

**ЦЕНТРАЛЬНЫЙ БАНК РОССИЙСКОЙ ФЕДЕРАЦИИ
(БАНК РОССИИ)**

**Унифицированные форматы
электронных банковских сообщений**

СПРАВОЧНИКИ

Версия 2023.4.0

Москва

2023

Содержание

1. ССЫЛКИ	3
2. ИСПОЛЬЗУЕМЫЕ СПРАВОЧНИКИ И КЛАССИФИКАТОРЫ	4
3. ТЕРМИНЫ, ОПРЕДЕЛЕНИЯ И СОКРАЩЕНИЯ	5
4. РЕКВИЗИТЫ ЭПС	9
5. РЕКВИЗИТЫ ЭСИС	10
6. РЕКВИЗИТЫ ПАКЕТА ЭПС (ЭСИС)	11
7. ТЕХНИЧЕСКИЕ СВЕДЕНИЯ ОБ УФЭБС	12
Оформление XML-документа	12
8. ОПИСАНИЕ ПРЕОБРАЗОВАНИЙ XML-ДОКУМЕНТА ДЛЯ ПРИВЕДЕНИЯ К НОРМАЛИЗОВАННОМУ ВИДУ	13
А. Удаление из XML-документа инструкций обработки	13
В. Удаление из элементов XML-документа атрибутов из пространства имен “http://www.w3.org/2001/XMLSchema-instance”	14
С. Упорядочение во всех элементах XML-документа префиксов пространств имен по заданному правилу	15
D. Удаление в каждом элементе XML-документа дочерних текстовых узлов, содержащих только пробельные символы	18
9. ОПИСАНИЕ ПРЕОБРАЗОВАНИЙ XML-ДОКУМЕНТА ДЛЯ ПРИВЕДЕНИЯ К КАНОНИЧЕСКОМУ ВИДУ	20
10. ОПИСАНИЕ АЛГОРИТМА КОДИРОВАНИЯ BASE64	24

1. Ссылки

В Альбомах УФЭБС использованы ссылки на следующие стандарты, нормативные документы и спецификации:

[Положение 762-П] Положение Банка России от 29.06.2021 № 762-П (ред. От 25.03.2022) “О правилах осуществления перевода денежных средств”. № 762-П;

[Положение 732-П] Положение о платежной системе Банка России: Положение Банка России от 24.09.2020 г. № 732-П;

ГОСТ ИСО 8601–2001 Система стандартов по информации, библиотечному и издательскому делу. Представление дат и времени. Общие требования;

Указание Банка России от 29 декабря 2016 г. N 4257-У " О перечне, формах, правилах и порядке составления и представления форм отчетности структурными подразделениями Банка России в Центральный банк Российской Федерации";

[XML] Extensible Markup Language (XML) 1.0 [Электронный ресурс]: W3C Recommendation 26 November 2008. – Fifth edition. – Режим доступа: <http://www.w3.org/TR/REC-xml>;

[XML-ns] Namespaces in XML [Электронный ресурс]: World Wide Web Consortium 14-January-1999. – Режим доступа: <http://www.w3.org/TR/REC-xml-names>;

[XML-schema] XML Schema Part 1: Structures [Электронный ресурс]: W3C Recommendation 02 May 2001. – First Edition. – Режим доступа: <http://www.w3.org/TR/2001/REC-xmlschema-1-20010502/>; XML Schema Part 2: Datatypes [Электронный ресурс]: W3C Recommendation 02 May 2001. – First Edition. – Режим доступа: <http://www.w3.org/TR/2001/REC-xmlschema-2-20010502/>;

[XML-c14n] Canonical XML Version 1.0 [Электронный ресурс]: W3C Recommendation 15 March 2001. – Режим доступа: <http://www.w3.org/TR/xml-c14n>;

[XPath] XML Path Language (XPath) Version 1.0 [Электронный ресурс]: W3C Recommendation 16 November 1999. – Режим доступа: <http://www.w3.org/TR/1999/REC-xpath-19991116>;

[base64] Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies [Электронный ресурс]: November 1996. – Режим доступа: <http://www.ietf.org/rfc/rfc2045.txt>;

[SOAP12] SOAP Version 1.2 Part 0: Primer [Электронный ресурс]: W3C Recommendation 24 June 2003. – Режим доступа: <http://www.w3.org/TR/soap12-part0/>; SOAP Version 1.2 Part 1: Messaging Framework [Электронный ресурс]: W3C Recommendation 24 June 2003. – Режим доступа: <http://www.w3.org/TR/soap12-part1/>; SOAP Version 1.2 Part 2: Adjuncts [Электронный ресурс]: W3C Recommendation 24 June 2003 – Режим доступа: <http://www.w3.org/TR/soap12-part2/>; SOAP Version 1.2 Specification Assertions and Test Collection [Электронный ресурс]: W3C Recommendation 24 June 2003. – Режим доступа: <http://www.w3.org/TR/2003/REC-soap12-testcollection-20030624/>;

[RFC2616] RFC 2616 Hypertext Transfer Protocol -- HTTP/1.1 [Электронный ресурс]: IETF RFC 2616 June 1999 – Режим доступа: <http://www.ietf.org/rfc/rfc2616.txt>;

[RFC2821] RFC 2821 Simple Mail Transfer Protocol [Электронный ресурс]: IETF RFC 2821 April 2001 – Режим доступа: <http://www.ietf.org/rfc/rfc2821.txt>;

[deflate] Request for Comments 1951. DEFLATE Compressed Data Format Specification version 1.3 [Электронный ресурс]: May 1996. – Режим доступа: <http://www.ietf.org/rfc/rfc1951>;

[zlib] Request for Comments 1950. ZLIB Compressed Data Format Specification version 3.3 [Электронный ресурс]: May 1996. – Режим доступа: <http://www.ietf.org/rfc/rfc1950>.

2. Используемые справочники и классификаторы

Перечисленные ниже обозначения справочников и классификаторов применяются в описании прикладных типов.

[БИК РФ] Справочник банковских идентификационных кодов участников расчетов на территории Российской Федерации (Справочник БИК РФ). Структура БИК, порядок его присвоения участникам платежной системы и клиентам Банка России, не являющимся участниками платежной системы, реквизиты Справочника банковских идентификационных кодов в платежной системе Банка России: Приложение 5 к Положению о платежной системе Банка России: Положение Банка России от 24.09.2020 г. № 732-П.

[ИНН] Справочник идентификационного номера налогоплательщика. Об утверждении порядка и условий присвоения, применения, а также изменения идентификационного номера налогоплательщика и форм документов, используемых при учете в налоговом органе юридических и физических лиц: Приказ МНС от 27.11.1998 г. № ГБ-3-12/309.

[КПП] Справочник кодов причины постановки на учет. Об утверждении порядка и условий присвоения, применения, а также изменения идентификационного номера налогоплательщика и форм документов, используемых при учете в налоговом органе юридических и физических лиц: Приказ МНС от 27.11.1998 г. № ГБ-3-12/309.

[УИС] Уникальный идентификатор составителя электронного сообщения.

[Шифр документа] Шифр документа. О правилах ведения бухгалтерского учета в Центральном банке Российской Федерации (Банке России): Положение Банка России от 01.01.2006 № 66-П, Приложение 1; О правилах ведения бухгалтерского учета в кредитных организациях, расположенных на территории Российской Федерации: Положение Банка России от 26.03.2007 №302-П.

[Лицевой счет] Номер счета. О правилах ведения бухгалтерского учета в Центральном банке Российской Федерации (Банке России): Положение Банка России от 01.01.2006 № 66-П, Приложение 1; О правилах ведения бухгалтерского учета в кредитных организациях, расположенных на территории Российской Федерации: Положение Банка России от 26.03.2007 №302-П.

[Очередность платежа] Очередность платежа. Гражданский Кодекс Российской Федерации, ст. 855, п.2;

[ГОСТ ИСО 8601–2001] ГОСТ ИСО 8601–2001 Система стандартов по информации, библиотечному и издательскому делу. Представление дат и времени. Общие требования.

