

Научная статья

УДК 930:388.14«1920-е»

DOI: 10.31143/2542-212X-2022-4-278-296

EDN: KSMSYL

## ИНФОРМАЦИОННЫЕ ДОКУМЕНТЫ ГПУ–ОГПУ 1920-х ГОДОВ: ОПЫТ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ТЕКСТОВ (НА МАТЕРИАЛАХ СЕВЕРНОГО КАВКАЗА)

Александр Юрьевич Рожков<sup>1</sup>, Тимофей Александрович Рунаев<sup>2</sup>

<sup>1,2</sup> Кубанский государственный университет, Краснодар, Россия

<sup>1</sup> avro14@mail.ru, <https://orcid.org/0000-0002-3064-2915>

<sup>2</sup> runaevt@yandex.ru, <https://orcid.org/0000-0002-0154-6959>

**Аннотация.** Использование массовых источников в историческом исследовании позволяет улучшить репрезентативность знаний об историческом периоде. Однако качественный анализ текстов имеет ограничения при работе с большим массивом данных. Поэтому в современной исторической науке появляется потребность в привлечении компьютерного («интеллектуального») анализа текстов для обработки массовых источников. Придерживаясь принципов дискурс-исторического подхода, авторы продемонстрировали возможность применения компьютерной программы RStudio для интеллектуального анализа объемного корпуса информационных сводок и обзоров ГПУ–ОГПУ 1920-х гг. на материалах Северного Кавказа. В результате установлено, что сводки и обзоры ГПУ–ОГПУ разных лет имели похожий терминологический, эмотивный и тематический состав. Несмотря на внутреннее единство информационных документов ОГПУ по Северному Кавказу, начиная с 1925 г. в них выявлено ежегодное расширение тематического поля. Эмоциональное нагнетание и усложнение «чекистского» дискурса настолько усилились к концу 1920-х, что терминологический состав документов 1928–1929 гг. стал значительно отличаться от состава текстов предыдущих лет. Это указывает на трансформацию регионального «чекистского» дискурса, который транслировался высшему руководству страны. Примененный подход позволяет минимизировать субъективность оценок и выводов исследователя.

**Ключевые слова:** 1920-е, ГПУ–ОГПУ, информационные сводки, обзоры политического состояния, Северный Кавказ, интеллектуальный анализ текста, программа RStudio

**Благодарности:** Исследование выполнено в рамках проекта РФФИ № 21–09–43110 «Взаимоотношения партийно-государственной власти и полиэтничного общества в ходе реализации национальной политики (на материалах Северо-Кавказского края и Крымской АССР, 1920-е гг.)».

**Для цитирования:** Рожков А.Ю., Рунаев Т.А. Информационные документы ГПУ–ОГПУ 1920-х годов: опыт интеллектуального анализа текстов (на материалах Северного Кавказа) // Электронный журнал «Кавказология». – 2022. – № 4. – С. 278-296. – DOI: 10.31143/2542-212X-2022-4-278-296. – EDN: KSMSYL.

---

© Рожков А.Ю., Рунаев Т.А., 2022

Original article

**INFORMATION DOCUMENTS OF THE GPU–OGPU IN 1920s:  
EXPERIENCE OF TEXT MINING  
(BASED ON MATERIALS FROM THE NORTH CAUCASUS)****Alexander Yu. Rozhkov<sup>1</sup>, Timofey A. Runaev<sup>2</sup>**<sup>1,2</sup> Kuban State University, Krasnodar, Russia<sup>1</sup> avro14@mail.ru, <https://orcid.org/0000-0002-3064-2915><sup>2</sup> runaevt@yandex.ru, <https://orcid.org/0000-0002-0154-6959>

**Abstract.** Using mass sources in historical research makes it possible to improve the representativeness of knowledge about the historical period. However, qualitative text analysis has limitations when working with a large array of data. Thus, in modern historical science is a need to involve computer («intellectual») analysis of texts for processing mass sources. Based on principles of the discourse-historical approach, authors indicated the possibility of using the RStudio computer program for intellectual analysis of the voluminous corpus of information reports and reviews of the GPU–OGPU of the 1920s on the materials of the North Caucasus. In the first stage, the contextualization of the studied sources was carried out. The authors found that reports and reviews of the GPU-OGPU of different years had similar terminological, emotive and thematic composition. While maintaining the internal unity of the OGPU information documents on the North Caucasus, an increase in the number of topics in summaries and reviews has been revealed since 1925. Emotional aggravation and complication of the «Chekist» discourse intensified so much by the end of the 1920s that terminological composition of the documents of 1928-1929 differed significantly from composition of texts of previous years. Thus, the text mining of the reports and reviews of the GPU–OGPU demonstrated the transformation of the regional «Chekist» discourse, which was broadcast to the top leadership of the country. The applied approach allows minimizing subjectivity of researcher's assessments and conclusions.

**Key words:** 1920s, GPU-OGPU, information summaries, political status reviews, North Caucasus, text mining, R Studio program

**Gratitude:** The study was carried out within the framework of the RFBR project No. 21-09-43110 "Relations of party-state power and multi-ethnic society during the realization of national politics (based on materials from the North Caucasus region and the Crimean Autonomous Soviet Socialist Republic, 1920s)."

**For citation:** Rozhkov A.Yu., Runaev T.A. Information documents of the GPU–OGPU in 1920s: experience of text mining (based on materials from the North Caucasus). IN: Electronic journal «Caucasology». – 2022. – № 4. – P. 278-296. – DOI: 10.31143/2542-212X-2022-4-278-296. – EDN: KSMSYL.

---

© Rozhkov A.Yu., Runaev T.A., 2022

**Введение**

Ставшие доступными после открытия советских архивов информационные документы ВЧК–ОГПУ–НКВД все активнее вводятся в научный оборот. Материалы сводок и обзоров спецслужб используются историками в качестве массового источника и для изучения различных процессов на Северном Кавказе [Баранов 2013; Наухацкий 2011; Рожков 2022 и др.].

Настоящим прорывом в источниковом обеспечении стал выход в свет многолетнего сборника документов «“Совершенно секретно”: Лубянка – Сталину о положении в стране (1922–1934 г.)», который дает возможность в помесечной динамике на протяжении 13 лет проследить информацию о положении в СССР, представленную руководством ГПУ–ОГПУ высшим органам власти. Достоинства сборника подробно описаны [Колодникова 2010; Кондрашин 2018]. Однако потенциал этого уникального издания используется не в полной мере. В основном разрозненные фрагменты сводок и обзоров ГПУ–ОГПУ задействуются исследователями как фактографический материал по частным темам, что затрудняет понимание основных тенденций как в социально-исторической динамике советского общества, так и в информационном потоке с «Лубянки». Практически нет работ, в которых тексты этих информационных документов исследовались бы системно с привлечением компьютерных программ. Задачей настоящей статьи является попытка в какой-то мере устранить этот пробел. Нас будут интересовать не отдельные цитаты из сводок и обзоров, а полные тексты информационных документов как массив данных с целью проведения их интеллектуального анализа с помощью компьютерной программы RStudio.

Поскольку публикация полных, без купюр, текстов обзоров ОГПУ в вышеназванном сборнике затрудняет качественный анализ, мы произвели выборку по одному из крупных регионов (Юго-Восточный, позднее Северо-Кавказский край) за период с 1922 по 1929 гг. Нижняя граница обусловлена начальной датой документов, размещенных в сборнике. Это был год трансформации ВЧК в ГПУ, период налаживания стройной системы информирования чекистами высшего руководства страны; завершился этот год образованием СССР. Верхняя граница означает начало «великого перелома», а также коренных перемен в системе информирования ОГПУ. С 1930 г. уже не составлялись ежемесячные обзоры, «Лубянка» перешла на систему многочисленных оперативных документов (спецсводок Секретно-оперативного управления (СОУ) ОГПУ, докладов и сообщений полпредов, докладных записок и справок информационного отдела (ИНФО) ОГПУ и т. д.). При этом наметился явный уход от детализации и от географического принципа построения сводок [Виноградов 2001: 73].

### Методология

Концептуально мы рассматриваем весь корпус текстов информационных сводок и обзоров ГПУ–ОГПУ, подвергнутый анализу, как «чекистский» дискурс, дискурс «института власти» [Согомонян 2012: 49]. Нам близка концепция М. Пешё о том, что дискурсивный процесс является частью идеологических классовых отношений, а смыслы слов меняются в зависимости от классовых позиций в политической борьбе [Скворцова 2007: 7–8]. В этой связи полагаем, что субъектом производства и трансляции «чекистского» дискурса выступали не конкретные люди (информаторы, редакторы и руководители ИНФО), а некая внесубъектная «матрица смыслов», «идеологическая формация» (по Пешё), которая автоматически управляет дискурсом субъекта и определяет способы производства дискурса. В изучаемый период эту «идеологическую формацию»

олицетворяли вполне реальные институции в виде ВЧК–ОГПУ и партийно-государственного аппарата, которые определяли, что можно говорить, а что нельзя [Кожемякин 2008: 10].

