Геопросторові бази даних.

План

- **Передумови виникнення об’єктно-орієнтованих баз даних.**
- **Порівняльна характеристика реляційних та об’єктно-орієнтованих баз даних.**
- **Поняття бази геопросторових даних: визначення, призначення та функції БГД.**
- **Етапи проектування баз геопросторових даних.**

ГІС і системи баз даних

"ГІС – це потужний набір інструментів для **збору**, **зберігання**, пошуку за бажанням, перетворення і відображення просторових даних з реального світу." (Burrough 1986)

"ГІС – це системи баз даних, в яких велика частина даних просторово індексуються, існує набір процедур для того, щоб відповісти на питання про просторові сутності в базі даних" (Сміт та ін., 1987)

Керування геоданими

ГІС дані

Географічні (просторові) дані

- *положення* (координати карти)
- *топологія* (відношення сусідства)
- *графічне подання* (компоновка карти)

Описові (атрибутивні) дані

- добре структуровані (таблиці)
- мультимедіа
- інші документи

Дані в ГІС поділяються на просторові і описові дані завдяки:
різне подання – карти / таблиці
різні вимоги – топологічні / логічні

Файлові структури і бази геоданих

Традиційний підхід: на основі файлів (SHP, DXF, MIF, ...) через відсутність підтримки просторових даних в системах баз даних

Обмеження підходу на основі файлів

Поділ та ізоляція

Копіювання даних

Залежність даних

Несумісність файлів

Основні запити / поширення прикладних програм

сьогодні:

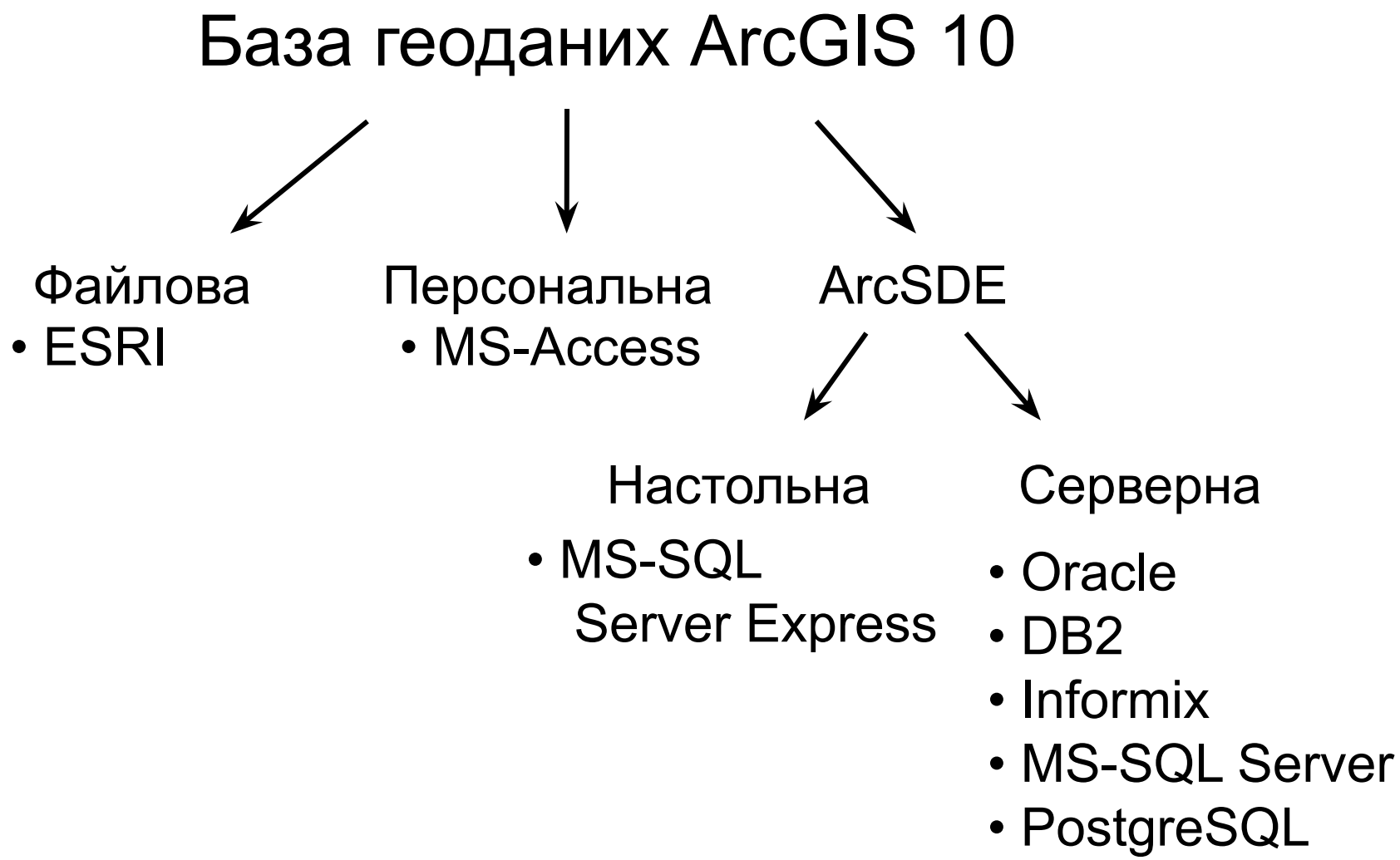
вбудовані реляційна / об'єктно-реляційна системи баз даних з підтримкою просторових даних (корпоративні ГІС)

