

«БОЛЬШИЕ ДАННЫЕ» В КРИМИНАЛИСТИКЕ: ОСНОВНЫЕ ПОНЯТИЯ И ДЕФИНИЦИЯ

В настоящей статье рассматриваются вопросы терминологии в сфере «больших данных», которая успешно и эффективно может быть использована в криминалистических целях. Дается сравнительный анализ общепринятого (заимствованного, публичного) и юридического понятийного аппарата. Автор анализирует зарубежные термины и отечественное законодательство сферы информации, данных и технологий. В статье рассматривается возможность и обосновывается необходимость использования в криминалистике уже сложившихся и операционализированных терминов, воспринятых и «определенных» российским законодательством, что даст понятийное единообразие.

Ключевые слова: «большие данные», цифровая криминалистика, общедоступная информация, открытые данные, электронно-цифровые следы.

S. A. Mashkov

«BIG DATA» IN CRIMINALISTIK: BASIC CONCEPTS AND DEFINITIONS

This article discusses terminology issues in the field of «big data», which can be successfully and effectively used for forensic purposes. A comparative analysis of the generally accepted (borrowed, public) and legal conceptual framework is given. The author analyzes foreign terms and domestic legislation in the field of information, data and technology. The article examines the possibility and justifies the need to use already established and operationalized terms in criminalistics, accepted and «defined» by Russian legislation, which will give conceptual uniformity.

Keywords: «big data», digital forensics, publicly available information, open data, electronic digital traces.

В сфере информационных технологий научное сообщество (в том числе криминалистическое) активно заимствует западную и международную терминологию. С одной стороны, это обоснованный процесс, но бывает, что это неоправданно в силу терминологической несостыковки с базовыми понятиями, закрепленными в нормах российского права, а также в силу того, что отечественная правовая система уже содержит термины – аналоги западным (при этом более точные и информативные). В то же время далеко не все и не всегда заимствованные понятия употребляют правильно и адекватно. При этом разные авторы одно и то же понятие могут обозначать разными терминами, и, наоборот, одно понятие определять разными содержательными дефинициями.

Российский законодатель (к его чести) очень оперативно и системно реагирует на запросы практики, а нормы права очень быстро закрепляют регла-

ментационные термины, которые операционализированы и представлены правоприменителю.

Сравним определения международной практики, российских стандартов в сфере «информации и данных» с терминами, закрепленными в российском законодательстве.

Начнем с базового понятия – понятия «информация» – англ. «information».

В международных стандартах (ISO 5127:201; ISO/IEC 10746-2:1996; ISO/IEC 2382:2015; ISO/IEC/IEEE 24765:2010) даются разные определения, которые сводятся к тому, что информация – это знания о предметах, фактах, идеях, событиях, вещах и понятиях, имеющие смысл (контекст, интерпретация), обработанные, организованные и связанные, представленные в виде данных, что позволяет ими обмениваться¹.

Межгосударственный стандарт «Система стандартов по информации, библиотечному и издательскому делу» (ГОСТ 7.0-99), введенный в действие непосредственно в качестве государственного стандарта Российской Федерации с 1 июля 2000 г.², дает следующее определение информации: сведения, воспринимаемые человеком и (или) специальными устройствами как отражение фактов материального или духовного мира в процессе коммуникации (п. 3.1.19).

Российское законодательство (Федеральный закон от 27 июля 2006 г. № 149-ФЗ «Об информации, информационных технологиях и о защите информации») дает очень краткую (лаконичную) и в тоже время объемную (лапидарную) дефиницию информации – «сведения (сообщения, данные) независимо от формы их представления»³. Данное понятие точное, универсальное, не содержит излишних признаков и характеристик.

При этом значение терминов «данные» и «сообщения» раскрыты в ГОСТ 7.0-99: данные – это информация, обработанная и представленная в формализованном виде для дальнейшей обработки (3.2.1.2.), а сообщение – это информация, переданная и/или полученная в процессе коммуникации (3.1.23.). Также Национальный стандарт Российской Федерации (ГОСТ Р ИСО/МЭК 20546-2019) «Информационные технологии. Большие данные. Обзор и словарь»⁴ дает следующее определение «данных» (data) – представление информации в формальном виде, пригодном для передачи, интерпретации или обработки (п. 3.1.5). Межгосударственный стандарт (ГОСТ 33707-2016 (ISO/IEC 2382:2015)). «Информационные технологии. Словарь»⁵ понимает под «данными» – представление информации в формализованном (формальном) виде, при-

¹ Википедия. URL: <https://ru.wikipedia.org/wiki/Информация> (дата обращения: 01.12.2024).

² О принятии и введении в действие межгосударственного стандарта : постановление Государственного комитета Российской Федерации по стандартизации и метрологии от 07.10.1999 № 334-ст // СПС «КонсультантПлюс».

³ Об информации, информационных технологиях и о защите информации : федер. закон от 27.07.2006 № 149-ФЗ // Собрание законодательства РФ. 2006. № 31 (1 ч.). Ст. 3448. (п. 1 ст. 2).

⁴ Национальный стандарт Российской Федерации (ГОСТ Р ИСО/МЭК 20546-2019) «Информационные технологии. Большие данные. Обзор и словарь». М. : Стандартинформ, 2021.

⁵ Межгосударственный стандарт (ГОСТ 33707-2016 (ISO/IEC 2382:2015)). Информационные технологии. Словарь (введен в действие Приказом Росстандарта от 22.09.2016 № 1189-ст). М. : Стандартинформ, 2021.

годном для коммуникации, передачи, интерпретации или обработки людьми и компьютерами (п.4.259).