В качестве методологической рамки выбран дискурс-исторический подход (*Discourse-Historical Approach – ДНА*), который основывается на принципе историчной контекстуализации дискурса. Здесь при интерпретации дискурса учитываются социальные, культурные, экономические, политические процессы исторических периодов, которые отражаются в дискурсах, зафиксированных в исторических текстах [Richardson 2017: 61–62]. Помимо исторического контекста важной переменной для ДНА является время, в котором тексты одного дискурса (или сами дискурсы) темпорально связаны друг с другом через темы, акторов, отсылки к одинаковым событиям, перенос системы аргументации [Wodak 2015: 279]. Дизайн исторического исследования в ДНА предполагает следующий алгоритм: анализ языка текстов → анализ интертекстуальности (поиск связей текстов исторических источников во времени, встроенных в дискурс) → анализ институциональных рамок, в которых создавались и циркулировали данные источники → анализ широкого исторического контекста, который влияет на содержание дискурса [Wodak, Meyer 2001: 67].

Мы подвергли тексты сводок и обзоров ГПУ–ОГПУ интеллектуальному анализу, дизайн которого построен на алгоритме ДНА, но в несколько иной последовательности. Вначале мы описываем исторический контекст и институциональные рамки, в которых создавались и циркулировали сводки и обзоры. Далее переходим непосредственно к анализу языка текстов и интертекстуальности.

### Исторический контекст

В 1920-е гг. Северный Кавказ площадью 293,7 тыс. кв. км представлял собой довольно сложный в политическом, экономическом, этнокультурном и социально-демографическом отношении макрорегион. По площади край был сопоставим с островной Англией или материковой Италией, в семь раз превосходил территории Дании и Швейцарии, в десять раз – Голландии [Северный Кавказ 1925: 57]. Край обладал рудными месторождениями цветных металлов, нефтяными и каменноугольными месторождениями. Большие площади равнинных пахотных земель, теплый климат, чернозем обеспечивали региону развитие сельского хозяйства. Выходы к Азовскому и Черному морям служили транспортными путями для торговли.

Структура и облик Юго-Востока России во многом изменились в результате административно-территориальной реформы (районирования) в 1923–24 гг. До районирования в границах Юго-Восточного края находились 10 административно-хозяйственных губернских единиц: 2 губернии, 2 области, 4 автономные национальные области, 1 автономная республика и 1 город на правах губернии [Северный Кавказ 1925: XIV]. В ходе реформы прежняя административная структура (область/губерния – округ/уезд – волость) сменяется укрупненной (край – округ+автономные области – район+национальный район) с краевым подчинением вместо центрального. В 1924 г. происходит слияние Кубано-Черноморской и Донской областей в Юго-Восточную, которая поглощает

затем Ставропольскую и Терскую губернии и, включая все горские автономии (кроме Дагестана), оформляется как Северо-Кавказский край, составленный из 21 окружных единиц – 14 округов, 1 автономного города и 6 национальных автономных областей [Цуциев 2007: 66].

По численности населения регион был равен взятым в отдельности Бельгии, Венгрии и европейской Турции. В конце 1926 г. в 66 городских и 15671 сельских поселениях края проживало около 8,4 млн человек (1/8 населения СССР), из них 19,8 % населения проживало в городах и 80,2 % на селе [Всесоюзная перепись... 1928: 2–13]. Регион населяли свыше 88 наций и народностей. Официально русские составляли в среднем 45,9 % населения, украинцы – 37,1 %, среди других этносов наиболее многочисленными были (в порядке убывания) чечены, армяне, осетины, кабардинцы, немцы, ингуши, черкесы, карачаи, белорусы, евреи, балкары, греки и др. [Всесоюзная перепись... 1928: 34–51].

Этносоциальные и земельные конфликты в регионе, унаследованные с до-революционных времен, обострились в ходе революции 1917 г., затяжной Гражданской войны и административно-территориальных преобразований в 1920-е гг. Вместе с тем стоит согласиться с В.С. Белозеровым, что период 1920–30-х гг. для Северного Кавказа оказался переходным от этапа колонизации к этапу интенсивного хозяйственного освоения и развития урбанизации [Белозеров 2005: 54], что во многом отражали сводки и обзоры ОГПУ.

### **Институциональные рамки**

С первых дней существования ВЧК в ее состав входило Информбюро, позже называемое Отделением обработки материалов в составе СОУ. В Особом отделе ВЧК с декабря 1919 г. существовал ИНФО с четырьмя отделениями [Капчинский 2005: 336]. После Кронштадтских событий ведомственная система информации преобразуется в государственную. В марте 1921 г. создан ИНФО ВЧК с непосредственным подчинением начальнику СОУ ВЧК. В задачи отдела входила систематизация и обработка поступающих с мест сводок по разной тематике. В ноябре 1922 г. упорядочивается классификация секретных сотрудников ГПУ (информаторы, осведомители, агенты). Те из них, кто давал материалы общего характера, были переведены в категорию информаторов и переданы в ИНФО. В тот период ИНФО включал три отделения: госинформации, секретной информации и иностранной информации. В ноябре 1925 г. ИНФО и Отдел политконтроля (ПК) объединились в Отдел информации и политконтроля, штат которого возрос до пяти отделений. С 1 декабря 1929 г. штаты отдела увеличились до восьми отделений [Лубянка 2003: 14–44].

Объект анализа – сводки (госинфсводки, спецполитсводки) и обзоры о политико-экономическом положении РСФСР (СССР). Информсводки ВЧК–ОГПУ предназначались для оперативного информирования высшего руководства партии. В 1921 г. появляется единая госинформсводка, рассылаемая в 30–40 адресов, которая давала общую картину политического положения по всем районам страны и по всему кругу проблем. Затем появились спецсводки ИНФО ГПУ – сводки по специальным вопросам. Спецполитсводки информировали о собственной деятельности ГПУ–ОГПУ против враждебных советскому строю по-



литических сил, организаций и лиц. Они рассылались восьми адресатам, начиная с Ф.Э. Держинского. Как правило, сводки ИНФО выходили ежедневно кроме воскресных дней и рассылались в тот же день вечером [Советская деревня 2000: 9–13]. Только в течение 1922 г. ИНФО выпустил 328 номеров госинфсводки, каждая от 33 до 36 экземпляров. На их основе в 1922 г. были подготовлены 1035 выпусков различных спецсводок [Виноградов 2001: 47–48].

С 1922 г. аппарат ИНФО ГПУ приступил к составлению ежемесячных Обзоров политического и экономического состояния (с 1925 г. – политического) республики (с 1923 г. – СССР), представлявших собой весьма обширные тексты (до 100 страниц и более) с анализом материалов не только информационных сводок, но и докладов, справок различных подразделений ГПУ–ОГПУ, в первую очередь Секретно-политического (СПО) и Контрразведывательного (КРО) отделов. Помимо информирования высшего руководства партии, обзоры были призваны «в сжатой форме дать правильную ориентацию местным органам ОГПУ по важнейшим вопросам» [Виноградов 2001: 62]. Обзоры структурировались по изначально определенным основным рубрикам (как правило, социальным группам и институтам). Первой всегда шла рубрика «Рабочие» (вариант – «Город»), затем «Крестьяне» («Деревня»), затем «Восточные национальные республики», далее – по ситуации «Армия», «Бандитизм», «Антисоветские организации», «Интеллигенция», «Духовенство» и «Эмиграция».

Источниками для сводок служили донесения секретных информаторов в учреждениях и на предприятиях; доклады уполномоченных ВЧК–ОГПУ; статистика происшествий; следственный и агентурный материал; материалы прессы; «народная молва» (вероятно, имелись в виду материалы перлюстрированной корреспонденции, а также донесения агентов о работе в толпе) [Лубянка 2003: 22; Рожков 1995]. По данным Л.П. Колодниковой, только пространственная информация за 1929 г., полученная аналитиками ИНФО ОГПУ, была собрана в 2300 административных единицах, находившихся в их поле зрения. «В числе этих административных единиц было 9 союзных республик, 11 автономных, 9 крупных административных единиц, 26 областей, 220 округов, 16 губерний, 537 сел, 532 района, множество станиц, хуторов, сел, деревень, поселков. Чекистская сеть только в 1929 г. информировала о том, что происходило в 18 кантонах, 7 аймаках, 56 аулах, 25 кишлаках, а также на 300-х отдельных административных участках, многочисленных станциях, станицах, а также в таких специфических административных единицах, как тумены, теми, сомоны, оазисы» [Колодникова 2010: 233].

