

Научная статья

УДК 004.41

doi: 10.34822/1999-7604-2022-3-39-45

АВТОМАТИЗИРОВАННЫЙ СБОР ДАННЫХ КЛЮЧЕВЫХ ФИНАНСОВЫХ ПОКАЗАТЕЛЕЙ ПРЕДПРИЯТИЙ IT-ОТРАСЛИ РЕГИОНА

**Сергей Олегович Крамаров^{1✉}, Виктор Андреевич Овсянников²,
Людмила Викторовна Сахарова³, Роман Сергеевич Усатый⁴,
Галина Викторовна Лукьянова⁵**

¹ Южный федеральный университет, Ростов-на-Дону, Россия

² Национальный исследовательский университет «Высшая школа экономики»,
Москва, Россия

^{3,4,5} Ростовский государственный экономический университет (РИНХ),
Ростов-на-Дону, Россия

¹ maoovo@yandex.ru ✉, <https://orcid.org/0000-0003-3743-6513>

² viktordee41@gmail.com, <https://orcid.org/0000-0003-1824-9536>

³ L_Sakharova@mail.ru, <https://orcid.org/0000-0002-4897-4926>

⁴ rs.usaty@gmail.com, <https://orcid.org/0000-0002-0993-7575>

⁵ lukyanova.g@yandex.ru, <https://orcid.org/0000-0001-6674-2240>

Аннотация. Рассматривается новое программное обеспечение для автоматизированного сбора данных, позволяющее на основе парсинга открытых сайтов за короткое время получать в Excel-формате данные о наиболее значимых финансово-экономических показателях фирм конкретной отрасли, что позволяет провести полноценный финансово-экономический анализ выбранной отрасли в данном регионе. Важным условием для сбора данных является наличие списка необходимых для изучения фирм на официальном сайте export-base.ru. В качестве основного инструмента написания программы был выбран язык программирования Python и ключевые библиотеки, такие как *request*, *selenium*, *beautifulsoup4* и *csv*. Рассматриваемое в работе ПО апробировано на примере быстрого получения комплексных данных IT-компаний Ростовской области для проведения финансово-экономического анализа IT-отрасли региона.

Ключевые слова: парсинг, скрейпинг, Python, автоматизированный сбор информации

Для цитирования: Крамаров С. О., Овсянников В. А., Сахарова Л. В., Усатый Р. С., Лукьянова Г. В. Автоматизированный сбор данных ключевых финансовых показателей предприятий IT-отрасли региона // Вестник кибернетики. 2022. № 3 (47). С. 39–45. DOI 10.34822/1999-7604-2022-3-39-45.

Original article

AUTOMATED DATA COLLECTION OF KEY FINANCIAL INDICATORS OF IT ENTERPRISES IN THE REGION

**Sergey O. Kramarov^{1✉}, Viktor A. Ovsyannikov², Lyudmila V. Sakharova³,
Roman S. Usaty⁴, Galina V. Lukyanova⁵**

¹ Southern Federal University, Rostov-on-Don, Russia

² Higher School of Economics, Moscow, Russia

^{3,4,5} Rostov State University of Economics, Rostov-on-Don, Russia

¹ maoovo@yandex.ru ✉, <https://orcid.org/0000-0003-3743-6513>

² viktordee41@gmail.com, <https://orcid.org/0000-0003-1824-9536>

³ L_Sakharova@mail.ru, <https://orcid.org/0000-0002-4897-4926>

⁴ rs.usaty@gmail.com, <https://orcid.org/0000-0002-0993-7575>

⁵ lukyanova.g@yandex.ru, <https://orcid.org/0000-0001-6674-2240>

Abstract. The article deals with a new software for automated data collection. Using parsing of open websites, this software allows obtaining a set of the most significant financial and economic indicators of firms for a particular industry in the region in a short time and presents it in an Excel document, resulting in a comprehensive financial and economic analysis of a given industry in the region. A list of required firms published at export-base.ru is a necessary condition for data collection. Python and such key libraries as *request*, *selenium*, *beautifulsoup4*, and *csv* are selected as the main tool for writing the program. The software developed is tested for quickly obtaining complex data from IT companies in the Rostov Oblast to conduct the financial and economic analysis of the IT industry in the region.

