НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГЕОГРАФИЧЕСКАЯ ИНФОРМАЦИЯ

Принципы оценки качества

Издание официальное

БЗ 6—2003/89

ГОССТАНДАРТ РОССИИ
Москва

Предисловие

1 РАЗРАБОТАН Государственным научно-внедренческим центром геоинформационных систем и технологий (Госгисцентр)

ВНЕСЕН Техническим комитетом по стандартизации ТК 394 «Географическая информация/геоматика»

2 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Постановлением Госстандарта России от 9 декабря 2003 г. № 359-ст

3 Настоящий стандарт идентичен международному стандарту ИСО 19113—2002 «Географическая информация — Принципы оценки качества», подготовленного Техническим комитетом ISO/TC 211 «Geographic information/Geomatics» (ИСО ТК/211 «Географическая информация/Геоматика»)

4 ВВЕДЕН ВПЕРВЫЕ

© ИПК Издательство стандартов, 2004

Настоящий стандарт не может быть полностью или частично воспроизведен, тиражирован и распространен в качестве официального издания без разрешения Госстандарта России

Содержание

1 Область применения 1
2 Соответствие 1
3 Нормативные ссылки 1
4 Определения 2
5 Принципы описания качества географической информации 3
6 Определение качества географической информации 6
7 Выдача информации о качестве 8
Приложение А Комплекс проверок набора данных 9
Приложение Б Концепция оценки качества данных и ее использование 11
Приложение В Примеры определения применимых элементов и подэлементов качества данных . 15

Введение

Наборы географических данных используются для различных целей.

При выборе наборов географических данных разработчики и пользователи руководствуются, прежде всего, их качеством. При этом возникают ситуации, требующие различных уровней качества таких данных. Некоторым пользователям для конкретных целей необходимы очень точные данные, в то время как другим пользователям достаточно иметь менее точные данные. Поэтому информация о качестве географических данных становится решающим фактором их использования. С развитием технологии появляются возможности сбора и использования наборов географических данных, качество которых может превосходить уровень, изначально задаваемый пользователями.

Описание качества географических данных облегчает выбор набора данных, лучше всего подходящего для конкретной цели. Информация о качестве географических данных позволяет разработчику или поставщику таких данных обосновать, насколько тот или иной набор данных отвечает критериям, определенным в спецификации его продукта, а пользователю данных — оценить способность продукта удовлетворять требованиям его конкретного приложения.

Главная задача настоящего стандарта — установить принципы описания качества географических данных и концепции для обработки информации об их качестве.

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГЕОГРАФИЧЕСКАЯ ИНФОРМАЦИЯ

Принципы оценки качества

Geographic information. Quality principles

Дата введения 2004—07—01

1 Область применения

Настоящий стандарт устанавливает принципы оценки качества географических данных, определяет подход к формированию и компоненты отчетной информации о качестве данных.

Стандарт предназначен для разработчиков данных, обеспечивающих информацию о качестве, для описания и оценки того, насколько удовлетворительно тот или иной набор данных отображает предметную область, определенную в спецификации продукта, а также для пользователей данных при определении качества данных, необходимого для их конкретного применения. Кроме того, стандарт может быть использован для определения области применения данных и формирования требований к их качеству.

Принципы оценки качества наборов географических данных могут быть использованы для определения, сбора и выдачи информации о качестве комплектов наборов данных, а также более мелких групп данных (поднаборов данных), являющихся подмножествами того или иного набора данных.

Настоящий стандарт применяется для:

- описания и выдачи информации о качестве;
- оценки качества наборов данных;
- разработки спецификации продукции и формирования требований пользователей;
- разработки схем применения данных.

Принципы оценки качества цифровых географических данных могут быть распространены на другие формы географических данных (карты, таблицы и текстовые документы).

Настоящий стандарт не преследует цель определения минимально приемлемого уровня качества географических данных.

2 Соответствие

Любой продукт, претендующий на соответствие настоящему стандарту, должен удовлетворять всем требованиям, определяемым в комплексе проверок набора данных, который приведен в приложении А.

3 Нормативные ссылки

В настоящем стандарте использованы ссылки на следующие стандарты:

ГОСТ Р ИСО 19105—2003 Географическая информация. Соответствие и тестирование

ИСО 19108—2002* Географическая информация — Временная схема

ИСО 19114—2003* Географическая информация — Процедуры оценки качества

ИСО 19115—2003* Географическая информация — Метаданные

* Оригиналы международных стандартов — во ВНИИКИ Госстандарта России.

Издание официальное

4 Определения

В настоящем стандарте использованы следующие термины с соответствующими определениями:

4.1 географическая информация: Информация об объектах и явлениях, содержащая в явном или неявном виде указание на их местоположение относительно Земли.

4.2 точность: Мера соответствия результата, полученного в ходе проверки, принятому эталонному значению.

Примечание — Результат проверки может быть результатом как наблюдения, так и измерения.

4.3 соответствие: Выполнение заданных требований (ГОСТ Р ИСО 19105).

4.4 уровень соответствия качества: Пороговое значение или набор пороговых значений для результирующих оценок качества данных, по которым определяют, насколько набор данных соответствует спецификации продукта или требованиям пользователя (ИСО 19114).

4.5 дата оценки качества данных: Дата или диапазон дат, указывающий, в какое время была произведена операция оценки качества данных.

4.6 элемент качества данных: Количественный компонент, определяющий качество набора данных.

Примечание — Применимость элемента качества данных к набору данных зависит от содержания набора данных и спецификации его продукта. Возможна ситуация, когда не все элементы качества данных будут применимы ко всем наборам данных.

4.7 процедура оценки качества данных: Операция (операции), используемая(-ые) при применении методов оценки и составлении соответствующих отчетов.

4.8 мера качества данных: Измеряемое значение подэлемента качества данных.

Пример — Выраженное в процентах количество правильных значений атрибута.

4.9 описательный элемент качества данных: Неколичественный компонент, определяющий качество набора данных.

Примечание — Неколичественная информация о качестве — назначение, использование и происхождение набора данных.

4.10 результирующая оценка качества данных: Значение или набор значений, получаемых в результате применения некоторой меры качества данных, или результат оценки полученного значения или набора значений по сравнению с заданным приемлемым уровнем качества.

Пример — Результирующая оценка качества данных «90» при типе значения качества данных «процент», выдаваемая для некоторого элемента качества данных, и его подэлемент качества данных «полнота, присутствие» — пример значения, получаемого в результате применения меры качества данных к области определения качества данных.

4.11 область определения качества данных: Область или характеристика(-и) данных, в отношении которых выдается информация о качестве.

Примечание — Область определения качества данных для некоторого набора данных может представлять собой комплект наборов данных, к которому такой набор данных относится, сам набор данных или меньшую группу данных, расположенную физически в пределах общих характеристик набора данных. Общими характеристиками могут быть определенный тип объекта, его атрибут или отношения, критерий сбора данных, источник данных, специфическая географическая или временная информация.

4.12 подэлемент качества данных: Компонент элемента качества данных, описывающий определенный аспект этого элемента качества данных.

4.13 тип значения качества данных: Тип значения для выдачи результирующей оценки качества данных.

Пример — «Булева переменная», «процент», «отношение».

Примечание — Тип значения качества данных всегда присутствует в результирующей оценке качества.

4.14 единица измерения качества данных: Единица измерения при описании результирующей оценки качества.

Пример — «Метр».

Примечание — Единица измерения присутствует только в результирующей оценке качества.

4.15 набор данных: Идентифицированная совокупность данных.

Примечание — Набором данных может быть также относительно малая группа данных, которая, хотя и имеет ограничения в виде пространственной протяженности или типа объекта, физически расположена в пределах более крупного набора данных. Теоретически набором данных могут быть отдельный объект или атрибут объекта, расположенные в пределах более крупного набора данных.

4.16 комплект набора данных: Совокупность наборов данных, имеющая спецификацию одного и того же продукта (ИСО 19115).

4.17 пространственный объект, объект: Абстрактное представление явления реального мира.

Примечание — Объект может быть представлен как тип объекта или как экземпляр (образец) типа объекта. Следует использовать либо тип объекта, либо экземпляр типа объекта в зависимости от назначения.

4.18 атрибут объекта: Характеристика объекта.

Примечание — Атрибут объекта имеет имя, характеризуется определенным типом данных и имеет область допустимых значений. Атрибут экземпляра типа объекта также имеет значение атрибута, принадлежащее области допустимых значений.

4.19 операция с объектом: Операция (действие), которая может происходить для объекта данного типа.

Пример — Операция над объектом типа «дамба» состоит в ее поднятии. Результатом этой операции является повышение уровня воды в резервуаре.

4.20 метаданные: Данные о данных (ИСО 19115).

4.21 спецификация продукта: Описание предметной области и требования к отображению предметной области набором данных.

