Информационные технологии

Электронные документы и задача обеспечения сохранности при обмене данными в цифровой экономике

А.В. Соловьев, И.А. Тарханов

Аннотация. В статье рассматривается задача обеспечения сохранности электронных документов при обмене данными между информационными системами в цифровой экономике. Целью данного исследования является создание теоретической основы управления безопасным информационным обменом электронных документов в динамически меняющейся программно-аппаратной среде их хранения и передачи.

Ключевые слова: *электронный документ, электронный архив, цифровая экономика, блокчейн, электронная подпись, распределенные реестры, системы хранения, деперсонализация данных.*

Введение

Популярный ныне термин «цифровая экономика» был введен в 1995 году американским информатиком из Массачусетского технологического института Николасом Негропонте. При этом подразумевалось, что речь идет не столько о разработке и продаже программного обеспечения, сколько об электронных товарах и услугах, производимых электронным бизнесом и электронной коммерцией. Однако фактически цифровая экономика затрагивает многие области, например, банковское дело, образование, здравоохранение. С легкой руки политиков в последнее время все чаще говорится о создании Электронного государства и Электронного правительства, т.е. о том, что общение граждан с государственными органами может быть перенесено на цифровую платформу.

Кроме того, задача безопасной передачи данных в распределенной среде актуальна и для коммерческих организаций. Большинство крупных финансовых, страховых и занятых в сфере образования и здравоохранения компаний имею развитый парк информационных систем (ИС) во множестве регионов. Создание единого информационного пространства холдинга – важнейшая задача для их дальнейшего развития, которая требует новых технологических и организационных подходов.

Теоретически современные информационные системы делают возможным развитие всех аспектов цифровой экономики. Однако нужно понять, как будет осуществляться передача информации по цифровым каналам между государством и отраслями цифровой экономики, государством и гражданами. Ответ на этот вопрос дает анализ современного состояния деловых процессов в экономике. Для любого элементарного делового процесса входами и выходами являются потоки информации, представляющие собой электронные документы (ЭлД) или сопровождающиеся ЭлД.

В первую очередь это деловые документы, т.е. которые регламентируют деятельность ведомств, организаций, учреждений, предприятий, фирм, а также документы, связанные с деятельностью юридических и физических лиц.

В настоящее время общая тенденция развития цифровых средств работы с ЭлД говорит о том, что в ближайшее время вытеснение бумажных документов станет массовым явлением. Уже сейчас, «де факто» во многих государственных и коммерческих организациях ЭлД начинают активно замещать документы «бумажные». «Де юро» ЭлД, согласно законодательным актам РФ (например, см. №379-ФЗ от 2014 года), становятся равнозначны привычным «бумажным» документам. Это означает, что в будущей цифровой экономике именно ЭлД будут циркулировать по каналам связи между ИС организаций, физическими лицами и государственными органами, органами здравоохранения и образования.

Так как общение организаций между собой, граждан с государством происходит именно с помощью документов, это порождает определенные вызовы, связанные с надежностью и безопасностью такого общения при помощи различных цифровых информационных технологий. Дополнительно учитывая тот факт, что надежность и безопасность ЭлД должна обеспечиваться в условиях изменяющейся программно-аппаратной цифровой среды их хранения и обработки, задача обеспечения сохранности ЭлД становится совсем нетривиальной.

Действительно, цифровая программно-техническая среда, т.е. набор технических (компьютеры, устройства хранения, записи, воспроизведения) и программных средств (операционные системы, средства создания и просмотра ЭлД) – это динамично меняющаяся во времени структура, работая с которой важно учитывать такие факторы как технологическое старение, износ, обновление программной среды, устаревание средств защиты информации и форматов ЭлД.

Можно легко показать, что жизненный цикл ЭлД превосходит сроки жизни оборудования и программного обеспечения. Согласно перечню типовых архивных документов, образующихся в научно-технической и производственной деятельности организаций (в редакции приказа Минкультуры от 28.04.2011 № 412) большинство контрактов, документов относящихся к переписке и авторскому праву хранятся не менее 5 лет с момента окончания их срока действия. А так называемые документы по личному составу должны храниться 75 лет. За это время несколько раз будет обновлено как системное ПО, так и цифровые носители данных, однако ЭлД в этих условиях должны остаться неизменными.

Данная противоречивая ситуация определяет необходимость решения важной научно-технической проблемы обеспечения сохранности ЭлД при электронном обмене данными между ИС в цифровой экономике, включая доступность, аутентичность (неизменность), интерпретируемость (читаемость) ЭлД, обеспечивая при этом разграничение доступа к данным и защиту конфиденциальных (в том числе персональных) данных.

Данная статья посвящена концептуальному решению данной проблемы.

1. Определения и основные понятия

Сохранность – свойство электронного документа существовать в качестве доступного и аутентичного свидетельства (доказательства) в произвольный момент времени.

Долговременная сохранность – обеспечение свойства сохранности ЭлД в течение срока не менее 5 лет.

Определение не претендует на «абсолютность». За минимальную границу срока долговременной сохранности (5 лет) взят максимальный срок хранения документов в оперативных архивах систем электронного документооборота (СЭД). В общем случае ЭлД могут сохраняться в течение десятилетий или даже столетий или «бессрочно» в зависимости от их важности.

Документ – структурированная информация, представляющая собой совокупность взаимосвязанных семантических блоков. Деловой документ, безусловно, имеет четкую структуру, форму и содержание.

Деловой документ – документ, регламентирующий деятельность организаций, учреждений, предприятий, фирм, а также связанный с деятельностью юридических и физических лиц.

Электронный документ – документ, семантические блоки которого и взаимосвязи между ними представлены в электронно-цифровой форме.

Семантические блоки – некоторые фрагменты документа, выделенные по смысловому содержанию, так как всякий реальный документ разбивается на взаимосвязанные части: разделы, подразделы, пункты и т.д.

Подробнее о математической модели ЭлД см. [16].

Аутентичный электронный документ – «электронный документ, точность, надежность и целостность которого сохраняются с течением времени» [5].

Доступность информации – «возможность реализации беспрепятственного доступа к информации субъектов, имеющих на это надлежащие полномочия» [6] и, одновременно, «избежание временного или постоянного сокрытия этой информации от пользователей, получивших права доступа» [7].

Доступность документа – «свойство документа, состоящее в том, что форма его представления обеспечивает физическую возможность измерения заданных параметров этого представления документа (содержания, атрибутов, технологии) заданными средствами в заданных точках за конечное время» [8].

Электронная подпись (ЭП) – «информация в электронной форме, которая присоединена к другой информации в электронной форме (подписываемой информации) или иным образом связана с такой информацией и которая используется для определения лица, подписывающего информацию» [9].

Блокчейн (*blockchain*) – «выстроенная по определенным правилам непрерывная последовательная цепочка блоков (связный список), содержащих информацию. Чаще всего копии цепочек блоков хранятся и независимо друг от друга (чрезвычайно параллельно), обрабатываются на множестве разных компьютеров» [13].

