

DOI 10.35775/PSI.2025.118.6.012

УДК 323.3

В.В. КОЗЕНЮК

аспирант кафедры политологии и этнополитики
Южно-Российского института управления – филиал
Российской академии народного хозяйства и государственной
службы при Президенте РФ, Россия, г. Ростов-на-Дону
E-mail: kozenyukv@mail.ru
<https://orcid.org/0009-0001-1767-2140>

ЦИФРОВЫЕ ИНСТРУМЕНТЫ В ИССЛЕДОВАНИЯХ ПОЛИТИЧЕСКИХ ЭЛИТ: ОЦЕНКА ЭФФЕКТИВНОСТИ И СОПОСТАВЛЕНИЕ С ТРАДИЦИОННЫМИ ПОДХОДАМИ

Введение. Современные исследования политических элит сталкиваются с необходимостью адаптации методологического аппарата к цифровой трансформации общества. Актуальность работы обусловлена растущим влиянием цифровых инструментов на политические процессы и отсутствием комплексных методик их интеграции в традиционные исследования элит. Автор подчеркивает разрыв между классическими подходами (структурно-функциональный анализ, биографические методы) и возможностями, которые предоставляют Big Data, машинное обучение и сетевая аналитика.

Цель. Цель исследования заключается в разработке алгоритма интеграции цифровых инструментов (на примере веб-скрейпинга) в изучение политической элиты, обеспечивающую повышение точности данных, автоматизацию рутинных задач и выявление скрытых паттернов власти. Работа направлена на преодоление ограничений традиционных методов за счет синтеза качественного анализа с технологиями обработки больших данных.

Методы. Исследование выполнено с применением междисциплинарной методологии. Были использованы методы системного и сравнительного анализов, технологий нейросетей для написания программного кода на языке программирования Python и технологии веб-скрейпинга на Python и R.

Результаты. Обозначены правовые, технические, этические ограничения при использовании технологии веб-скрейпинга, а также предложен поэтапный алгоритм автоматизированного сбора данных с помощью Python и R, учитывающий существующие ограничения в рамках научно-исследовательской деятельности.

Выводы. Интеграция цифровых инструментов значительно расширяет возможности исследования политических элит, обеспечивая многомерный анализ и прогнозирование динамики власти. Ключевые преимущества: скорость обработки данных, выявление латентных связей, минимизация субъективности. Однако сохраняются риски, связанные с этикой данных и алгоритмическими искажениями. Автор рекомендует

развивать междисциплинарные коллаборации и стандартизировать протоколы работы с цифровыми источниками.

Ключевые слова: политическая элита, цифровые инструменты.

Введение. За последнее десятилетие общество претерпело множество цифровых трансформаций, которые укоренились в повседневной жизни. Прежде всего, феномен цифровизации заключается в непрерывности, а следовательно, в недалеком будущем использование технологий станет неотъемлемым элементом существования любой системы.

В целом для успешного и эффективного функционирования любой современной системы внедрение информационно-коммуникационных технологий является обязательным. Использование цифровых технологий в науке позволяет работать с невиданными ранее масштабами данных, повышая точность и скорость исследований. Например, на данный момент существует значительное количество программного обеспечения, которое может решать широкий перечень задач: анализ данных и статистика (Python и R, SPSS, European Open Science Cloud), облачные технологии для совместной работы над проектами и исследованиями (YandexCloud, Google Drive, Dropbox, Trello, Slack), различные корпусные технологии (Text Encoding Initiative), научные сети и базы данных (ResearchGate, Open Access Journal Finder, SciFinder, Web of Science) и т.п.

Несмотря на внушительный список программного обеспечения, преобладающая часть цифровых инструментов заблокирована или ограничена в Российской Федерации ввиду санкционной политики Запада или нарушения российского законодательства.

Применение цифровых инструментов в политологических исследованиях имеет под собой множество преимуществ, которые в несколько раз тяжелее достичь при использовании аналоговых методов, несмотря на количество различных цифровых технологий с их применением возникают и некоторые вызовы: множество данных относятся к категории закрытых; необходимо обеспечение обновления инструментов ввиду динамичности политических процессов; интерпретация данных должна сочетать в себе количественные методы с качественным контекстом.

