

АВТОМАТИЗАЦИЯ СУДЕБНОЙ АВТОРОВЕДЧЕСКОЙ ЭКСПЕРТИЗЫ: ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ

Т. В. Романова

*Национальный исследовательский университет
«Высшая школа экономики – Нижний Новгород»*

А. Ю. Хоменко

*Национальный исследовательский университет
«Высшая школа экономики – Нижний Новгород»,
Центр экспертиз и исследований «ЕСИИ»*

FORENSIC TEXT ATTRIBUTION COMPUTERISATION: PROBLEMS AND PROSPECTS

T. V. Romanova

HSE Campus in Nizhny Novgorod

A. Yu. Khomenko

*HSE Campus in Nizhny Novgorod,
Center for Expertise and Research ESIN*

Аннотация: рассматривается процесс автоматизации судебной атрибуционной (автороведческой) экспертизы текстового материала. Одной из задач исследования является изучение тенденций, главенствующих в атрибуционной экспертизе, отечественной и иноязычной (преимущественно англоязычной), с целью анализа проблем и сложностей этого процесса, связанных с законодательной сферой. Посредством анализа норм существующего отечественного законодательства, а также некоторых прецедентов североамериканского и западноевропейского права, с одной стороны, и методико-технического состояния атрибуционной науки разных научных школ – с другой, делается вывод о перспективах судебной автороведческой экспертизы в России. Постулируется мысль о том, что на современном этапе развития науки и техники единственным путем развития для автороведческой экспертизы в нашей стране является путь интеграции квалификативных и количественных методов анализа.

Ключевые слова: судебная криминалистическая автороведческая экспертиза, атрибуционная экспертиза письменного текста, судебная экспертная деятельность.

Abstract: this article states the problem of computerisation in the field of forensic authorship attribution. One of the objectives of the study is to reveal the trends in authorship attribution expertise in domestic and foreign (mainly English-speaking science), in order to analyse the problems and difficulties of computerisation process related to the legislative sphere. By analysing the norms of existing domestic legislation, as well as some precedents of English-language law, and the methodological and technical state of attribution science in Russia and abroad, the authors reveal the prospects for forensic text attribution in the Russian Federation. The idea that at the present stage in science and technology the only way for forensic authorship attribution development in Russia is to integrate qualitative and quantitative analysis methods is established.

Key words: forensic authorship attribution examination, authorship attribution examination of written text, forensic expert activity.

Проблема компьютеризации социальных процессов на современном этапе развития науки обсуждается специалистами всех областей. Не чужда она и судебной экспертной деятельности. Судебная экспертиза может классифицироваться как криминалистическая и некриминалистическая, иными словами, иметь статус так называемой традиционной криминалистической экспертизы и нетрадиционной криминалистической экспертизы. Использование машинных средств анализа давно стало обыденным для большинства криминалистических исследований (трасологического, генетического и пр.), но еще не все экспертизы этого вида в России имеют компьютерное сопровождение. Так, судебная автороведческая экспертиза (наименование по классификации Минюста РФ¹) есть один из видов экспертных исследований, относящихся к исследованиям криминалистическим (некоторые определяют этот род экспертной практики как традиционную криминалистическую экспертизу², другие – как нетрадиционную³). Целью автороведческих экспертиз, проводимых в государственных экспертных учреждениях и вне таковых, является атрибуция материала на предмет определения его принадлежности конкретному автору, группе авторов или получения информации о личностях авторов, тем не менее этот вид экспертиз на данный момент имеет достаточно низкий уровень автоматизации. Вероятно, это связано в том числе с тем, что следственные органы на этапе предварительного следствия и суды на этапе судебного следствия нередко отказывают в проведении такого вида экспертиз⁴.

¹ Об утверждении Перечня родов (видов) судебных экспертиз, выполняемых в федеральных бюджетных судебно-экспертных учреждениях Минюста России, и Перечня экспертных специальностей, по которым представляется право самостоятельного производства судебных экспертиз в федеральных бюджетных судебно-экспертных учреждениях Минюста России : приказ от 20 апреля 2023 г. № 72. URL: <http://publication.pravo.gov.ru/Document/View/0001202304250001> (дата обращения: 16.07.2023).

² См.: *Моисеева Т. Ф.* Криминалистические экспертизы : изучение в рамках криминалистики и судебной экспертизы // Уголовное судопроизводство : проблемы теории и практики. 2016. № 1. С. 50–51.

³ См.: *Филиппов А. Г.* О понятии и классификации криминалистических экспертиз // Актуальные проблемы криминалистического обеспечения расследования преступлений. Труды Академии. М., 1996.

⁴ Решение суда по интеллектуальным правам от 4 декабря 2020 г. по делу № СИП-676/2019 ; Решение суда по

В рамках *текстовой* атрибуции для сугубо научных целей в современной лингвистике во всем мире активно используются автоматические методы анализа. Эти методы получают реализацию в виде программных продуктов за рубежом⁵ и в России:

– ПО (программное обеспечение) М. А. Марусенко: в основе – теория распознавания образов⁶;

– ПО «Атрибуция» В. Н. Захарова: в основе – грамматико-синтаксический анализ⁷;

интеллектуальным правам от 29 ноября 2019 г. по делу № СИП-695/2019 ; Апелляционное определение от 26 декабря 2018 г. № 203-АПУ18-25 ; и пр.

