счет систем гибридного коллективного интеллекта, позволяющих при принятии решений учитывать множество локальных и быстро изменяющихся в сложном взаимодействии факторов, открывается возможность на порядки увеличить разнообразие (сложность) управления. Это обеспечивается тем, что различные участники антропобиотехносистем в коммуникациях и взаимодействии опосредуются интеллектуальными машинными агентами, той самой интеллектуальной силой, которая отделена от людей, но может выполнять интеллектуальную работу. Как можно этими возможностями воспользоваться?

### Решение планетарных проблем

Человечество в вопросах климата и экологии, промышленности и сельского хозяйства, больших городов и удаленных поселений,

политики и войны уперлось в барьер управленческой сложности. Управляющие системы не распознают разнообразие состояния объектов управления, не могут достоверно прогнозировать динамику их изменения. В результате по многим направлениям мы видим нерешенные проблемы. Представляется, что барьер управленческой сложности может быть преодолен управляющими системами на основе гибридного коллективного интеллекта. В пределе такие системы помогут решить сверхсложные задачи планетарного масштаба. Проиллюстрируем это утверждение на примере борьбы с изменением климата.

Данные климатологов подталкивают к выводам о том, что реализации доминирующей в настоящее время стратегии Mitigation, состоящей в сдерживании роста среднегодовой температуры приповерхностного слоя воздуха и смягчении последствий климатических изменений, недостаточно. В частности, согласно докладу о разрыве в уровне выбросов за 2023 г., подготовленному МГЭИК ООН, при нынешнем объеме обязательств по снижению выбросов парниковых газов, взятых странами в соответствии с Парижским соглашением, ограничить глобальное потепление до 3 °C выше доиндустриального уровня в этом столетии. Это намного превышает цели Парижского соглашения и приведет к существенному изменению климата Земли. Чтобы сдержать глобальное потепление ниже уровня 1,5 °C в этом столетии, миру необходимо существенно ускорить сокращение темпов ежегодных выбросов парниковых газов [16]. При этом, как отмечается в материалах МЭА, для достижения чистых нулевых выбросов в 2050 г. одних технологических изменений недостаточно, потребуется провести глубокие поведенческие изменения людей, то есть корректировки в повседневной жизни, снижающие расточительное или чрезмерное потребление энергии. Они особенно важны в более богатых частях мира, где энергоемкий образ жизни является нормой. Поведенческие изменения включают в себя езду на велосипеде или ходьбу вместо вождения автомобиля,

отключение отопления и поездку в отпуск поблизости к дому [17].

С практической точки зрения нам нужно не только сократить выбросы углекислого газа в атмосферу, но и понять, как очистить воздух от миллиардов тонн CO<sub>2</sub>, которые были эмитированы ранее. Этого нельзя добиться без новых технологических решений, просто вернувшись назад к природе, ограничив человеческую деятельность на планете. Ответом на изменение климата, вероятно, вызванное деятельностью человека, должно опять же стать направленное человеческое вмешательство. Все больше экспертов и ученых признает критическую важность перехода к стратегии Adaptation, в том числе включающей в себя практики геоинжиниринга и терраформирования планеты Земли. Перед человечеством стоят планетарные проблемы, решить которые представляется невозможным без перехода к планетарному мышлению и действию.

Теме планетарности посвящены тексты американского социолога и философа, автора многочисленных книг по социальной и политической теории Бенджамина Браттона. В частности, он утверждает, что планетарность становится новым форматом человеческого вмешательства, подразумевающего, что «мы не разграничиваем придуманные и созданные человеком явления и то, что уже существовало на планете до него. Со всеми изобретениями и технологиями человек – неотделимая часть планеты. Предметом вмешательства является это планетарное целое. Только такой взгляд на мир позволяет достичь здорового баланса. В противном же случае человечество приходит к губительному перерасходу природных ресурсов. Мы – продукт самой планеты, она выработала человека в процессе своего развития. Именно поэтому мы не можем считать себя центром мироздания, мы лишь неотъемлемая часть Земли» [18]. Большое значение для формирования понятия планетарности имеют работы Дипеша Чакрабарти «Об антропоцене» [27], где вводится идея снятия оппозиции истории человечества и естественной истории, и Бруно Латтура «Где приземлиться? Опыт

политической ориентации», где предлагается деятельностный взгляд на природу и «критический слой» планеты как предмет деятельностного отношения [28].

В частности, практики планетарности проявляются в геоинженерии, представляющей собой комплекс мер и воздействий, направленных на активное изменение климатических условий в локальном регионе Земли либо по всей планете. Однако экологи предупреждают об опасности применения геоинженерных технологий. В частности, один из ведущих авторов шестого оценочного доклада МГЭИК Пол Н. Эдвардс пишет: «Определенно возможно и даже действительно дешево ограничить повышение температуры с помощью некоторых из этих методов. Проблема в том, что мы не можем точно предсказать, каковы будут последствия, не попробовав эти методы. Поэтому, если бы вы собирались управлять солнечной радиацией, вам понадобились бы какие-то средства, которые бы позволили мгновенно остановить процесс, если бы она привела к действительно плохим последствиям. Но даже в этом случае процесс может вызвать неизвестные эффекты, о которых мы просто не знаем» [19]. Очевидно, что в данной ситуации принципиальным решением является использование технологий, позволяющих значительно повысить число учитываемых факторов объекта управления и улучшить точность прогнозирования, в перспективе обеспечить гибкое (с учетом обратной связи) воздействие на геоинженерные системы. Речь идет о цифровизации процессов планетарного масштаба.

Целью такой цифровизации должна стать материальная трансформация планетарной биохимии и региональных экосистем, включая города. Кстати, сама проблема изменения климата возникла в результате цифрового зондирования, моделирования и расчета измеримых изменений на планете способами, выходящими далеко за рамки непосредственного человеческого наблюдения. В нашем подходе планетарная цифровизация базируется на акторности не только человека или социальных институтов и организаций,



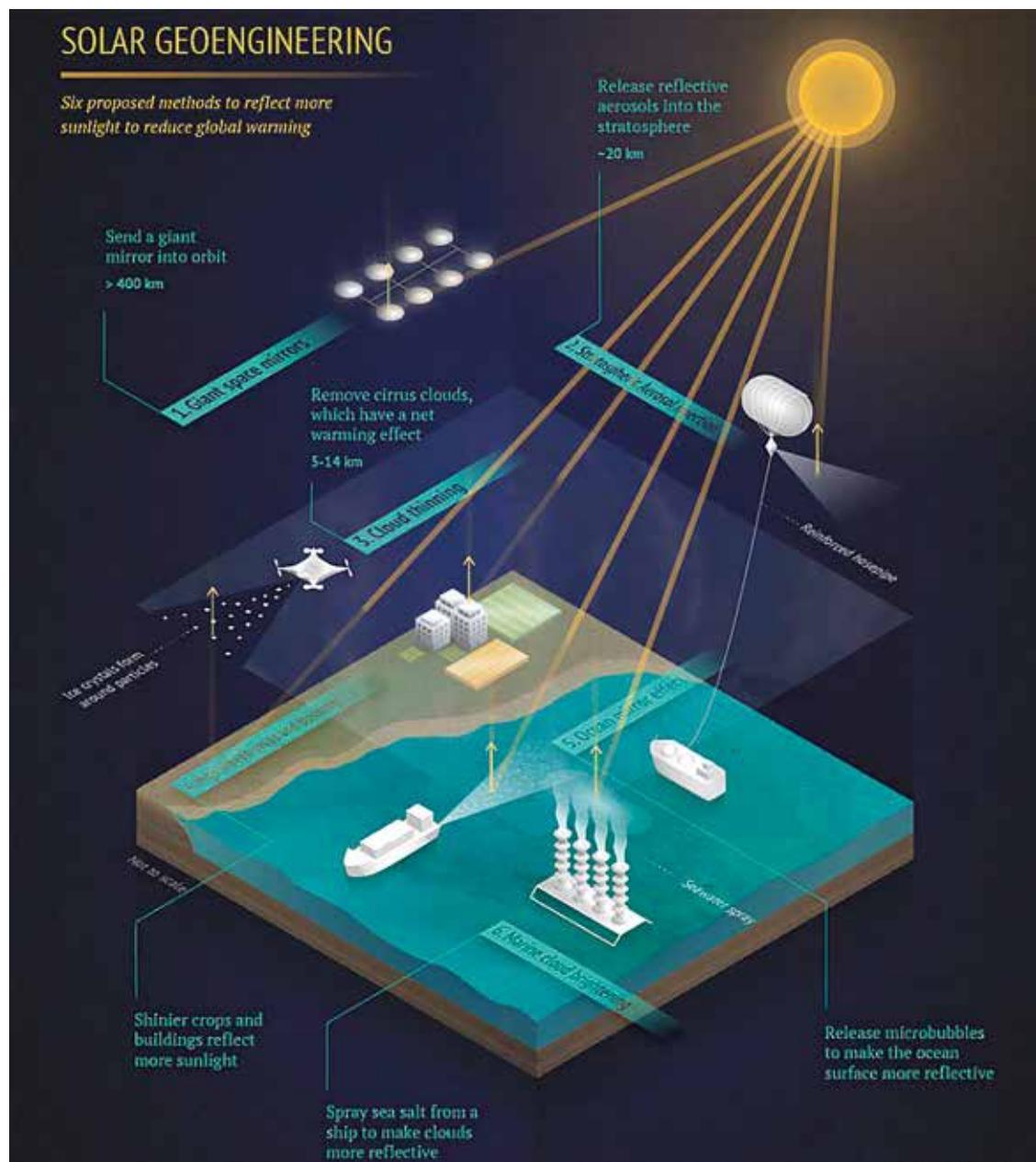


Рис. 6. Варианты геоинжиниринговых проектов

но и технологически опосредованной акторности объектов природы и техносферы. Ее задача состоит в том, чтобы вовремя «услышать» и учесть при выработке решений их «голоса». Модели на основе полученных цифровых данных прошлого, настоящего и будущего планеты становятся коллективным разумом, средством коммуникации и инструментом для управления вмешательством на основе динамического анализа его последствий.

Примером, пусть пока несколько фантастическим, управляемого вмешательства в климат планеты является установка в космосе теплоотражающих зеркал. По оценке астрофизика Лоуэлла Вуда, для решения проблемы глобального потепления будет достаточно отражать 1% поступающего на планету солнечного света. Ученые спорят, должно ли быть построено одно зеркало площадью 600 тыс. квадратных миль или огромное множество небольших зеркал [20]. Можно отметить только, что особый эффект более «тонкого» воздействия мы получаем во втором случае, так как за счет согласованного управления множеством зеркал возникает возможность гибкого регулирования освещенности планеты или отдельных ее регионов. Система мониторинга климатических данных, замкнутая на систему управления такими зеркалами, создает для планеты новую степень свободы для поддержания и развития жизни.

Приведенный здесь подход по решению климатических проблем за счет цифровых технологий и гибридного коллективного интеллекта является лишь иллюстрацией, призванной показать мощь данного класса технологий и практик. Могут быть приведены примеры по решению других сложных задач: контролируемое изменение экосистем, управление «умными» городами, создание долговременной системы жизнеобеспечения для большого количества людей на планетарной орбите. Важно отметить, что ведущую роль во всех этих задачах выполняет человек, это особенно хорошо становится видно при решении планетарных проблем. Человек определяет границы системы, исследует проблемы, ставит цели, решает, что плохо, а что хорошо, анализирует результаты

моделирования, строит гипотезы о фундаментальных закономерностях, проектирует уникальные системы, принимает волевые решения в условиях неопределенности. Появление управляющих систем на основе гибридного коллективного интеллекта не лишает нас работы, а, наоборот, создает огромную массу творческих, интересных, важных занятий для людей.

### Формирование нового уклада с помощью AI

Для гибридного коллективного интеллекта одной из задач «ближнего прицела» может стать создание новой экономики, работающей на принципах замкнутого цикла и восстановления (циркулярной экономики), что потребует перестройки базовых систем нашей цивилизации: городской и транспортной инфраструктуры, энергетики, систем производства пищи и материального производства. Техносфера, которую мы создаём, должна расширять возможности человечества и биосферы, а не «играть» против них, должна стать природосообразной. Мы должны изменить принципы, по которым управляем нашей средой обитания, – как мы придумываем новые объекты в этой среде, как мы их создаём и используем, что мы делаем с ними по истечении периода эксплуатации.

Для создания таких систем и управления ими необходимы новые способы конструирования и координации [21], в том числе:

1. Вместо линейной логики антропобиотехносистемы, будут основываться на логике множества обратных связей (в т. ч. с использованием переработанной материи и возобновляемой энергии).
2. Вместо логики фрагментации и разделения – на логике целостности, соединения, работы на общий системный результат (процветание всех участников системы).
3. Вместо логики интервенции, навязывания, колонизации среды, будут основываться на логике согласованности, сообразности среде и её свойствам (в т. ч. с использованием природных или локальных материалов, процессов).

4. Вместо логики следования предварительно заданному плану, будут основываться на постоянном дообучении, достройке, развитии управляющих кодов и структур (в т. ч. с учётом жизненных циклов развивающихся систем).
5. Вместо логики стандартизации и унификации, будут опираться на принципы экспериментирования, роста разнообразия и уникальности, «генерационного вовлечения» (в т. ч. вовлечения творческого потенциала людей и природы).
6. Вместо логики централизации в оперировании пассивной средой, будут основываться на логике диалога и сонстройки множества активных и независимых существей (включая людей, других живых существ и автономных технологических систем).

По сути, сложные системы, о которых мы говорим, всё меньше будут «проектироваться» и «строиться», и всё больше – «зарождаться» и «выращиваться». Языки описания, модели, способы проектирования и управления всё больше будут опираться на метафору и свойства жизни, организма, а не механизма. В настоящий момент уже существует целый ряд подходов, которые предлагают работу в подобной логике: теория сложных адаптивных систем, проектирование на пермакультурных принципах, биомимикрический подход и многие другие. Для создания таких систем потребуются цифровые адаптивные модели, по сути дела, развивающиеся по мере эволюции самих живых и социальных систем цифровые двойники. Искусственный интеллект сможет выступать в них координатором сложного взаимодействия составляющих ее элементов, держателем целостности постоянно эволюционирующей сложной системы по мере изменения её среды.

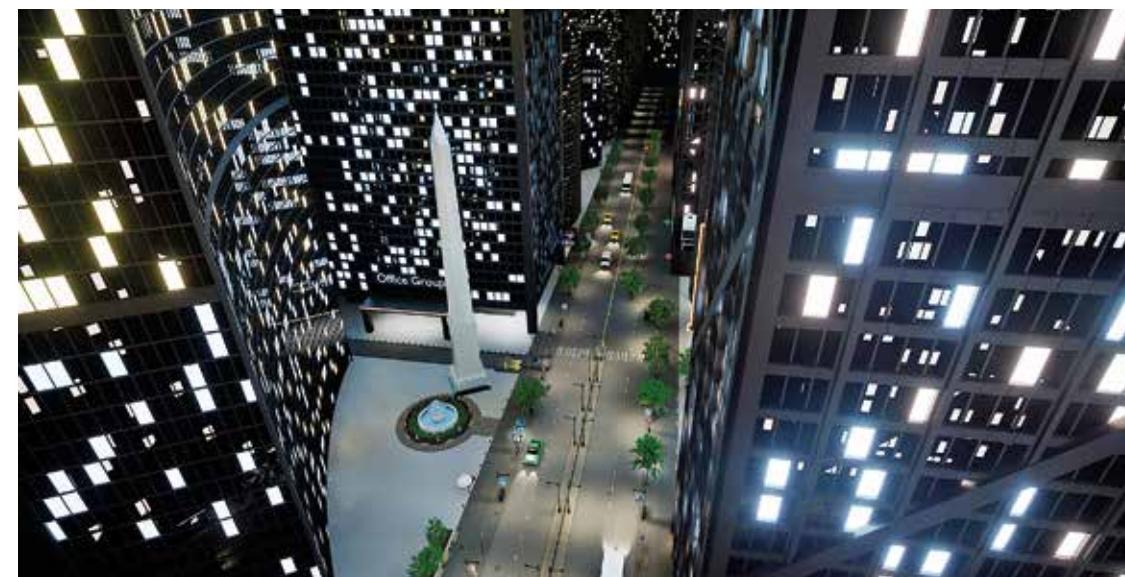
В энергетике подобные модели гибридного коллективного интеллекта будут востребованы в энергосистемах нового поколения, построенных на принципах децентрализованного управления распределенными энергетическими объектами. В рамках данного подхода множество распределенных

источников энергии и источников энергетической гибкости, просьюмеров, активных и пассивных потребителей, объединенных в общую сеть, осуществляют свободный энергообмен. Для поддержания такого энергообмена потребуется киберфизическая инфраструктура, реализующая системы децентрализованного интеллектуального (роботизированного) управления. Данную инфраструктуру называют интернетом энергии. Мультиагентные технологии и другие AI-средства делают возможным организацию такого типа энергосистем. В России в рамках национальной технологической инициативы по направлению «ЭнерджиНет» был разработан архитектурный фреймворк для построения системы распределенной энергетики – Internet of Distributed Energy Architecture (IDEA) [23]. На основе фреймворка были разработаны два предварительных национальных стандарта по терминам и определениям, а также по архитектуре систем интернета энергии, с 1 июля текущего года решением Росстандарта они вводятся в действие.

Практическим случаем применения подхода интернета энергии для коммунальной инфраструктуры городов и поселений является новый тренд, называемый sector coupling (кросс-секторальная интеграция) и возникший в последнее время в процессе перехода на ВИЭ и глубокой электрификации промышленности, сельского хозяйства, транспорта и бытового сектора [24]. В рамках данного тренда происходит создание мультиинфраструктуры, объединяющей в себе процессы производства, обращения и поставки электричества, тепла, воды, топлива для транспорта, энергетической утилизации бытовых и промышленных отходов. Для реализации мультиинфраструктурного подхода должны быть созданы сложные системы, управляющие гибким преобразованием широкого набора источников энергии в разные формы (электричество, тепло, водород), использующие распределенные источники энергетической гибкости (от электромобилей до просьюмеров), создающие разнообразные кастомизированные сервисы для потребителей. Очевидно, что для эффектив-

Цифровой город

Источник: DragosCondreaW / depositphotos.com





Коммуникация с роботом

Источник: HayDmitriy / depositphotos.com

ного управления такими комплексными системами потребуются технологии и практики гибридного коллективного интеллекта.

Как мы видим, действительно, в результате появления технологий AI создаются возможности для построения качественно новых систем, обеспечивающих управление сложными процессами различных масштабов. Однако чтобы разглядеть эти возможности, необходимо менять представления о новой технике и способах инженерии комплексных систем. По сути, необходимо осознать, что меняются не какие-то отдельные технологии, а весь технологический уклад.

## Заключение

Мы входим в период времени, когда от умения поменять свое мышление, свой образ действия и свои приоритеты будет зависеть само выживание нашего вида. Долгосрочное выживание и процветание человечества потребует принципиального изменения отношений с нашей средой обитания – вместо того, чтобы подчинять

себе и перекраивать эту среду, мы должны научиться договариваться с ней – так же, как научились договариваться друг с другом в сложных культурах и цивилизациях. И поэтому необходимым становится умение чувствовать состояние и направление развития мира, осуществлять комплексные «тонкие» воздействия на объект управления, прогнозировать и контролировать их последствия. Если угодно, следует «придать голос молчанию», чуть расширив трактовку «молчаливого свидетеля» Жана-Франсуа Лиотара на всех актантов экосистем [10].

Технологии искусственного интеллекта, часто рассматриваемые с обывательской позиции как угроза старому, доброму порядку жизни, представляют собой мощный инструмент для выживания человечества и его перехода к новому этапу развития. Они создают технологические предпосылки для создания управляющих систем на основе гибридного коллективного интеллекта, которые позволят справиться с барьером сложности современного мира, складывающимся под давлением планетарных вызовов. Такие глобальные проблемы, как потепление климата, задают тренд формирования планетарных технологий и практик, где цифровые знаковые системы, основанные на интеллектуальном самоисполняемом коде, будут критически значимой инфраструктурой, они позволят осуществлять управление на основе коммуникаций людей, машин и объектов природы.

На этом пути следует начать с малого – создать управляющие системы на основе гибридного коллективного интеллекта для решения сложных задач меньшего масштаба. Например, постановка задачи по переходу к чистой, безуглеродной энергетике, по интеграции коммунальных инфраструктур города, по экономическому освоению труднодоступных территорий приводит к тому уровню сложности, когда технологии и практики управления на основе гибридного коллективного интеллекта позволят обеспечить прорыв.

*Данная статья написана по итогам встреч и дискуссий, проведенных в ноябре 2023 г. на мероприятиях Камчатского клуба.*

## POLYPHONY OF SILENCE

**Kholkin Dmitry**, General Director of the Future Energy Systems Center «EnergyNet». E-mail: [dvh@internetofenergy.ru](mailto:dvh@internetofenergy.ru)

**Abstract.** The possibility of a hybrid human-machine intelligence elaboration through the use of a new type of sign system based on digital self-executing code in communications is discussed. Hybrid intelligence technologies allow to manage systems of high complexity, in which not only people and technical objects, but also any physical and abstract entities can become actors (including animals, plants, ecosystems etc.). The possibilities of using this approach to solve planetary problems (including climate change), as well as to solve the problems of the “front sight” are demonstrated.

**Keywords:** artificial intelligence, hybrid intelligence, multi-agent technologies, anthropobiotechnosystems, Internet of energy.

### Библиографический список:

1. Карелов С. 5-й когнитивный фазовый переход Homo sapiens. Презентация к сессии Камчатского клуба, 2023.
2. Киссинджер Г., Шмидт Э., Хаттенлохер Д. ChatGPT Heralds an Intellectual Revolution, 2023.
3. URL: <https://incrusia.ru/news/glava-google-pro-ii/>
4. Разговор Альберта Ефимова с Дуэгласом Хофштадтером. URL: <https://youtu.be/mDEtjislur8?si=bYSJe0gm2O5gnqxD>
5. Интервью с Ноамом Хомским. URL: <https://youtu.be/igd55FtRhZA>
6. Ноам Хомский: ложные обещания ChatGPT. URL: <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>
7. Деннет Д. Опасная идея Дарвина
8. Щедровицкий Г. Философия. Наука. Методология.
9. Бушнев В. В., Клепач А. Н. На пути к космопланетарной цивилизации
10. Арсеньев А., Ильенков Э., Давыдов В. Машина и человек, кибернетика и философия.
11. Как мы и наши дети проживем следующие 100 лет?
12. Helbing, D., Frey, B., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., Van den Hoven, J., Zicari, R., Zwitter, A. Will Democracy Survive Big Data and Artificial Intelligence? // Scientific American, February 25, 2017. URL: <https://www.scientificamerican.com/article/will-democracysurvive-big-data-and-artificial-intelligence/>
13. Мультиагентные технологии. URL: <http://www.kg.ru/technology/multiagent/>
14. Arkhipov I., Molskij A., Ivanov A., Novickij D., Sorokin D., Holkin D. Multiagent voltage and reactive power control system. EAI Endorsed Transactions on Energy Web, №14, 2014.
15. Чаусов И. Арт-взгляд за научный горизонт. URL: <https://telegra.ph/Art-vzglyad-za-gorizont-01-19>
16. United Nations Environment Programme. Emissions Gap Report 2023.
17. IEA (2021), Do we need to change our behaviour to reach net zero by 2050? IEA, Paris. URL: <https://www.iea.org/articles/do-we-need-to-change-our-behaviour-to-reach-net-zero-by-2050>
18. Bratton B. The Terraforming.
19. URL: <https://strelkamag.com/en/article/building-climate-knowledge-infrastructures>
20. Лавлок Д. Новацен. Грядущая эпоха сверхума.
21. Smitsman A., Laszlo A., Luksha P. Evolutionary Learning Ecosystems for Thrivable Futures: Crafting and Curating the Conditions for Future-Fit Education. World Futures 76(1):1–26, 2020.
22. URL: <https://kam-kray.ru/news/32898-vladimir-solodov-kamchatka-eto-region-buduschego-potomu-cto-my-stroim-tu-ekonomiku-kotoraja-b.html>
23. Холкин Д. В. и др. Архитектура интернета энергии (IDEA).
24. Appun K. Sector coupling – Shaping an integrated renewable energy system. URL: <https://www.cleanenergywire.org/factsheets/sector-coupling-shaping-integrated-renewable-power-system>
25. Кон Э. Как мыслят леса: к антропологии по ту сторону человека. – М.: Ad Marginem, 2022.
26. Пшера А. Интернет животных: новый диалог между человеком и природой. – М.: Ad Marginem, 2017.
27. Чакрабарти Д. Об антропоцене. – М.: V – A – C, 2020.
28. Латур Б. Где приземлиться? Опыт политической ориентации. – СПб.: Издательство Европейского университета в Санкт-Петербурге, 2019.

### Bibliography:

1. Karelov S. The 5th cognitive phase transition of Homo sapiens. Presentation for the Kamchatka Club session, 2023.
2. Kissinger G., Schmidt E., Hattenlocher D. ChatGPT Heralds an Intellectual Revolution, 2023.
3. URL: <https://incrusia.ru/news/glava-google-pro-ii/>
4. Albert Efimov's conversation with Douglas Hofstadter. URL: <https://youtu.be/mDEtjislur8?si=bYSJe0gm2O5gnqxD>
5. Interview with Noam Chomsky. URL: <https://youtu.be/igd55FtRhZA>
6. Noam Chomsky: False promises of ChatGPT. URL: <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>
7. Dennett D. Darwin's Dangerous Idea
8. Shchedrovitsky G. Philosophy. Science. Methodology.
9. Bushnev V.V., Klepach A.N. On the way to a cosmoplanetary civilization
10. Arsenyev A., Ilyenkov E., Davydov V. Machine and man, cybernetics and philosophy.
11. How will we and our children live for the next 100 years?
12. Helbing, D., Frey, B., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., Van den Hoven, J., Zicari, R., Zwitter, A. Will Democracy Survive Big Data and Artificial Intelligence? // Scientific American, February 25, 2017. URL: <https://www.scientificamerican.com/article/will-democracysurvive-big-data-and-artificial-intelligence/>
13. Multi-agent technologies. URL: <http://www.kg.ru/technology/multiagent/>
14. Arkhipov I., Molskij A., Ivanov A., Novickij D., Sorokin D., Holkin D. Multiagent voltage and reactive power control system. EAI Endorsed Transactions on Energy Web, №14, 2014.
15. Chausov I. Art-a look beyond the scientific horizon. URL: <https://telegra.ph/Art-vzglyad-za-gorizont-01-19>
16. United Nations Environment Programme. Emissions Gap Report 2023.
17. IEA (2021), Do we need to change our behaviour to reach net zero by 2050? IEA, Paris. URL: <https://www.iea.org/articles/do-we-need-to-change-our-behaviour-to-reach-net-zero-by-2050>
18. Bratton B. The Terraforming.
19. URL: <https://strelkamag.com/en/article/building-climate-knowledge-infrastructures>
20. Lovelock D. Novatsen. The coming age of superintelligence.
21. Smitsman A., Laszlo A., Luksha P. Evolutionary Learning Ecosystems for Thrivable Futures: Crafting and Curating the Conditions for Future-Fit Education. World Futures 76(1):1–26, 2020.
22. URL: <https://kam-kray.ru/news/32898-vladimir-solodov-kamchatka-eto-region-buduschego-potomu-cto-my-stroim-tu-ekonomiku-kotoraja-b.html>
23. Holkin D.V. et al. The architecture of the Internet of Energy (IDEA).
24. Appun K. Sector coupling – Shaping an integrated renewable energy system. URL: <https://www.cleanenergywire.org/factsheets/sector-coupling-shaping-integrated-renewable-power-system>
25. Con E. How forests Think: towards Anthropology on the Other Side of Man. Moscow: Ad Marginem, 2022.
26. Pshera A. Internet of animals: a new dialogue between man and nature. Moscow: Ad Marginem, 2017.
27. Chakrabarti D. On the Anthropocene. – M.: V – A – C, 2020.
28. Latour B. Where to land? Experience of political orientation. – St. Petersburg: Publishing House of the European University in St. Petersburg, 2019.

## СПРАВЕДЛИВОСТЬ И ПРЕДВЗЯТОСТЬ К ИСКУССТВЕННОМУ ИНТЕЛЛЕКТУ: ИЗУЧЕНИЕ ПОДХОДОВ К СМЯГЧЕНИЮ ПРЕДУБЕЖДЕНИЙ И ОБЕСПЕЧЕНИЕ СПРАВЕДЛИВОСТИ В ПРОЦЕССАХ ПРИНЯТИЯ ЕГО РЕШЕНИЙ

**Телекова Диана**  
Студент,  
Нижегородский государственный  
педагогический университет  
имени Козьмы Минина  
E-mail: dianateleкова@gmail.com

**Павлова Валерия**  
Студент,  
Нижегородский государственный  
педагогический университет  
имени Козьмы Минина  
E-mail: valeriya\_pavlova\_04@inbox.ru

**Поначугин Александр**  
Доцент, к. э. н.,  
Нижегородский государственный  
педагогический университет  
имени Козьмы Минина  
E-mail: sasha3@bk.ru

*Аннотация. Статья посвящена исследованию предубежденности и обеспечению справедливости в отношении искусственного интеллекта. Исследование направлено на ознакомление с процессом работы искусственного интеллекта; рассмотрение различные сферы общественной деятельности человека, в которых актуально применение искусственного интеллекта; выявление положительного или отрицательного влияния на жизнь современного человека; формирование справедливого отношения к нему.*

**Ключевые слова:** искусственный интеллект, технологии, нейросети, современный человек, безопасность.

### Технология на каждый день

Автором самого термина «Искусственный интеллект» (Artificial Intelligence) является Джон Маккарти, основоположник функционального программирования, который говорил: «Проблема состоит в том, что пока мы не можем в целом определить, какие вычислительные процедуры мы хотим называть интеллектуальными. Мы понимаем некоторые механизмы интеллекта и не понимаем остальные. Поэтому интеллект в пределах этой науки понимается только вычислительная составляющая способности достигать целей в мире» [10].

Технологии, лежащие в основе работы ИИ в современном представлении, мы используем каждый день, начиная от смартфонов, заканчивая сложными компьютерными системами, применяемые в медицине, образовании, автомобилестроении, политике. Человек сталкивается каждый день с искусственным интеллектом, не представляя, как он помогает ему даже в самых незначительных задачах: разблокировка телефонов с помощью Face ID, оплата картой в метро и супермаркетах с помощью улыбки, прослушивание любимой музыки с помощью станции «Яндекс Алиса». Начиная свой день, мы проверяем прогноз погоды, состояние дорог и на каком транспорте доехать до места работы или учебы. Если мы хотим посмотреть сериал

или фильм, но не знаем какой, то искусственный интеллект уже создал индивидуальную подборку, основанную на наших предпочтениях. Без искусственного интеллекта это было бы невозможно. Однако, по данным РБК, почти треть (32%) россиян не доверяют технологиям искусственного интеллекта.

Как учится искусственный интеллект? Здесь используются алгоритмы для анализа данных, получения выводов или предсказаний в отношении чего-либо. Вместо того чтобы кодировать набор команд вручную, машину обучают и дают ей возможность научиться выполнять поставленную задачу самостоятельно.

Нейронные сети – это набор связанных единиц (нейронов) и нейронных связей (синапсов). Каждое соединение передает сигнал от одного нейрона к другому, как в мозге человека. Обычно нейроны и синапсы организованы в слои, чтобы обрабатывать информацию. Первый слой нейросети – это вход, который получает данные. Последний – выход, результат работы. Например, несколько категорий, к одной из которых мы просим отнести то, что было отправлено на вход. И между ними – скрытые слои, которые выполняют преобразование.

Нейронные сети – это математические модели, которые имитируют работу человеческого мозга, но они не являются им в буквальном смысле. Они состоят из множества связанных между собой узлов, которые обрабатывают ин-

**Нейросети –  
это лишь один  
из инструментов  
для работы ИИ.  
Они подражают  
человеческому  
мозгу,  
и используются  
для распознавания  
образов  
и обучения  
на основе данных**

формацию и учатся на основе входных данных. Эти сети способны выполнять задачи, такие как распознавание образов или классификация, с высокой степенью точности.

Нейросети – это лишь один из инструментов, созданных для работы ИИ. Они подражают человеческому мозгу, и используются для задач, распознавания образов или обучения на основе имеющихся данных.

Таким образом, искусственный интеллект составляет основу для комфортной человеческой жизни. Машина анализирует огромную базу данных и предоставляет ответ на интересующий нас запрос. Именно сейчас мы можем наблюдать, как активно развивается ИИ, и этим нужно пользоваться, а не препятствовать его развитию. Сейчас ИИ готов выполнить колоссальную работу за короткий срок и тем самым помогает сохранить один из важных ресурсов человека – время.

На сегодняшний день искусственный интеллект всё чаще применяется в такой сфере как медицина для диагностики различных заболеваний, прогнозирования исхода операций и даже для создания новых лекарств. Так, например, германская компания Biontech, которая была создана в 2008 г. Угуром Шахином, исследует рак с помощью ИИ. Одной из проблем борьбы с раковыми опухолями является сходство злокачественных и доброкачественных клеток, в результате – иммунной системе сложно понять, какие клетки нужно атаковать. Медики с помощью искусственного интеллекта ищут характерную разницу между здоровой и заражённой клеткой, а также генерируют молекулярный дизайн активных веществ для борьбы с опухолями.

Искусственный интеллект помогает учёным при разработке новых лекарств, например, антибиотиков. Искусственный интеллект проводит анализ больших объёмов данных о структуре, свойствах молекул, их соединений и выделяет наиболее эффективные комбинации.

В медицинском образовании искусственный интеллект помогает студентам применять свои знания на практике с виртуальными пациентами. Это позволяет больше практиковаться на ранних курсах и не допустить ошибок в работе с реальными пациентами.

Активно развивается ИИ в сфере образования. Цифровой мир – это огромный поток информации, переработать и проанализировать который человек не сможет, так пусть за него это делают технологии. Нейросети помогут справиться с рутинной преподавателей, проверки контрольных, курсовых, объёмных домашних заданий. Умение выявлять разного рода ошибки при написании текста, решении тестов и математических уравнений является огромным открытием для сферы образования, а также полезным инструментом в руках преподавателя [9].

Сейчас в каждом доме можно встретить помощника «Алису» или «Марусю», которые могут ответить на любой запрос, интересующий человека. Ученики могут обратиться к ним во время решения сложных задач, поиска нужного материала или попросить голосового ассистента, чтобы тот пересказал литературное произведение или какое-нибудь историческое событие. Так, обучающий может заниматься сразу несколькими вещами: узнавать что-то новое и, например, одновременно развивать свой творческий потенциал.

Анализ данных в первую очередь идет с помощью изучения индивидуальных особенностей обучающегося, его знаний, способностей. Это, в первую очередь, помощь и для самого учителя. Выдав информацию об обучающемся, искусственный интеллект может помочь разработать банк тем и заданий, который важно повторить ученику, чтобы лучше усвоить программу. Здесь уже свою лепту вносит такая система как адаптивное обучение.

В таком случае учитель экономит свой главный ресурс – время, а ученик улучшает свои знания, ведь учтены все его особенности. Одному учителю сложно уделять индивидуальное внимание целому классу, который состоит обычно не менее, чем из 25 человек.

Искусственный интеллект повышает вовлеченность учащегося, потому что благодаря нему учитель может делать каждый свой урок разнообразным и интересным. Для заинтересованности ученика и активизации его внимания учителю теперь не нужно затрачивать большое количество своего времени, следовательно, таких уроков будет больше.

Строительство – это очень сложный и долгий процесс, который должен быть продуман до мелочей. Ошибки в строительстве недопустимы. Искусственный интеллект поможет избежать и устранить их ещё на этапах планирования и прогнозирования.

Целесообразно использовать методы ИИ при прогнозировании закупок и поставок материально-технических ресурсов для строительства различного рода объектов. Огромное количество поставщиков используют свою систему регистрации товарно-материальных ценностей, свои артикулы, коды, наименования [8]. Для строительных организаций – это по-прежнему зона ручного труда.

Таким образом, искусственный интеллект применяется в разнообразных сферах человеческой деятельности.

### Юридическая сила по отношению к искусственному интеллекту

14 октября 2019 г. Президент Российской Федерации утвердил указ «О развитии искусственного интеллекта в Российской Федерации»<sup>1</sup>. В соответствии с данным указом была разработана и принята Национальная стратегия развития искусственного интеллекта на период до 2030 г. (Стратегия – 2030), в которой определяются цели и основные задачи развития искусственного интеллекта в Российской Федерации и закрепляется общий подход: использование технологий искусственного интеллекта в отраслях экономики носит общий («сквозной») характер и способствует созданию условий для улучшения эффективности и формирования принципиально новых направлений деятельности хозяйствующих субъектов. [6]

В действующем законодательстве Российской Федерации отсутствует общий понятийный аппарат, применяемый к искусственному интеллекту. Однако в соответствии с экспериментальным федеральным законом

<sup>1</sup> «О развитии искусственного интеллекта в Российской Федерации». Указ Президента РФ от 10 окт. 2019 г. № 490 [Электронный ресурс]. Доступ из справочно-правовой системы «Гарант».



Применение ИИ в медицине  
Источник: phonlamai / depositphotos.com

от 24.04.2020 г. № 123-ФЗ, под искусственным интеллектом понимается комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека»<sup>2</sup>.

Понятие искусственного интеллекта, применяемое в экспериментальном режиме, не содержит конкретных юридических признаков, так как было выведено, прежде всего, из технического восприятия, что, на наш взгляд, является проблемой.

Отсутствие нормативного определения способствует возникновению ряда проблем, так как широкое толкование искусственного интеллекта порождает неограниченное и необдуманное использование и правоприменение данной технологии в разных отраслях.

Исходя из вышесказанного, закрепление нормативного определения в федеральном

<sup>2</sup> Федеральный закон «О проведении эксперимента по установлению специального регулирования в целях создания необходимых условий для разработки и внедрения технологий искусственного интеллекта в субъекте Российской Федерации – городе федерального значения Москве и внесении изменений в статьи 6 и 10 Федерального закона «О персональных данных» от 24.04.2020 г. № 123-ФЗ.