однокористувальницька персональна база геоданих без підтримки просторових даних (так само, як сховище даних)

GIS Системи баз даних з просторовими розширеннями

Назва продукту	Компанія	Операційні системи (Серверні)	Приклади використання
Oracle 11g Locator/Spatial	Oracle Corp.	UNIX, Linux, Windows	ArcSDE, GeoMedia, Autodesk MapGuide, MapInfo,, …
Dynamic Server 11 (IDS) Spatial DataBlade	Informix / IBM	UNIX, Windows, Linux	ArcSDE, MapInfo, Autodesk MapGuide
DB2 V9 Spatial Extender	IBM	UNIX, Linux, Windows	ArcSDE, MapInfo, GeoMedia, GeoServer, uDig
Ingres 2006 Object Management Extension (Spatial Data Types)	Ingres Corp.	Linux, Windows, VMS	ArcSDE, MapInfo
MS SQL Server 2008	Microsoft	Windows	Autodesk MapGuide, ArcSDE, GeoMedia, MapInfo …
SQL Anywhere 12 spatial query server	Sybase	UNIX, Linux,Windows	Autodesk MapGuide, MapInfo, ArcSDE
PostgreSQL 9 PostGIS 1.5	Refractions Open source	UNIX, Linux, Windows	GRASS, ArcSDE, gvSIG, QGIS, MapInfo…
MySQL 5 Spatial Extension	Oracle/ Open source	UNIX, Linux, OS/2, Windows	GRASS
SQLite SpatiaLite 2.3	Open Source	UNIX, Linux, Windows	QGIS, openJump, GeoTools, OGR

