Майя САВИЦКАЯ
Иван РОЩУПКИН

Процедуры управления качеством данных во многих банках строятся с нуля и вызывают методологические и технические затруднения, так как ранее подобных мероприятий банки в большинстве своем не проводили. А специалистов, умеющих оценивать качество данных в информационных системах, немного. Рассказываем, как «подступиться» к такой оценке, каким может быть алгоритм действий, как оценивать данные из внешних информационных систем, какие стандарты пригодятся для оценки.

Майя САВИЦКАЯ, *ООО «ФБК», менеджер Департамента аудиторских и консультационных услуг финансовым институтам*
Иван РОЩУПКИН, *ООО «ФБК», старший эксперт по информационной безопасности*

Как оценить качество данных в информационных системах по Положению № 716-П и зачем это нужно

Положением № 716-П[1] Банк России впервые ввел для кредитных организаций объемные требования к управлению риском информационных систем (далее — ИС) и к организации управления качеством данных в ИС. Требования и принципы управления данными в ИС и правила оценки их качества регулятор определил в главе 8 «Управление риском информационных систем» указанного нормативного акта.

Обеспечение качества данных в ИС — это важнейший элемент в управлении как риском ИС в частности, так и операционным риском в целом, оказывающий прямое влияние на обеспечение бесперебойной работы банковских процессов и операционную устойчивость кредитной организации.

Источники некачественных данных перечислены в табл. 1.

Ошибки могут появиться в данных на любом этапе сбора информации. В данных всегда больше ошибок, чем кажется. И главное —

[1] Положение Банка России от 08.04.2020 № 716-П «О требованиях к системе управления операционным риском в кредитной организации и банковской группе».

Как оценить качество данных в информационных системах по Положению № 716-П и зачем это нужно

Таблица 1

Источники некачественных данных

№	Источник	Описание
1	Ошибка персонала	Любые не соответствующие установленным регламентам или сложившимся практикам действия персонала, совершаемые без злого умысла по причине недостаточно четкого определения обязанностей, недостаточного обучения или квалификации персонала. Возникновению ошибок способствуют отсутствие дисциплинарного процесса и документирования процессов, отсутствие методов контроля, предоставление избыточных полномочий и др.
2	Социальный инжиниринг	Умышленные действия сторонних лиц, преследующих мошеннические цели, реализуемые посредством обмана, введения в заблуждение персонала и приводящие к ошибкам работников, несанкционированным изменениям и утрате информационных активов, нарушению конфиденциальности данных
3	Несанкционированный логический доступ	Несанкционированный логический доступ неавторизованных субъектов к информационным активам (компрометация пароля, предоставление пользователям/администраторам избыточных прав доступа, недостатки (отсутствие) механизмов аутентификации пользователей и администраторов, ошибки администрирования, оставление без присмотра программно-технических средств, вредоносные программы), что может привести к нарушению свойств информационных активов, сбоям, отказам, несанкционированным изменениям и уничтожению программных средств и информации
4	Несанкционированный физический доступ	Физический несанкционированный доступ неавторизованных лиц в контролируемую зону расположения технических средств и (или) информационных активов, что может привести к разрушению и уничтожению технических и программных средств, нарушению конфиденциальности, целостности, доступности информационных активов, нарушению непрерывности процессов
5	Выполнение вредоносных программ	Внедрение в систему и выполнение вредоносных программ вследствие халатности, низкой квалификации персонала (пользователей), уязвимостей программных средств. Возможные последствия: несанкционированный доступ к информационным активам, нарушение их свойств, сбои, отказы и уничтожение программных средств, информации
6	Использование программных средств и информации без гарантии источника	Использование в ИС организации непроверенных данных или нелицензионного программного обеспечения
7	Нарушение договорных обязательств сторонними (третьими) лицами	Невыполнение со стороны третьих лиц взятых на себя обязательств по качеству, составу, содержанию и (или) порядку оказания услуг, поставки программно-технических средств и т.д.
8	Ошибки в обеспечении безопасности ИС на стадиях жизненного цикла	Ошибки в обеспечении безопасности при разработке, эксплуатации, сопровождении и выводе из эксплуатации ИС
9	Сбои и отказы программно-технических средств	Нарушение работоспособности программно-технических средств, технические сбои, в том числе неполная загрузка данных, некорректная логика преобразований данных вследствие некорректного изменения параметров или свойств программных средств под влиянием внутренних процессов (ошибок) и (или) внешних воздействий