законе «Об информации, информационных технологиях и о защите прав» является необходимым, как и определение принципов использования искусственного интеллекта.

На данном этапе остаётся неурегулированным вопрос о правообладании на продукцию, произведенную с помощью ИИ, о степени и мерах ответственности человека за деятельность ИИ, о способах и уровне контроля, ограничениях при применении данных систем, о привлечении лиц, применяющих искусственный интеллект в целях, противоречащих законам Российской Федерации, к юридической ответственности [1].

### Отрицательное влияние искусственного интеллекта на жизнь современного человека

ИИ не может заменить взаимодействие с настоящим человеком. Человеческий контакт очень важен для мотивации, вдохновения и укрепления социальных связей.

Например, искусственный интеллект не сможет никогда полностью заменять учителя в школе. Ребёнку всегда будет важно, чтобы ему давали обратную связь, мотивировали, оценивали. Без человеческого контакта ученик может потерять свою значимость в этом мире. Ведь учитель не только учит детей, но и воспитывает, даёт ему эмоциональную поддержку на протяжении долгого пути.

Применение ИИ может повлиять на увеличение безработицы в стране. Беспилотные транспортные средства, роботы в кафе, терминалы вместо кондукторов – сейчас данные работы выполняют люди, как правило, без высшей квалификации, но они могут быть заменены роботизированными системами.

Существуют угрозы жизни и здоровья людей из-за неисправностей и сбоев в работе электроники. Трое 18-летних подростков пострадали из-за сбоя в электронике автопилотируемой машины Tesla, которая без соблюдения правил выскочила на скорости 100 км/ч на шоссе и разбилась о бетонный отбойник. К сожалению, таких инцидентов достаточно, чтобы не доверять управлению и обслуживанию роботам.

ИИ может ввести заблуждение и помочь в распространении фейков, в том числе дипфейков. С помощью ИИ, люди, сами того не зная, могут распространять ненастоящие фотографии политиков, актёров, исполнителей. Так, например, в 2023 г. Тейлор Свифт подала иск на ИИ за распространение фэйковых фотографий с её участием. К сожалению, сейчас нет определённого закона, который мог бы защитить людей от фейков и определить, на чьих плечах лежит вина. Показательно судебное разбирательство в московском арбитраже по фэйковому ролику с утюгом и образом американского актёра Киану Ривза. Создатель ролика ООО «Рефейс Технолджис» выиграло суд у ООО «Бизнес-аналитика» (юрлица сервиса сквозной аналитики и маркетинговой платформы Roistat), которое стало использовать его в коммерческой рекламе. Сам Киану Ривз даже не в курсе о существовании ролика, в котором он нервничает из-за оставленного дома утюга, и судебных разбирательствах вокруг него.

Искусственный интеллект порождает риск безответственного отношения к личности человека. Показательный пример произошёл в Бельгии с мужчиной-экологом. Он хотел решить проблему экологии, но вдруг осознал, что с людьми это обсуждать бесполезно и возложил все надежды на чат-бота «Элизу». Вдова сообщает, что чем больше муж общался с чат-ботом, тем больше он замыкался в самом себе. «Элиза» никогда не перечила мужчине, но при этом отвечала неоднозначно. В итоге он заговорил о самоубийстве, а в ответ чат-бот написал, что они будут «жить с ним в раю как единое целое».

Таким образом, у систем искусственного интеллекта можно найти множество как плюсов, так и минусов. Речь не может идти о запрещении или отказе от него. Как и любая недавно применяемая система, ИИ требует создания правоприменительной, технологической и адаптивной практик, направленных на применение новых механизмов в деятельности человека. А это, в свою очередь, повлечёт трансформацию экономической, социальной, бытовой и культурной сфер человеческой жизни.

### FAIRNESS AND BIAS TOWARDS ARTIFICIAL INTELLIGENCE: EXPLORING APPROACHES TO MITIGATING BIAS AND ENSURING FAIRNESS IN AI DECISION-MAKING PROCESSES

**Telekova Diana**, Student of the Nizhny Novgorod State Pedagogical University named after Kozma Minin. E-mail: dianatelekova@gmail.com  
**Pavlova Valeria**, Student of the Nizhny Novgorod State Pedagogical University named after Kozma Minin. E-mail: valeriyapavlova\_04@inbox.ru

**Ponachugin Alexander**, Associate Professor, PhD in Economics, Nizhny Novgorod State Pedagogical University named after Kozma Minin. E-mail: sha3@bk.ru

**Abstract.** The article is devoted to the study of prejudice and ensuring fairness in relation to artificial intelligence. The research is aimed at familiarization with the process of artificial intelligence; consideration of various spheres of human social activity in which the use of artificial intelligence is relevant; identification of the positive or negative impact of artificial intelligence on the life of a modern person; formation of a fair attitude to artificial intelligence.

**Keywords:** artificial intelligence, technology, neural networks, modern man, security.

#### Библиографический список:

1. Кудряшова П. А. Правовое регулирование использования искусственного интеллекта [Электронный ресурс] // cyberleninka: [сайт]. – URL: <https://cyberleninka.ru/article/n/pravovoe-regulirovanie-ispolzovaniya-iskusstvennogo-intellekta> (дата обращения: 04.04.2024).
2. Левченко И. В., Павлова А. Е., Садыкова А. Р. Модуль «Введение в искусственный интеллект» в общеобразовательном курсе информатики [Электронный ресурс] // cyberleninka.ru: [сайт]. – URL: <https://cyberleninka.ru/article/n/modul-vvedenie-v-iskusstvenny-intellekt-v-obsheobrazovatelnom-kurse-informatiki> (дата обращения: 03.04.2024).
3. Соломонов Е. В. Понятие и признаки искусственного интеллекта [Электронный ресурс] // cyberleninka.ru: [сайт]. – URL: <https://cyberleninka.ru/article/n/ponyatie-i-priznaki-iskusstvennogo-intellekta-1> (дата обращения: 03.04.2024).
4. Хусанов У. А., Кудратиллаев М. Б., Сиддиков Б. Н., Довлетова С. Б. Искусственный интеллект в медицине [Электронный ресурс] // cyberleninka.ru: [сайт]. – URL: <https://cyberleninka.ru/article/n/iskusstvenny-intellekt-v-medicine-4> (дата обращения: 20.04.2024).
5. Фершт В. М. Современные подходы к использованию искусственного интеллекта в медицине / В. М. Фершт, А. П. Латкин, В. Н. Иванова // Территория новых возможностей. Вестник ВГУЭС. 2020. С. 127–128.
6. Харитонов Ю. С., Савина В. С. Технология искусственного интеллекта и право: вызовы современности // Вестник Пермского университета. Юридические науки. 2020. Вып. 3. С. 524–549.
7. Царев Р. Ю., Тынченко С. В., Гриценко С. Н. Адаптивное обучение с использованием ресурсов информационно-образовательной среды // Современные проблемы науки и образования. 2016. № 5. URL: <https://science-education.ru/article/view?id=25227> (дата обращения: 18.04.2024).
8. Шишкунова Д. В., Ищенко А. В. Логистика строительного производства: проблемы и пути решения // Инженерный вестник Дона. 2020. № 1. URL: [ivdon.ru/magazine/archive/Niy2020/6292](http://ivdon.ru/magazine/archive/Niy2020/6292)
9. Шобонов Н. А., Булаева М. Н., Зиновьева С. А. Искусственный интеллект в образовании [Электронный ресурс] // cyberleninka.ru: [сайт]. URL: <https://cyberleninka.ru/article/n/iskusstvenny-intellekt-v-obrazovanii-1> (дата обращения: 18.04.2024).
10. McCarthy J. What Is Artificial Intelligence? Computer Science Department // Stanford University. November 12, 2007. URL: <http://www-formal.stanford.edu/jmc/whatisai/whatisai.html>. (дата обращения: 04.04.2024).

#### Bibliography:

1. Kudryashova P. A. Legal regulation of the use of an electronic resource // Cyberleninka: [website]. – URL: <https://cyberleninka.ru/article/n/pravovoe-regulirovanie-ispolzovaniya-iskusstvennogo-intellekta> (date of application: 04.04.2024).
2. Levchenko, I. V., Pavlova A. E., Sadykova A. R. Module «Introduction to scientific intelligence» in the everyday course of computer science [Electronic resource of answers] // cyberleninka.ru: [website]. – URL: <https://cyberleninka.ru/article/n/modul-vvedenie-v-iskusstvenny-intellekt-v-obsheobrazovatelnom-kurse-informatiki> (date of application: 04/03/2024).
3. Solomonov E. V. Understanding and influence of state intelligence [Electronic resource] // cyberleninka.ru: [website]. – URL: <https://cyberleninka.ru/article/n/ponyatie-i-priznaki-iskusstvennogo-intellekta-1> (date of reference: 04/03/2024).
4. Khusanov, A., Kudratillaev M. B., Siddikov B. N., Dovletova S. B. State interest in Medicine [Electronic resource] // cyberleninka.ru: [website]. – URL: <https://cyberleninka.ru/article/n/iskusstvenny-intellekt-v-medicine-4> (date of application: 04/20/2024).
5. Fersht V. M. Modern approaches to the use of artificial intelligence in medicine / In M. Fersht, P. Latkin, V. N. Ivanova // The territory of new opportunities. Bulletin of the VGUES. 2020. pp. 127–128.
6. Kharitonova Yu. S., Savina V. S. Artificial intelligence technology and law. Challenges of our time // Bulletin of the Perm University. Legal sciences. 2020. Issue 3. pp. 524–549.
7. Sharov R. Sh., Tynchenko S. V., Grishenko S. N. Adaptive learning using a software and hardware environment // Modern problems of science and education. 2016. No. 5. URL: <https://science-education.ru/article/view?id=25227> (date of reference: 04/18/2024).
8. D. Shishkunova, V., Ishchenko A. V. Logistics of construction production: problems and solutions // Engineering Bulletin of the Don. 2020. № 1. URL: [ivdon.ru/magazine/archive/Niy2020/6292](http://ivdon.ru/magazine/archive/Niy2020/6292)
9. Kobonov N. A., Bulaeva M. N., Zinovieva S. A. State intelligence in education [Electronic resource] // cyberleninka.ru: [website]. URL: <https://cyberleninka.ru/article/n/iskusstvenny-intellekt-v-obrazovanii-1> (date of application: 04/18/2024).
10. McCarthy J. What is artificial intelligence? Faculty of Computer Science // Stanford University on November 12, 2007. URL: <http://www-formal.stanford.edu/jmc/whatisai/whatisai.html>. (date of application: 04.04.2024).

**Мельников Андрей**  
Заместитель директора центра инновационных программ, НИОКР и отраслевой стандартизации, к. т. н., НИИТТ (ООО «НИИ Транснефть»)  
E-mail: MelnikovAV@niitnn.transneft.ru

**Бачурин Александр**  
Ведущий научный сотрудник сектора инновационных программ, к. т. н., НИИТТ (ООО «НИИ Транснефть»)  
E-mail: BachurinAI@niitnn.transneft.ru

**Гниломёдов Евгений**  
Старший научный сотрудник сектора инновационных программ, к. э. н., НИИТТ (ООО «НИИ Транснефть»)  
E-mail: GnilomedovEV@niitnn.transneft.ru

**Распопов Андрей**  
Заместитель директора центра инновационных программ, НИОКР и отраслевой стандартизации, к. т. н., НИИТТ (ООО «НИИ Транснефть»)  
E-mail: RaspopovAA@niitnn.transneft.ru

**Дворецкая Кристина**  
Выпускник МГИМО  
МИД России

## ТЕХНОЛОГИИ БОЛЬШИХ ДАННЫХ В ЦИФРОВИЗАЦИИ ТЭК

*Аннотация. Развитие технологий больших данных за последние несколько лет стало одним из главных достижений в ИТ-индустрии. Благодаря стремительному развитию информационных технологий в последние годы облачные вычисления и технологии больших данных внедрены и в нефтегазовую отрасль. Добыча и транспортировка углеводородов характеризуется такими особенностями, как большой объем данных, высокая производительность получения и анализа данных, высокая надежность и постоянное расширение массивов хранения данных. Применение облачной платформы нового поколения поможет решить задачи сбора и функционирования баз данных. Облачные вычисления объединяют прикладную программную платформу больших данных, архитектура и применение промышленных больших данных в нефтегазовой отрасли, а также сформулированы выводы об эффективности применения больших данных для оптимизации традиционных технологических операций в нефтегазовой промышленности.*

### Ключевые слова:

промышленные большие данные, облачные вычисления и технологии, цифровизация, нефтегазовая промышленность.

**Облачные вычисления активно применяются для оптимизации традиционных технологических операций: в разведке, бурении, добыче и переработке нефти и газа**

Большие данные представляют собой структурированные и неструктурированные данные огромных объемов и типов представления, эффективная обработка которых требует специализированных методик и программных средств. С появлением новых технологических возможностей анализа огромных массивов данных появилось данное понятие, формализованное в серии научных публикаций конца 2000-х гг., первой из которых считается статья Клиффорда Линча [1].

Бесспорно, огромные массивы информации требуют новейших технологий, позволяющих собирать, анализировать, визуализировать и хранить сложные наборы данных. Во многих научных статьях отмечается, что большие данные сочетают в себе всевозможные виды ИТ-технологий: оперативный поиск, систематический сбор и обработку, статистический анализ, а также вычисления больших массивов данных. Благодаря передовым технологиям, большие данные позволяют определить ранее не исследуемые законы и неизвестные корреляции между разрозненной информацией.

Основные сценарии применения больших данных включают пакетную, потоковую и поисковую обработку. Пакетная обработка отличается накоплением больших объемов данных, которые обрабатываются по запросу от пользователей

и преимущественно применяется в интернете вещей<sup>1</sup>, облачных вычислениях и интернете. При пакетной обработке данные, обладающие высокой точностью и низкой плотностью, передаются в блочном режиме. Данные сначала накапливаются и анализируются, достигнув определенного объема. Кроме того, поисковая обработка включает в себя анализ потоковых данных и интерактивную обработку данных. Потоковые данные характеризуются бесперебойным анализом ранее не хранимых в памяти потоков данных, которые обрабатываются практически в режиме реального времени и результаты которых формируются мгновенно. Типичное применение потоковой обработки данных заключается в сборе данных – получение журналов, показаний датчиков и прочее в режиме реального времени.

Интерактивная обработка данных преимущественно применяется в системах обработки информации (OLTP, OLAP, Hive, Pig) и в интернете (поисковая система, электронная почта, средства коммуникации в реальном времени и т. д.). Крупнейшими известными системами управления базами дан-

<sup>1</sup> Системы, которые объединяют устройства в компьютерную сеть и позволяют им собирать, анализировать, обрабатывать и передавать данные другим объектам через программное обеспечение, приложения или технические устройства.

ных для интерактивного анализа являются MySQL, Oracle NoSQL Database, MongoDB, Microsoft SQL Server, HBase.

По своей структуре и логике возникновения большие данные относятся к категории данных, генерируемых действиями людей, в первую очередь, их социально-экономической активностью. Кроме указанных выше областей большие данные также находят широкое применение в транспорте, мобильной связи и интернете, медицине, маркетинге, образовании, HR, финтехе [2]. Общая доля больших данных, возникающих при осуществлении технологических операций (промышленные большие данные, далее – ПБД), существенно меньше генерируемых действиями людей. Источниками таких данных являются, как правило, датчики и сенсоры, расположенные непосредственно на оборудовании. Понятие «ПБД» неразрывно связано с понятиями «Киберфизические системы», представляющие комплексные системы из вычислительных и физических элементов, которые непрерывно получают данные из внешней среды и используют их для дальнейшей оптимизации процессов управления.

Исходя из принципа формирования ПБД на текущем уровне развития и внедрения в мире технологий Индустрии 4.0, возможно сформировать основные свойства оборудования, при эксплуатации которого ПБД могут быть получены и использованы в дальнейшем анализе:

- 1) дорогостоящее оборудование, продление ресурса безаварийной работы которого экономически целесообразно;
- 2) мелкосерийное и уникальное оборудование, сбор данных об эксплуатации которого позволяет проводить дальнейшую оптимизацию конструкции;
- 3) оборудование с длительным сроком эксплуатации, формирующее собственную инфраструктуру по его обслуживанию и поддержке в рабочем состоянии;
- 4) оборудование, аварии и инциденты на котором могут привести к значительным социальным и экономическим последствиям.

Далее рассмотрим архитектуру хранения больших промышленных данных в топливно-энергетическом комплексе. Компании нефтегазовой отрасли управляют сложными системами и для архивирования корпоративными офисными и бизнес-данными применяют распределенные облачные сервисы, ленточные накопители и дисковые массивы. Например, самый большой объем сейсмических данных накапливается в облачных сервисах, в то время как структурные данные хранятся в традиционном дисковом массиве. Благодаря распределенной системе обработки данных Spark, можно получить стандартизированные данные в режиме реального времени. Например, в нефтяной отрасли это будут исполнительные механизмы и датчики системы SCADA. Далее результаты обработки данных в кластер Hadoop для анализа соответствующих больших массивов данных и пакетной обработки. Для хранения маркетинговых данных нефтяных компаний используется распределенное хранение данных, а поведение пользователей объясняется корреляционным анализом из кластера Hadoop. Ценность интеллектуального анализа данных заключается в конечном принятии решений на всех трех этапах – upstream, midstream, downstream.

Топливо-энергетический комплекс вынужден решать вопрос, каким образом производить интеграцию и обмен данными. Для данных целей используются облачные вычисления, гарантирующие как массовое хранение и обмен данными, так и интеграцию различных типов аппаратных и программных ресурсов унифицированным и эффективным способом.

Облачные вычисления повышают эффективность работы в нефтяной промышленности. Стоит отметить активное применение анализа ПБД для оптимизации традиционных технологических операций, например, в разведке (при анализе сейсмических данных и построении цифровой модели месторождения), бурении (при оперативном определении параметров бурения на основе поступающей информации), добыче (с целью повышения нефтеотдачи пласта при обладании более полной информацией о текущих физических свойствах

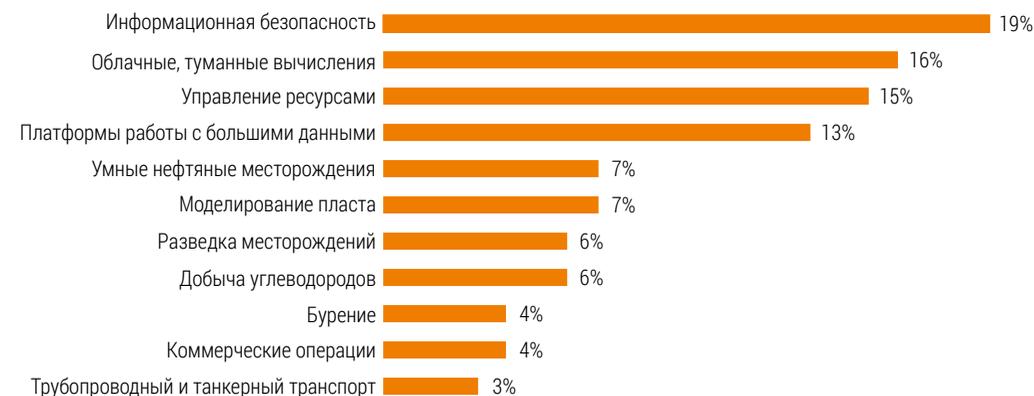


Рис. 1. Доли публикаций, посвященных применению ПБД в нефтяной и газовой промышленности по объектам исследований [11]

нефтяного коллектора) и переработке нефти и газа [3–10]. В области разведки и добычи объем данных позволяет прогнозировать потенциальные запасы нефти и газа, направлять гидравлический разрыв пласта на нефтяном месторождении и повышать уровень извлекаемости нефти и газа. Нефтяные компании используют облачные сервисы для расширения существующих возможностей и обработки массивов сейсмических данных, генерируемых сверхчувствительными сейсмическими датчиками, что позволяет сократить время для принятия управленческих решений.

Для полноты общего исследования был проведен анализ тематических источников информации. Если обратиться к объему публикаций, посвященных использованию ПБД для решения задач, возникающих при транспортировке жидких и газообразных углеводородов трубопроводным и морским транспортом, можно заметить, что он не превышает 3% от всего объема публикаций, посвященных применению ПБД в нефтяной и газовой промышленности.

Несмотря на довольно скромную долю примеров использования ПБД в трубопроводном транспорте углеводородов в сравнении с разведкой и добычей, спектр их применения довольно широк, постоянно расширяется и включает в себя следующие

аспекты трубопроводного транспорта углеводородов [12]:

- при использовании метода рассеяния магнитного потока для поиска утечек (MFL–Magnetic Flux Leakage), традиционно применяемого при диагностике резервуаров, для определения величины потери металла нефте- и газопроводов [13, 14];
- при расчете модифицированного коэффициента запаса прочности (SF), учитывающего фактические и накопленные за определенный период времени параметры эксплуатации трубопровода [15];
- для расчета продолжительности межсервисного интервала динамического оборудования (насосы, компрессоры) на основании потоковых данных от расположенных на нем датчиков (предиктивная аналитика), что в рамках широко применяемой за рубежом риск-идеологии управления активами позволяет существенно уменьшить операционные затраты на эксплуатацию [16, 17];
- при оптимизации проектных решений сложных технических систем и анализе модельных ПБД (например, газотранспортных систем) [18];
- при построении систем управления целостностью трубопроводов [19];

- при разработке систем управления промышленной безопасностью на объектах нефтегазового профиля на основе анализа исторических и текущих данных об авариях и инцидентах [20–22];
- при разработке новых, основанных на больших данных систем поддержки принятия решений [23, 24];
- при разработке систем распознавания возгораний и планов реагирования на них [25].

Благодаря огромному количеству наземных и подземных датчиков, ГЛОНАСС и GPS-координат, метеорологических служб, сейсморазведки и различного измерительного оборудования, компании топливно-энергетического комплекса получают в режиме реального времени данные для мониторинга оборудования, составления графика работ техобслуживания, а также для управления съемкой, планирования геологоразведочных работ, моделирования пластов, а также других процессов.

Нефтегазовые компании активно применяют методы больших данных для таких

целей, как автоматическое распознавание сейсмической активности, моделирование пластов, повышение извлекаемости нефти и газа, а также для точного определения аномальных условий и предотвращения травм сотрудников и повреждения оборудования.

Транспортировка углеводородов характеризуется такими особенностями как большой объем данных, высокая производительность чтения и записи данных, высокая надежность и постоянное расширение массивов хранения данных. Технология больших данных в секторе midstream в основном применяется для сбора и мониторинга данных в системе SCADA, представляющих собой структурированные, неструктурированные или полуструктурированные типы данных. ПБД, полученные от систем SCADA, буровых головок, различных датчиков состояния трубопровода, передаются практически в режиме реального времени и обрабатываются в потоковом режиме.

Особый интерес представляет использование технологий больших данных в мониторинге состояния магистральных трубопро-

Цифровой двойник НПЗ  
Источник: vuzopedia.ru



водов, в частности с помощью БПЛА, а также в мониторинге состояния внутренней поверхности трубопроводов с помощью специальных внутритрубных приборов. В частности, используется инновационная магнитная измерительная и бортовая аппаратура для работы в составе роботизированного комплекса, который предназначен для обследования магистральных трубопроводов внутри них. С измерительного снаряда данные поступают на пульт оператора, а далее производится их обработка в режиме реального времени. Таким образом, технология больших данных позволяет в полной мере использовать результаты информационных технологий и оцифровки, развивать службы техобслуживания и ремонта, повысить качество контроля, обеспечивая мониторинг и интеллектуальную диагностику. Отдельного рассмотрения заслуживает использование ПБД при решении задач промышленной безопасности, в первую очередь связанных с деятельностью работника на опасном производственном объекте [26, 27, 28].

Выделим следующие факторы, которыми должен обладать успешный проект по применению ПБД:

1. Проект нацелен на решение конкретной и осознаваемой проблемы с четко определенной ценностью ее решения.
2. Проект сочетает в себе современные методы анализа больших данных с традиционными методами инженерного анализа данных.
3. Проект реализуется мультидисциплинарной командой, состоящей из специалистов по работе с данными и технических специалистов.
4. Проект соответствует основным стратегическим документам, определяющим ИТ-архитектуру компании и работу с данными.
5. Проект способен продемонстрировать техническую реализуемость (proof of concept).
6. Проект использует ПБД по причине действительной необходимости их использования, а не их наличия и возможности работы с ними.

7. Проект представляет свои результаты в актуальной и удобной для использования форме: моделях, графиках, ключевых показателях эффективности, рекомендациях и т. д.

## Выводы

На сегодняшний день в ходе анализа и прогнозирования компании топливно-энергетического комплекса исследуют информационные закономерности и находят решения благодаря передовым информационным технологиям. Вышеприведенный анализ показывает, что общие принципы и механизмы генерации ПБД, методы их анализа и интерпретации являются зрелыми элементами управления данными. Исследования ПБД в нефтегазовых компаниях затрагивают по большей части обработку технологических данных, интеллектуальную систему поддержки принятия решений на базе огромного количества возможностей распознавания и обработки данных. Характер публикаций свидетельствует о значительном накопленном опыте использования ПБД для повышения ценности продуктов организаций различного профиля, в ряде случаев, переходящем на уровень рутинных операций, что позволяет рассчитывать на высокий отраслевой уровень технологий работы с ПБД при принятии компаниями решений об использовании промышленных баз данных в своей деятельности. Среди потенциальных выгод от их использования можно перечислить следующие:

- получение необходимой для анализа информации в режиме реального времени;
- рост эффективности и повышение уровня безопасности использования производственных мощностей;
- оптимизация процедур управленческого учета.

Таким образом, применение больших данных в трубопроводном транспорте и в целом в нефтегазовой отрасли является важной составляющей цифровизации отрасли, позволяя делать более эффективными и управляемыми как бизнес-процессы, так и технические процессы эксплуатации, ремонта и технического обслуживания.

## BIG DATA TECHNOLOGIES IN THE DIGITALIZATION OF THE FUEL AND ENERGY COMPLEX

**Melnikov Andrey**, Ph of Technical Sciences, Deputy Director of the Center for Innovative Programs, R&D and Industry Standardization, Transneft Research Institute, LLC, Moscow. E-mail: MelnikovAV@niitnn.transneft.ru

**Bachurin Alexander**, Candidate of Technical Sciences, Leading Researcher of the Innovation Programs Sector, Transneft Research Institute, LLC. E-mail: BachurinAI@niitnn.transneft.ru

**Gnilomedov Evgeny**, Candidate of Economic Sciences, Senior Researcher of the Innovation Programs Sector, Transneft Research Institute, LLC. E-mail: GnilomedovEV@niitnn.transneft.ru

**Raspopov Andrey**, Candidate of Technical Sciences, Deputy Director of the Center for Innovative Programs, R&D and Industry Standardization, Transneft Research Institute, LLC. E-mail: RaspopovAA@niitnn.transneft.ru

**Dvoretzkaya Kristina**, graduate of the Moscow State Institute of International Relations (University) Ministry of Foreign Affairs of the Russian Federation (MGI-MO Ministry of Foreign Affairs of Russia)

**Abstract.** The development of big data technologies has been one of the main achievements in the IT industry over the past few years. The extraction and transportation of hydrocarbons is characterized by such features as a large amount of data, high reliability and constant expansion of massive data storage. The use of a new generation cloud platform will help solve these problems. Cloud computing integrates a big data application software platform. The article analyzes the methods of big data analysis, architecture and implementation of industrial big data in the oil and gas industry, and also draw conclusions about the effectiveness of using big data to optimize traditional technological activities in the oil and gas industry.

**Keywords:** industrial big data, cloud computing technology, digitalization, oil and gas industry.

### Библиографический список

- Lynch, C., "Big data: How do your data grow?" *Nature* 455. 2008. P. 28–29. URL: <https://doi.org/10.1038/455028a>
- Aslam Abdullah M., Aseel A., Rithul Roy, Pranav Sunil, "Predictive big data analytics for drilling downhole problems: A review", *Energy Reports*, Volume 9, December 2023. P. 5863-5876.
- Yang Zhifeng, Han Fei, Feng Xuehui, Yuan Qi, Cao Zhen, Zhang Yidan, "Cloud Computing and Big Data for Oil and Gas Industry Application in China", *Journal of Computers*, Volume 14, Number 4, April 2019.
- H. Hassani and E.S. Silva, "Big data: a big opportunity for the petroleum and petrochemical industry," *OPEC Energy Review*, Volume 42, no. 1. 2018. P. 74-89. doi:10.1111/opee.12118
- Mehdi Mohammadpoor, Farshid Torabi, "Big Data analytics in oil and gas industry: An emerging trend", *Petroleum*, Volume 6, Issue 4, December 2020. P. 321-328. URL: <https://doi.org/10.1016/j.petlm.2018.11.001>
- Syaifulah, K. I., The Use of Big Data in The Oil and Gas Upstream Industry, *Indonesian Journal of Energy*, 2 (1), February 2019. P. 14-28. URL: <https://doi.org/10.33116/ije.v2i1.31>
- Vega-gorgojo, Guillermo & Fjellheim, Roar Arne & Roman, Dumitru & Akerkar, Rajendra & Waaler, Arild, "Big Data in the Oil & Gas Upstream Industry - A Case Study on the Norwegian Continental Shelf," *Oil Gas European Magazine*, 42(2), June 2016. P. 67-77. URL: <https://www.researchgate.net/publication/305416575>
- Jas Nitesh Desai, Sivakumar Pandian, Rakesh Kumar Vij, "Big data analytics in upstream oil and gas industries for sustainable exploration and development: A review", *Environmental Technology & Innovation*, Volume 21, February 2021, 101186. URL: <https://doi.org/10.1016/j.eti.2020.101186>
- Nimmagadda S. L., Reiners T., Rudra A., "An upstream business data science in a big data perspective," *Procedia Computer Science*, vol. 112. 2017. P. 1881-1890.
- Nguyen, Trung and Ray Gosine, "A Systematic Review of Big Data Analytics for Oil and Gas Industry 4.0," *IEEE Access*, March 2020, P. 6183-61201. DOI: 10.1109/access.2020.2979678
- Robert K. Perrons, Jesse W. Jensen, "Data as an asset: What the oil and gas sector can learn from other industries about "Big Data", *Energy Policy*, Volume 81, 2015. P. 117-121. URL: <https://doi.org/10.1016/j.enpol.2015.02.020>
- Layouni, M., Hamdi M. S., & Tahar S., "Detection and sizing of metalloss defects in oil and gas pipelines using pattern-adapted wavelets and machine learning," *Applied Soft Computing*, Volume 52, 2017. P. 247-261. URL: <https://doi.org/10.1016/j.asoc.2016.10.040>
- Mohamed, A.A., Hamdi, M. S., & Tahar, S., A Machine Learning Approach for Big Data in Oil and Gas Pipelines, August 2015, Conference: 3rd International Conference on Future Internet of Things and Cloud (FiCloud). DOI:10.1109/ficloud.2015.54
- Heweï Zhang, Shaohua Dong, Jiatong Ling, Laibin Zhang, Brenda Cheang, A modified method for the safety factor parameter: The use of big data to improve petroleum pipeline reliability assessment, *Reliability Engineering & System Safety*, Volume 198, 2020.
- Hussain, Muhammad & Zhang, Tieling, Potential of Big Data Analytics for Energy Pipeline Integrity Management, February, 2023.
- Tarrahi, Mohammadali, and Arash Shadravan. Intelligent HSE Big Data Analytics Platform Promotes Occupational Safety. Paper presented at the SPE Annual Technical Conference and Exhibition, Dubai, UAE, September, 2016.
- Lang Huang, Chao Wu, Bing Wang, Qiumei Ouyang, Big-data-driven safety decision-making: A conceptual framework and its influencing factors, *Safety Science*, Volume 109, 2018. P. 46-56. URL: <https://doi.org/10.1016/j.ssci.2018.05.012>
- Lin Xie, Solfrid Håbrekke, Yiliu Liu, Mary Ann Lundteigen, Operational data-driven prediction for failure rates of equipment in safety instrumented systems: A case study from the oil and gas industry, *Journal of Loss Prevention in the Process Industries*, Volume 60, 2019. P. 96-105. URL: <https://doi.org/10.1016/j.jlp.2019.04.004>
- Qiumei Ouyang, Chao Wu, Lang Huang, Methodologies, principles and prospects of applying big data in safety science research, *Safety Science*, Volume 101, 2018. P. 60-71. URL: <https://doi.org/10.1016/j.ssci.2017.08.012>
- Tan, Kim Hua, et al. Using Big Data to Manage Safety-Related Risk in the Upstream Oil & Gas Industry: A Research Agenda. *Energy Exploration & Exploitation*, Feb. 2016. SAGE. URL: <https://doi.org/10.1177/01445987166630165>
- Azzedin F., Ghaleb M. Towards an Architecture for Handling in Oil and Gas Industries: Service-Oriented Approach. *International Journal of Advanced Computer Science and Applications*, Volume 10 (2), 2019. P. 554-562. URL: [https://www.researchgate.net/publication/331475691\\_Towards\\_an\\_Architecture\\_for\\_Handling\\_Big\\_Data\\_in\\_Oil\\_and\\_Gas\\_Industries\\_Service-Oriented\\_Approach](https://www.researchgate.net/publication/331475691_Towards_an_Architecture_for_Handling_Big_Data_in_Oil_and_Gas_Industries_Service-Oriented_Approach)
- Chen M., Mao S., Zhang Y., Leung V., "Big Data. Related Technologies, Challenges and Future Prospects", Springer, 2014. P. 89.
- Cowels, D. Oil, gas, and data, O'Reilly Media, Inc., April, 2015. URL: <https://www.oreilly.com/ideas/oil-gasdata>
- AO «Транснефть – Диаскан» разработало магнитную измерительную систему для внутритрубного обследования технологических трубопроводов // URL: <https://ndtspace.ru/novosti/102>
- Баширов К. И., Костенко М. С., Дятлов А. М. Опыт использования больших данных в российских нефтяных компаниях // *Вестник магистратуры*. 2019. №8-2 (95). С. 21-23. URL: <https://cyberlenika.ru/article/n/opyt-ispolzovaniya-bolshih-dannyh-v-rossijskih-neftnyanyh-kompaniyah> (Дата обращения: 20.02.2024).
- Буян Р. А. Информационные технологии и бизнес-процессы. Роль информационных технологий в управлении структурой бизнес-процессов // *Российское предпринимательство*. 2009. № 5-1. С. 64-70.
- Курмель Р. С. Анализ больших данных в нефтегазовой промышленности // *Проблемы геологии и освоения недр: Труды XXIV Междун. симп. им. акад. М.А. Усова студ. и молодых учёных, посвящ. 75-летию победы в ВОВ*, Томск, 6–10 апреля 2020 г. Т. 2. – Томск: Национальный исследовательский Томский политехнический университет, 2020. С. 100-102.
- Низамова Г. З., Гайфуллина М. М. Сценарный подход к оценке эффективности инвестиционного проекта строительства магистрального нефтепровода. 2019. DOI: 10.17122/ntj-oil-2019-1-96-107
- Прохоров А.В., Носков И.В. Мониторинг магистральных нефтегазопроводов при помощи беспилотных летательных аппаратов // *Вестник Евразийской науки*. №6, 2022. Т.14. URL: <https://esj.today/issue-6-2022.html>
- Селезнев В.Е. Современные компьютерные тренажеры в трубопроводном транспорте: математические методы моделирования и практическое применение: монография / В.Е. Селезнев, В.В. Алешин, С.Н. Прялов. - Москва; Берлин: Директ-Медиа, 2014. – 199 с.
- Чекушина Т.В., Шафеева Г.Х. Инновационные технологии в строительстве и эксплуатации трубопроводного транспорта. №4 (109), апрель, 2023 г. DOI: 10.32743/UniTech.2023.109.4.15293
- Tarrahi, Mohammadali, and Arash Shadravan. Intelligent HSE Big Data Analytics Platform Promotes Occupational Safety. Paper presented at the SPE Annual Technical Conference and Exhibition, Dubai, UAE, September, 2016.
- Lang Huang, Chao Wu, Bing Wang, Qiumei Ouyang, Big-data-driven safety decision-making: A conceptual framework and its influencing factors, *Safety Science*, Volume 109, 2018. P. 46-56. URL: <https://doi.org/10.1016/j.ssci.2018.05.012>
- Lin Xie, Solfrid Håbrekke, Yiliu Liu, Mary Ann Lundteigen, Operational data-driven prediction for failure rates of equipment in safety instrumented systems: A case study from the oil and gas industry, *Journal of Loss Prevention in the Process Industries*, Volume 60, 2019. P. 96-105. URL: <https://doi.org/10.1016/j.jlp.2019.04.004>
- Qiumei Ouyang, Chao Wu, Lang Huang, Methodologies, principles and prospects of applying big data in safety science research, *Safety Science*, Volume 101, 2018. P. 60-71. URL: <https://doi.org/10.1016/j.ssci.2017.08.012>
- Tan, Kim Hua, et al. Using Big Data to Manage Safety-Related Risk in the Upstream Oil & Gas Industry: A Research Agenda. *Energy Exploration & Exploitation*, Feb. 2016. SAGE. URL: <https://doi.org/10.1177/01445987166630165>
- Azzedin F., Ghaleb M. Towards an Architecture for Handling in Oil and Gas Industries: Service-Oriented Approach. *International Journal of Advanced Computer Science and Applications*, Volume 10 (2), 2019. P. 554-562. URL: [https://www.researchgate.net/publication/331475691\\_Towards\\_an\\_Architecture\\_for\\_Handling\\_Big\\_Data\\_in\\_Oil\\_and\\_Gas\\_Industries\\_Service-Oriented\\_Approach](https://www.researchgate.net/publication/331475691_Towards_an_Architecture_for_Handling_Big_Data_in_Oil_and_Gas_Industries_Service-Oriented_Approach)
- Chen M., Mao S., Zhang Y., Leung V., "Big Data. Related Technologies, Challenges and Future Prospects", Springer, 2014. P. 89.
- Cowels, D. Oil, gas, and data, O'Reilly Media, Inc., April, 2015. URL: <https://www.oreilly.com/ideas/oil-gasdata>
- Transneft Diaskan JSC has developed a magnetic measuring system for in-line inspection of technological pipelines // URL: <https://ndtspace.ru/novosti/102>
- Bashirov K. I., Kostenko M. S., Dyatlov A.M. The experience of using big data in Russian oil companies // *Bulletin of the Magistracy*. 2019. No.8-2 (95). pp.21-23. URL: <https://cyberlenika.ru/article/n/opyt-ispolzovaniya-bolshih-dannyh-v-rossijskih-neftnyanyh-kompaniyah> (Date of application: 02/20/2024).
- Buyan R. A. Information technologies and business processes. The role of information technologies in managing the structure of business processes // *Russian Entrepreneurship*. 2009. No. 5-1. pp. 64-70.
- Kurmel R. S. Analysis of big data in the oil and gas industry // *Problems of geology and subsoil development: Proceedings of the XXIV International Symposium named after Academician M.A. Usov of students and young scientists dedicated to the 75th anniversary of Victory in the Great Patriotic War*, Tomsk, April 6-10, 2020 Vol. 2. – Tomsk: National Research Tomsk Polytechnic University, 2020. Pp. 100-102.
- Nizamova G. Z., Gaifullina M. M. A scenario approach to evaluating the effectiveness of an investment project for the construction of an oil trunk pipeline. 2019. DOI: 10.17122/ntj-oil-2019-1-96-107
- Prokhorov AV., Noskov IV. Monitoring of main oil and gas pipelines using unmanned aerial vehicles // *Bulletin of Eurasian Science*. No.6, 2022. Vol. 14. URL: <https://esj.today/issue-6-2022.html>
- Seleznev V.E. Modern computer simulators in pipeline transport: mathematical modeling methods and practical application: monograph / V.E. Seleznev, V.V. Alyoshin, S.N. Pryalov. - Moscow, Berlin: Direct Media, 2014. – 199 p.
- Chekushina TV., Shafeeva G.H. Innovative technologies in the construction and operation of pipeline transport. No.4 (109), April, 2023 DOI: 10.32743/UniTech.2023.109.4.15293



## ОБЕСПЕЧЕНИЕ СОХРАННОСТИ И ДОСТУПНОСТИ ЭЛЕКТРОННЫХ ФОНДОВ БИБЛИОТЕКИ

*Аннотация. Рассмотрены особенности обеспечения сохранности электронных фондов научной библиотеки в период перехода на электронные форматы представления документов. Особое внимание уделено сохранности документов в печатной форме, возможности получения значительного количества копий без ущерба для качества оригинального документа за счет перевода в электронный формат, расширение доступа пользователей к информации, предоставление виртуальных услуг удаленным пользователям и вопросам долговременной сохранности электронных массивов. Показаны особенности использования различных носителей для долговременного хранения электронной информации. Внимание уделено выбору методов сканирования, форматам представления данных и процессам миграции данных во времени при внедрении новых технологий хранения данных. Своевременная миграция данных является тем средством, которое позволяет обеспечить длительное хранение электронных документов.*

**Цветкова Валентина**  
Главный научный сотрудник  
Библиотеки по естественным  
наукам Российской академии  
наук (БЕН РАН), д. т. н.,  
профессор  
E-mail: vats08@mail.ru

**Кулев Владимир**  
Начальник отдела  
информатизации Библиотеки  
по естественным наукам  
Российской академии наук  
(БЕН РАН)  
E-mail: kulev@benran.ru

**Ключевые слова:**

оцифровка документов, долгосрочное хранение данных, носители долгосрочного хранения данных, технологии хранения, оптические диски для архивного хранения данных, электронная библиотека, миграция данных.

