第一章 GIS 应用基础

第 1 课　GIS 介绍

- 课程目标

了解 GIS 用于做什么，以及进行下一步学习必须理解的一些基本概念.

- 专题: 什么是 GIS?

概念：
GIS 的组成
地理数据的组成
地理要素的空间关系
GIS 是如何组织地理数据的？

- 专题：GIS 可以做什么?

概念：
- 数据采集
- 数据存储
- 数据查询
- 数据分析
- 数据显示

数据输出
练习
使用 ArcGIS 浏览地理数据

- 专题：什么是 GIS 数据库?

概念：
现实世界中地理要素的抽象
连接空间要素和属性
存储抽象对象
元数据

课程目标

在本课中，你将了解到:

- 组成 GIS 系统的五个部分是什么
- 地理数据的构成
- GIS 如何表示现实世界中的对象
- 什么是属性
- 空间关系和拓扑是如何影响到地理分析的
- GIS 是如何组织地理数据的
- 一个 GIS 应该具备哪 6 种基本功能

专题: 什么是 GIS?

GIS 被很多行业广泛使用, 公用事业, 商业, 法律实施、运输, 卫生保健、农业,与地区、省级, 和全国的政府部门。 各个行业将 GIS 应用于比如自然资源管理、土地利用规划、人口统计研究, 突发事件的车辆调度, 运营车辆管理,、环境评价及规划,等众多领域。基于因特网的 GIS 应用也在快速增长。

GIS 是一种通用的技术,但是它到底是什么? 有些什么功能? 首先,地理信息系统 (GIS) 是基于计算机的问题求解工具。GIS 以一种有助于我们理解问题和找到解决方案的方式对信息进行集成。描述现实世界中实体的数据被存储在数据库中并且被动态地连接到电子地图的相关区域，电子地图以图形化的方式表现现实世界中的实体。当数据库中的数据发生改变时，地图会作相应的更新以反映这种变化。

通常，人们出于各自的目的来使用 GIS： 数据制作、数据显示、分析和输出。你可以根据数据库中的数据显示地理空间要素 （在后面的内容中你将认识到这是一种强大的特性）。GIS 分析工具可以让你完成诸如此类的事情，比如：你的最佳客户到你的商店有多远，哪个地块可能处于洪水淹没区，哪一种土壤最适宜种某种作物。输出功能可以输出用于打印和具有绘图质量的地图，还有表格和图表。

你将详细了解到以下有关于 GIS 如何实现的有关概念.

什么是 GIS?

概念

- GIS 的组成
- 地理数据的组成
- 要素的空间关系
- GIS 是如何组织地理数据的

概念：GIS 的组成

GIS 由 5 个部分组成。即：用户, 数据, 硬件, 软件和过程.



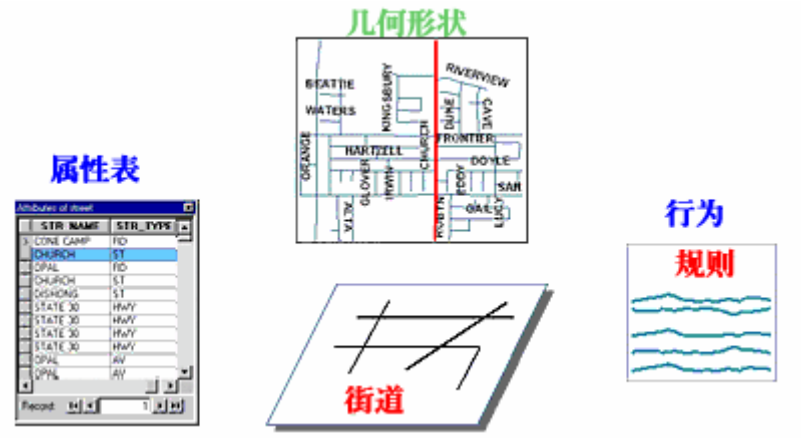
正如 GIS 这个词所提示的含义， GIS 是一个系统，其五个组成部分是生成你所希望的结果必须的。

