

ПРОТОКОЛ ОБРАБОТКИ НАБОРОВ ДАННЫХ ДЛЯ ИХ ПУБЛИКАЦИИ В ОТКРЫТЫХ ИСТОЧНИКАХ

Борисов Р.С., Ефименко А.А.*

Ключевые слова: открытые данные, структура набора данных, метаданные, паспорт данных, протокол, публикация наборов данных, конфиденциальность, классификационные признаки.

Аннотация.

Цель работы: совершенствование научно-методической базы регулирования опубликования в открытых источниках наборов данных, потенциально содержащих сведения, составляющие конфиденциальную информацию.

Методы: системно-правовой анализ внешних и атрибутивных свойств и организационных приемов обработки наборов данных, позволяющих обеспечить конфиденциальность информации при публикации в открытых источниках.

Результаты: выделены группы обладателей и потребителей наборов данных, определены их права, основные субъекты, а также накладываемые ими ограничения; сформирован организационно-технический регламент публикации наборов данных, включающий концептуальный вариант подхода к организации информационных процессов при публикации наборов данных; разработан итерационный протокол обработки наборов данных при их публикации, позволяющий выделить основные процедуры обработки, анализа и преобразования наборов данных.

DOI: 10.21681/1994-1404-2021-2-59-70

Введение

В настоящее время одной из важных общественных проблем является соблюдение определенного баланса между масштабами открытого распространения *полезной* информации и вместе с тем определенных ограничений на доступ к информации о деятельности государственных и муниципальных органов власти, предприятий промышленности, транспорта, связи, научных, образовательных, медицинских учреждений, общественных и других организаций. Основной задачей соответствующего *регламента опубликования наборов данных* является недопущение опубликования наборов данных, содержащих привилегированные сведения (включая персональные данные), составляющие *конфиденциальную* информацию и находящиеся в режиме тайны (государственной, личной, семейной, коммерческой, служебной, профессиональной [9, 14]), а также информацию, запрещенную к публикации — информацию, направленную на пропаганду войны, разжигание национальной, расовой или религиозной ненависти и вражды, а также иную информацию, за распространение которой предусмотрена уголовная

или административная ответственность¹; информацию, оскорбляющую человеческое достоинство и общественную нравственность, демонстрирующую явное неуважение к обществу, государству, официальным государственным символам Российской Федерации, Конституции Российской Федерации или органам, осуществляющим государственную власть в Российской Федерации².

Порядок обработки *персональных* данных регулируется Федеральным законом «О персональных данных» от 27 июля 2006 г. № 152-ФЗ и другими нормативными правовыми актами; порядок обработки сведений, составляющих *коммерческую тайну* — Федеральным законом «О коммерческой тайне» от 29 июля 2004 г. № 98-ФЗ; сведений, составляющих *служебную тайну* — Постановлением Правительства РФ от 3 ноября 1994 г. № 1233 (ред. от 06.08.20) «Об утверждении Положения о порядке обращения со служебной информацией ограниченного распространения в федеральных органах исполнительной власти, уполномоченном органе

¹ Пункт 6 ст.10 Федерального закона от 27 июля 2006 г. № 149-ФЗ (ред. от 09.03.21) «Об информации, информационных технологиях и о защите информации».

² Там же.

* **Борисов Роман Сергеевич**, кандидат технических наук, доцент кафедры информационного права, информатики и математики Российского государственного университета правосудия, г. Москва, Российская Федерация.
E-mail: bestseller@bk.ru

Ефименко Алексей Анатольевич, кандидат технических наук, доцент кафедры информационного права, информатики и математики Российского государственного университета правосудия, г. Москва, Российская Федерация.
E-mail: alex192@mail.ru

управления использованием атомной энергии и уполномоченном органе по космической деятельности».

Разнообразие профессиональной деятельности, в процессе которой должностные лица становятся обладателями конфиденциальной информации, влечет за собой многообразие видов информации из разряда *профессиональной тайны*: врачебная, юридическая, нотариальная тайны, тайна переписки и др. [1, 4, 6, 7, 9, 10]. Согласно российскому законодательству, в государственных организациях и коммерческих структурах должна быть разработана и утверждена система защиты государственной тайны и конфиденциальной информации (персональные данные, коммерческая, служебная и профессиональная тайна), в том числе перечни соответствующей информации.

Сложность анализа наборов данных и результатов исследований заключается в том, что при обработке (преобразовании) [12, 16] наборов данных, в паспортах которых указано, что данные не содержат конфиденциальную информацию (персональные данные, коммерческую, профессиональную и служебную тайны), в некоторых случаях выявляется наличие информации, которая является конфиденциальной [5, 8]. Чтобы исключить данную ситуацию, необходимо определить соответствующие преобразования данных (интеграция и агрегирование) и внедрить *регламент* обработки наборов данных, который позволит публиковать обрабатываемые данные в открытом доступе, не нарушая российского законодательства по обработке конфиденциальной информации с соблюдением авторских и смежных прав собственников информации.

Ключевые элементы, определяющие режим доступа к наборам данным

Выявление ключевых элементов наборов данных, содержащих элементы конфиденциальной информации, можно реализовать с помощью *экспертных* методов либо средствами интеллектуальной автоматизированной обработки данных.

В том случае, если структура наборов данных имеет *линейную* (табличную) структуру с ограниченным набором полей, можно ограничиться экспертными методами определения полей (атрибутов) данных, содержащих конфиденциальную информацию. Если наборы данных представлены в виде *неструктурированной* или *слабоструктурированной* информации, например, в виде текстов договоров, контрактов, анкет, заявлений и др., использование экспертных методов для определения возможной их публикации становится очень трудоемким и дорогостоящим процессом, нивелирующим все преимущества от опубликования наборов данных. Для инициации процедур анализа и преобразования таких данных необходимо специальное решение специального *экспертного совета*, подготовленное на основе мотивированного заключения *отдела управления данными*.

В случае использования экспертных методов экспертизе подвергаются *паспорта* наборов данных и все входящие в состав наборов данных — результаты исследований поля (атрибуты). Эксперт оценивает атрибуты и их содержание на наличие сведений конфиденциального характера. В зависимости от специфики наборов данных, в качестве экспертов могут выступать специалисты по отдельным категориям конфиденциальных данных, охватывающих те или иные отрасли права, входящие в классификатор [14], например, специалист по персональным данным или по вопросам коммерческой тайны и др. Для повышения качества распознавания атрибутов при анализе одних и тех же нормативных актов могут привлекаться несколько экспертов с последующей мажоритарной выборкой результатов их работы.