Смарт-контракты – фрагменты исполняемого, неизменяемого кода, хранящегося в блоках блокчейна, исполняемого при определенных условиях только лицами, имеющими определенные права доступа к смарт-контракту. В контексте данной статьи определение смарт-контракта отличается от более распространенных: «смарт-контракты – фрагменты кода, хранящиеся в блокчейне, с помощью которых можно обмениваться деньгами, собственностью, акциями или другими активами, не прибегая к услугам посредников» [14] либо «компьютерные программы, связанные с криптовалютами и другими способами хранения информации» [15].

Майнинг – вычислительная деятельность по поддержанию распределенной платформы и созданию новых блоков блокчейна. Производимые вычисления требуются для обеспечения аутентичности блоков.

Деперсонализация данных – представление персональных и/или конфиденциальных данных в виде, не позволяющем восстановить какую-либо информацию о субъекте персональных данных или о защищаемой конфиденциальной информации*.

2. Обзор современных средств обеспечения сохранности ЭлД при информационном обмене

Наиболее простым способом организации обмена ЭлД между организациями является создание отдельной внешней, по отношению ко всем участникам обмена базы данных или информационной системы. Но создание такой системы требует серьезных организационных и финансовых затрат. Также до недавнего времени такое решение имело ряд технических ограничений, связанных с отсутствием надежных СУБД, способных обеспечивать распределенное хранение и многопользовательский режим работы.

Первой предпосылкой для развития реальных распределенных систем стало появление распределенных и параллельных СУБД [17], которые за счет механизмов репликации, вертикального и горизонтального шардинга, обеспечили относительно недорогое хранение данных и прозрачный перенос данных между узлами СУБД. Но такие СУБД сами по себе не решают проблемы аутентичности ЭлД, проблемы разграничения доступа, а также неприменимы в динамично изменяющейся среде, где быстро меняются программные и аппаратные средства.

Появление программных реализаций связанных штампов времени (Linked timestamping) – сбора транзакций в блоки и связывания их при помощи хеш-функций в 90-е годы [18] и одновременно развитие технологий распределенных вычислений, позволяющих создавать сети из недорогих компьютеров, устойчивых к падениям и выведению узлов из строя путем захвата, позволило создать первый относительно безопасный распределенный реестр, в котором можно было хранить данные ЭлД.

Дальнейшее бурное развитие этого подхода в 2000-х годах привело к созданию блокчейн технологии, которая к принципу хранения Linked timestamping добавила применение современных средств шифрования и хэширования, а также введение специальных узлов (майнеров), которые занимаются регистрацией транзакций в общей сети, используя предназначенные для этого алгоритмы согласия (proof-of-work или proof-of-stake), гарантируя аутентичность зарегистрированных в сети данных.

Многочисленные реализации данной технологии (в первую очередь, при создании криптовалют) различаются размерами и структурой регистрируемых блоков, применением различных алгоритмов согласия для организации работы сети, наличием дополнительных функций, таких как смарт-контракты [19].

Каким образом сейчас обеспечивается безопасный обмен деловыми ЭлД? Самый яркий пример – это проект Peppol, созданный в рамках инициатив Европейского Союза (ЕС), система обмена ЭлД между представителями государств ЕС и их контрагентами. В каждой стране ЕС существует авторизованный Peppol-оператор, который осуществляет прием сообщений от других авторизованных операторов (не только из стран ЕС), и перенаправляет их конечным организациям. Внутри Peppol существует согласованный формат обмена данных, стандарт шифрования, основанный на открытых ключах, а также используется самостоятельная инфраструктура центров сертификации [20].

Другим более близким примером глобальной системы обмена ЭлД является российская система межведомственного электронного документооборота (МЭДО), которая обеспечивает обмен ЭлД

* Определение дано по №152 ФЗ «О персональных данных» от 26.07.2006 г.

между государственными органами и система межведомственного электронного взаимодействия (СМЭВ), которая обеспечивает гражданам РФ доступ к услугам в электронной форме через портал государственных услуг РФ (ЕПГУ) и другие региональные и федеральные сервисы.

Как в случае Peppol, так и в случае МЭДО/СМЭВ, весь обмен ложится на плечи операторов. Они же являются регуляторами этого обмена, устанавливают форматы и требования информационной безопасности обмена данными. При этом риск нарушения аутентичности передаваемых ЭлД находится на том же уровне, что и в случае традиционных систем с централизованным хранением данных.

3. Разработка постановки для задачи обеспечения сохранности ЭлД при электронном обмене данными между ИС в цифровой экономике

Задача обеспечения сохранности ЭлД при электронном обмене данными между ИС в цифровой экономике может быть сформулирована в общем виде следующим образом: требуется обеспечить неизменность свойства сохранности ЭлД, включая доступность, аутентичность (неизменность), интерпретируемость (читаемость) ЭлД, соблюдение требований информационной безопасности при информационном обмене между ИС, функционирующими в рамках цифровой экономики.

Чтобы не учитывать историю создания ЭлД до начала его размещения в БД исходной прикладной ИС и, в то же время искусственно не суживать возможные формы представления ЭлД, должны быть приняты следующие ***допущения***:

- аутентичность ЭлД на момент его создания и размещения в БД прикладной ИС одинакова и подтверждена (например, с помощью ЭП);
- ЭлД интерпретируем при создании и размещении в БД прикладной ИС;
- нет ограничений на форматы данных передаваемых документов;
- аппаратно-программная среда, в которой происходит обмен ЭлД, подвержена постоянному изменению: могут быть установлены новые версии операционных систем (ОС), и прикладного программного обеспечения, обеспечивающего интерпретацию ЭлД, изменяться технические устройства, обеспечивающие обработку и хранение ЭлД.

Представим ситуацию, в которой пользователь исходной прикладной ИС $\boldsymbol{IS_0}$ создает некоторый исходный ЭлД $\boldsymbol{d_0}$, обладающий исходной сохранностью $\boldsymbol{SV_0}$. Это может быть заявление о прикреплении к поликлинике, заявление при подаче документов в ВУЗ, конкурсная документация для электронных торгов и т.д. В ответ на этот ЭлД, а также вследствие движения данного ЭлД между пользователями ИС и/или между разными ИС, другими пользователями, не обязательно исходной ИС, создаются иные ЭлД $\boldsymbol{d_i}$, $\boldsymbol{i} \in [1,\boldsymbol{N}]$, где $\boldsymbol{N}$ – общее количество связанных документов. Все ЭлД $\boldsymbol{d_i}$, $\boldsymbol{i} \in [0,\boldsymbol{N}]$ являются связанными между собой и принадлежат множеству ЭлД $\boldsymbol{D}$. Связи ЭлД представляют собой граф $\boldsymbol{G}(\boldsymbol{V}, \boldsymbol{A})$, где $\boldsymbol{V} = \boldsymbol{D}$, $\boldsymbol{A} = \{ (\boldsymbol{d_i}, \boldsymbol{d_j}) \}$, $\boldsymbol{d_i} \in \boldsymbol{D}$, $\boldsymbol{d_j} \in \boldsymbol{D}$. Исходный ЭлД перемещается между пользователями ИС или разными ИС в соответствии с некоторым бизнес-процессом $\boldsymbol{BP_j}$, $\boldsymbol{j} \in [1,\boldsymbol{M}]$, где $\boldsymbol{M}$ – общее количество бизнес-процессов некоторого множества $\boldsymbol{BP}$.

