

УДК 004.9+624.131.37

DOI: [10.26102/2310-6018/2024.44.1.007](https://doi.org/10.26102/2310-6018/2024.44.1.007)

Система комплексного хранения данных геологических лабораторных испытаний

Н.Р. Тишин^{1,2✉}, О.Р. Озмидов², А.В. Пролетарский¹

¹Московский государственный технический университет им. Н.Э. Баумана, Москва, Российская Федерация

²АО «МОСТДОРГЕОТРЕСТ», Москва, Российская Федерация

Резюме. В статье рассматривается разработка нового подхода к хранению и организации результатов лабораторных опытов с учетом специфики их последующей обработки. Для решения поставленной задачи лабораторные опыты рассматриваются как структурированные данные с неструктурированными частями. При разработке системы была проанализирована специфика хранения и обработки данных лабораторных испытаний, после чего сформулированы основные требования к системе. Были определены основные модели данных, а также сущности базы. Для хранения структурированных данных выбрана стандартная реляционная модель данных, а хранение неструктурированной информации (такой как результаты опыта или параметры опыта) реализовано через поле BSON. Для решения задачи обеспечения защищенного доступа, а также создания API для системы был выбран асинхронный фреймворк FastAPI. Также рассмотрена реализация хранения дополнительных файлов опыта, которые находятся в объектном хранилище и связываются с опытом в реляционной модели через дополнительную сущность. Представленный подход отличается своей гибкостью к структуре хранимых лабораторных опытов, учитывает специфику геологических лабораторных испытаний, а также предоставляет возможности для комплексного метаанализа больших объемов данных. Система была протестирована и внедрена в технологический процесс геотехнической лаборатории АО «МОСТДОРГЕОТРЕСТ».

Ключевые слова: хранение данных геологических лабораторных испытаний, неструктурированные данные, система хранения результатов опытов, геоинформационная система, база данных, геологическая среда, информационный ресурс, инженерная геология.

Для цитирования: Тишин Н.Р., Озмидов О.Р., Пролетарский А.В. Система комплексного хранения данных геологических лабораторных испытаний. *Моделирование, оптимизация и информационные технологии*. 2024;12(1). URL: <https://moitvvt.ru/ru/journal/pdf?id=1495> DOI: 10.26102/2310-6018/2024.44.1.007

Integrated data storage system for geological laboratory experiments

N.R. Tishin^{1,2✉}, O.R. Ozmidov², A.V. Proletarsky¹

¹Bauman Moscow State Technical University, Moscow, the Russian Federation

²JSC MOSTDORGEOTREST, Moscow, the Russian Federation

Abstract. The article examines the development of a new approach to storing and organizing the results of laboratory experiments with consideration to the features of their subsequent processing. To solve this problem, laboratory experiments are considered as structured data with unstructured parts. During the development of the system, the features of storing and processing laboratory test data were analyzed, after which the basic requirements for the system were formulated. The main data models were defined as well as the database entities. A standard relational data model has been chosen for storing structured data, and the storage of unstructured information such as experiment results or experiment parameters is implemented through the BSON field. To solve the problem of providing secure access and creating

an API for the system, the asynchronous FastAPI framework was chosen. The implementation of storing additional experiment files, which are located in the object storage and are associated with the experiment in the relational model through an additional entity, is also considered. The presented approach is notable for its flexibility to the structure of stored laboratory experiments, takes into account the features of geological laboratory experiments and also provides opportunities for complex meta-analysis of large volume of data. The system was tested and implemented into the technological process of the geotechnical laboratory at JSC MOSTDORGEOTREST.

Keywords: storage of geological laboratory experiment data, unstructured data, experiment results storage system, geoinformation system, database, geological environment, information resource, engineering geology.

For citation: Tishin N.R., Ozmidov O.R., Proletarsky A.V. Integrated data storage system for geological laboratory experiments. *Modeling, Optimization and Information Technology*. 2024;12(1). URL: <https://moitvvt.ru/ru/journal/pdf?id=1495> DOI: 10.26102/2310-6018/2024.44.1.007 (In Russ.).

