

# Литературное наследие XIX–XX веков: классификация растровых изображений для интеллектуального анализа и тематического моделирования корпуса рукописных текстов<sup>1</sup>

Рассмотрены современные тенденции работы с цифровыми формами рукописного наследия по истории русской литературы второй половины XIX — середины XX в. Проанализирован процесс формирования виртуальных архивов как постепенное накопление *big data* научных исследований — нераспознанного информационного массива растровых документов, содержащего десятки тысяч цифровых форм архивных документов. Предложены новые подходы к классификации растровых изображений рукописных документов для работы в системах интеллектуального анализа, экспериментальные методы визуализации архивного документа, а также способы использования незадействованных ранее возможностей поискового аппарата. Большое внимание уделено архитектонике рукописи: переходу от графических элементов растрового изображения к семантическим, позволяющим применить элементы интеллектуального анализа данных (*data mining*) для нераспознанного массива данных.

**Ключевые слова:** рукописное наследие, цифровая форма, растровое изображение, новые методы, архитектоника рукописи, большие данные, интеллектуальный анализ данных.

The article examines the current trends in working with digital forms of handwritten heritage on the history of Russian literature of the second half of the 19 — mid-20 century. The process of forming virtual archives is analyzed as a gradual accumulation of the “big data” of scientific research — an unrecognized information array of raster documents containing tens of thousands of digital forms of archival documents. New approaches to classifying raster images of handwritten documents for use in intelligent analysis systems, experimental methods of visualization of archival documents, as well as previously unused capabilities of the search engine are proposed. Much attention is paid to the architectonics of the manuscript: the transition from graphic elements of a raster image to semantic ones, which allows the use of data mining elements for an unrecognized data array.

**Keywords:** manuscript heritage, digital form, bitmap image, new methods, manuscript architectonics, big data, data mining.

В своей книге «Дальнее чтение» [1] социолог литературы Франко Моретти назвал художественное наследие Великим Непрочтенным. «Вопрос даже не в том, что стоит исследовать, вопрос в том, как “я занимаюсь западноевропейской прозой”... Это не совсем так, потому что я занимаюсь лишь канонизированной ее частью, которая составляет меньше одного процента всей изданной литературы»<sup>2</sup> [Там же. С. 79]. К сожалению, архивные данные только подтверждают эту закономерность: по аналогии с приведенной метафорой рукописное наследие до сих пор остается «великим нераспознанным». Значительное время информация о составе и содержании фондов была доступна только в самом архивохранилище, и даже там в ряде случаев оставалась полностью «закрытой» для исследователей<sup>3</sup>. Не меньшей проблемой стала разобщенность рукописных коллекций: архивы находятся в разных странах, в разных городах. Преодолев значительные расстояния, специалист часто был вынужден изучать в читальном зале не оригиналы, а почти нечитаемые микрофотокопии рукописей. Особую актуальность это приобретает по отношению

<sup>1</sup> Исследование выполнено при финансовой поддержке РФФ (проект № 22-68-00066).

<sup>2</sup> Сама метафора восходит к работе Маргарет Коэн «Сентиментальное воспитание романа» [2]. В следующей главе «Медленного чтения» Моретти приводит более точную статистику: 0,5% прочитанного противостоят 99,5% Великого Непрочтенного [1. С. 79].

<sup>3</sup> В первую очередь затруднен доступ к документам, находящимся в хранилищах обособленного доступа (спецхране). Также крайне затруднен доступ к фондам, находящимся в процессе научного описания, который может занимать не одно десятилетие. Так, ключевое для изучения творчества Н.С. Гумилева архивное собрание П.Н. и С.П. Лукницких [3], переданное в Отдел рукописей Пушкинского Дома в 1997 г., было открыто для исследователей только в 2017 г. Цифровые копии рукописей Гумилева из указанного собрания размещены на портале «Автограф. XX век» [4].

к памятникам русской культуры второй половины XIX — середины XX в., 92% которых не только не опубликованы, но и не расшифрованы.

