



КОНЦЕПТУАЛЬНІ ЗАСАДИ ОЦІНЮВАННЯ ТА ЗАБЕЗПЕЧЕННЯ ЯКОСТІ ГЕОПРОСТОРОВИХ ДАНИХ

Рассмотрены современные подходы к вопросу формирования общей концепции качества геопространственных данных с учетом их информационно-модельной сущности и основного назначения для применения в геоинформационных системах принятия решений. Изложены основные элементы, меры, процедуры и методы оценивания качества данных, исходя из методологии серии международных стандартов ISO 19100.

The paper deals with the modern approaches to formation of general conception of geospatial data quality taking into account their information modelling essence and their basic assignment for application in geoinformation systems of decision-making. The basic elements, measures, procedures and methods of estimation of geospatial data quality according to methodology of ISO 19100 international standards are described.

Вступ. Категорія якості є однією з визначальних для оцінювання споживчих властивостей будь-якої продукції. За міжнародним стандартом ISO 9000-2001 [2], якість визначається як ступінь, до якого має прагнути сукупність власних характеристик продукції, щоб вона задовольняла встановлені вимоги.

Геопросторові дані (ГД) – це набори даних або бази даних про об'єкти реального світу, що мають певне місцеположення на Землі, зафіксоване у встановленій системі просторово-часових координат. Завдяки ГІС геопросторові дані перетворилися в новий, основний вид продукції топографо-геодезичного виробництва та суміжних сфер діяльності, пов'язаних з дослідженням, кадаструванням, моніторингом нерухомості, природних ресурсів, навколишнього природного середовища, інженерної інфраструктури тощо.

Специфіка ГД полягає в тому, що це не просто реєстр даних про об'єкти і явища реального світу. За своїм призначенням вони є основою для побудови геоінформаційних моделей місцевості в середовищі ГІС певного призначення або самі є такими моделями. Власне, з інформаційною, просторово-часовою та модельною сутністю ГД пов'язана множинність прикладних сфер їх застосування в геоінформаційних системах підтримки прийняття рішень (ГІС ППР) різного призначення, складність завдань забезпечення та оцінювання якості геоінформаційної продукції як на концептуальному, так і на практично-методичному і технологічному рівнях.

Питання якості геопросторових даних набуває особливого сенсу, коли йдеться про їх використання в ГІС ППР з метою попередження та швидкого реагування на надзвичайні ситуації, коли від повноти й актуальності даних залежить не тільки обсяг матеріальних втрат, але нерідко і життя людей.

Кожна предметна сфера висуває свої специфічні вимоги до структури та складу геоінформаційних моделей (набору об'єктів та їхніх атрибутів, точності визначення просторових властивостей об'єктів, топологічних і логічних відношень між

об'єктами, що мають бути відображені та реалізовані в моделях, тощо). Не варто забувати і про цифрові моделі місцевості, створення яких є трудомістким і вартісним процесом. Тому важливо знаходити підходи до створення багатоцільових моделей місцевості, розробляти загальні принципи та логічні схеми побудови прикладних моделей для забезпечення міжгалузевого обміну та інтероперабельності ГД, а також концептуальних засад, принципів, мір та методів оцінювання їх якості.

Вирішуючи останнє завдання, автори спиралися на власний багаторічний досвід виготовлення цифрової картографічної продукції, баз геопросторових даних (БГД) та прикладних ГІС, а також на зміст концепцій міжнародних стандартів комплексу ISO 19100: Географічна інформація/геоматика щодо якості географічної інформації та на публікації провідних фахівців у цій сфері.

Аналіз досліджень, публікацій та стану проблеми. Публікації з проблематики якості геопросторових даних віддзеркалюють етапи розвитку геоінформаційних технологій, ГІС та еволюцію їх архітектури – від локальних ГІС на основі цифрових карт, корпоративних ГІС на БГД до розподілених ГІС в Інтернет та інфраструктури геопросторових даних.

На ранніх етапах розвитку ГІС, включно до 1990-х років, основна увага приділялася якості цифрових карт, яка характеризувалася сукупністю властивостей цифрової картографічної продукції (ЦКП), що визначала рівень її придатності задовольняти певні потреби відповідно до призначення [1]. Наразі це уточнення у визначенні якості ЦКП ("задовольняти певні потреби відповідно до призначення" на відміну від "задовольняти вимоги" у визначенні за ISO 9000-2001) вказує на розуміння ГІС в контексті картографічного виробництва, в якому картографічні твори класифікуються за призначенням, тематичним змістом та масштабом. Основними ж показниками якості ЦКП було визначено характеристики відповідності цифрових даних та/або електронних карт, зображення яких генерувалося за цифровими даними, вихідним картографічним матеріалом, зокрема, за складом об'єктів, точністю їх просторового положення та геометричною формою тощо. Складалася типова



ситуація, за якої модель цифрової топографічної карти (ЦТК) повністю відповідала за змістом вихідній паперовій карті, забезпечувала її відтворення в електронному вигляді, але була непридатна для цілей геоінформаційного моделювання. Наприклад, цифрову топографічну карту не можна використати для пошуку оптимального маршруту через відсутність у ній топологічної сегментно-вузлової (графової) моделі вулично-дорожньої мережі або, з аналогічних причин, шар об'єктів гідрографії ЦТК непридатний для моделювання водотоків з метою оцінювання загрози повені. Таку модель ЦТК за картографічним призначенням можна використовувати як підоснову для введення та відображення прикладної цифрової моделі за аналогією використання паперової топографічної карти для нанесення ситуації певного тематичного змісту.

Розвиток ГІС у 2000-х роках характеризується такими кардинальними змінами, як тотальний перехід до цифрових технологій геодезичних вимірювань та збирання інформації про місцевість (GPS, цифрові знімки ДЗЗ, цифрова фотограмметрія, лідарні знімання тощо), втілення концепції інфраструктури ГД на основі систем керування базами геопросторових даних (СКБГД), геопорталів з базами метаданих, геоінформаційними сервісами в мережі Інтернет як основного засобу пошуку, отримання та спільного використання даних різного тематичного змісту й територіального охоплення, створених відповідно до міжнародних стандартів ISO у сфері географічної інформації / геоматики і технічних специфікацій відкритого геопросторового консорціуму OGC.

