

УДК 004.415.25

Н. Н. Бич

Гродненский государственный медицинский университет

**ПРОГРАММНЫЙ МОДУЛЬ МНОГОЧИСЛЕННОЙ
КОНВЕРТАЦИИ ФАЙЛОВ OBJ В ФОРМАТ CSV**

Актуальность данной работы обусловлена отсутствием соответствующего приложения как в отечественном мире разработок, так и в зарубежном. Наличие такой разработки поможет одновременно проанализировать содержимое многочисленного количества файлов. Основная цель работы – реализовать программный модуль для массового преобразования файлов формата obj в формат csv. Рассматриваемые файлы получены путем непрерывного сканирования конфокальным микроскопом гистологических образцов. В каждой точке образца считывается суперпозиция пиков спектра, формируются obj-файлы. Для удобства исследования полученные файлы необходимо преобразовать в простой для анализа формат csv. Разработка самостоятельного программного модуля на языке R является решением данной задачи. Выбранный язык – целая рабочая среда, куда уже встроены готовые методы статистического анализа и инструменты для визуализации. Разработанное приложение однозадачно, но его открытый код позволит в дальнейшем дописать дополнительные библиотеки для анализа полученных данных и рассматривать их не только как выборки небольших объемов, но и в формате BigData. Его плюсы заключаются в том, что он работает почти под любой операционной системой, включая Windows7, не имеет аналогов, может быть использован в дальнейшем как базовый модуль спроектированной и реализованной сложной информационной системы, позволяющей анализировать данные, обучаться, делать прогнозы.

Ключевые слова: программирование, конвертация, пик, файлы obj, формат csv.

Для цитирования: Бич Н. Н. Программный модуль многочисленной конвертации файлов obj в формат csv // Труды БГТУ. Сер. 3, Физико-математические науки и информатика. 2026. № 1 (302). С. 128–135.

DOI: 10.52065/2520-6141-2026-302-11.

N. N. Bich

Grodno State Medical University

**MULTIPLE CONVERSION SOFTWARE MODULE
OBJ FILES IN CSV FORMAT**

The relevance of this work is due to the lack of an appropriate application, both in the domestic world of development and abroad. The presence of such a development will help to temporarily analyze the contents of a large number of files. The main purpose of the work is to implement a software module for mass conversion of obj format files to csv format. The files in question were obtained by continuous scanning of histological samples with a confocal microscope. The super-position of the spectrum peaks is read at each point of the sample, and obj files are generated. For the convenience of research, the obtained files must be converted into an easy-to-analyze cvs format. The development of an independent software module in the R language is the solution to this problem. The chosen language is an entire working environment, where ready-made statistical analysis methods and visualization tools are already embedded. The developed application is single-tasking, but its open source code will allow additional libraries to be added in the future to analyze the data obtained, considering them not only as small samples, but also in BigData format. Its advantages are that it works under almost any operating system, including Windows7, has no analogues, and can be used as a basic module of a further designed and implemented complex information system that allows data analysis, learning, and making predictions.

Keywords: programming, conversion, peak, obj files, csv format.

For citation: Bich N. N. Multiple conversion software module obj files in csv format. *Proceedings of BSTU, issue 3, Physics and Mathematics, Informatics*, 2026, no. 1 (302), pp. 128–135 (In Russian).

DOI: 10.52065/2520-6141-2026-302-11.

Введение. В химии и биологии основной целью исследования спектров является анализ спектральных пиков. Результат решения данной задачи есть цифровая интерпретация химического состава или строения образца, представленная метрическими массивами данных.

