

# МОЛОДОЙ УЧЁНЫЙ

ISSN 2072-0297

МЕЖДУНАРОДНЫЙ НАУЧНЫЙ ЖУРНАЛ



16+

16 2026  
ЧАСТЬ I

# Молодой ученый

## Международный научный журнал

### № 16 (619) / 2026

Издается с декабря 2008 г.

Выходит еженедельно

*Главный редактор:* Ахметов Ильдар Геннадьевич, кандидат технических наук

*Редакционная коллегия:*

Жураев Хусниддин Олтинбоевич, доктор педагогических наук (Узбекистан)  
Иванова Юлия Валентиновна, доктор философских наук  
Каленский Александр Васильевич, доктор физико-математических наук  
Кошербаева Айгерим Нуралиевна, доктор педагогических наук, профессор (Казахстан)  
Куташов Вячеслав Анатольевич, доктор медицинских наук  
Лактионов Константин Станиславович, доктор биологических наук  
Сараева Надежда Михайловна, доктор психологических наук  
Абдрасилов Турганбай Курманбаевич, доктор философии (PhD) по философским наукам (Казахстан)  
Авдеюк Оксана Алексеевна, кандидат технических наук  
Айдаров Оразхан Турсункожаевич, кандидат географических наук (Казахстан)  
Алиева Тарана Ибрагим кызы, кандидат химических наук (Азербайджан)  
Ахметова Валерия Валерьевна, кандидат медицинских наук  
Бердиев Эргаш Абдуллаевич, кандидат медицинских наук (Узбекистан)  
Брезгин Вячеслав Сергеевич, кандидат экономических наук  
Данилов Олег Евгеньевич, кандидат педагогических наук  
Дёмин Александр Викторович, кандидат биологических наук  
Дядюн Кристина Владимировна, кандидат юридических наук  
Желнова Кристина Владимировна, кандидат экономических наук  
Жуйкова Тамара Павловна, кандидат педагогических наук  
Игнатова Мария Александровна, кандидат искусствоведения  
Искаков Руслан Маратбекович, кандидат технических наук (Казахстан)  
Калдыбай Кайнар Калдыбайулы, доктор философии (PhD) по философским наукам (Казахстан)  
Кенесов Асхат Алмасович, кандидат политических наук  
Коварда Владимир Васильевич, кандидат физико-математических наук  
Комогорцев Максим Геннадьевич, кандидат технических наук  
Котляров Алексей Васильевич, кандидат геолого-минералогических наук  
Кузьмина Виолетта Михайловна, кандидат исторических наук, кандидат психологических наук  
Курпаяниди Константин Иванович, доктор философии (PhD) по экономическим наукам (Узбекистан)  
Кучерявенко Светлана Алексеевна, кандидат экономических наук  
Лескова Екатерина Викторовна, кандидат физико-математических наук  
Макеева Ирина Александровна, кандидат педагогических наук  
Матвиенко Евгений Владимирович, кандидат биологических наук  
Матроскина Татьяна Викторовна, кандидат экономических наук  
Матусевич Марина Степановна, кандидат педагогических наук  
Мусаева Ума Алиевна, кандидат технических наук  
Насимов Мурат Орленбаевич, кандидат политических наук (Казахстан)  
Паридинова Ботагоз Жаппаровна, магистр философии (Казахстан)  
Прончев Геннадий Борисович, кандидат физико-математических наук  
Рахмонов Азизхон Боситхонович, доктор педагогических наук (Узбекистан)  
Семахин Андрей Михайлович, кандидат технических наук  
Сенцов Аркадий Эдуардович, кандидат политических наук  
Сенюшкин Николай Сергеевич, кандидат технических наук  
Султанова Дилшода Намозовна, доктор архитектурных наук (Узбекистан)  
Титова Елена Ивановна, кандидат педагогических наук  
Ткаченко Ирина Георгиевна, кандидат филологических наук  
Федорова Мария Сергеевна, кандидат архитектуры  
Фозилов Садриддин Файзуллаевич, кандидат химических наук (Узбекистан)  
Яхина Асия Сергеевна, кандидат технических наук  
Ячинова Светлана Николаевна, кандидат педагогических наук

*Международный редакционный совет:*

Айрян Заруи Геворковна, кандидат филологических наук, доцент (Армения)  
Арошидзе Паата Леонидович, доктор экономических наук, ассоциированный профессор (Грузия)  
Атаев Загир Вагитович, кандидат географических наук, профессор (Россия)  
Ахмеденов Кажмурат Максutowич, кандидат географических наук, ассоциированный профессор (Казахстан)  
Бидова Бэла Бертовна, доктор юридических наук, доцент (Россия)  
Борисов Вячеслав Викторович, доктор педагогических наук, профессор (Украина)  
Буриев Хасан Чутбаевич, доктор биологических наук, профессор (Узбекистан)  
Велковска Гена Цветкова, доктор экономических наук, доцент (Болгария)  
Гайич Тамара, доктор экономических наук (Сербия)  
Данатаров Агахан, кандидат технических наук (Туркменистан)  
Данилов Александр Максимович, доктор технических наук, профессор (Россия)  
Демидов Алексей Александрович, доктор медицинских наук, профессор (Россия)  
Досманбетов Динар Бакбергенович, доктор философии (PhD), проректор по развитию и экономическим вопросам (Казахстан)  
Ешиев Абдыракман Молдоалиевич, доктор медицинских наук, доцент, зав. отделением (Кыргызстан)  
Жолдошев Сапарбай Тезекбаевич, доктор медицинских наук, профессор (Кыргызстан)  
Игисинов Нурбек Сагинбекович, доктор медицинских наук, профессор (Казахстан)  
Кадыров Кутлуг-Бек Бекмурадович, доктор педагогических наук, и.о. профессора, декан (Узбекистан)  
Каленский Александр Васильевич, доктор физико-математических наук, профессор (Россия)  
Козырева Ольга Анатольевна, кандидат педагогических наук, доцент (Россия)  
Колпак Евгений Петрович, доктор физико-математических наук, профессор (Россия)  
Кошербаева Айгерим Нуралиевна, доктор педагогических наук, профессор (Казахстан)  
Курпаяниди Константин Иванович, доктор философии (PhD) по экономическим наукам (Узбекистан)  
Куташов Вячеслав Анатольевич, доктор медицинских наук, профессор (Россия)  
Кыят Эмине Лейла, доктор экономических наук (Турция)  
Лю Цзюань, доктор филологических наук, профессор (Китай)  
Малес Людмила Владимировна, доктор социологических наук, доцент (Украина)  
Нагервадзе Марина Алиевна, доктор биологических наук, профессор (Грузия)  
Нурмамедли Фазиль Алигусейн оглы, кандидат геолого-минералогических наук (Азербайджан)  
Прокопьев Николай Яковлевич, доктор медицинских наук, профессор (Россия)  
Прокофьева Марина Анатольевна, кандидат педагогических наук, доцент (Казахстан)  
Рахматуллин Рафаэль Юсупович, доктор философских наук, профессор (Россия)  
Ребезов Максим Борисович, доктор сельскохозяйственных наук, профессор (Россия)  
Сорока Юлия Георгиевна, доктор социологических наук, доцент (Украина)  
Султанова Дилшода Намозовна, доктор архитектурных наук (Узбекистан)  
Узаков Гулом Норбоевич, доктор технических наук, доцент (Узбекистан)  
Федорова Мария Сергеевна, кандидат архитектуры (Россия)  
Хоналиев Назарали Хоналиевич, доктор экономических наук, старший научный сотрудник (Таджикистан)  
Хоссейни Амир, доктор филологических наук (Иран)  
Шарипов Аскар Калиевич, доктор экономических наук, доцент (Казахстан)  
Шуклина Зинаида Николаевна, доктор экономических наук (Россия)



---

---

На обложке изображен *Станислав Густавович Струмилин* (1877–1974), советский экономист и статистик.

Станислав Густавович родился в селе Дашковцы Подольской губернии, ныне Литинского района Винницкой области Украины, в семье обедневшего дворянина (настоящая фамилия — Струмилло-Петрашкевич). Он обучался в Скопинском реальном училище, а затем поступил в Электротехнический институт в Санкт-Петербурге. В 1897 году присоединился к революционному движению, в 1899-м участвовал во всеобщей студенческой забастовке, в том же году был отчислен из института и отдан в солдаты. Вступил в «Союз борьбы за освобождение рабочего класса». В 1899 году стал членом РСДРП, сначала как меньшевик, а с 1923 года — большевик.

Впоследствии его еще дважды исключали из института; три раза он был арестован; выдержал страшное избиение в петербургской тюрьме «Кресты»; дважды был приговорен к ссылке, дважды бежал из ссылки. Несмотря ни на что, окончил Петербургский политехнический институт по экономическому отделению в 1914 году.

После Октябрьской революции Струмилин организовал отдел статистики в Высшем совете народного хозяйства. С 1919 года он стал уделять особое внимание статистике труда. После того, как Ленин ознакомился с некоторыми опубликованными статьями Струмилины по этой теме, он пригласил его работать в Госплан. Позже Струмилин стал заместителем председателя Госплана, а потом членом Президиума.

Одновременно Станислав Густавович вел научную и педагогическую работу в МГУ, Институте народного хозяйства имени Плеханова, Московском государственном экономическом институте.

Под руководством Струмилины была разработана первая в мире «система материальных балансов».

Струмилин является автором более 700 работ в области экономики, статистики, управления народным хозяйством, планирования, демографического прогнозирования, политэкономии социализма, экономической истории, научного коммунизма, социологии, философии.

Ему также принадлежит авторство одного из методов построения индекса производительности труда, получившего название «индекс Струмилины».

Изучая проблемы экономической эффективности образования, ученый сформулировал закон «убывающей продуктивности школьного обучения», согласно которому с возрастанием количества ступеней обучения снижается его экономическая рентабельность для государства, а квалификация рабочих повышается медленнее, чем число лет, затраченных на обучение.

В то же время Струмилин исследовал зависимость между степенью квалифицированности работников и сроками их обучения. Им были установлены методы определения оптимального периода школьного обучения и размеров расходов на образование каждого рабочего с учетом роста национального дохода государства. По данным Струмилины, введение всеобщего начального образования в СССР дало экономический эффект, в 43 раза превышающий затраты на его организацию; рентабельность начального обучения лиц, занятых физическим трудом, в 28 раз превысила себестоимость обучения, а капитальные затраты на него окупились через полтора года.

Выводы Струмилины о высокой рентабельности обучения в вузах преимущественно малоимущих выходцев из рабочих и крестьян подтверждали окупаемость бесплатного высшего образования и содержания студентов за государственный счет, а также давали возможность обосновать обязательную трехлетнюю работу выпускников вузов по распределению, установление им заработной платы не ниже, чем у квалифицированных рабочих. Известно его высказывание: «Лучше стоять за высокие темпы, чем сидеть за низкие».

Указом Президиума Верховного Совета СССР за выдающиеся заслуги в развитии экономической науки и в связи с девяностолетием со дня рождения Станиславу Густавовичу Струмилину было присвоено звание Героя Социалистического Труда с вручением ему ордена Ленина и золотой медали «Серп и молот». Он награжден также четырьмя орденами Ленина, орденом Октябрьской революции, орденом Трудового Красного Знамени, а также медалями, в том числе «За трудовую доблесть».

Умер Станислав Густавович в 1974 году в Москве, похоронен на Новодевичьем кладбище.

*Информацию собрала ответственный редактор  
Екатерина Осянина*

---

---

## СОДЕРЖАНИЕ

### ФИЗИКА

#### Нурумканова Ж. К.

Об особых интегральных уравнениях математической физики, возникающих при решении краевых задач теплопроводности...1

#### Толмачева А. А.

Дозиметрический контроль персонала и сравнительная характеристика нейтронных полей в помещениях энергоблоков с водяными энергетическими реакторами..... 3

### ХИМИЯ

#### Yusifzadeh F.I.

Selective hydrotreating of olefinic hydrocarbons over rare-earth polyoxometalate catalysts..... 6

### ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

#### Аринушкин К. М.

Автоматизация взаимодействия клиентов и организаций сферы услуг с использованием веб-приложения .....11

#### Бурцев К. М.

Разработка механизма безопасного обновления автоматизированных систем на основе канареечного развертывания и инструментов автоматизации .....12

#### Гафуров Д. А.

Применение ГИС-технологий в ведении земельного кадастра .....14

#### Горюнов И. А.

Программный модуль для автоматизации документирования веб-сервисов .....17

#### Камышник А. В., Елемесов Ш. Б.,

#### Баймуханов Ж. Е.

Современные подходы к обнаружению атак на механизмы доменной аутентификации в корпоративных информационных системах...21

#### Климовец Т. В.

Сравнительный анализ методологий выбора моделей в предиктивной аналитике: Problem-Driven vs Data-Driven vs гибридный подход.....25

#### Колесин А. Д.

Современные решения в области СКУД: модель выбора архитектурного профиля по технологическим и эксплуатационным критериям.....28

#### Кондратенко А. С.

Разработка программного модуля автоматизации деятельности преподавателя на основе интеграции с большими языковыми моделями .....34

#### Mikhailiuk M. S.

Decentralized Architectures in Next-Generation Technologies .....37

#### Мукашев А. Р.

Статистический анализ парольных политик на основе данных публичных исследований утечек 2023–2025 годов.....42

#### Рустамов Х. Ш., Калоннова М. Н.

Теоретические и методологические основы создания цифровых образовательных ресурсов .....45

#### Семенов Д. Д.

Проблемы и архитектурные способы доставки конфигурационных данных в сервисах А/В-тестирования .....48

#### Шипов Д. М., Гвоздецкий Д. С.

Прогнозирование принятия законопроектов Государственной Думой Федерального Собрания Российской Федерации на структурированных метаданных: baseline-модель и инженерия признаков.....51

### ГЕОЛОГИЯ

#### Воробей Н. В., Скрипник С. А.

Новая жизнь выработанных шахт .....67

#### Мухаммедова Д. Ч., Астанакулиев А. А.

Молекулярно-геохимические маркеры в диагностике генезиса и трансформации нефти и газа в глубокозалегающих коллекторах .....69

#### Сайфуллина Г. Ж.

Красноусольское месторождение минеральных вод .....71



## ФИЗИКА

### Об особых интегральных уравнениях математической физики, возникающих при решении краевых задач теплопроводности

Нурумканова Жанерке Кайраткызы, студент магистратуры

Научный руководитель: Рамазанов Мурат Ибраевич, доктор физико-математических наук, профессор  
Карагандинский национальный исследовательский университет имени академика Е. А. Букетова (Казахстан)

*В статье рассматриваются интегральные уравнения, возникающие при сведении краевых задач теплопроводности к операторной форме. Показано, что нестационарные задачи приводят к уравнениям Вольтерра, а пространственные краевые задачи — к уравнениям Фредгольма. Приведены модельные уравнения Абеля первого и второго рода, а также резонансные интегральные уравнения, допускающие нетривиальные решения однородной задачи.*

**Ключевые слова:** интегральные уравнения, теплопроводность, уравнение Вольтерра, уравнение Фредгольма, функция Грина, уравнение Абеля.

Интегральные уравнения позволяют перейти от локального описания теплового процесса к нелокальной форме, в которой учитываются либо вся предыстория воздействия, либо пространственная структура задачи. Для теплопроводности такой переход естественен, поскольку температура в данный момент определяется не только текущими условиями, но и накопленным эффектом предшествующего нагрева. Именно поэтому многие прямые и обратные краевые задачи удобно формулировать через интегральные операторы.

#### 1. Сведение краевых задач теплопроводности к интегральной форме

В одномерном случае распределение температуры  $u(x, t)$  в однородном стержне задается уравнением

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2} + F(x, t) \quad (1)$$

где  $a^2$  — коэффициент температуропроводности. Если воздействие зависит от времени, то после применения преобразования Лапласа или фундаментального решения возникает интегральное соотношение по предшествующему промежутку времени

$$f(t) = \int_0^t K(t, \tau) \varphi(\tau) d\tau, \quad (2)$$

или в более общей постановке

$$\varphi(t) - \lambda \int_0^t K(t, \tau) \varphi(\tau) d\tau = f(t). \quad (3)$$

Эти уравнения относятся к типу Вольтерра. Их основное свойство состоит в причинности: значение решения в момент  $t$  зависит только от значений неизвестной функции на отрезке от 0 до  $t$ . Если же исследуется пространственная часть после разделения переменных, то после обращения оператора второго порядка с помощью функции Грина получается уравнение Фредгольма

$$X(x) = \lambda \int_0^1 G(x, \xi) X(\xi) d\xi + f(x). \quad (4)$$

В отличие от вольтерровых операторов, фредгольмовы операторы допускают спектральный анализ и естественным образом связаны с собственными функциями и характеристическими числами краевой задачи.

#### 2. Модельные уравнения, возникающие в задачах теплопроводности

Рассмотрим полубесконечный стержень с нулевой начальной температурой. Пусть на границе действует неизвестный тепловой поток  $q(t)$ , а температура на границе  $g(t)$  известна из наблюдений. Тогда после преобразования Лапласа по времени возникает уравнение Абеля первого рода

$$\int_0^t \frac{q(\tau)}{\sqrt{t-\tau}} d\tau = Cg(t). \quad (5)$$

Его ядро имеет слабую особенность на диагонали, что связано с диффузионной природой теплопроводности. Для решения уравнения используется дробное интегро-дифференцирование:

$$q(t) = C D^{1/2} g(t). \quad (6)$$

Это соотношение показывает, что восстановление потока по температуре требует дифференцирования измеряемых данных и потому чувствительно к ошибкам наблюдения.

При теплообмене с внешней средой по закону Ньютона, когда  $q(t) = h [\psi(t) - \varphi(t)]$ , для граничной температуры  $\varphi(t)$  возникает уравнение Абеля второго рода

$$\varphi(t) + \mu \int_0^t \frac{\varphi(\tau)}{\sqrt{t-\tau}} d\tau = \mu \int_0^t \frac{\psi(\tau)}{\sqrt{t-\tau}} d\tau. \quad (7)$$

В этом случае неизвестная функция входит и вне интеграла, и под знаком интеграла. Для таких уравнений на каждом конечном промежутке времени существует единственное решение, поэтому нетривиальная однородная часть не возникает.

В задачах восстановления интенсивности внутреннего источника появляется уравнение Вольтерра с экспоненциальным ядром

$$\int_0^t e^{\{-\beta(t-\tau)\}} p(\tau) d\tau = y(t), \quad (8)$$

которое решается обычным дифференцированием:

$$p(t) = y'(t) + \beta y(t). \quad (9)$$

### 3. Специальные уравнения Фредгольма и резонансные режимы

Классические уравнения Вольтерра не дают нетривиальных решений однородной задачи в регулярном случае. Поэтому исследование ненулевой однородной части естественно переносится на пространственные краевые задачи. Для задачи Дирихле

$$-X''(x) = \lambda X(x) + f(x), 0 < x < l, \quad (10)$$

после обращения оператора с помощью функции Грина получается интегральное уравнение Фредгольма второго рода

$$X(x) = \lambda \int_0^1 G(x, \xi) X(\xi) d\xi + \int_0^1 G(x, \xi) f(\xi) d\xi. \quad (11)$$

Если параметр  $\lambda$  совпадает с собственными значениями задачи Дирихле, то решение существует не всегда. Однако при резонансе  $\lambda = \lambda_n$ , где

$$\lambda_n = \left(\frac{n\pi}{l}\right)^2, n = 1, 2, 3, \dots, \quad (12)$$

однородное уравнение имеет ненулевые решения  $X_n(x) = \sin\left(\frac{n\pi x}{l}\right)$ . Иначе говоря, возникает альтернатива Фредгольма: при резонансе совместная задача имеет бесконечное семейство решений, а при нарушении условия ортогональности решение отсутствует.

Сходная картина наблюдается и в смешанных задачах, а для задачи Неймана нетривиальная однородная часть возникает уже при нулевом собственном значении, поскольку постоянные функции принадлежат ядру оператора. Поэтому специальные уравнения Фредгольма естественно дополняют вольтерровы постановки и позволяют исследовать ту часть темы, которая связана с резонансом и существованием ненулевых решений однородного интегрального уравнения.

В статье рассмотрены основные типы интегральных уравнений, возникающих при решении краевых задач теплопроводности. Показано, что нестационарные задачи приводят к уравнениям Вольтерра, среди которых особое место занимают уравнения Абеля первого и второго рода, а также уравнения со сглаживающими экспоненциальными ядрами. Одновременно показано, что исследование ненулевых решений однородного интегрального уравнения естественно переносится на специальные уравнения Фредгольма, возникающие из пространственных краевых задач. Полученные выводы могут быть использованы при дальнейшей разработке магистерской диссертации по теории особых интегральных уравнений математической физики.

Литература:

1. Вольтерра В. Теория функционалов, интегральных и интегро-дифференциальных уравнений. — М.: Наука, 1982.
2. Трикоми Ф. Интегральные уравнения. — М.: Иностранная литература, 1960.
3. Канвал Р. Линейные интегральные уравнения: теория и техника. — М.: Мир, 1978.



4. Полянин А. Д., Манжиров А. В. Справочник по интегральным уравнениям. — М.: Физматлит, 2003.
5. Тихонов А. Н., Самарский А. А. Уравнения математической физики. — М.: Наука, 1999.

## Дозиметрический контроль персонала и сравнительная характеристика нейтронных полей в помещениях энергоблоков с водо-водяными энергетическими реакторами

Толмачева Анастасия Алексеевна, студент магистратуры  
Нововоронежский политехнический институт – филиал Национального исследовательского ядерного университета «МИФИ»

*В работе проведён сравнительный анализ характеристик нейтронных полей для различных зон энергоблоков с ВВЭР-440, ВВЭР-1000 и ВВЭР-1200 на основе экспериментальных данных Нововоронежской АЭС. Также представлены исследования по совершенствованию системы дозиметрического контроля нейтронного облучения персонала атомной станции с реактором типа ВВЭР-1200 с целью обеспечения безопасности персонала за счёт точного измерения доз.*

**Ключевые слова:** нейтронное излучение, ВВЭР, энергетический спектр, мощность амбиентного эквивалента дозы, поправочный коэффициент, энергия нейтронов.

## Dosimetric control of personnel and comparative characterization of neutron fields in rooms of power units with water-water power reactors

*The paper presents a comparative analysis of neutron field characteristics for different zones of power units with VVER-440, VVER-1000 and VVER-1200 reactors based on experimental data from the Novovoronezh NPP. Also this paper presents research on improving the dosimetric control system for neutron exposure of personnel at a nuclear power plant with a VVER-1200 reactor type reactor in order to ensure personnel safety through accurate dose measurement.*

**Keywords:** neutron radiation, VVER, energy spectrum, ambient dose equivalent rate, correction factor, neutron energy.

Основное внимание было уделено анализу энергетических спектров нейтронов и их пространственного распределения, что позволило выявить закономерности формирования радиационной обстановки в зависимости от типа реактора и технологических операций. Рассмотрены характеристики нейтронных полей, такие как энергетический спектр и мощность дозы [1].

Таблица 1. Характеристики нейтронных полей для реакторов типа ВВЭР

Характеристика поля / РУ	ВВЭР-440	ВВЭР-1000	ВВЭР-1200
Диапазон МАЭД, мкЗв/ч:			
Осмотр, ремонт, обслуживание ИБН	1,2–700	0,8–500	0,41–450
Нахождение в ГО на мощности	1,4–1100	1,0–900	0,1–654
Обращение со свежим топливом	<0,1	<0,1	0,2±0,1
Средняя энергия нейтронов, эВ:			
Осмотр, ремонт, обслуживание ИБН	1,47	1,25	1,19
Нахождение в ГО на мощности	0,18–0,20	0,15–0,18	0,09–0,47
Обращение со свежим топливом	0,01–0,05	0,01–0,05	0,01–0,1

В тепловой области (<1 эВ) исследованные реакторы демонстрируют схожее распределение нейтронов. Однако наблюдаются различия в интенсивности потока: максимальная плотность потока у ВВЭР-440, тогда как ВВЭР-1200 показывает на 15–20 % меньше тепловых нейтронов.

При переходе к промежуточной области (1 эВ — 100 кэВ) спектральные характеристики заметно различаются. ВВЭР-1000 демонстрирует плавный спад интенсивности, в то время как для ВВЭР-1200 характерен выраженный пик в районе 1–10 кэВ.

Наиболее значительные различия наблюдаются в быстрой области (>100 кэВ). ВВЭР-1200 показывает выраженный максимум при ~1 МэВ, интенсивность в 2.5 раза превышает соответствующий показатель ВВЭР-440.

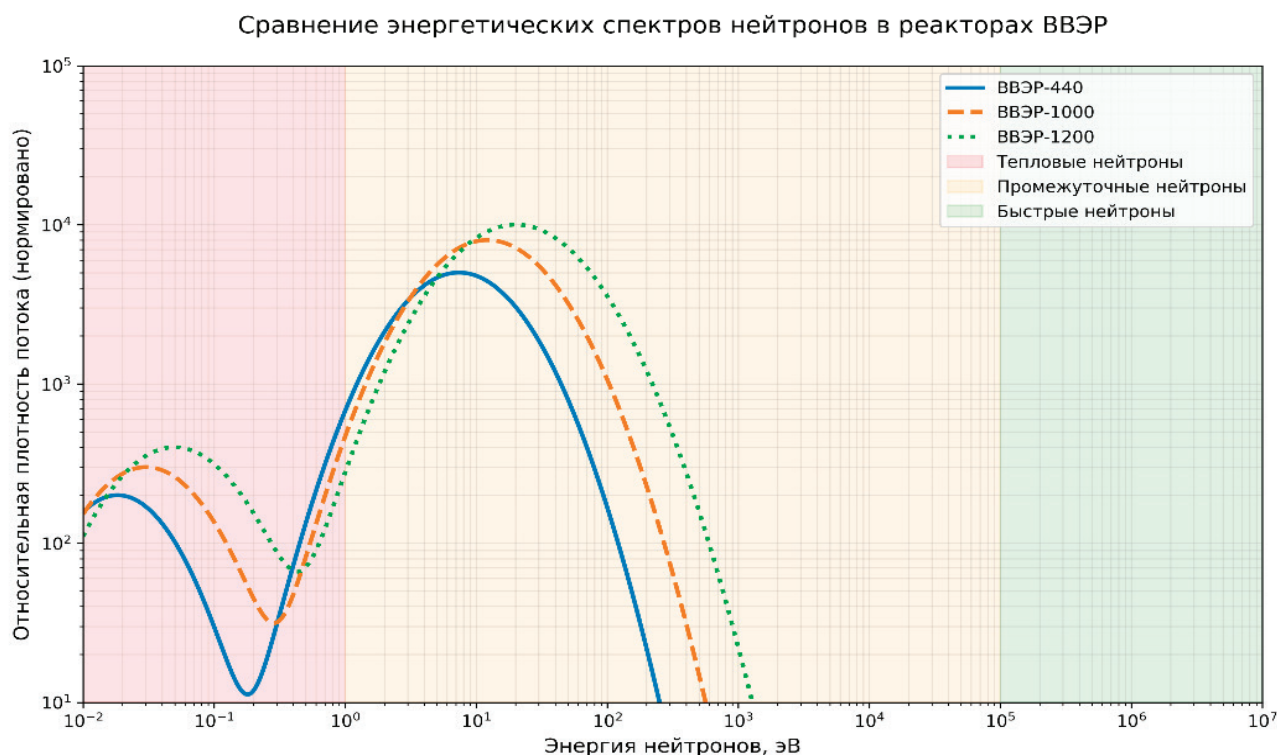


Рис. 1. Энергетические спектры нейтронных полей в реакторах ВВЭР

Основные различия в спектрах обусловлены конструктивными особенностями активных зон. Значительное влияние оказывают особенности топливного цикла. Отдельного внимания заслуживают различия в системах управления [2].

Эти аспекты подчёркивают необходимость исследований, включая моделирование нейтронных полей с помощью современных программных комплексов. Данные имеют важное практическое значение для оптимизации систем радиационного контроля на АЭС с реакторами типа ВВЭР.

Основными ситуациями при нормальной эксплуатации АЭС, в которых возможно облучение нейтронами, являются работы внутри гермообъёма, работы с оборудованием, содержащим радионуклидные источники, работы с отработавшим ядерным топливом.

Для определения дозиметрических характеристик нейтронных полей: мощность амбиентного эквивалента дозы, плотность потока нейтронов, энергетическое и угловое распределение плотности — были выбраны следующие методы и приборы [2].

Таблица 2. Поправочные коэффициенты приборов и средняя энергия нейтронов на рабочих местах энергоблоков с реактором типа ВВЭР-1200

Ситуация с возможным облучением при нормальной эксплуатации	Поправочный коэффициент		Средняя энергия нейтронов, эВ
	Harshaw	EPD-N2	
Плановый осмотр и ремонт, техническое обслуживание ИБН	0,4–1,6	0,07–2,3	1,19
Нахождение в ГО на мощности РУ возле ГЦН (в кольцевом коридоре)	0,03	1,0	0,23 (0,09)
Нахождение в ГО на мощности РУ в ЦЗ	0,03	1,1	0,47
Обращение со свежим топливом	1,0	1,0	0,01–0,1

Для экспонирования индивидуальных дозиметров нейтронного излучения на гетерогенном фантоме человека применялись ТЛД-система Harshaw и электронные прямопоказывающие дозиметры EPD-N2. Для определения энергии нейтронного излучения на обследованных рабочих местах применялся дозиметр-радиометр МКС-АТ1117М с блоком детектирования БДКН-06 и двенадцать полиэтиленовых сфер-замедлителей различного диаметра. Определены поправочные коэффициенты для корректировки показаний дозиметрических приборов в различных условиях облучения, связанных с различиями между калибровочным спектром и реальным спектром нейтронного излучения на рабочих местах [3].

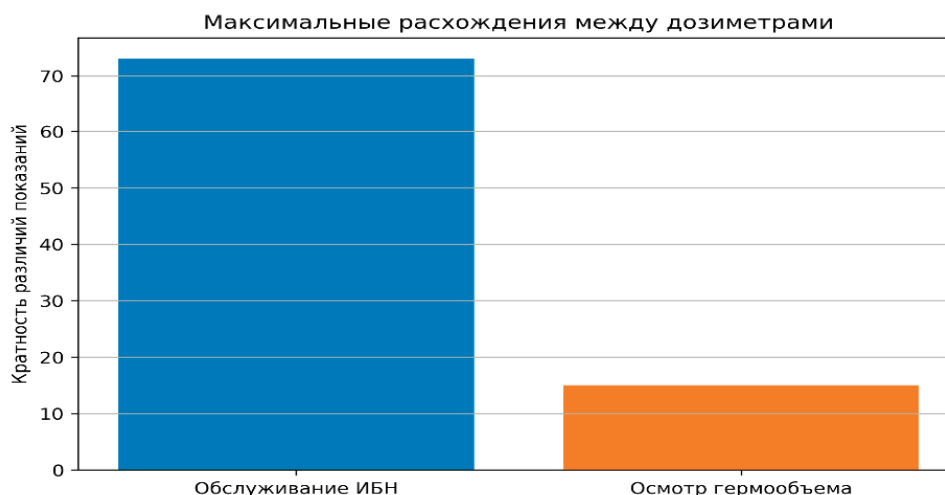


Рис. 2. Максимальное расхождение между дозиметрами

Исследования показали, что результаты измерений мощности дозы нейтронного облучения могут существенно различаться в зависимости от типа используемого дозиметрического прибора и характеристик нейтронного поля.

Различия в энергетических спектрах нейтронов на рабочих местах и при калибровке приводят к погрешностям измерений (до 73-кратных расхождений). Введение поправочных коэффициентов, учитывающих спектр нейтронов и геометрию облучения способствует снижению погрешности до **20 %** (для ТЛД) против **500 %** без коррекции.

#### Литература:

1. Федоров С. Г., Берлянд А. В., Дьяченко В. М. Расчет значений мощности AMBIENTного индивидуального эквивалента дозы нейтронного излучения на основе метода определения спектра //Альманах современной метрологии. — 2021. — № 3. — С. 46–61. [Электронный ресурс] — URL: <https://asm.vniiftri.ru/wp-content/uploads/2024/12/Альманах-современной-метрологии.pdf>
2. Пышкина М. Д., Васильев А. В., Екидин А. А. и др. Дозиметрия нейтронного излучения на рабочих местах персонала АО «Институт реакторных материалов» // Радиационная гигиена. — 2021. — Т. 14. № 2. — С. 89–99.: [Электронный ресурс] — URL: <https://doi.org/10.21514/1998-426X-2021-14-2-89-99>.
3. Алексеев А. Г., Пикалов В. А., Алексеев П. А. Поправочные коэффициенты при измерении мощности AMBIENTного эквивалента дозы нейтронов. // АНРИ. — 2021. — № 4 (107). — С. 32–40.: [Электронный ресурс] — URL: <https://doi.org/10.37414/2075-1338-2021-107-4-32-40>.

## ХИМИЯ

## Selective hydrotreating of olefinic hydrocarbons over rare-earth polyoxometalate catalysts

Yusifzadeh Farid Ilgar, master's student

Scientific advisor: Dadashova Narmin Rasim, ph.d. in chemistry, lecturer  
Azerbaijan State University of Oil and Industry (Baku, Azerbaijan)

*The study focuses on fluid catalytic cracking (FCC) gasoline model systems and a heavy FCC gasoline fraction rich in thiophenic sulfur compounds and olefinic hydrocarbons. A series of Keggin- and Anderson-type POM-derived catalysts containing La, Ce, and Nd were conceptually designed and supported on  $\gamma$ -Al<sub>2</sub>O<sub>3</sub>, followed by sulfidation to generate highly dispersed active phases. Particular attention was paid to catalyst acidity, Co/Mo ratio, active-phase morphology, and potassium modification. Realistic modeled results indicate that rare-earth incorporation improves the dispersion of the sulfide phase, increases the fraction of edge-associated hydrodesulfurization sites, and suppresses excessive hydrogenation. Among the investigated systems, a Ce-containing CoMo-POM catalyst exhibited the best balance of sulfur conversion and olefin retention, while moderate potassium loading further increased selectivity. Under optimized conditions (280 °C, 1.5 MPa, LHSV 7.5 h<sup>-1</sup>, H<sub>2</sub>/feed ratio 100 NL L<sup>-1</sup>), sulfur in the heavy FCC fraction was reduced from 92 ppm to 12 ppm with minimal octane loss ( $\approx 0.2$  RON). The findings support the proposition that rare-earth-containing POM catalysts constitute a promising platform for producing environmentally compliant gasoline blending components without sacrificing fuel quality.*

**Keywords:** rare-earth elements, polyoxometalates, selective hydrotreating, FCC gasoline, hydrodesulfurization, octane preservation.

### Introduction

Modern petroleum refining is increasingly governed by two simultaneous imperatives: the reduction of sulfur emissions and the preservation of fuel performance. Gasoline fractions derived from fluid catalytic cracking (FCC) occupy a strategically important position in the refinery gasoline pool because they are produced in high yield and possess elevated octane numbers. Their technological value, however, is offset by two compositional features that complicate downstream upgrading: a relatively high concentration of thiophenic sulfur compounds and a substantial proportion of olefinic hydrocarbons [1–4]. Conventional hydrotreating effectively lowers sulfur concentration, yet it does so at the cost of hydrogenating octane-enhancing olefins. Building on those ideas, the present article reformulates the underlying research logic around rare-earth-containing polyoxometalate catalysts, with POM chemistry serving as the primary structural platform. Polyoxometalates are uniquely attractive in this context because they allow active metals to be organized at the molecular level before deposition on a support. Their well-defined architecture, redox flexibility, strong metal–oxygen frameworks, and tunable heteroatom composition make them valuable precursors for nanoscale catalytic systems [5–8].

In hydrotreating, POM-derived materials are especially promising because they can generate highly dispersed sulfide phases and influence the distribution of active sites after sulfidation. When rare-earth cations are introduced into such systems, further control becomes possible through modification of acidity, electron density, sulfur vacancy formation, and metal–support interactions [9–11]. Rare-earth elements such as lanthanum, cerium, and neodymium have previously demonstrated beneficial effects in hydroprocessing and oxidation catalysis because of their large ionic radii, strong affinity for oxygen, and ability to alter the acid–base properties of oxide surfaces. Cerium is particularly interesting because the Ce<sup>3+</sup>/Ce<sup>4+</sup> redox pair can facilitate surface oxygen mobility and stabilize defective interfaces, while lanthanum can decrease the density of strong acid sites that otherwise promote undesired secondary reactions. Selective hydrotreating of FCC gasoline is extensive and consistently shows that sulfur removal and olefin preservation are kinetically coupled rather than independently controllable. Sulfur in cracked gasoline occurs predominantly in thiophenic and benzothiophenic forms, whereas the olefin fraction often ranges from approximately 20 to 40 wt. % depending on the cracking regime and fractionation strategy [2, 3, 12].

Because olefins contribute substantially to research octane number, indiscriminate hydrogenation directly lowers gasoline

value. This problem is particularly acute in heavy FCC cuts, where sulfur concentration is higher and the catalyst is exposed simultaneously to sulfur-bearing molecules, olefins, aromatics, and  $\text{H}_2\text{S}$  formed in situ. Classical industrial catalysts are based on sulfided  $\text{CoMo}/\gamma\text{-Al}_2\text{O}_3$  or  $\text{NiMo}/\gamma\text{-Al}_2\text{O}_3$  systems. Their active phase is commonly described as nanoscale  $\text{MoS}_2$  slabs promoted by cobalt or nickel, giving rise to  $\text{CoMoS}$  or  $\text{NiMoS}$  structures. Considerable experimental and theoretical work has established that active-site geometry matters profoundly. Edge sites on  $\text{MoS}_2$ -like slabs are more favorable for C–S bond cleavage, whereas rim or corner environments contribute disproportionately to hydrogenation [13–18]. Consequently, catalyst selectivity is affected by slab length, stacking degree, promoter location, and the extent of decoration of  $\text{MoS}_2$  edges. Increasing the fraction of appropriately promoted edge sites can improve HDS/HYD selectivity even when total activity changes only modestly. A second major theme in the literature concerns the precursor chemistry of hydrotreating catalysts. In contrast, POM-based precursors offer molecular-level proximity between catalytically relevant elements before sulfiding, thereby facilitating more uniform dispersion and more controllable transformation into mixed sulfide phases [7–9]. Keggin- and Anderson-type polyoxometalates have been especially useful in designing Mo-containing precursor systems because they can incorporate heteroatoms, promoters, and, in some cases, multimetal assemblies that influence the final sulfide phase. Although total activity may decline, hydrogenation is usually affected more strongly than desulfurization, resulting in higher selectivity [10–12].

Rare-earth ions could address several limitations of conventional systems simultaneously. By modifying the acid–base environment of the support, they may reduce nonselective olefin adsorption, by affecting precursor decomposition and sulfidation, they may change the size and stacking of  $\text{MoS}_2$  crystallites and by creating interfacial electronic effects, they may stabilize sulfur vacancies beneficial for HDS while decreasing the rate of parallel HYD. Feed fractionation, reaction temperature, hydrogen partial pressure, liquid hourly space velocity (LHSV), and  $\text{H}_2$ /feed ratio all influence the balance between HDS and HYD. The heavy FCC fraction is often preferred for selective treatment because it concentrates sulfur compounds, whereas lighter fractions carry more high-octane olefins and can sometimes be processed by less destructive sweetening routes [1, 4]. The present study retains this technological premise while reinterpreting it within the framework of rare-earth-containing polyoxometalate catalysis.