[ОКВ] Общероссийский классификатор валют. Общероссийский классификатор валют ОК 014-2000. Введен в действие с 01.07.2001 Постановлением ГК РФ по стандартизации и метрологии от 25.12.2000 г. № 405-ст. [С изм. и доп.].

3. Термины, определения и сокращения

В Альбомах УФЭБС используются следующие термины и определения:

Т а б л и ц а 1 – Термины и определения

Термин	Описание
Автоматизированная система эмиссионно-кассовых работ; АС ЭКР	Автоматизированная система, предназначенная для комплексной автоматизации эмиссионно-кассовых работ в территориальном учреждении Банка России и в подчиненных головных расчетно-кассовых центрах и расчетно-кассовых центрах.
Автономный режим подсистемы СБП	Режим работы подсистемы СБП в ночное время, выходные и праздничные дни (вне периода функционирования ЦК ПС БР), с которого начинается операционный день подсистемы СБП.
АРМ	Автоматизированное рабочее место
Владелец КА	ПБР, кредитная организация (филиал) или другой клиент Банка России, КА которого зарегистрирован в порядке, установленном договором между Банком России и его клиентом
Главный участник пула ликвидности	Владелец банковского (корреспондентского) счета (субсчета), открытого в Банке России и являющегося главным счетом пула ликвидности
Глобально-уникальный идентификатор (Globally Unique Identifier – GUID)	Шестнадцатибайтовый (128-битный) идентификатор, гарантированно уникальный. Вероятность генерации двух одинаковых GUID настолько мала, что ею можно пренебречь
ДОФР Банка России; ДОФР	Департамент операций на финансовых рынках Банка России
ЕКС	Единый казначейский счет
ЗК	Защитный код
КО	Кредитная организация
КО/ОК	Кредитные организации и другие клиенты Банка России
Код аутентификации; КА	Данные, используемые для удостоверения права распоряжения денежными средствами (для ЭС, содержащих распоряжения в электронном виде), контроля целостности и подтверждения подлинности ЭС
Конверт ВК	Конверт для сообщений формата ISO CTCCR в рамках осуществления валютного контроля
Конверт СФСФК	Конверт для специализированных финансовых сообщений по казначейскому обеспечению обязательств при банковском сопровождении государственных контрактов формата TSLC
Конверт ТО	Конверт для сообщений юридических лиц формата ISO TS в рамках торговых операций
Конверт ФС	Конверт для передачи финансовых сообщений
Конверт ЭССФ	Конверт для электронных сообщений в собственных форматах участников электронного обмена
Конверт ЮЛ	Конверт для сообщений юридических и иных лиц формата ISO
Косвенный участник	Организации, соответствующие следующим критериям: - организация может являться участником платежной системы в соответствии с частями 1 и 3 статьи 21 Федерального закона от 27 июня 2011 года № 161-ФЗ;

	<ul style="list-style-type: none"> - организация является клиентом кредитной организации (ее филиала) – прямого участника; - организации предоставлен доступ к услугам по переводу денежных средств с использованием распоряжений в электронном виде в соответствии с условиями договора корреспондентского счета (субсчета), заключенного между Банком России и кредитной организацией (ее филиалом) – прямым участником; - организация не является клиентом Банка России.
КП	Код пользователя
НСИ	Нормативно-справочная информация
НСПК	Национальная система платежных карт
ОД подсистемы СБП	Операционный день подсистемы СБП – период времени функционирования подсистемы СБП, в течение которого непрерывно осуществляется прием и обработка ЭС. Начинается с момента перехода подсистемы СБП в автономный режим и заканчивается переходом в автономный режим следующего ОД.
Пакет электронных сообщений	Одно или более ЭС (ЭПС/ЭСИС), с установленным КА, при этом на каждое ЭС в составе пакета КА не устанавливается
ПО	Программное обеспечение
Подразделение Банка России; ПБР	Главное управление Банка России, отделение, отделение – национальный банк главного управления Банка России, структурное подразделение центрального аппарата Банка России, расчетно-кассовый центр, кассовый центр, полевое учреждение Банка России (Положение 732-П)
Подчиненный участник пула ликвидности	Владелец банковского (корреспондентского) счета (субсчета), открытого в Банке России и являющегося подчиненным счетом пула ликвидности
ППО	Программная процедура обработки
Прямой участник	<p>Организации, соответствующие следующим критериям:</p> <ul style="list-style-type: none"> - организация является кредитной организацией (ее филиалом), Федеральным казначейством или его территориальным органом, иной организацией, которая может являться прямым участником в соответствии с частью 7 статьи 21 Федерального закона от 27 июня 2011 года № 161-ФЗ; - организация имеет банковский (корреспондентский) счет (субсчет) в Банке России в соответствии с договором банковского (корреспондентского) счета (субсчета).
САБС	Специализированная автоматизированная банковская система
СБП	Сервис быстрых платежей платежной системы Банка России
СВК	Унифицированная транспортная среда электронного взаимодействия территориальных учреждений Банка России с клиентами Банка России
СПФС	Система передачи финансовых сообщений Банка России
ТОФК	Территориальный орган Федерального казначейства
Транспортное сообщение (почтовое сообщение, HTTP-сообщение, сообщение WebSphere MQ)	Совокупность элементов информации, оформленных в соответствии с требованиями транспортного протокола (SMTP, HTTP, протокол IBM WebSphere MQ и другие), использующегося при обмене
Транспортный протокол	Протокол, реализующий сетевой сервис, например, электронная

	почта, web-доступ, гарантированная доставка сообщений (SMTP, HTTP, WMQ) и использующий протокол сетевого уровня (например, TCP или UDP)
ТС ЦК	Текущий остаток денежных средств на счете участника (для пула ликвидности – сумма денежных средств на всех счетах пула ликвидности) для осуществления перевода денежных средств с использованием сервиса срочного перевода, сервиса несрочного перевода и перераспределения ликвидности между ЦК ПС и СБП, без учета лимита внутрисдневного кредита и кредита овернайт. На значение ТС ЦК не влияют суммы денежных средств, забронированных под расходные операции по предъявленному денежному чеку и суммы денежных средств, зарезервированных для исполнения условных платежных распоряжений, требующих резервирования ликвидности.
ТУ	Территориальное учреждение Банка России
УФЭБС	Унифицированные форматы электронных банковских сообщений
Участники обмена ЭС с Банком России; участники	Кредитные организации, филиалы кредитных организаций и другие клиенты Банка России, заключившие с Банком России договор об обмене электронными сообщениями при осуществлении расчетов через расчетную сеть Банка России (далее Договор обмена)
ФК	Федеральное казначейство
ФОР	Фонд обязательных резервов
Форматы электронных сообщений	Упорядоченная последовательность символов, включенная в ЭС по единым правилам, представленная в формализованном виде, в установленной последовательности и размерности, используемая для передачи по каналам связи и для ее обработки
Централизованная компонента платежной системы Банка России; ПС БР	Автоматизированная система, предназначенная для автоматизации расчетов и создания общей информационной среды для функционального взаимодействия всех банковских субъектов (ПБР, коммерческих банков и их филиалов, некредитных организаций и других клиентов Банка России)
ЦОИ АС ЭКР	Центр обработки информации АС ЭКР
ЦОС	Центр обмена сообщениями
Электронное сообщение; ЭС	Совокупность данных, соответствующая установленному Банком России электронному формату, пригодная для однозначного восприятия его содержания, снабженная кодом аутентификации
Электронное платежное сообщение; ЭПС	Электронное сообщение, являющееся основанием для совершения операций по счетам кредитных организаций (филиалов) и других клиентов Банка России, открытым в подразделениях Банка России, с установленным КА и имеющий равную юридическую силу с расчетными документами на бумажных носителях, подписанными собственноручными подписями уполномоченных лиц и заверенными оттиском печати
Электронное служебно-информационное сообщение; ЭСИС	Электронное сообщение с установленным КА, обеспечивающее обмен информацией при совершении расчетов и проведении операций по счетам, открытым в подразделениях Банка России (запросы, отчеты, выписки из счетов, документы, связанные с предоставлением кредитов Банка России, и т.п.)
Электронные сообщения; ЭС	Электронные платежные сообщения и электронные служебно-информационные сообщения
CLS Bank International; CLS Bank; CLS Банк	Расчетный банк системы CLS