Отдельный элемент, подлежащий анализу экспертами — *паспорт набора данных*. Паспорт включает в себя дополнительную информацию, не входящую непосредственно в сам набор данных. Поля паспорта разработаны таким образом, что при правильном и полном их заполнении часть нагрузки экспертов по выявлению ключевых атрибутов, содержащих конфиденциальные данные, переносится на обладателей этих наборов данных. Кроме того, качественное заполнение паспорта позволит снизить квалификационные требования к компетенциям экспертов отдела управления данными.

Обладатель набора данных обязан самостоятельно заполнять паспорт информацией, которую он посчитает необходимой для правильной интерпретации режима доступа к своему набору данных.

Однако если наборы данных содержат слабоструктурированные данные, их «ручная» обработка становится нецелесообразной. Для обработки такой информации предлагается подход, основанный на применении однослойной *нейросети* [15] для обучения и распознавания классификационных признаков (ссылочных элементов) наличия конфиденциальных данных, основанный на ранжировании уровня конфиденциальности.

Взаимодействие субъектов обработки наборов данных

Одним из основных организационных механизмов, используемых в мировой практике при рассмотрении вопросов о возможности публикации документов государственных органов и организаций в открытом доступе, является создание специальных групп экспертов — *экспертных советов* (рис. 1), которым поручается рассмотрение этих вопросов с учетом существующих реестров и классификаций конфиденциальной информации и других нормативных правовых актов.

Эксперт-аналитик 1 (сектора оценки входящих данных), хорошо знающий классификатор конфиденциальных данных, изучает поступивший набор данных вместе с их паспортом и выявляет те законы и нормативные правовые акты (НПА), которые могут препят-

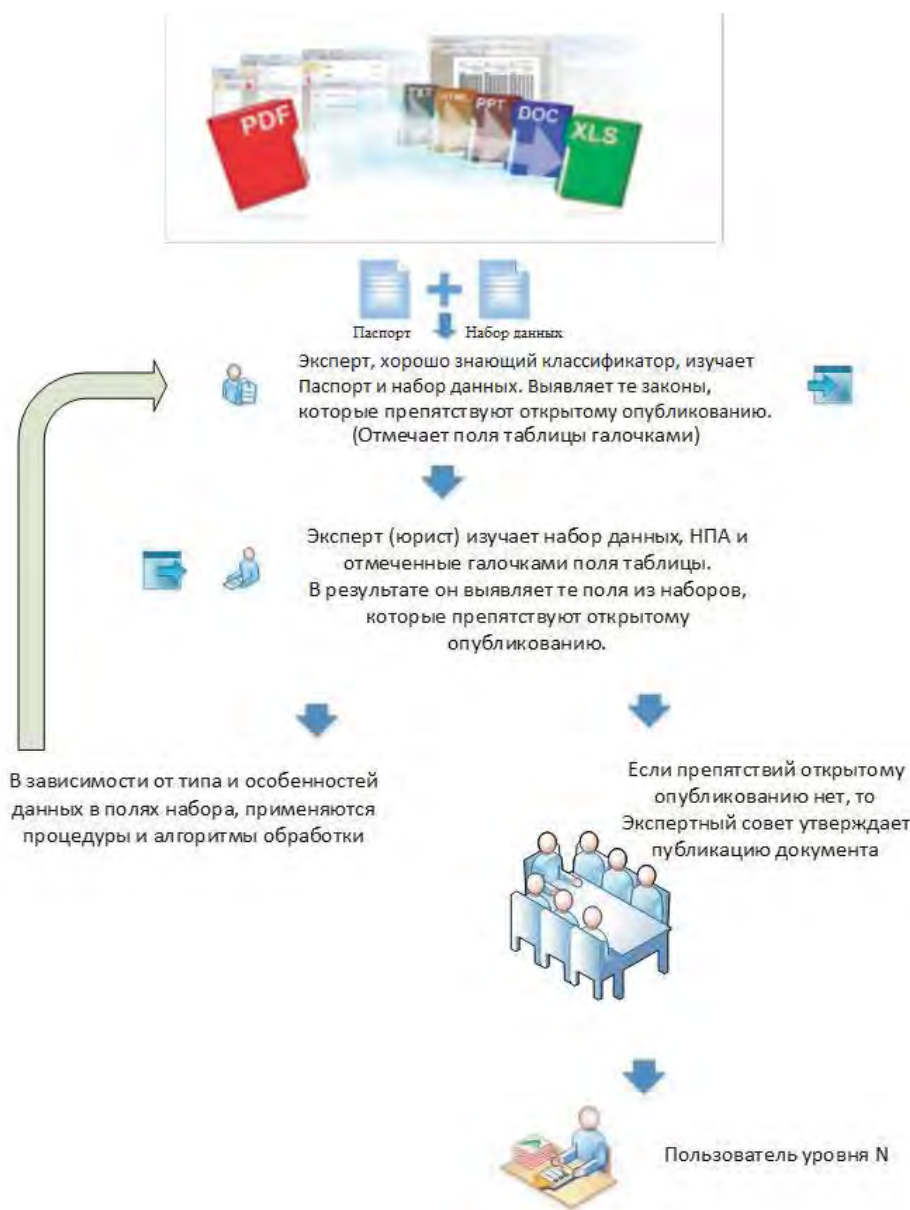


Рис. 1. Схема взаимодействия субъектов обработки наборов данных

ствовать открытому опубликованию (проставляет соответствующие отметки).

Эксперт-аналитик 2 (юрист), обладающий компетенциями по НПА, которые могут препятствовать открытому опубликованию и которые отмечены экспертом 1, выявляет те атрибуты (поля) из наборов, которые препятствуют открытому опубликованию.

В зависимости от типа и особенностей атрибутов данных (полей набора данных) к этим полям применяются процедуры и алгоритмы обработки [6], которые позволяют исключить конфиденциальные сведения, препятствующие открытой публикации поступившего набора данных. Формируется проект экспертного заключения, который передается на рассмотрение и утверждение в экспертный совет.