В процессе движения ЭлД от исходной ИС $\boldsymbol{IS_0}$ к другим ИС $\boldsymbol{IS_k}$, $\boldsymbol{k} \in [1,\boldsymbol{K}]$, где $\boldsymbol{K}$ – общее количество ИС задействованных в данном обмене в рамках цифровой экономики, с высокой вероятностью возникает необходимость разграничения прав доступа на передаваемую информацию ЭлД (персональные данные, коммерческая или государственная тайна и др.). В этом случае возникает задача деперсонализации данных ЭлД при электронном обмене данными ЭлД, т.е. введения некоторой функции $\boldsymbol{f}$: $\boldsymbol{d_i} \rightarrow \boldsymbol{d_i'}$, где $\boldsymbol{d_i'} \in \boldsymbol{D}$ – новый аутентичный, сохранный и интерпретируемый ЭлД с «деперсонализированными» данными. При этом «деперсонализация» данных определяется множеством требований $\boldsymbol{T} = \{ \boldsymbol{NTD_l} \}$, определяемых в нормативно-технической документации (НТД), которая составляет основу политик информационной безопасности, хранения и обработки персональных данных и т.д.

Кроме того, в процессе создания, хранения и движения документа на него влияет множество параметрических возмущений среды хранения и передачи $\boldsymbol{E} = \{ \boldsymbol{\varepsilon_q} \}$, дестабилизирующих $\boldsymbol{d_i}$.

В этих условиях необходимо, чтобы сохранность всех $\boldsymbol{d_i} \in \boldsymbol{D}$ была равна $\boldsymbol{SV_0}$.

Математическая постановка задачи может быть сформулирована в следующем виде:

Дано:

1. Множество ЭлД $\boldsymbol{D} = \{ \boldsymbol{d_i} \}$.
2. Множество параметрических возмущений среды хранения и передачи ЭлД $\boldsymbol{E} = \{ \boldsymbol{\varepsilon_q} \}$:

$\boldsymbol{\varepsilon_1}$ – нарушение аутентичности $\boldsymbol{D}$,
$\boldsymbol{\varepsilon_2}$ – нарушение интерпретируемости $\boldsymbol{D}$,
$\boldsymbol{\varepsilon_3}$ – изменения программно-аппаратной среды $\boldsymbol{D}$,
$\boldsymbol{\varepsilon_4}$ – нарушение надежности хранения и передачи $\boldsymbol{D}$,
$\boldsymbol{\varepsilon_5}$ – нарушение информационной безопасности $\boldsymbol{D}$.

3. Множество ИС $\boldsymbol{IS} = \{ \boldsymbol{IS}_k \}$ подверженное $\boldsymbol{E}$ и осуществляющее воздействие $\boldsymbol{f}$: $\boldsymbol{d}_i \rightarrow \boldsymbol{d}_i'$, где $\boldsymbol{d}_i' \in \boldsymbol{D}$ и организующих обработку, передачи и хранение $\boldsymbol{D}$ в соответствии с заданным множеством бизнес-процессов $\boldsymbol{BP}$.
4. Множество требований безопасности информации $\boldsymbol{T} = \{ \boldsymbol{NTD}_l \}$, определяемых в НТД.
5. Функция состояния ЭлД, определяющая принадлежность ЭлД к бизнес процессу: $\boldsymbol{state}(\boldsymbol{d}_i) = \boldsymbol{BP}_j$.

Найти:

1. Множество функций $\boldsymbol{\Phi} = \{ \boldsymbol{\Phi}_r \}$, обеспечивающих сохранность (включая доступность, аутентичность (неизменность), интерпретируемость (читаемость), соблюдение требований информационной безопасности при информационном обмене $\boldsymbol{d}_i$ между $\boldsymbol{IS}_k$) $\boldsymbol{D}$ на уровне $\boldsymbol{SV}_0$: $\boldsymbol{\Phi}_r(\boldsymbol{d}_i) = \boldsymbol{SV}_0$.
2. Множество технических решений $\boldsymbol{R} = \{ \boldsymbol{R}_r \}$, обеспечивающих сохранность (включая доступность, аутентичность (неизменность), интерпретируемость (читаемость), соблюдение требований информационной безопасности при информационном обмене $\boldsymbol{d}_i$ между $\boldsymbol{IS}_k$) $\boldsymbol{D}$ на уровне $\boldsymbol{SV}_0$: $\boldsymbol{R}_r(\boldsymbol{d}_i) = \boldsymbol{SV}_0$.

Утверждение 1. Можно утверждать, что $\boldsymbol{d}_i$ становится объектом управления, а задача обеспечения сохранности при электронном обмене данными между $\boldsymbol{IS}_k$ – это задача оптимального управления в условиях параметрических возмущений [10, 11]. Для обеспечения оптимального управления необходимо разработать технические решения для стабилизации объекта управления [12].

Утверждение 2. Задача контроля сохранности $\boldsymbol{d}_i$, как объекта управления, представляет собой задачу оптимального выбора по многим критериям.

Действительно, пусть сохранность $\boldsymbol{d}_i$, характеризуется функцией $\boldsymbol{\mu(t)}$, значение которой представляет собой вероятность сохранности $\boldsymbol{d}_i$ на уровне $\boldsymbol{SV}_0$ в произвольный момент времени $\boldsymbol{t}$. Тогда можно утверждать, что задача обеспечения сохранности $\boldsymbol{d}_i$ формулируется, как задача достижения максимума функции $\boldsymbol{\mu(t)}$ на произвольном временном интервале $\boldsymbol{t}$.

Т.е.: $\boldsymbol{M} = \mathbf{max}_{\,t \in [t0,\, \infty]}\, \boldsymbol{\mu(t)}$, где $\boldsymbol{t0}$ – момент времени создания $\boldsymbol{d}_i$ в $\boldsymbol{IS}_0$.

При этом функция $\boldsymbol{\mu(t)}$ представляет собой взвешенную аддитивную или мультипликативную свертку функций вероятностей сохранения аутентичности, интерпретируемости, устойчивости к изменениям программно-аппаратной среды хранения, надежности и информационной безопасности.

Например, $\mu(t) = \omega_1 \alpha(t) + \omega_2 \zeta(t) + \omega_3 \varphi(t) + \omega_4 \rho(t) + + \omega_5 \sigma(t)$.