Keywords: parsing, scraping, Python, automated data collection

For citation: Kramarov S. O., Ovsyannikov V. A., Sakharova L. V., Usatyi R. S., Lukyanova G. V. *Automated Data Collection of Key Financial Indicators of IT Enterprises in the Region // Proceedings in Cybernetics. 2022. No. 3 (47). P. 39–45. DOI 10.34822/1999-7604-2022-3-39-45.*

ВВЕДЕНИЕ

Исследование проводилось с целью разработки универсальной методологии сбора данных из открытых источников, предназначенных для финансово-экономического анализа конкретной отрасли в регионе на основе открытых данных показателей предприятий.

Несмотря на наличие большого количества программного обеспечения (далее – ПО), с помощью которого возможно проведение финансово-экономического анализа отдельных предприятий, отсутствуют программы, осуществляющие полный анализ конкретной отрасли. Препятствием для создания подобного ПО, помимо проблем с методиками расчета интегральных показателей, является то, что данные о фирмах должны удовлетворять нескольким требованиям одновременно:

- а) используются стандартные для финансового анализа данные;
- б) данные получены законным путем, из открытых источников, для выборки предприятий отрасли, позволяющей провести полноценный статистический и нечетко-множественный анализ;
- в) проверены и достоверны;
- г) представлены в единой форме за ряд лет;
- д) указаны на самой html-странице или в виде таблицы.

Кроме того, информация на сайте должна быть разрешена к копированию.

Обеспечить быстрый сбор и обработку информации в удобном виде из открытых источников данных достаточно сложно, потому что:

1. Бухгалтерская отчетность предприятий, которая представлена на ряде сайтов, находится в полузакрытом либо в малодоступном

виде. Встречаются также проблемы CAPTCHA (блокировки данных после ряда обращений, проверка cookie-файлов, задержка перед действиями, ограничения по количеству обращений в течение определенного времени и т. д.).

2. Многие предприятия постоянно меняют дизайн сайтов, где убирают информацию за предыдущие годы, а за текущий год вводят платный доступ, добавляют или убирают новые блоки информации, меняют названия запросов данных с сервера и др.

3. Информация часто представлена в малопригодном для парсига (нетиповом) виде, то есть данные могут находиться в JS-файле, для каждого предприятия организация xpath может быть разной, что часто мешает точно определить собранные данные.

В настоящем исследовании была поставлена задача разработки ПО, которое ориентировано на решение задач автоматического сбора данных, необходимых и достаточных для проведения анализа финансово-экономического состояния конкретной отрасли по необходимым показателям. Имеющиеся в настоящее время программы, осуществляющие автоматизированный сбор общедоступной информации из интернета без использования API сайтов (Scraping (далее – СКРП), или его российский вариант – Parsing (далее – ПРС)), используются для анализа данных в достаточно узких целях [1–2].

Разрабатываемое в данной статье ПО должно быть основано на предварительном автоматическом анализе данных из различных источников с учетом их полноты, достоверности и корректности, а также отвечать современным требованиям корректного ис-

пользования информации из сети с учетом имеющихся законов, ограничивающих парсинг [2–3].

В настоящее время в различных методах сбора данных из сети для аналитических исследований активно используют СКРП, который является одной из основных технологий получения веб-данных путем извлечения их со страниц веб-ресурсов, что помогает автоматизировать процесс сбора данных путем создания различных библиотек [2].

Однако при использовании СКРП в большинстве случаев решаются только конкретные, практико-ориентированные задачи, подразумевающие сбор типовых данных из одного источника, упорядоченных особым способом, на относительно коротком промежутке времени. Всякий раз, когда источники данных меняются, ПО должно создаваться заново. Как следствие, имеется крайне мало теоретических исследований, посвященных возможностям СКРП по извлечению из сети репрезентативных выборок, пригодных для полноценного статистического и эконометрического анализа данных, особенно за короткое время.

Между тем в настоящее время назрела необходимость в сборе данных для полноценного информационного анализа данных (далее – ИАД) сложных социальных, экономических и других систем на основе больших массивов данных за длительные промежутки времени. Особое значение такие исследования имеют для системного анализа при формировании региональных кластеров. Современные методы ИАД позволяют это сделать, а необходимость управления системами требует разработки соответствующего универсального исследовательского аппарата.

