

ГІС-ТЕХНОЛОГІЇ В КАРТОГРАФІЇ

УДК 528.92:91(075.8)

Бондаренко Е. Л.

Київський національний університет імені Тараса Шевченка

ПРОБЛЕМИ ІНТЕГРАЦІЇ ДАНИХ У ГЕОІНФОРМАЦІЙНОМУ КАРТОГРАФУВАННІ

Охарактеризовані інтеграційні можливості сучасних географічних інформаційних систем. Визначено види інтеграції даних, групи проблем, які виникають в процесі використання даних у геоінформаційному картографуванні та шляхи їх вирішення. Вказано на можливі рівні інтеграції даних при створенні картографічних моделей.

Ключові слова: інтеграція даних, географічна інформаційна система, геоінформаційне картографування, рівні інтеграції даних.

Вступ. Географічні інформаційні системи (ГІС), що є гнучким та універсальним засобом для вирішення різноманітних конструктивних задач, сьогодні все ширше застосовуються в різних сферах людської діяльності, зокрема, в науках про Землю, які мають практичне застосування. Насамперед, це стосується задач, пов'язаних з охороною навколишнього природного середовища, зокрема, моніторингом, оцінкою та управлінням природними ресурсами, аналізом соціально-економічного та екологічного стану територій та прогнозуванням його розвитку.

Серед важливих функцій ГІС доцільно виокремити функцію *інтеграції існуючих потоків даних в єдину систему географічної інформації* про територію, на основі якої може проводитись комплексна оцінка та прогноз стану природного середовища, інтелектуальна підтримка та оптимізація управлінських рішень, їх ув'язування з проблемами природокористування тощо.

Одним із характерних результатів функціонування ГІС є розробка на основі інтегрованої інформації багатоваріантних карт для відповідних проблемно-практичних потреб користувачів із застосуванням методів

просторового моделювання та аналізу даних.

Геоінформаційні програмні продукти по-різному виконують конструктивно закладену розробниками інтеграційну функцію для реалізації призначення та одержання практичних результатів. Актуальність же проблеми інтеграції даних пов'язана із широким використанням у середовищі ГС існуючих цифрових карт, інформації, що міститься в різноманітних базах просторових даних (створеної в різних програмних продуктах) і такої, що поширюється по телекомунікаційних мережах. Виходячи з цього, актуальність даної статті полягає у спробі визначення особливостей інтеграційних можливостей та рівнів інтеграції цифрових даних у багатофункціональних ГС, що розглядаються на прикладі програми MapInfo.

Аналіз останніх досягнень та публікацій. Проблема інтеграції цифрових даних тривалий час обговорюється в колі фахівців як в Україні, так і в країнах близького зарубіжжя. Про це свідчать, зокрема публікації [5 – 9 та ін.], в яких наголошується, що сучасні геоінформаційні технології створили умови для розвитку геоінформаційного картографування з новим видом продукції у вигляді баз просторових даних, розробка яких є результатом високих технологій і потребує створення інфраструктур просторових даних (на різних рівнях, зокрема для геоінформаційного еколого-географічного картографування, обґрунтованих в [6, 7]), за умов їх чіткої структуризації і географічно обґрунтованому застосуванні ГС-технологій при роботі з різнотипними даними.

Серед авторів, які досліджують проблеми інтеграції цифрових даних для геоінформаційного картографування доцільно виділити останні роботи О.В. Кошкарьова [3, 4], І.К. Лур'є [8], О.І. Мартиненка, Ю.Л. Бугаєвського, С.М. Шибалова [9], Ю.О. Карпінського, А.А. Лященко [5]. Всі ці праці можна віднести до робіт теоретичного плану.

Формулювання цілей. Незважаючи на низку вирішених теоретико-методологічних питань щодо інтеграції просторових даних, існують питання в практичному плані висвітлені ще недостатньо. Тому є доцільним:

1. Визначити інтеграційні можливості багатофункціональних ГС.
2. Вказати помилки, які можуть виникати при інтеграції даних в процесі геоінформаційного картографування та шляхи їх усунення.
3. Визначити можливі рівні інтеграції цифрових даних, що застосовуються в ГС.