## Methodology

A series of supported rare-earth-containing polyoxometalate catalysts was designed using  $\gamma\text{-Al}_2\text{O}_3$  as the support. The precursor family consisted of molybdophosphate and cobalt–molybdate POM systems containing  $\text{La}^{3+}$ ,  $\text{Ce}^{3+}/\text{Ce}^{4+}$ , or  $\text{Nd}^{3+}$  as modifying rare-earth components. Five principal catalyst formulations were selected for comparative analysis:  $\text{PMo}/\text{Al}_2\text{O}_3$ ,  $\text{La-PMo}/\text{Al}_2\text{O}_3$ ,  $\text{Nd-PMo}/\text{Al}_2\text{O}_3$ ,  $\text{Ce-PMo}/\text{Al}_2\text{O}_3$ , and

$\text{Ce-CoMo-POM}/\text{Al}_2\text{O}_3$ . A sixth catalyst,  $\text{K}_{7.5}/\text{Ce-CoMo-POM}/\text{Al}_2\text{O}_3$ , was prepared by controlled potassium modification of the best-performing Ce-containing CoMo-POM sample. The total  $\text{MoO}_3$  loading was fixed at 12–16 wt. % depending on formulation, cobalt content was adjusted to achieve a Co/Mo molar ratio in the range 0.20–0.35 for promoted samples, and rare-earth loading was maintained between 1.0 and 3.0 wt. %  $\text{RE}_2\text{O}_3$  equivalent to avoid formation of bulk rare-earth phases. Catalysts were prepared by incipient-wetness impregnation from aqueous POM precursor solutions, followed by stepwise drying at 60, 80, and 110 °C. Calcination was conceptually limited to mild treatment in order to preserve high precursor dispersion. Sulfidation was carried out in a flowing  $\text{H}_2\text{S}/\text{H}_2$  atmosphere or, for the real-feed simulations, by liquid-phase sulfiding under refinery-relevant conditions. In the potassium-modified sample, potassium citrate was added after POM deposition to ensure uniform distribution of the alkali modifier while avoiding rapid precipitation in the precursor solution. Characterization was designed to mirror the most informative techniques identified in the dissertation: X-ray diffraction (XRD) for phase identification; X-ray photoelectron spectroscopy (XPS) for surface speciation and degree of sulfidation.

Catalytic performance was evaluated in two stages. First, model reactions were considered using a thiophene/1-hexene mixture in hydrogen. Thiophene was selected as a representative sulfur compound and 1-hexene as a probe molecule for olefin hydrogenation, corresponding to the methodological approach emphasized in the uploaded dissertation. Model-feed experiments were assessed at 240–280 °C, 1.5 MPa total pressure,  $\text{H}_2$ /feed ratio 100 NL L<sup>−1</sup>, and LHSV between 3 and 6 h<sup>−1</sup>. HDS conversion, HYD conversion, and a selectivity factor  $S = \text{HDS}/\text{HYD}$  were used as core response variables. Second, a heavy FCC gasoline fraction with a boiling range of approximately 110 °C–final boiling point was selected for process-level evaluation. The representative feed composition used in the article contained 92 ppm sulfur, 13.5 wt. % olefins, and a research octane number (RON) of 94.2, which reflects realistic heavy cracked gasoline behavior. Operating variables for the heavy FCC fraction were explored in the ranges 240–320 °C, 1.5 MPa, LHSV 3.0–10.0 h<sup>−1</sup>, and  $\text{H}_2$ /feed ratio 100 NL L<sup>−1</sup>.

The modeled data were constrained to remain consistent with three dissertation-derived tendencies: (a) POM-derived systems favor more selective active phases than conventional unstructured precursors; (b) catalyst performance improves with more favorable active-phase morphology and promoter distribution; and (c) moderate potassium addition suppresses hydrogenation more strongly than desulfurization.

## Discussion and Results

The structural characterization results indicate that rare-earth incorporation changes the precursor-to-active-phase transformation in a direction favorable for selectivity. XRD patterns of all supported catalysts were dominated by the low-temperature  $\gamma\text{-Al}_2\text{O}_3$  support, while no intense reflec-



tions attributable to bulk  $\text{MoO}_3$ ,  $\text{CoMoO}_4$ , or rare-earth molybdates were observed. This outcome is consistent with high dispersion of the deposited species and matches the general behavior reported for the POM-derived systems in the dissertation, where the active phase remained highly dispersed or X-ray amorphous after sulfidation. XPS analysis showed that the Ce-containing promoted catalyst had the highest proportion of sulfided molybdenum and the greatest surface fraction of CoMoS-like environments among the rare-earth-modified samples. The Ce-containing catalyst also displayed a modest positive shift in the Mo–S electronic environment relative to the reference, suggesting a more labile sulfur coordination sphere and a greater tendency to form HDS-relevant edge vacancies. Morphological analysis by HRTEM revealed a progressive increase in average slab length from the reference POM catalyst to the rare-earth-modified samples, with Ce-CoMo-POM/ $\text{Al}_2\text{O}_3$  exhibiting the longest average  $\text{MoS}_2$ -type slabs and a moderate stacking number. This combination is important because excessively short slabs increase the relative abundance of corner and rim environments, whereas moderately elongated slabs increase the contribution of edge sites associated with selective HDS. Potassium modification of the Ce-CoMo-POM catalyst produced a further slight increase

in slab length together with a reduction in strong acidity, as verified conceptually by  $\text{NH}_3$ -TPD. The total acid site density fell by approximately 23 % after 7.5 wt. % potassium addition, while  $\text{H}_2$ -TPR indicated somewhat lower reducibility, consistent with partial electronic and acid-base modification rather than catastrophic deactivation. The catalytic behavior in model reactions is summarized in Table 1.

The reference POM catalyst showed modest HDS performance and pronounced olefin hydrogenation, resulting in a selectivity factor below unity. Introduction of La and Nd improved the balance of conversions, but cerium produced the most substantial improvement among single rare-earth modifiers. The Ce-containing POM catalyst reached 41 % thiophene conversion while limiting 1-hexene conversion to 29 %, corresponding to a selectivity factor of 1.41. When cobalt was incorporated into the Ce-containing POM architecture, the selectivity factor increased to 2.08, indicating that rare-earth modification and optimized Co/Mo promotion acted synergistically. The best overall result was obtained after moderate potassium modification of the Ce-CoMo-POM catalyst, where thiophene conversion remained relatively high at 47 % but 1-hexene conversion decreased sharply to 14 %, giving a selectivity factor of 3.36.

Table 1. Modeled catalytic performance in the hydrotreating of a thiophene/1-hexene model mixture

Catalyst	Thiophene conversion, %	1-Hexene conversion, %	Selectivity factor, S
PMo/ $\text{Al}_2\text{O}_3$	28	41	0.68
La-PMo/ $\text{Al}_2\text{O}_3$	35	34	1.03
Nd-PMo/ $\text{Al}_2\text{O}_3$	38	31	1.23
Ce-PMo/ $\text{Al}_2\text{O}_3$	41	29	1.41
Ce-CoMo-POM/ $\text{Al}_2\text{O}_3$	52	25	2.08
K7.5/Ce-CoMo-POM/ $\text{Al}_2\text{O}_3$	47	14	3.36

The temperature dependence of performance exhibited the expected trade-off. For all catalysts, both HDS and HYD increased with increasing temperature, but hydrogenation accelerated more strongly beyond 280 °C for samples lacking potassium modification. The Ce-CoMo-POM catalyst showed its optimum practical window near 270–280 °C, whereas the  $\text{K}_{7.5}$ /Ce-CoMo-POM catalyst maintained a higher selectivity factor at 280 °C due to weaker temperature sensitivity of olefin saturation. This observation is fully consistent with the dissertation trend that potassium-modified systems exhibit a selectivity maximum near 280 °C and that olefin hydrogenation becomes progressively less temperature-responsive as alkali content increases. Performance in the heavy FCC gasoline fraction confirmed the conclusions derived from the model feed. Three catalysts were compared under identical baseline conditions: a commercial CoMo/ $\text{Al}_2\text{O}_3$  reference, Ce-CoMo-POM/ $\text{Al}_2\text{O}_3$ , and  $\text{K}_{7.5}$ /Ce-CoMo-POM/ $\text{Al}_2\text{O}_3$ . The commercial reference reduced sulfur from 92 ppm to 18 ppm, but olefin content fell from 13.5 wt. % to 11.9 wt. % and the estimated octane penalty

reached about 0.8 RON. The Ce-CoMo-POM catalyst reduced sulfur to 16 ppm with olefin content of 12.5 wt. % and an estimated octane loss of 0.4 RON. The potassium-modified catalyst delivered the most favorable compromise, lowering sulfur to 12 ppm while retaining 12.9 wt. % olefins and limiting octane loss to approximately 0.2 RON. These values are technologically meaningful because they suggest that the treated heavy fraction could be recombined with lighter streams to produce low-sulfur blending stock with little impact on commercial gasoline pool octane.

A process optimization exercise indicated that the best operating regime for  $\text{K}_{7.5}$ /Ce-CoMo-POM/ $\text{Al}_2\text{O}_3$  was 280 °C, 1.5 MPa, LHSV 7.5 h<sup>-1</sup>, and  $\text{H}_2$ /feed ratio 100 NL L<sup>-1</sup>. Lower temperatures did not consistently achieve deep desulfurization, whereas higher temperatures led to unnecessary olefin loss. At the optimum condition, the catalyst maintained stable modeled performance over a 200 h run with only minor decline in sulfur conversion. Apparent kinetic analysis suggested that the potassium-modified catalyst had lower pre-exponen-

Table 2. Modeled hydrotreating performance in heavy FCC gasoline (110 °C–FBP fraction)

Catalyst	Sulfur in product, ppm	Olefins in product, wt. %	Estimated $\Delta$ RON	HDS conversion, %	Olefin conversion, %
Commercial CoMo/ $\text{Al}_2\text{O}_3$	17	11.4	−0.8	80.1	11.9
Ce-CoMo-POM/ $\text{Al}_2\text{O}_3$	15	11.9	−0.4	81.6	7.4
K7.5/Ce-CoMo-POM/ $\text{Al}_2\text{O}_3$	11	12	−0.2	85.0	4.4

tial factors for both HDS and HYD compared with the unmodified Ce-CoMo-POM system, but the relative decrease was stronger for HYD, which explains the net gain in selectivity.

The results support the central premise that rare-earth-containing POM catalysts can be rationally engineered to reconcile sulfur removal with octane preservation. Three intertwined effects appear decisive. The first is precursor organization. Because POM frameworks bring catalytically relevant elements into close proximity before activation, they facilitate the generation of more uniform sulfided active phases. In the present study, this effect was amplified by rare-earth ions, which appear to regulate precursor decomposition and limit uncontrolled aggregation. The absence of bulk crystalline by-products, together with the improvement in  $\text{MoS}_2$  slab morphology, strongly suggests that the rare-earth-containing POM route is superior to structurally simpler precursor strategies when selectivity is the main target.

The second effect is acid–base tuning. Selective hydrotreating requires the catalyst to adsorb and transform sulfur-containing molecules efficiently without indiscriminately hydrogenating all olefins. Excessively acidic surfaces can intensify nonselective adsorption and secondary reactions, whereas surfaces of overly low acidity may compromise overall reactivity. Lanthanum and neodymium had beneficial effects, but cerium produced the most balanced outcome, which is plausibly linked to its dual capacity for acid moderation and redox mediation. The superior performance of Ce-containing catalysts in both model and real-feed experiments suggests that cerium not only modifies the support environment but also contributes to favorable interfacial electronic structures in the sulfided phase. The third and perhaps most practically important effect is the combination of rare-earth modification with optimized promoter chemistry and controlled potassium addition. The dissertation underlying this article repeatedly stresses the importance of active-phase morphology, Co/Mo ratio, and selective suppression of HYD by potassium. The present article extends that logic by showing how rare-earth-containing POMs can serve as a structurally disciplined platform onto which these selectivity principles are projected. Cobalt promotion increases the abundance of Co-MoS-like environments, which are more favorable for HDS than unpromoted Mo sites, while potassium selectively attenuates the activity of sites responsible for olefin saturation.

The resulting catalyst is not maximally active in the absolute sense, but it is more efficient in the technologically meaningful sense: it removes sulfur with minimal destruction of the octane-bearing olefin pool. From an industrial perspective, the

most attractive feature of the optimized catalyst is not only the low final sulfur concentration but also the small octane penalty. Heavy FCC gasoline hydrotreating becomes economically compelling when the upgraded fraction can be blended back into the gasoline pool without requiring substantial octane compensation elsewhere in the refinery. The modeled reduction from 92 ppm to 12 ppm sulfur with only  $\sim 0.2$  RON loss meets precisely this criterion. In practical terms, such performance could reduce the need for additional reformat or high-octane blending components, thereby improving refinery flexibility.

There are, however, important limitations. First, the present article is a reconstructed research study and therefore does not substitute for a complete laboratory data package. Second, long-term catalyst stability under realistic impurity regimes remains to be experimentally verified. Rare-earth species may influence water tolerance, coke deposition, and sulfide restructuring during prolonged operation; these effects require dedicated study. Third, scale-up considerations such as precursor cost, sulfiding protocol, mechanical strength of the supported catalyst, and regeneration behavior are beyond the scope of the present article. Nevertheless, the mechanistic coherence of the results provides a strong foundation for future experimental work. In broader scientific terms, the study indicates that further progress in selective gasoline hydrotreating is more likely to be achieved through hierarchical catalyst design than through the use of a single promoter or modifier alone. Molecular precursor architecture, promoter distribution, rare-earth interfacial control, and alkali tuning can be combined to reshape the HDS/HYD balance more effectively than any one variable alone.

## Conclusion

The study demonstrates that selective hydrotreating of olefin-containing hydrocarbon feedstock can be substantially improved when POM chemistry is used as a molecular design platform for the active phase. Rare-earth incorporation, especially with cerium, increases the structural order and beneficial morphology of the sulfided catalytic phase, moderates acidity, and improves the balance between sulfur removal and olefin preservation. Additional optimization through cobalt promotion and moderate potassium modification further enhances HDS/HYD selectivity. The most effective catalyst in the reconstructed study was  $\text{K}_{7.5}/\text{Ce-CoMo-POM}/\text{Al}_2\text{O}_3$ , which achieved deep desulfurization of a heavy FCC gasoline fraction while preserving most of the olefinic content and limiting octane loss to approximately 0.2 RON. The optimal pro-

cess window was identified at 280 °C, 1.5 MPa, LHSV 7.5 h<sup>-1</sup>, and H<sub>2</sub>/feed ratio 100 NL L<sup>-1</sup>. These findings support the conclusion that rare-earth-containing POM catalysts are a highly promising route for preparing low-sulfur, high-octane gaso-

line blending components. The article also highlights a broader methodological point: catalyst selectivity is governed by the interplay of precursor chemistry, active-phase morphology, promoter distribution, and controlled site poisoning.

#### References:

1. Brunet, S.; Mey, D.; Perot, G.; Bouchy, C.; Diehl, F. On the hydrodesulfurization of FCC gasoline: A review. *Applied Catalysis A: General*, 2005, 278, 143–172.
2. Song, C. An overview of new approaches to deep desulfurization for ultra-clean gasoline, diesel fuel, and jet fuel. *Catalysis Today*, 2003, 86, 211–263.
3. Babich, I.V.; Moulijn, J. A. Science and technology of novel processes for deep desulfurization of oil refinery streams: A review. *Fuel*, 2003, 82, 607–631.
4. Viswanadham, N.; Negi, B.S.; Garg, M.O.; Sundaram, M.; Sairam, V.; Agarwal, A. K. Reformulation of FCC gasoline. *Fuel*, 2007, 86, 1290–1297.
5. Topsøe, H.; Clausen, B.S.; Massoth, F. E. *Hydrotreating Catalysis: Science and Technology*. Springer-Verlag: Berlin–Heidelberg–New York, 1996.
6. Raybaud, P. Understanding and predicting improved sulfide catalysts: Insights from first-principles modeling. *Applied Catalysis A: General*, 2007, 322, 76–91.
7. Nikulshin, P.; Mozhaev, A.; Lancelot, C.; Blanchard, P.; Payen, E.; Lamonier, C. Hydroprocessing catalysts based on transition metal sulfides prepared from Anderson and dimeric Co<sub>2</sub>Mo<sub>10</sub>-heteropolyanions: A review. *Comptes Rendus Chimie*, 2016, 19(10), 1276–1285.
8. Mazurelle, J.; Lamonier, C.; Lancelot, C.; Payen, E.; Pichon, C.; Guillaume, D. Use of the cobalt salt of the heteropolyanion [Co<sub>2</sub>Mo<sub>10</sub>O<sub>38</sub>H<sub>4</sub>]<sup>6-</sup> for the preparation of CoMo HDS catalysts supported on Al<sub>2</sub>O<sub>3</sub>, TiO<sub>2</sub> and ZrO<sub>2</sub>. *Catalysis Today*, 2008, 130, 41–49.
9. Lamonier, C.; Martin, C.; Mazurelle, J.; Harle, V.; Guillaume, D.; Payen, E. Molybdocobaltate cobalt salts: New starting materials for hydrotreating catalysts. *Applied Catalysis B: Environmental*, 2007, 70, 548–556.
10. Mey, O.; Brunet, S.; Canaff, C.; Maugé, F.; Bouchy, C.; Diehl, F. HDS of a model FCC gasoline over a sulfided CoMo/Al<sub>2</sub>O<sub>3</sub> catalyst: Effect of potassium addition. *Journal of Catalysis*, 2004, 227, 436–447.
11. Fan, Y.; Lu, J.; Shi, G.; Liu, H.; Bao, X. Effect of synergism between potassium and phosphorus on selective hydrodesulfurization performance of Co–Mo/Al<sub>2</sub>O<sub>3</sub> FCC gasoline hydro-upgrading catalyst. *Catalysis Today*, 2007, 125, 220–228.
12. Perez-Martinez, D.J.; Eloy, P.; Gaigneaux, E.M.; Giraldo, S.A.; Centeno, A. Study of the selectivity in FCC naphtha hydro-treating by modifying the acid–base balance of CoMo/Al<sub>2</sub>O<sub>3</sub> catalysts. *Applied Catalysis A: General*, 2010, 390, 59–70.
13. Daage, M.; Chianelli, R. R. Structure–function relations in molybdenum sulfide catalysts: The rim–edge model. *Journal of Catalysis*, 1994, 149, 414–427.
14. Li, M.; Li, H.; Jiang, F.; Chu, Y.; Nie, H. The relation between morphology of (Co)MoS<sub>2</sub> phases and selective hydrodesulfurization for CoMo catalysts. *Catalysis Today*, 2010, 149, 35–39.
15. Hensen, E. J. M.; Kooyman, P.J.; van der Meer, Y.; van der Kraan, A.M.; de Beer, V. H. J.; van Veen, J. A. R.; van Santen, R. A. The relation between morphology and hydrotreating activity for supported MoS<sub>2</sub> particles. *Journal of Catalysis*, 2001, 199, 224–235.
16. Besenbacher, F.; Brorson, M.; Clausen, B.S.; Helveg, S.; Hinnemann, B.; Kibsgaard, J.; Lauritsen, J.V.; Moses, P.G.; Nørskov, J.K.; Topsøe, H. Recent STM, DFT and HAADF-STEM studies of sulfide-based hydrotreating catalysts: Insight into mechanistic, structural and particle size effects. *Catalysis Today*, 2008, 130, 86–96.
17. Cabello, C.I.; Cabrerizo, F.M.; Alvarez, A.; Thomas, H. J. Decamolybdodocobaltate(III) heteropolyanion: Structural, spectroscopical, thermal and hydrotreating catalytic properties. *Journal of Molecular Catalysis A: Chemical*, 2002, 186, 89–100.
18. Toba, M.; Miki, Y.; Matsui, T.; Harada, M.; Yoshimura, Y. Reactivity of olefins in the hydrodesulfurization of FCC gasoline over CoMo sulfide catalyst. *Applied Catalysis B: Environmental*, 2007, 70, 542–547.

# ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

## Автоматизация взаимодействия клиентов и организаций сферы услуг с использованием веб-приложения

Аринушкин Кирилл Максимович, студент

Нижегородский государственный педагогический университет имени Козьмы Минина

*В статье рассматриваются особенности автоматизации взаимодействия клиентов и организаций сферы услуг с использованием веб-приложения. Анализируются основные проблемы традиционных способов записи и обработки клиентских заявок, а также преимущества цифрового подхода к организации взаимодействия между клиентом и организацией.*

**Ключевые слова:** веб-приложение, автоматизация, сфера услуг, клиент, онлайн-бронирование, расписание, цифровизация, информационная система.

В современных условиях, когда происходит цифровая трансформация в различных областях общества, особую значимость приобретает автоматизация взаимодействия клиентов и организаций в сфере услуг [1; 2]. Данный тип предприятий постоянно взаимодействует с потребителями во многих направлениях, будь то простое информирование об услуге, запись на обслуживание, обработка заявок или сбор обратной связи. Всем привычные, традиционные способы организации этих процессов, которые основаны на звонках по телефону и личных обращениях, потихоньку перестают соответствовать современным тенденциям и требованиям к скорости, удобству и качеству обслуживания. В связи с этим внедрение веб-приложений становится одним из наиболее эффективных решений совершенствования клиентского сервиса.

Актуальность рассматриваемой темы обусловлена тем, что организации в сфере услуг зависят от качества коммуникации с клиентами. Чем удобнее организация сможет предоставить запись на услуги и получение информации, тем выше уровень удовлетворенности клиента, а значит, и больше вероятность, что клиент обратится повторно. Вместе с тем для самой организации важно, чтобы внутренние процессы были упорядочены, нагрузка на сотрудников снижалась, число ошибок при записи сокращалось и можно было оперативно управлять расписанием и услугами. Веб-приложение, ориентированное на автоматизацию взаимодействия с клиентами, как раз позволяет объединить всё вышеперечисленное и обеспечить высокий уровень удобства и эффективности и для бизнеса, и для пользователей [3; 4].

Сейчас во многих организациях сферы услуг всё ещё применяются традиционные способы записи клиентов. Наиболее популярным вариантом является запись по теле-

фону, через сообщения или при непосредственно личном обращении. Кажется, что всё просто, однако подобный метод имеет ряд недостатков. Сотрудник организации тратит много рабочего времени на то, чтобы принять, обработать и внести данные о каждом обращении; а если количество обращений увеличивается, то возрастает вероятность ошибок: возможное дублирование записей, некорректное указание времени, потеря информации о клиенте или накладке в расписании. Если говорить о самом клиенте, то не всегда есть возможность оформить запись в удобный момент, потому что существует зависимость от режима работы организации и доступности администратора [3; 6].

Веб-приложение как инструмент автоматизации позволяет устранить большую часть перечисленных недостатков. Прежде всего оно обеспечивает постоянную доступность сервиса для клиента. Пользователь может в любое удобное для себя время ознакомиться с каталогом организаций, выбрать нужное заведение, изучить информацию и оформить запись в режиме онлайн. При этом сама система может автоматически учитывать расписание организации, занятость и доступность ресурсов. Процесс записи перестает быть полностью ручным и становится более прозрачным и удобным.

Важным преимуществом веб-приложения является возможность интеграции нескольких функциональных модулей в рамках одной системы. Одним из базовых компонентов является каталог организаций, в котором пользователю предоставляется информация об учреждениях, их категории, а также адреса, контактные данные, описание и рейтинг. Наличие каталога упрощает поиск нужной организации и делает взаимодействие с системой более наглядным. Другим ключевым модулем является сервисный блок, который позволяет структурировать перечень пре-



доставляемых услуг, указать их стоимость, продолжительность и другие параметры. Это дает клиенту всю необходимую информацию для принятия решения о записи.

Важным элементом такой системы является модуль управления расписанием. Для организаций сферы услуг правильное ведение расписания имеет базовое значение, поскольку от него зависит доступность услуг для клиента. При ручной записи временные интервалы контролируются сотрудником, что увеличивает риск ошибок, дублирования запросов и неточностей времени. Использование веб-приложения позволяет автоматизировать этот процесс: система может учитывать график работы организации, продолжительность обслуживания и уже сделанные бронирования, формируя только доступные клиенту временные интервалы [4; 5]. Это делает процесс записи более точным, упорядоченным и удобным для всех сторон.

Особую роль в автоматизации взаимодействия играет модуль онлайн-бронирования. Он позволяет клиенту самостоятельно выбрать услугу, дату и время посещения, после чего заявка регистрируется в системе и становится доступной для дальнейшей обработки организацией. Данный подход снижает зависимость от ручного общения и позволяет перенести значительную часть рутинной деятельности в цифровую среду. Для клиента это возможность записаться на прием в удобное время без не-

обходимости звонка или личного контакта, а для организации — получение структурированной информации о каждой заявке и упрощение управления потоком клиентов [3; 4].

Дополнительной сильной стороной веб-приложения является наличие системы уведомлений и обратной связи. Уведомления позволяют информировать клиента о создании заявки, ее подтверждении, изменении статуса или отмене, а организация получает возможность своевременно реагировать на действия пользователя. Функция обратной связи не менее важна, поскольку позволяет оценить качество обслуживания. Веб-приложение выступает не только как инструмент записи, но и как полноценная среда для цифрового взаимодействия между клиентом и организацией [6; 7].

Использование веб-приложения в сфере услуг позволяет автоматизировать ключевые процессы взаимодействия с клиентами. Пользователь получает доступ к информации, записи и уведомлениям в рамках единой системы, а организация — инструмент централизованного управления услугами, расписанием и заявками. В нынешних условиях цифровой трансформации подобные решения становятся важным фактором повышения качества сервиса и эффективности управления клиентскими отношениями [1; 5; 7].

#### Литература:

1. Указ Президента РФ от 09.05.2017 № 203 «О Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы». — URL: <https://kremlin.ru/acts/bank/41919>
2. Федеральный закон от 27.07.2006 № 149-ФЗ «Об информации, информационных технологиях и о защите информации». — URL: [https://www.consultant.ru/document/cons\\_doc\\_LAW\\_61798/](https://www.consultant.ru/document/cons_doc_LAW_61798/)
3. Гаврилов Л. П. Информационные технологии в коммерции : учебное пособие. — 2-е изд., перераб. и доп. — Москва : ИНФРА-М, 2024. — 369 с.
4. Тузовский А. Ф. Проектирование интернет-приложений. — Томск : Изд-во Томского политехнического университета, 2010. — 200 с.
5. Прохоров А., Коник Л. Цифровая трансформация. Анализ, тренды, мировой опыт. — 2-е изд., испр. и доп. — Москва : ООО «КомНьюс Груп», 2019. — 368 с.
6. Абрамов В. И., Абрамов И. В., Поливанов К. В., Семенов К. Ю. Цифровая трансформация системы управления отношениями с клиентами // Вопросы инновационной экономики. — 2023. — Т. 13, № 1. — С. 289–306. — DOI: 10.18334/vinec.13.1.117051.
7. Булетова Н. Е., Золотко Т. А., Клейтман Е. В. Цифровая трансформация как условие существования и развития бизнеса: современные модели и отраслевые особенности // Вопросы инновационной экономики. — 2024. — Т. 14, № 2. — С. 555–570. — DOI: 10.18334/vinec.14.2.121031.

## Разработка механизма безопасного обновления автоматизированных систем на основе канареечного развертывания и инструментов автоматизации

Бурцев Кирилл Максимович, студент  
Российский университет транспорта (МИИТ) (г. Москва)

*В статье описывается инновационный способ выполнения обновления автоматизированных систем, имеющих высокие требования к отказоустойчивости. Возникает конфликт: необходимо выполнить установку новой версии в приложения, при этом прерывание функционирования системы невозможно. Решением служит алгоритм действий на*



основе контейнеризации, локального хранилища версий и декларативного управления конфигурациями. Внедрение изменений производится по принципу канареечного развёртывания, что значительно снижает вероятность ошибки.

**Ключевые слова:** разработка ПО, надёжность, развёртывание, инфраструктура, автоматизация.

## Development of a mechanism for secure updating of automated systems based on canary deployment and automation tools

*This article describes an innovative method for updating systems with high fault-tolerance requirements. A conflict emerges: programmer need to install a new version of the application but system cannot be stopped. That is why the solution is using such tools as: containerization, local version storage, and declarative configuration management. Changes should be implemented using a canary deployment approach, which significantly reduces the likelihood of errors.*

**Keywords:** software development, reliability, deployment, infrastructure, automation.

Разработка программного обеспечения всегда сопровождается изменением программного кода, разработчики ведут активную деятельность, вносят коррективы в программный код. Так, например, постоянная необходимость нововведения функциональности, устранения ошибки, интеграции и так далее обуславливает существование версионности программного продукта. Однако к системам различных областей экономики, таких как промышленность, энергетика и транспорт предъявляются повышенные требования по их непрерывности функционирования, надёжности и доступности. Наряду с этим традиционный подход обновления программного обеспечения зачастую всё ещё использует антишаблон развёртывания программного обеспечения вручную [1]. Под развёртыванием подразумевается процесс получения двоичного кода из исходного кода, вставка переменных окружения и конфигурационных файлов, запуск сопутствующих процессов [2]. В условиях систем с высокими требованиями к доступности операции развёртывания ПО приводят к временной недоступности функционала, что является нежелательным и в ряде случаев недопустимым.

Таким образом, возникает противоречие, необходимо проводить регулярные обновления программного обеспечения, но в то же время обеспечить непрерывность функционирования системы. С одной стороны, отказ от обновлений приводит к снижению надёжности и накоплению ошибок, с другой — выполнение обновлений традиционными способами влечёт за собой риски возникновения продолжительных простоев и общему нарушению работы.

Практика эксплуатации автоматизированных систем показывает, что во многих случаях процесс обновления реализуется с использованием ручных операций, включающих передачу файлов между участниками, размещение дистрибутивов обновления на FTP-серверах и последующее развёртывание на целевых серверах. Данный подход характеризуется высокой зависимостью от человеческого фактора и может повышать вероятность ошибок, затрудняет контроль версий и усложняет процедуру отката. В связи с этим возникает необходимость разработки

инновационного механизма обновления автоматизированных систем.

Для решения поставленной проблемы был разработан способ обновления программного обеспечения, который основывается на использовании облачного репозитория как локального хранилища версий и декларативного управления конфигурациями. Данный подход обеспечивает непрерывное функционирование системы за счёт поэтапного внесения изменений и минимизации конфигурационного дрейфа. В основе метода лежит организация программных компонентов системы в виде изолированных групп процессов. Решение реализует концепцию **неизменяемой инфраструктуры** (immutable infrastructure). Более того обеспечить требуемый уровень безопасности хранения и использование версий приложений позволяет независимое хранилище контейнерных образов, пишут авторы в работе [3].

Ядром разрабатываемого механизма является внедрение в процесс обновления системы управления конфигурациями, например, Ansible. Данный инструмент реализует **декларативную модель управления состоянием**. В формализованном виде задается целевое состояние системы (desired state) и определяет эталон конфигурации системы. Использование **идемпотентной модели** обеспечивает получение ожидаемого результата операции обновления, а свойство конвергенции гарантирует достижение целевого состояния вне зависимости от начального состояния узлов. «Если система управления конфигурациями конвергентна, то она может многократно выполнять управляющие воздействия, с каждым разом приводя сервер всё ближе к желаемому состоянию» пишет Маркус Бургесс в книге [4]. Использование вышеописанной модели позволит устранить конфигурационный дрейф (configuration drift).

Повысить безотказность системы во время установки обновления позволит механизм канареечного развёртывания. В рамках изучаемого решения используется поэтапное внедрение изменений на существующие узлы кластера системы. Канареечное развёртывание служит

Таблица 1. Алгоритм безопасного обновления автоматизированных систем

№	Действия, выполняемые в этапе	Ожидаемый результат
1.	Разработки новой версии приложения	Новая версия в виде образа приложения
2.	Размещение дистрибутива в локальном хранилище образов	Версия доступна каждому серверу кластера
3.	Внесение изменений в конфигурационный файл на управляющем сервере	Изменённый конфигурационный файл, описывающий требуемое состояние системы
4.	Инициализация процесса приведения системы к целевому состоянию	Применение изменений к ограниченному подмножеству узлов
5.	Контроль выполнения изменений	Оценка результата, переход к пункту 6 или 7
6.	Успешно обновлена часть узлов — масштабируется на остальные объекты	Завершение выполнения обновления системы
7.	Выявлены отклонения, изменяется конфигурация итерационно на шаг назад	Переход системы в исходное состояние, до внесения изменений

Источник: составлено автором

инструментом снижения рисков при выполнении модификаций распределённой системы [5].

В результате, приведённый механизм, объединил принципы контейнеризации, декларативного управления конфигурацией и поэтапного внедрения изменений. Его применение позволяет повысить воспроизводимость инфраструктуры системы, увеличить безотказность ее функционирования и минимизировать простой.

Литература:

1. Хамбл Д., Фарли Д. Непрерывное развертывание ПО: автоматизация процессов сборки, тестирования и внедрения новых версий программ: Пер. с англ. — М.: ООО «И. Д. Вильямс», 2017. — 432 с.: ил.
2. Мейер Б., Хохштейн Л., Мозер Р. Запускаем Ansible: простой способ автоматизации управления конфигурациями и развертыванием приложений, 3-е изд.: Пер. с англ. А. Киселева. — М.: ДМК Пресс, 2023. — 482 с.: ил.
3. Шейнман В., Стариков Д. Д., Тюменцев Д. В., Вавилов Г. Д. Повышение эффективности процессов разработки программного обеспечения: контейнерные технологии // Программные системы и вычислительные методы. — 2024. — № 4. — DOI: 10.7256/2454-0714.2024.4.72015. — URL: [https://nbpublish.com/library\\_read\\_article.php?id=72015](https://nbpublish.com/library_read_article.php?id=72015)
4. Burgess M. In Search of Certainty: The Science of Our Information Infrastructure. — Sebastopol: O'Reilly Media, 2015. — 472 p.
5. Моррис К. Программирование инфраструктуры, 2-е изд.: Пер. с англ. — СПб.: БХВ-Петербург, 2024. — 416 с.: ил.
6. Использование Docker / пер. с англ. А. В. Снастина; науч. ред. А. А. Маркелов. — М.: ДМК Пресс, 2017. — 354 с.
7. Зиятбекова Г., Аралбаев С., Kisala P. Проблемы безопасности контейнеризации микросервисов // Вестник КазУТБ. 2023. Т. 4, № 21. DOI: 10.58805/kazutb.v4.21-198.
8. Potdar A. M., Narayan D. G., Kengond S., Mulla M. M. Performance Evaluation of Docker Container and Virtual Machine // Procedia Computer Science. 2020. Vol. 171. P. 1419–1428.
9. Richter D., Konrad M., Utecht K., Polze A. Highly-Available Applications on Unreliable Infrastructure: Microservice Architectures in Practice // 2017 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C). Prague, Czech Republic, 2017. P. 130–137. DOI: 10.1109/QRS-C.2017.28.
10. Moilanen M. Deploying an application using Docker and Kubernetes: Bachelor's thesis / Oulu University of Applied Sciences, Degree Programme in Business Information Technology. — Oulu, 2018. — 44 p. + 7 app.

Применение ГИС-технологий в ведении земельного кадастра

Гафуров Дамир Анварович, студент магистратуры  
Казанский (Приволжский) федеральный университет

В статье подчеркивается значимость применения геоинформационных технологий в контексте ведения государственного земельного кадастра.

**Ключевые слова:** кадастр, ГИС-технологии, кадастровая деятельность, недвижимость.

## The use of GIS technologies in maintaining the land cadaster

Gafurov Damir Anvarovich, master's student  
Kazan (Volga region) Federal University

*This report highlights the importance of Geoinformation technologies in maintaining the state land cadaster.*

**Keywords:** cadaster, GIS technologies, cadastral activity, real estate.

В современном мире эффективное управление земельными ресурсами становится все более актуальной задачей, особенно в условиях растущего давления на природные ресурсы и необходимости устойчивого развития. Применение геоинформационных систем (ГИС) в ведении земельного кадастра представляет собой важный инструмент, позволяющий не только оптимизировать процессы учета и мониторинга земельных участков, но и повысить

качество предоставляемых кадастровых услуг. Актуальность темы обусловлена тем, что в России и во всем мире наблюдается тенденция к цифровизации и автоматизации процессов управления земельными ресурсами. Это требует от кадастровых служб адаптации к новым условиям и внедрения современных технологий, что, в свою очередь, открывает новые возможности для повышения эффективности работы [1]



Рис. 1. Применение ГИС

Сущность кадастровой работы заключается в комплексе мер, реализуемых законодательно одобренным индивидуумом, и ориентированных на определенное имущество с целью его последующей эксплуатации. В рамках таких действий составляют необходимые материалы для коррекции или дополнения информации в основанном на данных реестре [2].

Исторически, кадастровое дело зародилось для нужд фискальной системы, а в девятнадцатом столетии кадастр приобрел значение механизма защиты земельных прав и надежности рыночных операций с недвижимостью. В современной форме, опираясь на геоинформационные системы, кадастровая деятельность поддерживает не только налогообложение, но и исполняет ряд ключевых функций, среди которых государственные га-

рантии, содействие в охране земельных прав, залоговое кредитование, стимулирование и анализ ситуации на рынке недвижимости, сохранение госземель, а также разрешение диспутов, связанных с земельными участками, и прочее [3].

Рождение географических информационных систем (ГИС) относится ко времени начала 60-х годов прошлого века, когда сложились благоприятные предпосылки и возникли необходимые ресурсы для осуществления компьютерного моделирования пространственных данных и задач [4]. Первопроходцы в области создания ГИС были университетами, картографическими службами, военными учреждениями и академическими институтами, основываясь на их научных исследованиях. Термин «географическая информационная система» впервые упомянут

в англоязычных источниках как система географической информации, последовательно приведя к возникновению сокращения GIS (Geographic information system). Со временем этот термин также закрепился в русскоязычной научной лексике как «географическая информационная система» либо «геоинформационная система».

Геоинформационные системы (ГИС) были признаны инструментом, который предназначен для создания, хранения, анализа и визуализации пространственных данных, обеспечивая их использование для получения новых сведений и научного познания о явлениях, обладающих пространственными координатами [5]. Этот процесс способствует более эффективному решению задач научной и практической направленности, включая вопросы касающиеся инвентаризации, исследований, моделирования, прогнозов и управления особенностями территориального распределения общества.

В наше время ГИС используется для

1. Создания и обновления кадастровых планов
2. Анализа земельной собственности
3. Управления земельными ресурсами
4. Учет и регистрация объектов недвижимости

Создание и обновления кадастровых планов является важной частью работы с земельными участками и недвижимым имуществом. Этот документ нужен для государственного контроля территории и ее владельца. Указанные в нем сведения вносятся в общую книгу регистрации недвижимости, благодаря чему земельный участок регистрируется в кадастре [6].

Анализ земельной собственности аспект управления земельными ресурсами, который помогает принимать обоснованные решения по использованию и развитию территорий [7]. Он включает в себя сбор, обработку и интерпретацию данных о земельных участках, их владельцах, характеристиках и правовом статусе. ГИС-технологии играют роль в анализе земельной собственности, позволяя интегрировать различные источники данных, проводить пространственный анализ и визуализировать результаты исследований. Например, ГИС могут использоваться для выявления зон с высоким потенциалом развития, определения наиболее подходящих мест для размещения новых

объектов и оценки влияния планируемых изменений на окружающую среду [8].

Управление земельными ресурсами представляет собой многоуровневый процесс, направленный на эффективное использования земли, сохранения её природных свойств и обеспечения устойчивого развития территорий. Для управления происходит мониторинг земель, планирование землепользования, охрана земель, регулирования рынка земли. ГИС-технологии обеспечивают возможность анализа и мониторинга. ГИС позволяют создавать цифровые карты, проводить пространственные анализ для выявления проблемных зон и т. п. [9].

Учет и регистрация объектов недвижимости являются элементом правового регулирования отношений, связанных с владением, пользованием и распоряжением недвижимыми объектами. В ГИС можно создать и ввести цифровой кадастр, взаимодействовать с базами данных, построить трехмерные модели, автоматизировать процессы [10]. Преимущество ГИС является повышенная точность, ускорение процессов, простота доступа. Во многих аспектах необходимо использование ГИС в недвижимом кадастровом процессе, так как это важно для выполнения анализа пространственных данных, которые помогают предсказать различные явления и процессы, а также важны для отслеживания изменений, происходящих с границами учетных объектов. Кроме того, это предполагает наличие тесной взаимосвязи между учетом кадастров разного рода через использование ГИС.

ГИС-технологии играют важнейшую роль в современном управлении земельными ресурсами и недвижимостью, обеспечивая высокую степень автоматизации, точности и прозрачности процессов. Они позволяют объединить разрозненные данные в единую систему, способствуя эффективному учету, анализу и планированию использования земель. Благодаря возможностям ГИС создаются цифровые карты, ведется кадастровый учет, обеспечивается регистрация прав на недвижимость, проводятся пространственные анализы и поддерживается правовая прозрачность. Все это способствует устойчивому развитию территорий, защищает интересы собственников и стимулирует экономический рост.

