



**ФЕДЕРАЛЬНАЯ СЛУЖБА ГЕОДЕЗИИ И КАРТОГРАФИИ  
РОССИИ**

**ГОСУДАРСТВЕННЫЙ НАУЧНО-ВНЕДРЕНЧЕСКИЙ ЦЕНТР  
ГЕОИНФОРМАЦИОННЫХ СИСТЕМ И ТЕХНОЛОГИЙ  
(ГОСГИСЦЕНТР)**

## **СТАНДАРТ ОТРАСЛИ**

**КАРТЫ ЦИФРОВЫЕ  
МЕТОДЫ ОЦЕНКИ КАЧЕСТВА ДАННЫХ.  
ОБЩИЕ ТРЕБОВАНИЯ**

**Москва  
ЦНИИГАиК  
2003**

**СТАНДАРТ ОТРАСЛИ**

**КАРТЫ ЦИФРОВЫЕ**  
**МЕТОДЫ ОЦЕНКИ КАЧЕСТВА ДАННЫХ.**  
**ОБЩИЕ ТРЕБОВАНИЯ**

**Москва**  
**ЦНИИГАиК**  
**2003**

## Предисловие

**1 РАЗРАБОТАН** и внесен на утверждение Государственным научно-внедренческим центром геоинформационных систем и технологий (Госгисцентр)

|                       |                |
|-----------------------|----------------|
| Директор,             |                |
| руководитель темы     | В.Г. Плешков   |
| Заместитель директора | А.В. Рогачев   |
| Начальник отдела      | Т.Н. Максимова |
| Начальник сектора     | Л.Я. Лимонтов  |

**2 ВНЕСЕН** НТУ Роскартографии

|               |                  |
|---------------|------------------|
| Начальник НТУ | В.Н. Александров |
|---------------|------------------|

**3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ** Приказом Роскартографии от 31 марта 2003 г. № 58-пр.

**4 ВВЕДЕН ВПЕРВЫЕ**

© Роскартография, 2003

## Содержание

|  |    |
|--|----|
| 1. Область применения  | 1  |
| 2. Нормативные ссылки  | 2  |
| 3. Определения   | 2  |
| 4. Общие положения   | 3  |
| 5. Метод оценки качества данных<br>изготовителем   | 4  |
| 6. Особенности метода оценки<br>качества данных пользователем                                      | 17 |
| 7. Общие требования к представлению<br>информации о результатах оценки<br>качества данных          | 22 |
| Приложение А (справочное).<br>Термины и определения, необходимые<br>для понимания текста стандарта | 25 |
| Приложение Б (информационное)<br>Классификация методов оценки<br>качества данных                   | 29 |
| Приложение В (информационное)<br>Оценка качества данных с использованием<br>выборки                | 32 |
| Приложение Г (обязательное)<br>Формирование отчетов о результатах<br>оценки качества данных        | 38 |
| Приложение Д (информационное)<br>Обобщение результатов оценки<br>качества данных                   | 45 |
| Приложение Е (обязательное)<br>Формирование метаданных о<br>качестве данных                        | 49 |
| Приложение Ж (справочное)<br>Библиография  | 53 |

**СТАНДАРТ ОТРАСЛИ**

---

**КАРТЫ ЦИФРОВЫЕ**

Методы оценки качества данных. Общие требования.

---

Дата введения 2003-06-01

**1 Область применения**

Настоящий стандарт распространяется на цифровые карты и планы (далее – цифровые карты) различного вида и масштаба, изготавливаемые по лицензии федерального органа исполнительной власти по геодезии и картографии.

Настоящий стандарт устанавливает методы оценки качества наборов данных, составляющих содержание цифровых карт (ЦК), последовательность выполнения операций оценки качества и общие требования к представлению результатов оценки качества.

Настоящий стандарт применяется при:

- оценке качества наборов данных ЦК их изготовителями и пользователями;
- согласовании между изготовителями и пользователями ЦК требований к качеству их данных;
- оценке качества наборов данных в процессе сертификации ЦК;
- разработке программного обеспечения для оценки качества наборов данных ЦК.

Настоящий стандарт должен применяться совместно с ОСТ 68-3.4.1.

## 2 Нормативные ссылки

В настоящем стандарте использованы ссылки на следующие стандарты:

|                                      |   |
|--------------------------------------|---|
| ГОСТ 15467 – 79                      | Управление качеством продукции. Основные понятия. Термины и определения     |
| ГОСТ 15895 – 77<br>(СТ СЭВ 547 – 77) | Статистические методы управления качеством продукции. Термины и определения |
| ГОСТ 15971-90                        | Системы обработки информации. Термины и определения                         |
| ГОСТ 20886-85                        | Организация данных в системах обработки данных. Термины и определения       |
| ГОСТ 28441-99                        | Картография цифровая. Термины и определения                                 |
| ГОСТ Р 51608-2000                    | Карты цифровые топографические. Требования к качеству                       |
| ОСТ 68-3.4.1-02                      | Карты цифровые. Оценка качества данных. Основные положения                  |

## 3 Определения

3.1 В настоящем стандарте использованы термины по ГОСТ 15467, ГОСТ 15895, ГОСТ 15971, ГОСТ 20886, ГОСТ 28441, термины, определения к которым даны в

ОСТ 68-3.4.1, а также термины и определения понятий, приведенные в Приложении А настоящего стандарта.

Часть терминов, принятая в данном стандарте, согласована с терминами и определениями в проектах стандартов ИСО 19113 «Географическая информация. Принципы оценки качества» и ИСО 19114 «Географическая информация. Процедуры оценки качества».

## 4 Общие положения

4.1 Настоящий стандарт регламентирует два основных направления оценки качества данных ЦК:

- оценка качества набора данных изготовителем;
- оценка качества набора данных пользователем.

*Примечания*

*1 Вновь создаваемый набор данных может включать в себя исходные наборы данных, прошедшие проверку по той или иной программе оценки качества. Независимо от этого изготовитель данных во время оценки исходных наборов должен установить соответствие качества данных нормативным требованиям.*

*2 Если пользователь в соответствии со спецификой своих потребностей вносит изменения в набор данных, который предполагается к использованию и другими пользователями, он обязан выполнить оценку качества набора по программе изготовителя данных.*

4.2 Настоящий стандарт рассматривает оценку качества набора данных как многоступенчатый контроль, включающий в себя ряд этапов в соответствии с 4.4 ОСТ 68-3.4.1.

Для снижения объёма продукции, бракуемой по завершении работ, промежуточные этапы контроля (контроль качества во время создания или обновления набора данных ЦК) должны осуществляться как контроль по альтернативному признаку.

4.3 Операции и процедуры оценки качества данных, устанавливаемые настоящим стандартом, не зависят от методов, технологий и технических средств получения данных ЦК.

4.4 Процесс оценки качества набора данных ЦК должен осуществляться в соответствии с положениями ОСТ 68-3.4.1. Отчётная документация о выполнении указанного процесса должна обеспечивать возможность определения достоверности оценки качества набора данных ЦК.

## **5 Метод оценки качества данных изготовителем**

Оценка качества набора данных ЦК должна осуществляться по предварительно составленной программе, которая может быть оформлена как раздел технического проекта по созданию (обновлению, модификации) ЦК, либо как самостоятельный технический проект по оценке качества ЦК. Программа должна включать в себя следующие операции:

- указание областей определения качества данных;
- выбор методов оценки качества данных для каждой из указанных областей определения качества и способов для их проведения;
- определение значений количественных показателей и мер качества данных;



- указание уровня качества;
- анализ результатов оценки качества данных.