4.22 качество: Совокупность характеристик продукта, отражающая его способность удовлетворять заявленному и подразумеваемому назначению.

4.23 предметная область: Отображение реального или гипотетического мира, которое включает все, что представляет интерес.

5 Принципы описания качества географической информации

5.1 Компоненты описания качества

Описание качества может быть применено к комплекту наборов данных, набору данных или поднабору данных.

Качество набора данных описывают с использованием элементов качества данных и описательных элементов качества данных.

Элементы качества данных вместе с подэлементами качества данных и дескрипторами подэлементов качества данных определяют, насколько тот или иной набор данных отвечает критериям, установленным в спецификации продукта, и обеспечивают количественную информацию о качестве.

Описательные элементы качества данных обеспечивают общую неколичественную информацию о качестве.

Примечание — Описательные элементы качества данных необходимы для оценки качества набора данных для конкретного приложения, отличного от того, для которого он предназначен.

Качество данных определяется количественной и неколичественной информацией.

Информация о качестве может содержать данные о ее достоверности и надежности. Такой тип информации отражается в отчете об оценке качества по ИСО 19114.

На рисунке 1 представлены компоненты описания и выдачи информации о качестве набора данных.

В приложении Б приведена концепция оценки качества данных, используемая для определения компонентов описания качества географических данных.

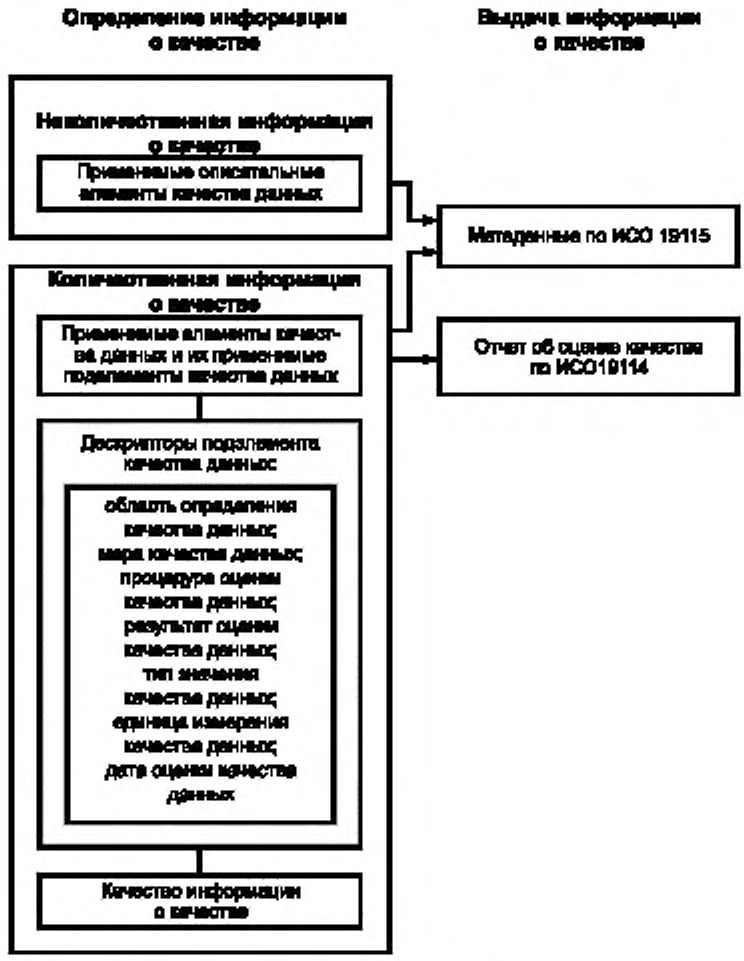

Рисунок 1 — Организация информации о качестве данных

5.2 Элементы качества данных и подэлементы качества данных

5.2.1 Элементы качества данных

Для описания того, насколько набор данных отвечает критериям, установленным в спецификации продукта, используют следующие элементы качества:

а) полноту — наличие и отсутствие объектов, их атрибутов и отношений;

б) логическую согласованность — степень соответствия логических правил структуры данных, атрибутов и отношений (структура может быть концептуальной, логической или физической);

в) позиционную точность — точность положения объектов;

г) временну́ю точность — точность временны́х атрибутов и временны́х отношений объектов;

д) тематическую точность — точность количественных атрибутов и корректность неколичественных атрибутов и классификаций объектов и их отношений.

Для описания количественных компонентов качества набора данных могут быть использованы дополнительные элементы качества, которые в настоящем стандарте не рассматриваются.

5.2.2 Подэлементы качества данных

С целью описания аспектов количественной оценки качества набора данных для элементов качества данных, приведенных в 5.2.1, используют следующие подэлементы качества данных:

а) полноту:

1) присутствие — избыточность данных в наборе данных,

2) отсутствие — отсутствие данных в наборе данных,

б) логическую согласованность:

1) концептуальную согласованность — соответствие правилам концептуальной схемы,

2) доменную согласованность — соответствие значений атрибутов области допустимых значений,

3) согласованность по формату — степень соответствия накопленных данных физической структуре набора данных,

4) топологическую согласованность — корректность представления закодированных топологических характеристик набора данных;

в) позиционную точность:

1) абсолютную или внешнюю точность — соответствие заявленных значений координат значениям координат, принятым или считающимся правильными,

2) относительную или внутреннюю точность — соответствие относительного положения объектов в наборе данных их соответствующим исходным положениям, принятым или считающимся правильными,

3) точность позиционирования данных относительно координатной сетки — соответствие значения позиционирования данных значениям, принятым или считающимся правильными;

г) временну́ю точность:

1) точность измерения времени — правильность временны́х данных об объекте (свидетельствующая об ошибках в измерении времени),

2) временну́ю согласованность — правильность порядка последовательности событий,

3) временну́ю достоверность — соответствие даты по отношению к времени;

д) тематическую точность:

1) правильность классификации — соответствие классификации объектов или их атрибутов предметной области (т. е. реальной местности или эталонному набору данных),

2) правильность неколичественных атрибутов,

3) точность количественных атрибутов.

Для любого элемента качества данных могут быть созданы дополнительные подэлементы качества данных.

5.2.3 Дескрипторы подэлемента качества данных

Полную информацию о качестве формируют для каждого подэлемента качества данных, используя семь следующих дескрипторов подэлемента:

- область определения качества данных;
- меру качества данных;
- процедуру оценки качества данных;
- результирующую оценку качества данных;
- тип значения качества данных;
- единицу измерения качества;
- дату оценки качества данных.

5.3 Описательные элементы качества данных

Для описания неколичественной информации о качестве набора данных используют следующие описательные элементы качества данных:

- назначение;
- использование;
- происхождение.

Элемент «Назначение» содержит информацию о целевом предназначении набора данных.

Примечание — Целевое предназначение набора данных не обязательно совпадает с его фактическим использованием. Фактическое использование определяется описательным элементом качества данных «Использование».

Элемент «Использование» описывает приложение (приложения), для которого (которых) использовался набор данных. Использование набора данных определяется разработчиком набора данных или иными конкретными пользователями данных.

Элемент «Происхождение» описывает историю набора данных и, в известной степени, содержит подробное описание жизненного цикла набора данных, начиная с его сбора и получения, последующего кодирования и преобразования в текущую форму данных.

Описание элемента «Происхождение» складывается из двух уникальных компонентов:

- источника информации для формирования набора данных;
- описания событий или преобразований, выполненных в процессе существования набора данных, включая периодическое или постоянное обновление.

Дополнительный описательный элемент (элементы) качества данных для описания неколичественной информации о качестве набора данных в настоящем стандарте не рассматривается.

6 Определение качества географической информации

6.1 Определение количественной информации о качестве

6.1.1 Общие положения

В общем процессе определения количественной информации о качестве некоторые подпункты могут быть неприменимы во всех случаях.

6.1.2 Определение применимых элементов качества данных

Все элементы качества данных, применимые к набору данных, должны быть определены. Некоторые элементы качества данных могут оказаться неприменимыми к конкретному типу набора данных.

Примечания

1 При определении применимости следует использовать спецификацию продукта набора данных.

Пример — Для набора данных, содержащего только объекты с пространственными ссылками в виде географических идентификаторов, элемент качества «позиционная точность» может отсутствовать.

2 В приложении В даны примеры определения применимых элементов качества данных.

6.1.3 Создание определяемых пользователем элементов качества данных

Новый элемент (элементы) качества данных может быть установлен в том случае, если элементы качества данных, приведенные в настоящем стандарте, в недостаточной степени характеризуют качество данных. Наименование и описание дополнительного элемента качества данных могут быть включены как часть информации о качестве набора данных.

6.1.4 Определение применимых подэлементов качества данных

Все применимые подэлементы качества данных для каждого элемента качества данных должны быть определены (не менее одного подэлемента для каждого применимого элемента качества данных). Некоторые подэлементы какого-то применимого элемента качества данных могут быть неприменимыми для конкретного типа набора данных.