МАТЕРИАЛЫ И МЕТОДЫ

Для решения поставленных в работе задач нами были проведены следующие исследования:

- сделан подробный анализ сайтов с данными о финансовом состоянии предприятий с целью выбора оптимального сайта, позволяющего получать для выборки предприятий заданного ОКВЭД сведения о таких показателях, как прибыль, сумма уплаченных нало-

- гов и социальных выплат (по видам), коэффициентов оценки финансового состояния и др.;

- выполнена разработка соответствующего ПО, осуществляющего автоматизированный сбор данных с отобранных сайтов посредством использования СКРП и ПРС;

- на примере ИТ-предприятий Ростовской области был проведен автоматический перевод данных в удобный для анализа формат Excel.

На первом этапе работы нами были проанализированы сайты, пригодные для получения необходимой информации: «Государственный информационный ресурс бухгалтерской (финансовой) отчетности», «За честный бизнес», «Спектр-аудит», «Бухгалтерская (финансовая) отчетность предприятий» и др. [4–7]. Однако все они, как правило, имеют ряд недостатков, перечисленных выше.

В качестве основного нами был выбран сайт TestFirm.ru, как наиболее удовлетворяющий вышеперечисленным требованиям [7]. Использование других сайтов проводилось нами также для извлечения дополнительной информации.

На следующем этапе работы нами был написан соответствующий код ПРС, который позволяет выполнять настройку создаваемого ПО под блоки изучаемых сайтов. В качестве инструмента для написания программы нами был выбран Python и его ключевые библиотеки (*request*, *selenium*, *beautifulsoup4*, *csv*) [8].

В качестве технических новшеств предложенного нами ПО можно отметить те, которые позволили не ограничивать работу во времени с сайтами, с момента начала работы с клавиатурой (в любом удобном виде). Это дало возможность решить проблему быстрого сбора информации после начала работы с клавиатурой и проводить работу без временных задержек и ограничений на количество обращений к ПО, проводить проверку ПРС, по ключевым словам, а также осуществлять упорядоченное сохранение собранных данных.

При разработке ПО нами была также создана новая технология использования СКРП для систематизации информации, размещенной на различных сайтах, при этом програм-

ма позволяет подстроить ПРС и под другие сайты.

Разработанное ПО было разделено нами на три блока:

1. Сбор URL-адресов.
2. Парсинг финансовых данных с сайта.
3. Отдельный парсер для налогов и страховых взносов.

Остановимся более подробно на каждом из них.

Для сбора URL-адресов была использована библиотека selenium для Python, а в качестве эмулированного браузера использовался Google Chrome. Сам код разрабатывался с добавлением *Chromedriver* (часть *Chromium*, которая отвечает за управление им с помощью внешних инструментов) [9] с последующим подключением необходимых электронных библиотек и созданием функции, которая отвечает за вставку ИНН конкретных фирм и их URL-адресов. При внесении ИНН фирм, которые рассматриваются в качестве поиска, используем библиотеку csv для сохранения наших адресов. Для сбора информации некоторых блоков был использован также xpath (язык запросов к элементам XML-документа). Разработан для орга-

низации доступа к частям документа XML в файлах трансформации XSLT, что является стандартом консорциума W3C из библиотеки requests [10–12].

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

В качестве примера был проведен анализ открытых данных предприятий Ростовской области, осуществляющих деятельность в области разработки компьютерного программного обеспечения, оказывающих консультационные услуги в данной области и другие сопутствующие услуги (код ОКВЭД 62). В этой базе данных имеются характеристики 447 предприятий, однако в результате парсинга были получены данные всего для 363, поскольку остальные не сдавали налоговую отчетность более года по данным ФНС.

Список предприятий Ростовской области, осуществляющих деятельность по разработке компьютерного программного обеспечения, оказанию консультационных услуг в данной области и других сопутствующих услуг, получен на сайте export-base.ru, откуда была сделана выгрузка названий компаний и их ИНН (рис. 1) [13].