Стандартной является ситуация, при которой оценка качества набора данных ЦК производится по всем аспектам, определяемым ОСТ 68-3.4.1. В том случае, когда оценка по какому-либо аспекту не выполняется, отчёт о результатах оценки качества набора данных должен содержать обоснование этого положения.

Этап, на котором выполняется оценка качества набора данных по тому или иному аспекту, определяется в период составления технического проекта.

*Пример – Оценка качества набора данных по аспекту «актуальность данных» целесообразна на этапе анализа качества исходных данных, так как показатель «соответствие нормативному сроку обновления» определяет актуальность и исходных данных, и конечного продукта – цифровой карты.*

## **5.1 Указание областей определения качества данных**

Областями определения качества данных для каждого из аспектов, указанных в ОСТ 68-3.4.1, являются:

- в процессе создания ЦК и по завершении этого процесса - весь набор данных ЦК;
- в процессе обновления ЦК – группа вновь введённых данных, включая данные по согласованию новых и ранее существовавших объектов, их свойств и отношений.

Области подразделяются на подобласти в соответствии с принципом «структуризации результатов оценки качества данных», описанным в 5.4.6 ОСТ 68-3.4.1.

*Примеры*

*1 При оценке качества семантических данных производится выделение подобластей определения качества для оценки однородных групп данных с учётом важности тех или иных объектов ЦК.*

*2 При оценке качества метрических данных производится выделение подобластей определения качества для следующих групп данных:*

- объекты, используемые в качестве планово-высотной основы;*
- объекты с четкими контурами;*
- объекты, описывающие растительный покров и грунты;*
- горизонтали и другие объекты, положение которых определяется с использованием метода интерполяции.*

## **5.2 Выбор методов оценки качества данных**

Методы оценки качества данных описываются в Приложении Б.

Для каждой из указанных областей (подобластей) определения качества изготовитель должен использовать один из следующих методов прямой оценки:

- метод внутренней оценки;
- метод внешней оценки.

Недопустимо использование изготовителем метода косвенной оценки данных.

Методы оценки качества данных устанавливаются для каждой из областей (подобластей) определения качества индивидуально. Установление метода оценки качества дан-

ных для каждой из областей (подобластей) определения качества должно выполняться на основе возможности использования автоматического контроля, автоматизированного контроля или контроля только при наличии дополнительных данных.

При создании или периодическом обновлении каждый набор данных ЦК должен быть проконтролирован по программе полной проверки в процессе создания (обновления) и по завершении этого процесса. Полная проверка требует исследования всего набора данных в целом. Каждая из указываемых областей определения качества данных должна включать в себя все группы данных, которые могут быть оценены с использованием применяемых показателей и мер качества данных.

При оценке качества данных в составе набора применяются как полный (100%) контроль, так и контроль по выборке данных.

100% контроль применяется для оценки качества данных в составе набора, которые могут быть проконтролированы по какому-нибудь аспекту (аспектам) в автоматическом режиме.

*Пример – По аспекту «синтаксическая корректность» должны быть проконтролированы все данные набора данных ЦК.*

Контроль по выборке используется для оценки качества данных в составе набора, которые могут быть проконтролированы по какому-нибудь аспекту (аспектам) только в автоматизированном режиме или с использованием дополнительных данных.

Объём выборки должен гарантировать установленный нормативными или техническими требованиями уровень достоверности результатов оценки качества данных ЦК.

Оценка качества данных с использованием выборки описывается в Приложении В.

### **5.3 Определение значений количественных показателей и мер качества данных**

В рамках требований, определяемых каждым из аспектов качества, регламентируемых ОСТ 68-3.4.1, изготовитель данных должен использовать такой комплекс количественных показателей, которые охватывают все параметры набора данных, влияющие на уровень его качества. Так как количественный показатель качества данных является комплексной оценкой данных, значение его определяется с использованием одной или более мер качества данных.

Ниже описан стандартный комплекс количественных показателей, критериев и мер качества данных, используемых для оценки качества набора данных ЦК.

*Примечание - Показатели и меры качества, включающие в себя такие понятия, как «число» или «количество», должны сопровождаться сведениями об объёме данных, проверенных с использованием указанной меры. Альтернативным вариантом является указание меры в относительных значениях, например, в процентах.*

#### **5.3.1 Оценка синтаксической корректности**

При оценке набора данных ЦК по аспекту «синтаксическая корректность» все данные следует считать однородными.

Критерий оценки качества: соответствие представления набора данных используемому формату.

Мера качества: несоответствие представления набора данных используемому формату.

*Примечание - При обнаружении любого числа нарушений синтаксической корректности изготовитель должен присвоить набору данных ЦК оценку качества «бракуется», определить источник нарушения и заново создать элементы ЦК, в которых были обнаружены нарушения.*

### 5.3.2 Оценка правильности логической структуры

При оценке набора данных ЦК по аспекту «правильность логической структуры» все данные следует считать однородными.

Критерий оценки качества: соответствие набора данных нормативно установленной логической структуре.

Мера качества: число отсутствующих типов элементов данных.

Мера качества: число избыточных типов элементов данных.

Мера качества: число неправильно присвоенных типов элементов данных.

*Пример – Вместо типа «точечный нестандартно ориентированный объект» элементу данных присвоен тип «линейный объект».*

Мера качества: число нарушений порядка размещения элементов данных различных типов.

*Пример – Элемент данных с типом «значение характеристики» предшествует элементу с типом «количественная характеристика».*

### 5.3.3 Оценка актуальности данных

При оценке набора данных ЦК по аспекту «актуальность» его содержание следует считать неоднородным и подлежащим разделению на группы однородных данных.

*Примеры*

*1 С учётом положений [1] данные могут быть сгруппированы на основе категорий важности объектов местности.*

*2 Пользователь набора данных ЦК может отнести к однородным группам данных объекты, необходимые и вспомогательные для решения его специфических задач.*

Критерий оценки качества: соответствие нормативному сроку обновления.

Мера качества: число изменений на местности, не описанных содержанием ЦК в установленные сроки.

### 5.3.4 Оценка соответствия состава данных

При оценке набора данных ЦК по аспекту «соответствие состава» его содержание следует считать неоднородным и подлежащим разделению на группы однородных данных.

#### 5.3.4.1 Оценка соответствия состава справочных данных

Критерий оценки качества: наличие всех элементов справочных данных.

Мера качества: число пропущенных элементов справочных данных.

Мера качества: число избыточных элементов справочных данных.

*Примечание – Так как справочные данные используются для контроля качества других данных набора ЦК, отсутствие каких-либо элементов этих данных недопустимо.*

#### 5.3.4.2 Оценка соответствия состава метрических данных

Критерий оценки качества: отсутствие или дублирование метрических данных в составе объекта.

Мера качества: число пропущенных метрических данных.

Мера качества: число избыточных метрических данных.

#### 5.3.4.3 Оценка соответствия состава семантических данных

Критерий оценки качества: отсутствие или дублирование семантических данных в составе объекта.

Мера качества: число пропущенных семантических данных.

Мера качества: число избыточных семантических данных.

### 5.3.5 Оценка логической согласованности данных

При оценке набора данных ЦК по аспекту «логическая согласованность» его содержание следует считать неоднородным и подлежащим разделению на группы однородных данных.