Майя САВИЦКАЯ
Иван РОЩУПКИН

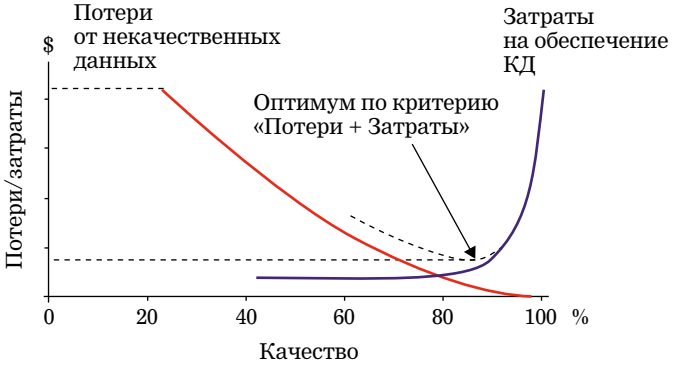
Окончание табл. 1

№	Источник	Описание
10	Нарушения функциональности криптографической системы	Случайное или намеренное неправильное управление криптографическими ключами, криптографическими протоколами и алгоритмами, программно-аппаратными средствами систем криптографической защиты информации, приводящее к потере конфиденциальности, целостности и доступности данных, нарушению бесперебойности приема-передачи информации, блокировке функционирования ИС
11	Нарушения функциональности архивной системы	Нарушение конфиденциальности и целостности архивных данных и (или) непредоставление услуг архивной системой (нарушение доступности) вследствие случайных ошибок пользователей или неправильного управления архивной системой, а также вследствие физических воздействий на компоненты архивной системы

достижение определенного качества данных и абсолютного значения не всегда экономически целесообразно (см. рисунок).

Рисунок

Оптимальное соотношение потерь и затрат на обеспечение качества данных



Способы и механизмы реализации требований Банка России к управлению риском ИС, в том числе контролю качества данных, кредитные организации разрабатывают самостоятельно и закрепляют во внутренних нормативных документах, утверждаемых их органами управления.

Банк России в абз. 6 п. 8.7.6 Положения № 716-П установил требование проводить не реже одного раза в год независимую оценку качества данных в ИС, обеспечивающих критически важные процессы. Такую оценку банки могут проводить как самостоятельно, так и с привлечением внешнего эксперта. Если проводится внутренняя оценка — то, чтобы она была действительно независимая,

Как оценить качество данных в информационных системах по Положению № 716-П и зачем это нужно

необходимо избегать конфликта интересов: проводить такую оценку не могут работники и (или) подразделения, обеспечивающие качество данных.

Обеспечение проведения оценки качества данных в ИС, обеспечивающих критически важные процессы, — обязанность должностного лица кредитной организации, несущего персональную ответственность за качество данных в ИС.

Оценка качества данных проводится по разработанной самостоятельно «Методике обеспечения качества данных в информационных системах, обеспечивающих критически важные процессы», в которой должны быть регламентированы в том числе:

— требования к качеству данных в ИС;

— классификатор возможных источников и причин образования в ИС данных, не соответствующих требованиям;

— показатели качества данных для оценки характеристик качества данных и их предельно допустимые значения;

— методы и алгоритмы расчета, правила измерения показателей качества данных;

— критерии оценки качества данных;

— порядок формирования отчетов о качестве данных.

В своей практике мы видим, что у кредитных организаций возникают следующие вопросы при разработке методологии оценки:

— Какие данные и в каких процессах оценивать?