Процесс проведения спектрального анализа сопровождается рядом сложностей, связанных с отсутствием или частичным наличием автоматизации обработки данных спектра. Процедура автоматического анализа постоянно прерывается вмешательством человека, что замедляет процесс обработки. Это объясняется отсутствием полного программного компонентного комплекса, включающего помимо модуля формирования спектрального изображения ряд дополнительных частных блоков, реализующих решение отдельных задач посредством удобной для пользователя программной графической оболочки. В качестве самостоятельных программных модулей выступают процедуры, выполняющие следующие функции: считывание данных из файлов, представление данных в формате, удобном для исследования, визуализация данных, их статистический анализ, интерпретация полученных результатов. Цель данной публикации – описать разработку программного блока, который является частью подготовительного этапа на пути основного исследования полученных данных. В частности, речь идет о трех первых опциях программного комплекса: считывание, форматирование, визуализация.

Основная часть. В данной работе исследованию подлежат детальные спектры люминесценции наночастиц CdSe/ZnS во многих точках гистологического образца [1]. Регистрация подобных сигналов была проведена с использованием конфокального микроскопа «NanoFinder 30». Микроскоп функционирует в режиме сканирования с диапазоном регистрации длин волн: 480–800 нм [1]. В качестве гистологических препаратов использовались следующие образцы: 8390, 13920, 25020, 32205 (все – пораженные); 27583, 28272, 51074 (дисплазия); 11589, 16286 (здоровые ткани) [1].

Тонирование образцов спектрами люминесценции наночастиц CdSe/ZnS, записанными в различных pH-окружениях (pH: 1, 3, 4, 6, 7, 9, 10) – так называемых эталонных спектрах, проводилось программно по оригинальной методике [1]. Автоматически для каждой отдельно взятой точки образца считывалась суперпозиция пиков. Затем она визуализировалась в виде картинка с графиком зависимости интенсивности отраженной световой волны от ее длины. Изображения спектров были закодированы и сохранены в одинаковом формате orj для их дальнейшей обработки.

Файлы с расширением orj либо orju (новая версия orj с обновленной структурой unicode для сохранения проектов) – это файлы авторского формата, чьи права принадлежат компании OriginLab. Такой вид файлов используется для хранения научных данных, а также их визуализации посредством электронных таблиц и графиков [2].

Если открыть orj-документ с помощью специальной программы просмотра, информация, в конкретном случае о спектрах, будет представлена данными, размещенными на двух вкладках. Первая вкладка содержит график зависимости интенсивности отраженного света на определенной длине волны, вторая – электронную таблицу с двумя массивами числовых показателей: абсцисс и ординат точек графика, соответствующих количеству измерений, полученных для отдельного образца в группе исследования (рис. 1).

Чтобы работать с показателями дальше, сохранив их предварительно в отдельное хранилище, они должны быть доступны программному пакету, например Ms Excel, Statistica, или программной среде со встроенным языком программирования для проведения над ними определенного анализа с последующей интерпретацией результатов. OriginLab для этого подходит не в полной мере. Поэтому данные, представленные orj-файлами, необходимо конвертировать в наиболее универсальный для таких манипуляций формат, например csv.

Рассматриваемый стандарт прост, совместим с различными приложениями и платформами, широко используется для обмена данными между разными программами, базами данных и электронными таблицами. Небольшой размер csv делает его эффективным для обработки больших наборов данных. Простая структура (легко считываемые значения, разделенные запятыми, представленные в виде строк) облегчает его синтаксический анализ [3].