#### 5.3.5.1 Оценка согласованности справочных данных

Критерий оценки качества: Взаимная согласованность значений элементов справочных данных.

Мера качества: число элементов справочных данных с противоречивыми значениями.

#### 5.3.5.2 Оценка согласованности метрических данных

Критерий оценки качества: согласованность взаимного положения объектов в плане.

Мера качества: число взаимно несогласованных пар объектов.

Критерий оценки качества: согласованность описания отдельных частей как линейных объектов, так и границ площадных объектов.

Мера качества: число нарушений согласованности.

Критерий оценки качества: согласованность взаимного положения объектов по высоте.

Мера качества: число ошибок согласования положения объектов по высоте.

Критерий оценки качества: согласованность описания положения объектов с границей ЦК.



Мера качества: число ошибок в положении объектов по отношению к границе ЦК.

#### 5.3.5.3 Оценка согласованности семантических данных

Критерий оценки качества: правильность идентификации объектов, характеристик, значений характеристик и отношений (отсутствие идентификации, недопустимая идентификация, повторная идентификация).

Мера качества: число ошибок идентификации.

Критерий оценки качества: согласованность идентификации по границам со смежными наборами данных.

Мера качества: число несогласованной идентификации элементов с данными смежных наборов.

#### 5.3.5.4 Оценка согласованности метрических и семантических данных

Критерий оценки качества: согласованность семантических данных о размерах объектов с характером локализации этих объектов.

Мера качества: число несогласующихся данных.

#### 5.3.6 Оценка точности данных

При оценке набора данных ЦК по аспекту «точность данных» его содержание следует считать неоднородным и подлежащим разделению на группы однородных данных.

### 5.3.6.1 Оценка точности справочных данных

Критерий оценки качества: точность указания значений элементов справочных данных.

Мера качества: число ошибочных значений.

### 5.3.6.2 Оценка точности метрических данных

Критерий оценки качества: точность абсолютного положения объектов в плане.

Мера качества: средние квадратические погрешности положения объектов в плане.

Мера качества: максимальные погрешности положения объектов в плане.

Критерий оценки качества: точность абсолютного положения объектов по высоте.

Мера качества: средние квадратические погрешности положения объектов по высоте.

Мера качества: максимальные погрешности положения объектов по высоте.

Критерий оценки качества: абсолютная точность взаимного положения объектов в плане.

Мера качества: средние квадратические погрешности расстояний.

Мера качества: максимальные погрешности расстояний.

*Примечание – Показатели «точность абсолютного положения объектов в плане» и «абсолютная точность взаимного положения объектов в плане» являются альтернативными, поэтому в процессе оценки качества должен быть выбран один из них в зависимости от наличия тех или иных внешних данных.*

Критерий оценки качества: точность представления рельефа местности.

Мера качества: средние квадратические погрешности представления высот.

Мера качества: максимальные погрешности представления высот.

Критерий оценки качества: точность представления физических полей.

Мера качества: средние квадратические погрешности представления физических полей

Мера качества: максимальные погрешности представления физических полей.

### 5.3.6.3 Оценка точности семантических данных

Критерий оценки качества: точность значений характеристик объектов.

Мера качества: число ошибочных значений характеристик объектов.

## 5.4 Указание уровня качества

Для выполнения заключительного этапа оценки качества набора данных ЦК необходимо на основе действующих нормативных или технических требований на ЦК различного вида и масштаба указать допустимые уровни качества данных для всех определяемых показателей качества.

*Пример - Уровень качества цифровых топографических карт, при создании которых в качестве исходных картографических материалов используются топографические карты тех же масштабов, устанавливают на основе ГОСТ Р 51608.*

При определении соответствия между результатами оценки качества и нормативными уровнями допустимых погрешностей следует учитывать, что не для каждой из мер качества должен существовать соответствующий ей нормативный уровень. В этом случае необходимо формировать комплексный показатель качества, для которого такой уровень установлен.

*Пример - Каждая из мер качества «число пропущенных метрических данных» и «число избыточных метрических данных» не имеют индивидуально установленного допустимого уровня. Для оценки качества набора данных с помощью указанных мер определяется комплексный показатель «соответствие состава метрических данных» по формуле:*

$$S = m + n,$$

*где S – соответствие состава метрических данных,  
m – число пропущенных метрических данных,  
n – число избыточных метрических данных.*

*Примечание - Следует учитывать, что метод оценки может влиять на уровень качества. Например, если используется контроль по выборке, то представительность выборки является одним из факторов, который следует учитывать при установлении уровня качества.*

## 5.5 Анализ результатов оценки качества данных

Значение показателя качества представляется в обобщенном виде в соответствии с 5.4 или таблицы значений. Например, таблицы распределения погрешностей планового положения объектов. Наряду со значениями комплексных показателей качества в отчет о результатах оценки качества данных помещается результат сравнения этих

значений с допустимыми уровнями погрешностей данных. Качество данных считается соответствующим установленному уровню качества, если полученные в результате анализа данных абсолютные значения перечисленных показателей принадлежат области допустимых значений.

Завершающим стандартным отображением результата сравнения по каждому из показателей является запись «принимается» или «бракуется».

В связи с тем, что объём отчёта о результатах оценки качества набора данных ЦК является, как правило, избыточным и для пользователей наборов данных, и для сведений, помещаемых в отраслевую (межотраслевую) базу метаданных, изготовитель данных должен выполнить процедуру обобщения данных о качестве набора данных ЦК.

Результат обобщения данных о качестве набора данных формируется в соответствии с положениями Приложения Е.

## **6 Особенности метода оценки качества данных пользователем**

В рамках своей деятельности пользователь наборов данных ЦК должен использовать программы оценки качества данных для изготовителя в тех случаях, когда им производится изменение наборов данных с перспективой дальнейшего их использования другими пользователями.

Для оценки того, в какой мере качество набора данных ЦК удовлетворяет требованиям конкретного пользователя, допускается использование программ оценки качества как изготовителей данных, так и разработанных самим пользователем. В

рамках последних допускается использование косвенного метода оценки качества набора данных.

### **6.1 Указание областей определения качества данных, исследуемых аспектов, применяемых показателей и мер качества данных**

При оценке качества данных пользователем допускается использование двух стратегий оценки:

- стратегии, основанной на упорядоченности областей определения;
- стратегии, основанной на упорядоченности показателей.

При осуществлении первой стратегии на первом этапе поочередно выбираются различные области определения качества. На втором этапе для очередной области последовательно применяются все показатели качества из числа возможных, после чего осуществляется переход к следующей области определения качества.

При реализации второй стратегии на первом этапе осуществляется последовательный перебор показателей качества. На втором этапе выбранный показатель качества применяется, если это возможно, к поочередно выбираемым областям определения качества, после чего осуществляется переход к следующему показателю качества.

Выбор стратегии, наиболее отвечающей целям оценки качества, осуществляется специалистом-исполнителем, выполняющим анализ качества набора данных. Не допускается в пределах одного описания качества данных смешивание двух стратегий.

При разработке собственной программы оценки качества набора данных ЦК пользователь должен определить, для ка-

ких аспектов будет применен прямой, а для каких - косвенный методы оценки данных в соответствии с Приложением Б.