ArcGIS і бази геоданих



Comparing the three types of Geodatabases

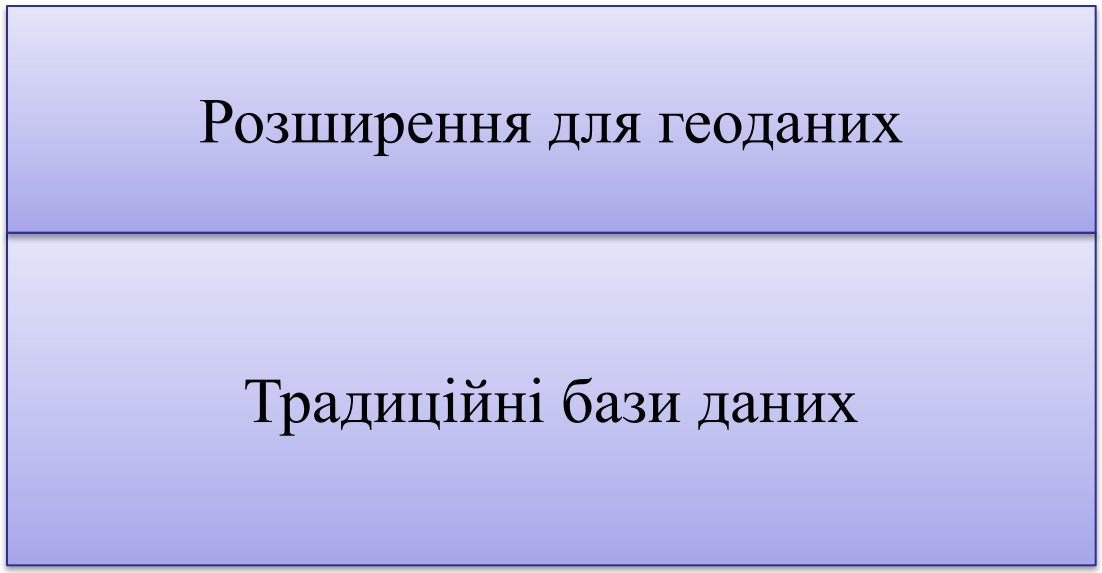
Key Characteristics	ArcSDE Geodatabase	File Geodatabase	Personal Geodatabase
Description	A collection of various types of GIS datasets held as tables in a relational database This is the recommended native data format for ArcGIS stored and managed in a relational database.	A collection of various types of GIS datasets held in a file system folder This is the recommended native data format for ArcGIS stored and managed in a file system folder.	Original data format for ArcGIS geodatabases stored and managed in Microsoft Access data files This is limited in size and tied to the Windows operating system.
Number of Users	Multiuser Many readers and many writers	Single user and small workgroups Many readers or one writer per feature dataset, standalone feature class, or table. Concurrent use of any specific file eventually degrades for large numbers of readers.	Single user and small workgroups with smaller datasets Some readers and one writer. Concurrent use eventually degrades for large numbers of readers.
Storage Format	• Oracle • Microsoft SQL Server • IBM DB2 • IBM Informix • PostgreSQL	Each dataset is a separate file on disk A file geodatabase is a file folder that holds its dataset files.	All the contents in each personal geodatabase are held in a single Microsoft Access file (.mdb).
Size Limits	Up to DBMS limits	One TB for each dataset. Each file geodatabase can hold many datasets. The 1 TB limit can be raised to 256 TB for extremely large image datasets. Each feature class can scale up to hundreds of millions of vector features per dataset.	Two GB per Access database Effective limit before performance degrades is typically between 250 and 500 MB per Access database file.
Versioning Support	Fully supported across all DBMSs; includes cross-database replication, updates using checkout and checkin, and historical archiving	Only supported as a geodatabase for clients who post updates using checkout and checkin and as a client to which updates can be sent using one-way replication.	Only supported as a geodatabase for clients who post updates using checkout and checkin and as a client to which updates can be sent using one-way replication.
Platforms	Windows, Unix, Linux, and direct connections to DBMSs that can potentially run on any platform on the user's local network	Cross-platform	Windows only
Security and Permissions	Provided by DBMS	Operating file system security	Windows file system security
Database Administration Tools	Full DBMS functions for backup, recovery, replication, SQL support, security, and so on	File system management	Windows file system management
Notes	Requires the use of ArcSDE technology ArcSDE for SQL Server Express is included with: • ArcEditor and ArcInfo • ArcGIS Engine • ArcGIS Server for Workgroups ArcSDE for all other DBMSs is included with ArcGIS Server for Enterprises.	Allows you to optionally store data in a read-only compressed format to reduce storage requirements	Often used as an attribute table manager (via Microsoft Access). Users like the string handling for text attributes.

ESRI online help

Що потрібно знати фахівцям про системи баз геоданих?

