

ЦИФРОВЫЕ РЕШЕНИЯ И ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Научно-практический рецензируемый журнал
Издается с 2025 г.

**Издатель: Финансовый университет
при Правительстве Российской Федерации,
Москва, Российская Федерация**

125167, Москва, Ленинградский пр-т, д. 53

Периодичность издания — 4 номера в год

Журнал ориентирован на научное обсуждение актуальных проблем
по научным специальностям:

- 1.2.1. «Искусственный интеллект и машинное обучение»
- 1.2.2. «Математическое моделирование, численные методы и комплексы программ»
- 2.3.6. «Методы и системы защиты информации, информационная безопасность»
- 5.2.2. «Математические, статистические и инструментальные методы в экономике»

Электронная версия журнала
находится в открытом доступе на сайте digitari.ru
Журнал публикует материалы на условиях лицензии
Creative Commons Attribution 4.0 International (CC BY 4.0).

DIGITAL SOLUTIONS AND ARTIFICIAL INTELLIGENCE TECHNOLOGIES

Scientific and practical peer-reviewed journal
Published since 2025

ЦИФРОВЫЕ РЕШЕНИЯ И ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Научно-практический журнал
Том 2, № 3, 2026

Издается с 2025 года

Периодичность: 4 раза в год

ISSN 3033-5477 (Print)

ISSN 3033-7097 (Online)

Свидетельство о регистрации:

Эл № ФС77-90338

от 01 ноября 2025 г.

Учредитель и издатель:

Финансовый университет
при Правительстве Российской
Федерации

125167, Российская Федерация,
Москва,

Ленинградский проспект, 49

+7 (499) 700-86-00 (доб. 1083)

<https://www.digitarin.ru/jour-aaizvekova@fa.ru>

Главный редактор

Феклин

Вадим Геннадьевич

«Цифровые решения
и технологии искусственного
интеллекта» — платформа
для публикации оригинальных
и актуальных исследований,
посвященных современным
технологиям в сфере
искусственного интеллекта
и цифровых решений.

Заведующий редакцией

В.И. Головатенко

Выпускающий редактор

А.А. Извекова

Переводчик **И.М. Евстратов**

Корректор **С.Ф. Михайлова**

Верстальщик **Е.А. Смирнова**

Отпечатано в ООО «Коми
Республиканская типография»,
г. Сыктывкар, ул. Савина, д. 81.

Тираж 500 экз.

(I завод — 40 экз.).

Формат 60 x 84 1/8.

Заказ № 26-4861

Подписано в печать:

24.08.2026

Выход в свет 24.09.2026

© Финансовый университет
при Правительстве Российской
Федерации, Москва

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И МАШИННОЕ ОБУЧЕНИЕ

Бородина А.Д., Кызыл-оол М.В., Кочкаров Р.А.

Формирование объединенного алфавита
древнетюркских диалектов для аннотации изображений. . . 6

Шелепаева А.Х.

Применение искусственного интеллекта
в сельском хозяйстве: мировая практика. 16

МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ, ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

Баранова Е.К., Лебедкина Я.В.

Обезличивание персональных данных
в банковской сфере: анализ методов
и выбор стратегии внедрения 26

Иванов А.В., Антулина В.А.

Организационно-технологический механизм безопасной
интеграции инфраструктуры доверия цифрового рубля
в экосистему кредитно-финансовой организации. 34

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, ЧИСЛЕННЫЕ МЕТОДЫ И КОМПЛЕКСЫ ПРОГРАММ

Tulupova O.P., Zholobova G.N., Enikeev F.U.

Data Engineering and Machine Learning for Solving
the Inverse Problem Based on CAE Simulation Results 41

Баротов Р.Н., Баротов Д.Н.

О количестве натуральных решений специфического
дискретного уравнения и его свойствах 51

МАТЕМАТИЧЕСКИЕ, СТАТИСТИЧЕСКИЕ И ИНСТРУМЕНТАЛЬНЫЕ МЕТОДЫ В ЭКОНОМИКЕ

Гречушкина Д.С., Гринева Н.В., Голосов П.Е.

Интеллектуальные системы контроля как инструмент
повышения энергоэффективности в энергосервисе 58

Борисова Л.Р., Фридман М.Н.

Анализ некоторых аспектов цифровизации в области
оказания образовательных услуг населению методами
статистического анализа и машинного обучения. 67

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Borodina A.D., Kyzyl-ool M.V., Kochkarov R.A.

Formation of a Unified Alphabet
of Ancient Turkic Dialects for Image Annotation. 6

Shelepaeva A.Kh.

The Use of Artificial Intelligence
in Agriculture: Global Practices 16

METHODS AND SYSTEMS OF INFORMATION PROTECTION, INFORMATION SECURITY

Baranova E.K., Lebedkina Ya.V.

Depersonalization of Personal Data in the Banking
Sector: Analysis of Methods and Selection
of an Implementation Strategy 26

Ivanov A.V., Aptulina V.A.

An Organizational and Technological Mechanism for the Secure
Integration of the Digital Ruble's Trust Infrastructure
into the Ecosystem of a Credit and Financial Institution. 34

MATHEMATICAL MODELING, NUMERICAL METHODS AND SOFTWARE PACKAGES

Tulupova O.P., Zholobova G.N., Enikeev F.U.

Data Engineering and Machine Learning for Solving
the Inverse Problem Based on CAE Simulation Results 41

Barotov R.N., Barotov D.N.

On the Number of Natural Solutions
of a Specific Discrete Equation and Its Properties 51

MATHEMATICAL, STATISTICAL AND INSTRUMENTAL METHODS IN ECONOMICS

Grechushkina D.S., Grineva N.V., Golosov P.E.

Application of Intelligent Energy Consumption
Control Systems in Energy Service Projects 58

Borisova L.R., Fridman M.N.

Analysis of Certain Aspects of Digitalization in the Field
of Providing Educational Services to the Population Using
Statistical Analysis and Machine Learning Methods. 67

DIGITAL SOLUTIONS AND ARTIFICIAL INTELLIGENCE TECHNOLOGIES

Scientific and practical journal

Vol. 2, No. 3, 2026

Published Since 2025

Publication Frequency:

4 times a year

ISSN 3033-5477 (Print)

ISSN 3033-7097 (Online)

Founder and Publisher:

Financial University

under the Government
of the Russian Federation

125167, Russian Federation,

Moscow,

49, Leningradsky Avenue

+7 (499) 700-86-00 (доб. 1083)

<https://www.digitarin.ru/jour>

aaizvekova@fa.ru

Digital Solutions and
Artificial Intelligence

Technologies is a platform
for publishing original and
relevant research on modern

technologies in the field
of artificial intelligence

and digital solutions.

Editor-in-Chief

Vadim G. Feklin

Head of Scientific Journals

Editorial Department

V.I. Golovatenko

Managing Editor

A.A. Izvekova

Translator

I.M. Evstratov

Proofreader

S.F. Mikhaylova

Design, make up

E.A. Smirnova

Printed in Komi Republican

Printing House LLC,

Syktyvkar, Savina St. 81

The edition is 500 copies.

(I factory – 40 copies).

Format is 60 x 84 1/8.

Order № 26-4861

Signed off to printing:

24.08.2026

Publication date 24.09.2026

© Financial University

under the Government

of the Russian Federation,

Moscow

ГЛАВНЫЙ РЕДАКТОР

ФЕКЛИН В.Г., кандидат физико-математических наук, доцент, декан факультета информационных технологий и анализа больших данных, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА

КОЧКАРОВ Р.А., доктор технических наук, доцент кафедры искусственного интеллекта, заместитель декана по научной работе факультета информационных технологий и анализа больших данных, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

ЧЛЕНЫ РЕДАКЦИОННОЙ КОЛЛЕГИИ

АНДРИЯНОВ Н.А., кандидат технических наук, доцент, доцент кафедры искусственного интеллекта, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

АФНАСЬЕВ А.А., доктор экономических наук, доцент, ведущий научный сотрудник ЦЭМИ РАН, заведующий кафедрой моделирования и системного анализа, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

БОГДАНОВ Е.А., Ph.D, заведующий кафедрой информационной безопасности, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

ВАСИЛЬЕВА Е.В., доктор экономических наук, доцент, заведующая кафедрой бизнес-информатики, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

ЖУКОВСКАЯ И.Е., доктор экономических наук, доцент, профессор кафедры бизнес-информатики, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

КОРОВИН Д.И., доктор экономических наук, доцент, профессор кафедры моделирования и системного анализа, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

КАШИРИН А.Б., директор Центра продвинутой аналитики АО «Альфа-Банк», заведующий базовой кафедрой Альфа-Банка, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

КОЗЬМИНЫХ С.И., доктор технических наук, доцент, профессор кафедры информационной безопасности, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

КОРОТЕЕВ М.В., кандидат экономических наук, доцент, заведующий кафедрой искусственного интеллекта, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

КОЧКАРОВ А.А., доктор технических наук, доцент, заместитель директора по инновационной работе ФИЦ Биотехнологии РАН, профессор кафедры искусственного интеллекта, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

МАКАРОВ В.Л., доктор физико-математических наук, профессор, академик РАН, научный руководитель ЦЭМИ РАН, научный руководитель факультета информационных технологий и анализа больших данных, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

МЕЛЬНИКОВ Д.А., доктор технических наук, доцент, профессор кафедры информационной безопасности, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

МИХАЙЛОВА С.С., доктор экономических наук, заведующая кафедрой математики и анализа данных, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

НЕИЗВЕСТНЫЙ С.И., доктор технических наук, старший научный сотрудник, профессор кафедры бизнес-информатики, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

ОСТАПЕНКО Г.А., доктор технических наук, профессор, проректор по цифровизации, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

ПЕТРОВСОВ Д.А., кандидат технических наук, доцент, заведующий кафедрой информационных технологий, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

ПРОКОПЧИНА С.В., доктор технических наук, профессор, профессор кафедры искусственного интеллекта, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

РЯБОВ П.Е., доктор физико-математических наук, доцент, профессор кафедры математики и анализа данных, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

СЛАВИН Б.Б., доктор экономических наук, профессор кафедры бизнес-информатики, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

ТИМОШЕНКО А.В., доктор технических наук, профессор, профессор кафедры информационной безопасности, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

ЦАРЕГОРОДЦЕВ А.В., доктор технических наук, профессор, главный научный сотрудник Института цифровых технологий, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

ЩЕТИНИН Е.Ю., доктор физико-математических наук, доцент, профессор кафедры искусственного интеллекта, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

POURMOHAMMADBAGHER L., Dr., Allameh Tabataba'i University (ATU), Iran

BAHRANI M., Dr., Allameh Tabataba'i University (ATU), Iran

PEYMANI M., Dr., Allameh Tabataba'i University (ATU), Iran

SHARMA CH., Prof., GNIOT Institute Of Management Studies, Greater Noida, India

JUNSHENG ZH., Prof., Dalian Neusoft University of Information, China

LIANZHUANG Q., PhD, Dalian Neusoft University of Information, China

KANG L., Prof., Dalian Neusoft University of Information, China

EDITOR-IN-CHIEF

FEKLIN V.G., Cand. Sci. (Phys.-Math.), Associate Professor, Dean of the Faculty of Information Technology and Big Data Analysis, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

DEPUTY EDITOR-IN-CHIEF

KOCHKAROV R.A., Dr. Sci. (Tech.), Associate Professor, Department of Artificial Intelligence, Deputy Dean for Research at the Faculty of Information Technology and Big Data Analysis, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

MEMBERS OF THE EDITORIAL BOARD

ANDRIANOV N.A., Cand. Sci. (Tech.), Associate Professor, Associate Professor of Artificial Intelligence, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

AFANASYEV A.A., Dr. Sci. (Econ.), Associate Professor, Leading Researcher at CEMI RAS, Head of the Department of Modeling and System Analysis, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

BOGDANOV E.A., PhD, Head of the Department of Information Security, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

VASILYEVA E.V., Dr. Sci. (Econ.), Associate Professor, Head of the Department of Business Informatics, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

ZHUKOVSKAYA I.E., Dr. Sci. (Econ.), Associate Professor, Professor, Department of Business Informatics, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

KOROVIN D.I., Dr. Sci. (Econ.), Associate Professor, Professor of the Department of Modeling and System Analysis, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

KASHIRIN A.B., Director of the Center for Advanced Analytics of Alfa-Bank JSC, Head of the Basic Department of Alfa-Bank, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

KOZMINYKH S.I., Dr. Sci. (Tech.), Associate Professor, Professor, Department of Information Security, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

KOROTEEV M.V., Cand. Sci. (Econ.), Associate Professor, Head of the Department of Artificial Intelligence, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

KOCHKAROV A.A., Dr. Sci. (Tech.), Associate Professor, Deputy Director for Innovation at the Institute of Biotechnology of the Russian Academy of Sciences, Professor, Department of Artificial Intelligence, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

MAKAROV V.L., Dr. Sci. (Phys.-Math.), Professor, Academician of the RAS, Scientific Director of the Central Research Institute of the Russian Academy of Sciences, Scientific Director of the Faculty of Information Technology and Big Data Analysis, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

MELNIKOV D.A., Dr. Sci. (Tech.), Associate Professor, Professor, Department of Information Security, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

MIKHAILOVA S.S., Dr. Sci. (Econ.), Head of the Department of Mathematics and Data Analysis, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

NEIZVESTNYI S.I., Dr. Sci. (Tech.), Senior Researcher, Professor, Department of Business Informatics, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

OSTAPENKO G.A., Dr. Sci. (Tech.), Professor, Vice-Rector for Digitalization, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

PETROSOV D.A., Cand. Sci. (Tech.), Associate Professor, Head of the Department of Information Technology, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

PROKOPCHINA S.V., Dr. Sci. (Tech.), Professor, Professor, Department of Artificial Intelligence, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

RYABOV P.E., Dr. Sci. (Phys.-Math.), Associate Professor, Professor of the Department of Mathematics and Data Analysis, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

SLAVIN B.B., Dr. Sci. (Econ.), Professor, Department of Business Informatics, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

TIMOSHENKO A.V., Dr. Sci. (Tech.), Professor, Professor, Department of Information Security, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

TSAREGORODTSEV A.V., Dr. Sci. (Tech.), Professor, Chief Researcher at the Institute of Digital Technologies, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

SHCHETININ E.Y., Dr. Sci. (Phys.-Math.), Associate Professor, Professor, Department of Artificial Intelligence, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

POURMOHAMMADBAGHER L., Dr., Allameh Tabataba'i University (ATU), Iran

BAHRANI M., Dr., Allameh Tabataba'i University (ATU), Iran

PEYMANI M., Dr., Allameh Tabataba'i University (ATU), Iran

SHARMA CH., Prof., GNIOT Institute Of Management Studies, Greater Noida, India

JUNSHENG ZH., Prof., Dalian Neusoft University of Information, China

LIANZHUANG Q., PhD, Dalian Neusoft University of Information, China

KANG L., Prof., Dalian Neusoft University of Information, China

DOI: 10.26794/3030-7097-2026-2-3-6-15
УДК 621.382.323(045)

Формирование объединенного алфавита древнетюркских диалектов для аннотации изображений

А.Д. Бородина, М.В. Кызыл-оол, Р.А. Кочкаров

Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

В статье представлена разработка объединенного алфавита древнетюркских диалектов для решения задач автоматизированного распознавания орхон-енисейских рунических надписей. Актуальность исследования обусловлена разрозненностью существующих классификаций рун и отсутствием единой системы их сопоставления, что затрудняет формирование совместимых наборов данных и снижает эффективность моделей компьютерного зрения. **Цель исследования** заключается в создании унифицированного алфавита, обеспечивающего корректную аннотацию изображений и возможность обучения многоалфавитных нейросетевых моделей для распознавания и сравнительного анализа рунических символов. **Методология** исследования базируется на комплексном анализе визуальных, фонетических и историко-филологических характеристик рун четырех источников: памятника Кюль-Тегину, рукописи Ирк Битиг, орхонских и енисейских надписей. В ходе работы выделены структурно устойчивые символы, стилиевые вариации и уникальные маркеры диалектов, на основе чего сформирован консолидированный набор классов рун. **Полученные результаты** создают основу для стандартизации разметки изображений, повышения устойчивости распознавания и дальнейшего развития многоалфавитных систем компьютерного зрения в исследованиях древнетюркской письменности.

Ключевые слова: орхон-енисейская письменность; руническая письменность; древнетюркский язык; систематизация алфавитов; распознавание символов; компьютерное зрение; искусственный интеллект; многоалфавитная модель

Для цитирования: Бородина А.Д., Кызыл-оол М.В., Кочкаров Р.А. Формирование объединенного алфавита древнетюркских диалектов для аннотации изображений. *Цифровые решения и технологии искусственного интеллекта*. 2026;2(3):6-15. DOI: 10.26794/3030-7097-2026-2-3-6-15

ORIGINAL PAPER

Formation of a Unified Alphabet of Ancient Turkic Dialects for Image Annotation

A.D. Borodina, M.V. Kyzyl-ool, R.A. Kochkarov

Financial University under the Government of the Russian Federation, Moscow, Russian Federation

ABSTRACT

This paper is devoted to the development of a unified alphabet of Ancient Turkic dialects for solving the problems of automated recognition of Orkhon–Yenisei runic inscriptions. The relevance of the study is determined by the fragmentation of existing rune classifications and the absence of a unified system for correlating symbols across different sources and regions, which complicates the formation of compatible datasets and reduces the efficiency of computer vision models. **The research** is to create a unified alphabet that ensures correct image annotation and enables the training of multi-alphabet neural network models for the recognition and comparative analysis of runic symbols. **The methodology** is based on a comprehensive analysis of visual, phonetic, and historical-philological characteristics of runes from four sources: the Kül Tegin monument, the Irk Bitig manuscript, Orkhon inscriptions, and Yenisei inscriptions. During the study, structurally stable symbols, stylistic variations, and unique dialect markers were identified, on the basis of which a consolidated set of rune classes was formed. **The obtained results** provide a foundation for the standardization of image annotation, improvement of recognition robustness, and further development of multi-alphabet computer vision systems in the study of Ancient Turkic writing.

Keywords: Orkhon-Yenisei script; runic writing; Ancient Turkic language; alphabet systematization; symbol recognition; computer vision; artificial intelligence; multi-alphabet model

For citation: Borodina A.D., Kyzyl-ool M.V., Kochkarov R.A. Formation of a unified alphabet of ancient Turkic Dialects for Image Annotation. *Digital Solutions and Artificial Intelligence Technologies*. 2026;2(3):6-15. DOI: 10.26794/3030-7097-2026-2-3-6-15



ВВЕДЕНИЕ

Рунические надписи Орхоно-Енисейского региона представляют собой уникальное культурное наследие древнетюркских народов. Большинство из них высечены на камнях, а тексты посвящены памятным событиям или важным людям. Также сохранилась и рукописная «Гадательная книга» (Ирк Битиг), содержащая поэтические миниатюры и выводы к ним. Эти памятники, возникшие в VI–X вв., являются важнейшими источниками для изучения истории, языка и культуры древних тюрков¹.

История изучения орхоно-енисейских рунических письменностей берет свое начало с работы датского лингвиста Вильгельма Томсена, впервые расшифровавшего орхонские надписи в 1893 г. Это открытие стало основой для дальнейших исследований и положило начало систематическому изучению этого вида письменности. В XX в. значительный вклад в эту область внес С.Е. Малов, чьи труды способствовали глубокому пониманию структуры и значения текстов, написанных этими письменностями [1]. Современные исследования продолжают развивать тему, применяя технологии машинного обучения². Однако, несмотря на значимость темы, исследование рунических надписей сталкивается с рядом проблем. Одной из ключевых является разрозненность существующих классификаций и отсутствие единой схемы систематизации [2, 3].

В последние годы интерес к орхоно-енисейским письменностям усилился в рамках цифровых гуманитарных исследований, где предпринимаются попытки автоматизации анализа памятников с использованием методов компьютерного зрения и машинного обучения³ [4, 5]. Однако большинство существующих работ опираются на локальные классификации символов, что ограничивает возможность объединения данных и переноса моделей между различными корпусами [6].

Современные методы компьютерного зрения и глубокого обучения позволяют решать задачи автоматического распознавания древних письменностей с высокой точностью, однако их эффективность напрямую зависит от качества и согласованности исходных данных⁴ [7]. В случае орхоно-енисейских рунических памятников основной проблемой остается отсутствие единой системы классификации символов. Такая разрозненность приводит к ряду последствий:

- во-первых, невозможно объединить существующие наборы данных для обучения единой модели, так как один и тот же символ в разных классификациях может иметь разные обозначения;
- во-вторых, модели компьютерного зрения, обученные на одном наборе, теряют обобщающую способность при применении к надписям из других регионов или источников;
- в-третьих, возникает риск ошибочной интерпретации редких или региональных вариантов начертания, что снижает точность распознавания.

Ранее нами была предложена система классификации на основе текста с памятника Кюль-Тегину⁵ [8]. Однако данный подход оказывается неприменимым в случаях, когда в надписях присутствуют руны, отсутствующие в исходной классификации. Для решения данной проблемы все символы, не вошедшие в нашу классификацию, отмечались одним классом. В результате в модель подавались как один класс руны с абсолютно разным начертанием. Это негативно сказывалось на процессе обучения и приводило к снижению валидационной точности модели [9].

Научная новизна настоящего исследования заключается в формировании объединенного алфавита, ориентированного на прикладные задачи аннотации изображений и обучения многоалфавитных моделей компьютерного зрения, а не на традиционную лингвистическую классификацию. Объединение разрозненных классификаций в единый алфавит позволяет решить эти проблемы. Единая система символов создает универсальную основу для аннотации изображений, что делает возможным формирование крупного, сбалансированного корпуса данных. Это, в свою очередь, обеспечивает стабильность и воспроизводимость обучения моделей, упрощает перенос знаний между различными источниками и позволяет строить многоалфавит-

¹ Кормушин И.В. Древние тюркские языки. Учебное пособие. М.: Восточная литература; 2006. 30 с. URL: <https://www.elibrary.ru/qrijsof>

² См., например: Пойкалайнен А.М., Сердюк М.С., Таратин А.А., Кочкаров Р.А. Распознавание тюркских рунических надписей на примере Ирк Битиг. 2024. URL: <https://t.me/Digitalisimus>; Исследование по цифровой обработке гуманитарных данных. URL: <https://core.ac.uk/download/pdf/74335732.pdf>; Войников А. Цифровые методы анализа письменных источников. URL: <http://biblio.darial-online.ru/text/Voinikov/06.pdf>; Куприянова Н. Применение цифровых методов в гуманитарных исследованиях. URL: <https://elib.sfu-kras.ru/bitstream/handle/2311/156100/kupriyanova.pdf?sequence=1&isAllowed=y>

³ Кочкаров Р.А., Бородина А.Д., Виноградова В.В. и др. Классификация и разметка символов рунических надписей Евразии методами машинного обучения. Отчет о НИР/НИОКР (итоговый). М.; 2025. URL: <https://www.elibrary.ru/upauqb>

⁴ Распознавание орхоно-енисейских рунических надписей методами машинного обучения. URL: <https://habr.com/ru/companies/fa/articles/907950/?ysclid=ms64e72r1540015961>

⁵ См. также: Задача озвучивания орхоно-енисейских рунических надписей. URL: <https://habr.com/ru/companies/fa/articles/1028000/?ysclid=ms648hbj357696357>

ные нейросети, способные учитывать региональные и временные особенности письменности [10].

В данном контексте целью настоящего исследования является создание объединенного алфавита четырех наречий древнетюркского языка для обучения многоалфавитной модели компьютерного зрения. Достижение этой цели позволит не только разметить изображения рун, но и в дальнейшем обучить модель компьютерного зрения для распознавания символов всех включенных диалектов.

ОБЗОР СУЩЕСТВУЮЩИХ КЛАССИФИКАЦИЙ

Существующие классификации орхоно-енисейских письменностей представляют собой множество систем, разработанных различными исследователями в течение XX в. Их количество превышает двадцать, и они часто противоречат друг другу, что создает значительные трудности для унификации данных и их дальнейшего анализа. Разрозненность классификаций приводит к путанице, усложняет процесс сопоставления данных и затрудняет междисциплинарное сотрудничество [2, 3]. Для решения этой проблемы в своих исследованиях мы решили сформировать рунические алфавиты из источников, имеющих точную историческую и географическую привязку: памятник Кюль-Тегину, Ирк Битиг («Гадательная книга»), а также орхонские и енисейские памятники. Разделив надписи с памятников на символы, мы создали наборы уникальных символов для каждого источника отдельно. Таким образом получилось четыре алфавита, которые далее будут называться Кюль-Тегин, Ирк Битиг, орхонский и енисейский [11–14]. Они отличаются стилем написания, но можно заметить идентичные или похожие руны (рис. 1).

Ирк Битиг, как и орхоно-енисейская письменность, представляет собой важный этап в развитии тюркской письменности. Книга была создана в VII–VIII вв., что делает ее современницей памятника Кюль-Тегину, который был установлен в 732 г. Тем не менее начертание рун в Ирк Битиг значительно отличается от остальных трех алфавитов. Отчасти причина этого может быть в разных способах написания — символы Кюль-Тегина выбиты на камне, когда Ирк Битиг был написан на бумаге. Отличия особо заметны на одинаковых рунах. На рис. 1 классы 006, 106, 206 и 306 объективно обозначают одну руну. При этом самые сильные отличия в начертании наблюдаются у Ирк Битиг.

Рунический состав алфавитов тоже различается. В каждом алфавите есть уникальные символы, не имеющие визуальных аналогов. Енисейская классификация самая многочисленная и состоит из 41 руны. Оставшиеся три алфавита содержат 38 символов.

Ирк Битиг	Енисейский	Орхонский	Кюль-Тегин
001 	101 	201 	301 
002 			302 
003 	103 	203 	303 
004 	104 	204 	304 
005 	105 	205 	305 
006 	106 	206 	306 
007 	107 	207 	307 

Рис. 1 / Fig. 1. Фрагмент таблицы письменностей диалектов / Fragment of the Table of Dialects' Scripts

Источник / Source: составлено авторами / Compiled by the authors.

Все вышеперечисленное затрудняет объединение этих диалектов в общую систему классификации. Одна и та же руна может обозначаться по-разному, а некоторые символы отсутствуют в отдельных системах. Без единого алфавита символов невозможно сформировать крупные обучающие выборки и проводить сравнение результатов между исследованиями⁶ [15, 16].

МЕТОДИКА ОБЪЕДИНЕНИЯ АЛФАВИТОВ

Цель создания объединенного алфавита — разметить наибольшее возможное количество символов на изображениях с руническими надписями. Следовательно, это будет прикладной инструмент машинного зрения, а не лингвистическая классификация. Алфавит должен включить не только все уникальные руны из четырех диалектов, но

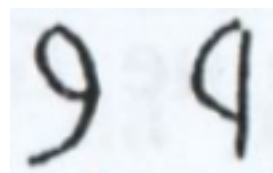


Рис. 2 / Fig. 2. Руны орхонского и енисейского наречия / Runes of the Orkhon and Yenisei Dialects

Источник / Source: составлено авторами / Compiled by the authors.

⁶ Воронцов К.В. Разметка данных для обучения нейросетевых моделей языка как способ формализации гуманитарных знаний. Доклад. XVIII научная конференция межрегиональной ассоциации «История и компьютер» «Историческая информатика как Historical Data Science». М.; 2022.

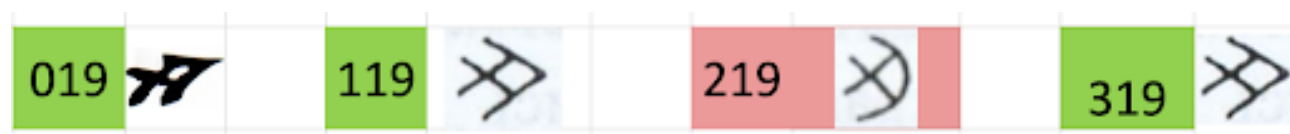


Рис. 3 / Fig. 3. Сравнение четырех наречий: Ирк Битиг, Енисейский, Орхонский, Кюль-Тегин / Comparison of Four Dialects: Irk Bitig, Yenisei, Orkhon, and Kül Tegin

Источник / Source: составлено авторами / Compiled by the authors.

и учесть все варианты начертания. Это значит, что если один и тот же звук или руна имеют два существенно отличающихся начертания, то такие варианты будут включены в систематизацию как отдельные классы. Такой подход позволит модели машинного обучения формировать устойчивые векторные представления для каждого символа и надежно определять их даже при визуальных вариациях начертания.

При создании общего алфавита также было важно не перегрузить его вариациями начертания. Если облик и форма символов, обозначающих один звук, имеет незначительные отличия (рис. 2), то определяли их в один класс.

После формирования объединенного алфавита нужно было выделить уникальные символы, которые могли бы указывать на определенный диалект. Самый простой вариант, когда в других алфавитах нет даже похожих символов. Например, класс 002 алфавита Ирк Битиг на рис. 1. В таком случае эта руна входит в общий алфавит отдельным классом, а также становится «маркером» диалекта Ирк Битиг. Если она встречается в надписи, то это может указывать на ее принадлежность именно к этому наречию. Также уникальными символами могут стать и те руны, аналоги которых есть во всех алфавитах. Например, класс 219 Орхонского алфавита хоть и похож на руны из других алфавитов, имеет заметное отличие. Выделение его в отдельный класс не только улучшит его точность распознавания, но также поможет определить надпись с ним к Орхонскому наречию (рис. 3).

В результате сопоставления (рис. 4) был сформирован новый объединенный алфавит, содержащий руны всех четырех наречий.

СОЗДАНИЕ НОВОГО ОБЪЕДИНЕННОГО АЛФАВИТА

Процесс формирования нового объединенного алфавита является центральным этапом исследования, поскольку именно он создает основу для дальнейшей разметки изображений. На данном этапе четыре независимых набора: Кюль-Тегин, Ирк Битиг, орхонский и енисейский — будут интегрированы в единую систему классов, отображающую как общие, так и различающиеся руны древ-

034	↓	134	↑ ↓	234	↑ ↓	334	↓
035	⦿	135	⦿ ⦿	235	⦿ ⦿	335	⦿
036	⦿				⦿	336	
037	≡	137	:	237	:	337	:
038	⦿	Знак числа - идет обычно между страницами					338
		139	◁ ▷	239	◁ ▷	339	◁
		140	М	240	М	340	М

Рис. 4 / Fig. 4. Визуальные различия рун четырех наречий: Ирк Битиг, Енисейского, Орхонского и Кюль-Тегин / Visual Differences Between Runes of Four Dialects: Irk Bitig, Yenisei, Orkhon, and Kül Tegin

Источник / Source: составлено авторами / Compiled by the authors.

нетюркской письменности. Ранее был собран набор классов для начертаний рун 4 диалектов (рис. 1), который мы собираемся объединить в один общий.

Объединение проводилось поэтапно. Сначала были сопоставлены все руны, встречающиеся в каждом из четырех источников. На основе визуального сходства, фонетических интерпретаций и историко-филологических описаний были выделены группы идентичных или функционально эквивалентных символов. Если руна из разных диалектов совпадала по структуре, но различалась стилистическими особенностями (толщина линий, углы штрихов, орнаментальные элементы), она включалась в общий класс как вариант начертания. Например, классы 105 и 204 на рис. 1. В случаях, когда символы демонстрировали устойчивые различия (структурные расхождения, различное число элементов, уникальные добавочные штрихи), они фиксировались как отдельные классы, которые будут упомянуты в следующем разделе. Следовательно, объединенный алфавит формировался не как простое пересечение четырех наборов, а как их интегративная схема, отражающая внутреннюю вариативность и устраняющая многократное дублирование классов.

Рассмотрим структуру нового алфавита и количество классов. В результате сопоставления был получен консолидированный набор рун, включающий:

- базовые руны, встречающиеся во всех четырех наборах;
- вариативные руны, имеющие отличающиеся начертания при сохранении общей функции;
- уникальные символы, характерные только для одного или двух источников и отсутствующие в остальных.

301	↘ ↗
302	✂
303	└
304	>
305	4 4
306	5
307	8 8
308	3 3
309	X
310	1 1
311	7 7
312	D
313	9 9
314	H
315	Y Y
316	R
317	✓
318	Y
319	→

Рис. 5 / Fig. 5. Объединенный набор рун /
Combined Set of Runes

Источник / Source: составлено авторами / Compiled by the authors.

Каждому элементу присвоен идентификатор класса, не зависящий от конкретного наречия. Такая структура позволяет использовать набор и для аннотации данных, и для написания программных алгоритмов, работающих с конкретными символами.

Итоговое количество классов определяется числом структурно различных рун. Для практического применения в задаче классификации каждый класс соответствует одному элементу пространства меток. Также мы сохранили вспомогательный класс неизвестной руны, которым можно пометить символы, не вошедшие в нашу классификацию.

Весь набор представлен в виде таблицы соответствий, в которой для каждого объединенного класса указаны:

- идентификатор руны, общей для четырех наборов;
- визуальные варианты начертания.

Стандартизация меток обеспечила совместимость разметки, корректный экспорт в форматы YOLO/CNN и единообразие аннотаций при работе с графическим корпусом. Фрагмент итогового общего алфавита представлен на рис. 5.

Итоги формирования общего алфавита: выделена 51 руна и один вспомогательный класс. В новом объединенном алфавите каждый символ формально принадлежит одному классу, даже если встречается в нескольких источниках. Теперь разметка текста становится однозначной. Каждый символ имеет непротиворечивую и стабильную метку.

ВЫДЕЛЕНИЕ УНИКАЛЬНЫХ ПРИЗНАКОВ НАРЕЧИЙ

После формирования объединенного алфавита стала возможна систематическая оценка распределения рун по четырем диалектам. Систематизация знаков позволила сопоставлять символы напрямую, не учитывая различия в стилевых вариациях, что послужило основой для выявления уникальных рун, характерных для определенных диалектов.

Для отнесения руны к категории уникальных использовались следующие критерии:

1. Структурная автономность: символ должен сохранять отличимую форму, не являющуюся стилевым вариантом другой руны в объединенном алфавите.

2. Локализованность: руна не присутствует хотя бы в одном из наборов.

При анализе алфавитов четырех диалектов была составлена таблица сопоставления. В ней указано, к каким наречиям относятся руны (рис. 6).

Там же отмечены уникальные руны для конкретного наречия (рис. 7) и выбывающие руны с другой вариацией начертания (рис. 8).

1		Енисейский, Орхонский (второй вариант)
2		Енисейский, Орхонский, Кюль-Тегин
3		Ирк Битиг
4		Енисейский, Орхонский, Кюль-Тегин
5		Енисейский, Орхонский (первый вариант)
6		Енисейский, охронский, Ирк Битиг

Рис. 6 / Fig. 6. Общие руны, с указанием к каким наречиям они относятся / General Runes, with an Indication of which Dialects they Belong to

Источник / Source: составлено авторами / Compiled by the authors.

7		Кюль-тегин
8		Енисейский

Рис. 7 / Fig. 7. Уникальные руны для одного наречия / Unique Runes for One Dialect

Источник / Source: составлено авторами / Compiled by the authors.

54		Енисейский — это просто перевернутый символ 4
----	---	---

Рис. 8 / Fig. 8. Руна с другой вариацией начертания / A Rune with an Alternative Form of Inscription

Источник / Source: составлено авторами / Compiled by the authors.

Выделение уникальных рун имеет ключевое значение для статистических и аналитических исследований в этой области. Они позволяют повысить точность определения происхождения надписей. Кроме того, наличие таких наборов дает возможность применять алгоритм, где сначала определяется сам текст, а затем по уникальным символам — предположительное наречие надписи.

Таким образом, выявление уникальных признаков наречий стало возможным благодаря созданию общего алфавита. Полученные результаты демонстрируют, что даже в пределах близкородственных письменностей сохраняется выраженная диалектная специфика.

ИТОГ СОСТАВЛЕНИЯ АЛФАВИТА

Проведенный анализ отдельных диалектов, составление общего алфавита и поиск уникальных символов показали, что уникальные признаки описывают отличительные особенности различных наречий:

- *Кюль-Тегин* характеризуется наличием нескольких сложных по структуре и хорошо стандартизированных рун, что связано с монументальным характером надписи и строгой эпиграфической нормой.
- *Ирк Битиг* демонстрирует набор упрощенных символов и специфических вариантов начертания, обусловленных рукописной традицией и большей свободой писца.
- *Орхонские памятники* содержат руны, устойчиво встречающиеся в надписях на камне, незначительно отличающиеся от енисейских вариантов формой штрихов и структурой внутренних элементов.
- *Енисейская письменность* выделяется присутствием нескольких редких, но строго локальных рун, отсутствующих в других наборах, что делает их ключевыми для атрибуции фрагментов.

В совокупности эти различия позволяют говорить о наличии выраженной региональной стратификации внутри рунической традиции.

ПРИМЕНЕНИЕ В ЗАДАЧАХ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Алфавитом, который получился в результате этого исследования, будет размечен корпус изображений рунических надписей. Это изображения реальных памятников, а также перенесенные на бумажные и цифровые носители тексты. Так как корпус включает изображения памятников из разных регионов и созданных в различные эпохи, то и надписи на них могут относиться к разным диалектам. Использование алфавита со всеми диалектами позволит разметить эти снимки наиболее полно. При разметке изображений каждая руна получает уникальную метку из объединенной системы, что делает данные совместимыми с современными архитектурами нейросетей, такими как YOLO и сверточные нейронные сети (CNN).

Модель YOLO используется на этапе детекции — она определяет координаты символов на изобра-

жении и присваивает им классы из единого набора [17]. Сверточные сети применяются на этапе классификации для уточнения принадлежности выделенных фрагментов к конкретным рунам. Унификация алфавита позволяет обучать обе модели на согласованных данных и устраняет проблему несовпадений меток между разными корпусами.

