



ПРЕДВАРИТЕЛЬНЫЙ
НАЦИОНАЛЬНЫЙ
СТАНДАРТ
РОССИЙСКОЙ
ФЕДЕРАЦИИ

ПНСТ
848—
2023

Искусственный интеллект БОЛЬШИЕ ДАННЫЕ

Обзор и требования по обеспечению сохранности данных

(ITU-T Y.3604(2020), Big data — Overview and requirements for data preservation,
MOD)

Издание официальное

Москва
Российский институт стандартизации
2023

Предисловие

1 ПОДГОТОВЛЕН Научно-образовательным центром компетенций в области цифровой экономики Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный университет имени М.В. Ломоносова» (МГУ имени М.В. Ломоносова) и Обществом с ограниченной ответственностью «Институт развития информационного общества» (ИРИО) на основе собственного перевода на русский язык англоязычной версии стандарта, указанного в пункте 4

2 ВНЕСЕН Техническим комитетом по стандартизации ТК 164 «Искусственный интеллект»

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 13 декабря 2023 г. № 94-пнст

4 Настоящий стандарт является модифицированным по отношению к международному стандарту МСЭ-Т Y.3604 (02/2020) «Большие данные. Обзор и требования по обеспечению сохранности данных» (Recommendation ITU-T Y.3604 (02/2020) «Big data — Overview and requirements for data preservation», MOD) путем дополнительных положений, выделенных курсивом, а также невключения отдельных структурных элементов для учета особенностей национальной стандартизации в области искусственного интеллекта и работы с большими данными, объяснение причин их внесения приведено во введении к настоящему стандарту.

Наименование настоящего стандарта изменено относительно наименования указанного международного стандарта для приведения в соответствие с ГОСТ Р 1.5—2012 (пункт 3.5).

Сопоставление структуры настоящего стандарта со структурой примененного в нем международного стандарта приведено в дополнительном приложении ДА

Правила применения настоящего стандарта и проведения его мониторинга установлены в ГОСТ Р 1.16—2011 (разделы 5 и 6).

Федеральное агентство по техническому регулированию и метрологии собирает сведения о практическом применении настоящего стандарта. Данные сведения, а также замечания и предложения по содержанию стандарта можно направить не позднее чем за 4 мес до истечения срока его действия разработчику настоящего стандарта по адресу: 119991, Российская Федерация, Москва, Ленинские горы, д. 1 и в Федеральное агентство по техническому регулированию и метрологии по адресу: 123112 Москва, Пресненская набережная, д. 10, стр. 2.

В случае отмены настоящего стандарта соответствующая информация будет опубликована в ежемесячном информационном указателе «Национальные стандарты» и также будет размещена на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет (www.rst.gov.ru)

© Оформление. ФГБУ «Институт стандартизации», 2023

Настоящий стандарт не может быть полностью или частично воспроизведен, тиражирован и распространен в качестве официального издания без разрешения Федерального агентства по техническому регулированию и метрологии

II

Содержание

1 Область применения	1
2 Термины и определения	1
3 Соглашения по терминологии.	2
4 Обзор сохранности больших данных	3
4.1 Общая концепция сохранности данных	3
4.2 Сохранность данных в экосистеме больших данных	3
4.3 Категории данных при сохранности больших данных	4
4.4 Пакеты данных при сохранности больших данных	5
4.5 Функциональные объекты для управления сохранностью больших данных	5
5 Функциональные требования к сохранности больших данных	6
5.1 Отбор данных, подлежащих сохранению	6
5.2 Хранение данных, подлежащих сохранению	6
5.3 Доступ к сохраненным данным	6
5.4 Управление политикой сохранности данных	6
6 Обеспечение безопасности.	7
Приложение А (справочное) Варианты использования методов сохранности больших данных	8
Приложение ДА (справочное) Сопоставление структуры настоящего стандарта со структурой примененного в нем международного стандарта.	13
Библиография	15

Введение

В настоящем стандарте представлена общая концепция сохранности данных, описан процесс сохранности в экосистеме больших данных, перечислены категории данных, пакеты данных и функциональные объекты для сохранности больших данных. Кроме того, приведены функциональные требования к сохранности больших данных и требования и рекомендации, которые необходимо учитывать при обеспечении безопасности больших данных, а также варианты использования методов сохранности больших данных.

Соблюдение требований настоящего стандарта является добровольным, за исключением определенных обязательных требований (например, необходимых для обеспечения функциональной совместимости или применимости), которые должны быть выполнены. Для выражения требований использованы такие слова, как «обязан», или другие императивные формулировки, например «должен», а также их отрицательные эквиваленты. Употребление данных слов не означает, что настоящий стандарт является обязательным для какой-либо из сторон.

В настоящий стандарт включены дополнительные положения для учета потребностей национальной экономики Российской Федерации и особенностей российской национальной стандартизации, выделенные курсивом, а именно:

- изменен элемент «Введение»;
- в 2.1 изменено определение термина в соответствии с требованиями ГОСТ Р ИСО/МЭК 20546—2021, статья 3.1.2;
- в раздел 2 «Термины и определения» добавлены пункты 2.2—2.7 для взаимопонимания между различными пользователями настоящего стандарта.