HTTP (Hypertext Transfer Protocol)	Гипертекстовый протокол передачи [RFC2616]
RTGS (Real Time Gross Settlement)	Валовые расчеты в режиме реального времени
Schema / Схема	Логическое и физическое определение элементов данных, физических характеристик и внутренних отношений
SMTP (Simple Mail Transfer Protocol)	Простой протокол передачи почтовых сообщений [RFC2821]
S.W.I.F.T.(Society of Worldwide Interbank Financial Telecommunications); SWIFT	Сообщество всемирных межбанковских финансовых телекоммуникаций
W3C (World Wide Web Consortium)	Консорциум всемирной сети, Интернет консорциум
WMQ	Система передачи сообщений с гарантированной доставкой WebSphere MQ (ранее MQSeries). Данная аббревиатура также используется для обозначения транспортного протокола гарантированной доставки сообщений, реализованного в системе WebSphere MQ
XML (extensible markup language)	Расширяемый (открытый) язык разметки
XML schema / XML-схема	Язык описания структуры документа. Предусматривает описание допустимой структуры документа и, возможно, типов данных в значениях атрибутов и содержанием элементов. В настоящее время существует несколько таких языков
XML-документ	Совокупность данных, соответствующая требованиям [XML]
XML namespaces/ пространства имен XML	Способ уточнения (квалификации) имен элементов и атрибутов путем сопоставления набора имен некоторой области применения, технически выраженной в виде идентификатора ресурса (URI)

4. Реквизиты ЭПС

Состав и порядок заполнения полей (реквизитов) ЭПС должны соответствовать составу и порядку заполнения полей (реквизитов) расчетных документов на бумажных носителях.

Исключение составляют следующие реквизиты:

- наименование и место нахождения банка плательщика и банка получателя (поля 10 и 13) в ЭПС отсутствуют (в копиях ЭПС на бумажном носителе указанные реквизиты заполняются в соответствии со Справочником БИК РФ);

- реквизиты «Срок плат.», «Наз.пл.» (поля 19, 20 соответственно) до указаний Банка России в ЭПС не используются;

- сумма прописью (поле 6) в ЭПС отсутствует (в копиях ЭПС на бумажном носителе реквизит формируется программным способом на основании цифрового значения суммы);

- наименование документа (поле 1), номер формы по ОКУД (поле 2), а также графические реквизиты (штампы, подписи и оттиски печати, проставляемые в полях 43, 44, 45, 46, 47, 69) в ЭПС отсутствуют;

- в поле «дата представления документов в банк получателя» (поле 48) проставляется значение даты из поля расчетного документа «отметки банка получателя» (поле 48).

Помимо реквизитов, содержащихся в расчетных документах на бумажных носителях, в ЭПС содержатся реквизиты, заполняемые в следующем порядке.

Платежный ордер и платежное поручение в виде ЭПС в случаях, когда они составляются на основании ЭПС (частичная оплата, зачисление средств по назначению после визуального контроля ЭПС, возврат средств по ЭПС и т.п.), содержат ссылочную группу реквизитов, позволяющую однозначно идентифицировать исходное ЭПС, на основании которого составлен данное ЭПС, и включающую в себя «Порядковый номер исходного ЭПС», «Дату составления исходного ЭПС», «Уникальный идентификатор составителя исходного ЭПС» (поля 203, 204, 205 соответственно).

Платежное требование в виде ЭПС при частичном акцепте помимо акцептованной суммы, подлежащей оплате, содержит сумму исходного расчетного документа, предъявленного к акцепту (поле 150).

Т а б л и ц а 2 – Группы реквизитов ЭПС

Группа	Реквизиты	Примечание
Призначная группа	Порядковый номер ЭС (уникальный для каждого участника или ПБР в течение дня); Дата составления ЭС; Уникальный идентификатор составителя ЭС.	Содержится всегда. Реквизит «Уникальный идентификатор составителя ЭС» для ЭС, исходящих из ПС БР, заполняется УИС ПС БР.
Ссылочная группа	Порядковый номер исходного ЭС; Дата составления исходного ЭС; Уникальный идентификатор составителя исходного ЭС.	Платежный ордер и платежное поручение в виде ЭПС в случаях, когда они составляются на основании ЭПС (частичная оплата, возврат средств по ЭПС и т.п.), содержат ссылочную группу реквизитов, позволяющую однозначно идентифицировать исходное ЭПС, на основании которого составлен данное ЭПС.

5. Реквизиты ЭСИС

Наряду с ЭПС участник при осуществлении безналичных расчетов через расчетную сеть Банка России осуществляет обмен **ЭСИС** с обслуживающим его ПБР в соответствии с регламентом и форматами, предусмотренными данным документом и Договором обмена.

Т а б л и ц а 3 – Группы реквизитов ЭСИС

Группа	Реквизиты	Примечание
Призначная группа	Порядковый номер ЭС; Дата составления ЭС; Уникальный идентификатор составителя ЭС.	Содержится всегда. Реквизит «Уникальный идентификатор составителя ЭС» для ЭС, исходящих из ПС БР, заполняется УИС ПС БР.
Дополнительная информация	Уникальный идентификатор получателя ЭС. Время составления ЭС.	Прикладная информация. Не используется в транспортной системе Банка России.
Ссылочная группа	Порядковый номер исходного ЭС; Дата составления исходного ЭС; Уникальный идентификатор составителя исходного ЭС.	Заполняются в соответствии со значениями реквизитов призначной группы ЭПС или ЭСИС, в ответ на который или с запросом о котором направляется данное ЭСИС
Содержательная часть	Код результата; Информация о проведенных операциях; Текст пояснения и т.д.	Устанавливается актами технического характера Банка России
Дополнительные реквизиты		Могут не заполняться

6. Реквизиты пакета ЭПС (ЭСИС)

Под **пакетом ЭПС (ЭСИС)** понимается одно или более ЭПС (ЭСИС), подписанных КА (порядок устанавливается нормативными актами Банка России), при этом каждое ЭПС (ЭСИС) в составе пакета не подписывается КА. Пакет ЭС имеет общую группу реквизитов, относящуюся ко всей совокупности ЭПС, входящих в состав пакета.

Банк России вправе устанавливать в Договоре обмена ограничение общего количества ЭПС, входящих в состав пакета ЭПС, а также его состав.

Пакет ЭПС имеет общую группу реквизитов, относящуюся ко всей совокупности ЭПС, входящих в состав пакета, и включающую «Общее количество ЭПС, входящих в состав пакета», «Общая сумма ЭПС, входящих в состав пакета»

7. Технические сведения об УФЭБС

УФЭБС разработаны в соответствии со следующими спецификациями:

- "Extensible Markup Language (XML) 1.0" – «Расширяемый язык разметки XML 1.0» [XML];
- "Namespaces in XML" – «Пространства имен в XML» [XML-ns];
- "XML Schema Part 1: Structures" и "XML Schema Part 2: Datatypes" – «XML схема. Часть 1 – структуры» и «XML схема. Часть 2 – типы данных» [XML-schema].

Оформление XML-документа

Электронные сообщения оформляются согласно [XML]. XML-документы передаются в кодировке WINDOWS-1251 или UTF-8.

XML-документы должны начинаться с декларации XML, которая определяет версию XML и применяемую кодировку согласно [XML].

```
<?xml version="1.0" encoding="WINDOWS-1251"?>
```

```
<?xml version="1.0" encoding="UTF-8"?>
```

Примечание – Декларация XML с указанием версии XML в XML-документе обязательна. В случае, если не указана применяемая кодировка, считается, что XML-документ составлен в кодировке UTF-8.

Для XML-документов объявляются описанные пространства имен и их префиксы (возможно применение пространства имен по умолчанию). Для элементов, содержащих значение КА или ЗК, возможно определение необходимых пространств имен непосредственно в самом элементе.