Для аспектов, по которым оценка данных выполняется на основе прямого метода оценки, используя количественные показатели и меры качества из числа приведенных в 5.3. В случае необходимости учёта специфических требований к набору данных допустимо введение показателей и мер качества, не указанных в 5.3.

При выборе мер качества, определяемых пользователем, должно быть проведено исследование на применимость выбранных мер оценки качества данных. Если выбранная пользователем мера неприменима, он должен определить возможность замены её мерами качества из состава мер, указанных в 5.3.

*Пример - Пользователем в состав набора данных введена новая группа площадных объектов. В результате этой операции возникают следующие ситуации.*

*1 Если отсутствуют внешние данные об этих объектах, то невозможно определить значение меры точности их положения.*

*2 Если известно, что границы вновь введённых объектов должны совпадать с границами объектов, которые находились в исходном наборе данных, то следует воспользоваться мерой качества, которая содержит число и значения погрешностей в согласовании планового положения вышеуказанных объектов.*

## **6.2 Косвенный метод оценки качества наборов данных**

Использование косвенного метода основано на положении о тождественности качества данных, имеющих одинаковое происхождение и назначение.

Для численного выражения результатов оценки качества данных по указанным показателям должны использоваться статистические данные о результатах практического использования наборов данных ЦК для решения задач, аналогичных тем, которые будут решаться на основе оцениваемого набора данных.

Настоящим стандартом для определения достоверности оценки качества нового набора данных устанавливаются следующие меры качества данных:

- вероятность соответствия качества по показателю “назначение”,
- вероятность соответствия качества по показателю “происхождение” и единый способ определения их численных значений, заключающийся в следующем.

Для показателя «назначение» («происхождение») данных из совокупности статистических данных об использовании наборов данных осуществляется выборка сведений об успешности использования наборов данных для решения аналогичных задач, назначение (происхождение) которых совпадает со значением анализируемого показателя (или достаточно близко к нему). При однократном использовании всех принадлежащих выборке наборов данных значение вероятности соответствия качества по соответствующему показателю вычисляется по формуле:



$$p = \frac{m}{m + n},$$

где  $p$  – значение вероятности соответствия качества по выборке,

$m$  – число наборов данных, соответствие уровню качества которых установлено успешным применением;

$n$  – число наборов данных, несоответствие уровню качества которых установлено в результате применения.

При многократных применениях хотя бы одного набора данных для решения аналогичных задач значение вероятности  $p$  соответствия качества необходимому уровню вычисляется по формуле, являющейся обобщением приведенной выше,

$$p = \frac{\sum_{i=1}^k m_i}{\sum_{i=1}^k (m_i + n_i)},$$

где  $k$  – число наборов данных в выборке;

$m$  – число успешно решенных задач с применением  $i$ -го набора данных;

$n$  – число нерешенных задач с применением  $i$ -го набора данных.

### 6.3 Оценка соответствия данных требованиям пользователя

Пользователь данных для каждой из применимых мер (комплексных показателей, вероятностей соответствия) качества должен указать с учётом положений, закреплённых

в его технической документации, требуемый уровень качества набора данных ЦК. В качестве параметров, определяющих уровни качества, должны указываться предельные значения показателей, допустимые с учетом тех или иных аспектов качества данных.

Для определения соответствия качества набора данных потребностям пользователя указываемые уровни качества должны сравниваться со значениями мер (комплексных показателей) качества, представленными в метаданных или в отчете о результатах оценки качества набора данных.

Если качество данных для выбранного аспекта изготовителем данных не сообщается, пользователь должен определить произвести оценку качества данных по самому набору данных, применяя прямой или косвенный методы оценки качества.

Полученные пользователем результаты оценки качества набора должны быть помещены в метаданные или отчет о качестве набора данных с указанием, для решения каких задач ориентировано его использование.

## **7 Общие требования к представлению информации о результатах оценки качества данных**

Настоящий стандарт определяет две формы отчетности о результатах оценки качества набора данных ЦК:

- отчет о результатах оценки качества набора данных;
- метаданные оценки качества набора данных.

7.1 Отчет о результатах оценки качества формируется изготовителем данных ЦК в обязательном порядке при

создании новых наборов данных, а также в процессе периодического обновления.

Отчёт должен представлять результаты оценки качества единичного набора данных ЦК или совокупности наборов данных.

Представление результатов оценки качества данных ЦК в виде отчета должно основываться на иерархическом принципе («от общего – к частному» или «сверху – вниз»).

С целью сокращения объемов информации о качестве наборов данных ЦК при их представлении необходимо в максимально возможной степени использовать принцип наследования значений показателей качества данных.

Отчет о результатах оценки качества данных пользователем должен создаваться при следующих условиях:

- когда существуют накопленные результаты по качеству данных, а набор данных имеет достаточно универсальный характер, то есть, если конкретное его использование со специфическими требованиями еще не определено;
- если показатели качества данных должны иметь тип «определенные пользователем».

Отчет о результатах оценки качества может создаваться в любом другом случае, но не взамен представления в форме метаданных.

Компоненты отчета в зависимости от методов оценки качества данных приведены в Приложении Г.

7.2 Порядок формирования метаданных для оценки качества данных приведен в Приложении Е.

Для использования в пределах одной организации допускается представление пользователем метаданных не в полном объеме. При минимальном объеме метаданных ре-

зультаты по качеству данных сообщаются только в виде «принимается – бракуется» по отношению к их соответствию нормативному уровню качества.

7.3 Показатели качества набора данных ЦК карт могут храниться и выдаваться пользователям в виде элементов справочных данных ЦК либо в виде метаданных, сопровождающих наборы данных ЦК.

Для информационных сообщений о наличии тех или иных наборов данных ЦК достаточны метаданные, состоящие из единичной оценки качества, характеризующей набор в описательном или количественном виде.

*Примеры:*

*1 Набор данных ЦК удовлетворяет требованиям, предъявляемым нормативно-технической документацией.*

*2 Набор данных ЦК с вероятностью 99.9% не содержит погрешностей, выходящих за рамки предельно допустимых значений.*

## ПРИЛОЖЕНИЕ А

### (справочное)

#### Термины и определения, необходимые для понимания текста стандарта

Для целей настоящего стандарта применяются приведенные ниже термины и определения. Часть из них уже приводилась в ОСТ 68-3.4.1 и повторяются здесь для удобства.

**А.1 набор данных цифровой карты:** Идентифицированная согласованная совокупность данных об участке земной поверхности.

*Пример - Одно- или многофайловая структура, содержащая совокупность данных на номенклатурный лист топографической или специальной карты, или данные на планшет крупномасштабной съемки, или совокупность картографических данных о районе субъекта Российской Федерации, о квартале населенного пункта и т.д.*

**А.2 группа данных:** Подмножество того или иного набора данных цифровых карт, обладающее целостностью в тематическом и/или пространственном плане

**А.3 элемент данных:** Порция данных, в контексте использования которой отсутствует способ выделения из неё порции, отличной от неё самой (по ГОСТ 20886).

*Пример - Примером элемента данных в составе набора данных ЦК являются код объекта ЦК, значение характеристики объекта, координата точки и т.д.*

**А.4 однородные данные:** Данные, объединенные общностью одного или нескольких принадлежащих им свойств.

**А.5 неоднородные данные:** Данные, не обладающие общими для них свойствами.

**А.6 метрические данные:** Разновидность картографических данных, отражающая в определенной системе координат пространственное положение объектов местности.