Виклад основного матеріалу. *Цифрові просторові дані* представляють собою інформацію в цифровій формі про об'єкти, які включають відомості про місце їх розташування та властивості, їхні просторові та непросторові атрибути. За структурою ці дані складаються з наступних характеристик: *позиційної*, що описує просторове положення (геометрію); *топологічної*, що визначає взаємну підпорядкованість окремих об'єктів множини;

непозиційної (характеризує тематичний зміст складової даних), які в сукупності дають повний опис просторових даних і становлять основу інформаційного забезпечення ГІС [10].

Різні види та типи цифрових просторових даних, що використовуються в процесі геоінформаційного картографування, потребують розробки методів їх сумісного використання. Це можливо через процедуру **інтеграції**, яка передбачає визначення певних правил щодо досягнення можливості одночасного та спільного використання геоінформаційною системою або запитом користувача системи декількох файлів як єдиного цілого [1, 2].

До інтеграційних можливостей сучасних багатофункціональних географічних інформаційних систем слід віднести:

- наявність функцій безпомилкового відкриття даних, створених у різних форматах;
- здатність геоінформаційної системи виправляти помилки в різноманітних типах даних, створених в різних форматах та геоінформаційних системах;
- наявність системи запитів для формування виразів, здійснення перетворень даних, об'єднання даних у процесі моделювання;
- існування процедури надання непросторовим даним просторової (координатної) прив'язки;
- наявність додаткових модулів ГІС – універсальних конвертерів форматів;
- здатність ГІС зберігати дані в обмінних форматах.

Аналіз перерахованих інтеграційних можливостей ГІС доводить, що інтеграція в залежності від *методу здійснення* може бути **логічною** та **фізичною**, а в залежності від *способу її проведення* **внутрішньою** та **зовнішньою**.

Логічна інтеграція допускає об'єднання даних на логічному рівні, не порушуючи їхньої фізичної організації, тобто такі дані знаходяться в окремих файлах бази даних і використовуються лише в окремому сеансі моделювання в процесі геоінформаційного картографування. Таким чином, можна стверджувати, що логічна інтеграція передбачає використання даних, які будуть знаходитися в так званому «невяному» вигляді. В будь-який момент такі дані можуть бути збереженими в окремі файли в пам'яті комп'ютера, що передбачає їх злиття до єдиного інформаційного масиву (наприклад, одного файлу), тобто проведення *фізичної інтеграції*.

Внутрішня інтеграція передбачає об'єднання та сумісне використання цифрових даних в процесі геоінформаційного картографування в середовищі географічної інформаційної системи.

Зовнішня інтеграція передбачає застосування спеціальних програм (модулів) для трансляції цифрових даних та їх конвертування з одного формату до іншого, як правило внутрішнього формату багатофункціональної

ГІС. Наприклад, зовнішня інтеграція в середовищі MapInfo передбачає використання модуля інтеграції – універсального транслятора даних, – який дозволяє її здійснення перед початком процесу геоінформаційного картографування. При цьому для перетворення даних до внутрішнього формату MapInfo вихідні дані можуть мати наступні формати створення: dxf, dwg (формати програми AutoCad), e00, shp (ESRI), dgn (MicroStation Design), sdts (Spatial Data Transfer Standard), vpf (Vector Product Format). Для конвертації даних з формату MapInfo можуть бути отримані дані у форматах: dxf, dwg (формати програми AutoCad), e00, shp (ESRI), dgn (MicroStation Design).

В процесі інтеграції об'єктивно виникають проблеми, які можна звести до наступних груп:

– *недоступність та/або відсутність інформації щодо походження джерел даних, методів створення, точності і вірогідності* (вказує на тенденцію нарощування потужності потоку інформації, що актуалізує забезпечення сумісності і порівнянності даних, їх взаємо обміну);

– *різна якість цифрових даних* (дані можуть містити помилки різного роду, які впливають з комбінації неточностей у джерелах даних, лімітованої точності пристроїв цифрування чи специфіки комп'ютерних систем, які використовуються).