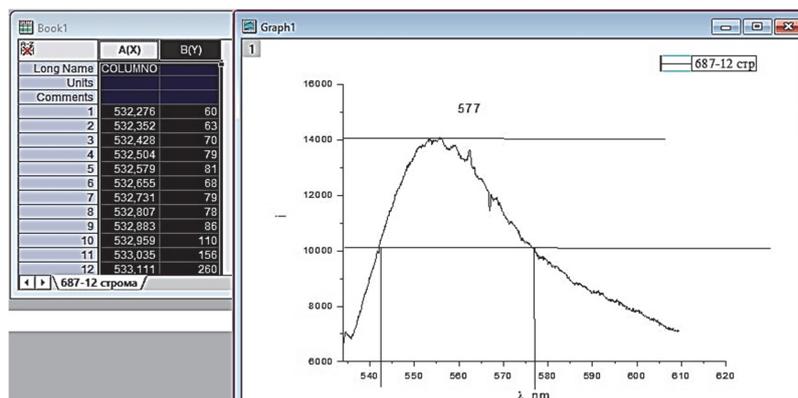


Рис. 1. Представление данных в OriginLab

Важность многочисленного автоматического преобразования orj-ресурсов в csv-формат обуславливается необходимостью работы с данными, хранящимися в них, как с единой генеральной совокупностью, как правило, большой размерности.

Существуют различные бесплатные конверторы, в том числе и среда просмотра OriginLab, позволяющие экспортировать orj-файлы в csv-стандарт. Недостатком всех этих приложений является возможность только единичного преобразования каждого отдельно взятого orj-объекта, а также их сепаративность относительно программ дальнейшей обработки данных в силу отсутствия между ними прямой связи. Это делает соответствующую процедуру монотонной и непродуктивной.

Новый вид orji результативнее в производительности относительно своего предшественника, более открыт для внешних программных средств его обработки, но, тем не менее, массовая конвертация таких файлов в csv-формат средой просмотра все так же невозможна.

Чтобы сделать этот процесс более продуктивным, целесообразно воспользоваться альтернативными программными продуктами. Для этого был сделан обзор существующих решений, соответствующих поставленной задаче. В результате не было выявлено каких-либо конкретных законченных вариантов, готовых к конечному использованию. Предлагаемые решения, например, [4], в основном представляют собой часть задачи и, как правило, содержат ошибки и недочеты, требуют доработки и могут использоваться лишь как основа для дальнейшей разработки.

В связи с этим было предложено разработать оригинальное приложение посредством языков программирования, имеющих помимо возможности подключения соответствующих библиотек для работы с orj-файлами, средства анализа данных.

Среда просмотра orj-файлов Origin имеет встроенный язык сценариев (LabTALK). LabTALK может быть расширен с помощью встроенного языка компиляции Origin C, в основе которого лежит C / C ++. Основная функция данного инструмента – управление программным обеспечением, что делает LabTALK неудобным и малоэффективным для обработки и анализа данных.

Эффективнее всего воспользоваться преимуществами таких языков программирования, как Python или R. Выбор сделан в пользу последнего, поскольку данное средство оказалось менее требовательным относительно совместимости версий операционной системы и языка программирования при решении задачи импорта orj-формата в формат csv.

Язык программирования R обладает богатой экосистемой библиотек, предоставленных пользователями со всего мира. Векторы применения этих библиотек разнообразны и охватывают такие актуальные инструменты для исследования, как статистика, машинное обучение, визуализация данных, анализ временных рядов и многое другое. Обширный функционал R-языка всегда легко может быть расширен соответствующей установкой дополнительных программных модулей. Мощный инструментарий R значительно сокращает время, необходимое для реализации сложных функциональных возможностей.

В данной публикации представлена разработка нового программного инструмента с открытым исходным кодом. Этот инструмент, как дополнительный модуль, позволяет пользователям выполнять быструю конвертацию многочисленных `orj`-файлов в формат `csv` через простой пользовательский интерфейс, что значительно ускоряет подготовку данных для последующего исследования. Измененные файлы сохраняются в требуемом формате в каталог, предусмотренный приложением по умолчанию. Так как файлы могут конвертироваться потоком, при сохранении им автоматически присваиваются исходные имена для того, чтобы было понятно, какой из файлов выборки был преобразован.

На рис. 2 представлена схема организации работы приложения.



Рис. 2. Схема организации работы приложения

Весь код программы, как было указано выше, написан на языке программирования R. Сценарий содержит два основных момента.