База скачана с https://export-base.ru Код ОКВЭД				ИНН	ОГРН	Стационарный телефон	Мобильный телефон	Дата регистрации	Город
Ростовская область									
62. Разработка компьютерного программного обеспечения, консультационные услуги в данной области и другие сопутствующие услуги									
62.01. Разработка компьютерного программного обеспечения									
ВС САППОРТ, ООО	62.01	Разработка компьютерного программного обеспечения	6165211290	1176196061546		+7 (906) 4160220	28.12.2017	Ростовна-Дону	
ФЛАГМАН, ООО	62.01	Разработка компьютерного программного обеспечения	6163106798	1116195002274	+7 (8632) 709169	+7 (928) 2293363	18.03.2011	Ростовна-Дону	
БУРАЛО ТЕХНОЛОГИИ, ООО	62.01	Разработка компьютерного программного обеспечения	6165209816	1176196046290			29.09.2017	Ростовна-Дону	
ИМПЕКС ГРУПП, ООО	62.01	Разработка компьютерного программного обеспечения	6147037843	1146191000229	+7 (863) 6065200	+7 (928) 6065200	18.02.2014	Каменск-Шахтинский	
ТАЛАН, ООО	62.01	Разработка компьютерного программного обеспечения	6143063225	1066143050797	+7 (863) 9248683		04.09.2006	Волгодонск	
ГЭНДАЛЬФ СЕРВИС, ООО ВЦ	62.01	Разработка компьютерного программного обеспечения	6163154865	1176196024641	+7 (863) 2191600, +7 (863) 2990446, +7 (863) 2236600, +7 (863) 2506897, +7 (863) 3000230, +7 (988) 5882968, +7 (863) 3001000, +7 (863) 3001006	+7 (951) 2563130	26.02.1997	Ростовна-Дону	
СПАРК ДЕВЕЛОПМЕНТ, ООО	62.01	Разработка компьютерного программного обеспечения	6164301047	1106164005639	+7 (863) 2993718	+7 (918) 5025607	28.12.2010	Ростовна-Дону	
АЛИН, ООО	62.01	Разработка компьютерного программного обеспечения	6165120653	1056165030085	+7 (863) 2450574		02.03.2005	Ростовна-Дону	
ГЭНДАЛЬФ, ООО	62.01	Разработка компьютерного программного обеспечения	6167084305	1066167006421	+7 (863) 2236500, +7 (863) 3001001, +7 (863) 3001000	+7 (928) 1713951	07.02.2006	Ростовна-Дону	
ГЭНДАЛЬФ, ООО ЦИТ	62.01	Разработка компьютерного программного обеспечения	6167132823	1166196062834	+7 (863) 2236500, +7 (863) 3001000	+7 (903) 4063130	23.11.1993	Ростовна-Дону	
ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ ЮФУ, НТЦ	62.01	Разработка компьютерного программного обеспечения	6163027810	1026103165241	+7 (8634) 312350, +7 (8634) 315941, +7 (8634) 318000, +7 (8634) 328725, +7 (8634) 371635, +7 (8634) 615459, +7 (8634) 362550		03.02.1994	Таганрог	

Рис. 1. Пример таблицы данных по предприятиям из export-base.ru по Ростовской области

Примечание: составлено авторами на основании данных, полученных в исследовании.

Разработанное нами ПО выполняет автоматически следующие действия: эмулирует браузер, вводит ИНН фирмы в поиск и копирует URL-адрес страниц в отдельный csv файл для дальнейшей работы с ними, собирает необходимую информацию (ИНН, название компании, ОКВЭД, доходы, расходы, налог на прибыль, земельный, социальное страхование и др.). Эти данные вносятся автоматически в таблицы Excel со всеми необходимыми финансовыми показателями предприятий. Такие таблицы удобны для редактирования и пригодны для работы дополнительного программного комплекса, осу-

ществляющего анализ конкретной отрасли на основе выбранных показателей.

Для такого сбора информации при использовании разработанного ПО требуется от 30 секунд до 1 минуты, а после на парсинг одной страницы уходит одна-две секунды, что приводит к значительной экономии времени при анализе от 100 и более фирм.

В итоге мы получаем таблицу Excel со всеми финансовыми показателями предприятий (рис. 2), которые удобны для дальнейшего редактирования и пригодны для работы программного комплекса, осуществляющего комплексный анализ ИТ-отрасли на основе выбранных показателей.