ПРЕДВАРИТЕЛЬНЫЙ НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

Искусственный интеллект

БОЛЬШИЕ ДАННЫЕ

Обзор и требования по обеспечению сохранности данных

Artificial intelligence. Big data. Overview and requirements for data preservation

Срок действия — с 2024—02—01
до 2027—02—01

1 Область применения

В настоящем стандарте приведен обзор и установлены требования к сохранности больших данных.

В настоящем стандарте рассматриваются следующие темы:

- обзор сохранности больших данных;
- функциональные требования к сохранности больших данных;
- варианты использования методов сохранности больших данных.

2 Термины и определения

В настоящем стандарте применены следующие термины с соответствующими определениями:

2.1

большие данные (big data): Большие массивы данных, отличающиеся главным образом такими характеристиками, как объем, разнообразие, скорость обработки и/или вариативность, которые требуют использования технологии масштабирования для эффективного хранения, обработки, управления и анализа.

Примечание — Термин «большие данные» широко применяется в различных значениях, например в качестве наименования технологии масштабирования, используемой для обработки больших массивов данных.

[ГОСТ Р ИСО/МЭК 20546—2021, статья 3.1.2]

2.2

данные (data): Представление информации в формальном виде, пригодном для передачи, интерпретации или обработки.

Примечание — Данные могут быть обработаны автоматически или вручную.

[ГОСТ Р ИСО/МЭК 20546—2021, статья 3.1.5]

2.3

изменчивость данных (data volatility): Характеристика данных, относящаяся к скорости их изменения с течением времени.

[ГОСТ Р ИСО/МЭК 20546—2021, статья 3.1.17]

2.4

массив данных (data set, dataset): Идентифицируемая совокупность данных, к которой можно получить доступ или скачать в одном или нескольких форматах.

[ГОСТ Р ИСО/МЭК 20546—2021, статья 3.1.11]

2.5

обработка данных (data processing): Систематическое выполнение операций с данными.

Примечания

1 Арифметические или логические операции с данными, объединение или сортировка данных или такие операции с текстом, как редактирование, сортировка, объединение, хранение, извлечение, отображение или печать.

2 Термин «обработка данных» не должен использоваться в качестве синонима для термина «обработка информации».

[ГОСТ Р ИСО/МЭК 20546—2021, статья 3.1.9]

2.6

тип данных (data type, datatype): Совокупность объектов данных установленной структуры и набора допустимых операций над этими объектами.

Примечания

1 Целочисленный тип данных имеет простую структуру, каждый экземпляр которой, обычно называемый значением, представляет собой одно из целых чисел из заданного диапазона, а допустимые действия включают в себя обычные арифметические операции над этими целыми числами.

2 Если толкование не вызывает сомнений, то вместо термина «тип данных» может быть использован термин «тип».

3 Тип данных: определение и термины, стандартизованные ИСО/МЭК [ИСО/МЭК 2382-15:1999].

[ГОСТ Р ИСО/МЭК 20546—2021, статья 3.1.12]

2.7

файл (file): Поименованная совокупность записей, рассматриваемая как единое целое.

[ГОСТ Р ИСО/МЭК 20546—2021, статья 3.1.21]

2.8 сохранность больших данных (big data preservation): Комплекс мероприятий, стратегий и политик, разработанный сервис-провайдером больших данных, для обеспечения доступности и использования ценной информации в экосистеме больших данных.

Примечание — Экосистема больших данных определяет необходимые действия для ролей, предоставляющих или использующих сервисы больших данных, а также взаимоотношения между ними (см. [1]).

2.9 сохранность данных (data preservation): Политики и мероприятия по обеспечению постоянного доступа к данным.

2.10 политика сохранности данных (data preservation policy): Набор правил для контроля мероприятий по сохранности данных.

3 Соглашения по терминологии

В настоящем стандарте:

- ключевые слова «требуется, чтобы» означают требование, которое должно строго соблюдаться и отклонение от которого не допускается, если будет сделано заявление о соответствии настоящему стандарту;

- ключевое слово «рекомендуется» означает требование, которое рекомендуется, но не является абсолютно необходимым. Таким образом, это требование не является обязательным для заявления о соответствии настоящему стандарту.

4 Обзор сохранности больших данных

В данном разделе представлен обзор сохранности больших данных: приведена общая концепция сохранности данных, а также проиллюстрирован процесс сохранности данных в экосистеме больших данных. В разделе также описаны категории данных, пакеты данных и функциональные объекты для сохранности больших данных.

4.1 Общая концепция сохранности данных

Сохранность данных — это политики и мероприятия, направленные на обеспечение постоянного доступа к данным. Обеспечение сохранности и целостности данных и их метаданных для дальнейшего использования является основополагающим принципом сохранности данных. Целью сохранности данных являются защита данных от их потери или разрушения и доступность данных в течение длительного времени.