Экспертный совет рассматривает полученный проект экспертного заключения и дает (или не дает) решение на публикацию в открытом контуре. Данные

публикуются в открытом доступе либо с этими наборами данных производятся другие действия, рекомендованные экспертным советом.

На предварительном этапе должны быть проанализированы все перечни конфиденциальной информации (персональные данные, коммерческая, служебная и профессиональная тайна), которые обладатель передаваемых наборов данных установил в своей организации. В соглашении об информационном взаимодействии с обладателем должны быть предусмотрены условия передачи таких перечней.

Анализ необходимо проводить на предмет выделения передаваемой в наборе данных информации, которая обладает свойством *кумулятивности* [9] — при достижении определенного объема может становиться конфиденциальной, например:

- временные интервалы (за квартал — не является конфиденциальной, за год — является);

- данные по административно-территориальному признаку (по области — не конфиденциальные, по федеральному округу или всей стране — конфиденциальные);
- данные по материнской компании — конфиденциальные, а по дочерним компаниям по отдельности — нет.

В каждом паспорте набора данных, полученного от обладателя набора данных, должны быть предусмотрены поля:

- поле, в котором отмечено, являются ли данные этого набора данных конфиденциальной информацией (персональные данные, коммерческая, служебная и профессиональная тайна — одновременно можно выбрать несколько категорий);
- поле, в котором отмечен признак возможности опубликовать данные и, если определена возможность опубликования, то в каком виде публикуется:

0 — разрешена публикация в передаваемом виде;

1 — разрешена публикация в обезличенном (для персональных данных) или деперсонифицированном (для коммерческой тайны) виде;

2 — разрешена публикация в анонимизированном виде;

3 — запрещена публикация.

В отдельных случаях необходим еще и анализ по перечням конфиденциальной информации и классификационным признакам, определяющим сведения, составляющие государственную тайну. Это необходимо, например, в случаях, когда в наборе данных содержится информация, поступающая от отдельных организаций, которая является открытой информацией, а в совокупности может быть закрытой (например, добыча или покупка редкоземельных элементов отдельной компанией и страной в целом).

При формировании как паспорта наборов данных, так и самих наборов данных следует учитывать *правило*: для каждого набора данных должна быть установлена периодичность обновления с учетом потребностей пользователей наборов данных, например, данных *судебной статистики* [2]. Для каждого набора данных в паспорте набора данных должна быть указана дата его актуальности. Частота обновления набора данных может быть следующей:

Оперативные — данные, поступающие в интерактивном режиме, оперативные данные о деятельности компании, текущая отчетность и др. Например, оперативные данные МЧС России о числе возгораний и природных пожаров в России или оперативные итоги голосования ЦИК России. Для оперативных данных частота обновления может быть установлена следующими значениями: чаще одного раза в день; ежедневно; еженедельно.

Долговременные данные — управленческие регламенты, например, должностные и рабочие инструкции, правила поведения персонала; законы, нормы, нормативы, стандарты, технологические карты производ-

ственных процессов и др. Для долговременных данных частота обновления может быть установлена следующими значениями: ежемесячно; ежеквартально; каждые полгода; ежегодно; по мере изменения данных.

Обладатель набора данных может указать срок хранения наборов данных, но для полноты предоставления услуг пользователям необходимо стремиться обеспечить хранение всех версий наборов данных.

Обладателей наборов данных можно отнести к двум группам: первичные и вторичные.

К *первичной* группе относятся обладатели, предоставляющие наборы данных — результаты исследований:

- федеральные органы исполнительной власти (ФОИВ) или другие органы — в части информации, предоставляемой:

- а) по запросу согласно ст. 7.1 210-ФЗ³ (в отношении информации о деятельности государственного органа);

- б) в соответствии с законами и (или) иными нормативными правовыми актами, принятыми на основании ст. 4 8-ФЗ⁴ (в отношении информации о деятельности государственного органа);

- в) по соглашениям об информационном взаимодействии, в том числе на основании ч. 4 ст. 10 149-ФЗ⁵ и с учетом предусмотренных в ч. 3 ст. 6 149-ФЗ прав обладателя информации (Российской Федерации), от имени которого выступает ФОИВ;

- г) в соответствии с положениями законодательства, определяющими особенности предоставления доступа к информации, содержащейся в отдельных государственных информационных системах и ресурсах, доступа к иной информации, не относящейся к информации о деятельности государственного органа;

- третьи лица (организации) — в части информации, предоставляемой по лицензионным договорам или по договорам оказания информационных услуг;

- пользователи (физические лица, потребители) — в части научных исследований, в которых используются результаты обработки данных по лицензионным договорам.

К *вторичной* группе относятся обладатели наборов данных, опубликованных в открытом доступе, а также эксперты отдела управления данными, выполняющие обработку исходной информации.

В паспортах, соответствующих наборам данных — результатам исследований, указываются разрешения и ограничения на пользование информацией, включая разрешения на обработку, распространение, опубликование персональной и другой информации.

³ Федеральный закон от 27 июля 2010 г. № 210-ФЗ (ред. от 30.12.20) «Об организации предоставления государственных и муниципальных услуг».

⁴ Федеральный закон от 9 февраля 2009 г. 8-ФЗ (ред. от 08.12.20) «Об обеспечении доступа к информации о деятельности государственных органов и органов местного самоуправления».

⁵ Федеральный закон от 27 июля 2006 г. № 149-ФЗ (ред. от 09.03.21) «Об информации, информационных технологиях и о защите информации».