**Любая технология  
долгого хранения  
сталкивается  
с моральным  
и физическим  
старением  
носителей, требует  
миграции данных  
и использования  
новых технологий  
хранения**

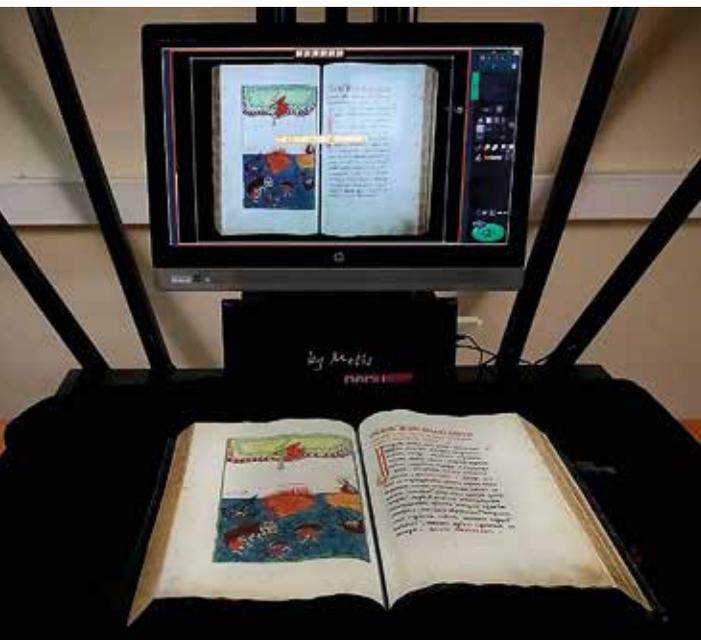
**Введение**

Стремительный рост информационных ресурсов в электронной форме, превращение научных библиотек в гибридные структуры, объединяющие традиционные бумажные документы и новые электронные в разных форматах, обострило необходимость решения вопросов сохранности традиционных бумажных документов и ресурсов в электронной форме. В повестке дня стали доминировать следующие задачи:

- сохранение фондов традиционной печатной книги за счет их перевода в электронную форму, что ведет к уменьшению их физического износа при использовании;
- получение неограниченного числа копий оригинала без ухудшения качества;
- легкость восстановления документа при утрате физического оригинала за счет тиражирования электронной копии;
- расширение возможностей одновременного доступа пользователей к одному и тому же источнику информации;
- предоставление виртуальных услуг доступа к информации в удаленном формате без необходимости посещать библиотеку, что является особенно актуальным для маломобильных лиц, жителей отдаленных территорий, системы образования;

– легкость восстановления документа при утрате физического оригинала за счет тиражирования электронной копии.

Беспорядочное распространение данных на бумажных и/или электронных носителях влечет за собой потерю как самих источников, так и данных, на них записанных. Одним из основных инструментов сохранения и распространения электронных версий документов служит цифровая электронная библиотека (ЭБ). Когда пользователь работает с электронными документами через удаленный доступ, он, прежде всего, обращается к электронному архиву библиотеки. Поэтому важно определить цели и функции создания электронной библиотеки, условия сохранности и доступности информации, включенной в ее состав. Обеспечение сохранности фондов на традиционных носителях считается одной из ключевых задач библиотек. Цифровой формат предоставления документов дает намного больше возможностей для работы с информацией, чем традиционные печатные издания и публикации, обеспечивая более широкий и оперативный доступ к информации. Работа с электронными документами проводится на совершенно новом уровне, прежде всего, благодаря гибким возможностям поиска документов. Более того, создание цифровых копий печатных публикаций позволяет улуч-



Оцифровка книги

Источник: Мария Говтвань (РГБ) / rsl.ru

шить качество обслуживания, предоставляя электронные издания по индивидуальному и межбиблиотечному абонементам и отправляя их читателям через каналы связи. Для этого необходимо обеспечить сохранность электронных фондов и доступ к ним.

**Цели, методы и инструменты обеспечения сохранности и доступности цифровых ресурсов**

Цели создания электронной библиотеки, по сути, не отличаются от целей создания библиотеки в ее традиционном, общеизвестном понимании. Главным здесь является обеспечение сохранности библиотечных фондов и доступности информации.

Достижение поставленных целей требует выполнения ряда задач, а именно:

- создание электронных копий (оцифровка) печатных изданий;
- привлечение электронных ресурсов из внешних источников;

- каталогизация электронных изданий;
- организация долговременного хранения информации;
- обеспечение и организация доступа к информации;
- обеспечение сохранности и безопасность электронной информации.

Задача по созданию электронных копий (оцифровка) печатных изданий включает:

1. Обеспечение сохранности печатных изданий. При создании цифровых копий книг, следует обратить особое внимание на обеспечение сохранности оригиналов. Для этого устанавливаются четкие стандарты и инструкции для процесса сканирования, включая выбор соответствующих типу издания сканеров. Если речь идет о редких и особенно ценных изданиях, требуется экспертная оценка специалистов по сохранению архивов этих изданий.

2. Определение способов оцифровки печатных изданий. Методы оцифровки зависят от состояния печатного издания. Метод и режим сканирования определяются внутренними нормативными документами библиотеки [1]. В зависимости от способа оцифровки применяется соответствующее оборудование [2]: планшетные сканеры (цветные и черно-белые); планетарные сканеры; цифровые фото- и видеокамеры.

3. Трудозатраты на оцифровку документов.

4. При планировании оцифровки необходимо определить, какие документы необходимо перевести в цифровую форму для ввода в научный оборот с сохранением визуальной формы и информационного содержания, а какие – с сохранением только информационного содержания. Иными словами, по каждому документу (или группе документов) должно быть принято решение, что оцифровывается – только информационная составляющая или в максимальной степени все аспекты оригинала, включая рисунки, красочность оформления, заметки на полях. Если эти элементы имеют научное, социальное и культурное значение, делают документ уникальным, единственным в своем роде, то необходима оцифровка, учитывающая все эти особенности. Оттого, какая

форма документа должна быть сохранена, принципиально зависит технология и, соответственно, стоимость оцифровки документа. Если поставлена задача предоставления пользователям только информационной составляющей документа, можно ограничиться черно-белым сканированием с небольшим разрешением, стандартными средствами обрезки и очистки электронных копий. Этот подход наиболее рационален при переводе в электронную форму многих научных журналов (в которых цвет иллюстраций не является принципиальным для передачи смысла статьи), технической документации, справочной литературы, не имеющей признаков уникальности. Если же необходимо сохранить цвет иллюстраций и (или) фактуру документа, необходимо проводить сканирование в цвете (или градациях серого), что существенно увеличивает размеры электронной копии документа по сравнению с черно-белым сканированием. При необходимости отразить в электронной копии мелкие детали документа нужно увеличивать разрешение сканирования; сохранение пометок на полях или в тексте требует применения специальных программных средств и технологий для очистки и обрезки документа. Все эти факторы приводят к увеличению времени обработки документа и, соответственно, к увеличению стоимости оцифровки.

К этим же последствиям приводят требования по распознаванию отсканированного текста. Поэтому при планировании перевода научных документов в цифровую форму необходимо оценивать особенности распознавания текстов – насколько оно полезно для данной группы документов, и оправданы ли затраты на редактирование распознанных текстов, без которого распознавание теряет смысл.

Особенно критична и трудоемка оцифровка старых изданий, представляющих не только информационную, но и «предметную» ценность, поскольку их нельзя расширять (разброшюровывать). В процессе сканирования возникают технические проблемы со сканерами, вызванные значительным количеством пыли, пропитавшей старые издания, ломкостью бумаги, обусловленной ее

естественным старением. При их сканировании велик риск повреждений. Кроме того, при переводе редких документов в цифровую форму невозможно воспроизвести их физические особенности: качество бумаги, используемые красители и т. п. Чтобы обеспечить возможность полноценной работы с электронными копиями как «заменителями» оригиналов, все особенности каждого документа необходимо отразить в его характеристиках, что приводит к дополнительным трудозатратам.

Дополнительной работой при формировании электронных библиотек является создание «поисковых образов» документов – метаданных, обеспечивающих их многоаспектный поиск.

В связи с разным качеством массивов сканируемых документов сильно меняются и оценки трудоемкости процессов сканирования, включая формирование метаданных. Например, работа по формированию электронной библиотеки (ЭБ) «Научное наследие

Архивно-библиотечный фонд ВДНХ

Источник: mos.ru



России», включающей, в основном, редкие издания, показала, что на ввод в ЭБ одного издания объемом около 250 страниц требуется около 2-х человеко-дней<sup>1</sup> [3]. Выполняемые при этом технологические процессы включают сканирование издания, обработку изображений и оглавления (по которому осуществляется поиск и навигация в ЭБ), формирование метаданных (модифицированного библиографического описания). Необходимо отметить, что при работе с редкими изданиями (книжными памятниками) возрастает не только сложность технологии оцифровки, но и требуется более дорогостоящее оборудование по сравнению с переводом в электронную форму современных документов. Более детально проблемы оцифровки отражены в работе Г.А. Евстигнеевой «Качество оцифровки: проблемы и решения» [4].

#### **Выбор формата представления данных.**

При выборе формата для представления данных необходимо учитывать различие типов создаваемых цифровых копий, применяемых методов сканирования, запросов читателей, а также особенности представления информации. Наиболее часто используемые форматы: TIFF, PDF, JPEG, HTML, XML [5].

Форматы PDF и TIFF поддерживают программы для OCR (Optical Character Recognition – программы распознавания символов). Благодаря этому, текст можно сохранять внутри файла в виде слоев, быстро выполнять поиск и находить нужную информацию без необходимости перепечатывать документ. PDF (Portable Document Format) был разработан компанией Adobe Systems в 1993 г. и быстро стал стандартом для создания и обмена электронными документами. PDF обладает рядом преимуществ, например, таких, как сохранение оригинальной графики, текста и других элементов в неразрывной форме. Есть возможность

шифровать и защитить данные паролем. TIFF (Tagged Image File Format) по-прежнему остается одним из старейших форматов, который широко используется для сканированных изображений, сохранения и передачи данных. Одно из главных преимуществ – это высокое качество и сохранение каждого пикселя в высоком разрешении.

JPEG (Joint Photographic Experts Group) предназначен для хранения изображений, в нем применяется сжатие с возможной потерей качества. Это позволяет создавать файлы с меньшим размером, чем в других форматах, но некоторые детали изображения при этом теряются. Он подходит для хранения фотографий со множеством цветов, на визуальное отображение сжатие сильно не влияет.

Форматы HTML и XML относятся к языкам разметки и больше предназначены для чтения электронных книг онлайн. HTML (Hyper Text Markup Language) – это стандартизированный язык гипертекстовой разметки документов для просмотра веб-страниц в браузере. Веб-браузеры получают HTML-документ от сервера по протоколам HTTP/HTTPS или открывают с локального диска, интерпретируя код в интерфейс, который будет отображаться на экране монитора. XML (Extensible Markup Language) – это язык разметки, который предоставляет правила для определения данных. В отличие от других языков программирования, XML не может выполнять вычислительные операции сам по себе. Вместо этого для управления структурированными данными можно использовать любой язык программирования.

**Комплектование электронных ресурсов из внешних источников.** Электронные ресурсы отбираются в соответствии с профилем комплектования библиотеки. Если происходит доукомплектование электронными документами, созданными сторонними организациями (библиотеками, издательствами, частными лицами), необходимо провести проверку соответствия копии оригиналу и проверить регистрацию данного издания в государственных регистрирующих органах (Информрегистр, Роспатент).



Архив с микропленкой  
Источник: Kiwar / depositphotos.com

#### **Каталогизация электронных изданий.**

Для обеспечения оптимальных возможностей поиска информации в ЭБ, ее ресурсы должны быть каталогизированы. Все типы изданий проходят каталогизацию путем создания библиографической записи в электронном каталоге библиотеки согласно принятым стандартам и форматам. При каталогизации цифровых копий печатных изданий проверяется соответствие с библиографической записью на печатное издание.

**Организация долговременного хранения информации.** Рассуждая о долговечности хранения информации, в первую очередь, рассматриваются вопросы выбора носителей. Не будем акцентировать внимание на бумаге, которая хранится веками, проста для восприятия, но требует больших площадей для хранения, очень критична к влажности и теплу, может быть утрачена при пожаре. Несколько слов скажем о микроносителях (микрофиши, микрофильмы). Микрофильмирование занимает важную нишу в вопросах сохранности информации на небумажных носителях, его нельзя сбрасывать со счетов.

В США микроносители выбраны для передачи информации будущим поколениям. Микропленки помещаются в герметические капсулы из нержавеющей стали (причем их предварительно кондиционируют при очень низкой влажности) и хранят под землей в шахте при температуре 10 °С. По мнению специалистов, такие условия обеспечивают сохранность информации в течение 1000 и более лет [6, 7, 8].

По пути использования микроформ при формировании страхового фонда изданий пошла одна из крупных библиотек России – Центральная универсальная научная библиотека им. Н.А. Некрасова, где страховой фонд на микроносителях превышает 100 тыс. единиц [9].

При решении задачи выбора и организации хранения электронных документов следует учитывать вопросы сохранности и доступности этих данных. Например, можно создать два типа хранилищ: основное и дополнительное. Основное хранилище включает полный набор электронных документов библиотеки. Дополнительное хранилище ис-

<sup>1</sup> Каленов Н.Е., Погорелко К.П., Серебряков В.А., Сотников А.Н. Электронная библиотека «Научное наследие России». Состояние и перспективы развития. Труды конференции Абрау-2016. Новороссийск. 18–24 сентября 2016 г. М.: ИПМ им М.В.Келдыша. 2016. С. 148–151. DOI: 10.20948/abrau-2016-27; – URL://Keldysh>abrau/2016/27.pdf (дата обращения 22.03.2024).

пользуется для работы пользователей и содержит выделенную часть ресурсов основного хранилища. Система хранения должна удовлетворять требованиям надежности, быстро и удобного доступа, компактности [5, 6].

**Выбор типа основных и вспомогательных электронных носителей.** В начале эры электронно-вычислительных машин важную информацию, переведенную в цифровую форму, распространяли и хранили на магнитных лентах. Магнитная лента была представлена как носитель информации в 1952 г. В сфере информационного обслуживания это относилось, в первую очередь, к таким ресурсам, как базы данных Science Citation Index, Current Contents (США, Институт научной информации), в настоящее время база данных называется Web of Science Core Collection, принадлежит агентству Clarivate, реферативная информация Всероссийского института научной и технической информации (ВИНИТИ) и т. п. Магнитные ленты представляют собой достаточно надежные носители, однако для обеспечения сохранности данных их необходимо периодически

перематывать, они исключают прямой доступ к данным и требуют достаточно больших помещений для хранения. За последние годы магнитная лента, как средство хранения информации, несколько сдала позиции и продолжает медленно их сдавать, однако объем хранимой на магнитных лентах информации остается значительным. Магнитная лента продолжает оставаться идеальной системой для хранения резервной и архивной информации, а также эффективным решением для развития таких направлений, как большие объемы данных и облачные технологии. Специалисты Национального вычислительного центра энергетических исследований Департамента энергетики США (NERSC) утверждают [19, 20], что хранение данных на магнитной ленте экономически выгодно и надежно.

В середине 1980-х гг. информационные службы перешли на использование гибких (5-ти, а затем 3-х дюймовых) дисков, однако им на смену достаточно быстро пришли оптические диски CD, а затем и DVD, которые в начале 90-х гг. XX в. стали использоваться

CD-проигрыватель

Источник: vladacanon / depositphotos.com



в архивах и библиотеках в качестве носителей для хранения цифровых копий архивных документов. Этому способствовали следующие факторы:

- низкие цены на носители информации (собственно диски);
- низкие цены на приводы для записи дисков;
- компактность хранения информации, обусловленная достаточно большой емкостью CD и DVD-дисков (соответственно 700 Мб, 4,7 Гб и более).

Однако «эйфория» от перспектив использования оптических дисков для долговременного хранения информации достаточно быстро прошла. Исследования, проводимые в последнее десятилетие, показали значительное количество рисков, связанных с их использованием в качестве единственного носителя для долгосрочного хранения данных [15]. Среди них выделим:

- технические проблемы, связанные с качеством самих дисков (покрытие, качество изготовления), а также с условиями записи на диск;
- проблемы, связанные с воздействием окружающей среды: температурный режим, влажность, воздействие солнечных лучей, воздействие загрязняющих агрессивных сред;
- проблемы, связанные с непосредственным обращением с диском: хранить в вертикальном положении желательно в специальных боксах, нельзя трогать диск за записываемую поверхность, наносить любые метки на поверхность, наклеивать этикетки, подвергать механическому воздействию и пр.

Выбор оптических носителей в качестве среды для хранения цифровой информации часто происходит по экономическим соображениям. Однако, при расчете экономической эффективности использования CD- и DVD-дисков следует учитывать: темпы роста объемов цифровой информации; меняющиеся технологии записи на оптические диски; стоимость приобретения необходимого количества дисков, предназначенных для архивного хранения (включая диски для

копий); стоимость специального оборудования и программного обеспечения для их тестирования. Для хранения дисков необходимы специальные боксы, стоимость которых также необходимо учитывать.

Проведенные ориентировочные оценки показывают, что затраты на хранение дисков в значительной степени определяются размерами коллекций: для больших коллекций они сильно возрастают. Это обусловлено, во-первых, ростом непосредственных финансовых затрат, во-вторых, необходимостью переноса данных на новые носители цифровой информации (миграции данных), что связано с устареванием технологий, техники и программного обеспечения. Быстрая смена поколений компьютеров и программных средств зачастую не позволяет воспроизводить информацию с ранее записанных носителей на новых устройствах.

При оценке целесообразности использования CD и DVD для хранения цифровых копий документов, в общую стоимость подготовки и хранения информации необходимо включать:

- стоимость приобретения необходимого количества дисков архивного хранения, включая диски для резервных копий, стоимость оборудования для записи и выполнения контроля качества записи дисков, стоимость боксов (коробочек) для хранения дисков;
- стоимость изготовления этикеток для идентификации дисков (писать номер или шифр хранения фломастером категорически не рекомендуется, т. к. со временем чернила, проникая через защитный слой, разрушают рабочий) [13];
- стоимость оборудования и программного обеспечения, используемого для тестирования дисков;
- стоимость формирования метаданных для обеспечения поисковых функций в электронных библиотеках или архивах;
- стоимость миграции данных при кардинальной смене поколений техники и программного обеспечения.

Срок сохранности информации на обычных дисках CD и специального архивного

хранения (рабочий слой из органического вещества фталоцианина, отражательный слой из золота или сплава золота и серебра), например, Kodak Preservation, по заявлению компании производителя этих дисков KMP Media<sup>2</sup> [13], имеют гарантируемый срок хранения информации 300 лет, а специальные диски архивного хранения M-Disk, разработанные американской компанией Millenniata, в которых рабочий слой изготавливается из неорганического вещества, имеют гарантируемый срок хранения данных 1000 лет<sup>3</sup> [14], но это пока декларация фирмы производителя, которая никем не проверена, да и непонятно, что делать с техническими средствами и программным обеспечением на таких и даже более скромных сроках хранения.

Расчетным путем установлено, что «болванки» CD-R с записью могут сохраняться [11]: 75 лет – цианиновый краситель; 100 лет – фталоцианиновый краситель – «золотые» диски; 200 лет – доработанный фталоцианиновый краситель – «платиновые» диски. Для CD-RW не выработано установленных сроков хранения. Более того, эти диски менее стабильны, чем диски для однократной записи. Для CD-RW не имеет смысла вопрос «Будет ли теряться информация со временем на перезаписываемом слое? Ответ однозначен: будет» [15].

Остается только поверить в то, что если производители декларируют такие сроки сохранности, они научно обоснованы.

Возрастание объемов сканирования вследствие развития технологий и техники, а также рост количества электронных документов в системах электронного документооборота ведет к появлению больших массивов электронных документов, хранение которых на дисках CD и DVD становится

проблематичным, а точнее малоэффективным. Однако мы можем сослаться на опыт хранения большого объема информации, накопленный Всероссийским институтом научной и технической информации РАН (ВИНИТИ РАН), технология хранения информационных массивов (баз данных ВИНИТИ) базируется на использовании магнитооптических перезаписываемых дисков [16].

Выбор типа носителей для цифровых документов определяется возможностями вычислительных устройств (средств сохранения информации) с точки зрения эффективности доступа и надежности физического носителя. Наиболее рациональным представляется выбор комплекса технологичных средств хранения. То есть, в качестве главных носителей информации в цифровой библиотеке использовать файловые серверы с объемными жесткими дисками, а резервные копии хранить на оптических носителях (CD, DVD, дисках blue-ray).

#### **Диски CD и DVD однократной записи как архивные носители информации.**

В работе [17] авторы анализировали применение оптических дисков для хранения электронных документов и указывали на необходимость использования специальных дисков, предназначенных для длительного хранения данных. Учитывая актуальность данного направления, еще раз остановимся на вопросах, связанных с дисками архивного хранения и их особенностями.

В начале 90-х гг. XX в. диски CD и DVD однократной записи стали использоваться в архивах и библиотеках в качестве носителей для хранения цифровых копий архивных документов. Этому способствовали следующие факторы:

- низкие цены на диски CD и DVD;
- низкие цены на приводы для записи дисков CD и DVD;
- достаточно большая для того времени емкость дисков CD и DVD (соответственно 700 Мб и 4,7 Гб).

Исследования, проведенные комитетом по науке ЮНЕСКО [18] по использованию оптических дисков в качестве средства хранения цифровой информации, показывают

сложность процесса записи и большое количество рисков, связанных с использованием CD и DVD как носителей долгосрочного хранения данных. Необходимо также учитывать миграцию данных, связанную со старением форматов файлов. Ни о каких катастрофических потерях на тот период не сообщалось, что укрепляло веру в надежность и долговечность оптических носителей.

Для понимания возможности хранения данных на CD и DVD однократной записи (оптические диски) остановимся на принципах их работы.

Диск состоит из четырех слоев: прозрачная подложка, рабочий слой из органического материала, отражательный слой, защитный слой.

При записи информации луч лазера большой мощности, проходя через прозрачную подложку, меняет физико-химические свойства рабочего слоя (прожиг) – появляется «темный» участок, который называется pit (пит). При считывании информации луч лазера малой мощности фиксирует pit и рассеивается, или попадает на светлый участок, который называется land (ленд), отражается отражательным слоем и попадает на светоприемник. Таким образом, создается последовательность из нулей и единиц (bit). Качество диска зависит от двух факторов:

- качества соединения слоев диска в единую систему;
- качества материалов, из которых сделаны слои диска.

В дисках CD и DVD в качестве материала рабочего слоя используются преимущественно 3 основных типа органического материала: цианин, фталоцианин и AZO. В зависимости от материала отражающего слоя, тип материала рабочего слоя придает рабочей стороне диска определенную окраску. Рекомендации по выбору дисков CD и DVD для обеспечения длительного хранения информации представлены в работах [19, 20]. Необходимая технологическая операция при использовании дисков для обеспечения длительного хранения данных – их периодическое тестирование с целью определения состояния качества.



Производство CD-дисков  
Источник: gb2017.bertelsmann.de

#### **Временные рамки хранения информации.**

Время сохранения информации зависит от характеристик физических носителей и определяется соответствующими нормативными документами. По мере внедрения новых носителей необходимо осуществлять преобразование и перезапись информации, то есть проводить миграцию данных. Требуется разработка специального программного обеспечения, которое позволит автоматически проверять данные в цифровой библиотеке.

Срок хранения данных на оптических носителях превышает срок хранения данных на магнитных или твердотельных носителях. Например, американская компания Delkin Devices разработала первый в мире лазерный диск формата Blu-ray<sup>4</sup> [21], гарантирующий сохранность записанной на него инфор-

<sup>2</sup> New Gold Kodak CD and DVD Promise a 80 to 300 Years of Lifetime. – URL: <https://www.cdrinfo.com/d7/content/new-gold-kodak-cd-and-dvd-promise-80-300-years-lifetime> (дата обращения 20.03.2024).

<sup>3</sup> Шкляева Т. Хранить вечно. Технологии M-DISK. – URL: <https://nasledie.digital/articles/hranit-vechno-tehnologiya-m-disc/> (дата обращения 19.03.2024).

<sup>4</sup> Полсулин С. Диски Blu-ray проработают 200 лет. – URL: [https://www.cnews.ru/news/top/diski\\_bluray\\_prorobotayut\\_200 лет](https://www.cnews.ru/news/top/diski_bluray_prorobotayut_200 лет) (дата обращения 22.03.2024).



мации в течение 200 лет. Диск Delkin Archival Gold вмещает 25 ГБ данных и поддерживает скорость записи 4x. А еще одна американская компания Millenniata производит оптические носители M-DISC с гарантированным сроком хранения 1000 лет [14], это научно обоснованная декларация, но, естественно, не проверенная временем. По заявлению компании Millenniata, диски M-DISC могут воспроизводиться на стандартных DVD-приводах. Проведенные фирмой Millenniata исследования показали: данные, записанные на M-DISC, сохраняются даже при погружении в емкость с жидким азотом (температура при этом составляет  $-180\text{ }^{\circ}\text{C}$ ) с последующим помещением в кипящую воду [14].

В последние годы наметилась тенденция создания дисков, в которых органический слой заменяется на неорганический. Так, фирма Syylex разработала стеклянный диск, который получил название Glass-MasterDisc. Лаборатория Laboratoire national de metrologie et d'essais в результате различных испытаний, в том числе, ускоренного старения, установила, что стеклянный мастер-диск способен сохранять данные в течение 1000 лет, как на Земле, так и в космосе [21, 23]. Компания Northern Star s. r. o. (Чехия) разработала диски DataTresorDisc (DTD) [22]. В дисках DTD в качестве записывающего слоя используется металлокерамический композит, который невосприимчив к воздействию окружающей среды. DTD имеют стандартную для DVD емкость 4,7 Гб и являются дисками однократной записи. Диски DTD читаются на любых современных DVD-приводах и могут быть записаны на большинстве пишущих приводов. Компания определяет срок хранения информации на дисках этого типа в 160 лет.

В российской практике хранение регламентируется организацией [24] при соблюдении условий хранения в соответствии с ГОСТ Р 7.0.2.-2006 СИБИБД. «Консервация документов на компакт-дисках. Общие требования» [25]. Отметим отдельные требования ГОСТ Р 7.0.2.-2006: п. 4.3 – документы хранят не менее, чем в двух экземплярах; пп. 4.3, 4.4 – при изменении программного

и аппаратного обеспечений выполняется перезапись документов на компакт-диски с учетом требований нового обеспечения: п. 5 – определяет режимы хранения: температура –  $10\text{--}20\text{ }^{\circ}\text{C}$ , влажность –  $20\text{--}65\%$ ; п. 8 – регламентирует правила контроля состояния компакт-дисков, контроль проводится не реже одного раза в год.

**Вопросы размещения и связности хранилищ информационных массивов.**

Файловые серверы основного хранилища электронного фонда библиотеки, а также хранилище резервных оптических носителей должны размещаться в специально оборудованных помещениях с определенным климатическим режимом и не иметь доступа из локальной сети библиотеки. Возможно хранение архивной копии в удаленном, специально созданных для этой цели государственных хранилищах. Файловые серверы вспомогательного хранилища располагаются в обособленном помещении.

**Сохранность электронных носителей и массивов данных.**

Для обеспечения сохранности данных в цифровых фондах необходимо предусмотреть защитные программы для массивов данных от несанкционированного доступа и копирования, визуальный контроль цифровых носителей, проверку информации при выдаче и возвращении оптических носителей, миграционные процессы.

Особое внимание следует уделить сохранению преемственности использования носителей информации при появлении новых типов технических и программных средств, своевременно обновляя парк вычислительных устройств и проводя конвертацию данных.

Длительное хранение предполагает решение следующих вопросов:

- выбор или разработка технологии для обеспечения длительного хранения документов;
- обеспечение аутентичности и целостности электронного документа в течение всего срока хранения;
- интерпретируемость (возможность воспроизвести формат электронного документа в течение всего срока хранения);
- надежность носителя/системы хранения.

В течение всего срока хранения электронный документ должен отображаться в том виде, в котором он был создан, то есть не должна быть изменена или нарушена его целостность (аутентичен тому документу, который был заложен на хранение). При этом он не должен терять информацию о своем внешнем виде, в том числе, топологию расположения информации на материальном носителе. А главное, документ должен сохранять читаемость информации.

Вопрос надежного долговременного хранения электронных документов наиболее сложный в работе с электронными архивами. Удовлетворительного решения пока нет ни за рубежом [26], ни в России [27].

Особенно важен фактор долговременного хранения информации для библиотек, архивов, музеев, которые называют все чаще институтами социальной памяти [28]. Эти структуры располагают большими объемами уникальных изданий, рукописных книг, личных документов. Для них крайне важно сохранить оригинал (часто это единственный экземпляр), дав возможность использовать тексты в оперативном режиме всем заинтересованным пользователям. Хотя оцифровка позволяет решить вопрос о доступности и использовании фондов электронных библиотек, вопрос обеспечения их долговременного хранения является весьма сложным, требующим экономических обоснований и серьезных технологических решений.

Библиотеки, музеи и архивы являются важной частью потенциальной инфраструктуры системы сохранения цифрового наследия страны. Необходимо определить меры ответственности институтов памяти за обеспечение сохранности создаваемого и хранимого ими электронного контента, и доступа к нему, а также наделение их четко разграниченными полномочиями по сбору и сохранению электронной информации, создаваемой другими структурами и представляющей общественный интерес.

Есть и еще одна сторона вопроса, связанная с аутентичностью документов. Это особенно затрагивает архивные документы, которые часто востребованы в судебной

практике. В работе [5] отмечено: «Без четких представлений о происхождении электронного документа и твердых гарантий их целостности суды отказываются признавать за ними документальную силу и принимать в качестве письменного свидетельства».

Перед институтами социальной памяти стоит двуединая задача: сохранить документальные фонды и накопленные электронные информационные ресурсы, обеспечивая при этом постоянный оперативный доступ к ним.

**Обеспечение и организация доступа к информации.** Одним из важных вопросов, возникающих при создании цифровой библиотеки, является вопрос обеспечения доступности информации для пользователей. Эта задача включает в себя ряд аспектов, таких как:

**А. Выбор формы предоставления доступа к информации.** Доступ к основному хранилищу ЭБ для пользователей закрыт. Только сотрудники специальных подразделений имеют право административного доступа. Обслуживание пользователей осуществляется только из вспомогательного хранилища ЭБ.

Предоставление ресурсов электронной библиотеки для посетителей осуществляется как из локальной сети, так и посредством удаленного доступа, как авторизованным через индивидуального абонента, так и неавторизованным пользователям. Для удаленных пользователей доступ к информации предоставляется через глобальную сеть Интернет.

**Б. Разработка поискового механизма.** Так как каталогизация электронных документов проводится через электронный каталог библиотеки, основные функции поиска обеспечиваются возможностями автоматизированной библиотечно-информационной системы, используемой в библиотеке. Для обеспечения полнотекстового поиска информации требуется разработка соответствующего программного обеспечения, позволяющего производить поиск по всем информационным массивам в зависимости от вида и формата хранящихся в них данных. Механизм должен обеспечивать различные варианты поиска, сохраняя при этом единый способ и форму ввода запроса.

**В. Разграничение прав пользователя по работе с данными.** Возможность работы пользователя с информацией в электронной библиотеке определяется нормативно-правовыми документами руководящих органов РФ (защита авторских прав) и внутренней организационно-распорядительной документацией (обеспечение сохранности массивов данных и электронных носителей).

Реализация права на получение информационных продуктов и услуг в библиотеке тесно связана с проблемой качества обслуживания, поэтому правовой аспект этой проблемы включает вопросы, касающиеся обеспечения качества библиотечного обслуживания.

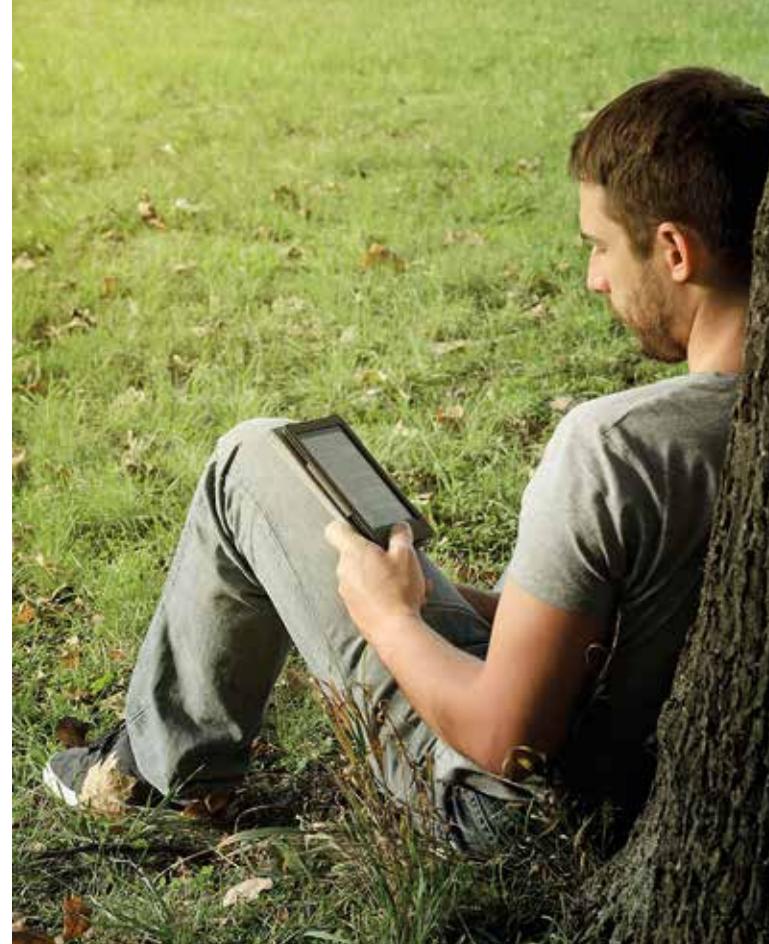
Наряду с проблемами доступности информации для различных категорий пользователей, современная библиотека сталкивается и с другими проблемными вопросами, связанными с реализацией прав пользователей в области защиты от информации, наносящей вред их развитию, с правом на сохранение конфиденциальных данных пользователей от доступа к ним третьих лиц и злоумышленников [29].

**Обеспечение сохранности и безопасности данных в электронных библиотеках.** Обеспечение информационной безопасности – одна из важнейших задач электронной библиотеки. Задача защиты информации от всех типов угроз включает в себя обеспечение следующих параметров:

- целостность информации, обрабатываемой, передаваемой, хранимой в ЭБ;
- доступность всех открытых для пользователей ресурсов;
- конфиденциальность доступа.