В XML-документе должна быть задана принадлежность всех элементов к определенному пространству имен (посредством задания пространства имен по умолчанию или явно с использованием префиксов). Атрибуты в XML-документе оформляются как, так называемые, неклассифицированные локальные компоненты, т.е. атрибуты определяются в составе элемента и не относятся ни к какому именованному пространству имен (префикс, определяющий именованное пространство имен, к атрибуту не применяется). В схеме принадлежность элементов и атрибутов к конкретному пространству имен задана глобально следующими значениями:

```
elementFormDefault="qualified" attributeFormDefault="unqualified".
```

Примечание – префикс пространства имен не несет смысловой нагрузки и используется только для привязки имен элементов и атрибутов к названию пространства имен.

Запрещается использовать в XML-документах следующие конструкции, которые явно не описаны в данном документе:

- комментарии,
- инструкции обработки,
- декларации типа документа (DTD),
- определения пространств имен, отличающихся от описанных,
- атрибуты из пространства имен "http://www.w3.org/2001/XMLSchema-instance"¹,
- атрибуты из пространства имен "http://www.w3.org/XML/1998/namespace",
- секции CDATA.

¹ Не допускается наличие атрибутов schemaLocation, noNamespaceSchemaLocation из пространства имен "http://www.w3.org/2001/XMLSchema-instance" с указанием местонахождения схемы (в качестве подсказки программе обработки).

8. Описание преобразований XML-документа для приведения к нормализованному виду

В данном приложении описываются преобразования XML-документа для приведения к нормализованному виду (нормализация). Данные преобразования применяются к XML-документу, содержащему защищаемые данные ЭС (пакета ЭС), и выполняются до канонизации XML-документа, содержащего защищаемые данные ЭС (пакета ЭС).

Идентификатор алгоритма, описывающего преобразования:

“urn:cbr-ru:dsig:v1.1#normalization”.

При описании алгоритма нормализации XML-документа для адресации частей XML-документа, содержащего защищаемые данные ЭС (пакета ЭС), используется язык [XPath].

XPath представляет XML-документ в виде дерева узлов. Узлы (node) бывают различных типов: узлы элементов, узлы атрибутов, узлы пространств имен, узлы инструкций обработки, узлы комментариев, текстовые узлы. При электронном обмене в системе безналичных расчетов запрещается использовать в XML-документах следующие узлы:

- узлы комментариев,
- узлы инструкции обработки,
- узлы пространств имен, отличающихся от описанных,
- узлы атрибутов из пространства имен <http://www.w3.org/2001/XMLSchema-instance>,
- текстовые узлы, если в родительском узле имеются дочерние узлы элементов.

Таким образом, при обработке, хранении информации из XML-документа применяется лишь часть узлов, а все прочие просто игнорируются. Информация, которая не используется при электронном обмене, не должна защищаться ЗК, поэтому перед формированием ЗК необходимо применить преобразования, которые позволяют удалить лишнюю информацию из XML-документа (неиспользуемые узлы). Это дает возможность защиты информации без учета особенностей разметки.

Ниже представлен список необходимых преобразований XML-документа для приведения к нормализованному виду:

- удаление из XML-документа инструкций обработки;
- удаление из элементов XML-документа атрибутов из пространства имен <http://www.w3.org/2001/XMLSchema-instance>;
- упорядочение во всех элементах XML-документа префиксов пространств имен по заданному правилу;
- удаление в каждом элементе XML-документа дочерних текстовых узлов, содержащих только пробельные символы.

Порядок выполнения преобразований имеет значение и должен быть именно такой, как указан.

А. Удаление из XML-документа инструкций обработки

Использование инструкций обработки запрещено в УФЭБС Банка России, поэтому они должны игнорироваться при приеме ЭС и не защищаются ЗК.

Инструкция обработки (processing instruction) имеет следующий синтаксис:

```
<?адресат[_содержимое]?>
```

где

- адресат – адрес, используемый для идентификации приложения, которому предназначается эта инструкция;
- содержимое – содержимое инструкции обработки, исключая адресата и любые пробельные символы, непосредственно следующие за ним.
- необязательная синтаксическая единица представляется в квадратных скобках;
- для обозначения символа пробела используется символ подчеркивания (_).

Для каждой инструкции обработки, за исключением инструкций обработки, помещенных в декларацию типа документа, создается соответствующий узел.

Алгоритм для удаления из XML-документа инструкций обработки:

Для каждого узла дерева, представляющего XML-документ, содержащий защищаемые данные ЭС (пакета ЭС), выполняется следующее преобразование:

- если данный узел является узлом инструкции обработки, он удаляется.

Пример

XML-документ до преобразования:

```
<?xml version="1.0" encoding="Windows-1251"?>
<?xml-stylesheet type="text/xsl" href="UniESID1.xslt"?>
<ED202 xmlns="urn:cbr-ru:ed:v2.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="urn:cbr-ru:ed:v2.0 UniDoc1_1.xsd"
  EDNo="8" EDDate="2003-04-14" EDAuthor="4525545000"
  EDReceiver="4525000000" InquiryCode="1">
  <EDRefID EDNo="7" EDDate="2003-04-14" EDAuthor="4525545000"/>
</ED202>
```

XML-документ после преобразования:

```
<?xml version="1.0" encoding="Windows-1251"?>
<ED202 xmlns="urn:cbr-ru:ed:v2.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="urn:cbr-ru:ed:v2.0 UniDoc1_1.xsd"
  EDNo="8" EDDate="2003-04-14" EDAuthor="4525545000"
  EDReceiver="4525000000" InquiryCode="1">
  <EDRefID EDNo="7" EDDate="2003-04-14" EDAuthor="4525545000"/>
</ED202>
```

Из XML-документа удалена инструкция обработки `<?xml-stylesheet type="text/xsl" href="UniESID1.xslt"?>`. Декларация XML `<?xml version="1.0" encoding="Windows-1251"?>` не является инструкцией обработки, поэтому оставлена в XML-документе.

В. Удаление из элементов XML-документа атрибутов из пространства имен “http://www.w3.org/2001/XMLSchema-instance”

Использование атрибутов из пространства имен “http://www.w3.org/2001/XMLSchema-instance” запрещено в УФЭБС Банка России. Атрибуты из данного пространства имен должны игнорироваться при обработке ЭС, и не защищаются ЗК.

Алгоритм для удаления из элементов XML-документа атрибутов из пространства имен “http://www.w3.org/2001/XMLSchema-instance”:

Для каждого узла дерева, представляющего XML-документ, содержащий защищаемые данные ЭС (пакета ЭС), выполняется следующее преобразование:

а) если данный узел является узлом элемента, для каждого его дочернего узла выполняется следующее преобразование:

1) если данный дочерний узел является узлом атрибута, причем данный атрибут принадлежит пространству имен http://www.w3.org/2001/XMLSchema-instance и локальное имя атрибута совпадает с одним из следующих идентификаторов: schemaLocation, noNamespaceSchemaLocation, type или nil, он удаляется.

Пример

XML-документ до преобразования:

```
<?xml version="1.0" encoding="Windows-1251"?>
<ED202 xmlns="urn:cbr-ru:ed:v2.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="urn:cbr-ru:ed:v2.0 UniDoc1_1.xsd"
  EDNo="8" EDDate="2003-04-14" EDAuthor="4525545000"
  EDReceiver="4525000000" InquiryCode="1">
  <EDRefID EDNo="7" EDDate="2003-04-14" EDAuthor="4525545000"/>
</ED202>
```

XML-документ после преобразования:

```
<?xml version="1.0" encoding="Windows-1251"?>
<ED202 xmlns="urn:cbr-ru:ed:v2.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  EDNo="8" EDDate="2003-04-14" EDAuthor="4525545000"
  EDReceiver="4525000000" InquiryCode="1">
  <EDRefID EDNo="7" EDDate="2003-04-14" EDAuthor="4525545000"/>
</ED202>
```

Из элемента ED202 удален атрибут xsi:schemaLocation из пространства имен "http://www.w3.org/2001/XMLSchema-instance".