В определениях Верховного Суда РФ также нередко встречаются упоминания о наличии в апелляционных жалобах информации об отказе в удовлетворении ходатайств о назначении автороведческих экспертиз судами общей юрисдикции: Апелляционное определение от 26 декабря 2018 г. № 203-АПУ18-25 («Суд необоснованно отказал в удовлетворении ходатайства о проведении почерковедческой и автороведческой экспертиз»); Апелляционное определение от 24 сентября 2014 г. № 41-АПУ14-32 («Считает необоснованным отказ суда в назначении автороведческой экспертизы по показаниям потерпевших М., Т., С., Л., Б., М., В., Ц., П., поскольку показания данных лиц являются по своему содержанию идентичными и вызывают сомнение в их достоверности»); Кассационное определение от 13 декабря 2011 г. № 48-011-120 («Необоснованно отклонены ходатайства защиты о назначении повторной судебно-психиатрической экспертизы Соколова, комплексной судебно-технической, автороведческой, комплексной повторной судебно-медицинской, ситуационно-баллистической, физико-химической экспертиз»); Кассационное определение от 19 октября 2006 г. № 5-006-135сп и пр. («Судом необоснованно отказано в удовлетворении ходатайства защиты о проведении автороведческой экспертизы по показаниям Дерцяна»).

⁵ См., например: *Coulthard M.* Author identification, idiolect, and linguistic uniqueness // *Applied Linguistics*. 2004. № 24(4). P. 431–447 ; *Stamatatos E.* Authorship attribution using text distortion // *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, 2017. Vol. 1. P. 1138–1149 ; *Wright D.* Using word n-grams to identify authors and idiolects : A corpus approach to a forensic linguistic problem // *International Journal of Corpus Linguistics*. 2017. № 22(2). P. 212–241. URL: <https://benjamins.com/#catalog/journals/ijcl.22.2.03wri/details> (дата обращения: 05.08.2019) и пр.

⁶ См.: *Марусенко М. А.* Атрибуция анонимных и псевдонимных литературных произведений методами распознавания образов. Л., 1990. 164 с.

⁷ См.: Программа систем поддержки атрибуции текстов статей Ф. М. Достоевского / В. Н. Захаров, А. А. Леонтьев, А. А. Рогов, Ю. В. Сидоров // Труды ПетрГУ. Петрозаводск, 2000. Вып. 9. Серия : Прикладная математика и информатика. С. 113–122.

- ПО «Атрибутор» А. Н. Тимашева: в основе – буквенные триады⁸;
- ПО «Авторовед» А. С. Романова: в основе – метод опорных векторов, наиболее частые в языке триграммы и слова⁹;
- ПО «Лингвоанализатор» Д. В. Хмелева: в основе – цепи А. А. Маркова¹⁰;
- ПО «Стилеанализатор» О. Г. Шевелёва: в основе – цепи А. А. Маркова, деревья решений, нейронные сети¹¹, а также более современные их аналоги:
- ПО «Lingster 3.0» Института криминалистики ФСБ России¹²;
- ПО «КАТ» Нижегородского государственного университета имени Н. И. Лобачевского¹³;
- ПО «TextAnalyst 2.0» Московского исследовательского центра¹⁴;
- база данных «Rusldiolect» лаборатории корпусной идиолектологии ВГПУ¹⁵;
- ПО «СМАЛТ» Петрозаводского государственного университета¹⁶ и пр.

⁸ См.: Тимашев А. Н. Атрибутор : Версия 1.01: Описание программы. М., 1999–2007. URL: http://www.textology.ru/atr_resum.html (дата обращения: 01.02.16).

⁹ См.: Authorship Identification of a Russian-Language Text Using Support Vector Machine and Deep Neural Networks / А. Romanov, А. Kurtukova, А. Fedotova, А. Shelupanov, V. Goncharov // Future Internet. 2021. Vol. 13, iss. 1–3. 16 p.

¹⁰ См.: Хмелев Д. В. Лингвоанализатор. М., 2002. URL: <http://www.rusf.ru/books/analysis/> (дата обращения: 16.11.2017).

¹¹ См.: Шевелёв О. Г. Методы автоматической классификации текстов на естественном языке : учеб. пособие. Томск, 2007. 144 с.

¹² См.: Комплексная методика производства автороведческих экспертиз : методические рекомендации / Рубцова И. И., Е. И. Ермолаева, А. И. Безрукова [и др.]. М, 2007. С. 104–107.

¹³ См.: Радбиль Т. Б., Маркина М. В. Вероятностно-статистические модели в производстве автороведческой экспертизы русскоязычных текстов // Политическая лингвистика. 2019. № 2 (74). С. 156–166.

¹⁴ См.: Ионова С. В., Огорелков И. В. Речевая диагностика личности по гендерному признаку в автороведении: квантитативный подход // Вестник Волгоград. гос. ун-та. Серия 2 : Языкознание. 2020. Т. 19, № 1. С. 115–127. 10.15688/jvolsu2.2020.1.10.