Потеря или разрушение данных происходит по разным причинам, в том числе из-за износа и старения физических носителей данных, исчезновения аппаратного или программного обеспечения, используемого для чтения данных, невозможности прочтения формата файлов, содержащих данные, и т. д.

Сохранность данных расширяет концепцию наличия или обладания данными или резервными копиями данных. Сохранность данных обеспечивает постоянный доступ к данным путем планирования стратегий резервного копирования и восстановления, предшествующих аварийным ситуациям или технологическим изменениям.

4.2 Сохранность данных в экосистеме больших данных

В экосистеме больших данных сохранность данных, как правило, включает в себя комплекс мероприятий, стратегий и политик, направленных на обеспечение доступности и использования ценной информации сервис-провайдером больших данных.

В дальнейшем термин «сохранность больших данных» использован для обозначения сохранности данных в экосистеме больших данных.

4.2.1 Проблемы сохранности больших данных

Следующие характеристики больших данных (см. 6.1 [1]) создают проблемы для сохранности больших данных:

а) объем: характеризует объемы собранных, сохраненных, проанализированных и визуализированных данных, которыми необходимо оперировать с помощью технологий больших данных. С этой характеристикой связаны следующие проблемы:

- 1) как масштабировать хранилища данных и
- 2) как определять данные, подлежащие сохранению.

Примечание — Данные, которые напрямую не используют, могут представлять собой ценность косвенно, например для аналитики. Поэтому может быть сложнее принять решение о том, что сохранять и когда предпринимать действия по обеспечению сохранности больших объемов данных;

б) разнообразие: характеризует различные типы и форматы данных, которые обрабатываются с помощью технологий больших данных. Для каждого типа данных может потребоваться отдельная программно-аппаратная среда для оперирования данными. С этой характеристикой связана следующая проблема:

- 1) как сохранять различные медиа и контекст их обработки;

с) скорость обработки данных: характеризует то, с какой скоростью происходит комплектование данных и с какой скоростью обрабатываются данные с помощью технологий больших данных для получения ожидаемых результатов. С этой характеристикой связана еще одна проблема сохранности больших данных:

- 1) как максимально автоматизировать деятельность по обеспечению сохранности больших данных.

4.2.2 Стратегии сохранности больших данных

Согласно рекомендации МСЭ (см. [1]) сервис-провайдер больших данных поддерживает мероприятия по управлению данными, связанные с происхождением, конфиденциальностью и безопасностью данных, а также с политикой хранения данных, владением данными и т. д. Хотя это явно не отмечено, сохранность данных также является одним из видов деятельности по управлению данными и обеспечивает постоянный доступ к данным в соответствии с политикой сохранности данных, которая представляет собой набор правил для контроля мероприятий по сохранности данных.

Для преодоления вышеуказанных трудностей сервис-провайдер больших данных может принять следующие стратегии сохранности больших данных:

- автоматизированные операции: операции по обеспечению сохранности больших данных включают в себя отбор и подготовку данных для сохранения, непосредственно сохранение и извлечение сохраненных данных. Из-за огромного объема и высокой скорости обработки больших данных сервис-провайдеру больших данных для повышения эффективности следует поддерживать эти операции и взаимосвязи между ними в автоматическом режиме;
- поддержка гибкой политики сохранности данных: сохранность больших данных должна быть основана на соответствующей политике. Какие данные следует сохранять, какой тип носителя следует использовать и как долго сохраненные данные следует сохранять — все это примеры положений, которые должны быть определены в гибких политиках сохранности данных. Сохранность больших данных, основанная на соответствующей политике, может помочь сервис-провайдеру больших данных преодолеть трудности, связанные с объемом и скоростью обработки данных;
- легкое масштабирование и иерархическое хранение данных: большие объемы данных нуждаются в хранилищах большой емкости. Для поддержания баланса экономичности и эффективности использования хранилища для сохранности данных сервис-провайдеру больших данных следует применять иерархическую систему хранения данных, в которой каждый уровень системы должен легко масштабироваться.

4.3 Категории данных при сохранности больших данных

Ниже приведены различные категории данных, учитываемые при сохранности больших данных:

- контент: необработанные данные, подлежащие сохранению;
- метаданные: данные о данных или элементах данных, включая их описания, а также данные о владельцах данных, путях и правах доступа и изменчивости данных. Общая концепция метаданных и их применение в экосистеме больших данных описаны в [2]. Для сохранности больших данных необходимы следующие типы метаданных:
 - дескриптивные метаданные: описывают структуру и семантику контента (см. 8.2.3 [2]). Они также включают в себя иные метаданные, которые могут потребоваться для понимания контента, такого как программное обеспечение, алгоритмы и т. д.,
 - нормативно-справочные метаданные: предоставляют информацию о доступе и для обеспечения поиска (например, тематическая категория и ключевые слова) контента (см. 8.2.3 [2]),
 - метаданные о происхождении: документируют историю контента (см. 8.2.6 [2]),
 - метаданные о целостности: обеспечивают проверку целостности данных для защиты контента от недокументированных изменений,
 - метаданные о правах доступа: содержат условия доступа к контенту (см. 8.2.3 [2]);
 - метаданные о правилах: правила, установленные сервис-провайдером больших данных, которыми следует руководствоваться и которые следует соблюдать при выполнении действий по сохранности больших данных. Репрезентативные правила сохранности больших данных включают правила отбора данных для хранения, которые определяют, какие виды данных сервис-провайдеру больших данных следует хранить, правила определения периода хранения данных, которые определяют, в течение какого периода сервис-провайдеру больших данных следует хранить данные, правила хранения сохраняемых данных, в соответствии с которыми определяют, на каком из уровней хранения должны размещаться эти данные и при каких условиях их следует переместить на другие уровни хранения и т. д.