**А.7 семантические данные:** Разновидность картографических данных, отражающая сущность и характеристики объектов местности.

**А.8 аспект качества данных:** Наиболее общее требование к качеству разнородных данных, выполнение которого является необходимым условием при использовании данных для решения тех или иных задач.

**А.9 критерий оценки качества данных:** Требование к свойствам множества однородных данных, сформулированное в виде однозначного условия, не допускающего различных трактовок

**А.10 показатель качества данных:** Характеристика одного или нескольких свойств однородных данных, обусловленных необходимостью использования этих данных для решения тех или иных задач.

**А.11 количественный показатель качества данных:** — Показатель качества, характеризующий данные цифровых карт с точки зрения соответствия предъявляемому к ним конкретному требованию в виде одной или нескольких количественно определенных мер качества данных.

*Примечание - Мера качества состоит из наименования, типа значения, единицы измерения и значения.*

*Пример – Типом значения меры качества может служить тип «логический», «целый», «вещественный» и др.*

**А.12 описательный показатель качества данных:** Метаданные, допускающие характеристику качества данных цифровых карт только в текстовом виде.

**А.13 область определения качества данных:** Подлежащее оценке качества множество данных, выбранное с учётом того или иного условия или набора условий.

*Примеры*

*1 Все точки в наборе данных.*

*2 Названия всех дорог в определенной географической зоне.*

*3 Набор данных номенклатурного листа цифровой топографической карты в целом.*

**А.14 уровень качества:** Пороговое значение или набор используемых пороговых значений, которые определяют, насколько хорошо набор данных удовлетворяет техническим требованиям к продукту.

**А.15 метод прямой оценки:** Метод оценки качества набора данных, основанный на проверке групп данных внутри набора данных.

**А.16 метод косвенной оценки:** Метод оценки качества набора данных, основанный на внешних знаниях.

*Пример - Примерами внешнего знания являются сведения об использовании наборов данных, методе их получения и т.п.*

**А.17 выборка:** Несколько отбираемых по установленным правилам из набора данных единиц информации, с помощью которых оценивается часть набора или набор в целом.

**А.18 контроль по альтернативному признаку:** Контроль по качественному признаку, в ходе которого каждую проверенную единицу относят к категории годных или дефектных, а последующее решение о контролируемой совокупности принимают в зависимости от числа обнаруженных в выборке или пробе дефектных единиц продукции или числа дефектов, приходящих на определённое число единиц продукции (по ГОСТ 15895).



## ПРИЛОЖЕНИЕ Б (информационное)

### Классификация методов оценки качества данных

#### Б.1 Введение в классификацию

Методы оценки качества данных, используемые в процедурах оценки наборов данных ЦК, делятся на два основных класса: прямая оценка и косвенная оценка. Эти два основных класса в дальнейшем делятся на подклассы по источнику информации, необходимой для выполнения оценки. Рис. Б1 представляет структуру этой классификации.

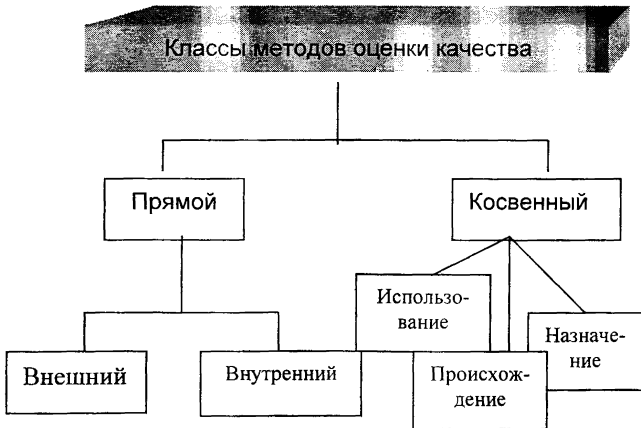


Рис. Б.1 Классификация методов оценки качества

Метод прямой оценки данных применяется в обязательном порядке изготовителем набора (наборов) данных ЦК. Пользователь может использовать как метод прямой оценки, так и метод косвенной оценки данных.

## Б.2 Метод прямой оценки

Метод прямой оценки набора данных ЦК подразделяется на внутренний метод и внешний метод.

Б.2.1 Применение метода прямой внутренней оценки качества данных должно предусматривать использование только той информации, которая входит в состав оцениваемого набора данных ЦК.

Б.2.2 Применение метода прямой внешней оценки качества данных должно предусматривать использование той информации, которая входит в состав оцениваемого набора данных ЦК, и применение эталонных данных, внешних по отношению к оцениваемому набору данных.

### *Примеры*

*1 Информация, необходимая для выполнения проверки на соответствие состава данных группы данных «названия дорог», требует наличия другого источника информации о названиях дорог (например, атласа автомобильных дорог).*

*2 Проверка точности метрических данных требует эталонного набора данных, например, каталога координат топо-геодезической основы или картографических данных более крупного масштаба, или выборочного полевого контроля.*

Б.2.3 В процессе оперативного обновления набора данных ЦК прямой метод применяется для оценки качества только тех данных, которые будут введены в состав набора

в процессе обновления, и данных, которые имеют пространственные или логические связи с вновь вводимыми данными.

### Б.3 Метод косвенной оценки

Косвенная оценка исследуемого набора данных ЦК производится на основе:

- информации о качестве наборов данных, ранее созданных для решения аналогичных задач.

- метаданных о происхождении оцениваемого набора данных, включающих информацию об исходных материалах, которые использованы для создания этого набора, о применяемых при его создании методах, технологиях, технических и программных средствах и другие сведения, позволяющие судить о пригодности его для предполагаемого использования.

#### *Примеры*

*1 Метаданные о наборе данных ЦК указывают, что содержащиеся в нём данные получены методом автоматизированной стереотопографической съёмки с использованием аэрофотоснимков, полученных при определенных условиях. Многократный опыт получения таких данных позволяет априори принять, что среднеквадратическая ошибка положения объектов относительно съёмочного обоснования, должна составлять не более 10 метров.*

*2 Метаданные, описывающие цель, ради которой создавались те или иные наборы данных, должны использоваться, когда определяется качество нового набора. Такой целью может быть поддержка конкретного требования пользователя или ориентация на нужды нескольких, но определенных пользователей.*

## **ПРИЛОЖЕНИЕ В**

### **(информационное)**

#### **Оценка качества данных с использованием выборки**

##### **В.1 Введение**

Важнейшими составляющими оценки качества данных с использованием выборки являются «группа» и «элемент данных». Группа является минимальной единицей, которая должна быть определена и проверена изготовителем данных в соответствии с техническими требованиями к конечному продукту. Если качество группы получило неудовлетворительную оценку, то все элементы данных в группе должны браковаться и заменяться на повторно созданные данные. Для статистически однородной группы применяется простая случайная выборка. Если группы статистически неоднородны, то простая случайная выборка с одним и тем же уровнем частоты выборки применяться не может.

##### **В.2 Общие положения**

При оценке качества данных с использованием выборки необходимо принимать во внимание следующее.

**В.2.1** Если набор данных описывает непрерывное пространство, включению в состав группы подлежат только элементы данных, находящиеся в границах области определения качества, независимо от того, частично или полностью объекты, которым принадлежат эти элементы, расположены в указанной области.

**В.2.2** Поскольку на качество данных оказывает влияние множество факторов, включая качество исходных данных и квалификацию операторов, изготовитель данных должен при определении группы гарантировать однородность данных с точки зрения качества.