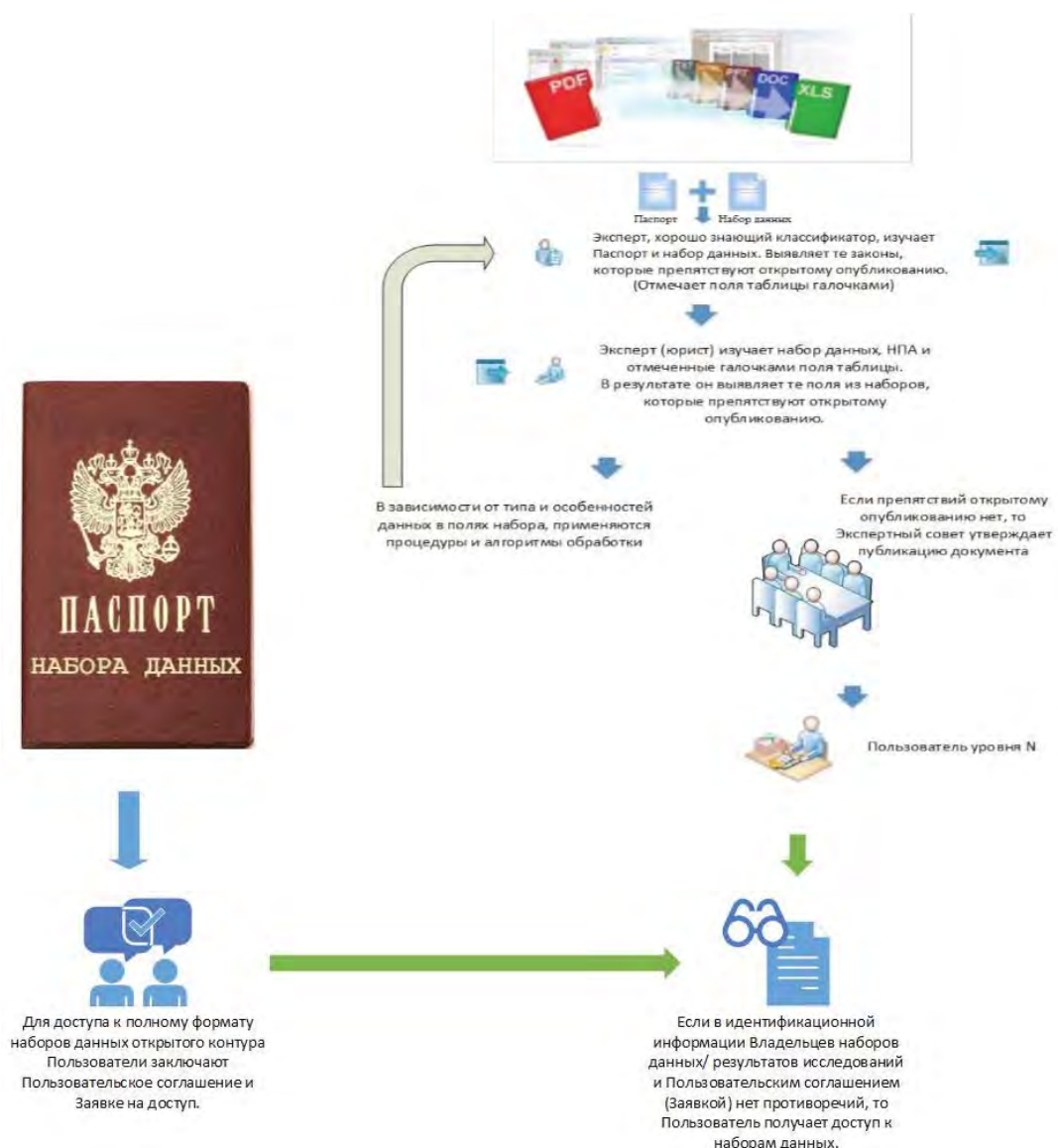


Рис. 2. Общая схема регламента публикации наборов данных

Потребителей наборов данных в зависимости от заключенного пользовательского соглашения можно условно разделить на следующие группы:

- лица, не заключающие пользовательского соглашения и имеющие возможность работы с общедоступной информацией, размещаемой в сети Интернет или получаемой по их запросам;
- лица, заключившие пользовательское соглашение на доступ к наборам данных, размещаемых в регламентируемых папках — указанные наборы данных до их передачи пользователям были подвергнуты преобразованиям, исключающим утечку конфиденциальных данных;
- эксперты, оказывающие услуги в рамках контрольного или экспертно-аналитического мероприятия;
- эксперты (специалисты), оказывающих иные услуги (например, участие в подготовке годовых планов, планов научно-исследовательских работ и др.);
- иные лица, получающие доступ к информации, в том числе рецензенты.

Для отдельных категорий пользователей доступ может быть ограничен.

Протокол публикации наборов данных

Регламент публикации наборов данных (рис. 2) от момента их получения от обладателя до момента их публикации определяет порядок подготовки и публикации общедоступной информации в формате открытых данных в соответствии с принципом открытости. Соответствующий итерационный протокол [7] обработки наборов данных при их публикации (рис. 3), позволяющий выделить основные процедуры обработки, анализа и преобразования наборов данных, а также определить цели, участников и ожидаемые результаты на каждом шаге, состоит в следующем:

Шаг 1. В отдел управления данными поступает паспорт набора данных и непосредственно набор данных. Если в паспорте отмечено, что набор данных не содержит конфиденциальных сведений (составляющих