Введение

На современном этапе развития геологических изысканий крайне важен вопрос хранения и обработки информации. Создание баз данных, формализация данных и применение компьютерной обработки дают изыскателям практические возможности перехода к метаанализу данных лабораторных испытаний, что, в свою очередь, уже позволяет как проверять отдельные данные, так и строить системы прогнозирования. При достаточно активной разработке глобальных геоинформационных систем (ГИС) [1-4], некоторые прикладные аспекты, такие как хранение результатов геологических лабораторных испытаний, остаются на достаточно низком уровне, и зачастую реализуются без использования системного подхода локально на предприятиях или в качестве дополнительного модуля для ГИС [5]. На текущий момент для хранения физических характеристик грунтов широко используется программный комплекс EngGeo [6, 7], который реализует реляционную модель физических свойств грунтов. Программный комплекс EngGeo хорошо зарекомендовал себя в качестве базы данных физических свойств, но в то же время в комплексе отсутствует возможность хранения данных по различным опытам, таким как определения прочностных и деформационных свойств, динамических свойств и других характеристик [8, 9], определяемых по стандартам ГОСТ 12248.3-2020 и ГОСТ Р 56353-2022.

Для решения поставленной задачи была разработана система комплексного хранения данных геологических лабораторных испытаний грунтов. При анализе специфики хранения и обработки данных испытаний были выделены следующие основные требования к системе:

1. Наряду с хранением текстовых данных результатов опытов требуется структура данных для хранения дополнительных файлов опыта, таких как файл прибора, и различные массивы данных, полученные с прибора. Это требование становится крайне актуальным с учетом выходов новых редакций ГОСТов по лабораторным испытаниям, а также разработкой новых конечно-элементных моделей грунтов и обеспечения их входными параметрами. Опыт показывает, что для построения различных статистических или прогнозирующих моделей требуется получение большого набора исходных данных, которые могли не определяться ранее, и дополнительная обработка уже проведенных опытов может решить эту задачу.

2. Система должна хранить результаты различных опытов, а также иметь настолько гибкую структуру, чтобы возможность добавления новых типов опытов не требовало изменения общей структуры.

3. Система должна обеспечивать безопасное хранение данных, исключая возможность несанкционированных подключений.

4. Система должна учитывать и другие базы опытов, такие как EngGeo, в целях обеспечения получения комплексных срезов данных.

5. Система должна иметь возможность встраиваться в более общие геоинформационные системы

Материалы и методы

Поскольку результаты лабораторных испытаний носят неструктурированный характер, но при этом сама структура отношений опытов и их привязок к скважинам и объектам отлично описываются в реляционной модели, в качестве основной системы управления базами данных (СУБД) для хранения результатов лабораторных испытаний была взята СУБД PostgreSQL. Преимуществами данной СУБД является поддержка транзакций, наличие полей для неструктурированных данных BSON с возможностью покрытия индексами по ключам, а также удовлетворительная производительность относительно NoSQL СУБД MongoDB [10].

Для исключения транзитивной функциональной зависимости и нормализации базы до нормальной формы Бойса-Кодда были выделены 5 основных сущностей: объект, скважина, образец, опыт и тип опыта. Для поддержки совместимости системы с базой EngGeo уникальные идентификаторы объектов, скважин и образцов назначаются пользователем и должны совпадать с аналогичными идентификаторами в базе EngGeo. Связь между опытами и образцами осуществляется как многие к одному. Схема реляционной модели представлена на Рисунке 1.