- 人员，是 GIS 中最重要的组成部分。开发人员必须定义 GIS 中被执行的各种任务，开发处理程序。 熟练的操作人员通常可以克服 GIS 软件功能的不足，但是相反的情况就不成立。最好的软件也无法弥补操作人员对 GIS 的一无所知所带来的负作用。
- 数据，精确的可用的数据可以影响到查询和分析的结果。
- 硬件，硬件的性能影响到处理速度，使用是否方便及可能的输出方式。
- 软件，不仅包含 GIS 软件，还包括各种数据库，绘图、统计、影像处理及其它程序。
- 过程，GIS 要求明确定义，一致的方法来生成正确的可验证的结果。

地理数据的组成：

地理数据由三部分构成:

- 几何图形，表示与现实世界中某个位置相关的地理要素。地理要素被抽象（表示）为点、线或多边形(面)。
- 属性，用于技术地理要素的特征。
- 行为， 意味着地理要素可以设定为只允许某种特定的编辑、显示或分析操作， 根据用户定义的使用环境，要素行为是很容易在地理数据模型(Geodatabase)中实现的。



这里，表示的是 GIS 中，城市街道数据。街道在几何特征上是线段，因为在 GIS 中它们被存储成线要素。每条街道由其名称来描述，这里街道名称就是属性。你也可以为街道设定行为。比如，你可以定义一个行为，它规定每个街道不能拥有 4 个以上的巷道。

要素的空间关系：

在一幅地图上，要素空间关系或者说它们在空间上的位置相对于其它要素的关系传递某种重要的信息。隐含在地图上的空间关系决定地图要传达给用户的是什么。比如，连通性和相邻性是两种空间关系。如下图，在美国地图上，80 号州际公路连接了旧金山和纽约市，旧金山与太平洋相邻，而纽约市则与大西洋相邻。



从此图上还可以确定其它要素关系，你可看到加州是处于美国境内，它与 80 号州际公路相交，它还与太平洋相邻。

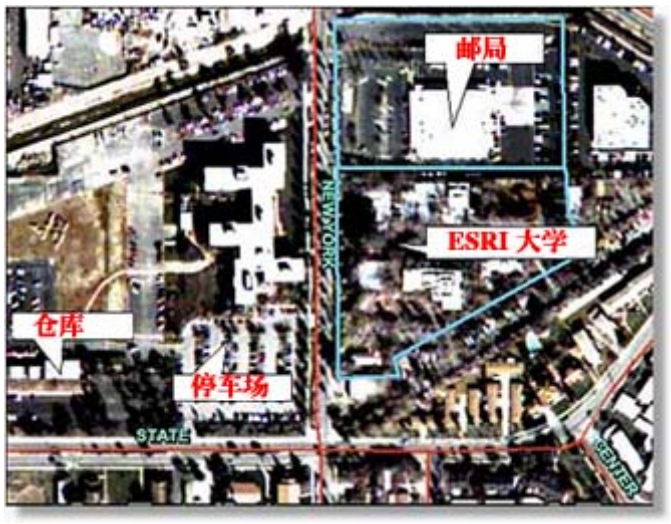
拓扑是一种用以确定要素空间关系及属性的数学 procedure。包括：

- 线段的连通性
- 线段的方向性
- 线段的长度
- 区域间的相邻性
- 区域的定义

拓扑使得大多数的地理分析成为可能，因为它允许我们回答具有要素空间关系的必须解决的

问题。比如，对下图中的地理要素，拓扑可以回答这类问题：

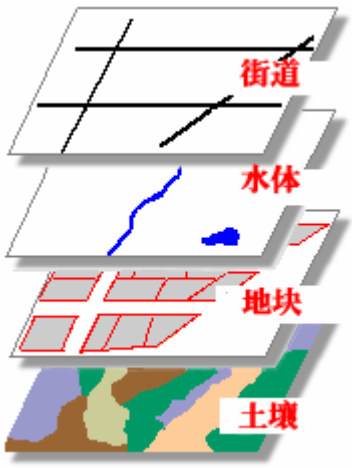
- 与 ESRI 大学相邻的是什么?
- 哪条街道与 State 这条街道相交?
- ESRI 停车场有多大面积?
- 从仓库到邮局最快的路线是哪条?



拓扑关系告诉我们邮局所处的那个地块与 ESRI 大学相邻，因为它们有一条公共边。

GIS 是如何组织地理数据的?