- як проектувати бази геоданих, тобто логічні елементи інформації
- як реалізувати проект, тобто технічні та структурні елементи СКБД
- як налаштувати базу даних для якісної експлуатації, тобто (просторова) індексація, керування транзакціями, керування версіями
- як працювати з базою геоданих, тобто використання SQL для отримання даних

Системи баз геоданих



проектування
структурування
настроювання
робота з
просторовими даними

проектування
структурування
настроювання
робота з
не просторовими
даними

Моделювання

У ГІС ми хочемо подати реальний світ, або, принаймні, частину реального світу, який нас цікавить. Ми будуємо модель такої частини, яка називається предметна область.

В результаті такого моделювання дані спрощуються і стають орієнтованими на комп'ютерне використання, наприклад:

- **координати як модель положення,**
- **полігон як модель земельної ділянки,**
- **ім'я власника як модель самої людини**.

Проектування бази геоданих

Метою проектування бази геоданих є забезпечення основи структур даних, де конкретні набори даних можуть бути збережені, як ортофотоплани, в космічні знімки або геологічні полігональні шари карти.

Проектування бази геоданих

В принципі, географічні дані не відрізняються від будь-якої іншої інформації
Геометрія це просто ще один атрибут

Додаткові правила:

- відокремлюються негеографічні дані від географічних
- відокремлюються грубі дані від тонких даних, наприклад, національні дані від регіональних
- відокремлюються дані відповідно до їх поданням, тобто векторні і растрові
- відокремлюються дані, засновані на формі геометрії, тобто точки, лінії, полігони

- І, нарешті, дані потрібно зберігати в конкретній базі даних, наприклад файл або персональна база геоданих ArcGIS
- Хоча концептуальне проектування має бути незалежним від будь-яких технічних проблем, проте потрібно пам'ятати, про структурні елементи конкретної бази геоданих

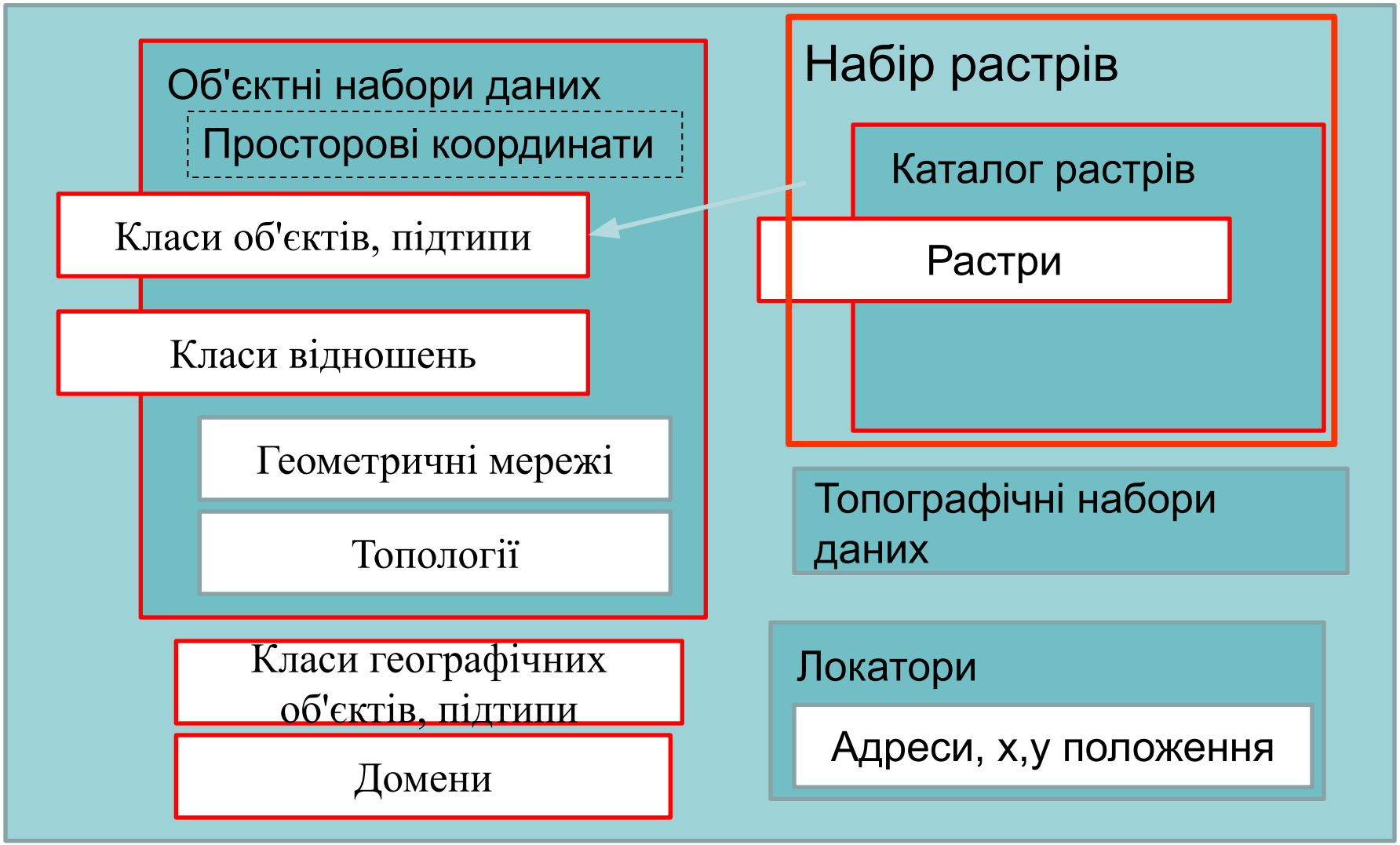
Що в середині бази геодних ArcGIS?