Или $\mu(t) = \alpha(t)^{\omega 1}\, \zeta(t)^{\omega 2}\, \varphi(t)^{\omega 3}\, \rho(t)^{\omega 4}\, \sigma(t)^{\omega 5}$, где $\sum \omega_i = 1$, $\omega_i > 0$, $\boldsymbol{i}$=[1,6];

$\alpha(t)$ – вероятность сохранения аутентичности ЭлД на произвольном временном интервале $\boldsymbol{t}$;

$\zeta(t)$ – вероятность сохранения интерпретируемости ЭлД на произвольном временном интервале $\boldsymbol{t}$;

$\varphi(t)$ – вероятность сохранения устойчивости к изменениям программно-аппаратной среды хранения ЭлД на произвольном временном интервале $\boldsymbol{t}$;

$\rho(t)$ – вероятность сохранения надежности, определяющая вероятность работоспособности программно-аппаратной среды хранения ЭлД на произвольном временном интервале $\boldsymbol{t}$;

$\sigma(t)$ – вероятность сохранения устойчивости программно-аппаратной среды хранения ЭлД к угрозам нарушения информационной безопасности на произвольном временном интервале $\boldsymbol{t}$.

4. Концептуальное техническое решение для задачи обеспечения сохранности ЭлД при электронном обмене данными между ИС в цифровой экономике

Попробуем сформулировать концептуальное техническое решение для поставленной задачи.

Итак, мы выяснили, что задача обеспечения сохранности требует стабилизации характеристик сохранности ЭлД $\boldsymbol{d}_i$ на исходном уровне $\boldsymbol{SV}_0$. При этом должно быть обеспечено разграничение прав доступа к данным ЭлД в соответствии с принятой моделью обеспечения информационной безопасности (ИБ) и в соответствие с $\boldsymbol{T} = \{ \boldsymbol{NTD}_l \}$.

Т.е. нужно обеспечить аутентичность, интерпретируемость и безопасность ЭлД при работе с ним внутри $\boldsymbol{IS}_0$, а также при передаче данных, в том числе и деперсонализированных $\boldsymbol{d}_i'$, к другим ИС, включенным в бизнес-процесс $\boldsymbol{BP}_j$.

В работе [1] достаточно подробно описано, как можно использовать средства ЭП для контроля аутентичности ЭлД, в том числе для обеспечения долговременной сохранности. Для этого должен использоваться алгоритм инвентаризации ЭП, т.е. перезаверение ЭП с извлечением данных об авторе из компонентов (сертификатов открытых ключей) старой ЭП и запись их в новую ЭП. Старая ЭП сохраняется. Таким образом, в результате инвентаризации ЭП возникают цепочки ЭП, проверка которых возможна в течение срока хранения и обработки ЭлД. Преимуществом использования усиленной квалифицированной ЭП, снабженной штампом времени (timestamp), является то, что в работе ИС появляется третье незаинтересованное

лицо – независимый удостоверяющий центр (УЦ), подтверждающий достоверность сертификатов ЭП. Кроме того, появляется возможность с помощью штампа времени через независимые его службы дополнительно удостовериться, что на определенный момент времени ЭлД был аутентичен.

Однако даже защищенный с помощью инвентаризации ЭП ЭлД остается подверженным риску полного бесследного уничтожения или же полной подмены (включая ЭП) лицами, имеющими полные права доступа к ЭлД.

Чтобы предотвратить подобную потерю документа, требуется наличие хорошо защищенного журнала операций, в котором бы отражались не только факт сохранения, подписания, инвентаризации ЭП, но и преобразование данных ЭлД, движение документа к пользователям ИС, в другие ИС или получения ЭлД от других ИС.

Так как структура крупной, территориально-распределенной ИС обычно повторяет географию размещения объектов ее информационной инфраструктуры в независимых, удаленных друг от друга центрах обработки данных (ЦОД), то и движение ЭлД происходит от одного ЦОД к другому в рамках разных ИС одной организации. Тогда и журнал операций становится распределенным реестром, так как операции над ЭлД могут выполняться разными ИС одной организации.

Следовательно, для такого журнала операций могут быть использованы технологии блокчейн, как раз отвечающих условию поддержания достоверности информации в отсутствии доверенного центра. Конечно, некоторые реализации блокчейна (например, биткоин) слишком тяжелы и ресурсоемки для корпоративной ИС. Поэтому лучше для этого случая использовать технологии эксклюзивного или приватного блокчейна, который не требует столь больших энергозатрат на майнинг, так как во-первых, действует в доверенной среде корпоративной ИС, а во-вторых для обеспечения достоверности блоков при небольшом количестве ЦОД ИС не требуются столь сложные криптопреобразования информации при подготовке блоков. Достаточно поддерживать актуальность на всех узлах распределенного реестра блоков и защитить его от изменений, даже со стороны административного персонала ИС. Администраторы ИС не должны менять блокчейн-журнал операций, даже получив контроль 51% майнеров (или средств криптозащиты), не должна быть дана возможность перезаписать блоки.

Так как данные ЭлД могут быть конфиденциальными, то нет смысла хранить сам ЭлД в блоках блокчейна, однако каждая операция с ЭлД должна фиксироваться в блокчейне с помощью создания очередного блока. В самом блоке должен сохраняться хеш ЭлД, возможно компоненты ЭП (сертификаты, списки отзыва сертификатов (СОС), штампы времени). Блок может содержать код (смарт-контракт, или ссылку на исполняемый код), позволяющий проверить ЭП (в том числе и цепочки ЭП, возникающие при их инвентаризации) и ЭлД [1]. Каждый блок подвергается криптопреобразованию с помощью сертифицированных средств криптозащиты информации (СКЗИ). Использование СКЗИ может быть альтернативой энергоемкому майнингу. В частности исполняемый код смарт-контратка блока должен быть доступен для исполнения только определенным лицам при проведении аудита ЭП (например, в процессе инвентаризации ЭП) и не должен быть никем изменен. Также в коде не должно содержаться действий для изменения данных ЭлД. Результаты аудита также должны записываться в блокчейн.

При обмене данными ЭлД различных ИС в рамках цифровой экономики появляется необходимость защитить ЭлД при передаче. Для этих целей также может быть использована технология блокчейн для фиксации фактов передачи ЭлД и подтверждения их аутентичности. Запись о передаче данных (но не сами данные) ЭлД также записываются в цепочки блоков, причем как на стороне ИС приемника информации, так и на стороне ИС источника информации. Состав блоков при этом аналогичен составу блоков приватного блокчейна для внутренних операций ИС.

Таким образом, получается, что кроме внутренних журналов операций, построенных с использованием технологии блокченй, также существует внешний журнал операций, в котором фиксируются все операции передачи данных между ИС.

Примерная схема предлагаемого концептуального решения приводится на рис.1 и 2.

При реализации данной схемы возникают две проблемы:

1. Нужно ли передавать ЭлД полностью, или в ЭлД не должны присутствовать данные, передача которых не допускается в соответствии с $\boldsymbol{T} = \{ \boldsymbol{NTD}_I \}$?
2. Должны ли все ИС, вовлеченные в процессы информационного обмена в рамках цифровой экономики видеть все факты передачи ЭлД от одной отдельной ИС к другой, так как подобная открытость подразумевается технологией блокчейн?