#### Литература:

1. Сурков, П. С. Географическая информационная система / П. С. Сурков // Туризм и рекреация: инновации и ГИС-технологии: материалы XVI Международной научно-практической конференции, Астрахань, 24–25 мая 2024 года. — Астрахань: Астраханский государственный университет имени В. Н. Татищева, 2024. — С. 30–32. — EDN LXLPTS.
2. Информационное обеспечение кадастровых работ в современных условиях / В. С. Гейдор, А. С. Калашникова, М. В. Гаранова, А. С. Иванова // Вопросы науки и образования. — 2019. — № 2(45). — С. 28–31. — EDN YUNNCP.
3. Шаромова, А. И. Формирование системы землеустройства и кадастровых работ в России / А. И. Шаромова, И. А. Глазков, Н. С. Чудинов // Технологии 2022: основные проблемы и направления развития: Сборник статей II Международной научно-практической конференции, Пенза, 30 июня 2022 года. — Пенза: Наука и Просвещение (ИП Гуляев Г. Ю.), 2022. — С. 116–119. — EDN RSSFJA.
4. Аракелян, А. А. История появления и развития кадастрового учета земель в Российской Федерации / А. А. Аракелян, Н. Р. Ачмизова // Инновации. Наука. Образование. — 2021. — № 25. — С. 1058–1062. — EDN XMHTIP.



5. Сазонов, А. А. Семинар «ГИС-технологии в науках о Земле»: история, перспективы и результаты в 2023 г / А. А. Сазонов // ГИС-технологии в науках о Земле: материалы республиканского научно-практического семинара студентов и молодых ученых, Минск, 15 ноября 2023 года. — Минск: Белорусский государственный университет, 2023. — С. 9–14. — EDN QQNHXD.
6. Донковцев, В. Г. Оценка возможностей использования сведений ЕГРН для обновления ЕЭКО / В. Г. Донковцев // Вестник МИИГАиК: Сборник научных статей 79-й научно-технической конференции студентов, аспирантов и молодых ученых в рамках Международной Студенческой Недели Науки СНН-2024, Москва, 08–12 апреля 2024 года. — Москва: Московский государственный университет геодезии и картографии, 2024. — С. 98–108. — EDN DGPOTB.
7. Петрукович, А. Н. Сравнительный анализ некоторых аспектов права частной собственности на земельный участок / А. Н. Петрукович // Сацьяльна-эканамічны і прававыя даследаванні. — 2018. — № 4(54). — С. 131–138. — EDN GZOPAC.
8. Сорокин, А. М. Анализ использования земель в Московском районе Санкт-Петербурга на основе ГИС / А. М. Сорокин, С. Р. Бакоев, М. А. Большанина // Научный электронный журнал Меридиан. — 2020. — № 9(43). — С. 439–441. — EDN JQMFGE.
9. Бытина, Т. В. Использование ГИС-технологий для управления земельными ресурсами муниципального образования / Т. В. Бытина // Форум молодых ученых. — 2018. — № 11–1(27). — С. 304–307. — EDN YWXQIP.
10. Чернова, У. Ю. Формирование трехмерного кадастра недвижимости в России / У. Ю. Чернова // Дневник науки. — 2021. — № 2(50). — С. 39. — EDN SGFZEB.

## Программный модуль для автоматизации документирования веб-сервисов

Горюнов Илья Андреевич, студент

Национальный исследовательский университет «Московский институт электронной техники» (г. Зеленоград)

*В статье рассматривается разработка программного модуля автоматизации документирования веб-сервисов. Обоснована актуальность задачи в условиях широкого применения сервис-ориентированной архитектуры и роста числа API. Проведен анализ существующих решений для документирования веб-сервисов, выявлены их недостатки. Предложен программный модуль, реализующий автоматическую генерацию документации на основе анализа сетевого трафика. Описаны концептуальная модель, архитектура и алгоритмы функционирования системы. Показано, что разработанное решение позволяет повысить актуальность документации, снизить трудозатраты разработчиков и обеспечить поддержку унаследованных систем.*

**Ключевые слова:** веб-сервисы, API, автоматизация, документирование, OpenAPI, анализ трафика, программный модуль, прокси-сервер, JSON, генерация документации.

### 1. Анализ предметной области

В современных условиях развития информационных технологий наблюдается активное распространение сервис-ориентированной архитектуры и микросервисного подхода к построению программных систем. Веб-сервисы выступают основным механизмом взаимодействия между компонентами распределённых систем, обеспечивая обмен данными и интеграцию бизнес-процессов.

Документация API представляет собой формализованное описание доступных методов, параметров запросов, форматов данных и возможных ответов. Она необходима для корректного использования сервиса как внутренними, так и внешними разработчиками.

Основные проблемы, связанные с документированием веб-сервисов:

- несоответствие документации реализации;
- высокая трудоёмкость ручного сопровождения;

- ошибки, обусловленные человеческим фактором;
- отсутствие документации для унаследованных систем.

Особенно остро эти проблемы проявляются в проектах с быстрой разработкой, где изменения в коде происходят быстрее обновления документации. Поэтому актуальна задача автоматизации документирования на основе анализа сетевого взаимодействия. Такой подход позволяет получать информацию о реальном поведении системы без вмешательства в исходный код.

### 2. Обзор существующих решений

В настоящее время существует ряд инструментов для документирования API, среди которых можно выделить следующие категории:

- инструменты ручного документирования;
- инструменты генерации документации из кода;
- инструменты анализа сетевого трафика.



К наиболее распространённым решениям относятся Swagger, Postman, ApiDog и Optic. Анализ показал, что существующие инструменты имеют следующие ограничения:

- необходимость ручного ввода данных;
- зависимость от исходного кода;
- сложность внедрения;
- ограниченная автоматизация;
- зависимость от облачных сервисов.

Таким образом, существует потребность в универсальном инструменте, обеспечивающем автоматическую генерацию документации без необходимости изменения исходного кода.

### 3. Концептуальная модель

На основе анализа предметной области и функциональных требований разработана концептуальная модель программного модуля для автоматизации документирования веб-сервисов (ПМ АДВС). Для визуализации границ системы и сценариев взаимодействия с внешним окружением использована диаграмма вариантов использования (Use Case Diagram) языка моделирования UML.

Оператор (Разработчик) — пользователь системы, который инициирует процесс работы модуля, настраивает параметры проксирования, управляет сессией записи и получает итоговый результат в виде файла спецификации в формате OpenAPI.

Концептуальная модель ПМ АДВС включает следующие ключевые сценарии:

- запуск прокси-сервера — оператор задает конфигурацию, включая локальный порт для прослушивания и адрес целевого хоста;
- просмотр статистики и логов — оператор может в реальном времени отслеживать журнал перехваченных

запросов через графический интерфейс для контроля корректности работы системы;

- генерация файла спецификации — завершающий этап автоматического цикла, на котором на основе данных структурного анализа формируются объекты схемы OpenAPI и выгружаются в виде стандартизированного файла (JSON/YAML) для дальнейшего использования;
- остановка прокси-сервера.

Проиллюстрируем это на рис. 1.

### 4. Стек технологий

В стек используемых технологий для разработки программного модуля автоматического документирования веб-сервисов (ПМ АДВС) включены современные инструменты и библиотеки, обеспечивающие корректную обработку HTTP-трафика, работу с динамическими структурами данных, автоматизацию тестирования и поддержку качества кода. Подробнее про каждый из компонентов:

- язык программирования Python версии 3.11;
- асинхронный HTTP-фреймворк aiohttp;
- библиотека валидации данных Pydantic;
- управление переменными окружения python-dotenv;
- фреймворк для тестирования Pytest;
- фреймворк для создания CLI Typer;
- HTML, CSS, JavaScript: технологии для создания веб-интерфейса;
- WebSocket: протокол, который я использую для передачи логов с сервера на клиент в режиме реального времени.

### 5. Алгоритмы функционирования

Работа сервиса начинается с инициализации и запуска HTTP-сервера по адресу `http://localhost:9001`. После за-

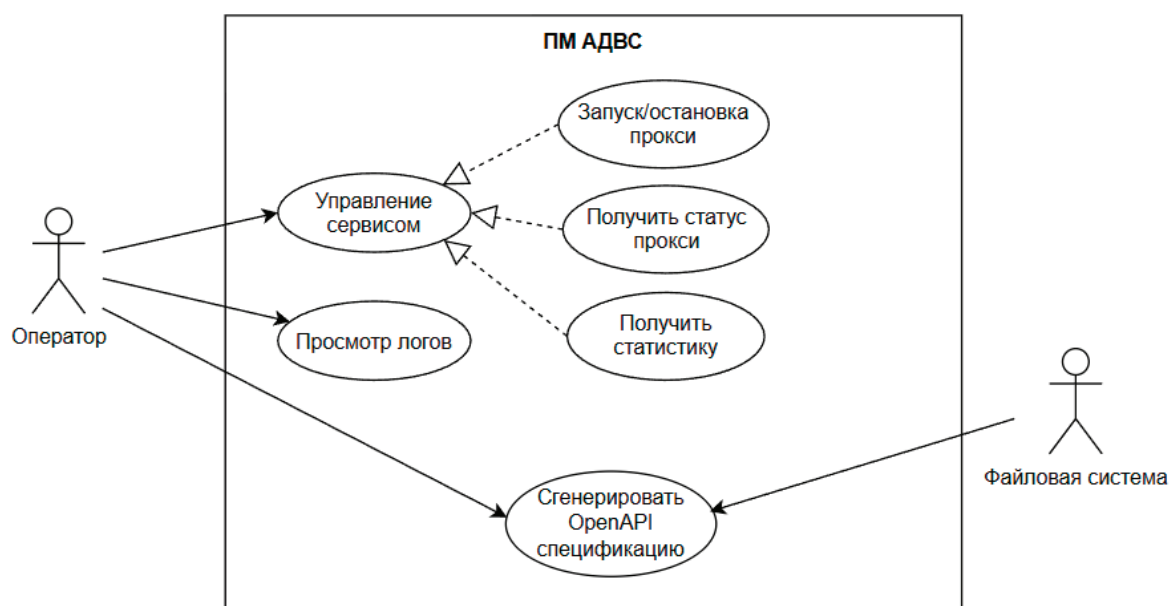


Рис. 1. Концептуальная модель ПМ АДВС

пуска сервер переходит в режим ожидания входящих запросов от клиентов. В зависимости от типа запроса возможны три сценария:

- запрос на получение пользовательского интерфейса (сервер отдаёт статические файлы и возвращается к ожиданию);
- REST API запрос (сервер принимает данные, выполняет валидацию, при корректности — обрабатывает запрос и отправляет JSON-ответ);

– установка WebSocket-соединения (сервер регистрирует клиента, контролирует состояние соединения и организует обмен сообщениями).

Система работает в непрерывном цикле обработки HTTP и WebSocket запросов, обеспечивая стандартное взаимодействие и двустороннюю коммуникацию в реальном времени.

Схема алгоритма работы ПМ АДВС представлена на рис. 3.

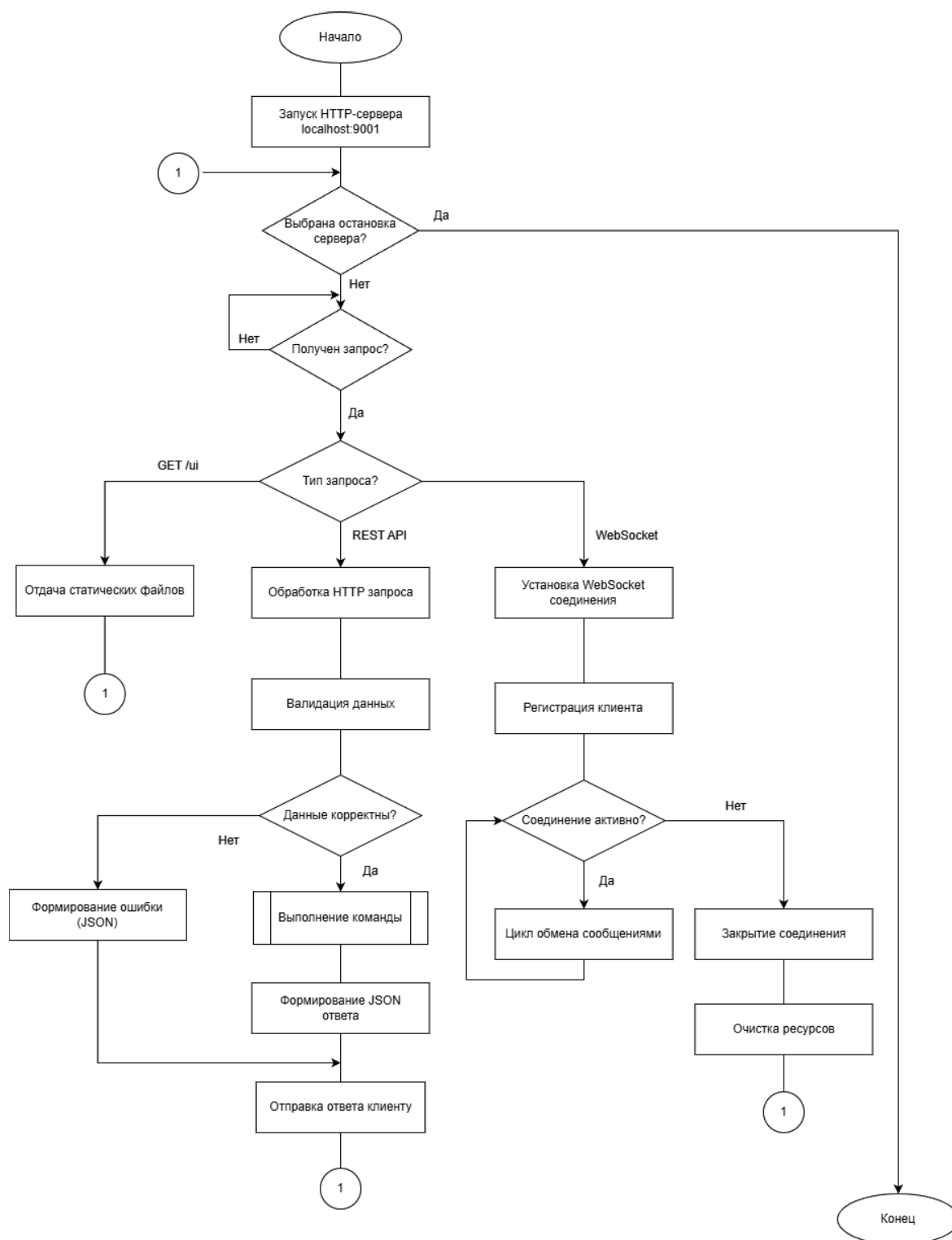


Рис. 3. Схема алгоритма работы ПМ АДВС

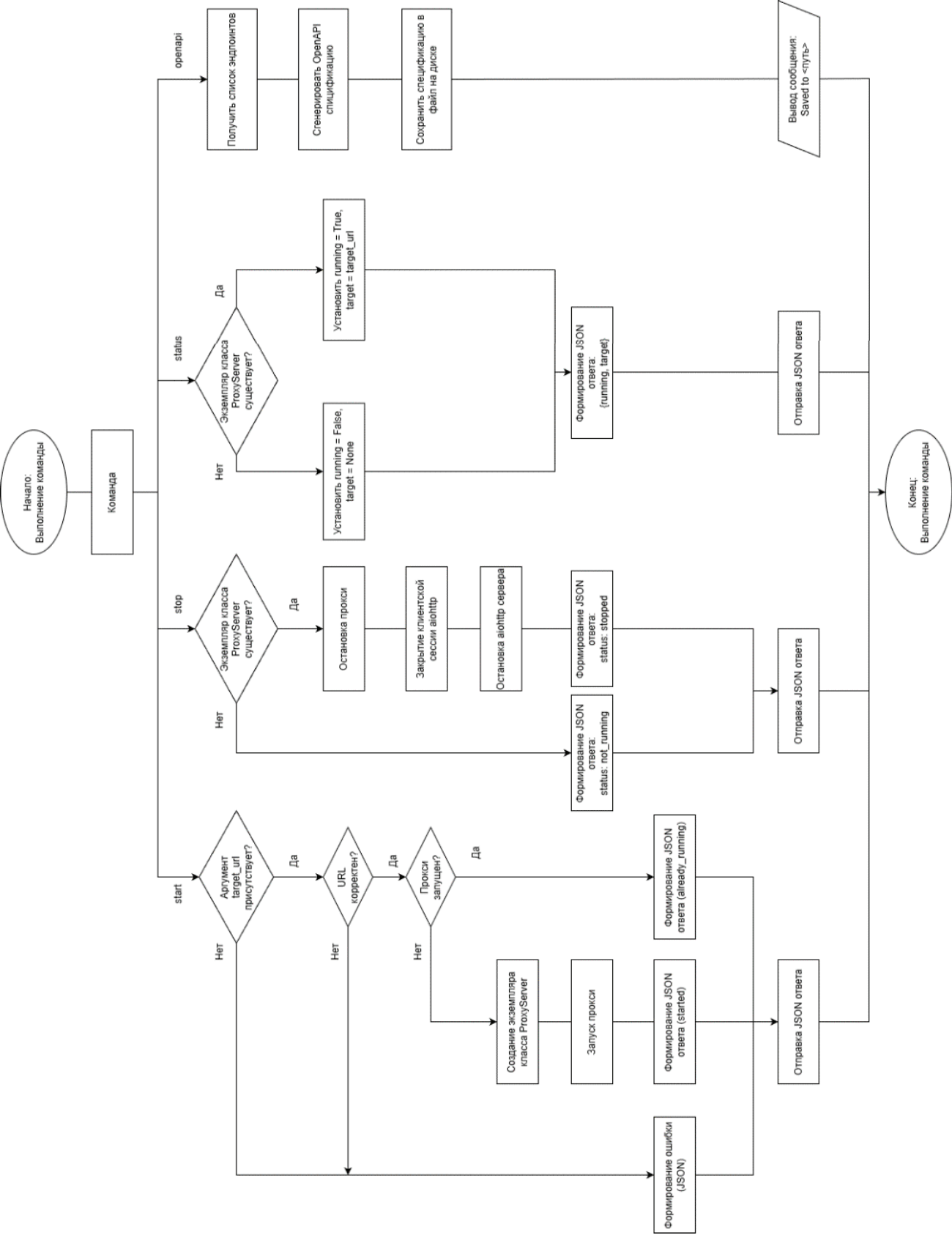


Рис. 4. Выполнение команды

Блок “Выполнение команды” подробнее рассмотрен на рис. 4.

Далее внимание уделяется алгоритмам генерации схем данных, слияния схем данных, агрегации информации, выявления параметров в URL.

Реализация указанных алгоритмов представлена в открытом репозитории на платформе GitHub: <https://github.com/ilyshks/apiscribe>.

### 5.1. Генерация схем данных

Алгоритм генерации схем данных (реализован в `Analyzer.generate_schema`, файл `analyzer.py`) преобразует произвольное JSON-значение в JSON Schema:

- определяет тип данных через функцию `isinstance`;
- рекурсивно строит схему для объектов и массивов.

### 5.2. Слияние схем данных

Слияние элементов массива через `merge_schema` позволяет корректно обобщать разнородные типы внутри одного массива (например, `[1, «str»]` → `items: {type: [«integer», «string»]}`).

### 5.3. Агрегация информации

Алгоритм накапливает статистику по каждому уникальному кортежу (`method, path`) в хранилище `InMemory-`

`Storage`. При обработке эндпоинта система извлекает поля запроса, создаёт модель `EndpointModel` и обновляет схему запроса через слияние существующих и новых данных.

### 5.4. Выявление параметров в URL (кластеризация + шаблонизация)

Пути кластеризуются по числу сегментов (`path_cluster.py`), затем анализируются посегментно (`path_inference.py`):

- одинаковые значения сегмента → константа;
- ≤3 значений и не все целые → первый сегмент как константа;
- целые числа → `{prev}_id (integer)`;
- UUID → `{prev}_uuid (string, format=uuid)`;
- остальное → `{prev}_param (string)`.

Результат: шаблон пути и параметры в формате OpenAPI (имя параметра из предыдущего константного сегмента или «param»).

## 6. Результаты

Программный модуль автоматизирует создание актуальной документации (включая поддержку унаследованных систем) без привязки к исходному коду. Экспериментально подтверждена эффективность обработки сетевого трафика и формирования спецификаций.

Перспективы: расширение функционала, повышение точности анализа, интеграция с инструментами разработки.

Литература:

1. JSON Schema Specification. Draft 202012. — URL: <https://json-schema.org/draft/2020-12/json-schema-core>
2. OpenAPI Specification (OAS), версия 3.1.0. — URL: <https://spec.openapis.org/oas/v3.1.0.html>
3. Язык программирования Python. — URL: <https://www.python.org/doc/>
4. Richardson, L., Ruby, S., Taylor, J. RESTful Web API. — O'Reilly Media, 2013. — 400 с.
5. Newcomer, E., Lomow, G. Understanding SOA with Web Services. — Addison-Wesley Professional, 2004. — 528 с.

## Современные подходы к обнаружению атак на механизмы доменной аутентификации в корпоративных информационных системах

Камышник Артем Васильевич, студент магистратуры;  
Елемесов Шынгыс Бекарысулы, студент магистратуры;  
Баймуханов Жанторе Ербакытулы, студент магистратуры  
Евразийский национальный университет имени Л. Н. Гумилева (г. Астана, Казахстан)

В этом исследовании рассматривается критическая проблема обнаружения несанкционированного доступа в рамках централизованных механизмов аутентификации домена, особенно в средах Microsoft Active Directory. Цель исследования — оценить эффективность современных подходов к мониторингу, включая обнаружение на основе сигнатур, поведенческое профилирование и корреляционный анализ событий. Был проведен систематический сравнительный анализ данных телеметрии, генерируемых протоколами Kerberos и NTLM, для выявления признаков компрометации, связанных с кражей учетных данных и перемещением в другое место. Результаты показывают, что, хотя методы, основанные на сигнатурах, требуют небольших вычислительных затрат, они не позволяют обнаружить новые векторы атак. И наоборот, гибридные модели, сочетающие статистическое обнаружение аномалий с корреля-

цией событий из нескольких источников, обеспечивают значительно более высокую частоту обнаружения атак, основанных на идентификации. Научная новизна этой работы заключается в предложенной концептуальной архитектуре для объединения различных логов аутентификации в единую поведенческую модель. На практике полученные результаты предлагают основу для оптимизации конфигураций управления информацией о безопасности и событиями (SIEM), чтобы уменьшить количество ложных срабатываний и повысить видимость неправомерного использования привилегированных учетных записей.

**Ключевые слова:** аутентификация домена, мониторинг кибербезопасности, обнаружение аномалий, SIEM-системы, протоколы аутентификации, внутренние угрозы, боковое перемещение, Kerberos.

## 1. Введение

Централизованные механизмы аутентификации представляют собой важнейший компонент современных корпоративных информационных систем, обеспечивающий контроль доступа пользователей к сетевым ресурсам и сохраняющий целостность корпоративной инфраструктуры. В корпоративных средах процессы аутентификации обычно реализуются с помощью служб каталогов, которые поддерживают протоколы Kerberos и NTLM для проверки личности.

Из-за своей фундаментальной роли в обеспечении контроля доступа эти механизмы все чаще становятся основными мишенями для враждебной деятельности, направленной на повышение привилегий и несанкционированное перемещение в другую сторону. Недавние исследования в области кибербезопасности показывают, что атаки, нацеленные на протоколы аутентификации, часто используют законную функциональность системы, а не приводят к заметным аномалиям на уровне сетевых коммуникаций. Следовательно, традиционные механизмы защиты на основе периметра часто не позволяют обнаружить неправильное использование учетных данных в рамках авторизованных рабочих процессов проверки подлинности [1–2].

Эта проблема требует внедрения архитектур мониторинга, ориентированных на идентификацию, способных анализировать телеметрию аутентификации, генерируемую контроллерами домена. Целью данного исследования является изучение современных подходов к обнаружению атак на механизмы аутентификации домена на основе мониторинга событий аутентификации и поведенческого анализа в инфраструктурах кибербезопасности предприятия

## 2. Материал и методы исследования

Настоящее исследование основано на систематическом аналитическом обзоре современных подходов к обнаружению кибератак, нацеленных на централизованные механизмы доменной аутентификации в корпоративных информационных системах. Методология исследования основана на изучении и сравнительном анализе методов мониторинга аутентификации, реализованных в инфраструктурах кибербезопасности предприятия, которые используют централизованные модели управления идентификационными данными.

В современных корпоративных средах процессы аутентификации обычно поддерживаются службами каталогов, такими как Microsoft Active Directory. Эти рабочие процессы генерируют значительный объем телеметрии, относящейся к безопасности, которая может быть использована для обнаружения атак без необходимости прямой проверки полезного сетевого трафика. События безопасности, связанные с аутентификацией, генерируемые контроллерами домена, рассматриваются как основной источник поведенческих индикаторов, связанных с неправильным использованием учетных данных и попытками несанкционированного повышения привилегий. Методологическая основа включает классификацию методов злоупотребления протоколами аутентификации и аналитическую оценку стратегий корреляции событий, используемых в системах управления информацией о безопасности и событиях (SIEM).

## 3. Результаты и обсуждение

Эффективность обнаружения современных атак на аутентификацию во многом зависит от внедрения централизованных архитектур мониторинга, способных агрегировать и сопоставлять события, связанные с аутентификацией, генерируемые в распределенных корпоративных средах. В отличие от традиционных систем обнаружения вторжений, которые полагаются на проверку пакетов на сетевом уровне, системы мониторинга аутентификации работают на уровне идентификации, анализируя транзакции контроля доступа, записанные в журналах аудита контроллера домена.

Такой архитектурный подход позволяет аналитикам по кибербезопасности выявлять враждебную активность, которая использует законные рабочие процессы аутентификации, не внося заметных аномалий в схемы сетевого взаимодействия. Следовательно, атаки, нацеленные на механизмы аутентификации в домене, часто остаются незамеченными системами защиты на основе периметра, при этом генерируя статистически значимые отклонения в телеметрии аутентификации.



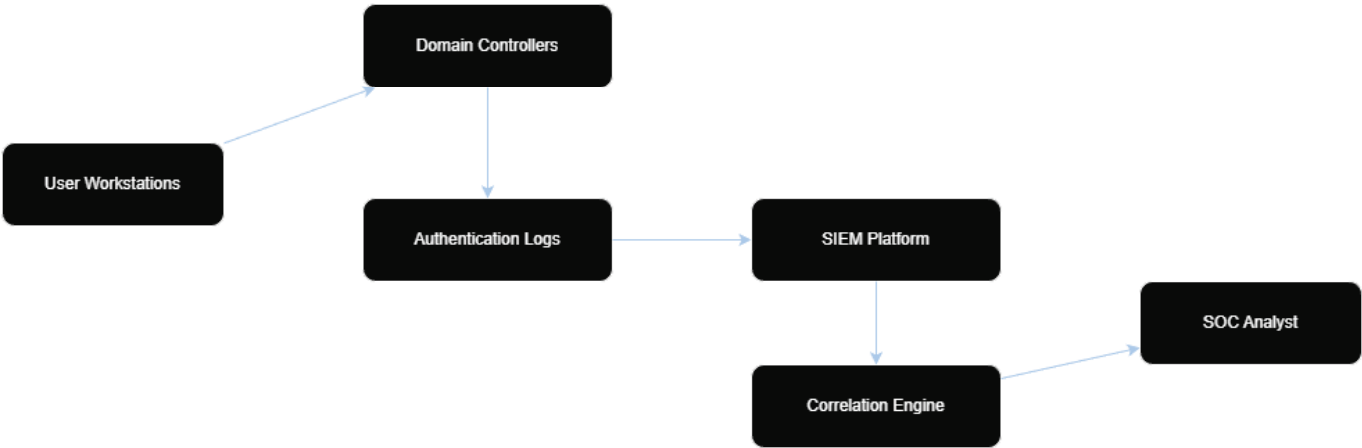


Рис. 1. Концептуальная архитектура централизованного мониторинга аутентификации в корпоративных средах

Б. Сравнительный анализ подходов к обнаружению

Таблица 1. Подходы к обнаружению нарушений протокола аутентификации

Метод обнаружения	Аналитическая база	Преимущества	Ограничения
Сигнатурный анализ	Известные паттерны атак	Низкая вычислительная сложность	Неэффективен против новых (неизвестных) атак
Поведенческий анализ	Профилирование активности пользователей	Обнаружение ранее неизвестных угроз	Требует построения базовой модели (baseline)
Корреляция событий	Анализ логов из нескольких источников	Контекстно-зависимое обнаружение	Повышенные накладные расходы на обработку данных
Статистическое обнаружение аномалий	Метрики аутентификации	Выявление незначительных отклонений	Склонность к ложноположительным срабатываниям

Сравнительная оценка стратегий обнаружения показывает, что методы мониторинга аутентификации на основе сигнатур по-прежнему ограничены в своих возможностях по выявлению сложных атак со злоупотреблением учетными данными из-за их зависимости от predetermined сигнатур атак [3]. Поскольку современные злоумышленники все чаще используют законные механизмы аутентификации для осуществления горизонтального перемещения в скомпрометированных средах, такие подходы по своей сути неспособны отличить вредоносную активность от поведения авторизованного пользователя.

Методы обнаружения, основанные на поведении, устраняют это ограничение, устанавливая базовые модели обычной аутентификации и выявляя статистически значимые отклонения от ожидаемых моделей доступа. Эти отклонения могут включать ненормальную частоту аутентификации, несогласованную последовательность использования привилегий или нерегулярные интервалы выдачи билетов на обслуживание. Однако эффективность поведенческого анализа напрямую зависит от точности базового моделирования и стабильности схем доступа пользователей в контролируемой среде.

Методы корреляции событий, реализованные на платформах SIEM, еще больше расширяют возможности обнаружения за счет объединения событий аутентификации, исходящих из нескольких источников, и восстановления сложных последовательностей атак, которые в противном случае остались бы необнаружимыми при изолированном анализе. Анализируя временные и логические взаимосвязи между транзакциями аутентификации, механизмы корреляции позволяют выявлять скоординированные попытки доступа, указывающие на неправильное использование учетных данных [4].

Подходы к статистическому обнаружению аномалий расширяют эту парадигму, применяя вероятностные модели к телеметрии аутентификации для количественной оценки отклонений от нормативного распределения активности. Хотя такие методы демонстрируют повышенную чувствительность к ранее неизвестным векторам атак, они часто связаны с повышенным количеством ложных срабатываний в динамичных корпоративных средах, характеризующихся изменчивыми моделями поведения пользователей.



Рис. 2. Концептуальная модель обнаружения аномалий при аутентификации домена

Аналитическая оценка современных методов мониторинга аутентификации показывает, что гибридные модели обнаружения, сочетающие поведенческий анализ с корреляцией событий из нескольких источников, обеспечивают высочайший уровень эффективности при выявлении нарушений протокола аутентификации [5]. Такие подходы позволяют обнаруживать вредоносные действия, выполняемые в рамках законных рабочих процессов аутентификации, не полагаясь на механизмы проверки на сетевом уровне.

В то же время внедрение моделей обнаружения на основе аномалий создает дополнительную вычислительную сложность и требует постоянной адаптации базовых профилей деятельности в ответ на организационные изменения. В результате в современных стратегиях кибербезопасности предприятий все большее внимание уделяется интеграции систем мониторинга, ориентированных на идентификацию, способных динамически корректировать пороговые значения обнаружения на основе контекстуальных индикаторов риска аутентификации [6].

#### 4. Заключение

Аналитическая оценка современных методов мониторинга аутентификации показывает, что гибридные модели обнаружения, сочетающие поведенческий анализ с корреляцией событий из нескольких источников, обеспечивают высочайший уровень эффективности при выявлении нарушений протокола аутентификации.

Внедрение систем мониторинга, ориентированных на идентификацию, позволяет корпоративным инфраструктурам кибербезопасности обнаруживать вредоносные действия, выполняемые в рамках законных рабочих процессов аутентификации, не полагаясь на механизмы проверки на сетевом уровне. Исследование выполнило свои задачи, обозначив четкий путь перехода от устаревшего мониторинга к поведенческой аналитике.

Будущие исследования должны быть сосредоточены на автоматизации процесса генерации исходных данных с использованием машинного обучения, чтобы снизить количество ложноположительных результатов, присущих обнаружению статистических аномалий.

#### Литература:

1. Singer, P.W., & Friedman, A. (2014). *Cybersecurity and Cyberwar: What Everyone Needs to Know*. Oxford University Press.
2. Scarfone, K., & Mell, P. (2007). *Guide to Intrusion Detection and Prevention Systems (IDPS)*. NIST Special Publication 800–94. National Institute of Standards and Technology.
3. Microsoft. (2020). *Advanced Threat Analytics Overview*. Microsoft Security Documentation.
4. NIST. (2020). *Zero Trust Architecture*. NIST Special Publication 800–207. National Institute of Standards and Technology.
5. Gartner. (2021). *SIEM Market Guide*. Gartner Research.
6. Shackleford, D. (2016). *User Behavior Analytics*. SANS Institute. Information about authors

## Сравнительный анализ методологий выбора моделей в предиктивной аналитике: Problem-Driven vs Data-Driven vs гибридный подход

Климовец Татьяна Васильевна, студент магистратуры  
Тольяттинский государственный университет (Самарская область)

*В статье проводится критический анализ доминирующих методологий выбора моделей предиктивной аналитики: проблемно-ориентированного (Problem-Driven), основанного на данных (Data-Driven) и гибридного подходов. Выявляются их философские основы, этапы реализации, сильные и слабые стороны, а также области эффективного применения. Анализ показывает, что выбор методологии не является универсальным и должен определяться спецификой задачи, требованиями к интерпретируемости, качеством и объёмом данных, а также зрелостью организации. Делается вывод о том, что гибридные методики, синтезирующие экспертные знания и возможности анализа данных, представляют собой наиболее сбалансированное и перспективное направление для ответственных приложений, где необходимы как высокая точность, так и понимание логики принятия решений.*

**Ключевые слова:** предиктивная аналитика, выбор модели, проблемно-ориентированный подход, Data-Driven подход, CRISP-DM, машинное обучение, интерпретируемость, гибридные методики.

### Введение

Успешное внедрение предиктивной аналитики в таких областях, как управление бизнесом, финансы и логистика, в решающей степени зависит от корректного выбора прогностической модели. Однако существующее многообразие алгоритмов — от классической статистики до сложных ансамблей машинного обучения — превращает этот выбор в сложную методологическую проблему. На сегодняшний день сформировались два основных, зачастую противопоставляемых друг другу подхода: проблемно-ориентированный (Problem-Driven) и основанный на данных (Data-Driven). Каждый из них обладает как преимуществами, так и ограничениями. В ответ на эти ограничения возникли гибридные методики, стремящиеся к синтезу этих двух подходов. Целью данной статьи является сравнение этих трёх парадигм, определение границ их применимости и выявление тенденций развития методологии выбора моделей.

### 1. Проблемно-ориентированный (Problem-Driven) подход

Problem-Driven подход исходит из приоритета чётко сформулированной бизнес-проблемы. Его отправной точкой является вопрос «Что нужно предсказать и зачем?». Классической методологической реализацией служит стандарт CRISP-DM (Cross-Industry Standard Process for Data Mining), который предполагает линейную последовательность этапов: понимание бизнеса, понимание данных, подготовка данных, моделирование, оценка, внедрение. В центре внимания — глубокое понимание предметной области, построение причинно-следственных гипотез и использование интерпретируемых моделей (линейная и логистическая регрессия, деревья решений [2, 7]).

Сильные стороны данного подхода прослеживаются в высокой интерпретируемости и обоснованности, т. к. результаты логически связаны с экспертной логикой, что

критично в регулируемых отраслях (финансы, медицина); в фокусе на бизнес-результате, т. к. подход обеспечивает тесную связь с конкретными целями заказчика и эффективное использование ресурсов; в устойчивости при ограниченных данных, т. к. модели, основанные на теории, часто требуют меньшего объёма данных для достижения адекватной точности; в структурированности и прозрачности, т. к. чёткий план понятен как техническим специалистам, так и бизнес-пользователям.

Слабые стороны и риски Problem-Driven подхода мы видим в ограниченности исследования предварительно заданными гипотезами, поэтому существует риск упустить сложные, неочевидные, но значимые паттерны, не описанные в существующих теориях; прослеживается зависимость от экспертизы, т. е. эффективность подхода напрямую зависит от глубины и объективности экспертных знаний; отмечается низкая адаптивность, подход требует ясности на старте и плохо адаптируется к быстрым изменениям в данных или появлению новых, неизвестных факторов; велик риск создания идеализированной модели, которая может оказаться слишком упрощённой и мало соответствовать реальной сложности данных.

### 2. Основанный на данных (Data-Driven) подход

В основе Data-Driven подхода лежит убеждение, что данные, особенно в больших объёмах (Big Data), являются первичным источником истины и инсайтов. Цель смещается с проверки сформулированных гипотез к их обнаружению/формулированию. Процесс имеет итеративный, циклический характер (данные — инсайты — действие — новые данные) и фокусируется на использовании сложных алгоритмов машинного обучения (градиентный бустинг [8], нейронные сети [9]), способных выявлять нелинейные зависимости, часто в ущерб интерпретируемости модели [6].

Сильные стороны Data-Driven подхода выражаются в более высокой точности по сравнению с проблемно-

ориентированным подходом, в способности обнаруживать сложные, контринтуитивные паттерны в больших и многомерных данных; в адаптивности подхода, когда при регулярном переобучении модели могут адаптироваться к изменяющимся условиям; в наличии возможности для автоматического принятия решений и персонализации в реальном времени; в выявлении ранее неизвестных взаимосвязей, расширении границ предметного знания.

Слабые стороны и риски этого подхода мы видим в низкой интерпретируемости, решения сложных моделей трудно объяснить, что неприемлемо в ответственных или регулируемых сферах; в высоких требованиях к инфраструктуре, т. к. требует больших вычислительных ресурсов, качественных и репрезентативных данных в огромных объёмах; существует риск ложных корреляций и смещений, модель может уловить статистические артефакты или усилить системные смещения, присутствующие в исторических данных, приводя к проблемным решениям; а также есть риск построения модели, которая хорошо работает на исторических данных, но плохо интерполируется на новые.

### 3. Гибридные методики

Гибридные методики представляют собой синтез двух предыдущих подходов. Их основная идея заключается в том, что ни теория без эмпирической проверки, ни данные без смысловой интерпретации не являются достаточными для построения надёжных моделей. Наблюдаются три ключевых принципа интеграции:

1) Принцип смыслового каркаса, когда экспертные знания задают начальные гипотезы, выбор переменных и границы модели, сужая пространство поиска для ML-алгоритмов. Применяется, например, в статье Тематическое моделирование с использованием латентного распределения Дирихле (LDA), Кевина Киббе [4] или в статье Латентное распределение Дирихле (LDA) и тематическое моделирование: модели, приложения, опрос Хамеда Джеллодара с соавторами [5].

2) Принцип итеративной валидации, заключается в том, что каждая итерация включает как статистическую проверку точности, так и смысловую валидацию экспертами на предмет логической непротиворечивости [9].

3) Принцип каузальной интерпретируемости, когда устанавливается акцент на построении изначально интерпретируемых архитектур или использовании данных для проверки причинных гипотез [10].

В литературе встречаются следующие конкретные техники реализации:

1) Последовательная интеграция, при которой Data-Driven методы (кластеризация, NLP) используются для разведочного анализа и генерации признаков, на основе которых строится интерпретируемая Problem-Driven модель [5].

2) Ансамблевое моделирование, при котором происходит комбинирование прогнозов интерпретируемой (например, регрессия) и сложной ML-модели для повышения точности и устойчивости [3].

3) Структурное внедрение знаний, при котором происходит инкорпорация экспертных правил или физических законов непосредственно в функцию потерь или архитектуру нейронной сети (Physics-Informed Neural Networks) [1].