GIS 将现实世界的相关信息以各种专题图层形成的集合进行组织和存储，这些图层可以按地理关系进行连接。 每个图层包含具有相似特征的一类要素，例如，处于相同的地理范围内的街道或城市。这个看似简单但确是非常强大和应用广泛的概念已被证明，在解决许多实际问题时具有无法估计的价值，这些问题比如是运输车辆的跟踪、记录规划应用的详细资料、全球大气环流建模。



这四个图层可能是某市地理数据集的组成部分。这些图层都包含有处于城市边界以内的要各类要素，但是每个图层只表示某种“专题”信息。

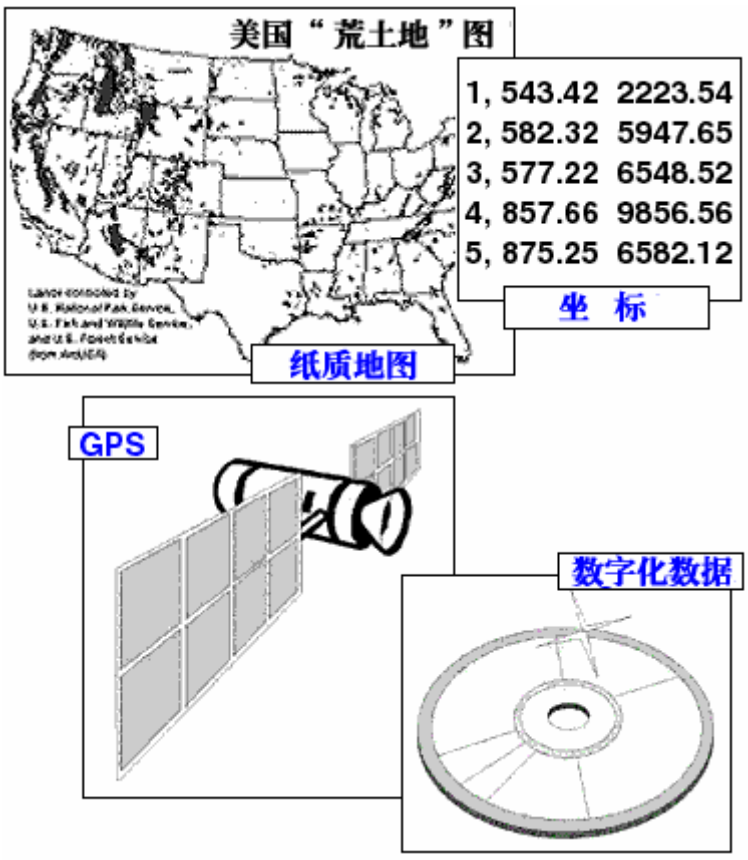
主题: GIS 有什么功能?

为了能够找到现实世界中实际问题的解决方案，任何地理信息系统都应该具备 6 种基本功能，这 6 种基本的功能是：

- 数据获取
- 数据存储
- 数据查询
- 数据分析
- 数据显示
- 数据输出

概念：数据获取

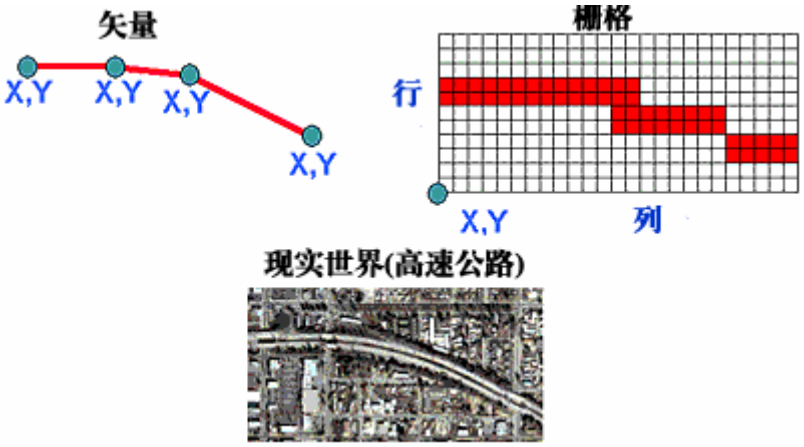
数据所描述的地理要素被存储在地理数据库中。数据库是地理信息系统中必不可少且花费较多的组成部分，,因此要认真考虑数据的的入口问题。 你将如何将现有的存在于纸上的数据转换成可获取的数据格式并存储在数据库中？对已经数字化的但没法使用的数据又该如何，是什么格式呢？GIS 必须提供各种途径用以输入地理(具有坐标的)数据和表格(属性)数据。数据输入方式越多，这个 GIS 的的用途越广。