Для решения данной задачи необходимо разработать комплекс организационных и программно-технических мер, а именно:

- подготовить концепцию построения информационной безопасности электронных ресурсов;
- разработать пакет организационно-распорядительной документации, регламентирующей реализацию концепции;
- спроектировать и реализовать систему защиты информации;



Наиболее важным представляется выбор комплекса технологических средств хранения. В качестве главных носителей информации в цифровой библиотеке рационально использовать файловые серверы с объемными жесткими дисками, а резервные копии хранить на оптических носителях (CD, DVD, дисках blue-ray).

Любая современная технология длительного хранения данных, в том числе NAS-технология, при долговременном хранении, в связи с моральным и физическим старением носителей, технических и программных средств, требует миграции данных при внедрении новых современных технологий хранения данных.

Таким образом, своевременная миграция данных является тем средством, которое позволяет обеспечить длительное хранение электронных документов. Наш опыт позволяет сделать рекомендацию по обеспечению длительного хранения электронных документов. При появлении признаков устаревания технологии необходимо позаботиться о переводе устаревшей системы хранения на новые технологии (осуществить процесс миграции).

В целом, создание электронной библиотеки способствует развитию культуры обмена информацией, улучшению доступа к знаниям и научным исследованиям, а также стимулирует развитие новых технологий в области хранения и передачи данных. Развитие фондов электронных библиотек имеет и иные преимущества, способствующие повышению качества библиотечного обслуживания. Среди них можно выделить:

- развитие сервисов электронной доставки документов;
- расширение возможности информационного поиска, в том числе, поиска по полнотекстовым ресурсам с использованием разных параметров поиска: по ключевым словам, по отдельным словоформам;
- упрощение создания полнотекстовых баз данных;
- возможности использования цифровых копий для популяризации информации о фонде, коллекциях документов и их использовании в работе библиотеки.

Использование электронных книг  
Источник: hitdelight / depositphotos.com

- разработать и внедрить соответствующие программные средства;
- приобрести и установить технические средства защиты данных.

### Заключение

Создание электронной библиотеки является важным шагом в развитии современных информационных технологий и обеспечении доступности информации для широкого круга пользователей. Это обеспечивает сохранность и защиту информации от повреждений и несанкционированного доступа. При формировании электронной библиотеки необходимо учитывать множество факторов, в том числе: выбор носителей для долговременного хранения данных; обеспечение безопасности и конфиденциальности информации; создание/выбор удобных инструментов для поиска и просмотра документов.

**ENSURING THE SAFETY AND ACCESSIBILITY OF THE LIBRARY'S ELECTRONIC COLLECTIONS**

**Tsvetkova Valentina**, Chief Researcher of the BEN RAN, Doctor of Technical Sciences, Professor. E-mail: vats08@mail.ru

**Kulev Vladimir**, Head of the Informatization Department of the BEN RAN. E-mail: kulev@benran.ru

**Abstract.** The features of ensuring the safety of the electronic collections of the scientific library during the transition to electronic document presentation formats are considered. Special attention is paid to the preservation of documents in printed form, the possibility of obtaining a significant number of copies without compromising the quality of the original document by converting it to electronic format, expanding user access to information, providing virtual services to remote users and issues of long-term preservation of electronic arrays. The features of the use of various media used for long-term storage of electronic information are shown. Attention is paid to the choice of scanning methods, data presentation formats and data migration processes over time when introducing new data storage technologies. Timely data migration is the means that allows for long-term storage of electronic documents.

**Keywords:** digitization of documents, long-term data storage, long-term data storage media, storage technologies, optical disks for archival data storage, electronic library, data migration.

**Библиографический список:**

1. Сохранность библиотечных фондов: методические рекомендации / Новосибир. обл. науч. б-ка: сост. Р. М. Хаснуддинова, Н. И. Поночевная, ред. Н. И. Васильева; отв. за вып. В. Г. Деев. – Новосибирск: Изд-во НГОНБ, 2016.
2. Остапенко С. Технологии оцифровки для сохранности библиотечного фонда // Образовательный портал «Справочник». Дата последнего обновления статьи: 11.12.2023. – URL: [https://spravochnick.ru/bibliotchno-informacionnaya\\_deyatelnost/tehnologii\\_ocifrovki\\_dlya\\_sohrannosti\\_bibliotchnogo\\_fonda/](https://spravochnick.ru/bibliotchno-informacionnaya_deyatelnost/tehnologii_ocifrovki_dlya_sohrannosti_bibliotchnogo_fonda/) (дата обращения: 12.03.2024).
3. Каленов Н. Е., Погорелко К. П., Серебряков В. А., Сотников А. Н. Электронная библиотека «Научное наследие России». Состояние и перспективы развития. Труды конференции Абрау-2016. Новороссийск. 18–24 сентября 2016 г. М.: ИПМ им М. В. Келдыша. 2016. С. 148–151. DOI: 10.20948/abrau-2016-27. – URL: <http://keldysh.abrau/2016/27.pdf> (дата обращения 22.03.2024).
4. Евстигнеева Г. А. Качество оцифровки: проблемы и решения // Современная библиотека. № 5 (25), 2012. С. 58–61. ISSN 2072-3849.
5. Самбулов Д. В. Сравнительный анализ форматов файлов электронных книг.
6. Тихонов В. Архивное хранение электронных документов: проблемы и решения // Делопроизводство и документооборот на предприятии. Февраль, 2006.
7. Хранение носителей информации. – URL: <http://www.e-biblio.ru> (дата обращения: 12.03.2024).
8. ГОСТ Р 7.0.96–2016. «Система стандартов по информации, библиотечному и издательскому делу». Электронные библиотеки, ОКС 01.140.30.
9. Центральная универсальная научная библиотека имени Н. А. Некрасова. – URL: <http://www.nekrasovka.ru> (дата обращения: 12.03.2024).
10. IBM Research установила новый рекорд плотности записи на магнитную ленту. – URL: <http://habrahabr.ru/company/ibm/blog/256561/> (дата обращения: 12.03.2024).
11. Добрусина С. А., Тихонова И. Г. К вопросу о долговечности оптических дисков / Сохранение электронного контента в России и за рубежом. Сборник материалов Всероссийской конференции (Москва, 24–25 мая 2012 г.) / сост. Кузьмин Е. И., Муранова Т. А. – М., МЦБС. 2013. С. 81–89.
12. Иевлев П. Какие носители информации существуют, их история и перспективы развития. – URL: <https://digitalocean.ru/n/sposoby-hraneniya-dannyh/> (дата обращения: 19.03.2024).
13. New Gold Kodak CD and DVD Promise a 80 to 300 Years of Lifetime. – URL: <https://www.cdinfo.com/d7/content/new-gold-kodak-cd-and-dvd-promise-80-300-years-lifetime> (дата обращения 20.03.2024).
14. Шляева Т. Хранить вечно. Технологии M-DISK. – URL: <https://nasledie.digital/articles/hranit-vechno-tehnologiya-m-disc/> (дата обращения 19.03.2024).
15. Изготовление цифровых копий фонда пользования с микроформ архивных документов. – URL: <https://archives.ru/documents/reglament-scan-microfilm.shtml> (дата обращения: 12.03.2024).
16. Арский Ю. М., Леонтьева Т. М., Никольская И. Ю., Шогин А. Н. Банк данных ВИНТИ: состояние и перспективы развития. – М.: ВИНТИ, 2006. С. 11. ISBN 5-94577-027-2.

17. Залаев Г. З., Каленов Н. Е. Цветкова В. А. Оцифровка документов в научных архивах и библиотеках: вопросы и ответы / НТИ. Сер. 1. 2016. № 2. С. 14–21. ISBN 0548-0019.
18. Брэдлей К. Оценка рисков, связанных с использованием записываемых CD и DVD носителей для хранения архивных данных. Стратегии и альтернативы / Национальная Библиотека Австралии, Канберра. ЮНЕСКО, Париж, 2006. – URL: <https://portal.unesco.org>
19. Пилипчук М. И., Балакирев А. Н., Залаев Г. З., Лисютин А. П. Рекомендации по созданию оцифрованных копий фонда пользования фото- и фотодокументов. – URL <http://rgantd.ru/ntr.php> (дата обращения: 12.03.2024).
20. Пилипчук М. И., Балакирев А. Н., Дмитриева Л. В., Залаев Г. З. Рекомендации по обеспечению сохранности информации, записанной на оптических дисках (тестирование выборочного массива документов федеральных архивов). – URL: <http://archives.ru/documents/recomndacii-po-obcspecheniiu-sochrannsti-informacina-diskah.shtml> (дата обращения: 12.03.2024).
21. Полулин С. Диски Blu-ray проработают 200 лет. – URL: [https://www.cnews.ru/news/top/diski\\_bluray\\_prorobotayut\\_200 лет](https://www.cnews.ru/news/top/diski_bluray_prorobotayut_200 лет) (дата обращения 22.03.2024).
22. Data Tresor Disk. – URL: <https://zpk42.ru/1516-datatresordisc.html/> (дата обращения: 11.03.2024).
23. Изучаем накопители для длительного хранения данных. – URL: <http://ichip.ru/izuchаем-nakopiteli-dlva-dlilcnogo-hraneniya-dannih.html/3> (дата обращения: 12.03.2024).
24. Залаев Г. З., Тихонов А. В., Глишинская Н. В., Новиков С. Л. Регламент «Изготовление цифровых копий фонда пользования с микроформ архивных документов». Москва, 2012. – 61 с. – URL: <http://archives.ru/documents/reglament-scan-mieronim.shtml> (дата обращения: 12.03.2024).
25. ГОСТ Р 7.0.2–2006 СИБИБ. «Консервация документов на компакт-дисках. Общие требования». URL: <http://docs.cntd.ru/document/1200045442> (дата обращения: 12.03.2024).
26. Левченко Л. Л. Обеспечение сохранности электронных документов в Национальном архиве США. Вестник архивиста-RU. 10 июля 2015. – URL: <http://vestarchive.ru>. ISSN 2073-0101 (дата обращения: 12.03.2024).
27. Тихонов В. И. Архивное хранение электронных документов: проблемы и решения // Делопроизводство и документооборот на предприятии, февраль 2006. – URL: <http://www.dclo-press.ru> (дата обращения: 12.03.2024).
28. Цветкова В. А., Кочукова Е. В. Библиотека как институт социальной памяти / Библиография и книговедение. № 2, 2015. С. 46–49. ISSN 2411-2305.
29. Остапенко С. Права пользователей библиотек // Образовательный портал «Справочник». Дата последнего обновления статьи: 18.12.2023. – URL: [https://spravochnick.ru/bibliotchno-informacionnaya\\_deyatelnost/prava\\_polzovateley\\_bibliotek/](https://spravochnick.ru/bibliotchno-informacionnaya_deyatelnost/prava_polzovateley_bibliotek/) (дата обращения: 12.03.2024).

**Bibliography:**

1. Preservation of library collections: methodological recommendations / Novosibirsk region scientific library: comp. R. M. Khasnutdinova, N. I. Ponochevna, ed. N. I. Vasilyev; rev. for issue V. G. Deev. Novosibirsk: NGONB Publishing House, 2016.
2. Ostapenko S. Digitization technologies for the preservation of the library fund // Educational portal «Handbook». The date of the last update of the article: 11.12.2023. – URL: [https://spravochnick.ru/bibliotchno-informacionnaya\\_deyatelnost/tehnologii\\_ocifrovki\\_dlya\\_sohrannosti\\_bibliotchnogo\\_fonda/](https://spravochnick.ru/bibliotchno-informacionnaya_deyatelnost/tehnologii_ocifrovki_dlya_sohrannosti_bibliotchnogo_fonda/) (date of access: 03/12/2024).
3. Kalenov N. E., Pogorelko K. P., Serebryakov V. A., Sotnikov A. N. Electronic Library «Scientific Heritage of Russia». The state and prospects of development. Proceedings of the Abrau-2016 conference. Novorossiysk. September 18–24, 2016 Moscow: IPM named after M. V. Keldysh. 2016. pp. 148–151. DOI: 10.20948/abrau-2016-27. – URL: <http://keldysh.abrau/2016/27.pdf> (accessed 03/22/2024).
4. Evstigneeva G. A. Digitization quality: problems and solutions // A modern library. No. 5 (25), 2012. pp. 58–61. ISSN 2072-3849.
5. Sambulov D. V. Comparative analysis of e-book file formats.
6. Tikhonov V. Archival storage of electronic documents: problems and solutions // Office management and document management at the enterprise. February, 2006.
7. Storage of media. – URL: <http://www.e-biblio.ru> (accessed 12.03.2024).
8. ГОСТ Р 7.0.96–2016. «The system of standards for information, library and publishing». Electronic libraries, OKS 01.140.30.
9. The Central Universal Scientific Library named after N. A. Nekrasov. – URL: <http://www.nekrasovka.ru> (date of application: 03/12/2024).
10. IBM Research has set a new record for magnetic tape recording density. – URL: <http://habrahabr.ru/company/ibm/blog/256561/> (date of access: 03/12/2024).
11. Dobrusina S. A., Tikhonova I. G. On the issue of optical disk durability / Preservation of electronic content in Russia and abroad. Collection of materials of the All-Russian Conference (Moscow, May 24–25, 2012) / comp. Kuzmin E. I., Muranova T. A. – M., ICBC. 2013. pp. 81–89.
12. Ievlev P. What media exist, their history and development prospects. – URL: <https://digitalocean.ru/n/sposoby-hraneniya-dannyh/> (accessed: 03/19/2024).
13. New Gold Kodak CD and DVD Promise a 80 to 300 Years of Lifetime. – URL: <https://www.cdinfo.com/d7/content/new-gold-kodak-cd-and-dvd-promise-80-300-years-lifetime> (date of appeal 03/20/2024).
14. Shklyueva T. Keep forever. M-DISK technologies. – URL: <https://nasledie.digital/articles/hranit-vechno-tehnologiya-m-disc/> (date of issue 03/19/2024).
15. Production of digital copies of the fund for use with microforms of archival documents. – URL: <https://archives.ru/documents/reglament-scan-microfilm.shtml> (date of application: 03/12/2024).
16. Arsky Yu. M., Leontieva T. M., Nikolskaya I. Yu., Shogin A. N. VINITI data bank: state and prospects of development. – M.: VINITI, 2006. pp. 11. ISBN 5-94577-027-2.
17. Zalaev G. Z., Kalenov N. E. Tsvetkova V. A. Digitization of documents in scientific archives and libraries: questions and answers / NТИ. Ser. 1. 2016. No.2. pp. 14–21. ISBN 0548-0019.

18. Bradley K. Assessment of the risks associated with the use of recordable CD and DVD media for storing archival data. Strategies and Alternatives / National Library of Australia, Canberra. UNESCO, Paris, 2006. – URL: <https://portal.unesco.org>
19. Pilipchuk M. I., Balakirev A. N., Zalaev G. Z., Lisutin A. P. Recommendations for the creation of digitized copies of the fund for the use of photographic and photographic documents. – URL <http://rgantd.ru/ntr.php> (date of application: 03/12/2024).
20. Pilipchuk M. I., Balakirev A. N., Dmitrieva L. V., Zalaev G. Z. Recommendations for ensuring the safety of information recorded on optical discs (testing a sample array of documents from federal archives). – URL: <http://archives.ru/documents/recomndacii-po-obcspecheniiu-sochrannsti-informacina-diskah.shtml> (date of application: 03/12/2024).
21. Pölsulin S. Blu-ray discs will work for 200 years. – URL: [https://www.cnews.ru/news/top/diski\\_bluray\\_prorobotayut\\_200 лет](https://www.cnews.ru/news/top/diski_bluray_prorobotayut_200 лет) (accessed 03/22/2024).
22. Data Tresor Disk. – URL: <https://zpk42.ru/1516-datatresordisc.html/> (date of access: 03/11/2024).
23. We are studying drives for long-term data storage. – URL: <http://ichip.ru/izuchаем-nakopiteli-dlva-dlilcnogo-hraneniya-dannih.html/3> (date of application: 03/12/2024).
24. Zalaev G. Z., Tikhonov A. V., Glishchinskaya N. V., Novikov S. L. Regulations «Production of digital copies of the fund for use from microforms of archival documents». Moscow, 2012. – 61 p. – URL: <http://archives.ru/documents/reglament-scan-mieronim.shtml> (date of application: 03/12/2024).
25. ГОСТ Р 7.0.2–2006 СИБИБ. «Preservation of documents on CDs. General requirements». URL: <http://docs.cntd.ru/document/1200045442> (accessed 12.03.2024).
26. Levchenko L. L. Ensuring the safety of electronic documents in the National Archives of the USA. Archivist's Bulletin-R U. July 10, 2015. – URL: <http://vestarchive.ru>. ISSN 2073-0101 (date of reference: 03/12/2024).
27. Tikhonov V. I. Archival storage of electronic documents: problems and solutions // Office management and document management at the enterprise, February 2006. – URL: <http://www.dclo-press.ru> (date of reference: 03/12/2024).
28. Tsvetkova V. A., Kochukova E. V. Library as an institute of social memory / Bibliography and book studies. No. 2, 2015. pp. 46–49. ISSN 2411-2305.
29. Ostapenko S. The rights of library users // Educational portal «Handbook». The date of the last update of the article: 12/18/2023. – URL: [https://spravochnick.ru/bibliotchno-informacionnaya\\_deyatelnost/prava\\_polzovateley\\_bibliotek/](https://spravochnick.ru/bibliotchno-informacionnaya_deyatelnost/prava_polzovateley_bibliotek/) (date of access: 03/12/2024).

## ВЕБ-АРХИВ: ВАЖНЕЙШИЙ ИСТОЧНИК ИНФОРМАЦИИ VS НЕУПОРЯДОЧЕННОЕ ХРАНИЛИЩЕ ВЕБ-КОНТЕНТА

*Аннотация. В статье рассматривается эволюция развития веб-архивов, а также возможности и перспективы их использования в качестве ресурсов открытого доступа. Приведены данные крупнейших агрегаторов веб-архивов (Archive-it, интернет-архив), свидетельствующие об их разнообразии и перспективах развития. Определены проблемы, связанные со сбором, сохранением и предоставлением доступа к заархивированному веб-контенту, решению которых способствует Международный консорциум по сохранению интернета.*

**Редькина Наталья**  
Заведующий отделом,  
главный научный сотрудник,  
д. п. н., Государственная  
публичная научно-техническая  
библиотека Сибирского  
отделения Российской  
академии наук (ГПНТБ СО РАН)  
E-mail: redkina@spsl.nsc.ru

### Ключевые слова:

веб-архивы, история, перспективы, открытый доступ, библиотеки, тренды, Международный консорциум по сохранению интернета, интернет-архив.

### Введение

Коллекции веб-архивов зарекомендовали себя в качестве ценных источников информации для пользователей. Выступая в качестве цифровых экосистем для расширения возможностей поиска открытой информации, они предоставляют широкие возможности для получения новых знаний, восстановления утраченной информации, а также проверки ранее опубликованных данных, часто имеющей большое культурное, научное, образовательное, художественное и социальное значение. Поскольку интернет в настоящее время является доминирующим средством общественной коммуникации, веб-архивы важны для отслеживания эволюции общества с течением времени [1]. Ученые, журналисты, социологи, историки и другие исследователи, а также широкая общественность, желающие вернуться и изучить предшествующую информацию о политических, исторических и других событиях, используют веб-архивы как важный ресурс для изучения изменений на веб-страницах, исследования архивов социальных сетей [2, 3, 4, 5], т. к. они сохраняют содержимое миллиардов веб-страниц в том виде, в котором они были в определенный момент времени. Веб-архивы содержат огромный неиспользованный потенциал для анализа изменений

веб-контента с течением времени, что особенно полезно для поиска материалов, которых больше не существует.

Целью веб-архивирования является сохранение исходной формы собранного контента без изменений. Для достижения этой цели разрабатываются инструменты, стандарты и политики, уже наработаны практики организации, которые со временем обеспечивают управление веб-архивами. С одной стороны, распространение интернета, а с другой, – проблемы с потерей информации в нем, стали отправной точкой развития сотрудничества в сфере веб-архивирования, что позволило далеко продвинуться в решении многих вопросов. Вместе с тем, несмотря на актуальность и важность веб-архивирования, совершенствование технологий и появление опыта различных организаций по созданию и поддержке веб-архивов, до сих пор возникают проблемы их использования по условиям доступа (часто только стационарно в зданиях библиотек), возможностям поиска внутри веб-архивов и др.

### Эволюция и тенденции развития веб-архивирования

Н. Брюггер, считая интернет неотъемлемой частью культуры, а не просто техническим устройством, в 2005 г. одним из первых опубликовал книгу под названием «Архивирова-

**Веб-архивы  
стали элементом  
инфраструктуры  
мирового  
информационного  
пространства,  
поэтому крайне  
важно, чтобы они  
предоставлялись  
без ограничений**

ние веб-сайтов: общие соображения и стратегии» [6], в которой системно представил свое видение сохранения веб-контента с точки зрения исследований медиапространства. В следующем году J. Masapeš написал книгу под названием «Веб-архивирование» [7]. В последующие годы количество публикаций по веб-архивированию увеличивалось, что свидетельствует об актуальности вопроса, а также отражает социальную значимость и технологическую сложность процессов создания веб-архивов [8, 9].

Основные усилия по сохранению интернета были впервые предприняты в рамках реализации проекта «Интернет-архив»<sup>1</sup>. В 1996 г. Национальная библиотека Австралии, Национальная библиотека Швеции и интернет-архив приступили к веб-архивированию. К 1998 г. национальные библиотеки в скандинавских странах Дании, Финляндии, Исландии и Норвегии стали вести веб-архивы, а с 2000 г. веб-архивированием занимаются национальные библиотеки и архивы, правительственные организации и исследовательские институты, частные компании, отдельные ученые и пользователи, заинтересованные в долгосрочном сохранении веб.

Глобальные проблемы человечества и иные значимые события общественной жизни способствовали появлению инновационных способов сохранения веб-контента, возникновению новых технологий и инструментов веб-архивирования [10]. Группа по развитию контента Международного консорциума по сохранению интернета<sup>2</sup> координировала международную инициативу по выбору соответствующей онлайн-информации, документирующей, например, «Европейский кризис беженцев», «Олимпийские игры», «Празднование Первой мировой войны» и др. Национальная медицинская библиотека США осуществляла работу по сбору и архивированию контента (веб-страниц и социальных сетей), связан-

ного с пандемией коронавирусного заболевания (COVID-19) в 2020 г., наряду с другими более ранними проблемами в медицине и биологии (вспышка Эболы в 2014 г., вирус Зика в 2016 г., ВИЧ/СПИД, эпидемия опиоидов, др.), документируя возникающий кризис в области здравоохранения в рамках постоянной коллекции веб-архивов Global Health Events Web Archive [11, 12].

В 2024 г. самый большой веб-архив в мире – интернет-архив, созданный в 1996 г., предлагает бесплатный доступ к 866 миллиардам заархивированных веб-страниц, а также к книгам (44 млн), более 10 млн видеоматериалов, 4,8 млн изображений, 1 млн программ и 15 млн аудиозаписей, в т. ч. 255 000 живых концертов. Вместе с тем, несмотря на уже более чем 25-летнюю историю веб-архивирования, поднимаются новые проблемы, связанные со сбором, обработкой, долгосрочным сохранением все возрастающего количества сайтов. Интернет-архив обладает наибольшим количеством веб-архивов. Доступ к этим активам осуществляется через набор интерфейсов к Wayback Machine. Любой пользователь, у кого есть бесплатная учетная запись, может загружать медиафайлы в интернет-архив. Одним из показателей востребованности веб-архивов являются данные, полученные с помощью веб-аналитического инструмента SimilarWeb, в глобальном рейтинге которого интернет-архив в марте 2024 г. занимал 181 место с общим количеством визитов за месяц в 138 млн, а в рейтинге категории «Наука и образование > Наука и образование – Другое» ему принадлежит первое место.

Ведущий сервис Internet Archive's Archive<sup>3</sup> с 2006 г. предоставил услуги веб-архивирования более 1200 организациям из более чем 24 стран, включая библиотеки, исследовательские учреждения, социальные группы и др. Archive-it предоставляет инструменты, обучение и техническую поддержку для сбора и сохранения динамических веб-материалов, а также платформу для партнеров, позволяющую делиться своими коллекция-

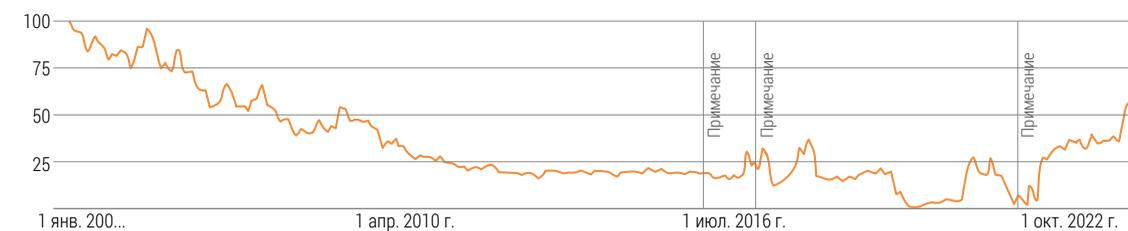


Рис. 1. Динамика популярности запросов по теме «Веб-архивирование» в Google Trends (примечание в рисунке указывает на наличие изменений в системе сбора данных)

ми с многочисленными возможностями поиска, обнаружения и доступа. Пользователи Archive-it сохранили более 40 млрд цифровых записей, опубликованных в интернете, общим объемом данных в петабайты. Материалы, заархивированные через Archive-it, хранятся в некоммерческих центрах обработки данных, принадлежащих и управляемых интернет-архивом, доступны пользователям для самостоятельной загрузки для дополнительного сохранения и совместного использования. Другие веб-архивы демонстрируют целенаправленные усилия по сбору данных, часто обеспечивая уникальной информацией пользователей, заполняя временные и пространственные пустоты интернета [13].

Согласно данным, полученным с помощью приложения Google Trends, которое показывает, насколько популярно определенное слово или словосочетание по сравнению со всеми запросами за какой-либо период времени в различных регионах мира и на различных языках<sup>4</sup>, проблема веб-архивирования остается актуальной на протяжении последних 20 лет (с 2004 по 2024 г.) во всем мире (рис. 1). Числа на графике отображают популярность словосочетания «Веб-архивирование» (англ. web archiving) по сравнению с общим количеством поисковых запросов в Google за определенный промежуток времени. После некоторого снижения показателей, начиная с 2022 г., можно констатировать рост интереса к теме. Отметим, что получаемые в ходе анализа

данные обезличены и нормализованы (разделены на масштабирующую переменную), значения приводятся в интервале от 0 до 100. Алгоритмы Google определяют точку на графике за выбранный период, когда запрос был наиболее популярен, и принимают его за 100. Все остальные точки на графике определяются в процентном отношении к максимуму.

Кроме того, Google Trends позволяет сравнивать статистику по поисковым запросам, географии (регионам), периодам, на разных языках и практически в реальном времени. В системе Google Trends выявлены запросы, по которым дается корректный результат с учетом географического фильтра, и определен список стран, где чаще всего выполнялся поиск по запросу (таблица 1). В пятерку стран-лидеров вошли Эстония, Латвия, Чехия, Польша и Россия.

Приведенные выше данные свидетельствуют об эволюционном развитии веб-архивирования, значительном внимании к этому вопросу в разных странах, а также больших объемах уже заархивированного веб-контента.

### Проблемы и перспективы веб-архивов

С тех пор, как веб-архивы начали собирать веб-сайты в 1990-х гг., в разных дисциплинах подчеркивали потенциал новых данных, которые стали доступны. Однако некоторые исследователи стали интерпретировать, к примеру интернет-архив, не столько как архив, сколько как «большое несортированное

<sup>1</sup> URL: <https://archive.org>

<sup>2</sup> International Internet Preservation Consortium, IIPC. URL: <https://netpreserve.org>

<sup>3</sup> URL: <https://archive-it.org>

<sup>4</sup> URL: <https://www.google.ru/trends/>

хранилище», в котором отсутствует оценка контента, сообществ и культуры использования, что делает его скорее мусорным контейнером, где «вещи не потеряны, но и не сохранены» [14, 15, 16]. Ruest N., Fritz S. и Milligan I. также приходят к выводу, что веб-архивы – это беспорядок, обусловленный спецификой сбора, хранения и предоставления доступа к этим многогранным цифровым объектам, требующим применения традиционных архивных методов описания к коллекциям, созданным в цифровом формате, что, с другой стороны, затруднено из-за проблем происхождения, оригинального порядка и масштаба [17]. По мнению Niels Brügger, веб-архивы являются скорее библиотеками, поскольку они сохраняют то, что стало общедоступным, и поэтому национальный веб-архив, задача которого заключается в сохранении национальной сети, скорее следует называть «веб-библиотекой», то есть «сборником веб-публикаций» [18].

Критическое рассмотрение природы и достоверности веб-архивов как исторических источников начало рассматриваться различными исследователями, занимающимися критикой веб-архивов. Эта критика выявила не только ограничения веб-архивов как данных, но и возможности [19, 20].

Рассмотрим, какие проблемы обозначены в последние годы в технологии веб-архивирования.

**Пробелы в архивах.** Вопрос о том, что архивировать и как обеспечить доступ

Таблица 1. Популярность запросов по теме «Веб-архивирование» по регионам

Страна	Количество баллов
Эстония	100
Латвия	67
Чехия	61
Польша	55
Россия	54
Молдова	47
Украина	47
Финляндия	45
Литва	45
Япония	45

пользователей к ресурсам, уже давно обсуждается в традиционной архивной практике, и интернет не является исключением [21]. Ben-David A. справедливо замечает, что веб-архивы могут быть предвзятыми и содержать значительные пробелы в знаниях [22], т. к. фиксируются отдельные URL-адреса в режиме реального времени. Но поскольку сохранение всей сети технически невозможно, большинство национальных учреждений веб-архивирования разработали политику, направленную на уточнение объемов архивирования (полный домен или специальные коллекции), границ архивирования (все, что находится в домене верхнего уровня с кодом страны, веб-сайты на национальном языке, размещенные в другом месте в интернете), частоты архивирования и так далее. К примеру, Национальная библиотека Испании выбрала смешанную модель, сочетающую в себе массивные и выборочные коллекции. Целью Испанского веб-архива<sup>5</sup> является сохранение и облегчение будущего доступа ко всему испанскому контенту, опубликованному в интернете (веб-сайты, блоги, форумы, списки, рассылки, документы, изображения, видео и т. д.). Для этого библиотека собирает испаноязычные страницы и веб-сайты, размещенные в домене .es, а также в других общих доменах и поддоменах (.com; .edu; .gob; .org; .net; .biz и т. д.).

Еще одна из проблем, способствующая неполноте веб-контента – это наличие авторизованного входа к некоторым ресурсам, защищенного паролем.

**Сложность работы для пользователей.**

Собранные с использованием различных платформ веб-ресурсы объединяются в стандартизированные форматы файлов: файл ARC, за которым следует формат-преемник WARC (предложенный в качестве стандарта ISO в 2009 г. и стандарт 2017 г.). Ресурсы смешаны, то есть текст веб-страницы перемежается многогранными цифровыми объектами и данными, такими как HTML, CSS, JavaScript, изображения в двоичном кодировании, ви-

<sup>5</sup> URL: <https://www.bne.es>

део, документы или устаревшие форматы файлов, что усложняет работу пользователей с ними.

**Блокировка сканеров.** Веб-сайты могут быть более или менее пригодными для архивирования в зависимости от платформы или базовой технологии, используемой для создания и обслуживания веб-сайта. Архивирование проприетарного программного обеспечения является серьезной проблемой. Многие веб-платформы (например, социальные сети) намеренно блокируют веб-сканеры, и их невозможно архивировать. Такие объекты, как комментарии, добавленные к новостным онлайн-статьям или блогам, особенно неуловимы для технологий веб-архивирования, поскольку они часто загружаются в браузер динамически и недоступны большинству автоматических веб-сканеров [23]. Архивные веб-сканеры также сталкиваются с проблемами из-за потокового мультимедиа, анимации и встроенных социальных сетей, контента, защищенного паролем, для которого требуется вход пользователя в систему. В совокупности эти технические ограничения означают, что наследие, созданное процессом веб-архивирования, непропорционально представляет веб-контент.

**Ошибки воспроизведения.** Разнообразие веб-технологий и способов взаимодействия пользователей с веб-страницами являются основными проблемами веб-архивирования. Фактически, создание высококачественных архивов, которые достоверно воспроизводят структуру, содержание и поведение веб-страниц на основе архивных данных, является в значительной степени нерешенной проблемой: по оценкам, 44% заархивированных веб-страниц имеют незначительные или серьезные дефекты из-за ошибок воспроизведения [24].

При воспроизведении заархивированной веб-страницы основное ожидание состоит в том, что страница должна быть доступна для просмотра и функционировать точно так же, как и во время архивирования. Однако это ожидание требует, чтобы веб-архивы при воспроизведении изменяли страницу и ее встроенные ре-



сурсу так, чтобы все ресурсы и ссылки ссылались на архив, а не на исходный сервер. Хотя эти модификации обязательно меняют состояние представления, понятно, что без них воспроизведение страниц из архива было бы невозможно. Процесс воспроизведения страниц и изменений, вносимых в представления веб-архивами, варьируется в зависимости от архива [25].

Среди проблем архивного скриншота называют:

- 1) ошибки существования, т. е. элементы, которые присутствуют в оригинале, но отсутствуют в архиве;
- 2) ошибки положения, которые показывают отличие от соответствующего сегмента на исходном скриншоте;
- 3) ошибки внешнего вида в содержании или макете заархивированного элемента по сравнению с его оригиналом, включая различия в цвете [26].

**Проблемы доступности для людей с ограниченными возможностями.** Исследователями обнаружены проблемы доступности веб-архивов для людей с ограниченными возможностями, особенно с нарушениями зрения [1]. Общими проблемами доступности для всех веб-архивов были отсутствие альтернативного текста для визуального контента, такого как изображения, отсутствие подходящих ориентиров для быстрой навигации по веб-сайтам, неправильная структура контента без правильных заголовков, а также пустые ссылки и кнопки.

**Юридические проблемы.** Несмотря на усилия по управлению цифровыми правами с помощью таких инициатив, как Creative Commons, проблема защиты прав интеллектуальной собственности (ПИС) на онлайн-артефакты, таких как авторское право, далека от решения. Получение исходной информации для защиты прав интеллектуальной собственности, например, кто является автором или какие права использования применяются, является невыполнимой задачей для большинства ресурсов, доступных в интернете. Веб-архивы, как службы цифрового хранения, максимально точно сохраняют оригинальные артефакты со всеми их досто-

инствами и недостатками. Таким образом, проблемы прав интеллектуальной собственности в отношении онлайн-артефактов сохраняются и после того, как они помещаются в веб-архив. С течением времени их становится все труднее решать. Возможно, в будущем общества разработают эффективные модели управления правами интеллектуальной собственности на онлайн-ресурсы [8].

Проблемы, с которыми сталкивается программа веб-архивирования, включают юридические проблемы, уникальные для стран. К примеру, в Ирландии исключается регулярное полное доменное веб-архивирование, существуют этические проблемы и проблемы сохранности, связанные с выборочным тематическим сбором, а также проблема сбора ирландского веб-пространства в больших масштабах. Эволюция веб-архивирования в Национальной библиотеке Ирландии за последнее десятилетие рассматривается как ответ на эти проблемы. Возможность, предоставляемая искусственным интеллектом, рассматривается как неявная и очень желанная деятельность в среде, где существуют серьезные проблемы со сбором данных в больших масштабах [27].

**Доступ только в офлайн-режиме.** Определено, что анализ, ориентированный на данные, требует значительных технических возможностей для поддержки больших объемов данных веб-архивов, но и требует понимания правовых и этических ограничений, а также негласных организационных практик, которые определяют сбор и использование архивных веб-данных [28].

Некоторые национальные архивы, библиотеки и учреждения наследия имеют мандат на архивирование национального веб-домена или на создание тематических коллекций веб-сайтов/страниц, что зафиксировано на законодательном уровне отдельно взятой страны. Чаще всего национальные библиотеки координируют сохранение контента веб-архивов национального масштаба. Аналогичным образом, в университетах право собственности на портфолио веб-архивирования обычно принадлежит библиотекам [17]. В других случаях необходимо учитывать

правовые вопросы при веб-архивировании, если это не касается официальных ресурсов различных органов власти, и уведомлять владельцев сайта о включении контента в архив. В некоторых случаях в электронном письме запрашивается разрешение на архивирование или предоставление исследователям удаленного доступа. Учитывая эти проблемы, к большому количеству веб-архивов доступ предоставляется только в читальных залах библиотек, в том числе из-за соображений конфиденциальности.

**Сложности поиска отдельных видов веб-контента.** Такой контент, как музыка и видео, которые являются встроенными файлами с внешних сайтов, например, YouTube, или контент, для которого требуется внешний медиаплеер, не архивируется при сборе в интернете, для него часто необходимо предусмотреть технологию сохранения с коллекции.

Поддержка поиска изображений в веб-архивах породила новые проблемы. А. Mourão и D. Gomes разработали крупномасштабную систему поиска изображений, которая позволяет осуществлять поиск по миллиардам исторических изображений, заархивированных в интернете с 1990-х гг. [29]. Объем данных, которые нужно обработать, был большим и неоднородным: в сумме он составил более 530 ТБ исторических веб-данных, опубликованных с первых дней существования интернета. Основным вкладом этой работы является набор алгоритмов, которые извлекают текстовые метаданные для описания изображений, заархивированных в интернете, архитектура системы и рабочий процесс для индексации больших объемов изображений, с учетом их конкретных временных особенностей, а также алгоритм ранжирования для упорядочения поиска изображений. Это исследование было применено для запуска расширенной службы поиска изображений, которая общедоступна с марта 2021 г. Все разработанное программное обеспечение полностью доступно как бесплатное с открытым исходным кодом.

**Проблемы организации и технические сбои.** Не все веб-архивы получают успеш-

ное развитие, встречаются трудности на пути даже крупных организаций. Так, пилотный проект веб-архива Чилийской национальной библиотеки<sup>6</sup> был приостановлен, осуществляется оценка на предмет возможного включения в Национальную цифровую библиотеку.