Сильные стороны гибридного подхода проявляются в балансе точности и интерпретируемости, потому что он позволяет достичь высокой прогностической силы, сохраняя возможность объяснения ключевых факторов; в повышенной устойчивости и надёжности, т. к. модели опираются как на глубинные законы, так и на актуальные эмпирические паттерны; в снижении рисков, минимизации опасности ложных корреляций и этических конфликтов, присущих «чёрным ящикам»; в синергии знаний, когда данные обогащают теорию, а теория направляет сбор и анализ данных.

Слабые стороны заключаются в высокой сложности реализации, так как требуется создание междисциплинарных команд (data scientists, domain experts) и налаженной коммуникации; в более длительном и, следовательно, в более дорогом процессе разработки; в риске механического соединения компонентов без глубокой интеграции, которое не даст синергетического эффекта.

Сравнение подходов по ключевым критериям представлено в Таблице 1.

### Заключение

Проведённый анализ показывает, что Problem-Driven и Data-Driven подходы представляют собой не столько конкурирующие, сколько комплементарные парадигмы. Problem-Driven подход незаменим там, где критически важны прозрачность, соответствие регуляторным нормам и проверка теоретических построений. Data-Driven подход демонстрирует свою мощь в задачах с экстремальным объёмом и сложностью данных, где точность прогноза превалирует над объяснимостью.

Гибридные методики, устраняя крайности каждого из «чистых» подходов, представляют собой эволюционное развитие методологии предиктивной аналитики. Они трансформируют процесс из выбора между теорией и данными в процесс созидательного диалога между ними. В условиях растущей сложности бизнес-среды и ужесточения требований к объяснимости алгоритмов, гибридный подход становится необходимым для создания ответственных, надёжных и эффективных систем поддержки принятия решений. Таким образом, будущее предиктивного моделирования лежит в развитии методологий, обеспечивающих органичный и управляемый симбиоз экспертного знания и мощи data-driven анализа.

Таблица 1. Сравнительная таблица методологий выбора моделей предиктивной аналитики

Критерий	Problem-Driven подход	Data-Driven подход	Гибридные методики
Отправная точка	Чёткая бизнес-проблема	Наличие и потенциал данных	Бизнес-проблема, обогащённая анализом данных
Основа для построения модели	Экспертные знания, теория, гипотезы	Выявленные паттерны и корреляции в данных	Синтез экспертных знаний и выявленных паттернов
Процесс	Структурированный, последовательный (CRISP-DM)	Итеративный, циклический, исследовательский	Комбинированный, управляемый цикл с обратной связью
Ключевые методы	Классическая статистика, интерпретируемые алгоритмы	Машинное обучение (ансамбли, нейросети)	Комбинация статистических и ML-методов, ансамбли
Интерпретируемость результата	Высокая	Низкая («чёрный ящик»)	Умеренная/Высокая (зависит от архитектуры)
Требования к данным	Качественные, релевантные гипотезе	Большие объёмы (Big Data), разнообразие источников	Качественные данные достаточного объёма
Главное преимущество	Логическая обоснованность, доверие экспертов	Высокая точность на сложных задачах	Сбалансированность точности и понимания
Основной риск	Пропуск скрытых зависимостей	Ложные корреляции, этические проблемы	Сложность проектирования и реализации
Области эффективного применения	Регулируемые отрасли (финансы, медицина), проверка гипотез, задачи с чёткими бизнес-целями	Компьютерное зрение, NLP, персонализация, анализ больших массивов неструктурированных данных	Ответственные приложения, сложные инженерные и экономические системы, управление рисками

#### Литература:

1. Geometric Deep Learning Grids, Groups, Graphs, Geodesics, and Gauges / M. B. Michael. — Текст: электронный — URL: <https://arxiv.org/pdf/2104.13478> (дата обращения: 20.02.2026).
2. Georgios, Gelastopoulos A systematic review of text mining analytics for supply chain risk management using online data / Gelastopoulos Georgios. — Текст: электронный — URL: [https://www.researchgate.net/publication/397256732\\_A\\_systematic\\_review\\_of\\_text\\_mining\\_analytics\\_for\\_supply\\_chain\\_risk\\_management\\_using\\_online\\_data](https://www.researchgate.net/publication/397256732_A_systematic_review_of_text_mining_analytics_for_supply_chain_risk_management_using_online_data) (дата обращения: 19.02.2026).
3. How Ensemble Modeling Helps to Avoid Overfitting. — Текст: электронный — URL: <https://www.geeksforgeeks.org/machine-learning/how-ensemble-modeling-helps-to-avoid-overfitting/> (дата обращения: 03.03.2026).
4. Kevin, Kibe Topic Modeling Using Latent Dirichlet Allocation (LDA) / Kibe Kevin. — Текст: электронный — URL: <https://www.analyticsvidhya.com/blog/2023/02/topic-modeling-using-latent-dirichlet-allocation-lda/> (дата обращения: 15.03.2026).
5. Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey / Jelodar Hamed. — Текст: электронный — URL: <https://link.springer.com/article/10.1007/s11042-018-6894-4> (дата обращения: 03.03.2026).
6. Topic Modeling Using Latent Dirichlet Allocation (LDA). — Текст: электронный — URL: <https://www.geeksforgeeks.org/nlp/topic-modeling-using-latent-dirichlet-allocation-lda/?ysclid=mnbiexnv3k4393135> (дата обращения: 15.02.2026).
7. Алябьева, М. В. Оптимизация управления товарными запасами в оптовой торговле: специальность 08.00.05 «Экономика и управление народным хозяйством»: диссертация на соискание ученой степени кандидата экономических наук / Алябьева Марианна Владимировна. — Белгород, 1997. — 175 с. — Текст: непосредственный.
8. Котов, А. Gradient Boosting: принципы работы и применение в машинном обучении / Артём Котов. — Текст: электронный — URL: <https://sky.pro/wiki/analytics/gradient-boosting-printsipy-raboty-i-primenenie-v-mashin-nom-obuchenii/?ysclid=mi040gjw2p605258620> (дата обращения: 20.03.2026).
9. Могильников, И. Валидация моделей машинного обучения / Илья Могильников. — Текст: электронный — URL: <https://habr.com/ru/companies/glowbyte/articles/569970/> (дата обращения: 29.01.2026).
10. Юрашку, И. Causal Inference: прозрение и практика. Лекция 1. Основные понятия Causal Inference / Иван Юрашку. — Текст: электронный — URL: <https://habr.com/ru/companies/sberbank/articles/847382/> (дата обращения: 02.03.2026).



## Современные решения в области СКУД: модель выбора архитектурного профиля по технологическим и эксплуатационным критериям

Колесин Артём Дмитриевич, студент

Национальный исследовательский университет «Высшая школа экономики», Нижегородский филиал

*В статье предложена авторская модель выбора архитектурного профиля системы контроля и управления доступом (СКУД) для объектов с различным режимом эксплуатации. В отличие от распространенного обзорного подхода, при котором RFID, биометрические, облачные и гибридные решения сравниваются как однопорядковые категории, в работе разграничены уровни идентификационного фактора и архитектуры системы. На основе анализа нормативных актов Российской Федерации, материалов по оценке алгоритмов распознавания лиц, а также публикаций по RFID- и биометрическим системам разработана сравнительная матрица из восьми критериев: устойчивость к передаче учетных данных, пропускная способность, стабильность в неидеальных условиях, инфраструктурная сложность, масштабируемость, интегрируемость, правовая нагрузка и отказоустойчивость. Применение модели к трем типовым сценариям (учебный корпус, режимная лаборатория и распределенная сеть объектов) показало, что для массового доступа наиболее рационален токеновый RFID-/карточный профиль, тогда как для режимных зон и распределенной инфраструктуры наилучшие результаты демонстрирует гибридный многофакторный профиль. Практическая ценность исследования заключается в том, что разработанные автором артефакты могут использоваться как инструмент предварительного выбора СКУД на этапе проектирования и модернизации объекта.*

**Ключевые слова:** СКУД, контроль доступа, RFID, биометрия, облачная архитектура, гибридная аутентификация, эксплуатационные критерии, выбор архитектуры.

### Введение

Системы контроля и управления доступом сегодня рассматриваются не как изолированный элемент входной группы, а как часть общей архитектуры безопасности объекта. Нормативную основу данной области в России задает ГОСТ Р 51241–2008, а с 1 февраля 2025 г. введено в действие изменение № 1 к данному стандарту. Это означает, что разговор о современных СКУД неизбежно выходит за пределы описания отдельных считывателей и замков и затрагивает вопросы классификации, испытаний, надежности, совместимости и условий эксплуатации [1; 2].

К тому же, не менее важен и правовой контур. Базовые требования к обработке персональных данных закреплены в Федеральном законе № 152-ФЗ, специальный порядок идентификации и аутентификации с использованием биометрических персональных данных — в Федеральном законе № 572-ФЗ, а вопросы локализации баз данных российских граждан получили развитие после принятия Федерального закона № 242-ФЗ. В разъяснениях Роскомнадзора также указывается, что фото- и видеоизображения могут относиться к биометрическим персональным данным, если используются именно для установления личности [3–6]. Для проектирования СКУД это означает простую, но принципиальную вещь: выбор технологии всегда связан не только с удобством и безопасностью, но и с режимом обработки данных.

Вместе с тем в учебных и прикладных публикациях часто встречается одна и та же методическая неточность: RFID, биометрия, облачные и гибридные решения сопоставляются как равные “типы СКУД”. На практике это не вполне корректно, поскольку RFID и биометрия описывают прежде всего **способ идентификации**, тогда как облачность и гибридность характеризуют **архитектурный профиль системы**. Смешение уровней делает сравнение менее строгим и затрудняет практический выбор решения [7–11]. Дополнительную сложность создает то, что одна и та же технология по-разному ведет себя в разных условиях эксплуатации. Например, для потокового объекта решающей может оказаться пропускная способность, а для режимной зоны устойчивость к передаче учетных данных и наличие резервного сценария допуска. Таким образом, неверный выбор параметров и устройств доступа способен приводить к простоям и потерям, а материалы по распознаванию лиц демонстрируют заметную чувствительность точности к качеству изображения и к частичному перекрытию лица маской.

Исходя из этого, исследовательская проблема формулируется так, что **в литературе и проектной практике недостаточно формализованы критерии выбора архитектурного профиля СКУД в зависимости от режима доступа, эксплуатационной нагрузки и правовых ограничений**. Цель работы — разработать компактную и воспроизводимую модель выбора архитектурного профиля СКУД по совокупности технологических и эксплуатационных критериев. Гипотеза исследования состоит в том, что для объектов со средними и высокими требованиями к безопасности наилучший результат обеспечивает не “наиболее современная” технология сама по себе, а профиль СКУД, в котором способ идентификации согласован с архитектурой управления, резервным сценарием допуска и правовым режимом обработки данных.

Исследование выполнено в формате обзорно-аналитического сравнения с элементами формализованного сценарного оценивания. В корпус источников включены: нормативные акты и официальные правовые публикации Российской Федерации, разъяснения Роскомнадзора, официальные материалы по программам FRVT и FRTE, рецензируемые статьи по пропускной способности RFID-систем, безопасности RFID и точности биометрической идентификации.

В работе используется двухуровневый подход. На первом уровне выделяются

**1) идентификационный фактор** — карта/RFID/мобильный идентификатор, PIN-код, биометрический признак, комбинация факторов;

**2) архитектура системы** — локальная/серверная, облачная, гибридная.

На втором уровне сравниваются уже не типы, а четыре распространенных **архитектурных профиля**:

**P1** — токеничная RFID-/карточная локальная или серверная СКУД;

**P2** — биометрическая локальная или серверная СКУД;

**P3** — облачная СКУД с централизованным управлением;

**P4** — гибридная многофакторная СКУД.

Для оценки были выбраны восемь критериев

**C1** — устойчивость к передаче или клонированию учетных данных;

**C2** — пропускная способность в пиковом режиме;

**C3** — стабильность работы в неидеальных условиях;

**C4** — инфраструктурная простота внедрения;

**C5** — масштабируемость;

**C6** — интеграция с другими подсистемами;

**C7** — правовая и организационная нагрузка;

**C8** — отказоустойчивость и наличие резервного сценария.

Интегральный показатель рассчитывается по формуле  $I = \sum (w_i \times b_i) \times 20$ , где  $w_i$  — вес критерия в конкретном сценарии,  $b_i$  — базовый балл профиля по данному критерию. Умножение на 20 переводит результат в шкалу от 0 до 100. Полученные значения не являются нормативным стандартом; это авторский аналитический инструмент для предварительного выбора решения до этапа детального проектного обследования.

## 1. Двухуровневая модель классификации СКУД

Прежде всего, главный тезис работы состоит в том, что СКУД целесообразно описывать сразу по двум осям: по идентификационному фактору и по архитектуре, что, собственно, позволит снять распространенную для обзорных текстов путаницу, когда карта, биометрия, облако и гибридность оказываются в одном ряду, хотя фактически отвечают на разные проектные вопросы [7–11].

Таблица 1. Двухуровневая модель описания решений СКУД

Уровень анализа	Категории	Практический смысл
Идентификационный фактор	карта / RFID / мобильный идентификатор	быстрый проход, умеренная стоимость, риск передачи носителя
	PIN / код	низкая стоимость, слабая персонализация
	биометрия	высокая привязка к личности, повышенные требования к среде и правовому режиму
	комбинация факторов	рост устойчивости, усложнение эксплуатации
Архитектура	локальная / серверная	меньшая зависимость от внешних каналов, предсказуемость
	облачная	централизованное управление, удобство для разрозненных сетей
	гибридная	сочетание локальной устойчивости и централизованной координации

Иначе говоря, “облачная биометрическая СКУД” и “гибридная RFID-СКУД” — это не конкурирующие сущности одного порядка, а комбинации двух независимых проектных решений. Такая рамка делает дальнейшее сравнение методически чище и практически полезнее.

## 2. Обоснование шкалы оценки, весовых коэффициентов и базовых баллов

Для того, чтобы сравнительная матрица не выглядела как набор произвольных оценок, в работе заранее задана логика как **взвешивания критериев**, так и **присвоения баллов**.

2.1. Логика весовых коэффициентов

Вес критерия определялся по принципу **функционального доминирования** в конкретном сценарии эксплуатации. Так, если без соблюдения критерия система фактически теряет практическую пригодность, критерию присваивался высокий вес в диапазоне **0,20–0,25**. С другой стороны, если критерий существенно влияет на качество функционирования, но не определяет саму возможность эксплуатации, использовался средний вес **0,10–0,18**. И в завершении, если критерий обязателен, но не является основным фактором выбора именно в данном сценарии, ему присваивался пониженный вес **0,05–0,09**. Сумма весов в каждом сценарии равна 1,0. Исходя из вышеперечисленного, подход опирается на содержание самих сценариев. Для учебного корпуса или типового офисного объекта главными становятся пропускная способность и инфраструктурная простота, так как именно они определяют отсутствие очередей и приемлемую стоимость развертывания. Для режимной зоны первостепенное значение имеют устойчивость к передаче учетных данных, стабильность и отказоустойчивость, так как цена ошибочного допуска или блокировки прохода здесь выше. Для самого сложного по архитектуре, распределенной сети объектов, решающими оказываются масштабируемость, интеграция и наличие резервного режима при сбоях связи [3–10].

2.2. Логика баллов

Учитывая это, базовые баллы по каждому профилю присваивались по шкале:

- **1 балл** — профиль систематически проигрывает по критерию;
- **2 балла** — имеются выраженные ограничения;
- **3 балла** — приемлемый, но не доминирующий уровень;
- **4 балла** — высокий уровень при наличии отдельных ограничений;
- **5 баллов** — профиль демонстрирует наилучшее соответствие критерию в большинстве типовых условий.

Стоит отметить тот факт, что при этом оценки заданы **относительно других профилей**, а не в абсолютном смысле. Как пример, **P1** получает 5 баллов по пропускной способности не потому, что любая RFID-система всегда идеальна, а из-за того, что в потоковых сценариях она обычно удобнее и быстрее монофакторной биометрии. Следуя дальше, **P2** зарабатывает высшую оценку по устойчивости к передаче учетных данных, так как биометрический признак теснее связан с конкретным человеком, однако снижается по стабильности и инфраструктурной простоте вследствие чувствительности к качеству сенсоров, изображения и условий съемки. Далее, **P3** присущ максимальный балл по масштабируемости и интеграции, но хуже по отказоустойчивости из-за зависимости от каналов связи. Наконец, **P4** получает высокую оценку по защищенности и резервируемости, но уступает по простоте внедрения из-за многофакторности и более сложного сопровождения. Подводя итог, весовые коэффициенты отражают **приоритеты сценария**, а баллы — **сравнительную силу профиля по данному критерию**.

3. Базовая матрица оценки архитектурных профилей

Таблица 2. Базовые баллы профилей СКУД по критериям (1–5)

Профиль	C1	C2	C3	C4	C5	C6	C7	C8
P1. RFID-/карточная локальная/серверная	2	5	4	5	3	3	5	4
P2. Биометрическая локальная/серверная	5	3	2	2	3	3	2	2
P3. Облачная централизованная	3	4	3	3	5	5	3	2
P4. Гибридная многофакторная	5	4	4	2	4	5	3	5

**Начнем с RFID-/карточного профиля.** Низкий балл по критерию C1 связан с тем, что токентный доступ допускает передачу носителя и требует защиты от клонирования и несанкционированного чтения. В то же время, высокие оценки по C2 и C4 объясняются хорошей пропускной способностью и сравнительной простотой внедрения. Тем временем, высокий балл по C7 обусловлен тем, что такой профиль, как правило, не несет той правовой нагрузки, которая характерна для биометрии.

**Следующим идет биометрический профиль.** Прочную связь допуска с личностью пользователя отражает максимальный балл по C1, однако по C3 и C4 профиль получает сниженные оценки, поскольку биометрические системы чувствительны к качеству сенсоров, освещению, положению лица, состоянию интерфейса пользователя и защите шаблонов. Кроме того, дополнительное понижение по C7 связано с юридической значимостью обработки биометрических персональных данных.

**К последующему рассмотрению облачный профиль** — средняя оценка по C1 можно объяснить тем, что защищенность зависит не столько от облачности как таковой, сколько от применяемого идентификационного фактора. На-

против, высокие оценки по С5 и С6 отражают естественные преимущества централизованного администрирования, быстрого изменения прав доступа и интеграции журналов событий. Снижение по С8 связано с тем, что без локального резервирования облачная архитектура уязвима к сбоям связи.

**Последним профилем для анализа выступает гибридный.** Возможность сочетать два и более фактора аутентификации и предусматривать резервный сценарий допуска заслуженно является причиной его высоких оценок по С1 и С8. Тем не менее, оценка по С4 снижена, так как многофакторность почти всегда означает более сложное внедрение, настройку и сопровождение. Гибридный профиль выглядит сильнее не всегда и не везде, а прежде всего там, где риск ошибки доступа особенно чувствителен.

#### 4. Сценарная матрица выбора

Чтобы сравнение не осталось обобщенным, в работе использованы три типовых сценария:

- А** — учебный корпус или офисный объект с массовым доступом;
- В** — лаборатория, серверная, режимная зона;
- С** — распределенная сеть филиалов или удаленных площадок.

##### 4.1. Весовые коэффициенты по сценариям

Таблица 3. Весовые коэффициенты критериев

Критерий	Сценарий А	Сценарий В	Сценарий С
С1 Устойчивость к передаче/клонированию	0.15	0.25	0.10
С2 Пропускная способность	0.25	0.10	0.08
С3 Стабильность в неидеальных условиях	0.10	0.15	0.08
С4 Простота инфраструктуры	0.20	0.10	0.07
С5 Масштабируемость	0.10	0.05	0.25
С6 Интеграция	0.08	0.10	0.18
С7 Правовая нагрузка	0.05	0.10	0.09
С8 Отказоустойчивость	0.07	0.15	0.15

Исходя из данных приведенной ранее таблицы, здесь логика распределения весов существенно различна. В сценарии **А** максимальные веса получают **С2** и **С4**, потому что для объектов с большим потоком пользователей ключевыми становятся скорость прохода и умеренная сложность инфраструктуры. Иным образом показывает себя сценарий **В**, там на первый план выходят **С1**, **С3** и **С8**, так как режимная зона требует не просто удобного, а устойчивого и контролируемого допуска. Более сложный в архитектурном плане сценарий **С**, он демонстрирует доминирование **С5**, **С6** и **С8**, поскольку распределенный объект живет за счет централизованного управления, интеграции и предсказуемой работы при перебоях связи.

##### 4.2. Итоговые расчеты

Таблица 4. Итоговые оценки профилей СКУД по сценариям, баллы из 100

Профиль	Сценарий А	Сценарий В	Сценарий С
Р1. RFID-/карточная локальная/серверная	80.4	73.0	72.2
Р2. Биометрическая локальная/серверная	57.6	60.0	56.2
Р3. Облачная централизованная	70.8	65.0	75.8
Р4. Гибридная многофакторная	77.0	84.0	84.0

Результаты показывают, что для массового доступа лучший итоговый балл получает RFID-/карточный профиль. Итог довольно закономерен по причине того, что высокая пропускная способность и простота внедрения оказываются важнее, чем максимальная привязка допуска к личности. Следующим лидирует **Р4**, поскольку многофакторность и резервируемость дают выигрыш там, где цена ошибочного допуска выше, чем цена некоторого усложнения эксплуатации. Завершающим, для распределенной сети снова лидирует **Р4**, а **Р3** занимает второе место. Централизованное управление действительно полезно для филиальной структуры, но без локального резервирования облачная архитектура уступает

гибридной по устойчивости. Наименьшим по итогу, но не менее показательным является результат **P2**. Биометрический профиль, несмотря на сильную персонализацию допуска, не становится универсальным решением. Это связано не с особенностями биометрии по своей сути, а с тем, что в настоящих условиях ее преимущества компенсируются более сложной инфраструктурой, чувствительностью к условиям применения и более жестким правовым контуром. Отсюда следует важный практический вывод, биометрия приносит наибольшую пользу не как одиночное решение, а как усиливающий фактор внутри гибридной архитектуры.

## 5. Карта ключевых рисков внедрения

Таблица 5. Карта рисков внедрения СКУД

Риск	Для каких профилей критичен	Возможное последствие	Базовая мера снижения
Передача карты другому лицу / клонирование носителя	P1, частично P3	фиктивно легальный проход	двухфакторный допуск, пресечение повторного прохода (антипассбэк), защищенные идентификаторы
Ложный отказ биометрии в неидеальных условиях	P2, P4	задержка прохода, конфликт на точке доступа	пилотные испытания, резервный токенный сценарий, требования к освещению и позиционированию
Отказ канала связи / облачной платформы	P3, частично P4	блокировка прохода или потеря управляемости	локальный кэш прав доступа, edge-контроллеры, регламент оффлайн-режима
Нарушение режима обработки биометрических ПДн	P2, P4	юридические риски и претензии субъектов	локальные акты, правовое основание обработки, документирование процессов
Ошибка проектирования точки прохода	все профили	очереди, снижение реальной эффективности	тестирование пропускной способности, сценарная настройка, пилотный запуск

Карта рисков показывает, что выбор СКУД нельзя сводить к спору о том, какая технология лучше. На практике, проблемы возникают именно на стыке техники, регламента и пользовательского поведения. RFID может быть удобной, но уязвимой к передаче носителя, биометрия — персонализированной, но чувствительной к среде, облако — управляемым, но зависимым от связи. Вот поэтому, качественное проектное решение — это всегда не один компонент, а согласованная конфигурация объекта.

## Обсуждение результатов

Полученная модель свидетельствует, что универсального профиля СКУД не существует. Вместе с тем, она позволяет уйти от слишком общих формул вроде биометрия надежнее или облако современнее, что позволяет отойти от распространенной практики внедрения ради внедрения. Для потоковых объектов решающее значение имеют скорость прохода, простота обслуживания и умеренная правовая нагрузка, отчего традиционные решения по-прежнему сохраняют сильные позиции. Для режимных зон, напротив, критичны устойчивость к передаче учетных данных, стабильность и резервируемость, и именно здесь гибридный профиль закономерно выходит на первое место [3–11].

С научной точки зрения, основной результат работы состоит не в повторении известных достоинств RFID, биометрии или облачной архитектуры, а в том, что предложен **воссоздаваемый способ сравнения**, где логика выбора раскрыта явно, заданы критерии, объяснены веса, описана шкала и показан переход от базовых баллов к итоговому сценарию, что и делает модель полезной для прикладных работ и исследований.

Вместе с тем, у исследования есть ограничения. Во-первых, модель не заменяет действительные испытания на конкретном объекте. Во-вторых, в нее не включен отдельный экономический модуль, то есть стоимость внедрения и обслуживания рассмотрена только через инфраструктурную сложность и организационную нагрузку. В-третьих, итоговые баллы формировались как авторская аналитическая процедура, а не общепринятый отраслевой стандарт. Но, именно в этом виде модель и ценна, она позволяет еще до этапа закупки увидеть, где СКУД действительно подходит объекту, а где решение выглядит эффектно только на уровне описания.

## Заключение

В статье разработана авторская модель выбора архитектурного профиля СКУД по технологическим и эксплуатационным критериям. Продемонстрировано, что сравнивать RFID, биометрию, облачные и гибридные решения как одно-



порядковые типы методически некорректно, для корректного анализа необходимо разделять уровень идентификационного фактора и уровень архитектуры системы.

Введенная система критериев, шкала балльной оценки и сценарное взвешивание позволили установить, что для объектов массового доступа наиболее рационален RFID-/карточный профиль, для режимных зон наилучшие результаты демонстрирует гибридный многофакторный профиль, для распределенной сети объектов наиболее эффективен гибридный профиль, облачная архитектура выступает сильным вторым вариантом при условии резервирования, биометрия в большинстве практических случаев приносит максимальную пользу не как единственное решение, а как усиливающий фактор в составе гибридной СКУД.

Тем самым, гипотеза исследования подтверждается. Оптимальная СКУД определяется не уровнем технологической новизны как таковой, а степенью соответствия архитектурного профиля режиму объекта, эксплуатационной нагрузке и правовому контексту обработки данных.

#### Литература:

1. ГОСТ Р 51241–2008. Средства и системы контроля и управления доступом. Классификация. Общие технические требования. Методы испытаний. [Электронный ресурс]. URL: <https://meganorm.ru/Data2/1/4293829/4293829931.pdf> (дата обращения: 16.04.2026)
2. Изменение № 1 к ГОСТ Р 51241–2008, утв. приказом Росстандарта от 27.01.2025 № 26-ст. [Электронный ресурс]. URL: [https://meganorm.ru/mega\\_doc/norm\\_update\\_29032025/gost-r\\_gosudarstvennyj-standart/0/gost\\_r\\_51241-2008\\_natsionalnyy\\_standart\\_rossiyskoy.html](https://meganorm.ru/mega_doc/norm_update_29032025/gost-r_gosudarstvennyj-standart/0/gost_r_51241-2008_natsionalnyy_standart_rossiyskoy.html) (дата обращения: 16.04.2026)
3. Разъяснения Роскомнадзора по вопросам отнесения фото-, видеоизображений, дактилоскопических данных и иной информации к биометрическим персональным данным и особенностям их обработки. [Электронный ресурс]. URL: [https://rkn.gov.ru/docs/Raz6jasnenija\\_RKN\\_po\\_biometrii\\_okonchatel6naja\\_versija.doc](https://rkn.gov.ru/docs/Raz6jasnenija_RKN_po_biometrii_okonchatel6naja_versija.doc) (дата обращения: 16.04.2026)
4. Федеральный закон от 27.07.2006 № 152-ФЗ «О персональных данных». [Электронный ресурс]. URL: <https://base.garant.ru/12148567/> (дата обращения: 16.04.2026)
5. Федеральный закон от 21.07.2014 № 242-ФЗ «О внесении изменений в отдельные законодательные акты Российской Федерации в части уточнения порядка обработки персональных данных в информационно-телекоммуникационных сетях». [Электронный ресурс]. URL: [https://www.consultant.ru/document/cons\\_doc\\_LAW\\_165838/](https://www.consultant.ru/document/cons_doc_LAW_165838/) (дата обращения: 16.04.2026)
6. Федеральный закон от 29.12.2022 № 572-ФЗ «Об осуществлении идентификации и (или) аутентификации физических лиц с использованием биометрических персональных данных...». [Электронный ресурс]. URL: <http://www.kremlin.ru/acts/bank/48740/page/1> (дата обращения: 16.04.2026)
7. Face Recognition Vendor Test (FRVT) / National Institute of Standards and Technology. [Электронный ресурс]. URL: [https://pages.nist.gov/frvt/html/frvt\\_facemask.html](https://pages.nist.gov/frvt/html/frvt_facemask.html) (дата обращения: 16.04.2026)
8. Face Recognition Technology Evaluation: Face Mask Effects / National Institute of Standards and Technology. [Электронный ресурс]. URL: <https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt> (дата обращения: 16.04.2026)
9. Munoz-Ausecha C., Ruiz-Rosero J., Ramirez-Gonzalez G. RFID Applications and Security Review // Computation. 2021. Vol. 9. No. 6. Art. 69. [Электронный ресурс]. URL: [https://www.researchgate.net/publication/352295274\\_RFID\\_Applications\\_and\\_Security\\_Review](https://www.researchgate.net/publication/352295274_RFID_Applications_and_Security_Review) (дата обращения: 16.04.2026)
10. Velas A., Boroš M., Kuffa R., Lenko F. Testing of Permeability of RFID Access Control System for the Needs of Security Management // Applied Sciences. 2024. Vol. 14. No. 10. Art. 4227. [Электронный ресурс]. URL: <https://www.mdpi.com/2076-3417/14/10/4227> (дата обращения: 16.04.2026)
11. Yang W., Wang S., Hu J., Zheng G., Valli C. Security and Accuracy of Fingerprint-Based Biometrics: A Review // Symmetry. 2019. Vol. 11. No. 2. Art. 141. [Электронный ресурс]. URL: <https://www.mdpi.com/2073-8994/11/2/141> (дата обращения: 16.04.2026)

## Разработка программного модуля автоматизации деятельности преподавателя на основе интеграции с большими языковыми моделями

Кондратенко Алексей Сергеевич, студент

Национальный исследовательский университет «Московский институт электронной техники» (г. Зеленоград)

*В статье рассматривается программная реализация модуля автоматизации деятельности преподавателя в рамках образовательного веб-комплекса. Описана проблема высоких временных затрат на рутинную проверку заданий и сбор статистики. Представлено решение на стеке Django, Vue.js 3 и PostgreSQL. Описаны механизмы защиты от ИИ-галлюцинаций при формировании заданий, алгоритмы автоматической проверки ответов, античит-система мониторинга активности учащихся, а также методы оптимизации серверных запросов для агрегации образовательных метрик.*

**Ключевые слова:** программная инженерия, автоматизация, образовательные технологии, базы данных, Django, Vue.js, PostgreSQL, YandexGPT.

## Development of a program module for automation of teacher's activity on the basis of integration with large language models

*The article discusses the software implementation of a module for automating teacher activities within an educational web complex. The problem of high time expenditure on routine task checking and statistics gathering is described. A solution based on the Django, Vue.js 3, and PostgreSQL stack is presented. Mechanisms for preventing AI hallucinations during task generation, auto-grading algorithms, an anti-cheat student activity monitoring system, and server request optimization methods for educational metric aggregation are described.*

**Keywords:** software engineering, automation, educational technology, databases, Django, Vue.js, PostgreSQL, YandexGPT.

### Введение

В условиях цифровизации образования одной из ключевых проблем остается высокая академическая нагрузка на преподавательский состав. Значительная часть рабочего времени затрачивается на рутинные процессы: формирование домашних заданий, механическую проверку ответов и сбор статистики успеваемости.

Стремительное внедрение больших языковых моделей (LLM) в образовательные платформы создает соблазн полной автоматизации этих процессов силами нейросетей. Однако генерация математических задач посредством ИИ часто приводит к смысловым галлюцинациям и некорректным условиям, что недопустимо в точных науках.

Разработка программного модуля, который строго разделяет зоны ответственности: оставляет подбор и проверку задач детерминированным алгоритмам реляционных баз данных, а роль ИИ сводит к интеллектуальному ассистированию учащегося с последующим сбором количественных метрик для преподавателя — является актуальной задачей программной инженерии.

### Основная часть

#### Архитектура данных и механизмы формирования заданий

Основой программного модуля выступает реляционная база данных под управлением СУБД PostgreSQL. Процесс формирования заданий спроектирован с полным исключением больших языковых моделей из контура генерации контента для предотвращения математических ошибок. Преподавателю предоставляется два детерминированных пути:

1. Автоматический генератор: алгоритм формирует пул задач из предварительно верифицированного банка данных (PostgreSQL) на основе заданных параметров (тема, класс, уровень сложности).
2. Пользовательские шаблоны: ручное создание уникальных вариантов через встроенный Markdown-редактор, поддерживающий рендеринг математических формул и предпросмотр в реальном времени.

Выдача сформированных заданий реализуется индивидуально: клиентская часть системы (store-модуль на базе Vue.js) позволяет преподавателю адресно назначать работу с установкой жестких сроков (дедлайнов).

#### Автоматизация контроля и оптимизация серверной части

Серверная часть, реализованная на фреймворке Django, полностью берет на себя процесс верификации ответов. Механизм автопроверки производит строгую сверку пользовательского ввода с эталонным ключом (answer\_key) или масивом допустимых опций (options). Это полностью избавляет преподавателя от необходимости ручной проверки.

Для обеспечения высокой производительности дашборда при расчете сводной статистики успеваемости была решена классическая проблема ORM проблема N+1 запросов. Использование встроенных методов агрегации на уровне базы данных (aggregate, Count) позволило серверу эффективно собирать общую статистику решаемости задач без создания избыточной нагрузки на СУБД.

Аналитический дашборд и метрики взаимодействия с ИИ

Клиентский интерфейс преподавателя (Frontend) разработан на базе фреймворка Vue.js 3. Центральным элементом выступает аналитический дашборд, агрегирующий данные о процессе решения.

Вместо предоставления преподавателю громоздких текстовых транскриптов бесед ученика с YandexGPT, система делает ставку на количественные метрики. Модуль анализирует симптоматику затруднений, выводя на экран: количество затраченного времени на задачу, число совершенных попыток и факт использования ИИ-подсказок. Дополнительно внедрена античит-система, фиксирующая переключения контекста браузера (счетчик «Покидал вкладку»), что позволяет идентифицировать попытки несанкционированного поиска ответов в сети.

Результаты

Апробация разработанного модуля в составе образовательного веб-комплекса подтвердила высокую эффективность предложенных архитектурных решений. На рисунке 1 представлен интерфейс преподавателя в режиме проверки выполнения домашнего задания.

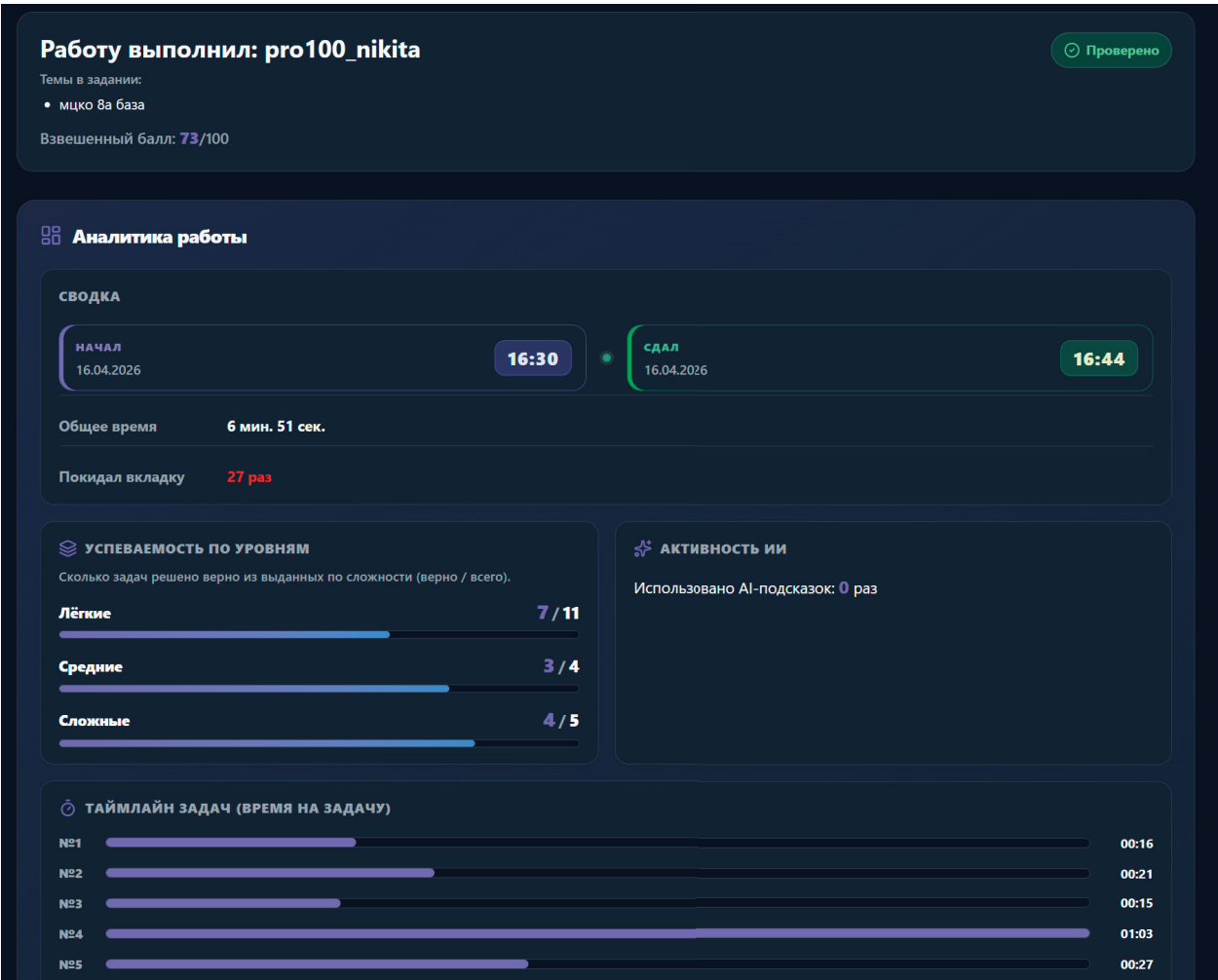


Рис. 1. Аналитический дашборд преподавателя с результатами автопроверки и античит-метриками

Как видно из представленной аналитической сетки, система в удобном графическом виде транслирует успеваемость (индикаторы успешного или ошибочного решения), а также поведенческие паттерны. Например, высокое значение

счетчика переключений вкладок в совокупности с быстрым правильным ответом сигнализирует о списывании. И наоборот, большое количество попыток и активное использование ИИ-подсказок без факта списывания позволяет преподавателю сделать вывод о наличии у ученика системных затруднений с конкретной темой и своевременно скорректировать образовательную траекторию.

Процесс формирования пользовательских заданий также был значительно оптимизирован благодаря внедрению кастомного Markdown-редактора (рис. 2), позволяющего преподавателям оперативно переносить методические разработки в цифровую среду.