Руководство ВЧК–ОГПУ уделяло серьезное внимание качеству первичной (на сленге чекистов – «сырой» и «полусырой») информации. Интеллектуальный уровень многих осведомителей, особенно в деревне, не способствовал получению качественной информации, на что указывает сводка Ростовского окружного ЧК в информационную часть ДонЧК «О ходе работ секции по борьбе с левыми антисоветскими партиями за первую половину января 1921 года»: «От 56 осведомителей получено 30 осведомлений неудовлетворительных, так как большинство осведомителей не понимают политического смысла. Проинструктировать не представляется возможным ввиду того, что в секции нет никаких

руководств и указаний, а поэтому необходимо: 1) секцию снабжать своевременно указаниями, инструкциями, распоряжениями и газетами; 2) указать штат работников секции; 3) непосредственно иметь связь с Политбюро, дабы работа протекала в тесном контакте и согласованности, для чего необходимо обязать Политбюро представлять доклады ежемесячно ко 2-му числу» [Советская деревня 2000: 9–10]. Между тем недостатки в работе информационного аппарата на местах выявлялись и позже. В феврале 1924 г. в циркуляре за подписью Г.Г. Ягоды отмечалось слабое проникновение в деревню и малые города, получение «казенной» информации от милиции, отделов исполкомов, малосодержательность «пухлых» госинфсводок с мест, которые отражают «жизнь губернии и особенно по рабоче-крестьянской периферии, нередко слабее, чем иная газета». Выход виделся в пересмотре и укреплении персонального состава инфаго (информационного аппарата губернских отделов), основным источником информации которых должны были стать оперативные, а не «казенные», данные. Начиная с 1924 г., постановка информационной работы ОГПУ положительно сказалась на повышении надежности сводок как исторического источника [Виноградов 2001: 58–61].

#### Анализ языка текстов

Даже извлеченные из сводок и обзоров ГПУ–ОГПУ тексты по Северному Кавказу характеризуются большим объемом, который усложняет исследование их содержания описательным методом. Решением этой проблемы может стать интеллектуальный анализ текста (*text mining*), нацеленный на извлечение скрытой информации из текстов, представляющих в виде неструктурированного массива данных [Jo 2019: 3]. Эта технология включает в себя процесс структурирования исходного текста с применением синтаксического анализа, использования некоторых лингвистических функций с последующей загрузкой в базу данных и интерпретацией результатов. Главная цель – превращение текста в структурированные данные, пригодные для анализа.

Одним из инструментов проведения такого анализа выступает компьютерная программа RStudio, являющаяся консолью языка программирования R, уже использованного в отечественной историографии [Рыбаков 2018]. Изучение сводок и обзоров ОГПУ по Северному Кавказу в рамках интеллектуального анализа текстов предполагает реализацию двух этапов: 1) препроцессинга – предварительной обработки текста и 2) применения методов *text mining*, в частности латентно-семантического анализа (LSA), анализа тональности текстов (*sentiment analysis*) и метода латентного размещения Дирихле (LDA). Эти процедуры позволяют увидеть, насколько документы ОГПУ разных лет похожи друг на друга, как изменялся их эмоциональный фон с течением времени и какие темы в них имели интертекстуальный характер, конструировавший «чекистский» дискурс. Для достижения планируемых результатов использован ряд программных пакетов (*lsa*, *SnowballC*, *tidytext*, *topicmodels*, *tm*), встраиваемых в RStudio.

Для составления массива текста для анализа из всех сводок и обзоров ГПУ–ОГПУ с 1922 по 1929 гг., опубликованных в сборнике «Совершенно сек-

ретно...», были отобраны полные фрагменты текстов, сфокусированные на Северном Кавказе. Поскольку структура информационных материалов в сборнике неоднородная (в 1922 обзоры чередуются с госинфсводками, в 1923 – со спецполитсводками, в 1927 – с информсводками), нам пришлось для машинной обработки все сводки за каждый месяц объединять в один блок, чтобы он по объему был соразмерен с текстами обзоров. Чтобы сделать корпус текстов «аккуратным» [Silge, Robinson 2017: 85], т. е. прошедшим препроцессинг, были пройдены все этапы предварительной подготовки. Тем самым была определена выборочная совокупность лексем, необходимая для проведения анализа. Если вначале корпус насчитывал 187489 слов, то после препроцессинга его объем составил 32115 терминов.

Применение методов интеллектуального анализа текста потребовало преобразования выборочной совокупности слов каждого года в матрицу «термин-документ» посредством пакета tm. Таким образом, тексты ОГПУ приобрели табличную форму, где столбцы – это тексты ОГПУ, распределенные по месяцам, а строки – это слова. В таком случае содержание ячейки – это частота упоминания конкретного слова в сводках и обзорах конкретного месяца.

### Анализ интертекстуальности

Латентно-семантический анализ позволяет увидеть повторное выражение мыслей в разных текстах, определив в них совместное появление слов через значение косинусного сходства [Wild 2016: 71, 76]. Установлено, что тексты ОГПУ по Северному Кавказу разных лет имеют практически идентичный словарный состав, о чем свидетельствуют значения косинусного сходства, полученные в ходе LSA: во всех сводках и обзорах ОГПУ, написанных в разные годы, значение косинусного сходства близится к единице. Иными словами, с 1922 по 1929 гг. ГПУ–ОГПУ обладало сложившимся терминологическим аппаратом, используемым в сводках и обзорах. Основой устоявшегося чекистского понятийного аппарата в обозначенный период являлись такие слова, как «кулак» (частота упоминаний = 866), «власть» (786), «хлеб» (721), «советский» (702), «собрание» (700), «рабочие» (662), «крестьяне» (650), «агитация» (574), «анти-советский» (552), «зажиточный» (551), «сельсовет» (504), «беднота» (501), «выступление» (412), «кулачество» (398), «средняк» (395), «налог» (377), «банда» (347), «перевыборы» (306), «слух» (305), «недовольство» (296), «война» (274), «коммунист» (271), «казак» (270), «земля» (248) и т. д. Однако, несмотря на поддержание общего курса применения терминов, с 1928 г. наблюдается дрейф словарного состава текстов ОГПУ. Именно после 1928 г. сводки и обзоры ОГПУ становятся менее похожими по словарному составу с аналогичными текстами предыдущих лет, но в то же время сохраняется их сходство с текстами 1929 г.

В условиях социальных изменений одной из стратегий дискурса является смягчение или усиление его содержания [Reisigl 2018: 52]. В машинном анализе текстов вектор дискурсивной стратегии позволяет определить анализ тональности текста, основывающийся на полярности, в границах которой слова имеют отрицательную, нейтральную или положительную направленность настроений



[Игнатов, Михальча 2021: 221–222]. Согласно принципу устройства использованного нами лексикона AFINN, адаптированного к русскому языку, отрицательная и положительная тональности текста находятся в диапазоне от  $-5$  до  $5$  соответственно [Silge, Robinson 2017: 13]. Проведенный анализ тональности показал, что текстам ГПУ–ОГПУ с 1922 по 1929 гг. присуще использование отрицательных терминов: в каждом году обзоры и сводки обладают преимущественно негативным содержанием (столбики слева), что подчеркивает отрицательную тональность «чекистского» дискурса (рис. 1).

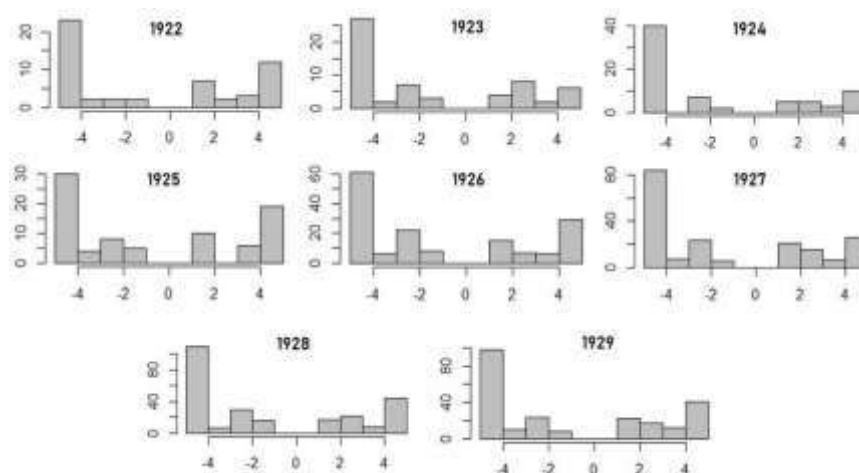


Рис. 1. Тональность текстов ГПУ–ОГПУ по годам  
(ось  $x$  – степень тональности; ось  $y$  – частота слов)