Как отмечал один из ведущих специалистов по распознаванию образов Л.М. Местецкий<sup>4</sup>, технологии дешифровки рукописных текстов успешно продвигаются в коммерческих приложениях: банковские чеки и счета, почтовые адреса, анкеты. Такого рода тексты имеют стандартный формат, малые размеры, используют небольшой словарь терминов. Но существует и обратная закономерность: компании, разрабатывающие подобные системы, лишены стимулов для создания более сложной и дорогой технологии для работы с культурным наследием, не имеющей ясной коммерческой перспективы. Несмотря на стремительно растущий научный и социальный запрос, для разработки технологии, ориентированной на вклад в культуру, нет ни соответствующих способов реализации, ни штатных специалистов для работы в области data mining, ни бюджетных средств. Подавляющее большинство современных исследований рукописных документов предполагает извлечение информации из этих источников вручную в процессе расшифровки и подготовки научной публикации. Даже учитывая точечную результативность подобных публикаций, для корпуса архивных материалов объемом в десятки и сотни тысяч страниц приведенная организация работ неэффективна и закономерно приводит к тиражированию одного и того же набора текстов. Создание современной методики анализа информации, потенциально содержащейся в рукописных документах, остается крайне актуальной научно-исследовательской задачей.

Не менее важной предпосылкой для проведения корпусных исследований рукописных документов стало активное создание архивно-музейным кластером цифровых копий документов (страховое копирование) и реконструкция архивных комплексов историко-культурного наследия, что обеспечивает возможность удаленного электронного анализа данных<sup>5</sup>. В начале XXI в. появились два принципиально новых способа работы с рукописным наследием. Наравне с подлинниками в исследовании и архивном деле стали использоваться электронные копии документов (собственно, с этого времени рукописное наследие и стало «цифровым»<sup>6</sup>). Вторым способом работы для текстолога стал виртуальный архив, представляющий собой открытый каталог и фонд рукописей, в любое время дня и ночи доступный для исследователя в интернете. Становясь частью электронного архива, источник приобретает новое качество — мобильность. Независимо от научной квалификации исследователя и состояния оригинала рукописи любой пользователь интернета получает к нему прямой доступ: может его изучать, комментировать, цитировать в своих работах и ссылаться на сам источник, а не только на публикацию в собрании сочинений<sup>7</sup>.

Развитие технологий обусловило практически неограниченный рост цифровых архивов. Буквально за несколько лет объем данных стал измеряться уже не в килобайтах, а в терабайтах ин-

<sup>4</sup> В этой связи наиболее актуальны работы Л.М. Местецкого, посвященные машинному распознаванию графем (визуальных образов букв) в процессе скорописи [5; 6]. Работы, выполненные в настоящем проекте, находятся в процессе публикации.

<sup>5</sup> Бесконтактная цифровая реставрация рукописных документов — одно из наиболее перспективных направлений научно-исследовательских разработок в сфере архивоведения. В статье Л.И. Борожкина «Digital history: применение цифровых медиа в сохранении культурного наследия?» виртуальная реконструкция и визуализация объектов культурного наследия рассматриваются как самостоятельное направление цифровой истории [7. С. 18]. Методика бесконтактной реконструкции объектов архитектурного наследия представлена в его работе «Виртуальная реконструкция исторического городского ландшафта: проблемы междисциплинарного синтеза и их решение» [8. С. 130–131].

<sup>6</sup> Речь идет об Электронном фонде пользования, который представляет собой совокупность электронных копий документов Архивного фонда РФ, записанных на цифровые носители и предназначенных для использования вместо подлинников документов, что обеспечивает сохранность документов и возможность формирования электронных ресурсов, обеспечивающих оперативность доступа к документу, в том числе с использованием интернет-технологий [9. П. 2.1].

<sup>7</sup> В качестве оптимального источника возможно рассматривать предложенный Н.А. Богомоловым и В.Л. Гайдук «интегрированный комментированный текст», находящийся в непосредственной связи с опубликованной в интернете цифровой копией рукописного автографа [10. С. 335].