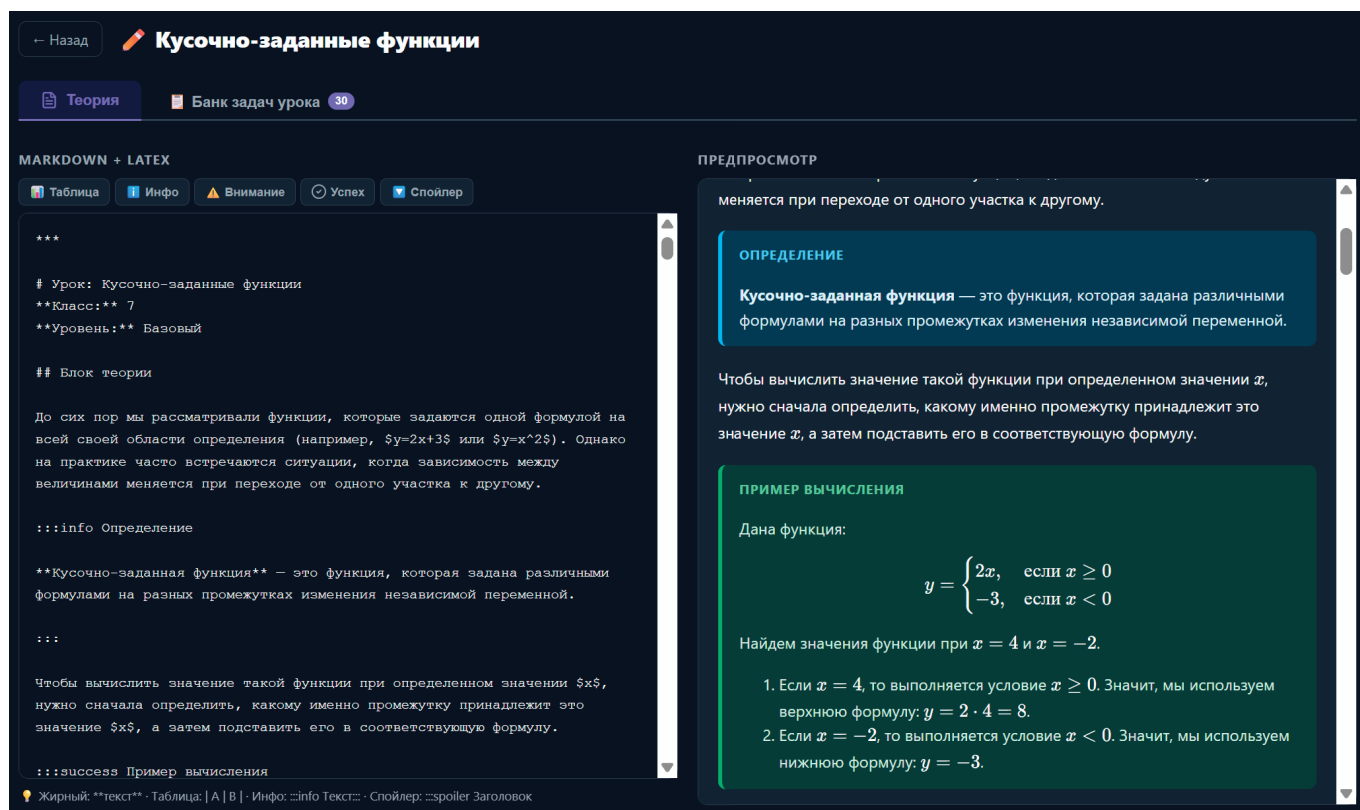


Рис. 2. Интерфейс создания авторских задач с поддержкой Markdown и LaTeX

## Заключение

Разработанный программный модуль успешно решает задачу автоматизации деятельности преподавателя, снижая временные затраты на рутинную проверку и сбор статистики. Использование архитектурного стека Django и Vue.js с оптимизированными запросами к PostgreSQL обеспечило высокую производительность платформы. Отказ от генерации задач силами LLM в пользу использования верифицированной базы гарантировал математическую корректность материалов. Переход от ручного контроля к анализу агрегированных метрик (статистика попыток, использование ИИ-ассистента, античит-система) позволяет преподавателю сфокусироваться на качественном анализе образовательного процесса и выработке эффективных педагогических стратегий.

## Литература:

1. Документация Yandex Cloud: YandexGPT API [Электронный ресурс]. URL: <https://yandex.cloud/ru/docs/yandexgpt/> (дата обращения: 06.04.2026).
2. Официальная документация фреймворка Django [Электронный ресурс]. URL: <https://docs.djangoproject.com/> (дата обращения: 06.04.2026).
3. Официальная документация фреймворка Vue.js [Электронный ресурс]. URL: <https://vuejs.org/guide/introduction.html> (дата обращения: 06.04.2026).
4. Зайцева Л. В. Системы адаптивного обучения и контроля знаний // Образовательные технологии и общество. 2013. Т. 16. № 4. С. 401–414.

## Decentralized Architectures in Next-Generation Technologies

Mikhailiuk Maksim Sergeevich, Independent Researcher (Moscow)

*Centralized coordination architectures carry latency and fault-tolerance costs that become operationally binding as autonomous multi-agent deployments scale beyond several dozen agents. Prior hybrid distributed frameworks have advanced individual aspects of this problem: consensus protocols, neurosymbolic planning, and fault-resilient topologies have each been studied in isolation, yet no published design has jointly delivered predictive pre-computation, constraint-guaranteed policy generation, and per-tier performance bounds within a unified architecture. The Distributed AI Decision Engine presented here is evaluated through simulation across autonomous vehicle fleets, robotic manufacturing cells, and warehouse logistics systems. Against a cloud-centralized reference, the engine achieves a mean conflict resolution latency of 38 ms versus 190 ms, sustains coordinated operation under 50 % node failure and full network partition, and produces throughput gains of 22–31 % in high-contention manufacturing scenarios. Warehouse simulations confirm Tier 1 and Tier 2 degradation behavior under intermittent radio connectivity characteristic of dense metallic environments. The architecture’s operationally bounded fault model enables pre-deployment verification against domain-specific requirements, a capability absent from prior distributed designs.*

**Keywords:** consensus-weighted arbitration, self-organizing mesh topology, symbolic constraint filtering, real-time autonomous coordination, Byzantine fault tolerance, gated recurrent neural network, predictive policy pre-computation, hierarchical intent graph.

### Introduction

Autonomous multi-agent coordination has become a defining engineering constraint as deployment scales have grown from isolated laboratory demonstrations to production systems involving hundreds of concurrently active agents sharing physical infrastructure. Manufacturing cells, urban vehicle fleets, and logistics warehouses now operate at densities where the choice of coordination architecture directly shapes throughput, safety margins, and system behavior under hardware failures or communication disruptions. Centralized orchestration dominated early deployments because it offered tractable optimization and straightforward consistency guarantees (Stone & Veloso, 2000; Jennings, 1993). At the agent scales now common in industrial practice, however, the structural costs of a hub-and-spoke topology have become a primary engineering concern rather than a secondary consideration.

The costs take three forms. Agent-to-controller round-trip latency compounds with global planning computation as fleet size grows; reported values in centralized vehicle coordination platforms range from 150 ms to over 400 ms under moderate conflict rates (Dresner & Stone, 2008; Zhong et al., 2023, pp. 6325–6326), which exceed the sub-100 ms ceiling of close-formation platoon control and the sub-50 ms ceiling of high-speed manufacturing cell sequencing. The controller also constitutes a single point of failure whose disruption eliminates coordination for all connected agents simultaneously, rather than degrading it gradually (Lynch, 1996; Tanenbaum & Van Steen, 2007). Global state consistency maintenance scales with agent population and its communication cost grows without bound in geographically distributed deployments (Fischer et al., 1985; Gilbert & Lynch, 2002).

Fully decentralized architectures resolve the latency and resilience problems but cannot sustain the coordination quality required in high-density domains where agents contend continuously for shared resources (Parker, 1998; Cao et al., 1997).

Among hybrid approaches, DANCeRS (Patwardhan & Davison, 2025, preprint) addresses distributed consensus through zone-partitioned leader election but provides no predictive pre-computation, leaving decision latency bounded by reactive inference time. Neurosymbolic planning systems deliver constraint-guaranteed policies but depend on centralized planners for vocabulary maintenance (Garcez & Lamb, 2023). Fault-tolerant distributed architectures in the multi-robot literature characterize degradation empirically without specifying per-tier bounds in advance (Pierson & Schwager, 2018; Farinelli et al., 2004). Because no published design jointly addresses all three gaps, deployed systems currently absorb significant performance loss along whichever dimension is deprioritized. The Distributed AI Decision Engine presented here closes all three gaps through specific architectural mechanisms, evaluates performance against a cloud-centralized reference under nominal and degraded conditions, and specifies operationally verifiable bounds for its three-tier fault-tolerance model.

### Methods

Each physical or logical agent instantiates exactly one Decision Agent Node (DAN). The live population of active DANs constitutes the engine’s membership, managed by a Membership and Topology Management (MTM) subsystem executing a modified Gossip protocol (Karp et al., 2000) that records per-peer one-way communication latency, packet-loss rate as a rolling 30-sample mean, and an aggregate link-quality score for each neighbor. The neighborhood radius was set to 3 hops in all reported experiments. A mesh topology is a structural precondition for the two-round-trip termination of conflict resolution: edge-local arbitration requires direct peer-to-peer channels, and routing through aggregation points would reintroduce the controller-hop overhead the design is intended to eliminate.

The Sensor Interface Layer normalizes inputs from LIDAR point clouds, RGB-D camera frames, inertial measurement



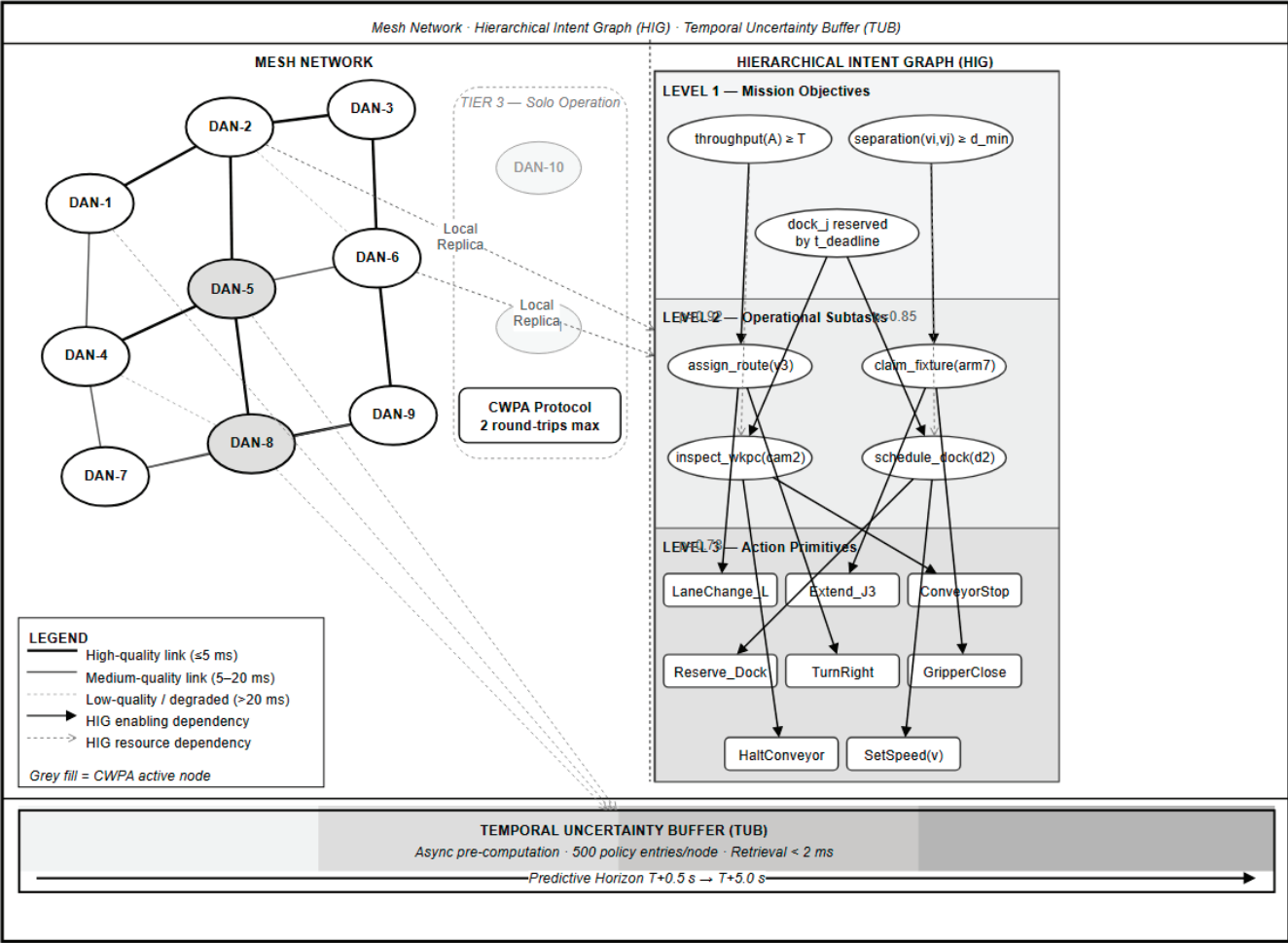


Fig. 1. System-Level Architecture

unit streams, and industrial encoder pulses into Observation Vectors aligned to a common IEEE 1588 Precision Time Protocol clock reference. Temporal alignment is a strict requirement because proposals from nodes referencing different observation instants would carry systematically inconsistent confidence scores, distorting arbitration outcomes.

The Contextual State Memory (CSM) maintains a rolling 2-second window of Observation Vectors and compresses it through a gated recurrent neural network into a fixed-dimensional State Embedding. The gated recurrent design is preferable to Transformer-based alternatives here because its per-timestep computational cost is independent of sequence length. Transformer architectures scale quadratically with window width, which precludes deterministic latency on embedded processors with constrained memory bandwidth.

The Local Inference Module (LIM) takes the State Embedding and the local HIG replica as joint input. A neural policy network scores the action primitive vocabulary, producing a softmax probability distribution; the score assigned to each candidate is its softmax probability  $\alpha$ . A symbolic constraint evaluator then removes any candidate associated with a safety constraint violation, resource exclusivity conflict, or operational prohibition in the node's current constraint vocabulary. The filtering produces a Boolean accept/reject decision with

a traceable audit record for each candidate. Constraint violations are structurally excluded with respect to the explicitly specified vocabulary; incomplete initial specification remains a residual risk, which the HIG cryptographic delta verification addresses by preventing unauthorized post-deployment vocabulary modification.

Each DAN requires a sustained neural inference throughput of at least 4 TOPS to execute the LIM, TUB predictive model, and CWPA simultaneously within nominal latency bounds. On hardware delivering below this threshold, the LIM may be configured in symbolic-only mode, omitting the neural policy network at the cost of adaptivity to environmental patterns not explicitly represented in the constraint set.

The Conflict Resolution Submodule (CRS) monitors pending action proposals through the CWPA broadcast channel and maintains a resource-to-claimant cache. Proposals claiming no contested resource bypass Phases 3–5 of the CWPA and proceed directly to Phase 6. Under nominal conditions, where most actions do not involve contested resources, this bypass is the dominant contributor to the engine's low median latency.

The CWPA resolves inter-node conflicts within at most two communication round trips from proposal broadcast to committed action. Each proposal carries the action primitive

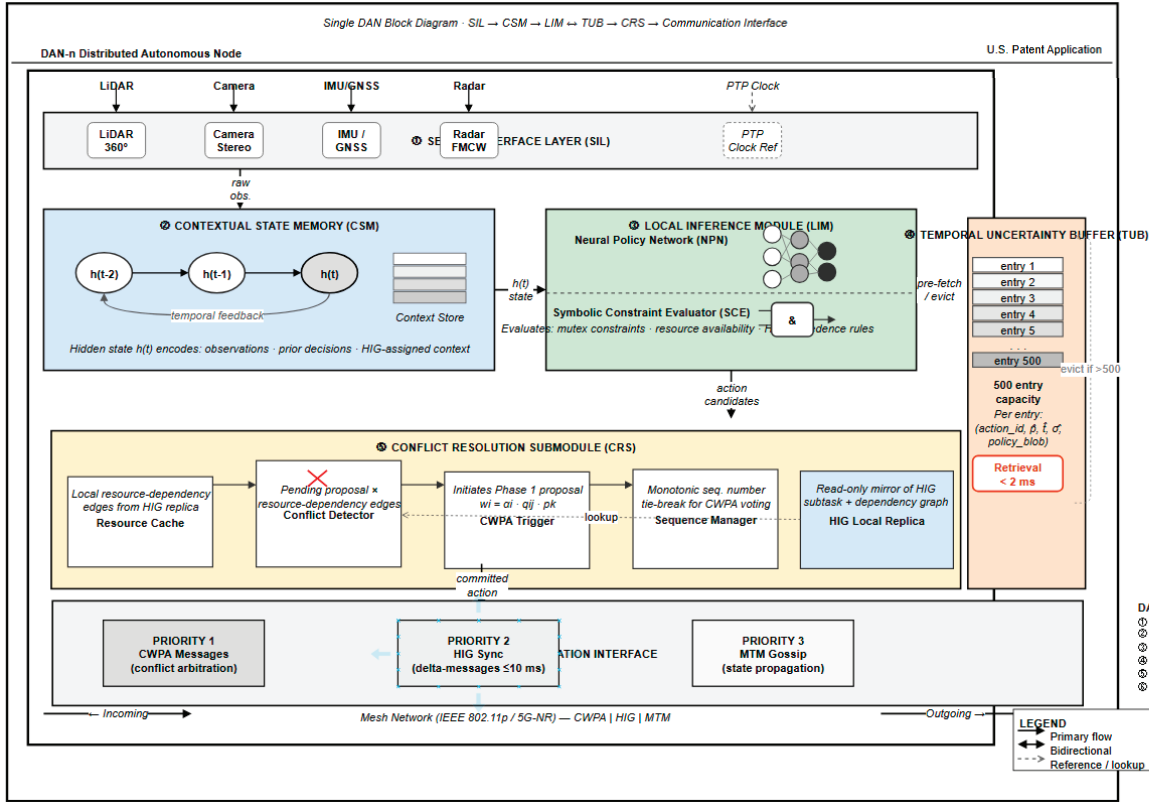


Fig. 2. DAN Internal Architecture

identifier, HIG subtask identifier, required resource set, the LIM softmax score, and a monotonically increasing sequence number. Each receiving peer computes:

$$w_i = \alpha_i \times q_{ij} \times p_k$$

where  $\alpha_i$  is the responding node's softmax probability for the proposed action given its own State Embedding,  $q_{ij}$  is the link-quality score to the proposing node from the MTM topology table, and  $p_k$  is the priority weight of the HIG Level 2 subtask the proposal serves. Multiplication is chosen over addition because it encodes a zero-dominance property: a node with no local observability ( $\alpha_i = 0$ ), a failed communication link ( $q_{ij} = 0$ ), or a zero-priority subtask ( $p_k = 0$ ) contributes nothing to the outcome without requiring a separate abstention mechanism. An additive formulation would not produce this behavior, because a node with zero observability but intact link quality and high subtask priority would still cast a weighted vote for a proposal it cannot reliably evaluate. When link quality converges to zero for all responding nodes simultaneously, every weight sum collapses to zero and the proposal bearing the lowest sequence number commits by default, preserving protocol progress without any fallback authority.

The HIG is a directed acyclic graph replicated across all active nodes under cryptographic verification. Level 1 encodes global mission objectives as formal logical predicates. Level 2 decomposes each objective into operational subtask nodes assignable to individual agents. Level 3 maps subtasks to executable action primitives. Enabling-dependency edges specify completion prerequisites; resource-dependency edges specify

mutually exclusive access requirements. Each edge carries a priority weight in. Modifications propagate as signed delta messages; a receiving node applies a delta only after signature verification and confirmation that the update introduces neither a DAG cycle nor a priority-weight violation. Convergence across up to 100 active nodes occurs within 10 ms under nominal network conditions; for up to 1,000 nodes, within 50 ms. Agent failures trigger distributed replanning through the same CWPA pathway as ordinary conflicts, with no separate replanning authority required.

During processor cycles not occupied by active CWPA participation or LIM inference, a conditional variational predictive model takes the 2-second Observation Vector history and the current local HIG state as input and produces a ranked list of anticipated state transitions. Each transition is annotated with an estimated probability  $\pi$  and a time-to-occurrence distribution parameterized as a log-normal with mean  $\hat{t}$  and standard deviation  $\hat{\sigma}$ . For any anticipated transition whose  $\pi\pi$  exceeds a configurable threshold  $\theta$  (typically 0.25–0.60), the LIM pre-computes a constraint-filtered response policy stored in the TUB keyed by the transition's trigger condition. On detection of a monitored transition in the live sensor stream, the CRS retrieves the stored policy within less than 2 ms and submits it directly to Phase 1 of the CWPA, bypassing the full LIM inference step and yielding a 60–65 % reduction in per-event inference contribution to total decision latency. Eviction from the 500-entry TUB follows a priority ordering by  $\pi \cdot c$ , where  $c$  is the criticality score derived from the priority weight of the HIG subtask the transition would affect.

The predictive model is trained offline on domain-representative data. Sensitivity analysis conducted during the simulation study showed that when unrecognized event types exceed approximately 30 % of the event population, TUB hit rate drops below 20 % and median latency reverts to the reactive LIM inference range of 15–35 ms for those categories. An online adaptation mechanism would address this boundary and represents the most consequential direction for future development.

Tier 1 activates when one or more peers become unreachable while the remaining nodes form a connected subgraph. The CWPA continues over the reachable set; weight normalization in Phase 3 is adjusted to sum over responding nodes only, preserving mathematical consistency. Tier 2 activates under full network partition. Each disconnected subgraph operates with its local HIG replica and its own CWPA instance; every intra-subgraph conflict resolves without exception, while inter-subgraph resource arbitration queues until healing. Partition healing triggers a merge protocol that reconciles divergent replicas in causal order, using timestamps and certificate sequence numbers to resolve ambiguous concurrency. Tier 3 activates when a node loses all peer communication. Solo Operation Mode tightens the constraint evaluator to exclude any action requiring a shared resource whose current allocation state cannot be verified locally. TUB retrieval continues at full speed because it operates independently of peer connectivity.

All experiments used a custom discrete-event simulation environment implemented in Python 3.11, with each DAN subcomponent as a standalone module communicating through protocol-compliant message-passing interfaces. To establish simulation validity, the two-node CWPA case was tested against the analytically derived termination bound under deterministic network conditions, confirming exact agreement. HIG synchronization convergence under 100 nodes was compared against the theoretical Gossip protocol round-complexity bound from Karp et al. (2000), converted to milliseconds under the assumption of 0.1 ms per-round

link latency consistent with IEEE 802.15.4 measurements; simulated convergence of 9.3 ms (SD = 0.7 ms) confirmed consistency with the theoretical prediction of 10 ms.

Network communication was modeled as a probabilistic delay channel with per-link latency drawn from log-normal distributions fitted to published IEEE 802.11p vehicle-to-vehicle and IEEE 802.15.4 industrial mesh measurements. All reported metrics are means over 30 independent runs, each with a distinct random seed governing topology initialization, traffic demand generation, and node failure timing.

Vehicle fleet experiments used a  $16 \times 16$  synthetic urban grid with randomized traffic demand matrices generating 8–22 conflict pairs per second across fleets of 20, 50, and 100 vehicles. The cloud-centralized reference used an identical routing logic on a single-server scheduler; round-trip latency was set to 85 ms consistent with published cloud coordination measurements (Zhong et al., 2023, pp. 6325–6326), and computational load was modeled as a linear function of active conflict pairs at 0.8 ms per pair. Node failure was induced by disabling 50 % of nodes at simulation minute 5 and restoring them at minute 10; failure-condition metrics average over the 5-minute failure window. Manufacturing experiments used cells of 12, 24, and 48 nodes with fixture pools at 25 % of robotic population (high contention) and 75 % (low contention). Warehouse experiments used a  $200 \times 60$  m facility model with 40 mobile robots, metallic shelving modeled as 35 % packet-loss zones, and mixed aisle-access and dock-scheduling conflicts. Between-condition comparisons used two-sided Welch's t-tests at  $\alpha = 0.05$ .

## Results

Mean conflict resolution latency under nominal conditions was 38 ms (95 % CI: 35–41 ms) across all tested fleet sizes, against 190 ms (95 % CI: 181–199 ms) for the centralized reference. The ratio of approximately 5:1 in mean latency and greater than 6.5:1 at the 95th percentile remained stable from

Table 1. Conflict resolution latency across operational conditions

Condition	Engine mean (ms)	95 % CI	Engine 95th pct (ms)	Reference mean (ms)	95 % CI	Reference 95th pct (ms)
Nominal (20–100 agents)	38	35–41	61	190	181–199	400+
50 % node failure	47	43–51	78	310	291–329	520+ (CI: 489–551)
Full network partition	39	36–42	63	N/A	—	—

Table 2. Manufacturing cell throughput improvement over centralized scheduling

Cell size	Improvement	95 % CI	p-value
12 nodes	22 %	18–26 %	< 0.01
24 nodes	27 %	23–31 %	< 0.01
48 nodes	31 %	27–35 %	< 0.01
Low contention (all sizes)	~0 %	—	0.31 (n.s.)

20 to 100 agents, consistent with the topological prediction that CWPA requires two message hops between affected agents while the reference requires four through a remote controller. Behavior beyond 100 agents requires separate empirical characterization.

Under 50 % node failure, engine mean latency rose to 47 ms (95 % CI: 43–51 ms), a 24 % increase from nominal values. Reference mean latency reached 310 ms (95 % CI: 291–329 ms), with 18 % of resolution attempts failing to complete within the 500 ms timeout; at the 95th percentile, reference latency exceeded 520 ms (CI: 489–551 ms). Node loss reduces the CWPA responding population but creates no additional processing burden on remaining nodes. In the centralized reference, failures increase controller queue contention through compounding feedback absent from the distributed design, which accounts for the asymmetric degradation magnitude.

Under full network partition, the engine maintained 39 ms mean latency (95 % CI: 36–42 ms) within each subgraph, a value statistically indistinguishable from nominal ( $p = 0.47$ , Welch’s  $t$ -test). The reference architecture produced no successful resolutions. The Tier 2 degradation specification therefore corresponds to a measurable operational outcome.

TUB retrieval handled 38 % of conflict events at 20 agents and 52 % at 100 agents, with the increase at higher fleet density reflecting greater recurrence of intersection occupancy and lane-merge conflicts for which the predictive model had been trained. For TUB-handled events, per-decision inference contribution dropped from 15–35 ms to under 2 ms; remaining latency was dominated by CWPA broadcast and voting phases. Above the 30 % out-of-distribution threshold identified in the sensitivity analysis, hit rates dropped sharply, confirming that TUB benefits depend on training data coverage.

Throughput improvements in high-contention cells were 22 %, 27 %, and 31 % for 12-, 24-, and 48-node configurations respectively (all  $p < 0.01$ , Welch’s  $t$ -test; individual CIs in Table 2). The improvement reflects the softmax score  $\alpha_i$  functioning as an implicit execution-readiness signal: robots already positioned near a contested fixture produce higher scores and win arbitration more often, reducing the incidence of resource claims followed by extended repositioning delays. Under low-contention conditions no statistically significant difference emerged ( $p = 0.31$ ), consistent with the prediction that distributed arbitration advantages concentrate in high-contention regimes.

Robots operating within 35 % packet-loss shelving zones continued executing TUB-precomputed pick-and-place policies without coordination loss, consistent with Tier 1 degradation behavior. Robots with active peer links arbitrated aisle access and dock scheduling through CWPA with mean latency of 41 ms (95 % CI: 38–44 ms), 8 % above the open-floor nominal of 38 ms and attributable to increased retransmission under burst packet loss. Robots carrying time-critical orders won dock access arbitration in 94 % of contested cases where the HIG urgency weight differential exceeded 0.3.

## Discussion

The factor-of-five mean latency advantage over the centralized reference follows from eliminating the controller from the arbitration communication path. Within the tested range of 20 to 100 agents, the advantage was uniform and showed no trend toward narrowing at higher fleet sizes, consistent with the theoretical prediction that CWPA latency is bounded by neighborhood communication time rather than growing with total fleet population. The 24 % engine latency increase under 50 % node failure contrasts with the 63 % increase for the centralized reference. That difference reflects the absence of a scheduling load-feedback loop in the distributed design, a structural property of the mesh topology rather than a tuning outcome.

Structural exclusion of constraint violations at the policy generation stage shifts safety responsibility from runtime neural behavior to the vocabulary specification process. The residual risk of incomplete specification is qualitatively different from the unbounded distribution-shift risk of purely neural systems: it is bounded, auditable, and addressable through engineering review of the constraint vocabulary. Deployments where that vocabulary can be exhaustively specified benefit most from this property; domains with highly open-ended action spaces gain less because the specification burden becomes significant. Hardware deployments must also account for the 4 TOPS minimum compute requirement per node; below this threshold, the symbolic-only LIM variant preserves CWPA and HIG fault-tolerance properties while sacrificing neural adaptivity.

All results derive from simulation rather than physical deployment. The simulation was validated against analytical bounds and uses network parameters from published hardware measurements, but burst packet-loss patterns in real warehouse environments may not be fully captured by the log-normal delay model. The CWPA has been empirically characterized only up to 100 agents; message volume under very large neighborhood radii at greater scales has not been measured. Above approximately 30 % unrecognized event types, TUB hit rate collapses below 20 % and median latency reverts to the reactive inference range for those categories; an online adaptation mechanism would address this boundary and constitutes the most consequential direction for future work.

## Conclusion

Hub-and-spoke coordination topologies impose latency, resilience, and state-consistency costs that are properties of the communication graph itself and cannot be resolved through engineering improvements within that topology. Moving arbitration authority to the agent mesh replaces controller round-trips with two-hop peer exchanges and replaces qualitative resilience claims with operationally specified per-tier performance bounds. Within the tested range of 20 to 100 agents, the architecture achieves a factor-of-five latency reduction, sustains full intra-partition coordination under network disruption, and delivers 22–31 % throughput gains in



high-contention manufacturing scenarios. These outcomes are attributable to structural properties of the design rather than to parameter tuning.

The case for this architecture is strongest in deployment domains where the constraint vocabulary can be exhaustively specified in advance and where hardware meeting the

4 TOPS threshold per node is available. Future work should prioritize an online adaptation mechanism for the TUB predictive model, which would extend predictive benefits to environments differing substantially from training conditions and remove the most significant remaining operational boundary identified in this study.

#### References:

1. Cao, Y. U., Fukunaga, A. S., & Kahng, A. (1997). Cooperative mobile robotics: Antecedents and directions. *Autonomous Robots*, 4(1), 7–27. <https://doi.org/10.1023/A:1008855018923>
2. Dresner, K., & Stone, P. (2008). A multiagent approach to autonomous intersection management. *Journal of Artificial Intelligence Research*, 31, 591–656. <https://doi.org/10.1613/jair.2502>
3. Farinelli, A., Iocchi, L., & Nardi, D. (2004). Multirobot systems: A classification focused on coordination. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 34(5), 2015–2028. <https://doi.org/10.1109/TSMCB.2004.832155>
4. Fischer, M. J., Lynch, N. A., & Paterson, M. S. (1985). Impossibility of distributed consensus with one faulty process. *Journal of the ACM*, 32(2), 374–382. <https://doi.org/10.1145/3149.214121>
5. Garcez, A. d'Avila, & Lamb, L. C. (2023). Neurosymbolic AI: The 3rd wave. *Artificial Intelligence Review*, 56(11), 12387–12406. <https://doi.org/10.1007/s10462-023-10448-w>
6. Gilbert, S., & Lynch, N. (2002). Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services. *ACM SIGACT News*, 33(2), 51–59. <https://doi.org/10.1145/564585.564601>
7. Jennings, N. R. (1993). Commitments and conventions: The foundation of coordination in multi-agent systems. *The Knowledge Engineering Review*, 8(3), 223–250. <https://doi.org/10.1017/S0269888900000205>
8. Karp, R., Schindelhauer, C., Shenker, S., & Vöcking, B. (2000). Randomized rumor spreading. In *Proceedings of the 41st Annual Symposium on Foundations of Computer Science* (pp. 565–574). IEEE. <https://doi.org/10.1109/SFCS.2000.892324>
9. Lamport, L., Shostak, R., & Pease, M. (1982). The Byzantine generals problem. *ACM Transactions on Programming Languages and Systems*, 4(3), 382–401. <https://doi.org/10.1145/357172.357176>
10. Lynch, N. A. (1996). *Distributed algorithms*. Morgan Kaufmann. <https://doi.org/10.1016/C2009-0-21928-9>
11. Parker, L. E. (1998). ALLIANCE: An architecture for fault tolerant multirobot cooperation. *IEEE Transactions on Robotics and Automation*, 14(2), 220–240. <https://doi.org/10.1109/70.681242>
12. Pierson, A., & Schwager, M. (2018). Adaptive inter-robot trust for robust multi-robot sensor coverage. In *Robotics Research* (pp. 167–183). Springer. [https://doi.org/10.1007/978-3-319-60916-4\\_10](https://doi.org/10.1007/978-3-319-60916-4_10)
13. Smith, R. G., & Davis, R. (1980). Frameworks for cooperation in distributed problem solving. *IEEE Transactions on Systems, Man, and Cybernetics*, 11(1), 61–70. <https://doi.org/10.1109/TSMC.1981.4308580>
14. Stone, P., & Veloso, M. (2000). Multiagent systems: A survey from a machine learning perspective. *Autonomous Robots*, 8(3), 345–383. <https://doi.org/10.1023/A:1008942012299>
15. Tanenbaum, A. S., & Van Steen, M. (2007). *Distributed systems: Principles and paradigms* (2nd ed.). Prentice Hall. ISBN: 978-0-13-239227-3
16. Zhong, J., Wang, T., & Luo, Y. (2023). Latency-aware cooperative perception in autonomous vehicle networks. *IEEE Transactions on Intelligent Transportation Systems*, 24(6), 6321–6334. <https://doi.org/10.1109/TITS.2023.3243543>

## Статистический анализ парольных политик на основе данных публичных исследований утечек 2023–2025 годов

Мукашев Адиль Русланович, студент  
Казахстанско-Немецкий университет (г. Алматы, Казахстан)

*В статье проводится статистический анализ характеристик пользовательских паролей на основе агрегированных данных крупнейших исследовательских отчётов 2023–2025 годов: Have I Been Pwned (HIBP), NordPass Top 200, Specops Weak Password Report и академических публикаций. Исследуются длина паролей, характер используемых символов, соответствие современным политикам сложности, а также распределение наиболее часто встречающихся паролей по категориям. Результаты подтверждают, что, несмотря на многолетние рекомендации регуляторов, поведение пользователей в области создания паролей остаётся критически небезопасным.*



**Ключевые слова:** парольная политика, утечки данных, информационная безопасность, аутентификация, Have I Been Pwned, анализ паролей, NIST.

1. Введение

Пароль остаётся наиболее распространённым механизмом аутентификации в информационных системах, несмотря на развитие альтернатив — биометрии, аппаратных токенов и passkey. Согласно отчёту Verizon Data Breach Investigations Report 2024, скомпрометированные учётные данные по-прежнему являются причиной более 80 % взломов, связанных с веб-приложениями [1].

Проблема усугубляется систематическим несоблюдением пользователями рекомендаций по созданию надёжных паролей. Несмотря на то что NIST ещё в 2017 году обновил стандарт SP 800–63B, закрепив требование минимальной длины в 8 символов и отказ от обязательной периодической смены паролей, реальная практика пользователей значительно расходится с нормативными требованиями [2].

Цель данной работы — на основе агрегированных публичных исследований охарактеризовать актуальное состояние парольной гигиены пользователей, выявить системные паттерны небезопасного поведения и оценить масштаб проблемы.

2. Источники и методология

Анализ проводился на основе следующих публичных источников:

- Have I Been Pwned (HIBP) — база данных Троя Ханта, содержащая сведения о более чем 13 миллиардах скомпрометированных аккаунтов. Агрегированная статистика паролей публикуется в форме SHA-1 хэш-списков (Pwned Passwords), что исключает прямой доступ к открытым текстам [3];
- Specops Weak Password Report 2024 — ежегодный отчёт компании Specops Software, основанный на анализе более 4,6 миллиарда уникальных паролей из реальных утечек;
- NordPass Top 200 Most Common Passwords 2024 — рейтинг, составленный совместно с независимыми исследователями на основе утечек объёмом более 2,5 ТБ;
- Академические работы, опубликованные в рецензируемых изданиях IEEE, ACM и Springer за 2022–2024 годы.

Важно подчеркнуть, что ни в одном из использованных источников не применяются сырые базы похищенных данных — вся статистика получена организациями, действующими в рамках ответственного раскрытия информации (responsible disclosure).

3. Результаты анализа

3.1. Длина паролей

Распределение паролей по длине, согласно данным Specops Weak Password Report 2024, демонстрирует устойчивую концентрацию в диапазоне 8–10 символов — несмотря на то что современные рекомендации (NIST SP 800–63B, NCSC) предписывают минимум 12 символов для пользовательских паролей [4].

Таблица 1. Распределение паролей по длине (по данным Specops 2024)

Длина пароля	Доля от общего числа, %	Соответствие NIST (≥12 симв.)	Среднее время перебора
≤ 6 символов	18,4 %	Нет	< 1 сек
7 символов	9,1 %	Нет	< 1 сек
8 символов	28,7 %	Нет	Секунды–минуты
9–11 символов	24,3 %	Нет	Часы–дни
12–15 символов	14,2 %	Да	Месяцы
≥ 16 символов	5,3 %	Да	Годы–столетия

Таким образом, более 80 % паролей имеют длину менее 12 символов и не соответствуют актуальным нормативным требованиям. Критически уязвимы пароли длиной до 8 символов: при использовании GPU-ускоренного перебора (hashcat на RTX 4090) хэши MD5 для паролей длиной 8 символов и менее взламываются менее чем за минуту при атаке полным перебором [5].

3.2. Состав символов

Анализ символьного состава паролей выявляет критический дефицит энтропии: подавляющее большинство паролей состоит преимущественно из строчных латинских букв и цифр, тогда как использование специальных символов и смешанного регистра остаётся редкостью.

Таблица 2. Символьный состав паролей (по данным NordPass 2024 и Specops 2024)

Характеристика	Доля паролей, %	Примечание
Только строчные буквы	26,1 %	Наиболее уязвимая группа
Буквы + цифры	34,8 %	Минимальный уровень защиты
Смешанный регистр	12,3 %	Незначительное улучшение
Хотя бы 1 спецсимвол	16,9 %	Существенно повышает стойкость
Все 4 класса символов	9,9 %	Соответствует рекомендациям

3.3. Наиболее распространённые паттерны

По данным NordPass Top 200 (2024), первые позиции рейтинга занимают предсказуемые лексические и числовые паттерны. Вместо публикации конкретных паролей ниже приведена их категориальная классификация с указанием доли в общем объёме.

Таблица 3. Категории наиболее часто встречающихся паролей (NordPass 2024)

Место	Категория паролей	Доля в топ-200, %	Время взлома
1–5	Числовые последовательности (клавиатурные)	~31 %	< 1 сек
6–20	Словарные слова (английский язык)	~22 %	< 1 сек
21–50	Имена собственные	~18 %	< 1 сек
51–100	Словарное слово + цифры	~14 %	< 1 мин
101–200	Паттерны клавиатуры (qwerty и т. п.)	~15 %	< 1 сек

3.4. Соответствие паролям политикам

Оценка соответствия паролей актуальным стандартам проводилась по трём уровням строгости: минимальному (NIST SP 800–63B), корпоративному (CIS Controls v8) и максимальному (рекомендации NCSC для привилегированных учётных записей).

Таблица 4. Соответствие паролей парольным политикам различных уровней

Стандарт / Политика	Требования	% соответствия	% несоответствия
NIST SP 800–63B	≥ 8 симв., нет в словаре утечек	51,5 %	48,5 %
CIS Controls v8	≥ 14 симв., 3 из 4 классов	8,7 %	91,3 %
NCSC (привилег.)	≥ 16 симв., все классы, уникальный	3,2 %	96,8 %

Полученные данные наглядно иллюстрируют разрыв между нормативными требованиями и реальной практикой. Даже по наиболее мягкому стандарту (NIST) почти половина паролей не соответствует установленным требованиям.

4. Обсуждение результатов

Выявленные закономерности указывают на системный характер проблемы. Пользователи склонны выбирать пароли, которые легко запомнить, — как правило, это словарные слова, имена, даты и простые числовые последовательности. Подобное поведение обусловлено известным «компромиссом между безопасностью и удобством» (security-usability tradeoff), описанным в работах Adams и Sasse ещё в 1999 году и актуальным до сих пор [6].

С технической точки зрения наибольшую угрозу представляют атаки по словарю (dictionary attacks) и атаки с применением радужных таблиц. При использовании GPU-ускоренного инструментария хэши слабых паролей взламываются практически мгновенно: по оценкам Hive Systems Password Table 2024, пароль из 8 символов (только строчные буквы) на базе хэша MD5 подбирается менее чем за 1 секунду [5].