Картина распределения количества слов с негативными коннотациями по годам (рис. 2) позволяет сделать два заключения. Во-первых, заметна тенденция к нагнетанию тональности дискурса в сводках и обзорах ГПУ–ОГПУ с 1922 по 1928 гг. Во-вторых, видно, что по количеству отрицательных терминов документы «Лубянки» заметно делятся на две группы: первая – с 1922 по 1925 гг.; вторая – с 1926 по 1929 гг. Тексты из второй группы наполнены негативным содержанием в три раза больше, чем документы из первой группы. В текстах первого периода самыми негативными терминами (со значением тональности  $-5$  при частоте упоминания  $\geq 2$ ) являются «бандитизм», «забастовка», «восстание», «нападение», «конфликт», «сопротивление», «террор», «столкновение», «бездеятельность», «зло», «ликвидация», «убийство», «драка», «избиение», «стычка», «ограбление». Во втором периоде, помимо перечисленных слов, появляются новые негативные термины (со значением тональности  $-5$  при частоте упоминания  $\geq 2$ ), которые характерны именно для текстов 1926–1929 гг.: «отказ», «произвол», «халатность», «сволочь», «захват», «скандал», «протест», «угроза», «антисемитизм» и т.д. Это свидетельствует о том, что в «чекистском» дискурсе, отраженном в сводках и обзорах ОГПУ по Северному Кавказу, к 1926 г. произошел перелом, который усиливал негативное впечатление о происшедших в стране событиях и процессах.

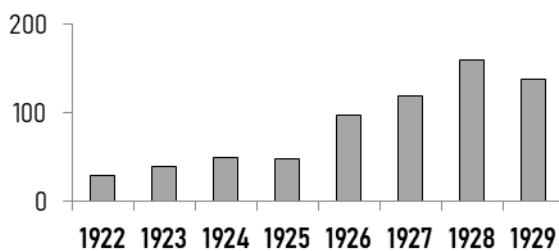


Рис. 2. Количество негативных слов в текстах ОГПУ

Между тем наряду с ростом негативной терминологии в текстах ОГПУ находились также слова с положительной коннотацией. Поэтому целесообразно учитывать соотношение негативной и позитивной тональностей в текстах. Благодаря сравнению средних значений тональностей документов ГПУ–ОГПУ видна динамика отражаемых в текстах настроений с учетом положительных слов:  $-0.815$  в 1922 г.,  $-1.708$  в 1923 г.,  $-1.944$  в 1924 г.,  $-0.730$  в 1925 г.,  $-1.205$  в 1926 г.,  $-1.529$  в 1927 г.,  $-1.347$  в 1928 г.,  $-1.167$  в 1929 г. Средние значения в очередной раз подчеркивают, что в обозначенный временной промежуток «чехистский» дискурс испытал две волны изменений тональности: 1922–1925 гг. и 1926–1929 гг. (рис. 3). При этом настроения в текстах второй группы немного нивелируются, т. е., в отличие от более ранних сводок и обзоров, их тональность не колеблется из года в год, а эмотивный набор терминов становится устойчивым.

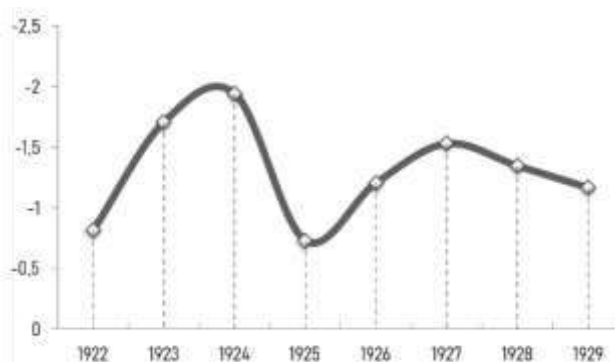


Рис. 3. Сравнение средних значений тональностей текстов ОГПУ

Даже за один месяц сводки и обзоры ОГПУ наполнены массой сведений, которые относятся к различным темам. Это накладывает ограничения на качественный анализ при выявлении тем в корпусе документов ОГПУ. В машинном анализе текстов для определения тем дискурса распространен метод латентного размещения Дирихле (*Latent Dirichlet allocation*), который помогает выявить скрытые темы, присутствующие во всех документах массива. В программе RStudio за выполнение LDA в данном исследовании отвечают два пакета: *topicmodels* и *tidytext*. На выходе LDA предлагает перечень тем с присущими им терминами, а также доли тем в каждом отдельном документе [Kwartler 2017: 155]. После определения скрытых тем осуществляется их аннотирование (посредством коэффициента корреляции терминов, входящих в темы, с другими словами), позволяющее обобщить корпус текстов и извлечь контекст, в котором существовали документы [Kumar, Paul 2016: 124].

Всего алгоритмом LDA нами было выявлено 20 тем, преобладавших в сводках и обзорах ОГПУ по Северному Кавказу (рис. 4). Их распределение во времени позволяет сделать несколько выводов по структуре «чекистского» дискурса. Во-первых, с годами заметно расширение тематического поля дискурса: к концу исследуемого периода темы, представленные в документах, становятся разнообразнее. Во-вторых, тематически наиболее связанными являются тексты 1928–1929 гг. В-третьих, имеются интертекстуальные темы, которые объединяют документы ОГПУ во времени и выступают основой «чекистского» дискурса. Среди таковых магистральных тем выделяются пять: «местная власть», «бандитизм», «социальная напряженность», «налогообложение» и «национальные движения».

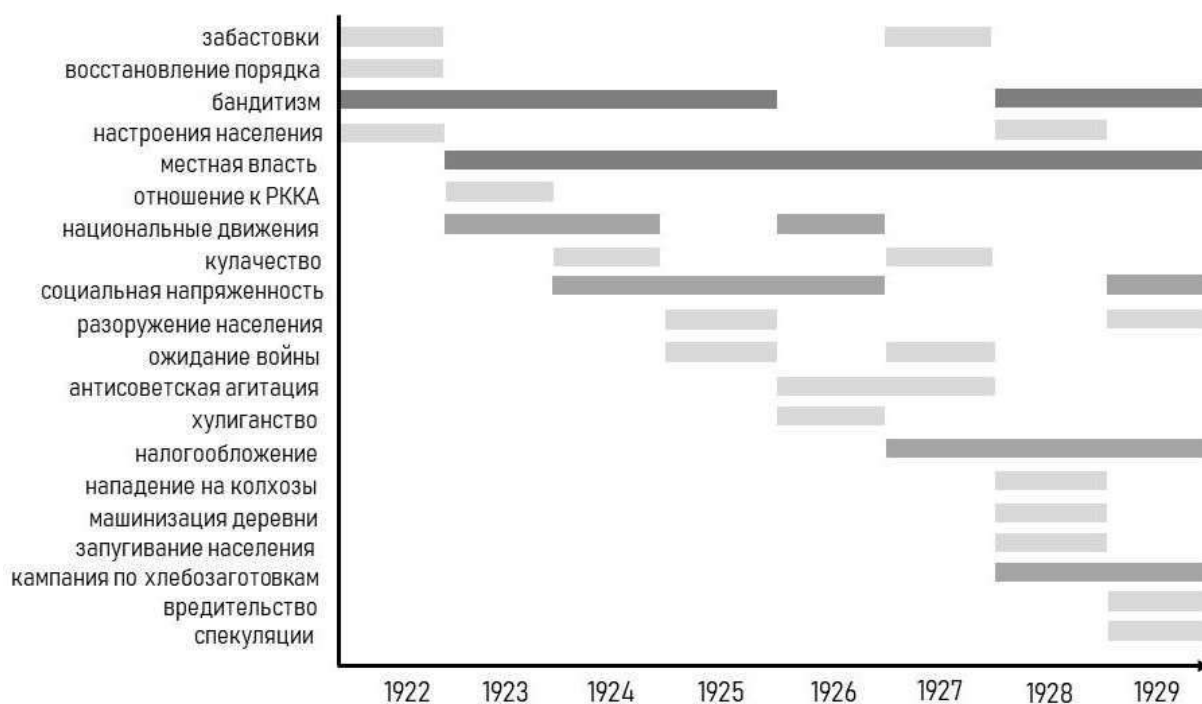


Рис. 4. Темы сводок и обзоров ОГПУ по Северному Кавказу, выявленные через LDA

Рамки статьи не позволяют представить подробное описание каждой из этих тем, поэтому иллюстративный материал дается только в первой из них, а по остальным четырем темам ограничимся текстовым изложением.