Описанный в работе подход помогает модели формировать устойчивые векторные представления и уверенно различать даже схожие по форме символы. Это повышает точность и обобщающую способность системы, делая ее пригодной для работы с надписями из разных регионов, эпох и носителей.

В исследовании распознавания рун алфавита Кюль-Тегин возникла техническая проблема из-за символов, не вошедших в набор. Они подходили для обучения модели обнаружения, но модель классификации, наоборот, сбивали. В итоге модель ошибочно относила часть известных рун к этому классу. Использование объединенного алфавита должно решить эту проблему. В процессе обучения классификатора предполагается отказ от использования класса неизвестной руны. Для предотвращения искажения процесса обучения будут закрашены. Ожидается, что использование только известных рун обеспечит достаточный объем данных для обучения, а ансамбль моделей детекции и классификации продемонстрирует более высокие показатели по сравнению с предыдущими экспериментами.

ВЫВОДЫ

Проделанная работа показала, что объединение четырех классификаций древнетюркских рун в единый алфавит действительно помогает упорядочить разрозненный материал и устранить противоречия, накопившиеся в разных исследовательских традициях. Полученные результаты согласуются с выводами современных исследований в области

цифровой эпиграфики и компьютерного зрения, в которых подчеркивается ключевая роль стандартизации данных для повышения устойчивости и воспроизводимости моделей машинного обучения. Единый набор символов делает разметку более предсказуемой и сокращает количество ошибок, которые обычно возникают при работе с несколькими несогласованными системами. Это важно для задач машинного обучения, где качество данных напрямую влияет на итоговую точность моделей.

Получившийся алфавит охватывает весь спектр рун четырех диалектов орхоно-енисейской письменности и одновременно фиксирует различия между вариантами начертания. Такой формат позволяет корректно применять систематизацию при обучении моделей и избежать ситуации, когда непохожие по форме руны оказываются в одном классе, а визуально близкие — в разных.

Отдельной задачей стало выделение рун, характерных только для одного наречия. Этот шаг помогает не только распознавать символы, но и учитывать особенности происхождения надписей. Такие маркеры могут служить инструментом атрибуции и обогатить аналитическую часть будущих исследований.

Подготовленный алфавит создает прочную основу для разработки многоалфавитной системы компьютерного зрения. Он упрощает обучение нейросетей на разнородных изображениях, повышает стабильность распознавания и помогает моделям уверенно работать с текстами, созданными в разных регионах и эпохах.

Следующим шагом станет расширение корпуса изображений и проверка созданной структуры в реальных задачах автоматического распознавания. Новый алфавит уже используется для аннотации изображений, которые станут основой для дальнейших исследований тюркской письменности с помощью технологий искусственного интеллекта.

СПИСОК ИСТОЧНИКОВ

1. Убрятова Е.И. Сергей Ефимович Малов: (К 75-летию со дня рождения). В сб.: Тюркологический сборник. М.: Издательство Академии наук СССР; 1951. URL: https://www.orientalstudies.ru/rus/images/pdf/journals/p_ts_1_1951_02_ubryatova.pdf
2. Кляшторный С.Г. Древнетюркские рунические памятники как источник по истории Средней Азии. М.; 1964. 215 с. URL: <https://www.elibrary.ru/qaujxn>
3. Кызласов И.Л. Рунические письменности евразийских степей. М.: Наука; 1994. 318 с. URL: <https://www.elibrary.ru/wjrref>
4. Пенская Е.Н. Машинное обучение и представление информации: новые возможности цифровых архивов. *Имагология и компаративистика*. 2025;23:380-389. URL: <https://doi.org/10.17223/24099554/23/18>
5. Примова Ф.М. Искусственный интеллект в филологии. *Международная конференция академических наук*. 2025;4(1)63-65. URL: <https://doi.org/10.5281/zenodo.14682779>

6. Рудакова С.В. Особенности взаимодействия филологии с искусственным интеллектом. В сб.: Общество. Наука. Инновации (НПК-2024). Киров: Вятский государственный университет; 2024:70-73. URL: <https://www.elibrary.ru/eeekohu>
7. Матвеева Т.В. Категориально-текстовый анализ речевого произведения: обоснование и применение метода. *Quaestio Rossica*. 2024;12(3):901-920. URL: <https://doi.org/10.15826/qr.2024.3.915>
8. Кочкаров Р.А., Бородина А.Д. Распознавание орхоно-енисейских рунических надписей методами машинного обучения. *Нелинейный мир*. 2025;23(3):96-106. URL: <https://doi.org/10.18127/j20700970-202503-12>
9. Кормушин И.В. О взглядах Э.Р. Тенишева на историю древних литературных тюркских языков. *Вестник Института языкознания РАН*. 2021;1-2 (30-31):45-53. URL: <https://www.elibrary.ru/mbprias>
10. Кочкаров У.Ю., Беляева В.Н., Кочкаров Р.А., Кочкаров А.А. Подготовка набора визуальных данных для машинного распознавания рунической письменности. В сб.: Проблематика и методология современных тюркологических исследований. Коллективная монография. М.: Издательство МБА; 2025:162-173. URL: <https://www.elibrary.ru/chxczmz>
11. Лебедев Ю.С., Попов П.В. Погребение VIII-IX вв. из Астраханской области и горшок с рунической надписью. *Российская археология*. 2023;1:178-186. URL: <https://doi.org/10.31857/S0869606323010130>
12. Мунчаев Р.М., Корневский С.Н., ред. Проблемы археологии Кавказа. Вып. 1. М.: ТAYC; 2012. 248 с. URL: <https://www.elibrary.ru/qpxfjv>
13. Кормушин И.В. Тюркские енисейские эпитафии. Монография. М.: Наука; 1997. 338 с. URL: <https://www.elibrary.ru/sfhmdj>
14. Томсен В.Л.П. Дешифровка орхонских и енисейских надписей. Ада Чир-Суу — Отечество: краеведческий альманах. 2018;4:129-134. URL: <https://www.elibrary.ru/ntspvz>
15. Володин А.Ю. Цифровые гуманитарные исследования в контексте датаизма. *Современная наука: Актуальные проблемы теории и практики. Серия: Гуманитарные науки*. 2024;7(2):6-9. URL: <https://doi.org/10.37882/2223-2982.2024.7-2.05>
16. Сухан Д.Д., Плюснина Е.А. Метаразметка и визуализация данных в корпусе текстов по корпусной лингвистике. В сб.: Компьютерная лингвистика и вычислительные онтологии. Вып. 8. СПб.: Университет ИТМО; 2024:45-60. URL: <https://doi.org/10.17586/2541-9781-2024-8-45-60>
17. Кочкаров Р.А., Бородина А.Д. Систематизация орхоно-енисейских рунических письменностей и перспективы их автоматизированного анализа методами машинного обучения. *Динамика сложных систем — XXI век*. 2025;19(5):97-103. URL: <https://doi.org/10.18127/j19997493-202505-11>

REFERENCES

1. Ubryatova E.I. Sergey Efimovich Malov: (On the 75th Anniversary of His Birth). In: Turkological Collection. Moscow: Publishing House of the USSR Academy of Sciences; 1951. URL: https://www.orientalstudies.ru/rus/images/pdf/journals/p_ts_1_1951_02_ubryatova.pdf (In Russ.).
2. Klyashtorny S.G. Ancient Turkic Runic Monuments as a Source on the History of Central Asia. Moscow; 1964. 215 p. URL: <https://www.elibrary.ru/qaujxn> (In Russ.).
3. Kyzlasov I.L. Runic Scripts of the Eurasian Steppes. Moscow: Nauka; 1994. 318 p. URL: <https://www.elibrary.ru/wjrref> (In Russ.).
4. Penskaya E.N. Machine Learning and Information Presentation: New Opportunities for Digital Archives. *Imagology and Comparative Studies*. 2025;23:380-389. (In Russ.). URL: <https://doi.org/10.17223/24099554/23/18>
5. Primova F.M. Artificial Intelligence in Philology. *International Conference of Academic Sciences*. 2025;4(1):63-65. (In Russ.). URL: <https://doi.org/10.5281/zenodo.14682779>
6. Rudakova S.V. Features of the Interaction of Philology with Artificial Intelligence. In: Society. Science. Innovations (NPC 2024). Kirov: Vyatka State University; 2024:70-73. URL: <https://www.elibrary.ru/eeekohu> (In Russ.).
7. Matveeva T.V. Categorical-textual analysis of a speech work: Justification and application of the method. *Quaestio Rossica*. 2024;12(3):901-920. (In Russ.). URL: <https://doi.org/10.15826/qr.2024.3.915>
8. Kochkarov R.A., Borodina A.D. Recognition of Orkhon-Yenisei Runic Inscriptions Using Machine Learning Methods. *Non-Linear World*. 2025;23(3):96-106. (In Russ.). URL: <https://doi.org/10.18127/j20700970-202503-12>
9. Kormushin I.V. On E.R. Tenishev's Views on the History of Ancient Literary Turkic Languages. *Bulletin of the Institute of Linguistics of the Russian Academy of Sciences*. 2021;1-2(30-31):45-53. URL: <https://www.elibrary.ru/mbprias> (In Russ.).

10. Kochkarov U.Yu., Belyaeva V.N., Kochkarov R.A., Kochkarov A.A. Preparation of a set of visual data for machine recognition of runic writing. In: Problems and Methodology of Modern Turkological Research. Monograph. Moscow: MBA Publishing House; 2025. URL: <https://www.elibrary.ru/chxcmz> (In Russ.).
11. Lebedev Yu.S., Popov P.V. Burial of the 8th-9th Centuries from the Astrakhan Region and a Pot with a Runic Inscription. *Russian Archaeology*. 2023;1:178-186. (In Russ.). URL: <https://doi.org/10.31857/S0869606323010130>
12. Munchaev R.M., Korenevsky S.N., ed. Problems of the Caucasus Archaeology. Vol. 1. Moscow: TAUS; 2012. 248 p. URL: <https://www.elibrary.ru/qpxfjv> (In Russ.).
13. Kormushin I.V. Turkic Yenisei epitaphs. Monograph. Moscow: Nauka; 1997. URL: <https://www.elibrary.ru/sfhmdj> (In Russ.).
14. Tomsen V.L.P. Deciphering the Orkhon and Yenisei Inscriptions. *Ada Chir Suu — Fatherland: Local History Almanac*. 2018;4:129-134. URL: <https://www.elibrary.ru/ntspvz> (In Russ.).
15. Volodin A. Yu. Digital Humanities Research in the Context of Dataism. *Modern Science: Current Issues in Theory and Practice. Series: Humanities Sciences*. 2024;7(2):6-9. (In Russ.). URL: <https://doi.org/10.37882/2223-2982.2024.7-2.05>
16. Suhan D.D., Plyusnina E.A. Meta-markup and data visualization in the corpus of texts on corpus linguistics. In: Computer linguistics and computational ontologies. Issue 8. St. Petersburg: ITMO University; 2024:45-60. (In Russ.). URL: <https://doi.org/10.17586/2541-9781-2024-8-45-60>
17. Kochkarov R.A., Borodina A.D. Systematization of the Orkhon-Yenisei Runic Scripts and Prospects for Their Automated Analysis Using Machine Learning Methods. *Dynamics of Complex Systems — 21st Century*. 2025;19(5):97-103. (In Russ.). URL: <https://doi.org/10.18127/j19997493-202505-11>

ИНФОРМАЦИЯ ОБ АВТОРАХ / ABOUT THE AUTHORS

Анастасия Дмитриевна Бородин — студентка бакалавриата, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

Anastasia D. Borodina — Bachelor's degree student, Financial University under the Government of the Russian Federation, Moscow, Russian Federation
<https://orcid.org/0009-0002-5126-9854>
225034@edu.fa.ru

Монгун-Ай Вячеславовна Кызыл-оол — студентка бакалавриата, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

Mongun-Ai V. Kyzyl-ool — Bachelor's degree student, Financial University under the Government of the Russian Federation, Moscow, Russian Federation
<https://orcid.org/0009-0006-5385-9603>

Автор для корреспонденции / Corresponding author:
221361@edu.fa.ru

Расул Ахматович Кочкаров — доктор технических наук, заместитель декана по научной работе, доцент кафедры искусственного интеллекта факультета информационных технологий и анализа больших данных, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

Rasul A. Kochkarov — Dr. Sci. (Tech.), Deputy Dean for Research, Associate Professor of the Department of Artificial Intelligence, Faculty of Information Technology and Big Data Analysis, Financial University under the Government of the Russian Federation, Moscow, Russian Federation
<https://orcid.org/0000-0003-3186-3901>
rkochkarov@fa.ru

Заявленный вклад авторов:

А.Д. Бородин — участие в систематизации алфавитов, анализе визуальных различий рун и подготовке экспериментальной части исследования.

М.В. Кызыл-оол — участие в анализе источников, формировании объединенного алфавита, подготовка иллюстративных материалов и текста статьи.

Р.А. Кочкаров — разработка общей концепции исследования, научное руководство, анализ результатов и подготовка финальной версии текста.

Authors' declared contributions:

A. D. Borodina — participation in alphabet systematization, analysis of visual differences in runes, and preparation of the experimental part of the study.

M. V. Kyzyl-ool — participation in source analysis, development of the unified alphabet, preparation of visual materials and manuscript drafting.

R. A. Kochkarov — development of the general research concept, scientific supervision, analysis of results, and preparation of the final version of the manuscript.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

Статья поступила 22.04.2026; после рецензирования 10.07.2026; принята к публикации 01.08.2026.

Авторы прочитали и одобрили окончательный вариант рукописи.

The article was submitted on 22.04.2026; revised on 10.07.2026; accepted for publication on 01.08.2026.

The authors read and approved the final version of the manuscript.

DOI: 10.26794/3030-7097-2026-2-3-16-25
УДК 004.89:63+631.15:004.89(045)

Применение искусственного интеллекта в сельском хозяйстве: мировая практика

А.Х. Шелепаева

Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

АННОТАЦИЯ

Статья посвящена анализу современных тенденций и практик применения искусственного интеллекта (ИИ) в сельском хозяйстве на основе рецензируемых научных источников 2022–2026 гг. Целью работы выступает комплексное исследование проблематики и ключевых трендов развития ИИ в аграрном секторе, а также систематизация количественных данных, подтверждающих эффективность внедряемых технологий. В статье рассматриваются основные направления использования ИИ: прогнозирование урожайности на базе нейросетевых алгоритмов (LSTM, CNN), мониторинг здоровья растений и точное внесение ресурсов с применением компьютерного зрения (ResNet, YOLO), роботизация производственных процессов, а также внедрение интеллектуальных систем в экономику и управление агропромышленным комплексом. На основе анализа эмпирических исследований систематизированы показатели эффективности: точность диагностики заболеваний растений достигает 99,2%, снижение использования пестицидов — до 90%, экономия водных ресурсов — до 46%, рост доходов хозяйств — на 15–20%. Особое внимание уделено сравнительному анализу барьеров внедрения ИИ в различных группах стран. В развитых государствах преобладают проблемы технологической фрагментации и дефицита квалифицированных кадров. Для стран БРИКС характерны инфраструктурные ограничения и разрыв между наукой и производством. В развивающихся странах ключевыми препятствиями выступают отсутствие базовой инфраструктуры, низкая платежеспособность мелких хозяйств и недостаточный уровень цифровой грамотности. В заключительной части определены слабоизученные аспекты, требующие дальнейших исследований: социально-психологические барьеры восприятия технологий, экологический след самих ИИ-решений, вопросы суверенитета и монетизации данных, экономическая эффективность для мелких хозяйств, кибербезопасность и этические дилеммы в селекции.

Ключевые слова: искусственный интеллект; агропромышленный комплекс; инновации в АПК; большие данные; сельскохозяйственные предприятия; цифровые решения; информационные системы

Для цитирования: Шелепаева А.Х. Применение искусственного интеллекта в сельском хозяйстве: мировая практика. *Цифровые решения и технологии искусственного интеллекта*. 2026;2(3):16-25. DOI: 10.26794/3030-7097-2026-2-3-16-25

ORIGINAL PAPER

The Use of Artificial Intelligence in Agriculture: Global Practice

A. Kh. Shelepaeva

Financial University under the Government of the Russian Federation, Moscow, Russian Federation

ABSTRACT

This article is devoted to the analysis of current trends and practices of the application of artificial intelligence (AI) in agriculture based on peer-reviewed scientific sources of 2022–2026. The purpose of the work is a comprehensive study of the problems and key trends of the development of AI in the agricultural sector, as well as the systematization of quantitative data confirming the effectiveness of the implemented technologies. The article discusses the main areas of AI use: crop yield forecasting based on neural network algorithms (LSTM, CNN), plant health monitoring and precise resource application using computer vision (ResNet, YOLO), robotization of production processes, and the introduction of intelligent systems into the economy and management of the agro-industrial complex. Based on the analysis of empirical studies, the article systematizes performance indicators: the accuracy of plant disease diagnosis reaches 99.2%, the reduction of pesticide use is up to 90%, water resources are saved by up to 46%, and farm income increases by 15–20%. Special attention is paid to a comparative analysis of the barriers to the implementation of AI in different groups of countries. In developed countries, the main problems are technological fragmentation and a shortage of qualified personnel. In BRICS countries, there are infrastructure limitations and a gap between science and production. In developing countries, the key obstacles are the lack of basic infrastructure, the low solvency of small farms, and a lack of digital literacy. The final part identifies understudied aspects that require further research, such as socio-psychological barriers to technology adoption,

the environmental impact of AI solutions, issues related to data sovereignty and monetization, economic efficiency for small-scale farmers, cybersecurity, and ethical dilemmas in breeding.

Keywords: artificial intelligence; agro-industrial complex; innovations in the agro-industrial complex; big data; agricultural enterprises; digital solutions; and information systems

For citation: Shelepaeva A.Kh. The use of artificial intelligence in agriculture: Global practice. *Digital Solutions and Artificial Intelligence Technologies*. 2026;2(3):16-25. DOI: 10.26794/3030-7097-2026-2-3-16-25

ВВЕДЕНИЕ

Сельское хозяйство, являясь основой продовольственной безопасности и экономического развития многих стран, стоит перед лицом беспрецедентных вызовов. Рост мирового населения, прогнозируемый на уровне 10 млрд человек к 2050 г., изменение климата, деградация почв, нехватка водных ресурсов и сокращение сельскохозяйственной рабочей силы требуют коренной трансформации отрасли [1, 2]. Традиционные методы ведения хозяйства более не способны в полной мере обеспечить растущий спрос на продовольствие при сохранении экологического баланса.

В этой связи технологии Индустрии 4.0, в особенности искусственный интеллект (ИИ), становятся ключевым драйвером перехода к «умному сельскому хозяйству» (Smart Farming) и сельскому хозяйству 4.0 [3, 4]. ИИ, включающий в себя машинное обучение, компьютерное зрение, робототехнику и анализ больших данных, открывает возможности для оптимизации всех этапов производственной цепочки: от селекции и прогнозирования урожайности до мониторинга здоровья растений и автоматизации сбора урожая [5].

Целью данного исследования является комплексный анализ проблематики и современных тенденций применения ИИ в сельском хозяйстве на основе рецензируемых научных источников 2022–2026 гг., а также систематизация количественных данных об эффективности этих технологий.

МЕТОДЫ ИССЛЕДОВАНИЯ

В работе были применены сетевые методы анализа научной литературы путем создания библиометрических карт. Основой для анализа послужила выборка из 22 рецензируемых публикаций на английском языке, охватывающих период 2022–2026 гг. и индексированных в международных базах научного цитирования. Применение сетевых методов анализа позволило автоматизировать обработку значительного массива публикаций по теме исследования, выявить скрытые закономерности и тематические кластеры.

В результате проведенного библиометрического анализа все релевантные теме исследования термины были распределены по трем основным кластерам, отражающим структуру современных

исследований в области применения искусственного интеллекта в сельском хозяйстве.

Первый кластер (технологический) объединил термины, связанные с современными технологиями сбора и обработки данных: Internet of Things (IoT), drones, sensors, robotics, precision agriculture, remote sensing, cloud computing, 3D modeling, digital twins. Данный кластер отражает инструментальную базу, обеспечивающую сбор первичной информации и физическое взаимодействие с объектами сельского хозяйства.

Второй кластер (методологический) включил термины, непосредственно связанные с методами искусственного интеллекта и анализа данных: machine learning, deep learning, neural networks, computer vision, LSTM, CNN, YOLO, ResNet, predictive analytics, big data. Этот кластер демонстрирует, какие алгоритмы и подходы наиболее востребованы для решения прикладных задач в агросекторе.

Третий кластер (прикладной) объединил термины, описывающие конкретные области применения и сопутствующие факторы: yield prediction, crop monitoring, disease detection, weed control, soil health, sustainability, climate change, risk management, food security, supply chain, green innovation. Этот кластер отражает спектр задач, на решение которых направлены разрабатываемые технологии.

Выявленная кластеризация подтверждает, что современные исследования в области ИИ в сельском хозяйстве развиваются на пересечении трех ключевых областей: развитие аппаратной базы (сенсоры, роботы), совершенствование алгоритмов обработки данных (нейросети) и адаптация этих инструментов для решения конкретных прикладных задач, преимущественно связанных с повышением продуктивности и устойчивости агросистем.

НАПРАВЛЕНИЯ И МЕТОДЫ ПРИМЕНЕНИЯ ИИ

Проведенный анализ научных публикаций дает возможность систематизировать существующие разработки по ряду магистральных направлений, характеризующихся высоким потенциалом внедрения ИИ-технологий. Авторами были выделены четыре основных раздела: прогнозирование, мониторинг, роботизация и процессы управления:

1. Прогнозирование урожайности и управление рисками. Одним из ключевых приложений ИИ является разработка прогнозных моделей урожайности. Исследователи активно применяют нейросетевые алгоритмы для анализа сложных нелинейных зависимостей между урожайностью и множеством факторов — погодными условиями, состоянием почвы (индексы NDVI), агротехническими приемами. С.В. Шайтура и др. [6] в своем исследовании использовали и сравнили три типа моделей: многослойный перцептрон (MLP), сверточные нейронные сети (CNN) и рекуррентные сети с долгой краткосрочной памятью (LSTM), показав превосходство последних в точности прогнозов на временных рядах. Shebanina et al. [7] также подчеркивают важность таких моделей для стратегического планирования в условиях климатической неопределенности, позволяя предприятиям разрабатывать адаптивные стратегии.

2. Мониторинг здоровья растений и точное внесение ресурсов. Компьютерное зрение и глубокое обучение лежат в основе систем диагностики болезней растений и идентификации сорняков. Например, модели ResNet и YOLO достигают высокой точности (до 99%) в классификации заболеваний сельскохозяйственных культур [1]. Эти данные позволяют перейти от сплошной обработки полей к адресному внесению пестицидов и удобрений. Такой подход, реализуемый с помощью дронов и сенсоров, не только повышает эффективность, но и снижает химическую нагрузку на окружающую среду, способствуя развитию устойчивого сельского хозяйства. Monsalve-Castro et al. [8] в своей работе на примере Колумбии подтверждают, что ИИ положительно связан с внедрением «зеленых» инноваций как в продуктах (экодизайн), так и в процессах (оптимизация использования ресурсов).

3. Роботизация и автоматизация процессов. Роботизация охватывает широкий спектр задач — от посева и прополки до сбора урожая. В обзоре Ashe & Mengistu [2] описываются роботизированные системы для сбора фруктов и овощей, которые используют компьютерное зрение для определения спелости плодов и аккуратного их сбора. Wildan J. [3] акцентирует внимание на интеграции роботов с IoT-сенсорами, что позволяет создавать полностью автономные системы управления фермой, включая мониторинг почвы, ирригацию и внесение удобрений.

4. Экономика и управление в АПК. ИИ трансформирует не только производственные, но и управленческие и маркетинговые процессы. Zhang et al. [9] на примере китайских агропредприятий показывают, что внедрение ИИ ведет к ро-

сту совокупной факторной производительности за счет оптимизации структуры трудовых ресурсов и межфирменной кооперации. В контексте маркетинга, Saravanan P. [10] предлагает концептуальную модель, где ИИ улучшает результаты сбыта через повышение рыночной аналитики и операционной эффективности, что особенно актуально для развивающихся экономик, таких как Индия. Caglar E. [11], анализируя данные венчурных инвестиций стран G7 и Турции, выявил значительную диспропорцию в финансировании: сектор социальных медиа и маркетинга привлекаеткратно большие объемы инвестиций, чем сельское хозяйство, что указывает на недооценку инвесторами агротехсектора.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ: КОЛИЧЕСТВЕННЫЕ ПОКАЗАТЕЛИ ЭФФЕКТИВНОСТИ

Анализ представленных исследований позволил систематизировать численные данные (табл. 1), демонстрирующие эффективность внедрения ИИ-технологий.

СУЩЕСТВУЮЩИЕ ПРОБЛЕМЫ ВНЕДРЕНИЯ ИИ В СЕЛЬСКОМ ХОЗЯЙСТВЕ В РАЗРЕЗЕ СТРАН

Обобщение эмпирических данных и результатов полевых исследований, представленных в научной литературе, позволяет не только количественно оценить преимущества ИИ, но и систематизировать барьеры, сдерживающие его повсеместное распространение. Анализ показывает, что инфраструктурные, экономические и социально-психологические препятствия имеют ярко выраженную страновую специфику. Для наглядного сопоставления этих различий все многообразие проблем было сгруппировано по пяти ключевым категориям (инфраструктурные, экономические, социально-образовательные, институциональные и культурно-психологические) в разрезе трех макрорегионов: развитых стран, стран БРИКС и развивающихся государств (табл. 2).

1. Развитые страны (США, Канада, страны ЕС, G7)

В этих странах сельское хозяйство характеризуется высокой степенью механизации, наличием крупных агрохолдингов и развитой цифровой инфраструктурой. Проблемы здесь носят иной характер, нежели в развивающихся экономиках:

- *Технологическая фрагментация и отсутствие стандартизации.* Несмотря на обилие ИИ-решений, они часто представляют собой закрытые проприетарные системы разных производителей (например,

**Сводные количественные показатели эффективности применения ИИ в сельском хозяйстве /
Summary of Quantitative Indicators of the Effectiveness of AI Application in Agriculture**

Технология / Направление / Technology / Direction	Ключевой показатель эффективности / Key Performance Indicator	Значение / Диапазон / Value / Range	Источник / Source
Модели диагностики болезней растений	Точность (Accuracy) распознавания заболеваний	85–99,2%	Bayrak & Olu (2025); Bulgen (2024); Topcu & Gunes (2024)
	Пример: ResNet101 для кукурузы	94,56%	Bulgen (2024)
	Пример: ResNet для классификации листьев	99,2%	Topcu & Gunes (2024)
	Пример: YOLOv8 для пшеницы (Precision)	95,4%	Bulgen (2024)
Прогнозирование урожайности (Yield Prediction)	Снижение ошибки прогноза (RMSE)	До 2,5 ц/га	Shaitura et al. (2025)
	Средняя абсолютная ошибка (MAPE)	5,5–7,5%	
	Преимущество модели LSTM перед MLP	Статистически значимо ($p < 0,05$)	
Оптимизация внесения химикатов	Снижение использования пестицидов/гербицидов	20–90%	Bayrak & Olu (2025); Cakmakci & Cakmakci (2023); Istudor et al. (2023)
	Экономия воды при «умном» орошении	20–46%	Monsalve-Castro et al. (2025); Routis & Roussaki (2023)
	Экономия удобрений (точное земледелие)	30%	Bayrak & Olu (2025)
Применение дронов (UAV)	Повышение производительности при опрыскивании	20–30% (к базовому уровню)	
	Экономия времени при опрыскивании	70% (по сравнению с ручным / наземным)	
	Снижение потерь при наземном внесении	54%	
	Увеличение осаджения на листву (canopy deposition)	309%	
Экономическая эффективность	Рост доходов фермеров при использовании ИИ-платформ	15–20%	Saravanan (2025)
	Рентабельность инвестиций (ROI) в ИИ	20–30%	Singh & Rana (2023); Saravanan (2025)
	Рост производительности (TFP) агропредприятий	Статистически значимый положительный эффект	Zhang et al. (2026)

Источник / Source: составлено автором / Compiled by author.

техника John Deere, платформы от IBM, стартапы вроде Blue River). Отсутствие единых стандартов обмена данными затрудняет создание целостной экосистемы «умной фермы».

- *Этические и правовые коллизии (Data Governance).* В ЕС и США остро стоит вопрос о праве собственности на данные, генерируемые техникой. Кому принадлежит информация о состоянии поля — фермеру, арендующему технику, или корпорации-производителю? Существующее законодательство (например, европейский GDPR) не полностью покрывает специфику агроданных, что создает правовую неопределенность.

- *Сопротивление малых и семейных ферм.* В то время как крупные агрохолдинги активно внедряют инновации, малые семейные фермы (характерные для Европы) демонстрируют сдержанность. Причины кроются не столько в финансовой недоступности, сколько в социокультурных факторах: приверженность традициям, нежелание терять автономию в пользу алгоритмов, недостаток доверия к «черному ящику» ИИ-решений.

- *Кадровый голод и переквалификация.* Парадоксально, но в странах с высоким уровнем образования остро не хватает агрономов, обладающих компетенциями в области Data Science. Возникает разрыв между аграрным и IT-образованием. Подготовка специалистов «двойного профиля» требует времени и системных изменений в образовательных программах.

2. Страны с переходной экономикой и БРИКС (Россия, Турция, Китай, ЮАР)

У государств с переходной экономикой наблюдается высокий потенциал роста, но внедрение ИИ сталкивается с инфраструктурными и институциональными ограничениями:

- *Неравномерность цифровой инфраструктуры.* Успешные пилотные проекты в передовых хозяйствах соседствуют с полным отсутствием даже базового интернета в отдаленных районах. Это создает «цифровые разрывы» внутри одной страны. Для России, например, характерна высокая концентрация ИИ-разработок в европейской части и их слабая диффузия в Сибирь и на Дальний Восток [6, 12].

- *Проблемы с качеством и доступностью данных.* Государственная статистика часто устаревает, а данные с полей не систематизированы. Отсутствуют единые национальные платформы для сбора и хранения агроданных, что делает разработку точных прогнозных моделей крайне затруднительной. В Китае эта проблема решается жестким государственным регулированием и созданием централизованных платформ, но это порождает вопросы о приватности.

- *Высокая стоимость адаптации западных технологий.* Зависимость от импортного ПО и оборудования (датчиков, дронов, чипов) делает ИИ-решения уязвимыми к валютным колебаниям и санкционным рискам. Разработка собственных аналогов требует времени и инвестиций, которые частный бизнес не всегда готов нести. Для Турции, активно стремящейся к технологическому суверенитету, эта проблема также актуальна.

- *Институциональные барьеры трансфера знаний.* Существует разрыв между академической наукой (университеты, НИИ) и реальным агробизнесом. Научные разработки часто не доходят до поля, а бизнес не доверяет «теоретическим» решениям. Механизмы коммерциализации и трансфера технологий работают неэффективно.

3. Развивающиеся страны и страны с малоземельным хозяйством (Индия, Тунис, Нигерия, Колумбия, Индонезия)

Здесь доминируют мелкие фермерские хозяйства, и проблемы носят наиболее острый, зачастую базовый характер.

- *Недоступность технологий для мелких фермеров.* Высокая стоимость оборудования делает его покупку невозможной для большинства. Даже доступ к смартфону, необходимому для работы с ИИ-приложениями, есть не у всех. Отсутствуют эффективные модели коллективного использования (кооперации) дорогостоящей техники (дронов, роботов) [13].

- *Низкий уровень цифровой и общей грамотности.* Многие фермеры не только не умеют пользоваться ИИ-инструментами, но и не доверяют информации с экрана, полагаясь на традиционные методы и советы старших. Языковой барьер (интерфейсы на английском) усугубляет ситуацию. Это превращает проблему внедрения из технологической в социально-образовательную [14, 15].

- *Отсутствие базовой инфраструктуры.* Проблемы с электроснабжением, нестабильный или отсутствующий мобильный интернет в сельской местности делают любые «умные» технологии бесполезными. Это фундаментальное ограничение, которое не решить простой поставкой оборудования [16].

- *Фрагментарность и неполнота данных.* Отсутствие систем учета, нерегулярный сбор статистики и большое количество теневых операций делают имеющиеся данные непригодными для обучения точных ИИ-моделей. Даже там, где технологии доступны, они могут работать некорректно из-за плохого качества входных данных [17, 18].

- *Риск усиления социального неравенства.* Существует обоснованное опасение, что ИИ-революция обойдет стороной массу мелких хозяйств, сконцентрировав выгоды у крупных экспортно ориентиро-



ванных производителей. Это может привести к обнищанию сельского населения и росту социальной напряженности.

- *Проблема доверия и культурного восприятия.*

В ряде культур (например, в некоторых регионах Индии или Африки) решения, принятые машиной, могут восприниматься как угроза традиционному укладу или авторитету старейшин, что создает мощное психологическое сопротивление, не объяснимое чисто экономическими причинами [19, 20].

Проведенный анализ позволяет не только систематизировать достигнутые результаты и выявленные барьеры, но и, что не менее важно для развития научного дискурса, очертить проблемное поле, остающееся за рамками современных исследований. Несмотря на обилие эмпирических данных о точности алгоритмов и экономической эффективности отдельных решений, целый ряд фундаментальных вопросов, лежащих на стыке технологий, социологии, экологии и этики, до сих пор не получил должного освещения в рецензируемой литературе. Дальнейшее развитие концепции «умного сельского хозяйства» требует перехода от сугубо технологической парадигмы к междисциплинарному осмыслению последствий цифровизации. Ниже представлены семь наиболее значимых, но слабоизученных проблемных областей.

1. Социально-психологические и культурные барьеры внедрения

Проблема «низкой цифровой грамотности» или «сопротивления фермеров» фигурирует во многих статьях, однако глубинные причины такого сопротивления, как правило, остаются за пределами анализа. Практически не изучено влияние традиционных культурных устоев, соотношение доверия к «машинному» знанию и традиционному опыту, а также возрастные и гендерные различия в восприятии ИИ. Отсутствуют исследования, посвященные тому, как психологический профиль фермера (склонность к риску, инновационность) влияет на скорость принятия решений о внедрении ИИ. Между тем без понимания этих факторов любые программы обучения и субсидирования рискуют остаться неэффективными.

2. Долгосрочные экологические риски самих ИИ-технологий

Большинство исследований фокусируется на том, как ИИ снижает экологический вред от сельского хозяйства за счет сокращения использования пестицидов и оптимизации водопотребления. Однако обратная сторона медали, как «экологический след» самого ИИ, практически полностью игнорируется. За пределами анализа остаются углеродный след от обучения сложных нейросетей, энергоемкость

Таблица 2 / Table 2

Классификация проблем внедрения ИИ в АПК по группам стран / Classification of AI Implementation Problems in the Agricultural Sector by Country Group

Категория проблем / Problem Category	Страна / Country		
	Развитые	БРИКС и переходные экономики	Развивающиеся
Инфраструктурные	Фрагментация данных, отсутствие стандартов	Неравномерность покрытия, зависимость от импорта	Отсутствие базовой инфраструктуры (интернет, электричество)
Экономические	Высокая стоимость для малых ферм	Высокая стоимость адаптации, валютные риски	Полная недоступность для мелких хозяйств
Социально-образовательные	Дефицит кадров «агроном-дата-саентист»	Разрыв между наукой и бизнесом	Низкая цифровая грамотность, языковой барьер
Институциональные и правовые	Неопределенность прав на данные (Data Governance)	Слабая госстатистика, отсутствие единых платформ	Отсутствие систем учета, теневая экономика
Культурно-психологические	Сопротивление на почве традиций	Региональные диспропорции в доверии к инновациям	Недоверие к «машинному» знанию, традиционный уклад

Источник / Source: составлено автором / Compiled by author.

производства и утилизации миллионов датчиков, дронов и роботов, а также неизбежный рост энергопотребления в сельской местности при массовом внедрении таких технологий. В работе Ashe & Mengistu [2] содержится лишь краткое упоминание о том, что обучение одной NLP-модели может сопровождаться выбросом 300 000 кг CO₂, а выбросы от обучения крупных ИИ-моделей удваиваются каждые 3–4 месяца. Однако эта проблема не рассматривается в контексте баланса между экологической пользой от применения ИИ и вредом от его же функционирования. Между тем «зеленая» революция не должна порождать «серые» проблемы, так как необходим комплексный анализ полного жизненного цикла AgriTech-продуктов, который сегодня отсутствует.

3. Проблема суверенитета и монетизации данных (Data Governance)

Вопросы сбора и анализа данных поднимаются в литературе достаточно часто, однако правовые и этические аспекты владения этими данными по-прежнему остаются «серой зоной». Zhang et al. [9] лишь вскользь упоминают о наличии институциональных трендов, а Rozhkova et al. [4] ограничиваются констатацией отсутствия документов, регулирующих земельные ресурсы, но проблема данных лежит гораздо глубже. Ключевые вопросы остаются без ответа: кто является истинным владельцем информации, генерируемой умными тракторами и дронами на поле фермера — сам фермер, производитель техники или государство? каким образом фермер может монетизировать свои данные и участвовать в распределении прибыли от их использования? каковы риски использования этих данных крупными корпорациями, например для спекулятивных операций на продовольственных рынках? Без ответа на эти вопросы невозможно создание справедливой и сбалансированной цифровой экосистемы в агропромышленном комплексе.