Решением второй проблемы может быть разделение внешних блокчейнов на контуры согласс-

но некоторым разработанным мандатам доступа. Это частично решит проблему, так как позволит разделить передаваемые данные по уровням конфиденциальности. Однако останутся проблемы: а) появление большого количества внешних блокчейнов разного уровня; б) открытость всех фактов передачи ЭлД внутри одного уровня.

Вариантом решения может быть организация хранения в одном блокчейне, при этом информацию в самих узлах цепочек блоков надо защищать мандатными метками или другими методами разграничения доступа.

Решение первой проблемы может быть технология «деперсонализации» данных при передаче ЭлД как внутри одной ИС между пользователями с разными правами доступа, так и между различными ИС.

Функцию «деперсонализации» данных $\boldsymbol{f}$: $\boldsymbol{d_i \rightarrow d_i'}$ должен осуществлять отдельный программный модуль, интегрированный в ИС конкретной организации, сертифицированный для работы со средствами СКЗИ. Входами данного программного модуля должны быть: ЭлД $\boldsymbol{d_i}$, требования защиты данных $\boldsymbol{T} = \{ \boldsymbol{NTD_l} \}$, шаблоны поиска персональных данных в структуре и тексте ЭлД. Для поиска персональных данных применяются интеллектуальный анализ данных ЭлД, анализ текста ЭлД по ключевым словам, полнотекстовый поиск. Кроме того, в структуре метаданных ЭлД должен быть предусмотрен атрибут «персональные данные», которым должны быть отмечены соответствующие данные ЭлД. Данный атрибут должен отмечаться при создании ЭлД, при инвентаризации ЭП ЭлД, при изменении $\boldsymbol{T} = \{ \boldsymbol{NTD_l} \}$, а также периодически.

На выходе модуля получается новый ЭлД $\boldsymbol{d_i'}$ не содержащий персональных данных, заверенный новой ЭП. Заверение ЭП необходимо для дальнейшего контроля аутентичности «деперсонализированной» копии ЭлД $\boldsymbol{d_i'}$, так как деперсонализация – это преобразование исходного документа $\boldsymbol{d_i}$. При передаче между ЭлД $\boldsymbol{d_i'}$ ИС, хеш нового ЭлД $\boldsymbol{d_i'}$ сохраняется во внешнем и приватном блокчейне, возможно со смарт-контрактом для проверки ЭП $\boldsymbol{d_i'}$, для обеспечения возможности аудита. Также в блоках помещаются компоненты ЭП (сертификат открытого ключа, СОС, штамп времени).

Для защиты данных при передаче между ИС необходимы защищенные каналы связи и передачи данных.

Если документ $\boldsymbol{d_i}$ в исходной ИС в процессе обработки подвергается изменению – необходима организация хранения версий ЭлД $\boldsymbol{d_{i(k)}}$. Каждая версия $\boldsymbol{k}$ ЭлД должна храниться как отдельный ЭлД $\boldsymbol{d_{i(k)}}$, также заверенный ЭП, проходящий периодическую инвентаризацию ЭП, связанный с