Несколькими годами ранее И.А. Быков в своем исследовании отмечал сильную преемственность в изучении искусственного интеллекта в политике [2. С. 23-33]. Внедрение цифровых технологий в контексте исследования политической элиты привносят значительный вклад в научный труд в междисциплинарном плане. Например, в своей работе П.К. Победин приводит несколько примеров применения цифровых инструментов в Индии и США: в 2020 году в Индии движения «Партнерство за открытость» использовали цифровые технологии для расследования коррупции в правительстве, по итогам которого было выявлено несколько случаев коррупции, а уже в 2021 году в США активисты движения «Прозрачного правительства» использовали цифровые инструменты визуализации данных для создания интерактивных карт и графиков, что позволило

населению получить полное представление о целевых расходах бюджетных средств [12]. Данные примеры наглядно показывают, какую роль играют информационные технологии в современном обществе и как они способны воздействовать на политические элиты.

Среди цифровых инструментов для исследования политической элиты можно выделить следующие:

- Bright Data, ScrapperAPI, Octoparse, Apify, Scrapy – для автоматизированного сбора данных;

- Tableau, Power BI, TimelineJS, QGIS, Yandex DataLens, Visiology – для анализа и визуализации данных;

- а также другие инструменты анализа текстов и дискурса, сетевого анализа, карьерного трекинга и биографического анализа, прогнозирования.

Применение цифровых технологий для анализа интернет-данных, по мнению Д. Фаррелла и Дж. Петерсена, на сегодняшний день являются проблематичными ввиду возможного нарушения приватности данных [19]. Однако, по мнению О.В. Вилковой применение цифровых инструментов, например, веб-скрейпинга является законным при соблюдении установленных российским законодательством ограничений: не нарушаются авторские права, не допускается сбор сведений, составляющих коммерческую тайну; не допускается ограничение конкуренции и т.д. [3].

В контексте элитологических и политических исследований применение такого инструмента, как веб-скрейпинг способно обеспечить структурирование и систематизацию большого объема данных автоматизированным способом, что позволит сократить время на сбор информации и уделить больше времени ее анализу.

Под «веб-скрейпингом» принято понимать автоматизированный процесс сбора данных с веб-страниц, который включает извлечение информации из различных источников в сети Интернет и ее сохранение в структурированном формате. Некоторые отождествляют данное определение с термином «парсинг». Однако данные определения не являются равнозначными. Под «парсингом» понимается часть процесса «веб-скрейпинга», связанная с обработкой и структурированием данных. Простыми словами «парсинг» предусматривает анализ формата исходных данных (HTML, XML, JSON) и извлечение из них конкретной информации (текста, чисел, ссылок).

Актуальность исследования прежде всего обуславливается наличием общедоступных цифровых технологий, которые открывают новые перспективы в области автоматизированного анализа большого количества данных в кратчайшие сроки, что может положительно влиять на количественные и качественные исследования в политической науке. В своей работе И.А. Лавров также отмечает, что «современная цифровая эпоха открывает перед исследователями новые возможности, т.к. политические элиты стало возможным изучать без непосредственного контакта с ее представителями, а также большое количество

информации о ней находится в открытом доступе в сети Интернет, различных интервью, социальных сетях и т.д. [8. С. 173-179].

Совмещая современные цифровые инструменты с уже сформировавшимися методологическими подходами, можно достичь значительной глубины в исследованиях политической элиты благодаря способности быстрой обработки большого количества данных по сравнению с аналоговыми методами. Например, симбиоз технологии веб-скрейпинга и библиографического метода, описанного в работе Л.Ю. Логуновой, позволит в рамках исследования представителей политической элиты структурировать большие объемы данных о карьерных траекториях, образовании и ключевых событиях политических акторов в кратчайшие сроки [9. С. 17-23].

Также следует обратить внимание на исследование М.Г. Бреслера, где проводится парсинг сети для выявления структуры коммуникаций сетевого общества, которое выделяет коммуникации между акторами сетевого общества [1]. Стоит заметить, что в современных условиях с применением цифровых инструментов появляется реальная возможность проведения аналогичного анализа в исследованиях коммуникативных связей политической элиты с последующей визуализацией полученных данных.

Следует подчеркнуть, что в работах российских авторов, опубликованных в последние годы, освещается широкий спектр вопросов близких к данной предметной области [4; 5; 6; 7; 10; 11; 13; 14; 15; 16; 17].

Однако проблему интеграции цифровых инструментов для расширения возможностей исследования политических элит нельзя назвать однозначно исчерпанной. В силу многих объективных обстоятельств изучение обозначенной темы продолжает сохранять высокий уровень актуальности.