Тема «Местная власть» по среднему значению составляет 15 % объема текста ежемесячно. LDA позволяет не только определить магистральные темы дискурса, но и увидеть их динамику, а также их смысловое наполнение. На рис. 5 видна динамика изменения доли темы «Местная власть» в изучаемый период: больше всего эта тема интересовала ОГПУ в марте 1923 г., в ноябре 1925 г., в январе 1926 г., в январе-феврале 1927 г. и в октябре 1929 г. Результат LDA показывает, что данная тема генерируется из таких терминов, как «население», «рабочий», «недовольство», «местный», «власть», «работа», «член», «группа», «группировка», «сельсовет», «председатель», «перевыборы», «зажиточный», «ВКП». Кроме того, данная тема имеет ассоциации (слова в текстах, которые обладают очень высокой степенью корреляции с терминами темы, где  $r \geq 0.9$ ), к

которым относятся слова «казацкий» (1.00), «незаможник» (0.99), «бандит» (0.98), «избиратель» (0.98), «комячейка» (0.98), «политработа» (0.98), «кулак» (0.97), «средняк» (0.97), «кандидатура» (0.97), «избирсобрание» (0.96), «недоверие» (0.96), «наступление» (0.96), «эсеровский» (0.95), «волисполком» (0.94), «кандидат» (0.94), «предисполком» (0.94), «неустойчивость» (0.94). По основным терминам и ассоциациям виден главный лейтмотив темы: в составе сельсоветов присутствуют группировки кандидатов из антисоветских элементов (кулаков, казаков, середняков, бандитов, эсеров), что создает неустойчивость работы органов власти. Поэтому в течение всего рассматриваемого периода ГПУ–ОГПУ на Северном Кавказе особо интересовали перевыборы в местные органы с последующим распределением власти между социально-политическими силами в регионе: «*В стц. Поповической на перевыборах образовавшаяся антисоветская группировка – бывший врангелевец, бывший белый бандит, бывший белый офицер, райстатистик и ветфельдшер – агитировали за выбор «наших» казаков-хлеборобов, а не коммунистов*» (апрель 1925) [«Совершенно секретно» 2002а: 265]; «*В стц. Ширванской кулаки совместно с бывшими полтурядниками и атаманами пытаются организовать к перевыборам, чтобы провести в Совет “своих людей”*» (декабрь 1926) [«Совершенно секретно» 2001с: 995]; «*На перевыборное собрание кулацко-зажиточные элементы явились с заранее обсужденным списком кандидатов в состав ККОВ, коих полностью и провели, провалив список комячейки...*» (январь 1928) [«Совершенно секретно» 2002b: 94] и т. д.

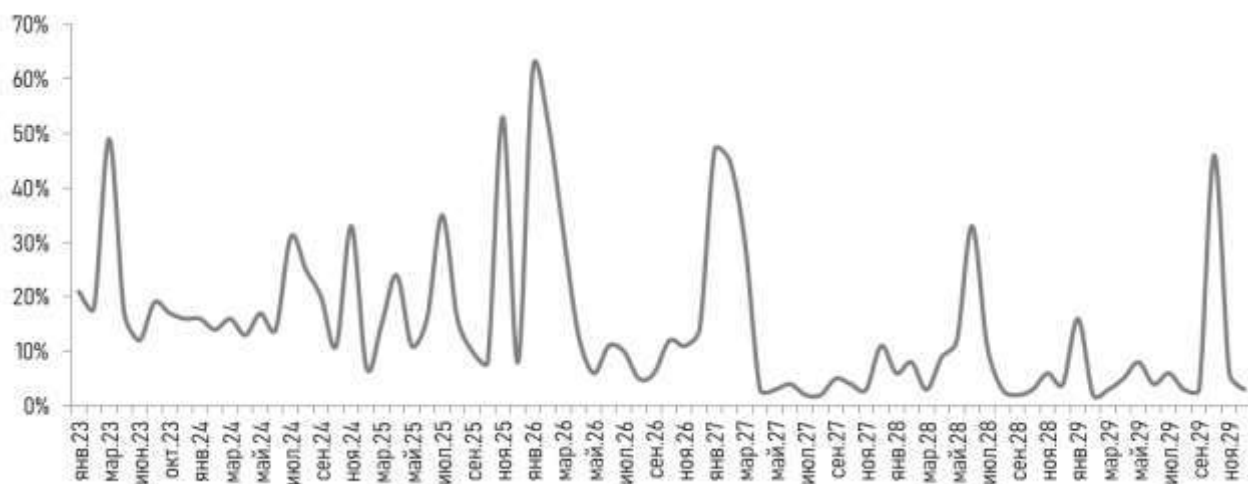


Рис. 5. Доля темы «местная власть» в каждом месяце сводок и обзоров ОГПУ

Тема «Бандитизм» в среднем составляет 18 % объема текста каждого месяца, в котором она фигурирует. Сильнее всего эта тема интересовала ОГПУ в декабре 1922 г. и в декабре 1925 г. (66 % и 62 % соответственно от содержания обзоров). Ее основу составляют такие термины, как «агитация», «антисоветский», «банда», «бандит», «выступление», «восстание», «затруднение», «кулацкий», «хлебозаготовка». К ассоциациям темы ( $r \geq 0.9$ ) относятся: «угрозыск» (0.99), «демонстрации» (0.96), «выступление» (0.95), «вредитель» (0.94), «воровской» (0.94), «де-

бош» (0.94), «уголовщина» (0.94), «разоружение» (0.94), «религиозный» (0.94), «борьба» (0.91), «безработный» (0.91), «амнистия» (0.9), «казак» (0.9), «партизане» (0.9), «теракт» (0.9). Видно, что термины и ассоциации предлагают очень пестрый портрет бандита на Северном Кавказе: под понятие бандита попадали различные группы населения (казаки, кулаки, религиозные деятели, партизаны, безработные и т. д.), которые выступали против советской власти. Таким образом, тема бандитизма в дискурсе ОГПУ приобретала разные формы (политическую, национально-политическую, уголовно-бытовую, национально-бытовую): «*На Юго-Востоке бандитизм до сего времени определенно политический, не встречая сочувствия у станичников, чрезвычайно быстро вырождается в уголовщину...*» (декабрь 1922) [«Совершенно секретно» 2001a: 488]; «*Политбандитизм на Юго-Востоке мало активен и фактически носит организационный характер (установление связи с бандитами и с населением) <...> В горских округах края по-прежнему сильно развит уголовно-бытовой и национально-политический бандитизм*» (март 1924) [«Совершенно секретно» 2001b: 70]; «*Национально-бытовой бандитизм в ряде районов продолжает сохранять устойчивость. В Кабардино-Балкарской, Адыгейско-Черкесской, Осетинской и Ингушской областях общая численность банд поднялась с ноября по декабрь с 34-х человек до 41 человека*» (декабрь 1925) [«Совершенно секретно» 2002a: 723].

Тема «Социальная напряженность» в среднем составляет 17 % от содержания ежемесячных документов, в которых она присутствует. Наибольший интерес социальная напряженность представляла в октябре 1924 г., в мае 1925 г., в мае 1926 г., в июне 1929 г. (61 %, 58 %, 53 %, 46 % соответственно от содержания обзоров). Тема формируется на следующих терминах: «русский», «деревня», «казак», «иностранец», «землеустройство», «земля», «село», «сельсовет», «толпа», «хозяйство», «убийство». Среди ассоциаций ( $r \geq 0.9$ ) присутствуют: «почва» (0.98), «конфликт» (0.98), «дагестанцы» (0.96), «безработица» (0.95), «конокрадство» (0.95), «расслоение» (0.95), «спор» (0.94), «раскол» (0.94), «расправа» (0.93), «Кубань» (0.92). Тем самым основной лейтмотив темы – конфликты между казаками, иностранцами и горским населением Северного Кавказа, возникшие на почве распределения земельных участков: «*В Дагестане массовый характер принимают споры на земельной почве между дагестанцами и городскими евреями<sup>1</sup>*» (октябрь 1924) [«Совершенно секретно» 2001b: 250]; «*В ст. Балаковской на собрании по вопросу о землеустройстве один из казаков заявил: “Мы, казаки, ждали Врангеля и ждем, и Врангель все-таки придет, а тогда мы вас всех, иностранцев, перевешаем, и земли вам все-таки не будет”*» (июнь 1925) [«Совершенно секретно» 2002a: 402]; «*Столкновение на почве пользования пастбищами имело место на границе Самурского округа Дагестана и Азербайджана*» (май 1929) [«Совершенно секретно» 2004: 256].

Тема «Налогообложение» в среднем ежемесячно занимает 15 % объема обзоров, в которых она фигурирует. Больше всего интерес ОГПУ был прикован к ней с сентября по ноябрь 1927 г. (64 %, 60 % и 62 % от содержаний обзоров соответственно). Тема генерируется из ряда терминов: «зажиточный», «Кавказ», «кампа-

---

<sup>1</sup> Авт.: так в источнике.