**В.2.3** При проверке на наличие пропуска элементов данных размер группы равен числу позиций, представленных в группе, плюс число пропущенных элементов данных. Однако, так как информация о пропущенных элементах данных до проверки обычно отсутствует, размер группы может определяться числом имеющихся элементов данных до тех пор, пока количество пропусков не будет больше указанного в нормативно-технической документации.

**В.2.4** В случаях, когда для оценки качества данных необходимы дополнительные измерения (например, съемка на местности), проверка каждой случайно выбранной позиции выборки является неэффективной и дорогостоящей. В таких случаях, более эффективным может быть установление пространственно ограниченной зоны выборки и проверка всех элементов данных в зоне этой выборки. Однако, если какие-либо погрешности в зоне выборки имеют систематический характер, то точность оценки с использованием такой выборки снижается. Если это имеет место, необходимо создавать достаточно малые зоны выборки при увеличении их числа.

### **В.3 Объём выборки**

Размер одного и того же набора данных и, соответственно, объём выборки могут оцениваться по-разному в зависимости от элементов данных, которые будут участвовать в оценке качества. С учетом этого определение объёма выборки должно быть закончено, если определены элементы

данных, которые будут участвовать в выборке. В табл. В.3 приводятся примеры способов определения объема выборки.

Таблица В.3

Примеры определения объемов выборки

| Набор данных          | Размер набора данных                                   | Объем выборки для оценки   |
|-----------------------|--|--|
| Объекты дорожной сети | Число объектов дорожной сети.                          | Число объектов дорожной сети, выраженное в процентах от общего числа объектов.     |
| Охваченная площадь    | Значение площади, включающей в себя набор данных       | Площадь, включающая в себя выборку, выраженная в процентах от общей площади.       |
| Длина дуг             | Общая длина дорог, составляющих набор данных           | Длина выбранных дорог, выраженная в процентах от общей длины.                      |
| Вершины углов         | Общее число вершин углов, описывающих положение дороги | Число вершин углов в выборке, выраженное в процентах от общего числа вершин углов. |

Для усиления представительности выборки её объем рекомендуется определять с использованием совокупности различных критериев.

*Пример - Выборка должна включать 10% от площади, охваченной набором данных, и содержать не менее 5% от общей длины дорог в наборе данных.*

С учетом необходимости максимально автоматизировать процесс оценки качества набора данных ЦК приведенные ниже методы следует рассматривать как предлагаемые алгоритмы программного формирования выборок.

### В.3.1 Методы формирования выборки на основе объектов

Выбор объектов для построения выборки набора данных должен производиться, используя случайный процесс, предполагающий одинаковые условия создания данных для всего набора данных. В некоторых случаях случайный процесс может не создавать удовлетворительную выборку либо из-за того, что однородность можно обнаружить лишь для групп в составе набора данных, либо из-за того, что не проверяются дополнительные ограничения для выборки.

*Пример - Если формирование выборки осуществляется путем случайного выбора объектов, то существует риск возникновения выборки, сосредоточенной на малой площади (которая может быть или не быть приемлемой).*

Метод полуслучайной выборки может использоваться, чтобы гарантировать контроль и оценку наиболее важных элементов данных в составе набора данных ЦК при объемах выборки, ограниченных необходимостью сократить затраты на проведение вышеуказанных процессов.

*Примечание - Полуслучайный выбор состоит из случайного выбора объектов одного типа и затем анализа данных по объектам различных типов из окружения первых.*

### В.3.2 Методы формирования выборки на основе пространственного распределения

Методы формирования выборки считаются действующими на основе пространственного распределения, когда выбор объектов или компонентов объектов, которые должны включаться в выборку, производится с учётом местоположения объекта (части объекта). Этот тип формиро-

вания выборки может использоваться в качестве первого этапа с последующим использованием метода формирования выборки на основе объектов.

Пространственное распределение выборок может иметь нерегулярный, полурегулярный и регулярный характер. Варианты таких распределений проиллюстрированы на Рис. В.1, В.2 и В.3.

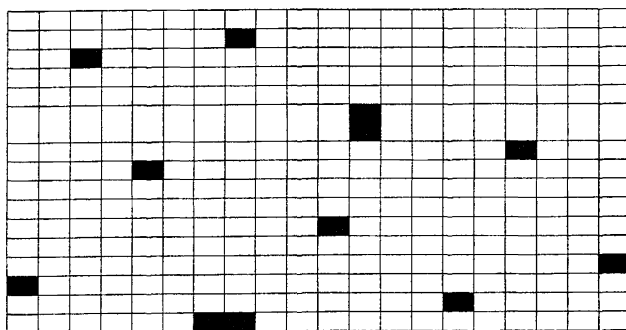


Рис. В.1 Пример нерегулярного размещения пространственных выборок (залитые прямоугольники). При использовании этого метода формирования выборки допустимо варьировать не только положение выборки, но и её объём.



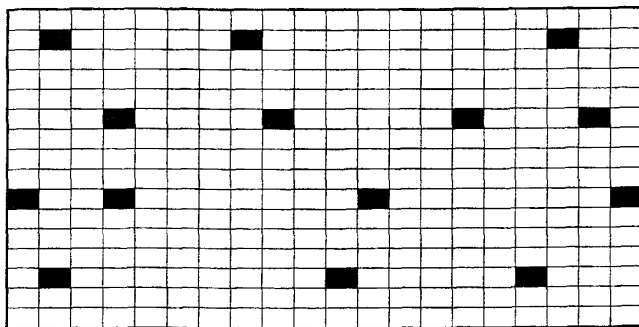


Рис. В.2 Пример полуслучайного размещения пространственных выборок (все выборки размещены на параллельных линиях, проходящих на одинаковом расстоянии друг от друга).

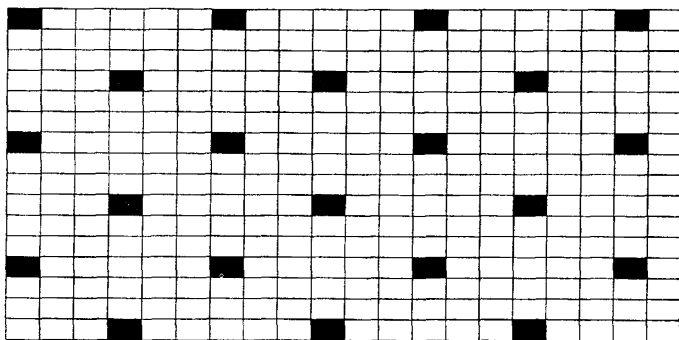


Рис. В.3 Пример регулярного размещения пространственных выборок.

## Приложение Г (обязательное)

### Формирование отчетов о результатах оценки качества данных

Настоящее приложение регламентирует содержание отчетов о результатах оценки качества наборов данных ЦК, когда требуется представление информации в указанной форме в соответствии с 7.1.

Таблица Г.1 устанавливает компоненты отчета о результатах оценки качества набора данных методом прямой оценки, таблица Г.2 – компоненты отчета о результатах оценки качества набора данных методом косвенной оценки.