Цель исследования заключается в выявлении преимуществ, недостатков и ограничений автоматизированного сбора интернет-данных с помощью технологий веб-скрейпинга в контексте исследования политической элиты, а также в разработке алгоритма анализа данных с помощью веб-скрейпинга и нейросетей.

Материалы и методы. Исследование выполнено с применением междисциплинарной методологии. Были использованы методы системного и сравнительного анализов, технологий нейросетей для написания программного кода на языке программирования Python и технологии веб-скрейпинга на Python и R.

Результаты и обсуждения. При применении «веб-скрейпинга» существуют определенные ограничения на его использование, которые можно подразделить на следующие группы: правовые, технические, этические.

Среди правовых ограничений в сфере применения автоматизированного сбора данных можно выделить:

1. Копирование текстов, изображений, видео и других объектов авторского права без разрешения правообладателя запрещено (ст. 1270 ГК РФ) (1. Ст. 1270).

Например, извлечение фотографий с логотипами или уникального контента может привести к судебным искам.

2. Если сайт признается базой данных, созданной с существенными затратами, извлечение всей или существенной части ее содержимого без согласия правообладателя нарушает ст. 1334 ГК РФ (1. П. 3. ст. 1334). Даже многократное извлечение незначительных частей данных может быть незаконным, если это противоречит нормальному использованию базы (п. 3 ст. 1335.1 ГК РФ) (1. П. 3. ст. 1335.1).

3. Сбор и обработка персональных данных (например, имен, email-адресов, номеров телефонов) без согласия субъекта данных запрещены законом «О персональных данных» (№ 152-ФЗ) (3). Это особенно актуально при скрейпинге социальных сетей или форумов.

4. Нарушение правил, указанных в пользовательских соглашениях сайтов, может трактоваться как неправомерное действие. Например, пользовательское соглашение интернет-ресурса явно запрещает автоматизированный сбор данных и скрейпинг становится основанием для судебных исков, даже если данные общедоступны. В деле Avito против Яндекс суд учитывал соблюдение условий размещения объявлений.

5. Обход технических средств защиты (например, CAPTCHA, блокировок IP) может квалифицироваться как «несанкционированный доступ к компьютерной информации» по ст. 272 УК РФ (2. Ст. 10). Это актуально, если скрейпинг приводит к повреждению систем или нарушению работы сайта.

Среди технических ограничений в сфере применения автоматизированного сбора данных можно выделить:

1. Этические и технические нормы, изложенные в файле robots.txt (например, <https://yandex.ru/robots.txt>). Хотя несоблюдение правил, указанных в файле robots.txt, не влечет юридической ответственности, суды могут рассматривать это как недобросовестное поведение.

2. Интенсивные запросы, ведущие к замедлению работы сайта, могут трактоваться как злоупотребление ресурсами, что подпадает под ст. 10 ГК РФ о запрете злоупотребления правом (1. Ст. 10).

Среди этических ограничений в сфере применения автоматизированного сбора данных можно выделить вопросы использования данных без явного согласия, которые впоследствии могут вызвать споры и разбирательства.

На сегодняшний день создание программного кода не составляет труда в связи с развитием технологий нейросетей. Однако исследователь, который поставил задачу проведения автоматизированного сбора данных, должен обладать определенными компетенциями в области информационных технологий и уверено пользоваться профессиональным компьютером. Для того чтобы автоматизированный сбор данных произошел успешно, необходимо следовать определенному алгоритму, который учитывает ограничения, связанные с таким сбором данных (рисунок 1).