Примечания

1 Для определения применимости подэлемента качества следует использовать спецификацию продукта.

2 В приложении В даны примеры определения применимых подэлементов качества.

6.1.5 Создание определяемых пользователем подэлементов качества данных

Новый подэлемент (подэлементы) качества данных может быть установлен в том случае, если подэлементы качества данных, перечисленные в настоящем стандарте, в недостаточной степени характеризуют аспект качества. Наименование и описание дополнительного подэлемента качества данных могут быть включены как часть информации о качестве набора данных.

6.1.6 Использование дескрипторов подэлемента качества данных

6.1.6.1 Область определения качества данных

Для каждого применимого подэлемента качества данных должно быть установлено не менее одной области определения качества. Областью определения качества данных может быть комплект наборов данных, к которому принадлежит набор данных, набор данных или более мелкая группа данных, расположенная в пределах набора данных, обладающая общностью характеристик (поднабор данных). Если область определения качества данных не может быть установлена, ею должен быть сам набор данных.

Примечание — Область (области) определения качества следует устанавливать с использованием спецификации продукта и неколичественной информации о качестве, содержащейся в описательных элементах качества.

В рамках одного и того же набора данных качество может быть различным. Для каждого применимого подэлемента качества данных может быть установлено несколько областей определения качества данных, что позволяет более полно описывать количественную информацию о качестве. Область определения качества данных должна быть адекватно описана. Для описания области определения качества данных используют:

- уровень (комплект данных, к которому принадлежит набор данных, собственно набор данных или поднабор данных);
- типы объектов (перечни типов объектов, атрибутов объектов и отношений) или конкретные объекты (перечни экземпляров объектов, значений атрибутов и отношений объектов);
- географическую область;
- временну́ю область (заданные временны́е пределы и точность временны́х пределов).

6.1.6.2 Мера качества данных

Для каждой области определения качества данных должна быть предусмотрена одна мера качества данных. Для каждой меры качества данных должно быть указано наименование (если оно существует), вид критерия, применимого к оценке данных в конкретной области определения качества, а также граничные или предельные параметры.

Примечание — Примерами граничных или предельных параметров являются доверительные интервалы и частота (вероятность) ошибок.

Качество набора данных измеряют с использованием различных критериев. Одной меры качества данных может быть недостаточно для полной оценки качества данных, задаваемых областью их определения, и получения меры качества для всех возможных вариантов использования такого набора данных. Комбинация мер качества данных может давать полезную информацию. Для данных, принадлежащих к одной предметной области, может быть предусмотрено несколько мер качества данных.

Примечание — В ИСО 19114 приведены примеры наименований и дескрипторов типов мер качества данных.

6.1.6.3 Процедура оценки качества данных

Для каждой меры качества данных необходимо предусмотреть одну процедуру оценки качества данных. Процедура оценки качества данных должна описывать методологию, используемую для применения меры качества к данным, задаваемым некоторой областью определения качества данных, или содержать ссылку на документацию, в которой такая методология описывается.

Примечания

1 Примерами такой документации являются опубликованные статьи или принятые промышленные стандарты.

2 В ИСО 19114 приведена структура процедур оценки качества данных, применимых к наборам данных, и тип информации, которая должна быть отражена в процедуре оценки качества данных.

6.1.6.4 Результирующая оценка качества данных

Для каждой меры качества данных должна быть определена результирующая оценка качества данных. Результирующей оценкой качества данных являются:

- значение или набор значений, полученных в результате применения некоторой меры качества данных к данным, заданным областью определения качества данных;
- результат сравнения значения или набора значений, полученных при применении некоторой меры качества к данным, заданным областью определения качества, с некоторым приемлемым уровнем соответствия качества. Этот тип результирующей оценки качества данных определяют как «удовлетворительная/неудовлетворительная».

Могут предусматриваться оба типа результирующих оценок качества данных, определяемых в настоящем стандарте.

Примечание — Определение уровней соответствия качества приведено в ИСО 19114.

6.1.6.5 Тип значения качества данных

Для каждой результирующей оценки качества данных должен быть предусмотрен один тип значения качества данных.

Примечание — Типом значения качества данных «удовлетворительно/неудовлетворительно» будет «булева переменная».

6.1.6.6 Единица значения качества данных

Для каждой результирующей оценки качества данных должна быть предусмотрена одна единица значения качества данных.

6.1.6.7 Дата оценки качества данных

Для каждой меры качества данных должна быть предусмотрена одна дата оценки качества данных в соответствии с требованиями ИСО 19108.

6.2 Определение неколичественной информации о качестве

6.2.1 Определение применимых описательных элементов качества данных

Элемент качества «Назначение» набора данных всегда должен быть применим.

Элемент качества «Использование», заявленный разработчиком набора данных, должен быть применим.

Элемент качества «Происхождение» набора данных всегда должен быть применим. Если информация о происхождении неизвестна, должно быть дано объяснение отсутствия такого рода информации.

Для поднабора данных, являющегося предметом анализа, может быть приведена информация о происхождении, отличающаяся от информации о происхождении остальной части набора данных. Такая информация о происхождении для поднаборов данных может быть предусмотрена как часть неколичественной информации о качестве набора данных для более полного представления информации о качестве.

6.2.2 Создание дополнительных описательных элементов качества данных

Новый описательный элемент (элементы) качества данных может быть использован в том случае, если описательные элементы качества данных, перечисляемые в настоящем стандарте, в недостаточной степени характеризуют область общей неколичественной оценки качества. Наименование и определение дополнительного описательного элемента качества включают как часть информации об их качестве.

7 Выдача информации о качестве

7.1 Выдача количественной информации о качестве

Количественную информацию о качестве выдают в виде метаданных в соответствии с требованиями ИСО 19115.

Кроме того, количественную информацию о качестве выдают с использованием отчета об оценке качества в соответствии с требованиями ИСО 19114.

7.2 Выдача неколичественной информации о качестве

Неколичественную информацию о качестве выдают как метаданные в соответствии с требованиями ИСО 19115.

Примечание — Неколичественную информацию о качестве не выдают в отчете об оценке качества по ИСО 19114.

ПРИЛОЖЕНИЕ А
(обязательное)

Комплекс проверок набора данных

А.1 Комплекс проверок

А.1.1 Общие положения

Все проверки набора являются предварительными.

А.1.2 Проверка компонентов

а) Назначение проверки — определение соответствия посредством удостоверения в том, что в описании качества использованы компоненты качества.

б) Методы проверки:

1) исследование описания качества и удостоверение в том, что элементы качества данных (подэлементы качества данных и дескрипторы подэлементов качества данных) использовались для обеспечения количественной информации о качестве;

2) исследование описания качества и удостоверение в том, что описательные элементы качества данных использовались для обеспечения неколичественной информации о качестве.

в) См. 5.1 настоящего стандарта.

А.1.3 Проверка правильности описания качества

а) Назначение проверки — определение соответствия посредством удостоверения в правильности описания качества.

б) Методы проверки:

1) исследование описания качества и удостоверение в том, что используемые элементы и подэлементы качества данных перечислены в настоящем стандарте или относятся к дополнительным и описывают компонент или аспект количественной информации о качестве, который в настоящем стандарте не определен;

2) исследование описания качества и удостоверение в том, что дескрипторы подэлементов качества данных, определяемые в настоящем стандарте, использовались для количественного описания качества;

3) исследование описания качества и удостоверение в том, что используемые описательные элементы качества данных перечислены в настоящем стандарте или относятся к категории элементов, определяемых пользователем, и описывают область неколичественной информации о качестве, которая в настоящем стандарте не определена.

в) См. 5.2 и 5.3 настоящего стандарта.

А.1.4 Проверка применимости количественной информации о качестве

а) Назначение проверки — определение соответствия посредством удостоверения в применимости количественного описания качества.

б) Метод проверки — определение требований из спецификации продукта, относящихся к количественной информации о качестве, и их использования для установления применимых элементов и подэлементов качества данных. Сравнение применимых подэлементов качества данных с подэлементами качества данных, использованными в описании качества, для удостоверения в том, что все подэлементы качества данных в рассматриваемом наборе данных были определены и использовались в описании качества.

Примечание — Соответствие подтверждается, если неприменимые подэлементы качества данных дополнительно используют для количественного описания качества. Однако неприменимые подэлементы качества данных не могут подвергаться дальнейшей проверке на соответствие.

в) См. 6.1 настоящего стандарта.

А.1.5 Проверка применимости неколичественной информации о качестве

а) Назначение проверки — определение соответствия посредством удостоверения в применимости неколичественного описания качества.

б) Метод проверки — удостоверение в том, что применимые описательные элементы качества данных используют для неколичественного описания качества.

в) См. 6.2 настоящего стандарта.