Можливість за допомогою геоінформаційних сервісів Інтернет отримувати ГД, що створюються для різних потреб різними виробниками і користувачами, підсилила об'єктивну потребу в методах та засобах досягнення інтероперабельності прикладних геоінформаційних моделей, а також в уніфікованих процедурах і сервісах забезпечення та оцінювання якості даних.

Група експертів EuroGeographics ще у 2007 р. підготувала настанову про впровадження базових стандартів якості географічної інформації комплексу ISO 19100 в національних картографічних та кадастрових агентствах (НККА) [5]. У цій настанові, зокрема, йшлося про те, що більшість НККА країн Європи завершила векторизацію паперових топографічних карт, а їх оперативне пооб'єктне чергування здійснюється в цифрових форматах у базах топографічних даних (БТД) з використанням різних джерел первинної топографічної інформації. Тому в базах даних інформація для різних об'єктів мала різну точність, різні періоди проведення вимірювань тощо. Ці особливості в сучасній діяльності НККА експерти назвали "новою парадигмою виробництва", що включає використання багатьох джерел ГД у партнерстві з вишукувальними підрозділами різних відомств ("виробництво багатьох типів") та посилює роль системи управління якістю.

В умовах формування та розвитку інфраструктури геопросторових даних (ІГД) зміни у виробництві й використанні ГД характеризуються такими двома загальними тенденціями [10]:

1) еволюцією від виготовлення цифрової картографічної продукції до процесів моделювання (від даних виробників до даних користувачів; від централізованої до децентралізованої структури вироблення й підтримки даних; від створення локальних баз даних до обміну даними і федералізації банків даних);

2) переходом від розроблення концепцій до їх реалізації (зокрема, від координування процесу формування компонентів інфраструктури до керування її функціонуванням; від однорівневої до багаторівневої участі та від існуючих до нових організаційних структур вироблення і постачання геоінформаційної продукції).

На підтвердження цього у публікаціях [7, 10] наводяться приклади надання нових послуг з незалежного аудиту якості ГД, який забезпечує професійне оцінювання третьою стороною якості даних для замовника. Аудит якості визначається у цих дослідженнях як системний метод, за яким ідентифікують, збирають та аналізують інформацію про набори геопросторових даних (НГД) з точки зору задоволення потреб користувачів, для об'єктивного оцінювання та/або прийняття рішення стосовно використання наборів даних для визначених замовником цілей.

У більшості публікацій з проблематики якості геопросторових даних в останні роки йдеться про такі основні напрями дослідження:

1. Розвиток концепції визначення якості даних, уніфікація принципів та моделей комплексної системи управління якістю ГД [3-7, 10].

2. Розроблення та впровадження міжнародних стандартів з якості даних, що входять до комплексу стандартів ISO 19100: Географічна інформація/геоматика. Це, зокрема, стандарти ISO 19113:2002 Quality principles (Принципи якості), ISO 19114:2003 Quality evaluation procedures (Методи оцінювання якості), ISO/CD TS 19138:2006 Data quality measures (Вимірювання якості даних), ISO 19115:2003 Geographic information – Metadata (Географічна інформація – Метадані) та проект міжнародного стандарту ISO/DIS 19157:2011 Geographic information – Data quality (Географічна інформація – Якість даних), в якому переглядаються та узагальнюються положення стандартів ISO 19113:2002, ISO 19114:2003 та ISO/TS 19138:2006, на заміну яких він розробляється [5-7].

3. Обґрунтування та уніфікація мір і процедур оцінювання якості, від адекватності вибору яких значною мірою залежить достовірність оцінки якості ГД [4-6, 11, 12].

4. Створення засобів автоматизованого контролю та оцінювання якості. Розробники інструментальних ГІС та СКБГД намагаються інтегрувати інструменти для оцінювання якості у свої продукти. Наприклад, у складі ArcGIS реалізовано



підсистему Data Reviewer для автоматизованого контролю якості ГД [13]; у просторових опціях СКБД Oracle Spatial, починаючи з 10-ї версії, постачаються програмні засоби Topology and Network Data Model, які за попередньо встановленими правилами дозволяють контролювати топологічну цілісність ГД [8]; постійно розширюється бібліотека функцій контролю топології ГД в середовищі СКБД PostgreSQL+PostGIS [9].

У подальших своїх міркуваннях автори цієї статті відштовхуються від основних концептуальних засад згаданих вище міжнародних стандартів, розглядають шляхи їх практичного впровадження в системі оцінювання й забезпечення якості ГД.

Виклад основного матеріалу. До системи оцінювання якості будь-якої продукції входять такі типові завдання: формулювання принципів оцінювання та обґрунтування номенклатури показників якості; розроблення методів і процедур визначення показників якості; оптимізація типорозмірів параметричних рядів виробів; обґрунтування принципів побудови узагальнених показників та умов їхнього застосування у завданнях стандартизації та управління якістю. Ці завдання належать до сфери *кваліметрії* (від латинського *qualis* – який за якістю та *..метрія*) – наукової дисципліни, що вивчає та розвиває методи кількісного оцінювання якості різних об'єктів. Загальним методичним базисом забезпечення якості геопросторових даних є міжнародні стандарти ISO 19113, ISO 19114, ISO 19138 та ISO/DIS 19157:2011, в яких розглядаються принципи, методи визначення та оцінювання якості НГД в процесі їх виробництва.

Система комплексного управління якістю геоінформаційної продукції та надання геоінформаційних послуг (мал. 1) ґрунтується на чотирьох основних концептуальних підходах, які відображають особливості змісту показників та організаційних аспектів забезпечення оцінювання якості ГД, а саме [7, 10]: виробничо-орієнтований (*production-centred*); планувально- або проектно-орієнтований (*planning-centred*); клієнт-орієнтований (*customer-centred*); системно-орієнтований (*system-centred*) підходи.

Виробничо-орієнтований підхід належить до ключових у комплексній системі управління якістю географічної інформації, оскільки саме від забезпечення якості на етапах збирання інформації та формування НГД найбільшою мірою залежить у майбутньому рівень бездефект-

ності й кондиційності геоінформаційної продукції та послуг.

Проектно-орієнтований підхід додатково передбачає розгляд питань якості ГД з точки зору їхньої придатності для створення та/або розвитку певних прикладних систем, зокрема, з урахуванням придатності даних для використання в середовищі інструментальних ГІС, що плануються до застосування в цих прикладних системах.