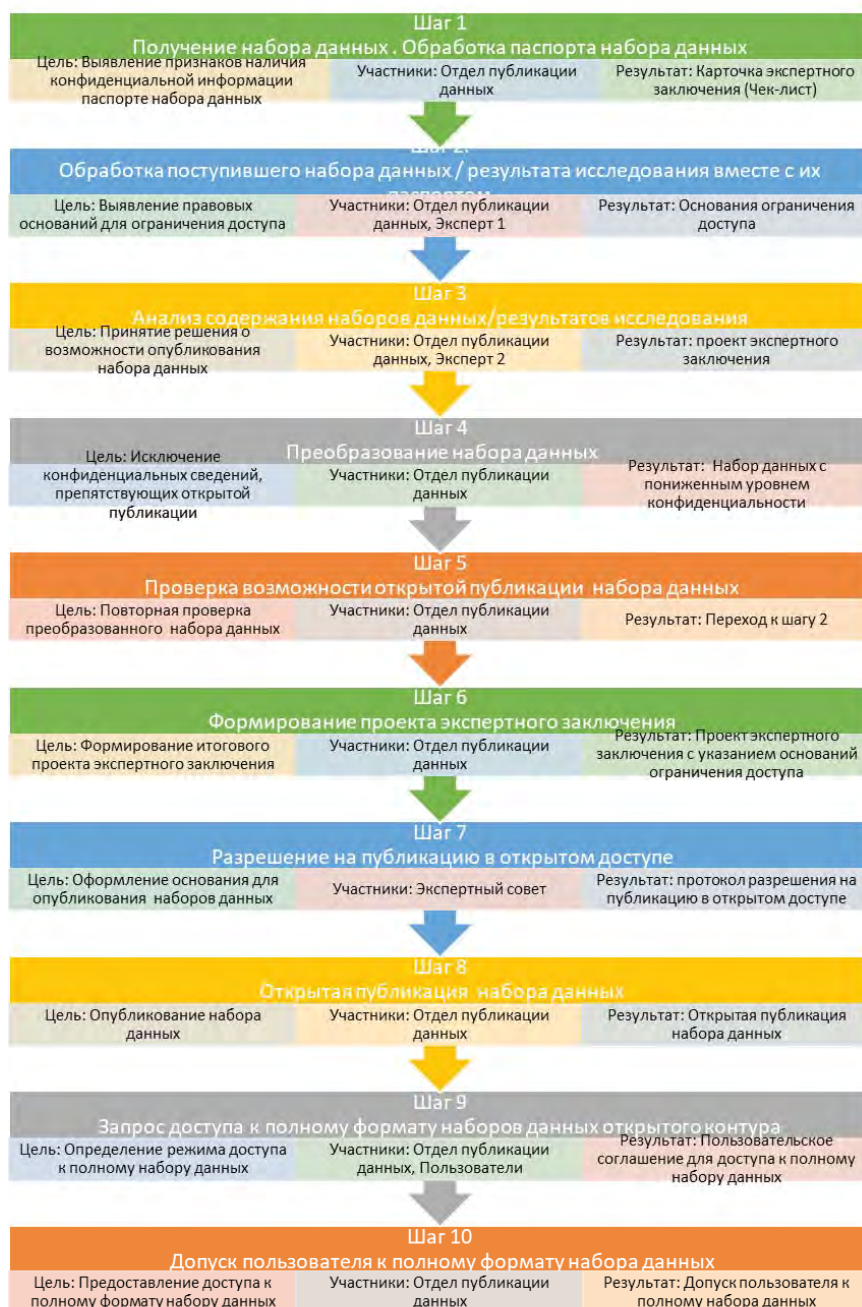


Рис. 3. Протокол обработки наборов данных при их публикации

государственную тайну или других сведений ограниченного распространения), то отдел управления данными принимает набор данных. В противном случае эксперт-аналитик готовит проект экспертного заключения и передает его в экспертный совет для рассмотрения. Экспертный совет в утвержденном протоколе определяет дальнейшие действия, и набор данных — результат исследования поступает в режимное делопроизводство.

В отдел управления данными поступает набор данных от обладателя набора данных, который предварительно заключил соглашение об информационном взаимодействии. Эксперт отдела управления данными (сектор оценки входящих данных) заполняет карточку экспертного заключения по шаблону, представленному на рис. 4.

Шаг 2. На основе разработанного классификатора эксперт-аналитик 1 (отдел управления данными, сектор оценки входящих данных) изучает поступивший набор данных вместе с паспортом (который представляет обладатель набора данных) и выявляет нормативные правовые акты, которые могут препятствовать открытому опубликованию (делает соответствующие отметки). Если таких оснований не выявлено, то переход к **Шагу 6**, иначе — к **Шагу 3**.

Шаг 3. Эксперт-аналитик 2 (юрист, обладающий компетенциями по НПА, которые отмечены на **Шаге 2**) также изучает набор данных по отношению к тем НПА, которые отмечены на предыдущем шаге.

В случае выявления в структуре и элементах наборов данных сведений, составляющих конфиденциаль-

Наименование Обладателя набора данных	
Дата, номер Пользовательского соглашения	
Наименование набора данных	
Формат набора данных	Текстовый, графический, табличный, иной.
Период, за который предоставляется набор данных	
Дата предоставления Паспорта и набора данных	
Способ получения набора данных	
периодичность обновления / срок хранения	
Категория информации в Паспорте (общедоступная информация/сведения ограниченного доступа) предоставленная Обладателем	Общедоступная информация - не содержит государственной, служебной, коммерческой и иной охраняемой законом тайны, а также сведений, отнесенных к информации ограниченного доступа Обладателем набора данных.
Результат обследования набора данных экспертом Отдела управления данными	Совпадают сведения, представленные в наборе данных с Паспортом. Вывод о необходимых действиях.
Необходимые действия эксперта	1) Если ответ эксперта положительный, данная карточка подписывается экспертом и представляется в Экспертный совет, который принимает решение о дальнейших действиях с набором данных. (См. Шаг 8 регламента). 2) Если сведения, представленные в наборе данных, не совпадают с Паспортом или в результате обследования эксперт выявил сведения, (по классификационным признакам) попадающие под ограничения открытого опубликования. Эксперт прописывает конкретные действия по устранению данных ограничений.
Результат действия эксперта (проведенные преобразования)	1) « – » если преобразования не проводились 2) Описываются, какие преобразования были проведены. Эксперт, после вторичной проверки набора данных, делает вывод о возможности публикации в открытом доступе.
Эксперт Отдела управления данными, предоставившей заключение по набору данных (ФИО, сектор, должность)	
Подпись, дата (Эксперта)	

Рис. 4. Карточка-шаблон экспертного заключения набора данных для представления в экспертный совет

ную информацию, принимается решение о запрете публикации или необходимости преобразования наборов данных для исключения из них конфиденциальной информации, а также необходимости маркировки и наложения ограничений для пользователей наборов данных. Все выявленные в ходе проведения проверки

основания для ограничения публикации, в виде кодов разработанного классификатора, включаются в проект экспертного заключения о возможности открытого опубликования.

Шаг 4. В зависимости от типа и особенностей атрибутов данных (полей набора данных) к этим полям при-

меняются процедуры и алгоритмы обработки (преобразования), которые позволяют исключить конфиденциальные сведения, препятствующие открытой публикации поступившего набора данных.

Вид применяемых к набору данных преобразований, понижающих класс конфиденциальности данных, зависит от содержания самих данных и разрешения обладателя набора данных на публикацию в том или ином виде (разрешена публикация, разрешена с той или иной степенью обезличивания или деперсонификации, разрешена публикация только в анонимизированном виде, запрещена публикация).

Для набора данных, содержащих персональные данные, применяются преобразования обезличивания с разной степенью обезличивания или анонимизации. Необходимо заметить, что к методам обезличивания персональных данных не относятся криптографические методы.

Для набора данных, представленных в табличной форме, содержащего, например, статистическую или бухгалтерскую информацию, применяются преобразования деперсонификации, анонимизации или обфускации (от англ. *obfuscate* — запутывать).