формации. С созданием data-центров публикация тысяч и десятков тысяч цифровых изображений не представляет никаких технических затруднений<sup>8</sup>. Тем не менее с постоянным накоплением цифры органически связаны и слабые стороны электронной публикации. При ориентировании на сложившиеся стандарты академического исследования<sup>9</sup> цифровой архив просто физически не может подготовить полноценный научно-справочный аппарат для стремительно растущего объема источников. Постепенно независимо от позиции его создателей виртуальный архив ставит вопрос об альтернативных стандартах публикации<sup>10</sup>. Авторское исследование информационного массива (вступительная статья, научная передача текста, создание реального, текстологического и контекстного комментария) функционально заменяется разработкой новой модели управления информацией, обеспечивающей экстерриториальный доступ к наследию русской литературы и его исследование на современном научном и технологическом уровне.

**Т**ехнологически цифровое собрание представляет собой нераспознанный корпус растровых изображений, состоящий из множества объектов — электронных форм архивных документов. В подобный корпус может входить от нескольких десятков тысяч (персональный архив) до сотен тысяч цифровых изображений (коллекция, музейное собрание). Понимаемый как единый массив, виртуальный архив идеально подходит под уже ставшее классическим определение «больших данных»: структурированные и неструктурированные данные большого объема и значительного разнообразия. В качестве определяющих характеристик массивов big data традиционно выделяют «три V»: объем (volume), скорость приращения (velocity) и многообразие (variety). Действительно, и накопленный объем цифровых изображений, и темпы роста виртуальных архивов полностью укладываются в предложенный формат. Третий и важнейший критерий — разнообразие и унификация частного — отсылает к «Структуре научных революций» Томаса Куна: числа, собранные при отсутствии каких-либо ожидаемых закономерностей, почти наверняка останутся просто числами. В поисках «ожидаемой закономерности», которая смогла бы превратить принципиально нерешаемые задачи в исследования, доступные парадигме «нормальной науки», оправданно обращение к стандартам описания архивных документов. Разработанный в 1994 г. Основной международный стандарт архивного описания (ISAD, Оттава) [19] за два последующих десятилетия получил несколько редакций, в последнюю из которых были включены цифровые копии. International Standard Archival Description предусматривает

<sup>8</sup> В 2010–2013 гг. в постоянно функционировавших цифровых архивах «Объединенный электронный архив И.А. Бунина» [11], «Виртуальный архив Анны Ахматовой» [12] и «Объединенный архив Вяч. Иванова» [13] ежегодный рост информационного массива составлял около 1000 электронных изображений. В 2014 г. начал работать Цифровой архив русской литературы «Автограф. XX век» [14]. В 2014–2021 гг. его экспонентный рост составил 4000 изображений. В 2022 г. ежегодное пополняемое количество изображений было увеличено до пяти тысяч. В конце 2022 г. было запланировано открытие портала «“Стенограмма”: политика и литература. Цифровой архив литературных организаций 1920–30-х годов» (ИМЛИ РАН) [15], электронный массив которого составляет более 15 тыс. изображений. Лучше всего этот тип научного комментария охарактеризовал А.В. Лавров в одной из своих недавних публикаций: «Архивную единицу можно было ввести в оборот без солидного идеологического сопровождения, просто охарактеризовав конкретную ситуацию, конкретное лицо, персонажа. Вот тогда у нас и появился новый тип комментария — развернутого, с большими дополнительными этажами материалов».

<sup>9</sup> Первым его освоил Гарик Суперфин: его стиль работы — к письму в четыре строки давать комментарий в четыре страницы. Образец такой работы — письма Ахматовой к Брюсову (совместно с Романом Тименчиком), где пишем несколько строчек, а работа на 20 страниц [16]. Или более академично: «В течение последнего столетия были осуществлены или продолжают осуществляться под эгидой Академии наук несколько десятков собраний сочинений русских классиков. За это время установлены и отработаны в ходе практической деятельности базовые текстологические принципы, которым призваны удовлетворять издания академического типа» [17. С. 7].