Примечание — Например, для автоматизации процесса отбора данных для хранения сервис-провайдер больших данных может устанавливать правило отбора данных для хранения, определяющее, какие виды данных подлежат хранению. Параметр данного правила может быть связан с типом данных, размером массива данных, ключевыми словами в наименовании файла и т. д.

4.4 Пакеты данных при сохранности больших данных

Пакеты данных содержат контент и связанные с ними метаданные. При сохранности больших данных выделяют три типа пакетов данных:

- пакет отбора данных (DSP), который содержит данные, подлежащие хранению, и связанные с ними метаданные (как правило, дескриптивные метаданные и метаданные о правах доступа), предоставляемые системой управления выбором данных (см. 4.5);
- пакет сохраненных данных (DPP), который содержит сохраненные данные и метаданные, генерируемые и поддерживаемые управлением хранением сохраненных данных (см. 4.5);
- пакет доступа к данным (DAP), который содержит запрошенные сохраненные данные и связанные с ними метаданные (как правило, включающие дескриптивные метаданные), предоставляемые системой управления доступом к данным (см. 4.5).

Отношения между DPP и DAP могут быть «один к одному», «один ко многим», «многие к одному», «многие ко многим». Например, в соответствии с различными условиями запроса данные DAP могут поступать из одного или нескольких DPP. Отношения между DPP и DSP такие же многозначные, как и отношения между DPP и DAP.

4.5 Функциональные объекты для управления сохранностью больших данных

Функциональные объекты для управления сохранностью больших данных включают управление: политикой сохранности данных, отбором данных, хранением сохраненных данных и доступом к данным (см. рисунок 4.1).

Управление политикой сохранности данных.

Управляет политикой сохранности данных для обеспечения доступности и пригодности к использованию сохраненных данных.

Управление отбором данных.

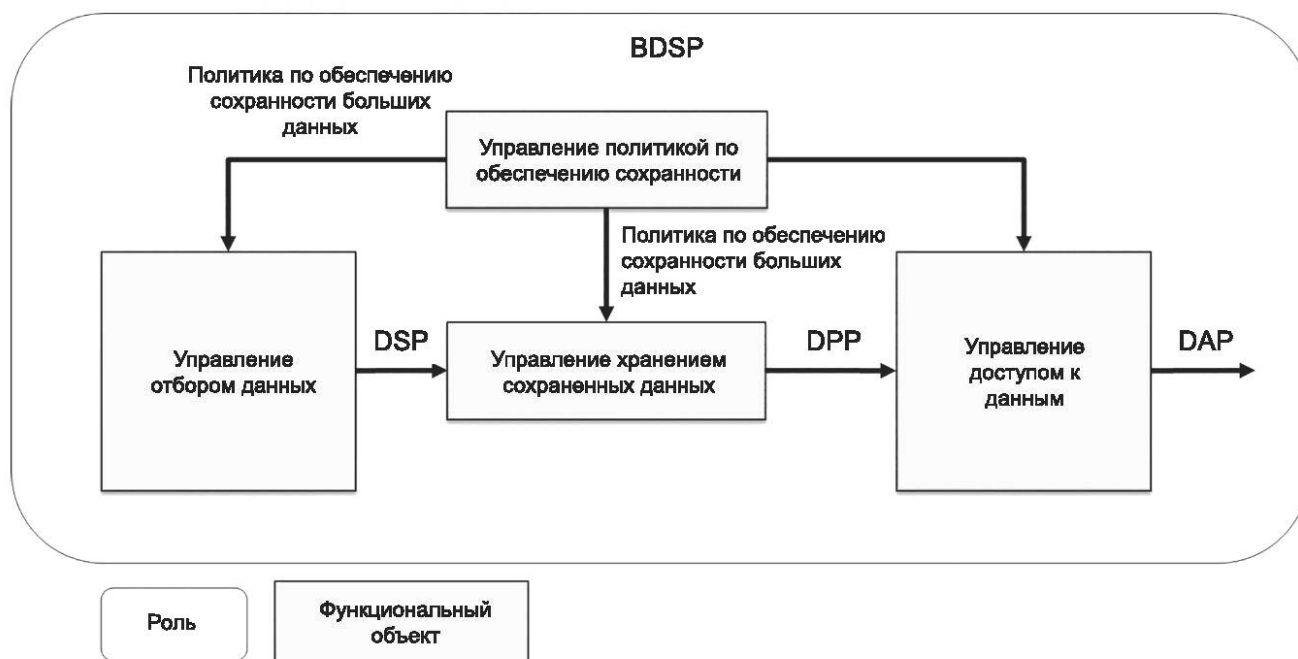
Обеспечивает отбор данных, подлежащих сохранению, и подготовку DSP к хранению в соответствии с установленными правилами.