Следующее понятие, получившее широкое распространение в 2002 года, – «открытый доступ» (англ. «open access»), под которым понимались «принципы открытого доступа к научно-исследовательским работам, то есть свободный, ничем не ограничиваемый доступ в сети Интернет к научно-исследовательским публикациям, подразумевающий возможность любого их использования»¹.

В отечественном законодательстве дано соответствующее понятие «доступ к информации – возможность получения информации и ее использования»².

Сегодня широко используется термин «большие данные» – англ. «big data», который возник и стал применяться в мире с 2008 года³. Им «обозначают структурированные и неструктурированные данные огромных объёмов и значительного многообразия», которые невозможно эффективно обработать с использованием традиционных методов, и для этого применяют специальные технологии и программное обеспечение: системы распределенных хранилищ данных; технологии обработки потоков данных; методы машинного обучения; алгоритмы анализа, позволяющие определять ценность и значение информации и извлекать её из огромных объемов данных⁴ [1]. Определяющими характеристиками больших данных (5V) являются [1; 2]: объём (англ. «volume») – огромные физические объёмы информации, которые включают в себя терабайты, петабайты и даже эксабайты данных; разнообразие (англ. «variety») – возможность одновременной обработки различных типов и форматов данных, включая неструктурированные данные (текст, изображения, видео, аудио) и структурированные данные (таблицы и базы данных), а также данные в реальном времени и данные с географическими координатами; скорость (англ. «velocity») – быстрое время обработки и анализа данных в реальном времени (высокоскоростные технологии обработки) даже с учетом того, что информация может поступать со скоростью нескольких тысяч транзакций в секунду (скорость прироста); истинность (англ. «veracity») – качество данных, включая точность, надежность и актуальность; ценность (англ. «value») – способность извлекать ценную информацию и знания из больших данных и использовать их в исследованиях.

Национальный стандарт Российской Федерации (ГОСТ Р ИСО/МЭК 20546-2019) «Информационные технологии. Большие данные. Обзор и словарь» дает следующее определение термина «большие данные» – большие массивы данных, имеющих такие характеристики как объём, разнообразие, скорость обработки и/или вариативность, и требующие использования технологий

¹ Budapest Open Access Initiative. The original Declaration and guidelines to make research free and available to anyone with internet access and promote advances in the sciences, medicine, and health. URL: <https://www.budapestopenaccessinitiative.org> (дата обращения: 01.12.2024).

² Об информации, информационных технологиях и о защите информации : федер. закон от 27.07.2006 № 149-ФЗ // Собрание законодательства РФ. 2006. № 31 (1 ч.). Ст. 3448. (п. 6 ст. 2).

³ Volume 455 Issue 7209, 4 September 2008. URL: <https://www.nature.com/nature/volumes/455/issues/7209> (дата обращения: 01.12.2024).

⁴ Preimesberger, Chris Hadoop, Yahoo, 'Big Data' Brighten BI Future / EWeek (15 августа 2011). URL <https://www.eweek.com/storage/hadoop-yahoo-big-data-brighten-bi-future> (дата обращения: 01.12.2024).

масштабирования для эффективного хранения, обработки, управления и анализа (п. 3.1.2).

Законодательство Российской Федерации использует понятие «общедоступная информация», которая определяется как «общеизвестные сведения и иная информация, доступ к которой не ограничен, в том числе, информация, размещаемая ее обладателями в сети Интернет в формате, допускающем автоматизированную обработку без предварительных изменений человеком в целях повторного ее использования (общедоступная информация, размещаемая в форме открытых данных)»¹. Также действующее законодательство использует понятие «открытые данные», которыми считается «информация, размещаемая ее обладателями в сети Интернет в формате, допускающем автоматизированную обработку без предварительных изменений человеком в целях повторного ее использования»².

Следовательно, «большие данные» по российскому законодательству являются общедоступной информацией, размещаемой в форме открытых данных.

Таким образом, мы видим, что действующее законодательство Российской Федерации, содержит и определяет ряд понятий, аналоги которых заимствованы из зарубежной практики. Они также операционализированы в терминологической документации. Эти понятия могут быть использованы в разных сферах человеческой деятельности, в том числе, в сфере криминалистики. При этом необходимость попыток дать им иные определения излишни. Существующие дефиниции специальные, профессионально определенные и применимые во всех отраслях знаний.

Список использованной литературы

1. Наташкин А. Что такое Big Data: как собирают и где применяют большие данные? / А. Наташкин. – URL: <https://lenta.ru/articles/2023/11/27/chto-takoe-big-data/?ysclid=lwccvccbor627380342> (дата обращения: 01.12.2024).

2. Канаракус К. Машина Больших Данных / К. Канаракус // Сети. – 2011. – № 4. – URL: <https://www.osp.ru/nets/2011/04/13010802> (дата обращения: 01.12.2024).

Информация об авторе

Машков Сергей Александрович – кандидат юридических наук, доцент кафедры криминалистики, судебных экспертиз и юридической психологии, Байкальский государственный университет, Российская Федерация, г. Иркутск.

Author

Mashkov Sergey Alexandrovich – Candidate of Law, Ass. Professor of the Department of Criminalistics, Forensic Examinations and Legal Psychology, Baikal State University, the Russian Federation, Irkutsk.

¹ Об информации, информационных технологиях и о защите информации : федер. закон от 27.07.2006 № 149-ФЗ // Собрание законодательства Российской Федерации. 2006. № 31 (1 ч.). Ст. 3448. Ч. 1. Ч. 4, ст. 7.

² «Типовые условия использования общедоступной информации, размещаемой в информационно-телекоммуникационной сети «Интернет» в форме открытых данных» (утв. протоколом заочного голосования Правительственной комиссии по координации деятельности открытого правительства от 19 сентября 2016 г. № 6) // СПС «КонсультантПлюс».