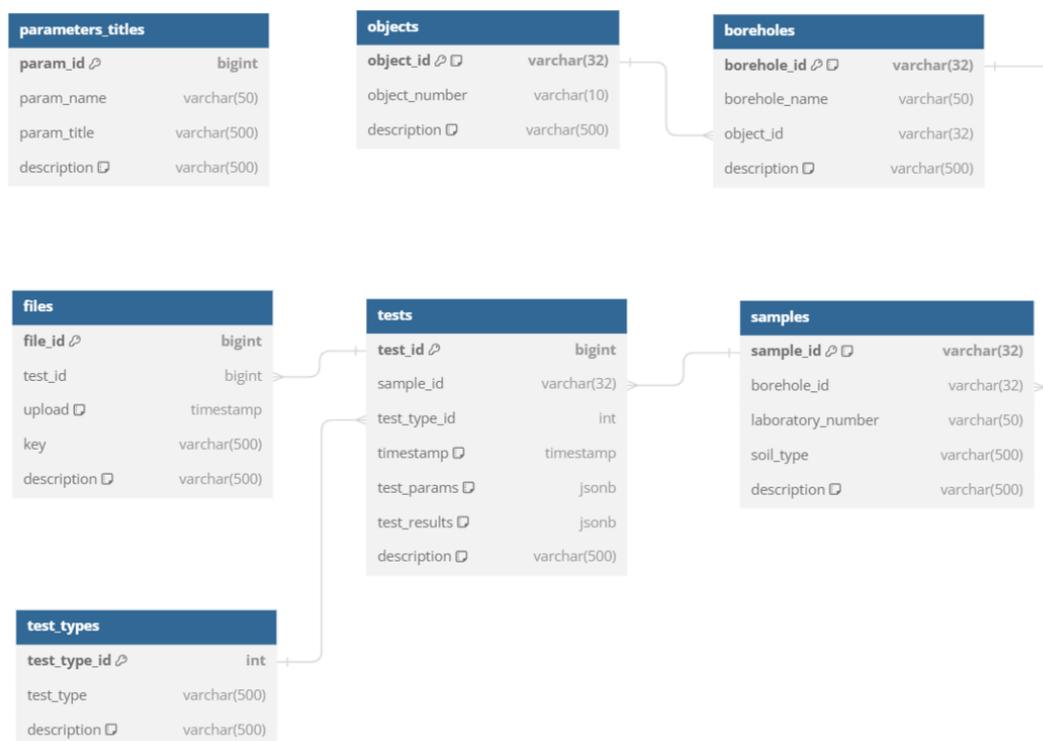


Рисунок 1 – Схема реляционной модели данных
Figure 1 – Diagram of the relational data model

Для реализации сервиса доступа к базе данных был использован фреймворк fastapi. Данный выбор обусловлен высокой скоростью разработки и читаемостью кода данного фреймворка, асинхронной реализацией, а также возможностью валидации данных на уровне приложения [11]. Поскольку обновление поля BJSON происходит полностью, на уровне приложения операция обновления реализована как внесение дополнений или обновления данных внутри поля BJSON по соответствующим ключам:

```
async def update(self, test_id: int, test_data: TestUpdate) -> tables.Tests:
```

```
...
test_params = test.test_params
test_results = test.test_results
data = test_data.to_dict()
if data.get("test_params", None):
    test_params.update(data["test_params"])
    data["test_params"] = test_params
if data.get("test_results", None):
    test_results.update(data["test_results"])
    data["test_results"] = test_results
q = update(
    tables.Tests
).where(
    tables.Tests.test_id == test_id
).values(
    **data
)
...
```

Реализация хранения дополнительных файлов происходит с помощью объектного хранилища и дополнительной сущности файлов, содержащей ключ файла в объектном хранилище и реализованной отношением многие к одному в реляционной модели. Поскольку одним из требований к системе является разграничение доступа, то связь между пользователем и объектным хранилищем должна осуществляться через приложение с сокрытием данных о URL и имени бакета. Для этого в приложении реализована зависимость для асинхронного доступа к объектному хранилищу:

```
from aiobotocore.session import get_session
class S3Service:
```

```
def __init__(self, client):  
    self.client = client  
  
    async def upload(self, key: str, data: bytes):  
        return await self.client.put_object(  
            Bucket=configs.bucket,  
            Key=key,  
            Body=data  
        )  
  
    async def delete(self, key: str):  
        return await self.client.delete_object(  
            Bucket=configs.bucket,  
            Key=key  
        )  
  
    async def get(self, key: str):  
        return await self.client.get_object(  
            Bucket=configs.bucket,  
            Key=key  
        )  
  
    async def get_s3_service():  
        session = get_session()  
        async with session.create_client(  
            's3',  
            endpoint_url=configs.endpoint_url,  
            region_name=configs.region_name,  
            aws_secret_access_key=configs.aws_secret_access_key,  
            aws_access_key_id=configs.aws_access_key_id  
        ) as client:  
            yield S3Service(client)
```

Поскольку система должна иметь возможность обмена данными с различными геоинформационными системами, в приложение добавлен сервис с реализацией REST API. Для удовлетворения требованиям безопасности добавлена поддержка авторизации через JWT-токен. Общая схема системы хранения результатов опытов геологических испытаний представлена на Рисунке 2.

Для удобства просмотра информации по сохраненным результатам испытаний, а также для доступа к загруженным файлам, помимо API, в систему добавлен веб-интерфейс для просмотра данных лабораторных испытаний (Рисунок 3).