Рекомендуется применять инструменты визуализации наборов данных, что не исключает требования публикации машиночитаемых данных. При визуальном представлении наборов данных, содержащих географические координаты в качестве атрибутов объектов набора данных, необходимо обеспечить корректное отображение данных объектов на карте местности в соответствии с координатными данными.

Для визуализации табличных данных необходимо применять инструменты, позволяющие осуществить предварительный просмотр данных в веб-браузере в табличном виде.

Наборы данных, которые нельзя преобразовать (видео-, фото- и аудиоинформация, рисунки в разных форматах представления данных и др.), поступают на анализ экспертам, которые должны определить, можно ли преобразовать данные в машиночитаемые форматы для применения преобразований, понижающих класс конфиденциальности. Например, информацию в наборе данных в форматах *pdf*, *jpeg* и др. необходимо преобразовать в табличный вид с помощью сканирования с распознаванием.

В паспорте набора данных делается отметка о произведенных преобразованиях.

Шаг 5. Открытая публикация возможна только в том случае, если в результате проведения повторной проверки не обнаружено сведений конфиденциального характера. Иначе переход к **Шагу 2**.

Шаг 6. В отделе управления данными формируется проект экспертного заключения, который передается на рассмотрение и утверждение в экспертный совет.

Шаг 7. Экспертный совет рассматривает полученный проект экспертного заключения и, при необходимости, сам набор данных и в утвержденном протоколе дает (или не дает) разрешение на публикацию в открытом доступе.

Шаг 8. Данные представляются в открытом доступе в виде элемента каталога с аннотацией, кратко описывающей набор данных. Иначе с этим набором данных производятся другие действия, рекомендованные экспертным советом.

Шаг 9. Для доступа к набору данных в открытом контуре пользователи заключают пользовательское соглашение, в котором фиксируются возможности, предоставляемые (определяемые) обладателями наборов данных, в частности, устанавливается:

1. *Режим использования наборов данных.*

Определяется, может ли пользователь:

- загружать данные;
- использовать, в том числе цитировать, данные только для научных и образовательных целей;
- использовать, в том числе цитировать, данные с указанием источника;
- предоставлять результаты обработки данных;
- создавать программные или справочно-информационные продукты (кроме научных и образовательных) с использованием полученных данных (*вторичное использование*).

2. *Режим доступа в отношении:*

- наборов данных;
- отдельных (конкретных) наборов данных, определенных к регламентируемому доступу;
- сведений в полях паспорта набора данных.

Кроме того, в заявке на доступ, оформляемой вместе с пользовательским соглашением, определяются области (тематика) и классы наборов данных, к которым пользователь хочет получить доступ.

Шаг 10. Если в идентификационной информации обладателей наборов данных — результатов исследований и пользовательским соглашением (заявкой) нет противоречий, то пользователь допускается к открытому контуру. В противном случае возврат к **Шагу 9**.

Если набор данных получен в результате преобразования одного или нескольких наборов данных, ранее уже рассмотренных отделом управления данными, и в паспортах исходных наборов данных есть отметка, что при достижении определенного объема они могут стать конфиденциальными, то полученный в результате исследования набор данных проверяется по всем перечням конфиденциальной информации организаций, от которых получены исходные наборы данных, а также вышестоящих организаций (материнской компании, соответствующего министерства), а в случае необходимости и по перечням информации, содержащей сведения, составляющие государственную тайну. В паспорте полученного набора данных указывается, из каких наборов данных был получен этот набор данных, и экспертами отдела управления данными определяется, в каком виде можно его публиковать. Полученные наборы данных — результаты исследований подвергаются преобразованиям по анонимизации, деперсонификации, обфускации или обезличиванию, о чем в паспорте делается соответствующая отметка.

Запрос пользователя на предоставление наборов данных

Если пользователю нужны наборы данных, он должен подать запрос для их получения. Запрос будет направлен в отдел управления данными в сектор оценки данных.

Эксперт отдела управления данными заполняет карточку запроса по шаблону, представленному ниже. При заключении пользовательского соглашения пользователю предлагается заполнить запрос (рис. 5). Если пользователь не заполняет *обязательные* поля (отмеченные * — звездочкой), то эксперт связывается с ним для уточнения существа запроса.

Карточка - шаблон Запроса для представления Пользователю набора данных / результатов исследований	
Место работы* Должность *	
Ф.И.О. *	
Адрес электронной почты *	
Дата, номер Пользовательского соглашения / дата обращения	При наличии (или проставляется Экспертом после заключения) / дд.мм.гг.
Вид деятельности (по категории) *	– Государство – Общественные организации – Бизнес – Наука и образование – СМИ – другое:
Какие данные Вы хотели бы получить	
Наименование набора данных	
Формат набора данных	Текстовый, графический, табличный, иной.
Период, за который предоставляется набор данных	
К каким областям относятся данные *	☐ Образование ☐ Электроника ☐ Досуг и отдых ☐ Туризм ☐ Метеоданные ☐ Экономика ☐ Государство ☐ Картография ☐ Здоровье ☐ Транспорт ☐ Безопасность

	☐ Строительство ☐ Торговля ☐ Экология ☐ Спорт ☐ Культура
Сталкивались ли Вы ранее с проблемами доступа к этим данным? *	☐ Да ☒ Нет
Как Вы собираетесь использовать эти данные	☐ в бизнесе ☐ в личных целях ☐ в общественной работе ☐ для исследовательской ☐ другое:
Какую пользу, на ваш взгляд, принесёт публикация запрашиваемых данных *	
Дата предоставления Паспорта и набора данных / результатов исследования	дд.мм.гг.
Способ получения набора данных	
Результат обследования набора данных экспертом Отдела публикации данных и исследований	Совпадают сведения, представленные в наборе данных с Паспортом. Вывод о необходимых действиях.
Подпись, ФИО, дата Эксперта	
Подпись, ФИО, дата Пользователя	

Рис. 5. Карточка-шаблон запроса для предоставления пользователю набора данных

Предоставляемые сведения будут сохранены и использованы для аргументации необходимости опубликования набора данных — результатов исследований.

После публикации материалов пользователь получает соответствующее уведомление на указанный им адрес электронной почты. Этот адрес может быть использован также для связи с целью уточнения деталей запроса.