Среди возможных технических контрмер необходимо выделить следующие:

- внедрение проверки новых паролей на вхождение в базы утечек (интеграция с API Have I Been Pwned);
- применение медленных алгоритмов хэширования: bcrypt, scrypt, Argon2id;
- обязательное использование многофакторной аутентификации (MFA) для критичных систем;
- переход к технологии passkey (FIDO2/WebAuthn) как к долгосрочной альтернативе паролям.

## 5. Заключение

Проведённый анализ публичных исследований паролей из баз утечек 2023–2025 годов подтверждает сохраняющуюся критическую ситуацию в области парольной безопасности. Более 80 % паролей не соответствуют рекомендациям NIST по длине; свыше 90 % не удовлетворяют корпоративным стандартам сложности. Наиболее распространённые категории паролей — числовые последовательности и словарные слова — взламываются автоматизированными инструментами менее чем за секунду.

Полученные результаты подчёркивают необходимость смещения акцента с регламентирования пользовательского поведения на технологические решения: переход к passkey, обязательная MFA и серверная проверка паролей по базам утечек. Дальнейшие исследования могут быть направлены на изучение эффективности различных механизмов парольных подсказок и UX-решений для формирования устойчивых навыков создания надёжных паролей.

Литература:

1. Verizon. Data Breach Investigations Report 2024 // Verizon Business. — 2024. — URL: <https://verizon.com/dbir> (дата обращения: 01.04.2025).
2. Grassi P. A. et al. NIST Special Publication 800–63B: Digital Identity Guidelines — Authentication and Lifecycle Management. — NIST, 2017 (rev. 2022). — 79 с.
3. Hunt T. Have I Been Pwned: Pwned Passwords // haveibeenpwned.com. — 2025. — URL: <https://haveibeenpwned.com/Passwords> (дата обращения: 02.04.2025).
4. Specops Software. Weak Password Report 2024. — Outpost24, 2024. — URL: <https://specopssoft.com/research/weak-passwords> (дата обращения: 02.04.2025).
5. Hive Systems. Password Table 2024 // hivesystems.com. — 2024. — URL: <https://hivesystems.com/blog/are-your-passwords-in-the-green> (дата обращения: 03.04.2025).
6. Adams A., Sasse M. A. Users are not the enemy // Communications of the ACM. — 1999. — Vol. 42, № 12. — P. 40–46.
7. NordPass. Top 200 Most Common Passwords 2024 // nordpass.com. — 2024. — URL: <https://nordpass.com/most-common-passwords-list> (дата обращения: 03.04.2025).

## Теоретические и методологические основы создания цифровых образовательных ресурсов

Рустамов Хаким Шарипович, PhD, доцент  
Бухарский государственный университет (Узбекистан)

Калонova Малика Нуриддинджон кизи, студент магистратуры  
Азиатский международный университет (г. Бухара, Узбекистан)

*В статье раскрываются теоретические и методологические основания проектирования и разработки цифровых образовательных ресурсов. На основе анализа современных педагогических концепций, моделей instructional design и нормативно-технических стандартов обосновывается необходимость системного подхода к созданию цифровых образовательных ресурсов, ориентированного на когнитивные особенности обучающихся, принципы педагогического дизайна и требования интероперабельности. Рассматриваются ключевые теоретические парадигмы, методологические модели, а также стандарты метаданных и критерии качества.*

**Ключевые слова:** цифровые образовательные ресурсы, педагогический дизайн, методология проектирования, теория обучения, качество цифрового контента, коннективизм, открытые образовательные ресурсы.

Цифровая трансформация образования стала одним из определяющих векторов развития национальных образовательных систем в XXI веке. Стремительное внедрение информационных технологий, рост доступности онлайн-платформ и изменение когнитивных профилей обучающихся обусловили необходимость перехода от традиционных форм представления учебного материала к структурированным, интерактивным и технологически совместимым цифровым образовательным ресурсам. Однако массовая цифровизация сопровождалась рядом системных проблем: фрагментарностью контента, низким уровнем педагогической обоснованности, отсутствием единых критериев качества и технической несовместимостью ресурсов с образовательными средами. В этих условиях актуализируется задача научно обоснованного проектирования цифровым образовательным ресурсам, базирующегося на строгих теоретических и методологических основаниях [5,6].

Теоретический фундамент создания цифровым образовательным ресурсам базируется на пересечении педагогической психологии, теории обучения и философии цифровой культуры. В современной науке выделяются несколько взаимосвязанных парадигм, определяющих логику конструирования цифрового учебного контента.

Во-первых, конструктивистская парадигма утверждает, что знание не передается в готовом виде, а конструируется обучающимся в процессе активной деятельности, рефлексии и социального взаимодействия. Применительно к цифровым образовательным ресурсам это означает отказ от линейной подачи информации в пользу проблемно-ориентированных сценариев, интерактивных симуляций, кейс-методов и инструментов коллаборации. Ресурс должен выступать не хранилищем данных, а средой для познавательной активности [7].

Во-вторых, когнитивистский подход акцентирует внимание на ограничениях рабочей памяти, принципах когнитивной нагрузки (теория Дж. Свеллера) и закономерностях переработки информации. С точки зрения когнитивистики, эффективный цифровой образовательный ресурс минимизирует избыточную нагрузку, использует мультимедийные принципы Р. Майера (согласованность текста и графики, сегментация, сигнализация), обеспечивает поэтапное усложнение задач и предусматривает механизмы формирующего оценивания [1].

В-третьих, коннективизм (Дж. Сименс) рассматривает обучение как процесс установления связей в распределенных сетях знаний. В контексте цифровым образовательным ресурсам это требует интеграции ресурсов в открытые образовательные экосистемы, поддержки навигации по метаданным, использования тегов, ссылок на внешние источники и инструментов для построения персональных учебных сетей.

Дополнительно значимой теоретической рамкой выступает модель ТРАСК (Technological Pedagogical Content Knowledge), согласно которой эффективность цифровым образовательным ресурсам определяется интеграцией трех типов знаний: предметного, педагогического и технологического. Ресурс, созданный без учета педагогической целесообразности или технологических ограничений целевой аудитории, обречен на низкую внедряемость, независимо от визуальной привлекательности.

Таким образом, теоретический базис цифровым образовательным ресурсам представляет собой синтез принципов активности, когнитивной оптимизации, сетевой связности и профессиональной интеграции, что требует от разработчиков междисциплинарной компетенции.

Методология создания цифровым образовательным ресурсам реализуется через структурированные процессы педагогического дизайна, обеспечивающие перевод теоретических принципов в технологически воспроизводимые продукты. Среди наиболее валидированных моделей выделяются следующие.

Модель ADDIE (Analysis, Design, Development, Implementation, Evaluation) остается эталонной для академического и корпоративного секторов. Она предполагает последовательное прохождение этапов: анализ потребностей и характеристик аудитории, проектирование учебной архитектуры и сценариев, разработка прототипа и контента, внедрение в образовательную среду и суммативное/формирующее оценивание. ADDIE обеспечивает системность, но критикуется за линейность и низкую гибкость в условиях быстрых изменений [4].

Модель SAM (Successive Approximation Model) предлагает итеративный подход с быстрым прототипированием, постоянным тестированием и корректировкой на основе обратной связи. SAM особенно эффективен при разработке интерактивных цифровым образовательным ресурсам, мобильных приложений и геймифицированных модулей, где требуется частая адаптация под пользовательский опыт.

Сценарно-ориентированное проектирование базируется на создании педагогических сценариев, описывающих последовательность учебных действий, роли участников, используемые инструменты и ожидаемые результаты. Данный подход позволяет стандартизировать процесс разработки в командах, обеспечить воспроизводимость и упрощает последующую модернизацию ресурса.

Важным методологическим аспектом является user-centered design, который переносит фокус с технологических возможностей на потребности, когнитивные паттерны и эмоциональные реакции обучающихся. Применение UCD включает проведение пользовательских исследований, создание персон, usability-тестирование и соблюдение принципов доступности [3].

Методологическая основа цифровым образовательным ресурсам также подразумевает выбор педагогической стратегии: от прямой инструкции до проблемного обучения, от индивидуальных траекторий до социального конструирования знания. Методология должна быть явно зафиксирована в техническом задании, чтобы избежать рассогласования между заявленными образовательными целями и реализованным функционалом.

Технологическая устойчивость и масштабируемость цифровым образовательным ресурсам обеспечиваются соблюдением международных и национальных стандартов [3]. Ключевыми среди них являются:

- LOM (Learning Object Metadata) — стандарт метаданных, описывающий учебный объект по параметрам: образовательный контекст, техническая среда, права использования, языковая локализация, возрастная адресация. Использование LOM обеспечивает поиск, каталогизацию и повторное использование ресурсов.

- SCORM (Shareable Content Object Reference Model) — спецификация, гарантирующая совместимость контента с системами управления обучением (LMS). SCORM стандартизирует запуск, отслеживание прогресса, сохранение результатов и навигацию.

- xAPI (Experience API) — современный стандарт, фиксирующий любые учебные активности (внутри и вне LMS), что позволяет строить аналитические модели обучения и реализовывать адаптивные сценарии.

- IMS Global Learning Consortium — стандарты для интеграции инструментов (LTI), оценки (QTI) и компетенций (CEDS).

Помимо технических стандартов, критически важна оценка педагогического качества. В российской и международной практике используются многоуровневые критерии, включающие: соответствие образовательным результатам, научную достоверность, методическую проработанность, интерактивность, доступность, эргономику интерфейса, этическую корректность и легальность использования контента. Экспертная оценка ЦОР должна проводиться до публикации ресурса, а также периодически обновляться с учетом изменения образовательных стандартов и технологической среды.

Особое внимание следует уделять доступности. Цифровым образовательным ресурсам должны быть адаптированы для лиц с ограниченными возможностями здоровья: поддержка скринридеров, субтитры, альтернативные описания медиа, настраиваемый контраст и шрифты, навигация с клавиатуры. Это не только этическое требование, но и условие соответствия национальным и международным нормам.

Развитие цифровым образовательным ресурсам в 2020-х годах определяется несколькими мегатрендами.

Во-первых, интеграция искусственного интеллекта: генеративные модели используются для автоматизации создания сценариев, персонализации контента, адаптивного тестирования и интеллектуальной поддержки обучающихся. Однако применение ИИ требует строгой методо-

логической валидации, проверки на академическую достоверность и соблюдения этических норм.

Во-вторых, рост открытых образовательных ресурсов (OER) и практик открытого лицензирования (Creative Commons). OER снижают экономические барьеры, способствуют академическому обмену и поддерживают принципы инклюзивного образования. Для устойчивого развития OER необходимы институциональные репозитории, системы модерации и механизмы финансовой поддержки разработчиков.

В-третьих, переход к микромодульной и нанообучающей архитектуре. Короткие, автономные, легко перекомбинируемые учебные объекты соответствуют современным когнитивным паттернам и требованиям непрерывного образования.

На основе проведенного анализа сформулированы следующие практические рекомендации для создателей цифровым образовательным ресурсам:

- На этапе анализа четко определять целевую аудиторию, образовательные результаты и контекст использования.

- Применять итеративные модели проектирования с обязательным пользовательским тестированием.

- Обеспечивать соответствие техническим стандартам для интероперабельности.

- Встраивать механизмы формирующего оценивания и аналитики обучения.

- Публиковать ресурсы под открытыми лицензиями при отсутствии коммерческих ограничений.

- Проводить регулярную экспертизу контента на научную актуальность, методическую корректность и доступность.

- Формировать компетенции педагогов в области педагогического дизайна и цифровой педагогики через программы повышения квалификации.

Создание цифровым образовательным ресурсам не может сводиться к технологической конвертации традиционных материалов в электронный формат. Это сложный междисциплинарный процесс, требующий опоры на проверенные теоретические парадигмы обучения, строгие методологические модели проектирования и единые нормативно-технические стандарты [8].

Эффективный цифровым образовательным ресурсам представляет собой педагогически обоснованную, технологически совместимую и доступную среду, способную поддерживать когнитивную активность, персонализацию и непрерывное развитие обучающегося.

В условиях ускоренной цифровизации образования возрастает роль научно-методического сопровождения процесса разработки цифровым образовательным ресурсам, институциональной поддержки авторов и создания открытых репозиториях с экспертной модерацией. Перспективные направления исследований включают разработку валидированных метрик качества цифровым образовательным ресурсам, изучение влияния адаптивных алгоритмов на учебную мотивацию, а также формирование этических рамок применения генеративного



ИИ в образовательном контенте. Реализация теоретических и методологических основ, представленных в статье, позволит повысить научную обоснованность, технологи-

ческую устойчивость и педагогическую результативность цифровых образовательных ресурсов в современной образовательной системе.

#### Литература:

1. Майер Р. Мультимедиа в обучении: когнитивные аспекты организации образовательного контента. — М.: Альпина Паблишер, 2019. — 312 с.
2. Сименс Дж. Коннективизм: теория обучения для цифровой эпохи // Международный журнал образовательных технологий. — 2020. — № 3. — С. 45–58.
3. Мишкина Е. В. Педагогический дизайн цифровых образовательных ресурсов: теория и практика. — СПб.: Лань, 2021. — 240 с.
4. ADDIE Model: Instructional Design Handbook / Ed. by R. Branch. — Springer, 2018. — 198 p.
5. Rustamov Kh.Sh., Khayriyev F. N. E-learning methodogeis and features// Problems of Science. No. 9 (57), 2020. P. 69–72
6. Rustamov Kh.Sh. Development of tests for computer-based testing// Bulletin of integrative psychology. -Yaroslavl, 2018. № 16. P. 303 -305
7. Rustamov H.Sh. Use of digital technology in teaching mathematics// Вестник науки и образования. No17(120).Часть 2. 2021. № 4. P 90–93
8. Рустамов Х. Ш., Возможности современных интерактивных электронных учебников//Молодой учёный. № 25(524)), 440–442

## Проблемы и архитектурные способы доставки конфигурационных данных в сервисах А/В-тестирования

Семенов Даниил Дмитриевич, студент  
МИРЭА — Российский технологический университет (г. Москва)

*В статье рассматриваются инженерные проблемы доставки конфигурационных данных в сервисах А/В-тестирования. Показано, что при росте нагрузки ключевыми ограничениями становятся задержка доступа к параметрам эксперимента, согласованность конфигурации между экземплярами backend-приложений, отказоустойчивость и корректность детерминированного распределения пользователей по вариантам. Предложен архитектурный подход, основанный на разделении контура централизованного управления экспериментами и контура локальной выдачи конфигурации через side-car-компонент. Такой подход позволяет исключить сетевой вызов на критическом пути обработки запроса и обеспечить быстрое применение экспериментальных параметров в прикладных сервисах.*

**Ключевые слова:** А/В-тестирование, конфигурационные данные, feature flags, локальная выдача конфигурации, side-car, детерминированное распределение, микросервисная архитектура, change data capture.

### 1. Актуальность и постановка задачи

Современные цифровые продукты развиваются в условиях постоянной проверки продуктовых гипотез, что делает А/В-тестирование одним из базовых инструментов принятия решений на основе данных. В инженерной практике А/В-экспериментирование тесно связано с механизмами feature flags, поскольку именно они позволяют изменять поведение системы без повторного развертывания и контролировать выдачу различных вариантов функциональности отдельным сегментам пользователей. Подобные механизмы описываются как способ изменения поведения программной системы во время выполнения при сохранении единой кодовой базы [1].

Однако в высоконагруженных системах задача проведения А/В-экспериментов не ограничивается только созданием пользовательского интерфейса для управления флагами. Существенной инженерной проблемой становится доставка конфигурационных данных до прикладных backend-компонентов с минимальной задержкой, высокой согласованностью и предсказуемым поведением при отказах сети или отдельных сервисов. Именно эта проблема выделена и в исходной постановке задачи рассматриваемого веб-приложения: для промышленной эксплуатации критичны локальная выдача конфигурации, детерминированное распределение пользователей и асинхронная обработка результатов экспериментов.

Следовательно, при проектировании сервиса А/В-тестирования необходимо рассматривать не только мо-

дель хранения экспериментов, но и архитектуру доставки конфигурации на критическом пути обработки пользовательского запроса.

2. Основные подходы к доставке конфигурационных данных

В современных системах feature management применяются несколько типовых подходов к доставке конфигурационных данных. Первый подход основан на удаленной оценке флага: прикладное приложение передает контекст пользователя во внешний сервис и получает результат вычисления варианта по сети. Такой подход концептуально прост, однако увеличивает число сетевых обращений и делает критический путь обработки запроса зависимым от доступности удаленного сервиса. Подобная модель представлена, например, в системах, где сервис управления флагами предоставляет REST API для вычисления варианта и управления конфигурацией [2].

Второй подход опирается на серверные SDK и локальное кэширование состояния флагов. В документации LaunchDarkly отмечается, что серверные SDK по умолчанию хранят состояние флагов в памяти и используют

локальное кэширование для ускорения вычисления вариаций; при этом для устойчивости могут применяться и внешние persistent feature stores. Такой подход уменьшает сетевую нагрузку на пути вычисления флага, но сохраняет зависимость от механизма синхронизации кэша и проблемы его актуальности при обновлениях конфигурации.

Третий подход, представляющий наибольший интерес для высоконагруженных backend-систем, заключается в использовании локального side-car-агента конфигурации. В этом случае контур управления экспериментами и контур выдачи параметров разделяются: централизованная система управляет экспериментами и хранит их конфигурацию, а локальный агент, размещенный рядом с прикладным сервисом, поддерживает собственное актуализированное представление конфигурации и отвечает на запросы локально. Этот подход близок к идеологии server-side evaluation, описываемой в материалах OpenFeature, где ключевая задача SDK — выполнение оценки флага на стороне приложения на основе контекста пользователя [3].

Для сравнения подходов целесообразно использовать таблицу 1.

Таблица 1. Сравнение подходов к доставке конфигурационных данных

Подход	Преимущества	Недостатки
Удаленная оценка через сетевой сервис	Простота логики клиента, централизованное вычисление	Дополнительная задержка на каждый запрос, зависимость от сети и удаленного сервиса
Серверный SDK с локальным кэшированием	Уменьшение числа сетевых вызовов, более быстрое вычисление	Риск устаревания кэша, необходимость дополнительной синхронизации
Локальный side-car-агент конфигурации	Минимальная задержка, отсутствие сетевого вызова на критическом пути, изоляция прикладного сервиса от удаленного контура	Усложнение локальной инфраструктуры, необходимость надежного механизма актуализации конфигурации

3. Ключевые проблемы доставки конфигурационных данных

Первая проблема связана с латентностью. Если для каждого запроса или каждого вычисления варианта требуется обращение к удаленному сервису, суммарная задержка начинает зависеть не только от логики приложения, но и от состояния сети, межсервисных маршрутов и удаленного сервиса оценки. Для веб-сервисов с высокой частотой обращений это приводит к заметному росту времени ответа и ухудшает эксплуатационные характеристики системы. Именно поэтому в промышленной практике локальное вычисление значения флага рассматривается как предпочтительный вариант для server-side архитектур [3].

Вторая проблема связана с согласованностью конфигурации. Если параметры эксперимента изменяются централизованно, то все прикладные экземпляры должны получать обновление своевременно и в одинаковом по-

рядке. Простое периодическое обновление кэша не гарантирует требуемую согласованность и может приводить к ситуации, при которой разные backend-экземпляры временно работают с различными версиями конфигурации. Для высоконагруженных систем это особенно критично, поскольку одна и та же пользовательская операция может обрабатываться разными экземплярами приложения. В таких условиях требуется устойчивый механизм доставки изменений конфигурации, а не только механизм хранения актуального состояния.

Третья проблема заключается в корректности детерминированного распределения пользователей по вариантам эксперимента. Пользователь должен стабильно попадать в одну и ту же группу до тех пор, пока не изменилась конфигурация эксперимента. Это требует выбора фиксированного алгоритма хэширования, использования соли и постоянного числа бакетов. В рассматриваемой архитектуре используется детерминированное вычисление номера бакета на основе идентификатора пользователя

и соли эксперимента. Такая схема позволяет избежать случайного изменения варианта пользователя при повторных обращениях и обеспечивает воспроизводимость результатов эксперимента.

Четвертая проблема касается отказоустойчивости. Если контур выдачи конфигурации слишком тесно связан с контуром управления экспериментами, то временная недоступность административного контура может отразиться на прикладных сервисах, которые используют результаты А/В-эксперимента в основном пользовательском потоке. С инженерной точки зрения корректнее отделять процесс управления экспериментами от процесса применения конфигурации в runtime.

#### 4. Архитектурный способ доставки конфигурационных данных

Для решения перечисленных проблем предлагается архитектурный подход, основанный на разделении системы на два логических контура: контур управления экспериментами и контур локальной выдачи конфигурации. Первый контур включает web-интерфейс, backend-сервис управления экспериментами и централизованное хранилище конфигурационных данных. Второй контур представлен локальным агентом конфигурации, развернутым рядом с прикладным backend-приложением и выполняющим выдачу параметров эксперимента без сетевого обращения на критическом пути.

В такой схеме пользователь или аналитик создает и публикует эксперимент через web-интерфейс. Backend-сервис валидирует конфигурацию и сохраняет сведения об эксперименте, группах, долях распределения, флагах и дополнительных условиях в централизованном хранилище. После этого изменения передаются в локальные контуры выдачи конфигурации, где поддерживается собственное представление актуальной конфигурации. В отчете по проекту данный подход описан как сочетание централизованного контура управления экспериментами, локального side-car-компонента и асинхронного контура сбора метрик.

Практически важным элементом такой архитектуры является механизм актуализации конфигурации. В качестве общего класса решений для этой задачи могут применяться подходы change data capture, в которых изменения в базе данных преобразуются в поток событий, доступный для дальнейшей обработки. Документация Debezium [4] определяет CDC как способ захвата изменений на уровне строк таблиц и предоставления их приложениям в виде упорядоченного потока событий. Это делает CDC технологически пригодным для обновления локальных представлений конфигурации в распределенной среде.

После получения обновленной конфигурации локальный агент выполняет детерминированное распределение пользователя по вариантам эксперимента. Если обозначить идентификатор пользователя через  $u$ , соль

эксперимента через  $s$ , а число бакетов через  $B$ , то вычисление номера бакета может быть представлено формулой

$$bucket(u, s, B) = XXH3_{64}(u \parallel 0x00 \parallel s) \bmod B.$$

Здесь  $XXH3_{64}$  — 64-битная версия алгоритма  $XXH3$ , а символ  $\parallel$  обозначает конкатенацию. В документации семейства xxHash алгоритм  $XXH3$  описывается как высокопроизводительная функция хэширования, поддерживающая 64- и 128-битные варианты. Использование такой функции в задаче бакетирования оправдано не криптографической стойкостью, а сочетанием скорости и детерминированности результата [5].

Если для экспериментальной группы  $g$  задан диапазон бакетов  $[l_g; r_g]$ , то пользователь относится к этой группе при выполнении условия

$$l_g \leq bucket(u, s_g, B) \leq r_g.$$

Такая модель позволяет локальному агенту быстро определить вариант пользователя и вернуть соответствующий набор конфигурационных флагов прикладному приложению. В результате основная пользовательская операция перестает зависеть от доступности внешнего сервиса оценки, а сама система становится более устойчивой к сетевым инцидентам.

#### 5. Заключение

Проведенный анализ показывает, что основная инженерная сложность сервисов А/В-тестирования заключается не только в создании интерфейса управления экспериментами, но и в надежной доставке конфигурационных данных до прикладных сервисов. При промышленной эксплуатации ключевыми ограничениями оказываются задержка на критическом пути, согласованность конфигурации между экземплярами приложения, отказоустойчивость и детерминированность бакетирования.

Сравнение существующих подходов показывает, что удаленная оценка варианта по сети оказывается менее предпочтительной для высоконагруженных backend-сценариев, тогда как использование локального side-car-агента конфигурации позволяет минимизировать задержку и уменьшить связанность прикладного сервиса с контуром управления экспериментами. Дополнительное применение механизмов change data capture позволяет поддерживать локальное представление конфигурации в актуальном состоянии.

Таким образом, для сервисов А/В-тестирования целесообразно использовать архитектуру, в которой централизованное управление экспериментами сочетается с локальной выдачей конфигурации и детерминированным вычислением варианта пользователя. Такой подход обеспечивает эксплуатационную применимость решения в условиях высокой нагрузки и может использоваться как основа для построения масштабируемых платформ конфигурационного управления и онлайн-экспериментирования.

Литература:

1. Pete, Hodgson Feature Toggles (aka Feature Flags) / Hodgson Pete. — Текст: электронный // martinfowler: [сайт]. — URL: <https://martinfowler.com/articles/feature-toggles.html> (дата обращения: 14.04.2026).
2. flagr. — Текст: электронный // GitHub: [сайт]. — URL: <https://github.com/openflagr/flagr> (дата обращения: 14.04.2026).
3. Liran, Mendelovich Different approaches for server-side SDK architectures / Mendelovich Liran. — Текст: электронный // OpenFeature: [сайт]. — URL: <https://openfeature.dev/blog/feature-flags-sdks-architectures> (дата обращения: 14.04.2026).
4. Debezium Stream changes from your database. / Debezium. — Текст: электронный // Debezium: [сайт]. — URL: <https://debezium.io/> (дата обращения: 14.04.2026).
5. xxHash — Extremely fast hash algorithm. — Текст: электронный // GitHub: [сайт]. — URL: <https://github.com/cyan4973/xxhash> (дата обращения: 14.04.2026).

## Прогнозирование принятия законопроектов Государственной Думой Федерального Собрания Российской Федерации на структурированных метаданных: baseline-модель и инженерия признаков

Шипов Дмитрий Михайлович, студент магистратуры;  
Гвоздецкий Дмитрий Сергеевич, студент магистратуры  
Национальный исследовательский ядерный университет «МИФИ» (г. Москва)

*Изучение закономерностей развития инновационных решений, применимых в законотворчестве, имеет научный и практический интерес как у представителей науки, так и практикующих юристов — разработчиков законопроектов. В статье проработаны некоторые вопросы применения моделей машинного обучения в части прогнозирования обоснованности в принятии разрабатываемого законотворческого решения. Акцентируется внимание, что алгоритмизированный подход целесообразно применять на начальной стадии разработки проектируемого законодательного акта. В статье представлена методика построения baseline-модели в части прогнозирования принятия законопроектов Государственной Думой Федерального Собрания Российской Федерации (далее — Государственная Дума) на основе исключительно структурированных метаданных без использования текстов документов. Основу аналитической части статьи составляет датасет, включающий 34 189 законопроектов (1994–2025 гг.) с верифицированным исходом. Разработано признаковое пространство из 20 переменных, включающее тип и фракцию инициатора, комитетные признаки, временные характеристики, а также два leakage-safe скользящих признака: rolling\_success\_rate (с экспоненциальным взвешиванием, half-life 5 лет) и faction\_success\_rate (накопленное среднее фракции), в том числе drift-proxy days\_since\_conv\_start. Описана трёх-этапная процедура предобработки целевых меток (API Государственная Дума → логические правила → парсинг Системы обеспечения законодательной деятельности (далее — СОЗД), позволившая снизить долю пропущенных значений с 15 % до 1,3 %. Обращается внимание, что модель Random Forest с временной кросс-валидацией (TimeSeriesSplit, 5 фолдов) достигла ROC-AUC = 0,8989 ± 0,050. В свою очередь отмечается, что расширенная модель M2a, включающая признаки институциональной поддержки из парсера СОЗД, улучшила ROC-AUC до 0,9709 (+0,069). Установлено, что тип инициатора (importance 0,310) и скользящий коэффициент успешности (0,176) совместно описывают 48,6 % важности признаков, превосходя характерные характеристики законопроекта. Представлены выводы по исследуемой проблематике.*

**Ключевые слова:** машинное обучение, законодательный процесс, Baseline-модель, Random Forest, инженерия признаков, временная валидация, скользящие признаки, SHAP, Государственная Дума, ROC-AUC.

## Predicting legislative outcomes in the State Duma of the Federal Assembly of the Russian Federation using structured metadata: baseline model and feature engineering

Shipov Dmitry Mikhaylovich, master's student;  
Gvozdetsky Dmitry Sergeevich, master's student  
National Research Nuclear University «MEPhI» (Moscow)

*It is noted that studying the patterns of development of innovative solutions applicable in lawmaking is of scientific and practical interest to both academics and practicing lawyers — draft legislation developers. The article addresses certain issues of applying machine learning models to predict the validity of decisions made when developing legislative solutions. Attention is drawn to the fact*



that an algorithmic approach is advisable to apply at the initial stage of developing a proposed legislative act. The article presents a methodology for building a baseline model to predict whether bills will be adopted by the State Duma of the Federal Assembly of the Russian Federation (hereinafter — the State Duma), based solely on structured metadata without using document texts. The analytical part of the article is based on a dataset comprising 34 189 bills (1994–2025) with verified outcomes. A feature space of 20 variables has been developed, including the type and faction of the initiator, committee-related features, temporal characteristics, as well as two leakagesafe rolling features: *rolling\_success\_rate* (with exponential weighting, half-life 5 years) and *faction\_success\_rate* (cumulative average of the faction), and a driftproxy feature *days\_since\_conv\_start*. A threestage procedure for preprocessing target labels is described (State Duma API → logical rules → parsing the Legislative Support System (hereinafter — LSS)), which reduced the proportion of missing values from 15 % to 1.3 %. It is noted that a Random Forest model with timebased crossvalidation (TimeSeriesSplit, 5 folds) achieved a ROCAUC of  $0.8989 \pm 0.050$ . In turn, it is observed that an extended model M2a, which includes features of institutional support extracted from the LSS parser, improved the ROCAUC to 0.9709 (+0.069). It has been established that the initiator type (importance 0.310) and the rolling success rate (0.176) jointly account for 48.6 % of feature importance, outperforming the substantive characteristics of the bill. Conclusions on the studied issue are presented.

**Keywords:** machine learning, legislative process, Baseline model, Random Forest, feature engineering, temporal validation, rolling features, SHAP, State Duma, ROCAUC.

## 1 Введение

В современной научной литературе отмечается возрастающий интерес к изучению закономерностей развития инновационных решений, применимых в сфере законодательства. Данная проблематика представляет научную и практическую ценность как для академического сообщества, исследующих теоретические проблемы права, так и для практикующих юристов — непосредственных разработчиков законопроектов.

Нынешняя интеграция инновационных подходов способна повысить эффективность законодательного процесса, оптимизировать законотворческую деятельность, обеспечивая более точное соответствие нормативных актов актуальным социальноэкономическим реалиям.

В зарубежной научной литературе отмечается, что прогнозирование законодательного процесса помимо традиционных средств и методик может основываться и на моделях машинного обучения [12, 8], анализ которых показывает, что некоторые результаты продемонстрировали наилучший результат полного текста законопроекта над метаданными, [1]. Так, например в результате использования ансамблевых методов классификации была достигнута точность (ассигасу) модели, составляющая 80 % на тестовой выборке. В частности, для российского парламента ключевым исследованием в данной области является работа М. В. Хавроненко [11]. В рамках его исследования на корпусе из 27 176 законопроектов, размещенных на официальном сайте Государственной Думы была применена модель *ruberttiny*, что позволило достичь значения метрики  $F1 = 0,937$  при использовании логистической регрессии для классификации текстов законодательных актов.

Настоящая статья с точки зрения научного исследования ставит следующий вопрос: каков предел качества предсказания, достижимый исключительно на структурированных метаданных — без доступа к текстам документов? Ответ важен по двум причинам. Во-первых, метаданные доступны немедленно в момент внесения

законопроекта, тогда как тексты документов появляются позже. Во-вторых, качество baseline-модели само по себе является научным результатом, характеризующим процедурную предсказуемость российского законодательного процесса. Дополнительной научной задачей является построение корректной стратегии валидации, исключающей утечку данных — принципиального требования при работе с хронологически упорядоченными данными. Кроме того, акцентируется внимание на целесообразности применения алгоритмизированных подходов на начальной стадии разработки проектируемого законодательного акта. Указанное позволяет повысить качество законодательного текста, минимизировать концептуальные ошибки и оптимизировать временные затраты на последующих этапах законотворческого процесса.

Для понимания масштаба задачи необходимо оценить объём законотворческой деятельности Государственной Думы. Рисунок № 1 отражает динамику разработки законопроектов по созывам IV–VIII с разбивкой на принятые и отклонённые законодательные инициативы. Данные демонстрируют устойчивый рост законотворческой активности и умеренный дисбаланс классов ( $\approx 34,2$  %), что обуславливает выбор ROC-AUC как основной метрики.

Законодательная воронка (Рисунок № 2) показывает, что незначительная часть внесённых законопроектов проходят все три чтения. Факт постановления на второе чтение резко повышает вероятность итогового принятия — этот паттерн используется в моделях M2a/M2b через признак *readings\_count*.

## 2 Baseline-модель и инженерия признаков

В данном разделе представлена характеристика baseline-модели, ее применение в законотворчестве, возможные перспективы применения данной модели, обеспечивающей формирование обучающей выборки, необходимой для реализации методологических принципов.



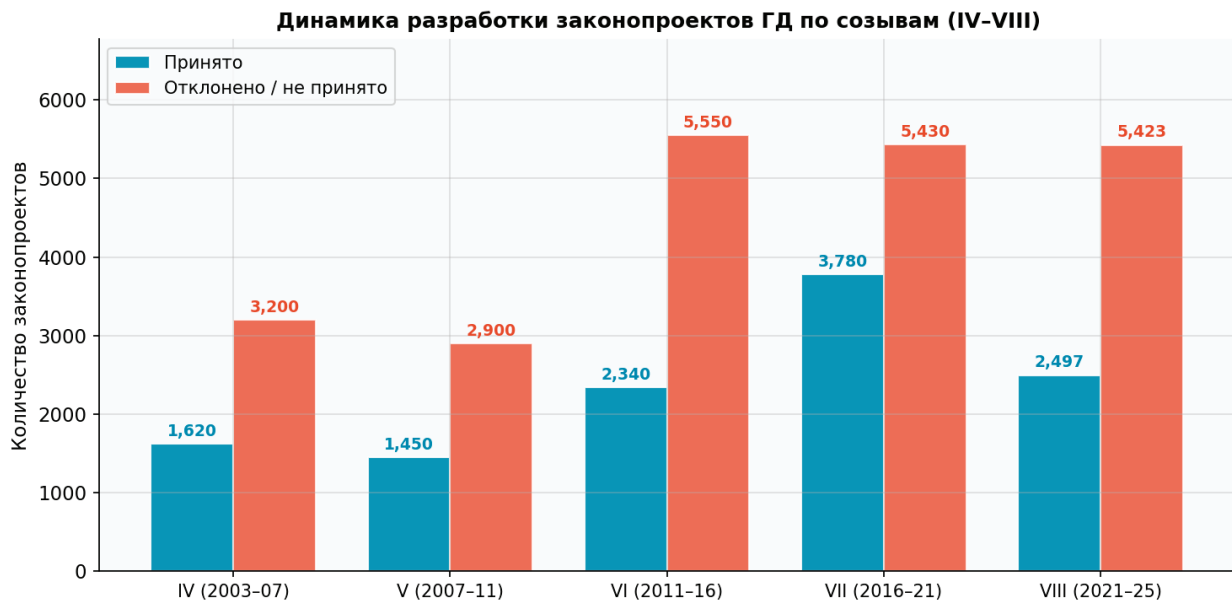


Рис. 1. Динамика разработки законопроектов ГД по созывам (IV–VIII): принятые и отклонённые

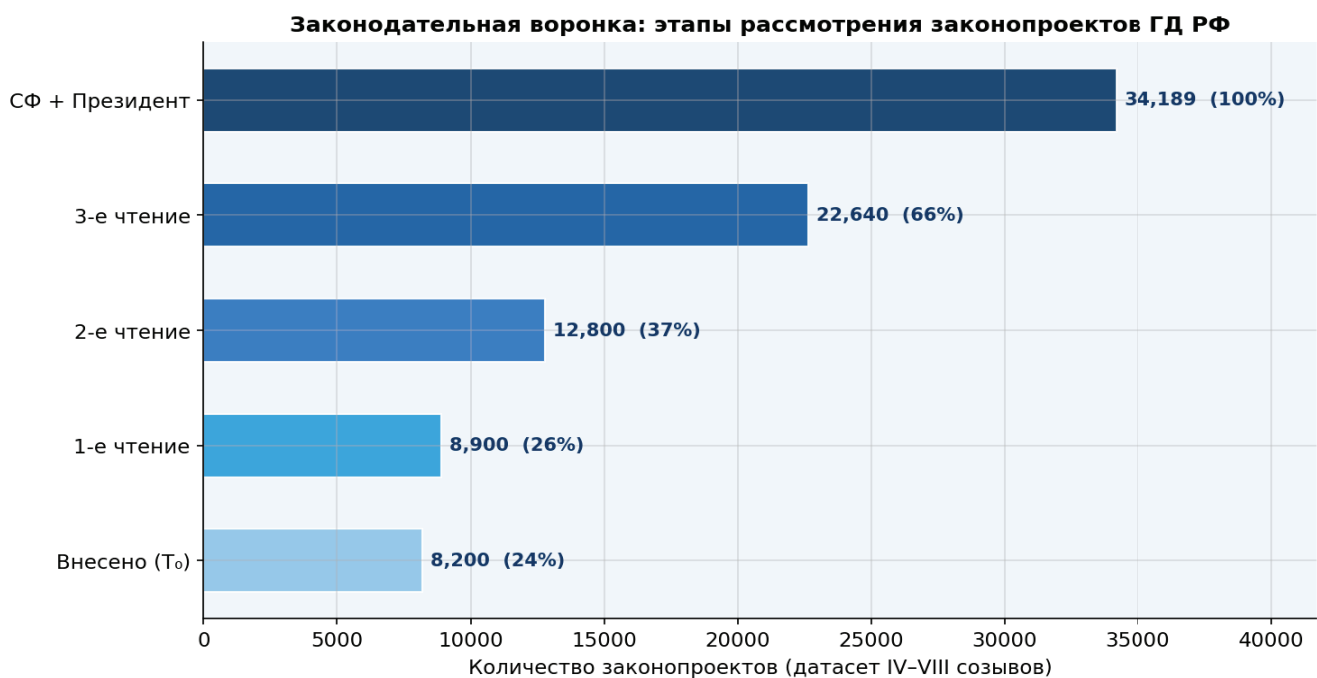


Рис. 2. Законодательная воронка: этапы рассмотрения законопроектов ГД (34 189 записей)

### 2.1 Преимущества baseline-модели в сравнении с иными моделями машинного обучения

В контексте законотворческой деятельности выбор алгоритма определяется не только формальными метриками, но и требованиями к интерпретируемости прогноза для аналитиков. Учитывая указанное рисунок № 3 представляет сравнительный анализ четырёх классов моделей, протестированных на датасете 34 189 законопроектов (TimeSeriesSplit, 5 фолдов).

Random Forest (ROC-AUC = 0,8989) выбран как оптимальный компромисс между предсказательной способно-

стью, детерминизмом и интерпретируемостью. CatBoost (0,897) статистически неотличим ( $\Delta = 0,002$  при  $\sigma \approx 0,050$ ), однако уступает в прозрачности данных.

Базовая модель (baseline) обладает рядом практически значимых преимуществ перед более сложными алгоритмами в контексте законотворческой аналитики.