С. Упорядочение во всех элементах XML-документа префиксов пространств имен по заданному правилу

В УФЭБС Банка России значение префикса пространства имен не несет информативной нагрузки и может быть любым. Для исключения необходимости сохранять сведения об особенностях разметки задаются правила для упорядочения и приведения к определенному виду префиксов пространств имен во всех элементах XML-документа, содержащего защищаемые данные ЭС (пакета ЭС).

Преобразование приводит XML-документ к виду, удовлетворяющему следующим требованиям:

- Используемые префиксы пространств имен задаются в форме n1, n2 и т.п. Положительное целое число, следующее за n, именуется индексом префикса.
- Пространства имен никогда не наследуются.
- Пространства имен по умолчанию никогда не применяются.

В элементе должны быть объявлены только те пространства имен, которым принадлежат элемент и его атрибуты. Объявления прочих пространств имен необходимо **удалить**, чтобы предотвратить конфликт имен.

Примечание – Данный подход выбран для обеспечения контекстно-независимого представления: форма элемента не зависит от места появления в XML-документе.

Алгоритм для упорядочения во всех элементах XML-документа префиксов пространств имен по заданному правилу:

Для каждого узла элемента дерева, представляющего XML-документ, содержащий защищаемые данные ЭС (пакета ЭС), выполняется следующее преобразование:

- a) составляется список названий пространств имен (URI):
 - 1) элементами списка являются названия пространств имен, которым принадлежат элемент и его атрибуты;
 - 2) список названий пространств имен сортируется в лексикографическом порядке, все повторяющиеся названия пространств имен удаляются;
 - 3) согласно отсортированному порядку каждому названию пространства имен из списка присваивается префикс n1, n2, ..., nXX (индексы префиксов всегда задаются последовательными целыми числами, начиная с 1);
- b) часть имени элемента, представляющая префикс пространства имен, приводится в соответствие со списком названий пространств имен;

с) для каждого дочернего узла данного узла элемента выполняется следующее преобразование:

- 1) если данный дочерний узел является узлом пространства имен, он удаляется;
- 2) если данный дочерний узел является узлом атрибута, часть имени атрибута, представляющая префикс пространства имен, приводится в соответствие со списком названий пространств имен.

д) к узлу элемента добавляются дочерние узлы пространств имен согласно списку названий пространств имен.

Примечание – В каждом элементе объявляются все пространства имен, которым принадлежит данный элемент и его атрибуты. Данный подход выбран для упрощения алгоритма.

Примеры

1 Упорядочение префиксов пространств имен в XML-документе

XML-документ до преобразования:

```
<?xml version="1.0" encoding="Windows-1251"?>
<ED202 xmlns="urn:cbr-ru:ed:v2.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  EDNo="8" EDDate="2003-04-14" EDAuthor="4525545000"
  EDReceiver="4525000000" InquiryCode="1">
  <EDRefID EDNo="7" EDDate="2003-04-14" EDAuthor="4525545000"/>
</ED202>
```

XML-документ после преобразования:

```
<?xml version="1.0" encoding="Windows-1251"?>
<n1:ED202 xmlns:n1="urn:cbr-ru:ed:v2.0"
  EDNo="8" EDDate="2003-04-14" EDAuthor="4525545000"
  EDReceiver="4525000000" InquiryCode="1">
  <n1:EDRefID xmlns:n1="urn:cbr-ru:ed:v2.0" EDNo="7" EDDate="2003-04-14"
  EDAuthor="4525545000"/>
</n1:ED202>
```

Элемент ED202 принадлежит пространству имен "urn:cbr-ru:ed:v2.0", объявленному как пространство имен по умолчанию, его атрибуты принадлежат не специфицированному пространству имен. Следовательно, список названий пространств имен состоит из одного элемента "urn:cbr-ru:ed:v2.0", которому присваивается префикс n1. Данное пространство имен объявляется в элементе, ему назначается префикс n1. Префикс n1 проставляется перед именем элемента ED202, который принадлежит пространству имен "urn:cbr-ru:ed:v2.0".

Элемент EDRefID принадлежит пространству имен "urn:cbr-ru:ed:v2.0", объявленному как пространство имен по умолчанию в родительском элементе ED202, его атрибуты принадлежат не специфицированному пространству имен. Следовательно, список названий пространств имен состоит из одного элемента "urn:cbr-ru:ed:v2.0", которому присваивается префикс n1. Данное пространство имен объявляется в элементе, ему назначается префикс n1. Префикс n1 проставляется перед именем элемента EDRefID, который принадлежит пространству имен "urn:cbr-ru:ed:v2.0".

Объявление пространства имен http://www.w3.org/2001/XMLSchema-instance удалено, т.к. ни элемент, ни его атрибуты не принадлежат данному пространству имен.

Данный пример достаточен для понимания процедуры упорядочения префиксов пространств имен в рамках УФЭБС. Поскольку пространства имен по умолчанию не используются, а объявляются только пространства имен, которым принадлежит элемент и его атрибуты, а атрибуты в XML-документе определяются в составе элемента, список названий пространств имен всегда состоит из одного элемента, которому присваивается префикс n1. Таким образом в рамках УФЭБС процедура упорядочения префиксов пространств имен во всех элементах сводится к объявлению в каждом элементе единственного пространства имен с префиксом n1 и проставлению данного префикса перед именем элемента.

2 Упорядочение префиксов пространств имен в абстрактном XML-документе

XML-документ до преобразования (строки пронумерованы для удобства комментирования):

```
1 <x:doc xmlns:x="http://w3.org/2"
2     xmlns:y="http://w3.org/1"
3     xmlns:z="http://w3.org/3"
4     xmlns:w="http://w3.org/4"
5     xmlns:n1="http://w2.org/1">
6   <x:e a="a"/>
7   <x:e x:a="x:a"/>
```



```

8      <e x:a="x:a"/>
9      <e x:a="x:a" a="a" xmlns="http://w2.org/2"/>
10     <e x:a="x:a" y:a="y:a"/>
11     <e x:a="x:a" x:b="x:b"/>
12     <x:e x:a="x:a" x:b="x:b" xmlns:v="http://w3.org/6"
xmlns:z="http://w3.org/5">
13         <z:e x:a="x:a" x:b="x:b" n1:c="n1:c"/>
14     </x:e>
15 </x:doc>

```

XML-документ после преобразования:

```

1      <n1:doc xmlns:n1="http://w3.org/2">
6          <n1:e a="a" xmlns:n1="http://w3.org/2"/>
7          <n1:e n1:a="x:a" xmlns:n1="http://w3.org/2"/>
8          <e n1:a="x:a" xmlns:n1="http://w3.org/2"/>
9          <n1:e a="a" n2:a="x:a" xmlns:n1="http://w2.org/2"
xmlns:n2="http://w3.org/2"/>
10         <e n1:a="y:a" n2:a="x:a" xmlns:n1="http://w3.org/1"
xmlns:n2="http://w3.org/2"/>
11         <e n1:a="x:a" n1:b="x:b" xmlns:n1="http://w3.org/2"/>
12         <n1:e n1:a="x:a" n1:b="x:b" xmlns:n1="http://w3.org/2">
13             <n3:e n1:c="n1:c" n2:a="x:a" n2:b="x:b"
xmlns:n1="http://w2.org/1" xmlns:n2="http://w3.org/2"
xmlns:n3="http://w3.org/5"/>
14         </n1:e>
15     </n1:doc>

```

Таблица 3 демонстрирует процедуру упорядочения префиксов пространств имен для данного примера.

Пространства имен "http://w3.org/1", "http://w3.org/3", "http://w3.org/4" и "http://w2.org/1", объявленные в элементе x:doc (строки 1-5), не подчиняются правилу упорядочения префиксов пространств имен (элемент не принадлежит этим пространствам имен) и удаляются, чтобы избежать конфликта имен (после упорядочения пространству имен "http://w3.org/2" будет назначен префикс n1, который уже назначен пространству имен "http://w2.org/1").

Пространства имен "http://w3.org/6" и "http://w3.org/5", объявленные в элементе x:e (строка 12), не подчиняются правилу упорядочения префиксов пространств имен и удаляются.