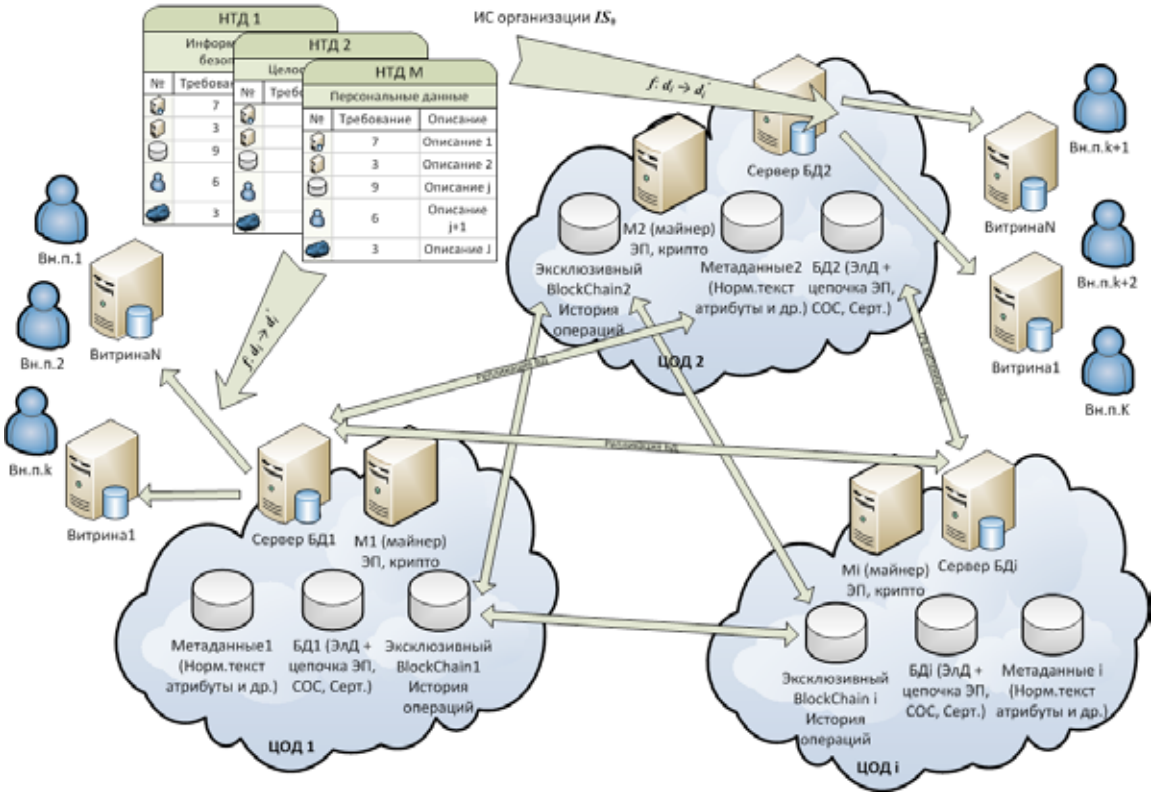

Рис. 1. Схема использования приватного блокчейна

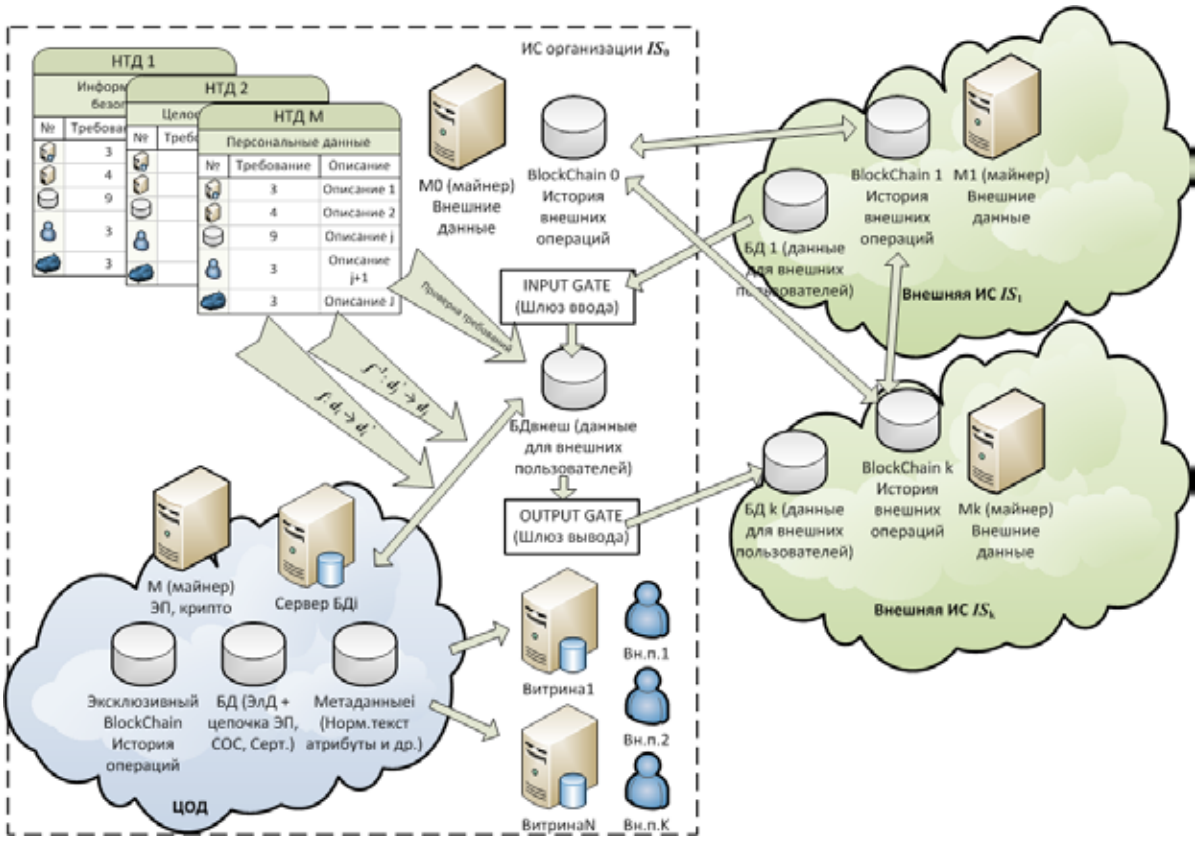

Рис. 2. Схема передачи данных между ИС

исходным ЭлД. Все факты преобразования ЭлД также отмечаются в блокчейн журнале операций.

Редко когда в ИС ЭлД не бывает связан с другими ЭлД. Следовательно, возникает необходимость хранения связей между ЭлД, как в исходной ИС, так и при передаче ЭлД между разными ИС.

Хранение связей можно организовать следующим образом.

1. Неявное хранение связей. При этом факт связывания двух ЭлД представляется как операция над ЭлД. В блоки приватного блокчейна записывается информация о характере связи и хеши двух ЭлД. В этом случае для получения цельной картины связей необходимо будет по блокчейну восстановить граф связей документов в рамках единого процесса. Если во внешние ИС необходимо передать несколько связанных ЭлД, то во внешний блокчейн также должна уходить информация о связях этих ЭлД. Возможным недостатком такого способа является отсутствие единой картины взаимосвязей, возможное длительное время на построение взаимосвязей.
2. Явное хранение связей. В этом случае связи между ЭлД можно представить как отдельный ЭлД, представляющий собой граф связей (по сути граф хешей) нескольких ЭлД. В блокчейне хранится хеш этого отдельного ЭлД. При его изменении создается новый ЭлД (старые ЭлД сохраняются как версии), в приватный блокчейн записывается новый хеш нового ЭлД, описывающего связи. В данном случае на любой момент времени в системе существует граф связей документов. Возможным недостатком данного способа хранения является то, что при передаче во внешнюю ИС нескольких связанных документов, вместе с ними исходная ИС должна будет создать отдельный ЭлД связей этих документов. Его также будет необходимо передать во внешний блокчейн, что потребует дополнительного времени.
3. Хранение связей с помощью смарт-контракта. В этом случае при создании связи между ЭлД в ИС проводится модификация смарт–контракта, связанного с блоком. Поскольку используется объектный подход к описанию сущностей при работе со смарт-контрактами, то связь может быть представлена простой структурой. В этом случае достигается компактность хранения, однако факт изменения смарт-контракта должен также отражаться в блокчейне. Должны храниться версии смарт-контрактов. Также должна быть обеспечена защита аутентичности смарт-контракта и его версий.

4. Хранение связей во внешней БД. Тогда дополнительно к ИС и блокчейну создается отдельная БД, хранящая только ключи-хэши. В этом случае достигается независимость хранения связей ЭлД. Однако создается третья система, которая требует также информационного обмена с ИС и блокчейном.

Описанное концептуальное решение предназначено для решения в общем виде проблемы обеспечения сохранности ЭлД при электронном обмене данными между ИС в условиях цифровой экономики. Тем не менее, для окончательного решения проблемы, особенно при необходимости обеспечить сохранность ЭлД на сроках хранения в десятки лет и более необходимо получить ответ как минимум на следующие вопросы:

1. Как сохранить блокчейн-журнал операций работоспособным при смене программно-аппаратной среды на длительном сроке хранения?
2. Как с помощью блокчейн можно обеспечить подтверждение аутентичности на длительном сроке хранения при включении в блоки смарт-контрактов, элементов ЭП? Ведь на длительном сроке хранения смарт-контракты (блоки кода) могут оказаться неисполняемыми (неинтерпретируемыми), а элементы ЭП – не подлежащими проверке из-за технологического устаревания СКЗИ.
3. Как обеспечить взаимосвязь и хранение блоков при стремительном росте их количества? При миграции блокчейн-журнала операций с одной аппаратной платформы на другую?
4. Как обеспечить интерпретируемость блокчейна в целом и аутентичность блоков в частности? При длительном сроке хранения возможен взлом/подмена блоков из-за роста мощностей компьютеров и устарения средств СКЗИ. Аналогичная проблема для ЭП рассмотрена в работе [1].

5. Возможное практическое применение концептуального решения в условиях цифровой экономики

Рассмотренное концептуальное решение с использованием технологии блокчейн подходит для решения следующих проблем практической организации информационного обмена данными:

- Необходимо организовать обмен данными между внутренними пользователями системы и внешними по отношению к ней пользователями и ИС.
- Требуется гарантия неизменности данных при обмене, так как существует серьезная угроза их несанкционированной подмены в случае мошенничества.
- Централизованное хранение всех данных внутри одной ИС требует внушительных материальных и ресурсных затрат. Например, полное резервное копирование данных требует увеличения существующих мощностей в 3 и более раз, данные плохо систематизируются (нет единого классификатора, описание структуры) или полная синхронизация (обновление) данных из всех источников требует слишком много времени.
- Данные часто меняются или перемещаются между внутренними и внешними пользователями и ИС. Необходимо контролировать и регулировать данный процесс, отлеживать полную цепочку изменений данных.