<sup>10</sup> Эта проблематика была подробно рассмотрена в выступлениях Н.В. Корниенко, Д.С. Московской и М.Л. Спивак на Международной конференции «Литературный цифровой архив как культурная практика и социальный опыт», проходившей 20 октября 2022 г. в ИМЛИ РАН. При этом в качестве двух взаимосвязанных задач цифрового архива рассматривались «систематизация больших архивных материалов» и «обеспечение условий к дальнейшему введению этого материала в научно-исследовательский оборот» [18].

выделение общих признаков по шести основным (страна, хранилище, шифр<sup>11</sup>, подлинность, авторство, датировка) и 26 дополнительным параметрам, варьируемым в каждой стране или даже административном округе. За исключением фондирования, включенные в международный стандарт информационные поля соответствуют архитектонике рукописного документа. В беловике и черновике, письме и дневнике мы можем выделить информационное поле даты, заголовка, подписи, рисунка и корпуса документа, т.е. собственно текста. Эти поля, универсальные для любой рукописи, могут стать стабильными ориентирами, на которых базируется управление любым (и здесь уже не столь важно, распознанным или растровым) информационным массивом.

«Новая оптика» рукописного текста позволит применить к нему интеллектуальный анализ данных, опирающийся на предложенные Франко Моретти количественные (численные) признаки. Наиболее последовательно выраженный количественный метод корпусных исследований предполагает анализ количественных показателей для уяснения качественной модели. Предложенный метод, названный Игорем Пильщиковым «новым количественным формализмом»<sup>12</sup>, вводит в массив больших данных «функцию измерений», позволяющую на основе скрупулезного подсчета частотности прийти к выводам о жанровой принадлежности произведения или определить степень близости жанров в синхронном срезе [20. С. 138–140]. В отличие от наиболее популярного подхода — распознавания всех символов текста, не реализуемого в настоящее время программным обеспечением любого класса, — новая методика сосредоточена на количественном анализе архитектурных элементов рукописного текста и их вариативных сочетаний. Графический анализ таких элементов системы, как ключевые слова (сходство изображений-слов как растровых паттернов), в комбинации с дополнительными элементами (выделение области даты и места создания, форма, размеры и расположение обращения, подписи, заголовка, расположение пунктуационных знаков и даже строфика) позволит перейти от количественных элементов к качественным и исходя из их сочетания сделать выводы о составе и содержании рукописного источника. При использовании системы автоматизированной навигации поиск нужной информации в растровом массиве нераспознанного текста будет занимать не месяцы и годы гуманитарных исследований, а несколько минут.

Используя навигацию по растровому массиву, можно задействовать дифференцированные программные модели для решения самых разнообразных задач: тематической сегментации текстов, классификации и категоризации документов, многоязычного информационного поиска, тематической визуализации больших текстовых коллекций и др. Существуют модели,

<sup>11</sup> Именно здесь заложен основополагающий принцип фондирования — от общего (фонда) к частному (единице хранения и листу).

<sup>12</sup> Обращения к литературе как к гипертексту И.А. Пильщиков связывает не только с работой Франко Моретти «Дальнее чтение» [1], но и с работами представителей русской формальной школы и близких к ней ученых (Б.В. Томашевского, Б.И. Ярхо), указанными им в качестве источника исследований Моретти. В статье «Русский количественный формализм 1910–30-х гг. как предшественник Digital Humanities» Пильщиков отмечал «удивительный факт: русские первопроходцы количественной поэтики и статистического стиховедения — Андрей Белый, Борис Исаакович Ярхо и Борис Викторович Томашевский — начали и долгое время вели свою работу независимо друг от друга» [20. С. 130]. Показательно, что для характеристики работ Моретти Т.Д. Венедиктова заимствует термины из «Структуры научных революций» Томаса Куна: «Здесь перед нами — “нормальный” (в куновском смысле) литературовед-зарубежник, автор диссертации и монографии о творчестве Т.С. Элиота, уже готовый, впрочем, к рискованной творческой эскапade» [21. С. 82]. О применении distant reading в библиотечном деле см. в работе В.П. Леонова: «Ф. Моретти, начав с поисков точного метода, приходит не к большому данным, с чем литературоведу и библиографу еще нужно научиться иметь дело, а к методологическому повороту в самом литературоведении и, полагаю, библиографоведении. На этом повороте мы получаем объекты и данные нового типа, надстроенные над языком пристального чтения, которые нужно не понимать, но на которые нужно смотреть. <...> Не исключаю также, что взаимодействие пристального, дальнего и масштабированного чтения в рамках цифровой и автоматизированной обработки текстов будет способствовать формированию нового исследовательского объекта — точного библиографоведения» [22. С. 65].