А.1.6 Проверка исключительности

а) Назначение проверки — определение соответствия посредством удостоверения в том, что определяемые пользователем элементы (подэлементы) качества данных, используемые в описании качества, являются исключительными, и о каждом элементе (подэлементе) обеспечивается достаточная информация.

б) Методы проверки:

1) исследование всех определяемых пользователем элементов качества данных и удостоверение в том, что каждый из них характеризует тот компонент количественной информации о качестве, который не приведен в настоящем стандарте;

2) исследование всех определяемых пользователем подэлементов качества данных и удостоверение в том, что каждый из них характеризует тот аспект количественной информации о качестве, который не приведен в настоящем стандарте;

3) исследование всех определяемых пользователем описательных элементов качества данных и удостоверение в том, что каждый из них характеризует некоторую область неколичественной информации о качестве, которая не приведена в настоящем стандарте, а имя и описание определяемого пользователем элемента составляют часть описания качества.

в) См. 6.1.3, 6.1.5 и 6.2.2 настоящего стандарта.

А.1.7 Проверка правильности использования дескрипторов подэлементов качества данных

а) Назначение проверки — определение соответствия посредством удостоверения в том, что дескрипторы подэлементов качества данных правильно использовались в описании качества.

б) Метод проверки — сравнение настоящего стандарта и информации о качестве, представляемой для каждого применимого подэлемента качества данных (включая определяемые пользователем подэлементы качества данных), для определения того, что правила использования дескрипторов подэлементов качества данных были соблюдены.

в) См. 6.1.6 настоящего стандарта.

А.1.8 Проверка выдачи информации о качестве в виде метаданных

а) Назначение проверки — определение соответствия посредством удостоверения в том, что описание качества выдано в виде метаданных.

б) Метод проверки — удостоверение в том, что количественная и неколичественная информация о качестве была выдана в виде метаданных в соответствии с ИСО 19115.

в) См. раздел 7 настоящего стандарта.

А.1.9 Проверка выдачи количественной информации о качестве в виде отчета об оценке качества

а) Назначение проверки — определение соответствия посредством удостоверения в том, что количественная информация о качестве выдана в виде отчета об оценке качества.

б) Метод проверки — удостоверение в том, что количественная информация о качестве в отчете об оценке качества соответствует требованиям ИСО 19114.

в) См. 7.1 настоящего стандарта.

ПРИЛОЖЕНИЕ Б
(рекомендуемое)

Концепция оценки качества данных и ее использование

Б.1 Основные положения

Набор данных определяют как идентифицированную совокупность данных. Эти данные описывают объекты реального мира, которые характеризуются пространственными, атрибутивными и временными аспектами. Процесс моделирования потенциально бесконечного разнообразия реального мира в виде некоторой идеальной структуры включает в себя переход от объектов реального мира к некоторой предметной области. При установлении качества набора (или части набора) данных определяют соответствие предметной области заранее подготовленной спецификации.

Б.2 Концепция оценки качества данных

Поскольку набор данных, как правило, создается не для какого-то конкретного приложения, а для нескольких возможных приложений, качество набора данных может быть определено только оценкой элементов качества данных и описательных элементов качества данных. Элементы качества данных оценивают расхождение между созданным набором данных и предметной областью (т. е. правильным набором данных будет считаться тот, который соответствует спецификации продукта). Описательные элементы качества данных обеспечивают общую неколичественную информацию. Назначение определяет причины создания набора данных и его предназначение. Использование определяет тип приложения, для которого набор данных использовался. Происхождение описывает историю набора данных.

Концепция оценки качества данных предоставляет структурированный набор положений, средств и методов, необходимых для разработчиков и пользователей данных. Разработчик данных получает средства для определения того, насколько информация, используемая для создания того или иного набора данных, отражает его предметную область. Разработчик данных может проверить, насколько набор данных отвечает критериям, заданным в спецификации его продукта. Пользователи данных получают средства для оценки наборов данных применительно к той или иной предметной области, определяемой на основании требований пользователя данных. Пользователи данных могут оценивать качество для удостоверения в том, что набор данных удовлетворяет требованиям того или иного приложения (рисунок Б.1).

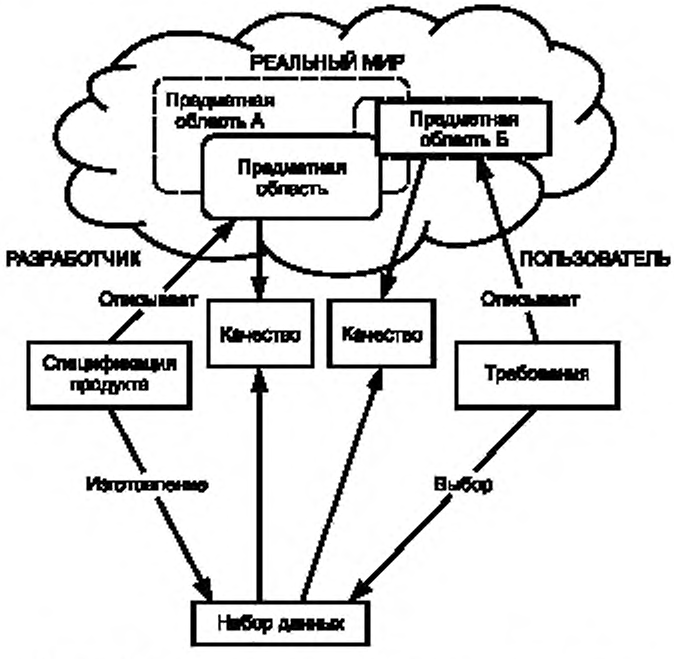

Рисунок Б.1 — Структура концепции оценки качества данных

Б.3 Структура наборов данных и компонентов для описания качества

Набор данных может принадлежать комплекту наборов данных. Все наборы данных, принадлежащие некоторому комплекту наборов данных, зачастую обладают одним и тем же качеством. Концепция оценки качества данных различает комплекты наборов данных и позволяет замену и выдачу для набора данных информации о качестве соответствующего комплекта.

Набор данных может рассматриваться как содержащий большое, хотя и конечное, число более мелких групп данных. Предполагается, что более мелкие группы данных, объединенные принадлежностью к одному и тому же типу объекта, атрибуту объекта или отношению, критерию сбора или к одной и той же географической области, имеют одинаковое качество. Более мелкая группа данных может состоять из одного экземпляра типа объекта, значения атрибута или отношения и, теоретически, концепция оценки качества данных позволяет каждому экземпляру типа объекта, значению атрибута и виду отношения набора данных иметь собственное качество. Качество более мелких групп данных может не совпадать с качеством остальной части набора данных, к которому они принадлежат. Концепция оценки качества данных допускает выдачу информации о качестве набора данных и дополнительно отличной от нее информации о качестве более мелких групп, которая определяется спецификой области определения качества данных.

Примечание — Для разработчика данных спецификация продукта описывает предметную область и содержит параметры для построения набора данных. Для пользователя данных требования описывают предметную область, которая может совпадать или не совпадать с предметной областью набора данных. Истинное качество набора данных определяется тем, насколько хорошо оно представляет предметную область.

Выделяют два уникальных компонента информации о качестве данных для описания качества набора данных — количественные и неколичественные компоненты качества. Элементы качества данных являются количественными компонентами информации о качестве, описательные элементы качества данных — неколичественными компонентами.

Элементы качества данных позволяют оценить, насколько тот или иной набор данных отвечает критериям, определенным в спецификации его продукта. Аспекты элементов качества данных называют подэлементами качества данных. Подэлементы качества данных оценивают или проверяют различными способами. Концепция оценки качества данных подразумевает, что не все элементы и подэлементы качества данных, а также не все средства их оценки и проверки могут быть применимы к конкретному типу набора данных. Кроме того, некоторые подэлементы качества данных являются применимыми к набору данных и могут быть оценены или проверены для него, в то время как другие применимы к более мелким группам данных, принадлежащим одному и тому же набору данных, и могут быть оценены или проверены для них.

Настоящий стандарт определяет элементы качества данных непосредственно как средства выявления и выдачи раздельных видов информации о качестве. Тем не менее, как правило, подэлементы качества данных взаимосвязаны. Например, ошибка в координатах может повлечь за собой, как минимум, ошибки двух типов — позиционную и топологическую. Значения подэлементов качества данных в соглашении на продукт и способ их трактовки находятся в компетенции разработчика спецификации продукта.

Если элементы качества данных позволяют оценивать, насколько набор данных отвечает критериям, определенным в спецификации его продукта, то описательные элементы позволяют дополнительно оценить набор данных для конкретного приложения, поскольку содержат информацию о назначении, использовании и происхождении набора данных.