Структурні елементи:

- Класи географічних об'єктів (без геометрії)
- Класи інформаційних об'єктів
- Класи відношень
- Домени

- Геометричні мережі
- Топології
- Растри
- TIN
- Локатори

Структурні елементи бази геоданих ArcGIS



- # Набори даних об'єктів:

  Колекція класів об'єктів, які мають спільну систему координат. Геометричні мережі і топологічні правила можуть бути визначені тільки всередині наборів даних. Набір даних об'єктів можна також використовувати для логічного структурування геометричних даних.

- # Класи об'єктів:

  Набір об'єктів того ж типу геометрії і з тим же набором атрибутів

  – Клас Простих об'єктів: без явних топологій
  – Клас Топологічних об'єктів: з явною топологією (наприклад, геометричні мережі)

- # Підтипи:

  Підтип буде мати той же набір атрибутів, як відповідного класу об'єктів або класу об'єктів, у той час як домен або значення за замовчуванням з одного або декількох атрибутів можуть відрізнятися

Підтип

Підтипи підмножина об'єктів в класі об'єктів або об'єктів в таблиці, які поділяють одні й ті ж атрибути. Вони використовуються в якості методу для класифікації даних.

Підтипи дозволяють робити наступне:

- **Збільшувати** продуктивність бази геоданих, представляючи різні об'єкти реального світу як підмножину об'єктів в даному класі об'єктів, а не створюючи нових класів об'єктів для кожного об'єкта. Наприклад, вулиці як клас вулиці можна розділити на три підтипи: місцеві, регіональні, і магістральні вулиці.
- **Встановлювати** значення за замовчуванням, яке буде автоматично застосовуватися при створенні нових об'єктів. Наприклад, підтип квартальна вулиця може бути створена і визначена таким чином, що кожен раз, коли цей тип вулиці додають до класу об'єктів, його атрибут обмеження швидкості автоматично встановлюється на 40 км на годину.
- **Застосовувати** закодовані або ранжовавні домени об'єктів, що дозволяє обмежити введення інформації в припустимий набір значень. Наприклад, в мережі водопостачання, підтипу водопровід може мати домен кодованих для атрибуту матеріал труби, що обмежують їх бути виготовлені з чавуну або міді.
- **Створення** правил для взаємодії між іншими підтипами і класами для підтримки цілісності мережі. Наприклад, у водопровідній мережі, гідрант можна підключити до бокового гідранта, але не до будинкового водопровідного вводу.