выявляющие ключевые фразы, учитывающие морфологию слов и синтаксическую структуру предложений, а также различные характеристики (модальности) документов: авторство, ссылки и др., отслеживающие изменения тем во времени, строящие иерархические отношения между темами. В настоящее время сформулированы две наиболее актуальные задачи, решаемые современными программными средствами. Реализация первой позволит, используя машинное видение, устанавливать датировку растрового массива эпистолярных документов без тотальной расшифровки текста. Предварительно определяется поле даты, и только после этого установленная «дата» по каждой позиции проверяется на совпадение с числовыми паттернами. Вторая задача — автоматизированное моделирование и последующий поиск ключевых слов при работе с растровым массивом машинописного текста. Решение этих задач станет первым шагом в управлении большими данными растровых изображений культурного наследия и приведет исследователя к качественным результатам «нового квантитативного формализма» рукописного текста.

### Литература

1. *Moretti F.* Дальнее чтение / пер. с англ. А. Вдовина, О. Собчука, А. Шели; науч. ред. пер. И. Кушнарева. М.: Изд-во Института Гайдара, 2016. 352 с.
2. *Cohen M.* The sentimental education of the novel. Princeton: Princeton University Press, 1999. 219 p.
3. Собрание П.Н. и С.П. Лукницких // РО ИРЛИ РАН. Ф. 754. Оп. 1.
4. Собрание П.Н. и С.П. Лукницких // Автограф. XX век: портал. URL: <http://gumilev.literature-archive.ru/ru/digital-archive/stihotvoreniya-i-poemy> (01.12.2022).
5. *Липкина А.Л., Местецкий Л.М.* Классификация букв в изображениях на основе медианного представления // Geometric Modeling. Computer Graphics in Education. 2018. URL: <https://www.graphicon.ru/html/2018/papers/362-368.pdf> (01.12.2022).
6. *Lipkina A., Mestetskiy L.* Grapheme approach to recognizing letters based on medial representation // Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. 2019. Vol. 4. No. 1. P. 351–358.
7. *Бородкин Л.И.* Digital history: применение цифровых медиа в сохранении культурного наследия? // Методологические проблемы исторической информатики: информационный бюллетень. 2012. Т. 1. № 1. С. 14–21.
8. *Бородкин Л.И.* Виртуальная реконструкция исторического городского ландшафта: проблемы междисциплинарного синтеза и их решение // Историко-культурное наследие в цифровом измерении: материалы Междунар. науч. конф., г. Пермь, 20–22 октября 2021 г. Пермь, 2021. 210 с. URL: <http://www.psu.ru/files/docs/science/books/sborniki/istoriko-kulturnoe-nasledie-v-cifrovom-izmerenii.pdf> (01.12.2022).
9. *Юмашева Ю.Ю.* Методические рекомендации по электронному копированию архивных документов и управлению полученным информационным массивом. М.: ВНИИДАД, 2012. 125 с. URL: [https://archives.gov.ru/documents/rekomend\\_el-copy-archival-documents.shtml](https://archives.gov.ru/documents/rekomend_el-copy-archival-documents.shtml) (01.12.2022).
10. *Богомолов Н.А., Гайдук В.Л.* Валерий Брюсов: дневник 1890 г. / предисл. Н.А. Богомолова, подгот. текста и примеч. В.Л. Гаидук и Н.А. Богомолова // Studia Literarum. 2020. Т. 5. № 3. С. 328–357.
11. Объединенный электронный архив Ивана Бунина. URL: <http://www.bunin-rgali.ru> (01.12.2022).
12. Виртуальный архив Анны Ахматовой. URL: <http://www.akhmatova-rgali.ru> (01.12.2022).
13. Объединенный архив Вяч. Иванова. URL: <http://www.ivanov-rgali.ru/> (01.12.2022).