— Как определить владельцев данных?

— Как отобрать критические данные, оказывающие влияние на процесс? Или нужно работать со всеми данными?

— Какие ИС включать в оценку?

— Какие параметры оценки использовать и как установить их целевые значения?

Если проводится внутренняя оценка — то, чтобы она была действительно независимая, необходимо избегать конфликта интересов: проводить такую оценку не могут работники и (или) подразделения, обеспечивающие качество данных.

Как «подступиться» к процессу оценки качества данных в ИС

Для начала выстраивания методики оценки предлагаем использовать следующий примерный алгоритм действий.

Очевидно, что для того, чтобы чем-либо управлять и что-либо оценивать, надо сначала идентифицировать объект управления и определить критерии оценки, позволяющие делать однозначные качественные выводы о таком объекте.

Итак, давайте сначала *определим объект оценки «качество данных в информационных системах, обеспечивающих критически важные процессы»*. Объект состоит из четырех компонентов:

Майя САВИЦКАЯ
Иван РОЩУПКИН

1. Критически важные процессы.
2. Информационные системы.
3. Данные в ИС.
4. Качество данных в ИС.

Компонент 1 — «Критически важные процессы»

Данный элемент определяется достаточно просто: список таких процессов 1-го уровня определен регулятором в абз. 3 п. 4.1.1 Положения № 716-П.

Это процессы, которые обеспечивают:

1) выполнение операций, указанных в п. 1–4 и 9 ч. 1 ст. 5 Закона № 395-1[1]:

— привлечение денежных средств физических и юридических лиц во вклады (до востребования и на определенный срок);

— размещение указанных в предыдущем пункте привлеченных средств от своего имени и за свой счет;

— открытие и ведение банковских счетов физических и юридических лиц;

— осуществление переводов денежных средств по поручению физических и юридических лиц, в том числе банков-корреспондентов, по их банковским счетам;

— осуществление переводов денежных средств без открытия банковских счетов, в том числе электронных денежных средств (за исключением почтовых переводов);

2) ведение бухгалтерского учета;

3) представление отчетности в Банк России;

4) поддержание ликвидности;

5) выполнение операций на финансовых рынках;

6) выполнение кассовых операций;

7) работу онлайн-сервисов дистанционного обслуживания и доступ к осуществлению операций;

8) соблюдение Закона № 152-ФЗ[2];

9) соблюдение Трудового кодекса РФ;

10) соблюдение Закона № 395-1;

11) другие процессы, которые определены кредитной организацией самостоятельно и прерывание функционирования которых оказывает влияние на выполнение обязательств перед клиентами и контрагентами.

Список критически важных процессов 1-го уровня определен в абз. 3 п. 4.1.1 Положения № 716-П. Из них выделяются те, которые образуют структуру технологических процессов, требующих обеспечения информационного взаимодействия, обработки и хранения информации с помощью информационных систем (Приложение 1 к Положению № 787-П).

[1] Федеральный закон от 02.12.1990 № 395-1 «О банках и банковской деятельности».

[2] Федеральный закон от 27.07.2006 № 152-ФЗ «О персональных данных».

Как оценить качество данных в информационных системах по Положению № 716-П и зачем это нужно

Далее для определения процессов, которые важны для бесперебойного обслуживания клиентов и обеспечения операционной надежности банка, целесообразно выделить из этих процессов те, которые образуют структуру технологических процессов банка, требующих обеспечения информационного взаимодействия, обработки и хранения информации с помощью ИС. Список таких технологических процессов закрытый и определен в Приложении 1 к Положению № 787-П[1]. Технологических процессов всего 13 — это процессы, которые обеспечивают:

1) привлечение денежных средств физических лиц во вклады;

2) привлечение денежных средств юридических лиц во вклады;

3) размещение привлеченных во вклады денежных средств физических и (или) юридических лиц от своего имени и за свой счет;

4) осуществление переводов денежных средств по поручению физических лиц по их банковским счетам;