¹⁵ См.: Литвинова Т. А., Громова А. В. Компьютерные технологии в судебной автороведческой экспертизе : проблемы и перспективы использования // Вестник Волгоград. гос. ун-та. Серия 2 : Языкознание. 2020. Т. 19, № 1. С. 77–88. DOI: 10.15688/jvolsu2.2020.1.7

¹⁶ См.: Рогов А. А., Кулаков К. А., Москин Н. Д. Программная поддержка в решении задачи атрибуции текстов // Программная инженерия. 2019. № 5. С. 234–240.

Принципы судебного автороведения в силу специфики сферы деятельности несколько отличаются от атрибуции текстов в рамках решения собственно научных задач. Обосновано это в первую очередь законодательством Российской Федерации: Федеральный закон от 31 мая 2001 г. № 73-ФЗ «О государственной судебно-экспертной деятельности в Российской Федерации» (далее – Федеральный закон № 73-ФЗ) и все кодексы РФ, закрепляющие процессуальные нормы (Уголовно-процессуальный кодекс (далее – УПК), Арбитражный процессуальный кодекс (далее – АПК), Гражданский процессуальный кодекс (далее – ГПК), Кодекс об административных правонарушениях (далее – КоАП), Кодекс административного судопроизводства (далее – КАС)), устанавливают личную ответственность экспертов, связанную с даваемым заключением. «Заключение эксперта – письменный документ, отражающий ход и результаты исследований, проведенных экспертом [курсив наш. – Авт.]»¹⁷. Ответственность эта не может быть переложена на машину, эксперт должен критически анализировать результаты, выдаваемые программным обеспечением (далее – ПО), при условии его наличия, и давать «обоснованное и объективное заключение»¹⁸ «в пределах соответствующей специальности, всесторонне и в полном объеме»¹⁹. При условии неисполнения требований законодательства эксперт несет уголовную ответственность за свое решение по ст. 307 Уголовного кодекса РФ перед гражданским обществом и государством²⁰.

Закрепленность личной ответственности эксперта является одной из причин для неиспользования полностью автоматических техник атрибуционного анализа в судебной деятельности в России. Но эта преграда не единственная. Для отечественного законодательства, например УПК, ГПК, АПК, КоАП, равно как и для Федерального закона № 73-ФЗ (ст. 8), характерен постулат о том, что эксперт при ответе на вопросы должен находиться строго в пределах компетенции, накладываемой объемом его специальных знаний:

– УПК: «Эксперт вправе <...> 4) давать заключение в пределах своей компетенции [курсив

¹⁷ Рос. газета. 2001. № 256. С. 9.

¹⁸ Там же. С. 8.

¹⁹ Там же.

²⁰ Уголовный кодекс Российской Федерации от 13 июня 1996 г. № 63-ФЗ. URL: http://www.consultant.ru/document/cons_doc_LAW_10699/ (дата обращения: 03.05.2020).

наш. – Авт.], в том числе по вопросам, хотя и не поставленным в постановлении о назначении судебной экспертизы, но имеющим отношение к предмету экспертного исследования»²¹;

– ГПК: «В случае, если поставленные вопросы выходят за пределы специальных знаний эксперта [курсив наш. – Авт.] <...> эксперт обязан направить в суд, назначивший экспертизу, мотивированное сообщение в письменной форме о невозможности дать заключение»²²;

– АПК: «Эксперт вправе отказаться от дачи заключения по вопросам, выходящим за пределы его специальных знаний [курсив наш. – Авт.], а также в случае, если представленные ему материалы недостаточны для дачи заключения»²³;

– КоАП: «Эксперт имеет право отказаться от дачи заключения, если поставленные вопросы выходят за пределы его специальных познаний [курсив наш. – Авт.] или если предоставленных ему материалов недостаточно для дачи заключения»²⁴;

– КАС: «Если поставленные перед экспертом вопросы выходят за пределы его специальных знаний, <...> либо если современный уровень развития науки не позволяет ответить на поставленные вопросы, эксперт обязан представить в суд мотивированное сообщение в письменной форме о невозможности дать заключение [курсив наш. – Авт.]»²⁵.

«Профессиональная компетенция эксперта (от лат. *compro* – добываюсь, соответствую, подхожу) складывается из комплекса знаний в области теории, методологии и практики экспертизы

определенного рода, вида»²⁶. Специалисты, проводящие судебные автороведческие экспертизы, как правило, имеют базовое лингвистическое или филологическое образование и переподготовку по специальности «Исследование продуктов речевой деятельности» и/или «Исследование письменной речи» (по классификации Минюста РФ)²⁷. Данный базис не предполагает наличия специальных знаний в области больших данных (big data), теории вероятностей машинного обучения и нейронных сетей, математической статистики, теории распознавания образов, теории векторного пространства и пр., дисциплин, необходимых для освоения и понимания программных средств, использующих наиболее удачные алгоритмы автоматической идентификации автора письменного текста. Отсюда следует, что законодательство Российской Федерации в области судебной экспертизы на базовом уровне (посредством постулатов процессуальных кодексов, федеральных законов, ведомственных инструкций и приказов) ограничивает использование сугубо компьютерных средств при производстве автороведческих экспертиз, т. е. эксперт не вправе использовать только ПО для формирования вывода по исследованию, не применяя интерпретативный (ручной) анализ материала в соответствии с рекомендованными методиками. Более того, эксперт, естественно, не вправе использовать ПО, принцип работы которого не понимает ввиду отсутствия специальных знаний в области статистики, математики, теории вероятности и пр.