При **клієнт-орієнтованому підході** якість геопросторових даних оцінюється як складова комплексної геоінформаційної продукції (прикладної ГІС), яку отримують кінцеві користувачі для вирішення своїх прикладних задач. Варто зауважити, що кінцевими користувачами якості прикладної ГІС, як правило, сприймається інтегрально, при цьому невизначеність даних може бути причиною недоліків програмних засобів і навпаки.

При **системно-орієнтованому підході** якість НГД оцінюється в контексті його інтероперабельності та придатності для розміщення й використання в технологічному середовищі ГД, у т. ч. й разом з іншими наборами даних та/або уніфікованими геоінформаційними сервісами. Важливою компонентою при оцінюванні якості НГД з точки зору системно-орієнтованого підходу є повнота, конкретність та достовірність метаданих про набір даних загалом та про його якість зокрема.

Якість може мати різну інтерпретацію залежно від стадії життєвого циклу (ЖЦ) геопросторових даних. У табл. 1 [5] узагальнено концепти якості для трьох основних стадій ЖЦ НГД.

Концептуально процес визначення якості ГД



Мал. 1. Структура концептуальних підходів до забезпечення якості ГД у комплексній системі управління якістю надання геоінформаційних послуг



Таблиця 1. Інтерпретація концептів якості для різних стадій життєвого циклу набору геопросторових даних

Фаза ЖЦ НГД	Документація про якість	Мета заходів системи управління якістю	Процедури визначення якості	Рівень компонентів НГД
Підготовка виробництва	Технічні вимоги та вхідна модель якості	Визначення вимог до якості	Вивчення вимог користувача	Рівень класів об'єктів
Виробництво	База даних з документами про історію процесу	Оцінювання відповідності специфікації. Зазначення очікуваної якості в базі даних	Ретельний контроль	Екземпляри об'єктів (дата, точність координат тощо)
Використання	Метадані, звіт про контроль якості	Вимірювання відповідності вимогам якості	Оцінювання. Звітування	Рівень бази даних

можна розглядати як оцінювання відмінності реально вироблених даних від певного ідеального еталонного набору, в якому немає будь-яких помилок, що виникають у ході формування БГД.

Розподіл помилок розглядається за фазами ЖЦ НГД:

- *збирання даних*: неточності польових вимірювань, неточність приладів, неточність ведення записів, помилки при аналізі даних, отриманих дистанційно;
- *введення даних*: помилки цифрування, нечіткість природних контурів об'єктів;
- *збереження даних*: числова неточність, просторова неточність (для растрових даних);
- *оброблення даних*: неправильні класифікаційні інтервали, помилки при створенні полігонів;
- *формування кінцевої продукції*: похибки масштабування (для растрових даних), обмеження кінцевого формату даних;
- *використання даних*: неправильне розуміння структури та вмісту, неправильне використання даних.

Більшість дослідників цієї теми розрізняють внутрішню та зовнішню якість ГД [3,4].

Під **внутрішньою якістю** розуміють рівень відповідності між створеним та "ідеальним" набором, який мав би бути виготовлений (тобто дані, виготовлені без помилок). Такий набір є відображенням реального світу на певну дату відповідно до специфікації на продукцію, що встановлює набір правил та вимог переходу від реального світу до моделі даних. Специфікація, наприклад, включає перелік об'єктів, що мають бути відображені, тип геометрії для кожного класу об'єктів, атрибути, що мають описувати ці об'єкти, а також допустимі значення для атрибутів. Описати внутрішню якість можна різними способами, але в основному використовують критерії, визначені у міжнародному стандарті ISO 19113, а саме *повнота, логічна узгодженість, позиційна, часова й тематична точність*.

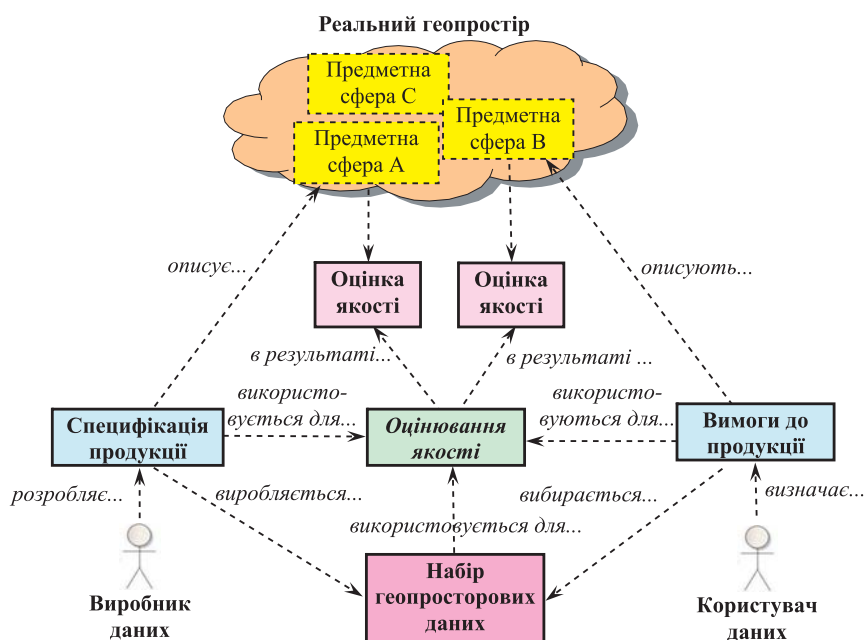
Під **зовнішньою якістю** мають

на увазі рівень відповідності готового продукту потребам чи очікуванням користувача. Така якість не є абсолютною, а тому один і той же НГД може мати різну зовнішню якість для різних користувачів. Зовнішню якість часто визначають як придатність продукту для використання. Оскільки саме поняття зовнішньої якості відрізняється для різних користувачів, то не існує єдиних стандартизованих критеріїв для її опису. І. Бедард та Д. Вальєре [3] виділяють шість характеристик для опису зовнішньої якості ГД:

- *призначення*: цільова функція набору даних;
- *охоплення*: період часу й територія, на яку створено дані;
- *походження*: методи та процеси оброблення, використані для отримання кінцевих даних;
- *точність*: відповідність тематичної, часової та просторової точності вимогам користувача;
- *легітимність*: відповідність створених даних стандартам; гарантії від постачальника даних;
- *доступність*: зручність для користувача в отриманні даних (вартість, формат, конфіденційність, авторські права і т. д.).