Первый заключается в том, чтобы иметь возможность дальнейшей перспективной модификации программы. Для этого она нуждается в подвижном инструменте обработки показателей. В связи с этим рационально было дополнить существующий сценарий интерактивными эффектами `web`-приложения с графическим интерфейсом пользователя (сокращенно `UI`) – огромное преимущество для небольшого приложения. Решить подобную задачу помогло подключение в коде необходимой библиотеки (фреймворка) `library(shiny)`. Данная библиотека представляет собой тщательно отобранный набор функций пользовательского интерфейса, которые генерируют скрипты `HTML`, `CSS` и `JavaScript`, необходимые для выполнения обычных задач и позволяющие запускать приложение непосредственно из браузера, не требуя при этом дополнительных знаний подробностей установки R, что существенно упрощает процесс инсталляции приложения для конечного потребителя.

`UI` и `Server` – два базовых модуля фреймворка `library(shiny)` (рис. 3, 4). Первый под руководством второго отвечает за сборку объектов, представляющих внешний вид приложения. Таким образом, `UI` – внешний вид программы, `Server` – ее техническая сторона, реализующая логику работы интерфейса, обеспечивая процесс взаимодействия с пользователем. Оба модуля непосредственно плотно коррелирует друг с другом и наделяют программу такой ключевой особенностью, как реактивность. Все расчеты, манипуляции, происходящие в серверной части, напрямую зависят от значений (`inputs`), введенных пользователем в интерактивном режиме. Изменения во внешней части вызывают реакцию сервера, который в свою очередь через выходные элементы (`outputs`) обновляет интернет-страницу, избегая ручного вмешательства [5, 6].

Второй момент основывается на чтении файлов `orj`. Для этого в коде программы используется библиотека `library(ropj)` [7]. Данный пакет экспортирует единственную полезную в данном контексте функцию `read.orj`:

```
read.orj(sample_file, encoding = "latin1", tree = FALSE),
```

способную прочитать файлы программы OriginLab и вернуть список объектов, из которых он состоит.

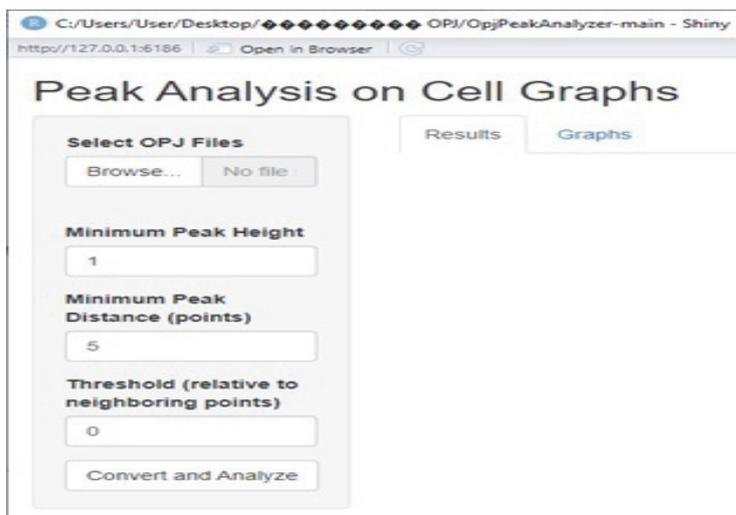


Рис. 3. Модуль UI – интерфейс программы

Функция включает несколько параметров. Ключевым из них является аргумент – `file`. В нашем случае он имеет имя `sample_file` и представляет собой путь к файлу, содержащему в данной работе электронную таблицу. В других вариантах файл может рассматриваться как массив или запись. Вторым аргументом функции `read.opj` – `encoding`, соответствует выбранной кодировке кортежей внутри открываемого файла [7]. Данный параметр в соответствии с кодовой страницей символов ANSI по умолчанию равен безопасному значению `latin1`.

```
server <- function(input, output, session) {
  output$column_select_ui <- renderUI({
    req(input$file_opj)
    sample_file <- input$file_opj$datapath[1]
    opj_data <- tryCatch({
      read.opj(sample_file, encoding = "latin1", tree = FALSE)
    }, error = function(e) {
      showNotification(paste("Error reading file:", basename(sample_file), "-", e$message), type = "error")
      return(NULL)
    })
    if (is.null(opj_data)) return(NULL)
    first_obj <- opj_data[[1]]
    if (!is.data.frame(first_obj)) return(NULL)
    cols <- names(first_obj)
    tagList(
      selectInput("x_col", "Select X Column", choices = cols, selected = cols[1]),
      selectInput("y_col", "Select Y Column", choices = cols, selected = cols[2]),
      helpText("Phantom coordinates will be automatically removed before analysis.")
    )
  })
}
```