### References

1. *Moretti F.* Dal'nee chtenie / per. s angl. A. Vdovina, O. Sobchuka, A. Sheli; nach. red. per. I. Kushnareva. Moscow: Izd-vo Instituta Gaidara, 2016. 352 s.
2. *Cohen M.* The sentimental education of the novel. Princeton: Princeton University Press, 1999. 219 p.
3. Sobranie P.N. i S.P. Luknitskikh // RO IRLI RAN. Ф. 754. Оп. 1.
4. Sobranie P.N. i S.P. Luknitskikh // Avtograf. XX vek: portal. URL: <http://gumilev.literature-archive.ru/ru/digital-archive/stihotvoreniya-i-poemy> (01.12.2022).
5. *Lipkina A.L., Mestetskiy L.M.* Klassifikatsiia bukv v izobrazheniiah na osnove mediannogo predstavleniia // Geometric Modeling. Computer Graphics in Education. 2018. URL: <https://www.graphicon.ru/html/2018/papers/362-368.pdf> (01.12.2022).
6. *Lipkina A., Mestetskiy L.* Grapheme approach to recognizing letters based on medial representation // Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. 2019. Vol. 4. No. 1. P. 351–358.
7. *Borodkin L.I.* Digital history: Primenenie tsifrovyykh media v sokhraneni kul'turnogo nasledii? // Metodologicheskie problemy istoricheskoi informatiki: informatsionnyi biulleten'. 2012. T. 1. No. 1. S. 14–21.
8. *Borodkin L.I.* Virtual'naia rekonstruktsiia istoricheskogo gorodskogo landshafta: problemy mezhdistsiplinarnogo sinteza i ikh reshenie // Istoriko-kul'turnoe nasledie v tsifrovom izmerenii: materialy Mezhdunar. nach. konf., g. Perm, 20–22 oktiabria 2021 g. Perm, 2021. 210 s. URL: <http://www.psu.ru/files/docs/science/books/sborniki/istoriko-kulturnoe-nasledie-v-cifrovom-izmerenii.pdf> (01.12.2022).
9. *Yumasheva Yu.Yu.* Metodicheskie rekomendatsii po elektronnomu ko-pirovaniu arkhivnykh dokumentov i upravleniu poluchennym informatsionnym massivom. Moscow: VNIIDAD, 2012. 125 s. URL: [https://archives.gov.ru/documents/rekomend\\_el-copy-archival-documents.shtml](https://archives.gov.ru/documents/rekomend_el-copy-archival-documents.shtml) (01.12.2022).
10. *Bogomolov N.A., Gaiduk V.L.* Valerii Briusov: dnevnik 1890 g. / predisl. N.A. Bogomolova, podgot. teksta i primеч. V.L. Gaiduk i N.A. Bogomolova // Studia Literarum. 2020. T. 5. No. 3. S. 328–357.
11. Ob"edinennyi elektronnyi arkhiv Ivana Bunina. URL: <http://www.bunin-rgali.ru> (01.12.2022).
12. Virtual'nyi arkhiv Anny Akhmatovoi. URL: <http://www.akhmatova-rgali.ru> (01.12.2022).
13. Ob"edinennyi arkhiv Viach. Ivanova. URL: <http://www.ivanov-rgali.ru/> (01.12.2022).

14. Автограф. XX век: электронный архив русской литературы. URL: <http://literature-archive.ru> (01.12.2022).

15. Объединенный цифровой архив рукописей Ф.М. Достоевского. URL: <https://dostoevskyarchive.pushdom.ru/about> (01.12.2022).

16. Лавров А. «У нас все — целина: куда ни копни, все впервые» // *Arzamas*. 2022. 3 августа. URL: <https://arzamas.academy/mag/1108-lavrov> (01.12.2022).