Загальна структура концепції для визначення якості геопросторових даних за ISO 19113 (мал. 2) відображує ролі виробників і користувачів у процесі створення та оцінювання якості НГД.

НГД створюється для певного застосування. Якість набору може бути оцінена на основі знань про елементи якості даних, а в деяких випадках й



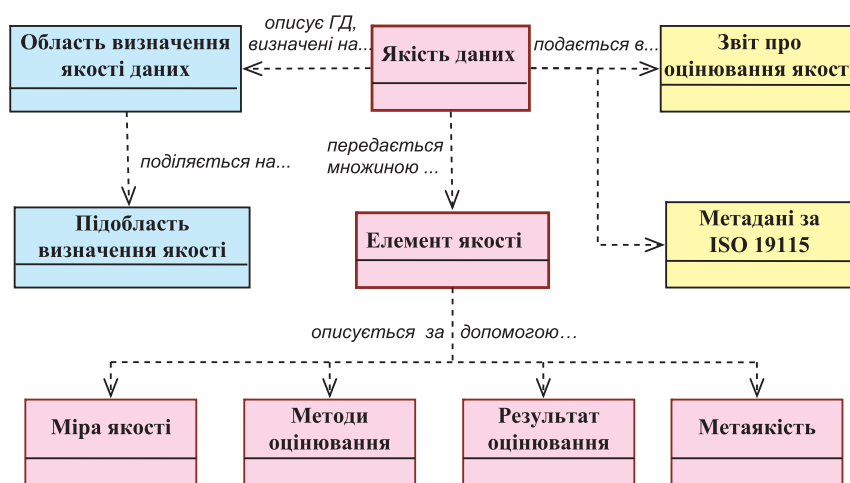
Мал. 2. Загальна структура концепції оцінювання якості ГД за ISO 19113



опосередковано на підставі декількох елементів, наприклад, призначення та походження даних, що описуються в метаданих згідно з ISO 19115. Елементи якості даних оцінюють рівень відповідності набору даних предметній сфері, що є частиною геопростору, а фактично технічним умовам (специфікації) виробника на створення НГД для певної мети.

Користувач даних оцінює якість НГД як рівень його відповідності вимогам застосування для геоінформаційного моделювання певної предметної сфери, яка може не збігатися з цільовим призначенням набору даних. А отже, оцінка якості виробника даних може не відповідати оцінці якості потенційного користувача даних. Важливо, щоб така оцінка в контексті мети оцінювання достовірно відображала рівень відповідності НГД специфікації виробника або вимогам користувача. Природно, що при зміні технічних умов виробника або вимог користувача має бути проведено нове оцінювання якості набору даних. Також важливо із застереженням порівнювати різні оцінки якості НГД, якщо його цільове призначення за специфікацією виробника не відповідає предметній сфері потенційного користувача.

За ISO 19113 передбачається розроблення в складі технічних вимог вхідної моделі якості НГД (мал. 3), яка в подальшому використовується в процедурах оцінювання переважно внутрішньої якості продукції.



Мал. 3. Концептуальна модель якості геопросторових даних

Згідно з цією моделлю, якість ГД передається за допомогою елементів якості. Для кожного елемента визначаються міри та методи його оцінювання. Міра якості дає кількісну характеристику елемента якості. Методи оцінювання якості визначають підходи до перевірок даних та обчислення мір. Окрім цього, кожен елемент якості описується елементами метаякості, що характеризують ступінь довіри до результату оцінювання та містять обґрунтування доцільності застосування обраної міри якості та методу оцінювання для конкретного елемента.

Інформація про якість даних подається в спеціальному звіті про оцінювання якості та в метаданих. Звіт про якість складається обов'язково як результат оцінювання за процедурами та методами, визначеними у вхідній моделі якості. Найчастіше такий звіт використовується в подальшому для виправлення помилок, виявлених в НГД під час контролю та оцінювання якості. Інформація про якість також описується у відповідних розділах та елементах метаданих, уніфікованих за ISO 19115. Метадані включаються до НГД як кінцевого продукту та можуть використовуватись споживачами наборів у процесі оцінювання зовнішньої якості даних.

Поєднання елементів якості та меж їх застосування встановлює область визначення якості. Область описує просторові та часові рамки застосування кожного елемента. Нею може бути набір даних, серія наборів або фрагмент даних, обмежений однією або декількома характеристиками: клас об'єкта, екземпляр об'єкта, територіальне охоплення, часове охоплення.

Загалом набір даних розглядається як такий, що містить велике, хоча й скінченне число менших груп даних. Припускається, що менші групи даних, об'єднані за належністю до того самого типу об'єкта, типу атрибуту об'єкта або відношення, критерію збору або до однієї і тієї ж географічної області, мають однакову якість. Менша група даних може складатися з одного екземпляра типу об'єкта, значення атрибуту або відношення. Теоретично, за концепцією оцінювання якості даних, допускається, що кожен екземпляр типу об'єкта, значення атрибуту й вид відношення набору даних може мати окремі елементи якості. Якість менших груп даних може не збігатися з якістю іншої частини набору даних, до якого вони належать. Концепція оцінювання якості даних допускає видачу інформації про якість набору загалом та додатково інформацію про якість менших груп, що визначається специфікою профілю ГД.

Елементи якості геопросторових даних. Згідно з міжнародним стандартом ISO 19113, якість НГД вказується з використанням кількісних та описових елементів.

Кількісні елементи якості геопросторових даних (див. табл. 2) дають змогу оцінити, наскільки той або інший набір даних відповідає критеріям, вказаним у специфікації на відповідну продукцію. Аспекти елементів якості даних називають піделементами. Піделементи оцінюють або перевіряють різними способами. Як уже зазначалося, за концепцією оцінювання якості даних не всі елементи й піделементи, а також не всі засоби їхнього оцінювання та перевірки можуть підходити конкретному набору даних. Крім того, деякі піделементи можуть застосовуватись до всього