5) осуществление переводов денежных средств по поручению юридических лиц, в том числе банков-корреспондентов, по их банковским счетам, за исключением переводов по распоряжениям участников платежной системы (для переводов денежных средств по распоряжениям участников платежной системы — в соответствии с Положением № 607-П[2]);

6) открытие и ведение банковских счетов физических лиц;

7) открытие и ведение банковских счетов юридических лиц;

8) осуществление переводов денежных средств без открытия банковских счетов, в том числе электронных денежных средств (за исключением почтовых переводов);

9) выполнение операций на финансовых рынках;

10) выполнение кассовых операций;

11) работу онлайн-сервисов дистанционного обслуживания и доступа к осуществлению операций;

12) размещение и обновление биометрических персональных данных в единой биометрической системе;

13) идентификацию и (или) аутентификацию с использованием биометрических персональных данных физических лиц, в том числе с применением информационных технологий без их личного присутствия.

[1] Положение Банка России от 12.01.2022 № 787-П «Об обязательных для кредитных организаций требованиях к операционной надежности при осуществлении банковской деятельности в целях обеспечения непрерывности оказания банковских услуг».

[2] Положение Банка России от 03.10.2017 № 607-П «О требованиях к порядку обеспечения бесперебойности функционирования платежной системы, показателям бесперебойности функционирования платежной системы и методикам анализа рисков в платежной системе, включая профили рисков».

Майя САВИЦКАЯ
Иван РОЩУПКИН

По каждому процессу необходимо определить подразделение, ответственное за осуществление операций и сделок и за результаты процесса (далее — ответственное подразделение).

Перечень «Технологический процесс/критичный процесс/ответственное подразделение» является элементом системы управления операционным риском и закрепляется во внутренних документах.

Такой перечень можно представить, например, в виде матриц (табл. 2 и 3).

Таблица 2

Матрица 1 «Критичные процессы в рамках технологических процессов»

Технологический процесс (Приложение к Положению № 787-П)	**Критичные процессы (п. 4.1.1 Положения № 716-П)**			
	привлечение денежных средств физических и юридических лиц во вклады (до востребования и на определенный срок)	**размещение указанных в предыдущем пункте привлеченных средств от своего имени и за свой счет**	**открытие и ведение банковских счетов физических и юридических лиц**	**осуществление переводов денежных средств по поручению физических и юридических лиц, в том числе банков-корреспондентов, по их банковским счетам**
	Ответственное подразделение			
Технологический процесс, обеспечивающий привлечение денежных средств физических лиц во вклады	Дополнительный офис			
Технологический процесс, обеспечивающий привлечение денежных средств юридических лиц во вклады	Управление депозитных операций			
Технологический процесс, обеспечивающий размещение привлеченных во вклады денежных средств физических и (или) юридических лиц от своего имени и за свой счет		Кредитный департамент		
Технологический процесс, обеспечивающий осуществление переводов денежных средств по поручению физических лиц по их банковским счетам				Отдел расчетов
Технологический процесс, обеспечивающий осуществление переводов денежных средств по поручению юридических лиц, в том числе банков-корреспондентов, по их банковским счетам, за исключением переводов по распоряжениям участников платежной системы (для переводов денежных средств по распоряжениям участников платежной системы — в соответствии с Положением № 607-П				Отдел расчетов

Как оценить качество данных в информационных системах по Положению № 716-П и зачем это нужно

Таблица 3

Матрица 2 «Иные критичные процессы»

Критичные процессы (п. 4.1.1 Положения № 716-П)	Ответственное подразделение
Ведение бухгалтерского учета	Управление бухгалтерского учета
Представление отчетности в Банк России	Отдел формирования отчетности
Поддержание ликвидности	Казначейство
Соблюдение Трудового кодекса РФ	Управление по работе с персоналом

Компонент 2 — «Информационные системы»

При отборе ИС для оценки качества данных важно понимать, что оценке подлежат не все используемые кредитной организацией ИС, а только те, которые обеспечивают осуществление критически важных процессов.