Карточка-шаблон запроса для представления пользователю набора данных — результатов исследований является обязательным приложением к пользовательскому соглашению.

При передаче наборов данных для публикации с обладателем наборов данных должно быть заключено соглашение об информационном взаимодействии. Эксперт отдела управления данными заполняет карточку экспертного заключения по шаблону, представленному в приложении данного соглашения.

Соглашение об информационном взаимодействии должно включать следующие обязательные пункты:

- описаны возможные форматы наборов данных, которые будут поступать — с предпочтением машиночитаемых форматов;
- обладатель наборов данных передает перечни сведений персональных данных, коммерческой, профессиональной или служебной тайны организации — обладателя набора данных;
- обладатель наборов данных сообщает о пунктах перечня сведений, составляющих государственную тайну, которые определены его организации федеральным регулятором в области защиты государственной тайны;
- в паспорте набора данных должно быть отражено, включает ли передаваемый набор данных сведения, содержащие конфиденциальную информацию (персональные данные, коммерческую, профессиональную или служебную тайну);
- если передаваемый набор данных содержит сведения, содержащие конфиденциальную инфор-

мацию, то должны быть ссылки на пункты соответствующего перечня сведений;

- если передаваемый набор данных содержит сведения, представляющие собой охраняемые *результаты интеллектуальной деятельности* [11, 13] (объекты авторских или смежных прав), должна быть ссылка на лицензионные договоры с правообладателями;
- в паспорте набора данных должно быть отражено, в каком виде разрешается публикация информации: в неизменном виде, в обезличенном (степень обезличивания данных), деперсонифицированном или анонимизированном;
- определен порядок взаимодействия эксперта отдела управления данными в том случае, если передаваемый обладателем набор данных после рассмотрения будет признан экспертом как содержащий сведения, составляющие личную (персональные данные), коммерческую, профессиональную или служебную тайну, например, по перечням вышестоящей организации в корпорации.

В приложении к соглашению об информационном взаимодействии должна быть представлена карточка-шаблон экспертного заключения набора данных для представления в экспертный совет.

Заключение

Таким образом, рассмотренный протокол и организационно-технический регламент обработки наборов данных для их публикации в открытых источниках является исходным для разработки комплекса нор-

мативных документов, регламентирующих процессы и методики определения степени конфиденциальности, разработки регламентов обработки и публикации данных, а также проектов функциональной структуры экспертного совета и отдела управления данными для однозначного определения возможности публикации в открытом доступе.

Определены и описаны группы обладателей и потребителей данных и их права на выполнение различных действий с наборами данных. Для каждой из категорий обладателей и потребителей наборов данных определены основные субъекты, а также накладываемые ими ограничения.

Сформирована организационно-функциональная модель публикации наборов данных, включающая концептуальный вариант подхода к организации информационных процессов экспертного совета, принимающего решения о публикации наборов данных. Определены базовые элементы функциональной структуры отдела управления данными, реализующего организационно-функциональную модель публикации наборов данных.

Протокол и регламент публикации наборов данных рекомендуется использовать в составе методического и в инфраструктуре информационно-программного обеспечения существующих и перспективных информационных платформ организаций, принимающих, обрабатывающих и публикующих наборы данных с целью повышения уровня публичной открытости наборов данных, с одной стороны, и соблюдения обоснованных ограничений на доступ к ним и обеспечения их информационной безопасности, с другой.

Литература

1. Белин А.В., Ващекин А.Н. О проблеме дифференциации коммерческой, служебной и иной охраняемой законом тайны // Российское правосудие. 2019. № 6. С. 32—41.
2. Борисов Р.С. Эффективный алгоритм управления переработкой судебной статистической информации // Правовая информатика. 2018. № 1. С. 15—22. DOI: 10.21681/1994-1404-2018-1-15-22.
3. Борисов Р.С. Аппаратно-ориентированная модель взаимодействия открытых систем // Электротехнические и информационные комплексы и системы. 2011. Т. 7. № 1. С. 21—25.
4. Ващекин А.Н., Дзедзинский А.В. Правовое регулирование отношений в цифровом пространстве // Правосудие. 2020. № 2. С. 108—114.
5. Ефименко А.А. Государственный финансовый контроль: принципы и проблемы реализации // Труды Междунар. науч.-практ. конф. «Роль бухгалтерского учета и налогообложения в финансовом развитии бизнеса» (19 октября 2017 г.) / АПНИ. М. : Агентство персп. науч. исслед., 2017. С. 46—49.
6. Ефименко А.А. Интегрированная интеллектуальная технология оптимизации параллельных алгоритмов в высокопроизводительных вычислительных системах // Труды Междунар. науч.-практ. конф. «Современные тенденции в науке, технике, образовании» (31 января 2016 г.). Смоленск : Новаленко, 2016. С. 59—61.
7. Коваленко А.О. Протокол рациональной переработки и правовые режимы судебной информации // Правовая информатика. 2019. № 2. С. 49—56. DOI 10.21681/1994-1404-2019-2-49-56.
8. Ловцов Д.А. Проблема гарантированного обеспечения информационной безопасности крупномасштабных автоматизированных систем // Правовая информатика. 2017. № 3. С. 66—74. DOI: 10.21681/1994-1404-2017-3-66-74.
9. Ловцов Д.А. Системология правового регулирования информационных отношений в инфосфере. М. : РГУП, 2016. 316 с. ISBN 978-5-93916-505-1.
10. Ловцов Д.А. Основы технологии эффективного двухуровневого правового регулирования информационных отношений в инфосфере // Правовая информатика. 2018. № 2. С. 4—14. DOI: 10.21681/1994-1404-2018-2-4-14.

11. Ловцов Д.А. Эффективные методы защиты результатов интеллектуальной деятельности в инфосфере глобальных телематических сетей // Открытое образование. 2016. № 5. С. 85—88.
12. Ловцов Д.А. Информационная теория эргасистем : монография. М. : РГУП, 2021. 314 с. ISBN 978-5-93916-887-8.
13. Ловцов Д.А., Богданова М.В. Экономико-правовое регулирование оборота результатов интеллектуальной деятельности предприятий промышленности России // Экономика, статистика и информатика. Вестник УМО. 2013. № 1. С. 53—56.
14. Ловцов Д.А., Федичев А.В. Архитектура национального классификатора правовых режимов информации ограниченного доступа // Правовая информатика. 2017. № 2. С. 35—54. DOI: 10.21681/1994-1404-2017-2-35-54.
15. Осовский С. Нейронные сети для обработки информации. М. : Финансы и статистика, 2004. 344 с.
16. Федосеев С.В. Применение современных технологий больших данных в правовой сфере // Правовая информатика. 2018. № 4. С. 50—58. DOI: 10.21681/1994-1404-2018-4-5-58.