Таблиця 2. Кількісні елементи і піделементи якості за ISO 19113

Назва елемента і піделемента якості, їх ідентифікатори в метаданих	Опис змісту
Повнота даних <i>DQ_Completeness</i>	Наявність чи відсутність об'єктів, їх атрибутів та відношень
Надлишковість <i>DQ_Commission</i>	Надлишкові дані, наявні в наборі даних
Відсутність <i>DQ_Omission</i>	Дані, відсутні в наборі даних
Логічна узгодженість даних <i>DQ_LogicalConsistency</i>	Ступінь відповідності логічним правилам структури даних, атрибутики та відношень
Концептуальна узгодженість <i>DQ_ConceptualConsistency</i>	Відповідність правилам концептуальної схеми
Доменна узгодженість <i>DQ_DomainConsistency</i>	Відповідність значень домену
Форматна узгодженість <i>DQ_FormatConsistency</i>	Ступінь відповідності зберігання даних фізичній структурі набору даних
Топологічна узгодженість <i>DQ_TopologicalConsistency</i>	Правильність експліцитно закодованих топологічних характеристик набору даних
Точність місцезположення об'єктів <i>DQ_PositionalAccuracy</i>	Точність місцезположення об'єктів
Абсолютна чи зовнішня точність <i>DQ_AbsoluteExternalAccuracy</i>	Близькість значень координат значенням, вказаним у звіті й прийнятим як правильні
Відносна чи внутрішня точність <i>DQ_RelativeInternalAccuracy</i>	Близькість відносних місцезположень об'єктів, вказаних у наборі даних, відповідним місцезположенням, прийнятим як правильні
Точність місцезположення даних у комірках <i>DQ_GridDEDDataPositionAccuracy</i>	Близькість значень місцезположення, вказаних для даних у комірках, значенням, що прийняті як правильні
Тематична точність даних <i>DQ_ThematicAccuracy</i>	Точність числових атрибутів, правильність нечислових атрибутів, класифікації об'єктів та їх відношень
Правильність класифікації <i>DQ_ClassificationCorrectness</i>	Порівняння класів об'єктів та їх атрибутів з певною предметною сферою (тобто відповідність базовим концептам предметної сфери або концептам еталонного набору даних)
Правильність нечислових атрибутів <i>DQ_NonQuantitativeAttributeCorrectness</i>	Правильність нечислових атрибутів
Точність числових атрибутів <i>DQ_QuantitativeAttributeAccuracy</i>	Точність числових атрибутів
Часова точність даних <i>DQ_TemporalQuality</i>	Точність часових атрибутів та часових відношень об'єктів
Точність часового вимірювання <i>DQ_AccuracyOfATimeMeasurement</i>	Правильність часових зв'язків елемента (звітування про помилку у вимірюванні часу)
Часова узгодженість <i>DQ_TemporalConsistency</i>	Правильність впорядкованих подій або послідовностей, якщо про них звітують
Часова відповідність <i>DQ_TemporalValidity</i>	Коректність даних з плином часу

набору даних, можуть бути оцінені або перевірені для нього, а інші застосовуються до менших груп даних, що належать до набору даних вищого порядку та можуть бути оцінені або перевірені для них.

Стандарт ISO 19113 визначає елементи якості даних безпосередньо як засоби виявлення і вироблення різних видів інформації про якість. Як правило, піделементи якості даних взаємозалежні. Наприклад, помилка в координатах може викликати, як мінімум, помилки двох типів – позиційну й топологічну. Значення піделементів, передбачене в угоді на продукт, та спосіб їх трактування належать до компетенції розробника специфікації на продукт.

Описові елементи якості даних забезпечують загальну не кількісну інформацію про якість. Вони дозволяють додатково оцінити придатність набору даних для конкретного застосування, містять інформацію про його призначення, використання й походження.

Елементи "Призначення" та "Використання" описують сферу застосування набору даних. Використання набору даних визначається його розробником або користувачами даних.

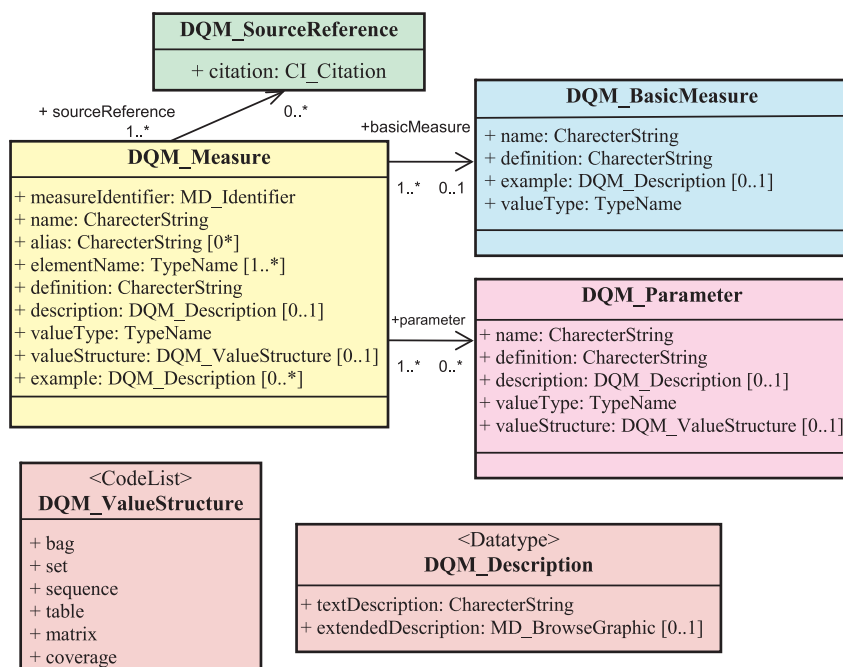
Елемент "Походження" описує історію формування набору даних і певною мірою його життєвий цикл, починаючи з процесів збирання, наступного кодування й перетворення у поточний формат даних.

Міри якості геопросторових даних. Міра якості є кількісною характеристикою ГД. Уніфікація мір якості здійснюється з метою досягнення сумісності та порівнюваності кількісної інформації про якість різних наборів даних. Однією з основних вимог до мір якості є однозначність їх визначення та коректність методів обчислення.

У проекті міжнародного стандарту ISO/DIS 19157:2011 пропонується набір стандартизованих мір якості, які дозволяють оцінювати практично всі кількісні елементи та піделементи якості ГД.