你可以将多种数据源中的数据转换到 GIS 中。

数据存储

有两种基本的模型用于地理数据的存储：矢量模型和栅格模型。一个 GIS 应该同时具备存储这两种地理数据的能力。



要素在 GIS 中如何显示，取决于它们是以什么格式进行存储。

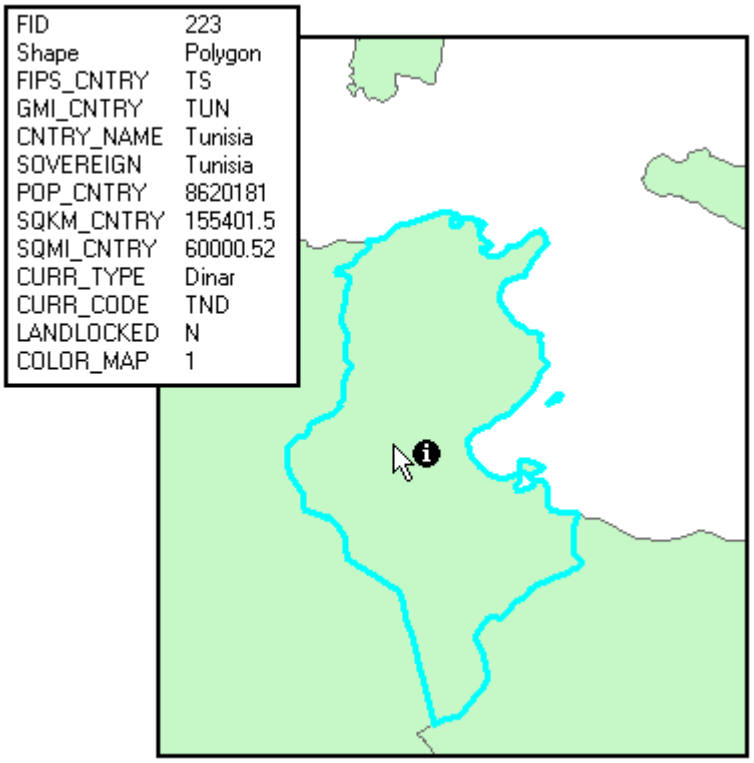
在矢量数据模型中，地理要素表示的方式与它们在地图上显示的方式相同，以点、线段和多边形的来表示。使用格网或 x,y (笛卡尔) 坐标系表现这些要素在现实世界中的位置。

栅格数据模型用像元格网表示要素。各个像元覆盖要素的位置，并且被赋值。栅格像元的大小决定对某一要素刻画的详细程度。栅格格式适合进行空间分析，同时也便于以格网形式存在的数据的存储。栅格数据不适合于地块管理一类的应用，这种情况需要知道离散要素的边界在哪里。

数据查询

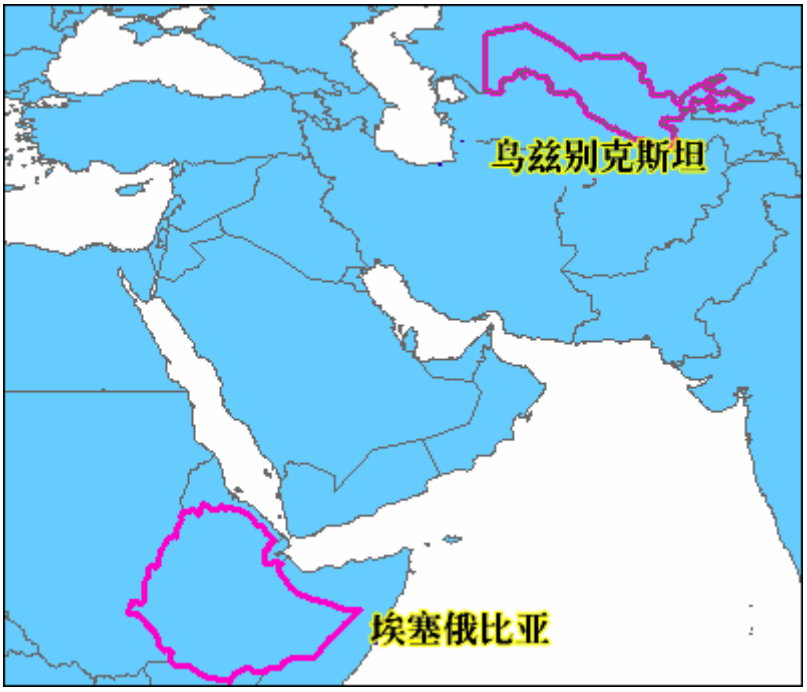
GIS 必须提供根据要素位置或属性查找特定要素的工具。查询，通常是以逻辑语句和表达式来构造，查询用于从地图上或与地图相关的数据库记录中选择要素。