Кроме законодательства, применение машинных средств анализа текста с целью установления его авторства ограничено и отечественной научной традицией, связанной с широким распространением интерпретирующей исследовательской парадигмы в языкознании вообще и в судебной лингвистике в частности. Так, базисом судебных атрибуционных методик является

²¹ Уголовно-процессуальный кодекс Российской Федерации от 18 декабря 2001 г. № 174-ФЗ. URL: http://www.consultant.ru/document/cons_doc_LAW_34481/ (дата обращения: 03.05.2020).

²² Гражданский процессуальный кодекс Российской Федерации от 14 ноября 2002 г. № 138-ФЗ. URL: http://www.consultant.ru/document/cons_doc_LAW_39570/ (дата обращения: 03.05.2020).

²³ Арбитражный процессуальный кодекс Российской Федерации от 24 июля 2002 г. № 95-ФЗ. URL: http://www.consultant.ru/document/cons_doc_LAW_37800/ (дата обращения: 03.05.2020).

²⁴ Кодекс Российской Федерации об административных правонарушениях от 30 декабря 2001 г. № 195-ФЗ. URL: http://www.consultant.ru/document/cons_doc_LAW_34661/ (дата обращения: 03.05.2020).

²⁵ Кодекс административного судопроизводства Российской Федерации от 8 марта 2015 г. № 21-ФЗ (ред. от 30.12.2021). URL: http://www.consultant.ru/document/cons_doc_LAW_176147/ (дата обращения: 03.05.2022).

²⁶ Энциклопедия судебной экспертизы / под ред. Т. В. Аверьяновой, Е. Р. Россинской. М., 1999. С. 177.

²⁷ Об утверждении Перечня родов (видов) судебных экспертиз, выполняемых в федеральных бюджетных судебно-экспертных учреждениях Минюста России, и Перечня экспертных специальностей, по которым представляется право самостоятельного производства судебных экспертиз в федеральных бюджетных судебно-экспертных учреждениях Минюста России : приказ от 20 апреля 2023 г. № 72. URL: <http://publication.pravo.gov.ru/Document/View/0001202304250001> (дата обращения: 16.07.2023).

учение С. М. Вула²⁸, которые дополняются разработками А. Ю. Комиссарова²⁹, Е. И. Горошко³⁰, Е. И. Галяшиной³¹. Одной из релевантных ведомственных методик на данный момент является «Комплексная методика производства автороведческих экспертиз» И. И. Рубцовой, Е. И. Ермолаевой, А. И. Безруковой и др.³²

Практика автоматизации в текстовой атрибуции сейчас перенимается в России из западноевропейской и североамериканской школ, где автороведение традиционно связано с количественными, стилометрическими методами анализа (от Л. Кэмпбела³³ в 1867 г. до современности, например, М. Коппела, Дж. Шлера³⁴, Д. Райта³⁵ и пр.). Несмотря на это, в указанных школах имеет место сходная с отечественной традиция использования собственно лингвистических, качественных методов и методик как инструментария определения авторства³⁶, а судебные атрибуционные методики основаны на теории идиолекта³⁷.

Суды Америки в некоторых штатах и Соединенного королевства (только после разрешения

суда по конкретному делу) принимают атрибуционные исследования собственно количественного содержания³⁸, проведенные с помощью ПО. Можно привести несколько примеров таких дел: Апелляционный суд Лондона (1991) – дело Томаса Мак Кроссена (Court of Appeal, London 1991: *The Queen vs. Thomas McCrossen*); Лестерский Королевский суд (1992) – дело Фрэнка Бэка (Leicester Crown Court 1992: *The Queen vs. Frank Beck*). Но использование полностью автоматизированных исследований в судебной атрибуции Запада, скорее, исключение, чем правило. В России, как было указано, они не используются вовсе. Вообще, назначения автороведческой экспертизы по текстовому материалу в современной отечественной практике редки. Автороведческие экспертизы часто назначаются по музыкальным, художественным произведениям, произведениям искусства³⁹ и значительно реже собственно по текстовому материалу⁴⁰. При рассмотрении уголовных дел частота назначения автороведческой экспертизы по текстовому материалу возрастает⁴¹, тем не менее, на наш взгляд, она недостаточна.

²⁸ См.: Вул С. М. Судебно-автороведческая идентификационная экспертиза : методические основы : методическое пособие. Харьков, 2007. 64 с.

²⁹ См.: Комиссаров А. Ю. Криминалистическое исследование письменной речи : учеб. пособие. М., 2000. 126 с.

³⁰ См.: Горошко Е. И. Судебно-автороведческая классификационная экспертиза : проблемы установления пола автора документа // Теория и практика судебной экспертизы и криминалистики. Харьков, 2003. Вып. 3. С. 221–226.

³¹ См.: Галяшина Е. И., Ермолова Е. И. Лингвокриминалистические возможности авторизации письменных и устных текстов : Междунар. науч. конф. М., 2005. С. 20–22.

³² См.: Комплексная методика производства автороведческих экспертиз: Методические рекомендации / Рубцова И. И., Ермолаева Е. И., Безрукова А. И. [и др.]. М., 2007. 192 с.

³³ См.: Campbell L. *The Sophistries and Polilicus of Plato*. Oxford, 1867. 170 p.