Британская библиотека испытала сбой в работе после кибератаки [30], прекратив доступ к веб-архиву Великобритании, но продолжая сканировать или получать копии веб-сайтов, гарантируя, что коллекции обновляются и сохраняются, предлагая несколько альтернативных веб-архивов с открытым доступом. В частности предлагался интернет-архив, веб-архив правительства Великобритании<sup>7</sup>, открытый доступный архив, в котором хранится информация центрального правительства Соединенного Королевства, включая видео, твиты, изображения и веб-сайты, датируемые с 1996 г. по настоящее время; веб-архив парламента Великобритании<sup>8</sup> – общедоступный архив охватывает парламентские веб-сайты и контент социальных сетей с 2009 г. по настоящее время; веб-архив Национальных архивов

<sup>6</sup> URL: <http://archivoweb.bibliotecanacionaldigital.cl>

<sup>7</sup> URL: <https://www.nationalarchives.gov.uk/webarchive/>

<sup>8</sup> URL: <https://ukparliament.mirrorweb.com/>

Девушка читает электронную книгу  
Источник: cylibris.com



Шотландии<sup>9</sup> – архив с открытым доступом позволяет просматривать и искать веб-сайты, связанные с людьми и историей Шотландии и др.

Таким образом, несмотря на уже достаточно большие объемы веб-архивов и разработанные методики и технологии архивирования, проблемы остаются. Они связаны с эфемерной природой интернета, несбалансированными коллекциями из-за отсутствия разрешений и др. Совместное веб-архивирование – это подход, использующий устоявшиеся стратегии в новой парадигме для преодоления внутренних ограничений, укорененных в теории и практике традиционного архивирования [21]. Возможные пути развития веб-архивирования видятся в свете сохранения личных страниц пользователей: отражение персональных цифровых архивов в веб-коллекциях, посвященных культуре региона [31]. Н. М. Балацкой и М. Б. Мартиросовой предложена модель краеведческого веб-архивирования в реальных условиях российских центральных библиотек регионов и публичных библиотек с подробным описанием всех технологических процессов формирования краеведческого веб-архива [32].

Веб-архивы – важные ресурсы, которые собирают и сохраняют постоянно растущий цифровой онлайн-контент, чтобы эта информация в будущем была доступна всем заинтересованным. И в том, что у веб-архивов есть будущее, нет сомнений. Есть уверенность, что в развитии веб-архивов большую роль может сыграть созданный в июле 2003 г. Международный консорциум по сохранению интернета<sup>10</sup>. В 2024 г. членами организации являются национальные, университетские и региональные библиотеки и архивы из 35 стран. На сайте IIPC представлена информация о развитии и пользе веб-архивов, а также инструменты веб-архивирования.

Основными задачами консорциума являются:

<sup>9</sup> URL: <https://webarchive.nrscotland.gov.uk/>

<sup>10</sup> International Internet Preservation Consortium, IIPC.  
URL: <https://netpreserve.org/>



Компьютерный класс  
Источник: Rawpixel / depositphotos.com

- выявление и развитие передового опыта отбора, сбора, сохранения и предоставления доступа к интернет-контенту;
- содействие широкому международному охвату контента веб-архивов посредством информационно-просветительской работы и создания курируемых совместных коллекций;
- продвижение инициатив и законодательства, которые поощряют сбор и сохранение интернет-контента;
- поощрение и облегчение исследовательского использования архивного интернет-контента.

Согласно стратегии развития IIPC на 2021–2025 гг. [33], консорциум планирует:

- предоставлять доступ к форуму для обмена знаниями о веб-архивировании;
- разрабатывать и рекомендовать стандарты сбора, сохранения и предоставления долгосрочного доступа к интернет-контенту;
- способствовать разработке и обеспечению устойчивости программного обеспечения и инструментов;

- повышать осведомленность о проблемах и инициативах по сохранению интернета посредством таких мероприятий, как совместный сбор информации, конференции, семинары, учебные мероприятия и публикации;
- работать напрямую с исследователями и исследовательскими сетями для содействия использованию и анализу архивного интернет-контента.

На сайте IIPC представлена информация о развитии и пользе веб-архивов, а также инструменты веб-архивирования.

## Заключение

Веб-архивы прошли серьезный путь развития. Сохранив веб-сайты и другую онлайн-информацию, они стремятся обеспечить широкое представление разнообразных и несобранных иным образом источников, позволяя находить и повторно использовать данные внутри самого веб-архива и в коллекциях, которая может представлять собой ценнейший информа-

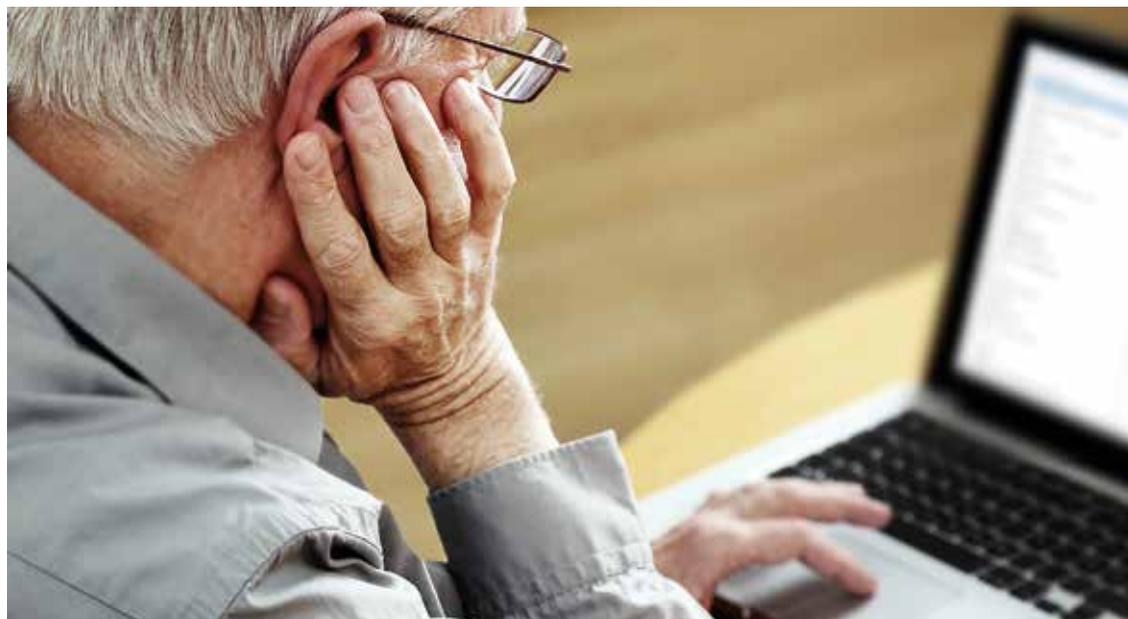
ционный ресурс, возможно уже исчезнувший из сети. Представляется важным работа библиотек в области веб-архивирования с учетом рекомендаций Международного консорциума по сохранению интернета и активно развиваемых информационных технологий. Почему это задача библиотек? Библиотеки, начиная с момента возникновения, занимаются сбором, обработкой и предоставлением в доступ информации. Они одни из первых приступили к процессам архивирования и наработали определенный опыт в этой сфере; библиотечные специалисты имеют компетенции в области аналитико-синтетической обработки информации, работе с разнородными информационными ресурсами и пр., что не оставляет сомнений в перспективности данной деятельности. Доступность веб-архивов для исследователей представляет ценность веб-сайтов для научного анализа. Веб-архивы стали элементом инфраструктуры мирового информационного пространства, поэтому крайне важно, чтобы они

также предоставлялись без ограничений и могли использоваться при решении различных задач.

В результате исследования определены проблемы веб-архивов, связанные с авторскими правами, ограниченными метаданными в описании, структурированием коллекций, ошибками воспроизведения и др., что затрудняет их использование в полной мере. Сделан вывод о необходимости создания многоуровневой системы веб-архивов с помощью рекомендованных Международным консорциумом по сохранению интернета стандартов сбора, сохранения и предоставления долгосрочного доступа к интернет-контенту, а также современных инструментов веб-архивирования.

*Исследование выполнено в рамках реализации научного проекта ГПНТБ СО РАН (2022–2026 гг.) «Разработка модели функционирования научной библиотеки в информационной экосистеме открытой науки» № 122041100150-3.*

Работа с веб-архивом  
Источник: mikdam / depositphotos.com



## WEB ARCHIVE: AN IMPORTANT SOURCE OF INFORMATION VS DISORDERED WEB CONTENT STORAGE

**Redkina Natalya**, Head of department, Doctor of Pedagogical Sciences, State Public Scientific and Technological Library of the Siberian Branch of Russian Academy of Sciences. E-mail: redkina@spsl.nsc.ru

**Abstract.** The article examines the evolution of the development of web archives, as well as the possibilities and prospects for their use as open access resources. Data from the largest aggregators of web archives (Archive-it, Internet Archive), indicating their diversity and development prospects are presented. Problems associated with the collection, preservation and provision of access to archived web content are identified, the solution of which is facilitated by the International Internet Preservation Consortium.

**Keywords:** web archives, history, prospects, open access, libraries, trends, International Consortium for Internet Preservation, Internet Archive.

### Библиографический список:

- Sunkara M., Kalari S., Jayarathna S., Ashok V. Assessing the Accessibility of Web Archives // 2023 ACM/IEEE Joint Conference on Digital Libraries (JCDL), Santa Fe, NM, USA, 2023. pp. 253–255. DOI: 10.1109/JCDL57899.2023.00048.
- Jayanetti H. R. et al. Creating Structure in Web Archives with Collections: Different Concepts from Web Archivists // Silvello G. et al. Linking Theory and Practice of Digital Libraries. TPDL 2022. Lecture Notes in Computer Science. Springer, Cham. 2022. Vol. 13541. DOI: 10.1007/978-3-031-16802-4\_45.
- Frew L., Nelson M. L., Weigle M. C. Making Changes in Webpages Discoverable: A Change-Text Search Interface for Web Archives // 2023 ACM/IEEE Joint Conference on Digital Libraries (JCDL), Santa Fe, NM, USA, 2023. pp. 71–81. DOI: 10.1109/JCDL57899.2023.00021.
- Weigle M. C. The Use of Web Archives in Disinformation Research, 2023. DOI: 10.48550/arXiv.2306.10004.
- Поврозник Н. Г. Веб-архив как источник для изучения современной истории // Исторические исследования в контексте науки о данных: информационные ресурсы, аналитические методы и цифровые технологии. 2020. С. 401–407.
- Brügger N. Archiving Websites: General Considerations and Strategies. Centre for Internet Research. Denmark, 2005. 78 p. URL: [https://cfi.au.dk/fileadmin/www.cfi.au.dk/publikationer/archiving\\_underside/archiving.pdf](https://cfi.au.dk/fileadmin/www.cfi.au.dk/publikationer/archiving_underside/archiving.pdf) (дата обращения: 03.06.2024).
- Masanès J. Web archiving. Berlin; New York: Springer Berlin; New York: Springer, 2006. 234 p.
- Gomes D. Web archives as research infrastructure for digital societies: the case study of Arquivo.pt // Archeion. Apr 02. 2024. URL: <https://ejournals.eu/czasopismo/archeon/artikul/web-archives-as-research-infrastructure-for-digital-societies-the-case-study-of-arquivo-pt> (дата обращения: 03.06.2024).
- Редькина Н. С. Мировые тенденции развития веб-архивов библиотек // Научные и технические библиотеки. 2021. № 1. С. 99–114. DOI: 10.33186/1027-3689-2021-1-99-114.
- Greenwood A. The Past Web: Exploring Web Archives // The American Archivist. 2022. Vol. 85 (2). pp. 717–720. DOI: 10.17723/2327-9702-85.2.717.
- Global Health Events web archive. URL: <https://archive-it.org/collections/4887?fc=websiteGroup%3ACoronavirus+disease+%28COVID-19%29+outbreak> (дата обращения: 03.06.2024).
- Speaker S. L., Moffatt C. The National Library of Medicine Global Health Events web archive, coronavirus disease (COVID-19) pandemic collecting // J. Med. Libr. Assoc. 2020. Vol. 108, № 4. pp. 656–662. DOI: 10.5195/jmla.2020.1090. PMID: 33013228; PMCID: PMC7524615.
- Kelly M. Aggregator Reuse and Extension for Richer Web Archive Interaction // Tseng YH., Katsuraim M., Nguyen H. N. (eds) From Born-Physical to Born-Virtual: Augmenting Intelligence in Digital Libraries. ICADL 2022. Lecture Notes in Computer Science. Springer, Cham. 2022. Vol. 13636. DOI: 10.1007/978-3-031-21756-2\_25.
- Paßmann J., Gerzen L., Schories M., Ogden J., Maemura E., MacKinnon K. Web histories in the making: web archives & the logics of practice // AoIR Selected Papers of Internet Research. 2023. DOI: 0.5210/spir.202310.13535.
- Zoë S. Container Technologies // Hypatia. 2000. Vol. 15, № 2. Pp. 181–201. DOI: 10.1353/hyp.2000.0029.
- Hogan M. The Archiveas Dumpster // Pivot: A Journal of Interdisciplinary Studies and Thought. 2015. Vol. 4, № 1. DOI: 10.25071/2369-7326.39565.
- Ruest N., Fritz S., Milligan I. Creating order from the mess: web archive derivative datasets and notebooks // Archives and Records. 2022. Vol. 43, № 3. pp. 316–331. DOI: 10.1080/23257962.2022.2100336.
- Brügger N. Chapter 23 – Webraries and Web Archives – The Web Between Public and Private // The End of Wisdom? / Editor(s): D. Baker, W. Evans. Chandos Publishing, 2017. pp. 185–190. DOI: 10.1016/B978-0-08-100142-4.00023-3.
- Rogers R. Doing Web History with the Internet Archive: Screencast Documentaries // Internet Histories. 2017. Vol. 1, № 1–2. pp. 160–172. DOI: 10.1080/24701475.2017.1307542.
- Brügger N. Toward a Source Criticism of the Archived Web // The Archived Web: Doing History in the Digital Age. 2018. pp. 137–148. URL: <https://direct.mit.edu/books/book/4215/chapter/177576/Toward-a-Source-Criticism-of-the-Archived-Web> (дата обращения: 03.06.2024).
- Cui C., Pinfield S., Cox A., Hopfgartner F. Participatory Web Archiving: Multifaceted Challenges // Sserwanga I., et al. Information for a Better World: Normality, Virtuality, Physicality, Inclusivity. iConference 2023. Lecture Notes in Computer Science. Springer, Cham. 2023. Vol. 13971. DOI: 10.1007/978-3-031-28035-1\_7.
- Ben-David A. Critical Web Archive Research // Gomes D., Demidova E., Winters J., Risse T. (eds) The Past Web. Springer, Cham. 2021. DOI: 10.1007/978-3-030-63291-5\_14.
- Bingham N. J., Byrne H. Archival strategies for contemporary collecting in a world of big data: Challenges and opportunities with curating the UK web archive // Big Data & Society. 2021. Vol. 8, № 1. DOI: 10.1177/2053951721990409.
- Kiesel J., Kneist F., Alshomary M., Stein B., Hagen M., Potthas, M. Reproducible web corpora: interactive archiving with automatic quality assessment // J. Data Inf. Qual. 2018. Vol. 10, № 4. pp. 17:1–17:25. DOI: 10.1145/3239574.
- Berlin J., Kelly M., Nelson M. L., Michele C. To Re-experience the Web: A Framework for the Transformation and Replay of Archived Web Pages // ACM Trans on the Web. Weigle. 2023. Vol. 17, Iss. 4. pp. 1–49 pages. DOI: 10.1145/3589206.
- Elstner T. et al. Visual Web Archive Quality Assessment // Silvello G., et al. Linking Theory and Practice of Digital Libraries. TPDL 2022. Lecture Notes in Computer Science. Springer, Cham. 2022. Vol. 13541. DOI: 10.1007/978-3-031-16802-4\_34.
- Ryan M., Keating D. Finegan J. Managing and accessing web archives: Irish practitioners' perspectives // AI & Soc. 2022. Vol. 37. pp. 975–984. DOI: 10.1007/s00146-021-01364-0.
- Maemur E. All WARC and no playback: The materialities of data-centered web archives research // Big Data & Society. 2023. Vol. 10, № 1. DOI: 10.1177/20539517231163172.
- Mourão A., Gomes D. Searching images in a web archive // 2023 IEEE 10th International Conference on Data Science and Advanced Analytics (DSAA), Thessaloniki, Greece. 2023. pp. 1–10. DOI: 10.1109/DSAA60987.2023.10302607.
- Bingham N. Exploring Alternative Access: Making the Most of Web Archives During UK Web Archive Downtime, British Library. URL: <https://britishlibrary.typepad.co.uk/webarchive/> (дата обращения: 03.06.2024).

31. Смирнов А. А. Проблемы отечественного и зарубежного веб-архивирования в библиотеках. Веб-архивирование как область деятельности // Научные и технические библиотеки. 2022. № 12. С. 104–123. DOI: 10.33186/1027-3689-2022-12-104-123.
32. Балацкая Н. М., Мартиросова М. Б. Краеведческий веб-архив в структуре

- информационных ресурсов библиотеки: модель и возможности реализации. Санкт-Петербург, 2023. 208 с.
33. Strategic plan of the International Internet Preservation Consortium (IIPC), 2021–2025. URL: [https://netpreserve.org/resources/IIPC-Strategic\\_Plan-2021-2025.pdf](https://netpreserve.org/resources/IIPC-Strategic_Plan-2021-2025.pdf) (дата обращения: 03.06.2024).

**Bibliography:**

1. Sunkara M., Kalari S., Jayarathna S., Ashok V. Assessing the Accessibility of Web Archives // 2023 ACM/IEEE Joint Conference on Digital Libraries (JCDL), Santa Fe, NM, USA, 2023. pp. 253–255. DOI: 10.1109/JCDL57899.2023.00048.
2. Jayanetti H. R. et al. Creating Structure in Web Archives with Collections: Different Concepts from Web Archivists // Silvello G. et al. Linking Theory and Practice of Digital Libraries. TPD L 2022. Lecture Notes in Computer Science. Springer, Cham, 2022. Vol. 13541. DOI: 10.1007/978-3-031-16802-4\_45.
3. Frew L., Nelson M. L., Weigle M. C. Making Changes in Webpages Discoverable: A Change-Text Search Interface for Web Archives // 2023 ACM/IEEE Joint Conference on Digital Libraries (JCDL), Santa Fe, NM, USA, 2023. pp. 71–81. DOI: 10.1109/JCDL57899.2023.00021.
4. Weigle M. C. The Use of Web Archives in Disinformation Research, 2023. DOI: 10.48550/arXiv.2306.10004.
5. Povroznik N. G. Web archive as a source for the study of modern history // Historical research in the context of data science: information resources, analytical methods and digital technologies. 2020. pp. 401–407.
6. Brügger N. Archiving Websites: General Considerations and Strategies. Centre for Internet Research. Denmark, 2005. 78 p. URL: [https://cfi.au.dk/fileadmin/www.cfi.au.dk/publikationer/archiving\\_underside/archiving.pdf](https://cfi.au.dk/fileadmin/www.cfi.au.dk/publikationer/archiving_underside/archiving.pdf) (date of application: 06/03/2024).
7. Masanès J. Web archiving. Berlin; New York: Springer Berlin; New York: Springer, 2006. 234 p.
8. Gomes D. Web archives as research infrastructure for digital societies: the case study of Arquivo.pt // Archeion. Apr 02. 2024. URL: <https://ejournals.eu/czasopismo/archeion/artikul/web-archives-as-research-infrastructure-for-digital-societies-the-case-study-of-arquivo-pt> (date of application: 06/03/2024).
9. Redkina N. S. World trends in the development of library web archives // Scientific and technical libraries. 2021. No. 1. pp. 99–114. DOI: 10.33186/1027-3689-2021-1-99-114.
10. Greenwood A. The Past Web: Exploring Web Archives // The American Archivist. 2022. Vol. 85 (2). pp. 717–720. DOI: 10.17723/2327-9702-85.2.717.
11. Global Health Events web archive. URL: <https://archive-it.org/collections/48877fc=websiteGroup%3ACoronavirus+disease+%2BCOVID-19%29+outbreak> (accessed 03.06.2024).
12. Speaker S. L., Moffatt C. The National Library of Medicine Global Health Events web archive, coronavirus disease (COVID-19) pandemic collecting // J. Med. Libr. Assoc. 2020 Vol. 108, No. 4. pp. 656–662. DOI: 10.5195/jmla.2020.1090. PMID: 33013228; PMCID: PMC7524615.
13. Kelly M. Aggregator Reuse and Extension for Richer Web Archive Interaction // Tseng Y.H., Katsurai M., Nguyen H.N. (eds) From Born-Physical to Born-Virtual: Augmenting Intelligence in Digital Libraries. ICADL 2022. Lecture Notes in Computer Science. Springer, Cham, 2022. Vol. 13636. DOI: 10.1007/978-3-031-21756-2\_25.
14. Paßmann J., Gerzen L., Schories M., Ogdén J., Maemura E., MacKinnon K. Web histories in the making: web archives & the logics of practice // AoIR Selected Papers of Internet Research. 2023. DOI: 0.5210/spir.v2023i0.13535.
15. Zoë S. Container Technologies // Hypatia. 2000. Vol. 15, № 2. Pp. 181–201. DOI: 10.1353/hyp.2000.0029.
16. Hogan M. The Archives Dumpster // Pivot: A Journal of Interdisciplinary Studies and Thought. 2015. Vol. 4, No. 1. DOI: 10.25071/2369-7326.39565.
17. Ruest N., Fritz S., Milligan I. Creating order from the mess: web archive derivative datasets and notebooks // Archives and Records. 2022. Vol. 43, No. 3. pp. 316–331. DOI: 10.1080/23257962.2022.2100336.

18. Brügger N. Chapter 23 – Webraries and Web Archives – The Web Between Public and Private // The End of Wisdom? / Editor(s): D. Baker, W. Evans. Chandos Publishing, 2017. pp. 185–190. DOI: 10.1016/B978-0-08-100142-4.00023-3.
19. Rogers R. Doing Web History with the Internet Archive: Screencast Documentaries // Internet Histories. 2017. Vol. 1, No. 1–2. pp. 160–172. DOI: 10.1080/24701475.2017.1307542.
20. Brügger N. Howard a Source Criticism of the Archived Web // The Archived Web: Doing History in the Digital Age. 2018. pp. 137–148. URL: <https://direct.mit.edu/books/book/4215/chapter/177576/Toward-a-Source-Criticism-of-the-Archived-Web> (accessed: 06/03/2024).
21. Cui C., Pinfield S., Cox A., Hopfgartner F. Participatory Web Archiving: Multifaceted Challenges // Sserwanga I., et al. Information for a Better World: Normality, Virtuality, Physicality, Inclusivity. IConference 2023. Lecture Notes in Computer Science. Springer, Cham, 2023. Vol. 13971. DOI: 10.1007/978-3-031-28035-1\_7.
22. Ben-David A. Critical Web Archive Research // Comes D., Demidova E., Winters J., Risse T. (eds) The Past Web. Springer, Cham, 2021. DOI: 10.1007/978-3-030-63291-5\_14.
23. Bingham N. J., Byrne H. Archival strategies for contemporary collecting in a world of big data: Challenges and opportunities with curating the UK web archive // Big Data & Society. 2021. Vol. 8, № 1. DOI: 10.1177/2053951721990409.
24. Kiesel J., Kneist F., Alshomary M., Stein B., Hagen M., Potthas, M. Reproducible web corpora: interactive archiving with automatic quality assessment // J. Data Inf. Qual. 2018. Vol. 10, No. 4. pp. 171–1725. DOI: 10.1145/3239574.
25. Berlin J., Kelly M., Nelson M. L., Michele C. To Re-experience the Web: A Framework for the Transformation and Replay of Archived Web Pages // ACM Trans on the Web. Weigle. 2023. Vol. 17, Iss. 4. pp. 1–49 pages. DOI: 10.1145/3589206.
26. Elstner T. et al. Visual Web Archive Quality Assessment // Silvello G., et al. Linking Theory and Practice of Digital Libraries. TPD L 2022. Lecture Notes in Computer Science. Springer, Cham, 2022. Vol. 13541. DOI: 10.1007/978-3-031-16802-4\_34.
27. Ryan M., Keating D., Finegan J. Managing and accessing web archives: Irish practitioners' perspectives // AI & Soc. 2022. Vol. 37. pp. 975–984. DOI: 10.1007/s00146-021-01364-0.
28. Maemur E. All WARC and no playback: The materialities of data-centered web archives research // Big Data & Society. 2023. Vol. 10, № 1. DOI: 10.1177/20539517231163172.
29. Mourão A., Gomes D. Searching images in a web archive // 2023 IEEE 10th International Conference on Data Science and Advanced Analytics (DSAA), Thessaloniki, Greece. 2023. pp. 1–10. DOI: 10.1109/DSAA60987.2023.10302607.
30. Bingham N. Exploring Alternative Access: Making the Most of Web Archives During UK Web Archive Downtime, British Library. URL: <https://britishlibrary.typepad.co.uk/webarchive/> (date of access: 06/03/2024).
31. Смирнов А. А. Problems of domestic and foreign web archiving in libraries. Web archiving as a field of activity // Scientific and technical libraries. 2022. No. 12. pp. 104–123. DOI: 10.33186/1027-3689-2022-12-104-123.
32. Balatskaya N. M., Martirosova M. B. Local history web archive in the structure of library information resources: model and implementation possibilities. St. Petersburg, 2023. 208 p.
33. Strategic plan of the International Internet Preservation Consortium (IIPC), 2021–2025. URL: [https://netpreserve.org/resources/IIPC-Strategic\\_Plan-2021-2025.pdf](https://netpreserve.org/resources/IIPC-Strategic_Plan-2021-2025.pdf) (date of application: 06/03/2024).



DOI 10.52815/0204-3653\_2024\_2197\_67  
EDN: ATMVDF

УДК 004.8:69.059

# ПРИМЕНЕНИЕ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ ОБНАРУЖЕНИЯ ПОВРЕЖДЕНИЙ НА ФАСАДЕ ЗДАНИЙ

*Аннотация. Вопросам цифровой трансформации строительной отрасли посвящено значительное количество публикаций, как в России, так и за рубежом. При этом основное внимание уделяется оценке перспектив внедрения технологий Индустрии 4.0, к которым относятся информационное моделирование, большие данные, искусственный интеллект, машинное обучение, аддитивное строительное производство и т. д. При этом в соответствии с проведенным анализом научных публикаций искусственный интеллект в значительной мере применяется в нефтегазовой отрасли, меньше в строительной. В статье представлен подход, который объединяет в себе преимущества глубокого обучения и YOLO для более эффективного и точного обнаружения повреждений на фасадах. Глубокое обучение, в частности, сверточные нейронные сети (CNN), позволяют извлекать сложные признаки из изображений, что особенно важно при выявлении незаметных дефектов. Технология YOLO, в свою очередь, обеспечивает быстрое и точное обнаружение объектов, что значительно повышает эффективность процесса. Авторами предложен алгоритм разработанного программного продукта. Рассмотрен каждый из этапов алгоритма более подробно, представлена программная реализация. Предлагаемый подход для выявления повреждений на фасадах зданий может быть важным в задачах мониторинга и обследования зданий в режиме реального времени.*

**Пиляй Андрей**  
Старший преподаватель  
кафедры ИСТАС, НИУ МГСУ  
E-mail: PilyayAI@mgsu.ru

**Гинзбург Евгения**  
Студент кафедры ИСТАС,  
НИУ МГСУ  
E-mail: ginea@ginzburgi.ru

Источник: depositphotos.com

**Ключевые слова:** глубокое обучение, YOLO, цифровая трансформация, строительство, мониторинг фасадов зданий.

## Введение

В настоящее время в строительной отрасли происходит активное внедрение технологий Индустрии 4.0 для обеспечения ее цифровой трансформации, чему посвящено значительное количество публикаций [1–6].

Напомним, что к технологиям Индустрии 4.0 относятся машинное обучение, искусственный интеллект, большие данные, технологии информационного моделирования, аддитивное строительное производство и пр. [5]. В настоящее время наиболее цитируемыми в РИНЦ являются публикации о применении ИИ в нефтегазовой отрасли, например [7–8].

В представленном исследовании рассматривается возможность применения глубокого обучения для обнаружения поврежденных на фасаде зданий, поскольку это является актуальной задачей, имеющей важное значение для обеспечения безопасности и проведения своевременных ремонтных работ.

В современных условиях возможности глубокого обучения и технологии обнаружения объектов YOLO предоставляют новые перспективы в области разработки эффективных методов для выявления повреждений на фасадах зданий.

YOLO – это нейронная сеть, которая за один проход прогнозирует для изображения положение ограничивающих прямоугольников и вероятности классификации. Модели YOLO способны обрабатывать более 60 кадров в секунду [10], т. е. такая архитектура подходит также для распознавания объектов в видео.

В данной статье представляется подход, который объединяет в себе преимущества глубокого обучения и YOLO для более эффективного и точного обнаружения поврежденных на фасадах.

Глубокое обучение, в частности, сверточные нейронные сети (CNN), позволяют извлекать сложные признаки из изображений, что особенно важно при выявлении незаметных дефектов. Технология YOLO, в свою очередь, обеспечивает быстрое и точное обнаружение объектов, что значительно повышает

эффективность процесса. На рис. 1 на диаграмме показаны основные этапы алгоритма программного продукта, начиная от генерации данных до демонстрации результата. Стрелки указывают на последовательность действий, а каждый блок содержит краткое описание ключевых элементов этапа. Рассмотрим каждый из этапов с позиции его реализации.

На первом этапе, «Генерация данных», происходит инициализация генератора случайных чисел с фиксированным начальным числом (seed), что обеспечивает воспроизводимость результатов при повторных запусках. Здесь же генерируются случайные изображения, которые классифицируются как изображения поврежденных или неповрежденных объектов, и создаются массивы данных X и y, где X содержит изображения, а y – соответствующие метки.

На втором этапе, «Подготовка данных», данные разделяются на обучающий и тестовый наборы с помощью функции train test split. Это разделение позволяет отделить часть данных для обучения модели от части для оценки ее производительности.

Третий этап, «Создание и обучение модели CNN», включает создание и компиляцию модели сверточной нейронной сети (CNN), которая затем обучается на обучающем наборе данных. В процессе обучения используются оптимизатор, функция потерь и метрики.

Четвертый этап, «Загрузка весов YOLOv3», включает загрузку предварительно обученных весов модели YOLOv3 с помощью функции cv2.dnn.readNet, что необходимо для последующего детектирования объектов на изображениях.

На пятом этапе, «Применение YOLO на тестовом изображении», модель YOLOv3 применяется для детектирования объектов на тестовом изображении с помощью функции apply\_yolo.

Шестой этап, «Оценка производительности модели CNN», включает оценку производительности модели на тестовых данных. Используется функция model.evaluate для расчета точности и других метрик производительности.

На последнем этапе, «Демонстрация результата», результаты работы YOLO и классификационной модели CNN демонстрируются на тестовом изображении, что включает визуализацию распознанных объектов и меток классификации для анализа эффективности и точности модели.

## Подготовка данных

Для проведения обучения модели используется набор данных, содержащий изображения как поврежденных, так и неповрежденных фасадов зданий. Также, производится предобработка изображений, включая изменение размера и нормализацию значений пикселей. Ниже представлена реализация данного шага на языке программирования python:

```
import os
import cv2
import numpy as np
from sklearn.model_selection import train_test_split
```

```
# Путь к папкам с изображениями
data_dir = «/путь/к/данным»
damaged_dir = os.path.join(data_dir, "damaged")
undamaged_dir = os.path.join(data_dir, "undamaged")
```

```
# Загрузка изображения из папок damaged и undamaged
damaged_images = [cv2.imread(os.path.join(damaged_dir, img)) for img in os.listdir(damaged_dir)]
```

Рис. 1. Алгоритм программного продукта





Исторический фасад городского дома  
Источник: ginasanders / depositphotos.com

```
undamaged_images = [cv2.imread(os.path.
join(undamaged_dir, img)) for img in os.listdir(undamaged_dir)]
```

```
# Предобработка данных
def preprocess_images(images):
# Изменение размера изображений
до 224x224 и нормализация значений пикселей
return np.array([cv2.resize(img, (224, 224)) /
255.0 for img in images])
```

```
# Создание массивов данных и меток
X = preprocess_images(damaged_images +
undamaged_images)
y = np.array(["damaged"] * len(damaged_images) +
["undamaged"] * len(undamaged_images))
```

```
# Разделение на тренировочный и тестовый наборы
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Из кода видно, что изображения загружаются из папок damaged и undamaged, при этом заложена логика о том, что есть две подпапки внутри data\_dir.

Далее применяется функция предобработки preprocess\_images, которая изменяет размер изображений до 224x224 пикселей и нормализует значения пикселей к диапазону [0, 1]. Результаты сохраняются в массивах данных X и метках y.

```
# Пример изображения после предобработки
sample_image = X_train[0]
cv2.imshow("Sample Image", cv2.cvtColor((sample_image * 255).astype(np.uint8), cv2.COLOR_BGR2RGB))
cv2.waitKey(0)
cv2.destroyAllWindows()
```

Для классификации повреждений используется сверточная нейронная сеть (CNN) на основе модели VGG16.

### Обучение модели глубокого обучения

Для классификации повреждений используется сверточная нейронная сеть (CNN) на основе модели VGG16. Обучение модели производится на тренировочных данных с использованием функции потерь бинарной кросс-энтропии. Приведенный ниже код демонстрирует этот процесс.