ния», «налог», «новый», «обложение», «сельсовет», «хозяйство». К ассоциациям темы ( $r \geq 0.9$ ) относятся: «досрочно» (0.95), «заработки» (0.95), «закон» (0.9). Тем самым нежелание зажиточного крестьянства (кулачества) на Северном Кавказе уплачивать сельхозналог – основной лейтмотив темы: *«Зажиточно-кулацкий элемент по селениям ведет агитацию против ЕСХН, указывая, что Дагправительство обложило население налогом больше, чем следует и что якобы в Терском округе налог взимается вдвое меньше»* (ноябрь 1926) [«Совершенно секретно» 2001с: 898]; *«В с. Ново-Благодаринское четверо зажиточных агитируют за неплату налога, заявляя, что тогда правительство пойдет на уступки и предоставит большие скидки»* (октябрь 1927) [«Совершенно секретно» 2003: 623]; *«20 октября в Каменно-Бродском районе со стороны кулаков и зажиточных усилились за последнее время тенденции к посылке ходоков в центр с ходатайствами о снижении сельхозналога и жалобами на местную власть, всячески оттягивая сдачу налога»* (октябрь 1928) [«Совершенно секретно» 2002b: 547].

Тема «Национальные движения» (нередко проявляющаяся в виде сообщений о конфликтах на межнациональной почве), когда она встречается в ежемесячных обзорах, в среднем занимает 18 % объема. Тема наиболее выражена в мае 1924 г. и в мае 1926 г. (46 % и 55 % соответственно). Примечательно, что к концу 1920-х гг. вопрос о национальных движениях на Северном Кавказе в текстах ОГПУ постепенно теряет свою актуальность. Данная тема состоит из совокупности терминов, среди которых – «политический», «слух», «чечен», «черкесский», «русский», «земельный», «вопрос». Ассоциациями темы ( $r \geq 0.9$ ) являются: «турецкий» (0.98), «антагонизм» (0.97), «арест» (0.97), «брожения» (0.97), «возмущение» (0.97), «контрреволюция» (0.97), «монархист» (0.96), «политбандитизм» (0.97), «политсекты» (0.97), «террор» (0.97), «спор» (0.94), «шейх» (0.91), «вооруженный» (0.9), «ликвидация» (0.9). Таким образом, ключевым лейтмотивом темы является возникновение на Северном Кавказе очагов национально-освободительных (повстанческих) сил, направленных против советской власти: *«В Чечне состоялся созванный турецким инструктором тайный съезд мулл, шейхов и влиятельных лиц по вопросу создания национального собрания»* (июнь 1923) [«Совершенно секретно» 2001а: 904]; *«Положение Д[А]ССР за отчетный период характеризуется непрекращающимся антисоветским движением, все более ярко принимающим религиозные формы. Большую активность в этой области проявляют шейхи Али Гаджи Акушинский и Магомад Али Аджи»* (август 1924) [«Совершенно секретно» 2001b: 178]; *«Движение за объединение Осетии среди национальной интеллигенции и торговых кругов Северной Осетии развивается уже давно. <...> Аналогичное, возглавляемое националистической интеллигенцией движение наблюдается в Чечне и Ингушетии (стремление к объединению обеих областей)»* (июнь 1926) [«Совершенно секретно» 2001с: 388].

### Выводы

Интеллектуальный анализ текстов сводок и обзоров ГПУ–ОГПУ показал, что «чекистский» дискурс, сфокусированный на Северном Кавказе, в течение исследуемого периода обладал единым терминологическим аппаратом с прева-

лированием постоянно усиливающегося негативного эмоционального фона. При этом нагнетание оценок настроений и сгущение красок при описании советской реальности сопровождалось обилием тем, интересовавших органы ГПУ–ОГПУ. К концу десятилетия темы становились более разнообразными и конкретными, что свидетельствует о регулярном расширении дискурса, представленного в сводках и обзорах ОГПУ, и / или об осложнении социальных отношений в региональной общности, которые фиксировала и осмысливала советская власть.

«Чекистский» дискурс в масштабах Северного Кавказа хронологически претерпел два этапа развития (1922–1925 гг. и 1926–1929 гг.), связанные между собой одними и теми же магистральными темами («местная власть», «бандитизм», «социальная напряженность», «налогообложение», «национальные движения»), которые поддерживали связь всего «чекистского» дискурса в рамках региона. Компьютерный анализ выявил, что эти «сквозные» темы зачастую не совпадают с принятой тогда структурой обзоров, что было затруднительно установить интуитивной интерпретацией чекистских текстов. Заслуживает внимания тот факт, что дискурс ОГПУ 1926–1929 гг., несмотря на высокий уровень негативного эмоционального фона, отсутствие резких перепадов настроений в сводках и обзорах, наличие магистральных тем внутренне неоднороден. Последние два года (1928–1929 гг.) рассматриваемого периода существенно отличаются от предыдущих лет по сходству словарного и тематического состава, что позволяет говорить о начале тенденции изменения дискурса в сводках и обзорах ОГПУ последующего десятилетия.

### СПИСОК ИСТОЧНИКОВ И ЛИТЕРАТУРЫ

Баранов 2013 – *Баранов А.В.* Политические настроения земледельцев казачьего Юга России в условиях «расширения» НЭПа 1924–1926 гг. (по материалам информационных сводок ОГПУ) // *Новейшая история России.* – 2013. – № 3. – С. 112–125.

Белозеров 2005 – *Белозеров В.С.* Этническая карта Северного Кавказа. – М.: ОГИ, 2005. – 298 с.

Виноградов 2001 – *Виноградов В.К.* Об особенностях информационных материалов ОГПУ как источника по истории советского общества // «Совершенно секретно»: Лубянка – Сталину о положении в стране (1922–1934 гг.). Сб. документов в 10 т. / отв. ред. А.Н. Сахаров, В.С. Христофоров. – М.: ИЦ ИРИ РАН, 2001. – Т. 1. – С. 31–73.

Всесоюзная перепись... 1928 – Всесоюзная перепись населения 1926 года. – М.: ЦСУ СССР, 1928. – Т. 9. – 228 с.

Игнатов, Михальча 2021 – *Игнатов Г., Михальча Р.* Текст майнинг. Интеллектуальный анализ текста: дизайн исследований, сбор данных и методы анализа. – Харьков: Гуманитарный Центр, 2021. – 344 с.

Капчинский 2005 – *Капчинский О.* Госбезопасность изнутри. Национальный и социальный состав. – М.: Яуза; Эксмо, 2005. – 383 с.

Кожемякин 2008 – *Кожемякин Е.А.* Дискурс-анализ как междисциплинарная методология: исторический аспект // *Научные ведомости.* – 2008. – № 15. – С. 5–12.

Колодникова 2010 – *Колодникова Л.П.* Проблемы истории советского общества 20-х годов XX столетия по рассекреченным документам архивов «Лубянки» // *Труды Института российской истории.* Вып. 9 / отв. ред. А.Н. Сахаров, ред.-коорд. Е.Н. Рудая. – М.; Тула: Гриф и К, 2010. – С. 225–238.

Кондрашин 2018 – *Кондрашин В.В.* Лубянка – Сталину о положении в сельском хозяйстве СССР в 1932–1934 гг. // Известия высших учебных заведений. Поволжский регион. – 2018. – № 1. – С. 58–69. – DOI 10.21685/2072-3024-2018-1-6.

Лубянка 2003 – Лубянка: Органы ВЧК-ОГПУ-НКВД-НКГБ-МГБ-КГБ. 1917–1991. Справочник / под ред. акад. А.Н. Яковлева; авторы-сост. А.И. Кокурин, Н.В. Петров. – М.: МФД, 2003. – 768 с.

Наухацкий 2011 – *Наухацкий В.В.* О роли органов ОГПУ–НКВД в становлении колхозной системы // Правовая политика и правовая жизнь. – 2011. – № 1. – С. 115–118.

Рожков 1995 – *Рожков А.Ю.* Первая смерть вождя: Болезнь и кончина В.И. Ленина в общественном восприятии // Россия XXI (Москва). – 1995. – № 5–6. – С. 130–157.

Рожков 2022 – *Рожков А.Ю.* Информационные документы ОГПУ о социальном самочувствии и политической активности молодежи в северокавказской деревне (1924–1929 гг.) // Проблемы и перспективы аграрной модернизации СССР–России (XX – первые десятилетия XXI вв.): сборник статей Международной научно-практической конференции. – Саратов: ФГБОУ ВО Саратовский ГАУ, ООО «ЦеСАин», 2022. – С. 148–156.