³⁴ См.: Koppel M., Schler J. Exploiting Stylistic Idiosyncrasies for Authorship Attribution // Proceedings of IJCAI'03 Workshop on Computational Approaches to Style Analysis and Synthesis. 2003. Vol. 69. P. 72–80.

³⁵ См.: Wright D. Using word n-grams to identify authors and idiolects: A corpus approach to a forensic linguistic problem // International Journal of Corpus Linguistics. 2017. No 22(2). P. 212–241. URL: <https://core.ac.uk/download/pdf/84587040.pdf> (дата обращения: 05.07.2019).

³⁶ См.: McMenamin G. R. *Forensic Linguistics : advances in forensic stylistics*, 2002. 331 p.

³⁷ См.: Coulthard M. Author identification, idiolect, and linguistic uniqueness // Applied Linguistics. 2004. No. 24(4). P. 431–447.

³⁸ См.: Juola Patrick. Authorship Attribution // Foundations and Trends in Information Retrieval. 2006. Vol. 1, no. 3. P. 233–334.

³⁹ Постановление суда по интеллектуальным правам от 4 марта 2019 г. по делу № А63-22578/2017. URL: <https://www.consultant.ru/cons/cgi/online.cgi?req=doc&base=SIP&n=47458#Vx4EAKT6469uEwb6>; Постановление суда по интеллектуальным правам от 18 июня 2019 г. по делу № А40-224162/2017. URL: <https://www.consultant.ru/cons/cgi/online.cgi?req=doc&base=SIP&n=50624#O8vEAKTOk5PTp1VC1>; Постановление суда по интеллектуальным правам от 13 января 2020 г. по делу № А57-15203/2018. URL: <https://www.consultant.ru/cons/cgi/online.cgi?req=doc&base=SIP&n=57313#0i2FAKT04aVc7Ndw>; и пр.

⁴⁰ Определение от 20 июля 2020 г. по делу № СИП-250/2017 о приостановлении производства по делу и проведении экспертизы. URL: <https://www.consultant.ru/cons/cgi/online.cgi?req=doc&base=SIP&n=65485#oGCFAKT4cCjdSonC1>; Кассационное определение от 30 июня 2021 г. № 223-УД21-10-А6 «Анализируя информацию, размещенную в социальной сети "ВКонтакте" <...> заключения комплексной психолого-лингво-религиоведческой судебной экспертизы от 10 июня 2019 г., а также автороведческой судебной экспертизы от 15 ноября 2019 г.». URL: <https://legalacts.ru/sud/kassatsionnoe-opredelenie-sudebnoi-kollegii-po-delam-voennosluzhashchikh-verkhovnogo-suda-rossiiskoi-federatsii-ot-30062021-n-223-ud21-10-a6/>

⁴¹ Постановление следователя отдела № 6 СЧ ГСУ ГУ МВД России по Свердловской области лейтенанта юстиции Сайфетдинова Р. Р. от 5 сентября 2018 г. по уголовному делу № 11801650081000303; Постановление стар-

На данный момент в англоязычной судебной лингвистике главным событием автоматической обработки текста для целей определения его авторства и иных индивидуальных особенностей письменной языковой личности является, пожалуй, серия мероприятий PAN в рамках Conference and Labs of the Evaluation Forum или Cross-Language Evaluation Forum (<https://pan.webis.de/>), в которых принимают участие и отечественные ученые, а именно Т. А. Литвинова из RusProfilingLab⁴². Следует, правда, отметить, что RusProfilingLab – на данный момент практически единственная организация в России, на профессиональном уровне и постоянной основе занимающаяся разработкой открытых, общедоступных автоматических атрибуционных алгоритмов для русского языка, в том числе в рамках решения судебно-экспертных задач. Для нужд закрытых ведомственных организаций улучшением уже имеющегося ПО «Авторовед» занимается А. С. Романов и его коллеги из Томского государственного университета систем управления и радиоэлектроники⁴³.

Несмотря на наличие сильной традиции интерпретирующей лингвистики как в русскоязычной, так и в англоязычной судебных атрибуционных школах, преимущественное использование качественных методов обусловлено не столько традициями этой области лингвистики, сколько закреплённой законодательно (как в России, так и за рубежом) личной ответственностью эксперта (как перед гражданским обществом, так и перед государством) при составлении заключения. Важно, что ни на за рубежом, ни в России сейчас не существует в полном объеме верифицированных и повсеместно рекомендованных методик автоматического (машинного) атрибуционного анализа, использующего только статисти-

стику, извлеченную из текста. Это мотивировано сложностью объекта анализа: тексты, предоставляемые на исследование, могут различаться объемом, функциональным стилем, метаданными, налагающими отпечаток на их структуру, и пр. Отсутствие единого, общепринятого и повсеместно рекомендованного автоматического исследовательского алгоритма текстовой атрибуции при постулатах современного законодательства в России не дает эксперту на данном этапе возможности использовать собственно статистические методы без применения интерпретирующего подхода. Тем не менее цифровизация экспертных методов неизбежна, поэтому, на наш взгляд, *путем развития судебного автороведения атрибуции в России является путь интеграции* машинных, автоматических методов анализа текста, использующих количественную основу, и интерпретирующего, квалификационного исследования, реализуемого непосредственно экспертом, в одном программном комплексе. Такие попытки, безусловно, существуют⁴⁴, и видится логичным их продолжение.