Поэтому чтобы понять, какие же ИС отобрать для оценки, целесообразно построить перечень «Критичный процесс/информационная система».

И в проведении такого отбора поможет описанная в соответствии с п. 8.5 Положения № 716-П во внутренних документах архитектура ИС, которая как раз и будет содержать взаимосвязи ИС с обслуживаемыми ею критичными процессами.

По итогам анализа и отбора информации по элементам 1 и 2 получается матрица (табл. 4).

Таблица 4

Матрица «Критичный процесс/ответственное подразделение/информационные системы»

Критичный процесс	Ответственное подразделение	Информационные системы		
		ИС 1 — «Диасофт»	ИС 2 — Retail-bank	ИС 3 — «1С Персонал»
Открытие и ведение банковских счетов физических и юридических лиц	Дополнительный офис	X	X	
	Операционное управление	X		
Соблюдение Трудового кодекса РФ	Управление по работе с персоналом			X

Майя САВИЦКАЯ
Иван РОЩУПКИН

Компонент 3 — «Данные в информационных системах»

Положение № 716-П не дает толкования понятия «данные в информационных системах». Поэтому обратимся к стандартам информационных технологий, устанавливающим это понятие. Итак, данные в информационных системах — это «информация, представленная в формализованном виде, пригодном для передачи, интерпретации или обработки людьми или компьютерами»[1] и «формы представления информации, с которыми имеют дело информационные системы и их пользователи»[2].

Хорошие определения для конструктивного подхода к решению нашей задачи. Остановимся на понятии «формализованный вид»: то есть, чтобы информация стала данными, надо ее формализовать. При формализации она приобретает свойства, характеристики, атрибуты, дающие возможность выявить и зафиксировать существенные стороны объектов, — то есть становится данными.

Важно определить, какие из данных, хранящихся в информационной системе, существенны для корректной реализации технологического процесса и (или) критичного для кредитной организации процесса, а какие не окажут влияния на их осуществление.

Таким образом, данные, хранящиеся в ИС, могут быть использованы как людьми, так и другими ИС и должны быть представлены в виде, необходимом для дальнейшего использования информации: обработки, хранения и принятия решений. Особенно если данные идут по цепочке технологического процесса, в котором задействовано несколько ИС, и данные в каждом звене такой цепочки влияют на формирование и вывод данных в следующем звене.

Очевидно, что степень влияния каждого вида данных и информации на возможность реализовать процесс, обслуживаемый ИС, в установленный срок, без искажения данных, информации и цели такого процесса разная.

То есть важно определить, какие из данных, хранящихся в ИС, существенны для корректной реализации технологического процесса и (или) критичного для кредитной организации процесса, а какие не окажут влияния на их осуществление. Например, атрибут карточки клиента «пол» при его неверном указании не окажет никакого влияния на итог функционирования процесса «обслуживание текущих счетов физических лиц», а вот искажение атрибута «дата рождения» может повлиять на корректность идентификации клиента и затормозить процесс управления денежными средствами.

[1] ГОСТ 33707-2016 (ISO/IEC 2382:2015) «Информационные технологии. Словарь». ГОСТ 34.321-96 Межгосударственный стандарт «Информационные технологии. Система стандартов по базам данных. Эталонная модель управления данными».

[2] ГОСТ Р ИСО/МЭК 10746-2-2000 «Информационная технология. Взаимосвязь открытых систем. Управление данными и открытая распределенная обработка».

Как оценить качество данных в информационных системах по Положению № 716-П и зачем это нужно

Без знаний особенности функционирования бизнес-процесса, его цели, информации, утеря или искажение которой может способствовать неверной или неточной реализации процесса, сбою в его функционировании или принятию неверного решения, определить это невозможно.

И как раз выводы по компонентам 1 и 2 позволяют установить владельцев такой информации — ответственные подразделения.