Таблица Г.1

Отчет о результатах оценки качества набора данных ЦК  
методом прямой оценки

| №<br>п/п | Содержание раздела,<br>подраздела или пункта<br>отчёта  | Усло-<br>вие<br>запол-<br>нения | Макси-<br>мальная<br>частота<br>появле-<br>ний | Способ<br>заполне-<br>ния                          |
|----------|---|---------------------------------|--|--|
| 1        | 2   | 3                               | 4  | 5  |
| 1        | Информация по иденти-<br>фикации отчета   | О                               | 1  | Свободный<br>текст                                 |
| 2        | Нормативная документа-<br>ция, регламентирующая<br>диапазоны или уровни до-<br>пустимых значений по-<br>грешностей данных | О                               | 1  | Реквизиты<br>НД или тех-<br>нических<br>требований |

| 1   | 2  | 3 | 4 | 5               |
|-----|--|---|---|-----------------|
| 3   | Метаданные оцениваемого набора данных  | О | М | Свободный текст |
| 4   | Анализируемый аспект качества данных: «синтаксическая корректность»                                      | У | 1 | Раздел отчёта   |
| 4.1 | Область (подобласть) определения качества данных по указанному аспекту                                   | О | 1 | Свободный текст |
| 4.2 | Способ проведения контроля (полный контроль, контроль по выборке и по тем или иным критериям)            | О | 1 | Свободный текст |
| 4.3 | Метаданные группы данных   | О | М | Свободный текст |
| 4.4 | Описание анализируемых показателей   | О | М | Перечисление    |
| 4.5 | Метод оценки качества (прямой внутренний, прямой внешний).   | О | М | Свободный текст |
| 4.6 | Используемый критерий (-и) оценки качества данных  | О | М | Свободный текст |
| 4.7 | Используемая мера (-ы) оценки качества данных (имя, тип значения, единицы измерения)                     | О | М | Таблица         |
| 4.8 | Диапазон или уровень допустимых значений погрешностей данных (тип значения, значение, единицы измерения) | О | М | Таблица         |

## ОСТ 68-3.4.2-03

| 1    | 2  | 3 | 4 | 5   |
|------|--|---|---|---|
| 4.9  | Процедура оценки качества  | О | М | Свободный текст (при автоматическом контроле указываются реквизиты программных средств) |
| 4.10 | Оценка качества данных   | О | М | Значение оценки или таблица значений  |
| 4.11 | Результат оценки качества данных   | О | 1 | Принимается (бракуется)   |
| 4.12 | Дата проведения оценки качества данных   | О | 1 | Дата  |
| 5    | Анализируемый аспект качества данных: «правильность логической структуры» (выполняются все операции, аналогичные указанным в 4.1 – 4.12 таблицы) | У | 1 | Раздел отчёта   |
| 6    | Анализируемый аспект качества данных: «актуальность» (выполняются все операции, аналогичные указанным в 4.1 – 4.12 таблицы)                      | У | 1 | Раздел отчёта   |

| 1  | 2   | 3 | 4 | 5  |
|----|---|---|---|--|
| 7  | Анализируемый аспект качества данных: «соответствие состава» (выполняются все операции, аналогичные указанным в 4.1 – 4.12 таблицы)       | У | 1 | Раздел отчёта  |
| 8  | Анализируемый аспект качества данных: «логическая согласованность» (выполняются все операции, аналогичные указанным в 4.1 – 4.12 таблицы) | У | 1 | Раздел отчёта  |
| 9  | Анализируемый аспект качества данных: «точность» (выполняются все операции, аналогичные указанным в 4.1- 4.12 таблицы)                    | У | 1 | Раздел отчета  |
| 10 | Анализируемый аспект качества данных «соответствие целям использования»   | О | 1 | Возможно/невозможно использовать по прямому назначению<br>Возможно/невозможно использовать в области, отличной от прямого назначения |

*Примечания*

*1 Столбец «Условие заполнения» определяет необходимость отчетности по тому или иному разделу. Существуют три кода обязательств:*

*- обязательный (О) означает, что этот ввод информации необходим всегда;*

*-условный (У) означает, что ввод информации производится для всех подразделов раздела, когда удовлетворяется установленное требование; в противном случае ввод информации необязателен (Н);*

*2 Столбец «Максимальная частота появлений» определяет число появлений этого подраздела в пределах соответствующего раздела. «1» указывает на однократное появление подраздела, а “ М” – на многократное.*

*3 Столбец «Способ заполнения» заполняется для каждого элемента отчета. “Свободный текст” указывает, что на содержание сообщения не накладывается никаких ограничений.*

Таблица Г.2

**Отчет о результатах оценки качества набора данных ЦК методом косвенной оценки**

| <b>№ п/п</b> | <b>Содержание раздела, подраздела или пункта отчёта</b>                                   | <b>Условие заполнения</b> | <b>Максимальная частота появлений</b> | <b>Способ заполнения</b> |
|--------------|---|---------------------------|---------------------------------------|--------------------------|
| <b>1</b>     | <b>2</b>  | <b>3</b>                  | <b>4</b>                              | <b>5</b>                 |
| 1            | Перечень источников данных для косвенного метода оценки качества                          | О                         | 1                                     | Перечисление             |
| 2            | Идентификация первичных документов, используемых как основа для косвенного метода оценки  | Н                         | 1                                     | Перечисление             |
| 3            | Метаданные наборов данных, используемых в процессе косвенной оценки.                      | О                         | 1                                     | Перечисление             |
| 4            | Информация о результатах оценки качества данных, используемых в процессе косвенной оценки | О                         | 1                                     | Перечисление             |
| 5            | Значение уровня качества оцениваемого набора данных                                       | О                         | 1                                     | Свободный текст          |

| 1 | 2   | 3 | 4 | 5               |
|---|---|---|---|-----------------|
| 6 | Дата определения значения уровня качества   | О | 1 | Дата            |
| 7 | Любая дополнительная информация, включая промежуточные результаты на любых этапах оценки качества набора данных | Н | 1 | Свободный текст |

*Примечание – Расшифровка значений, помещаемых в таблице, аналогична той, которая используется в примечаниях к табл.Г.1.*



## Приложение Д

### (информационное)

#### Обобщение результатов оценки качества данных

##### Д.1 Введение

При наличии отчета о результатах оценки качества набора данных ЦК его качество может быть представлено одним или более обобщенными показателями качества данных. Обобщенный показатель ориентирован на получение его в автоматическом режиме и с использованием разных методов.

*Примечание – Обобщённый показатель является абстрактной величиной, используемой как оценка качества группы данных (группой данных может в тех или иных случаях являться и набор данных в целом).*

Следует учитывать, что при таком способе оценки качества набор данных может получить общую положительную оценку, даже если один или более показателей качества не соответствуют нормативным или техническим требованиям. В любом случае, смысл обобщенной оценки должен быть определён ещё до выполнения операций по определению обобщённого показателя качества.

Ниже приведены примеры методов, которые используются для создания обобщённого показателя качества.

##### Д.2 Однозначная оценка качества данных

Показателю качества данных каждой группы, вовлеченному в вычисления, придается логическое значение  $V_i$ ,

равное единице (1), если значение показателя соответствует требованиям, и нулю (0), если нет. Обобщенный показатель качества определяется по формуле:

$$ADQ = V_1 + V_2 + V_3 + \dots + V_n$$

где  $ADQ$  – обобщенный показатель качества набора данных;

$n$  - число групп определения качества данных.

Если  $ADQ=n$ , то качество набора данных считается полностью соответствующим требованиям. Если  $ADQ<n$ , то качество считается несоответствующим требованиям. Метод не указывает на конкретные показатели, которые не соответствуют установленным требованиям.