- Геометричні мережі:

  Спеціальна структура даних у вигляді топологічного графа (модель вузол-ребро) для моделювання реальних мереж (річки, дороги, трубопроводи)

- Топології:

  Топологічні правила, які можуть бути визначені для класів об'єктів, наприклад «клас об'єктів земельні ділянки не повинні мати будь-яких прогалин чи дублювання"

- Класи географічних об'єктів:

  Таблиця База даних без геометрії, наприклад, з додатковими атрибутивними даними для класу об'єктів.

- Класи відношень:

  Реалізація відношень між Класами географічних та / або інформаційних. ArcGIS не використовує функціональні можливості основних СКБД, але керує інформацією по собі в додаткових (система) таблиць.

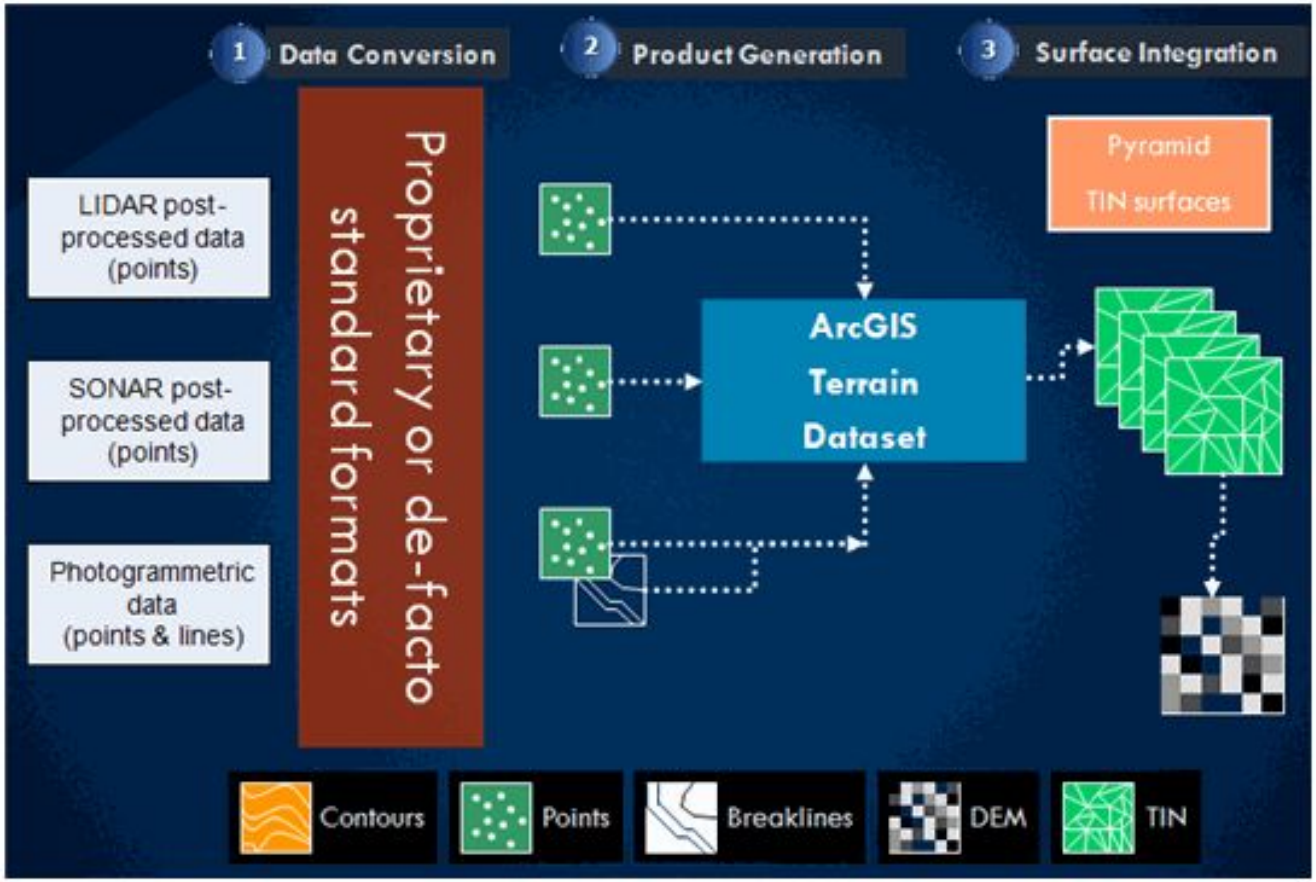
- Домени:

  Діапазон допустимих значень атрибута

  – Домен діапазону значень: Діапазон, наприклад, між 100 і 200
  – Домен кодованих значень: Список всіх можливих значень, наприклад, шосе, дорога, магістраль

- Мозаїка растрів:
  Набір даних Мозаїка дозволяє зберігати, керувати, переглядати і робити запити від малих до великих колекцій растрових даних і зображень.
- Каталог растрів:
  Колекція наборів растрових даних, які візуалізуються як один шар
- Набір растрів:
  Растрове зображення, яке складається з щонайменше одного або декількох смуг частот
- Растр
  один растр
- Локатор:
  Геокодування і розділення за Адресами

- Топографічні набори даних:
  TIN моделі з різною роздільною здатністю
  TIN: Нерегулярна тріангуляційна мережа, набір трикутників рівно з одним Z-значення для кожного вузла



Основні етапи проектування баз геоданих

- уявлення про реальність (тобто фіксація ПО)
- формалізація ПО без будь-яких конвенції чи обмежень щодо реалізації (тобто створення концептуальної моделі даних)
- подання концептуальної моделі, яка відображає те, як дані записуються в комп'ютер (тобто створення логічної моделі даних)
- структура файлу, яка э поданням структури даних в пам'яті комп'ютера (наприклад, створення фізичної моделі даних)
- загальноприйняті правила при роботі з даними (тобто створення моделі маніпулювання даними, іноді називають функціональна модель)
- загальноприйняті правила і процедури для відображення і подання просторових даних (тобто настройки графічної моделі)

11 кроків проектування баз геоданих

1.	**Визначення інформаційних продуктів, які можна створювати і керувати з ГІС.**
2.	**Визначення ключових тем даних на основі інформаційних вимог.**
3.	**Вказання діапазонів масштабів і просторових подань по кожній темі даних на кожній шкалі.**
4.	**Розкладання кожного подання на один або декілька географічних наборів даних.**
5.	**Визначення табличної структури бази даних і поведінки для описових атрибутів.**
6.	**Визначення просторової поведінки, просторових відношень, і правил цілісності для даних.**
7.	**Пропозиція проекту бази геоданих.**
8.	**Проектування редагування технологічних процесів і візуалізація властивостей.**
9.	**Розподіл обв'язків зі створення та підтримання кожного шару даних.**
10.	**Створення робочого прототипу. Перегляд і уточнення дизайну.**
11.	**Документування проекту бази геоданих.**

http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=Geodatabase_design_steps

Проектування баз геоданих

- Мови моделювання
  - ERM/EERM
  - UML
  - DSL (Предметно-орієнтована мова) (мова програмування, призначена для вирішення завдань конкретної предметної галузі, наприклад SQL-DDl логічна модель БД, XSD для XML документів та інші.)

UML i DSL

To summarize, we would recommend using UML and UML-based tools for:

- Sketching.
- White boarding.
- Documentation.
- Conceptual drawings that do not directly relate to code.

We would recommend precisely defined DSLs and DSL-based tools for:

- Precise abstractions from which code is generated.
- Precise abstractions that map to variation points in frameworks and components.
- Precise mappings between DSLs.
- Conceptual drawings that have precisely specifiable mappings to other DSLs or to code artifacts.

(Microsoft:
http://msdn.microsoft.com/en-u/library/ms379623(VS.80).aspx

Modeling Languages

- Sketching the world with ERM
- From ERM to UML
- The ArcGIS Diagrammer as an example of a DSL-tool