Управление хранением сохраненных данных.

Обеспечивает хранение данных, подлежащих сохранению, и поддержку сохраненных данных, включая обновление и миграцию в соответствии с установленными правилами.

Управление доступом к данным.

Обеспечивает обработку запросов на доступ к сохраненным данным и предоставление запрошенных данных отправителю запроса с помощью DAP.



Примечание — На рисунке 4.1 представлена одна из ролей BDSP.

Рисунок 4.1 — Функциональные объекты для управления сохранностью больших данных

5 Функциональные требования к сохранности больших данных

Данный раздел содержит функциональные требования, применимые к сохранности больших данных.

5.1 Отбор данных, подлежащих сохранению

Требования к отбору данных, подлежащих сохранению, включают следующее:

- рекомендуется, чтобы сервис-провайдер больших данных поддерживал отбор данных, подлежащих сохранению, из исходных данных в соответствии с установленными правилами отбора данных для хранения.

Примечания

1 Исходные данные включают в себя как данные, собранные сервис-провайдером больших данных вовне, так и результаты обработки данных, сгенерированные внутри сервис-провайдера больших данных.

2 Правила по отбору данных для сохранения определяют, какие виды данных сервис-провайдеру больших данных следует сохранять. Параметр данного правила может быть связан с типом данных, размером массива данных, ключевыми словами в наименовании файла и т. д.;

- требуется, чтобы сервис-провайдер больших данных извлекал метаданные из данных, подлежащих сохранению.

Примечание — Метаданные, извлеченные из данных, подлежащих сохранению, как правило, включают дескриптивные метаданные, метаданные о целостности и метаданные о правах доступа;

- требуется, чтобы сервис-провайдер больших данных поддерживал генерацию DSP, который состоит из контента и метаданных.

5.2 Хранение данных, подлежащих сохранению

Требования к хранению данных, подлежащих сохранению, включают следующее:

- требуется, чтобы сервис-провайдер больших данных поддерживал преобразование DSP в DPP;
- рекомендуется, чтобы сервис-провайдер больших данных поддерживал выбор надлежащего уровня хранения для DPP в соответствии с установленными правилами хранения сохраняемых данных.

Примечание — Правила хранения сохраняемых данных определяют, на каком уровне должны храниться эти данные и при каких условиях их следует перенести на другие уровни хранения;

- рекомендуется, чтобы сервис-провайдер больших данных проводил мониторинг статистики доступа к DPP для управления иерархией хранилища;

- требуется, чтобы сервис-провайдер больших данных поддерживал удаление сохраненных данных в соответствии с установленными правилами определения периода хранения данных.

Примечание — Правила определения периода сохранения уточняют, как долго сервис-провайдер больших данных должен хранить данные, подлежащие сохранению.

5.3 Доступ к сохраненным данным

Требования обеспечения доступа к данным, подлежащим сохранению:

- требуется, чтобы сервис-провайдер больших данных поддерживал проверку полномочий лица, запрашивающего сохраненные данные, при обеспечении доступа к сохраненным данным;

- требуется, чтобы сервис-провайдер больших данных поддерживал поиск по базе сохраненных данных на условиях фильтрации, предоставляемой лицом, запрашивающим сохраненные данные;

- требуется, чтобы сервис-провайдер больших данных поддерживал преобразование DPP в DAP.

5.4 Управление политикой сохранности данных

Требования к управлению политикой сохранности данных:

- рекомендуется, чтобы сервис-провайдер больших данных поддерживал установление правил по отбору данных для сохранения, чтобы указывать, какие виды данных он должен сохранять;

- рекомендуется, чтобы сервис-провайдер больших данных поддерживал установление правил определения периода сохранения данных, чтобы указывать, как долго он должен сохранять данные;

- рекомендуется, чтобы сервис-провайдер больших данных поддерживал установление правил хранения сохраняемых данных, чтобы определять, на каком уровне следует хранить эти данные и при каких условиях их следует перенести на другие уровни хранения.

6 Обеспечение безопасности

Необходимо учитывать соответствующие требования по обеспечению безопасности (см. [3], [4] и соответствующие рекомендации (см. [1], [2]), в том числе требования по контролю доступа, аутентификации, конфиденциальности данных, политике хранения данных, безопасности сети, целостности данных, доступности и защите персональных данных.

Приложение А
(справочное)

Варианты использования методов сохранности больших данных

А.1 Шаблон варианта использования

Разработанные в приложении А варианты использования должны быть основаны на унифицированном формате для читабельности и четкого формирования материала.