4. Экономика внедрения для мелких и средних хозяйств

Практически во всех работах упоминается «высокая стоимость технологий», однако детальный экономический анализ для различных масштабов производства отсутствует. За пределами исследований остается расчет реального срока окупаемости инвестиций (ROI) для мелкого фермера в развивающейся стране. Не изучены альтернативные бизнес-модели доступа к технологиям, такие как аренда дронов, кооперативное использование оборудования или модель pay-per-use, которые могли бы сделать ИИ доступнее. Эмпирические данные NABARD о рентабельности в 20–30% за 3–5 лет, упомянутые в работе Saravanan [10], скорее всего, представляют

собой исключение, чем правило, и не подкреплены сравнительными исследованиями эффективности различных моделей доступа. Между тем без понимания экономики внедрения для разных категорий хозяйств ИИ-революция рискует остаться уделом крупных агрохолдингов, лишь усугубляя социальное неравенство на селе.

5. Интеграция разнородных источников данных (Data Fusion)

Хотя принципы работы искусственного интеллекта с данными подробно описаны в литературе [3, 19], проблема интеграции гетерогенных данных в реальном времени остается практически не исследованной. До сих пор неясно, каким образом можно объединить спутниковые снимки низкого разрешения, данные с наземных сенсоров, метеорологические сводки и исторические показатели урожайности в единую, самообучающуюся систему, способную функционировать в режиме реального времени. Отсутствуют проработанные методы калибровки и валидации столь разнородных данных. Предложенная González-Rodríguez et al. [21] платформа AP4CF представляет собой шаг в этом направлении, однако она, скорее всего, формализует проблему, нежели предлагает ее инженерное решение. Между тем качество любой ИИ-модели напрямую зависит от качества и согласованности входных данных, и без решения этой проблемы даже самые совершенные алгоритмы будут давать сбои при практическом применении.

6. Кибербезопасность и уязвимость «умных ферм»

С ростом числа подключенных устройств и повсеместным внедрением интернета вещей сельскохозяйственные фермы становятся потенциальной мишенью для кибератак, однако эта тема лишь бегло упоминается в единичных работах [22]. В литературе полностью отсутствует анализ возможных сценариев атак, таких как подмена данных с сенсоров с целью уничтожения урожая или отключение систем жизнеобеспечения на животноводческих фермах. Не исследованы методы защиты для распределенных сетей с ограниченными вычислительными ресурсами, к которым относятся большинство периферийных устройств. Не существует также анализ систем страхования киберрисков, специфичных для агропромышленного комплекса. Между тем растущая зависимость от ИИ создает принципиально новые векторы уязвимости, реализация которых может привести к катастрофическим последствиям не только для отдельного хозяйства, но и для продовольственной безопасности целого региона.

7. Этические дилеммы в селекции и биотехнологиях

Использование искусственного интеллекта в генной инженерии и селекции открывает огромные



возможности для создания новых сортов растений, однако оно порождает целый ряд этических вопросов, которые пока остаются за рамками научного дискурса. В первую очередь не определено, кто должен нести ответственность за непредвиденные последствия использования нового сорта, «спроектированного» нейросетями. Не исследованы механизмы предотвращения монополизации генетических ресурсов крупными корпорациями и риски сокращения биоразнообразия при массовом переходе к ИИ-оптимизированным культурам. Остается открытым вопрос о необходимости и формах общественного контроля за подобными технологиями. От этих вопросов, лежащих на стыке науки, этики и философии, и от ответов на них напрямую зависит будущее биоразнообразия и продовольственной независимости целых регионов.

В целом, существующие исследования создают прочный фундамент, демонстрируя ЧТО ИИ может сделать. Следующий шаг — понять, КАК внедрять эти технологии ответственно, справедливо и безопасно для всех участников агропродовольственной системы.

ОБСУЖДЕНИЕ

Представленные количественные данные убедительно свидетельствуют о высоком потенциале ИИ в решении ключевых проблем сельского хозяйства. Во-первых, исключительно высокая точность моделей диагностики (до 99%) и прогнозирования позволяет перейти от реактивного к проактивному управлению, предотвращая потери урожая от болезней и неблагоприятных погодных условий. Во-вторых, экономический эффект от оптимизации ресурсов является двойным: прямое снижение затрат на закупку удобрений и пестицидов (до 30–90%) и косвенный экологический эффект от снижения химической нагрузки на почвы и водоемы. Это напрямую коррелирует с принципами устойчивого развития.

Особого внимания заслуживают результаты применения дронов. Помимо очевидной экономии времени (до 70%), данные указывают на качественное улучшение процесса обработки: увеличение осаждения действующего вещества на растениях более чем в 3 раза при одновременном снижении потерь при наземном сносе. Это демонстрирует, что ИИ и роботизация не просто автоматизируют, но и технологически совершенствуют агрономические процессы.

Однако, несмотря на впечатляющие результаты, внедрение ИИ сталкивается с рядом системных барьеров. Во-первых, это высокая стоимость внедрения и цифровое неравенство. Объемы венчурных инвестиций в AgriTech несопоставимы с фи-

нансированием ИТ-сектора, а для малых и средних фермерских хозяйств, особенно в развивающихся странах, такие технологии остаются недоступными без государственной поддержки. Во-вторых, это дефицит качественных данных и отсутствие унифицированных стандартов для их сбора и обмена. В-третьих, существует социальный аспект: необходимость переобучения кадров и преодоления психологического барьера у фермеров при переходе от традиционных методов к цифровым, при этом существуют институциональные барьеры, мешающие эффективному трансферу знаний между университетами и агробизнесом.

ВЫВОДЫ

В ходе комплексного анализа научных источников 2022–2026 гг. получены следующие основные результаты, подтверждающие ключевую роль искусственного интеллекта в трансформации сельскохозяйственной отрасли:

1. Систематизированы направления применения ИИ. Выявлено, что технологии ИИ охватывают все этапы производственной цепочки АПК: от селекции и прогнозирования урожайности до мониторинга посевов и автоматизации сбора урожая. Наибольший потенциал демонстрируют нейросетевые алгоритмы (LSTM, CNN), компьютерное зрение (ResNet, YOLO) и роботизированные системы, интегрированные с IoT.

2. Количественно подтверждена высокая эффективность ИИ-технологий. В результате обработки эмпирических данных установлено:

- Точность диагностики заболеваний растений достигает 99,2%, что позволяет минимизировать потери урожая на ранних стадиях.
- Применение алгоритмов прогнозирования снижает ошибку оценки урожайности (MAPE) до 5,5–7,5%, обеспечивая эффективное стратегическое планирование.
- Внедрение систем точного земледелия обеспечивает сокращение использования пестицидов на 20–90%, экономию водных ресурсов на 20–46% и снижение затрат на удобрения до 30%.
- Использование дронов повышает производительность полевых работ на 20–30%, сокращает время обработки на 70% и увеличивает эффективность осаждения средств защиты растений на 309% при одновременном снижении потерь от сноса на 54%.
- Экономическая отдача от внедрения ИИ выражается в росте доходов фермерских хозяйств на 15–20% и рентабельности инвестиций (ROI) на уровне 20–30%.

3. Классифицированы барьеры внедрения ИИ в разрезе групп стран. Впервые проведено срав-

нительное исследование, показавшее, что характер проблем существенно различается в зависимости от уровня развития экономики:

- В развитых странах доминируют проблемы технологической фрагментации, дефицита кадров «двойного профиля» и правовой неопределенности в отношении данных.

- В странах БРИКС ключевыми ограничениями являются неравномерность цифровой инфраструктуры, зависимость от импортных технологий и разрыв между наукой и производством.

- В развивающихся странах фундаментальными препятствиями выступают отсутствие базовой инфраструктуры (интернет, электричество), низкая платежеспособность мелких хозяйств и недостаточный уровень цифровой грамотности.

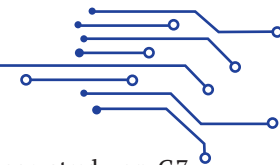
4. Выявлены перспективные направления для дальнейших исследований. Установлено, что наименее изученными остаются следующие проблемные области:

- социально-психологические и культурные механизмы принятия ИИ-решений фермерами;
- оценка совокупного экологического следа самих ИИ-технологий (производство оборудования, энергопотребление, утилизация);
- модели монетизации данных и правовые механизмы регулирования цифрового суверенитета в АПК;
- экономическая эффективность альтернативных бизнес-моделей (кооперация, аренда, pay-per-use) для малых форм хозяйствования;
- методы интеграции гетерогенных данных в реальном времени и вопросы кибербезопасности «умных ферм».

Иначе говоря, в исследовании не только обобщены современные достижения в области применения ИИ в сельском хозяйстве, но и определены системные барьеры, преодоление которых является необходимым условием для перехода к устойчивому и высокотехнологичному агропроизводству.

СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Bayrak T., Olu M. Applications of artificial intelligence in smart agriculture: Plant health, drone technology, and digital communication. *Journal of Agricultural Production*. 2025;6(3):157-166. URL: <https://doi.org/10.56430/japro.1767174>
2. Ashe G., Mengistu D. Review on future farming using AI. *Asian Science Bulletin*. 2025;3(1):82-91. URL: <https://doi.org/10.3923/asb.2025.82.91>
3. Wildan J.F. A review: Artificial intelligence related to agricultural equipment integrated with the internet of things. *Journal of Advanced Technology and Multidisciplinary*. 2023;02(02):47-60. URL: <https://doi.org/10.20473/jatm.v2i2.51440>
4. Rozhkova A.V., et al. Prospects for the use of artificial intelligence in agriculture. *IOP Conference Series: Earth and Environmental Science*. 2022;1076(1):012051. URL: <https://doi.org/10.1088/1755-1315/1076/1/012051>
5. Kali S. Artificial intelligence in plant modelling and simulation. In: *Plant Digital Twins: The Fusion of Biology and Artificial Intelligence*; 2025. URL: <https://doi.org/10.55938/wlp.v2i7.310>
6. Шайтура С.В., Семичевская Н.П., Шайтура Н.С. Прогнозирование урожайности в регионах юга России с использованием инструментов искусственного интеллекта. *Цифровые решения и технологии искусственного интеллекта*. 2025;1(4):76-85. URL: <https://doi.org/10.26794/3033-7097-2025-1-4-76-85>
Shaitura S.V., Semichevskaya N.P., Shaitura N.S. Predicting crop yields in the southern regions of Russia with artificial intelligence tools. *Digital Solutions and Artificial Intelligence Technologies*. 2025;1(4):76-85. (In Russ.). URL: <https://doi.org/10.26794/3033-7097-2025-1-4-76-85>
7. Shebanina O., Tyshchenko S., Parkhomenko O., Khylo I., Krainii V. Application of artificial intelligence to improve the economic efficiency of land use management in the agricultural sector. *Ekonomika APK*. 2025;32(1):82-90. URL: <https://doi.org/10.32317/ekon.apk/1.2025.82>
8. Monsalve-Castro C., Ramirez Molina R.I., Fajardo Ortiz E.J., Soto Polo G.P. Artificial intelligence as a driver in adopting green innovation in the agricultural sector: evidence from an emerging economy. *Cogent Business & Management*. 2025;12(1):2575261. URL: <https://doi.org/10.1080/23311975.2025.2575261>
9. Zhang Q., Wang Q., Liang Y., Zhang K., Hu H. Artificial intelligence and the sustainable development of agricultural enterprises: a total factor productivity perspective. *Frontiers in Sustainable Food Systems*. 2026;10:1768115. URL: <https://doi.org/10.3389/fsufs.2026.1768115>
10. Saravanan P., Suguna Dr.M. A study on problems and prospects of agricultural marketing in the era of artificial intelligence and smart farming with special reference to Tamil Nadu. *International Journal of Accounting and Economics Studies*. 2025;12(4):554-561. URL: <https://doi.org/10.14419/jf9c9x58>



11. Caglar E. The impact of sectors on agriculture based on artificial intelligence data: a case study on G7 countries and Türkiye. *International Journal of Agriculture, Environment and Food Sciences*. 2024;8(3):486-494. URL: <https://doi.org/10.31015/jaefs.2024.3.1>
12. Slavova G., Todorova M. Opportunities for the application of artificial intelligence in the agricultural sector of Bulgaria. In: *The interdisciplinary approach in economic and social sciences*. 2024:323-333. URL: <https://doi.org/10.3897/ap.10.e0323>
13. Vedapathak D.M.D. Artificial intelligence (AI) and agricultural biotechnology. *The International Journal of Commerce Management and Business Law in International Research*. 2025;2(2):13-16. URL: <https://doi.org/10.5281/zenodo.15709415>
14. Harfouche Z., Benattallah A. Evaluating the efforts of utilizing artificial intelligence to enhance agricultural entrepreneurship: A case study of Tunisia. *Current Perspectives in Social Sciences*. 2025;29(4):732-753. URL: <https://doi.org/10.53487/atasobed.1734836>
15. Aladejebi O., Amao-Taiwo B., Oshinowo B. How artificial intelligence is revolutionizing agriculture in Nigeria? *Archives of Business Research*. 2026;14(02):139-157. URL: <https://doi.org/10.14738/abr.1402.20078>
16. Raihan A. A review of artificial intelligence for sustainable agriculture. *Proceedings of the International Conference on Climate-Aware Agriculture*. 2024. URL: <https://www.researchgate.net/publication/386989366>
17. Wahyuni R., Utomo S.W.A., Ghifari Z. Artificial neural network prediction model for agricultural commodity production using backpropagation algorithm. *Knowbase: International Journal of Knowledge in Databases*. 2025;5(01):52-68. URL: <https://doi.org/10.30983/knowbase.v5i1.9530>
18. Singh V., Choudhary S., Kumar H. Utilizing artificial intelligence technologies in the Indian agriculture sector. 2024. URL: <https://doi.org/10.13140/RG.2.2.24666.56001>
19. Yadav S. Introduction to plant digital twins: Concept and evolution. In: *Plant Digital Twins: The Fusion of Biology and Artificial Intelligence*. Wisdom Leaf Press; 2025:1-25. URL: <https://doi.org/10.55938/wlp.v2i7.308>
20. Kumar A., et al. Role of artificial intelligence in vegetable production: A review. *Journal of Scientific Research and Reports*. 2024;30(9):950-963. URL: <https://doi.org/10.9734/jsrr/2024/v30i92423>
21. González-Rodríguez V.E., Izquierdo-Bueno I., Cantoral J.M. et al. Artificial intelligence: A promising tool for application in phytopathology. *Horticulturae*. 2024;10:197. URL: <https://doi.org/10.3390/horticulturae10030197>
22. Istudor N., Ignat R., Deaconu M., Constantin M. Risk-management in the era of artificial intelligence in agriculture. *Proceedings of the International Conference on Agriculture and Food Engineering*. 2023. URL: <https://doi.org/10.24818/CAFEE/2023/12/08>

ИНФОРМАЦИЯ ОБ АВТОРЕ / ABOUT THE AUTHOR

Альбина Хатмулловна Шелепаева — кандидат педагогических наук, доцент кафедры бизнес-информатики, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

Albina Kh. Shelepaeva — Cand. Sci. (Education), Assoc. Prof. of the Department of Business Informatics, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

<https://orcid.org/0000-0002-4678-9671>

akshelepaeva@fa.ru

Конфликт интересов: автор заявляет об отсутствии конфликта интересов.

Conflicts of Interest Statement: The author has no conflicts of interest to declare.

Статья поступила 26.01.2026; после рецензирования 19.05.2026; принята к публикации 22.07.2026.

Автор прочитала и одобрила окончательный вариант рукописи.

The article was submitted on 26.01.2026; revised on 19.05.2026 and accepted for publication on 22.07.2026.

The author read and approved the final version of the manuscript.

DOI: 10.26794/3030-7097-2026-2-3-26-33
УДК 004.056.5(045)

Обезличивание персональных данных в банковской сфере: анализ методов и выбор стратегии внедрения

Е.К. Баранова, Я.В. Лебедкина

Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

АННОТАЦИЯ

В статье рассматривается актуальная проблема обезличивания персональных данных в кредитно-финансовых организациях (КФО) в контексте ужесточения регуляторных требований, в частности Приказа Роскомнадзора № 140. Проведен анализ классических и современных методов обезличивания (маскирование, псевдонимизация, генерация синтетических данных и др.) с точки зрения их применимости в банковской среде, где критически важны сохранение формата данных, референциальная целостность и высокая производительность. Дана сравнительная оценка трех современных отечественных решений: DataMask, «Гарда Data Masking» и N1 AI. Выявлены их системные ограничения, включая потенциальное несоответствие новым нормативным актам, недостаточную гибкость для уникальных банковских сущностей и высокую стоимость владения. На основе проведенного анализа обосновывается стратегическая целесообразность разработки специализированного внутрикорпоративного инструмента обезличивания, который гарантирует полное соответствие регулированию, максимальную адаптивность к бизнес-процессам банка и долгосрочную экономическую эффективность. Для кредитно-финансовой организации характерна обработка обширного массива структурированных персональных данных (ПДн), совмещающих признаки общих, иных данных и сведений, составляющих банковскую тайну. Данные сведения, как правило, хранятся в реляционных базах данных в формализованном виде (отдельные поля для номера паспорта, телефона, счета и т.д.), что, с одной стороны, упрощает их автоматизированный поиск, а с другой — требует применения строгих мер защиты, к которым относится и безопасное обезличивание для использования в непроизводственных средах. Это определяет ключевые требования к рассматриваемой системе: необходимость точного обнаружения полей различных категорий финансово-идентифицирующей информации и применение алгоритмов маскировки, обеспечивающих необратимость преобразования при сохранении структурной целостности и формата данных для корректной работы тестовых систем.

Ключевые слова: обезличивание персональных данных; псевдонимизация; маскирование данных; банковские информационные системы; информационная безопасность; защита информации

Для цитирования: Баранова Е.К., Лебедкина Я.В. Обезличивание персональных данных в банковской сфере: анализ методов и выбор стратегии внедрения. *Цифровые решения и технологии искусственного интеллекта*. 2026;2(3):26-33. DOI: 10.26794/3030-7097-2026-2-3-26-33

ORIGINAL PAPER

Depersonalization of Personal Data in the Banking Sector: Analysis of Methods and Selection of an Implementation Strategy

Е.К. Baranova, Ya.V. Lebedkina

Financial University under the Government of the Russian Federation, Moscow, Russian Federation

ABSTRACT

This article examines the pressing issue of personal data anonymization in credit and financial institutions (CFIs) in the context of tightening regulatory requirements, in particular, Roskomnadzor Order No. 140. It analyzes classic and modern anonymization methods (masking, pseudonymization, synthetic data generation, etc.) for their applicability in a banking environment where preserving data formats, referential integrity, and high performance are critical. A comparative assessment of three modern domestic solutions is provided: DataMask, Garda Data Masking, and N1 AI. Their systemic limitations are identified, including potential non-compliance with new regulations, insufficient flexibility for unique banking entities, and high cost of ownership. Based on this analysis, the strategic feasibility of developing a specialized internal anonymization tool is substantiated. This tool guarantees full regulatory compliance, maximum adaptability to the bank's business processes, and long-term economic efficiency. A credit and financial institution is characterized by processing an extensive array of structured personal data (PD), which combines the features of general, other data, and information that

constitutes bank secrecy. This information is typically stored in relational databases in a formalized format (separate fields for passport number, phone number, account number, etc.), which, on the one hand, simplifies their automated search, but on the other hand, requires strict protection measures, including secure anonymization for use in non-production environments. This determines the key requirements for the system under consideration: the need to accurately detect fields of various categories of financial and identifying information, and the use of masking algorithms that ensure irreversible conversion while preserving the structural integrity and data format for the correct operation of test systems.

Keywords: anonymization of personal data; pseudonymization; data masking; banking information systems; information security; information protection

For citation: Baranova E.K., Lebedkina Ya.V. Depersonalization of personal data in the banking sector: Analysis of methods and selection of an implementation strategy. *Digital Solutions and Artificial Intelligence Technologies*. 2026;2(3):26-33. DOI: 10.26794/3030-7097-2026-2-3-26-33

ВВЕДЕНИЕ

В современной цифровой экономике данные стали ключевым активом, а их защита — одним из главных приоритетов, особенно для кредитно-финансовых организаций. Обработывая огромные массивы персональных данных (ПДн), банки несут строгую юридическую ответственность за их конфиденциальность. Федеральный закон «О персональных данных»¹ устанавливает базовые принципы обработки ПДн, а обезличивание признается одним из основных методов приведения их в соответствие с законом. Согласно статье 3 Федерального закона «О персональных данных», обезличивание — это действия, в результате которых становится невозможным без использования дополнительной информации определить принадлежность персональных данных конкретному субъекту.

Однако на практике перед ИТ-специалистами и архитекторами банков встает сложный комплекс задач. С одной стороны, необходимо безусловно выполнять регуляторные требования, которые постоянно ужесточаются (яркий пример — детализированный Приказ Роскомнадзора № 140², введенный с 01.09.2025 г.). С другой — обезличенные данные должны сохранять свою прикладную ценность для тестирования программного обеспечения, анализа бизнес-процессов и работы с данными в производственных средах. Традиционное решение — покупка готового коммерческого продукта — все чаще наталкивается на специфику высоко динамичной и сложно устроенной банковской ИТ-инфраструктуры.

Цель данной статьи — проанализировать существующие методы и рыночные решения для

обезличивания данных, оценить их применимость в условиях конкретных требований банковского сектора и обосновать выбор стратегии внедрения. Работа построена на сопоставлении функциональных возможностей, нормативного соответствия и экономических аспектов, что позволит сформировать объективный взгляд на проблему и предложить прагматичный путь ее решения.

КЛАССИФИКАЦИЯ МЕТОДОВ ОБЕЗЛИЧИВАНИЯ

Современные подходы и нормативные акты, такие как Приказ Роскомнадзора № 140, позволяют систематизировать методы обезличивания по их принципу действия и области применения (см. рисунок).

Наибольшее распространение на практике получили статические методы. Данные методы предполагают необратимое преобразование данных в рамках отдельного набора, например, копии производственной базы, с созданием нового, обезличенного массива информации. К ним относятся:

- **Маскирование.** Этот метод заключается в частичном или полном сокрытии оригинального значения путем замены символов на статические или динамические шаблоны. Различают несколько видов:

- а) маскирование по формату: замена части символов на фиксированные знаки (например, «*» или «X»). Например, номер паспорта «4012 123456» преобразуется в «XXXX XXXXX6». Данный метод прост в реализации, но может снизить полезность данных для тестирования бизнес-логики, зависящей от формата;

- б) детерминированное маскирование: замена исходного значения на одно и то же псевдонимизированное значение во всех системах и наборах данных. Например, ФИО «Иванов Сергей» всегда преобразуется в «Петров Андрей». Это критически важно для сохранения референциальной целостности между связанными таблицами базы данных;

- в) случайное маскирование: замена на случайное, но правдоподобное значение из заданного домена (словаря). Например, замена реального имени на случайное из списка распространенных имен.

¹ Федеральный закон от 27.07.2006 № 152-ФЗ (ред. от 24.02.2025) «О персональных данных». URL: <http://www.kremlin.ru/acts/bank/24154>

² Приказ Роскомнадзора от 19.06.2025 № 140 «Об утверждении требований к обезличиванию персональных данных и методов обезличивания персональных данных, за исключением случаев, указанных в п. 9.1 ч. 1 ст. 6 Федерального закона от 27.07.2006 № 152-ФЗ «О персональных данных»». URL: <http://publication.pravo.gov.ru/document/0001202508010002>

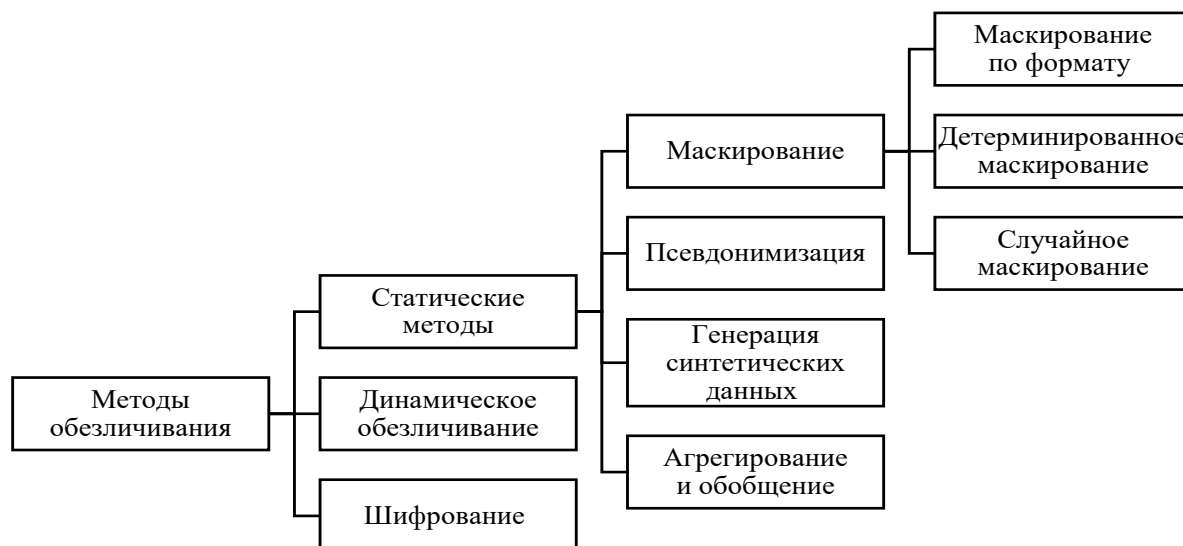
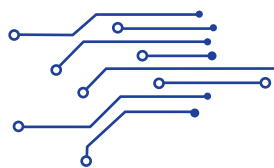


Рис. / Fig. Классификация методов обезличивания / Classification of Depersonalization Methods

Источник / Source: составлено авторами / Compiled by the authors.

- **Псевдонимизация.** В отличие от простого маскирования, псевдонимизация часто предполагает использование алгоритмов шифрования или хеширования с ключом («солью») для замены прямого идентификатора на устойчивый псевдоним. Исходные данные могут быть восстановлены только при наличии ключа, хранящегося отдельно. Это основной метод, соответствующий концепции GDPR (General Data Protection Regulation). Пример: замена e-mail «client@domain.com» на детерминированный хеш «a1b2c3d4e5».

- **Генерация синтетических данных** предполагает создание полностью искусственного, но статистически похожего на реальный набор данных с помощью специальных алгоритмов. Синтетические данные не содержат реальной информации о клиентах, что полностью исключает риски, однако их создание, адекватно отражающее все сложности и аномалии реальных данных, является нетривиальной задачей.

- **Агрегирование и обобщение.** Данные заменяются на обобщенные категории или статистические показатели. Например, точный возраст заменяется на возрастную группу «25–35 лет», а конкретная сумма транзакции — на диапазон «от 10 000 до 50 000 руб.». Метод эффективен для аналитических целей, но бесполезен для функционального тестирования приложений.

Динамическое обезличивание предполагает маскировку данных «на лету» в момент выполнения запроса к базе данных, без изменения самих хранимых данных. Правила маскирования применяются на уровне СУБД или промежуточного программного слоя в зависимости от прав доступа пользователя. Например, оператор колл-центра ви-

дит только последние четыре цифры номера карты, в то время как сотрудник бэк-офиса имеет доступ к полному номеру. Этот метод оптимален для разграничения доступа в производственных системах, но не подходит для создания обезличенных копий для тестирования.

Криптографическое преобразование данных является надежным методом защиты. Для целей обезличивания может применяться форматированное шифрование, которое позволяет сохранить исходный формат и тип данных (например, шифрует номер кредитной карты, оставляя валидный с точки зрения контрольной суммы номер). Восстановление данных возможно только при наличии ключа дешифрования.

При разработке системы обезличивания для кредитно-финансовых организаций (КФО) необходимо учитывать специфические требования [1–3].

1. **Сохранение формата и валидности:** обезличенные номера паспортов, ИНН, банковских карт должны сохранять контрольные суммы и структуру для успешного прохождения валидации в тестовых контурах.

2. **Обеспечение референциальной целостности:** связи между таблицами (например, «client_id» в таблицах «clients» и «accounts») не должны быть разорваны после обезличивания.

3. **Детерминированность:** при необходимости проведения согласованного тестирования одинаковые исходные данные должны маскироваться в одинаковые значения при каждом запуске процесса.

4. **Необратимость:** в соответствии с Приказом № 140 процесс должен исключать возможность восстановления ПДн без специальных средств, хранящихся отдельно.



5. Производительность: алгоритмы должны эффективно обрабатывать большие объемы данных, характерные для КФО.

Таким образом, для реализации скрипта обезличивания — наиболее целесообразным представляется применение комбинации статических методов, соответствующих Приказу Роскомнадзора № 140. Ядро системы будет составлять маскирование по формату или псевдонимизацию по словарям для атрибутов, требующих сохранения правдоподобия (ФИО, адреса, номера документов)³ [4]. Такой подход позволит обеспечить баланс между требованиями безопасности, функциональной полезности обезличенных данных и производительности.

АНАЛИЗ СУЩЕСТВУЮЩИХ РЕШЕНИЙ И ИХ ПРИМЕНИМОСТЬ В БАНКЕ

В условиях ужесточения регуляторных требований и роста угроз информационной безопасности рынок предлагает КФО ряд готовых решений для обезличивания персональных данных. Выбор стратегии — внедрение коммерческого продукта или разработка собственного инструмента — требует тщательного анализа существующих предложений [5–7], а именно — анализ трех актуальных отечественных решений: DataMask (Secret Technologies)⁴, «Гарда Data Masking» (Гарда Технологии)⁵ и N1 AI (ООО «Умные решения»)⁶. Целью анализа является оценка их функциональной полноты, соответствия специфическим требованиям банковского сектора и обоснование экономической и технической целесообразности разработки специализированного внутрикорпоративного решения.

Современный рынок решений для защиты данных [8–10] можно условно разделить на два класса: системы, ориентированные на работу со структурированными данными в СУБД, и платформы, использующие искусственный интеллект (ИИ) для обработки неструктурированной информации.

1. DataMask (Secret Technologies)

Решение позиционируется как универсальный инструмент для автоматического обнаружения и маскирования более 30 типов чувствительных данных в реляционных базах данных. К его декларируемым преимуществам относятся:

- автоматизация, т.е. сквозной процесс обнаружения и маскирования;
- обеспечение референциальной целостности данных между связанными таблицами;
- поддержка различных СУБД (PostgreSQL, Oracle, MS SQL) и возможность встраивания в конвейеры CI/CD через API;
- реалистичность, однородность, необратимость.

2. «Гарда Data Masking» (Гарда Технологии)

Более зрелый продукт, представленный как отраслевой стандарт для безопасной передачи данных третьим лицам. Его отличительные черты:

- автоматическое выявление обширного спектра данных: от персональных (ФИО, паспорт) до финансовых (номера счетов, карт) и юридических (реквизиты организаций);
- заявленное соответствие широкому кругу стандартов, включая Федеральный закон «О персональных данных» GDPR, PCI DSS, СТО БР ИББС;
- поддержка пользовательских алгоритмов маскирования и ролевой модели доступа.

3. N1 AI (ООО «Умные решения»)

Данное решение представляет собой принципиально иной подход, основанный на технологиях искусственного интеллекта и машинного обучения, сфокусированный на обработке текстовых данных. Основные особенности:

- фокус на неструктурированные данные: предназначен для обезличивания текстов, документов (PDF, Word), датасетов, что актуально для сканов договоров, корреспонденции, расшифровок переговоров;
- использование NLP-моделей (Natural Language Processing), обученных на русском языке, с заявленной точностью распознавания сущностей свыше 99%, включая обработку опечаток;
- предлагает как полное замещение, так и псевдонимизацию с идентификаторами, а также детектирование позиций ПДн;
- предоставление REST API для обезличивания в реальном времени и возможность дообучения модели на данных заказчика.

Для системной оценки необходимо провести сравнение решений по ключевым критериям [11, 12], критически важным для банка. Результаты представлены в *таблице*.

Проведенный анализ выявляет ряд системных ограничений готовых решений в контексте высокоспециализированной среды КФО:

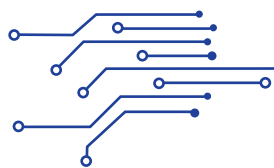
1. Проблема гарантированного соответствия актуальному регулированию. Приказ Роскомнадзора № 140 устанавливает детализированные, юридически обязательные требования к методам и необратимости обезличивания. Большинство

³ Шаблинский И.Г. Правовое регулирование информационных отношений в сфере обработки персональных данных. Учебное пособие для вузов. М.: Издательство Юрайт; 2024.

⁴ Диджитал Маскинг. URL: <https://d-mask.ru>

⁵ Garda.ai. Маскирование данных. URL: <https://garda.ai/products/data-protection/datamasking>

⁶ Портал о технологиях больших данных N1.AI. URL: <https://n1ai.ru/bigdata>



Сравнительный анализ решений для обезличивания данных /
Comparative Analysis of Solutions for Data Anonymization

Критерий анализа / The Criterion of Analysis	DataMask	«Гарда Data Masking»	N1 AI	Ключевые требования для банка / Key Requirements for the Bank
Основной объект обработки	Структурированные данные в СУБД		Неструктурированные/ текстовые данные (документы, текст)	Структурированные данные в СУБД
Ключевая технология	Детерминированные алгоритмы, словари, правила	Детерминированные алгоритмы, шаблоны	Модели искусственного интеллекта (NLP)	Детерминированные алгоритмы, словари, правила, шаблоны
Соответствие Приказу РКН № 140	Не заявляется производителем. Требуется валидации		Требуется глубокой валидации. Объяснимость и необратимость ИИ- методов — сложная задача	Гарантированное и документально подтвержденное соответствие Приказу Роскомнадзора № 140
Покрывание специфических банковских сущностей	Ограничено предустановленными шаблонами. Расширение возможно через вендора	Широкий, но фиксированный набор. Пользовательские правила требуют настройки	Возможно через дообучение модели, что требует экспертизы Data Science	Полное и адаптивное. Необходима возможность мгновенного добавления правил для новых продуктов
Гибкость и скорость адаптации	Зависит от вендора. Обновление правил — часть процесса поддержки	Настройка параметров доступна, изменение ядра — у вендора	Высокая за счет дообучения, но требует данных и вычислительных ресурсов	Максимальная независимость от цикла разработки вендора. Возможность изменений силами команды банка
Экономическая модель	Высокие CAPEX (лицензии) и OPEX (ежегодная поддержка, обновления)	Индивидуальный расчет, масштабируемая стоимость	Лицензирование/ подписка, возможны затраты на дообучение и инфраструктуру для ИИ	Первоначальные инвестиции в разработку с последующей минимальной стоимостью владения

Источник / Source: составлено авторами / Compiled by the authors.

коммерческих решений были разработаны до его выхода. Их адаптация под новые требования носит для банка характер «черного ящика» и требует дополнительных затрат на валидацию. Для ИИ-решений, подобных N1 AI, доказательство необратимости и объяснимости алгоритмов представляет отдельную научно-практическую проблему.

2. Бизнес-процессы банка порождают уникальные категории конфиденциальных данных, не входящие в стандартные шаблоны:

- номера внутренних кредитных договоров, заявок, депозитных счетов со сложной структурой;

- коды валютных операций, признаки бенефициарного владения;
- структурированные описания залогового имущества.

Эффективное и оперативное управление правилами для таких сущностей в рамках «коробочного» продукта затруднено.

3. Инфраструктура банка — сложный организм с унаследованными системами, строгими регламентами переноса данных между изолированными контурами (Prod/Dev/Test). Интеграция стороннего продукта, даже с API, часто превращается в длительный и дорогой проект



по адаптации банка под логику работы вендора, а не наоборот.

Учитывая выявленные системные ограничения коммерческих продуктов, разработка собственной автоматизированной системы обезличивания представляется для банка не только целесообразной, но и стратегически необходимой.

ВЫВОДЫ

Проведенное исследование позволяет сделать ряд значимых выводов относительно обезличивания данных в банковской сфере. Современные нормативные требования, особенно Приказ Роскомнадзора № 140, задают высокую планку для методов обезличивания, делая акцент на их необратимости и доказательной базе. Ни один из рассмотренных классических методов (маскирование, псевдонимизация, агрегирование) не является универсальным; их эффективное применение требует комбинированного подхода, тщательно настроенного под типы обрабатываемых данных и целевые сценарии использования.

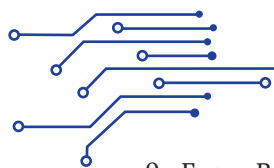
Сравнительный анализ ведущих отечественных коммерческих решений (DataMask, «Гарда Data Masking», N1 AI) демонстрирует, что, несмо-

тря на заявленную функциональность, они в полной мере не отвечают уникальным запросам крупного банка. Ключевыми проблемами остаются негибкость в адаптации под специфические банковские сущности, сложности с гарантированным соответствием последним редакциям нормативных актов, а также высокие совокупные затраты на интеграцию и владение [13, 14].

Таким образом, для кредитно-финансовой организации, обладающей развитой ИТ-экспертизой, наиболее рациональным и перспективным путем представляется разработка собственной автоматизированной системы обезличивания. Такой подход обеспечит не просто выполнение формальных требований регулятора, но и создаст стратегическое конкурентное преимущество. Он даст банку полный контроль над алгоритмами, позволит мгновенно реагировать на изменения в продуктовой линейке и регулировании, а в долгосрочной перспективе окажется экономически более выгодным, чем зависимость от внешних вендоров. Реализация подобного проекта требует значимых первоначальных инвестиций, но формирует устойчивую основу для безопасной и эффективной работы с данными во всех контурах информационной системы.