Рис. 4. Модуль Server – реализация интерфейса

Следующий показатель `tree`, относящийся к булевому типу, контролирует структуру возвращаемого списка [7]. Если его значение по умолчанию равно `FALSE`, то на выходе содержимое файла будет представлено плоским списком, а его идентификаторы соответствовать коротким именам объектов. В противном случае, если `tree` принимает значение `TRUE`, то список становится рекурсивной древовидной структурой данных с вложенным доступом к его объектам посредством указания имен всех узлов-элементов, предшествующих данному. Последний вариант имеет определенные сложности в работе, поэтому в рассматриваемом коде в качестве значения параметра `tree` используется значение `false`.

Определенно, функция `read.opj` дополнительно сопровождается рядом сценариев, обеспечивающих проверку корректности выполнения процесса считывания данных из файла. При этом она является ключевой и единственной функцией, предоставляющей доступ

к ресурсам орj-файла, что позволяет перейти к следующему важному шагу, а именно сохранить полученные объекты в подходящем для анализа данных формате csv с помощью функции `write.csv(obj, file = csv_path, row.names = FALSE, fileEncoding = "UTF-8")`, где параметр `obj` – объект орj файла, в нашем случае координаты точек спектра, представленные таблицей. Значение, которое принимает параметр `obj` с помощью функции `write`, пишется в формат csv. Логическая переменная `row.names` принимает значение `false`, так как наименование строк нам не требуется. В качестве атрибута `fileEncoding` взята кодировка «UTF-8», на тот случай, если в программе будут использоваться языки, не предусмотренные операционной системой.

Функция `write` аналогично функции `read` также сопровождается дополнительным кодом, соответствующим проверке на адекватность полученного результата.

После конвертации в приложении доступна демонстрация значений csv-файла на графике.

Иллюстрация результатов конвертации в первую очередь необходима для визуального сравнения данных, полученных в итоге исследования здоровых и больных биоклеток.

Есть возможность выбрать переменные, для которых нужно построить и отобразить зависимость. Также в программе (рис. 5) представлены первоначальные настройки для иллюстрации пиков спектра на графике:

- Minimum Peak Height – минимальная высота пика, отображаемая на рисунке;
- Minimum Peak Distance (points) – наименьшее допустимое расстояние между точками на графике;
- Threshold (relative to neighboring points) – пороговое значение, которое возможно принять за пик.