Рыбаков 2018 – *Рыбаков Р.В.* Использование приложения R в исторических исследованиях на примере анализа писем молодежи 1920-х гг. // Омский научный вестник. Серия: Общество. История. Современность. – 2018. – № 3. – С. 16–22. – DOI: 10.25206/2542-0488-2018-3-16-22.

Северный Кавказ 1925 – Северный Кавказ после районирования (итоги и выводы). В 2 т. – Ростов н/Д: Изд-во «Советский Юг», 1925. – Т. 1. – 541 с.

Скворцова 2007 – *Скворцова О.Ф.* Основные теоретико-методологические подходы к интерпретации дискурса // Антиномии. – 2007. – № 7. – С. 5–34.

«Совершенно секретно» 2001a – «Совершенно секретно»: Лубянка – Сталину о положении в стране (1922–1934 гг.). Сб. документов в 10 т. / отв. ред. А.Н. Сахаров, В.С. Христофоров. – М.: ИЦ ИРИ РАН, 2001a. – Т. 1. – 1103 с.

«Совершенно секретно» 2001b – «Совершенно секретно»: Лубянка – Сталину о положении в стране (1922–1934 гг.). Сб. документов в 10 т. / отв. ред. А.Н. Сахаров, В.С. Христофоров. – М.: ИЦ ИРИ РАН, 2001b. – Т. 2. – 511 с.

«Совершенно секретно» 2002a – «Совершенно секретно»: Лубянка – Сталину о положении в стране (1922–1934 гг.). Сб. документов в 10 т. / отв. ред. А.Н. Сахаров, В.С. Христофоров. – М.: ИЦ ИРИ РАН, 2002a. – Т. 3. – 967 с.

«Совершенно секретно» 2001c – «Совершенно секретно»: Лубянка – Сталину о положении в стране (1922–1934 гг.). Сб. документов в 10 т. / отв. ред. А.Н. Сахаров, В.С. Христофоров. – М.: ИЦ ИРИ РАН, 2001c. – Т. 4. – 1191 с.

«Совершенно секретно» 2003 – «Совершенно секретно»: Лубянка – Сталину о положении в стране (1922–1934 гг.). Сб. документов в 10 т. / отв. ред. А.Н. Сахаров, В.С. Христофоров. – М.: ИЦ ИРИ РАН, 2003. – Т. 5. – 798 с.

«Совершенно секретно» 2002b – «Совершенно секретно»: Лубянка – Сталину о положении в стране (1922–1934 гг.). Сб. документов в 10 т. / отв. ред. А.Н. Сахаров, В.С. Христофоров. – М.: ИЦ ИРИ РАН, 2002b. – Т. 6. – 795 с.

«Совершенно секретно» 2004 – «Совершенно секретно»: Лубянка – Сталину о положении в стране (1922–1934 гг.). Сб. документов в 10 т. / отв. ред. А.Н. Сахаров, В.С. Христофоров. – М.: ИЦ ИРИ РАН, 2004. – Т. 7. – 767 с.

Советская деревня 2000 – Советская деревня глазами ВЧК–ОГПУ–НКВД. 1918–1939. Документы и материалы. В 4 т. / под ред. А. Береловича, В. Данилова. – М.: РОССПЭН, 2000. – Т. 1. – 864 с.

Согомонян 2012 – *Согомонян В.* Что такое дискурс власти? // 21-й век. – 2012. – № 1. – С. 34–51.

Цуциев 2007 – *Цуциев А.А.* Атлас этнополитической истории Кавказа (1774–2004). – М.: «Европа», 2007. – 128 с.

- Jo 2019 – *Jo T.* Text Mining. Concepts, Implementation, and Big Data Challenge. – Cham: Springer, 2019. – 386 p.
- Kumar, Paul 2016 – *Kumar A., Paul A.* Mastering Text Mining with R. – Birmingham: Packt Publishing Ltd, 2016. – 237 p.
- Kwartler 2017 – *Kwartler T.* Text Mining in Practice with R. – Hoboken: John Wiley & Sons, 2017. – 307 p.
- Reisigl 2018 – *Reisigl M.* Discourse-Historical Approach // *The Routledge Handbook of Critical Discourse Studies.* – London; New York: Routledge, 2018. – P. 44-59.
- Richardson 2017 – *Richardson J.* British Fascism. A Discourse-Historical Analysis. – Stuttgart: ibidem Press, 2017. – 307 p.
- Silge, Robinson 2017 – *Silge J., Robinson D.* Text mining with R. A Tidy Approach. – Sebastopol: O'Reilly Media, 2017. – 184 p.
- Wild 2016 – *Wild F.* Learning Analytics in R with SNA, LSA, and MPIA. – Berlin: Springer, 2016. – 275 p.
- Wodak 2015 – *Wodak R.* Critical Discourse Analysis, Discourse-Historical Approach // *The International Encyclopedia of Language and Social Interaction.* – Chichester: John Wiley & Sons, Inc., 2015. – P. 275-289. – DOI:10.4135/9780857028020.d6.
- Wodak, Meyer 2001 – *Wodak R., Meyer M.* Method of Critical Discourse Analysis. – London; Thousand Oaks; New Delhi: SAGE Publications, 2001. – 200 p.

#### REFERENCES

- BARANOV A.V. *Politicheskie nastroyeniya zemledel'cev kazach'ego Juga Rossii v usloviyah «rasshirenija» NJePa 1924–1926 gg. (po materialam informacionnyh svodok OGPU).* IN: *Novejshaja istorija Rossii [The Modern History of Russia].* – 2013. – № 3. – P. 112-125. (In Russ.).
- BELOZEROV V.S. *Jetnicheskaja karta Severnogo Kavkaza [Ethnic map of the North Caucasus].* – Moscow: OGI, 2005. – 298 p. (In Russ.).
- VINOGRADOV V.K. *Ob osobennostjah informacionnyh materialov OGPU kak istochnika po istorii sovetskogo obshhestva [About the features of the OGPU information materials as source on the history of Soviet society].* IN: *«Sovershenno sekretno»: Lubyanka – Stalinu o polozhenii v strane (1922–1934 gg.) [«Top secret»: Lubyanka to Stalin on the situation in the country (1922–1934). Collection of documents in 10 vols.]*. – Moscow: IRI RAN, 2001. – Vol. 1. – Iss. 1. – P. 31-73. (In Russ.).
- Vsesojuznaja perepis'... 1928 – Vsesojuznaja perepis' naselenija 1926 goda [All - Union Population Census of 1926].* – Moscow: TsSU SSSR, 1928. – Vol 9. – 228 p. (In Russ.).
- IGNATOV G., MIKHALCHA R. *Tekst majning. Intellektual'nyj analiz teksta: dizajn issledovanij, sbor dannyh i metody analiza [Text mining: research design, data collection and analysis methods].* – Kharkov: Gumanitarnyi Tsent, 2021. – 344 p. (In Russ.).
- JO T. Text Mining. Concepts, Implementation, and Big Data Challenge. – Cham: Springer, 2019. – 386 p. (In English)
- KAPCHINSKIJ O. *Gosbezopasnost' iznutri. Nacional'nyj i social'nyj sostav [State security from the inside. National and social composition].* – Moscow: Yauza; Eksmo, 2005. – 383 p. (In Russ.).
- KOZHEMJAKIN E.A. *Diskurs-analiz kak mezhdisciplinarnaja metodologija: istoricheskij aspekt [Discourse analysis as an interdisciplinary methodology: historical aspect].* IN: *Nauchnye vedomosti [Scientific Vedomosti].* – 2008. – № 15. – P. 5-12. (In Russ.).
- KOLODNIKOVA L.P. *Problemy istorii sovetskogo obshhestva 20-h godov XX stoletija po rassekrechennym dokumentam arhivov «Lubjanki» [Problems of the History of Soviet Society in the 20s of the XXth century according to declassified documents of the Lubyanka archives].* IN: *Trudy Instituta rossijskoi istorii [Proceedings of the Institute of Russian History] / Edited by A.N. Sakharov, E.N. Rudaya.* – Moscow; Tula: Grif i K, 2010. – Vol. 9. – P. 225-238. (In Russ.).
- KONDRASHIN V.V. *Lubyanka – Stalinu o polozhenii v sel'skom hozjajstve SSSR v 1932–1934 gg. [Lubyanka to Stalin about the situation in agriculture of the USSR in 1932-1934].* IN:



*Izvestiya vysshikh uchebnykh zavedenii. Povolzhskii region* [News of higher educational institutions. Volga Region]. – 2018. – № 1. – P. 58-69. (In Russ.).