Отдельно хотим обратить ваше внимание на ответственность за данные. Процесс поддержания качества данных схож с бюджетным управлением — это сквозной процесс в рамках операционных процессов организации. Должен быть сформирован институт владельцев данных (Data Stuard). Их задача — обеспечивать процессы управления качеством в зоне своей ответственности. Важно, что зона ответственности определяется только природой данных и не зависит от ИС, где эти данные используются.

Компонент 4 — «Качество данных в информационных системах»

Согласно ГОСТу[1], качество данных — степень, с которой набор характеристик, присущих данным, отвечает требованиям, то есть потребностям или ожиданиям, которые установлены, предполагаются или являются обязательными для пользователя данных — человека или ИС. Некачественными данными признаются данные, характеристики которых не соответствуют и (или) не достигают установленных для них требований, что делает невозможным для пользователей данных полагаться на них, тогда как данные с высоким качеством — это надежная база для любого процесса и операции, способствующая «правильной» интерпретации информации и данных и их целевому использованию. Итак, объект оценки понятен: мы определили критичные для банка процессы, информационные системы, их обслуживающие, и данные, которые существенны для процесса с использованием этих ИС. Можно приступать к разработке показателей для измерения качества данных.

Измеряем качество данных

Как мы выяснили ранее, оценка качества данных производится путем сравнения их характеристик с целевыми (допустимыми) значениями показателей, при недостижении которых данные считаются некачественными.

[1] ГОСТ Р ИСО 8000-2-2019 «Качество данных. Словарь».

Майя САВИЦКАЯ
Иван РОЩУПКИН

Положение № 716-П определяет семь базовых характеристик качества данных:

1) точность и достоверность;
2) полнота;
3) актуальность;
4) согласованность;
5) доступность;
6) контролируемость;
7) восстанавливаемость.

Точность и достоверность

Это отсутствие синтаксических и семантических ошибок в данных, а также их соответствие реальным и статистически наиболее вероятным значениям свойств, характеристик и параметров, зафиксированных в данных.

Данные признаются качественными с точки зрения точности и достоверности, если соответствие эталонным значениям превышает 98%, а доля некорректных записей, выявленных в исследуемой совокупности данных, менее 2%.

Для оценки этой характеристики могут использоваться частные показатели:

— корректность записей (доля значений, не соответствующих корректным, выявленных по результатам обработки инцидентов, связанных с качеством данных);

— соответствие эталонным данным/первоисточникам.

Например, эталонное значение даты рождения клиента банка — «дд.мм.гггг», причем элемент «дд» должен находиться в диапазоне от 1 до 31, «мм» — от 1 до 12, «гггг» — в диапазоне «19гг – текущий год». При искажении эталонного значения и записи даты рождения как 36.13.1863, 1/11/2026 или 23.02.1 этот атрибут карточки клиента не будет «считываться» пользователем и будет исключен из массива отбора информации по его запросу.

О качественном состоянии данных по этой характеристике обычно можно сделать вывод, если соответствие эталонным значениям превышает 98%, а доля некорректных записей, выявленных в исследуемой совокупности данных, менее 2%.

Полнота

Это достаточность данных для функционирования банковских процессов. Для оценки этой характеристики могут использоваться частные показатели:

— доля незаполненных полей (атрибутов), обязательных к заполнению;

— доля объектов учета, данные о которых не включены в ИС, в общей массе объектов учета;

— доля записей с избыточными (например, полученными в результате дублирования) значениями.

Например, ИС позволяет записать карточку кредита без указания даты начисления процентов. При запуске автоматической утилиты для расчета и формирования проводок по начислению процентов на остатки по ссудному счету такой кредит не будет отобран системой для обработки, процент за пользование ссудой клиенту не начислится и банк недополучит запланированный доход.

О качественном состоянии данных по этой характеристике обычно можно сделать вывод, если доля учтенных в ИС данных составляет не менее 98%, а доля незаполненных или дублирующих атрибутов не более 2%.