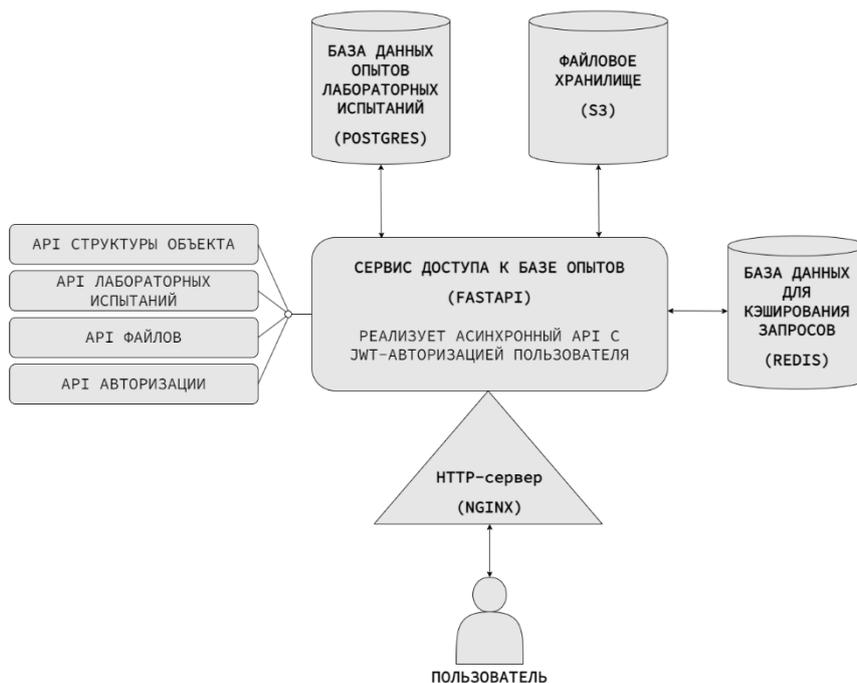


Рисунок 2 – Общая структура системы хранения данных лабораторных испытаний
Figure 2 – General structure of the laboratory test data storage system

Объекты				Опыты				
Номер объекта	Описание	ID		ID	Скважина	Л/Н	Тип опыта	Дата
615-23	None	964746934784478FA753CA58BB53240		88	147	147-2	Резонансная колонка	26.08.2022
588-23	None	09C11129257448B83EA34F579846045		83	132	132-4	Резонансная колонка	26.08.2022
762-23	None	B6060D31025C49358381FF2576B208C		84	132	132-5	Резонансная колонка	26.08.2022
555-22	None	A094227A750E490DEB79A93998F188B68		85	135	135-5	Резонансная колонка	26.08.2022
				86	136	136-3	Резонансная колонка	26.08.2022
				87	145	145-3	Резонансная колонка	26.08.2022

Результаты		Параметры		Результаты		Файлы	
Л/Н						Название	Описание
147-2		e_Q54		EQ_13145		tests_db/files/88/file	дирт файл словарь, содержащий shear_strain и G_опыт
		p_ref_01		G0_54.77			
				threshold_shear_strain_156			

Рисунок 3 – Веб-интерфейс для просмотра данных лабораторных испытаний
Figure 3 – Web interface for viewing laboratory test data

Результаты

На текущий момент система комплексного хранения данных геологических лабораторных испытаний является основной системой хранения и аналитической

обработки лабораторных опытов в компании АО «МОСТДОРГЕОТРЕСТ». Опыт практического использования системы показывает, что современные методы хранения и обработки данных позволяют создавать эффективные комплексные системы для хранения данных лабораторных геологических испытаний с учетом их специфики. Изначально организация системы разрабатывалась как OLAP без серьезных постоянных нагрузок сложными запросами, поэтому дополнительные индексы, кроме автоматически создаваемых БД PostgreSQL, не создавались.

Обсуждение

При сложных запросах по ключам неструктурированных данных СУБД PostgreSQL позволяет создавать специализированный индекс GIN, увеличивающий скорость запросов к неструктурированным данным. Дальнейшее развитие системы представляет собой построение аналитических сервисов для получения и обработки сложных срезов данных, включающих в себя как физические параметры, так и механические и динамические параметры грунтов.