普通的 GIS 查询就是确定在一个指定位置上存在什么。在这类查询中，用户知道他感兴趣的要素在哪里，但是想要知道与之相关的特性。这对于 GIS 来说是比较容易实现的，因为地图中的地理要素与存储在数据库中的属性是关联的。



在 GIS 中，你可以在地图中的地理要素上点击以查看数据库中和要素关联的属性。

另一类 GIS 查询是确定满足特定条件的要素所处的位置。在这种情况下，用户知道要素的重要特征，他想知道具有这些特征的那些要素的具体位置。



假定你想查找人口数大于 2000 平万公里的内陆国家。你应该使用这个限定条件创建一个查询表达式。 一旦 GIS 找到符合查询限定条件的要素，将会在地图上高亮显示这些要素。

数据分析

地理分析通常涉及不只一个地理数据集，它要经过一系列步骤才能完成，得到最终结果。GIS 必须能够分析多个数据集的空间关系以回答并提供问题的解决方案。

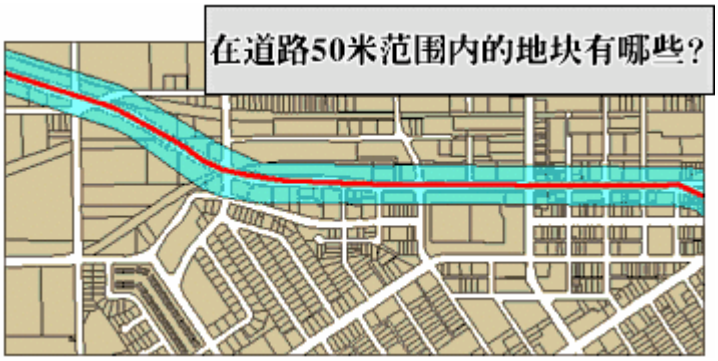
有多种类型的地理分析，但是本教程不可能涉及所有类型。下面将描述两种普通的地理分析。

邻近分析：

邻近分析根据要素间的距离回答诸如以类型的问题：：

- 在距自来水总管 100m 以内的范围内有多少房屋？
- 这距某商店 10KM 的范围内顾客的总数是多少？
- 某个水井 500m 范围内紫花苜蓿（牧草）作物的面积是多少？