Типи помилок, що виникають, наприклад, через неточності в оригіналах даних, картографічних джерел, помилкових дій оператора при ручній векторизації графічних об'єктів, не можуть бути виправлені без одержання інших джерел даних. Особливо це стосується помилок позиціонування, які часто не можуть бути виявлені навіть при співставленні різних джерел. Більшість помилок векторизації усувається на основі візуального контролю та інтерактивних дій оператора.

На якість даних впливають: помилки реєстрації і визначення контрольних точок, перетворення координат, особливо коли невідома проекція вихідного джерела; помилки обробки даних, неправильний логічний підхід, генералізація і проблеми інтерпретації; математичні помилки; втрата точності представлення через невисоку точність обчислень; векторно-растрові та растрово-векторні перетворення.

Технологія створення цифрових карт в країні сьогодні визначається тимчасовими, не затвердженими на загальнодержавному рівні, не узгодженими, не завжди професійно складеними інструкціями і технічними завданнями або відомчими інструкціями. При цьому часто не враховується складність і неоднозначність інтеграції даних, що представляються на карті позамасштабними та оригінальними умовними знаками. Все це призводить до помилок у цифрових картах, а іноді до їх повної непридатності у використанні, до їх ненадійності як джерел інформації.

Таким чином, часто можливості ГІС-технологій щодо визначення

розташування і просторових взаємодій складових реальних геосистем на основі інформації, отриманої з різних джерел, визначають їх придатність для інтеграції інформації в просторовому контексті. У той же час, технічно легке комбінування таких даних може призвести до помилкових результатів при аналізі через їхню різну якість і непорівнянність.

Незважаючи на те, що інтерактивність ГІС-технологій дозволяє сполучити методи створення і використання карт, автоматичні і картографічні методи класифікації і генералізації об'єктів і явищ, вказані помилки і неточності призводять до суб'єктивізму і помилковості отриманих результатів.

Перевага геоінформаційних методів полягає в можливості оцінити придатність даних з метою їх спільного використання і здійснити їхню інтеграцію на основі виконання просторового аналізу за допомогою процедур накладання шарів, побудови буферних зон. Але при цьому необхідно притримуватись основного правила при інтеграції інформації: якість даних повинна бути визначена раніше, в процесі їх накопичення, а не в процесі їх застосування. Тільки при такій умові зазначені технології можуть істотно полегшити їхнє коректування для поставленої задачі.

Основними проблемами, що виникають при спільному використанні різнотипних даних є визначення реального відображення положення границь у різних цифрових джерелах, а також часові параметри даних. При цьому необхідно вирішити два питання:

- наскільки правильно цифрові структури, що представляються в базах даних, відбивають реальну ситуацію (моделюють реальність);
- наскільки точно алгоритми, що використовуються, дозволяють розрахувати правильне значення результату співставлення даних.

Окресливши основні проблеми, що виникають при інтеграції даних та шляхи їх вирішення далі слід розглянути можливі рівні інтеграції даних.

В залежності від обсягів та особливостей представлення цифрових даних їхню інтеграцію можна здійснювати на різних рівнях, що відповідають складності даних. Це відповідно рівні інтеграції за ступенем складності, розпочинаючи з найпростішого: файловий, профілю бази даних, бази даних.

Найпростішим є *файловий* (що відповідає системному рівню представлення) *рівень інтеграції*, який може бути розглянутий на прикладі застосування багатofункціональної ГІС MapInfo 8.0. До переваг цієї системи відноситься властивість сумісності з іншими ГІС, що є однією з інтеграційних їх можливостей. Оскільки часто файли баз даних можуть створюватися та зберігатися в різних форматах (наприклад, *.dbf програми dBase, *.xls (MS Excel), *.mdb (MS Access) тощо) існує можливість в середовищі MapInfo файли цих форматів відкривати без додаткових трансляцій, про що вже було вказано вище.