Таблица 4 – Упорядочение префиксов пространств имен во всех элементах абстрактного примера

Строки	Элемент и его атрибуты до и после упорядочения префиксов			Отсортированный список наименований пространств имен с назначенными префиксами	
	Имя до	Пространство имен, которому принадлежит элемент/атрибут	Имя после	Наименование	Префикс
1-5, 15	doc	"http://w3.org/2"*	n1:doc	"http://w3.org/2"	n1
6	x:e	"http://w3.org/2"*	n1:e	"http://w3.org/2"	n1
	a	не специфицировано	a		
7	x:e	"http://w3.org/2"*	n1:e	"http://w3.org/2"	n1
	x:a	"http://w3.org/2"*	n1:a		
8	e	не специфицировано	e	"http://w3.org/2"	n1
	x:a	"http://w3.org/2"*	n1:a		
9	e	"http://w2.org/2"*	n1:e	"http://w2.org/2"	n1
	a	не специфицировано	a	"http://w3.org/2"	n2
	x:a	"http://w3.org/2"*	n2:a		
10	e	не специфицировано	e	"http://w3.org/1"	n1
	x:a	"http://w3.org/2"*	n2:a	"http://w3.org/2"	n2
	y:a	"http://w3.org/1"*	n1:a		
11	e	не специфицировано	e	"http://w3.org/2"	n1
	x:a	"http://w3.org/2"*	n1:a		

Строки	Элемент и его атрибуты до и после упорядочения префиксов			Отсортированный список наименований пространств имен с назначенными префиксами	
	Имя до	Пространство имен, которому принадлежит элемент/атрибут	Имя после	Наименование	Префикс
	x:b	"http://w3.org/2"*	n1:b		
12, 14	x:e	"http://w3.org/2"*	n1:e	"http://w3.org/2"	n1
	x:a	"http://w3.org/2"*	n1:a		
	x:b	"http://w3.org/2"*	n1:b		
13	z:e	"http://w3.org/5"***	n3:e	"http://w2.org/1"	n1
	x:a	"http://w3.org/2"*	n2:a	"http://w3.org/2"	n2
	x:b	"http://w3.org/2"*	n2:b	"http://w3.org/5"	n3
	n1:c	"http://w2.org/1"*	n1:c		
* пространство имен объявлено в родительском элементе doc. ** пространство имен объявлено по умолчанию. *** пространство имен объявлено в родительском элементе x:e.					

D. Удаление в каждом элементе XML-документа дочерних текстовых узлов, содержащих только пробельные символы

Символьные данные группируются в узлы текста. В каждый из таких узлов помещается столько символьных данных, сколько возможно: с текстовым узлом ни до, ни после не может соседствовать какой-либо другой текстовый узел, имеющий того же родителя. Строковым значением текстового узла являются эти самые текстовые данные. Текстовый узел всегда содержит по крайней мере один символ данных.

Использование для передачи информации текстовых узлов в элементах в УФЭБС Банка России разрешается только в некоторых листовых (не содержащих дочерних) элементах, причем в таких элементах запрещена передача содержимого, состоящего только из пробельных символов. Во всех прочих текстовых узлах значащая информация содержаться не может (контролируется XML-схемой). Если текстовые узлы используются при форматировании XML-документа для улучшения читаемости, то они игнорируются при обработке ЭС, и не защищаются ЗК.

Алгоритм для удаления из элементов XML-документа дочерних текстовых узлов, содержащих только пробельные символы:

Для каждого узла дерева, представляющего XML-документ, содержащий защищаемые данные ЭС (пакета ЭС), выполняется следующее преобразование:

а) если данный узел является текстовым узлом, который содержит только пробельные символы, он удаляется.

Примечания

1 Пробельный символ (white space) состоит из одного или нескольких символов пробела (#x20), возврата каретки (#xD), конца строки (#xA) или табулятора (#x9). Пробельные символы могут служить для выделения разметки для лучшей читаемости.

2 Не удаляются текстовые узлы, содержащие хотя бы один символ, не относящийся к пробельным.

Пример

XML-документ до преобразования:

```
<?xml version="1.0" encoding="Windows-1251"?>
<n1:ED202 xmlns:n1="urn:cbr-ru:ed:v2.0"
  EDNo="8" EDDate="2003-04-14" EDAuthor="4525545000"
  EDReceiver="4525000000" InquiryCode="1">
  <n1:EDRefID xmlns:n1="urn:cbr-ru:ed:v2.0" EDNo="7" EDDate="2003-04-14"
  EDAuthor="4525545000"/>
</n1:ED202>
```

XML-документ после преобразования:

```
<?xml version="1.0" encoding="Windows-1251"?>
<n1:ED202 xmlns:n1="urn:cbr-ru:ed:v2.0">
```

```
EDNo="8" EDDate="2003-04-14" EDAuthor="4525545000"  
EDReceiver="4525000000" InquiryCode="1"><n1:EDRefID xmlns:n1="urn:cbr-  
ru:ed:v2.0" EDNo="7" EDDate="2003-04-14" EDAuthor="4525545000"/></n1:ED202>
```

Нормализация пробельных символов внутри тэгов зависит от парсера, обработавшего XML-документ. В данном примере пробельные символы внутри открывающего тэга <n1:ED202> оставлены без изменений.

9. Описание преобразований XML-документа для приведения к каноническому виду

В данном приложении описываются преобразования XML-документа для приведения к каноническому виду (канонизация). Данные преобразования применяются к XML-документу, содержащему защищаемые данные ЭС (пакета ЭС), и выполняются после нормализации XML-документа, содержащего защищаемые данные ЭС (пакета ЭС).

Идентификатор алгоритма, описывающего преобразования:

“<http://www.w3.org/2001/10/xml-c14n>”.

При описании алгоритма канонизации XML-документа для адресации частей XML-документа, содержащего защищаемые данные ЭС (пакета ЭС), используется язык [XPath].

XML-документ, содержащий защищаемые данные ЭС (пакета ЭС), после того, как к нему была применена процедура нормализации, необходимо канонизировать, чтобы его двоичное представление не зависело от конкретного парсера, обработавшего XML-документ, и от операционной системы.

Стандартная канонизация [XML-c14n] определяет список изменений в XML-документе, после применения которых может быть установлена логическая эквивалентность XML-документа с другим XML-документом в канонической форме.

Комментарии запрещены в УФЭБС и не должны защищаться ЗК. В связи с этим в рамках УФЭБС используется канонизация без комментариев.

Список изменений в XML-документе при канонизации без комментариев:

- XML-документ представляется в кодировке UTF-8;
- символы перевода строки нормализуются в #xA; до разбора XML-документа;
- нормализуются значения атрибутов;
- ссылки на символы и разобранные сущности заменяются соответствующим текстом;
- секции CDATA заменяются содержащимися в них символами;
- удаляются декларации XML `<?xml >` и декларации типа документа (DTD);
- пустые элементы преобразовываются в пару открывающих/закрывающих тегов (элемент вида `<Ename/>` к виду `<Ename><Ename/>`);
- нормализуются пустые пространства вне корневого элемента XML-документа, а также внутри открывающих и закрывающих тегов;
- остаются все пробельные символы в символьном содержимом (исключая символы, удаленные во время нормализации перевода строки);
- ограничительные символы атрибутов заменяются двойными кавычками;
- специальные символы в значениях атрибутов и символьном содержимом заменяются ссылками на символы;
- избыточные декларации пространств имен удаляются из каждого элемента (декларация пространства имен считается избыточной, если ближайший родительский элемент в канонической форме имеет эквивалентное объявление пространства имен в своей области видимости);
- атрибуты по умолчанию, описанные в DTD, добавляются к каждому элементу;
- внутри каждого элемента устанавливается лексикографический порядок следования для деклараций пространств имен и атрибутов (атрибуты упорядочиваются по наименованиям имен пространств и по локальным именам);
- комментарии удаляются.

Алгоритм для приведения XML-документа, содержащего защищаемые данные ЭС (пакета ЭС) к каноническому виду:

Примечание – данный алгоритм приведен для XML-документа в целом, а не для подмножества его узлов. Особенности алгоритма применительно к подмножеству узлов (document subset), представлены в оригинальной спецификации [XML-c14n].