KUMAR A., PAUL A. *Mastering Text Mining with R*. – Birmingham: Packt Publishing Ltd, 2016. – 237 p. (In English)

KWARTLER T. *Text Mining in Practice with R*. – Hoboken: John Wiley & Sons, 2017. – 307 p. (In English)

*Lubjanka: Organy VChK-OGPU-NKVD-NKGB-MGB-KGB. 1917–1991. Spravochnik* [Lubyanka: Organs of the VCHK-OGPU-NKVD-NKGB-MGB-KGB. 1917-1991. Directory] / Edited by A.N. Yakovlev. – Moscow: MFD, 2003. – 768 p. (In Russ.).

NAUHACKIJ V.V. *O roli organov OGPU–NKVD v stanovlenii kolhoznoj sistemy* [On the role of the OGPU–NKVD organs in the formation of the collective farm system]. IN: *Pravovaja politika i pravovaja zhizn'* [Legal policy and legal life]. – 2011. – № 1. – P. 115-118. (In Russ.).

REISIGL M. *Discourse-Historical Approach // The Routledge Handbook of Critical Discourse Studies*. – London; New York: Routledge, 2018. – P. 44-59. (In English)

RICHARDSON J. *British Fascism. A Discourse-Historical Analysis*. – Stuttgart: ibidem Press, 2017. – 307 p. (In English).

ROZHKOVA A.JU. *Pervaja smert' vozhdja: Bolezni i konchina V.I. Lenina v obshchestvennom vospriyatii* [The first death of the leader: Illness and death of V.I. Lenin in public perception]. IN: *Rossija XXI* [Russia XXI]. – Moscow, 1995. № 5-6. – P. 130-157. (In Russ.).

ROZHKOVA A.JU. *Informacionnye dokumenty OGPU o social'nom samochuvstvii i politicheskoi aktivnosti molodezhi v severokavkazskoi derevne (1924-1929 gg.)* [Information documents of the OGPU on the social well-being and political activity of youth in the North Caucasian village (1924-1929)]. IN: *Problemy i perspektivy agrarnoi modernizatsii SSSR–Rossii (XX – pervye desjatiletiya XXI vv.): sbornik statej Mezhdunarodnoj nauchno-prakticheskoi konferencii* [Problems and prospects of agrarian modernization of the USSR–Russia (XX - the first decades of the XXI century.): collection of articles of the International scientific and practical Conference]. – Saratov: FGBOU VO Saratovskij GAU, OOO «CeSAin», 2022. – P. 148-156. (In Russ.).

RYBAKOV R.V. *Ispolzovanie prilozhenija R v istoricheskii issledovanijah na primere analiza pisem molodezhi 1920-h gg.* [The use of apps R in historical research on the example of the analysis of youth letters of the 1920s.]. IN: *Omskii nauchnyi vestnik. Seriya: Obshchestvo. Istoriya. Sovremennost* [Omsk Scientific Bulletin. Series: Society. History. Modernity]. – 2018. – № 3, P. 16-22. (In Russ.).

*Severnyj Kavkaz 1925 – Severnyj Kavkaz posle rajonirovanija (itogi i vyvody). V 2 t.* [The North Caucasus after zoning (results and conclusions). In 2 vols]. – Rostov n/D: Sovetskii Yug, 1925. – Vol. 1. – 541 p. (In Russ.).

SILGE J., ROBINSON D. *Text mining with R. A Tidy Approach*. – Sebastopol: O'Reilly Media, 2017. – 184 p. (In English).

SKVORCOVA O.F. *Osnovnye teoretiko-metodologicheskie podhody k interpretacii diskursa* [Basic theoretical and methodological approaches to the interpretation of discourse]. IN: *Antinomii* [Antinomies]. – № 7. – P. 5-34. (In Russ.).

«*Sovershenno sekretno*»: *Lubjanka – Stalinu o polozhenii v strane (1922–1934 gg.). Sb. dokumentov v 10 t.* [«Top secret»: Lubyanka to Stalin on the situation in the country (1922-1934). Collection of documents in 10 volumes] / Edited by A.N. Sakharov, V.S. Hristoforov. – M.: IC IRI RAN, 2001. – Vol. 1. – 1103 p. (In Russ.).

«*Sovershenno sekretno*»: *Lubjanka – Stalinu o polozhenii v strane (1922–1934 gg.). Sb. dokumentov v 10 t.* [«Top secret»: Lubyanka to Stalin on the situation in the country (1922-1934). Collection of documents in 10 volumes] / Edited by A.N. Sakharov, V.S. Hristoforov. – M.: IC IRI RAN, 2001. – Vol. 2. – 511 p. (In Russ.).

«*Sovershenno sekretno*»: *Lubjanka – Stalinu o polozhenii v strane (1922–1934 gg.). Sb. dokumentov v 10 t.* [«Top secret»: Lubyanka to Stalin on the situation in the country (1922-1934). Collection of documents in 10 volumes] / Edited by A.N. Sakharov, V.S. Hristoforov. – M.: IC IRI RAN, 2002. – Vol. 3. – 967 p. (In Russ.).



«Sovershenno sekretno»: Lubyanka – Stalinu o polozhenii v strane (1922–1934 gg.). *Sb. dokumentov v 10 t.* [«Top secret»: Lubyanka to Stalin on the situation in the country (1922-1934). Collection of documents in 10 volumes] / Edited by A.N. Sakharov, V.S. Hristoforov. – M.: IC IRI RAN, 2001. – Vol. 4. – 1191 p. (In Russ.).

«Sovershenno sekretno»: Lubyanka – Stalinu o polozhenii v strane (1922–1934 gg.). *Sb. dokumentov v 10 t.* [«Top secret»: Lubyanka to Stalin on the situation in the country (1922-1934). Collection of documents in 10 volumes] / Edited by A.N. Sakharov, V.S. Hristoforov. – M.: IC IRI RAN, 2003. – Vol. 5. – 798 p. (In Russ.).

«Sovershenno sekretno»: Lubyanka – Stalinu o polozhenii v strane (1922–1934 gg.). *Sb. dokumentov v 10 t.* [«Top secret»: Lubyanka to Stalin on the situation in the country (1922-1934). Collection of documents in 10 volumes] / Edited by A.N. Sakharov, V.S. Hristoforov. – M.: IC IRI RAN, 2002. – Vol. 6. – 795 p. (In Russ.).

«Sovershenno sekretno»: Lubyanka – Stalinu o polozhenii v strane (1922–1934 gg.). *Sb. dokumentov v 10 t.* [«Top secret»: Lubyanka to Stalin on the situation in the country (1922-1934). Collection of documents in 10 volumes] / Edited by A.N. Sakharov, V.S. Hristoforov. – M.: IC IRI RAN, 2004. – Vol. 7. – 767 p. (In Russ.).

*Sovetskaja derevnja 2000 – Sovetskaja derevnja glazami VChK–OGPU–NKVD. 1918–1939. Dokumenty i materialy. V 4 t.* [Soviet Village 2000 – Soviet village through the eyes of the VChK–OGPU–NKVD. 1918–1939. Documents and materials. In 4 vol.] / Edited by A. Berelovich, V. Danilov. – Moscow: ROSSPJeN, 2000. – Vol. 1. – 864 p. (In Russ.).

SOGHOMONYAN V. *Chto takoe diskurs vlasti?* [What is the discourse of power?] *IN: 21 vek* [21st century]. – 2012. – №. 1. – P. 34-51. (In Russ.).

CUCIEV A.A. *Atlas jetnopoliticheskoi istorii Kavkaza (1774–2004)* [Atlas of the Ethnopolitical History of the Caucasus (1774-2004)]. – Moscow: «Europe», 2007. – 128 p. (In Russ.).

WILD F. *Learning Analytics in R with SNA, LSA, and MPIA.* – Berlin: Springer, 2016. – 275 p. (In English).

WODAK R. *Critical Discourse Analysis, Discourse-Historical Approach* // *The International Encyclopedia of Language and Social Interaction.* – Chichester: John Wiley & Sons, Inc., 2015. – P. 275-289. (In English).

WODAK R., MEYER M. *Method of Critical Discourse Analysis.* – London; Thousand Oaks; New Delhi: SAGE Publications, 2001. – 200 p. (In English).

### **Информация об авторах**

А.Ю. Рожков – доктор исторических наук, профессор.

Т.А. Рунаев – кандидат социологических наук.

### **Information about the authors**

A.Yu. Rozhkov – Doctor of Sciences (History), Professor.

T.A. Runaev – Candidate of Sciences (Sociology).

**Вклад авторов:** все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

**Contribution of the authors:** the authors contributed equally to this article. The authors declare no conflicts of interests.

Статья поступила в редакцию 22.09.2022 г.; одобрена после рецензирования 20.10.2022 г.; принята к публикации 15.12.2022 г.

The article was submitted 22.09.2022; approved after reviewing 20.10.2022; accepted for publication 15.12.2022.