Рецензент: **Омельченко Виктор Валентинович**, доктор технических наук, профессор, заслуженный деятель науки и техники Российской Федерации, государственный советник Российской Федерации 1-го класса, советник секретариата научно-технического совета АО «ВПК «НПО машиностроения», Российская Федерация, г. Москва.
E-mail: omvv@yandex.ru

A DATA SETS PROCESSING PROTOCOL FOR THEIR PUBLICATION IN OPEN SOURCES

Roman Borisov, Ph.D. (Technology), Associate Professor at the Department of Information Technology Law, Informatics and Mathematics of the Russian State University of Justice, Moscow, Russian Federation.
E-mail: bestseller@bk.ru

Aleksei Efimenko, Ph.D. (Technology), Associate Professor at the Department of Information Technology Law, Informatics and Mathematics of the Russian State University of Justice, Moscow, Russian Federation.
E-mail: alex192@mail.ru

Keywords: open data, data set structure, metadata, data passport, protocol, data sets publication, confidentiality, classification attributes.

Abstract.

Purpose of the paper: improving the scientific and methodological base for regulating the publication in open sources of data sets possibly containing confidential information.

Methods used: system and legal analysis of external and attributive properties and organisational techniques for processing data sets allowing to ensure the confidentiality of information when published in open sources.

Results obtained: groups of data sets owners and consumers are identified, their rights, main subjects, as well as restrictions imposed by them are determined. An organisational and technical regulation for data sets publication is set up including a conceptual version of the approach to information processes organisation when publishing data sets. An iterative protocol for processing data sets when they are published is developed making it possible to identify the main procedures for data sets processing, analysis and transformation.

References

1. Belin A.V., Vashchekin A.N. O probleme differentsiatsii kommercheskoi, sluzhebnoi i inoi okhraniaemoi zakonom tainy. Rossiiskoe pravosudie, 2019, No. 6, pp. 32-41.
2. Borisov R.S. Effektivnyi algoritm upravleniia pererabotkoi sudebnoi statisticheskoi informatsii. Pravovaia informatika, 2018, No. 1, pp. 15-22. DOI: 10.21681/1994-1404-2018-1-15-22.
3. Borisov R.S. Apparato-orientirovannaia model' vzaimodeistviia otkrytykh sistem. Elektrotekhnicheskie i informatsionnye kompleksy i sistemy, 2011, t. 7, No. 1, pp. 21-25.
4. Vashchekin A.N., Dziedzinskii A.V. Pravovoe regulirovanie otnoshenii v tsifrovom prostranstve. Pravosudie, 2020, No. 2, pp. 108-114.
5. Efimenko A.A. Gosudarstvennyi finansovyi kontrol': printsipy i problemy realizatsii. Trudy Mezhdunar. nauch.-prak. konf. "Rol' bukhgalterskogo ucheta i nalogoblozheniia v finansovom razvitii biznesa" (19 oktiabria 2017 g.). APNI. M. : Agentstvo persp. nauch. issled., 2017, pp. 46-49.

6. Efimenko A.A. Integrirovannaya intellektual'naya tekhnologiya optimizatsii parallel'nykh algoritmov v vysokoproizvoditel'nykh vychislitel'nykh sistemakh. Trudy Mezhdunar. nauch.-prak. konf. "Sovremennye tendentsii v nauke, tekhnike, obrazovanii" (31 ianvaria 2016 g.). Smolensk : Novalenso, 2016, pp. 59-61.
7. Kovalenko A.O. Protokol ratsional'noi pererabotki i pravovye rezhimy sudebnoi informatsii. Pravovaia informatika, 2019, No. 2, pp. 49-56. DOI: 10.21681/1994-1404-2019-2-49-56 .
8. Lovtsov D.A. Problema garantirovannogo obespecheniia informatsionnoi bezopasnosti krupnomasshtabnykh avtomatizirovannykh sistem. Pravovaia informatika, 2017, No. 3, pp. 66-74. DOI: 10.21681/1994-1404-2017-3-66-74 .
9. Lovtsov D.A. Sistemologiya pravovogo regulirovaniia informatsionnykh otnoshenii v infosfere. M. : RGUP, 2016. 316 s. ISBN 978-5-93916-505-1.
10. Lovtsov D.A. Osnovy tekhnologii effektivnogo dvukhurovneвого pravovogo regulirovaniia informatsionnykh otnoshenii v infosfere. Pravovaia informatika, 2018, No. 2, pp. 4-14. DOI: 10.21681/1994-1404-2018-2-4-14 .
11. Lovtsov D.A. Effektivnye metody zashchity rezul'tatov intellektual'noi deiatel'nosti v infosfere global'nykh telematicheskikh setei. Otkrytoe obrazovanie, 2016, No. 5, pp. 85-88.
12. Lovtsov D.A. Informatsionnaya teoriia ergasistem : monografiia. M. : RGUP, 2021. 314 s. ISBN 978-5-93916-887-8.
13. Lovtsov D.A., Bogdanova M.V. Ekonomiko-pravovoe regulirovanie oborota rezul'tatov intellektual'noi deiatel'nosti predpriatii promyshlennosti Rossii. Ekonomika, statistika i informatika. Vestnik UMO, 2013, No. 1, pp. 53-56.
14. Lovtsov D.A., Fedichev A.V. Arkhitektura natsional'nogo klassifikatora pravovykh rezhimov informatsii ogranichenogo dostupa. Pravovaia informatika, 2017, No. 2, pp. 35-54. DOI: 10.21681/1994-1404-2017-2-35-54 .
15. Osovskii S. Neironnye seti dlia obrabotki informatsii. M. : Finansy i statistika, 2004. 344 s.
16. Fedoseev S.V. Primenenie sovremennykh tekhnologii bol'shikh dannykh v pravovoi sfere. Pravovaia informatika, 2018, No. 4, pp. 50-58. DOI: 10.21681/1994-1404-2018-4-5-58 .