Заключение

Текущий этап развития прикладных наук и технологий позволяет успешно решать такие задачи, как создание и обслуживание Системы комплексного хранения данных геологических лабораторных испытаний. В работе были определены основные требования к системе, выбраны основные модели хранения данных. В процессе тестирования системы были получены результаты работы, удовлетворяющие современным потребностям хранения и обработки данных. Выбор реляционной модели позволил получить нормализованную структуру данных, а дополнительные возможности СУБД PostgreSQL по хранению неструктурированных данных позволили добавить в систему такие данные, как результаты геологических испытаний, с возможностью их обновления без изменения основной структуры данных.

Использование асинхронного фреймворка FastAPI позволило решить такие задачи, как защищенный доступ к API системы хранения данных, а также добиться высокой производительности обработки запросов, а реализация доступа к объектному хранилищу S3, написанная с использованием данного фреймворка, позволила реализовать в системе дополнительную возможность хранения дополнительных файлов, таких как лог-вайлы прибора, результаты обработки лог-файла и другие.

СПИСОК ИСТОЧНИКОВ

1. Красильников П.А., Хронусов В.В., Барский М.Г. Принципы создания и ведения базы данных инженерно-геологической информации. *Геология и полезные ископаемые Западного Урала*. 2018;(18):252–257.
2. Dos Santos J.V., Thiesen S., dos Reis Higashi R.A. Geological-Geotechnical database from standard penetration test investigations using geographic information systems. *Management of Information Systems*. 2018:245–256. DOI: 10.5772/intechopen.74208.
3. Анищик В.В. и др. Базы данных как основа работы геологического направления. *ПРОНЕФТЬ. Профессионально о нефти*. 2022;(3):76–79.
4. Красильников П.А. Принципы формирования инженерно-геологических баз данных при разработке месторождений полезных ископаемых. *Ученые записки Крымского федерального университета имени В.И. Вернадского. География. Геология*. 2019;5(3):345–357.

5. Левшина М.И., Токаревский П.А. Автоматизированное хранение и обработка данных инженерно-геологических изысканий. *Перспективы развития инженерных изысканий в строительстве в Российской Федерации*. 2016:44–47.
6. Красильников П.А. Обзор современных программных продуктов, используемых в инженерной геологии. *Геология и полезные ископаемые Западного Урала*. 2020;(3):51–60.
7. Болдырев Г.Г., Дивеев А.А. К вопросу использования информационных систем при изысканиях и проектировании оснований фундаментов зданий и сооружений. *Независимый электронный журнал "GeoИнфо"*. 2020;(2). URL: <https://geoinfo.ru/product/boldyrev-gennadij-grigorevich/k-voprosu-ispolzovaniya-informacionnyh-sistem-pri-izyskaniyah-i-proektirovanii-osnovanij-fundamentov-zdaniij-i-sooruzhenij-42530.shtml> (дата обращения: 25.12.2023).
8. Строкова Л.А. Определение параметров для численного моделирования поведения грунтов. *Известия Томского политехнического университета*. 2008;313(1):69–74. URL: <https://cyberleninka.ru/article/n/opredelenie-parametrov-dlya-chislennogo-modelirovaniya-povedeniya-gruntov> (дата обращения: 25.12.2023).
9. Строкова Л.А. Научно-методические аспекты создания расчетных моделей грунтовых оснований. *Известия Томского политехнического университета*. 2010;316(1):151–156. URL: <https://cyberleninka.ru/article/n/nauchno-metodicheskie-aspekty-sozdaniya-raschetnyh-modeley-gruntovyh-osnovanij> (дата обращения: 25.12.2023).
10. Lindvall J., Sturesson A. *A comparison of latency for MongoDB and PostgreSQL with a focus on analysis of source code*. Dissertation; 2021.
11. Peralta J.H. *Microservice APIs: Using Python, Flask, FastAPI, OpenAPI and More*. New York, Simon and Schuster; 2023. 411 p.