A8	✕	✓	<i>f_x</i>	ООО «ГЭНДАЛЬФ»
	A	B	C	
1	ООО «Флагман»	(Малое предприятие)	6163106798	
2	ООО «БУРАЛО ТЕХНОЛОГИИ»	(Микропредприятие)	6165209816	
3	ООО «ИМПЕКС ГРУПП»	(Микропредприятие)	6147037843	
4	ООО «Талан»	(Микропредприятие)	6143063225	
5	ООО ВЦ «ГЭНДАЛЬФ сервис»	(Малое предприятие)	6163154865	
6	ООО «Спарк Девелопмент»	(Микропредприятие)	6164301047	
7	ООО «АЛИН»	(Микропредприятие)	6165120653	
8	ООО «ГЭНДАЛЬФ»	(Малое предприятие)	6167084305	
9	ООО ЦЕНТР ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ «ГЭНДАЛЬФ»		6167132823	
10	ООО «ПРОГА»	(Микропредприятие)	6167040234	
11	ООО ТОРГОВО-ПРОИЗВОДСТВЕННАЯ КОМПАНИЯ «РОСТСТРОЙПАНЕЛЬ»	(Микропредприятие)	6165206685	
12	ООО «РУФОРМАТ»	(Микропредприятие)	6165216813	
13	ЗАО «ТАМГА»	(Малое предприятие)	6167046035	
14	ООО «АСТОРМЕДИА»	(Микропредприятие)	6163152297	
15	ООО «ДА-ВИНЧИ»	(Микропредприятие)	6165128998	
16	ООО «Информационные Технологии»	(Микропредприятие)	6154140351	
17	ООО «ЮГ-ИНТЕГРАТОР»	(Микропредприятие)	6161042796	
18	ООО «АЛМУР»	(Микропредприятие)	6161004945	

Рис. 2. Excel-таблица финансовых показателей предприятий ИТ-отрасли Ростовской области

Примечание: составлено авторами на основании данных, полученных в исследовании.

На рис. 3 приведен пример анализа финансовых показателей предприятий ИТ-отрасли Ростовской области.

Предложенное программное обеспечение ориентировано на решение научно-исследовательской задачи сбора данных, необходимых и достаточных для проведения анализа финансово-экономического состояния отрасли по конкретным показателям, что отличает его от распространенных в настоящее время программ скрейпинга для анализа данных в практических, сиюминутных целях. Как след-

ствие, разработанное программное обеспечение основано на предварительном анализе источников данных с учетом их полноты, достоверности и корректности.

Благодаря разработке данного ПО сделан шаг в направлении развития технологий использования скрейпинга для решения долгосрочных научно-исследовательских задач и систематизации информации, размещенной на различных сайтах. Блоки кода позволяют подстроить парсер под другие сайты.

	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN
	Налог, взимаемый в связи с применением упрощенной системы налогообложения	Неналоговые доходы	Налог на добавленную стоимость	Налог на прибыль	Земельный налог	Транспортный налог	Налог на имущество организаций	Задолженность и перерасчеты по отмененным доходам	Страховые и другие взносы на обязательное пенсионное страхование	Страховые взносы на обязательное медицинское страхование	Страховые взносы на обязательное социальное страхование
1											
8		0 руб.	0 руб.	0 руб.				0 руб.	0 руб.		
9	3,4 млн. руб.	0 руб.	0 руб.	0 руб.			0 руб.	0 руб.	2,23 млн. руб.	1,11 млн. руб.	550 тыс. руб.
10			0 руб.	0 руб.			0 руб.	0 руб.	0 руб.	0 руб.	20 руб.
11	40,2 тыс. руб.		0 руб.	0 руб.					37,3 тыс. руб.	9,11 тыс. руб.	4,69 тыс. руб.
12	0 руб.	0 руб.	126 тыс. руб.	8,28 тыс. руб.					1,81 тыс. руб.	422 руб.	240 руб.
13	71,9 тыс. руб.								161 тыс. руб.	37,6 тыс. руб.	21,1 тыс. руб.
14	0 руб.		35,3 млн. руб.	404 тыс. руб.	222 тыс. руб.	7,85 тыс. руб.	5,05 тыс. руб.	0 руб.	5,76 млн. руб.	1,96 млн. руб.	402 тыс. руб.
15	16,4 тыс. руб.	0 руб.							54,1 тыс. руб.	12,5 тыс. руб.	7,14 тыс. руб.
16		0 руб.	419 тыс. руб.	68,2 тыс. руб.		1,78 тыс. руб.	0 руб.	0 руб.	50 тыс. руб.	14,2 тыс. руб.	5,19 тыс. руб.
17	10 тыс. руб.								16,3 тыс. руб.	3,8 тыс. руб.	2,16 тыс. руб.
18	12,5 тыс. руб.								16,4 тыс. руб.	3,81 тыс. руб.	2,16 тыс. руб.
19	890 тыс. руб.		0 руб.	0 руб.			0 руб.	0 руб.	519 тыс. руб.	120 тыс. руб.	68,5 тыс. руб.
20									0 руб.		
21	12,6 тыс. руб.	0 руб.							28,8 тыс. руб.	6,69 тыс. руб.	3,8 тыс. руб.
22	12,6 тыс. руб.								90 тыс. руб.	26,9 тыс. руб.	11,8 тыс. руб.
23	188 тыс. руб.		0 руб.	0 руб.			0 руб.		71,4 тыс. руб.	18,5 тыс. руб.	8,41 тыс. руб.
24			0 руб.	0 руб.			0 руб.		0 руб.		
25	0 руб.								0 руб.	0 руб.	0 руб.
26	24,2 тыс. руб.	0 руб.							148 тыс. руб.	37,2 тыс. руб.	18,3 тыс. руб.
27		0 руб.	2,49 млн. руб.	0 руб.	0 руб.	2,56 тыс. руб.	24,3 тыс. руб.	0 руб.	1,69 млн. руб.	896 тыс. руб.	384 тыс. руб.
28	63,1 тыс. руб.		0 руб.						99,4 тыс. руб.	24,8 тыс. руб.	12,2 тыс. руб.