Т а б л и ц а А.1 — Шаблон варианта использования

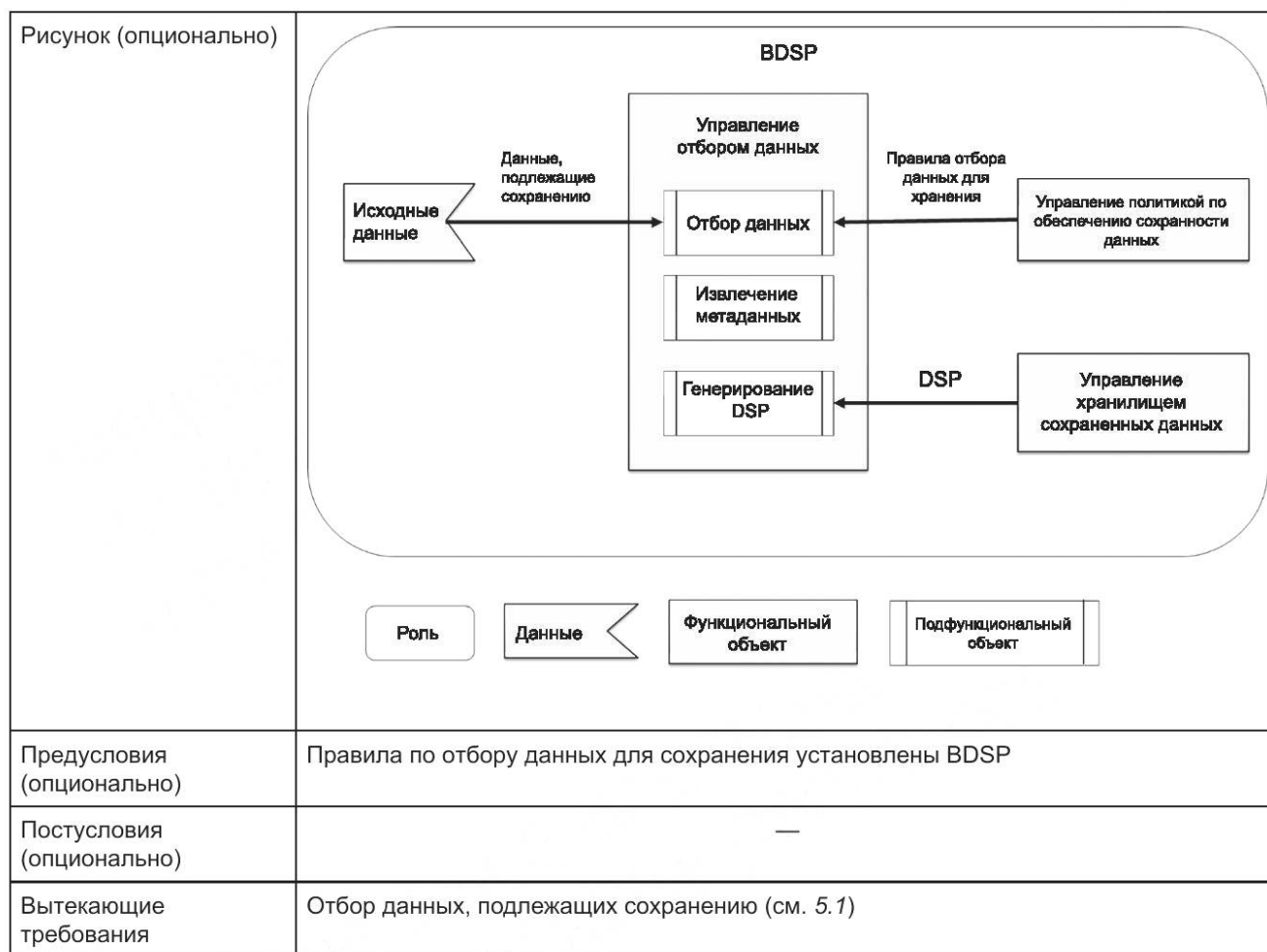
Название	П р и м е ч а н и е — Название варианта использования
Описание	П р и м е ч а н и е — Описание варианта использования
Роли	П р и м е ч а н и е — Роли, задействованные в варианте использования
Рисунок (опционально)	П р и м е ч а н и е — Рисунок для пояснения варианта использования, создаваемый при необходимости
Предусловия (опционально)	П р и м е ч а н и е — Необходимые предусловия, которые должны быть соблюдены до начала варианта использования
Постусловия (опционально)	П р и м е ч а н и е — Постусловие, которое будет выполнено после завершения текущего варианта использования
Вытекающие требования	П р и м е ч а н и е — Требования основаны на вариантах использования, подробное описание которых представлено в соответствующей части

А.2 Вариант использования: отбор данных, подлежащих сохранению

Т а б л и ц а А.2 — Отбор данных, подлежащих сохранению

Название	Отбор данных, подлежащих сохранению
Описание	Для подготовки данных, подлежащих сохранению, BDSP выбирает данные, в соответствии со следующими процедурами: - BDSP отбирает данные для сохранения (т. е. контент) из исходных данных в соответствии с правилами по отбору данных для сохранения, определяющими, какие именно виды данных подлежат сохранению. П р и м е ч а н и я 1 Исходные данные включают в себя как все данные, собранные BDSP извне, так и результаты обработки данных, сгенерированные BDSP внутри. 2 Правила отбора данных для сохранения определяют, какие именно виды данных следует сохранять. Параметры данных правил могут быть связаны с типом данных, размером массива данных, ключевыми словами в наименовании файла и т. д.; - BDSP извлекает из данных, подлежащих сохранению, некоторые метаданные (например, дескриптивные метаданные, метаданные о целостности и метаданные о правах доступа); - BDSP генерирует DSP, состоящий из контента и метаданных
Роли	BDSP