При вирішенні, зокрема, задач екологічного напрямку в процесі геоінформаційного картографування та наявності баз даних, наприклад, по підприємствах-забруднювачах навколишнього середовища, що зберігаються в форматі *.xls, база даних містить інформацію про код та назву підприємства, назву населеного пункту, де знаходиться підприємство.

Для нанесення відповідних даних на географічну основу та моделювання електронної карти їм необхідно надати географічної прив'язки, тобто відповідним записам непросторових даних треба призначити координати або прив'язати нові дані до інших вже існуючих просторово визначених об'єктів. Сутність зводиться до визначення місцеположення для нових об'єктів шляхом співставлення одного з атрибутів в новій базі даних з аналогічним записом в таблиці вже існуючої карти (географічної основи). Процес знаходження пар таких відповідностей в теорії геоінформатики та геоінформаційного картографування називається *геокодуванням* [3].

Досвід роботи автора в геоінформаційній системі MapInfo показує, що процес геокодування краще виконувати через процедуру об'єднання непросторових даних з просторовими за ключовими полями в реляційній базі даних, що дозволяє в повній мірі застосовувати усі відомі в картографії способи картографічного зображення на відміну від класичної процедури геокодування (використання відповідної функції, яка дозволяє наносити дані, які мають точковий характер прив'язки зображувальних засобів і обмежуються лише способами картодіаграми, локалізованих діаграм, локалізованих значків). Результатом файлового рівня інтеграції даних є надання непросторовим об'єктам, які містяться в окремих файлах бази даних просторових координат з метою нанесення їх на картографічну модель.

Середнім і значно складнішим рівнем інтеграції даних є інтеграція на *рівні профілю бази даних* (асоціюється з геометричним рівнем представлення), сутність якого може бути розглянуто на прикладі бази даних (профілю) адміністративно-територіального поділу країни і заснована на проведенні процесу геокодування із застосуванням кодів класифікатора об'єктів адміністративно-територіального устрою України (КОАТУУ) в модулі BNDCOD, фізично реалізованому у форматі MS Access.

Для здійснення інтеграції даних на рівні профілю бази даних та для уніфікації обміну інформацією потрібно мати єдиний класифікатор населених пунктів. Структура вхідних баз даних може бути довільною. Назви полів, за якими буде проводитись кодування, зазначаються інтерактивно. В результаті роботи початкова таблиця доповнюється стовпчиком з кодом КОАТУУ. При цьому програма дозволяє здійснювати визначення коду КОАТУУ за назвами області, району, сільради, населеного пункту та типом населеного пункту (назви можуть бути подані однією з трьох мов – українською, російською або англійською); автоматичне кодування для

даних з повним набором розпізнавальної інформації; автоматичне кодування для даних, які відносяться до однієї області або одного району або одного населеного пункту; інтерактивне кодування для даних, які не містять достатньо інформації для здійснення автоматичного кодування; можливість зазначення альтернативних назв населених пунктів; можливість здійснення кодування у випадку, коли ідентифікаційні поля містять зайву інформацію; можливість одночасної роботи з великою кількістю таблиць.

Інтеграція на третьому, найскладнішому рівні, *рівні бази даних*, що асоціюється з географічним рівнем представлення застосовується підхід, який базується на використанні еталонних некартографічних баз даних просторових об'єктів.

Задача інтеграції з використанням цього рівня виникає при наявності кількох різних баз даних, що містять інформацію про одні і ті ж об'єкти; при побудові інтегрованих тематичних баз даних, що поєднують інформацію, отриману з різних джерел, різними організаціями та в різні періоди часу; при переході від локальних рішень до систем принаймні рівня підрозділу чи підприємства.

Реалізація даного рівня інтеграції зводиться до послідовного вирішення наступних задач, що включають:

- розробку еталонної некартографічної бази даних просторових об'єктів;
- розробку вимог до локальних тематичних баз даних та обмінних форматів тематичних даних;
- розробку процедури розповсюдження еталонних баз даних та процедури внесення змін;
- розробку картографічної бази даних.