Рис. 3. Финансовые показатели предприятий IT-отрасли Ростовской области
 Примечание: составлено авторами на основании данных, полученных в исследовании.

ЗАКЛЮЧЕНИЕ

Разработанное программное обеспечение для автоматизированного сбора данных позволяет на основе парсинга открытого сайта TestFirm.ru в компактной форме Excel-файла за кратчайшее время получать сведения о значимых финансово-экономических показателях фирм заданной отрасли по интересующему региону. Новизна разработанного ПО состоит в том, что оно ориентировано не на решение текущих практических задач анализа информации в сети, а на сбор информации для полноценного научного исследования. Как следствие, при разработке программного обеспечения выполнены следующие требования: собраны стандартные данные, используемые в финансовом анализе; данные получены законным путем, из открытых источников; данные получены для выборки предприятий отрасли, позволяющей провести полноценный статистический и нечетко-множественный анализ; данные прове-

рены и достоверны, представлены в единой форме за ряд лет; данные указаны на самой html-странице или в виде таблицы; информация на сайте разрешена к копированию.

Разработанное ПО имеет четко продуманную структуру, оптимизированное быстродействие и отвечает современным требованиям корректного использования информации из сети с учетом имеющихся в российском законодательстве законов, ограничивающих парсинг. В качестве технических новшеств предложенного ПО следует отметить следующие: нажатие кнопок и сбор информации после нажатия, сокращение времени задержки для обхода ограничений на количество нажатий по времени, проверка парсинга по ключевым словам, а также упорядоченное сохранение собранных данных.

Условием сбора данных является наличие списка фирм, использованного в данном исследовании, на официальном сайте export-base.ru.

Список источников

1. Krotov V., Silva L. Legality and Ethics of Web Scraping // Proceedings of the 24th Americas Conference on Information Systems, New Orleans, 2018. URL: https://www.researchgate.net/publication/324907302_Legality_and_Ethics_of_Web_Scraping (дата обращения: 30.08.2022).
2. Рожкова М. А. Базы данных и сервисы онлайн-классифайдов: пользование базой и использование информации // Журн. суда по интеллектуал. правам. 2019. № 26. С. 25–32.

References

1. Krotov V., Silva L. Legality and Ethics of Web Scraping // Proceedings of the 24th Americas Conference on Information Systems, New Orleans, 2018. URL: https://www.researchgate.net/publication/324907302_Legality_and_Ethics_of_Web_Scraping (accessed: 30.08.2022).
2. Rozhkova M. A. Databases and Services of Online Classifieds // Zhurn. suda po intellektual. pravam. 2019. No. 26. P. 25–32. (In Russian).