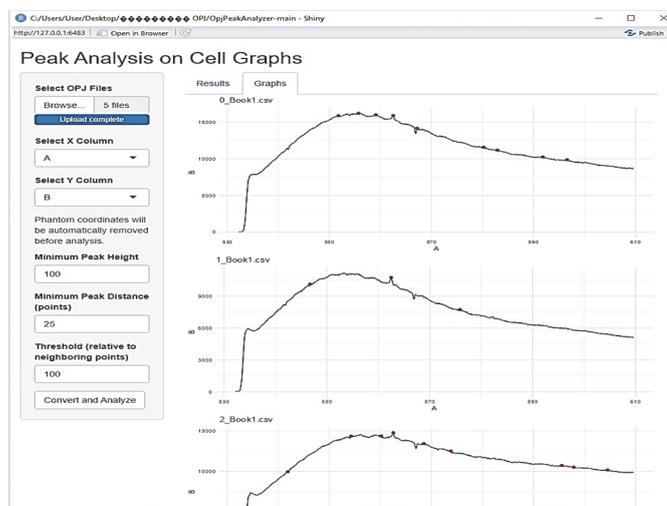


Рис. 5. Графическое представление значений csv-файла

Для идентификации пиков в массиве данных в коде задействована библиотека `library` (`grasta`), которая имеет в своем составе большое количество функций, в частности функции оптимизации. Для нахождения некоторых агрегированных значений пиков, их дополнительных числовых показателей (высот, расстояний, частот), осуществления над ними фильтрации в программе применяется библиотека `library` (`dplyr`).

Непосредственное построение графиков с несколькими фасадами осуществляется с помощью мощной и популярной библиотеки `library` (`ggplot2`). Основная ее функция `ggplot` имеет несколько параметров, отвечающих за набор отображаемых данных, тип и стиль графика.

Пики спектров, представленные в csv-формате, в перспективе позволят получить их характеристики (расстояния между соседними пиками, максимумы и минимумы среди пиков, высоты максимумов и минимумов, расстояние между ними) для дальнейшей обработки

и статистического анализа, чтобы понять, есть ли значимая разница в результатах детального исследования спектров люминесценции наночастиц гистологических образцов здоровых и больных биоклеток.

Заключение. В качестве заключения отметим, что приложение может работать в Linux, Windows и macOS. Оно не является многофункциональным, но его практическая значимость высока. Механизм преобразования указанного формата полностью автоматизирован, что дает возможность сократить время подготовки показателей биосигналов, представить их в удобной для обработки форме и планировать очередные шаги реализации следующих модулей целого программного комплекса, позволяющих провести наглядную оценку частоты зарегистрированного сигнала, так и более развернутую обработку, включающую вычисление важных параметров, на основе большого диапазона методов – от статистического до спектрального анализа [8].

Список литературы

1. Копыцкий А. В., Хильманович В. Н. Определение распределения рН гистологического среза, выдержанного в водном растворе квантовых точек CDSE/ZNS, по их спектрам люминесценции // Медицинский университет: современные взгляды и новые подходы: материалы Респ. науч.-практ. конф. с междунар. участием, посвященной 65-летию Гродненского гос. мед. ун-та, Гродно, 28–29 сент. 2023 г. Гродно, 2025. С. 11–12.
2. Шпаков К. В. Обработка данных в OriginPro // Data process manual (rus).md: сайт. 2020. URL: [https://github.com/shpakovkv/OriginPro/blob/master/Data process manual \(rus\).md](https://github.com/shpakovkv/OriginPro/blob/master/Data%20process%20manual%20(rus).md) (дата обращения: 12.12.2024).
3. Васильев А. Н. Экспорт файла JSON в CSV с помощью Python // Программирование на Python: сайт. 2021. URL: <https://pythonim.ru/osnovy/eksport-fayla-json-v-csv-s-pomoschyu-python?ysclid=m5qhk3yuxl396343004> (дата обращения: 12.12.2024).
4. Areed0615 Spectrophotometer-data-compiler // github: сайт. 2025. URL: <https://github.com/Areed0615/Spectrophotometer-data-compiler> (дата обращения: 12.12.2024).
5. Введение в машинное обучение // Учебник R: сайт. 2021. URL: <https://nweb42.com/books/r-lang/vvedenie-v-shiny/> (дата обращения: 12.12.2024).
6. Тихонов А. Н. Визуализация статистики производительности оборудования с R – Shiny // Хабр: сайт. 2015. URL: <https://habr.com/ru/articles/253419/> (дата обращения: 12.12.2024).
7. Garriga M. Package «Rorj» // CRAN: сайт. 2025. URL: <https://cran.r-project.org/web/packages/Rorj/Rorj.pdf> (дата обращения: 12.12.2024).
8. Бич Н. Н., Кынкурогов А. А., Скируха А. А. Разработка программного модуля многочисленной конвертации файлов OPJ в формат CSV // Модернизация высшего образования в сторону цифровизации проблемы, решения, перспективы: материалы II Респ. науч.-практ. конф. с междунар. участием, Гродно, 12 марта. 2025 г. Гродно, 2025. С. 11–12.