Перечисленные проблемы характерны для взаимодействия связанных ИС, хранящими и обменивающимися огромным количеством данных. Рост количества данных обуславливается как ростом производительности вычислительной техники, так и все более глубоким охватом различных сфер деятельности и производства информационными технологиями.

Тем самым, авторы считают, что данное решение подойдет для создания распределенных сервисов подтверждения квалификации, проверки качества дипломов о высшем образовании, лицензий на осуществление профессиональной деятельности и сроков их действия. При реализации непосредственно проверки на открытом внешнем блокчейне, эксклюзивные внутренние блокчейны могу стать частью систем внутреннего документооборота (например, при автоматизации процесса приема на работу) или HR-систем при проведении собеседований, учетных системы ВУЗ-ов и центров профессиональной подготовки, учетных систем государственных институтов и ведомств, занимающихся сертификацией и лицензированием разного рода деятельности.

Другим интересным аспектом применения описанного концептуального решения является создание распределенных систем управления знаниями, применяемых в судебной практике, страховании и медицине.

Заключение

В данной статье описана постановка задачи обеспечения сохранности ЭлД при электронном обмене данными между ИС в цифровой экономике. Рассмотрены существующие и перспективные способы организации такого электронного обмена.

Предложено концептуальное решение для обмена деловыми ЭлД, обеспечивающее их аутентичность с помощью технологии блокчейн..

Стоит отметить, что потенциал технологии распределенных реестров высоко оценен на государственном уровне. Именно блокчейн наряду с искусственным интеллектом, большими данными и квантовыми вычислениями вошел в список основных сквозных технологий в рамках национальной технической инициативы, которые определяют ключевые научно-технические направления, оказывающие существенное влияние на развитие рынков в цифровой экономике [21]. Основные усилия исследователей, которые занимаются интеграцией блокчейн и существующих информационных систем, направлены на обеспечение дополнительной безопасности и разработке методов деперсонализации данных при распределенном хранении.

В данной статье авторы предприняли попытку не только рассмотреть недостатки предложенного решения и способы их устранения, но и определили ряд новых актуальных задач, которые ранее не рассматривались. Это задачи обеспечения долговременной сохранности ЭлД и управления связями между ЭлД при использовании блокчейн.

В дальнейшем авторы планируют серию статей, посвященных практическому применению решений поставленной задачи, а также разработку детальных постановок для частных проблем и задач, возникающих при передаче цифровых данных.

Разработка единого комплексного подхода к обеспечению сохранности при электронном обмене данными между ИС в цифровой экономике позволит создавать действительно распределенные программные решения, которые с одной стороны будут обеспечивать надежность и неизменность данных, а с другой не будут требовать таких организационных и финансовых затрат как существующие проекты электронного обмена ЭлД.

Литература

1. *Соловьев А.В.* Решение проблем оценки и сохранения аутентичности электронных документов при долговременном хранении / А.В. Соловьев // Системы высокой доступности. №4. т.10. М.: Радиотехника. 2014. С. 99-106.
2. *Соловьев А.В.* Методология организации долговременного хранения электронных деловых документов / А.В. Соловьев // Труды XXI Международной научно-практической конференции «Документация в информационном обществе: нормативно-методическое обеспечение управления документами» (Москва, РГАСПИ, 18-19 ноября 2014 г.). М.: ВНИИДАД. 2015. С. 320-324.
3. *Соловьев А.В.* Электронные архивы: методологический подход к решению проблемы катастрофоустойчивости при долговременном хранении / Акимова Г.П., Пашкин М.А., Пашкина Е.В., Соловьев А.В., Соловьев Д.В. // Труды ИСА РАН. Том 64. Вып. 3. 2014. С. 91-98.
4. *Соловьев А.В.* Электронные архивы: проблема определения понятия и характеристик электронного документа, как объекта долговременного хранения / А.В. Соловьев, А.С. Богданов // Информационные технологии и вычислительные системы. №4. 2016. С. 24-32.
5. *ГОСТ Р 54989-2012* /ISO TR 18492:2005 Обеспечение долговременной сохранности электронных документов (вступил в силу с 01.05.2013).
6. *Решение Совета глав государств СНГ*: «О Концепции сотрудничества государств-участников Содружества Независимых Государств в сфере обеспечения информационной безопасности и о Комплексном плане мероприятий по реализации Концепции сотрудничества государств-участников Содружества Независимых Государств в сфере обеспечения информационной безопасности на период с 2008 по 2010 год» / [Электронный ресурс] – 2008 – Режим доступа: http://official.academic.ru/6177/%D0%94%D0%BE%D1%81%D1%82%D1%83%D0%BF%D0%BD%D0%BE%D1%81%D1%82%D1%8C_%D0%B8%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%86%D0%B8%D0%B8.
7. *Финансовый словарь «Финам»* [Электронный ресурс] – 2015 – Режим доступа: http://dic.academic.ru/dic.nsf/fin_enc/22465.
8. *ГОСТ Р 52292-2004.* Информационная технология. Электронный обмен информацией. Термины и определения. Утвержден и введен в действие Приказом Федерального агентства по техническому регулированию и метрологии от 29 декабря 2004 г. № 135-ст.
9. *Федеральный закон* Российской Федерации от 6 апреля 2011 г. № 63-ФЗ «Об электронной подписи».
10. *Емельянов С.В.* Системы автоматического управления с переменной структурой. М.: Наука. 1967 336 с.
11. *Емельянов С.В., Костылева Н.Е., Матич Б.Л., Миловидов Н.Н.* Системное проектирование средств автоматизации. М.: Машиностроение. 1978.
12. *Емельянов С.В.* Новые типы обратной связи. М.: Наука. Физматлит. 1997. – 352 с.

13. *Melanie Swan.* Blockchain: Blueprint for a New Economy. – O'Reilly Media, Inc., 2015. – 152 p. – ISBN 978-1-4919-2047-3. В русском переводе *Мелани Свон*. Блокчейн: Схема новой экономики. – Олимп-Бизнес, 2016. – 240 с.
14. *Абелян В.* Что такое смарт-контракты и чем они так хороши? [Электронный ресурс] // Rusbase. Технологии, аналитика, обзор рынков. – 24.08.2017 – Режим доступа: https://rb.ru/opinion/kontrakty-umny-i-horoshi/.
15. *Зеньков А.* Все, что нужно знать об умных контрактах [Электронный ресурс] // Rusbase. Технологии, аналитика, обзор рынков. – 29.09.2017 – Режим доступа: https://rb.ru/story/smart-contract/.
16. *Соловьев А.В.* Электронные архивы: разработка математической модели электронного документа при долговременном хранении / А.В. Соловьев // Информационные технологии и вычислительные системы, №1. 2017. С. 46-61.
17. *Valduries P.* Parallel Database Systems: Open Problems and New Issues. Distributed and Parallel Databases, April 1993, 1(2), pp. 137-165.
18. *Haber S., Stornetta W. S.* (1991). «How to time-stamp a digital document». Journal of Cryptology. 3 (2). doi:10.1007/BF00196791.
19. *Anderson, L., Holz, R., Ponomarev, A., Rimba, P., & Weber, I.* (2016). New kids on the block: an analysis of modern blockchains (2016).
20. *OpenPeppol*. Transport Infrastructure ICT Services-Components. Trust Network Certificate Policy. [Электронный ресурс] Version: 2.00. 07.07.2014. Режим доступа: https://github.com/OpenPEPPOL/documentation/blob/master/TransportInfrastructure/ICT-Transport-Trust_Network_Certificate_Policy-2.00.pdf
21. *Сквозные технологии НТИ*. [Электронный ресурс] – 20.10.2017 – Режим доступа: http://www.nti2035.ru/technology/.