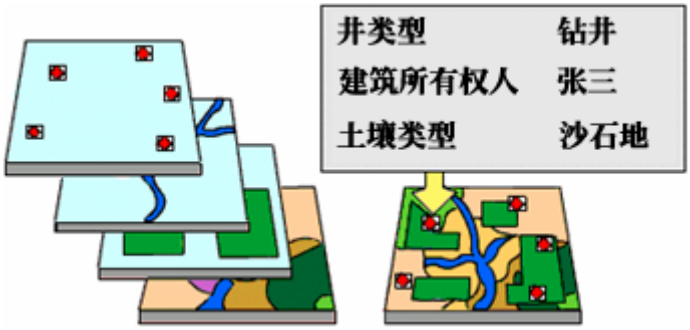
GIS 通常使用称为创建缓冲区的操作来确定要素间的邻近关系。



在道路的每边创建一个 50m 的缓冲区 ，从而找出距离道路 50m 以内的地块。

叠加分析

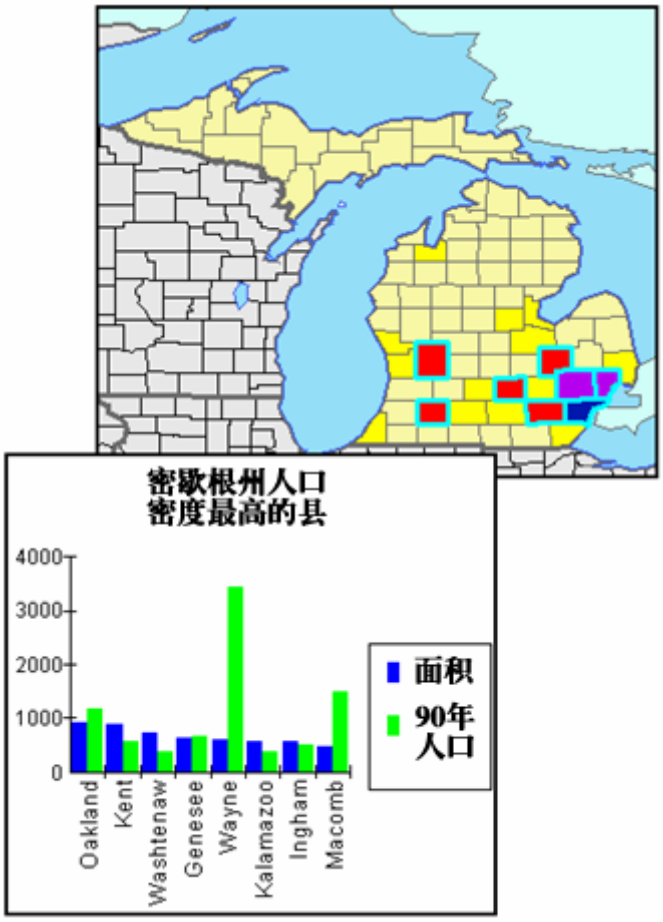
将不同的数据层进行综合，涉及到称为“叠加”的操作过。最简单的情形是，这可以是一种可视化的操作，但是分析操作要求对一个或多个数据图层进行综合（也就是，合并为数据库中的一个图层）。 叠加分析可以用于将土壤类型数据、坡度数据及植被数据或土地所有权数据集成，从而确定土地的价格和税额。



叠加分析结果可以有助于确定合适的土地利用

数据展示

GIS 同样需要具备用不同符号来显示地理要素的功能。在多数情况下，地理分析的最终结果最好的表达形式是一幅可视的地图、图表或报表。

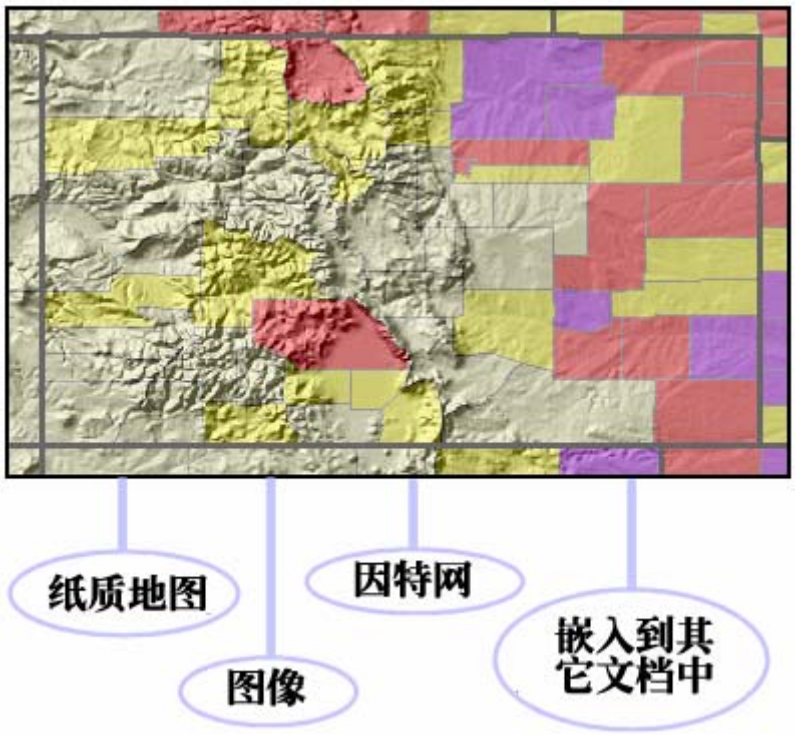


使用 GIS，用户可以用不同的方式显示数据以满足要求。

地图是表现地理信息最有效的形式。尽管制图人员已经制作了成千上万不计其数的地图，GIS 提供的新手段和令人激动的一些工具还是有助于提高制图行业的技巧和艺术性。地图可以与报表、图表、相片、多媒体等集成。.