Основной целью исследования является разработка интегративной методики атрибуции текстового материала, в том числе формализация моделей атрибуции языковой личности с целью адаптации алгоритма: а) для машинной реализации; б) для широкого круга лингвистов, в том числе судебных. Примером такого алгоритма может служить, например, прототип атрибуционного ресурса «Хором», апробированный авторами статьи⁴⁵ и доступный в сети Интернет: <http://khorom-attribution.ru/#/>.

Библиографический список

Баранов А. Н. Введение в прикладную лингвистику : учеб. пособие. М. : Эдиториал УРСС, 2001. 360 с.

шего следователя Следственного отдела ОМВД России по г. Ноябрьску капитана юстиции Никифоровой Е. А. от 15 июня 2018 г. по уголовному делу № 11701711492002633 ; Постановление старшего следователя по ОВД СЧ СУ УТ МВД России по ПФО подполковника юстиции Тютнева Ф. В. от 22 февраля 2017 г. по уголовному делу № 11701000150103930 ; и пр.

⁴² Overview of the RusProfiling PAN at FIRE Track on Cross-genre Gender Identification in Russian / T. Litvinova, F. Rangel, P. Rosso et al. Working notes of FIRE 2017 Forum for Information Retrieval Evaluation, 2017. URL: <http://ceur-ws.org/Vol-2036/T1-1.pdf> (дата обращения: 05.07.2019).

⁴³ См.: Authorship Identification of a Russian-Language Text Using Support Vector Machine and Deep Neural Networks...

⁴⁴ См., например: Баранов А. Н. Введение в прикладную лингвистику : учеб. пособие. М., 2001. 360 с. ; Ионова С. В., Огорелков И. В. Речевая диагностика личности по гендерному признаку в автороведении: количественный подход // Вестник Волгоград. гос. ун-та. Серия 2: Языкознание. 2020. Т. 19, № 1. С. 115–127. DOI: 10.15688/jvolsu2.2020.1.10.

⁴⁵ См.: Khomenko A., Baranova Yu., Romanov A., Zadornov K. Linguistic Modeling as a Basis for Creating Authorship Attribution Software // Computational Linguistics and Intellectual Technologies : Proceedings of the International Conference «Dialogue 2021», 2021. URL: <http://www.dialog-21.ru/media/5315/khomenkoaplusetal048.pdf> (дата обращения: 23.06.2021).

Вул С. М. Судебно-автороведческая идентификационная экспертиза : методические основы : метод. пособие. Харьков : ХНИИСЭ, 2007. 64 с.

Галяшина Е. И., Ермолова Е. И. Лингвокриминалистические возможности авторизации письменных и устных текстов : Междунар. науч. конф. М., 2005. С. 20–22.

Горошко Е. И. Судебно-автороведческая классификационная экспертиза : проблемы установления пола автора документа // Теория и практика судебной экспертизы и криминалистики. Харьков : Право, 2003. Вып. 3. С. 221–226.

Ионова С. В., Огорелков И. В. Речевая диагностика личности по гендерному признаку в автороведении : квантитативный подход // Вестник Волгоград. гос. ун-та. Серия 2 : Языкознание. 2020. Т. 19, № 1. С. 115–127. DOI: 10.15688/jvolsu2.2020.1.10

Комиссаров А. Ю. Криминалистическое исследование письменной речи : учеб. пособие. М. : ЭКЦ МВД России, 2000. 126 с.

Комплексная методика производства автороведческих экспертиз : методические рекомендации / И. И. Рубцова, Е. И. Ермолаева, А. И. Безрукова [и др.]. М. : ЭКУ МВД России, 2007. С. 104–107.

Литвинова Т. А., Громова А. В. Компьютерные технологии в судебной автороведческой экспертизе : проблемы и перспективы использования // Вестник Волгоградского государственного университета. Серия 2 : Языкознание. 2020. Т. 19, № 1. С. 77–88. DOI: 10.15688/jvolsu2.2020.1.7

Марусенко М. А. Атрибуция анонимных и псевдонимных литературных произведений методами распознавания образов. Л. : Изд-во Ленингр. ун-та, 1990. 164 с.

Моисеева Т. Ф. Криминалистические экспертизы : изучение в рамках криминалистики и судебной экспертизы // Уголовное судопроизводство : проблемы теории и практики. 2016. № 1. С. 50–51.

Программа систем поддержки атрибуции текстов статей Ф. М. Достоевского / В. Н. Захаров, А. А. Леонтьев, А. А. Рогов, Ю. В. Сидоров // Труды ПетрГУ. Петрозаводск, 2000. Вып. 9. Серия : Прикладная математика и информатика. С. 113–122.

Радбиль Т. Б., Маркина М. В. Вероятностно-статистические модели в производстве автороведческой экспертизы русскоязычных текстов // Политическая лингвистика. 2019. № 2 (74). С. 156–166.

Рогов А. А., Кулаков К. А., Москин Н. Д. Программная поддержка в решении задачи атрибуции текстов // Программная инженерия. 2019. № 5. С. 234–240.

Россинская Е. Р. Судебная экспертиза в гражданском, арбитражном, административном и уголовном процессе. 3-е изд., доп. М. : Норма : ИНФРА-М, 2014. 735 с.