Актуальность

Это способность данных адекватно отражать состояние объектов предметной области в текущий момент. Для оценки этой характеристики могут использоваться частные показатели:

— доля актуальных данных в составе ИС (массива данных);

— коэффициент готовности данных (отношение времени пребывания в актуальном/нормативном состоянии к времени пребывания в неактуальном или неопределенном состоянии);

— коэффициент длительности обработки, проверки ошибок, согласования и внесения данных в информационной системе.

Например, справочники об официальных курсах иностранных валют подгружаются в АБС нерегулярно или с опозданием, соответственно при волатильности курса банк может получить незапланированный убыток.

О качественном состоянии данных по этой характеристике обычно можно сделать вывод, если коэффициент готовности данных не менее 98%, а доля актуальных данных не менее 95%.

Согласованность

Это взаимная непротиворечивость данных, хранящихся в ИС, других источниках и носителях информации. Для оценки могут использоваться частные показатели:

— доля использования альтернативных обозначений или сокращений для сущностей, снабженных стандартизованным обозначением;

— доля нестандартных наименований объектов учета.

Например, указание номера кредитного договора в различных источниках разными способами — в карточке кредита в модуле

Майя САВИЦКАЯ
Иван РОЩУПКИН

«Кредиты» АБС 1/2023, в бумажном договоре с клиентом 1/23, в наименовании счета бухгалтерского учета в учетной АБС 1.2023 — может привести к неидентификации трех источников как относящихся к одному кредиту и искажению как аналитического, так и синтетического бухгалтерского учета.

Согласованность данных особенно важна, если данные передаются по цепочке информационных систем, «вливаясь» из одной в другую. Так, довольно часто аналитический учет по розничным операциям, операциям с пластиковыми картами или биржевым операциям происходит в отдельных информационных системах, а общая сумма попадает в синтетический учет главной учетной АБС.

О качественном состоянии данных по этой характеристике обычно можно сделать вывод, если доля соответствующих ошибок меньше 1%.

Согласованность данных особенно важна, если данные передаются по цепочке информационных систем, «вливаясь» из одной в другую. О качественном состоянии данных по характеристике согласованности можно сделать вывод, если доля соответствующих ошибок меньше 1%.

Доступность

Это возможность использования данных при функционировании процессов. Для оценки могут использоваться частные показатели:

— время доступности;
— время простоя ИС.

О качественном состоянии данных по этой характеристике обычно можно сделать вывод, если время простоя не превышает допустимого порогового значения и не происходит деградации технологического процесса. Для формирования вывода о доступности данных в ИС целесообразно использовать информацию о фактах реализации риска ИС, отраженных в базе событий операционного риска. Обычно недопустимой считается ситуация, когда время доступности данных в ИС менее 98% от запланированного времени функционирования процесса, обслуживающегося этой ИС.

Контролируемость

Это возможность контролировать качество и происхождение данных. Оценка этой характеристики может зависеть от наличия в ИС функционала отражения:

— источников данных;
— истории создания;
— истории изменения;
— истории преобразования;
— истории удаления;
— истории хранения;
— истории передачи данных.

Как оценить качество данных в информационных системах по Положению № 716-П и зачем это нужно

О качественном состоянии данных по этой характеристике обычно можно сделать вывод, изучив ИС на предмет наличия журналов изменений и их атрибутов, наличия логов и возможности получить оперативный доступ к этой информации.

Восстанавливаемость

Это возможность сохранять установленный уровень функциональности и качества данных после их утраты, повреждения или изменения в результате сбоев или других нарушений функционирования ИС. Для оценки этой характеристики может использоваться показатель RPO (recovery point objective) — целевая точка восстановления, который определяется допускаемым уровнем потери данных в случае прерывания операций.

О качественном состоянии данных по этой характеристике обычно можно сделать вывод, исходя из соблюдения RPO. Также для формирования вывода о «работоспособности» процесса восстановления данных желательно использовать информацию о тестировании плана обеспечения непрерывности и восстановления деятельности кредитной организации.