REFERENCES

1. Krasilnikov P.A., Hronusov V.V., Barskii M.G. Principles of creation and modelling engineering geology database. *Geologiya i poleznye iskopaemye Zapadnogo Urala*. 2018;(18):252–257. (In Russ.).
2. Dos Santos J.V., Thiesen S., dos Reis Higashi R.A. Geological-Geotechnical database from standard penetration test investigations using geographic information systems. *Management of Information Systems*. 2018:245–256. DOI: 10.5772/intechopen.74208.
3. Anishhik B.B. et al. Geological databases, as the basis for operation in geological direction. *PRONEFT. Professionalno o nefti = PROneft. Professionally about Oil*. 2022;(3):76–79. (In Russ.).
4. Krasilnikov P.A. Principles of creation and modelling engineering geology database of mineral deposits. *Uchenye zapiski Krymskogo federalnogo universiteta imeni V.I. Vernadskogo. Geografiya. Geologiya*. 2019;5(3):345–357. (In Russ.).
5. Levshina M.I., Tokarevskii P.A. Avtomatizirovannoe khranenie i obrabotka dannyh inzhenerno-geologicheskikh izyskanii. *Perspektivy razvitiya inzhenernykh izyskanii v stroitelstve v Rossiiskoi Federatsii*. 2016:44–47. (In Russ.).
6. Krasilnikov P.A. Review of modern software products used in engineering geology. *Geologiya i poleznye iskopaemye Zapadnogo Urala*. 2020;(3):51–60. (In Russ.).
7. Boldyrev G.G., Diveev A.A. K voprosu ispolzovaniya informatsionnykh sistem pri izyskaniyah i proektirovanii osnovanii fundamentov zdaniij i sooruzhenii. *Nezavisimyy elektronnyy zhurnal "GeoInfo"*. 2020;(2). URL: <https://geoinfo.ru/product/boldyrev-gennadij-grigorevich/k-voprosu-ispolzovaniya-informacionnyh-sistem-pri-izyskaniyah-i->

- [proektirovanii-osnovanij-fundamentov-zdaniy-i-sooruzhenij-42530.shtml](#) (accessed on 25.12.2023). (In Russ.).
8. Strokova L.A. Opredelenie parametrov dlya chislennogo modelirovaniya povedeniya gruntov. *Izvestiya Tomskogo politekhnicheskogo universiteta*. 2008;313(1):69–74. URL: <https://cyberleninka.ru/article/n/opredelenie-parametrov-dlya-chislennogo-modelirovaniya-povedeniya-gruntov> (accessed on 25.12.2023). (In Russ.).
 9. Strokova L.A. Nauchno-metodicheskie aspekty sozdaniya raschetnykh modelei gruntovykh osnovanii. *Izvestiya Tomskogo politekhnicheskogo universiteta*. 2010;316(1):151–156. URL: <https://cyberleninka.ru/article/n/nauchno-metodicheskie-aspekty-sozdaniya-raschetnykh-modeley-gruntovykh-osnovanii> (accessed on 25.12.2023). (In Russ.).
 10. Lindvall J., Stuesson A. *A comparison of latency for MongoDB and PostgreSQL with a focus on analysis of source code*. Dissertation; 2021.
 11. Peralta J.H. *Microservice APIs: Using Python, Flask, FastAPI, OpenAPI and More*. New York, Simon and Schuster; 2023. 411 p.

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Тишин Никита Романович, аспирант, Московский государственный технический университет им. Н.Э. Баумана, Москва, Российская Федерация.
e-mail: tnick1502@mail.ru
ORCID: [0009-0008-2484-591X](https://orcid.org/0009-0008-2484-591X)

Nikita R. Tishin, Postgraduate Student, Bauman Moscow State Technical University, Moscow, the Russian Federation.

Озмидов Олег Ростиславович, кандидат геолого-минералогических наук, академик Российской Академии Естественных Наук, научный руководитель геотехнической лаборатории, АО «МОСТДОРГЕОТРЕСТ», Москва, Российская Федерация.
e-mail: ozmidov@mail.ru

Oleg R. Ozmidov, Candidate of Geological and Mineralogical Sciences, Academician at the Russian Academy of Natural Sciences, Scientific Director of the Geotechnical Laboratory, JSC MOSTDORGEOTREST, Moscow, the Russian Federation.

Пролетарский Андрей Викторович, доктор технических наук, декан факультета Информатика и системы управления, Московский государственный технический университет им. Н.Э. Баумана, Москва, Российская Федерация.
e-mail: pav@bmstu.ru
ORCID: [0000-0002-6932-0016](https://orcid.org/0000-0002-6932-0016)

Andrey V. Proletarsky, Doctor of Technical Sciences, Dean of the Faculty of Informatics and Control Systems, Bauman Moscow State Technical University, Moscow, the Russian Federation.

Статья поступила в редакцию 22.12.2023; одобрена после рецензирования 22.01.2024; принята к публикации 31.01.2024.

The article was submitted 22.12.2023; approved after reviewing 22.01.2024; accepted for publication 31.01.2024.