Тимашев А. Н. Атрибутор : Версия 1.01 : Описание программы. М., 1999–2007. URL: http://www.textology.ru/atr_resum.html (дата обращения: 01.02.16).

Филиппов А. Г. О понятии и классификации криминалистических экспертиз // Актуальные проблемы криминалистического обеспечения расследования преступлений : труды Академии управления МВД России. М. : Изд-во МВД России, 1996.

Хмелев Д. В. Лингвоанализатор. М., 2002. URL: <http://www.rusf.ru/books/analysis/> (дата обращения: 16.11.2017).

Шевелёв О. Г. Методы автоматической классификации текстов на естественном языке : учеб. пособие. Томск : ТМЛ-Пресс, 2007. 144 с.

Энциклопедия судебной экспертизы / под ред. Т. В. Аверьяновой, Е. Р. Россинской. М., 1999. С. 177.

Authorship Identification of a Russian-Language Text Using Support Vector Machine and Deep Neural Networks / A. Romanov, A. Kurtukova, A. Fedotova et al. // Future Internet. 2021. Vol. 13, iss. 1. 3. 16 p.

Campbell L. The Sophistries and Polilicus of Plato. Oxford : Clarendon, 1867. 170 p.

Coulthard M. Author identification, idiolect, and linguistic uniqueness // Applied Linguistics. 2004. No. 24 (4). P. 431–447.

Juola P. Authorship Attribution // Foundations and Trends in Information Retrieval. 2006. Vol. 1, no. 3. P. 233–334.

Koppel M., Schler J. Exploiting Stylistic Idiosyncrasies for Authorship Attribution // Proceedings of IJCAI'03 Workshop on Computational Approaches to Style Analysis and Synthesis. 2003. No. 69. P. 72–80.

Linguistic Modeling as a Basis for Creating Authorship Attribution Software / A. Khomenko, Yu. Baranova, A. Romanov, K. Zadvornov // Computational Linguistics and Intellectual Technologies : Proceedings of the International Conference “Dialogue 2021”. Moscow, June 16–19, 2021. URL: <http://www.dialog-21.ru/media/5315/khomenkoaplusetal048.pdf>

McMenamin G. R. Forensic Linguistics : advances in forensic stylistics, 2002. 331 p.

Overview of the RusProfiling PAN at FIRE Track on Cross-genre Gender Identification in Russian / T. Litvinova, F. Rangel, P. Rosso et al. // Working notes of FIRE 2017 – Forum for Information Retrieval Evaluation. Bangalore, India, 8th–10th December 2017. 2017. Vol. 2036. P. 1–7. URL: <http://ceur-ws.org/Vol-2036/T1-1.pdf>

Stamatatos E. Authorship attribution using text distortion // Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics. 2017. Vol. 1, Long Papers. P. 1138–1149.

Wright D. Using word n-grams to identify authors and idiolects : A corpus approach to a forensic linguistic problem // International Journal of Corpus

Linguistics. 2017. No. 22 (2). P. 212–241. URL: <https://benjamins.com/#catalog/journals/ijcl.22.2.03wri/details>

References

Authorship Identification of a Russian-Language Text Using Support Vector Machine and Deep Neural Networks / A. Romanov, A. Kurtukova, A. Fedotova et al. // *Future Internet*. 2021. Vol. 13, iss. 1. 3. 16 p.

Baranov A. N. Introduction to Applied Linguistics : Study Guide. Moscow : Editorial URSS, 2001. 360 p.

Campbell L. The Sophistries and Polilicus of Plato. Oxford : Clarendon, 1867. 170 p.

Coulthard M. Author identification, idiolect, and linguistic uniqueness // *Applied Linguistics*. 2004. No. 24 (4). P. 431–447.

Encyclopedia of Forensic expertise / eds. T. V. Aver'yanova, E. R. Rossinskaya. Moscow, 1999. P. 177.

Filippov A. G. On the concept and classification of forensic expertise // Actual problems of forensic support investigation of crimes : Proceedings of the Academy. Moscow : Publishing House of the Ministry of Internal Affairs of Russia, 1996.

Galyashina E. I., Ermolova E. I. Linguistic and criminalistic authoship opportunities for written and oral texts. Moscow, 2005. P. 20–22.

Goroshko E. I. Forensic classification examination : problems of establishing the gender of the document author // *Theory and practice of forensics examination*. Khar'kov : Pravo, 2003. Vol. 3. P. 221–226.

Integrated methodology for the production forensic authorship examinations : Methodological recommendations / I. I. Rubcova, E. I. Ermolaeva, A. I. Bezrukova et al. Moscow : EKV MVD of Russia, 2007. 192 p.

Ionova S. V., Ogorel'kov I. V. Personality speech diagnostics in author identification based on gender parameter : quantitative approach // *Science Journal of Volgograd State University*. Linguistics. 2020. Vol. 19, no. 1. P. 115–127. DOI: <https://doi.org/10.15688/jvol-su2.2020.1.10>

Juola P. Authorship Attribution // *Foundations and Trends in Information Retrieval*. 2006. Vol. 1, no. 3. P. 233–334.

Khmelev D. V. Lingvoanalyzer. Moscow, 2002. URL: <http://www.rusf.ru/books/analysis/>

Komissarov A. Yu. Forensic study of written language. Moscow : EKV MVD of Russia, 2000. 126 p.