Б.4 Выдача информации о качестве

Б.4.1 Когда следует выдавать информацию о качестве

Наборы данных постоянно создают, обновляют и объединяют, в результате чего качество или компонент качества набора данных может изменяться. На информацию о качестве набора данных могут повлиять следующие условия:

а) некоторый объем данных удаляется из набора данных, изменяется или добавляется к набору данных;

б) изменяется спецификация набора данных (описание и требования к набору данных);

в) изменяется реальный мир.

Изменение набора данных может происходить довольно часто, так как многие наборы данных не являются статическими. Происходит увеличение обмена информацией, использование наборов данных в различных целях и обусловленное этим обновление и уточнение наборов данных для удовлетворения различных целей. Если информация о качестве некоторого набора данных изменяется в результате его модификации, то качество набора данных должно быть заново оценено и уточнено.

При создании набора данных необходимо обеспечить полное представление обо всех применимых количественных и описательных элементах качества данных. Исключением является элемент качества данных «Использование», который должен быть описан после того, как набор данных создан. Первоначально могут сообщаться сведения об использовании набора данных его разработчиком (в предположении, что разработчик данных действительно использует такой набор данных). В дальнейшем пользователи данных должны предоставлять информацию об использовании набора данных, если оно отличается от его предназначения. Благодаря таким сообщениям появится возможность постоянного обновления этого описательного элемента качества данных для отражения в нем непредвиденных вариантов использования.

По мере использования некоторого набора данных спецификация его будет обновляться для того, чтобы будущие изменения этого набора лучше отвечали фактическим потребностям. Если спецификация изменяется, изменяется также и качество текущего набора данных. Информация о качестве для набора данных всегда должна отражать текущий набор данных с учетом текущей спецификации.

Изменения реального мира происходят постоянно. Это могут быть и природные явления, например движение земной коры или эрозия почвы, и, главным образом, — результат деятельности человека. Последние, как правило, происходят быстро. По этой причине время сбора данных является важным показателем при оценке качества набора данных. В некоторых случаях даже скорость изменения данных представляет интерес.

Количество информации о качестве и требования к ее хранению могут превышать соответствующие параметры набора данных. Важно представлять информацию о качестве в компактном, легко воспринимаемом и легко отыскиваемом виде.

При наличии больших объемов информации о качестве комплектов наборов данных, наборов данных или групп данных она должна помещаться в файл или архив метаданных.

Б.4.2 Выдача информации о качестве в виде метаданных

Б.4.2.1 Выдача количественной информации о качестве в виде метаданных

Количественная информация о качестве набора данных может быть получена для множества областей определения качества. Области определения качества данных могут включать комплект наборов данных, набор данных и группу данных, входящую в набор данных.

Количественная информация о качестве может описывать качество комплекта наборов данных в целом, если все наборы данных одного комплекта эквивалентны и лучше всего оцениваются на уровне комплекта. Информация о качестве может храниться в виде метаданных вместе с комплектом наборов данных. В этом случае метаданные каждого набора данных должны содержать указатель на него. Эта же информация о качестве может повторяться в метаданных наборов данных и быть их частью. Если информация о качестве некоторого набора данных уникальна и отличается от информации о качестве комплекта наборов данных, такой набор данных должен иметь уникальные метаданные о качестве, а подстановка другой информации не рекомендуется.

Количественная информация о качестве может различаться для набора данных и поднаборов, определяемых областью определения их качества. Объем записываемой количественной информации о качестве частично зависит от числа установленных областей определения качества данных. Количественную информацию о качестве записывают, как правило, для той или иной группы, являющейся объектом анализа, только в том случае, если она отличается от количественной информации о качестве, выдаваемой на «более высоком» уровне. Выдача информации о качестве обычно начинается с верхних уровней набора данных и постепенно переходит на более низкие уровни (таблица Б.1).

Таблица Б.1 — Иерархический принцип выдачи количественной информации о качестве

Область определения качества	Задаваемая позиционная точность (оценка качества данных в отношении абсолютной точности)	Определенная позиционная точность (результирующая оценка качества данных в отношении абсолютной точности)
Набор данных	1,35	1,35
Дороги	1,10	1,10
Водотоки	1,35	Не выдается
Железные дороги	1,20	1,20
Трубопроводы	1,80	1,80
Примечание — Типом значения качества для результирующих оценок качества в приведенном примере является «расстояние», единицей измерения качества — «метры».		

Для того чтобы минимизировать усилия на создание, хранение и интерпретацию данных, в этом примере предполагается, что информация о качестве будет выдаваться в виде метаданных только для самого набора данных и данных, принадлежащих к таким областям определения качества, как «дороги», «железные дороги» и «трубопроводы». Информация о качестве для данных, входящих в область определения качества «водотоки», будет опускаться.

В ИСО 19115 не регламентирована запись количественной информации о качестве в виде метаданных для экземпляров типов объектов, значений атрибутов или отдельных видов отношений объектов. Чтобы обойти это ограничение, задают некоторую определенную группу, состоящую из типа объекта, значения атрибута или отдельных видов отношений. Если количественная информация о качестве единичных объектов отличается от количественной информации о качестве объектов одного типа, она может быть описана в виде атрибута такого объекта. Иерархический принцип выдачи информации может быть также применен к однотипным объектам и характеристикам объектов. Предполагается, что количественная информация о качестве при единичном ее проявлении будет выдаваться только в том случае, когда она отличается от значений более высокого уровня.

Поскольку способ определения параметров качества в рамках наборов данных зависит от набора данных, рекомендации по фактическому установлению информации о качестве для объектов, характеристик и отношений объектов не предлагаются.

На рисунке Б.2 представлена иерархическая структура набора данных, для которого количественная информация о качестве выдается в виде метаданных, а также предлагается метод выдачи (как часть файла метаданных, архива метаданных или в виде атрибута в рамках набора данных)

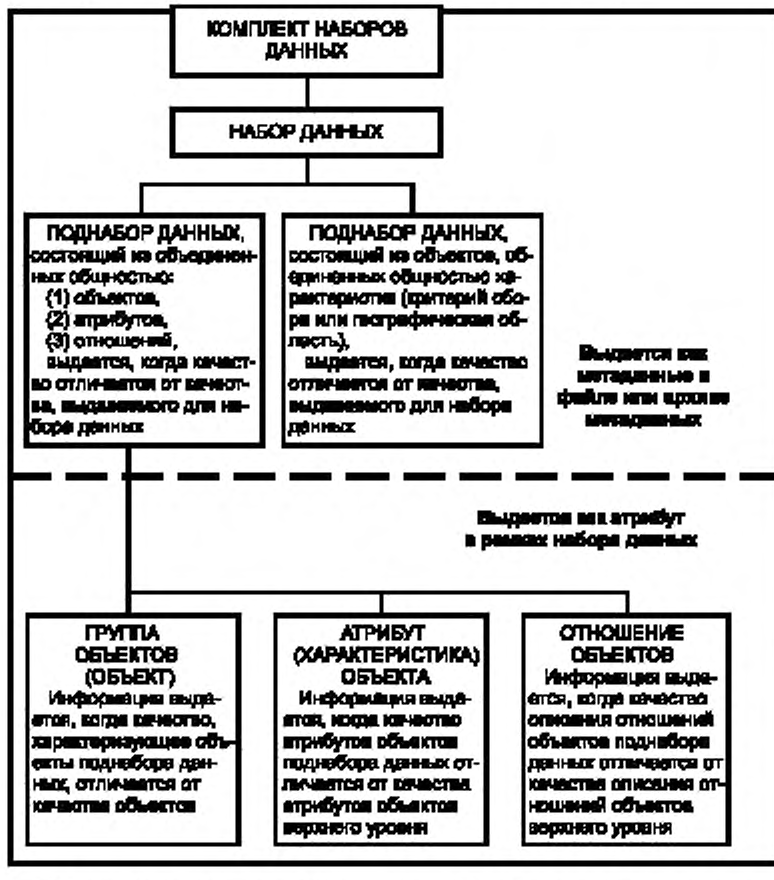

Рисунок Б.2 — Предлагаемая схема выдачи количественной информации о качестве в виде метаданных

В ИСО 19114 рассматривается проблема избыточного представления количественной информации о качестве и необходимость сжатия такой информации, а также проблема получения более детальной количественной информации о качестве посредством использования отчета об оценке качества.

Б.4.2.2 Выдача неколичественной информации о качестве в виде метаданных

Информацию о качестве таких описательных элементов качества данных, как «Назначение» и «Использование», считают существенной, относящейся только к набору данных, и записывают только для набора данных. Если набор данных принадлежит комплекту наборов данных и содержит такие же описательные элементы качества данных, как «Назначение» и «Использование», настоящий стандарт позволяет выдавать информацию о назначении и использовании комплекта наборов данных для набора данных.

Информацию о качестве описательного элемента качества данных «Происхождение» считают существенной информацией о наборе данных. Однако информация о происхождении набора данных может отличаться от информации о происхождении входящих в него поднаборов данных, установленных областью определения качества. Предполагается, что информация о происхождении поднаборов данных из области определения качества должна быть включена в описание, если она отличается от информации о происхождении, включенной в описание для набора данных.