Подходы к оценке данных из внешних информационных систем

При оценке важно не забывать об информационных системах третьих лиц (подрядчиков, контрагентов, участников банковской группы), обеспечивающих критически важные процессы банка и структуры информационного обмена между их элементами и элементами его собственных ИС. Особенно это актуально при оценке таких характеристик качества данных, как согласованность и доступность.

При отборе данных для оценки важно определить, какие данные банк генерирует самостоятельно, а какие заимствует из ИС третьих лиц, так как эти ИС банку не принадлежат и не контролируются им, соответственно он не может оценить качество данных в этих системах.

Для этого надо ввести понятие потока данных (именно в рамках Положения № 716-П; не путаем с определением потока данных для программирования). Поток данных — это передача и принятие той информации, которая нам нужна для обеспечения бизнес-процессов.

В ходе анализа потока данных необходимо принять стратегическое решение, какие данные, полученные из других ИС, мы считаем эталонными и требующими сверки данных из ИС банка. То есть мы пропускаем через все наши системы внешние данные как эталонные и по итогам сравнения будем вносить изменения в наши данные.

При отборе данных для оценки важно определить, какие данные банк генерирует самостоятельно, а какие заимствует из ИС третьих лиц. Для этого надо ввести понятие потока данных — это передача и принятие той информации, которая нам нужна для обеспечения бизнес-процессов.

Майя САВИЦКАЯ
Иван РОЩУПКИН

Или мы считаем, что наши данные эталонные. Но это спорный подход, ведь больше данных извне мы получаем из систем, которые создавались для определенной цели — собирать и агрегировать определенные данные и передавать нам.

Целесообразно ввести допустимый процент некачественных данных в получаемых извне эталонных данных, например 1–2% в рамках оцениваемых характеристик данных. Поскольку открытой информации о качестве данных из внешних систем у нас нет, такой подход позволит минимизировать наши риски и обращаться к поставщикам внешних ИС с вопросом о качестве поставляемых ими данных. Также целесообразно отнести используемые банком ИС третьих лиц к возможным источникам и причинам образования некачественных данных.

Какие стандарты пригодятся при оценке

Разрабатывая методику оценки качества данных в ИС, целесообразно использовать нормы и принципы стандартов, названных в табл. 5.

Таблица 5

Стандарты для оценки качества данных в информационных системах

№	Стандарт	Описание
1	ISO/TS 8000 «Качество данных» (ГОСТ Р 56214-2014/ISO/TS 8000-1:2011)	Стандарты комплекса ISO 8000 определяют параметры характеристик, которые могут быть проверены любой организацией в цепочке передачи данных для определения соответствия этой информации требованиям ISO 8000. Стандарты обеспечивают совершенствование качества информации, используемой как самостоятельно, так и в рамках систем управления качеством. В стандартах комплекса ISO 8000 представлены технические характеристики качества данных, применяемых на протяжении всего жизненного цикла продукции, и рассматриваются различные виды данных, включая основные данные, данные транзакций и данные о продукции
2	ISO/IEC 25012:2008 «Разработка программного обеспечения — Требования и оценка качества программного продукта (SQuaRE) — Модель качества данных»	ISO/IEC 25012:2008 определяет общую модель качества данных для данных, хранящихся в структурированном формате в компьютерной системе. Разделы: — «Менеджмент качества» (2500n); — «Модель качества» (2501n); — «Измерение качества» (2502n); — «Требования к качеству» (2503n); — «Оценка качества» (2504n)

Как оценить качество данных в информационных системах по Положению № 716-П и зачем это нужно

Очевидно, что оценка качества данных в ИС полезна и для других целей кроме комплаенса. И если говорить о выгодах бизнеса от такой оценки, то уверенность менеджмента в качестве данных в ИС, используемых для осуществления критичных бизнес-процессов, служит своего рода гарантией того, что при принятии управленческих решений и формировании финансовых данных банк защищен от риска искажения, неполноты, утраты или манипулирования данными. Базируемые на качественных данных стратегии и модели развития банка будут более достоверными и актуальными, а бизнес-планы — более точными.