Соловьев Александр Владимирович. Заместитель директора ИСА ФИЦ ИУ РАН по научной работе. Окончил в 1994 г. МФТИ им. Н.Э. Баумана. Доктор технических наук. Количество печатных работ: 71. Область научных интересов: системный анализ, системы управления базами данных, теория надежности, математическое моделирование, электронный документооборот, электронный архив, долговременное хранение электронных документов. E-mail: soloviev@isa.ru

Тарханов Иван Александрович. Старший научный сотрудник ИСА ФИЦ ИУ РАН. Окончил в 2005 г. МФТИ. Кандидат технических наук. Количество печатных работ: 23. Область научных интересов: электронный документооборот, электронный архив, моделирование бизнес процессов, информационная безопасность. E-mail: tarkhanov@isa.ru

Electronic documents and the problem of ensuring security in the exchange of data in the digital economy

A.V. Solovyev, I. A. Tarkhanov

Abstract. The article deals with the problem of ensuring the safety of electronic documents in the electronic exchange of data between information systems in the digital economy. The purpose of this study is to create a theoretical basis for managing the secure information exchange of electronic documents in a dynamically changing software and hardware environment for their storage and transmission.

Keywords: *electronic document, electronic archive, digital economy, blockchain, electronic signature, distributed registers, storage systems, depersonalization of data.*

References

1. *Solovyev A.V.* The problems of assessment and conservation of authenticity of electronic documents for long term storage / A.V. Solovyev // High availability systems. – 2014 – №4, Part.10 – P.99-106.
2. *Solovyev A.V.* The methodology of the organization for long-term storage of electronic business documents / A.V. Solovyev // Proceedings of the XXI International scientific and practical conference "Documentation in information society: regulatory and methodological support document management" (Moscow, 18-19 November 2014) – 2015 – P.320-324.
3. *Solovyev A.V.* Electronic archives: methodological approach to the problem of fault tolerance for long term storage / G.P. Akimova, E.V. Pashkina, M.A. Pashkin, A.V. Solovyev, D.V. Solovyev // Proceedings of Institute of system analysis RAS (ISA RAS) – 2014 – T.64, Part.3 – P.91-98.
4. *Solovyev A.V.* Electronic archives: the problem of definition and characteristics of the electronic

document as an object of long-term storage / A.V. Solovyev, A.S. Bogdanov // Information technology and computer systems. – 2016 – №4 – P.24-32.
5. *GOST R 54989-2012* /ISO TR 18492:2005 Ensuring long-term preservation of electronic documents (entered into force with 01.05.2013).
6. *The decision of the Council of CIS* heads of state: "On the concept of cooperation of States-participants of the Commonwealth of Independent States in the sphere of information security and Complex plan of measures on realization of the concept of cooperation of States-participants of the Commonwealth of Independent States in the sphere of ensuring information security for the period from 2008 to 2010" [Electronic resource] – 2008 – Access mode: http://official.academic.ru/6177/%D0%94%D0%BE%D1%81%D1%82%D1%83%D0%BF%D0%BD%D0%BE%D1%81%D1%82%D1%8C_%D0%B8%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%86%D0%B8%D0%B8
7. *Financial dictionary "Finam"* [Electronic resource] – 2008 – Access mode: http://dic.academic.ru/dic.nsf/fin_enc/22465
8. *GOST R 52292-2004* Information technology. The electronic exchange of information. Terms and definitions. Approved and put into effect by the Federal Agency for technical regulation and Metrology 29 december 2004. № 135-st.
9. *Federal law of the Russian Federation* 6 april 2011. №63-FZ «On the electronic signature».
10. *Emelyanov S.V.* Automatic control systems with variable structure. – M.: Science, 1967 – 336p.
11. *Emelyanov S.V., Kostileva N.E., Matich B.L., Milovidov N.N.* System design automation. – M.: Mechanical engineering, 1978.
12. *Emelyanov S.V.* New types of feedback. – M.: Science, Fizmatlit, 1997 – 352p.
13. *Melanie Swan.* Blockchain: Blueprint for a New Economy. – O'Reilly Media, Inc., 2015. – 152 p. – ISBN 978-1-4919-2047-3.
14. *Abelyan V.* What are smart contracts and what are they so good at? [Electronic resource] // Rusbase. Technology, analytics, market review. – 24.08.2017 – Access mode: https://rb.ru/opinion/kontrakty-umny-i-horoshi/.
15. *Zenkhov A.* All you need to know about smart contracts [Electronic resource] // Rusbase. Technology, analytics, market review. – 29.09.2017 – Access mode: https://rb.ru/story/smart-contract/.
16. *Solovyev, A.V.* Electronic archives: development of a mathematical model of an electronic document for long-term storage / A.V. Solovyev // Information technology and computer systems, №1, 2017, P.46-61.
17. *P. Valduries.* Parallel Database Systems: Open Problems and New Issues. Distributed and Parallel Databases, April 1993, 1(2), pp. 137-165.
18. *Haber S.; Stornetta W. S.* (1991). «How to time-stamp a digital document». Journal of Cryptology. 3 (2). doi:10.1007/BF00196791
19. *Anderson L., Holz R., Ponomarev A., Rimba P., & Weber I.* (2016). New kids on the block: an analysis of modern blockchains (2016).
20. *OpenPeppol.* Transport Infrastructure ICT Services-Components. Trust Network Certificate Policy. [Electronic resource] Version: 2.00. 07.07.2014. Access mode: https://github.com/OpenPEPPOL/documentation/blob/master/TransportInfrastructure/ICT-Transport-Trust_Network_Certificate_Policy-2.00.pdf
21. *Cross-cutting technologies of NTI.* [Electronic resource] – 20.10.2017 – Access mode: http://www.nti2035.ru/technology/.

Solovyev Alexandr Vladimirovich, Deputy Director ISA FRC CSC RAS. Moscow, prosp. 60-let Oktyabrya, 9. BMSTU 1994. Number of publications: 71. Area of scientific interests: system analysis, database management systems, reliability theory, mathematical modeling, electronic document management, electronic archive, long-term storage of electronic documents. E-mail: soloviev@isa.ru

Tarkhanov Ivan Alexandrovich, Senior Researcher of ISA FRC CSC RAS. Moscow, prosp. 60-let Oktyabrya, 9. MIPT 2005. Number of publications: 23. Area of scientific interests: electronic document management, electronic archive, business process modeling, information security. E-mail: tarkhanov@isa.ru.