### Д.3 Взвешенная оценка качества данных

Показателю качества данных для каждой группы, во-влеченному в вычисления, придается логическое значение  $V_i$ , равное единице (1), если значение показателя соответствует требованиям, и нулю (0), если нет. Кроме того, на основании значимости показателя для оценки качества в целом каждому из них присваивается весовое значение  $W$  в интервале между 0,0 и 1,0 включительно. Сумма всех весов должна равняться 1,0. Выбор весов является субъективным решением, принимаемым изготовителем или пользователем данных. Как правило, значения весов устанавливаются на основе предварительной оценки качества. Обобщенное качество определяется по формуле:

$$ADQ = V_1 \times W_1 + V_2 \times W_2 + V_3 \times W_3 + \dots + V_n \times W_n$$

где  $n$  - число групп определения качества данных.

Метод обеспечивает получение значения величины, показывающей, насколько набор данных близок к полному соответствию. Метод не указывает на конкретные показатели, которые не соответствуют установленным требованиям

#### Д.4 Достаточная оценка качества набора данных

Этот метод должен использоваться, если объём оценки, представленной в отчёте по оценке качества набора данных ЦК больше, чем нужно, чтобы удовлетворить технические требования к набору данных. Для получения достаточной оценки набора данных привлекаются только значимые для этого конкретного случая значения мер и/или показателей качества. На основе отобранных мер и/или показателей обобщённая оценка получается с использованием метода однозначной или взвешенной оценки качества набора данных. Применение этого метода должно сопровождаться обязательным указанием мер и/или показателей, использованных для определения значения достаточной оценки.

#### Д.5 Оценка данных на основе максимальных или минимальных значений показателя качества данных

Показателю качества каждой группы данных придается значение  $V_i$ , основанное на значимости данного показателя для набора данных. Способ расчета значения  $V_i$ , является решением, принятым изготовителем или пользователем данных. Основанием для подобного решения должны являться результаты оценки качества отдельных групп данных.

Обобщенный показатель качества определяется одним из двух уравнений:

$$ADQ = MAX(V_i, i=1...n)$$

или

$$ADQ = MIN(V_i, i=1...n)$$

где  $n$  является числом определений качества данных.

Полученный таким образом показатель является дополнительным к обобщенному показателю, полученному одним из методов, описанных в Д.2 – Д.4, и указывает, где имеет место наибольшее соответствие или несоответствие нормативным требованиям. Однако, этот метод не сообщает о ситуации с другими показателями качества данных.

## **Приложение Е**

### **(обязательное)**

#### **Формирование метаданных о качестве данных**

##### **Е.1 Введение**

Настоящее приложение описывает метаданные, требующиеся для представления отчетности о качестве набора данных ЦК в соответствии с 7.2. Метаданные в целом или элементы метаданных рассматриваются как обобщенный(ые) показатель(и) качества набора данных. Способ получения обобщенных показателей изготовителем или пользователем данных выбирается из числа тех, что описаны в Приложении Д. Обоснование выбора того или иного способа обобщения возлагается на исполнителя процесса контроля и оценки качества данных.

Следует учитывать, что показатели качества набора данных, причисляемые к метаданным, имеют различную степень обобщения, в зависимости от назначения создаваемых метаданных, объема исходной информации для их формирования и требуемой степени детальности метаданных на стадии оформления отчета о результатах оценки качества данных ЦК.

Метаданные должны содержать сведения о качестве данных по каждому из аспектов качества, указанных соответствующим нормативно-техническим документом.

Несмотря на разнообразие показателей качества, метаданные должны иметь компактный характер. Если значения показателей качества, определённые для отдельных групп данных, удовлетворяют одни и те же требования, то

они должны быть обобщены на более высоком уровне для этих групп данных. Дополнительные уровни метаданных определяются и сохраняются только в тех случаях, когда показатели качества отдельных групп данных обладают значениями, отличными от показателей на верхнем уровне.

## Е.2 Форма представления метаданных

Таблица Е.2 устанавливает порядок и форму представления метаданных о качестве набора ЦК.

Таблица Е.2

| № п/п | Содержание метаданных о качестве набора данных   | Условие заполнения | Максимальная частота появлений | Способ заполнения                   |
|-------|--|--------------------|--------------------------------|-------------------------------------|
| 1     | 2  | 3                  | 4                              | 5                                   |
| 1     | Описание набора данных   | У                  | 1                              | Свободный текст                     |
| 2     | Нормативно-технический документ, определяющий требования к набору данных                   | О                  | 1                              | Официальное название                |
| 3     | Метод определения значений обобщённых показателей  | О                  | 1                              | Свободный текст                     |
| 4     | Наименование аспекта качества данных   | О                  | Н                              | Свободный текст                     |
| 4.1   | Количественные показатели качества данных, на основе которых получен обобщённый показатель | О                  | Н                              | Перечень использованных показателей |

| 1   | 2   | 3 | 4 | 5                         |
|-----|---|---|---|---------------------------|
| 4.2 | Область определения качества обобщённым показателем                                     | О | Н | Свободный текст           |
| 4.3 | Значение обобщённого показателя   | О | Н | Число с десятичной точкой |
| 4.4 | Допустимый уровень качества данных для исследуемого аспекта                             | О | Н | Свободный текст           |
| 4.5 | Результат оценки качества по исследуемому аспекту                                       | О | Н | Принимается/бракуется     |
| 4.6 | Дата проведения оценки качества набора данных по исследуемому аспекту                   | О | Н | Дата                      |
| 5   | Информация о результате оценки качества всего набора данных ЦК                          | О | 1 | Свободный текст           |
| 5.1 | Результат оценки уровня качества набора данных  | О | 1 | Число с десятичной точкой |
| 5.2 | Допустимый уровень качества при оценке качества всего набора данных                     | О | 1 | Число с десятичной точкой |
| 5.3 | Информация о возможности применения набора данных в соответствии с целями использования | О | 1 | Свободный текст           |
| 5.4 | Дата проведения оценки качества набора данных   | О | 1 | Дата                      |

*Примечания*

*1 Расшифровка обозначений, помещенных в таблице, аналогична той, которая используется в Приложении Г.*

*2 Пункт 1 заполняется только в тех случаях, когда сведения о качестве набора данных ЦК представляются отдельно от собственно данных, например, для информационных или рекламных целей.*

*3 Множественность заполнения данных в 4.1 – 4.6 связана с тем, что обобщённая оценка данных предварительно производится для каждого из исследуемых аспектов качества.*

*4 Допустимо одновременное использование более, чем одного метода определения обобщённого показателя качества. В этом случае для каждого очередного метода таблица Е.2 пополняется собственными для него данными по 3 – 5.*



## **Приложение Ж** **(справочное)**

### **Библиография**

- [1] Основные положения по созданию и обновлению топографических карт масштабов 1:10 000, 1:25 000, 1:50 000, 1:100 000, 1:200 000, 1:500 000, 1:1 000 000, М., РИО ВТС, 1984

---

Подписано в печать  
26.05.03  
Формат 60х90/16  
Бумага типографская  
Печать офсетная  
Усл. печ. л. 3,75  
Усл. кр. отт. 3,88  
Уч. изд. л. 3,60

Тираж 400  
Заказ 21-03

ЦНИИГАиК  
125413, Москва,  
Онежская ул., 26