СПИСОК ИСТОЧНИКОВ

1. Альмухаметова Э.И., Шаймарданов И.Ф. Актуальность и методы обезличивания персональных данных. В сб.: Передовые инженерные решения в нефтепереработке и нефтехимии. Салават: Уфимский государственный нефтяной технический университет; 2025:261-263. URL: <https://elibrary.ru/ffypro>
2. Алексеев Д.В., Тумбинская М.В. Современные методы обезличивания персональных данных: анализ эффективности и направления развития. В сб.: Цифровые системы и модели: теория и практика проектирования, разработки и использования. Казань: Казанский государственный энергетический университет; 2025:2036-2039. URL: <https://elibrary.ru/wldthv>
3. Колбасов В.В. Персональные данные: обезличивание персональных данных, обрабатываемых в информационных системах персональных данных. *Некоммерческие организации в России*. 2017;1:59-62. URL: <https://elibrary.ru/zegwvv>
4. Кучин И.Ю. Защита конфиденциальности персональных данных с помощью обезличивания. *Вестник Астраханского государственного технического университета. Серия: Управление, вычислительная техника и информатика*. 2010;2:158-162. URL: <https://elibrary.ru/muyqhf>
5. Полуянова Е.В. Актуальные вопросы, связанные с обезличиванием персональных данных в Российской Федерации. *Тенденции развития науки и образования*. 2024;114-6:36-38. URL: <https://doi.org/10.18411/trnio-10-2024-243>
6. Чепик П.И. Обезличивание персональных данных как способ повышения их защиты при обработке в информационных системах. В сб.: Безопасные информационные технологии. М.: Московский государственный технический университет им. Н.Э. Баумана; 2023:150-153. URL: <https://elibrary.ru/tktuem>
7. Алексеев Д.В., Тумбинская М.В. Современные методы обезличивания персональных данных: анализ эффективности и направления развития. В сб.: Цифровые системы и модели: теория и практика проектирования, разработки и использования. Казань; 2025:2036-2039. URL: <https://elibrary.ru/wldthv>
8. Борисов А.В. Применение имитационного компьютерного моделирования к задаче обезличивания персональных данных. Оценка состояния и Основные положения. *Программирование*. 2023;4:58-74. URL: <https://doi.org/10.31857/S0132347423040040>



9. Бесов В.В., Данилова Э.З., Мекуриа Д.В. Методы обезличивания пользовательских данных: анализ применимости в условиях современных требований к приватности. *Телекоммуникации и информационные технологии*. 2025;12(1):84-89. URL: <https://elibrary.ru/xjkqtn>
10. Мищенко Е.Ю., Соколов А.Н. Алгоритмы реализации методов обезличивания персональных данных в распределенных информационных системах. *Доклады Томского государственного университета систем управления и радиоэлектроники*. 2019;22(1):66-70. URL: <https://doi.org/10.21293/1818-0442-2019-22-1-7-66-70>
11. Малявко А.А., Реутов В.В., Коротких И.В., Шперлинг В.К. Алгоритмы, методы и подходы к обезличиванию и обогащению данных, в том числе персональных. *Безопасность цифровых технологий*. 2022;4(107):9-26. URL: <https://doi.org/10.17212/2782-2230-2022-4-9-26>
12. Гусева Т.М., Карабанова А.Н., Блимготов Т.К. Защита персональных данных путем их обезличивания. В сб.: *Теория и практика применения новых информационных технологий*. Ставрополь: Издательство «АГРУС»; 2020:103-106. URL: <https://elibrary.ru/jssuvn>
13. Маршалко Г.Б., Дали Ф.А., Савиных А.Н. Сравнительный анализ методов обезличивания персональных данных. В сб.: *Методы и технические средства обеспечения безопасности информации*. СПб.; 2022:161-163. URL: <https://elibrary.ru/wppbj>
14. Учаева Е.С., Фот Ю.Д. Исследование методов обезличивания персональных данных. В сб.: *Актуальные вопросы обеспечения комплексной безопасности*. Оренбург; 2025:1570-1574. URL: <https://elibrary.ru/eevae>

REFERENCES

1. Almukhametova E.I., Shaimardanov I.F. Relevance and methods of depersonalization of personal data. In: *Advanced engineering solutions in oil refining and petrochemistry*. Salavat: Ufa State Petroleum Technical University; 2025:261-263. URL: <https://elibrary.ru/ffyp> (In Russ.).
2. Alekseev D.V., Tumbinskaya M.V. Modern methods of depersonalization of personal data: analysis of effectiveness and development directions. In: *Digital Systems and Models: theory and practice of design, development and use*. Kazan: Kazan State Power Engineering University; 2025:2036-2039. URL: <https://elibrary.ru/wldthv> (In Russ.).
3. Kolbasov V.V. Personal data: depersonalization of personal data processed in personal data information systems. *Non-profit organizations in Russia*. 2017;1:59-62. URL: <https://elibrary.ru/zegwv> (In Russ.).
4. Kuchin I.Y. Protection of the confidentiality of personal data through depersonalization. *Bulletin of the Astrakhan State Technical University. Series: Management, Computer Engineering and Computer Science*. 2010;2:158-162. URL: <https://elibrary.ru/muyqhf> (In Russ.).
5. Poluyanov E.V. Topical issues related to depersonalization of personal data in the Russian Federation. *Trends in the Development of Science and Education*. 2024;114-6:36-38. (In Russ.). URL: <https://doi.org/10.18411/trnio-10-2024-243>
6. Chepik P.I. Depersonalization of personal data as a way to increase their protection during processing in information systems. In: *Secure Information technologies*. Moscow: Bauman Moscow State Technical University; 2023:150-153. URL: <https://elibrary.ru/tktuem> (In Russ.).
7. Alekseev D.V., Tumbinskaya M.V. Modern methods of depersonalization of personal data: analysis of effectiveness and development directions. In: *Digital Systems and Models: Theory and practice of design, development and use*. Kazan; 2025:2036-2039. URL: <https://elibrary.ru/wldthv> (In Russ.).
8. Borisov A.V. Application of simulation computer modeling to the problem of depersonalization of personal data. Assessment of the condition and the main provisions. *Programming*. 2023;4:58-74. (In Russ.). URL: <https://doi.org/10.31857/S0132347423040040>
9. Besov V.V., Danilova E.Z., Mekuria D.V. Methods of depersonalization of user data: an analysis of applicability in the context of modern privacy requirements. *Telecommunications and Information Technology*. 2025;12(1):84-89. URL: <https://elibrary.ru/xjkqtn> (In Russ.).
10. Mishchenko E.Yu., Sokolov A.N. Algorithms for implementing methods of depersonalization of personal data in distributed information systems. *Reports of Tomsk State University of Control Systems and Radio Electronics*. 2019;22(1):66-70. (In Russ.). URL: <https://doi.org/10.21293/1818-0442-2019-22-1-7-66-70>
11. Malyavko A.A., Reutov V.V., Korotkov I.V., Shperling V.K. Algorithms, methods and approaches to depersonalization and enrichment of data, including personal data. *Digital Technology Security*. 2022;4(107):9-26. (In Russ.). URL: <https://doi.org/10.17212/2782-2230-2022-4-9-26>

12. Guseva T.M., Karabanova A.N., Blimgotov T.K. Protection of personal data by depersonalization. In: Theory and practice of the application of new information technologies. Stavropol: AGRUS Publishing House; 2020:103-106. URL: <https://elibrary.ru/jssuvm> (In Russ.).
13. Marshalko G.B., Dali F.A., Savinykh A.N. Comparative analysis of methods of depersonalization of personal data. In: Methods and technical means of ensuring information security. St. Petersburg; 2022:161-163. URL: <https://elibrary.ru/wppbj> (In Russ.).
14. Uchaeva E.S., Fot Yu.D. Research of methods of depersonalization of personal data. In: Current issues of ensuring integrated security. Orenburg; 2025:1570-1574. URL: <https://elibrary.ru/eevae> (In Russ.).

ИНФОРМАЦИЯ ОБ АВТОРАХ / ABOUT THE AUTHORS

Елена Константиновна Баранова — доцент кафедры информационной безопасности, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

Elena K. Baranova — Assoc. Prof., Department of Information Security, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

<http://orcid.org/0000-0003-4650-2623>

Автор для корреспонденции / Corresponding author:

ekbaranova@fa.ru

Яна Витальевна Лебедкина — студент бакалавриата кафедры информационной безопасности, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

Yana V. Lebedkina — bachelor student, Department of Information Security, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

<http://orcid.org/0000-0002-539-0457>

228958@edu.fa.ru

Заявленный вклад авторов:

Е.К. Баранова — разработка общей концепции статьи, обзор публикаций по заявленной теме.

Я.В. Лебедкина — классификация методов обезличивания ПДн; анализ существующих решений и их применимость в банке.

Authors' declared contributions:

E.K. Baranova — development of the general concept of the article, and review of publications on the stated topic.

Ya.V. Lebedkina — classification of methods for anonymizing personal data; analysis of existing solutions and their applicability in the bank.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

Статья поступила 22.04.2026; после рецензирования 10.06.2026; принята к публикации 01.08.2026.

Авторы прочитали и одобрили окончательный вариант рукописи.

The article was submitted on 22.04.2026; revised on 10.06.2026; accepted for publication on 01.08.2026.

The authors read and approved the final version of the manuscript.



DOI: 10.26794/3030-7097-2026-2-3-34-40
УДК 004.056+004.738.5(045)

Организационно-технологический механизм безопасной интеграции инфраструктуры доверия цифрового рубля в экосистему кредитно-финансовой организации

А.В. Иванов, В.А. Апулина

Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

АННОТАЦИЯ

Внедрение цифрового рубля формирует перед кредитно-финансовыми организациями (КФО) вызовы, обусловленные необходимостью интеграции внешней инфраструктуры доверия во внутренние автоматизированные банковские системы (АБС). Проблема заключается в отсутствии типовых решений, регламентирующих безопасное сопряжение двух разнородных доменов безопасности. Новизна работы состоит в разработке целостного подхода, синтезирующего технологические меры защиты и организационные процедуры в единый механизм, реализуемый на стороне КФО. Целью статьи является формирование указанного механизма, обеспечивающего минимизацию рисков несанкционированного доступа, утечки информации и нарушения непрерывности бизнес-процессов. Методологическую основу составляют системный анализ архитектурных решений, синтез методов моделирования угроз и риск-ориентированный подход. В результате обоснован двухконтурный механизм защиты, включающий технологическое ядро (аппаратная криптография, однонаправленный шлюз) и организационный каркас (политика нулевого доверия, встроенный комплаенс). Доказано, что предложенный механизм снижает совокупную критичность векторов атак на этапе интеграции на 35–40% по шкале CVSS и формирует основу для тиражируемого типового решения.

Ключевые слова: цифровой рубль; информационная безопасность; кредитно-финансовая организация; инфраструктура доверия; киберустойчивость; платформа цифровой валюты центрального банка; организационно-технологический механизм

Для цитирования: Иванов А.В., Апулина В.А. Организационно-технологический механизм безопасной интеграции инфраструктуры доверия цифрового рубля в экосистему кредитно-финансовой организации. *Цифровые решения и технологии искусственного интеллекта*. 2026;2(3):34-40. DOI: 10.26794/3030-7097-2026-2-3-34-40

ORIGINAL PAPER

An Organizational and Technological Mechanism for the Secure Integration of the Digital Ruble's Trust Infrastructure into the Ecosystem of a Credit and Financial Institution

A.V. Ivanov, V.A. Aptulina

Financial University under the Government of the Russian Federation, Moscow, Russian Federation

ABSTRACT

The introduction of the digital ruble poses challenges for credit and financial institutions (CFIs) driven by the need to integrate external trust infrastructure into internal automated banking systems (ABS). The problem lies in the absence of standardized solutions regulating the secure interconnection of two heterogeneous security domains. The novelty of this work consists in the development of a holistic approach synthesizing technological protection measures and organizational procedures into a unified mechanism implemented on the CFI side. The objective of the article is to develop this mechanism, ensuring the minimization of risks of unauthorized access, information leakage, and disruption of business process continuity. The methodological framework is based on a systematic analysis of architectural solutions, a synthesis of threat modeling methods, and a risk-oriented approach. As a result, a two-circuit protection mechanism is substantiated, comprising a technological core (hardware cryptography, unidirectional gateway) and an organizational framework (zero-trust policy, in-line compliance). It has been proven that the proposed mechanism reduces the aggregate criticality of attack vectors at the integration stage by 35–40% according to the CVSS scale and forms the basis for a replicable standard solution.

Keywords: digital ruble; information security; credit and financial institution; trust infrastructure; cyber resilience; central bank digital currency platform; organizational and technological mechanism

For citation: Ivanov A.V., Aptulina V.A. An organizational and technological mechanism for the secure integration of the digital ruble's trust infrastructure into the ecosystem of a credit and financial institution. *Digital Solutions and Artificial Intelligence Technologies*. 2026;2(3):34-40. DOI: 10.26794/3030-7097-2026-2-3-34-40

© Иванов А.В., Апулина В.А., 2026



ВВЕДЕНИЕ

Актуальность настоящего исследования обусловлена переходом российской финансовой системы к принципиально новой парадигме расчетов, связанной с внедрением цифрового рубля¹. Концепция цифрового рубля², разработанная Банком России, предполагает создание гибридной розничной модели, где регулятор является эмитентом и оператором платформы, а кредитно-финансовые организации (КФО) выступают в роли посредников, обеспечивающих бесшовное взаимодействие клиентов с платформой. Это ставит перед службами информационной безопасности (ИБ) КФО задачу интеграции внешней инфраструктуры доверия, представленной узлом платформы цифрового рубля, в сложившийся за десятилетия высокозащищенный ландшафт автоматизированных банковских систем (АБС) [1, 2].

Анализ научной литературы и регуляторных документов выявил существенный пробел в области организационно-технологических решений для «последней мили» — сегмента между ядром платформы Банка России и системами дистанционного банковского обслуживания (ДБО) КФО. Проблема заключается в том, что есть четкие требования регулятора (стандарты Открытого API, требования к безопасности и интероперабельности), формулирующие строгие правила для ядра системы, и есть уже действующие системы ДБО самих КФО, часто устаревшие и не обладающие достаточными техническими мощностями. На практике возникает серьезная проблема — как технически и организационно связать эту стандартизированную платформу с разнообразными системами ДБО самих КФО и обеспечить безопасность, надежность и своевременность доставки сообщений при таких разнородных стеках?

Разрыв в технологиях порождает ряд практических сложностей:

- разрозненные системы ДБО КФО не всегда умеют бесшовно взаимодействовать с новыми API-интерфейсами и сервисами ядра;
- нестандартизированные шлюзы и конвертеры могут становиться «узким местом», снижая скорость и стабильность доставки финансовых сервисов;
- внедрение новых сервисов требует глубокой реорганизации бэкэнда КФО, и не все организации готовы к таким изменениям;
- если «ядро» работает нестабильно, это напрямую влияет на качество и доступность финансовых

услуг для клиентов КФО, особенно в удаленных или малодоступных районах.

Большинство работ, например [3, 4], фокусируются на макроуровне: архитектуре платформы Центрального банка, моделях программируемости и смарт-контрактах для конечных пользователей. При этом вопросы организации защищенного канала, методов разграничения доступа к API платформе, интеграции Hardware Security Modules (HSM) и процедур обеспечения непрерывности бизнеса при обновлении сертификатов на стороне КФО остаются слабо освещенными [5, 6]. Практика «пилотного» внедрения в 2023–2024 гг. показала, что основными вызовами являются не столько производительность, сколько обеспечение отказоустойчивости и конфиденциальности на стыке двух доменов безопасности.

Новизна настоящего исследования заключается в синтезе единого механизма, объединяющего требования Банка России и внутренние политики КФО по управлению доступом, в эшелонированную модель защиты. Авторами предлагается авторский подход к формированию контура «нулевого доверия» (Zero Trust) для потока данных от платформы цифрового рубля до автоматизированного рабочего места (АРМ) операциониста. Данное исследование закрывает пробел между теоретическими моделями цифрового рубля и их практической безопасной реализацией в гетерогенной среде КФО [7].

ОСОБЕННОСТИ АРХИТЕКТУРЫ ИНФРАСТРУКТУРЫ ДОВЕРИЯ ЦИФРОВОГО РУБЛЯ

Платформа цифрового рубля для КФО представляет собой не просто очередную шлюз, а доверенную вычислительную среду, развернутую в периметре банка, но операционно подконтрольную Банку России. Ключевым технологическим элементом здесь является узел доступа, обеспечивающий криптографически верифицированное взаимодействие с АБС [8]. Анализ модели угроз для данного узла показывает, что критическими активами являются: ключи электронной подписи клиента, управляемые на стороне КФО, и API-интерфейсы инициирования операций³. В отличие от традиционной модели ДБО, где банк полностью контролирует фронтальную и бэк-офисную части, здесь возникает разделенная ответственность. Компрометация узла КФО не должна приводить к возможности генерации легитимной транзакции в обход воли клиента, что достигается применением многосторонних защищенных вы-

¹ Положение Банка России от 03.08.2023 № 820-П «О платформе цифрового рубля». URL: <https://base.garant.ru/407513623/>

² URL: https://cbr.ru/Content/Document/File/120075/concept_08042021.pdf

³ Финтех: понятие, особенности и влияние на сферу банковских и финансовых услуг. URL: <https://fintech.globus-ltd.ru/blog/fintekh-ponyatie-osobennosti-i-vliyanie-na-sferu-finansovyh-uslug/>



числений и обязательной аппаратной поддержкой криптографических операций (HSM) [5].

Сетевое взаимодействие предполагает отказ от классической шины данных предприятия в пользу строго изолированного контура. Любое обращение к смарт-контрактам платформы из АБС инициируется через брокера доверенных сообщений, гарантирующего атомарность и целостность обмена [1]. Это требует внедрения на стороне КФО специализированной инженерной инфраструктуры, способной поддерживать жесткие тайминги и SLA, предъявляемые регулятором к доступности сервиса, при этом исключая возможность скрытого туннелирования трафика из внешнего периметра во внутренние учетные системы.

ОРГАНИЗАЦИОННО-ТЕХНОЛОГИЧЕСКИЙ МЕХАНИЗМ БЕЗОПАСНОЙ ИНТЕГРАЦИИ

Для минимизации рисков интеграции предлагается двухконтурный механизм, базирующийся на принципах многоуровневой эшелонированной защиты и нулевого доверия. Структурная схема предлагаемого механизма представлена на *рис. 1*.

Первый контур — технологический, реализуется на уровне программно-аппаратных средств и включает:

- Криптографический шлюз с аппаратной неизвлекаемостью ключей (HSM). Операции подписи распоряжений клиента на перевод цифровых рублей выполняются исключительно в доверенной среде криптомодуля, исключая появление приватного ключа в оперативной памяти серверов общего назначения.

- Однонаправленный диод данных (Data Diode). Для канала доставки реестров от платформы Банка России к витрине данных КФО предлагается физическое обеспечение однонаправленности потока, что делает невозможной эксфильтрацию данных из ядра АБС даже при полной компрометации самого узла платформы.

Контроль целостности и HWID-привязка. Запуск программных компонентов платформы разрешен только на поименованном серверном оборудовании с верифицированным TPM-модулем, что предотвращает клонирование узла в виртуальную⁴ среду злоумышленника⁵.

⁴ Исследование АФТ: обеспечение доверия клиентов, ИИ, меняющий финансовый рынок и фокус на новые риски — приоритеты 2025 года. URL: <https://www.fintechru.org/press-center/news/issledovanie-aft-obespechenie-doveriya-klientov-ii-menyayushchiy-finansovyy-rynok-i-fokus-na-novye-r/?ysclid=mshk3ysl2g502134327>

⁵ Левдонский А. Возможности реализации мер защиты информации в соответствии с приказом ФСТЭК России № 17 средствами операционной системы специального назначения Astra Linux Special Edition, очередные обновления

Второй контур — организационный, формирующий процедурные рамки:

- Внедрение модели Zero Trust на уровне приложений. Каждый запрос от операциониста КФО на проведение операции с цифровым рублем (например, разблокировка счета или комплаенс-проверка) требует отдельного токена доступа, подписанного ключом должностного лица и ограниченного по времени и контексту операции.

- Процедуры комплаенса, интегрированные в поток транзакций (In-line Compliance). До подписания транзакции клиента или банка данные проходят автоматизированную проверку на соответствие «антиотмывочному» законодательству (ПОД/ФТ) в изолированном сегменте. Это снижает риск легализации преступных доходов на этапе интеграции новой валюты в продуктовую линейку.

- Управление жизненным циклом сертификатов (CLM). Разрабатывается строгий регламент плановой и аварийной смены ключей электронной подписи, синхронизированный с аналогичными регламентами Банка России⁶. Критичным параметром является время кросс-сертификации при компрометации одного из доменов.

1.7 и 4.7. URL: <https://wiki.astralinux.ru/pages/viewpage.action?pageId=181666089&ysclid=mshk1aknp758331379>

⁶ См. раздел сайта Банка России «Стандарты Банка России». URL: https://cbr.ru/information_security/Gubzi_docs/?ysclid=ms_hnesjw13605227481. Стандарт «Обеспечение информационной безопасности организаций банковской системы Российской Федерации. Общие положения» СТО БР ИББС-1.0. (5-я ред., 2024 г.). Под «СТО БР ИББС» подразумевается целый комплекс стандартов Банка России по информационной безопасности. В 2024 г. в этот комплекс добавился новый профильный стандарт именно для платежных систем — СТО БР БФБО-1.8–2024 «Безопасность финансовых (банковских) операций. Обеспечение безопасности финансовых сервисов при проведении дистанционной идентификации и аутентификации. Состав мер защиты информации» — действует с 1 июля 2024 г. и имеет прямое отношение к платежным системам и дистанционным сервисам. СТО БР ИББС-1.0–2014 «Обеспечение информационной безопасности организаций банковской системы Российской Федерации. Общие положения» — базовый документ комплекса, задает системный подход: описывает, как выстроить всю систему ИБ в банке, включая процессы, роли людей, защиту на всех этапах жизненного цикла систем, криптографическую защиту, управление доступом, мониторинг и реагирование на инциденты. ГОСТ Р 57580.1–2017 «Безопасность финансовых (банковских) операций» — стандарт детализирует требования к защите конкретных платежных процессов — например, сервисов срочного и несрочного перевода, а также Системы быстрых платежей (СБП). Положение Банка России от 09.06.2012 № 382-П (ред. от 07.05.2018) «О требованиях к обеспечению защиты информации при осуществлении переводов денежных средств и о порядке осуществления Банком России контроля за соблюдением требований к обеспечению защиты информации при осуществлении переводов денежных средств» закрепляет обязанность операторов платежных систем и других участников (банки, процессинговые центры) регулярно оценивать, насколько их меры ИБ соответствуют установленным требованиям, и документировать результаты.



Источник / Source: составлено авторами / Compiled by the authors.

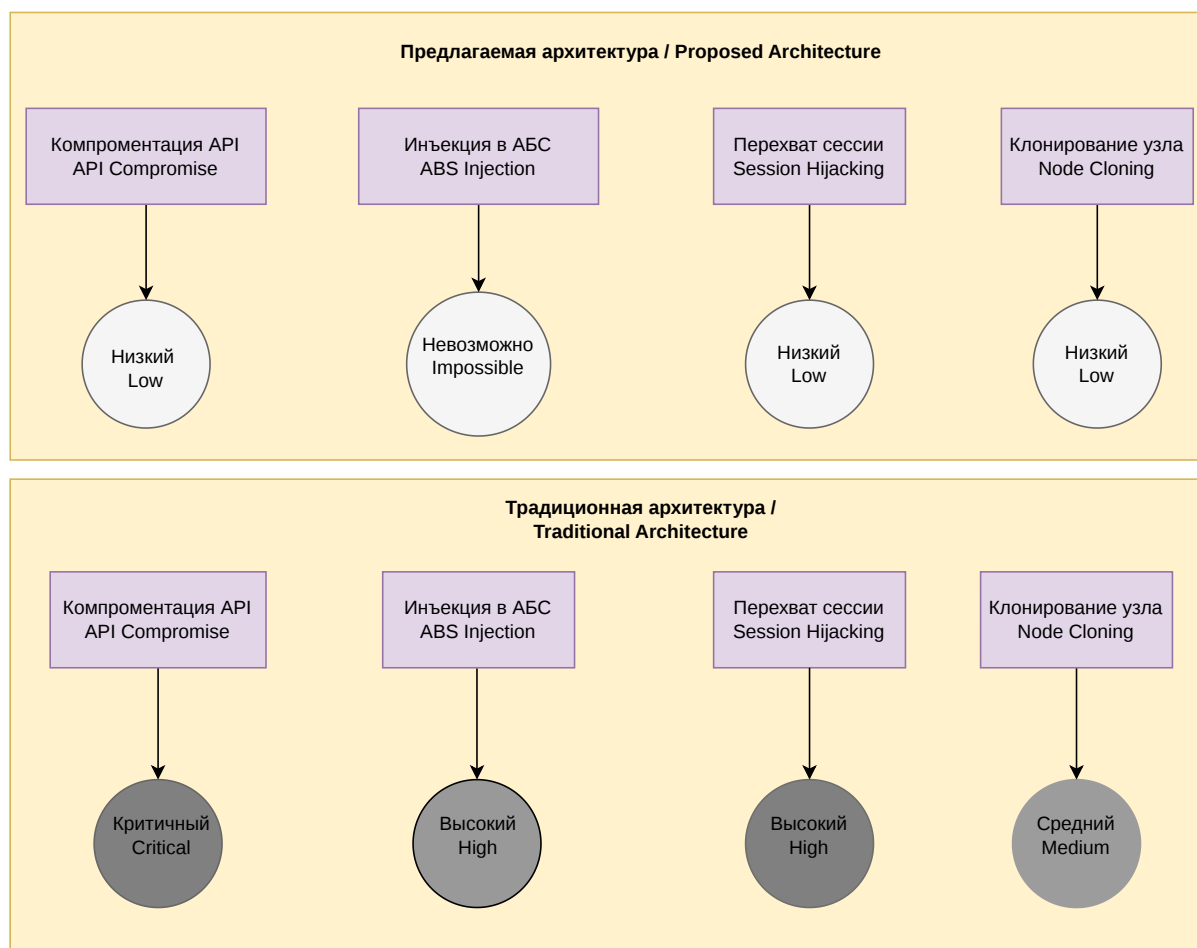


Рис. 2 / Fig. 2. Сравнительный анализ векторов атак на традиционную и предлагаемую архитектуру интеграции / Comparative Analysis of Attack Vectors Against Traditional and proposed Integration Architectures

Источник / Source: составлено авторами / Complied by the authors.

Взаимодействие контуров реализуется через матрицу доступа, где технологическая возможность инициировать транзакцию (например, API-вызов) проверяется на соответствие организационной политике (роль сотрудника, время, сумма) в режиме реального времени.

ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ И ОЦЕНКА КИБЕРУСТОЙЧИВОСТИ ПРЕДЛОЖЕННОГО МЕХАНИЗМА

В рамках пилотного обследования инфраструктуры одной из КФО была произведена оценка киберустойчивости предлагаемого механизма. Моделирование проводилось на стенде, имитирующем взаимодействие промышленного контура АБС с тестовой средой платформы цифрового рубля [2]. Ключевым показателем для оценки был выбран вектор атаки на компрометацию сессии легитимного пользователя через уязвимости на стороне интеграционного шлюза КФО. Результаты тестирования показали, что использование однонаправлен-

ного шлюза для загрузки реестров в АБС сокращает поверхность атаки на 45% по сравнению с традиционным двусторонним REST-API-взаимодействием (рис. 2). Это достигается за счет физической невозможности инъекции команд в ядро платформы из корпоративной сети банка. Внедрение механизма контекстной аутентификации (привязка сессии к HWID APM и HSM-ключу сотрудника) позволяет предотвратить атаки типа «Pass-the-Hash» и «Pass-the-Token» на уровне операционной системы операциониста. Вероятность успешной реализации сложной целевой атаки (APT) с целью несанкционированной эмиссии токенов цифрового рубля на баланс подставного лица была снижена с теоретически возможной до крайне низкой за счет обязательной многофакторной аутентификации должностных лиц КФО при каждой попытке выполнить критическую операцию [6, 7].

Особое внимание уделено такому аспекту, как отказоустойчивость. Предлагаемый механизм предполагает «горячий» резерв HSM-модулей и кластерную



конфигурацию шлюзов, что позволяет выдержать деградацию одного из узлов без потери доступности сервиса для конечных пользователей.

ВЫВОДЫ

Проведенное исследование позволяет заключить, что интеграция платформы цифрового рубля в КФО является не рядовой ИТ-задачей, а комплексной проблемой обеспечения национальной платежной безопасности, требующей синтеза передовых организационных и технологических мер. Авторами разработан и обоснован механизм безопасной интеграции, ключевыми элементами которого являются двухконтурная архитектура, реализация принципа нулевого доверия и строгое управление жизненным циклом криптографических ключей на базе отечественных HSM [5].

В отличие от фрагментированных подходов, данное решение обеспечивает сквозную защиту от рабочего места операциониста до смарт-контракта в платформе Банка России.

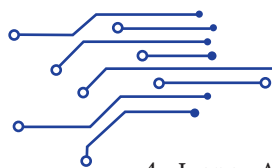
Теоретическая значимость работы состоит в развитии концепции киберустойчивости для систем CBDC на операционном уровне КФО [1]. Практическая ценность заключается в возможности тиражирования предложенного механизма в виде типового проектного решения, что существенно ускорит и обезопасит процесс масштабирования цифрового рубля. Перспективы дальнейших исследований авторов связаны с адаптацией механизма под трансграничные расчеты в цифровых валютах центральных банков и разработкой методики нагрузочного тестирования средств защиты без остановки промышленной эксплуатации.

СПИСОК ИСТОЧНИКОВ

1. Иванов А.В., Царегородцев А.В., Валеев М.В. Технологическое решение по формированию инфраструктуры доверия в системе защищенности цифрового рубля. *Известия ЮФУ. Технические науки*. 2025;3(245):233-245. URL: <https://doi.org/10.18522/2311-3103-2025-3-233-245>
2. Иванов А.В., Царегородцев А.В., Валеев М.В. Кибербезопасность цифрового рубля в системе деструктивных событий цифрового пространства. *Инновации и инвестиции*. 2025;8:555-559. URL: <https://www.elibrary.ru/frgqvo>
3. Иванов А.В., Царегородцев А.В., Валеев М.В. Кибербезопасность платформы цифрового рубля в условиях возрастания рисков информационной безопасности. *Computational Nanotechnology*. 2025;12(3):47-57. URL: <https://doi.org/10.33693/2313-223X-2025-12-3-47-57>
4. Иванов А.В., Царегородцев А.В., Валеев М.В. Ключевые проблемы и технологические риски перехода российской банковской системы на расчеты в цифровых рублях. *Защита информации. Инсайд*. 2025;5(125):6-11. URL: <https://www.elibrary.ru/romupn>
5. Иванов А.В., Царегородцев А.В., Валеев М.В. Инфраструктурная защищенность системы цифрового рубля: рекомендации и предложения. *Защита информации. Инсайд*. 2025;6(126):2-7. URL: <https://www.elibrary.ru/eucppe>
6. Иванов А.В., Царегородцев А.В. Ключевые компоненты оценки готовности российской банковской системы к внедрению цифрового рубля. *Вопросы региональной экономики*. 2025;3(64):207-218. URL: <https://www.elibrary.ru/aetymd>
7. Царегородцев А.В. Кибербезопасность цифрового рубля в системе деструктивных событий цифрового пространства. *Цифровые решения и технологии искусственного интеллекта*. 2025;1(3):42-52. URL: <https://www.elibrary.ru/frgqvo>
8. Auer R., Cornelli G., Frost J. Rise of the Central Bank Digital Currencies: Drivers, Approaches and Technologies. *International Journal of Central Banking*. 2022;18(1):225-268. URL: <http://dx.doi.org/10.2139/ssrn.3724070>
9. Мельников А.Д., Будников Д.А., Коннова И.Г., Кубаев А.В. Рекомендации по созданию трастовой инфраструктуры в интересах системы цифрового рубля. *Безопасность информационных технологий*. 2024;31:43-63. URL: <http://dx.doi.org/10.26583/bit.2024.3.01>

REFERENCES

1. Ivanov A.V., Tsaregorodtsev A.V., Valeev M.V. Technological solution for building a trust infrastructure in the digital ruble security system. *Izvestiya SFedU. Technical Sciences*. 2025;3(245):233-245. (In Russ.). URL: <https://doi.org/10.18522/2311-3103-2025-3-233-245>
2. Ivanov A.V., Tsaregorodtsev A.V., Valeev M.V. Cybersecurity of the digital ruble in the system of destructive events in the digital space. *Innovations and Investments*. 2025;8:555-559. (In Russ.). URL: <https://www.elibrary.ru/frgqvo>
3. Ivanov A.V., Tsaregorodtsev A.V., Valeev M.V. Cybersecurity of the digital ruble platform in the context of increasing information security risks. *Computational Nanotechnology*. 2025;12(3):47-57. (In Russ.). URL: <https://doi.org/10.33693/2313-223X-2025-12-3-47-57>



4. Ivanov A. V., Tsaregorodtsev A. V., Valeev M. V. Key problems and technological risks of the Russian banking system transitioning to settlements in digital rubles. *Information Security. Inside*. 2025;5(125):6-11. (In Russ.). URL: <https://www.elibrary.ru/romupn>
5. Ivanov A. V., Tsaregorodtsev A. V., Valeev M. V. Infrastructure security of the digital ruble system: recommendations and proposals. *Information security. Inside*. 2025;6(126):2-7. (In Russ.). URL: <https://www.elibrary.ru/eycppe>
6. Ivanov A. V., Tsaregorodtsev A. V. Key components of assessing the readiness of the Russian banking system to implement the digital ruble. *Issues of Regional Economics*. 2025;3(64):207-218. (In Russ.). URL: <https://www.elibrary.ru/aetymd>
7. Tsaregorodtsev A. V. Cybersecurity of the digital ruble in the system of destructive events in the digital space. *Digital solutions and artificial intelligence technologies*. 2025;1(3):42-52. (In Russ.). URL: <https://www.elibrary.ru/frgqvo>
8. Auer R., Cornelli G., Frost J. Rise of the Central Bank Digital Currencies: Drivers, Approaches and Technologies. *International Journal of Central Banking*. 2022;18(1):225-268. (In Russ.). URL: <http://dx.doi.org/10.2139/ssrn.3724070>
9. Melnikov A. D., Budnikov D. A., Konnova I. G., Kubaev A. V. Recommendations for Creating Trust Infrastructure in the Interests of the Digital Ruble System. *Information Technology Security*. 2024;3:43-63. (In Russ.). URL: <http://dx.doi.org/10.26583/bit.2024.3.01>

ИНФОРМАЦИЯ ОБ АВТОРАХ / ABOUT THE AUTHORS

Анатолий Викторович Иванов — доктор социологических наук, профессор кафедры информационной безопасности факультета информационных технологий и анализа больших данных, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

Anatoly V. Ivanov — Dr. Sci. (Soc.), Prof. at the Department of Information Security, Faculty of Information Technology and Big Data Analysis, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

<https://orcid.org/0000-0002-0316-1518>

Автор для корреспонденции / Corresponding author:

aivanov@fa.ru

Виктория Александровна Аптулина — студентка кафедры информационной безопасности факультета информационных технологий и анализа больших данных, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

Victoria A. Aptulina — Student at the Department of Information Security, Faculty of Information Technology and Big Data Analysis, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

<https://orcid.org/0009-0003-7134-9007>

228102@edu.fa.ru

Заявленный вклад авторов:

А.В. Иванов — разработка организационно-технологического механизма, написание раздела «Практическая реализация».

В.А. Аптулина — анализ архитектуры инфраструктуры доверия, формирование выводов.

Authors' declared contributions:

A. V. Ivanov — development of the organizational and technological mechanism, writing the “Practical Implementation” section.

V. A. Aptulina — analysis of the trust infrastructure architecture, drawing conclusions.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

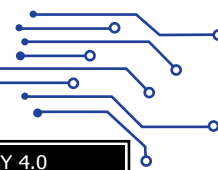
Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

Статья поступила 16.02.2026; после рецензирования 02.07.2026; принята к публикации 31.07.2026.

Авторы прочитали и одобрили окончательный вариант рукописи.

The article was submitted on 16.02.2026; revised on 02.07.2026 and accepted for publication on 31.07.2026.

The authors read and approved the final version of the manuscript.

DOI: 10.26794/3030-7097-2026-2-3-41-50
UDC 004.85:004.94:539.3(045)

Data Engineering and Machine Learning for Solving the Inverse Problem Based on CAE Simulation Results

O.P. Tulupova¹, G.N. Zholobova², F.U. Enikeev³^{1,2} Financial University under the Government of the Russian Federation, Moscow, Russian Federation;³ Ufa State Petroleum Technological University, Ufa, Russian Federation

ABSTRACT

The use of machine learning in solving direct forecasting problems is becoming an alternative to the use of comprehensive finite element modeling. However, the solutions of inverse problem by means of machine learning is conventionally not considered in the literature. Traditional approaches are based on iterative calculations that require multiple runs of models and do not work in real time, which limits the creation of digital doubles. **The objective of this study** is to solve the inverse problem by developing and verifying a Gaussian regression (GPR) model for determining the strain rate sensitivity index of superplastic material from the results of bulge tests based on the results of finite element calculations. The results of finite element modeling of the superplastic forming process in the ANSYS CAE package were used as a data source. 120 finite element simulations were performed, 40 for each of the three pressure modes. Two GPR models have been developed. One is for defining K with the Matern kernel, the second is for m with the RBF kernel. It is shown that replacing the RBF kernel with the Matern kernel ($\nu = 2,5$) reduces the median relative uncertainty of the forecast for parameter K from 50 to 16%. The following quality metrics were obtained on the test sample: $R^2 = 0.91$ for K and $R^2 = 0.92$ for m , the relative RMS error is 20% for K and 7% for m . **The developed model can be used** for non-destructive quality control, integration into digital twins, and optimization of superplastic forming processes.