3. Государственный информационный ресурс бухгалтерской (финансовой) отчетности. URL: <https://bo.nalog.ru> (дата обращения: 30.08.2022).
4. Всероссийская система проверки контрагентов. URL: <https://zachestnyibiznes.ru> (дата обращения: 30.08.2022).
5. Спектр-аудит. URL: <https://spektr-audit.com> (дата обращения: 30.08.2022).
6. Бухгалтерская (финансовая) отчетность предприятий. URL: <https://e-ecolog.ru/buh> (дата обращения: 30.08.2022).
7. Сравнение финансового состояния фирмы с отраслевыми показателями и конкурентами. URL: <https://www.testfirm.ru> (дата обращения: 30.08.2022).
8. Mitchell R. Web Scraping with Python. 2nd Ed. 2018. 288 p.
9. Chromedriver. URL: <https://chromedriver.chromium.org/downloads> (дата обращения: 30.08.2022).
10. Adams M. D. Principled Parsing for Indentation-Sensitive Languages: Revisiting Landin's Offside Rule // ACM SIGPLAN Notices. Vol. 48, Is. 1. P. 511–522. (дата обращения: 30.08.2022).
11. Ramu C., Gemünd C., Gibson T. J. Object-Oriented Parsing of Biological Databases with Python // Bioinformatics. 2000. Vol. 16, Is. 7. P. 628–638. URL: <https://academic.oup.com/bioinformatics/article/16/7/628/2279282?login=false> (дата обращения: 30.08.2022).
12. Chapagain A. Hands-On Web Scraping with Python: Perform Advanced Scraping Operations Using Various Python Libraries and Tools Such as Selenium, Regex, and Others. 2019. 189 p. URL: https://books.google.ru/books?hl=ru&lr=&id=j-6iDwAAQBAJ&oi=fnd&pg=PP1&dq=parsing+python+selenium&ots=J9F2ScFdG8&sig=y0U5WR-xLhvNR3bGpJdAOyDrXR8&redir_esc=y#v=onepage&q=parsing%20python%20selenium&f=false (дата обращения: 30.08.2022).
13. ExportBase. URL: <https://export-base.ru/> (дата обращения: 30.08.2022).
3. Gosudarstvennyi informatsionnyi resurs bukhgalterskoi (finansovoi) otchetnosti. URL: <https://bo.nalog.ru> (accessed: 30.08.2022). (In Russian).
4. Vserossiiskaia sistema proverki kontragentov. URL: <https://zachestnyibiznes.ru> (accessed: 30.08.2022). (In Russian).
5. Spektr-audit. URL: <https://spektr-audit.com> (accessed: 30.08.2022). (In Russian).
6. Bukhgalterskaia (finansovaia) otchetnost predpriatii. URL: <https://e-ecolog.ru/buh> (accessed: 30.08.2022). (In Russian).
7. Sravnenie finansovogo sostoiianiia firmy s otraslevymi pokazateliami i konkurentami. URL: <https://www.testfirm.ru> (accessed: 30.08.2022). (In Russian).
8. Mitchell R. Web Scraping with Python. 2nd Ed. 2018. 288 p.
9. Chromdriver. URL: <https://chromedriver.chromium.org/downloads> (accessed: 30.08.2022).
10. Adams M. D. Principled Parsing for Indentation-Sensitive Languages: Revisiting Landin's Offside Rule // ACM SIGPLAN Notices. Vol. 48, Is. 1. P. 511–522. (accessed: 30.08.2022).
11. Ramu C., Gemünd C., Gibson T. J. Object-Oriented Parsing of Biological Databases with Python // Bioinformatics. 2000. Vol. 16, Is. 7. P. 628–638. (accessed: 30.08.2022).
12. Chapagain A. Hands-On Web Scraping with Python: Perform Advanced Scraping Operations Using Various Python Libraries and Tools Such as Selenium, Regex, and Others. 2019. 189 p. URL: https://books.google.ru/books?hl=ru&lr=&id=j-6iDwAAQBAJ&oi=fnd&pg=PP1&dq=parsing+python+selenium&ots=J9F2ScFdG8&sig=y0U5WR-xLhvNR3bGpJdAOyDrXR8&redir_esc=y#v=onepage&q=parsing%20python%20selenium&f=false (accessed: 30.08.2022).
13. ExportBase. URL: <https://export-base.ru/> (accessed: 30.08.2022). (In Russian).

Информация об авторах

С. О. Крамаров – доктор физико-математических наук, профессор.

В. А. Овсянников – магистрант.

Л. В. Сахарова – доктор физико-математических наук, доцент.

Р. С. Усатый – аспирант.

Г. В. Лукьянова – кандидат технических наук, доцент.

Information about the authors

S. O. Kramarov – Doctor of Sciences (Physics and Mathematics), Professor.

V. A. Ovsyannikov – Master's Degree Student.

L. V. Sakharova – Doctor of Sciences (Physics and Mathematics), Associate Professor.

R. S. Usatyi – Postgraduate.

G. V. Lukyanova – Candidate of Sciences (Engineering), Associate Professor.