Висновки та перспективи подальших розробок. Розглянута інтеграційна функція багатфункціональних ГІС дозволить більш повно використовувати можливості вказаних програм для геоінформаційного картографування різноманітних об'єктів та явищ і підвищити якість електронних карт та їх графічних копій.

Охарактеризовано та проаналізовано помилки, які можуть виникати при інтеграції даних в процесі геоінформаційного картографування та шляхи їх усунення.

Визначено можливі рівні інтеграції цифрових даних, що застосовуються в ГІС та практичні результати, які можна одержати в залежності від складності даних.

Подальші дослідження повинні бути зорієнтовані на моніторинг геоінформаційних систем з визначенням їх додаткових інтеграційних можливостей, що без сумніву приведе до збагачення теоретико-методологічних основ геоінформаційного картографування з досягненням нових практичних результатів.

Рецензент – доктор географічних наук, професор В. О. Шевченко

Література:

1. *Бондаренко Е.Л.* Геоінформаційне еколого-географічне картографування. – К.: Фітосоціоцентр, 2007. – 272 с.
2. *Бондаренко Е.Л.* Наукові принципи створення та ефективного функціонування універсальних ГІС // Часопис картографії: зб. наук. пр. – 2011. – Вип. 1. – С. 70–78.
3. *Геоинформатика.* Толковый словарь основных терминов / Ю. Б. Баранов, А. М. Берлянт, Е. Г. Капралов и др. – М.: ГИС-Ассоциация, 1999. – 204 с.
4. *Геоинформатика: Учеб. для студ. вузов / Е.Г. Капралов, А.В. Кошкарёв, В.С. Тикунов и др.; Под ред. В.С. Тикунова.* – М.: Издательский центр «Академия», 2005. – 480 с.
5. *Карпінський Ю.О., Лепетюк Б.Д., Лященко А.А.* Формування національної інфраструктури просторових даних – основа для розвитку ГІС в Україні // Матеріали Міжнародної науково-практичної конференції «ГІС-форум 2001», Київ, 18 – 20 грудня 2001 р. – К., 2001. – С. 8 – 18.
6. *Кошкарёв А.В.* Инфраструктуры пространственных данных. – М.: ГИС-обозрение. – 2000. – № 3, 4. – С. 5 – 10.
7. *Кошкарёв А.В.* Инфраструктуры пространственных данных. – М.: ГИС-обозрение. – 2001. – № 1. – С. 28 – 32.
8. *Лурье И.К.* Методы интеграции пространственных данных // Картографический метод и возможности компьютерных систем. – Варшава, 2001. – С. 36 – 48.
9. *Мартыненко А.И., Бугаевский Ю.Л., Шибалов С.Н.* Основы ГИС: теория и практика. – М.: Астра семь, 1995. – 100 с.
10. *Шевченко В.О., Бондаренко Е.Л.* Критерії визначення якості цифрових просторових даних для геоінформаційного картографування // Картографія та вища школа: Збірник наукових праць. Вип. 10. – 2005. – С. 41 – 48.

Э. Л. Бондаренко

ПРОБЛЕМЫ ИНТЕГРАЦИИ ДАННЫХ В ГЕОИНФОРМАЦИОННОМ КАРТОГРАФИРОВАНИИ

Рассмотрены интеграционные возможности современных географических информационных систем. Определены виды интеграции данных, группы проблем, которые возникают в процессе использования данных в геоинформационном картографировании, пути их решения. Указано на возможные уровни интеграции данных при создании картографических моделей.

Ключевые слова: интеграция данных, географическая информационная

система, геоинформационное картографирование, уровни интеграции данных.

E. Bondarenko

PROBLEMS OF DATA INTEGRATION IN GEOINFORMATION MAPPING

Integration possibilities of modern geographical intelligence systems are considered. Sorts of integration of the data, groups of problems, which arise during usage of the data in geoinformation mapping, paths of its solution are defined. It is indicated on possible levels of data integration at creation of cartographical models.

Keywords: integration of the data, a geographical intelligence system, geoinformation mapping, levels of data integration.

Надійшла до редакції 17 травня 2011 р.