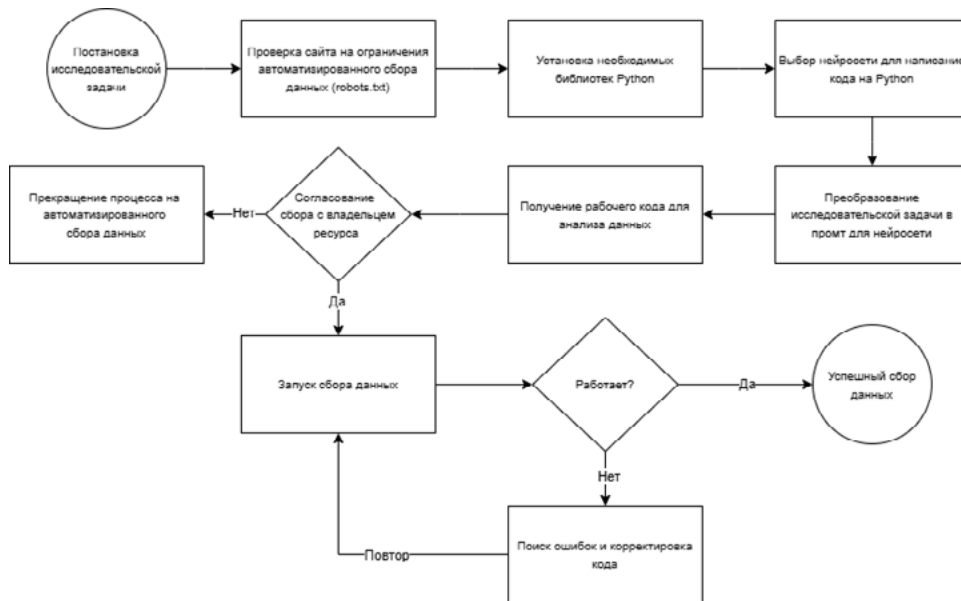


Рис. 1. Алгоритм веб-скрейпинга с помощью Python

После постановки исследовательской задачи необходимо провести предварительную подготовку для того, чтобы сбор данных прошел успешно:

1. Вначале необходимо удостовериться, что интернет-ресурс, с которого предполагается автоматизированный сбор данных не против подобного метода сбора. Для этого необходимо обратиться к стандарту исключений для роботов непосредственно на интернет-ресурсе (например, <http://duma.gov.ru/robots.txt>). Есть два варианта вывода информации: отсутствие такого файла (следовательно, ограничения к автоматизированному сбору не предъявлены) или вывод на экран ограничений автоматизированного сбора данных к конкретным разделам, например:

```

User-agent: *
Host: дума.gov.ru
Disallow: /search/
Disallow: /systems/law/?name=
Disallow: /analytics/tv/
    
```

Исходя из представленного текстового файла robots.txt можно заключить, что категорически запрещен (Disallow) автоматизированный сбор трех разделов, но этот факт не отменяет соблюдение других ограничений, установленных российским законодательством.

2. Для корректной работы программного кода на языке программирования Python требуется установка дополнительных библиотек (Beautiful Soup, Scrapy,

Selenium, Requests и др.). Установка библиотек производится через командную строку операционной системы.

3. На данный момент существует множество нейросетей (YandexGPT, GigaChat, GigaCode и др.) в открытом доступе, которые в первую очередь можно подразделить на платные и бесплатные. Почти каждая нейросеть вне зависимости от тарифного плана способна справиться с написанием программного кода для веб-скрейпинга. Иногда может потребоваться и корректировка кода, однако и здесь нейросеть подскажет, в чем может заключаться ошибка и автоматически ее исправит.

4. Качество программного кода будет зависеть в том числе от того, как будет интерпретирована исследовательская задача. Именно здесь нужно четко и лаконично сформулировать запрос для нейросети.

5. В настоящее время Федеральным законом «О персональных данных» запрещается сбор общедоступных персональных данных и для такого сбора потребуется личное согласие на обработку персональных данных. Однако ст. 18 и ст. 6 Федерального закона «О персональных данных» допускают обработку данных для исследовательских целей, в том числе научной деятельности, если при этом не нарушаются права и законные интересы субъектов персональных данных (данные должны быть обезличены, не использоваться в целях продвижения товаров, работ, услуг, а также политической агитации). В целях по недопущению возможных конфликтных ситуаций, следует заблаговременно связаться с владельцем интернет-ресурса, обозначить предполагаемые цели и задачи исследования, а также предоставить программный код для анализа его работы, который не причинит вред техническому оборудованию интернет-ресурса.

6. Если владелец ресурса против автоматизированного сбора данных, то следует поискать альтернативу или отказаться от данного метода, чтобы избежать возможного конфликта.

7. Если владелец ресурса выражает согласие на автоматизированный сбор данных, то можно приступать непосредственно к процессу веб-скрейпинга.

8. В случае некорректного вывода данных необходимо скорректировать программный код с помощью нейросети и повторить процедуру веб-скрейпинга.

9. Если выгрузка данных прошла успешно, то можно считать процесс «веб-скрейпинга» завершенным и переходить к следующим этапам исследования.

Веб-скрейпинг имеет свои преимущества и недостатки, как и любой другой метод или технология. Среди преимуществ можно выделить следующее:

- высокая скорость и точность обработки большого количества данных;
- автоматическое структурирование и вывод информации в виде электронных таблиц;
- возможность обработки больших объемов данных из множества источников одновременно;

- невысокие финансовые затраты по сравнению с оплатой труда людей, покупкой лицензионных баз данных, а также платными API;

- постоянная актуальность информации при сборе динамичных данных.