Кожна міра якості (*DQM_Measure*) в ISO/DIS 19157:2011 описується такими компонентами (мал. 4):

- ідентифікатор міри якості (*measureIdentifier*);
- назва міри якості (*name*);
- псевдонім (*alias*);
- назва елемента якості (*elementName*), до якого застосовується міра;



Мал. 4. Структура опису мір якості геопросторових даних за ISO/DIS 19157

- базова міра якості (*basicMeasure*);
- визначення (*definition*): фундаментальний концепт міри якості (якщо міра базується на одній з базових мір, то дається її визначення);
- опис (*description*): опис міри якості, включаючи всі методи обчислень і формули, необхідні для застосування міри;
- параметр (*parameter*): змінна, яку використовує міра якості (включає назву, визначення і тип даних параметра);
- тип значення (*valueType*): один з типів даних, що використовується для отримання результату міри (визначається за ISO/TS 19103:2005);
- структура значення (*valueStructure*), якщо результат включає декілька значень;
- посилання на джерело (*sourceReference*): посилання на документацію міри (якщо для міри якості додаткова інформація міститься у зовнішньому джерелі, то вказується посилання на це джерело);
- приклад застосування міри (*example*).

В описі базових мір якості виділяють два класи:

- 1) міри, що базуються на підрахунку помилкових або правильних об'єктів;
- 2) міри, що базуються на моделюванні невизначеності вимірювань статистичними методами.

У першому класі розрізняють шість базових мір

якості (табл. 3), що ґрунтуються на різних методах підрахунку помилкових чи правильних об'єктів. Числові дані, отримані в результаті вимірювань, мають певну точність, а тому для оцінювання ступеня невизначеності якоїсь вимірної величини рекомендується використовувати статистичні методи.

У міжнародному стандарті ISO 19138 та в проекті ISO/DIS 19157 застосування базових мір якості конкретизовано для усіх кількісних елементів та піделементів якості ГД, що в підсумку дозволило ідентифікувати понад 80 окремих мір якості. Однак через специфіку якості даних цей список не може бути повним. З розвитком ГС вимоги до якості постійно зростають, тому природним є розроблення додаткових мір якості.

Процес, процедури та методи оцінювання якості. Оцінювання якості ГД здійснюється на різних

стадіях ЖЦ продукції. Воно має різні цілі для кожної стадії. До основних стадій ЖЦ НГД можна віднести: розроблення технічних вимог (специфікації), виробництво, постачання, використання та оновлення.

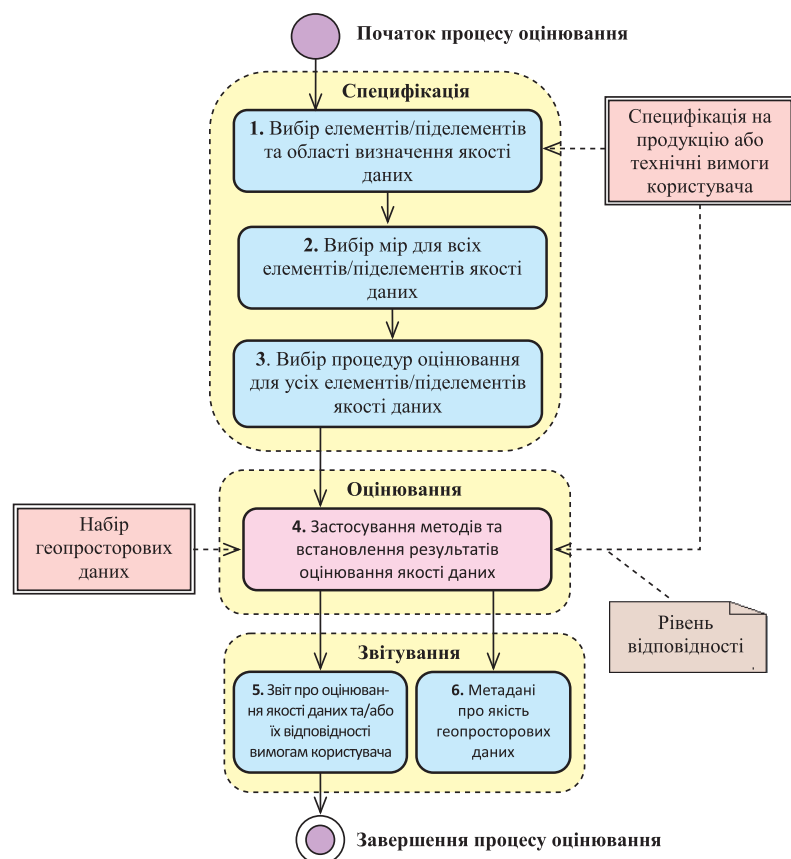
Процес оцінювання якості – це послідовність етапів, операцій та процедур, виконання яких дозволяє отримати результат як сукупність елементів якості для визначеної області (набору даних, окремих екземплярів об'єктів, їх атрибутів або відношень). Загалом процес складається із шести основних етапів (мал. 5).

Підкреслимо важливість перших трьох етапів, результатом яких є фактично специфікація моделі якості для процесу оцінювання якості конкретного набору даних з обґрунтуванням вибору елементів/піделементів, мір якості для них та процедур і методів оцінювання. Ця модель має узгоджуватися із вхідною моделлю якості, що зазначається у специфікації на продукцію або в технічних вимогах користувача.

Процедури оцінювання якості визначають порядок застосування одного або декілька методів оцінювання для певного елемента/піделемента якості даних. Методи оцінювання якості поділяються на два основних класи: прямі та непрямі. *Прямі* методи ґрунтуються на порівнянні оцінних

Таблиця 3. Базові міри якості, основою яких є підрахунок кількості помилок

Назва міри якості	Визначення	Приклад	Тип значення
Індикатор помилки	Показник того, що об'єкт виявлено як помилковий	False	Логічне
Індикатор правильності	Показник того, що об'єкт виявлено як правильний	True	Логічне
Кількість помилок	Загальна кількість помилкових об'єктів у наборі	11	Ціле
Кількість правильних об'єктів	Загальна кількість правильних об'єктів у наборі	15 571	Ціле
Відсоток помилок	Відношення кількості помилкових об'єктів до їх загальної кількості	0,0189	Дійсне
Відсоток правильних об'єктів	Відношення кількості правильних об'єктів до їх загальної кількості	0,9811	Дійсне



Мал. 5. Загальна схема процесу оцінювання якості геопросторових даних

даних з еталонними, *непрямі* – на використанні довідкової інформації про дані, наприклад, про їх походження.