References

1. Kopytskiy A. V., Khil'manovich V. N. Determination of the PH distribution of a histological section aged in an aqueous solution of CDSE/ZNS quantum dots by their luminescence spectra *Meditsinskiy universitet: sovremennyye vzglyady i novyye podkhody: materialy Rosp. nauch.-prakt. konf. s mezhdunar. uchastiyem, posvyashchennoy 65-letiyu Grodnenskogo gos. med. un-ta* [Medical University: modern views and new approaches: proceedings of the Republican scientific and practical conference with international participation dedicated to 65-anniversary of the Grodno State Medical University]. Grodno, 2025, pp. 11–12 (In Russian).
2. Shpakov K. V. Data processing in OriginPro [*Data process manual (rus).md*]: website. 2020. Available at: [https://github.com/shpakovkv/OriginPro/blob/master/Data process manual \(rus\).md](https://github.com/shpakovkv/OriginPro/blob/master/Data%20process%20manual%20(rus).md) (accessed 12/12/2024) (In Russian).

3. Vasil'yev A. N. Exporting a JSON file to CSV. *Programmirovaniye na Python: sayt* [Python Programming]: website, 2021. Available at: <https://pythonim.ru/osnovy/eksport-fayla-json-v-csv-s-pomoschu-python?ysclid=m5qhk3yyxl396343004> (accessed 12.12.2024) (In Russian).
4. Areed0615 Spectrophotometer-data-compiler. *github*: website. 2025. Available at: <https://github.com/Areed0615/Spectrophotometer-data-compiler> (accessed 12.12.2024) (In Russian).
5. Introduction to Machine Learning. *Uchebnik R* [Uchebnik R]: website, 2021. Available at: <https://nweb42.com/books/r-lang/vvedenie-v-shiny/> (accessed on 12.12.2024) (In Russian).
6. Tikhonov A. N. Visualization of Equipment Performance Statistics with R – Shiny. *Khabr: sayt* [Khabr]: website. 2015. Available at: <https://habr.com/ru/articles/253419/> (accessed 12.12.2024) (In Russian).
7. Garriga M. Package “Ropj”. *CRAN*: website, 2025. Available at: <https://cran.r-project.org/web/packages/Ropj/Ropj.pdf> (accessed 12.12.2024) (In Russian).
8. Bich N. N., Kynkurogov A. A., Skirukha A. A. Development of a software module for multiple conversion of OPJ files into CSV format. *Modernizatsiya vysshego obrazovaniya v storonu tsifrovizatsii problemy, resheniya, perspektivy: materialy II Resp. nauch.-prakt. konf. s mezhdunar. uchastiyem* [Modernization of higher education towards digitalization problems, solutions, prospects: proceedings of the II Republican Scientific and practical conference with international participation]. Grodno, 2025, pp. 11–12.

Информация об авторе

Бич Наталья Николаевна – кандидат физико-математических наук, доцент, доцент кафедры медицинской и биологической физики. Гродненский государственный медицинский университет (ул. М. Горького, 80, 230000, г. Гродно, Республика Беларусь). E-mail: nadaliya@mail.ru. SPIN-код: 5435-9008.

Information about the author

Bich Natalia Nikolaevna – PhD (Physics and Mathematics), Associate Professor, Assistant Professor, the Department of Medical and Biological Physics. Grodno State Medical University (80 M. Gorky str., 230000, Grodno, Republic of Belarus). E-mail: nadaliya@mail.ru. SPIN code: 5435-9008.

Поступила после доработки 14.12.2025