Окончание таблицы А.2



А.3 Вариант использования: создание уровней хранения для сохраненных данных

Таблица А.3— Создание уровней хранения для сохраненных данных

Название	Создание уровней хранения для сохраненных данных
Описание	Для хранения данных, подлежащих сохранению, BDSP использует иерархическую систему хранения, имеющую три уровня: онлайнное хранение, когда данные доступны немедленно; почти онлайнное хранение, когда данные доступны не сразу, но достаточно быстро; автономное хранение, когда данные полностью отключены от сети и требуют наибольшего времени для восстановления доступа к ним. Выбор уровня хранения данных, подлежащих сохранению, зависит от функции управления уровнями хранения и основывается на правилах хранения сохраняемых данных. Например, существует правило хранения сохраняемых данных, определяющее, что объем контента менее 10 Гбайт следует хранить на уровне онлайнного хранения, в противном случае его следует хранить на уровне почти онлайнного хранения. Если в течение 3 мес доступ к этим данным запрошен менее трех раз, то их следует переместить на более низкий уровень хранения, например с уровня онлайнного хранения на уровень почти онлайнного хранения. DSP с объемом контента 6 Гбайт будет храниться в соответствии со следующими процедурами: - BDSP преобразует DSP в DPP; - в соответствии с описанными выше правилами хранения сохраняемых данных DPP следует хранить в онлайнном хранении;

Окончание таблицы А.3

Описание	- BDSP будет продолжать мониторинг статистики доступа к DPP, чтобы через 3 мес определить необходимость его последующего перемещения на более низкий уровень хранения; - BDSP в соответствии с правилами хранения сохраняемых данных будет отслеживать, когда следует удалить сохраненные данные
Роли	BDSP
Рисунок (опционально)	
Предусловия (опционально)	Правила хранения сохраняемых данных установлены BDSP. Для управления хранением сохраненных данных из системы управления приемом данных предоставлен DSP
Постусловия (опционально)	В соответствии с правилами сохранности данных DPP хранится на соответствующем уровне хранения
Вытекающие требования	Хранение данных, подлежащих сохранению (см. 5.2)

А.4 Вариант использования: доступ к сохраненным данным

Таблица А.4 — Доступ к сохраненным данным

Название	Доступ к сохраненным данным
Описание	Лицо, запрашивающее сохраненные данные, намеревается получить от BDSP доступ к некоторым сохраненным данным. Доступ к сохраненным данным будет осуществлен в соответствии со следующими процедурами: - лицо, запрашивающее сохраненные данные, направляет в систему управления доступом к данным запрос (в соответствии с фильтрационными критериями), чтобы указать, к каким данным оно намеревается получить доступ;

Окончание таблицы А.4

Описание	- система управления доступом к данным проверяет, имеет ли данное лицо полномочия на доступ к запрашиваемым данным, и затем отправляет запрос на получение сохраненных данных в систему управления хранением сохраненных данных; - система управления хранением сохраненных данных осуществляет поиск запрошенных сохраненных данных в соответствии с фильтрационными критериями и отправляет полученный(ые) DPP в систему управления доступом к данным; - система управления доступом к данным преобразует DPP в DAP и отправляет обратно лицу, запрашивающему сохраненные данные
Роли	BDSP
Рисунок (опционально)	The diagram illustrates the process of managing data retention policy within the BDSP system. It shows the interaction between the 'Управление доступом к сохраненным данным' (Access Management) and 'Управление хранением сохраненных данных' (Data Retention Management) components. A user, 'Лицо, запрашивающее сохраненные данные' (Person requesting retained data), sends a 'Запрос сохраненных данных' (Request for retained data) to the Access Management component. This component performs a 'Проверка полномочий отправителя запроса' (Check sender's authority) and then sends a 'Запрос на сохраненные данные' (Request for retained data) to the Retention Management component. The Retention Management component performs a 'Поиск сохраненных данных' (Search for retained data) and returns a 'DPP' (Data Protection Policy) to the Access Management component. The Access Management component then converts the 'DPP' into a 'DAP' (Data Access Policy) and sends it back to the user. Legend: - Роль (Role) - Функциональный объект (Functional object) - Подфункциональный объект (Sub-functional object)
Предусловия (опционально)	Данные, к которым необходимо получить доступ, сохранены BDSP
Постусловия (опционально)	—
Вытекающие требования	Доступ к сохраненным данным (см. 5.3)