Прямі методи за джерелом інформації, необхідної для проведення порівняння, в свою чергу поділяються на внутрішні та зовнішні. Всі дані, необхідні для виконання внутрішнього прямого методу, є внутрішніми відносно оцінюваного набору даних. Наприклад, для перевірки топологічної узгодженості меж земельних ділянок потрібні лише дані про межі. Зовнішні методи потребують еталона для порівняння. Приміром, для перевірки повноти атрибутів інформації про назви населених пунктів потрібно мати додаткове джерело інформації про ці назви, а для перевірки позиційної точності потрібно мати координати еталонних (контрольних) точок.

Непрямі методи ґрунтуються на зовнішніх відомостях. Наприклад, інформацію про позиційну точність можна одержати, якщо відомі методи вимірювання координат точок. Проте методи непрямого оцінювання рекомендується застосовувати лише у випадках, коли не можна використати прямі методи.

Як для внутрішніх, так і для зовнішніх методів можливе застосування вибіркового або повного контролю. Повний контроль передбачає перевірку всіх об'єктів з набору даних. Його доцільно використовувати для невеликих наборів або у випадках, коли контроль можна здійснити автоматизо-

ваними засобами за допомогою комп'ютерних програм. В інших випадках доцільно проводити вибірко-вий контроль порівнянням певної частини набору даних з еталоном. Так, як правило, перевіряється близько 5 % даних.

На завершальних етапах процесу оцінювання якості зазвичай виводиться узагальнений показник якості даних – ADQR (Aggregated Data Quality Result), в якому агрегуються окремі результати оцінювання елементів/піделементів якості.

До різновидів узагальненого показника якості даних належить [6]:

1. *Стовідсоткова відповідність/невідповідність*, що визначається за формулою

$$ADQR = \prod v_i, \quad i = 1, 2, \dots, n,$$

де v_i – логічне значення елемента/піделемента якості; n – кількість елементів/піделементів узагальнення.

Кожен елемент v_i має логічне значення відповідності: 1 – відповідає вимогам, 0 – їм не відповідає. Якщо всі елементи відповідають вимогам, то узагальнений результат буде: "100 %-ва відповідність", якщо ж хоч один елемент не задовольняє вимог, то результат буде "Не відповідає";

2. *Зважена відповідність/невідповідність*, що визначається для відносних кількісних оцінювань елементів якості за формулою

$$ADQR = \sum v_i w_i, \quad i = 1, 2, \dots, n,$$

де v_i – відносна оцінка елемента/піделемента якості; w_i – коефіцієнт ваги, що встановлюється для кожного i -го елемента/піделемента в залежності від його значимості; n – кількість елементів/піделементів узагальнення.

Результат відображатиме ступінь наближення одержаної якості до еталонної.

3. *Максимум/мінімум* – показник, що відображує максимальну/мінімальну ступінь наближення оцінюваних даних до еталона (або вимог), але в аспекті певного елемента якості; формально описується як

$$ADQR = \text{MAX}(v_1, v_2, v_3, \dots, v_n) \text{ або}$$

$$ADQR = \text{MIN}(v_1, v_2, v_3, \dots, v_n).$$

Система забезпечення якості геопросторових даних. Відомо, що контроль якості продукції на кінцевій стадії виробничого процесу дає неадекватну оцінку. Він не може ефективно впливати на поліпшення якості продукції, а тому необхідна система наскрізного контролю якості всіх елементів на всіх етапах виробничого процесу. Наскрізний контроль охоплює загалом стан виробництва та визначається за ISO 8402 як підхід "Повна



якість". Реалізація даного підходу вимагає використання системи управління якістю QMS (Quality Management System), зокрема визначення повноважень, відповідальності та порядку взаємодії всіх учасників та всіх основних процесів і процедур, важливих для досягнення якісного виконання завдань організації та проведення довгострокової сталої політики якості.

Систему забезпечення якості геопросторових даних можна визначити як сукупність організаційно-технологічних заходів, методик та спеціального програмного забезпечення для контролю якості даних на усіх етапах збирання, введення, оброблення і використання ГД, що реєструються, накопичуються та оновлюються в БГД й експортуються у зовнішні системи як НГД та/або цифрові й електронні карти, створені на основі вмісту БГД. Контроль якості даних є обов'язковою складовою технологічної ланки в реалізації будь-якого способу отримання, введення, реєстрування та виведення ГД.

Зважаючи на широке застосування цифрових технологій збирання даних, інформаційно-модельну сутність ГД та великі обсяги реальних наборів даних, очевидно є значимість створення автоматизованих систем вироблення даних і контролю їх якості. Фактично йдеться про створення єдиної інформаційної інфраструктури виробництва даних, в якій каталогізуються уніфіковані міри та методи оцінювання якості, створюються специфікації з детально вписаними вимогами до якості даних і формалізованими описами моделей оцінювання якості, фіксуються метадані про всі об'єкти й атрибути на всіх стадіях вироблення даних, функціонують бази даних каталогів геопросторових об'єктів та їх атрибутів з формалізованим описом класифікаторів, відношень об'єктів та доменів значень атрибутів, різноманітні бази довідкових і кадастрових даних, топографічного моніторингу місцевості, доступні дані та метадані про вироблені набори даних з результатами оцінювання їхньої якості тощо.

Створення такої інформаційної інфраструктури та автоматизованих систем виробництва геопросторових даних високої якості, безперечно, можна віднести до найважливіших та актуальних завдань геоінформатики.

Висновки. Якість географічної інформації завжди була важливою проблемою в геоматиці як сфері, що займається вимірюванням просторових властивостей об'єктів реального світу, формуванням знань про навколишнє середовище. Її значення зросло в останні роки з розвитком інтернет-технологій, підвищенням доступності географічних даних і систем, а також в результаті використання географічних даних у цифрових форматах від багатьох виробників для різних застосувань.

Розробляючи або застосовуючи певні методи чи процедури оцінювання якості геопросторових даних, завжди необхідно пам'ятати про їх призначення (цільовий контекст даних), для чого їх ви-

робили та де планується використовувати. Такий підхід дуже важливий, оскільки ГД, як підкреслювалося вище, – це інформаційна модель географічного середовища, що визначається потребами певної предметної сфери та яка створюється за відповідними прикладними вимогами до структури й складу об'єктів моделі.