Алгоритм оперирует набором узлов (node-set) [XPath], полностью представляющих канонизируемый XML-документ. Обработчик XML-документа (парсер) при подготовке набора узлов должен выполнить следующие задачи:

- нормализация концов строк в соответствии со спецификацией [XML];
- нормализация значений атрибутов в соответствии со спецификацией [XML];
- замещение секций CDATA содержащимися в них символами в соответствии со спецификацией [XML];
- разрешение символьных ссылок и ссылок на разобранные сущности в соответствии со спецификацией [XML].

Набор узлов [XPath], подготовленный парсером, преобразуется в массив байтов путем генерации соответствующих UCS символов для каждого узла в наборе в порядке появления в XML-документе. После этого результат перекодируется в UTF-8. Каждый узел обрабатывается только один раз. Обработка узла элемента включает обработку всех дочерних узлов, для которых данный узел элемента является предком. Таким образом после того, как каноническая форма для элемента будет сгенерирована, узел данного элемента и все узлы, для которых данный узел является предком, удаляются из множества узлов. Узел элемента не удаляется из набора узлов до тех пор, пока не будут обработаны все его дочерние узлы.

Текст, генерируемый для каждого узла, зависит от типа узла.

Корневой узел. Результат обработки складывается из результатов обработки каждого дочернего узла в порядке появления в XML-документе. Каноническая форма не включает ни декларацию XML, ни декларацию типа документа.

Узел элемента. В результате обработки генерируется следующий текст: открывающая угловая скобка (<), квалифицированное имя элемента, результат обработки узлов пространств имен, относящихся к текущему узлу элемента, результат обработки узлов атрибутов данного элемента, закрывающая угловая скобка (>), результат обработки дочерних узлов элементов в порядке появления в XML-документе, открывающая угловая скобка (<), прямой слэш (/), квалифицированное имя элемента и закрывающая угловая скобка (>).

- Узлы пространств имен, относящихся к текущему узлу элемента обрабатываются в лексикографическом порядке. Если первый узел пространства имен в списке не является узлом пространства имен по умолчанию (узел пространства имен по умолчанию имеет пустой URI, пустое локальное имя, но не пустое содержимое), а ближайший элемент-предок текущего узла элемента содержит узел пространства имен по умолчанию, тогда результат обработки узлов пространств имен элемента следует за следующим текстом: пробел, атрибут, отменяющий объявление пространства имен по умолчанию (xmlns="").

- Узлы атрибутов обрабатываются в лексикографическом порядке.

Узел пространства имен. Узел пространства имен игнорируется, если ближайший элемент-предок родительского узла элемента имеет узел пространства имен с тем же локальным именем и значением, как и текущий узел пространства имен. В противном случае узел пространства имен обрабатывается так же, как и узел атрибута, за исключением назначения локального имени xmlns узлу пространства имен по умолчанию, если он существует (узел пространства имен по умолчанию имеет пустой URI и пустое локальное имя).

Узел атрибута. В результате обработки генерируется следующий текст: пробел, квалифицированное имя атрибута, знак равенства, открывающие двойные кавычки, модифицированное строковое значение и закрывающие двойные кавычки. Модифицированное строковое значение представляет собой строковое значение атрибута, нормализованное в соответствии со спецификацией [XML]. Следующие символы модифицированного строкового значения заменяются ссылками на символы:

- амперсанд (&) -> &
- открывающая угловая скобка (<) -> <
- двойные кавычки (") -> "

– пробельные символы (#x9, #xA, and #xD) -> символьными ссылками (причем, символьные ссылки представляются в шестнадцатеричном формате, в верхнем регистре, без ведущих нулей).

Текстовый узел. В результате обработки генерируется текст, представляющий строковое значение текстового узла, концы строк которого нормализованы в соответствии со спецификацией [XML]. Следующие символы строкового значения текстового узла заменяются ссылками на символы:

- амперсанд (&) -> &
- открывающая угловая скобка (<) -> <
- закрывающая угловая скобка (>) -> >
- символ возврата каретки (#xD) -> 

Узел инструкции обработки. В результате обработки генерируется следующий текст: открывающие символы инструкции обработки (<?), имя адресата инструкции обработки, ведущий пробел, строковое значение, если оно не пусто, и закрывающие символы инструкции обработки (?>). Если строковое значение пусто, ведущий пробел не добавляется.

Примечание – в результате нормализации набор узлов XML-документа, содержащего ЭС (пакет ЭС), не должен содержать узлы инструкций обработки.

Узлы комментариев игнорируются при канонизации без комментариев.

Примеры

1 Канонизация XML-документа

XML-документ до преобразования:

```
<?xml version="1.0" encoding="Windows-1251"?>
<n1:ED202 xmlns:n1="urn:cbr-ru:ed:v2.0"
  EDNo="8" EDDate="2003-04-14" EDAuthor="4525545000"
  EDReceiver="4525000000" InquiryCode="1"><n1:EDRefID xmlns:n1="urn:cbr-
ru:ed:v2.0" EDNo="7" EDDate="2003-04-14" EDAuthor="4525545000"/></n1:ED202>
```

XML-документ после преобразования:

```
<n1:ED202 xmlns:n1="urn:cbr-ru:ed:v2.0" EDAuthor="4525545000" EDDate="2003-04-
14" EDNo="8" EDReceiver="4525000000" InquiryCode="1"><n1:EDRefID
EDAuthor="4525545000" EDDate="2003-04-14" EDNo="7"></n1:EDRefID></n1:ED202>
```

После преобразования в XML-документе:

- удалена декларация XML <?xml version="1.0" encoding="Windows-1251"?>;
- нормализованы пустые пространства внутри открывающего тэга <n1:ED202>;
- удалена избыточная декларация имен пространств "urn:cbr-ru:ed:v2.0" в элементе n1:EDRefID;
- пустой элемент n1:EDRefID преобразован в пару открывающих/закрывающих тегов: <n1:EDRefID></n1:EDRefID>;
- внутри элементов n1:ED202 и n1:EDRefID установлен лексикографический порядок следования для деклараций пространств имен и атрибутов.

2 Канонизация абстрактного XML-документа

XML-документ до преобразования (строки пронумерованы для удобства комментирования, применено форматирование XML-документа для удобства восприятия):

```
1 <n1:doc xmlns:n1="http://w3.org/2">
2   <n1:e a="a" xmlns:n1="http://w3.org/2"/>
3   <n1:e n1:a="x:a" xmlns:n1="http://w3.org/2"/>
4   <e n1:a="x:a" xmlns:n1="http://w3.org/2"/>
5   <n1:e a="a" n2:a="x:a" xmlns:n1="http://w2.org/2"
        xmlns:n2="http://w3.org/2"/>
6   <e n1:a="y:a" n2:a="x:a" xmlns:n1="http://w3.org/1"
        xmlns:n2="http://w3.org/2"/>
7   <e n1:a="x:a" n1:b="x:b" xmlns:n1="http://w3.org/2"/>
8   <n1:e n1:a="x:a" n1:b="x:b" xmlns:n1="http://w3.org/2">
9     <n3:e n1:c="n1:c" n2:a="x:a" n2:b="x:b"
          xmlns:n1="http://w2.org/1"
          xmlns:n2="http://w3.org/2">
```

```

10         xmlns:n3="http://w3.org/5"/>
11     </n1:e>
12 </n1:doc>

```

XML-документ после преобразования:

```

1  <n1:doc xmlns:n1="http://w3.org/2">
2      <n1:e a="a"></n1:e>
3      <n1:e n1:a="x:a"></n1:e>
4      <e n1:a="x:a"></e>
5      <n1:e xmlns:n1="http://w2.org/2" xmlns:n2="http://w3.org/2" a="a"
n2:a="x:a"></n1:e>
6      <e xmlns:n1="http://w3.org/1" xmlns:n2="http://w3.org/2"
n1:a="y:a" n2:a="x:a"></e>
7      <e n1:a="x:a" n1:b="x:b"></e>
8      <n1:e n1:a="x:a" n1:b="x:b">
9          <n3:e xmlns:n1="http://w2.org/1" xmlns:n2="http://w3.org/2"
xmlns:n3="http://w3.org/5" n1:c="n1:c" n2:a="x:a" n2:b="x:b"></n3:e>
10     </n1:e>
11 </n1:doc>

```

После преобразования в XML-документе:

- нормализованы пустые пространства внутри открывающих тэгов (строки 5, 6, 9);
- удалены избыточные декларации имен пространств (строки 2, 3, 4, 7, 8);
- пустые элементы преобразованы в пары открывающих/закрывающих тегов (строки 2-7, 9);
- внутри элементов установлен лексикографический порядок следования для деклараций пространств имен и атрибутов (строки 5, 6, 9).