17. Лавров А.В. Тексты и комментарии: из материалов к истории русской литературы первой трети XX века. СПб.: Пушкинский Дом, 2018. 528 с.

18. Гусейнов В.Н. Литературный архив как культурная практика и социальный опыт: материалы Междунар. науч.-практ. конф. // *Вестник Московского университета*. Сер. 9: Филология. 2023 (в печати).

19. ISAD (G): основной международный стандарт архивного описания / пер. с англ.; гл. ред. пер. Е.Д. Жабко. 2-е изд. СПб.: Президентская библиотека имени Б.Н. Ельцина, 2011. 247 с.

20. Пильщиков И.А. Семь бесед о филологии и Digital Humanities: интервью и дискуссии (2015–2021). М.: МГУ, 2022. 192 с.

21. Венедиктова Т.Д. Хитроумный путешественник // *Новое литературное обозрение*. 2018. № 2. С. 82–88.

22. Леонов В.П. Дальнее чтение как стратегия точного библиографоведения // *Научные и технические библиотеки*. 2019. № 10. С. 56–67.

14. Avtograf. XX vek: elektronnyi arkhiv russkoi literatury. URL: <http://literature-archive.ru> (01.12.2022).

15. Ob"edinennyy tsifrovoy arkhiv rukopisei F.M. Dostoevskogo. URL: <https://dostoevskyarchive.pushdom.ru/about> (01.12.2022).

16. Lavrov A. "U nas vse — tselina: kuda ni kopni, vse v pervye" // *Arzamas*. 2022. 3 avgusta. URL: <https://arzamas.academy/mag/1108-lavrov> (01.12.2022).

17. Lavrov A.V. Teksty i kommentarii: iz materialov k istorii russkoi literatury pervoi treti XX veka. St. Petersburg: Pushkinskii Dom, 2018. 528 s.

18. Guseinov V.N. Literaturnyi arkhiv kak kul'turnaia praktika i sotsial'nyi opyt»: materialy Mezhdunar. nauch.-prakt. konf. // *Vestnik Moskovskogo universiteta*. Ser. 9: Filologiya. 2023 (v pechati).

19. ISAD (G): osnovnoi mezhdunarodnyi standart arkhivnogo opisaniia / per. s angl.; gl. red. per. E.D. Zhabko. 2-e izd. St. Petersburg: Prezidentskaia biblioteka imeni B.N. El'tsina, 2011. 247 s.

20. Pilshchikov I.A. Sem' besed o filologii i Digital Humanities: interv'iu i diskussii (2015–2021). Moscow: MGU, 2022. 192 s.

21. Venediktova T.D. Khitroumnyi puteshestvennik // *Novoe literaturnoe obozrenie*. 2018. No. 2. S. 82–88.

22. Leonov V.P. Dal'nee chtenie kak strategiiia tochnogo bibliografovedeniia // *Nauchnye i tekhnicheskie biblioteki*. 2019. No. 10. S. 56–67.



**Пенская Елена Наумовна,**

доктор филологических наук, профессор  
Национальный исследовательский университет  
«Высшая школа экономики»;  
руководитель группы Центра междисциплинарных исследований  
Московский физико-технический институт

**Penskaya Elena N.,**

Doctor of Philology, Professor  
National Research University "Higher School of Economics";  
Head of the Group of the Center for Interdisciplinary Research  
Moscow Institute of Physics and Technology

ORCID: 0000-0003-2469-584X

e-mail: e.penskaya@gmail.com



**Хачатурян Любовь Валерьевна,**

кандидат культурологии, доцент  
Национальный исследовательский университет  
«Высшая школа экономики»;  
старший научный сотрудник  
Центра междисциплинарных исследований  
Московский физико-технический институт

**Khachatryan Lyubov V.,**

Candidate of Culturology, Associate Professor  
National Research University "Higher School of Economics";  
Senior Scientific Researcher of the Center for Interdisciplinary Research  
Moscow Institute of Physics and Technology

ORCID: 0000-0002-2689-5186

e-mail: rgali2010@yandex.ru