(Настоящий стандарт рекомендует использовать для набора данных информацию о происхождении комплекта наборов данных, в который входит набор данных, чтобы не повторять информацию.)

ПРИЛОЖЕНИЕ В
(справочное)

Примеры определения применимых элементов и подэлементов качества данных

В.1 Пример 1 — Цифровая карта мира

(Digital Chart of the World — DCW)

В.1.1 Общее представление

Пример иллюстрирует оценку разработчиком данных соответствующей количественной информации о качестве определением применимых элементов и подэлементов качества данных на основе спецификации продукта. Когда какой-то подэлемент качества данных определяют как применимый, спецификацию продукта дополнительно используют для установления соответствующей (-их) области (-ей) определения качества данных. Для каждой установленной области определения качества данных задают соответствующую меру качества данных (также с использованием спецификации продукта).

Пример также включает оценку разработчиком данных неколичественной информации о качестве посредством сбора информации для применимых описательных элементов качества данных.

Пример не включает фактическую выдачу соответствующей информации о качестве в виде метаданных или с использованием отчета об оценке качества.

В.1.2 Базовая информация

Набор данных: Цифровая карта мира (Digital Chart of the World - DCW)
Спецификация продукта: Военная спецификация MIL-D-89009, 13 апреля 1992 г.
Описание продукта: (извлечено из спецификации продукта и излагается в сжатом виде)

DCW — глобальная база данных общего назначения, предназначенная для поддержки приложений географической информационной системы (ГИС). База данных DCW содержит пять библиотек, записанных на четырех компакт-дисках. Библиотека BROWSE содержит данные о мире в масштабе примерно 1:31000000, которые могут быть отображены.

Четыре библиотеки подробных данных (по одной на каждом компакт-диске) содержат данные для (1) Северной Америки, (2) Европы и Северной Азии, (3) Южной Америки, Африки и Антарктиды, а также для (4) Южной Азии и Австралии в масштабе 1:1000000.

В базе данных DCW, организованной по тематическим слоям, используется формат VPF (Vector Product Format).

Библиотека BROWSE содержит восемь тематических слоев; каждая из четырех библиотек подробных данных содержит по семнадцать тематических слоев:

1. Политико-административное деление и океаны
2. Урбанизированные территории
3. Железные дороги
4. Автомобильные дороги
5. Наземные и подземные коммуникации
6. Водные объекты на суше
7. Дополнительная информация по слою 6 — малоразмерные объекты
8. Рельеф с изолиниями, заливками и высотными отметками
9. Дополнительная информация по слою 8 — промежуточные изолинии, малоразмерные контуры
10. Земельные угодья, характеристики земной поверхности
11. Природные объекты в океанах
12. Геологические и геоморфологические объекты на суше
13. Аэропорты
14. Объекты общего назначения, производственные объекты, достопримечательности
15. Транспортная сеть
16. Растительность
17. Качественные и количественные характеристики данных по листам карты, атрибутивная информация.

Каждый объект в DCW определяется атрибутами и закодированными комбинациями значений атрибутов.

В.1.3 Оценка соответствующей количественной информации о качестве

Примечание — Разработчик данных не нашел все элементы качества данных и подэлементы качества данных применимыми, поскольку не все они определяются в спецификации продукта. В таблице В.1 перечислены только применимые элементы и подэлементы качества данных.

Таблица В.1 — Соответствие количественной информации о качестве

Пример 1

Номер пункта спецификации продукта	Содержание	Применимый элемент (подэлемент) качества	Область (-и) определения качества данных
4.1.2.a	Анализ произвольно выбранных фрагментов в различных слоях карты для удостоверения в полноте объектового состава и атрибутов данных (для всех 17 областей определения качества)	Полнота Присутствие	17 областей определения качества, каждая из которых состоит из одного тематического слоя
4.1.2.g	Анализ окончательного формата данных поставщика до преобразования в формат VPF и запись окончательной частоты появления объектов в каждом тематическом слое		
3.1.2	Проверка того, что цифровая информация, первоначально приводимая в примечаниях, таблицах и графиках исходного материала, — Операционной навигационной карты (Operational Navigational Chart — ONC), включается в область значений атрибутов, относящихся к информации о качестве данных, а также в таблицы качества данных	Полнота Присутствие	Набор данных
3.3	Единицей измерения для базы данных DCW является английская система мер	Логическая согласованность	Набор данных
4.1.2.g	Анализ окончательного формата данных поставщика до преобразования в формат VPF. Проверяют имена и определения всех атрибутов	Доменная согласованность	
3.2.1	В качестве базовой принимается Всемирная геодезическая система (World Geodetic System — WGS-84) (область определения качества —набор данных)	Логическая согласованность Согласованность по формату	1. Набор данных 2. Объекты, закодированные как точки и полигоны 3. Текстовые строки
3.9	Если на исходном материале (ONC) полигоны меньше минимально различимого размера в 3,05 мм (0.12″) и состоят из одного ребра, они представлены точечным объектом (область определения качества — закодированные объекты)		
4.1.2.d	Проверка соответствия качества должна производиться как средство выявления и исправления проблем, связанных с размещением текста. Например, необходимо проверять, чтобы текстовые строки не налагались одна на другую (область определения качества — текстовые строки)		
4.1.2.d	Анализ всех отображенных тематических данных для удостоверения во взаимной согласованности положения объектов между слоями-покрытиями (область определения качества — набор данных)	Логическая согласованность Топологическая согласованность	1. Набор данных 2. Все соединения
4.1.2.g	Проверка правильности топологии (область определения качества — набор данных)		
3.9.d	Правила соединения применяют к объектам типа «Дороги и железные дороги» в базе данных DCW. Там, где на ONC дороги и железные дороги разрываются текстом, разрывы в сети шириной больше 0.1″ (2,54 мм) кодируют как специальные соединительные элементы с кодом атрибута «Тип», равным 8, и кодом атрибута «Статус», равным 4, 6 или 9. Более узкие разрывы дорог и железных дорог имеют код атрибута «Тип», равный 1 или 2, а код атрибута «Статус», равный 1, 2, 3 или 5. Примечание — Правила соединения не применяют, если разрывы обусловлены природными препятствиями (область определения качества — соединения)		

Окончание таблицы В.1

Номер пункта спецификации продукта	Содержание	Применимый элемент (подэлемент) качества	Область (-и) определения качества данных
3.1.2	Абсолютная вертикальная точность DCW идентична точности оттисков ONC и Авиационной навигационной карты (Jet Navigation Chart — JNC) при линейной погрешности 90 % по отношению к уровню моря (область определения качества — набор данных)	Позиционная точность Абсолютная точность (вертикальная)	1. Набор данных 2. Все горизонтали 3. Все отметки высот
3.1.2.a	Точность положения горизонталей, взятых с источника данных (ONC), составляет половину высоты сечения рельефа на исходном материале или ±150 м (±500 футов) (область определения качества — горизонтали)		
3.1.2.b	Точность отметок высоты, взятых с источника данных (ONC), составляет ±30 м (±100 футов) (область определения качества — отметки высот)		
3.1.1.1	Абсолютная горизонтальная точность DCW для всех объектов, взятых с ONC, составляет 2040 м (6700 футов), округленных до 5 м при 90 %-ной повторяющейся погрешности, в WGS-84. Абсолютная горизонтальная точность DCW для всех объектов, взятых с JNC, составляет 4270 м (14006 футов) при 90 %-ной повторяющейся погрешности (применяется для всех трех отмеченных областей определения качества. Регион Антарктики является отдельной областью определения качества, поскольку он имеет уникальный источник данных. Выделение дорог в отдельную область определения качества объясняется тем, что дороги являются объектами с четко выраженными контурами, что предполагает более высокую точность их оценки)	Позиционная точность Абсолютная точность (горизонтальная)	1. Все регионы мира, за исключением Антарктики 2. Антарктика 3. Дороги
3.4.5	Каждый объект DCW определяется атрибутами и комбинациями закодированных значений атрибутов	Тематическая точность Правильность классификации	Набор данных
4.1.2.c	Частоты кодов атрибутов автоматически табулируются для выявления неправильных кодов и необычного поведения в таблицах атрибутов. После этого строят графики данных, которые подвергают всесторонней проверке с точки зрения точности кодов атрибутов. В большинстве случаев необходимо строить второй и третий графики для проверки исправления выявленных ошибок		

Составленная разработчиком данных сводка соответствующей количественной информации о качестве представлена в таблице В.2

Таблица В.2 — Сводка соответствующей количественной информации о качестве

Пример 1

Элемент качества данных	Подэлемент качества данных	Соответствие	Число установленных областей определения качества данных
Полнота	Присутствие	Да	1
	Отсутствие	»	1
Логическая согласованность	Концептуальная согласованность	Нет	—
	Доменная согласованность	Да	1
	Согласованность по формату	»	3
	Топологическая согласованность	»	2

Окончание таблицы В.2

Элемент качества данных	Подэлемент качества данных	Соответствие	Число установленных областей определения качества данных
Позиционная точность	Абсолютная или внешняя точность	Да	3 — вертикальная; 3 — горизонтальная
	Относительная или внутренняя точность	Нет	—
	Точность позиционирования относительно координатной сетки	»	—
Временная точность	Точность временных измерений	Нет	—
	Согласованность по времени	»	—
	Временное соответствие	»	—
Тематическая точность	Правильность классификации	Да	1
	Правильность количественных атрибутов	Нет	—
	Точность количественных атрибутов	»	—

Состав неколичественной информации о качестве

Назначение: DCW — глобальная база данных общего назначения, предназначенная для поддержки приложений ГИС — MIL-D-89009, 3.4.1 Описание продукта.