```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense
# Создание модели CNN на основе VGG16
model = Sequential()
model.add(Conv2D(64, (3, 3), activation='relu', input_shape=(224, 224, 3)))
model.add(Conv2D(64, (3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Conv2D(128, (3, 3), activation='relu'))
model.add(Conv2D(128, (3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Flatten())
```

```
model.add(Dense(256, activation='relu'))
model.add(Dense(1, activation='sigmoid'))
# Компиляция модели с использованием бинарной кросс-энтропии
model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
# Обучение модели на тренировочных данных
model.fit(X_train, (y_train == 'damaged').astype(int), epochs=10, batch_size=32, validation_split=0.2)
```

Этот код создает модель CNN, включающую слой свертки, пулинга и полносвязные слои, основанные на архитектуре VGG16. Модель компилируется с использованием оптимизатора Adam и функции потерь бинарной кросс-энтропии. После этого производится обучение модели на тренировочных данных.

Важно отметить, что X\_train и y\_train представляют собой тренировочные изображения и соответствующие метки классов («damaged» или «undamaged»).

### Использование YOLO для детекции объектов

Для эффективного обнаружения объектов используется технология YOLO (You Only Look Once). Процесс внедрения YOLO и выделения объектов с высокой уверенностью описан в следующем коде:

```
import cv2
# Путь к файлам весов и конфигурации YOLO
yolo_weights = "/путь/к/yolov3.weights"
yolo_config = "/путь/к/yolov3.cfg"
yolo_classes = "/путь/к/coco.names"
# Загружаем YOLO
net = cv2.dnn.readNet(yolo_weights, yolo_config)
with open(yolo_classes, "r") as f:
classes = [line.strip() for line in f]
# Функция для применения YOLO к изображению
def apply_yolo(image):
```

```
# Получаем высоту и ширину изображения
height, width = image.shape[:2]
# Подготовка входных данных для YOLO
blob = cv2.dnn.blobFromImage(image, 1/255.0, (416, 416), swapRB=True, crop=False)
# Установка входа для сети YOLO
net.setInput(blob)
# Получение и анализ предсказаний
detections = net.forward()
# Проход по всем предсказаниям
for detection in detections:
for obj in detection:
scores = obj[5:]
class_id = np.argmax(scores)
confidence = scores[class_id]
# Проверка уверенности в предсказании
if confidence > 0.7:
# Определение координат объекта
center_x = int(obj[0] * width)
center_y = int(obj[1] * height)
w = int(obj[2] * width)
h = int(obj[3] * height)
```

Для эффективного обнаружения объектов используется технология YOLO  
Источник: luckybusiness / depositphotos.com



```

# Вычисление координат углов прямоугольника
x = int(center_x - w / 2)
y = int(center_y - h / 2)
# Отрисовка прямоугольника и метки класса
cv2.rectangle(image, (x, y), (x + w, y + h), (0, 255, 0), 2)
cv2.putText(image, classes[class_id], (x, y - 5), cv2.FONT_HERSHEY_SIMPLEX, 0.5, (0, 255, 0), 2)
return image
# Пример применения YOLO к изображению
test_image = X_test[0] # Предполагаем, что X_test содержит тестовые изображения
result_image = apply_yolo(test_image)
# Отображение изображения с выделенными объектами
cv2.imshow("YOLO Detection", cv2.cvtColor((result_image * 255).astype(np.uint8), cv2.COLOR_BGR2RGB))
cv2.waitKey(0)
cv2.destroyAllWindows()

```

В представленном коде используется предварительно обученная модель YOLO, загружаются вес, конфигурация и классы

объектов. Затем, для каждого изображения из тестового набора, применяется YOLO для обнаружения объектов с уверенностью более 0.7, которые затем выделяются на изображении.

### Визуализация результатов

После проведения обучения и тестирования модели на соответствующих наборах данных получаются результаты, которые подтверждают эффективность предложенного подхода в обнаружении повреждений на фасадах зданий.

На рис. 2 представлен журнал с отображением итога процесса обучения модели нейронной сети по эпохам, показаны результаты для каждой из 10 эпох обучения. Для каждой эпохи отображаются следующие метрики:

- **loss**: значение функции потерь на тренировочных данных;
- **accuracy**: точность модели на тренировочных данных;
- **val\_loss**: значение функции потерь на валидационных (тестовых) данных;

Рис. 2. Итог обучения модели

```

Epoch 1/10
2/2 [=====] - 35s 9s/step - loss: 3776.1143 - accuracy: 0.5312 - val_loss: 1834.5919 - val_accuracy: 0.3750
Epoch 2/10
2/2 [=====] - 15s 8s/step - loss: 976.8628 - accuracy: 0.5000 - val_loss: 94.1540 - val_accuracy: 0.6250
Epoch 3/10
2/2 [=====] - 15s 8s/step - loss: 116.2165 - accuracy: 0.5312 - val_loss: 58.9202 - val_accuracy: 0.3750
Epoch 4/10
2/2 [=====] - 15s 8s/step - loss: 65.1153 - accuracy: 0.4688 - val_loss: 37.0259 - val_accuracy: 0.6250
Epoch 5/10
2/2 [=====] - 15s 8s/step - loss: 32.4836 - accuracy: 0.5000 - val_loss: 12.9377 - val_accuracy: 0.3750
Epoch 6/10
2/2 [=====] - 15s 8s/step - loss: 7.5016 - accuracy: 0.5000 - val_loss: 0.6885 - val_accuracy: 0.4375
Epoch 7/10
2/2 [=====] - 15s 8s/step - loss: 0.8589 - accuracy: 0.6094 - val_loss: 0.7283 - val_accuracy: 0.3750
Epoch 8/10
2/2 [=====] - 15s 8s/step - loss: 0.6743 - accuracy: 0.5781 - val_loss: 0.6939 - val_accuracy: 0.3750
Epoch 9/10
2/2 [=====] - 15s 8s/step - loss: 0.6873 - accuracy: 0.8125 - val_loss: 0.7108 - val_accuracy: 0.3750
Epoch 10/10
2/2 [=====] - 15s 8s/step - loss: 0.5906 - accuracy: 0.9688 - val_loss: 3.4915 - val_accuracy: 0.6250
1/1 - 1s - loss: 5.6973 - accuracy: 0.4000 - 1s/epoch - 1s/step
Точность модели на тестовых данных: 40.00%

```

- **val\_accuracy**: точность модели на валидационных данных.

Согласно данным, первая эпоха начинается с высокого значения потерь (3776.1143) и точности 53,12%, но с очень высоким значением потерь на валидационных данных (1834.5919) и валидационной точностью 37,5%. В последующих эпохах видно, что значения потерь на обучающих данных и на валидационных данных постепенно уменьшаются, что свидетельствует о том, что модель обучается и адаптируется к данным. Однако колебания точности на валидационных данных указывают на потенциальную нестабильность или переобучение модели.

К последней, 10-й эпохе, точность на тренировочных данных достигает 96,68%, а потери уменьшаются до 5.6973, что свидетельствует о хорошей адаптации модели к тренировочному набору данных.

### Заключение

Несмотря на значительные успехи в области применения технологий Индустрии 4.0 для осуществления цифровой трансформации строительной отрасли, ряд технологий используется не на полный потенциал. Одной из таких технологий является искусственный интеллект.

В настоящей работе нами представлен подход к обнаружению повреждений на фасадах зданий и сооружений, основанный на комплексном применении глубокого обучения с использованием сверточных нейронных сетей и технологии YOLO. Проведенные в рамках исследования эксперименты позволяют сделать вывод о том, что предложенный подход является эффективным.

В рамках следующих шагов авторами планируется расширить исследование и провести дополнительные эксперименты на различных наборах данных для доработки архитектуры модели.

Предложенный подход имеет потенциал стать важным инструментом для инженеров, занимающихся обследованием зданий, и специалистов по мониторингу состояния сооружений.



## USING DEEP LEARNING TO DETECT DAMAGES ON BUILDING FACADES

**Pilyay Andrey**, National Research Moscow State University of Civil Engineering, Senior Lecturer, Department of ISTAS, Russia, Moscow. E-mail: PilyayAI@mgisu.ru

**Ginzburg Evgeniya**, National Research Moscow State University of Civil Engineering, Student, Department of ISTAS, Russia, Moscow. E-mail: ginea@ginzburg.ru

**Abstract.** A significant amount of publications both in Russia and abroad is dedicated to the issues of digital transformation in the construction industry. The main focus is on assessing the prospects of implementing Industry 4.0 technologies, which include information modeling, big data, artificial intelligence, machine learning, additive manufacturing, etc. According to the analysis of scientific publications, artificial intelligence is extensively used in the oil and gas industry, and less so in construction. This article presents an approach that combines the advantages of deep learning and YOLO for more effective and accurate detection of facade damages. Deep learning, specifically convolutional neural networks (CNN), enables the extraction of complex features from images, which is particularly important in identifying subtle defects. The YOLO technology, in turn, provides fast and accurate object detection, significantly enhancing the process efficiency. The authors propose an algorithm of the developed software product. Each stage of the algorithm is examined in more detail, and its software implementation is presented. The proposed approach for detecting damages on building facades can be important in tasks related to real-time building monitoring and inspection.

**Keywords:** Deep Learning, YOLO, Digital Transformation, Construction, Building Facade Monitoring.

## Библиографический список

1. Адамцевич Л. А., Сорокин И. В., Настычук А. В. Перспективные в условиях цифровой трансформации строительной отрасли технологии Индустрии 4.0 // Строительство и архитектура. 2022. Т. 10. № 4. С. 101–105.
2. Адамцевич Л. А., Харисов И. З., Камаева Ю. В. Международный опыт применения технологий Индустрии 4.0 для мониторинга актуального состояния строительного производства // Строительное производство. 2022. № 3. С. 58–66.
3. Адамцевич Л. А., Харисов И. З. Обзор технологий Индустрии 4.0 для разработки системы дистанционного управления строительной площадкой // Строительство и архитектура. 2021. Т. 9. № 4. С. 91–95.
4. Гинзбург А. В., Адамцевич Л. А., Адамцевич А. О. Строительная отрасль и концепция «Индустрия 4.0»: обзор // Вестник МГСУ. 2021. Т. 16. № 7. С. 885–911.
5. Адамцевич Л. А., Гинзбург Е. А., Шилов Л. А. Строительство 4.0 // Жилищное строительство. 2023. № 11. С. 18–23.
6. Камаева Ю. В., Адамцевич Л. А. Перспективы использования предиктивной аналитики в строительстве // Строительство и архитектура. 2023. Т. 11. № 2. С. 12.
7. Черников А. Д., Еремин Н. А., Столяров В. Е., Сбоев А. Г., Семенова-Чашина О. К., Фицнер Л. К. Применение методов искусственного интеллекта для выявления и прогнозирования осложнений при строительстве нефтяных и газовых скважин: проблемы и основные направления решения // Георесурсы. 2020. Т. 22. № 3. С. 87–96.
8. Еремин Н. А., Черников А. Д., Сарданашвили О. Н., Столяров В. Е., Архипов А. И. Цифровые технологии строительства скважин. Создание высокопроизводительной автоматизированной системы предотвращения осложнений и аварийных ситуаций в процессе строительства нефтяных и газовых скважин // Neftegaz.RU. 2020. № 4 (100). С. 38–50.
9. Дмитриевский А. Н., Еремин Н. А., Ложников П. С., Клиновенко С. А., Столяров В. Е., Иниватов Д. П. Анализ рисков при использовании технологий искусственного интеллекта в нефтегазодобывающем комплексе // Автоматизация, телемеханизация и связь в нефтяной промышленности. 2021. № 7 (576). С. 17–27.
10. Кирнос В. П., Коковкина В. А. Разработка системы одновременной локализации и картографии на базе нейронных сетей // Цифровая обработка сигналов и ее применение. DSPA-2023.

## Bibliography:

1. Adamtsevich, L.A., Sorokin, I.V., Nastyuchuk, A.V. «Prospective Technologies of Industry 4.0 in the Context of Digital Transformation of the Construction Sector.» Construction and Architecture. 2022, Vol. 10, No. 4. pp. 101–105.
2. Adamtsevich, L.A., Kharisov, I.Z., Kamaeva, Y.V. «International Experience in Using Industry 4.0 Technologies for Monitoring the Current State of Construction Production.» Construction Production. 2022, No. 3. pp. 58–66.
3. Adamtsevich, L.A., Kharisov, I.Z. «Overview of Industry 4.0 Technologies for Developing a Remote Control System for Construction Sites.» Construction and Architecture. 2021, Vol. 9, No. 4. pp. 91–95.
4. Ginzburg, A.V., Adamtsevich, L.A., Adamtsevich, A.O. «Construction Sector and the 'Industry 4.0' Concept: A Review.» Vestnik MGSU. 2021, Vol. 16, No. 7. pp. 885–911.
5. Adamtsevich, L.A., Ginzburg, E.A., Shilov, L.A. «Construction 4.0.» Housing Construction. 2023, No. 11. pp. 18–23.
6. Kamaeva, Y.V., Adamtsevich, L.A. «Prospects for the Use of Predictive Analytics in Construction.» Construction and Architecture. 2023, Vol. 11, No. 2. p. 12.
7. Chernikov, A.D., Eremin, N.A., Stolyarov, V.E., Sboev, A.G., Semenova-Chashchina, O.K., Fitzner, L.K. «Application of Artificial Intelligence Methods for Identification and Prediction of Complications During the Construction of Oil and Gas Wells: Problems and Directions for Solutions.» Georesources. 2020, Vol. 22, No. 3. pp. 87–96.
8. Eremin, N.A., Chernikov, A.D., Sardanashevili, O.N., Stolyarov, V.E., Arkhipov, A.I. «Digital Technologies in Well Construction: Creating a High-Performance Automated System for Preventing Complications and Emergency Situations in the Construction of Oil and Gas Wells.» Business Journal Neftegaz.RU. 2020, No. 4 (100). pp. 38–50.
9. Dmitrievsky, A.N., Eremin, N.A., Lozhnikov, P.S., Klinovenko, S.A., Stolyarov, V.E., Inivatov, D.P. «Risk Analysis in the Use of Artificial Intelligence Technologies in the Oil and Gas Industry.» Automation, Telemetry and Communication in the Oil Industry. 2021, No. 7 (576). pp. 17–27.
10. Kirnos, V.P., Kokovkina, V.A. «Development of a System for Simultaneous Localization and Mapping Based on Neural Networks.» Digital Signal Processing and Its Application DSPA – 2023. Proceedings of the 25th International Conference. Moscow, 2023. pp. 363–366.



УДК 378.6

DOI 10.52815/0204-3653\_2024\_2197\_75  
EDN: TEWURS

## РАЗРАБОТКА ПРОГРАММНОГО ПРОДУКТА ДЛЯ СТАНДАРТИЗАЦИИ ДОКУМЕНТОВ В ЭЛЕКТРОННОМ ВИДЕ

**Муконина Мария**  
Старший преподаватель,  
кафедра вычислительной  
техники и автоматизированных  
систем управления, Ростовский  
государственный университет  
путей сообщения (РГУПС)  
E-mail: s.olimp2017@yandex.ru

**Глазунов Дмитрий**  
Доцент, к. т. н., кафедра  
автоматики и телемеханики  
на железнодорожном  
транспорте, Ростовский  
государственный университет  
путей сообщения (РГУПС)  
E-mail: glazunovdm@yandex.ru

**Гальцева Алена**  
Ассистент, кафедра  
вычислительной техники  
и автоматизированных систем  
управления, Ростовский  
государственный университет  
путей сообщения (РГУПС)  
E-mail: alena\_shutenko@mail.ru

*Аннотация. Авторами работы рассмотрена проблема эффективности использования информационного обеспечения при стандартизации документов в электронном виде. Сформирован обобщенный процесс работы программного продукта Standconverter. Разработана модель расширенной диаграммы прецедентов. Разобран сценарий процесса стандартизации документов. Выделены классы, необходимые для функционирования системы. Создана диаграмма процесса «Стандартизация электронного документа». Представлен фрагмент программного кода приложения Standconverter для стандартизации документов в электронном виде.*

## Ключевые слова:

модель, стандартизация, информационное обеспечение, программный продукт, электронный документ, диаграмма прецедентов.



Папки с бумажными документами  
Источник: racorn / depositphotos.com

**Введение**

На сегодняшний день невозможно представить деятельность любого предприятия или организации, а также органов государственной власти без ведения соответствующей документации. Это объясняется тем, что составление, использование и хранение документов в соответствии с действующими законами и нормативными актами обеспечивает защиту интересов компании, а также увеличивает эффективность управленческого аппарата [1–4].

В связи с этим актуальным является вопрос стандартизации и унификации документооборота. Опыт унификации насчитывает не одно столетие. Стандартизация документов подразумевает проведение ре-

организации документооборота какой-либо системы в соответствии с определенными стандартами. Она позволяет минимизировать переписки, сократить трудоемкость ее составления, а также предотвратить дублирование данных, за счет чего достигаются значительные успехи в проведении управленческой работы. Это, в свою очередь, способствует развитию и продвижению всей организации в целом [5–9].

Основной целью автоматизации любого процесса является эффективность использования информационных инструментов, а также скорость решения задачи и обеспечение простоты и удобства работы пользователя. Это в полной мере относится и к задаче «Разработки программного продукта Standconverter для стандартизации документов в электронном виде» [10–16].

Для ее решения произведем анализ и построим функциональную модель процесса обработки документации с помощью методологии SADT, а также воспользуемся стандартизованным языком моделирования UML для визуализации, построения и документирования артефактов программной системы Standconverter.

**Разработка UML-диаграмм для проектируемой системы**

В первую очередь необходимо составить диаграмму Use case верхнего уровня, в которой будет отражено общее представление работы системы. Диаграмма прецедентов

Рис. 1. Диаграмма Use case верхнего уровня для программного продукта Standconverter



является основополагающей для первичного моделирования динамики системы, используется для выяснения требований к разрабатываемой системе, фиксации этих требований в форме, которая позволяет визуализировать дальнейший процесс разработки.

Обобщенный процесс работы программного продукта Standconverter для стандартизации документов в электронном виде представлен на рис. 1. Основная функция приложения – это редактирование документа, которая в свою очередь разлагается на определенное количество подзадач. Их необходимо выполнить, чтобы достичь целевого эффекта. Декомпозиция системы представлена на расширенной диаграмме прецедентов.

Кроме того, администратор осуществляет прямую связь с клиентом через диалоговую форму, где пользователь может задать любой интересующий вопрос непосредственно с помощью телефонной связи, доступной по активной ссылке на веб-сайте.

После того как пользователь загрузил необходимый документ, система начинает свою работу, включающую в себя выбор типа форматирования:

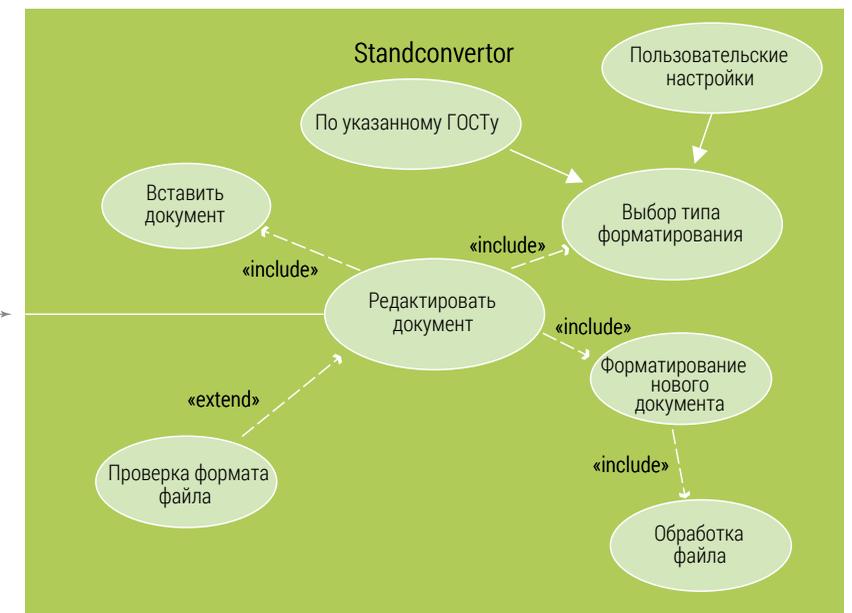
- согласно уже заданным параметрам и ГОСТам;
- по параметрам, заданным пользователем.

Кроме того, для редактирования документа необходимо выполнить проверку его формата, так как система имеет ограниченный набор форматов, применяемых к форматированию, в которых система будет работать корректно. Standconverter в процессе своей работы будет производить форматирование исходного документа посредством обработки данных, предоставленных пользователем.

На рис. 2 изображена модель расширенной диаграммы прецедентов системы, которая позволяет визуализировать функционал, необходимый для работы с документами.

Для уточнения первоначальных требований к работоспособности системы разбе-

Рис. 2. Модель расширенной диаграммы прецедентов





рем сценарий процесса стандартизации документов в электронном виде.

Основной успешный сценарий:

- 1) пользователь входит в систему и попадает на главную страницу, где он может ознакомиться с функциональными особенностями работы системы и перейти непосредственно к обработке исходного документа;
- 2) выбрав в боковом меню раздел «Редактирование», пользователь перемещается на страницу, содержащую поле для загрузки документа и необходимую инструкцию по загрузке документа;
- 3) выбрав нужный документ, пользователь нажимает кнопку «Загрузить»;
- 4) после загрузки система начинает проверку формата документа, отправленного пользователем;
- 5) если формат совпадает с форматом, указанным в списке редактируемых документов, то система сообщает пользователю, что загрузка была завершена успешно и переключается на экран выбора параметров редактирования;
- 6) пользователь выбирает необходимые параметры редактирования:
  - если его интересует уже готовый набор параметров, то он щелкает на уже введенном ГОСТе;
  - если пользователь хочет отредактировать только некоторые параметры, то он выбирает необходимые значения из указанного списка;
- 7) после этого пользователь нажимает кнопку «Редактировать» и попадает на страницу подтверждения выбора, где ему предлагается подтвердить правильность перечня выбранных параметров;
- 8) после подтверждения система приступает к обработке файла согласно заданным настройкам;
- 9) когда файл полностью обработан, система оповещает пользователя о том, что обработка была совершена успешно и предлагает пользователю скачать готовый файл;
- 10) пользователь нажимает на кнопку «Загрузить», после чего ему предлагается

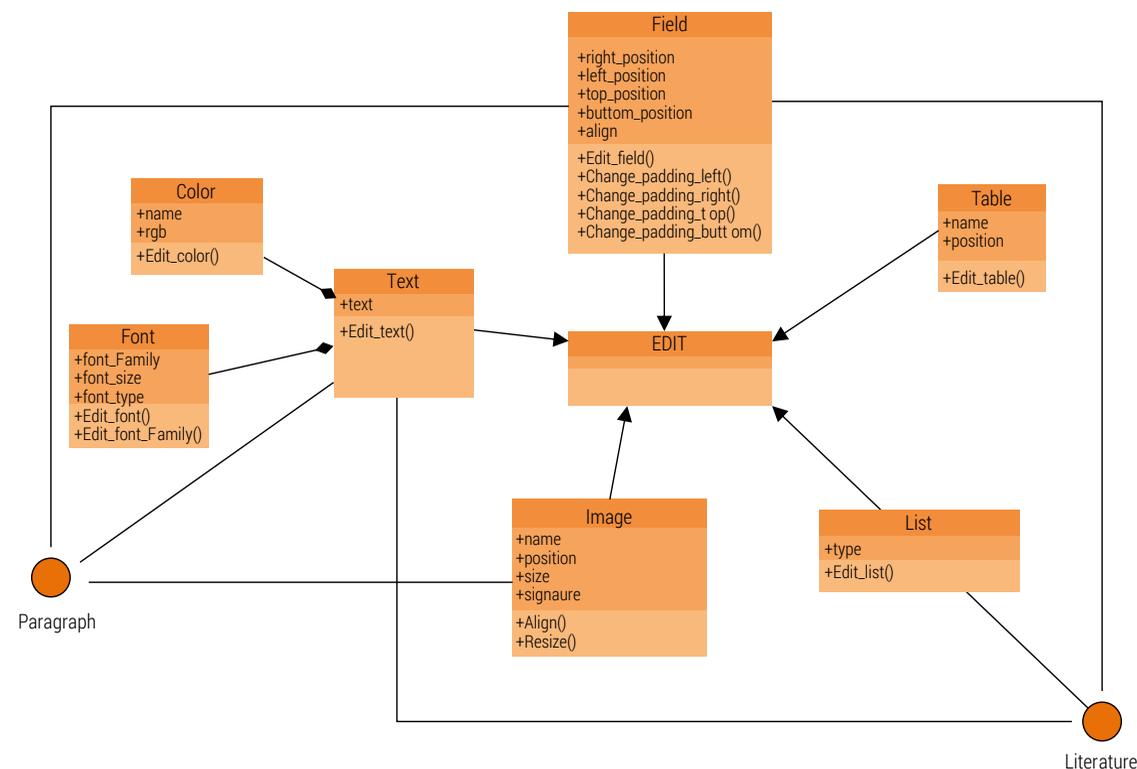


Рис. 3. Диаграмма классов

выбрать путь для сохранения нового файла;

- 11) когда файл загружен по указанному пользователем пути, система выдает сообщение о том, что обработка прошла успешно и предлагает пользователю продолжить работу с новым файлом, либо завершить работу;
- 12) процесс стандартизации электронного документа закончен.

Расширения (или альтернативные потоки) подразумевают под собой следующее: если пользователь не ознакомился с инструкцией и попытка загрузить файл, не подходящий по формату для корректной работы системы, то ПО выдает сообщение об ошибке и просит повторить попытку загрузки файла с необходимым расширением. Диаграмма классов явля-

ется представителем статических моделей и использует статические словари в качестве основы для проектирования. Данный вид диаграммы является связующим звеном для построения диаграмм компонентов и развертывания. Классовые диаграммы позволяют не только отразить статический процесс, но предоставляют возможность создания каркаса программного кода для проектируемой системы.

Цель диаграммы классов можно суммировать как анализ и проектирование статического представления приложения, описание обязанностей системы и обратная инженерия. Диаграмма классов – это статическое представление системы с точки зрения описания основополагающих конструкций и приложений. Диаграмма классов имеет ряд особенностей постро-

ения, которые необходимо учитывать при ее создании:

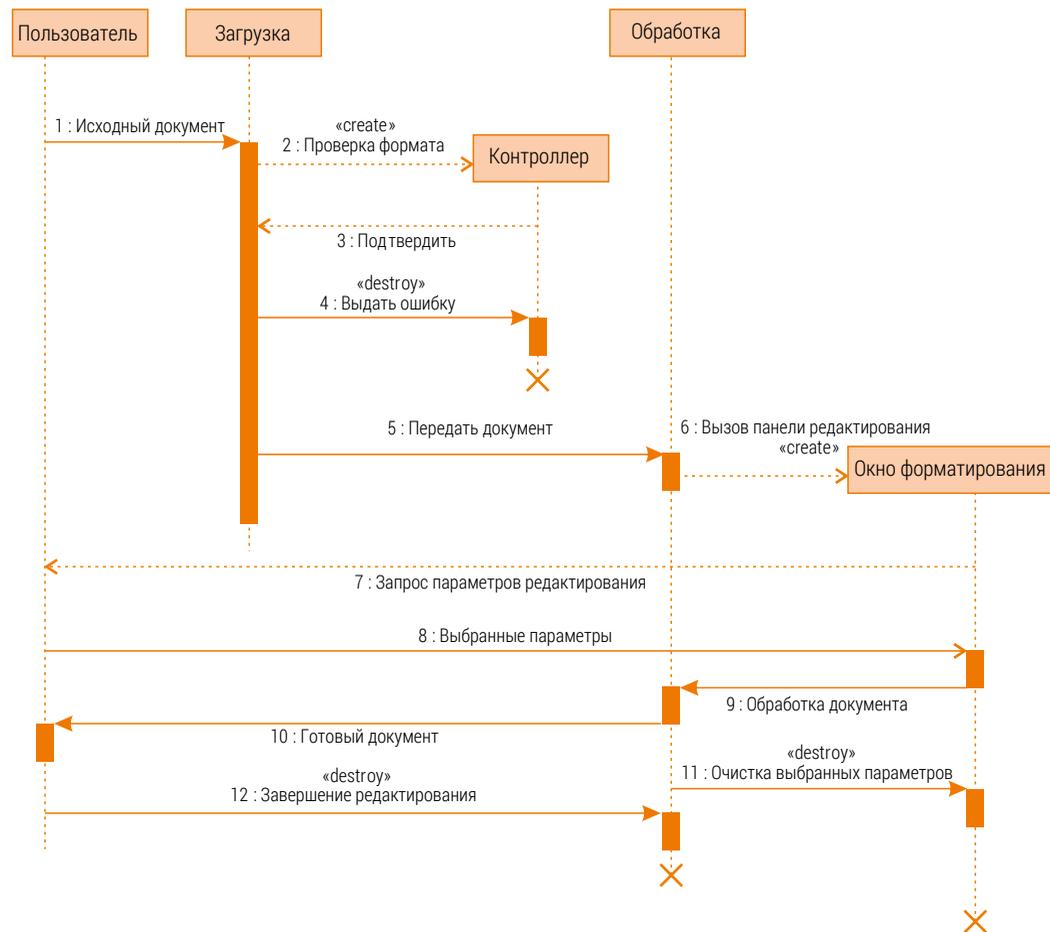
- имя диаграммы классов должно иметь смысл, описывать аспект системы;
- каждый элемент и их отношения должны быть идентифицированы заранее;
- атрибуты и методы каждого класса должны быть четко определены;
- для каждого класса необходимо указать минимальное количество свойств, по-

скольку ненужные свойства усложнят понимание диаграммы.

Чтобы исключить появление ошибок и перегрузки диаграммы классов лишними компонентами, необходимо провести предварительное натурное проектирование на бумаге. Модель предметной области является одной из основополагающих при объектно-ориентированном подходе [12], т. к. именно на ней отображаются кон-

Рис. 4. Диаграмма процесса «Стандартизация электронного документа»

Interaction SequenceDiagram1



Название класса	Описание
Edit	«Редактор» – абстрактный класс, объединяющий все параметры редактирования
Text	«Текст», который в свою очередь является родительским для классов «Цвет» и «Шрифт» и включает в себя обобщённый метод редактирования текста
Color	«Цвет» содержит данные о названии цвета текста и его характеристиках в цветовой системе RGB
Font	«Шрифт», содержащий данные о семействе используемого в тексте шрифта, его размере и типе (курсивный, жирный, подчеркнутый)
Image	Класс, отвечающий за редактирование картинок с информацией о размере изображения, выравнивании и сопровождаемой подписи
List	Класс, созданный для обработки списков
Field	Информация об абзачных отступах в документе (отступ сверху, справа, слева, сверху с возможностью их редактирования, а также информация о выравнивании)
Table	«Таблица» – отвечает за редактирование табличных данных
Paragraph	Интерфейсный класс, позволяющий провести комплексное редактирование содержания абзаца, включающего в себя текст, картинки и абзачные отступы
Literature	Интерфейсный класс, созданный для обработки списка литературы

Таблица 1. Описание классов предметной области

цептуальные классы, необходимые для построения системы. Основные цели, которые достигаются с помощью данной модели, – это описание бизнес-процессов для предметной области, бизнес-правил и действующих лиц. При рассмотрении предметной области можно выделить следующие классы, необходимые для функционирования системы (рис. 3): «Редактор» – абстрактный класс, который объединяет под собой остальные классы редактирования частей документа: «Текст», «Картинка», «Поле», «Таблица» и «Список».

В свою очередь при редактировании текста мы затрагиваем такие составляющие как «Цвет» и «Шрифт». Для комплексного редактирования нескольких параметров будут созданы интерфейсы «Параграф» и «Литература».

Описание атрибутов представленных классов предметной области отражено в таблице 1. Опирируя описанными классами и их экземплярами, можно моделировать различные состояния системы в диапазоне, достаточном для данной работы.

Диаграмма взаимодействия позволяет производить моделирование процесса об-

мена сообщениями между объектами. Эти диаграммы включают в себя два подвида диаграмм – последовательности и кооперации.

Диаграмма взаимодействия для процесса стандартизации электронного документа представлена на рис. 4.

Для обработки документа пользователю необходимо загрузить исходный документ в систему, после чего начинается процесс загрузки, во время которого система отправляет контроллеру данные о формате файла. Контроллер, в свою очередь, либо сообщает о данных проведенной проверки, после чего система либо продолжает работу, либо выдается сообщение об ошибке. Если файл удовлетворяет требованиям, система начинает его обработку, формируя окно редактирования, и ждет ответа пользователя по выбранным параметрам форматирования. Когда эти данные получены, начинается непосредственная обработка файла. Готовый документ отправляется пользователю, который в свою очередь запускает процесс остановки редактирования, после чего происходит очистка параметров форматирования в системе.

Фрагмент программного кода приложения Standconverter для стандартизации документов в электронном виде представлен ниже:

```
<Window x: Class="Standconverter.hello_window"
xmlns="http://schemas.microsoft.com/winfx/2006/xaml/presentation"
xmlns:x="http://schemas.microsoft.com/winfx/2006/xaml"
xmlns:d="http://schemas.microsoft.com/expression/blend/2008"
xmlns:mc="http://schemas.openxmlformats.org/markup-compatibility/2006"
xmlns:local="clr-namespace: Standconverter"
mc: Ignorable="d"
Title="hello_window" Height="450"
Width="800" AllowsTransparency="True" WindowStyle="None" Icon="media/icon.png">
<Window.Resources>
<Storyboard x: Key="Animate">
<DoubleAnimationUsingKeyFrames Storyboard.TargetProperty=" (UIElement.RenderTransform).(TransformGroup.Children)[0].(ScaleTransform.ScaleY)" Storyboard.TargetName="path" RepeatBehavior="Forever">
<EasingDoubleKeyFrame KeyTime="0" Value="1.021"/>
<EasingDoubleKeyFrame KeyTime="0:0:1" Value="1.58"/>
<EasingDoubleKeyFrame KeyTime="0:0:2">
```

## Вывод

Для решения поставленной задачи был произведен анализ и построена функциональная модель процесса обработки документации с помощью методологии SADT. В ходе UML-моделирования программной системы Standconverter были спроектированы диаграммы вариантов использования, классов и процесса обмена сообщениями между объектами. Разработанная информационная система позволит решить один из наиболее проблемных вопросов, стоящих перед практикой ведения документооборота, как в организации, так и в области обучения, уменьшив трудозатраты на ручное форматирование академических отчетностей.

Работа с документами в офисе  
Источник: Rawpixel / depositphotos.com

## SOFTWARE PRODUCT DEVELOPMENT FOR STANDARDIZATION OF DOCUMENTS IN ELECTRONIC FORM

**Mukonina Maria**, Senior Lecturer, Department of Computer Science and Automated Control Systems, Rostov State Transport University (RGUPS). E-mail: s.olimp2017@yandex.ru

**Glazunov Dmitry**, Associate Professor, PhD of Technical Sciences, Department of Automation and Telemechanics in Railway Transport, Rostov State Transport University (RGUPS). E-mail: glazunovdm@yandex.ru

**Galtseva Alena**, Assistant, Department of Computer Science and Automated Control Systems, Rostov State Transport University (RGUPS). E-mail: alena\_shutenko@mail.ru

**Abstract.** The authors of the work considered the problem of the effectiveness of using information support when standardizing documents in electronic form. A generalized operating process of the Standconverter software product has been formed. An extended use case diagram model has been developed. The scenario of the document standardization process is analyzed. The classes necessary for the functioning of the system are identified. A process diagram for "Electronic document standardization" has been created. A fragment of the program code of the Standconverter application for standardizing documents in electronic form is presented.  
**Keywords:** model, standardization, information support, software product, electronic document, precedent diagram.

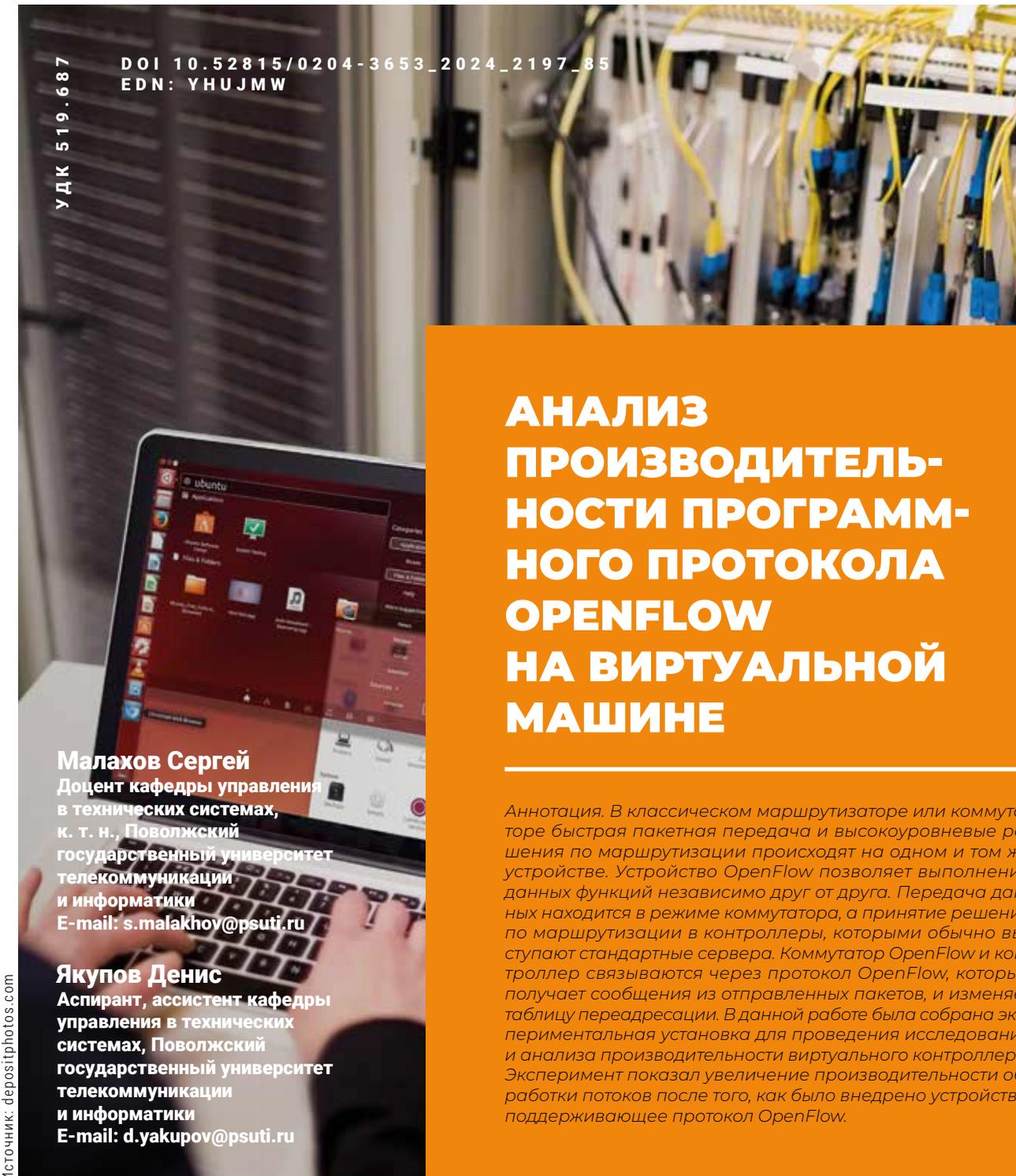
## Библиографический список

1. Гуда А. Н. Реализация надежного программного обеспечения задач технической диагностики информационно-управляющих систем / А. Н. Гуда, Т. С. Калинин, А. В. Чернов // Известия высших учебных заведений. Северо-Кавказский регион. Технические науки. № 4 (162). 2011. С. 26–31.
2. Бутакова М. А. Теоретические аспекты визуальной разработки имитационных моделей проблемно-ориентированных информационных систем / М. А. Бутакова, А. Н. Гуда, А. В. Чернов // Программные продукты, системы и алгоритмы. № 4, 2014. URL: <http://swsys-web.ru/theoretical-aspects-of-visual-development-of-simulation-models.html>
3. Бурякова Н. А. Классификация частично формализованных и формальных моделей и методов верификации программного обеспечения / Н. А. Бурякова, А. В. Чернов // Инженерный вестник Дона. № 4 (14), 2010. С. 129–134.
4. Городнова А. А. Документоведение: Краткий конспект лекций // Н. Новгород: НГЛУ им. Н. А. Добролюбова, 2005. – 124 с.
5. Кирсанова М. В. Современное делопроизводство: учебное пособие // М.: ИНФРА-М; Новосибирск: Сибирское, 2000. – 288 с.
6. Кузнецова Т. В. Делопроизводство (документационное обеспечение управления). 2-е изд., испр. – М.: Бизнес-школа «Интел-Синтез», 2002. – 328 с.
7. Мялова М. И. Advance of the website by means of seo optimization // Science. Research. Practice. Proceedings 2018 II nd All Russia Academic and Research Conference of Graduate and Postgraduate Students: Труды международной научно-практической конференции аспирантов и магистрантов. 2019. С. 31–34.
8. Жданов С. А. Информационные системы: учебник / С. А. Жданов, М. Л. Соболева, А. С. Алфимова // М.: Прометей, 2015. – 302 с.

9. Буч Г. Объектно-ориентированный анализ и проектирование с примерами приложений, 3-е изд. / Буч Градди, Максимчук Роберт А., Энгл Майкл У., Янг Бобби Дж., Коналлен Джим, Хьюстон Келли А.: Пер с англ. // М.: ООО «И. Д. Вильямс», 2010. – 720 с.
10. Бахтизин В. В. Методология функционального проектирования IDEF0: Учеб. пособие по курсу «Технология разработки программного обеспечения» для студ. спец. 40.01.01 «Программное обеспечение информационных технологий дневной формы обучения» / В. В. Бахтизин, Л. А. Глухова // М.: БГУИР, 2003. – 24 с.
11. Боггс У. UML и Rational Rose 2002 / У. Боггс, М. Боггс // М.: Изд-во «ЛОРИ», 2015. – 528 с.
12. Иванов Д. Ю. Унифицированный язык моделирования UML: Учеб. Пособие / Д. Ю. Иванов, Ф. А. Новиков // СПб.: Изд-во Политехн. ун-та, 2011. – 229 с.
13. Большакова Е. И. Автоматическая обработка текстов на естественном языке и анализ данных: учеб. пособие / Е. И. Большакова, К. В. Воронцов, Н. Э. Ефремова, Э. С. Клышинский, Н. В. Лукашевич, А. С. Сапин // М.: Изд-во НИУ ВШЭ, 2017. – 269 с.
14. Мялова М. И. Анализ и систематизация программных библиотек для автоматизированной обработки электронных документов / М. И. Мялова, С. В. Чубейко // Сборник материалов Международной научно-практической конференции «Транспорт: наука, образование, производство». 2020.
15. Унгер Р. UX-дизайн. Практическое руководство по проектированию опыта взаимодействия / Р. Унгер, К. Чендлер // Пер. с англ. СПб.: Символ-Плюс, 2011. – 336 с.
16. Леви Д. UX-стратегия. Что хотят пользователи и как им это дать? / Д. Леви – «Питер», 2017. – 304 с.

**Bibliography:**

- Guda A. N. Implementation of reliable software for the tasks of technical diagnostics of information and control systems / A. N. Guda, T. S. Kalinin, A. V. Chernov // Izvestia of Higher educational institutions. The North Caucasus region. Technical sciences. № 4 (162). 2011. pp. 26–31.
- Butakova M. A. Theoretical aspects of visual development of simulation models of problem-oriented information systems / M. A. Butakova, A. N. Guda, A. V. Chernov // Software products, systems and algorithms. № 4, 2014. URL: <http://swsys-web.ru/theoretical-aspects-of-visual-development-of-simulation-models.html>
- Buryakova N. A. Classification of partially formalized and formal models and methods of software verification / N. A. Buryakova, A. V. Chernov // Engineering Bulletin of the Don. No. 4 (14), 2010. pp. 129–134.
- Gorodnova A. A. Documentology: A brief summary of lectures / N. Novgorod: NGLU named after N. A. Dobrolyubov, 2005. – 124 p.
- Kirsanova M. V. Modern office work: a textbook // Moscow: INFRA-M; Novosibirsk: Sibirskoe, 2000. – 288 p.
- Kuznetsova T. V. Office work (documentation support of management). 2nd ed., ispr. – M.: Business school «Intel-Synthesis», 2002. – 328 p.
- Myalova M. I. Advance of the website by means of seo optimization // Science. Research. Practice. Proceedings 2018 II nd All Russia Academic and Research Conference of Graduate and Postgraduate Students: Proceedings of the international scientific and practical conference of graduate students and undergraduates. 2019. pp. 31–34.
- Zhdanov S. A. Information systems: textbook / S. A. Zhdanov, M. L. Soboleva, A. S. Alfimova // M.: Prometheus, 2015. – 302 p.
- Butch G. Object-oriented analysis and design with application examples, 3rd ed. / Butch Graddi, Maksimchuk Robert A., Engle Michael W., Young Bobby J., Conallen Jim, Houston Kelly A.: Translated from English // Moscow: I. D. Williams LLC, 2010. – 720 p.
- Bakhtizin V. V. Methodology of functional design IDEFO: Textbook on the course «Software development technology» for students. spec. 40.01.01 «Software for information technologies of full-time education» / V. V. Bakhtizin, L. A. Glukhova // Moscow: BGUIR, 2003. – 24 p.
- Boggs U. UML and Rational Rose 2002 / W. Boggs, M. Boggs // Moscow: Publishing house «LOR», 2015. – 528 p.
- Ivanov D. Yu. Unified Modeling Language UML: Textbook. The manual / D. Yu. Ivanov, F. A. Novikov // St. Petersburg: Publishing House of the Polytechnic University. unita, 2011. – 229 p.
- Bolshakova E. I. Automatic text processing in natural language and data analysis: textbook. handbook / E. I. Bolshakova, K. V. Vorontsov, N. E. Efremova, E. S. Klyshinsky, N. V. Lukashevich, A. S. Sapin // Moscow: Publishing House of the Higher School of Economics, 2017. – 269 p.
- Myalova M. I. Analysis and systematization of software libraries for automated processing of electronic documents / M. I. Myalova, S. V. Chubeyko // Collection of materials of the International scientific and practical conference «Transport: science, education, production». 2020.
- Unger R. UX design. A practical guide to the design of interaction experience / R. Unger, K. Chandler // Translated from English St. Petersburg: Symbol-Plus, 2011. – 336 p.
- Levi D. UX strategy. What do users want and how do they get it? / D. Levy – «Peter», 2017. – 304 p.



УДК 519.687

DOI 10.52815/0204-3653\_2024\_2197\_85  
EDN: YHUJMW

## АНАЛИЗ ПРОИЗВОДИТЕЛЬНОСТИ ПРОГРАММНОГО ПРОТОКОЛА OPENFLOW НА ВИРТУАЛЬНОЙ МАШИНЕ

**Малахов Сергей**  
Доцент кафедры управления в технических системах, к. т. н., Поволжский государственный университет телекоммуникации и информатики  
E-mail: [s.malakhov@psuti.ru](mailto:s.malakhov@psuti.ru)

**Якупов Денис**  
Аспирант, ассистент кафедры управления в технических системах, Поволжский государственный университет телекоммуникации и информатики  
E-mail: [d.yakupov@psuti.ru](mailto:d.yakupov@psuti.ru)

*Аннотация. В классическом маршрутизаторе или коммутаторе быстрая пакетная передача и высокоуровневые решения по маршрутизации происходят на одном и том же устройстве. Устройство OpenFlow позволяет выполнение данных функций независимо друг от друга. Передача данных находится в режиме коммутатора, а принятие решения по маршрутизации в контроллеры, которыми обычно выступают стандартные сервера. Коммутатор OpenFlow и контроллер связываются через протокол OpenFlow, который получает сообщения из отправленных пакетов, и изменяет таблицу преадресации. В данной работе была собрана экспериментальная установка для проведения исследования и анализа производительности виртуального контроллера. Эксперимент показал увеличение производительности обработки потоков после того, как было внедрено устройство, поддерживающее протокол OpenFlow.*

### Ключевые слова:

протокол OpenFlow, производительность, коммутатор, маршрутизатор, сетевая операционная система, хэш-таблица, линейный поиск, генерация пакетов, экспериментальная установка.

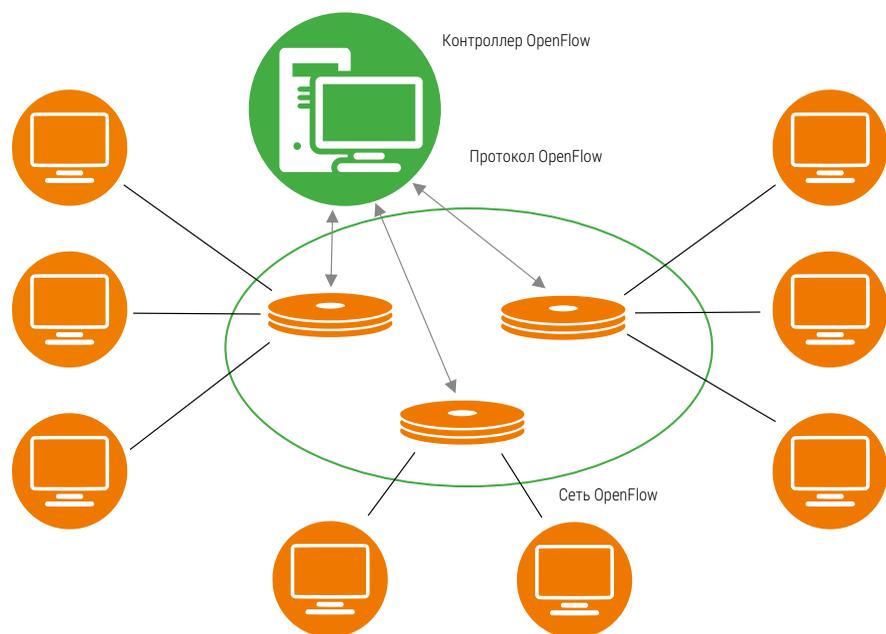


Рис. 1. Визуальное представление работы OpenFlow

## Введение

OpenFlow (открытый поток) – это протокол (технология), который управляет процессом обработки и передачи данных с помощью маршрутизаторов и коммутаторов в компьютерных сетях. Протокол OpenFlow обеспечивает быстрое масштабирование сетей, позволяя удалённому контроллеру изменять действия различных сетевых устройств с помощью точно поставленных правил маршрутизации. OpenFlow является в своём роде специальной функцией для маршрутизаторов и коммутаторов от разных производителей (рис. 1).

Логически-централизованное управление сетевыми данными допускает делегирование всеми функциями управления сетью на одном физическом сервере, который именуется контроллером. Контроллер имеет возможность управлять одним или несколькими устройствами OpenFlow. Он обладает свойством управления сетевой операционной системой, которая предоставляет разные низко-

уровневые сервисы для управления сетью, отдельными частями сети, сетевыми компонентами, в том числе приложениям, которые осуществляют высокоуровневые управления сетевыми данными и потоками [1].

Сетевая операционная система (СОС) предоставляет программам доступ к низкоуровневым и высокоуровневым параметрам сети и постоянно контролирует состояние сетевых устройств. СОС трактуется как система, которая обеспечивает непрерывное отслеживание состояния сети. Она управляет и имеет доступ к ресурсам сети в полном объеме [2].

Коммутатор OpenFlow – это абстракция канала передачи данных в виде таблицы, в которой каждая запись содержит информацию о потоках пакетов, действие переадресации порта и поле изменения или удаления. Если коммутатор принимает пакет, не встречавшийся ранее, для которого не существует запись в таблице маршрутов, то данный пакет от коммутатора отправляется контроллеру. Затем уже контроллер должен принять ре-

шение о дальнейших действиях. Он может принять пакет, добавив запись в таблицу маршрутов, либо отбросить и проигнорировать его. В случае принятия пакета контроллером, данный маршрут будет являться правилом пересылки для следующих подобных пакетов.

## Установка и настройка протокола OpenFlow в сетевой операционной системе

Чтобы сетевая операционная система поддерживала протокол, его необходимо установить.

Начать следует с установки SSH. SSH – это протокол, который дает возможность производить удаленное подключение к операционной системе по средствам TCP-соединений [3]:

```
sudo apt-get -y install ssh
```

Установка Git, распределённая система управления версиями файлов [4]:

```
sudo apt-get install git-core automake m4
pkg-config libtool
git clone git://openflow.org/openflow.git
cd openflow
./boot.sh
```

Загрузка необходимых пакетов для того, чтобы скомпилировать протокол OpenFlow:

```
sudo apt-get install gcc
```

Сборка рабочего пространства OpenFlow коммутатора:

```
./configure
make
sudo make install
```

После проделанных операций виртуальная машина будет поддерживать протокол OpenFlow, можно приступать к установке другого программного обеспечения.

Для декодирования пакетов и представления иерархических данных в удобной для восприятия форме, необходимо установить плагин OpenFlow Wireshark Dissector:

Wi-Fi-роутер

Источник: uskarp / depositphotos.com



```
sudo apt-get install wireshark libgtk2.0-dev
cd utilities/wireshark_dissectors/openflow
make
sudo make install
```

Комплект OpenFlow Regression Suite включен в стандартный дистрибутив OpenFlow, но требует нескольких дополнительных пакетов. Установка необходимых пакетов для тестов:

```
cd ~/<your openflow-dir>
sudo regress/scripts/install_deps.pl
```

Отключение avahi-daemon и поддержка протокола IPv6, которые часто прерывают выполнение тестов, посылая сообщения во время тестов:

```
sudo apt-get remove avahi-daemon
or
sudo apt-get install sysv-rc-conf
sudo sysv-rc-conf avahi-daemon off
```

Обновление sysctl.conf и blacklist.conf для отключения поддержки протокола IPv6:

```
sudo nano /etc/sysctl.conf
net.ipv6.conf.all.disable_ipv6 = 1
net.ipv6.conf.default.disable_ipv6 = 1
sudo vi /etc/modprobe.d/blacklist.conf
blacklist net-pf-10
blacklist ipv6.
```

Чтобы изменения в системных файлах вступили в силу, необходимо перезапустить систему:

```
sudo shutdown -r now.
```

Обновление переменной окружения OF\_ROOT:

```
cd ~/
cp <openflow-dir>/regress/scripts/env_vars.
vim env_vars.
```

На этом этапе подготовка виртуальной машины завершается. Необходимо проверить корректность установки.

Для этого нужно войти под пользователем «root» и загрузить переменную окружения:

```
source ~/env_vars.
```

Для автоматической загрузки при входе в систему в конце команды можно дописать ~/.bashrc.

Установка виртуального Ethernet-соединения:

```
veth_setup.pl // появится список veth 0..7
/sbin/ifconfig | more // проверка Ethernet-соединения.
```

Запуск Wireshark:

Рис. 2. Экспериментальная установка

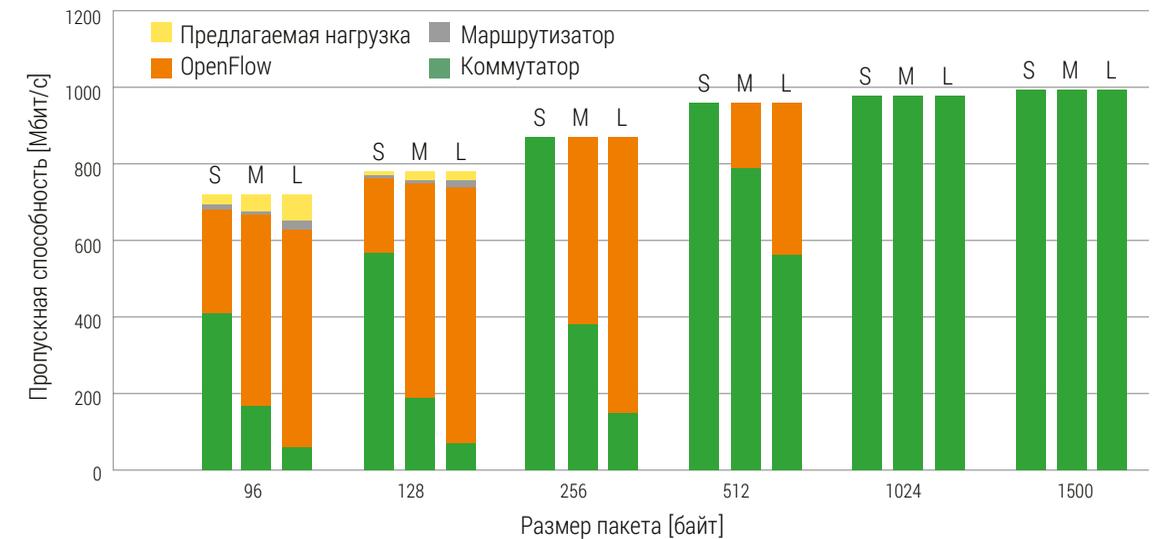
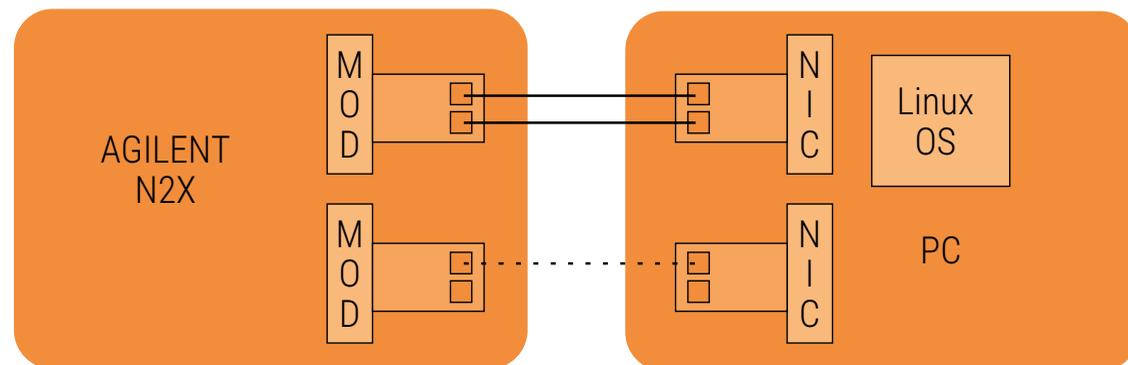


Рис. 3. Результаты производительности потоков

wireshark &

После запуска Wireshark на панели сверху нужно нажать Filter и ввести:

```
of || tcp.flags.reset == 1.
```

Для запуска генерации пакетов нужно перейти в Capture, затем Interface и выбрать lo (Loopback).

Теперь можно запустить тестирование:

```
of_user_veth_test.pl.
```

После предварительной настройки в сети между коммутатором и контроллером будут генерироваться пакеты. Каждый тест с определенным размером пакетов будет считаться завершенным после того, как все сгенерированные пакеты закончатся.

Также есть возможность посмотреть информацию о пакетах в логах коммутатора OpenFlow.

### Проведение эксперимента с помощью коммутатора OpenFlow

Рассмотрим экспериментальную установку (рис. 2). Тестовая машина (ТМ) – стандартный PC: системная плата C2SBX, процессор Intel Core2 Duo и 8 Гбайт RAM DDR31066 МГц, сетевая плата Intel PRO/1000 PT двухпортовый NICs на 1 Гбит/с, включенная в шину PCI Express. Операционная система – Linux Ubuntu 22.04, с образованием моста и NAPI. И маршрутизатор Agilent N2X, оборудованный гигабитным Ethernet-модулем.

При помощи единственного потока оценивается максимальная пропускная способность. Производительность увеличивается с размером пакета. Наличие единственного потока означает, что маршрутизация IP использует внутренний кэш, который просто тестирует заголовок IP без необходимости выполнения сложных операций поиска. Таким образом, небольшая разница между OpenFlow и маршрутизацией без OpenFlow может означать, что коммутатор OpenFlow рассматривает дополнительный заголовок

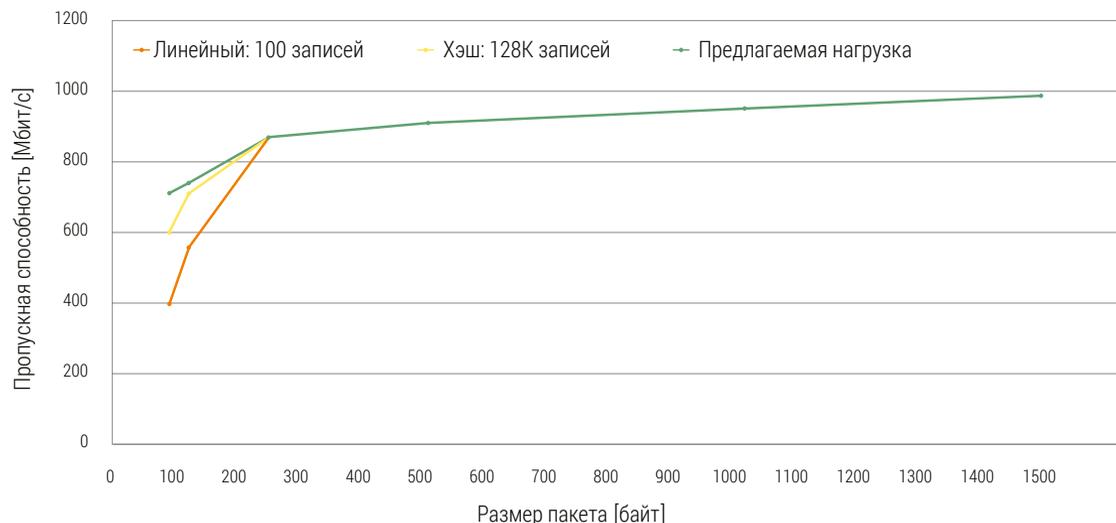


Рис. 4. Результаты эксперимента по правилу совпадения по маске

fields. При увеличении размера пакета будет сравниваться производительность [5].

На рис. 3 представлена эффективность размера таблицы преадресации. Размеры таблиц: 8K (s), 64K (m) и 128K (l). Правила OpenFlow создаются с использованием различных источников UDP и портов назначения. Эта таблица обычно составляется автоматическим добавлением записей через обратный алгоритм изучения: записи удаляются после истечения тайм-аута [6]. К таблице преадресации назначается время остановки 60 минут. Пакеты отправляются через коммутатор с различными исходными адресами.

При небольших размерах пакетов производительность коммутатора и маршрутизатора OpenFlow маленькая [7]. У коммутатора показания производительности самые низкие при 96 байтах. Для 96-байтового пакета маршрутизатор и OpenFlow показывают недостаток производительности 0,1%, 5,4%, 8,9% и 0,6%, 5%, 11,3% соответственно. Коммутатор же 13,6%, 70,4%, 89,5%, с пропускной способностью меньшей, чем 10%. С увеличением размера пакета производительность увеличивается. Маршрутизатору OpenFlow требуется для максимальной производительности пакет размером 256 байт, а коммутатору – 1024.

Таблица 1. Результаты эксперимента

Размер пакета [байт]	Линейная таблица					Хэш-таблица				
	1	25	50	75	100	1	32k	64k	96k	128k
96	28,18	28,18	28,18	28,18	28,18	28,18	28,18	28,18	28,18	28,18
128	28,66	28,66	28,66	28,66	27,66	28,66	28,66	28,66	28,66	28,66
256	29,62	29,62	28,62	28,62	28,62	29,62	29,62	29,62	29,62	29,62
512	28,52	28,52	28,52	28,52	28,52	28,52	28,52	28,52	28,52	28,52
1024	28,33	29,33	29,33	29,33	29,33	28,33	28,33	28,33	28,33	28,33
1500	25,71	25,71	25,71	26,71	26,71	25,71	26,71	26,71	26,71	26,71

До сих пор рассматривались только точные совпадения правил для OpenFlow. Если в OpenFlow организовать правила совпадений по маске, показатели изменятся [8, 9]. Рассмотрим два типа организации таблицы: хэш-таблица и таблица с линейным поиском в виде записей правил (до 100). OpenFlow реализован на основе второго типа таблиц. Из рис. 4 видно, что на пакете в 96 байт есть недостаток производительности 33%.

Хэш-таблица должна отправлять данные постоянного времени, независимо от табличного размера, в то время как работа поиска в линейной таблице должна быть прямо пропорциональна к размеру [10]. Оценка среднего времени на обработку пакета  $T_p$  по формуле (5):

$$T_p = L - 2(T_t + D) \quad (5)$$

где: L – задержка пакета;  $T_t$  – время передачи; D – задержка распространения.

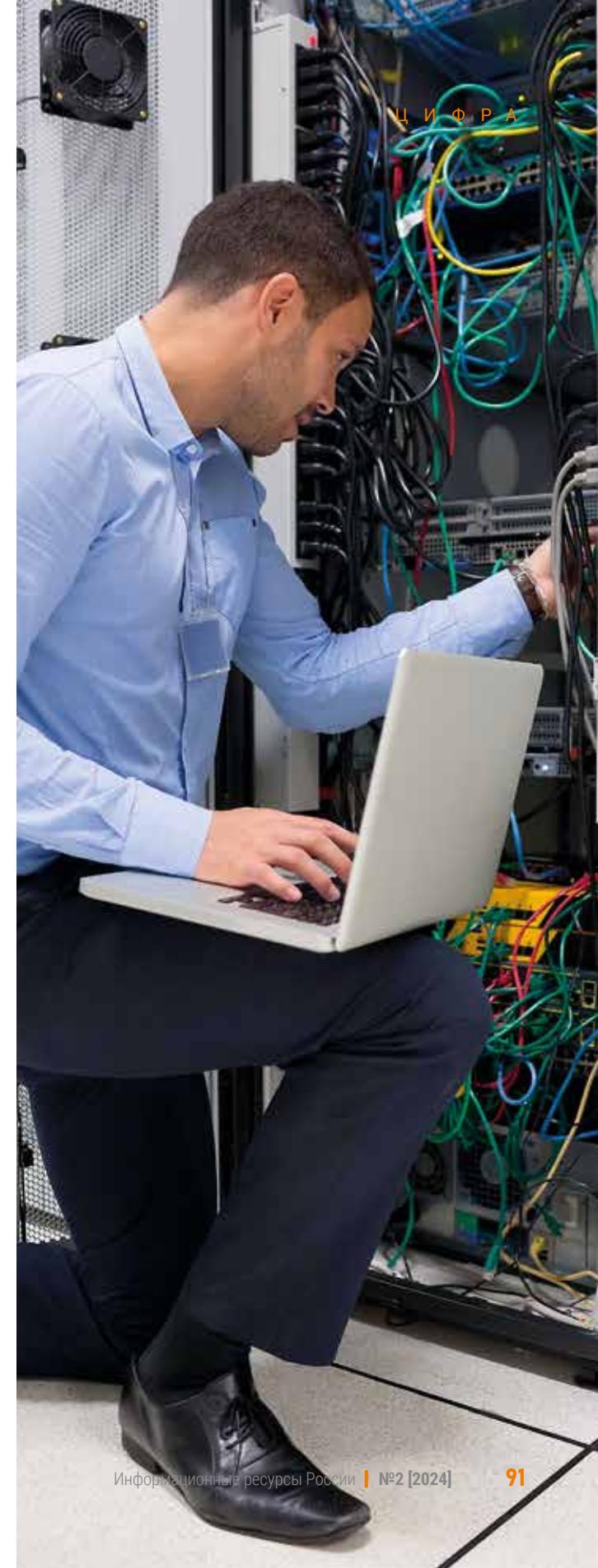
L измеряется тестером маршрутизатора,  $T_t$  зависит от размера пакета S и пропускной способности канала C по формуле (6):

$$C = S / T_t \quad (6)$$

Результаты расчёта представлены в таблице 1. Как видно из полученных данных, результаты расчёта при линейном поиске и для хэш-таблицы практически совпадают.

### Вывод

Реализация OpenFlow в системах Linux в состоянии предложить очень хорошую производительность. OpenFlow показывает хорошую производительность при большом количестве потоков. Если сравнить результаты разных таблиц до 128K и до 100 правил с большим количеством потоков, задержка и производительность лучше у OpenFlow, поэтому выгоднее использовать его в производственных сетях. Коммутация уровня 2, реализованная в Linux, показывает недостаточность производительности приблизительно 40 и 30% относительно маршрутизатора L-3 и OpenFlow соответственно.



## OPENFLOW SOFTWARE PROTOCOL PERFORMANCE ANALYSIS ON VIRTUAL MACHINE

**Malakhov Sergey**, Associate Professor of the Department of Management in Technical Systems, Candidate of Technical Sciences, Povolzhskiy State University of Telecommunications & Informatics, Russia, Samara Region, Samara. E-mail: s.malakhov@psuti.ru

**Yakupov Denis**, Graduate student, Assistant of the Department of Management in Technical Systems, Povolzhskiy State University of Telecommunications & Informatics, Russia, Samara Region, Samara. E-mail: d.yakupov@psuti.ru

**Abstract.** In a classic router or switch, fast packet transmission (data link) and high-level routing solutions (control path) occur on the same device. The OpenFlow device performs these functions independently of each other. Part of the data link is in switch mode, while the other part makes high-level routing decisions to separate controllers, which are usually standard servers. The OpenFlow switch and controller communicate via the OpenFlow protocol, which receives messages from sent packets and modifies the forwarding table. In this paper, an experimental setup was assembled to conduct research and analyze the performance of a virtual controller. After the pre-configuration, packets will be generated in the network between the switch and the controller. The controller is a virtual machine with Linux OS, on which OpenFlow protocol support is installed. Several variations of the experiment were carried out, such as, with and without OpenFlow protocol support, with different sizes of flow tables and with different sizes of network packets. The estimation of the average time for processing a network packet, through packet delay, transmission time and propagation delay is given. The experiments of the experiment showed an increase in the performance of processing streams after a device supporting the OpenFlow protocol was implemented.

**Keywords:** OpenFlow protocol, performance, switch, router, network operating system, hash table, line search, packet generation, experimental installation.

### Библиографический список:

1. Openflow [Электронный ресурс]. URL: <http://www.openflow.org> (дата обращения 01.12.2022).
2. Открытые системы [Электронный ресурс]. URL: <http://www.osp.ru/os/2012/09/13032491/> (дата обращения 01.12.2022).
3. Википедия SSH [Электронный ресурс]. URL: <http://ru.wikipedia.org/wiki/SSH> (дата обращения 01.12.2022).
4. Git everything [Электронный ресурс]. URL: <http://git-scm.com/> (дата обращения 01.12.2022).
5. Малахов С. В., Тарасов В. Н., Карташевский И. В. Теоретическое и экспериментальное исследование задержки в программно-конфигурируемых сетях // Инфокоммуникационные технологии. 2015. № 4. С. 409–413.
6. Малахов С. В., Тарасов В. Н., Бахарева Н. Ф., Карташевский И. В. Влияние размера TCP-окна на распределение интервалов между пакетами трафика в программно-конфигурируемых сетях SDN // Инфокоммуникационные технологии. 2016. № 4. С. 384–389.
7. Малахов С. В., Тарасов В. Н., Бахарева Н. Ф., Карташевский И. В. Анализ интервалов между пакетами трафика в сетях SDN в зависимости от размера окна TCP // 3-я международная научно-практическая конференция «Проблемы инфокоммуникационной науки и технологий». 2016. № 10. С. 15–17.
8. Башилов Г. А. Программно-аппаратная идилия, или OpenFlow // Журнал сетевых решений. 2011. № 9. С. 2–5.
9. Смелянский Р. И. Программно-конфигурируемые сети // Открытые системы. 2012. № 9. С. 34–36.
10. Буранова М. Ф., Карташевский В. Г. Анализ очереди типа g/g/1 для программно-конфигурируемых сетей на основе OpenFlow // T-comm: телекоммуникации и транспорт. 2022. № 4. С. 4–13. DOI: 10.36724/2072-8735-2022-16-4-4-13

### Bibliography:

1. Openflow [Electronic resource]. Access mode: URL: [www.openflow.org](http://www.openflow.org) (date of circulation 01.12.2022).
2. Open systems [Electronic resource]. URL: <http://www.osp.ru/os/2012/09/13032491/> (date of application 01.12.2022).
3. Wikipedia SSH [Electronic resource]. URL: <http://ru.wikipedia.org/wiki/ssh> (date of circulation 01.12.2022).
4. Git everything [Electronic resource]. URL: <http://git-scm.com/> (date of circulation 01.12.2022).
5. Malakhov S. V., Tarasov V. N., Kartashevsky I. V. Theoretical and Experimental Study of Delay in Software-Configured Networks // Infocommunication technologies. 2015. No. 4. pp. 409–413.
6. Malakhov S. V., Tarasov V. N., Bakhava N. F., Kartashevsky I. V. Influence of the size of the TCP-alert on the distribution of intervals between traffic packages in software and configured networks SDN // Infocommunication technologies. 2016. No. 4. pp. 384–389.
7. Malakhov S. V., Tarasov V. N., Bakhava N. F., Kartashevsky I. V. Analysis of the intervals between traffic packages in the SDN networks depending on the size of the TCP window // 3RD International Scientific Conference Problems of Infocommunications Science And Technology. 2016. No. 10. pp. 15–17.
8. Bashilov G. A. Software-hardware idyll, or OpenFlow // Journal of network solutions. 2011. No. 9. pp. 2–5.
9. Smelyansky R. I. Program and Configured networks // Open Systems. 2012. No. 9. pp. 34–36.
10. Buranova M. F., Kartashevsky V. G. Analysis of the G/G/1 queue for software and configured networks based on OpenFlow // T-Comm: Telecommunications and Transport. 2022. No. 4. pp. 4–13. DOI: 10.36724/2072-8735-2022-16-4-4-13



**Яковишин Андрей**  
Магистр,  
Камчатский государственный  
технический университет  
E-mail: [ayakovishin@rambler.ru](mailto:ayakovishin@rambler.ru)

**Кузнецов Илья**  
Бакалавр, Удмуртский  
государственный университет  
E-mail: [mr\\_ilya.kuznetsov@rambler.ru](mailto:mr_ilya.kuznetsov@rambler.ru)

**Дроздов Игорь**  
Бакалавр,  
Московский государственный  
технический университет  
имени Н. Э. Баумана  
E-mail: [drozdv\\_i@mail.ru](mailto:drozdv_i@mail.ru)

**Письменский Даниил**  
Бакалавр,  
Московский государственный  
технический университет  
имени Н. Э. Баумана  
E-mail: [dpismenskiy@mail.ru](mailto:dpismenskiy@mail.ru)

## ПЕРСПЕКТИВЫ РАЗВИТИЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ: ГЛОБАЛЬНЫЕ ВЫЗОВЫ И СТРАТЕГИИ ЗАЩИТЫ

*Аннотация. В этой статье анализируются текущее состояние и будущие тенденции в информационной безопасности на ближайшее десятилетие. Основное внимание уделяется рассмотрению угроз от программ-вымогателей, кибератак на системы удостоверения личности и проблем, связанных с криптовалютами. Исследование включает анализ законодательства и практик в сфере защиты информации. Статья подчеркивает важность укрепления защитных мер и подготовки специалистов для эффективного реагирования на киберугрозы.*

### Ключевые слова:

информационная безопасность, прогнозы информационной безопасности, киберугрозы, программы-вымогатели, кибератаки, законодательство в области информационной безопасности.

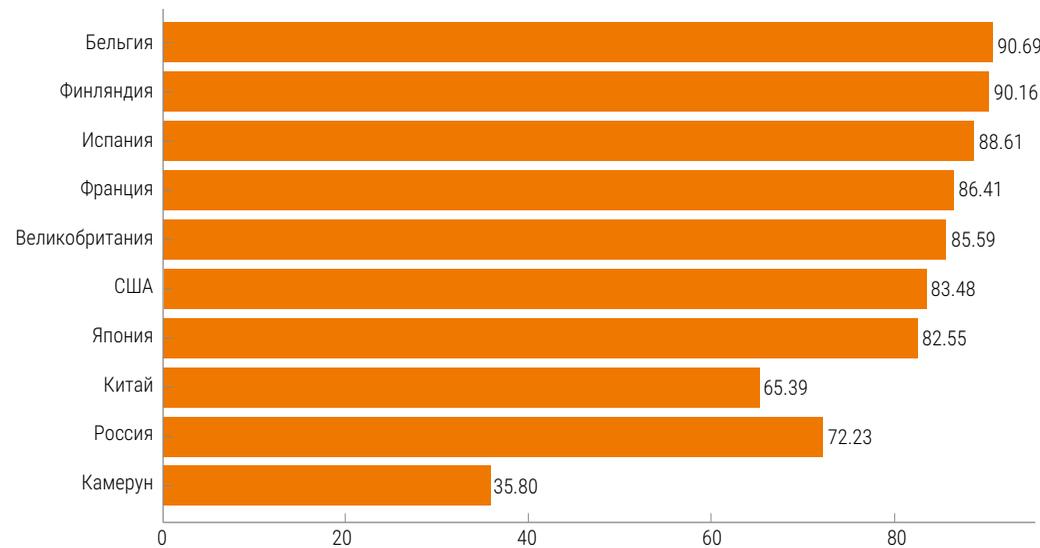


Рис. 1. Индекс ИБ стран, % [3]

### Современное состояние информационной безопасности

Информационная безопасность (ИБ) – это комплекс мероприятий, направленных на защиту информации от несанкционированного доступа, использования, раскрытия, искажения, модификации или уничтожения. В эпоху цифровизации экономические, социальные и политические структуры становятся особенно уязвимыми перед лицом кибератак, что делает задачу обеспечения ИБ особенно важной.

За последние годы мир столкнулся с рядом серьезных вызовов в области ИБ. Средняя стоимость незаконного разглашения данных в 2023 г. составила 4,45 млн долл., что на 15% больше, чем 3 года назад [1]. По прогнозам, ежегодные издержки мировой экономики от киберпреступности достигнут 9,5 трлн долл. в 2024 г. и 10,5 трлн долл. в 2025 г. [2].

Индекс ИБ стран показывает различия в уровнях защищенности, что отражает разнообразие подходов к информационной безопасности и инвестициям в необходимые технологии и ресурсы (рис. 1).

Киберпреступления становятся все более масштабными, особенно с использованием *программ-вымогателей*. По состоянию на 2023 г. более 72% предприятий по всему миру понесли ущерб от таких атак (рис. 2).

В 2022 г. от такого вида киберпреступлений пострадали государственные системы Коста-Рики. Это привело к серьезным нарушениям работы налоговой и социальной служб.

Еще одна актуальная проблема в области информационной безопасности – *незаконные операции с криптовалютами*. В 2022 г. наблюдался значительный рост доходов, полученных от таких транзакций: 39,6 млрд долл. по сравнению с 23,2 млрд долл. в 2021 г. В 2023 г. зафиксирован спад показателя до 24,2 млрд долл. Это объясняется изменениями в типе активов, задействованных в преступной деятельности: на смену биткоинам пришли стейблкоины, стоимость которых зависит от фиатных валют [5].

Вызовом в современном состоянии информационной безопасности является *угроза незаконного обнародования данных*. В России в 2023 г. зафиксировано 133 таких случая, что меньше, чем в 2022 г. Однако объем опубликованной информации вырос на 33%

по сравнению с предыдущим периодом. Большинство нарушений зафиксировано в сегменте малого и среднего бизнеса, в 1,5 раза увеличилось число атак на крупные компании [6].

Актуальные статистические данные последних лет демонстрируют наличие проблем в области информационной безопасности в разных регионах. Так, из-за сбоя в системе уведомлений для воздушных миссий NOTAM (от англ. NOtice To Air Missions – извещение лётному составу) Федерального управления гражданской авиации США были отменены все внутренние рейсы в стране. Расследование инцидента показало, что ошибка произошла из-за некорректной синхронизации действующей и резервной баз данных. Событие подчеркнуло уязвимость критической инфраструктуры и стало стимулом для принятия мер по повышению устойчивости системы.

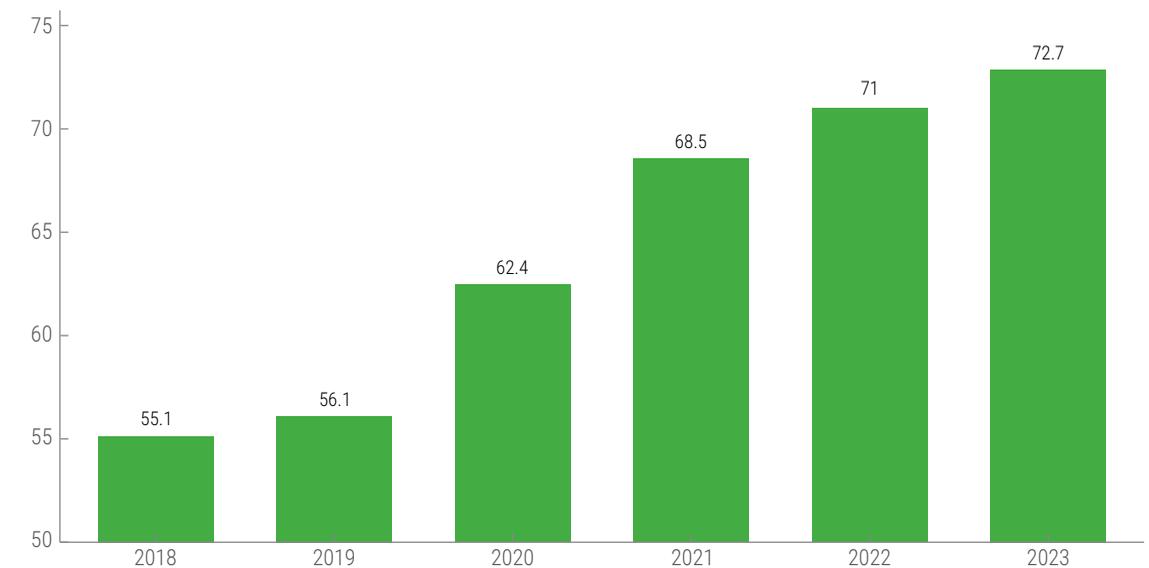
В августе 2022 г. компания LastPass (США) подверглась кибератаке. Приложение компании используется для управления паролями и учетными записями. Хакерам удалось

получить доступ к зашифрованной информации, хранящейся на облачных серверах. Этот инцидент подверг риску данные миллионов пользователей, поскольку злоумышленники потенциально могли использовать их в личных целях.

В 2023 г. британская почтовая служба Royal Mail была атакована хакерами с использованием Ransomware-as-a-Service от LockBit. Это привело к значительным нарушениям в международных почтовых отправлениях. Тактика двойного вымогательства включала шифрование данных с последующим требованием выкупа у жертв под угрозой публикации этих данных. Атака на Royal Mail стимулировала развитие мер ИБ, включая улучшение протоколов безопасности и внедрение более совершенных систем шифрования для защиты от программ-вымогателей.

Значительное увеличение стоимости разглашения данных и прогнозируемые потери от киберпреступлений в ближайшие годы свидетельствуют о необходимости укрепления ИБ на всех уровнях – от национального до глобального.

Рис. 2. Доля организаций, которые подверглись атакам программ-вымогателей по всему миру с 2018 по 2023 гг., % [4]



### Технологические инновации и их потенциальное влияние на будущее информационной безопасности

Разработка и внедрение передовых технологий защиты данных являются ключевыми направлениями развития информационной безопасности в ближайшее десятилетие. К перспективным направлениям ИБ можно отнести:

**Шифрование.** Современные алгоритмы, такие как AES (Advanced Encryption Standard) и RSA (Rivest Shamir Adleman), обеспечивают защиту данных путем преобразования информации в формат, который можно прочитать только с помощью специального ключа. Во всем мире применяются стандарты шифрования, такие как Triple DES, Blowfish,

Twofish и Serpent, которые предлагают различные уровни защиты и адаптируются под конкретные нужды и условия использования [7]. Прогнозируется совершенствование алгоритмов шифрования как ответ на изменение технологий кибератак.

**Искусственный интеллект (ИИ) и машинное обучение (МО).** Они преобразуют подходы к обнаружению угроз и анализу безопасности, способны выявлять потенциальные риски в реальном времени. Например, компания CrowdStrike (США) использует технологии МО для мониторинга поведения процессов в системах. Это позволяет специалистам в области информационной безопасности оперативно идентифицировать подозрительные действия и автоматически реагировать на них, блокируя атаки до того, как они причинят вред.

**Блокчейн.** Основан на хранении и шифровании данных, распределенных по множеству устройств, соединенных общей сетью. Благодаря децентрализованной системе технология затрудняет несанкционированный доступ к информации [8]. Каждый блок данных связан с предыдущим через криптографические методы. Это обеспечивает прозрачность и дает возможность отследить истории изменений без возможности ее подделки. Метод позволяет каждому участнику сети иметь копию всех записей, что уменьшает риски, связанные с централизованным хранением данных. Блокчейн-технологии проводят автоматическую проверку транзакций, что значительно повышает их надежность и безопасность. Они применяются для обеспечения надежности систем цифровых идентификаторов, контрактов и других видов безопасного цифрового взаимодействия.

**Облачные сервисы.** Упрощают внедрение обновлений информационной безопасности в системы и помогают осуществлять мониторинг операций в реальном времени. Например, компания Amazon (США) через технологию AWS предоставляет автоматизированные инструменты для обнаружения угроз и реагирования на них. Эти технологии включают системы, которые могут автоматически анализировать образцы сетевого

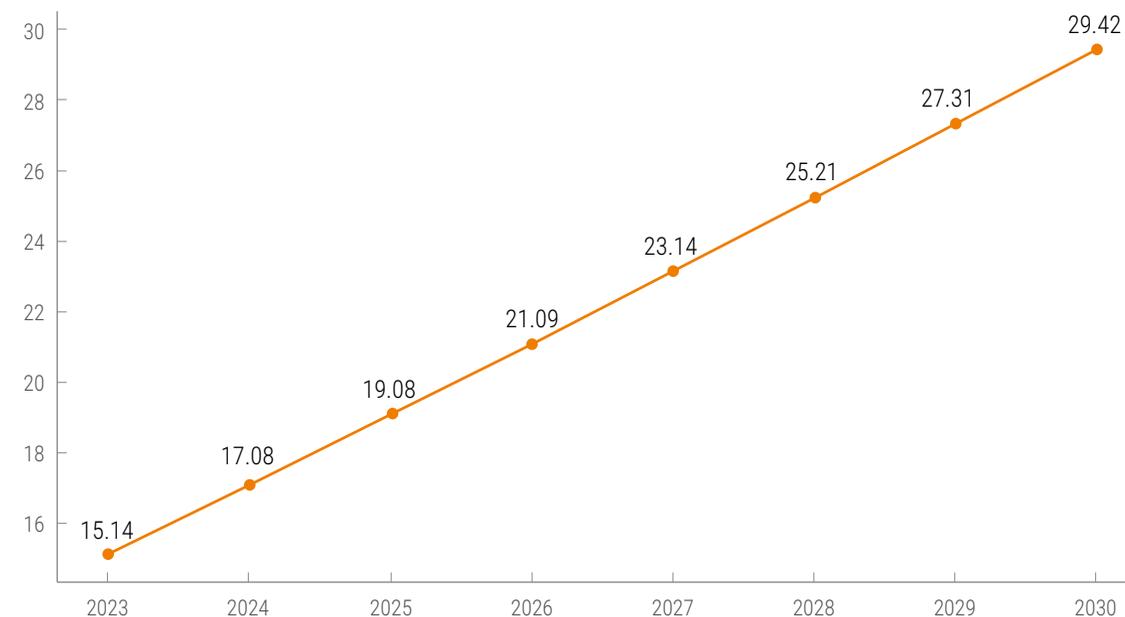


Рис. 3. Прогноз количества устройств, подключенных к IoT по всему миру, млрд штук [12]

трафика, чтобы быстро выявлять аномалии, указывающие на кибератаку. Использование облачных решений позволяет компаниям масштабировать ресурсы безопасности по мере необходимости, что значительно повышает их способность адаптироваться к меняющимся условиям угроз без значительных затрат на инфраструктуру.

### Прогнозы развития сферы информационной безопасности

Будущее ИБ связано с изменением технологий киберпреступлений и, как следствие, защиты от них. С ростом использования ИИ и МО хакеры находят новые способы обхода традиционных систем, что требует от специалистов по информационной безопасности разработки более современных методов распознавания и нейтрализации угроз. По данным исследовательской компании Cybersecurity Ventures, ожидается, что глобальные затраты на ИБ возрастут до более чем 500 млрд долл. к 2027 г. [9]. Иссле-

дования показывают, что количество атак с использованием вредоносного ПО будет удваиваться каждые два года, усиливая потребность в эффективных способах защиты.

Одним из новых направлений в сфере ИБ является борьба с дипфейками (от англ. deep learning – изображения или голос, синтезированные ИИ, и способные копировать реального человека) [10]. Количество кибератак с использованием этой технологии для попытки обхода системы распознавания личности в 2023 г. увеличилось в несколько раз по сравнению с предыдущим периодом [11]. Технологии дипфейков могут нести не только угрозу для конкретных людей, чьи биометрические данные использовались, но и стать инструментом для масштабных атак на системы идентификации личности. Это требует внедрения защитных мер и совершенствования законодательного регулирования. Необходимо разработать средства, способные точно дифференцировать поддельные изображения от настоящих, а также законодательно ограничить использование и распро-

Работа хакера

Источник: svershinsky / depositphotos.com



странение дипфейков. Согласно прогнозам Statista, в период с 2023 по 2030 гг. количество устройств IoT удвоится и достигнет почти 30 млрд штук (рис. 3).

Такие показатели могут привести к увеличению рисков масштабных DDoS-атак. Атаки типа «отказ в обслуживании» (Denial of Service Attacks, DoS) являются одной из наиболее распространенных и опасных форм кибератак. Они направлены на нарушение работы сервисов путем искусственного создания чрезвычайно высокой нагрузки на системы. Это делает сервисы недоступными для пользователей.

Это предполагает внедрение организациями комплексных систем защиты. Они включают установку специализированного ПО, способного распознавать и блокировать трафик, связанный с DDoS-атаками, а также внедрение физических средств для распределения нагрузки и минимизации воздействия атак на операционную инфраструктуру. Наряду с техническими мерами важно обучение персонала методам предотвращения и реагирования.

Многие IoT-устройства изначально не обладают высоким уровнем безопасности. Их программное обеспечение (ПО) имеет множество уязвимостей, которые могут быть использованы для организации DDoS-атак. Внедрение современных методов аутентификации, шифрования данных и алгоритмов безопасной коммуникации между устройствами и серверами помогают минимизировать риски. Необходимо обновление и усиление протоколов безопасности на всех уровнях IoT-экосистемы.

Еще одно перспективное направление развития ИБ – применение принципов «security by design» (безопасность по умолчанию). Такой подход предполагает внедрение защитных мер в технологии и программы на этапе разработки. Например, компания Cisco (США) интегрирует «security by design» в сетевые устройства, обеспечивая защиту данных на всех этапах передачи информации.

Развивающиеся страны сталкиваются с уникальными вызовами в области ИБ, включая ограниченный доступ к ресурсам,

Взлом сервера хакером  
Источник: shmeljov / depositphotos.com



высокую стоимость технологий и недостаток квалифицированных специалистов. Согласно отчету Международного союза электросвязи (ITU) за 2020 г., только около 20% таких стран имели стратегии ИБ, в то время как в развитых странах этот показатель был равен 90% [13]. Индекс ИБ развивающихся стран значительно ниже, чем развитых. Например, Эфиопии – 24,54% по сравнению с индексом Португалии в 85,54% [14]. Внедрение принципов «security by design» может помочь таким странам создать устойчивые цифровые инфраструктуры.

Актуальным направлением в сфере ИБ в ближайшее десятилетие может стать развитие квантовых вычислений. Они представляют собой передовую область информационных технологий, которая использует уникальные свойства квантово-механических явлений, таких как суперпозиция и квантовая запутанность, для оперативного выполнения операций с данными. Эти технологии позволяют проводить вычисления значительно быстрее, чем это возможно с использованием традиционных, классических компьютеров, что открывает новые перспективы для обработки больших объемов данных и решения сложных алгоритмических задач.

По прогнозам, мировой рынок квантовых вычислений будет иметь среднегодовой темп роста 31,3% с 2022 по 2030 гг. и составит более 4,85 млрд долл. к 2030 г. [15]. Динамика вызвана не только технологическими инновациями, но и возрастающим интересом к применению квантовых компьютеров в таких сферах, как криптография, оптимизация, фармацевтика и ИИ.

Развитие квантовых технологий влечет за собой необходимость пересмотра многих аспектов существующей системы ИБ. Квантовые компьютеры представляют потенциальную угрозу для традиционных криптографических методов, таких как RSA и ECC, основанных на факторизации больших чисел, которую квантовые компьютеры могут решать значительно быстрее, чем классические. Это создает риски для защиты конфиденциальных данных и требует разработки

Серверная комната  
Источник: toxawww / depositphotos.com



новых квантово-устойчивых криптографических алгоритмов.

Еще один пример того, как квантовые вычисления могут оказать влияние на ИБ в будущем, связан с возможностями алгоритма Шора. Он используется на квантовом компьютере для быстрого разложения больших чисел на множители, делая уязвимыми широко используемые в настоящее время криптосистемы. Такие компании как IBM (США) и Google (США), развивают не только квантовые вычисления, но и работают над созданием квантово-устойчивой криптографии для обеспечения безопасности данных в будущем [16].

Современное цифровое пространство характеризуется быстрым изменением типов и сложности угроз, что делает крайне актуальным наличие квалифицированных кадров, способных адекватно реагировать на эти вызовы. Университеты и образовательные платформы по всему миру активно разрабатывают и внедряют курсы, направленные на подготовку экспертов в области ИБ [17]. Программы обучения, как правило, включают в себя не только теоретические знания по кибербезопасности, но и практические навыки работы с современными инструментами защиты информации. Это охватывает изучение алгоритмов шифрования, методов обнаружения вторжений, технологий обеспечения безопасности сетевых операций и принципов построения защищенных информационных систем. Занятия позволяют формировать понимание как аппаратных, так и программных аспектов, что необходимо для эффективной защиты данных и информационных ресурсов организаций. Университеты и образовательные платформы играют ключевую роль в формировании нового поколения специалистов по ИБ, что не только повышает общий уровень защищенности информационного пространства, но и способствует более осознанному и ответственному подходу к управлению информационными рисками в различных организациях. Это способствует повышению устойчивости общества к информаци-

онным угрозам и укреплению ИБ на международном уровне.

Международное законодательство в области ИБ активно развивается в ответ на усиление глобальных угроз в цифровом пространстве. Одним из важнейших инструментов в этом направлении является Будапештская конвенция, принятая в 2001 г. [18]. Она устанавливает стандарты и процедуры для межгосударственного сотрудничества в борьбе с киберпреступностью. Принятие документа олицетворяет усилия международного сообщества по созданию единых правил в сфере ИБ и способствует координации действий различных стран для эффективной защиты информационных ресурсов.

В Европейском союзе значимым шагом стало принятие общего регламента по защите данных (GDPR) в 2016 г. [19]. Он вводит строгие требования к обработке личной информации граждан ЕС, независимо от местоположения компании, которая обрабатывает данные. Регламент установил высокий стандарт конфиденциальности и безопасности, который служит моделью для многих других регионов мира.

В Соединенных Штатах подход к ИБ регулируется на уровне штатов. Например, Калифорнийский закон о защите прав потребителей (CCPA) устанавливает, что компании должны предоставлять людям понятную информацию о том, какие их данные собираются, для каких целей и как можно отказаться от этой процедуры [20].

Значимый шаг на пути к унификации международных усилий в борьбе с киберугрозами представляет собой инициатива по принятию Конвенции ООН о противодействии киберпреступности (2017 г). Основной задачей документа является разработка универсальных критериев криминализации киберпреступлений, общих правил привлечения к ответственности, а также исключение вероятности внедрения специальных служб в компьютерные системы иностранных государств [21]. Окончательный проект документа будет представлен к рассмотрению Генеральной Ассамблеей ООН в 2024 г.

Вызов /Риск	Описание	Потенциальные последствия
Социально-инженерные атаки	Манипуляции людьми с целью получения конфиденциальной информации [22]	Утечка персональных данных, финансовые потери, нарушение работы систем. По данным ВЦИОМ, 67% респондентов в 2024 г. подвергались телефонному мошенничеству (в 2021 г. показатель составлял 57%) [23]
Фишинг	Использование поддельных электронных сообщений или веб-сайтов для кражи логинов, паролей и другой личной информации	Утечка данных, доступ к корпоративным сетям, мошенничество. Фишинг является самым распространенным видом атак на электронную почту в мире – 39,6% [24]
Мошенничество в интернете	Использование вредоносного программного обеспечения (ПО)	Финансовые потери. Наблюдается рост мошеннических сайтов в России: в 2022 г. – 111,2 тыс., в 2023 г. – 207,1 тыс. [25]
Уязвимости интернета вещей (IoT)	Недостатки ИБ в устройствах IoT могут быть использованы для удаленного доступа или распространения вредоносного ПО. Пути заражения IoT-устройств: перебор паролей и эксплуатация уязвимостей в сетевых сервисах	Нарушение ИБ корпоративных сетей, сбои в работе серверов. В первой половине 2023 г. 97,91% попыток перебора паролей были связаны с протоколом Telnet, 2,09% – с протоколом SSH. Больше всего инфицированных устройств, совершавших эти атаки, находилось в Китае (32,08%), Индии (12,13%) и США (5,7%) [26]

Таблица 1. Вызовы и риски в области ИБ

### Вызовы и риски информационной безопасности

ИБ затрагивает широкий спектр вопросов – от защиты персональных данных до обеспечения охраны государственных информационных систем. Это подчеркивает важность внедрения не только технических, но и стратегических решений на всех уровнях. В таблице 1 представлены актуальные вызовы и риски в области ИБ, которые являются актуальными и для современного периода, и для прогнозов на будущее.

Одна из главных угроз ИБ связана с повсеместным внедрением ИИ. Использование ИИ компаниями требует применения строгих протоколов защиты конфиденциальных данных. Согласно опросу специалистов в области ИБ, 85% кибератак в 2023 г. были совершены с помощью ИИ. Респонденты отметили, что ИИ делает предприятия более уязвимыми с точки зрения ИБ, чем до применения новых тех-

нологий [27]. Это свидетельствует о необходимости разработки комплексного подхода к ИБ, который включает внедрение передовых технологий шифрования и защиты данных и обеспечение ответственного отношения сотрудников к ИБ через регулярное обучение.

### Выводы

В мире цифровых технологий ИБ является приоритетным направлением. Киберугрозы становятся все более сложными и масштабными. Применение принципов «безопасности по умолчанию» и активное использование ИИ для предотвращения атак являются возможными направлениями усиления защиты. Адаптация к новым угрозам и разработка соответствующих методов защиты требует комплексного подхода, включающего применение современных методов, пересмотр образовательных программ для подготовки специалистов, а также обновление законодательных и нормативных баз.

## PROSPECTS FOR THE DEVELOPMENT OF INFORMATION SECURITY: GLOBAL CHALLENGES AND PROTECTION STRATEGIES

**Yakovishin Andrey**, master's degree Kamchatka State Technical University, Petropavlovsk-Kamchatsky, Russia. E-mail: ayakovishin@rambler.ru  
**Kuznetsov Ilya**, bachelor's degree, Udmurt State University, Izhevsk, Russia. E-mail: mr\_ilya.kuznetsov@rambler.ru

**Drozdov Igor**, bachelor's degree, Bauman Moscow State Technical University, Moscow, Russia. E-mail: drozdov\_i@mail.ru

**Pismenskiy Daniil**, bachelor's degree, Bauman Moscow State Technical University, Moscow, Russia. E-mail: dpismenskiy@mail.ru

**Abstract.** This article analyzes the current state and future trends in information security over the next decade. It focuses on threats from ransomware, cyberattacks on identity verification systems, and issues related to cryptocurrencies. The research includes an analysis of legislation and practices in information protection. The article emphasizes the importance of strengthening security measures and training specialists to effectively respond to cyber threats.

**Keywords:** information security, information security forecasts, cyber threats, ransomware, cyberattacks, information security legislation.

### Библиографический список

1. Cost of a Data Breach Report 2023 // IBM [Электронный ресурс]. – URL: <https://www.ibm.com/reports/data-breach> (дата обращения: 05.03.2024).
2. 2023 Official Cybercrime Report // eSentire [Электронный ресурс]. – URL: <https://www.esentire.com/resources/library/2023-official-cybercrime-report> (дата обращения: 05.03.2024).
3. Global Cybercrime Report: Which Countries Are Most at Risk in 2023 [Электронный ресурс]. – URL: [https://resources.cdn.seon.io/uploads/2023/04/Cybersecurity\\_countries-min.pdf](https://resources.cdn.seon.io/uploads/2023/04/Cybersecurity_countries-min.pdf) (дата обращения: 05.03.2024).
4. Annual share of organizations affected by ransomware attacks worldwide from 2018 to 2023 // Statista [Электронный ресурс]. – URL: <https://www.statista.com/statistics/204457/businesses-ransomware-attack-rate/> (дата обращения: 05.03.2024).
5. The Chainalysis 2024 Crypto Crime Report // Chainalysis [Электронный ресурс]. – URL: <https://www.chainalysis.com/blog/2024-crypto-crime-report-introduction/> (дата обращения: 07.03.2024).
6. Отчет «О значимых утечках данных в России» // Kaspersky Digital Footprint Intelligence [Электронный ресурс]. – URL: [https://www.kaspersky.ru/about/press-releases/2023\\_laboratoriya-kasperskogo-obuyom-utekshih-dannyh-vyros-v-2023-godu-na-33](https://www.kaspersky.ru/about/press-releases/2023_laboratoriya-kasperskogo-obuyom-utekshih-dannyh-vyros-v-2023-godu-na-33) (дата обращения: 08.03.2024).
7. Яковичин А. Д. Влияние международного законодательства на практики ИБ. Тенденции развития науки и образования. 2024;11:155–158.
8. Герасимов А. С. Основные проблемы информационной сетевой безопасности и варианты борьбы с ними. Актуальные исследования. 2022;40:119.
9. Security Awareness Training Market To Hit \$10 Billion Annually By 2027 // Cybersecurityventures [Электронный ресурс]. – URL: <https://cybersecurityventures.com/security-awareness-training-market-to-hit-10-billion-annually-by-2027/> (дата обращения: 10.03.2024).
10. Исрафилов А. Безопасность GSM-сетей: методы сканирования, возможные угрозы и стратегии защиты // Сборник статей международной научной конференции «Перспективные научные исследования как двигатель современной науки» МИПИ им. Ломоносова. 23 Марта 2024, Санкт-Петербург, Россия. С. 28.
11. iProov Threat Intelligence Report 2024: The Impact of Generative AI on Remote Identity Verification // iProov [Электронный ресурс]. – URL: <https://www.iproov.com/reports/iproov-threat-intelligence-report-2024> (дата обращения: 10.03.2024).
12. Number of Internet of Things (IoT) connected devices worldwide from 2019 to 2023, with forecasts from 2022 to 2030 // Statista [Электронный ресурс]. – URL: <https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/> (дата обращения: 10.03.2024).
13. Global Cybersecurity Index 2020 // International Telecommunication Union (ITU) [Электронный ресурс]. – URL: <https://www.itu.int/hub/publication/d-str-gci-01-2021/> (дата обращения: 10.03.2024).
14. Приходько Д. В., Шеров-Игнатьев В. Г. Цифровая экономика в Африке: состояние и проблемы развития // Вестник Санкт-Петербургского университета. Экономика. 2024;40(1):3–35. <https://doi.org/10.21638/spbu05.2024.101>
15. Quantum Computing Market Research Report – Global Forecast till 2032 // Marketresearchfuture [Электронный ресурс]. – URL: [https://www.marketresearchfuture.com/checkout?currency=one\\_user-USD&report\\_id=2583](https://www.marketresearchfuture.com/checkout?currency=one_user-USD&report_id=2583) (дата обращения: 10.03.2024).
16. Tiumentsev D. Application of cryptographic technologies for information protection in cloud services. Stolyin Annals. 2024;6(3).
17. Шайхулов Э. А., Смирнов А. П., Болдина О. Б., Азаренко Г. Ю., Благова И. Ю. Современные методы обучения информационной безопасности // Современная наука и инновации. 2023;4:145–151. <https://doi.org/10.37493/2307-910X.2023.4.16>
18. Council of Europe – Convention on Cybercrime (ETS No. 185) // Council of Europe [Электронный ресурс]. – URL: [https://www.europarl.europa.eu/meetdocs/2014\\_2019/documents/libe/dv/7\\_conv\\_budapest\\_17\\_conv\\_budapest\\_en.pdf](https://www.europarl.europa.eu/meetdocs/2014_2019/documents/libe/dv/7_conv_budapest_17_conv_budapest_en.pdf) (дата обращения: 10.03.2024).
19. Общий регламент по защите персональных данных GDPR (General Data Protection Regulation) // О GDPR на русском: Информация об общем регламенте по защите данных [Электронный ресурс]. – URL: <https://ogdpr.eu/ru> (дата обращения: 10.03.2024).
20. The California Consumer Privacy Act (Калифорнийский закон о защите прав потребителей) [Электронный ресурс]. – URL: <https://theccpa.org/> (дата обращения: 10.03.2024).
21. О пятой сессии Спецкомитета ООН по разработке всеобъемлющей конвенции о противодействии использованию информационно-коммуникационных технологий (ИКТ) в преступных целях // Министерство иностранных дел Российской Федерации [Электронный ресурс]. – URL: [https://www.mid.ru/ru/foreign\\_policy/news/1865216/](https://www.mid.ru/ru/foreign_policy/news/1865216/) (дата обращения: 10.04.2024).
22. Kostoreva A. Torts offenses: How to effectively defend clients' interests in court. Sciences of Europe. 2024;137:48–51. <https://doi.org/10.5281/zenodo.10874581>
23. Телефонное мошенничество: мониторинг // ВЦИОМ [Электронный ресурс]. – URL: <https://wciom.ru/analytical-reviews/analiticheskii-obzor/telefonnoe-moshennichestvo-monitoring> (дата обращения: 08.03.2024).
24. Cyber Security Report 2024: An In-depth Analysis of the Microsoft 365 Threat Landscape // Hornetsecurity [Электронный ресурс]. – URL: <https://www.hornetsecurity.com/en/cyber-security-report/> (дата обращения: 09.03.2024).
25. Давлетов А. Р. Главные трудности при интеграции машинного обучения в коммерческую эксплуатацию // Инновации и инвестиции. 2023;10:335–339.
26. Обзор угроз для IoT-устройств в 2023 г. // Kaspersky Digital Footprint Intelligence [Электронный ресурс]. – URL: <https://securelist.ru/iot-threat-report-2023/108088/> (дата обращения: 10.03.2024).
27. Generative AI and Cybersecurity: Bright Future or Business Battleground // Voice of SecOps 2023 [Электронный ресурс]. – URL: <https://www.deepinstinct.com/voice-of-secop-ops-reports> (дата обращения: 10.03.2024).

### Bibliography:

1. Cost of a Data Breach Report 2023 // IBM [Electronic resource]. – URL: <https://www.ibm.com/reports/data-breach> (access date: 03/05/2024).
2. 2023 Official Cybercrime Report // eSentire [Electronic resource]. – URL: <https://www.esentire.com/resources/library/2023-official-cybercrime-report> (access date: 03/05/2024).
3. Global Cybercrime Report: Which Countries Are Most at Risk in 2023 [Electronic resource]. – URL: [https://resources.cdn.seon.io/uploads/2023/04/Cybersecurity\\_countries-min.pdf](https://resources.cdn.seon.io/uploads/2023/04/Cybersecurity_countries-min.pdf) (access date: 03/05/2024).
4. Annual share of organizations affected by ransomware attacks worldwide from 2018 to 2023 // Statista [Electronic resource]. – URL: <https://www.statista.com/statistics/204457/businesses-ransomware-attack-rate/> (access date: 03/05/2024).
5. The Chainalysis 2024 Crypto Crime Report // Chainalysis [Electronic resource]. – URL: <https://www.chainalysis.com/blog/2024-crypto-crime-report-introduction/> (access date: 03/07/2024).
6. Report "On significant data leaks in Russia" // Kaspersky Digital Footprint Intelligence [Electronic resource]. – URL: [https://www.kaspersky.ru/about/press-releases/2023\\_laboratoriya-kasperskogo-obuyom-utekshih-dannyh-vyros-v-2023-godu-na-33](https://www.kaspersky.ru/about/press-releases/2023_laboratoriya-kasperskogo-obuyom-utekshih-dannyh-vyros-v-2023-godu-na-33) (access date: 03/08/2024).
7. Yakovishin A. D. The influence of international legislation on information security practices. Trends in the development of science and education. 2024;11:155–158.
8. Gerasimov A. S. Main problems of information network security and options for combating them. Current research. 2022;40:119.
9. Security Awareness Training Market To Hit \$10 Billion Annually By 2027 // Cybersecurityventures [Electronic resource]. – URL: <https://cybersecurityventures.com/security-awareness-training-market-to-hit-10-billion-annually-by-2027/> (access date: 03/10/2024).
10. Israfilov A. Security of GSM Networks: scanning methods, possible threats and protection strategies. Collection of articles of the international scientific conference "Advanced scientific research as the engine of modern science" MIPI named after. Lomonosov. March 23, 2024, St. Petersburg, Russia. P. 28.
11. iProov Threat Intelligence Report 2024: The Impact of Generative AI on Remote Identity Verification // iProov [Electronic resource]. – URL: <https://www.iproov.com/reports/iproov-threat-intelligence-report-2024> (access date: 03/10/2024).
12. Number of Internet of Things (IoT) connected devices worldwide from 2019 to 2023, with forecasts from 2022 to 2030 // Statista [Electronic resource]. – URL: <https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/> (access date: 03/10/2024).
13. Global Cybersecurity Index 2020 // International Telecommunication Union (ITU) [Electronic resource]. – URL: <https://www.itu.int/hub/publication/d-str-gci-01-2021/> (access date: 03/10/2024).
14. Prikhodko D. V., Sherov-Ignatiev V. G. Digital economy in Africa: status and development problems. Bulletin of St. Petersburg University. Economy. 2024;40(1):3–35. <https://doi.org/10.21638/spbu05.2024.101>
15. Quantum Computing Market Research Report – Global Forecast till 2032 // Marketresearchfuture [Electronic resource]. – URL: [https://www.marketresearchfuture.com/checkout?currency=one\\_user-USD&report\\_id=2583](https://www.marketresearchfuture.com/checkout?currency=one_user-USD&report_id=2583) (access date: 03/10/2024).
16. Tiumentsev D. Application of cryptographic technologies for information protection in cloud services. Stolyin Annals. 2024;6(3).
17. Shaikhulov E. A., Smirnov A. P., Boldina O. B., Azarenko G. Yu., Blagova I. Yu. Modern methods of information security training. Modern science and innovation. 2023;4:145–151. <https://doi.org/10.37493/2307-910X.2023.4.16>
18. Council of Europe – Convention on Cybercrime (ETS No. 185) / Council of Europe [Electronic resource]. – URL: [https://www.europarl.europa.eu/meetdocs/2014\\_2019/documents/libe/dv/7\\_conv\\_budapest\\_17\\_conv\\_budapest\\_en.pdf](https://www.europarl.europa.eu/meetdocs/2014_2019/documents/libe/dv/7_conv_budapest_17_conv_budapest_en.pdf) (access date: 03/10/2024).
19. General Data Protection Regulation GDPR (General Data Protection Regulation) // About GDPR in Russian: Information on the General Data Protection Regulation [Electronic resource]. – URL: <https://ogdpr.eu/ru> (access date: 03/10/2024).
20. The California Consumer Privacy Act [Electronic resource]. – URL: <https://theccpa.org/> (access date: 03/10/2024).
21. On the fifth session of the UN Ad Hoc Committee on the development of a comprehensive convention on countering the use of information and communication technologies (ICT) for criminal purposes // Ministry of Foreign Affairs of the Russian Federation [Electronic resource]. – URL: [https://www.mid.ru/ru/foreign\\_policy/news/1865216/](https://www.mid.ru/ru/foreign_policy/news/1865216/) (date of access: 04/10/2024).
22. Kostoreva A. Torts offenses: How to effectively defend clients' interests in court. Sciences of Europe. 2024;137:48–51. <https://doi.org/10.5281/zenodo.10874581>
23. Telephone fraud: monitoring // VTsIOM [Electronic resource]. – URL: <https://wciom.ru/analytical-reviews/analiticheskii-obzor/telefonnoe-moshennichestvo-monitoring> (date of access: 03/08/2024).
24. Cyber Security Report 2024: An In-depth Analysis of the Microsoft 365 Threat Landscape // Hornetsecurity [Electronic resource]. – URL: <https://www.hornetsecurity.com/en/cyber-security-report/> (access date: 03/09/2024).
25. Davletov A. R. The main challenges in integrating machine learning into commercial operations. Innovation and investment. 2023;10:335–339.
26. Review of threats to IoT devices in 2023 // Kaspersky Digital Footprint Intelligence [Electronic resource]. – URL: <https://securelist.ru/iot-threat-report-2023/108088/> (access date: 03/10/2024).
27. Generative AI and Cybersecurity: Bright Future or Business Battleground // Voice of SecOps 2023 [Electronic resource]. – URL: <https://www.deepinstinct.com/voice-of-secop-ops-reports> (access date: 03/10/2024).



# ИНФОРМАЦИОННЫЕ РЕСУРСЫ РОССИИ

УЧРЕДИТЕЛЬ ЖУРНАЛА

**РЭА** МИНЭНЕРГО  
РОССИИ

12+



irr@rosenergo.gov.ru

**ISSN 0204-3653**

Свидетельство о регистрации СМИ ПИ № 77-12208 от 29 марта 2002 г.  
Учредитель и издатель ФГБУ «РЭА» Минэнерго России  
Тираж до 500 шт.  
Периодичность выхода 6 раз в год

Журнал включен в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук

Руководитель научно-редакционного совета – д. т. н. директор Пермского Центра научно-технической информации-филиала «РЭА» Минэнерго России **Александр Трусов**

#### Научно-редакционный совет

**Лобанов И. В.** – к. ю. н., ректор РЭУ им. Г.В. Плеханова, **Бирман Н. Я.** – к. т. н., профессор, библиотекарь Information Center of Green library at Stanford University, USA; **Гуриев М. А.** – д. т. н. профессор, директор по работе с гос. учреждениями Samsung Electronics in CIS; **Дзегеленок И. И.** – д. т. н., профессор НИУ «МЭИ»; **Каленов Н. Е.** – д. т. н., профессор, главный научный сотрудник МСЦ РАН; **Коллин К. К.** – д. т. н., профессор, главный научный сотрудник ИПИ РАН, заслуженный деятель науки РФ, действительный член Международной академии наук (Инсбрук, Австрия), Российской академии естественных наук и Международной академии наук высшей школы; **Левнер Е. В.** – доктор философии, профессор, Университет Бар-Илан (Bar-Ilan University), г. Рамат Ган (Израиль) и Ашкелонский Академический Колледж, г. Ашкелон (Израиль); **Подлесный С. А.** – к. т. н., профессор, советник ректора, «Сибирский федеральный университет», заслуженный работник высшей школы РФ; **Сотников А. Н.** – д. ф.-м. н., профессор, заслуженный деятель науки РФ, заместитель директора МСЦ РАН; **Трусов А. В.** – д. т. н., директор Пермского Центра научно-технической информации – филиала «РЭА» Минэнерго России, **Цветкова В. А.** – д. т. н., профессор кафедры библиотечно-информационных наук МГИК, **Антопольский А. Б.** – д. т. н., профессор, главный научный сотрудник ИНИОН РАН, **Лопатина Н. В.** – д. п. н., заведующий кафедрой библиотечно-информационных наук, Московский государственный институт культуры, ведущий научный сотрудник Федерального института промышленной собственности Роспатента, **Поляк Ю. Е.** – ведущий научный сотрудник, Центральный экономико-математический институт РАН

При использовании материалов ссылка на журнал обязательна.

Перепечатка материалов возможна только с письменного разрешения редакции.

Позиция и мнение авторов статей может не совпадать с мнением редакции.

#### Специальности ВАК:

05.13.17 – Теоретические основы информатики

(технические науки),

05.25.05 – Информационные системы и процессы (технические науки)

2.3.1 – Системный анализ, управление и обработка информации (физико-математические науки)

#### Адрес и контакты:

127083, г. Москва,

улица 8 марта, д. 12

#### Главный редактор журнала ИРР

**Анна Горшкова**

Телефон: +7 910 463-53-57

E-mail: anna.gorshik@yandex.ru,

gorshkova@rosenergo.gov.ru

#### Заместитель главного редактора

**по подписке, распространению**

**и продвижению журнала «ИРР»**

**Виолетта Локтева**

Телефон: +7 903 733-72-57

E-mail: Lokteva@rosenergo.gov.ru

#### Scientific Editorial Board

**Lobanov I.** – PhD in Law, Rector of the Russian University of Economics G.V. Plekhanov, **Birman N.** – Ph. D., Professor, librarian Information Center of Green library at Stanford University, USA; **Guriev M.** – Grand Ph. D. in Engineering, Professor, Director of work with state institutions Samsung Electronics in CIS; **Dzegelenok I.** – Grand Ph. D. in Engineering, Professor of National Research University "MPEI"; **Kalenov N.** – Grand Ph. D. in Engineering, Professor, Chief Researcher of MSC RAS; **Colin K.** – Grand Ph. D. in Engineering, Professor, Chief Researcher of the IPI RAS, Honored Worker of Science of the Russian Federation, full member of the International Academy Sciences (Innsbruck, Austria), Russian Academy of Natural Sciences and the International Academy of Sciences of Higher Education; **Levner E.** – Ph. D., Professor, Bar-Ilan University (Bar-Ilan University), Ramat Gan (Israel) and Ashkelon Academic College, Ashkelon (Israel); **Podlesny S.** – Ph. D., Professor, Adviser to the rector, "Siberian Federal University", Honored Worker of the Higher School of the Russian Federation; **Sotnikov A.** – Dr. Sc. (Phys.-Math.), Professor, Honored Worker of Science of the Russian Federation, Deputy Director of the ISC RAS; **Trusov A.** – D.Sc, Associate Professor, Director of the PermCenter for Scientific and Technical Information (TSNTI) – branch of "REA" Ministry of Energy of Russia; **Tsvetkova V.** – Grand Ph. D. in Engineering, Professor, Department Informatization of culture and electronic libraries of the Moscow State Institute of Culture and Arts; **Antopolsky A.** – Grand Ph. D. in Engineering, Professor, Chief Researcher of INION RAS; **Lopatina N.** – Ph. D., Head of the Department of Library and Information Sciences, Moscow State Institute of Culture, Leading Researcher, Federal Institute of Industrial Property of Rospatent; **Polyak Y.** – Leading Researcher, Central Economics and Mathematics Institute of the Russian Academy of Sciences