З розвитком ІГД проблеми, пов'язані з якістю даних, позначаються на розвитку всіх сфер, в яких використовується географічна інформація. Хоча багато хто асоціює термін "якість" у геоматиці тільки з просторовою точністю зібраних даних, проте це поняття охоплює набагато ширший спектр елементів і впливає на весь процес збирання, управління, комунікації та використання географічних даних. Геопросторові дані як сучасні складні інформаційні моделі реального світу потребують створення адекватних автоматизованих систем їх виробництва, керування, контролю й забезпечення якості.

Важливе значення для забезпечення якості широкого спектра геопросторових даних є вироблення та постачання високоякісних наборів топографічних даних і на їх основі базових геопросторових даних ІГД, оскільки саме вони складають фундамент просторових моделей практично усіх інших профільних ГД. Підтвердженням цього положення може служити, зокрема, значна увага, що приділяється якості даних у специфікаціях на вироблення БНГД Європейської інфраструктури геопросторових даних INSPIRE [14].

У спеціальних стандартах з якості даних комплексу ISO 19100 уніфіковано принципи, елементи, міри, процедури та методи оцінювання якості ГД. Ці стандарти складають методичну основу реалізації сучасної системи оцінювання та забезпечення якості ГД на стадії їх виробництва. Методологія проектно-, клієнт- і системно-орієнтованого підходів до управління якістю геопросторових даних та геоінформаційних послуг знаходиться ще у стані становлення, усвідомлення проблем і пошуку напрямів їх вирішення.

Література

1. *ГОСТ Р 51608-2000*. Карты цифровые топографические. Требования к качеству [Текст]. – М.: ИПК "Издательство стандартов", 2000.
2. *ДСТУ ISO 9000-2001*. Системи управління якістю. Основні положення та словник [Текст]. – К.: Держстандарт України, 2001.
3. *Bedard, Y.* Qualité des données à référence spatiale dans un contexte gouvernemental: Rapport de recherche [Text] / Y. Bedard, D. Valliere. – Université Laval, Québec, Canada. – 1995. – 127 p.
4. *Fundamentals of Spatial Data Quality* [Text] / Edit by R. Devillers, R. Jeansoulin // ISTE Publishing. – 2006. – 309 p.
5. *Guidelines for Implementing the ISO 19100 Geographic Information Quality Standards in National Mapping and Cadastral Agencies*. EuroGeographics Expert Group on



Quality [Text] / Edited by Antti Jakobsson, Jørgen Giversen // EuroGeographics. – 2007. – 68 p. [http://www.eurogeographics.org/documents/Guidelines_ISO_19100_Quality.pdf].

6. *ISO/DIS 19157*: Geographic information – Data quality [Text]. – International Organization for Standardization, 2011.

7. *Jakobsson, A.* The Role of Quality in Spatial Data Infrastructures [Text] / A. Jakobsson, L. Tsoulos // In : Proceedings of the 23rd International Cartographic Conference, Moscow, Russia [Cd-Rom]. – 2007.

8. *Kothuri, R.* Pro Oracle Spatial for Oracle Database 11g [Text] / R. Kothuri, A. Godfrind, E. Beinart. – Apres, New York, 2007. – 787 p.

9. *Obe, R.O.* PostGIS in Action [Text] / R.O. Obe, L.S. Hsu. – Manning Publications, 2011. – 450 p.

10. *Oort, P.A.J. Van.* Spatial data quality: from descrip-

tion to application: PhD Thesis [Text]. – Wageningen: Wageningen Universiteit, 2006. – 125 p.

11. *The accuracy* of spatial databases [Text] / Edited by M. Goodchild and S. Gopal // London, New York, Philadelphia: Taylor & Francis e-Library, 2005. – 193 p.

12. *Zhang, J.* Uncertainty in Geographical Information [Text] / J. Zhang, M. F. Goodchild // London, Taylor and Francis. – 2002. – 192 p.

Інтернет-джерела

13. *ArcGIS Data Reviewer*. – <http://www.esri.com/software/arcgis/extensions/arcgis-data-reviewer/index.html>

14. *INSPIRE DS-D2.6*. Methodology for the development of data specifications, v3.0. – http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.6_v3.0.pdf

Надійшла 20.06.12

* * *

КАЛЕНДАР ПОДІЙ

Назва заходу	Дата і місце проведення	Веб-сайт / контактна адреса
27-ма сесія Групи експертів ООН з географічних назв	30 липня – 10 серпня 2012 р. м. Нью-Йорк (США)	http://unstats.un.org/unsd/geoinfo/UNGEGN/default.html
10-та конференція зі стандартизації географічних назв	31 липня – 9 серпня 2012 р. м. Нью-Йорк (США)	http://unstats.un.org/unsd/geoinfo/UNGEGN/default.html
32-й Міжнародний географічний конгрес (IGC Cologne 2012)	26-30 серпня 2012 р. м. Кельн (Німеччина)	http://www.igc2012.org/
XXII Конгрес Міжнародного товариства фотограмметрії та дистанційного зондування (ISPRS)	25 серпня – 1 вересня 2012 р. м. Мельбурн (Австралія)	http://www.isprs2012-melbourne.com/
Генеральна асамблея Eurogeographics	3-5 вересня 2012 р. м. Гельсінкі (Фінляндія)	http://www.eurogeographics.org/event/eurogeographics-general-assembly-2012
XXI Міжнародна наукова конференція на тему: "Інновації у географічній освіті й картографії"	10-13 вересня 2012 р. м. Харків (Україна)	Кафедра фізичної географії та картографії, Харківський національний університет імені В.Н. Каразіна, площа Свободи, 4, м. Харків, 61022
V Всеукраїнська науково-практична конференція "Національне картографування"	13-14 вересня 2012 р. м. Київ (Україна)	http://www.ignau.org.ua/conference.htm
INTERGEO 2012	9-11 жовтня 2012 р. м. Ганновер (Німеччина)	http://www.intergeo.de/de/index.html
Франкфуртський книжковий ярмарок	9-15 жовтня 2012 р. м. Франкфурт (Німеччина)	http://www.buchmesse.de
XXXV Всепольська картографічна конференція	25-26 жовтня 2012 р. м. Познань (Польща)	http://www.kartografia.amu.edu.pl/konferencje/XXXV_OKK/XXXV_OKK_zaproszenie.pdf
25-та Міжнародна картографічна конференція з історії картографії	30 червня – 5 липня 2013 р. Гельсінкі (Фінляндія)	ichc2013.fi
26-та Міжнародна картографічна конференція	25-30 серпня 2013 р. Дрезден (Німеччина)	http://icaci.org/icc2013/