Использование: Использование № 1 **Разработка баз данных:** ESRI (Environmental Science Research Institute) применял DCW для разработки базы данных ArcWorld в целях использования вместе с фирменными пакетами программного обеспечения ARC/INFO и Arcview. Этот вариант использования найден в Интернете.

Использование № 2 **Разработка комплекта электронных карт:** Комплект электронных карт в формате WHEAT был подготовлен на основе DCW; серии электронных карт в масштабе 1:1000000, исходным материалом для создания которых послужили ONC. Содержание базы данных DCW было выбрано для изысканий и исследований по использованию природных ресурсов и освоению территорий в развивающихся странах, а также создания образцов наборов данных в формате WHEAT. Этот вариант использования найден в Интернете.

Использование № 3 **Разработка трехмерных представлений для цифровых электронных карт:** Процедура решения задачи для конкретного примера (карта Холи-Ленд, Великобритания) заключалась в следующем:

1. Подготовка цифровой модели местности (ЦММ) Холи-Лэнд в ARC/INFO из набора данных DCW
2. Совмещение ЦММ с растровыми изображениями местности
3. Создание цифровой матрицы рельефа на территорию
4. Использование средства REGISTER для ввода космических снимков, полученных со спутника Landsat-TM, в ARC/INFO
5. Совмещение снимков с Landsat-TM с ЦММ путем их трансформирования программным модулем SURFACED FAULTS ARC/INFO
6. Задание азимута и угла наблюдения для построения перспективной блок-диаграммы территории с помощью программного модуля (опции) SURFACE OBSERVER RELATIVE
7. «Драпировка» рельефа космическим изображением с использованием средства SURFACE DRAPE. Формирование трехмерного представления Холи-Лэнд. Этот вариант использования найден в Интернете.

Происхождение: Источник: Содержимое базы данных DCW основывается, преимущественно, на объектовом составе комплекта ONC в масштабе 1:1000000 (все регионы, за исключением Атлантического региона), подготовленного Картографическим управлением Министерства обороны США в период с 1974 г. по 1991 г.

Этап процедуры: Стабильные позитивы создавались из негативов оригинальных репродукций (до 35 негативов на лист ONC) и либо оцифровывались сканированием с последующей векторизацией растровых записей, либо вручную в векторную форму. Векторные данные были затем дополнены атрибутивными данными с использованием программного обеспечения ARC/INFO. Преобразование в географические координаты было выполнено с использованием картографической сетки для каждого листа. Цифровая информация по границам прилегающих листов была сведена для создания крупных

региональных наборов данных. Региональные наборы данных затем разделялись на фрагменты 5×5 и преобразовывались из формата программного обеспечения ARC/INFO в VPF. После этого данные переписывались на компакт-диски. Контроль качества выполняла независимая группа на каждом этапе этой процедуры. Обработка была завершена в январе 1991 года.

Источник: Содержимое базы данных DCW только для Атлантического региона основывается на объектовом составе комплекта JNC в масштабе 1:2000000, подготовленного Картографическим управлением Министерства обороны США в период с 1974 г. по 1991 г.

Этап процедуры: См. «Этап процедуры» для комплекта ONC

Источник: Информация об аэропортах для DCW была получена из файла цифровой информации об авиарейсах (Digital Aeronautical Flight Information File) DAFIF, подготовленного Картографическим управлением Министерства обороны США и опубликованного в 1991 г. DAFIF включал информацию об аэропортах, содержащую: (1) название, (2) ICAO, (3) положение, (4) высоту над уровнем моря и (5) тип. DAFIF был выпущен на магнитной ленте.

Этап процедуры: DAFIF был оцифрован непосредственно в файлы VPF сотрудниками ESRI. Обработка была завершена в январе 1991 г.

Источник: Данные радиометра AVHRR (Advance Very High Resolution Radimeter), представленные EROS (EROS Data Centre), использовались для определения шести типов растительности, покрывающей континентальную часть США. Данные дистанционного зондирования представляют собой изображения в масштабе 1:000000 и системе WGS-84.

Этап процедуры: Каждодневно в течение двух недель на протяжении всего вегетационного периода в США AVHRR-изображения анализировались и усреднялись. Эти изображения, темпы их изменения, сведения о высоте и прочая информация использовались для получения единого классификационного изображения поверхности континентальной части США.

Этап процедуры: Растровые файлы изображений были преобразованы в векторные полигоны, аппроксимированы (удалены неровности), «утончены» (все полигоны площадью менее 2 км^2 удалены) и привязаны к существующим в DCW полигонам (водные поверхности, участки застройки). Результирующий файл был фрагментирован и конвертирован в слой растительности в формате VPF. Вся разработка выполнялась ESRI и была завершена в январе 1991 г.

B.2 Пример 2 — Цифровая карта местности

(Digital Terrain Map — DTM)

B.2.1 Общее представление

Пример иллюстрирует оценку разработчиком данных соответствующей количественной информации о качестве определением применимых элементов и подэлементов качества данных для области определения качества данных, представляющей собой набор данных, с использованием спецификации продукта.

Пример не включает фактическую выдачу соответствующей информации о качестве в виде метаданных или использование отчета об оценке качества.

B.2.2 Базовая информация

Набор данных: Гидрографическая цифровая карта бассейна

Спецификация продукта: Спецификации ГИС для использования вместе с гидрографическими планами бассейнов. Национальный институт водных ресурсов, 1998 г.

Описание продукта: DTM должна быть подготовлена с использованием топографической карты в масштабе 1:25000, созданной Национальным картографическим управлением и имеющей прямоугольную сетку координат. Размер ячеек должен быть равен 25 м, а расстояние от начала координат такой ячейки должно быть кратным 25 м.

DTM должна обеспечивать возможность гидрологического моделирования.

DTM может быть разделена на разные файлы с использованием гидрографических пределов бассейна и границ страны (информация предоставлена Национальным институтом водных ресурсов).

Примечание — Разработчик данных установил единственную область определения качества данных для рассматриваемого набора. Разработчик данных не нашел все элементы и подэлементы качества данных применимыми (в Спецификации продукта нет непосредственной ссылки на элементы и подэлементы качества данных). В таблице B.3 перечислены только применимые элементы и подэлементы качества данных.

Таблица В.3 — Соответствие количественной информации о качестве

Пример 2

Содержание соответствующих пунктов спецификации продукта	Применимый элемент (подэлемент) качества данных
DTM может быть разделена на разные файлы с использованием гидрографических пределов бассейна и границ страны (информация предоставлена Национальным институтом водных ресурсов)	Полнота Присутствие
DTM должна полностью охватывать соответствующую гидрографическую площадь своего бассейна	Полнота Отсутствие
В спецификации продукта нет ссылки на это требование, тем не менее разработчик данных указал на необходимость проверки, что в наборе данных нет высот более 2000 м	Логическая согласованность Доменная согласованность
Размер ячеек координатной сетки должен быть равен 25 м, а расстояние от начала координат сетки должно быть кратным 25 м. DTM должна подходить для гидрологического моделирования	Логическая согласованность Согласованность по формату
DTM должна обеспечивать гидрологическое моделирование	Логическая согласованность Топологическая согласованность
Спецификация продукта не содержит прямой ссылки на позиционную точность. Можно предположить, однако, что, поскольку использовались данные с топографической карты в масштабе 1:25000, погрешности высоты над уровнем моря должны быть менее 4 м	Позиционная точность Абсолютная или внешняя точность

Составленная разработчиком данных сводка соответствующей количественной информации о качестве приведена в таблице В.4.