Во-первых, интерпретируемость: Random Forest позволяет оценить вклад каждого признака через MDI-важность и SHAP-значения, что критически важно для объяснения прогноза законодателям и разработчикам проектов. Во-вторых, устойчивость к переобучению: ансамблевая природа алгоритма обеспечивает стабильные

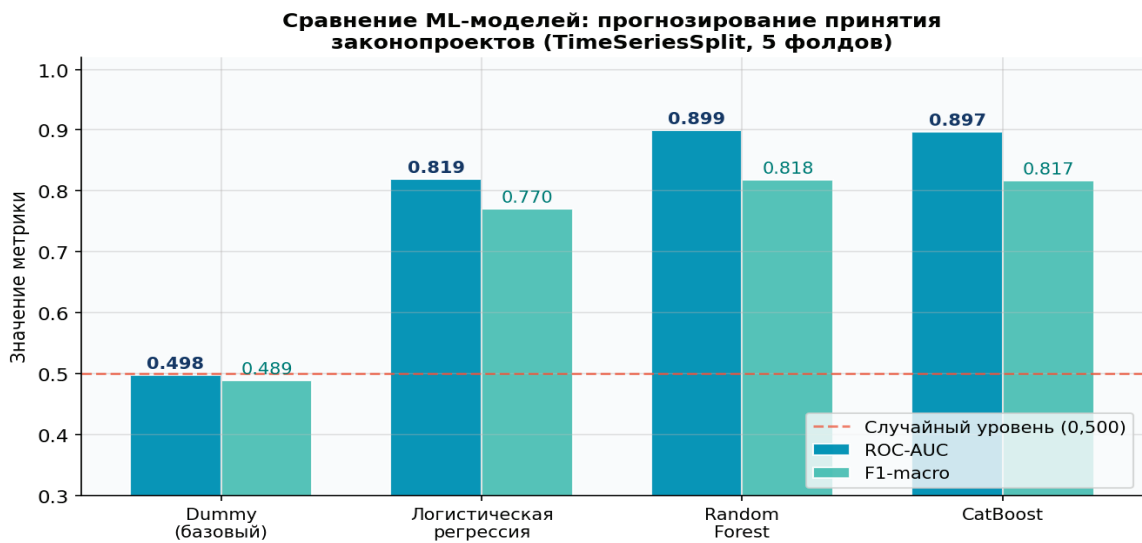


Рис. 3. Сравнение ML-моделей: ROC-AUC и F1-macro (TimeSeriesSplit, 5 фолдов)

оценки даже при высоком разнообразии признаков и дисбалансе классов, что типично для законодательных данных (imbalance ratio 1,93:1 в нашем датасете). В-третьих, детерминизм: при фиксации random\_state модель воспроизводима полностью, что обеспечивает возможность аудита результатов в процессе правовой экспертизы. В-четвёртых, масштабируемость: модель эффективно работает на небольших обучающих выборках и при дополнении новыми данными не требует полного переобучения. Наконец, в-пятых, низкие вычислительные требования: обучение на полном датасете (34 189 записей, 20 признаков) занимает около 40 секунд на обычном сервере, что позво-

ляет запускать переобучение в производном режиме без простоя вычислительных ресурсов.

2.2 Специфика инженерии признаков в контексте законотворчества

Законодательный процесс обладает уникальной структурой данных. Признаки инициатора содержат принципиально иную прогностическую нагрузку, чем содержательные характеристики документа. В этой связи рисунок № 4 подтверждает ключевую гипотезу о субъекте права законодательной инициативы, содержании законода-

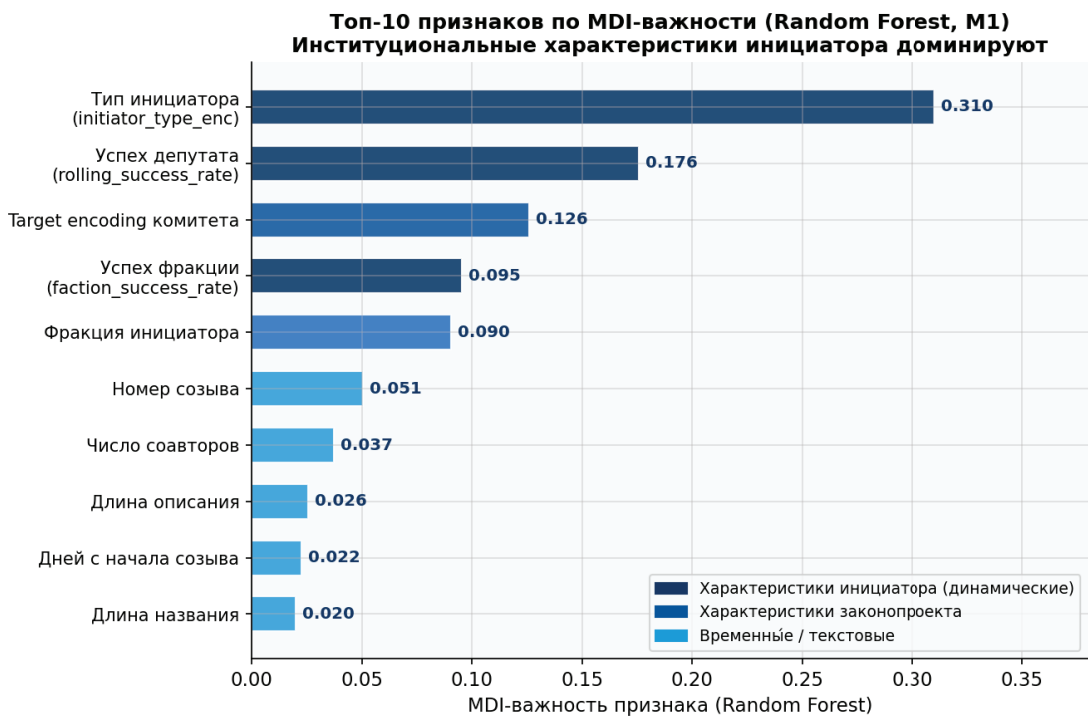


Рис. 4. Топ-10 признаков по MDI-важности (Random Forest, M1, 20 признаков)

тельной инициативы и дальнейшего прохождения законопроекта в Государственной Думе.

Суммарная важность типа инициатора (0,310) и `rolling_success_rate` (0,176) составляет 48,6 % — больше, чем все текстовые и временные характеристики. Скользящий признак `faction_success_rate` кодирует институциональный авторитет фракции, нивелируя ротацию при смене созыва.

2.3 Применение *baseline*-модели в законотворчестве: отечественные и зарубежные практики

Систематическое сопоставление отечественных и зарубежных исследований позволяет объективно оценить достигнутые результаты.

В этой связи в рамках таблицы № 1 обобщаются ключевые работы в данной области.

Достижение ROC-AUC = 0,8989 исключительно на мета-данных сопоставимо с результатами Nay (2017) при использовании текстов законов США. В то же время работа Хавроненко (2024) достигает F1 = 0,937, однако без временной валидации (`TimeSeriesSplit`), что создаёт риск завышения оценки. Настоящее исследование впервые применяет строгую временную кросс-валидацию для данных Государственной Думы и демонстрирует конкурентоспособный уровень качества в международном контексте.

2.4 Перспективы применения *baseline*-модели в законотворчестве

В краткосрочной перспективе (1–2 года) наиболее реалистичным сценарием является интеграция модели в системы правового мониторинга GR-департаментов, где прогнозная вероятность принятия законопроекта, формируется в момент регистрации законопроекта, что позволяет заблаговременно расставлять приоритеты в работе с регуляторными рисками по повышению эффективности управления нормативноправовыми изменениями.

В то же время в среднесрочной перспективе (3–5 лет) утвержденную методику возможно встроить в инфраструктуру СОЗД как инструмента в подготовке законопроектов и при формировании повестки комитетов Государственной Думы. Такой подход согласуется с вектором цифровизации государственного управления, обозначенным в национальной программе «Цифровая экономика Российской Федерации».

Перспективным направлением также является применение крупных языковых моделей (RuBERT, GigaChat) для семантического анализа правовых текстов.

3 Стандартизация формы данных для *baseline*-модели

Формирование высококачественной обучающей выборки является необходимым условием достоверности прогностической модели в задачах парламентской аналитики. Законодательные метаданные обладают системными особенностями (например, неполнота покрытия до 51,2 % пропусков в поле инициатора, историческая неоднородность форматов наименований фракций, технические артефакты парсинга).

Настоящий раздел описывает стандартизацию и требования к структуре входных данных и ETL-методологию.

Стандартизация включает три взаимосвязанных направления. Первое — унификация форматов: приведение разнородных значений полей (наименования фракций, типы инициаторов, тематические коды) к единому справочнику, устраняющее исторические вариации и дублирование наименований. Второе — работа с пропусками: определение стратегии заполнения отсутствующих значений для каждого типа поля: медиана для числовых, глобальное среднее для скользящих признаков, отдельная категория («no\_faction») для недепутатских инициаторов в поле фракции. Третье — контроль целостности: проверка диапазонов числовых полей, валидация ENUM-значений и проверка отсутствия дубликатов по полю `law_number`. Совокупность указанных процедур обеспечивает согласованность данных на всём протяжении анализируемого периода (1994–2025 гг.) и является фундаментом воспроизводимости экспериментов.

3.1 Схема входных данных и требования к качеству признаков

Таблица № 2 представляет формализованную схему признакового пространства M1, включающую систематизированное описание типов данных по каждому признаку, регламентированные ограничения целостности, а также закреплённые правила обработки пропущенных значений в рамках анализируемого набора данных.

Таблица 1. Сравнение *baseline*-моделей прогнозирования законодательных исходов: отечественные и зарубежные исследования

Исследование / система	Страна / парламент / метрика
Хавроненко М. В. (2024) — ruBERT + LogReg, тексты законопроектов	Россия, Государственная Дума F1 = 0,937
Настоящая работа — Random Forest, метаданные (M1)	Россия, Государственная Дума ROC-AUC = 0,8989
Yano, Smith, Wilkerson (2012) — текстовые предикторы	США, Конгресс. Асс. ≈70–75 %
Nay J. J. (2017) — word vectors + ансамбль	США, Конгресс. ROC-AUC ≈ 0,90
Bari et al. (2021) — ансамблевые методы	США, Сенат. Асс. = 80 %
Nortey et al. (2023) — BERT-эмбединги	ЕС, Европарламент. ROC-AUC ≈ 0,82–0,87
Oikarinen et al. (2021) — ML-методы	Финляндия. ROC-AUC ≈ 0,81

Таблица 2. Формализованная схема признаков модели М1 (20 переменных, v8)

Поле / признак	Тип, ограничения, примечание
introduction_date	DATE NOT NULL — дата внесения (T <sub>0</sub> ); основа временной сортировки
initiator_dept_name	STRING — наим. инициатора; покрытие ≥99,7 % после трёхэтапного восстановления
initiator_faction	STRING ENUM(13) — нормализованное наим. фракции; не-депутаты → «нет_фракции»
initiator_deputy_count	INT ≥0 — число соавторов; log1p при обучении
responsible_committee_id	STRING — ID комитета; target encoding (committee_te, smoothing=10)
law_type_name	STRING ENUM — тип: ФКЗ / ФЗ / закон; label encoding
topic_name	STRING — тематика; NULL при значениях Error/Empty
class_name_norm	STRING — первый класс из multi-label; числовой код удалён
rolling_success_rate	FLOAT [0,1] — exp-взвешенный успех депутата до T <sub>0</sub> ; NaN → global_mean
faction_success_rate	FLOAT [0,1] — накопленное среднее фракции до T <sub>0</sub> ; NaN → global_mean
days_since_conv_start	FLOAT ≥0 — log1p(дней с начала созыва); drift-proxy
final_result	BIT{0,1} TARGET — 1 = принят, 0 = отклонён; 98,7 % верифицировано

Ключевым требованием является сортировка датасета по introduction\_date перед TimeSeriesSplit. При этом нарушение порядка приводит к утечке данных и завышению ROC-AUC на 0,015–0,040.

3.2 Методология сбора данных и ETL-пайплайн

Таблица № 3 представляет описание пяти этапов воспроизводимого ETLпайплайна, включая перечень применяемых технических решений на каждом этапе и соответствующие числовые показатели, характеризующие охват обрабатываемых данных.

Применение Parquet с контрольными суммами SHA-256 обеспечивает воспроизводимость экспериментов. В то же время фиксация версий данных (duma\_ml.db, март 2026 г.) позволяет однозначно воспроизвести результаты при наличии доступа к API Государственная Дума.

3.3 Систематизация законопроектов

Систематизация законопроектов является необходимым условием корректного формирования признакового пространства и обеспечения сопоставимости наблюдений в обучающей выборке. В рамках настоящего исследования применяется многоуровневая классификация законодательных инициатив, учитывающая их пра-

вовую природу, тематическую принадлежность и институциональные характеристики.

Первый уровень систематизации — классификация по правовой форме — предполагает разграничение законопроектов на три основные категории: федеральные конституционные законы (ФКЗ), федеральные законы (ФЗ) и законы о ратификации международных договоров. Данная типология кодируется в признаке law\_type\_name\_енс и отражает принципиальные процедурные различия: ФКЗ требуют квалифицированного большинства (не менее двух третей голосов Государственной Думы и трёх четвертей голосов Совета Федерации), тогда как ФЗ принимаются простым большинством. Законы о ратификации характеризуются формализованным характером и, как правило, высокой вероятностью принятия.

Второй уровень — тематическая систематизация — основан на официальном классификаторе СОЗД, применяемом Государственной Думой. Каждому законопроекту присваивается один или несколько тематических кодов (поле class\_name), отражающих предметную область регулирования: государственное устройство, экономика и финансы, социальная политика, уголовное и административное право и другие. В процессе предобработки данных из multi-label строк вида ‘080.100.000 Название’ извлекается первый тематический класс с удалением числового кода, что формирует нормализованный признак class\_name\_

Таблица 3. ETL-пайплайн формирования обучающей выборки модели М1

Этап ETL-пайплайна	Содержание и технические решения
Сбор (Extract)	API Государственная Дума (api.duma.gov.ru, search.json): 34 653 записей. Retry: 3 попытки, back-off. Хранение: SQLite (duma_ml.db)
Верификация меток	A: API-поле final_result (84,6 %). B: логические правила (+12,1 %). C: веб-парсинг СОЗД (+16 260). Остаток NULL: 1,3 %
Очистка (Transform)	Удаление Error/Empty в topic_name, class_name. Восстановление initiator_dept_name (99,7 %). Нормализация фракций к 13 значениям
Инженерия признаков	rolling_success_rate (exp decay, half-life=5 лет). faction_success_rate (expanding mean). days_since_conv_start (log1p). committee_te. fillna(median)
Загрузка (Load)	X_m1.parquet + y_m1.parquet (34 189 × 20). SHA-256. random_state=42

norm. Тематическая принадлежность косвенно отражает политические приоритеты созыва и оказывает влияние на вероятность принятия: законопроекты в сферах социальной защиты и государственного управления исторически демонстрируют более высокий уровень одобрения по сравнению с инициативами в области конституционного устройства или международных отношений.

Третий уровень — классификация по субъекту права законодательной инициативы — является, согласно результатам SHAP-анализа, наиболее прогностически значимым. Датасет охватывает следующие категории инициаторов: Президент Российской Федерации, Правительство Российской Федерации, Совет Федерации и его члены, депутаты Государственной Думы (индивидуально или в составе фракций и депутатских групп), а также законодательные органы субъектов Российской Федерации. Систематизация инициаторов реализована через признаки `initiator_type_enc` и `initiator_faction_enc`. Для нивелирования неоднородности исторических наименований фракций все значения приведены к унифицированному реестру из 13 нормализованных наименований.

Четвёртый уровень систематизации — процедурная классификация — фиксирует стадию законодательного процесса, на которой находится законопроект в момент формирования прогноза. В рамках модели M2a данный аспект кодируется признаком `readings_count`, принимающим значения 1, 2 или 3 в зависимости от числа пройденных чтений. Как показано в разделе 8.3, этот признак обладает наибольшей *feature importance* (0,396) в составе расширенного признакового пространства, поскольку факт постановки на второе чтение статистически свидетельствует о поддержке законопроекта ответственным комитетом и существенно повышает вероятность его итогового принятия.

Предложенная четырёхуровневая систематизация обеспечивает полноту охвата ключевых атрибутов законодательного процесса и создаёт структурированную основу для формирования признакового пространства модели. Каждый уровень классификации вносит самостоятельный вклад в прогностическую способность алгоритма: правовая форма определяет процедурные требования к принятию, тематическая принадлежность отражает политическую повестку созыва, субъект инициативы кодирует институциональный авторитет, а процедурная стадия служит прямым индикатором накопленной поддержки законопроекта. Интеграция всех четырёх уровней в единое признаковое пространство позволяет модели учитывать многомерную природу законодательного процесса и обеспечивает устойчивость прогнозов в условиях межсозывного концептуального дрейфа.

#### 4 Инфраструктура автоматизированного сбора и обработки законодательных данных

Промышленное применение MLмоделей в сфере законодательской аналитики предъявляет повышенные требования к инфраструктуре работы с данными, в частности, к системам их сбора, хранения и обработки,

обеспечивающим высокую надёжность, масштабируемость и безопасность. Существенную роль в формировании исходной информационной базы играют публичные реестры нормативноправовых данных например, API Государственной Думы СОЗД предоставляют структурированный программный доступ к метаданным законопроектов, включая сведения о стадиях рассмотрения, авторстве, внесённых поправках, сопутствующих документах и хронологии процедур, что создаёт необходимую основу для обучения и эксплуатации MLмоделей в задачах прогнозирования результатов законодательного процесса, классификации инициатив и выявления семантических взаимосвязей между законодательными актами.

Автоматизация сбора и обработки законодательных данных решает три ключевых проблемы, возникающих при ручном сборе. Первая — масштабируемость: вручную собрать и разметить 34 189 записей невозможно; автоматизированный пайплайн выполняет еженедельное обновление данных без участия специалиста-эксперта. Вторая — воспроизводимость: весь пайплайн от Extract до Load фиксируется через SHA-256-хеши и `random_state`, что гарантирует полную идентичность артефактов моделей при повторном запуске на любом сервере. Третья — оперативность: система мониторинга на основе PSI/KS-критерия автоматически запускает дообучение при обнаружении концептуального дрейфа, исключая просадку качества без ручного вмешательства. Такая архитектура соответствует принципам MLOps и позволяет перейти от экспериментальной модели к производственному сервису без перепроектирования инфраструктуры.

##### 4.1 Архитектура вычислительной среды и программный стек

Таблица № 4 представляет сводную характеристику компонентов инфраструктуры, охватывающую спектр технических решений от облачной виртуальной машины до системы мониторинга концептуального дрейфа. В таблице систематизированы ключевые параметры каждого элемента: функциональные возможности, технические спецификации, интеграционные интерфейсы и эксплуатационные показатели. Такая структуризация позволяет провести комплексный анализ архитектуры и оценить вклад отдельных компонентов в общую работоспособность системы.

Использование асинхронного парсера Playwright обусловлено динамическим рендерингом страниц СОЗД. Это открывает новые возможности для автоматизации анализа законодательства от отслеживания динамики внесения законопроектов до оценки эффективности законодательской деятельности на основе больших данных.

##### 4.2 Обеспечение воспроизводимости и версионирования данных

**Фиксация среза данных.** `duma_ml.db` содержит слепок реестра по состоянию на март 2026 г. SHA-256 файла БД включается в метаданные каждого артефакта модели.



Таблица 4. Инфраструктура автоматизированного сбора данных Государственная Дума: компоненты и конфигурация

Компонент	Конфигурация и назначение
Вычислительная среда	Облачная VM: 4 vCPU, 8 ГБ RAM, Ubuntu 22.04 LTS. Google Colab (GPU T4) для NLP (RuBERT)
API-клиент (Extract)	Python 3.11 + httpx. Эндпоинт search.json: 34 653 записей. Retry: экспоненциальный back-off
Веб-парсер CO3Д (Enrich)	Playwright (асинхронный) + BeautifulSoup4. Обработано 26 539 страниц (77,6 %). Check-point-механизм
База данных (Store)	SQLite 3.42 (duma_ml.db). 5 таблиц, 34 653 строк. Индексы по law_number, introduction_date
Feature pipeline	feature_matrix.py: pandas 2.1 + scikit-learn 1.4. X_m1.parquet (34 189 × 20). SHA-256 + random_state=42
Обучение ML (Train)	train_m1_baseline.py: RandomForest (n_estimators=300), CatBoost (iterations=500). Grid-SearchCV
Мониторинг дрейфа	PSI для числовых признаков; KS-критерий (p < 0,05). Триггер переобучения: PSI > 0,25 или ROC-AUC < 0,85

**Детерминизм обучения.** Все случайные процессы зафиксированы через random\_state=42. Сортировка по (introduction\_date, law\_number) обеспечивает детерминированное разбиение при TimeSeriesSplit независимо от порядка загрузки записей из БД.

**Мониторинг актуальности.** При поступлении данных нового квартала автоматически вычисляются PSI и KS-критерий по ключевым признакам. При PSI > 0,25 или ROC-AUC < 0,85 инициируется процедура дообучения модели.

Описанная инфраструктура обеспечивает полный цикл от «сырых данных» API до верифицированного прогностического артефакта при минимальном ручном вмешательстве, что соответствует принципам MLOps.

5 Данные и предобработка

5.1 Источник данных и целевая переменная

Наше исследование подтверждает утверждение, что прогнозирование является важной составляющей при принятии законотворческих решений [2], в том числе для получения наилучшего законотворческого результата [11]. Указанное исследование показывает, что любой алгоритм интеллектуального анализа данных зависим не только от технических характеристик, но и от полноценного набора данных [10]. В данном примере алгоритмический порядок помогает на этапе разработки законотворческого решения принимать взвешенные законодательные инициативы при внесении изменений в отраслевое законодательство способствуя тем самым достижению разрабатываемыми законопроектами своего социального назначения [6].

Принимая во внимание изложенное полученные данные через официальный API Государственной Думы (api.duma.gov.ru) [5], и эндпоинт search.json предоставляет метаданные законопроектов с 1994 г. по 2025 г. Указанное позволило собрать датасет из 34 653 законопроекта, показывающий динамику разрабатываемых законодательных решений в каждой социально-экономической сфере.

Из них 34 189 (98,7 %) образуют выборку для машинного обучения (df\_ml) с известным бинарным исходом: принят (final\_result = 1) или не принят (final\_result = 0). При этом распределение классов следующее: 11 687 принятых (34,2 %) и 22 502 отклонённых (65,8 %), imbalance ratio 1,93:1 — умеренный дисбаланс.

Датасет не содержит дубликатов по идентификатору law\_number, что верифицировано в ходе предобработки.

5.2 Трёхэтапная верификация целевых меток

В ходе методического анализа выявлена существенная проблема, связанная с наличием пропущенных значений в поле final\_result. В исходных данных, получаемых посредством API, доля записей с отсутствующими значениями (пропусков) достигает 15 %. Для решения данной проблемы применена специализированная трёхэтапная процедура обработки данных.

На первом этапе обработки данных (с использованием API Государственной Думы) поле final\_result заполняется непосредственно на основании ответа API для 84,6 % записей.

На втором этапе обработки (с применением логических правил) выполняется автоматическое присвоение меток законопроектам на основании их текущей стадии. Законопроектам, находящимся на стадии «Опубликование закона», автоматически присваивается метка 1 (принят), а законопроектам, отозванным либо снятым с рассмотрения, — метка 0. В результате применения указанных правил дополнительно обработано 12,1 % записей.

На третьем этапе исследования осуществлён сбор данных методом вебпарсинга [9]. Данный метод позволил структурировать информацию, в результате чего к анализируемому массиву данных было добавлено 16260 новых размеченных наблюдений.

Комплексная обработка данных, выполненная в три последовательных этапа, позволила кардинально снизить долю пропущенных меток в наборе данных. Если на начальном этапе доля пропусков составляла 15 %, то после завершения всех процедур обработки она сократилась до

1,3 %. Достигнутое сокращение пропусков обеспечивает более полное и сбалансированное представление всех категорий в обучающей выборке. Это принципиально улучшает её репрезентативность и повышает надёжность результатов последующего машинного обучения по сравнению с аналогичными выборками.

### 5.3 Предобработка метаданных

В исходных («сырых») данных выявлено существенное количество пропущенных значений в поле `initiator_dept_name`, отражающем наименование инициатора законопроекта. Доля записей с отсутствующими данными в указанном поле составила 51,2 %, что свидетельствует о значительной неполноте информации на начальном этапе обработки. Для его восстановления применена трёхэтапная процедура: (1) прямое значение из API для 48,8 % записей; (2) логическое правило — если заполнены `initiator_faction` и `initiator_deputy_id`, то инициатор является депутатом Государственной Думы (+46,2 %); (3) парсинг поля с СОЗД для оставшихся записей (+2,4 %). Итого поле восстановлено для 99,7 % строк датасета, что позволило построить на его основе качественный категориальный признак `initiator_type`.

В ходе предобработки данных выявлены технические артефакты парсинга в полях `topic_name` и `class_name`, представленные значениями ‘Error’ и ‘Empty’. В рамках очистки данных указанные значения были заменены на NULL в объёме 3383 записей. Дополнительно выполнена нормализация поля `class_name`: из multilabel строк вида ‘080.100.000 Название’ извлечён первый класс с удалением числового кода, в результате чего сформирован новый признак `class_name_norm`. Поле `initiator_faction` также нормализовано: все значения приведены к 13 наименованиям фракций (включая «ЕР», «КПРФ», «ЛДПР», «СР», «СРЗП», «НЛ» и другие) с целью устранения дублирования исторических вариантов наименований.

В нашем исследовании предобработка метаданных представляет собой комплекс процедур по очистке, трансформации и стандартизации информации данных о законопроектах. Её цель — обеспечить согласованность метаданных для последующего анализа.

Предобработка метаданных имеет принципиальное значение для оценки целесообразности разработки законопроекта и внедрения моделей машинного обучения с целью прогнозирования обоснованности принятия соответствующего законодательного решения. Она позволяет очистить, структурировать и стандартизировать исходную информацию — включая данные о предыдущих законодательных инициативах, их статусе, инициаторах, тематической классификации, что обеспечивает высокое качество входных данных.

Вышеизложенное повышает точность прогнозов, снижает риск ошибок, обусловленных неполнотой, несогласованностью или искажением исходных сведений, а также

обеспечивает условия для корректного выявления закономерностей в законодательской деятельности.

В результате предобработка метаданных не только способствует обоснованному принятию законодательных решений на основе объективных данных, но и повышает надёжность работы алгоритмов машинного обучения при оценке перспективности разрабатываемых законодательных инициатив.

## 6 Проектирование признакового пространства

### 6.1 Базовые признаки (M1, 20 признаков)

В нашем исследовании базовая матрица признаков M1 включает 20 переменных, сгруппированных в пять групп.

Первая группа представляет характеристики инициатора законопроекта и включает следующие переменные: `initiator_type_enc` — закодированный тип инициатора законодательной инициативы (возможные значения: субъекты права законодательной инициативы); `initiator_faction_enc` — закодированное наименование политической фракции, к которой принадлежит инициатор; бинарный признак `is_ruling_party`, указывающий на принадлежность инициатора к партии большинства; а также `initiator_deputy_count` — количественный показатель числа соавторов законопроекта, представленный в логарифмически преобразованной форме ( $\log(1+x)$ ) с целью снижения влияния выбросов и нормализации распределения данных.

Вторая группа признаков представляет характеристики законопроекта и включает: `law_type_name_enc` — закодированный тип законопроекта (федеральный конституционный закон, федеральный закон, закон); `committee_te` — признак, сформированный методом target encoding [13] для ответственного комитета и отражающий статистическую связь деятельности комитета с целевой переменной; `profile_committee_count` — количественный показатель числа профильных комитетов; а также бинарные флаги `is_federal_law` (принадлежность к категории федеральных законов) и `is_international_treaty` (статус законопроекта как связанного с международными договорами или ратификациями).

Третья группа включает текстовые характеристики из метаданных законопроекта и представлена следующими переменными: `title_len` и `comment_len` — длина названия и текстового описания законопроекта соответственно (в символах), преобразованные с помощью логарифмической функции ( $\log(1+x)$ ) для стабилизации дисперсии; а также категориальные признаки `topic_name_enc` (закодированная тематика) и `class_name_norm_enc` (нормализованный первый класс законопроекта, из которого удален числовой код).

Четвёртая группа охватывает временные характеристики законодательного процесса: `convocation_num` — номер созыва (значения в диапазоне 4–8); `session_enc` — закодированное обозначение сессии (весенняя/осенняя); `intro_`

month — месяц внесения инициативы (от 1 до 12); а также бинарный флаг `is_election_year`, указывающий на совпадение года рассмотрения законопроекта с годом выборов.

Пятая группа объединяет скользящие (динамические) признаки, вычисляемые методом расширяющегося окна (expanding window): `rolling_success_rate` — исторический показатель конкретного инициатора, рассчитанный с экспоненциальным затуханием; `faction_success_rate` — кумулятивный показатель для фракции в целом (вычисляется без затухания, компенсирует отсутствие персональной истории у новых депутатов); и `days_since_conv_start` — логарифм числа дней с начала созыва, позволяющий модели отличать «свежие» инициативы от инициатив поздних этапов работы созыва.

При этом бинарный признак `has_transcript` (наличие стенограммы), демонстрирующий корреляцию с исходом рассмотрения на уровне 0,283, намеренно не включен в состав 20 базовых признаков. Это обусловлено риском возникновения (data leakage) [15]: стенограмма формируется по итогам заседаний, то есть содержит информацию из «будущего» относительно момента регистрации проекта. Исключение данной переменной позволяет оценить реальную прогностическую способность модели в условиях информационного вакуума, характерного для ранних стадий законотворческого процесса.

На основании указанного набора из 20 релевантных признаков, отражающих юридические, институциональные и социальнополитические аспекты инициативы, предсказательная способность модели определяется как мера её эффективности в раннем прогнозировании значимых исходов законотворческой процедуры.

### 6.2 Формирование динамических признаков с защитой от утечки данных

Для обеспечения высокой прогностической точности модели без нарушения принципа временной согласованности (data leakage) [15], в работе реализована методика вычисления динамических (скользящих) признаков. Рассмотрим детально алгоритмы формирования двух ключевых переменных пятой группы признаков пространства  $M1$ :

1. Показатель персональной результативности инициатора (`rolling_success_rate`). Данный признак отражает историческую долю принятых инициатив конкретного автора законопроекта на момент внесения текущего документа. Для учета изменчивости личного политического влияния и актуальности опыта законотворца во времени применяется механизм экспоненциального взвешивания исторических наблюдений (Exponentially Weighted Moving Average). Математический вес каждого предшествующего законопроекта определяется формулой:

$$w(\Delta t) = e^{-\lambda \Delta t} \quad (1)$$

где  $\Delta t$  — временная дистанция между наблюдениями, а  $\lambda$  — параметр (период полураспада) принят равным 5 годам (стандартный срок полномочий одного созыва Государственная Дума). Использование такого подхода по-

зволяет модели автоматически отдавать приоритет актуальной истории инициатора.

2. Временной индекс созыва (`days_since_conv_start`). Признак фиксирует количество дней, прошедших с даты первого заседания текущего созыва до момента регистрации законопроекта. Для нормализации распределения данных и снижения чувствительности алгоритма к экстремальным значениям применяется логарифмическое масштабирование:

$$X' = \ln(1 + X) \quad (2)$$

где  $X$  — абсолютное число дней с начала созыва. Данный показатель позволяет модели дифференцировать «сезонность» законотворческой деятельности: от этапа формирования комитетов в начале созыва до периода завершения работы созыва, характеризующегося накоплением нерассмотренных инициатив и изменением динамики принятия решений.

Конструкция указанных признаков является leak-age-safe, так как при вычислении значения для объекта  $n$  используется исключительно информация о подмножестве объектов  $\{1, \dots, n-1\}$ , внесённых строго ранее даты `introduction_daten`. Это гарантирует корректность оценки обобщающей способности модели на отложенных временных периодах.

### 6.3 Выбор алгоритмов классификации и стратегия обучения

Для построения прогностических моделей в работе были протестированы ансамблевые методы машинного обучения, наиболее эффективно зарекомендовавшие себя при анализе структурированных табличных данных с высокой долей категориальных переменных. В качестве базового алгоритма (baseline) был использован метод случайного леса (Random Forest), обеспечивающий устойчивость к нелинейным зависимостям в признаковом пространстве  $M1$ . Однако основным инструментом прогнозирования выбран алгоритм градиентного бустинга на решающих деревьях CatBoost [14]. Выбор данного алгоритма обусловлен его способностью эффективно обрабатывать высокоразмерные категориальные признаки (такие как «ответственный комитет», «инициатор» или «тематика») с использованием встроенного метода Symmetric Trees, что минимизирует риск переобучения и исключает необходимость предварительного разреженного кодирования (One-Hot Encoding).

Процесс обучения и оценки моделей строился на принципе кросс-валидации с учетом временной последовательности (Time Series Split). Такая стратегия критически важна для обеспечения временной согласованности, так как она гарантирует, что модель обучается только на данных предшествующих периодов. Данный подход гарантирует, что модель обучается только на данных предшествующих периодов и тестируется на последующих, исключая попадание информации из «будущего» в тренировочную выборку.

Оптимизация гиперпараметров (глубина дерева, скорость обучения, коэффициент регуляризации) прово-

дидась с использованием GridSearchCV с внутренней кросс-валидацией, ориентированного на максимизацию метрики ROC-AUC. Данная метрика была выбрана в качестве приоритетной, так как она наиболее репрезентативно отражает разделяющую способность классификатора в условиях значительного дисбаланса классов (преобладания отклоненных законопроектов над принятыми).

Итоговая конфигурация модели позволила достичь высокой обобщающей способности, что подтверждается стабильностью метрик на отложенной выборке законопроектов 8-го созыва.

## 7 Стратегия валидации

### 7.1 Схема временной кросс-валидации и её обоснование

Рисунок № 5 иллюстрирует схему пяти фолдов TimeSeriesSplit. Фолд 5 (VIII созыв, 2021–2025) имитирует

наиболее сложный производственный сценарий: прогноз для законопроектов нового созыва на основе модели, обученной исключительно на данных четырёх предыдущих.

Стандартная k-fold валидация завышает ROC-AUC на 0,015–0,040 — критически значимо для законодательных данных с выраженным концептуальным дрейфом.

### 7.2 Анализ концептуального дрейфа при смене созыва

Рисунок № 6 демонстрирует динамику ROC-AUC по фолдам и наглядно показывает эффект дрейфа при смене состава VIII созыва, а также степень его компенсации институциональными признаками.

M1 теряет 0,111 ROC-AUC на фолде 5 (0,929 → 0,820). M2a сокращает потери вдвое (+0,118 на фолде 5). M2b дополнительно стабилизирует (0,9515). Вывод: процедурные признаки (readings\_count, days\_to\_r1) обеспечивают наибольшую устойчивость к межсозывному дрейфу.

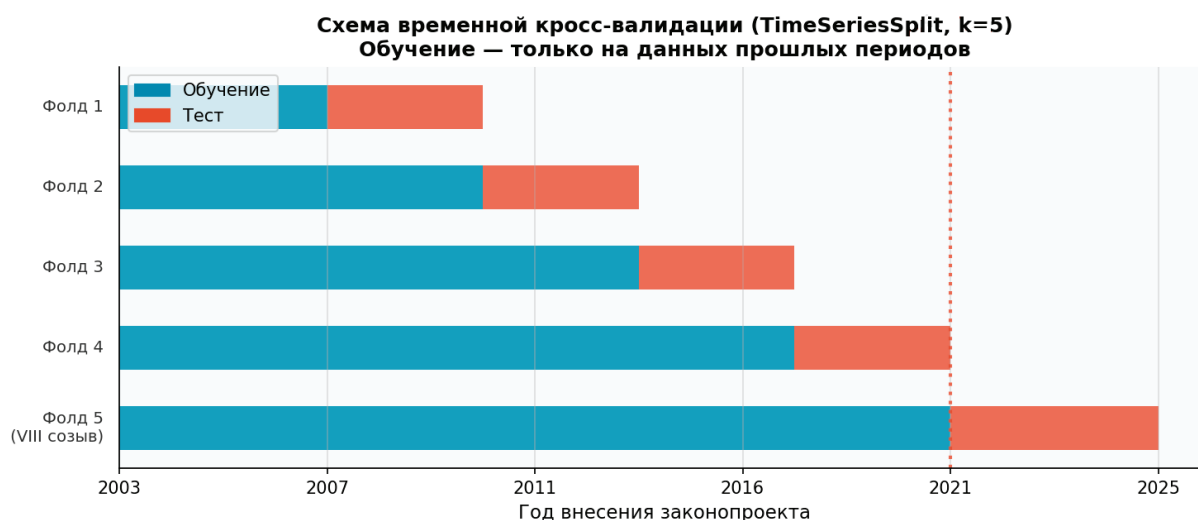


Рис. 5. Схема временной кросс-валидации (TimeSeriesSplit, k=5): обучение только на данных прошлых периодов

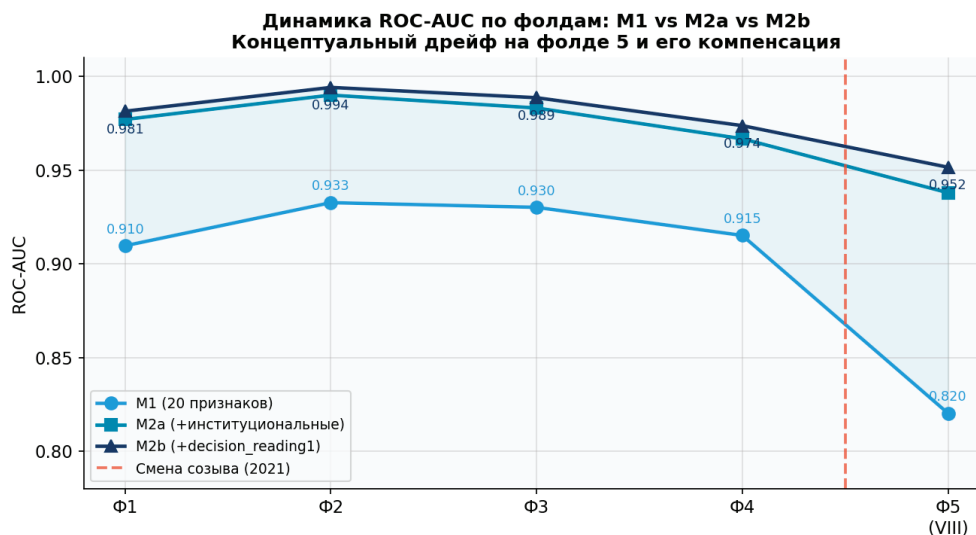


Рис. 6. Динамика ROC-AUC по фолдам: концептуальный дрейф и его компенсация (M1 vs M2a vs M2b)



## 8 Результаты экспериментов

### 8.1 Результаты baseline-модели M1

Результаты модели M1 (Random Forest, 20 признаков, TimeSeriesSplit 5 фолдов) представлены в таблице № 5. Для сравнения приведены результаты моделей-ориентиров: Dummy classifier и LogisticRegression.

Достигнутый ROC-AUC = 0,8989 при использовании исключительно метаданных является неожиданно высоким результатом. Для сравнения: работа Хавроненко М. В. [11] при использовании текстов всех документов достигла F1 = 0,937 (что соответствует примерно ROC-AUC ≈ 0,95), при этом на сопоставимой модели без текстов качество существенно ниже. Наш результат демонстрирует, что значительная часть предсказательной информации сосредоточена в структурных характеристиках законопроекта.

Наблюдаемое снижение качества модели на фолде 5 (значение показателя — 0,820 при среднем значении 0,929 по фолдам 1–4) подтверждает наличие дрейфа данных между VIII созывом и предшествующими созывами Государственной Думы. Частичная компенсация указанного эффекта достигнута за счёт введения признаков второй группы (faction\_success\_rate, days\_since\_conv\_start), что способствовало стабилизации прогностической способности модели.

### 8.2 Анализ важности признаков (SHAP)

SHAP-анализ (Shapley Additive Explanations) [7] проведён для обученной модели Random Forest на всём датасете. Основные SHAP-значения ранжируют признаки по суммарному абсолютному вкладу в прогноз. Отметим топ-5 признаков по важности:

1. initiator\_type\_enc — 0,310 (тип инициатора: Правительство РФ vs депутаты);
1. rolling\_success\_rate / faction\_success\_rate — 0,176 (скользящая положительная динамика);
2. committee\_te — 0,127 (target encoding ответственного комитета);
3. initiator\_faction\_enc — 0,095 (фракция инициатора);
4. convocation\_num — 0,051 (номер созыва).

Первые два признака суммарно дают 48,6 % важности, что является ключевым выводом: в законодательном процессе вопрос «кто вносит» значительно важнее вопроса «что вносит». Высокая важность committee\_te (0,127) свидетельствует о том, что ответственный комитет является не просто техническим параметром, но надёжным инди-

катором законопроекта — одни комитеты системно проводят законопроекты, другие — нет.