Keywords: inverse problem; machine learning; Gaussian process regression (GPR); ANSYS; finite element modeling (FEM); CAE simulation; data engineering; superplastic forming (SPF)

For citation: Tulupova O.P., Zholobova G.N., Enikeev F.U. Data engineering and machine learning for solving the inverse problem based on CAE simulation results. *Digital Solutions and Artificial Intelligence Technologies*. 2026;2(3):41-50. DOI: 10.26794/3030-7097-2026-2-3-41-50

ОРИГИНАЛЬНАЯ СТАТЬЯ

Инженерия данных и машинное обучение для решения обратной задачи по результатам CAE-моделирования

О.П. Тулупова¹, Г.Н. Жолобова², Ф.У. Еникеев³^{1,2} Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация;³ Уфимский государственный нефтяной технический университет, Уфа, Российская Федерация

АННОТАЦИЯ

Применение машинного обучения в решении прямых задач прогнозирования становится альтернативой применению ресурсоемкого конечно-элементного моделирования. Однако применение машинного обучения для решения обратной задачи в литературе практически не рассматривается. Традиционные подходы основаны на итерационных расчетах, которые требуют многократных прогонов моделей и не работают в реальном времени, что ограничивает создание цифровых двойников. **Цель исследования** — решение обратной задачи путем разработки и верификации модели гауссовской регрессии (GPR) для определения параметров K и m по геометрии готового изделия на основе конечно-элементных данных ANSYS. В качестве источника данных использовались результаты конечно-элементного моделирования процесса сверхпластической формовки в CAE-пакете ANSYS. Было выполнено 120 конечно-элементных симуляций, по 40 для каждого из трех режимов давления. Общий объем данных, полученных из решений ANSYS, составил 7571 строку. Были разработаны две GPR-модели. Одна для определения K с ядром Matern, вторая для m с ядром RBF. Масштабирование данных выполнялось методом RobustScaler. Показано, что замена RBF-ядра на ядро Matern ($\nu = 2,5$) снижает медианную относительную неопределенность прогноза для параметра K с 50 до 16%. На тестовой выборке получены следующие метрики качества: $R^2 = 0,91$ для K и $R^2 = 0,92$ для m , относительная среднеквадратичная ошибка составляет 20% для K и 7% для m . Выполненный корреляционный анализ выявил наиболее значимое влияние давления формовки на параметр K ($r = -0,32$). **Разработанная модель может быть использована**



для неразрушающего контроля качества, встраивания в цифровые двойники и оптимизации режимов сверхпластической формовки.

Ключевые слова: обратная задача; машинное обучение; регрессия гауссовского процесса (GPR); ANSYS; конечно-элементное моделирование (КЭ-моделирование); CAE-система; инженерия данных; процесс сверхпластической формовки (СПФ)

Для цитирования: Тулупова О.П., Жолобова Г.Н., Еникеев Ф.У. Инженерия данных и машинное обучение для решения обратной задачи по результатам CAE-моделирования. *Цифровые решения и технологии искусственного интеллекта*. 2026;2(3):41-50. DOI: 10.26794/3030-7097-2026-2-3-41-50

INTRODUCTION

In recent years, machine learning (ML) methods have found wide application in various fields, such as data analysis and control of technological processes. Their strength lies in their ability, by learning from experimental data, to identify complex nonlinear relationships between input process characteristics and output quality parameters [1, 2]. Therefore, predictive models built on these methods demonstrate high accuracy and can serve as an effective alternative to traditional approaches.

In particular, the significant potential of ML methods is revealed in the tasks of modeling and optimizing complex technological processes of metal forming by pressure. One of the promising areas for applying ML models is superplastic forming (SPF) — a technological process that makes it possible to produce parts of complex shape without welded seams due to the material's ability to undergo significant (up to hundreds of percent) strains without fracture [3–5].

For example, in [6], a method for intelligent design of superplastic forming process parameters based on an improved backpropagation neural network is proposed, which confirms the effectiveness of ML for complex technological tasks.

Numerical simulation of SPF is usually carried out using the finite element method (FEM), for example, in ANSYS or ABAQUS [7–9]. In this approach, an important aspect is the accuracy of describing the material deformation under superplastic conditions: the more accurate the model, the more reliable the calculation results. The key factor here is the parameters of the constitutive equations, in particular, the strain rate sensitivity index m and the strength coefficient K , which are introduced into the finite element model and determine the material response under external loads. The main difficulty when working with the constitutive parameters K and m is that they vary depending on many factors, such as temperature, grain size, and prior heat treatment. Therefore, there are no universal fixed values for these parameters. Direct methods for determining them (e.g., tensile tests) are labor-intensive and destroy the sample.

Earlier, in [10], the direct problem of predicting the geometric characteristics of a hemisphere obtained during the SPF process was solved. Within that work,

a Gaussian process regression (GPR) model was built based on finite element modeling in ANSYS, making it possible to predict the thickness at the pole s and the height of the inner ($UY1$) and outer ($UY111$) dome surfaces from the given parameters K , m , pressure p , and time t :

$$X_{direct} = [K, m, p, t] \rightarrow Y_{direct} = [s, UY1, UY111].$$

In [10], two pressure levels were considered: 0.29 MPa and 0.56 MPa. For the present study, in order to analyze the effect of pressure on the forming process in more detail, a third level — 0.9 MPa has been added to expand the applicability domain of the model and improve prediction accuracy over a wider range of loading regimes.

The ML model presented in [10] achieved high accuracy ($R^2 > 0.99$) on the validation set and demonstrated the possibility of replacing computationally expensive finite element modeling.

Unlike the direct problem, where the input parameters are known, the inverse problem of determining the constitutive parameters K and m from the geometry of the produced part remains less studied. Solving the inverse problem is important for creating digital twins, as it would allow refining the constitutive parameters without destructive testing, based on a single finished part.

Most of the problems considered that are solved using machine learning algorithms [11–13] are direct problems, whereas inverse problems are traditionally solved by iterative finite element calculations, which require significant computational resources and do not allow real-time operation.

The study conducted by the authors in [11] demonstrated the possibility of using ANNs for modeling complex technological relationships. In [11], a neural network was used to predict a set of mechanical properties of a cast alloy taking into account chemical composition, manufacturing technology, and heat treatment parameters. The study by authors [12] showed that a hybrid PSO-BP neural network predicts the thickness distribution of a part more accurately, especially in rounded areas, than the BP backpropagation algorithm, which confirms the possibility of using machine learning to analyze strain non-uniformity in SPF. In [13], the authors replaced direct finite ele-

ment calculations with a GPR surrogate model, which allowed them to reduce computational costs by a factor of 7–10 while maintaining high accuracy. Thus, direct problems of predicting geometry, properties, or thickness distribution in SPF are already being actively solved using ANN, PSO-BP, and GPR, whereas the inverse problem of reconstructing the constitutive parameters — the strain rate sensitivity index m and the strength coefficient K — from the geometry of the finished part remains virtually untouched in the literature, which defines its scientific and practical significance.

The aim of this work is to solve the inverse problem of superplastic forming by developing and verifying a GPR model for determining the parameters K and m from the geometry of the finished hemisphere based on ANSYS finite element data.

CONSTITUTIVE RELATIONS AND THEIR IMPLEMENTATION IN THE ANSYS CAE PACKAGE

In the classical formulation, superplastic forming is described by a power law relating the flow stress σ to the strain rate $\dot{\epsilon}$:

$$\sigma = K \cdot \dot{\epsilon}^m, \quad (1)$$

which is widely represented in the scientific literature [14–16].

In finite element modeling within the ANSYS CAE package, a different form of notation is used, adapted for the numerical calculation of deformation problems. ANSYS employs the Norton model, where the strain increment $\Delta\epsilon_{cr}$ over a time step Δt is expressed through the equivalent stress σ and the absolute temperature T :

$$\Delta\epsilon_{cr} = C_7 \sigma^{C_8} e^{-C_{10}/T} \Delta t. \quad (2)$$

C_7 is a scaling factor, C_8 is the stress sensitivity exponent, C_{10} is the temperature constant.

When modeling an isothermal SPF process, the temperature constant C_{10} can be set to zero ($C_{10} = 0$), which simplifies the exponential term to 1. Under constant temperature conditions, equation (2) reduces to:

$$\Delta\epsilon_{cr} = C_7 \sigma^{C_8} \Delta t. \quad (3)$$

This makes it possible to convert the parameters K and m into the ANSYS parameters C_7 and C_8 using the following formulas:

$$C_7 = \frac{1}{K^{1/m}}, C_8 = \frac{1}{m}, C_{10} = 0. \quad (4)$$

Since the original parameters K and m are more intuitive, the rest of this work will be based on them.

METHODOLOGY OF DATA GENERATION FOR SOLVING THE INVERSE PROBLEM

To build a model for determining the constitutive parameters K and m (1) from the hemisphere geometry, a dataset linking the input geometric features to the target parameters is required.

Because conducting physical experiments is expensive and time-consuming, the data were taken from finite element simulations of the superplastic forming process performed in the ANSYS CAE package.

To optimize computational resources while maintaining the required accuracy, an axisymmetric model was used. This approach transforms a three-dimensional problem into a two-dimensional one by exploiting the symmetry of the geometry and boundary conditions. It is widely adopted for modeling SPF of axisymmetric parts [15, 16] and provides acceptable accuracy with a significant reduction in computation time.

In ANSYS, the simulation began with constructing a geometric model of the blank and the cylindrical die (Fig. 1). The blank was defined as a deformable shell, while the die was represented as a rigid tool that determines the final shape of the part.

The parameters of the finite element model in ANSYS [8] were as follows:

1. Part geometry: sheet blank with initial thickness $s_0 = 1.2$ mm, round shape, diameter $D_s = 114$ mm; cylindrical die with diameter of $D_0 = 100$ mm.

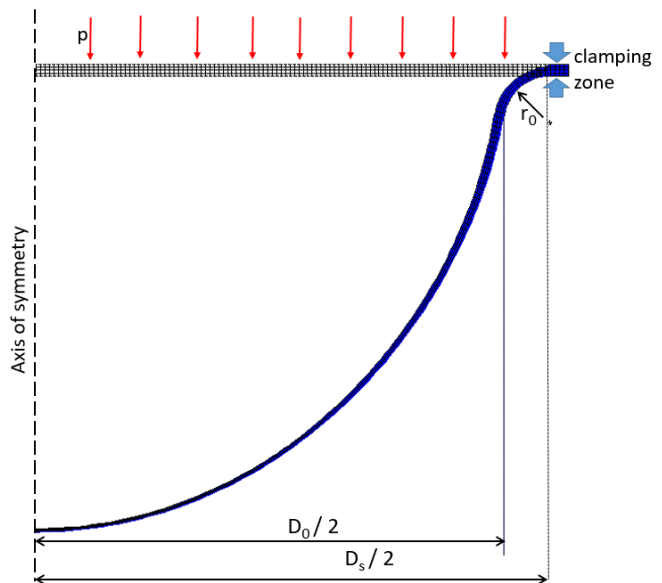
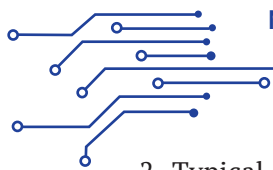


Fig. 1. Axisymmetric Finite Element Model with Mesh Overlay Before and After the Forming Process: D_0 – Diameter of the Cylindrical Die, D_s – Diameter of the Blank (Deformable Zone), r_0 – Entry Radius of the Die, p – Uniformly Applied Pressure

Source: author's calculations.



2. Typical mechanical properties of the alloy: Young's modulus $E = 70$ GPa, Poisson's ratio $\mu = 0.34$.

3. A 2D axisymmetric element PLANE183 was selected for the blank. CONTA172 and TARGE169 elements were used on the contact surfaces.

4. The blank mesh consisted of 2112 4-node elements.

5. Boundary conditions:

a) a uniform pressure p was applied to the top surface of the blank. Three experimental pressure levels were considered: $p_1 = 0.29$ MPa, $p_2 = 0.56$ MPa, and $p_3 = 0.9$ MPa;

b) blank constraints: a clamp is located in the contact zone between the blank and the die, therefore zero displacements were prescribed for both axes (All DOF = 0), and a symmetry condition $UX = 0$ was set on the axis of symmetry;

c) SPF process temperature: $T = 450$ °C.

6. For the mathematical model of the constitutive relationship, the Norton model was applied, following equation (2) and the parameter conversion (4).

The following parameters were varied to generate the training dataset:

1. The strain rate sensitivity index m was varied in the range typical for the superplastic state: $m = 0.3-0.7$.

2. The strength coefficient K was calculated for each value of m using equation (1).

3. Calibration of ξ and σ was performed by inverse modeling in ANSYS. The pressure histories $p(t)$ for each experimental case were set according to the input data. The parameters ξ and σ were varied within $\xi = 0.001-0.10$ s⁻¹ and $\sigma = 20-55$ MPa. The calibration criterion was the agreement between the calculated time to reach the experimental dome height and the actually measured time.

4. Forming pressure p : 0.29, 0.56 and 0.9 MPa following the experimental data from [8].

To capture the process dynamics and enlarge the training dataset without additional computational cost, intermediate forming parameters were recorded. For each ANSYS solution, the following quantities were saved at each time step:

1. Time t , s from the start of the process.

2. Current thickness s , m at the pole of the hemisphere.

3. $UY1$, m current height on the inner surface of the hemisphere.

4. $UY111$, m current height on the outer surface of the hemisphere.

The values of m within its typical range were sampled using the Latin Hypercube Sampling (LHS) method. LHS is a design of experiments technique that distributes points uniformly in a multidimensional parameter space with a limited number of simula-

tions. It was chosen because, for the same number of simulations, LHS produces more accurate models than the Monte Carlo method [17].

A total of 120 finite element simulations were performed, 40 for each of the three pressure levels. The total dataset obtained from the ANSYS solutions comprised 7571 rows.

For the inverse problem, the following data structure was adopted:

Input ($X_{inverse}$) = [$s, UY1, UY111, t, p$] → the Output ($Y_{inverse}$) = [K, m].

In this study, pressure p was included as an input feature, while only K and m were the target variables. If identifying p becomes necessary, the data structure can be modified by moving p to the target variables — a task that deserves a separate study.

DATA ENGINEERING AND FEATURE TRANSFORMATION WITH SCALING

Building a training dataset for the inverse problem involves several steps aimed at the correctness and efficiency of machine learning. The data from ANSYS were converted to CSV format and underwent the following steps:

1. To improve model stability and avoid loss of significance when dealing with very small numbers — for example, the height of the inner dome surface $UY1$ is on the order of 10^{-12} m, while the thickness s is down to 10^{-3} m — all linear dimensions were converted from meters to millimeters.

2. The superplastic forming process at the initial stage is characterized by an exponential growth of strains, followed by a slowdown. The time ranges from fractions of a second to thousands of seconds — on a linear scale, the data are distributed extremely unevenly. A logarithmic transformation of time, $\log t = \ln(t + \varepsilon)$ ($\varepsilon = 10^{-8}$), is applied in order to:

a) compress the dynamic range from 9 orders of magnitude to about 20 units;

b) maintain sensitivity to changes in geometry at the beginning of the process, when this is particularly important.

3. To evaluate the generalisation ability of the model, the data were randomly split into three parts:

a) 60% — training set;

b) 20% — validation set;

c) 20% — test set.

4. The input features differ in units of measurement and scales. For example, $UY1$ ranges from $3.7e-09$ to 49.7 mm, while pressure p ranges from 0.29 to 0.9 MPa. Therefore, scaling was applied using the RobustScaler method, which is robust to outliers and uses the median and interquartile range: $X_{scaled} = (X - median) / (Q75 - Q25)$, where $Q75$ and $Q25$ are the upper and lower quartiles.



Scaling was performed separately for the input features X and the target variables Y . To prevent data leakage from the test set into the training process, the scaling parameters (median and quartiles) were determined solely from the training set. These parameters were then applied to the validation and test sets.

5. When exporting data from ANSYS at the end of each simulation, duplicate rows were found, indicating that the limiting strain state had been reached. At that point, the geometric parameters stopped changing. To avoid overfitting, the model and to reduce the computational load, these duplicates were removed during the preprocessing stage. As a result, from 7571 original rows, 5505 unique values were selected (about 2066 duplicates removed, which is approximately 27% of the data).

JUSTIFICATION FOR CHOOSING THE MACHINE LEARNING

To solve the inverse problem of determining the constitutive parameters (1) from the geometry of the finished part, several machine learning methods were considered: artificial neural networks [18, 19], random forest [20], and Gaussian process regression (GPR) [13, 21, 22].

In this work, the GPR method was chosen to solve the inverse problem. This method offers several advantages confirmed in the scientific literature:

1. In [13], the authors demonstrated the effectiveness of using GPR under limited computational resources. They showed that applying GPR can reduce the number of direct finite element calculations by 70–90% compared to traditional methods, while maintaining high prediction accuracy ($R^2 > 0.96$). This result was achieved due to the ability of GPR to adaptively interpolate between training points and to quantify uncertainty, which is important for determining the constitutive parameters K and m .

2. The study [21] demonstrated that the GPR method has a built-in function for quantifying prediction uncertainty, which is essential for determining the constitutive parameters K and m from the geometry of the finished part. This method allows not only obtaining point estimates but also forming confidence intervals that characterise the reliability of identification.

3. The research presented in [22] showed that GPR effectively models complex processes whose outcomes depend on many heterogeneous factors, such as constitutive parameters, geometric features, and temperature regimes. The present study faces a similar situation: the geometry of the hemisphere after superplastic forming (SPF) is determined by five input parameters (s , $UY1$, $UY111$, $\log t$, p), which are in complex nonlinear relationships with the target characteristics K and m .

CONFIGURING THE GPR MODEL

The key element of a GPR is the kernel (covariance function). It determines how smooth the relationships between the data will be. In this study, two kernel configurations were tested.

First kernel configuration. Initially, it was assumed that the SPF process yields smooth relationships between the input parameters and the output characteristics. Therefore, the squared exponential kernel (RBF) was chosen for training, with the addition of a WhiteKernel to account for the errors in simulation and data export.

However, the analysis of the model training results showed that for the parameter K , the median relative uncertainty was about 50%, which is approximately twice the desired level.

Second kernel configuration. To improve the prediction accuracy for the parameter K , the Matern kernel ($\nu = 2.5$) was chosen, which corresponds to moderately smooth functions, as it is less sensitive to local variations.

For predicting the parameter m (where the relationships are smoother), the RBF kernel was retained.

For both kernels, a WhiteKernel component was added with a noise level of $1e-3$ (the value was determined experimentally).

The refined kernel configuration — Matern for K , RBF for m — yielded good results. $R^2 = 0.91$ for K and $R^2 = 0.92$ for m . For K , the median relative uncertainty decreased from 50 to 16%. Further hyperparameter optimization made no sense, as additional tuning no longer provided a significant improvement in quality.

RESEARCH RESULTS

The developed GPR model performed well in predicting the parameters of the constitutive relation (1). Table 1 presents the statistical characteristics of the target variables on the test set.

The coefficient of variation (CV) is calculated using the formula:

$$CV = \frac{\sigma}{\mu} \times 100\%,$$

where σ is the standard deviation of the actual values on the test set, μ is the mean actual value on the test set.

A CV value of 65% for the parameter K (see Table 1) indicates a significant spread of values of this parameter around the mean $\mu = 282.1 \text{ MPa} \cdot \text{s}^m$ in the original dataset. This is explained by the fact that the model was trained on a wide range of parameters of the constitutive relation (1): the values of the parameter m (0.3–0.7) and the corresponding K values were varied, as well as three pressure regimes p (0.29,

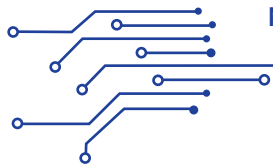


Table 1

Statistical Characteristics of the Target Variables in the Test Sample

Parameter	Mean (μ)	Standard deviation (σ)	Coefficient of variation (CV), %
K	282.1 МПа·с ^м	183.5	65.0
m	0.488	0.120	24.6

Source: Complied by the authors.

Table 2

Prediction Quality Metrics

Parameter	Set	R ²	RMSE	MAE	Relative RMSE, %
K	Validation	0.9346	44.9934	12.9657	—
	Test	0.9056	56.3923	16.1770	20.0
m	Validation	0.9363	0.0293	0.0116	—
	Test	0.9175	0.0343	0.0133	7.0

Source: Complied by the authors.

0.56 and 0.9 MPa). Table 2 presents the main quality metrics of the developed GPR model.

The relative RMSE (relative root mean square error) is calculated as:

$$\text{Relative RMSE} = \frac{\text{RMSE}}{\mu} \times 100\%,$$

where RMSE is the root mean square error, μ is the mean actual value of the target variable on the test set.

As can be seen from Figs 2 and 3, the actual and predicted values for K and m lie along the diagonal ($y = x$), indicating high prediction quality (see Table 2).

The scatter of points is slightly larger for K and denser for m , which corresponds to the difference in R².

Analysis of prediction quality for the parameter m :

1. The R² value exceeds 0.92 on the test set, indicating high prediction quality.

2. The low RMSE (0.0343) and MAE (0.0133) values also indicate high accuracy. The small difference between the validation and test sets confirms the absence of overfitting.

Analysis of prediction quality for the parameter K :

1. An R² value of ≈ 0.91 indicates good prediction quality of the developed model. The R² values on the validation (0.9346) and test (0.9056) sets are close,

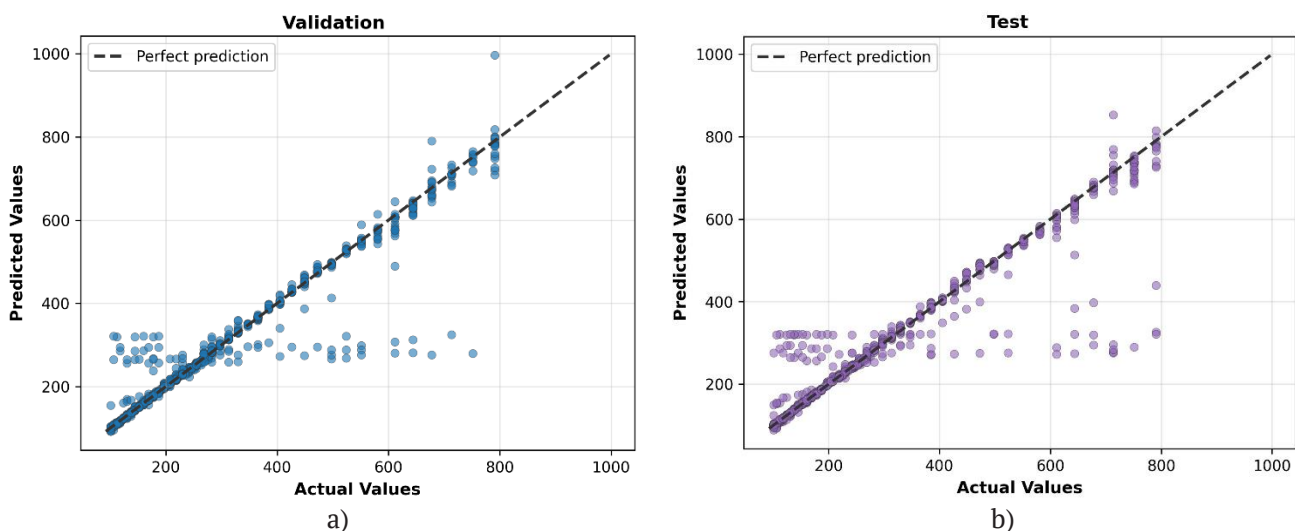


Fig. 2. Comparison of Actual and Predicted K Values (a – Validation set, b – Test set)

Source: author's calculations.

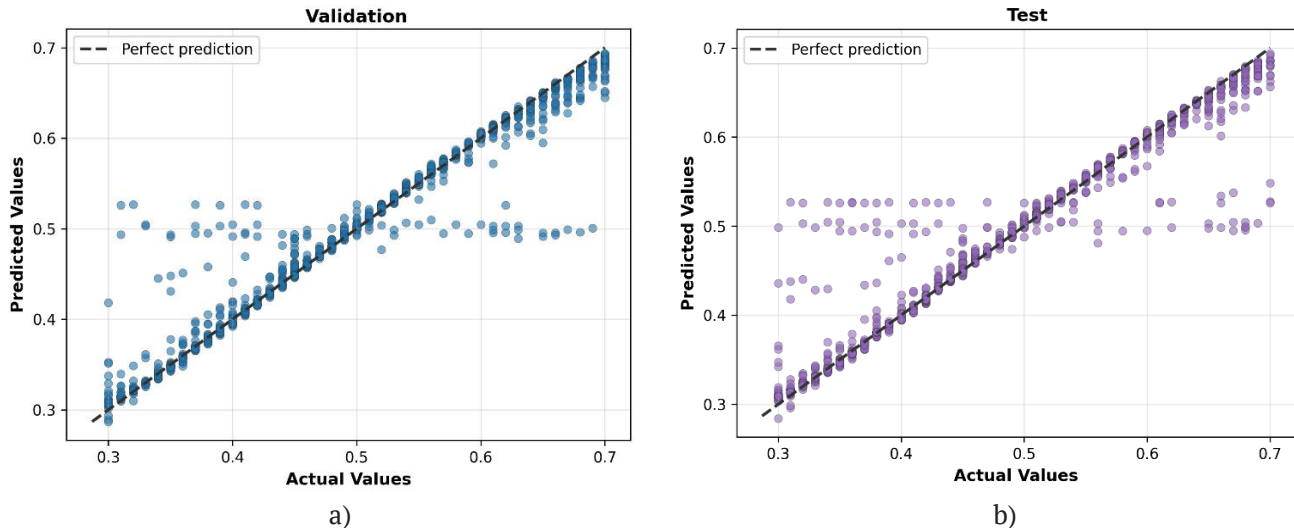


Fig. 3. Comparison of Actual and Predicted m Values (a – Validation set, b – Test set)

Source: author's calculations.

which indicates a good generalisation ability of the model and the absence of overfitting.

2. The obtained RMSE (≈ 56.4) and MAE (≈ 16.2) values require justification taking into account the scale of the data.

The mean value of the parameter K (see Table 1) on the test set is $\mu = 282.1 \text{ MPa}\cdot\text{s}^m$, and the standard deviation is $\sigma = 183.5$. Thus, the relative prediction error is:

$$\frac{RMSE}{\mu} \times 100\% = \frac{56.4}{282.1} \times 100\% \approx 20\%.$$

Thus, despite the high variability of the original data ($CV = 65\%$), the GPR model provides a relative prediction error for the parameter K of about 20%, which is a satisfactory result for the inverse problem of identifying the parameters of the constitutive relation (1).

To identify relationships between the parameters, a linear correlation analysis was performed. Its results are presented in Table 3 and Fig. 4.

Analysis of the correlation results (see Table 3 and Fig. 4):

1. Among all factors, pressure p has the most noticeable effect on K ($r = -0.32$). This is consistent with the physical principle of superplastic forming: the higher the strain rate, the lower the value of K should be.

2. The use of a logarithmic time scale made it possible to establish a weak correlation of the parameter K with the logarithm of time $\log t$ and thickness s (correlation coefficient $|r| \approx 0.12$). This result would not have been achievable using a linear time scale.

3. The values of $UY1$ and $UY111$ show little correlation ($|r| < 0.06$) with the target parameters K and m ,

Table 3

Correlations Between Input Features and Target Parameters (K, m)

Feature	With	
	K	m
Pressure p	-0.3199	-0.0091
Logarithm of time $\log t$	+0.1214	-0.0370
Thickness at the dome pole s	-0.1151	-0.0544
Dome height on the inner surface of the hemisphere $UY1$	+0.0587	+0.0062
Dome height on the outer surface of the hemisphere $UY111$	+0.0594	+0.0067

Source: Compiled by the authors.

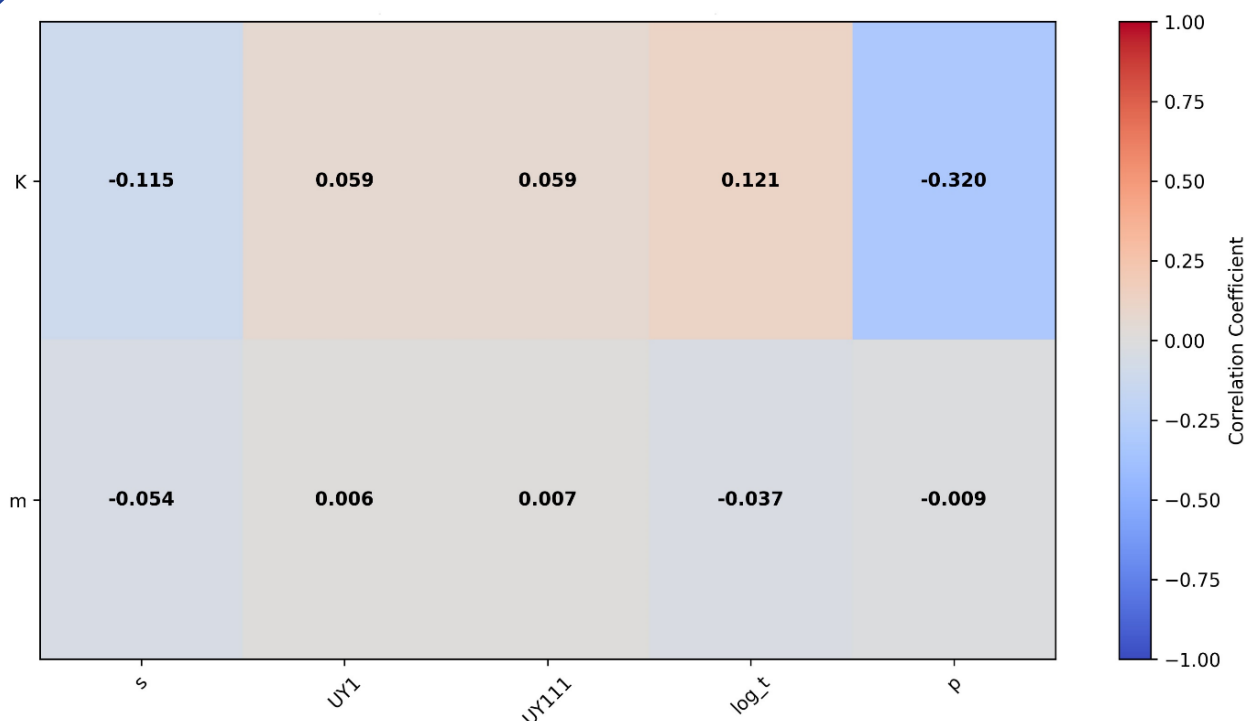
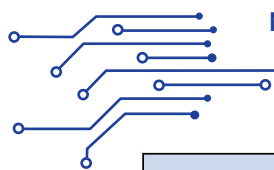


Fig. 4. Correlation Between Features and Target Variables, Data in mm Fors, UY1, UY111

Source: author's calculations.

which is justified in the case of solving the inverse problem, since they are not input characteristics.

4. The correlation results show that the parameter m correlates weakly with other features, which may be due to the fact that its value is almost independent of the input data within a linear model. Therefore, the choice of a nonlinear machine learning method, in particular GPR, to identify complex relationships is well justified.

CONCLUSIONS

To solve the inverse problem of determining K, m from the geometry of the finished hemisphere, a Gauss-

ian regression (GPR) model has been developed and tested.

The key result in determining the value of parameter K is the replacement of the RBF kernel model with Matern, which reduced the median relative uncertainty of the forecast from 50 to 16%. The correctness of choosing GPR to solve the inverse problem is further confirmed by the correlation analysis, which revealed the nonlinear nature of the relationships between the input geometric parameters and the target variables.

The model developed may be used for non-destructive quality control, integration into digital counterparts, and optimization of superplastic forming processes.

REFERENCES

1. Prates P.A., Pereira A.F. Recent advances and applications of machine learning in metal forming processes. *Metals*. 2022;12(8):1342. URL: <https://doi.org/10.3390/met12081342>
2. Kausik A. K., Rashid A. B., Baki R. F., Maktum M. M. J. Machine learning algorithms for manufacturing quality assurance: A systematic review of performance metrics and applications. *Array*. 2025;26:100393. URL: <https://doi.org/10.1016/j.array.2025.100393>.
3. Gordienko A. V., Beliakova T. A. Superplastic pressure forming of a metal sheet into a long die with an isosceles trapezoid cross section. *Lett Mater*. 2025;15(3):184-191. URL: <https://doi.org/10.48612/letters/2025-3-184-191>
4. Aksenov S. A., Mikolaenko V. Accurate determination of uniaxial flow behaviour of superplastic materials. *Eur J Mech A/Solids*. 2025;109:105469. URL: <https://doi.org/10.1016/j.euromechsol.2024.105469>
5. Mahajan A., Badheka V. Engineering applications of superplasticity of metals: review. In: *Advances in Materials Engineering. ICFAMMT 2024. Lecture Notes in Mechanical Engineering*. Singapore: Springer; 2025:245-257. URL: https://doi.org/10.1007/978-981-97-7114-1_20
6. Guo X., Yang W., Zhang Q., Pan J., Xiong C., Wu L. Precision design method for superplastic forming process parameters based on an improved back propagation neural network. *Processes*. 2025;13(10):3070. URL: <https://doi.org/10.3390/pr13103070>

7. Kim K.J., Lee J.W. Analysis of effects on parameters of super-plastic forming process using computer simulation. *J Mech Sci Technol.* 2025;39(9):5191-5196. URL: <https://doi.org/10.1007/s12206-025-0824-y>
8. Tulupova O.P., Gumerova C., Ganieva V.R., Kruglov A.A., Enikeev F.U. Improving the accuracy of finite element modeling of superplastic forming processes. *Lett Mater.* 2022;12(2):142-147. URL: <https://doi.org/10.22226/2410-3535-2022-2-142-147>
9. Giuliano G., Polini W. FEM analysis of superplastic-forming process to manufacture a hemispherical shell. *Appl Sci.* 2025;15(14):8080. URL: <https://doi.org/10.3390/app15148080>
10. Tulupova O.P., Zholobova G.N. Gaussian regression for predicting product geometry based on CAE simulation data. *Digital Solutions and Artificial Intelligence Technologies.* 2025;4(1):43-50. URL: <https://doi.org/10.26794/3033-7097-2025-1-4-43-50>
11. Mishra S.K., Brahma A., Dutta K. Prediction of mechanical properties of Al-Si-Mg alloy using artificial neural network. *Sādhanā.* 2021;46:139. URL: <https://doi.org/10.1007/s12046-021-01660-x>
12. Sun M., Xiong C., Zhou Y., et al. Thickness distribution prediction of superplastic forming parts based on machine learning and computer aided engineering. In: 2024 4th International Symposium on Artificial Intelligence and Intelligent Manufacturing (AIIM). IEEE; 2024:827-830. URL: <https://doi.org/10.1109/AIIM64537.2024.10934610>
13. Hu S., Yang Y., Zhang Q. Optimizing design and performance analysis of prestress-impact synergistic plastic forming equipment based on Gaussian process regression surrogate modeling. *Front Mater.* 2026;13:1794762. URL: <https://doi.org/10.3389/fmats.2026.1794762>
14. Murzina G.R., Ganieva V.R., Kruglov A.A., Enikeev F.U. Modeling of the process of superplastic forming of hemispherical shells from blanks of different profiles. *Lett Mater.* 2021;11(4):548-552. URL: <https://doi.org/10.22226/2410-3535-2021-4-548-552>
15. Tulupova O.P., Kruglov A.A., Enikeev F.U., Murzina G.R., Ganieva V.R. Effect of the strain rate sensitivity index and variable profile blanks in a welded envelope on the thickness distribution of spherical shells manufactured by free superplastic forming. *Lett Mater.* 2026;16(2):166-172. URL: <https://doi.org/10.48612/letters/2026-2-166-172>
16. Giuliano G., Polini W. A new approach to characterize superplastic materials from free-forming test and inverse analysis. *Appl Sci.* 2024;14(23):11113. URL: <https://doi.org/10.3390/app142311113>
17. Zogu P., Schäfer M. Reliability assessment of composite column according to Monte Carlo simulation and Latin hypercube sampling. *E3S Web Conf.* 2024;550(1):01037. URL: <https://doi.org/10.1051/e3sconf/202455001037>
18. Lourenço R., Andrade-Campos A., Georgieva P. The use of machine-learning techniques in material constitutive modelling for metal forming processes. *Metals.* 2022;12:427. URL: <https://doi.org/10.3390/met12030427>
19. Munzone F., Hazrati J., Hakvoort W., van den Boogaard T. Comparative study of artificial neural network and physics-informed neural network application in sheet metal forming. In: 27th International ESAFORM Conference on Material Forming (ESAFORM 2024). 2024:2278-2288. URL: <https://doi.org/10.21741/9781644903131-251>
20. Salman H.A., Kalakech A., Steiti A. Random forest algorithm overview. *Babylonian J Mach Learn.* 2024:69-79. URL: <https://doi.org/10.58496/BJML/2024/007>
21. Wang J. An intuitive tutorial to Gaussian process regression. *Comput Sci Eng.* 2023;25(4):4-11. URL: <https://doi.org/10.1109/MCSE.2023.3342149>
22. Zhang Z., Ikeura R., Hayakawa S., Wang Z. Ship hull steel plate deformation modeling based on Gaussian process regression. *J Mar Sci Eng.* 2024;12(12):2267. URL: <https://doi.org/10.3390/jmse12122267>

ABOUT THE AUTHORS / ИНФОРМАЦИЯ ОБ АВТОРАХ

Olga P. Tulupova — Cand. Sci. (Tech.), Assoc. Prof. of Artificial Intelligence Department of the Faculty of Information Technology and Big Data Analysis, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

Ольга Павловна Тулупова — кандидат технических наук, доцент кафедры ИТ факультета информационных технологий и анализа больших данных, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

<https://orcid.org/0000-0002-8621-0724>

Corresponding author / Автор для корреспонденции:

optulupova@fa.ru



Galina N. Zholobova — Cand. Sci. (Tech.), Assoc. Prof., Deputy Head of Artificial Intelligence Department for academic affairs of the Faculty of Information Technology and Big Data Analysis, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

Галина Николаевна Жолобова — кандидат технических наук, доцент кафедры ИТ, заместитель заведующего кафедрой информационных технологий по учебной работе факультета информационных технологий и анализа больших данных, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

<https://orcid.org/0009-0006-6082-1347>

nzholobova@fa.ru

Farid U. Enikeev — Dr. Sci. (Tech.), Prof. Department of Computer Science and Engineering Cybernetics, Ufa State Petroleum Technological University, Ufa, Russian Federation

Фарид Усманович Еникеев — доктор технических наук, профессор кафедры вычислительной техники и инженерной кибернетики, Уфимский государственный нефтяной технический университет, Уфа, Российская Федерация

<https://orcid.org/0000-0002-0362-4121>

Corresponding author / Автор для корреспонденции:

kobros@yandex.ru

Authors' declared contributions:

O.P. Tulupova — concept, theory, physical interpretation, implementation, programming.