Среди недостатков можно выделить следующее:

- правовые ограничения на автоматический сбор данных (законодательство, авторские права, правила использования сайтов и др.);

- технические сложности в виде различной защиты сайтов, что может усложнять сбор данных;

- высокий риск блокировки при обнаружении подозрительной активности;

- возможны неточности при обработке неструктурированных или плохо оформленных данных.

Сравнительная характеристика веб-скрейпинга по сравнению с традиционными методами сбора данных:

- API предоставляют структурированные данные, но часто платные или с ограниченным функционалом;

- ручной сбор является достаточно точным, но неэффективным для обработки большого объема данных;

- покупка баз данных, быстрый способ, но информация может быть устаревшей или дорогой.

Преимущества применения современных цифровых технологий в исследованиях политической элиты однозначны, особенно при включении их вместе с аналоговыми методами и существующими методологическими подходами, что также подчеркивает М. Хэнд в своем исследовании [18. Р. 205-229].

Заключение. Несмотря на преимущества применения цифровых инструментов в элитологических и политических исследованиях становится вопрос об их интеграции совместно с существующими методами и подходами. На начальных этапах необходимо создание инфраструктуры в образовательных учреждениях и научно-исследовательских институтах для проведения дальнейших исследований. Таким образом, можно выделить следующие элементы, которые могут способствовать успешной интеграции таких технологий:

- замещение зарубежных цифровых инструментов отечественными аналогами для создания собственной бесперебойной технической инфраструктуры, что актуально ввиду современных геополитических конфликтов и санкций в отношении России;

- повышения уровня профессиональных компетенций сотрудников в области цифровых технологий и компьютерной грамотности (внедрение курсов по работе с цифровыми инструментами и т.п.);

- разработка единой официальной методики, которая учитывала бы все технические и теоретические аспекты в рамках применения цифровых инструментов в научно-исследовательской сфере;

- создание облачной платформы для научных коллабораций и совместной работы исследовательского сообщества;

- разработка протоколов для работы с персональными данными для обеспечения конфиденциальности и этической составляющей;

Внедрение цифровых технологий требует критического подхода к этике, прозрачности методологии и осознания ограничений. Будущее политической науки связано с интеграцией цифровых инструментов, междисциплинарным сотрудничеством и разработкой регуляторных стандартов для ответственного использования технологий. Именно поэтому необходимо внедрения цифровых инструментов в образовательные учреждения и научно-исследовательские институты уже в настоящее время.

ПРИМЕЧАНИЯ:

- (1) Гражданский кодекс Российской Федерации (часть четвертая) от 30.11.1994 № 51-ФЗ (ред. от 22.07.2024) // Собрание законодательства Российской Федерации. 1994. № 32.
- (2) Уголовный кодекс Российской Федерации от 13.06.1996 № 63-ФЗ (ред. от 28.02.2025).
- (3) Федеральный закон «О персональных данных» от 27.07.2006 № 152-ФЗ (ред. от 08.08.2024).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК:

1. **Бреслер М.Г.** Сетевой принцип формирования элит в информационном/цифровом обществе // Вопросы элитологии. 2024. № 2. DOI: 10.46539/elit.v5i2.192.
2. **Быков И.А.** Искусственный интеллект как источник политических суждений // Журнал политических исследований. 2020. Т. 4. № 2. DOI: 10.12737/2587-6295-2020-23-33. EDN FCGCZO.
3. **Вилкова О.В.** К вопросу о научной осмысленности применения веб-скрейпинга как метода сбора данных в социологических исследованиях // Вестн. Том. гос. ун-та. Философия. Социология. Политология. 2020. № 54. DOI: 10.17223/1998863X/54/16.
4. **Дронов А.И.** Политическая коммуникация выборных элит с населением посредством социальных медиа при принятии управленческих решений // Вопросы политологии. 2025. № 4.
5. **Исмаилова А.С., Голубева О.А., Григорян Д.К.** Проблемы электронного правительства в России // Евразийский Союз: вопросы международных отношений. 2025. № 1.
6. **Калакутский А.В.** Информационные технологии и их роль в политическом управлении // Вопросы политологии. 2024. № 5.
7. **Каримов Э.Р.** Цифровые платформы и публичная политика: анализ роли и влияния // Вопросы национальных и федеративных отношений. 2024. № 9.