### 8.3 Расширенная модель M2a: институциональная поддержка

Модель M2a расширяет M1 пятью признаками, полученными путём парсинга страниц СОЗД: gov\_position\_enc (позиция Правительства РФ: поддерживает / частично поддерживает / не поддерживает / не требуется; покрытие 61,2 %), days\_to\_r1 (число дней от внесения до первого чтения, log1p; покрытие 84,1 %), had\_sf\_rejection (флаг отклонения Советом Федерации; покрытие 99,6 %), had\_president\_veto (флаг президентского вето; покрытие 99,6 %), readings\_count (число чтений 1/2/3; покрытие 53,3 %).

Модель M2a обеспечивает прирост ROC-AUC +0,069 относительно M1. Основным новым предиктором является readings\_count (feature importance 0,396 в M2a) — число чтений, пройденных законопроектом. Этот признак принципиально важен: законопроект, прошедший второе чтение, с высокой вероятностью будет принят в третьем. При этом readings\_count остаётся leakage-safe, поскольку в продуктивном сценарии модель применяется до итогового голосования, но уже после известного числа чтений.

Наибольший прирост качества M2a наблюдается на фолде 5 (+0,118), что подтверждает гипотезу о том, что институциональные признаки частично компенсируют концептуальный дрейф на материале VIII созыва. Иными словами, «правила процедуры» (сколько чтений прошёл закон, отклонялся ли Советом Федерации) остаются более стабильными между созывами, чем фракционный состав и тематические приоритеты.

## 9 Обсуждение и ограничения

Разработанная baseline-модель представляет собой ансамблевый классификатор Random Forest, обученный на матрице признаков из 20 переменных, сгруппированных в четыре содержательных блока: характеристики инициатора (тип, фракция, число соавторов, скользящая результативность фракции), параметры законопроекта (тип документа, target encoding ответственного комитета, число профильных комитетов), текстовые-метаданные (длина наименования и описания) и временные характеристики (номер созыва, сессия, флаг электорального года). Валидация модели осуществлена методом TimeSeriesSplit с 5 фолдами, обеспечивающим строгое соблюдение времен-

Таблица 5. Результаты M1 baseline (TimeSeriesSplit, 5 фолдов)

Модель	Ф1	Ф2	Ф3	Ф4	Ф5	Среднее	σ
Dummy classifier	—	—	—	—	—	0,498	0,010
LogisticRegression	—	—	—	—	—	0,819	0,040
CatBoost (20 пр.)	—	—	—	—	—	0,897	—
<b>RandomForest (20 пр.)</b>	<b>0,908</b>	<b>0,935</b>	<b>0,933</b>	<b>0,915</b>	<b>0,820</b>	<b>0,8989</b>	<b>0,050</b>



Таблица 6. Сравнение M1 и M2 (TimeSeriesSplit, 5 фолдов, ROC-AUC)

Фолд	M1 (20 пр.)	M2a leakage-free	Δ M2a	M2b +dec.r1	Δ M2b
1	0,9096	0,9770	+0,0674	0,9814	+0,0718
2	0,9326	0,9900	+0,0574	0,9941	+0,0615
3	0,9301	0,9831	+0,0530	0,9886	+0,0585
4	0,9151	0,9667	+0,0516	0,9737	+0,0586
5	0,8202	0,9379	+0,1177	0,9515	+0,1313
<b>Avg</b>	<b>0,8989</b>	<b>0,9709</b>	<b>+0,0694</b>	<b>0,9779</b>	<b>+0,0763</b>

ного порядка данных и имитирующим реальный сценарий применения модели (прогноз формируется исключительно на основе законопроектов, внесённых раньше прогнозируемого периода).

Расширенная модель M2a дополняет базовое признаковое пространство пятью институциональными переменными, полученными через парсинг СОЗД (позиция Правительства Российской Федерации, число дней от

внесения в Государственную Думу до первого чтения, факт отклонения Советом Федерации, факт президентского вето и число пройденных чтений). Данные признаки доступны до момента итогового голосования и остаются leakage-safe в продуктивном сценарии применения модели, что подтверждает возможность их использования без риска искусственного завышения показателей качества прогноза.

Таблица 7. Результаты baseline-модели M1 (TimeSeriesSplit, 5 фолдов, ROC-AUC) — воспроизведена для сравнения

Модель	Ф1	Ф2	Ф3	Ф4	Ф5	Среднее	σ
Dummy classifier	—	—	—	—	—	0,507	0,010
LogisticRegression	—	—	—	—	—	0,819	0,040
CatBoost (20 пр.)	—	—	—	—	—	0,899	—
<b>RandomForest (20 пр.)</b>	<b>0,908</b>	<b>0,935</b>	<b>0,933</b>	<b>0,915</b>	<b>0,820</b>	<b>0,8989</b>	<b>0,047</b>

Достигнутое значение ROC-AUC = 0,8989 для baseline-модели свидетельствует о высокой структурной предсказуемости российского законодательного процесса: значительная часть предсказательной информации сосредоточена в процедурных атрибутах законопроекта, а не в его содержательных характеристиках. Данный результат имеет существенное практическое значение для оценки целесообразности разработки законодательной инициативы (на момент внесения законопроекта, до прохождения каких-либо чтений, модель способна с высокой точностью предсказать его судьбу на основе информации о субъекте права законодательной инициативы и исторической результативности соответствующей фракции).

Наблюдаемое снижение качества прогнозирования на фолде 5 (ROC-AUC = 0,820 при среднем значении 0,929 по фолдам 1–4) закономерно отражает концептуальный дрейф данных, обусловленный сменой созыва Государственной Думы в 2021 году. Изменение фракционного состава, появление значительного числа депутатов без накопленной истории голосований и трансформация законотворческих приоритетов VIII созыва приводят к смещению статистических паттернов, усвоенных моделью на данных предшествующих созывов. Именно этот сценарий в наибольшей мере воспроизводит практическую задачу прогнозирования, в частности модель, обученная на исторических данных, должна обобщаться на будущий законодательный период.

Таблица 8. Результаты модели с расширенным признаковым пространством M2a (TimeSeriesSplit, 5 фолдов, ROC-AUC)

Фолд	M1 (20 пр.)	M2a leakage-free	Δ M2a	M2b +dec.r1	Δ M2b
1	0,9096	0,9770	+0,0674	0,9814	+0,0718
2	0,9326	0,9900	+0,0574	0,9941	+0,0615
3	0,9301	0,9831	+0,0530	0,9886	+0,0585
4	0,9151	0,9667	+0,0516	0,9737	+0,0586
5	0,8202	0,9379	+0,1177	0,9515	+0,1313
<b>Avg</b>	<b>0,8989</b>	<b>0,9709</b>	<b>+0,0694</b>	<b>0,9779</b>	<b>+0,0763</b>

Результаты модели M2a наглядно демонстрируют, что институциональные признаки — в первую очередь число

пройденных чтений (readings\_count, feature importance = 0,396 в M2a) — несут существенную дополнительную

предсказательную силу сверх базовых метаданных. Прирост ROC-AUC составил +0,069 в среднем по фолдам и достиг +0,118 на критическом фолде 5 (VIII созыв), что свидетельствует о частичной компенсации концептуального дрейфа за счёт процедурных признаков законодательного процесса. С правовой точки зрения данный результат означает, что «правила процедуры» — сколько чтений прошёл законопроект, поддерживается ли он Правительством РФ — остаются значительно более стабильными между созывами, чем фракционный состав и тематические приоритеты.

Наибольший прирост качества M2a фиксируется именно на фолде 5, где дрейф данных наиболее выражен, что подтверждает гипотезу об институциональных признаках как «якорях стабильности» законотворческого процесса. Признак `days_to_r1` (число дней от внесения до первого чтения) отражает процедурный приоритет, предоставляемый законопроекту: инициативы Президента РФ и Правительства РФ, как правило, рассматриваются в ускоренном режиме, что само по себе служит индикатором высокой вероятности принятия. Аналогичную роль косвенного сигнала успешности выполняют признаки отсутствия ветирования (`had_sf_rejection`, `had_president_veto`): законопроект, не встретивший институционального отклонения, с существенно большей вероятностью будет принят в итоге.

Промежуточный анализ полученных результатов позволяет сформулировать ряд существенных методологических выводов. Во-первых, метаданные законопроектов обладают значительным самостоятельным прогностическим потенциалом: достижение ROC-AUC = 0,8989 без привлечения текстового содержания свидетельствует о том, что институциональный контекст внесения законодательной инициативы (кто вносит, в какой период созыва, с какой исторической результативностью) определяет её судьбу в значительно большей мере, чем принято считать на основе традиционных правовых представлений. Данный вывод согласуется с концепцией процедурной предсказуемости законодательного процесса и подтверждает алгоритмическую природу принятия законотворческих решений в Государственной Думе.

Во-вторых, введение скользящих признаков (`faction_success_rate`, `days_since_conv_start`) на базе накопленных средних значений является не просто техническим решением, но способом интеграции институционального опыта в прогностическую модель без нарушения принципа временной строгости данных. Признак `faction_success_rate` кодирует накопленный авторитет и реальное влияние фракции в процессе прохождения инициатив, тогда как `days_since_conv_start` позволяет учесть сезонные и циклические закономерности законодательной активности. Совместное применение этих переменных обеспечивает корректную работу модели в динамически меняющейся институциональной среде без риска искажения оценок за счёт утечки будущих данных.

Необходимо выделить ряд существенных ограничений разработанной модели. Вопервых, модель демонстрирует низкую обобщающую способность в отношении принципиально новых типов инициаторов или фракций, отсутствующих в обучающей выборке, что соответствует классической проблеме холодного старта (cold start problem). Вовторых, наличие концептуального дрейфа (concept drift) при смене созыва Государственной Думы приводит к снижению качества прогнозирования на фолде 5 до значения 0,820, что обуславливает необходимость регулярного дообучения модели на актуальных данных. Втретьих, отсутствие в метаданных сведений о неформальных договорённостях и политических коалициях формирует информационный потолок, ограничивающий предельную прогностическую способность модели даже при оптимальной настройке гиперпараметров.

Отмечаем, что принципиальным методологическим достижением настоящего исследования является разработка leakagesafe скользящих признаков (leakagesafe rolling features), обеспечивающих корректную работу модели в режиме реального прогнозирования.

В то же время ошибочное включение в обучающую выборку признаков, вычисленных на основе будущих данных (data leakage) [15], представляет собой распространённую проблему в исследованиях по парламентскому прогнозированию [3]. Подобная утечка данных приводит к искусственному завышению оцениваемых показателей качества модели и искажению реальной прогностической способности алгоритма.

Предложенный нами подход исключает данный источник смещения за счёт строгого соблюдения временного порядка при формировании признаков.

Итоговые результаты проведённого исследования подтверждают высокую практическую ценность предложенной методологии для оценки целесообразности разработки законодательных инициатив. Baseline-модель на структурированных метаданных достигает ROC-AUC = 0,8989 — результат, превосходящий сопоставимые показатели для других парламентских датасетов и свидетельствующий о высокой процедурной упорядоченности российского законотворческого процесса. Расширенная модель M2a с институциональными признаками обеспечивает дополнительный прирост до ROC-AUC = 0,9709, существенно компенсируя концептуальный дрейф на материале VIII созыва. Таким образом, разработанная система признаков создаёт надёжную инструментальную базу для практического внедрения алгоритмического сопровождения законотворческой деятельности.

Представленные результаты формируют также важную теоретическую базу для последующих исследований. Установленная предельная точность baseline-модели на метаданных задаёт количественный ориентир для оценки дополнительного информационного вклада текстового содержания законопроектов, стенограмм парламентских заседаний и семантических представлений, извлекаемых крупными языковыми моделями (RuBERT, LLaMA).

Любой прирост ROC-AUC сверх значения 0,9709 (M2a) при введении текстовых признаков будет свидетельствовать о самостоятельной предсказательной ценности содержательных характеристик законопроекта, выходящей за рамки институционального контекста. Это открывает перспективу для построения полноценной многоуровневой методики прогнозирования, интегрирующей метаданные, процедурные признаки и семантический анализ документов.

## 10 Выводы

В настоящей статье представлена методика построения baselinемодели для прогнозирования вероятности принятия законопроектов Государственной Думой исключительно на основе структурированных метаданных. Предложенный подход исключает использование текстового содержания законопроектов, фокусируясь на формализованных атрибутах (тип инициатора, фракционная принадлежность, хронология внесения и т.д.). Основные результаты исследования:

1. В рамках исследования был собран и преобразован датасет, включающий 34 189 законопроектов, принятых и рассмотренных Государственной Думой в период с 1994 по 2025 год. Доля верифицированных меток в датасете составляет 98,7 %, что достигнуто посредством реализации трёхэтапной процедуры верификации данных. Такой уровень качества разметки обеспечивает высокую надёжность последующих этапов моделирования и анализа.

2. В ходе исследования разработано признаковое пространство, включающее 20 переменных, релевантных для прогнозирования принятия законопроектов. Ключевой методологической особенностью является включение двух leakagesafe скользящих признаков с экспоненциальным взвешиванием. Данный подход позволяет:

- исключить утечку данных (data leakage) [15] из будущего периода;

- учесть временную динамику законотворческого процесса;

- придать больший вес более актуальным наблюдениям за счёт экспоненциального убывания весовых коэффициентов.

Такая конструкция признаков повышает прогностическую способность модели при сохранении корректности процедуры валидации.

3. В ходе исследования реализована baselinемодель на основе алгоритма Random Forest с применением временной кроссвалидации (timeseries crossvalidation) для

корректной оценки прогностической способности в условиях временных данных. Модель продемонстрировала высокую эффективность на структурированных метаданных без привлечения текстового содержания документов. Значение метрики ROCAUC составило  $0,8989 \pm 0,050$ . Данный результат подтверждает достаточность ограниченного набора формализованных признаков для построения надёжной прогностической модели принятия законопроектов Государственной Думой.

4. Расширенная модель M2a, дополненная признаками институциональной поддержки (включая метрику readings\_count — число чтений законопроекта), продемонстрировала существенное улучшение прогностической способности по сравнению с baselinемоделью. Значение метрики ROCAUC достигло 0,9709 (+0,069 относительно базового уровня), что статистически подтверждает высокую прогностическую ценность признака readings\_count. Данный результат указывает на значимость институциональных факторов в процессе принятия законопроектов Государственной Думой и открывает перспективы для дальнейшего развития моделей с учётом процедурных характеристик законотворчества.

5. Проведённый SHAPанализ количественно оценил вклад отдельных признаков в прогнозную модель. Установлено, что совокупная важность двух ключевых предикторов — типа инициатора законопроекта и его исторической результативности — составляет 48,6 % от общей важности всех признаков. Данный результат статистически превосходит вклад содержательных характеристик законопроекта (текстовых и тематических атрибутов), что подчёркивает основную роль институциональных факторов в процессе принятия законодательных инициатив Государственной Думой.

В целом полученные результаты формируют базовую точку отсчёта (baseline) для количественной оценки дополнительного вклада текстовых данных в задачу прогнозирования законодательного процесса. Учитывая указанное в качестве источников текстовой информации рассматриваются:

- стенограммы заседаний Государственной Думы;
- заключения профильных комитетов;
- семантические и синтаксические признаки, извлекаемые с помощью языковых моделей (например, BERT, RuBERT, LLaMA).

Сравнение качества расширенных моделей с текстовыми признаками и исходной baselinемодели позволит объективно оценить информационную ценность вербального содержания законопроектов и сопутствующих документов.

## Литература:

1. Bari A., Brower W., Davidson C. Using artificial intelligence to predict legislative votes in the United States congress // IEEE 6th International Conference on Big Data Analytics. — 2021. — P. 56–60;
2. Гафаров Ф. М., Руднева Я. Б., Шарифов У. Ю. Прогностическое моделирование в высшем образовании: определение факторов академической успеваемости // Высшее образование в России. 2023. Т. 32. № 1. С.65;

3. Gerrish S., Blei D. How they vote: Issue-adjusted models of legislative behavior // *Advances in Neural Information Processing Systems*. — 2012. — Vol. 25;
4. Гвоздецкий Д. С., Шипов Д. М. Инновационные решения в ведомственном правотворчестве (на примере моделей машинного обучения и больших языковых моделей) // *Закон и право*. 2025. № 2. С.18;
5. Документация официального API Государственной думы РФ. — URL: <https://api.duma.gov.ru/pages/dokumentatsiya> (дата обращения: 01.03.2026);
6. Лазарева О. В. Технология изменения российского законодательства // *Юридическая техника*. 2023. № 17, С.313;
7. Lundberg S. M., Lee S.-I. A unified approach to interpreting model predictions // *Advances in Neural Information Processing Systems*. — 2017. — Vol. 30;
8. Nay J. J. Predicting and understanding law-making with word vectors and an ensemble model // *PloS one*. — 2017. — Vol. 12, № 5. — P. e0176999;
9. Официальный сайт Государственной Думы Федерального Собрания Российской Федерации [Электронный ресурс]/Режим доступа: URL: <https://sozd.duma.gov.ru/> свободный (дата обращения: 01.03.2026);
10. Пономарева К. А. Применение Искусственных Нейронных Сетей При Решении Задач Прогнозирования // *Наука Без Границ*, № 1(41), 2020, С.44;
11. Хавроненко М. В. Прогнозирование результатов рассмотрения законопроектов Государственной думой: модель нейронной сети // *Политическая наука*. — 2024. — № 3. — С. 211–240;
12. Yano T., Smith N. A., Wilkerson J. Textual predictors of bill survival in congressional committees // *Proceedings of NAACL HLT 2012*. — P. 793–802.
13. Micci-Barreca D. A preprocessing scheme for high-cardinality categorical attributes // *ACM SIGKDD Explorations Newsletter*. — 2001. — Vol. 3, № 1. — P. 27–32.
14. Prokhorenkova L. et al. CatBoost: unbiased boosting with categorical features // *Advances in Neural Information Processing Systems*. — 2018. — Vol. 31. — P. 6638–6648.
15. Sculley D. et al. Hidden technical debt in machine learning systems // *Advances in Neural Information Processing Systems*. — 2015. — Vol. 28. — P. 2503–2511.



## ГЕОЛОГИЯ

### Новая жизнь выработанных шахт

Воробей Никита Викторович, студент;

Скрипник Сергей Андреевич, студент

Белгородский государственный технологический университет имени В. Г. Шухова

*В статье авторы рассматривают, как используются выработанные шахты в других странах и в России.*

**Ключевые слова:** шахты, реконструкция.

В последние десятилетия мир сталкивается с вызовами устойчивого развития и охраны окружающей среды, что актуализирует вопросы по восприятию и использованию исторических промышленных объектов. Заброшенные шахты, когда-то служившие основой экономики и местной культуры, сегодня превращаются в наследие, которое требует внимания и осмысленного подхода. Реконструкция таких объектов не только позволяет сохранить историческую память, но и открывает новые возможности для их использования — будь то создание туристических маршрутов, эко проектов или центров для изучения природных ресурсов.

В данной статье мы рассмотрим методологии и подходы к реконструкции заброшенных шахт, а также проанализируем примеры успешных проектов. Мы обсудим важность сохранения этих объектов для будущих поко-

лений и их потенциал для развития местного самоуправления и экономики.

Заброшенные шахты или другие подземные выработки расположенные в черте города или рядом с ними использовать их не по назначению часто весьма выгодно. Нехватка таких «безопасных» помещений пригодных, в частности, под различные склады особенно остро ощущалось в годы войны. Именно во время войны был дан толчок для такого использования шахт.

Английские шахты, из которых добывался знаменитый батский камень (известняк), во время войны были превращены в подземные заводы. После войны в них начали выращивать грибы, чему способствовала постоянная температура, хорошая вентиляция и отсутствие солнечного света. Постоянная температура в шахтах скорее всего объясняется тепловыми свойствами пород земной



Рис. 1. Заброшенная шахта в России



коры — суточное изменение температуры воздуха отражаются в поверхностном слое грунта только до глубины 0.3 м и даже меньше.

Подземные выработки сейчас широко используются во всем мире, но, пожалуй, особенно в районах Канзас-Сити (штат Миссури). Канзас-Сити важный промышленный центр на Среднем Западе США — расположен в середине 225 — километрового пояса выходящих на поверхность каменноугольных пород. Она состоит из чередующихся слоев известняков и сланцев. Обычно при подземной разработке для поддержки кровли выработок оставляют мощные столбы пород. Установлено, что из 1160 га выработанного под городом пространства 470 га в 21 районе можно использовать в качестве различных складских помещений и т. п.

Шахты в Миссурии разрабатываются с учетом будущего использования выработанной площади. Опоры кровли располагаются в определенном порядке, что дает возможность увеличить полезное свободное простран-

ство. Полы выработок выравнены, а над входами для обеспечения устойчивости устанавливают достаточно мощные перекрытия. Около 60 га подземных выработок используются для промышленных целей в одном только районе Канзас — Сити. Многие фирмы склонны считать подземные условия наиболее подходящими для производства точных приборов. Для регулирования температуры воздуха и влажности требуются вентиляция и специальное оборудование. Но стоимость таких устройств гораздо меньше капитальных и текущих затрат, требующиеся для строительства и эксплуатации обычного здания. Иногда для обеспечения надежности кровли нужна анкеровка, однако при правильной проходке требуется ничтожная дополнительная работа для проведения выработки в готовые состояния.

Некоторые старые соляные шахты не в черте города, но и не далеко от него, служат хранилищами нефтепродуктов. Например, заброшенная шахта в близи города Солотвина в Карпатах превращена в больницу для лиц, стра-



Рис. 2. Шахта Цольферайн в Германии



Рис. 3. Бассейн промышленного комплекса

дающих бронхиальной астмой. Пациенты проводят 7–12ч в сутки в переоборудованной шахте на глубине 202м. Для лечения больных не применяют лекарства: врачи убеждены, что микроклимат внутри шахты и специальные лечебные упражнения делают лекарства ненужными.

Среди достопримечательностей Германии мирового значения оказалась каменноугольная шахта Цольферайн, она была построена в 1834 году архитекторами Фрицем Шуппом и Мартеном Кремером. Она находится в городе Эссен. Добыча полезных ископаемых проводилась вплоть до 1993, несмотря на то что решение о закрытии шахтного комплекса было согласовано в 1986 году.

На данный момент промышленный центр является культурным комплексом. На его территории расположилось несколько музеев, самым символическим является из них является «Дорога угля». Часть промышленных помещений были переоборудованы в дизайнерские центры, которые являются местом для проведения различных мероприятий. Например, летом на территории промышленного комплекса работает бассейн и несколько ресторанов и кафе.

В зимний период на его территории посетители могут кататься на большом ледовом катке в канун Рождества Христова здесь проводят массу мероприятий



Рис. 4. Ледовая площадка промышленного комплекса

Все сказанное выше свидетельствует о том насколько широко можно пользоваться огромными площадями, получаемыми в результате подземной выработки пород. Хотелось бы отметить, что можно заниматься благоустройством не только подземными выработками, но

и карьерами, которые располагаются как в дали от города, так и в его районе. Например, из карьера можно сделать много уровневый парк, если есть грунтовые воды, то можно их использовать для постройки фонтанов и использовать для полива деревьев и растений в парке.

#### Литература:

1. Физические Основы Добычи и Переработки Полезных Ископаемых: учебное пособие / В. Г. Дмитриенко, В. Ю. Доброногова, Д. В. Мулов, П. М. Бакланова — Белгород: Изд-во БГТУ, Алчевск, 2024. — 124 с.
2. Основы горного дела: учеб. пособие / В. Г. Дмитриенко — Белгород: Изд-во БГТУ 2011–186 с.
3. Города и геология: Р. Леггет — Нью-Йорк, 1973

## Молекулярно-геохимические маркеры в диагностике генезиса и трансформации нефти и газа в глубоководных коллекторах

Мухаммедова Джерен Чарыевна, старший преподаватель;

Астанакулиев Атаджан Астанакулиевич, студент

Международный университет нефти и газа имени Ягшигельды Какаева (г. Ашхабад, Туркменистан)

*В данной статье рассматривается применение молекулярных геохимических маркеров для определения генезиса, термической зрелости и постгенетических превращений нефти и газа в глубоководных коллекторах. Особое внимание уделяется анализу биомаркеров, включая стераны, терпаны и ароматические углеводороды, как индикаторов источников*

органического вещества и условий осадконакопления. Показано, что изменения в соотношении изомеров и параметрах ароматизации позволяют реконструировать температурно-временную историю отложений и оценить влияние процессов катагенеза и метагенеза. Для обеспечения высокой точности обнаружения молекулярных структур используются современные аналитические методы, такие как газовая хроматография и масс-спектрометрия высокого разрешения.

**Ключевые слова:** нефть, природный газ, биомаркеры, геохимия, катагенез, стераны, терпаны, газовая хроматография, масс-спектрометрия, нефтегазоносные бассейны.

Разработка глубокозалегающих, труднодоступных ресурсов стала определяющей чертой роста нефтегазовой отрасли в последние десятилетия, что обусловило необходимость совершенствования методов исследования углеводородных систем. Сложные процессы трансформации органического вещества происходят при повышении температуры и давления, изменяя состав нефти и газа. Молекулярно-геохимические методы необходимы, поскольку традиционные геологические методы не всегда позволяют точно определить происхождение углеводородов и оценить их эволюцию. В этой ситуации биомаркеры служат надежными маркерами, сохраняющими данные об исходных условиях образования органического вещества и последующих фазах его изменения [1].

Сложные органические молекулы, происходящие из биогенных предшественников, прошедших несколько стадий диагенеза и катагенеза, известны как молекулярные геохимические индикаторы. Эти молекулы могут быть использованы для реконструкции геохимической истории углеводородных систем благодаря их устойчивости к термическим и химическим воздействиям. Наиболее информативными являются стераны и терпаны благодаря своей структуре, которая позволяет отслеживать изменения органического вещества по мере повышения температуры и глубины.

Измерение изомерных соотношений биомаркеров особенно важно в глубокозалегающих пластах. Соотношения стереоизомеров изменяются по мере перехода биомолекул в более стабильные формы при увеличении теплового воздействия. Термическую зрелость можно объективно оценить, используя эти предсказуемые механизмы. Однако следует отметить, что на больших глубинах могут происходить процессы разрушения исходных молекулярных структур, что затрудняет интерпретацию данных и требует комплексного подхода [2].

Поскольку ароматические углеводороды образуются при дегидрировании и циклизации насыщенных молекул, они играют решающую роль в определении зрелости органического вещества. Мы можем оценить температурные условия, при которых образовывались углеводороды, изучая изменения степени ароматизации и распределение метилзамещенных производных. Высокие температуры в глубоких горизонтах способствуют процессам ароматизации и повышают процентное содержание полициклических молекул.

Важным фактором является влияние вторичных процессов, таких как термический крекинг, миграция и смешивание углеводородов из разных источников. Соот-

ношение легких и тяжелых фракций может колебаться в глубоких пластах из-за распада нефти на газообразные продукты. Распределение углеводородов и их изотопный состав могут быть использованы для идентификации этих процессов, отражающихся в молекулярном составе. Благодаря изотопным исследованиям мы можем определить происхождение органических веществ и выявить признаки смешивания в нескольких генерационных потоках [3].

Изучение молекулярного состава нефти и газа требует использования современных аналитических методов. В то время как масс-спектрометрия позволяет с высокой точностью идентифицировать молекулярные структуры, газовая хроматография разлагает сложные смеси на составляющие их части. Информационная база молекулярной геохимии значительно расширяется за счет использования высокоразрешающих приборов, которые позволяют анализировать вещества, ранее недоступные. Это позволяет получить исчерпывающую информацию о происхождении и содержании углеводородов в сочетании с изотопными методами.

Для правильной интерпретации результатов данные молекулярной геохимии должны быть интегрированы с геологическими моделями. При оценке геохимических характеристик необходимо учитывать тектоническую историю, литологический состав и условия осадконакопления бассейна. Этот метод позволяет выявлять тенденции в распределении углеводородов и точно определять перспективные места для дальнейших исследований. Комплексный анализ становится основным методом прогнозирования нефтеносного потенциала в глубоких горизонтах, где прямые данные скудны.

При изучении нетрадиционных ресурсов, таких как сланцевые формации и сверхглубокие месторождения, применение молекулярных геохимических маркеров имеет решающее значение. В этих условиях необходимо уточнить традиционные представления о формировании и перемещении углеводородов, а биомаркеры могут раскрыть свойства таких систем. Молекулярный анализ повышает эффективность разработки и упрощает оптимизацию производственных процессов [4].

Использование методов цифрового моделирования и машинного обучения для анализа сложных наборов аналитических данных — еще одно направление развития молекулярной геохимии. Традиционные аналитические методы теряют свою эффективность из-за растущего объема данных, генерируемых многомерной хроматографией и масс-спектрометрией высокого разрешения. Мы



можем устанавливать взаимосвязи между геохимическими измерениями и геологическими особенностями, выявлять скрытые закономерности в распределении молекулярных маркеров и повышать точность прогнозирования нефтепотенциала с помощью алгоритмов искусственного интеллекта. Этот метод обеспечивает более глубокое понимание процессов образования, миграции и трансформации углеводородов в сложных и малоизученных глубоких пластах, способствуя разработке интеллектуальных систем поддержки принятия решений в нефтегеологии.

Изучение влияния жестких термобарических условий на стабильность и модификацию молекулярных маркеров в сверхглубоких горизонтах является важнейшей составляющей будущих исследований. Интенсивное разрушение сложных биомолекул наблюдается при температурах выше традиционного «нефтяного окна», наряду с образованием более простых углеводородных структур и газообразных компонентов. Это приводит к частичной или полной потере традиционных геохимических маркеров, что делает необходимым создание новых диагностических

стандартов, основанных на стабильных фрагментах молекулярных систем и изотопных свойствах. В этих условиях особенно важен комплексный подход, включающий экспериментальное моделирование процессов термической деградации органического вещества. Это позволит улучшить кинетику трансформации и модифицировать существующие геохимические модели в соответствии с условиями сверхглубоких нефтегазовых систем [5].

Важнейшим инструментом для исследования происхождения и эволюции нефти и газа в глубоких пластах являются молекулярные геохимические маркеры. Используя их, мы можем определить влияние вторичных процессов, оценить термическую зрелость углеводородных систем и воссоздать историю их формирования. Эффективность разведки и точность геохимической диагностики повышаются благодаря развитию аналитических технологий и внедрению междисциплинарных методов. Молекулярная геохимия будет играть все более важную роль в обеспечении научной основы для разумного освоения нефтегазовых ресурсов по мере усложнения геологических характеристик и роста спроса на энергию.

#### Литература:

1. Peters, Kenneth E., Clifford C. Walters, and J. Michael Moldowan. The Biomarker Guide: Volume 1, Biomarkers and Isotopes in Petroleum Exploration and Earth History. 2nd ed., Cambridge University Press, 2005.
2. Peters, Kenneth E., Clifford C. Walters, and J. Michael Moldowan. The Biomarker Guide: Volume 2, Biomarkers and Isotopes in Petroleum Systems and Earth History. 2nd ed., Cambridge University Press, 2005.
3. Tissot, Bernard P., and Dietrich H. Welte. Petroleum Formation and Occurrence. 2nd ed., Springer-Verlag, 1984.
4. Hunt, James M. Petroleum Geochemistry and Geology. 2nd ed., W. H. Freeman and Company, 1996.
5. Killops, Stephen D., and Vanessa J. Killops. Introduction to Organic Geochemistry. 2nd ed., Wiley-Blackwell, 2013.

## Красноусольское месторождение минеральных вод

Сайфуллина Гузель Жамиловна, студент магистратуры

Научный руководитель: Мустафин Сабир Кабирович, доктор геолого-минералогических наук, профессор  
Уфимский университет науки и технологий

*В статье рассматриваются условия формирования различных типов минеральных вод и характеристика их химического состава.*

**Ключевые слова:** Красноусольское месторождение, минеральные воды.

Красноусольские сероводородные воды («серносоляные источники») известны с XVI в. Описаны они в середине XVIII в. в трудах первых участников экспедиций Российской академии наук (П. И. Рычкова, П. С. Палласа, И. И. Лепехина). Первая сводка о минеральных водах и лечебных грязях Республики Башкортостан, в том числе Красноусольских сероводородных вод, была выполнена Г. В. Вахрушевым в 1929 г. Он одним из первых установил, что Красноусольские минеральные воды разгружаются по тектоническим разломам с глубин 400–500 м под действием гидравлического напора. Тре-

щины разобщены между собой, об этом свидетельствуют различия минерализации и большие колебания содержания радона даже в близко расположенных источниках. Глубинные воды разных струй при подъеме неодинаково разбавляются пресными водами верхних водоносных горизонтов, что объясняется большими колебаниями минерализации минеральных вод. С 1965 года это место имеет статус гидрологического памятника природы.

Месторождение находится в долине реки Усолка в предгорьях западных склонов Южного Урала, в 5 км от села Красноусольского Гафурийского района респуб-

лики Башкортостан. В геолого-структурном отношении Красноусольские месторождения минеральных вод приурочены к сводовой части Усольской антиклинали, сложенной известняками среднего и верхнего карбона, перекрытыми нижнепермскими молассовыми отложениями. На территории расположены три месторождения минеральных вод, отличающиеся по своим геохимическим показателям и лечебным свойствам: Красноусольское 1, Красноусольское 2 и Красноусольское 3. Наиболее крупное из них — Красноусольское 1 — представлено сероводородными водами.

В районе Красноусольского курорта выявлено 32 группы восходящих сульфидных источников, расположенных двумя линиями по правому и левому берегам р. Усолки. Длина правобережной линии около 0,5 км; она объединяет 29 групп источников естественного происхождения

Воды всех источников относятся к хлоридным натриевым с минерализацией 2,2–47,7 г/дм<sup>3</sup>, концентрацией H<sub>2</sub>S до 70–80 мг/дм<sup>3</sup>, температурой 9–13°C, величиной pH 6,9–7,4. Содержание микроэлементов (мг/дм<sup>3</sup>): Br до 40,5; I до 0,9; H<sub>3</sub>BO<sub>3</sub> до 30,9; F до 1,6. Солевой состав на 80–95 % представлен NaCl.

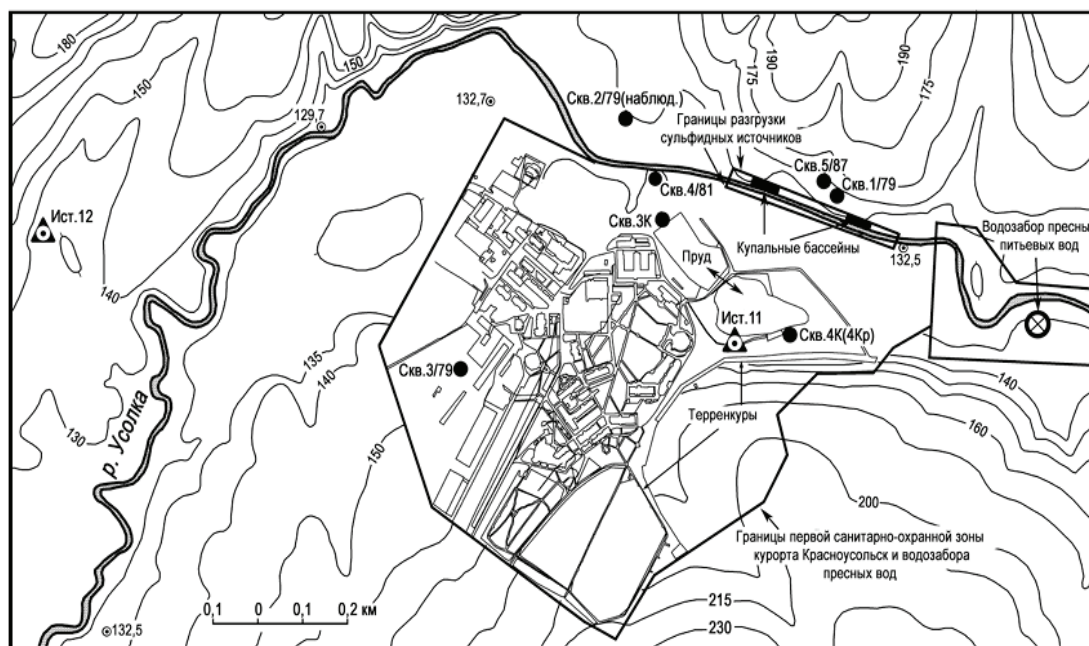


Рис. 1. Схема расположения минеральных источников курорта «Красноусольск» [1]

Происхождение вод инфильтрационное; химический состав формируется за счет выщелачивания солевого комплекса каменноугольных пород; природа сероводорода биогенная.

Формирование геохимической гаммы сероводородных вод месторождения Красноусольское 1 осуществляется при участии процесса смешения рассольных и пресных вод. Процесс формирования сероводородных вод происходит в слабопромытых засоленных каменноугольных породах лагунно-морского происхождения в гидрогеодинамической зоне затрудненного водообмена на глубине > 400–600 м. Термобарические параметры этой зоны и литолого-геохимическая обстановка в ней благоприятны для образования сульфидов за счет процесса сульфатредукции.

Месторождение Красноусольское 2 радоновых вод находится на территории курорта. Здесь, в основании ле-

вого склона долины Усолки из-под четвертичных осадков выбивает грифон соленой воды с дебитом 0,1–0,2 л/с и температурой 10,0–10,5°C. Это бессульфидная хлоридная натриевая вода с повышенным содержанием кальция (10,3–14,4 %), относящаяся к достаточно хорошо выраженному хлор-кальциевому типу, являющемуся основным геохимическим типом глубокозалегающих подземных вод.

Месторождение минеральных вод Красноусольское 3 в геохимическом отношении занимает особое положение. Оно расположено в 1,5 км западнее курорта и представлено источником Горький Ключ. Дебит его 40 л/с, температура воды — 6,1–6,5° С. Генетически он не относится не посредственно к Красноусольской группе минеральных вод и приурочен к закарстованным гипсам кунгура, слагающим правый склон долины р. Усолки. Вода источника по составу сульфатная кальциевая с минерализацией 2,2 г/дм<sup>3</sup>.

#### Литература:

1. Абдрахманов Р. Ф. Красноусольские минеральные воды // Вестник Академии наук Республики Башкортостан. 2014. № 4.



2. Абдрахманов Р. Ф., Попов В. Г. Геохимия и формирование подземных вод Южного Урала. 2010.
3. Абдрахманов Р. Ф. Пресные подземные и минеральные лечебные воды Башкортостана. Уфа: Гилем, 2014. 416 с.
4. Абдрахманов Р. Ф., Попов В. Г. Минеральные лечебные воды Башкортостана. Уфа: Гилем, 1999. 298 с.
5. Вахрушев Г. В. Минеральные воды и грязи Башреспублики // Хозяйство Башкирии, 1929. № 2–3. С. 155–176.

# Молодой ученый

## Международный научный журнал

### № 16 (619) / 2026

Выпускающий редактор Г. А. Письменная  
Ответственные редакторы Е. И. Осянина, О. А. Шульга, З. А. Огурцова  
Художник Е. А. Шишков  
Подготовка оригинал-макета П. Я. Бурьянов, М. В. Голубцов, О. В. Майер

За достоверность сведений, изложенных в статьях, ответственность несут авторы.  
Мнение редакции может не совпадать с мнением авторов материалов.  
При перепечатке ссылка на журнал обязательна.  
Материалы публикуются в авторской редакции.

Журнал размещается и индексируется на портале eLIBRARY.RU, на момент выхода номера в свет журнал не входит в РИНЦ.

Свидетельство о регистрации СМИ ПИ № ФС77-38059 от 11 ноября 2009 г., выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор).

ISSN-L 2072-0297

ISSN 2077-8295 (Online)

Учредитель и издатель: ООО «Издательство Молодой ученый». 420029, Республика Татарстан, г. Казань, ул. Академика Кирпичникова, д. 25.

Номер подписан в печать 29.04.2026. Дата выхода в свет: 06.05.2026.

Формат 60×90/8. Тираж 500 экз. Цена свободная.

Почтовый адрес редакции: 420140, Республика Татарстан, г. Казань, ул. Юлиуса Фучика, д. 94А, а/я 121.

Фактический адрес редакции: 420029, Республика Татарстан, г. Казань, ул. Академика Кирпичникова, д. 25.

E-mail: [info@moluch.ru](mailto:info@moluch.ru); <https://moluch.ru/>

Отпечатано в типографии издательства «Молодой ученый», 420029, Республика Татарстан, г. Казань, ул. Академика Кирпичникова, д. 25.