А.5 Вариант использования: управление политикой сохранности данных

Таблица А.5 — Вариант использования: управление политикой сохранности данных

Название	Управление политикой сохранности данных
Описание	BDSP реализует управление сохранностью данных на основе политики сохранности данных. BDSP может задавать политику сохранности данных через управление планированием сохранения. Политика сохранности данных состоит: - из правил отбора данных для сохранения, согласно которым определяют, какие виды данных BDSP следует сохранить. Параметры этих правил могут быть связаны с типом данных, размером массива данных, ключевыми словами в наименовании файла и т. д.; - правил определения периода сохранения, согласно которым определяют, в течение какого периода BDSP следует хранить данные;

Окончание таблицы А.5

Описание	- правил хранения сохраняемых данных, согласно которым устанавливают, на каком уровне следует хранить эти данные и при каких условиях их следует переместить на другие уровни
Роли	BDSP
Рисунок (опционально)	BDSP Управление политикой по обеспечению сохранности данных Политика по обеспечению сохранности данных Правила отбора данных для хранения Правила определения периода хранения Правила функционирования хранилищ для сохраняемых данных Роль Правило Функциональный объект
Предусловия (опционально)	—
Постусловия (опционально)	—
Вытекающие требования	Управление политикой сохранности данных (см. 5.4)

Приложение ДА
(справочное)

**Сопоставление структуры настоящего стандарта со структурой примененного
в нем международного стандарта**

Таблица ДА.1

Структура настоящего стандарта			Структура международного стандарта ITU-T Y.3604 (02/2020)		
Разделы	Пункты	Подпункты	Разделы	Пункты	Подпункты
Раздел 2 Термины и опре- деления	2.1	—	Раздел 2 Нормативные ссылки	—	—
	2.2			—	
	2.3			—	
	2.4			—	
	2.5			—	
	2.6			—	
	2.7			—	
	2.8			—	
	2.9			—	
	2.10			—	
Раздел 3 Соглашения по терминологии	—	—	Раздел 3 Термины и определения	3.1 Термины, опреде- ленные в других документах	3.1.1
	—	—		3.2 Термины, опреде- ленные в настоя- щем документе	3.2.1
		—			3.2.2
		—			3.2.3

Окончание таблицы ДА.1

Структура настоящего стандарта			Структура международного стандарта ITU-T Y.3604 (02/2020)		
Разделы	Пункты	Подпункты	Разделы	Пункты	Подпункты
Раздел 4 Обзор сохранности больших данных	4.1	—	Раздел 4 Сокращения	—	—
	4.2	4.2.1		—	—
		4.2.2		—	—
	4.3	—		—	—
	4.4	—		—	—
4.5	—	—	—	—	—
Раздел 5 Функциональные требования	5.1	—	Раздел 5 Соглашения по терминологии	—	—
	5.2	—		—	—
	5.3	—		—	—
	5.4	—		—	—
Раздел 6 Обеспечение безопасности	—	—	Раздел 6 Обзор сохранности больших данных	6.1	—
	—	—		6.2	6.2.1
		—			6.2.2
	—	—		6.3	—
	—	—		6.4	—
—	—	6.5	—		
—	—	—	Раздел 7 Функциональные требования	7.1	—
	—	—		7.2	—
	—	—		7.3	—
	—	—		7.4	—
—	—	—	Раздел 8 Обеспечение безопасности	—	—
Приложения	А		Приложение	I	
	ДА			—	
Примечания					
1 Раздел 2 «Нормативные ссылки» удален для взаимопонимания между различными пользователями настоящего стандарта.					
2 Настоящий стандарт дополнен терминологическими статьями 2.2—2.7.					
3 Раздел 4 «Сокращения» удален из настоящего стандарта в связи с отсутствием необходимости, а также для соответствия требованиям ГОСТ 1.5—2001 (см. 3.10.1).					
4 По тексту настоящего стандарта применены ссылки на национальные стандарты вместо международных стандартов, а также даны ссылки на стандарты из раздела «Библиография»					

Библиография

- [1] МСЭ-Т Y.3600 (2015) Большие данные. Требования и возможности, основанные на облачных вычислениях
- [2] МСЭ-Т Y.3603 (2019) Большие данные. Требования и концептуальная модель метаданных для каталога данных
- [3] МСЭ-Т Y.2201 Требования и возможности для сетей последующих поколений МСЭ-Т
- [4] МСЭ-Т Y.2701 Требования к безопасности для сетей последующих поколений версии 1

Ключевые слова: большие данные, категория данных, пакет данных, сохранность данных, функциональный объект, функциональное требование, стратегия

Редактор *Л.С. Зимилова*
Технический редактор *И.Е. Черепкова*
Корректор *Р.А. Ментова*
Компьютерная верстка *А.Н. Золотаревой*

Сдано в набор 14.12.2023. Подписано в печать 26.12.2023. Формат 60×84%. Гарнитура Ариал.
Усл. печ. л. 2,32. Уч.-изд. л. 1,90.

Подготовлено на основе электронной версии, предоставленной разработчиком стандарта

Создано в единичном исполнении в ФГБУ «Институт стандартизации» для комплектования Федерального информационного фонда стандартов,

117418 Москва, Нахимовский пр-т, д. 31, к. 2.
www.gostinfo.ru info@gostinfo.ru