10. Описание алгоритма кодирования Base64

В данном приложении описывается алгоритм кодирования [base64]. Данный алгоритм применяется к значению КА или ЗК, а также к ЭС (пакету ЭС) при формировании/проверке КА.

Идентификатор алгоритма, описывающего преобразования:

“<http://www.ietf.org/rfc/rfc2045#base64>”.

Примечание – в приложении приведен обзор алгоритма. Особенности алгоритма представлены в оригинальной спецификации [base64].

Алгоритм [base64] разработан для представления последовательности октетов в форме, которая не предназначена для чтения человеком. Алгоритмы кодирования и декодирования просты, однако закодированные данные занимают больший объем по сравнению с исходными данными примерно на 33%.

Для кодирования используется 65 символов US-ASCII, позволяющие представить 6-битовую последовательность в виде печатного символа (65-й символ используется в качестве заполнителя).

Алгоритм кодирования представляет 24-битовую группу входных бит в виде выходной строки из 4-х символов. Обработка входного потока происходит слева направо, 24-битовая группа формируется путем конкатенации трех 8-битовых входных групп. Полученные подобным образом 24 бита трактуются как четыре объединенных 6-битовых группы, каждая из которых транслируется в отдельный символ из алфавита Base64.

Каждая 6-битовая группа используется в качестве индекса для массива из 64 печатных символа. Символ, полученный по индексу, помещается в выходную строку. Таблица F.1 демонстрирует символы, используемые при кодировании, которые составляют алфавит Base64.

Т а б л и ц а 5 – Алфавит Base64

Индекс	Символ	Индекс	Символ	Индекс	Символ	Индекс	Символ
0	A	17	R	34	i	51	z
1	B	18	S	35	j	52	0
2	C	19	T	36	k	53	1
3	D	20	U	37	l	54	2
4	E	21	V	38	m	55	3
5	F	22	W	39	n	56	4
6	G	23	X	40	o	57	5
7	H	24	Y	41	p	58	6
8	I	25	Z	42	q	59	7
9	J	26	a	43	r	60	8
10	K	27	b	44	s	61	9
11	L	28	c	45	t	62	+
12	M	29	d	46	u	63	/
13	N	30	e	47	v	Заполнит ель	=
14	O	31	f	48	w		
15	P	32	g	49	x		
16	Q	33	h	50	y		

Рисунок 1 демонстрирует работу алгоритма кодирования [base64].

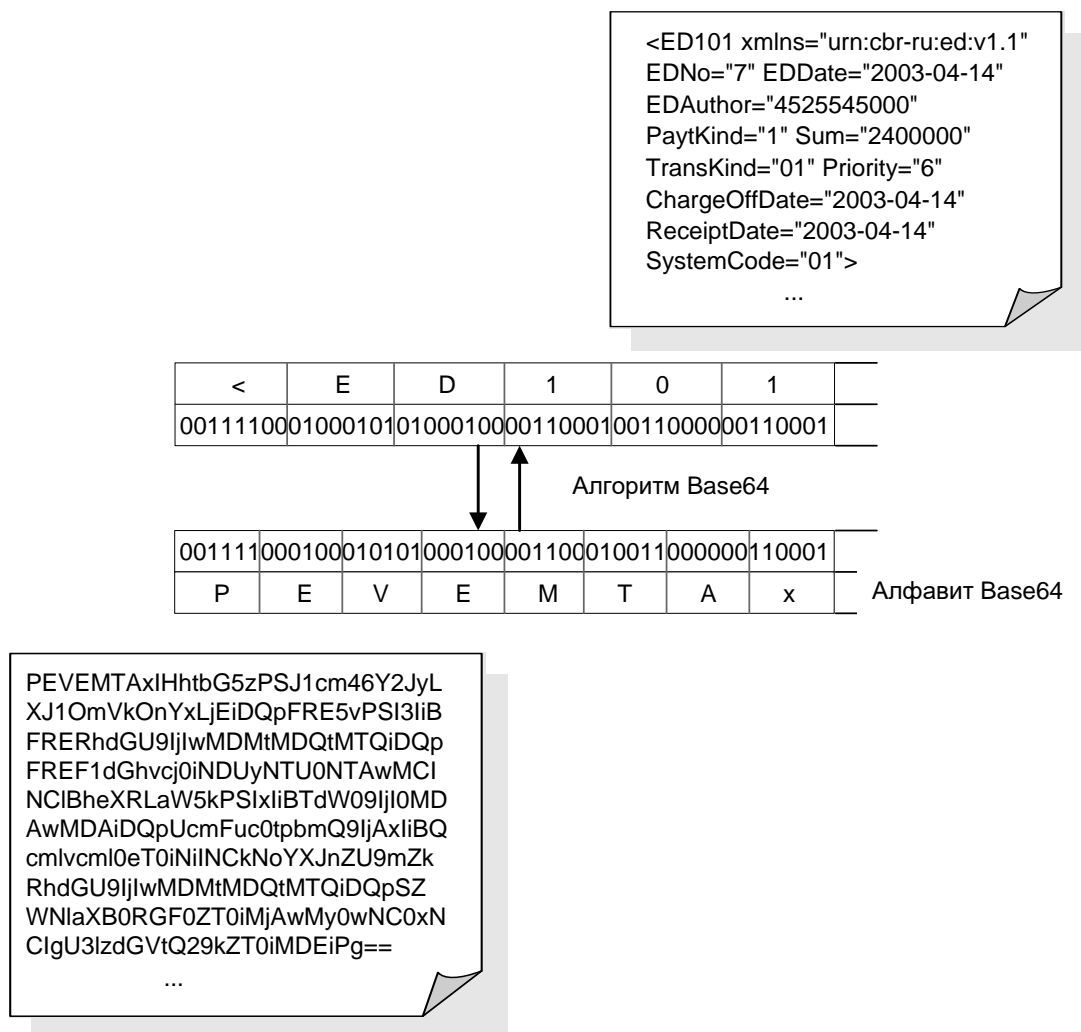


Рисунок 1– Демонстрация работы алгоритма кодирования Base64

Выходной поток может представляться строками не более, чем по 76 символов в каждой. Все символы конца строк и прочие символы, не представленные в алфавите Base64 должны игнорироваться при раскодировании.

Если последняя группа бит во входном потоке содержит менее 24 бит, требуется специальная обработка: справа добавляются нулевые биты до тех пор, пока не будет сформировано целое число 6-битовых групп. Если в результате получилось менее четырех 6-битовых групп, недостающие группы заменяются символами заполнителями (=). Поскольку входной поток всегда содержит целое число октетов (8-битовых групп), возможны три ситуации:

- Финальная группа во входном потоке содержит точно 24 бита. В этом случае финальная группа в выходном потоке содержит 4 символа без символов-заполнителей.
- Финальная группа во входном потоке содержит точно 8 бит. В этом случае финальная группа в выходном потоке содержит 2 символа, за которыми следуют два символа-заполнителя (=).
- Финальная группа во входном потоке содержит точно 16 бит. В этом случае финальная группа в выходном потоке содержит 3 символа, за которыми следует один символ-заполнитель (=).

В связи с тем, что символ "=" используется только в качестве заполнителя, появление такого символа может служить признаком того, что достигнут конец данных. Однако не следует забывать, что если количество входных октетов кратно трем, символ "=" в выходном потоке никогда не появится.

Любые символы вне алфавита Base64 должны игнорироваться в данных, закодированных по алгоритму [base64].