Koppel M., Schler J. Exploiting Stylistic Idiosyncrasies for Authorship Attribution // Proceedings of IJCAI'03 Workshop on Computational Approaches to Style Analysis and Synthesis. 2003. No. 69. P. 72–80.

Linguistic Modeling as a Basis for Creating Authorship Attribution Software / A. Khomenko, Yu. Baranova, A. Romanov, K. Zadvornov // *Computational Linguistics and Intellectual Technologies* :

Proceedings of the International Conference “Dialogue 2021”. Moscow, June 16–19, 2021. URL: <http://www.dialog-21.ru/media/5315/khomenkoaplusetal048.pdf>

Litvinova T. A., Gromova A. V. Current Problems of Forensic Authorship Analysis and the Possibility of Their Solution with the Use of Computer Methods : Problems and Prospects // *Science Journal of Volgograd State University*. Linguistics. 2020. Vol. 19, no. 1. P. 77–88.

Marusenko M. A. Attribution of Anonymous and Pseudonymous Texts as a Typical Pattern Recognition Problem. Leningrad : Leningr. University publ., 1990. 164 p.

McMenamin G. R. Forensic Linguistics : advances in forensic stylistics. 2002. 331 p.

Moiseeva T. F. Forensic examinations: a study within the framework of forensic science and forensic examination // *Criminal justice : problems of theory and practice*. 2016. No. 1. P. 50–51.

Overview of the RusProfiling PAN at FIRE Track on Cross-genre Gender Identification in Russian / T. Litvinova, F. Rangel, P. Rosso et al. // Working notes of FIRE 2017 – Forum for Information Retrieval Evaluation. Bangalore, India, 8th–10th December 2017. 2017. Vol. 2036. P. 1–7. URL: <http://ceur-ws.org/Vol-2036/T1-1.pdf>

Radbil T. B., Markina M. V. Probabilistic-Statistical Models in Conducting Authoring Expertise of Russian Texts // *Political Linguistics*. 2019. No. 2 (74). P. 156–166.

Rogov A. A., Kulakov K. A., Moskin N. D. Software support in solving the problem of text attribution // *Programmnyaya Ingeneria*. 2019. Vol. 10, no. 5. P. 234–240.

Rossinskaya E. R. Forensic examination in civil, arbitration, administrative and criminal proceedings / Russian Federal Center for Forensic Science under the Ministry of Justice of the Russian Federation. 3rd ed., add. Moscow : Norma : INFRA-M, 2014. 735 p.

Shevelyov O. G. Methods of automatic classification of texts in natural language : Textbook. Tomsk : TML-Press, 2007. 144 p.

Stamatatos E. Authorship attribution using text distortion // Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, 2017. Vol. 1, Long Papers. P. 1138–1149.

Timashev A. N. Atributor : Version 1.01 : Description of the program. Moscow, 1999–2007. URL: http://www.textology.ru/atr_resum.html

Vul S. M. Forensic Attribution Identification Examination : Methodological Basics : Methodological Manual. X. : HNIISE, 2007. 64 p.

Wright D. Using word n-grams to identify authors and idiolects: A corpus approach to a forensic linguistic problem // *International Journal of Corpus*

Linguistics. 2017. No. 22 (2). P. 212–241. URL: <https://benjamins.com/#catalog/journals/ijcl.22.2.03wri/details>

Zaharov V. N. The program of supporting systems for the attribution of articles by F. M. Dostoevsky // The works of PetrSU. Petrozavodsk, 2000. Vol. 9. P. 113–122.

Национальный исследовательский университет
«Высшая школа экономики – Нижний Новгород»

Романова Т. В., доктор филологических наук, профессор, руководитель департамента фундаментальной и прикладной лингвистики факультета гуманитарных наук

E-mail: tvromanova@hse.ru

Национальный исследовательский университет
«Высшая школа экономики – Нижний Новгород»,
Центр экспертиз и исследований «ЕСИН»

Хоменко А. Ю., кандидат филологических наук, доцент департамента фундаментальной и прикладной лингвистики факультета гуманитарных наук, эксперт-авторовед

E-mail: akhomenko@hse.ru

Поступила в редакцию: 02.05.2022

Для цитирования:

Романова Т. В., Хоменко А. Ю. Автоматизация судебной автороведческой экспертизы: проблемы и перспективы // Вестник Воронежского государственного университета. Серия: Право. 2023. № 4 (55). С. 288–296. DOI: <https://doi.org/10.17308/law/1995-5502/2023/4/288-296>

HSE Campus in Nizhny Novgorod

Romanova T. V., Doctor of Philological Sciences, Professor, Head of the School of Fundamental and Applied Linguistics of the Faculty of Humanities

E-mail: tvromanova@hse.ru

HSE Campus in Nizhny Novgorod, Center for Expertise and Research «ESIN»

Khomenko A. Yu., Candidate of Philological Sciences, Associate Professor of the School of Fundamental and Applied Linguistics of the Faculty of Humanities, forensic expert in text author attribution

E-mail: akhomenko@hse.ru

Received: 02.05.2022

For citation:

Romanova T. V., Khomenko A. Yu. Forensic text attribution computerisation: problems and prospects // Proceedings of Voronezh State University. Series: Law. 2023. № 4 (55). P. 288–296. DOI: <https://doi.org/10.17308/law/1995-5502/2023/4/288-296>.