G.N. Zholobova — optimization, conclusions.

F.U. Enikeev — setting goals and objectives, translation.

Заявленный вклад авторов:

О.П. Тулупова — концепция, теория, физическая интерпретация, реализация, программирование.

Г.Н. Жолобова — оптимизация, выводы.

Ф.У. Еникеев — постановка целей и задач, перевод.

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

The article was submitted on 02.04.2026; revised on 13.06.2026 and accepted for publication on 22.08.2026.

The authors read and approved the final version of the manuscript.

Статья поступила 02.04.2026; после рецензирования 13.06.2026; принята к публикации 22.08.2026.

Авторы прочитали и одобрили окончательный вариант рукописи.

О количестве натуральных решений специфического дискретного уравнения и его свойствах

Р.Н. Баротов¹, Д.Н. Баротов²¹ Худжандский государственный университет им. акад. Б. Гафурова, Худжанд, Республика Таджикистан;² Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

АННОТАЦИЯ

В статье исследуется дискретное уравнение в натуральных числах вида $x = \text{count}(d, x) + n$, где n — натуральное число; $\text{count}(d, x)$ — количество вхождений цифры $d \in \{0, 1, \dots, 9\}$ в десятичную запись числа x . **Цель работы** — получить априорные оценки решений уравнения и проанализировать зависимость числа решений от параметров d и n . **Основной результат исследования** — установлена двусторонняя априорная оценка для решений уравнения: $n \leq x \leq n + \lfloor \lg n \rfloor + 2$. Из нее следует: при фиксированных d и n число натуральных решений ограничено; уравнение разрешимо не для всех значений n . Для цифр $d \in \{9, 8, \dots, 2\}$ и любого натурального n доказано, что число решений не превышает 2. Для $d = 1$ и любого натурального n показано, что число решений не превышает 3. Для $d = 0$ конструктивно обосновано, что с ростом n число натуральных решений, каждое из которых может быть представлено в десятичной записи с использованием всего четырех цифр, может неограниченно возрастать. **Теоретическая и практическая значимость.** Результаты имеют теоретико-математический характер и могут быть применены: при исследовании аналогичных дискретных уравнений в натуральных числах; в анализе математических игр-головоломок, описываемых подобными уравнениями; как дополнение к известным результатам в области диофантовых уравнений.

Ключевые слова: количество вхождений цифры d в число x ; неалгебраическое диофантово уравнение; мощность множества; теория алгоритмов; оценка сложности; доказательства существования решений

Для цитирования: Баротов Р.Н., Баротов Д.Н. О количестве натуральных решений специфического дискретного уравнения и его свойствах. *Цифровые решения и технологии искусственного интеллекта*. 2026;2(3):51-57. DOI: 10.26794/3030-7097-2026-2-3-51-57

ORIGINAL PAPER

On the Number of Natural Solutions of a Specific Discrete Equation and Its Properties

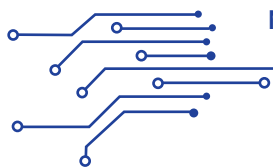
R.N. Barotov¹, D.N. Barotov²¹ Khujand State University named after academician B. Gafurov, Khujand, Republic of Tajikistan;² Financial University under the Government of the Russian Federation, Moscow, Russian Federation

ABSTRACT

The article studies a discrete equation in natural numbers of the form $x = \text{count}(d, x) + n$, where n is a natural number; $\text{count}(d, x)$ is the number of occurrences of the digit $d \in \{0, 1, \dots, 9\}$ in the decimal notation of the number x . The objective of the work is to obtain a priori estimates of solutions to the equation and to analyze the dependence of the number of solutions on the parameters d and n . The main result of the study is the establishment of a two-sided a priori estimate for solutions to the equation: $n \leq x \leq n + \lfloor \lg n \rfloor + 2$. It follows from this estimate that for fixed d and n , the number of natural solutions is limited; the equation is not solvable for all values of n . For digits $d \in \{9, 8, \dots, 2\}$ and any natural n , it is proved that the number of solutions does not exceed 2. For $d = 1$ and any natural n , it is shown that the number of solutions does not exceed 3. For $d = 0$, it is constructively substantiated that with increasing n , the number of natural solutions, each of which can be represented in decimal notation using only four digits, can increase indefinitely. Theoretical and practical significance. The results are of a theoretical-mathematical nature and can be applied in the study of similar discrete equations in natural numbers; in the analysis of mathematical puzzle games described by similar equations; they also complement known results in the field of Diophantine equations.

Keywords: number of occurrences of the digit d in the number x ; non-algebraic Diophantine equation; cardinality of a set; theory of algorithms; complexity estimation; proofs of the existence of solutions

For citation: Barotov R.N., Barotov D.N. On the number of natural solutions of a specific discrete equation and its properties. *Digital Solutions and Artificial Intelligence Technologies*. 2026;2(3):51-57. DOI: 10.26794/3030-7097-2026-2-3-51-57



ВВЕДЕНИЕ

Дискретные уравнения в натуральных числах представляют собой важный класс математических задач, сочетающих элементы теории чисел, комбинаторики и дискретной математики. Их исследование не только углубляет понимание структуры натуральных чисел, но и находит приложения в криптографии, теории алгоритмов, а также в анализе игровых механик и головоломок.

В настоящей работе изучается специфическое дискретное уравнение в натуральных числах вида

$$x = \text{count}(d, x) + n, \quad (1)$$

где: $x, n \in N$ — натуральные числа; $d \in \{0, 1, \dots, 9\}$ — цифра десятичной системы счисления; $\text{count}(d, x)$ — количество вхождений цифры d в десятичную запись числа x .

Актуальность исследования обусловлена двумя ключевыми факторами:

1. **Прикладная мотивация.** Уравнение (1) возникает при математическом описании нетривиальной игры-головоломки [1], где требуется восстановить число x по заданному «смещенному» количеству определенной цифры в его записи. Анализ таких уравнений позволяет формализовать правила игры, оценить сложность уровней и разработать алгоритмы генерации математических игр и задач.

2. **Теоретическая новизна.** Легкость уравнений типа (1) обманчива, именно поэтому, по мнению авторов, они не были достаточно изучены в научной литературе. Отсутствуют общие / конструктивные методы оценки числа решений, анализа их распределения и зависимости от параметров d и n . Это создает пробел в теории диофантовых уравнений, особенно в части неалгебраических дискретных соотношений.

Цель работы — проведение комплексного анализа уравнения (1), включающего получение двусторонней априорной оценки для решений; установление ограничений на число натуральных решений при фиксированных d и n ; исследование асимптотического поведения числа решений при $d = 0$ и росте n .

Основные задачи исследования:

1. Доказать двустороннюю оценку $n \leq x \leq n + \lfloor \lg n \rfloor + 2$ для любого решения x уравнения (1).

2. Показать, что при $d \in \{1, 2, \dots, 9\}$ число решений не превышает трех, причем равенство 3 возможно только при $d = 1$.

3. Конструктивно обосновать неограниченный рост числа решений при $d = 0$ с увеличением n , построив для любого $k \in N$ значение n_k , для которого уравнение имеет не менее k решений, каждое из которых может быть представлено в десятичной записи

с использованием всего четырех цифр 0, 1, 2 и 9.

4. Ввести и использовать аппарат дискретных производных $[\Delta_d(m)]$ и операций конкатенации чисел для анализа монотонности и структуры решений.

Методологическая база опирается на:

- комбинаторные методы подсчета цифр в десятичной записи;
- техники оценки роста функций с логарифмическими слагаемыми;
- аппарат теории множеств — анализ мощности $\text{card}(S_d(n))$;
- конструктивные доказательства существования решений через рекурсивное построение числовых последовательностей.

Научная новизна работы заключается в: первой систематической постановке и решении уравнения (1); получении неулучшаемой априорной оценки решений; доказательстве принципиального различия в поведении числа решений для $d = 0$ и $d \neq 0$.

Практическая значимость результатов состоит в: возможности их применения при разработке математических игр и головоломок с числовыми правилами; в задачах кодирования и сжатия данных, где важно учитывать частоту цифр в числах; как элемента теории дискретных уравнений в курсах высшей математики и теоретической информатики.

Структура статьи соответствует логике исследования: после постановки задачи вводятся необходимые обозначения, затем последовательно доказываются ключевые утверждения об оценках и числе решений, завершается работа выводами и перспективами дальнейших исследований.

ПОСТАНОВКА ЗАДАЧИ

Основной мотивацией авторов к изучению уравнения (1) является то, что, во-первых, оно возникает при математическом описании некоторой нетривиальной игры (головоломки) [1], а во-вторых, результаты, полученные в ходе его анализа, дополняют результаты, полученные в научных работах [2–17]. В результате исследования устанавливается общая двусторонняя априорная оценка решения дискретного уравнения (1), из которой следует, что, во-первых, при фиксированных d и n число его натуральных решений ограничено, а во-вторых, оно разрешимо не при каждом n . Выясняется, что для любой цифры $d \in \{9, 8, \dots, 1\}$ дискретная функция, равная $m - \text{count}(d, m)$, не убывает по m и на базе этого обосновывается, что для каждого натурального n и каждой цифры $d \in \{9, 8, \dots, 1\}$ число натуральных решений урав-



нения вида (1) не более трех, причем если оно равно 3, то $d = 1$. Также конструктивно аргументируется, что при $d = 0$ с ростом n количество натуральных решений, каждое из которых может быть представлено в десятичной записи с использованием четырех цифр, может расти неограниченно.

ИСПОЛЬЗУЕМЫЕ ОБОЗНАЧЕНИЯ

Пусть $S_d(n) = \{x \in \mathbb{N} : x = \text{count}(d, x) + n\}$ — множество натуральных решений уравнения $x = \text{count}(d, x) + n$, $\text{card}(S_d(n))$ — мощность (ко-

личество элементов) множества $S_d(n)$, где, как отмечено выше, n — натуральное число; $\text{count}(d, x)$ — число вхождений цифры $d \in \{0, 1, \dots, 9\}$ в десятичную запись x .

Пусть $\Delta_d(m) = 1 - \text{count}(d, m+1) + \text{count}(d, m)$ — (дискретная) производная (дискретной) функции, равной $m - \text{count}(d, m)$, по m .

Пусть $\text{concat}(a, b) = \overline{a_{p-1} \dots a_1 a_0 b_{q-1} \dots b_1 b_0}$ — натуральное число, полученное путем конкатенации натуральных чисел $a = \overline{a_{p-1} \dots a_1 a_0}$ и $b = \overline{b_{q-1} \dots b_1 b_0}$.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ

Начнем изложение с общей двусторонней априорной оценки решения рассматриваемого дискретного уравнения.

Утверждение 1. Если $x^* \in \mathbb{N}$ — решение уравнения $x = \text{count}(d, x) + n$, то справедлива следующая двусторонняя оценка

$$n \leq x^* \leq n + \lceil \lg(n) \rceil + 2. \quad (2)$$

Доказательство. Пусть $x^* = \text{count}(d, x^*) + n$. Левая часть неравенства (2) очевидна, так как ввиду определения функции $\text{count}(d, x)$ имеем

$$x^* = \text{count}(d, x^*) + n \geq 0 + n = n.$$

Докажем справедливость правой части неравенства (2). Для этого покажем, что имеет место следующее неравенство

$$n + \lceil \lg(n) \rceil + 2 + k - \text{count}(d, n + \lceil \lg(n) \rceil + 2 + k) > n \quad \forall k \in \mathbb{N}. \quad (3)$$

Действительно,

$$\begin{aligned} & n + \lceil \lg(n) \rceil + 2 + k - \text{count}(d, n + \lceil \lg(n) \rceil + 2 + k) > \\ & > n + \lg(n) + 1 + k - \text{count}(d, n + \lceil \lg(n) \rceil + 2 + k) \geq \\ & \geq n + \lg(n) + 1 + k - \lceil \lg(n + \lceil \lg(n) \rceil + 2 + k) \rceil - 1 \geq \\ & \geq n + \lg(n) + k - \lg(n + \lceil \lg(n) \rceil + 2 + k) \geq \\ & \geq n + \lg(n) + k - \lg(n + \lg(n) + 2 + k) = n + k - \lg\left(\frac{n + \lg(n) + 2 + k}{n}\right) = \\ & = n + k - \lg\left(1 + \frac{\lg(n) + 2}{n} + \frac{k}{n}\right) \geq n + k - \lg\left(1 + 2 + \frac{k}{n}\right) \geq \\ & \geq n + k - \lg\left(1 + 2 + \frac{k}{1}\right) = n + k - \lg(k + 3) > n + 0, 3 > n. \end{aligned}$$

Из неравенства (3) очевидным образом следует правая часть соотношения (2). Утверждение 1 доказано.

Замечание 1. Двусторонняя априорная оценка (2) в общем случае является неулучшаемой, так как, например, если $n = 9$ и $d = 1$, то граничные натуральные числа $x_1 = n = 9$ и $x_2 = n + \lceil \lg(n) \rceil + 2 = 9 + \lceil \lg(9) \rceil + 2 = 11$ являются решениями уравнения $x = \text{count}(d, x) + n$.

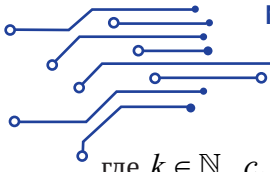
Аргументируем, что для ненулевой цифры d дискретная функция, равная $m - \text{count}(d, m)$, является неубывающей по m .

Утверждение 2. Имеет место следующее неравенство

$$\Delta_d(m) \geq 0 \quad \forall m \in \mathbb{N} \text{ и } \forall d \in \{1, 2, \dots, 9\}. \quad (4)$$

Доказательство. Пусть $c_{k-1} \dots c_1 c_0$ — десятичная запись числа m , т.е.

$$m = c_{k-1} \cdot 10^{k-1} + \dots + c_1 \cdot 10 + c_0,$$



где $k \in \mathbb{N}$, $c_{k-1} \neq 0$ и $c_0, c_1, \dots, c_{k-1} \in \{0, 1, \dots, 9\}$. Пусть $c_k = 0$. Тогда число m можно представить в виде

$$m = c_k \cdot 10^k + c_{k-1} \cdot 10^{k-1} + \dots + c_1 \cdot 10 + c_0 = \overline{c_k c_{k-1} \dots c_1 c_0}.$$

Пусть $s = \min \{j \in \{0, 1, \dots, k\} : c_j \neq 9\}$ — наименьший индекс, такой что $c_s \neq 9$. Тогда

$$\begin{aligned} \Delta_d(m) &= 1 - (\text{count}(d, m+1) - \text{count}(d, m)) = \\ &= \begin{cases} 1 - (\text{count}(d, m) + \delta(c_s) - \text{count}(d, m)), & \text{если } 1 \leq d < 9 \\ 1 - (\text{count}(d, m) - s + \varepsilon(c_s) - \text{count}(d, m)), & \text{если } d = 9 \end{cases} = \\ &= \begin{cases} 1 - \delta_d(c_s), & \text{если } 1 \leq d < 9 \\ s + 1 - \varepsilon(c_s), & \text{если } d = 9 \end{cases} \geq 0, \end{aligned}$$

так как

$$\delta_d(c_s) = \begin{cases} 1, & \text{если } c_s = d - 1 \\ -1, & \text{если } c_s = d \\ 0, & \text{если } c_s \notin \{d-1, d\} \end{cases}, \quad \varepsilon(c_s) = \begin{cases} 1, & \text{если } c_s = 8 \\ 0, & \text{если } c_s \neq 8 \end{cases}, \quad 0 \leq s \leq k.$$

Утверждение 2 доказано.

Замечание 2. Если $d \in \{1, 2, \dots, 9\}$, то $\Delta_d(m) = 0$ тогда и только тогда, когда выполняется одно из следующих двух условий:

- 1) выполняются неравенство $d \neq 9$ и равенство $c_s = d - 1$;
- 2) соблюдаются равенства $d = 9$, $s = 0$ и $c_s = 8$.

Несмотря на то что, ввиду неулучшаемой в общем случае оценки (2), число потенциальных кандидатов на решение неограниченно растет с ростом n , на основании утверждения 2 и замечания 2 обобщим, что если $d \in \{1, 2, \dots, 9\}$, то число натуральных решений уравнения вида (1) не более трех.

Утверждение 3. Имеет место следующее неравенство

$$0 \leq \text{card}(S_d(n)) \leq 3 \quad \forall n \in \mathbb{N} \text{ и } \forall d \in \{1, 2, \dots, 9\}. \quad (5)$$

Доказательство. Левая часть неравенства (5), ввиду определения $\text{card}(S_d(n))$, очевидна, а правую часть неравенства (5) докажем от противного. Пусть существуют $n^* \in \mathbb{N}$ и $d^* \in \{1, 2, \dots, 9\}$ такие, что $\text{card}(S_{d^*}(n^*)) > 3$. Тогда существуют четыре натуральных числа $x_1^* < x_2^* < x_3^* < x_4^*$ такие, что $x_1^*, x_2^*, x_3^*, x_4^* \in S_{d^*}(n^*)$. Отсюда, в силу утверждения 2 и неравенства $x_1^* < x_2^* < x_3^* < x_4^*$, получаем:

$$\begin{aligned} 0 = n^* - n^* &= x_4^* - \text{count}(d^*, x_4^*) - (x_1^* - \text{count}(d^*, x_1^*)) \geq \\ &\geq x_1^* + 3 - \text{count}(d^*, x_1^* + 3) - (x_1^* - \text{count}(d^*, x_1^*)) = \\ &= \Delta_{d^*}(x_1^* + 2) + \Delta_{d^*}(x_1^* + 1) + \Delta_{d^*}(x_1^*). \end{aligned}$$

Для завершения доказательства осталось показать, что

$$\Delta_{d^*}(x_1^* + 2) + \Delta_{d^*}(x_1^* + 1) + \Delta_{d^*}(x_1^*) > 0. \quad (6)$$

Относительно d^* рассмотрим три случая.

Случай 1. Пусть $d^* = 9$. Тогда, в силу утверждения 2 и замечания 2, выполняется $\Delta_{d^*}(x_1^* + 1) + \Delta_{d^*}(x_1^*) > 0$ и, следовательно, верно неравенство (6).

Случай 2. Пусть $9 > d^* > 1$. Тогда, в силу утверждения 2 и замечания 2, выполняется соотношение (6), так как справедливо $\Delta_{d^*}(x_1^* + 1) + \Delta_{d^*}(x_1^*) > 0$.

Случай 3. Пусть $d^* = 1$. Тогда, в силу утверждения 2 и замечания 2, неотрицательная величина $\Delta_{d^*}(x_1^* + 1) + \Delta_{d^*}(x_1^*)$ может равняться 0, но величина $\Delta_{d^*}(x_1^* + 2) + \Delta_{d^*}(x_1^*)$ строго положительна и, следовательно, выполняется неравенство (6). Утверждение 3 доказано.



Замечание 3. Неравенство (5) в общем случае неулучшаемо, так как ввиду утверждения 1 нетрудно заметить, что $S_1(1) = \emptyset$ и $S_1(9) = \{9, 10, 11\}$.

Замечание 4. Если в случае $d \neq 0$ множество $S_d(n)$ непустое, то его элементами являются последовательные натуральные числа.

Далее конструктивно обоснуем, что с ростом n количество специальных натуральных решений уравнения $x = \text{count}(0, x) + n$ может расти неограниченно, а именно, конструктивно докажем, что для каждого натурального k найдется натуральное число n_k такое, что число натуральных решений уравнения $x = \text{count}(0, x) + n_k$, каждое из которых может быть представлено в десятичной записи с использованием всего лишь цифр 0, 1, 2 и 9, не меньше k .

Утверждение 4. Для каждого натурального k найдется натуральное число n_k такое, что число натуральных решений уравнения $x = \text{count}(0, x) + n_k$, каждое из которых может быть представлено в десятичной записи с использованием только цифр 0, 1, 2 и 9, не меньше k .

Доказательство. Аргументируем, что если набор из m натуральных чисел $\{x_1, x_2, \dots, x_m\}$, удовлетворяющих условию

$$10^{s_m-1} \leq x_1 < x_2 < \dots < x_m \leq 10^{s_m} - 1,$$

является подмножеством множества $S_0(n_m)$, где $m, s_m, n_m \in \mathbb{N}$, то новый набор из $m+1$ натуральных чисел

$$\left\{ 2 \cdot 10^{s_m+n_m+1} - 1, \text{concat}(2 \cdot 10^{n_m+1}, x_1), \right. \\ \left. \text{concat}(2 \cdot 10^{n_m+1}, x_2), \dots, \text{concat}(2 \cdot 10^{n_m+1}, x_m) \right\},$$

удовлетворяющих условию

$$10^{s_{m+1}-1} \leq 2 \cdot 10^{s_m+n_m+1} - 1 < \text{concat}(2 \cdot 10^{n_m+1}, x_1) < \\ < \text{concat}(2 \cdot 10^{n_m+1}, x_2) < \dots < \text{concat}(2 \cdot 10^{n_m+1}, x_m) \leq 10^{s_{m+1}} - 1,$$

является подмножеством множества $S_0(n_{m+1})$, где $s_{m+1} = s_m + n_m + 2$, $n_{m+1} = 2 \cdot 10^{s_m+n_m+1} - 1$. Действительно,

$$2 \cdot 10^{s_m+n_m+1} - 1 - \text{count}(0, 2 \cdot 10^{s_m+n_m+1} - 1) = 2 \cdot 10^{s_m+n_m+1} - 1 - 0 = \\ = 2 \cdot 10^{s_m+n_m+1} - 1; \\ \text{concat}(2 \cdot 10^{n_m+1}, x_i) - \text{count}(0, \text{concat}(2 \cdot 10^{n_m+1}, x_i)) = \\ = 2 \cdot 10^{n_m+1} \cdot 10^{s_m} + x_i - \text{count}(0, 2 \cdot 10^{n_m+1}) - \text{count}(0, x_i) = \\ = 2 \cdot 10^{s_m+n_m+1} - (n_m + 1) + x_i - \text{count}(0, x_i) = \\ = 2 \cdot 10^{s_m+n_m+1} - (n_m + 1) + n_m = 2 \cdot 10^{s_m+n_m+1} - 1 \quad \forall i \in \{1, 2, \dots, m\}.$$

Теперь, учитывая $S_0(1) = \{1\}$, если начиная с $n_1 = 1$, $s_1 = 1$, $x_1 = 1$ последовательно повторить $k-1$ раз вышеизложенную идею построения $(m+1)$ натуральных s_{m+1} -значных решений уравнения $x = \text{count}(0, x) + n_{m+1}$ на основе m натуральных s_m -значных решений уравнения $x = \text{count}(0, x) + n_m$, то получим натуральное число n_k и k натуральных чисел, каждое из которых может быть представлено в десятичной записи с использованием только цифр 0, 1, 2 и 9, являющихся решениями уравнения $x = \text{count}(0, x) + n_k$. Утверждение 4 доказано.

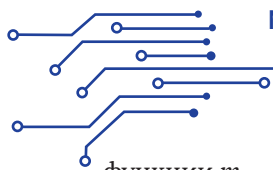
ВЫВОДЫ

В ходе исследования дискретного уравнения вида $x = \text{count}(d, x) + n$ получены следующие ключевые результаты:

1. Априорная оценка решений. Доказана двусторонняя неулучшаемая в общем случае оценка для любого решения x : $n \leq x \leq n + \lceil \lg n \rceil + 2$.

Это позволяет ограничить область поиска решений и служит основой для дальнейшего анализа.

2. Ограниченность числа решений при $d \neq 0$. Для любой цифры $d \in \{1, 2, \dots, 9\}$ и любого натурального n установлено, что число натуральных решений уравнения не превышает трех; случай с тремя решениями возможен только при $d = 1$. Выводы опираются на анализ монотонности дискретной



функции $m = \text{count}(d, m)$ и свойства ее дискретной производной $\Delta_d(m)$.

3. Неограниченный рост числа решений при $d = 0$. Показано, что при $d = 0$ с увеличением n количество решений может расти неограниченно. Для любого натурального k конструктивно построено значение n_k , при котором уравнение имеет **не менее k решений, составленных с использованием только цифр 0, 1, 2 и 9**. Это демонстрирует принципиальное отличие поведения уравнения при $d = 0$ от случая $d \neq 0$.

Конструктивные методы анализа. В работе разработан и применен аппарат: дискретных производных ($\Delta_d(m)$) для исследования монотонности; операции конкатенации чисел для построения множеств решений при $d = 0$. Эти инструменты могут быть полезны при изучении аналогичных дискретных уравнений.

Теоретическая и практическая значимость. Результаты заполняют пробел в теории диофан-

товых уравнений, касающийся неалгебраических дискретных соотношений; дают математический аппарат для анализа числовых головоломок и игр, где требуется восстановление числа по частоте цифр; могут найти применение в задачах кодирования и сжатия данных, где важна статистика цифр в числах.

Перспективы дальнейших исследований: обобщение результатов на другие позиционные системы счисления; изучение уравнений с несколькими цифрами d одновременно; разработка эффективных алгоритмов поиска решений на основе полученных оценок; анализ сложности вычисления функции $\text{count}(d, x)$ в контексте теории алгоритмов.

Таким образом, проведенное исследование не только решает поставленную задачу, но и открывает новые направления для изучения дискретных уравнений в натуральных числах с комбинаторными ограничениями.

REFERENCES

1. Barotov D.N. On the Analysis of a System of Equations Containing a Parameter n and Describing a Special State of a Certain Table of Numbers. *Mathematics*. 2025;14:119. URL: <https://doi.org/10.3390/math14010119>
2. Allouche J.-P., Shallit J. Automatic Sequences: Theory, Applications, Generalizations. *Cambridge University Press*. 2003. URL: <https://doi.org/10.1017/CBO9780511546563>
3. Allouche J.-P., Shallit J. The Ring of k -regular Sequences. *Theoretical Computer Science*. 1992;98:163–197. URL: [https://doi.org/10.1016/0304-3975\(92\)90001-V](https://doi.org/10.1016/0304-3975(92)90001-V)
4. Alekseyev M.A., Johnson D., Sloane N.J.A. On Kaprekar's Junction Numbers. 2021. arXiv:2112.14365. URL: <https://doi.org/10.48550/arXiv.2112.14365>.
5. Gupta S. S. On Some Special Numbers. In: Exploring the Beauty of Fascinating Numbers. Singapore: Springer Nature Singapore; 2025:527–565. URL: https://doi.org/10.1007/978-981-97-2465-9_22
6. Haggarty R. Discrete Mathematics for Computing. Harlow, England; New York: Addison Wesley; 2002. 304 p.
7. Cai Zeyu. An Effective Method for Solving a Class of Transcendental Diophantine Equations. 2025. URL: <https://doi.org/10.48550/arXiv.2510.11753>
8. Pooja Devi, Anjana Rajoria. Research on Solutions for Algebraic and Transcendental Diophantine Equations. *Mathematical Statistician and Engineering Applications*. 2022;71(4);12560-12577. URL: <https://doi.org/10.17762/msea.v71i4.2291>
9. Özkoç Öztürk A., Tekcan A. Integer Solutions of a Special Diophantine Equation. 2011. 1389. URL: <https://doi.org/10.1063/1.3637759>
10. Stewart Ia., Tall D. Algebraic number theory and fermat's last theorem, fourth edition. NW: CRC Press Taylor & Francis Group; 2015. 338 p.
11. Kuncak V., Rinard M., Marnette B. On Algorithms and Complexity for Sets with Cardinality Constraints. *Computing Research Repository — CORR*. 2005;9:1-20. URL: <http://drops.dagstuhl.de/opus/volltexte/2006/512>
12. Fogelklou O., Tucker W., Kreiss G. & Siklosi M. A computer-assisted proof of the existence of solutions to a boundary value problem with an integral boundary condition. *Communications in Nonlinear Science and Numerical Simulation*. 2010.16(3). URL: <https://doi.org/10.1016/j.cnsns.2010.07.008>
13. Øksendal B. Stochastic Differential Equations: An Introduction with Applications. Springer-Verlag; 2000. URL: <https://doi.org/10.1007/978-3-662-03185-8>
14. Garifullin R.N. Classification of semidiscrete equations of hyperbolic type. The case of third-order symmetries. *Theoretical and Mathematical Physics*. 202;217(2):1767–1776. URL: <https://doi.org/10.1134/S0040577923110119>

15. Berg Ja., Sheombarsing R. Rigorous numerics for ODEs using Chebyshev series and domain decomposition. *Journal of Computational Dynamics*. 2021;8(3). URL: <https://doi.org/10.3934/jcd.2021015>
16. Cheban D. One-dimensional monotone nonautonomous dynamical systems. 2024. *Science China Mathematics*. 67(2):281-314. URL: <https://doi.org/10.1007/s11425-021-2084-x>
17. Barotov D.N., Barotov R.N., Mikhaylova S.S., Feklin V.G. On the Equation $x = \text{count}(d, x) + n$ and Its Application. *Mathematics*. 2026. 14(6):972. URL: <https://doi.org/10.3390/math14060972>

ИНФОРМАЦИЯ ОБ АВТОРАХ / ABOUT THE AUTHORS

Рузбой Нумонжонович Баротов — преподаватель кафедры математического анализа им. профессора А. Мухсинова факультета математики, Худжандский государственный университет им. акад. Б. Гафурова, Худжанд, Республика Таджикистан

Ruziboy N. Barotov — Lecturer, Department of Mathematical Analysis named after Professor A. Mukhsinov, Faculty of Mathematics, Khujand State University named after academician Bobojon Gafurov, Khujand, Republic of Tajikistan

<https://orcid.org/0000-0003-3729-6143>
ruzmet.tj@mail.ru

Достонжон Нумонжонович Баротов — старший преподаватель кафедры математики и анализа данных факультета информационных технологий и анализа больших данных, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

Dostonjon N. Barotov — Senior Lecturer, Department of Mathematics and Data Analysis, Faculty of Information Technology and Big Data Analysis, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

<https://orcid.org/0000-0001-5047-7710>

Автор для корреспонденции / Corresponding author:
DNBarotov@fa.ru

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.
Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

Статья поступила 19.12.2025; после рецензирования 12.03.2026; принята к публикации 16.03.2026.

Авторы прочитали и одобрили окончательный вариант рукописи.

The article was submitted on 19.12.2025; revised on 12.03.2026 and accepted for publication on 16.03.2026.

The authors read and approved the final version of the manuscript.



Интеллектуальные системы контроля как инструмент повышения энергоэффективности в энергосервисе

Д.С. Гречушкина¹, Н.В. Гринева², П.Е. Голосов³

^{1,2} Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация;

³ РАНХиГС при Президенте Российской Федерации, Москва, Российская Федерация

АННОТАЦИЯ

Статья посвящена анализу интеллектуальных систем контроля энергопотребления как инструмента повышения энергоэффективности при реализации энергосервисных проектов. Цель исследования заключается в выявлении роли и возможностей таких систем, а также обосновании необходимости адаптации существующих решений для создания единого цифрового контура управления. Методологическую основу исследования составляют принципы системного анализа, методы классификации информационных систем, анализ нормативно-правовой базы энергосервисной деятельности. Эмпирическую базу составили данные Государственного доклада о состоянии энергосбережения в РФ за 2023 г. и аналитические обзоры рынка АСКУЭ. В результате исследования выявлено, что ни один из существующих классов систем в отдельности не обладает полным функционалом, необходимым для сквозного управления энергосервисным проектом. Доля интеллектуальных приборов учета в России достигает 72%. Проанализированы итоги реализации энергосервисных контрактов в 2023 г. Выводы исследования указывают на то, что ключевым направлением развития является адаптация и интеграция существующих систем для формирования единого цифрового контура управления энергосервисными проектами. Технологическая база для цифровизации уже сформирована, однако существующие решения требуют адаптации под комплексные задачи бизнеса.

Ключевые слова: энергосервис; цифровизация энергетики; интеллектуальные системы; мониторинг энергопотребления; АСКУЭ; адаптация систем; энергоэффективность; энергосервисные контракты

Для цитирования: Гречушкина Д.С., Гринева Н.В., Голосов П.Е. Интеллектуальные системы контроля как инструмент повышения энергоэффективности в энергосервисе. *Цифровые решения и технологии искусственного интеллекта*. 2026;2(3):58-66. DOI: 10.26794/3030-7097-2026-2-3-58-66

ORIGINAL PAPER

Application of Intelligent Energy Consumption Control Systems in Energy Service Projects

D.S. Grechushkina¹, N.V. Grineva², P.E. Golosov³

^{1,2} Financial University under the Government of the Russian Federation, Moscow, Russian Federation;

³ RANEPA, Moscow, Russian Federation

ABSTRACT

This article analyzes intelligent energy consumption control systems as a tool for improving energy efficiency in the implementation of energy service projects. The aim of the study is to identify the role and capabilities of such systems and to substantiate the need for adapting existing solutions to create a unified digital management framework. The methodological basis of the research consists of the principles of system analysis, methods for classifying information systems, and analysis of the regulatory framework for energy service activities. The empirical base includes data from the State Report on the State of Energy Saving in the Russian Federation for 2023 and analytical reviews of the ASKUE market. The study reveals that none of the existing classes of systems individually possesses the full functionality necessary for end-to-end management of an energy service project. The share of intelligent metering devices in Russia reaches 72%. The results of energy service contracts implementation in 2023 are analyzed. The findings indicate that the key direction of development is the adaptation and integration of existing systems to form a unified digital management framework for energy service projects. The technological base for digitalization has already been formed, but existing solutions require adaptation for complex business tasks.

Keywords: energy service; digitalization of energy; intelligent systems; energy consumption monitoring; ASKUE; system adaptation; energy efficiency; energy service contracts

For citation: Grechushkina D.S., Grineva N.V., Golosov P.E. Application of intelligent energy consumption control systems in energy service projects. *Digital Solutions and Artificial Intelligence Technologies*. 2026;2(3):58-66. DOI: 10.26794/3030-7097-2026-2-3-58-66



ВВЕДЕНИЕ

Современный этап развития экономики характеризуется активным внедрением цифровых технологий во все отрасли хозяйственной деятельности. Энергетический сектор, являясь системообразующим элементом национальной экономики, также претерпевает процессы цифровой трансформации. Повышение энергетической эффективности рассматривается как стратегическая задача, направленная на снижение издержек, повышение конкурентоспособности и обеспечение устойчивого развития¹.

Энергосервисная деятельность выступает ключевым инструментом реализации политики энергосбережения. В отличие от традиционных инвестиционных проектов, энергосервисный контракт предполагает возврат инвестиций за счет фактически достигнутой экономии энергоресурсов, что закреплено Федеральным законом «Об энергосбережении и о повышении энергетической эффективности»². Таким образом, контроль фактического эффекта требует непрерывного анализа потребления энергии и выявления отклонений от нормативных значений. Согласно требованиям стандарта ISO 50001 системы энергетического менеджмента должны опираться на измеримые показатели и аналитические данные для принятия управленческих решений³.

В научной литературе проблематика инновационного развития энергетического бизнеса рассматривается в работах Л.Д. Гительмана, М.В. Кожевникова и Б.Е. Ратникова, которые анализируют стратегии цифровой трансформации и интеллектуализации управления в энергетике [1]. Е.П. Грабчак в своей монографии воспринимает цифровую трансформацию электроэнергетики как комплексный процесс внедрения цифровых технологий, включая интеллектуальные системы учета и управления⁴. Вопросы развития энерго-

сервисной деятельности и повышения энергетической эффективности обсуждаются в работе Д.В. Аверкина, который анализирует механизм энергосервиса на примере Москвы [3]. Тема автоматизации и применения информационных технологий в энергетике, включая интеллектуальные системы мониторинга и облачные технологии, прослеживается в обзоре А.А. Егорова по итогам профильной конференции [4]. Применение современных цифровых технологий, включая облачные и граничные вычисления, для создания интеллектуальных энергосистем изучается в работах И.Н. Колосок и Е.С. Коркиной, Л.В. Глухова, Е.В. Пупкова, А.О. Бучнева [5–8].

Согласно данным аналитических исследований рынка, доля интеллектуальных приборов учета в общей структуре регистрируемых устройств в России достигает 72%⁵. При этом уровень внедрения автоматизированных систем коммерческого учета (АСКУЭ) варьируется в зависимости от категории потребителей: наиболее высокие показатели наблюдаются в промышленном секторе, тогда как в бюджетной сфере и жилищно-коммунальном хозяйстве потенциал цифровизации остается значительным.