Таблица В.4 — Сводка соответствующей количественной информации о качестве

Пример 2

Элемент качества данных	Подэлемент качества данных	Соответствие
Полнота	Присутствие Отсутствие	Да »
Логическая согласованность	Концептуальная согласованность Согласованность по формату Топологическая согласованность	Нет Да »
Позиционная точность	Абсолютная или внешняя точность Относительная или внутренняя точность Точность позиционирования относительно координатной сетки	Да Нет »
Временная точность	Точность временных измерений Согласованность по времени Временное соответствие	Нет » »
Тематическая точность	Правильность классификации Правильность количественных атрибутов Точность количественных атрибутов	Нет » »

В.3 Пример 3 — Набор данных о землепользовании

В.3.1 Общее представление

Пример иллюстрирует оценку разработчиком данных соответствующей количественной информации о качестве определением применимых элементов и подэлементов качества данных для области определения качества данных, представляющей собой набор данных, с использованием Спецификации продукта.

Пример не включает фактическую выдачу соответствующей информации о качестве в виде метаданных или использование отчета об оценке качества.

Пример включает компиляцию соответствующей неколичественной информации о качестве.

В.3.2 Базовая информация

Набор данных: Набор данных о землепользовании, состоящий из нескольких поднаборов (по одному на округ), подготовлен для поддержки общих проектов географических информационных систем

Спецификация продукта: Требования к продукту для создания в цифровой форме копий карт землепользования, подготовленных в аналоговом формате и утвержденных согласно установленным законом процедурам.

Описание продукта: Набор данных содержит поднаборы данных о землепользовании, созданные в результате оцифровки исходного картографического материала и дополненные статистическими данными.

Каждый поднабор соответствует какому-то округу, имеющему официально установленные границы, данные о которых были представлены разработчику.

Объекты каждого поднабора — это полигоны, классифицированные в соответствии с системой классификации объектов землепользования.

В.3.3 Оценка соответствующей количественной информации о качестве

Примечание — Разработчик данных установил единственную область определения качества данных для рассматриваемого набора данных. Разработчик данных не нашел все элементы качества данных и подэлементы качества данных применимыми (в спецификации продукта нет непосредственной ссылки на элементы качества данных и подэлементы качества данных). В таблице В.5 перечислены только применимые элементы и подэлементы качества данных.

Таблица В.5 — Соответствие количественной информации о качестве

Пример 3

Содержание соответствующих пунктов спецификации продукта	Применимый элемент (подэлемент качества данных)
Число пропущенных полигонов; площадь, соответствующая пропущенным полигонам; выявление неполного покрытия площади, определяемой границами округов, из-за неправильного задания границ округов	Полнота Присутствие
Идентификация дорог должна соответствовать именам, приведенным в предоставленном файле	Полнота Отсутствие
Объектами в поднаборах являются замкнутые контуры, закодированные в соответствии с принятой системой кодирования	Логическая согласованность Доменная согласованность
Расхождения между векторами свыше 1 мм в масштабе исходного материала не должны превышать 10 %	Позиционная точность Абсолютная или внешняя точность

Составленная разработчиком данных сводка соответствующей количественной информации о качестве приведена в таблице В.6.

Таблица В.6 — Сводка соответствующей количественной информации о качестве

Пример 3

Элемент качества данных	Подэлемент качества данных	Соответствие
Полнота	Присутствие Отсутствие	Да »
Логическая согласованность	Концептуальная согласованность Доменная согласованность Согласованность по формату Топологическая согласованность	Нет Да Нет »
Позиционная точность	Абсолютная точность Относительная точность Точность позиционирования относительно координатной сетки	Да Нет »
Временная точность	Точность временных измерений Согласованность по времени Временное соответствие	Нет » »
Тематическая точность	Правильность классификации Правильность количественных атрибутов Точность количественных атрибутов	Нет » »

Состав неколичественной информации о качестве

Назначение: Набор данных о землепользовании предназначен для поддержки управления землепользованием и его планирования; он должен как можно в большей степени соответствовать оригинальному аналоговому формату, утвержденному законным порядком.

Использование: Для управления землепользованием в каждом округе использовались свои исходные картографические материалы.

Происхождение: Источник: Карты землепользования в аналоговом формате, составленные с использованием в качестве основы топографических карт масштабов 1:25000 и 1:10000.

Этап процедуры: Исходные материалы сканировались. Изображения привязывались с использованием не менее девяти точек с известными координатами.

В.4 Пример 4 — Трехмерная база данных сети дорог

(A 3 Dimensional (3 D) Road Network Database)

В.4.1 Общее представление

Пример иллюстрирует оценку разработчиком данных соответствующей количественной информации о качестве определением применимых элементов и подэлементов качества данных для области определения качества данных, представляющей собой набор данных, с использованием спецификации продукта.

Пример не включает фактическую выдачу соответствующей информации о качестве в виде метаданных или использование отчета об оценке качества.

В.4.2 Базовая информация

Набор данных: Векторная трехмерная база данных сети дорог. Набор данных предназначен для использования в качестве источника для создания новых баз данных, таких как (1) двумерная база данных сети дорог с атрибутами, характеризующими состояние дорожного покрытия (со ссылками на расстояние, измеренное в пространстве (3D), от начала каждого сегмента), и (2) двумерная база данных сети дорог со ссылками на скорости и (3) профили дорог.

Спецификация продукта: Требования к продукту, содержащему набор трехмерных векторных ломаных линий, описывающих сеть национальных дорог.

Описание продукта: Каждая пространственная ломаная линия соответствует дуге в сети дорог. Каждая дуга запоминается в отдельном слое под именем, соответствующим ее идентификации. Идентификация дорог должна выполняться согласно именам, перечисляемым в файле, предоставленном разработчику данных.

Позиционные погрешности, как плановые, так и высотные, не должны превышать 2 м. Трехмерные ломаные линии должны соединяться друг с другом в точках пересечения и расходиться только в точках пересечения или в тех случаях, когда изменяется идентификация.

В.4.3 Оценка соответствующей количественной информации о качестве

Примечание — Разработчик данных установил единственную область определения качества данных для рассматриваемого набора. Разработчик данных не нашел все элементы и подэлементы качества данных применимыми (в Спецификации продукта нет непосредственной ссылки на элементы качества данных и подэлементы качества данных). В таблице В.7 приведены только применимые элементы и подэлементы качества данных.

Таблица В.7 — Соответствие количественной информации о качестве

Пример 4

Содержание соответствующих пунктов спецификации продукта	Применимый элемент качества данных (подэлемент качества данных)
Никакого особого требования; ссылка на полноту, отсутствие полноты	Полнота Присутствие
Набор данных должен содержать все дороги, классифицируемые как Национальные дороги	Полнота Отсутствие
Идентификация дорог должна выполняться в соответствии с именами, перечисляемыми в файле, предоставленном разработчику данных	Логическая согласованность Доменная согласованность
Трехмерные ломаные линии должны соединяться друг с другом в точках пересечения и расходиться только в точках пересечения или в тех случаях, когда изменяется идентификация	Логическая согласованность Топологическая согласованность
Трехмерные ломаные линии должны соединяться друг с другом в точках пересечения и расходиться только в точках пересечения или в тех случаях, когда изменяется идентификация	Позиционная точность Абсолютная точность

Составленная разработчиком данных сводка соответствующей количественной информации о качестве приведена в таблице В.8.

Таблица В.8 — Сводка соответствующей количественной информации о качестве

Пример 4

Элемент качества данных	Подэлемент качества данных	Соответствие
Полнота	Присутствие Отсутствие	Да »
Логическая согласованность	Концептуальная согласованность Доменная согласованность Согласованность по формату Топологическая согласованность	Нет Да Нет Да
Позиционная точность	Абсолютная точность Относительная точность Точность позиционирования относительно координатной сети	Да Нет »
Временная точность	Точность временных измерений Согласованность по времени Временное соответствие	Нет » »
Тематическая точность	Правильность классификации Правильность неколичественных атрибутов Точность количественных атрибутов	Нет » »

УДК 622.1:528:002:006.354 ОКС 35.240.70 Т43

Ключевые слова: географическая информация, набор данных, принципы оценки качества данных, концепция оценки качества данных, элементы качества данных, подэлементы качества данных, количественная информация о качестве, описательная информация о качестве, формирование информации о качестве, выдача информации о качестве, комплекс проверок качества набора данных

Редактор *Р.С. Федорова*
Технический редактор *В.Н. Прусакова*
Корректор *М.В. Бучная*
Компьютерная верстка *Л.А. Круговой*

Изд. лиц. № 02354 от 14.07.2000. Сдано в набор 22.01.2004. Подписано в печать 25.02.2004. Усл. печ. л. 3,26. Уч.-изд. л. 2,80. Тираж 200 экз. С 921. Зак. 210.

ИПК Издательство стандартов, 107076 Москва, Колодезный пер., 14.
http://www.standards.ru e-mail: info@standards.ru
Набрано в Издательстве на ПЭВМ
Отпечатано в филиале ИПК Издательство стандартов — тип. «Московский печатник», 105062 Москва, Лялин пер., 6.
Плр № 080102