8. **Лавров И.А.** Цифровая социология и современные методы изучения политической элиты // Вестник университета. 2019. № 6. DOI: 10.26425/1816-4277-2019-6-173-179.
9. **Логунова Л.Ю.** Биографический метод в исследовании личности: методология и архитектоника // Вестник Кемеровского государственного университета. Серия: Политические, социологические и экономические науки. 2016. № 1. DOI: <http://doi.org/10.21603/2500-3372-2016-1-17-23>.
10. **Медведев Н.П., Слизовский Д.Е.** Современные политические элиты: о векторе направленности российских исследований // Вопросы национальных и федеративных отношений. 2024. Т. 14. № 7 (112).
11. **Новиков А.А., Плужникова Н.Н.** Имидж современного политического лидера и социально-коммуникативные технологии его конструирования // Вопросы политологии. 2025. № 2.
12. **Победин П.К.** Формирование политических институтов и процессов с помощью цифровых технологий и искусственного интеллекта // Постсоветский материк. 2024. № 2 (42). DOI: 10.48137/23116412_2024_2_51.
13. **Поярко Р.А.** К вопросу о развитии цифровых компетенций государственных гражданских служащих Российской Федерации // Вопросы политологии. 2024. № 5.
14. **Старостин А.М., Понеделков А.В.** Цивилизационный формат репрезентации современных российских элит // Вопросы национальных и федеративных отношений. 2024. № 2.
15. **Степанов К.В.** Негативно-девиантные и позитивно-девиантные политические элиты как эпифеномены развития и трансформации современного российского общества // Вопросы политологии. 2024. № 10.
16. **Чемшит Д.А.** Предпосылки модернизационных кризисов политических систем в процессах рекрутирования элит // Вопросы политологии. 2024. № 9.
17. **Яценко А.С., Майстат М.А.** Влияние цифровизации на эффективность государственного управления: мировой опыт // Евразийский Союз: вопросы международных отношений. 2025. № 1.
18. **Hand M.** Digitization and memory: Researching practices of adaption to visual and textual data in everyday life // Big data? Qualitative approaches to digital research. In M. Hand & S. Hillyard (Eds.). Studies in Qualitative Methodology. Bingley, UK: Emerald Group Publishing Limited. 2014. Vol. 13. DOI: 10.1108/S1042-319220140000013013.
19. **Petersen J., Farrell D.** Online Research Methods // The Blackwell Encyclopedia of Sociology. New York: John Wiley & Sons, 2016. DOI: 10.1002/9781405165518.wbeos0628.

V.V. KOZENYUK

Postgraduate student of the Department of Political Science and Ethnopolitics, South-Russian Institute of Management of Russian Presidential Academy of National Economy and Public Administration, Rostov-on-Don, Russia
<https://orcid.org/0009-0001-1767-2140>

DIGITAL TOOLS IN THE RESEARCH OF POLITICAL ELITES: EVALUATION OF EFFECTIVENESS AND COMPARISON WITH TRADITIONAL APPROACHES

Introduction. Modern studies of political elites are faced with the need to adapt the methodological apparatus to the digital transformation of society. The relevance of the work is due to the growing influence of digital tools on political processes and the lack of comprehensive methods for their integration into traditional elite studies. The author highlights the gap between classical approaches (structural and functional analysis, biographical methods) and the opportunities offered by Big Data, machine learning, and network analytics.

Purpose. The purpose of the research is to develop an algorithm for integrating digital tools (using the example of web scraping) into the study of the political elite, which improves data accuracy, automates routine tasks and identifies hidden patterns of power. The work is aimed at overcoming the limitations of traditional methods by synthesizing qualitative analysis with big data processing technologies.

Methods. The research was carried out using an interdisciplinary methodology. Methods of system and comparative analysis, neural network technologies for writing software code in the Python programming language, and web scraping technologies in Python and R.

Results. The legal, technical, and ethical limitations of using web scraping technology are outlined, and a step-by-step algorithm for automated data collection using Python and R is proposed, taking into account existing limitations in the framework of research activities.

Conclusions. The integration of digital tools significantly expands the possibilities of studying political elites, providing multidimensional analysis and forecasting of power dynamics. Key advantages: data processing speed, identification of latent connections, minimization of subjectivity. However, risks associated with data ethics and algorithmic distortions remain. The author recommends developing interdisciplinary collaborations and standardizing protocols for working with digital sources.

Key words: political elite, digital tools.