При этом в научных исследованиях недостаточно внимания уделено системному анализу роли интеллектуальных систем именно «как инструмента» повышения энергоэффективности в энергосервисе, отсутствует комплексная оценка их возможностей и ограничений в контексте подтверждения экономии энергоресурсов. Данное исследование направлено на восполнение этого пробела.

Вместе с тем анализ практики реализации энергосервисных проектов показывает, что учет энергопотребления часто осуществляется на базе разрозненных программно-аппаратных комплексов. Отсутствие единой архитектуры и интеллектуальных механизмов анализа затрудняет формирование объективной картины потребления и создает риски конфликтов между сторонами контракта.

В этой связи возникает необходимость адаптации комплексной интеллектуальной системы

¹ Распоряжение Правительства РФ от 09.06.2020 № 1523-р «Об Энергетической стратегии РФ на период до 2035 года». URL: <http://static.government.ru/media/acts/files/1202006110003.pdf>

² Федеральный закон от 23.11.2009 № 261-ФЗ «Об энергосбережении и о повышении энергетической эффективности и о внесении изменений в отдельные законодательные акты Российской Федерации» (с изм. и доп.). URL: <https://base.garant.ru/12171109/?ysclid=msfsfaq5m435188327>

³ ГОСТ Р ИСО 50001–2023. Системы энергетического менеджмента. Требования и руководство по применению. М.: Российский институт стандартизации; 2023. 48 с. URL: <https://dqs-russia.ru/wp-content/uploads/2025/07/gost-r-iso-50001-2023.pdf?ysclid=msfsyc922v372507064>

⁴ Федеральный закон от 23.11.2009 № 261-ФЗ «Об энергосбережении и о повышении энергетической эффективности и о внесении изменений в отдельные законодательные

акты Российской Федерации» (с изм. и доп.). URL: <https://base.garant.ru/12171109/?ysclid=msfsfaq5m435188327>

⁵ Анализ рынка АСКУЭ в России, странах СНГ, Средней и Юго-Восточной Азии: переход на интеллектуальные системы с беспроводной передачей данных. MegaResearch. 2025. URL: https://www.megaresearch.ru/news_in/rynkiaskue-v-rossii-i-stranah-sng-sredney-i-yugo-vostochnoy-azii-perehod-na-intellektualnye-sistemy-s-besprovodnoy-peredachey-dannyh?ysclid=msfue1t2b8820909488



контроля энергопотребления, способной обеспечить автоматизированный сбор, обработку и анализ данных, а также подтверждение достигнутой экономии.

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ЦИФРОВИЗАЦИИ ЭНЕРГОСЕРВИСА

Особенностью энергосервисных проектов является их высокая зависимость от качества управления информацией. На всех этапах жизненного цикла энергосервисного контракта формируются и обрабатываются значительные объемы данных о потреблении энергетических ресурсов, режимах работы оборудования и условиях эксплуатации объектов. В этой связи энергосервисная деятельность обоснованно рассматривается как объект цифровизации и автоматизации управления. В традиционной модели энергосервисной деятельности учет потребления энергоресурсов часто осуществляется с использованием разрозненных средств измерения и локальных программных решений. Такие системы, как правило, ориентированы на фиксирование показаний приборов учета и формирование отчетности, но не обеспечивают комплексного анализа и поддержки управленческих решений. В условиях реализации энергосервисных контрактов этого оказывается

недостаточно, поскольку требуется не только учет, но и постоянный мониторинг, прогнозирование и анализ отклонений от плановых показателей.

Интеллектуальные системы контроля энергопотребления представляют собой следующий этап эволюции систем учета, интегрирующих технологии интернета вещей (IoT), обработки больших данных и аналитические платформы. Основные процессы энергосервисной деятельности и возможности их автоматизации с применением таких систем представлены в *табл. 1*.

Анализ *табл. 1* показывает, что большинство процессов энергосервисной деятельности требуют использования информационных технологий, а их ручная реализация приводит к росту трудоемкости и снижению точности расчетов. Интеллектуальные системы позволяют автоматизировать эти процессы, обеспечивая непрерывность контроля и повышение достоверности результатов.

На *рисунке* представлена обобщенная схема энергосервисной деятельности как объекта цифровизации, отражающая место интеллектуальных систем в управлении энергосервисными проектами.

Таким образом, интеллектуальные системы контроля выступают связующим звеном между техническим уровнем учета энергоресурсов и управленческим уровнем принятия решений,

Таблица 1 / Table 1

Основные процессы энергосервисной деятельности и возможности их цифровизации /
Key Processes of Energy Service Activities and Their Digitalization Possibilities

Процесс / Process	Содержание / Content	Возможности автоматизации / Automation Capabilities
Энергоаудит	Сбор исходных данных, анализ потребления	Использование информационных систем для хранения и обработки данных
Формирование базовой линии	Расчет базового уровня энергопотребления	Алгоритмизации расчетов, снижение субъективности
Мониторинг потребления	Учет фактического потребления ресурсов	Интеграция с автоматизированной системой контроля и учета энергоресурсов и интеллектуальными счетчиками
Контроль экономии	Сравнение базовых и фактических показателей	Автоматический расчет достигнутой экономии
Отчетность	Формирование отчетов для сторон договора	Генерация отчетов в информационной системе в установленном формате

Источник / Source: составлено авторами / Compiled by the authors.



Рис. / Fig. Энергосервисная деятельность как объект цифровизации управления / Energy Service Activities as an Object of Management Digitalization

Источник / Source: составлено авторами / Compiled by the authors.

обеспечивая прозрачность реализации энергосервисных контрактов и создавая информационную основу для объективного подтверждения достигнутой экономии.

АНАЛИЗ СОВРЕМЕННЫХ СИСТЕМ КОНТРОЛЯ ЭНЕРГОПОТРЕБЛЕНИЯ

В энергосервисной деятельности применяются различные типы информационных систем, которые можно классифицировать по функциональному назначению. К ним относятся автоматизированные системы коммерческого учета энергоресурсов (АСКУЭ), системы диспетчеризации (SCADA), системы энергетического менеджмента (EMS), а также аналитические платформы класса Business Intelligence (BI).

По данным Государственного доклада⁶ о состоянии энергосбережения и повышения энергетической эффективности в Российской Федерации за 2023 год, рынок энергосервисных контрактов демонстрирует устойчивую динамику. В 2023 г. было заключено 536 контрактов общей стоимостью 53,5 млрд руб. Накопленным итогом за период 2016–2023 гг. заключено 5814 контрактов на сумму 236,9 млрд руб.⁷ Распределение контрактов по отраслям представлено в табл. 2.

⁶ URL: https://www.economy.gov.ru/material/directions/investicionnaya_deyatelnost/povyshenie_energoeffektivnosti/gosudarstvennyy_doklad/gosudarstvennyy_doklad_o_sostoyanii_energoberezheniya_i_povyshenii_energeticheskoy_effektivnosti_v_rossiyskoy_federacii_za_2024_god.html?ysclid=msfsozwxsg622934010

⁷ Государственный доклад о состоянии энергосбережения и повышения энергетической эффективности в Российской Федерации в 2023 году. URL: https://www.economy.gov.ru/material/directions/investicionnaya_deyatelnost/povyshenie_energoeffektivnosti/gosudarstvennyy_doklad/gosudarstvennyy_doklad_o_sostoyanii_energoberezheniya_i_povyshenii_energeticheskoy_effektivnosti_v_rossiyskoy_federacii_za_2024_god.html?ysclid=msfv6l6sbs553269250

Анализ табл. 2 свидетельствует о сохранении высокого спроса на энергосервис в бюджетной сфере (79,9% контрактов) и в сфере модернизации уличного освещения. При этом наибольшая стоимость контрактов приходится на электроэнергетику (84,8%), что обусловлено высокой капиталоемкостью проектов в данной отрасли.

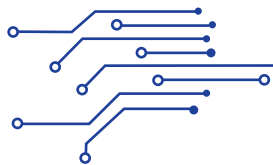
Согласно данным аналитических исследований, доля интеллектуальных приборов учета в общей структуре регистрирующих устройств в России составляет 72%⁸. Наиболее высокие показатели внедрения автоматизированных систем учета наблюдаются в промышленном секторе, тогда как в бюджетной сфере и жилищно-коммунальном хозяйстве сохраняется значительный потенциал для цифровизации.

Классификация информационных систем, применяемых в энергосервисных проектах, представлена в табл. 3.

Анализ табл. 3 показывает, что ни один из существующих классов систем в отдельности не обладает полным функционалом, необходимым для сквозного управления энергосервисным проектом. Это обуславливает необходимость адаптации и интеграции существующих решений для создания единой интеллектуальной системы, объединяющей все требуемые функции.

Современные интеллектуальные системы контроля энергопотребления представляют собой многоуровневые архитектуры, включающие уровни сбора данных на базе IoT-устройств, передачи

⁸ Анализ рынка АСКУЭ в России, странах СНГ, Средней и Юго-Восточной Азии: переход на интеллектуальные системы с беспроводной передачей данных. MegaResearch. 2025. URL: https://www.megaresearch.ru/news_in/rynki_askue-v-rossii-i-stranah-sng-sredney-i-yugo-vostochnoy-azii-perehod-na-intellektualnye-sistemy-s-besprovodnoy-peredachey-dannyh?ysclid=msfue1t2b8820909488

Распределение энергосервисных контрактов по отраслям, 2023 г. /
Distribution of Energy Service Contracts by Industry, 2023

Отрасль / Industry	Количество / Quantity		Стоимость / Cost	
	ед.	%	млрд руб.	%
Бюджетная сфера	428	79,9	4,9	9,2
Жилищная сфера	3	0,5	0,1	0,2
Коммунальное хозяйство	8	1,5	0,3	0,6
Уличное освещение	82	15,3	2,8	5,2
Электроэнергетика	15	2,8	45,4	84,8
Общий итог	536	100	53,5	100

Источник / Source: составлено авторами по данным Минэнерго России / Compiled by the authors based on data from the Ministry of Energy of the Russian Federation. URL: https://www.economy.gov.ru/material/directions/investicionnaya_deyatelnost/povyshenie_energoeffektivnosti/gosudarstvennyy_doklad/gosudarstvennyy_doklad_o_sostoyanii_energoberezeniya_i_povyshenii_energeticheskoy_effektivnosti_v_rossiyskoy_federacii_za_2024_god.html?ysclid=msfv6l6sbs553269250

с использованием промышленных протоколов, хранения в time-series базах данных, аналитической обработки и визуализации. Однако для эффективного применения в энергосервисе требуется их адаптация под конкретные задачи бизнеса.

Современные интеллектуальные системы контроля энергопотребления представляют собой многоуровневые архитектуры, включающие:

- **уровень сбора данных** — интеллектуальные приборы учета (smart meters), датчики, контроллеры с поддержкой промышленных протоколов (Modbus, BACnet, M-Bus, IEC 61850, MQTT);
- **уровень передачи и интеграции** — шлюзы, брокеры сообщений, обеспечивающие надежную передачу данных;
- **уровень хранения** — специализированные time-series базы данных (InfluxDB, TimescaleDB, ClickHouse) для эффективной работы с временными рядами;
- **уровень аналитической обработки** — модули M&V, алгоритмы обнаружения аномалий, методы машинного обучения;
- **уровень визуализации и отчетности** — дашборды, адаптированные под различные роли пользователей.

Ключевым преимуществом таких систем является возможность не только фиксировать потребление, но и проводить его углубленный анализ с использованием современных методов обработки данных.

МЕТОДЫ И ТЕХНОЛОГИИ АНАЛИЗА ДАННЫХ В ЭНЕРГОСЕРВИСЕ

Для успешной реализации энергосервисных проектов недостаточно просто установить приборы учета — важно уметь правильно анализировать полученные данные. Именно анализ позволяет ответить на главный вопрос: достигнута ли заявленная экономия энергоресурсов?

Ключевым элементом аналитической обработки является построение базовой линии энергопотребления (baseline) — эталонного уровня, который характеризовал бы объект при отсутствии энергосберегающих мероприятий. Формализованная модель базовой линии может быть представлена следующим образом:

$$BL = f(X_1, X_2, \dots, X_n) + \varepsilon,$$

где BL — базовая линия энергопотребления; $X_1, X_2, \dots, X_n$ — факторы, влияющие на энергопотребление (температура наружного воздуха, режим работы оборудования, производственные графики); f — функция зависимости; ε — погрешность модели.

Достигнутая экономия рассчитывается как разность между скорректированной базовой линией и фактическим потреблением:

$$E = BL_{corr} - A,$$

где E — экономия; BL_{corr} — скорректированная базовая линия; A — фактическое потребление.

Классификация информационных систем, применяемых в энергосервисных проектах / Classification of Information Systems Used in Energy Service Projects

Тип информационных систем / Type of Information Systems	Назначение / Purpose	Роль в энергосервисе / Role in the Energy Service	Уровень внедрения / Implementation Level, %
АСКАУ	Коммерческий учет	Подтверждение фактического потребления	68
Системы диспетчеризации	Оперативный контроль	Мониторинг режимов работы	42
EMS	Управление энергопотреблением	Оптимизация и экономия	35
BI-системы	Анализ и визуализация данных	Оценка эффективности проектов	28
Интегрированные платформы (IoT + аналитика)	Комплексный контроль и анализ	Сквозное управление энергосервисом	15

Источник / Source: составлено авторами на основе / Compiled by the authors based on URL: https://www.megaresearch.ru/news_in/rynki-askue-v-rossii-i-stranah-sng-sredney-i-yugo-vostochnoy-azii-perehod-na-intellektualnye-sistemy-s-besprovodnoy-predachey-dannyh?ysclid=msfue1t2b8820909488

Для обеспечения единообразия расчетов применяется международный протокол IPMVP (International Performance Measurement and Verification Protocol)⁹, определяющий четыре подхода к верификации экономии: частичное измерение (А), полное измерение (В), анализ объекта в целом (С) и калиброванное моделирование (D). Выбор подхода зависит от специфики объекта и требований сторон контракта.

РЕЗУЛЬТАТЫ И ПЕРСПЕКТИВЫ РАЗВИТИЯ

Анализ статистических данных показывает устойчивый рост рынка энергосервисных контрактов: накопленная стоимость контрактов за 2016–2023 гг. достигла 236,9 млрд руб.¹⁰ Структура распределения контрактов демонстрирует существенную диспропорцию между количеством и стоимостью проектов в различных отраслях. При

доминировании бюджетной сферы по количеству контрактов (79,9%) ключевая доля стоимости приходится на электроэнергетику (84,8%), что объясняется высокой капиталоемкостью проектов по модернизации систем учета и распределения электроэнергии.

Выявленный уровень внедрения интеллектуальных приборов учета (72%) создает технологическую базу для развития систем аналитической обработки данных¹¹. Как отмечают И.Н. Колосок и Е.С. Коркина, использование облачных и граничных технологий позволяет обрабатывать большие объемы данных в реальном времени, что критически важно для управления современными энергообъектами и создания интеллектуальных систем контроля [5]. Однако наличие приборов учета и базовых АСКУЭ само по себе не гарантирует эффективного управления энергопотреблением в рамках энергосервисных контрактов.

Классификация систем, представленная в табл. 3, демонстрирует функциональные ограничения каждого класса. АСКУЭ, являясь наиболее распространенным классом, обеспечивают только

⁹ International Performance Measurement and Verification Protocol (IPMVP). Vol. 1. Washington, D.C: Efficiency Valuation Organization; 2023. 142 p. URL: <https://evo-world.org/en/products-services-mainmenu-en/protocols/ipmvp>

¹⁰ Государственный доклад о состоянии энергосбережения и повышения энергетической эффективности в Российской Федерации в 2023 году. URL: https://www.economy.gov.ru/material/directions/investicionnaya_deyatelnost/povyshenie_energoeffektivnosti/gosudarstvennyy_doklad/gosudarstvennyy_doklad_o_sostoyanii_energobrezheniya_i_povyshenii_energeticheskoy_effektivnosti_v_rossiyskoy_federacii_za_2024_god.html?ysclid=msfv6l6sbs553269250

¹¹ Анализ рынка АСКУЭ в России, странах СНГ, Средней и Юго-Восточной Азии: переход на интеллектуальные системы с беспроводной передачей данных. MegaResearch. 2025. URL: https://www.megaresearch.ru/news_in/rynki-askue-v-rossii-i-stranah-sng-sredney-i-yugo-vostochnoy-azii-perehod-na-intellektualnye-sistemy-s-besprovodnoy-predachey-dannyh?ysclid=msfue1t2b8820909488



учет, но не позволяют рассчитывать экономию. EMS-системы ориентированы на оптимизацию, но требуют интеграции с приборами учета. VI-платформы обладают мощными аналитическими возможностями, но не имеют прямого доступа к данным. Данное обстоятельство подтверждает основной тезис исследования: ни одно из существующих решений в отдельности не покрывает все потребности энергосервисного бизнеса.

Ключевым выводом исследования является понимание того, что технические предпосылки для создания эффективных систем контроля уже сформированы. Основная задача сегодня заключается не в разработке принципиально новых технологий, а в адаптации и интеграции существующих решений под комплексные задачи энергосервисного бизнеса.

Бизнесу требуется система, которая в едином интерфейсе обеспечивает:

- непрерывный сбор данных с любых типов приборов учета независимо от производителя и используемых протоколов;
- автоматическое формирование базовых линий с учетом внешних факторов (температура, режимы работы, производственные графики);
- расчет экономии по методологии IPMVP с возможностью выбора подхода верификации;
- генерацию отчетов для заказчика в требуемом формате с соблюдением нормативных требований;
- интеграцию с бухгалтерскими и ERP-системами для автоматизации расчетов и формирования закрывающих документов;
- разграничение прав доступа для разных участников процесса (заказчик, исполнитель, регулирующие органы).

Полученные выводы о необходимости развития энергосервисной деятельности подтверждаются в исследованиях Д.В. Аверкина, который на практическом примере г. Москвы демонстрирует актуальность внедрения энергосервисных механизмов для повышения энергоэффективности [3]. Полученные выводы о необходимости интеграции интеллектуальных систем в управление энергопотреблением согласуются с исследованиями Л.Д. Гительмана и соавторов, которые подчеркивают клю-

чевую роль цифровых стратегий в повышении эффективности энергетического бизнеса [1].

Проведенный анализ открывает ряд направлений для дальнейших исследований. Необходима разработка методик оценки экономической эффективности внедрения интеллектуальных систем, позволяющих количественно обосновывать целесообразность инвестиций в адаптацию и интеграцию существующих решений с учетом специфики различных отраслей и типов объектов. Требуется унификация требований к системам контроля под различные категории потребителей и бизнес-процессы, обеспечивающая их масштабируемость и адаптивность к отраслевой специфике при сохранении базового функционального ядра. Перспективным направлением является исследование применения методов машинного обучения для прогнозирования энергопотребления, что позволит не только подтверждать достигнутую экономию, но и заблаговременно выявлять риски отклонений.

ВЫВОДЫ

Интеллектуальные системы контроля энергопотребления являются необходимым инструментом повышения энергоэффективности в энергосервисных проектах. Их применение обеспечивает непрерывный мониторинг и объективное подтверждение достигнутой экономии энергоресурсов.

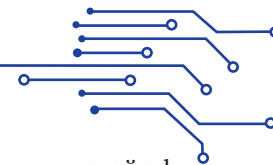
Существующие классы систем автоматизации обладают функциональными ограничениями. АСКУЭ обеспечивают только учет, EMS ориентированы на оптимизацию, VI-платформы дают аналитику, но ни одно из решений не покрывает все задачи энергосервисного проекта.

Технологическая база для цифровизации энергосервиса уже сформирована. Основная задача сегодня — адаптация и интеграция существующих систем в единый цифровой контур управления, объединяющий весь необходимый бизнесу функционал.

Практическая значимость работы заключается в возможности использования полученных выводов при выборе направлений адаптации интеллектуальных систем под задачи энергосервисных компаний.

СПИСОК ИСТОЧНИКОВ

1. Гительман Л.Д., Кожевников М.В., Ратников Б.Е. Инновационные стратегии энергетического бизнеса: путеводитель с активной цифровой компонентой для руководителей и кадрового резерва. Екатеринбург: Автограф; 2024. 196 с. URL: <https://elibrary.ru/bdeysl>
2. Грабчак Е.П. Цифровая трансформация электроэнергетики. Монография. Москва: Русайнс; 2024. 210 с. URL: <https://elibrary.ru/bllugl>



3. Аверкин Д.В. Развитие механизма энергосервиса как направления повышения энергетической эффективности и процесса энергосбережения на предприятиях (на примере г. Москва). *Экономика: вчера, сегодня, завтра*. 2023;13(5-1):280-286. URL: <https://doi.org/10.34670/AR.2023.59.33.112>
4. Егоров А.А. XII научно-практическая конференция «Автоматизация и информационные технологии в энергетике 2023» в рамках деловой программы Международного форума «Электрические сети России 2023»: обзор. Ч. 1. *Автоматизация и ИТ в энергетике*. 2023;12(173):22-35. URL: <https://elibrary.ru/auvzql>
5. Колосок И.Н., Коркина Е.С. Применение облачных и граничных технологий при развитии интеллектуальной энергосистемы. *Информационные и математические технологии в науке и управлении*. 2024;4:73-84. URL: <https://doi.org/10.25729/ESI.2024.36.4.008>
6. Глухова Л.В., Казиева Б.В., Казиев В.М., Шерстобитова А.А. Цифровая экономика и энергосервисный договор. *Вестник Волжского университета им. В.Н. Татищева*. 2024;21(53):18-26. URL: https://doi.org/10.51965/2076-7919_2024_2_1_19
7. Пупкова Е.В., Дулесов А.С., Дулесова Н.В. Анализ данных мониторинга электропотребления и выявление сверхпотребления электроэнергии на основе предлагаемого алгоритма. *Современные инновации, системы и технологии*. 2024;4(2):280-290. URL: <https://doi.org/10.47813/2782-2818-2024-4-2-0280-0290>
8. Бучнев А.О., Рыбина К.А. Интеллектуальные системы учета электроэнергии как фактор развития «умного города». *Государственная служба*. 2025;27(4(156)):43-51. URL: <https://elibrary.ru/eupauy>

REFERENCES

1. Gitelman L.D., Kozhevnikov M.V., Ratnikov B.E. Innovative strategies of the energy business: a guide with an active digital component for managers and the personnel reserve. Yekaterinburg: Autograph; 2024. URL: <https://elibrary.ru/bdeysl> (In Russ.).
2. Grabchak E.P. Digital transformation of the electric power industry. Monograph. Moscow: Rusains; 2024. URL: <https://elibrary.ru/bllugl> (In Russ.).
3. Averkin D.V. Development of the energy service mechanism as a direction for improving energy efficiency and the process of energy saving at enterprises on the example of Moscow. *Economics: Yesterday, Today and Tomorrow*. 2023;13(5-1):280-286. (In Russ.). URL: <https://doi.org/10.34670/AR.2023.59.33.112>
4. Egorov A.A. XII scientific and practical conference “Automation and Information Technologies in the energy sector 2023” within the framework of the business program of the International Forum “Electric Networks of Russia 2023”: Overview. Part 1. *Automation and IT in the Energy Sector*. 2023;12(173):22-35. URL: <https://elibrary.ru/auvzql> (In Russ.).
5. Kolosok I.N., Korkina E.S. Application of cloud and edge technologies in the development of an intelligent energy system. *Information and Mathematical Technologies in Science and Management*. 2024;4:73-84. (In Russ.). URL: <https://doi.org/10.25729/ESI.2024.36.4.008>
6. Glukhova L.V., Kazieva B.V., Kaziev V.M., Sherstobitova A.A. Digital economy and energy service contract. *Bulletin of the V.N. Tatishchev Volga State University*. 2024;21(53):18-26. (In Russ.). URL: https://doi.org/10.51965/2076-7919_2024_2_1_19
7. Pupkova E.V., Dulesov A.S., Dulesova N.V. Analysis of power consumption monitoring data and detection of overconsumption of electricity based on the proposed algorithm. *Modern Innovations, Systems and Technologies*. 2024;4(2):280-290. (In Russ.). URL: <https://doi.org/10.47813/2782-2818-2024-4-2-0280-0290>
8. Buchnev A.O., Rybina K.A. Intelligent electricity metering systems as a factor in the development of a “smart city”. *Public Service*. 2025;27(4(156)):43-51. (In Russ.). URL: <https://elibrary.ru/eupauy>

ИНФОРМАЦИЯ ОБ АВТОРАХ / ABOUT THE AUTHORS

Дарья Сергеевна Гречушкина — студентка магистратуры, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

Daria S. Grechushkina — master’s degree student, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

<https://orcid.org/0009-0002-3720-1407>

Автор для корреспонденции / Corresponding author:

248936@fa.ru



Наталья Владимировна Гринева — кандидат экономических наук, доцент, доцент кафедры информационных технологий, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

Natalia V. Grineva — Cand. Sci. (Econ.), Assoc. Prof., Assoc. Prof. of the Department of Information Technology, Financial University under the Government of the Russian Federation, Moscow, Russian Federation

<https://orcid.org/0000-0001-7647-5967>

ngrineva@fa.ru

Павел Евгеньевич Голосов — кандидат технических наук, директор Института общественных наук РАНХиГС, Москва, Российская Федерация

Pavel E. Golosov — Cand. Sci. (Tech.), Director of the Institute of Social Sciences, RANEPA, Moscow, Russian Federation

<https://orcid.org/0000-0003-4313-0887>

Автор для корреспонденции / Corresponding author:

golosov-pe@ranepa.ru

Заявленный вклад авторов:

Д.С. Гречушкина — аннотация, сбор и анализ материалов, подготовка текста статьи, оформление таблиц и рисунков.

Н.В. Гринева — разработка общей концепции статьи, постановка цели и задач исследования, корректировка выводов.

П.Е. Голосов — введение, заключение, список литературы.

Authors' declared contributions:

D.S. Grechushkina — abstract, collection and analysis of materials, preparation of the text of the article, design of tables and figures.

N. V. Grineva — development of the overall concept of the article; setting the goal and objectives of the study, adjusting the conclusions.

P.E. Golosov — introduction, conclusion, references.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

Статья поступила 19.06.2026; после рецензирования 02.07.2026; принята к публикации 31.07.2026.

Авторы прочитали и одобрили окончательный вариант рукописи.

The article was submitted on 19.06.2026; revised on 02.07.2026 and accepted for publication on 31.07.2026.

The authors read and approved the final version of the manuscript.

Анализ некоторых аспектов цифровизации в области оказания образовательных услуг населению методами статистического анализа и машинного обучения

Л.Р. Борисова, М.Н. Фридман

Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

АННОТАЦИЯ

Цифровая трансформация во всех сферах жизни значительно ускорилась после COVID-19. Во всем мире и в России, в частности, в последние пять лет наблюдается существенный прогресс в области профессиональных навыков в цифровых технологиях. Это одновременно вызов и возможности, когда вопросы интеграции и роста пересекаются. В нашей стране цифровизация в сфере государственных услуг населению развивается опережающими темпами. России необходимы образовательные системы, соответствующие цифровому веку, а также дополнительные программы для обучения и переподготовки той части взрослого населения, которая давно прекратила учебу. Это двусторонняя задача, касающаяся как базовых цифровых навыков, так и современного понимания грамотности. Цель развития цифровых образовательных услуг населению — формирование доступной и эффективной системы обучения для всех категорий граждан. Задача исследования — выявить объективными методами математического анализа данных (без какого-либо использования субъективных оценок) наличие или отсутствие взаимосвязи между социально-экономическими параметрами развития региона Российской Федерации и платными образовательными услугами населению. В работе использованы методы кластеризации и классификации однородных объектов по широкому набору социально-экономических и финансовых показателей, важных для исследования развития цифровизации в сфере услуг для населения. Получено четкое разделение регионов на две большие, практически равные группы по этому набору показателей. Представлен сравнительный анализ методов иерархической кластеризации и классической кластеризации с использованием метода К-средних. Отдано предпочтение в пользу иерархической кластеризации. Получены непротиворечивые результаты с использованием обоих методов. Сделан вывод о важности географического фактора и однородности в экономическом развитии регионов, сгруппированных в кластеры. Москва выделяется по уровню развития цифровизации по сравнению с другими регионами, что подтверждают оба метода кластеризации. Использование методов машинного обучения позволило сделать вывод о большей значимости абсолютных показателей по сравнению с относительными для классификации регионов по уровню развития цифровых образовательных услуг населению.

Ключевые слова: методы машинного обучения; цифровизация; статистика; регрессия; иерархическая кластеризация; метод К-средних; дендрограмма; услуги

Для цитирования: Борисова Л.Р., Фридман М.Н. Анализ некоторых аспектов цифровизации в области оказания образовательных услуг населению методами статистического анализа и машинного обучения. *Цифровые решения и технологии искусственного интеллекта*. 2026;2(3):67-78. DOI: 10.26794/3030-7097-2026-2-3-67-78

ORIGINAL PAPER

Analysis of Certain Aspects of Digitalization in the Field of Providing Educational Services to the Population Using Statistical Analysis and Machine Learning Methods

L.R. Borisova, M.N. Fridman

Financial University under the Government of the Russian Federation, Moscow, Russian Federation

ABSTRACT

Digital transformation in all areas of life has accelerated significantly since COVID-19. Significant progress in the field of professional skills in digital technologies has been observed throughout the world and in Russia, in particular, in the last five years. It is a challenge and an opportunity when issues of integration and growth coincide with each other.



Digitalization in the field of public services is developing rapidly in our country. Russia needs educational systems appropriate to the digital age, as well as additional programs for the education and retraining of that part of the adult population that has long abandoned their studies. This is a two-pronged task, dealing with both basic digital skills and a modern understanding of literacy. The goal of developing digital educational services for the population is to create an accessible and effective learning system for all categories of citizens. The paper uses clustering and classification methods to group homogeneous objects based on a wide range of socio-economic and financial indicators that are important for studying the development of digitalization in the services provided to the population. The objective of the study is to identify, using objective methods of mathematical data analysis (without any use of subjective assessments), the presence or absence of a relationship between the socio-economic parameters of a region's development in the Russian Federation and paid educational services provided to the population. The regions are clearly divided into two large, almost equal groups based on this set of indicators. A comparative analysis of hierarchical clustering and classical clustering using the K-means method is presented. Hierarchical clustering is preferred over classical clustering. Consistent results were obtained using both methods. The importance of the geographical factor and homogeneity in the economic development of regions grouped into clusters was concluded. Moscow stands out for its digitalization development compared to other regions, as evidenced by both clustering methods.

Keywords: machine learning methods; digitalization; statistics; regression; hierarchical clustering; K-means method; dendrogram; services

For citation: Borisova L.R., Fridman M.N. Analysis of certain aspects of digitalization in the field of providing educational services to the population using statistical analysis and machine learning methods. *Digital Solutions and Artificial Intelligence Technologies*. 2026;2(3):67-78. DOI: 10.26794/3030-7097-2026-2-3-67-78

ВВЕДЕНИЕ

Стремительный технический и технологический процесс бросает вызов не только различным отраслям экономики, но также и сфере оказания услуг населению.

Пандемия, наряду с общими тенденциями технического прогресса, стимулировала развитие цифровизации в различных областях оказания услуг населению во всех развитых странах. Россия оказалась одной из самых передовых в применении цифровых технологий для оказания широкого спектра социальных услуг: взаимодействие с государственными органами, с медицинскими и образовательными учреждениями, услуги в области культуры, туризма и спорта, в банковской и финансовой сферах, в области социальной и психологической помощи, в организации досуга и даже поиска партнеров [1, 2].

Достаточно вспомнить, как практически без потерь вузы и школы перешли на онлайн-обучение в разгар эпидемии. В области медицинских услуг Россия оперативно перевела взаимоотношения пациента и медицинского учреждения на полную цифровизацию [3], в отличие, например, от Германии, где цифровые истории болезней начали вводить только как факультативное предложение в 2021 г., и в тот момент на него подписалось менее 1% населения страны¹. Сегодня в области взаимодействия с государственными органами редко кто обходится без сайта «Госуслуги». Правда, следует отметить, что некоторые исследования указывают на некоторый откат в сторону оффлайн-предпочтений после

окончания пандемии² и с причинами этого следует разобраться.

В чем заключается особая важность цифровизации в области сферы услуг для России? В первую очередь это необходимо для обеспечения доступности и качества широкого спектра услуг во всех районах громадной территории страны, вне зависимости от их удаленности от центра, уровня их экономического развития. Это необходимо также для формирования нового, современного рынка услуг, что приводит в конечном итоге к ускорению соответствующих процессов в экономике [4]. Оуществляется обратная связь, взаимодействие спроса и предоставления услуг, причем бизнес должен проводить широкий анализ и прогнозирование спроса на услуги, используя современный математический аппарат работы с большими данными.

В научной среде достаточно работ, указывающих на важность использования цифровизации для повышения качества образования, где приводится анализ динамики цифровизации общего образования в России и сделан вывод, что развитию цифровизации в деятельности школ будет способствовать разработка единых требований в данной области [5, 6]. Но пока регионы отличаются по уровню использования школьниками электронного обучения³.

² Щербakov P.A. Цифровизация государственных услуг. М.: ИСИЭЗ НИУ ВШЭ; 2025. URL: <https://issek.hse.ru/news/1039474722.html>; Попов Е.В. Цифровые навыки россиян. М.: ИСИЭЗ НИУ ВШЭ; 2025. URL: <https://issek.hse.ru/news/1038822258.html>

³ Отчет Счетной палаты РФ «Анализ реализации мер в сфере цифровизации образования». URL: <https://ach.gov.ru/upload/iblock/0d8/b8pg1tjsapfepfg6r0xvur3tk1mrn0el.pdf?ysclid=mt8j4vee1f579362498>

¹ URL: <https://ausnews.de/digital-healthcare/>



Как уже было отмечено, широкая географическая и экономическая дисперсность является серьезной проблемой в задаче унификации качества и количества предоставляемых населению услуг в России.

Особый интерес для авторов представляет развитие цифровизации образовательных услуг, целью которой является формирование доступной и эффективной системы обучения для всех категорий граждан. Актуальность этого вопроса в условиях России не вызывает сомнений. Благодаря цифровой среде обучение становится возможным независимо от географического положения и финансовых возможностей обучающихся, что особенно важно для удаленных регионов с ограниченной инфраструктурой и экономически слаборазвитых регионов.

Предметом исследования было выявление определенных закономерностей в распределении элементов цифровизации по ряду показателей с помощью формирования территориальных кластеров с целью выработки некоторых рекомендаций для оптимизации процесса цифровизации в области предоставления платных образовательных услуг населению.

МАТЕРИАЛЫ И МЕТОДЫ

База данных. В исследовании использованы открытые данные сайта Росстата⁴ для 85 регионов Российской Федерации (РФ). Данные для Архангельской и Тюменской областей, Чукотского, Ямало-Ненецкого и Ханты-Мансийского автономных округов рассмотрены отдельно. Данные для четырех новых регионов Российской Федерации не представлены в статистике Росстата, поэтому их в статье не учитывали.

Используемые для машинного обучения данные в основном относятся к периоду 2023–2024 гг. Исключения составили два показателя: глобальные информационные сети и системы электронного документооборота, используемые организациями. Эти данные относятся к 2019 г. Оба показателя являются безразмерными и представляют собой процент от всех обследованных организаций.

Всего был проанализирован 31 показатель: 30 зависимых показателей приведены в *табл. 1* и 1 независимый показатель (группирующий признак для методов машинного обучения) — уровень платных образовательных услуг населению за июль 2024 г. Последние 7 показателей также относятся к этому же периоду.

Методы анализа. Были использованы иерархическая кластеризация и метод К-средних

(K-means) — одни из наиболее распространенных подходов в кластерном анализе. Иерархическая кластеризация может быть предпочтительнее в ряде случаев, поскольку позволяет исследователю самостоятельно определить количество кластеров по итогам анализа (например, «разрезав» полученную дендрограмму на нужном уровне), тогда как для К-средних число кластеров необходимо задать до запуска алгоритма⁵. Дендрограмма при использовании иерархической кластеризации была построена с применением функции «hclust» из пакета «cluster» языка программирования R. Двумерные диаграммы, иллюстрирующие разбиение 85 регионов России на кластеры, были построены с использованием программного комплекса компании Azforus [7]. Для классификации 95 регионов РФ по показателю цифровизации образовательных услуг также был применен метод статистически взвешенных синдромов (СВС) [8]. Метод СВС основан на принятии коллективных решений по системам синдромов двумерных областей признаков пространства, в котором преобладают объекты одного из распознаваемых классов. Области задаются с помощью границ, которые находятся через оптимальные разбиения интервалов значений признаков. Подсчет взвешенных голосов проводится как по одномерным, так и по двумерным разбиениям. Подробно метод СВС описан в лекциях разработчика метода О.В. Сенько⁶ и в работе [9].

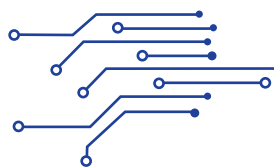
РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ И ИХ ОБСУЖДЕНИЕ

Результаты кластерного анализа. Методы кластерного анализа позволяют выделить в многомерном пространстве показателей (признаков) компактные группы близких друг другу объектов. Кластеры отличаются друг от друга по высоте (от англ. height), характеризующей степень различия регионов по исследуемым показателям. Данный подход позволяет выявить скрытые связи и закономерности, которые можно не заметить, если использовать только методы базовой статистики. Отметим, что в предыдущей работе по теме этот метод успешно применен для группировки стран Азиатско-Тихоокеанского региона по значениям инклюзивного экономического роста [10]. Там же подробно описано использование иерархической кластеризации.

⁵ James G., Witten D., Hastie T., Tibshirani R. An Introduction to Statistical Learning: With Applications in R. Textbook. Springer; 2021.

⁶ Сенько О.В. Математические основы теории прогнозирования. URL: http://www.machinelearning.ru/wiki/images/c/cd/MOTP13_6_bw.pdf?ysclid=mt8disakcn735931691

⁴ URL: <https://data.rosstat.org>



Показатели, используемые для исследования / Indicators Used for the Study

№	Показатели, используемые для исследования / Indicators Used for the Study
	ОРГАНИЗАЦИИ (% от общего числа обследованных) использовавшие
1	Технологии искусственного интеллекта
2	Интернет вещей
3	ПК
4	Серверы
5	Локальные вычислительные сети
6	Глобальные информационные сети
7	Облачные сервисы
8	Технологии сбора, обработки и анализа больших данных
9	Цифровые платформы
10	Мобильный и фиксированный интернет
11	Широкополосной интернет
12	Веб-сайт
13	Системы электронного документооборота
14	Электронный обмен между своими и внешними информационными системами документооборота
Удельный вес домашних хозяйств (% от всех домохозяйств)	
15	Имевших персональный компьютер
16	Имевших доступ к сети Интернет
17	Имевших доступ к широкополосному интернету
18	% населения, использующих интернет
19	% подключенных абонентских устройств мобильной связи на 1000 населения
20	Численность активных абонентов фиксированного широкополосного интернета (на 100 человек)
21	Численность активных абонентов мобильного широкополосного интернета (на 100 человек)
22	Ежемесячные среднедушевые денежные доходы, тыс. руб.
23	Ежемесячные среднедушевые потребительские расходы, тыс. руб.
Платные услуги населению за июль 2024 г. в рублях на чел.	
24	Телекоммуникационные
25	Физической культуры и спорта
26	Почтовой связи и курьеров
27	Фитнес-центров и спортивных клубов
28	Медицинские
29	Ветеринарные
30	Юридические

Источник / Source: составлено авторами по материалам Росстата / Compiled by the authors based on the materials of Rosstat.
URL: <https://rosstat.gov.ru/>



Методы кластерного анализа применены для оценки российских регионов по 31 показателю (30 показателей из табл. 1) и еще одному показателю, который в дальнейшем будет использоваться в качестве зависимой переменной при применении методов машинного обучения. 31-й показатель — это платные образовательные услуги населению за июль 2024 г. в рублях на человека. На рис. 1 представлена дендрограмма иерархической кластеризации.

Проанализируем дендрограмму рис. 1. Для удобства оценки номера регионов с их соответствующей расшифровкой приведены в табл. 2, в которую также записаны кластеры (10), четко разделяемые по высоте. Удобно для анализа разрезать дендрограмму горизонтальной линией на уровне $height = 10$. Тогда число кластеров должно равняться 10 или незначительно отличаться. Количество кластеров выбраны неслучайно. В работе [11] было проанализировано похожее число показателей (32), количество объектов для кластерного анализа — 45 и представлено подробное обоснование, почему опти-

мальное количество кластеров должно находиться в диапазоне 6–12. Как отмечено в работе [12], оптимальное количество кластеров определяется по дендрограмме. Поскольку в исследовании использовали готовую программу из библиотеки программ алгоритмического языка R, то не ставили задачу рассчитать индекс силуэта для оценки качества кластеризации [13].

Проанализируем разделение регионов на 10 кластеров (то же самое число выберем для возможности сравнения однотипных данных) с использованием самого распространенного метода кластеризации — K-средних, в котором число кластеров фиксируется исследователем до проведения кластеризации. Представим двумерную диаграмму из кластерного анализа на рис. 2, где четко прослеживается деление всех 85 регионов на 10 кластеров, отличающихся цветом и геометрическими фигурами (ромбы, квадраты, кружочки, звездочки, крестики) для визуализации кластеризации. Чтобы результаты двух различных методов (иерархической кластеризации и K-средних) можно было сравнить

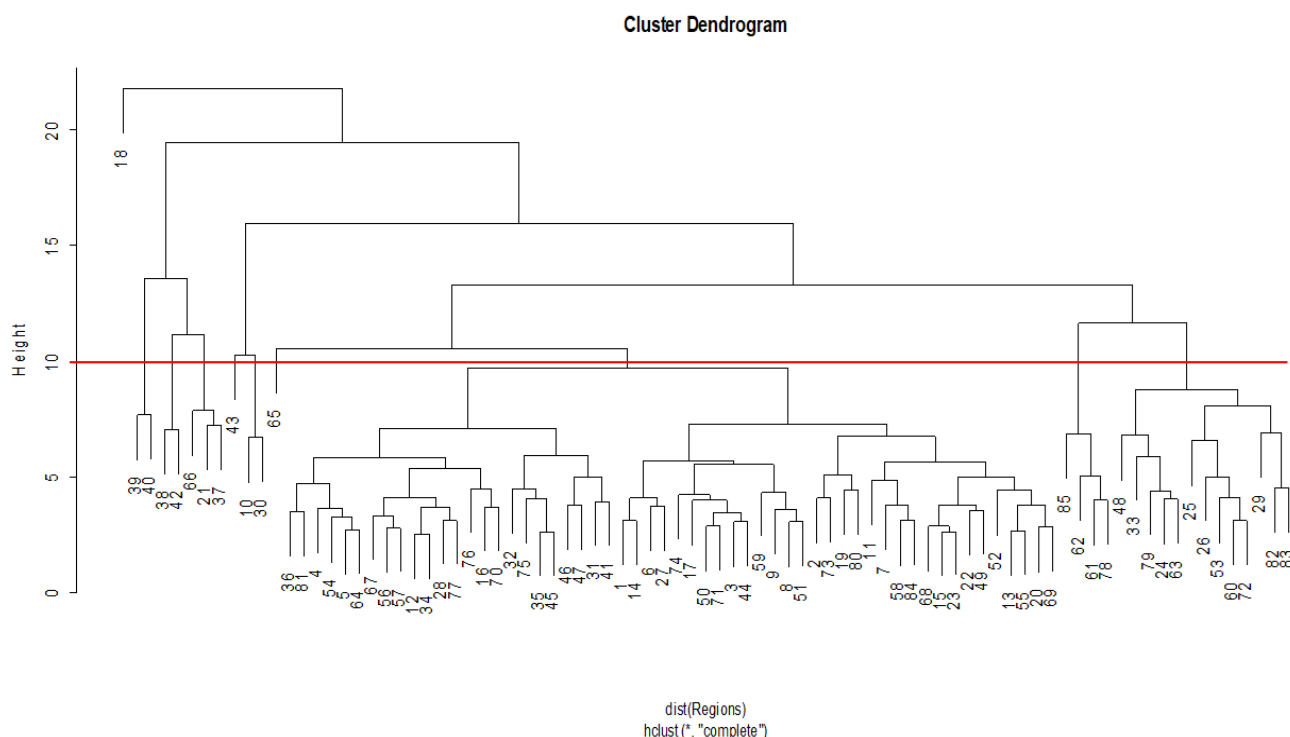
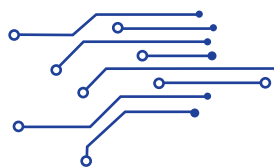


Рис. 1 / Fig. 1. Кластерная дендрограмма, соответствующая разделению РФ по показателям цифровизации услуг населению / Cluster Dendrogram Corresponding to the Division of the Russian Federation by Indicators of Digitalization of Services to the Population

Источник / Source: составлено авторами / Compiled by the authors.

Примечание / Note: Height — высота (ось Y), характеризующая уровень отличия наблюдений (кластеров); ось X — рисунок дендрограммы с условными номерами регионов, расшифрованными в табл. 2 / Height — height (Y-axis), which characterizes the level of difference between observations (clusters); X — axis is a dendrogram drawing with conditional numbers of regions, deciphered in Table 2.

**Нумерация и состав кластеров из регионов по методу иерархической классификации /
Numbering and Composition of Clusters from the Regions Using the Hierarchical Classification Method**

Номер / Number		Название региона / Region name
Кластер / Cluster	Регион / Region	
1	18	Москва
2	39, 40	Республики Ингушетия и Кабардино-Балкария
3	38, 42	Республики Дагестан и Северная Осетия (Алания)
4	66, 21, 37	Республика Тыва, Ненецкий автономный округ, Севастополь
5	43, 10, 30	Чеченская Республика, Московская обл., Республика Адыгея
6	65	Республика Алтай
7	36, 81, 4, 54, 5, 64, 67, 56, 57, 12, 34, 28, 77, 76, 16, 70, 32, 75, 35, 45, 46, 47, 31, 41	Ростовская, Амурская, Воронежская, Оренбургская, Ивановская, Челябинская обл., Республика Хакасия, Самарская, Саратовская, Рязанская, Астраханская, Псковская обл.; Забайкальский край, Республика Саха (Якутия), Тульская, Иркутская обл.; Республики Крым и Бурятия, Волгоградская обл., Республики Башкортостан, Марий Эл, Мордовия, Калмыкия, Карачаево-Черкесская
8	1, 14, 6, 27, 74, 17, 50, 71, 3, 44, 59, 9, 8, 51, 2, 73, 19, 80, 11, 7, 58, 84, 68, 15, 23, 22, 49, 52, 13, 55, 20, 69	Белгородская, Тамбовская, Калужская, Новгородская, Томская, Ярославская обл., Республика Чувашия; Кемеровская, Владимирская обл., Ставропольский край, Курганская, Липецкая, Курская обл., Пермский край, Брянская, Омская обл.; Республика Карелия, Хабаровский край, Орловская, Костромская, Ульяновская обл., Еврейская автономная обл., Алтайский край, Тверская, Волгоградская, Архангельская обл., Республика Удмуртия, Кировская, Смоленская, Пензенская обл., Республика Коми, Красноярский край
9	85, 62, 61, 78, 48	Чукотский, Ямало-Ненецкий, Ханты-Мансийский автономные округа, Камчатский край, Татарстан
10	33, 79, 24, 63, 25, 26, 53, 60, 72, 29, 82, 83	Краснодарский, Приморский края, Калининградская, Тюменская, Ленинградская, Мурманская, Нижегородская, Свердловская, Новосибирская обл., Санкт-Петербург, Магаданская и Сахалинская обл.

Источник / Source: составлено авторами / Compiled by the authors.

между собой, регионы, вошедшие в 10 кластеров с использованием метода К-средних, представим в табл. 3.

Результаты машинного обучения. Выявление скрытых статистически достоверных связей возможно с применением методов классификации. В качестве целевой функции использован показатель объема платных образовательных услуг населению за июль 2024 г. в рублях на человека.

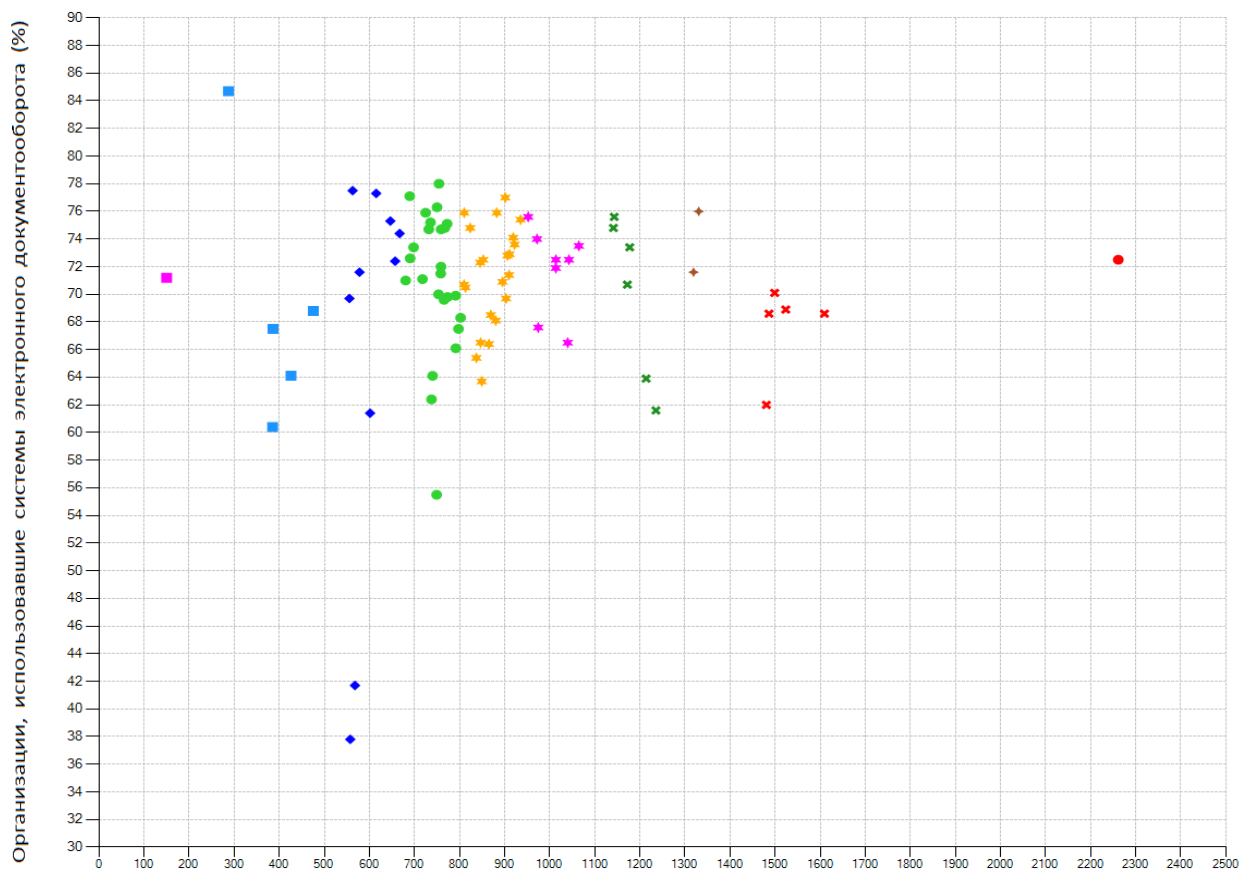
Для деления 85 регионов на две группы было вычислено медианное значение группирующего показателя, составившее 295,88 руб. в месяц на человека. Количество регионов, в которых этот показатель оказался ниже медианного уровня, равно 44, а выше медианного уровня — 41.

Был применен метод распознавания — статистически взвешенные синдромы (СВС), ранее описанный в работе [8]. Зависимые показатели ис-

пользовались те же, что и для кластерного анализа. Их количество составляет 30, и они представлены в табл. 1.

Результаты распознавания характеризуются наличием ошибок при использовании скользящего контроля. В процессе работы программы случайным образом были отобраны 18 регионов из 85 анализируемых, причем количество регионов в каждой группе составило по 9. Значения параметров классификации представлены в табл. 4.

Отметим, что Accuracy score — это классический для машинного обучения метод оценки точности модели, подразумевающий процент совпадения всех предсказанных и реальных результатов. ROC AUC score представляет из себя площадь под графиком рабочей характеристики приемника. График позволяет оценить качество бинарной клас-



Платные телекоммуникационные услуги населению за июль 2024 г. руб. на чел.

Рис. 2 / Fig. 2. Зависимость удельного веса организаций, использующих электронный документооборот (%) от ежемесячных платных телекоммуникационных услуг населению (в руб. за июль 2024 г.) / Dependence of the Share of Organizations Using Electronic Document Management (%) on Monthly paid Telecommunications Services for the Population (in rubles for July 2024)

Источник / Source: составлено авторами / Compiled by the authors.

сификации, отображает соотношение между долей объектов от общего количества носителей признака, верно классифицированных как несущих признак, и долей объектов от общего количества объектов, не несущих признака, ошибочно классифицированных как несущих признак. Точность линейной классификации можно основывать на 4 типах оценки результата:

TP — (True Positive) истинно-положительное решение;

TN — (True Negative) истинно-отрицательное решение;

FP — (False Positive) ложно-положительное решение;

FN — (False Negative) ложно-отрицательное решение.

Точность рассчитывается по формуле $\frac{TP}{TP + FP}$, полнота — $\frac{TP}{TP + FN}$, чувствительность — $\frac{TP}{TP + FN}$, специфичность — $\frac{TN}{TP + FP}$.

Было зарегистрировано всего четыре ошибки, а именно: три региона, относящиеся к 1-й группе, — Калужская и Иркутская области, Республика Саха (Якутия) (со значением группирующего показателя ниже медианного значения), при использовании скользящего контроля были отнесены ко 2-й группе, а Воронежская область, наоборот, была распознана как регион с показателем платных ежемесячных услуг населению ниже медианы, хотя в июле 2024 г. этот показатель был выше медианы, причем существенно — почти в 2 раза.

Распознавание, выполненное с использованием 30 зависимых показателей, оказалось точнее по основному показателю качества — площади под ROC-кривой, — чем в нашем предыдущем исследовании, когда использовали шесть показателей [9]. Но в предыдущей работе и цель была другая: исследовали влияние образовательных, социальных, информационных, медицинских услуг на показатель «расходы на информационные услуги

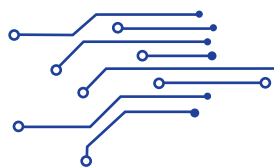


Таблица 3 / Table 3

**Нумерация и состав кластеров из регионов Российской Федерации по методу К-средних /
Numbering and Composition of Clusters from the Russian Federation Regions Using the K-means
Method**

Номер / Number		Название региона / Region name
Кластер / Cluster	Регион / Region	
1	18	Москва (красный кружок на рис. 2)
2	39, 66, 32, 21, 10	Республики Ингушетия, Тыва, Крым, Ненецкий автономный округ, Московская обл. (голубые квадраты)
3	31, 43, 30, 41, 38, 65, 46, 49, 50, 47	Республики Калмыкия, Чеченская, Адыгея, Карачаево-Черкессия, Дагестан, Алтай, Марий Эл, Удмуртия, Чувашия, Мордовия (синие ромбы)
4	84, 5, 14, 44, 58, 48, 34, 28, 40, 57, 42, 1, 8, 27, 6, 23, 17, 15, 4, 52, 73, 37, 56, 16	Еврейская автономная обл., Ивановская и Тамбовская обл., Ставропольский край, Ульяновская обл., Татарстан, Псковская обл., Кабардино-Балкария, Саратовская обл., Северная Осетия (Алания), Белгородская, Курская, Новгородская, Калужская, Ярославская, Тверская, Воронежская, Кировская, Омская обл., Севастополь, Самарская, Тульская обл. (зеленые кружочки)
5	75, 64, 11, 3, 13, 67, 54, 70, 74, 53, 24, 35, 69, 36, 7, 54, 71, 12, 2, 20, 51, 45, 68	Республика Бурятия, Челябинская, Орловская, Владимирская, Смоленская обл., Республика Хакасия, Пензенская, Томская, Иркутская, Нижегородская, Калининградская, Волгоградская обл., Красноярский край, Ростовская, Костромская, Оренбургская, Кемеровская, Рязанская, Брянская обл., Республика Коми, Пермский край, Башкирия, Алтайский край (оранжевые звездочки)
6	81, 63, 79, 19, 72, 59, 22, 9	Амурская, Тюменская обл., Приморский край, Карелия, Новосибирская, Курганская, Архангельская, Липецкая обл. (лиловые звездочки)
7	33, 76, 26, 83, 77, 80	Краснодарский край, Республика Саха (Якутия), Мурманская, Сахалинская обл., Забайкальский и Хабаровский края
8	82, 60	Магаданская и Свердловская обл. (коричневые крестики)
9	85, 62, 61, 78, 29	Чукотский, Ямало-Ненецкий, Ханты-Мансийский автономные округа, Камчатский край, Санкт-Петербург (красные крестики)
10	26	Ленинградская обл. (лиловый квадрат)

Источник / Source: составлено авторами / Complied by the authors.

Таблица 4 / Table 4

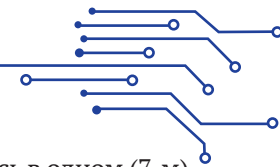
**Результаты классификации по методу статистически взвешенных синдромов /
The Results of Classification by the Method of Statistically Weighted Syndromes**

Правильность / Accuracy	Точность / Precision	Чувствительность / Sensitivity	Специфичность / Specificity	ROCAUC
0,78	0,86	0,67	0,89	0,90

Источник / Source: составлено авторами по результатам исследования / Complied by the authors based on the results of the study.

и сервисы в мае 2024 года». Неудивительно, что показатель чувствительности (табл. 4) для нашего исследования равен 0,67, в отличие от 0,78 для предыдущего исследования, в котором все показатели были однотипными и относились к одному временному периоду.

Как уже было отмечено выше, на сайте Росстата не все данные по удельному весу организаций, характеризующихся тем или иным уровнем цифровизации, были за один год (2023). За этот год были собраны данные на сайте Росстата в сборнике, опубликованном на сайте в 2024 г. Поскольку распозна-



вание только по 23 показателям было недостаточно качественным (в смысле показателей качества распознавания, представленных в табл. 2; данные не приводятся), то добавили информационные показатели ежемесячных трат населения за один выделенный месяц 2024 г.

Интересно посмотреть, изменятся ли существенно результаты исследования, если из анализа исключить влияние двух показателей за 2019 г. (организации, использовавшие глобальные информационные сети, и учреждения, применявшие системы электронного документооборота) на результирующий признак — уровень платных образовательных услуг населению за июль 2024 г. Показатель РОК АУК при этом не изменился. Его значение — 0,90123, что свидетельствует в пользу использования методов машинного обучения для классификации. Ошибки распознавания при использовании контрольной выборки в 20% также остались на прежнем уровне: алгоритм ошибочно отнес Калужскую, Иркутскую области и Якутию ко второму классу, а Воронежскую область — к первому классу.

Остальные результаты не приводятся, так как цель исследования заключалась в поиске наиболее информативных показателей, влияющих на платные ежемесячные образовательные услуги, — на примере зависимости 31 показателя от выбранного зависимого показателя. Месяц июль для группирующего показателя тоже был выбран неслучайно, так как к этому времени уже закончился образовательный процесс в школах, учреждениях СПО и высшего образования. Представлял интерес анализ трат населения на образовательные услуги, так как эти траты в основном были связаны с расходами на повышение квалификации в информационной сфере, цифровизацию разных услуг и использование искусственного интеллекта.

Полученное значение показателя ROC AUC, равное 0,9 (близкое к 1), свидетельствует в пользу использования методов машинного обучения для классификации регионов по показателям, важным для развития цифровизации во всех сферах человеческой жизни.

Отметим, что выполненный в исследовании кластерный анализ продемонстрировал интересные результаты: регионы, отличающиеся по основным экономическим показателям, оказались в одном кластере. Например, Чеченская Республика, Республика Адыгея и Московская область — в 5-м кластере. Санкт-Петербург оказался в одном кластере с самой западной областью РФ — Калининградской — и самыми восточными областями — Магаданской и Сахалинской (10-й кластер). При этом близкие по уровню развития республики Поволжья — Марий Эл,

Мордовия, Башкортостан — оказались в одном (7-м) кластере. Эти результаты свидетельствуют о том, что географический фактор очень важен для развития цифровизации в разных сферах человеческой жизни. Результаты, полученные при помощи метода К-средних (табл. 3), не противоречат этому выводу.

ВЫВОДЫ

Нельзя забывать, что интеллектуальный анализ данных требует больших затрат энергии. Такие факторы, как географический и энергетический, находятся в поле зрения не только властей такой протяженной страны, как Российская Федерация, но и других стран, в частности таких, у которых есть сходные по климату географические зоны. В связи с этим представляет интерес статья [14], в которой было показано, что децентрализация, приватизация и цифровизация — это движущие силы на местном и национальном уровнях. Они меняют распределение инвестиций и расстановку приоритетов.

Полученные в работе результаты свидетельствуют, что регионы по услугам цифровизации не сильно отличаются друг от друга, на что указывает большое количество кластеров. Это подчеркивает, что в настоящее время практически во всех регионах Российской Федерации доступны услуги интернета, в том числе в сфере образования. В статье [15] описано использование иерархической кластеризации для поиска и выявления аномальных наблюдений, что важно для возможности управления развитием социально-экономических систем. Работа интересна и тем, что в ней применяются методы иерархической и классической классификации.

Обсуждая результаты кластерного анализа, полученные в настоящей работе, можно сделать вывод, что методы иерархической кластеризации и метод К-средних не противоречат друг другу. Разумеется, если представить диаграмму для пары других признаков, используемых для кластерного анализа по методу К-средних, то изменится и наполнение кластеров (табл. 3), поэтому иерархическая кластеризация предпочтительнее для анализа с большим числом показателей.

Отметим, что нормализация показателей в кластерном методе проводится.

Методы машинного обучения позволяют выявить наиболее информативные признаки, используемые не для кластеризации, а для классификации изучаемых объектов, в частности регионов РФ по группирующему признаку — затраты на образовательные услуги. В работе получено распознавание с высоким качеством (ROC AUC = 0,9), что свидетельствует в пользу использования данных методов для количественного анализа цифровизации в сфере



услуг населению. Самыми информативными признаками стали абсолютные количественные, а не удельные показатели.

При проверке разделения всех регионов на две группы использование теста Манна–Уитни–Уилкоксона продемонстрировало наименьшие значения основного статистического показателя p -value (0,005 и меньше) для следующих показателей: платные телекоммуникационные услуги, услуги физической культуры и спорта, фитнес-центров и спортивных клубов, медицинские и ветеринарные услуги населению за июль 2024 г. Также статистически значимым оказался единственный удельный показатель — процент организаций, использующих серверы.

Интересно отметить, что ежемесячный среднедушевой доход и ежемесячные среднедушевые расходы не вошли в число наиболее значимых показателей, что свидетельствует в пользу того, что человек — объект безразмерной бесконечности, и цифровизация не в состоянии с вероятностью, близкой к 1, выявить набор самых важных показателей, так как за рамками исследования всегда остается много качественных индикаторов, которые невозможно формализовать.

Наибольшее значение p -value (0,7) было для показателя — процент организаций, использующих технологии искусственного интеллекта, что говорит о недостаточном внедрении методов машинного обучения, основанных на регрессионных моделях. Разумеется, в будущем данные методы найдут применение при анализе массивов данных, не превышающих 100 объектов. Так же широко распространились нейросетевые методы при анализе выборок с числом элементов, превышающим миллион объектов, для качества распознавания, близкого к 100% (ROC AUC $\approx$ 1).

Полученные результаты свидетельствуют о необходимости комплексного подхода к цифровизации образовательных услуг, учитывающего как количественные, так и качественные показатели развития цифровой инфраструктуры.

Проведенный в статье анализ цифровизации услуг на территории Российской Федерации, включая образовательные услуги населению, подтвердил, что цифровизация в целом оказывает положитель-

ное влияние на различные сферы жизни, а цифровизация образовательных услуг обеспечивает доступность обучения и его качество для всех категорий граждан, независимо от их географического положения и финансовых возможностей. Это особенно важно для удаленных и экономически слабо развитых регионов. Исследование показало, что регионы с высоким уровнем цифровизации образовательных услуг характеризуются широким применением электронного документооборота; значительной долей организаций, использующих цифровые платформы; существенным уровнем доступа населения к интернету.

Для оптимизации процесса внедрения цифровых решений в сфере образовательных услуг во всех регионах страны необходимо нарастить и форсировать развитие инфраструктуры широкополосного доступа в интернет в отдаленных регионах; поощрять освоение цифровых технологий в образовательных учреждениях; разработать программы профессионального совершенствования для педагогов в области цифровых технологий; задействовать методы машинного обучения для анализа и прогнозирования спроса на образовательные услуги.

Метод статистически взвешенных синдромов не требует проверки используемых данных на нормальный закон распределения, а также проверки мультиколлинеарности и исключения из дальнейшего анализа сильно коррелируемых признаков, поэтому может использоваться при анализе безразмерных и размерных величин при цифровизации образовательных услуг в дальнейшем.

Интересно, что выполненный на тестовой выборке контроль полученного решающего правила (модели) показал неустойчивое положение с точки зрения влияния всего комплекса анализируемых относительных и абсолютных признаков на показатель платных услуг в области образования в четырех совершенно разных с точки зрения географии и экономического развития регионах — Якутии, Калужской, Иркутской и Воронежской областях.

Полученный результат свидетельствует в пользу дальнейшего использования методов машинного обучения в области цифровизации.

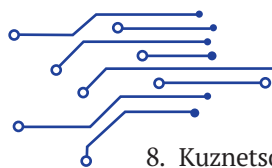
СПИСОК ИСТОЧНИКОВ

1. Сафонов С.А., Зандер Е.В. Оценка цифровой трансформации социальной сферы и региональные особенности цифровизации социальных услуг. *Цифровая экономика*. 2022;(4):53–62. URL: <https://doi.org/10.34706/DE-2022-04-07>
2. Кучумов И.В. Теоретические основы методов кластеризации данных в интеллектуальном анализе. *Цифровая экономика*. 2024;(1):69–77. URL: <https://doi.org/10.34706/DE-2024-01-08>.
3. Афан А.И., Полозова Д.В., Гордеева А.А. Цифровая трансформация государственной системы здравоохранения России: возможности и противоречия. *Цифровое право*. 2021;2(4):20–39. URL: <https://doi.org/10.38044/2686-9136-2021-2-4-20-39>

4. Нырков А.С. Влияние цифровизации на доступность и качество социальных услуг в Российской Федерации. *Молодой ученый*. 2025;20(571):522-524. URL: <https://www.elibrary.ru/xsfosr>
5. Назаров В.Л., Жердев Д.В., Буйначева А.В. Актуальные проблемы цифровой трансформации среднего образования. *Образование и наука*. 2023;25(4):109-166. URL: <https://doi.org/10.17853/1994-5639-2023-4-109-166>
6. Тищенко А.С., Токарева Г.С. Цифровая трансформация общего образования: анализ изменений. *Вестник Российского университета дружбы народов. Серия: Информатизация образования*. 2026;23(1):105-119. URL: <http://doi.org/10.22363/2312-8631-2026-23-1-105-119>
7. Сенько О.В., Кузнецова А.В., Воронин Е.М., Кравцова О.А., Борисова Л.Р., Кирилюк И.Л., и др. Методы интеллектуального анализа данных в исследованиях эпидемии COVID-19. *Журнал Белорусского государственного университета. Математика. Информатика*. 2022;1:83-96. URL: <https://doi.org/10.33581/2520-6508-2022-1-83-96>
8. Кузнецова Ю.О., Борисова Л.Р., Кузнецова А.В., Сенько О.В. Прозрачный интерфейс для прогноза в машинном обучении. В кн.: *Аналитика и управление данными в областях с интенсивным использованием данных*. М.: Федеральный исследовательский центр «Информатика и управление» Российской академии наук; 2017:493-495. URL: <https://elibrary.ru/ougrxt>
9. Кузнецова А.В., Борисова Л.Р., Демина И.А. Методы машинного обучения для прогнозирования течения заболевания на примере развития тяжелой степени пневмонии. *Цифровые решения и технологии искусственного интеллекта*. 2025;1(1):6-19. URL: <https://www.elibrary.ru/rzxxxy>
10. Матвиевский С.С., Борисова Л.Р. Кластеризация стран Азиатско-Тихоокеанского региона по значениям инклюзивного экономического роста. *Вестник университета*. 2023;1:112-121. URL: <https://doi.org/10.26425/1816-4277-2024-1-112-121>
11. Борисова Л.Р. Сравнительный анализ регионов Российской Федерации методами машинного обучения по набору показателей электронных услуг и сервисов. *Цифровая социология*. 2024;7(40):33-43. URL: <https://doi.org/10.26425/2658-347X-2024-7-4-33-43>
12. Репина С.И. Проверка качества кластеров с помощью силуэтного анализа. *Экономика и социум*. 2024;9(124):958-975. URL: <https://doi.org/10.5281/zenodo.13918451>
13. Rousseeuw P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*. 1987;20(1):53-65. URL: [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
14. Romanello M. et al. The 2022 report of the Lancet Countdown on health and climate change: health at the mercy of fossil fuels. *The Lancet 400 (suppl)*. 2022. URL: [https://doi.org/10.1016/S0140-6736\(22\)01540-9](https://doi.org/10.1016/S0140-6736(22)01540-9)
15. Кисляков А.Н., Поляков С.В. Иерархические методы кластеризации в задаче поиска аномальных наблюдений на основе групп с нарушенной симметрией. *Управленческое консультирование*. 2020;5:116-127. URL: <https://doi.org/10.22394/1726-1139-2020-5-116-127>

REFERENCES

1. Safonov S.A., Zander E.V. Assessment of digital transformation of the social sphere and regional features of digitalization of social services. *The Digital Economy*. 2022;(4):53-62. (In Russ.). URL: <https://doi.org/10.34706/DE-2022-04-07>
2. Kuchumov I.V. Theoretical foundations of data clustering methods in intellectual analysis. *The Digital Economy*. 2024;(1):69-77. (In Russ.). URL: <https://doi.org/10.34706/DE-2024-01-08>
3. Afyan A.I., Polozova D.V., Gordeeva A.A. Digital transformation of the Russian state healthcare system: opportunities and contradictions. *Digital Law*. 2021;2(4):20-39. (In Russ.). URL: <https://doi.org/10.38044/2686-9136-2021-2-4-20-39>
4. Nyrkov A.S. The impact of digitalization on the availability and quality of social services in the Russian Federation. *A Young Scientist*. 2025;20(571):522-524. (In Russ.). URL: <https://www.elibrary.ru/xsfosr>
5. Nazarov V.L., Zherdev D.V., Buinacheva A.V. Current problems of digital transformation in secondary education. *The Education and science journal*. 2023;25(4):109-166. (In Russ.) URL: <https://doi.org/10.17853/1994-5639-2023-4-109-166>
6. Tishchenko A.S., Tokareva G.S. Digital Transformation of General Education: Analysis of Changes. *RUDN Journal of Informatization in Education*. 2026;23(1):105-119. (In Russ.). URL: <http://doi.org/10.22363/2312-8631-2026-23-1-105-119>
7. Senko O.V., Kuznetsova A.V., Voronin E.M., Kravtsova O.A., Borisova L.R., Kirilyuk I.L., et al. Methods of data mining in research on the COVID-19 epidemic. *Journal of the Belarusian State University. Mathematics. Computer science*. 2022;1:83-96. (In Russ.). URL: <https://doi.org/10.33581/2520-6508-2022-1-83-96>



8. Kuznetsova Yu.O., Borisova L.R., Kuznetsova A.V., Senko O.V. Transparent interface for forecasting in machine learning. In: Data Analytics and Management in Data-intensive Areas. Moscow: Federal Research Center "Informatics and Management" of the Russian Academy of Sciences; 2017:493-495. (In Russ.). URL: <https://elibrary.ru/ougrxt>
9. Kuznetsova A.V., Borisova L.R., Demina I.A. Machine learning methods for predicting the course of the disease using the example of the development of severe pneumonia. *Digital Solutions and Artificial Intelligence Technologies*. 2025;1(1):6-19. (In Russ.). URL: <https://www.elibrary.ru/rzxxxyo>
10. Matviyevsky S.S., Borisova L.R. Clusterization of the countries of the Asia-Pacific region according to the values of inclusive economic growth. *Bulletin of the University*. 2023;1:112-121. (In Russ.). URL: <https://doi.org/10.26425/1816-4277-2024-1-112-121>
11. Borisova L.R. Comparative analysis of the Russian regions using machine learning methods for a set of indicators of electronic services. *Digital Sociology*. 2024;7(40):33-43. (In Russ.). URL: <https://doi.org/10.26425/2658-347X-2024-7-4-33-43>
12. Repina S.I. Quality verification of clusters using silhouette analysis. *Economics and Society*. 2024;9(124):958-975. (In Russ.). URL: <https://doi.org/10.5281/zenodo.13918451>
13. Rousseeuw P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*. 1987;20(1):53-65. (In Russ.). URL: [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
14. Romanello M. et al. The 2022 report of the Lancet Countdown on health and climate change: health at the mercy of fossil fuels. *The Lancet 400 (suppl)*. 2022. URL: [https://doi.org/10.1016/S0140-6736\(22\)01540-9](https://doi.org/10.1016/S0140-6736(22)01540-9)
15. Kislyakov A.N., Polyakov S.V. Hierarchical clustering methods in the problem of searching for anomalous observations based on groups with broken symmetry. *Management Consulting*. 2020;5:116-127. (In Russ.). URL: <https://doi.org/10.22394/1726-1139-2020-5-116-127>

ИНФОРМАЦИЯ ОБ АВТОРАХ / ABOUT THE AUTHORS

Людмила Робертовна Борисова — кандидат физико-математических наук, доцент кафедры математики и анализа данных факультета информационных технологий и анализа больших данных, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

Lyudmila R. Borisova — Cand. Sci. (Phys. and Math.) Assoc. Prof., Department of Mathematics and Data Analysis, Faculty of Information Technology and Big Data Analysis, Financial University under the Government of the Russian Federation, Moscow, Russian Federation
<http://orcid.org/0000-0002-5757-0341>

Автор для корреспонденции / Corresponding author:
lrborisova@fa.ru

Мира Нисоновна Фридман — доцент кафедры математики и анализа данных факультета информационных технологий и анализа больших данных, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

Mira N. Fridman — Assoc. Prof., Department of Mathematics and Data Analysis, Faculty of Information Technology and Big Data Analysis, Financial University under the Government of the Russian Federation, Moscow, Russian Federation
<http://orcid.org/0000-0001-9314-5986>
mnfridman@fa.ru

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Статья поступила 16.02.2026; после рецензирования 02.07.2026; принята к публикации 31.07.2026.

Авторы прочитали и одобрили окончательный вариант рукописи.

The article was submitted on 16.02.2026; revised on 02.07.2026 and accepted for publication on 31.07.2026.

The authors read and approved the final version of the manuscript.