数据输出

由于在 GIS 数据生产上开销巨大，因此我们有充分的理由将最终结果与其他人共享。获取借助于 GIS 生成的可视化结果(地图、图表、报表)并将其输出为可分发的格式是实现共享的方法之一，但这还不够。GIS 可提供的输入选项越多，就更有可能为目标用户提供有用和正确的信息。.



GIS 可视化结果，比如：纸质地图，还可以通过各种方式进行分发。

主题：什么是 GIS 数据库?

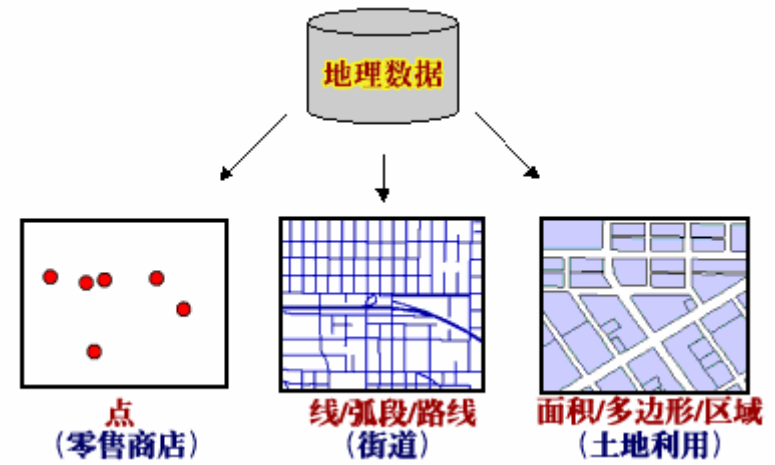
地理数据库 GIS 的核心。 它的完整性和准确性影响到以 GIS 数据库为核心的各种应用。地理数据库是现实世界中各种对象的空间及属性信息的集合。 最好的结果应该如此，数据库经过很好设计，为一个和多个应用提供有效的服务并且依照组织良好的文档和管理程序进行维护。

在 GIS 数据库中，通过抽象，现实世界中的对象被数字化。 如何组织及存储这些数字化的要素取决于几个因素，包括数据将如何应用、将用于什么软件环境、要素自身的特性。地理要素的属性通常是以某种方式与要素进行关联。在这部分的主题中，有专门的内容对此作详细介绍。

概念

对现实世界中实体的抽象

GIS 用户必须以某种方式将现实世界中的实体抽象为几何对象。现实世界中的实体可以抽象为三种主要的图形：点、线、面(多边形)。 这些图形经常被称为几何对象、几何要素、要素类型、或要素类。



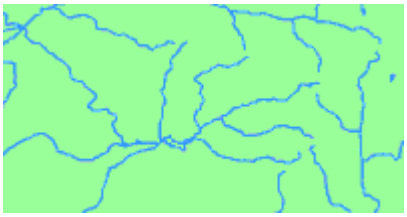
每个要素类有其惟一的一系列特征。点由一个坐标（或 X,Y ）确定。线是由一系列坐标点构成且必定有一定的长度。多边形由一条或多条相连的线段组成，且起点和终点的坐标完全相同，多边形必定具备一定的周长和面积。

如何表示几何对象影响到信息分析和结果显示的方式。我们与现实世界中各种对象的相互作用是多种多样的，因此，你可以以不同的方式来表示它们

考虑一个具体的例子：河流，河流是自然要素，通常可进行运输，作为划定行政分界线或国界的依据，而且是表面模型中重要的要素。 下面，你将了解到在 GIS 中表示河流众多形式中的几种。

一系列线段形成一个河流网络

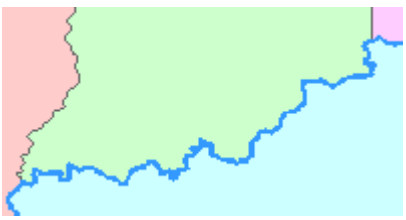
线段的每一段都具有流向、流量及其它的河流所具备的属性。你可以应用一个线性网络模型来进行水文分析或河运分析。



河流表示为线段

两个区域间的边界

河流可以界定国行政界线，比如国省界和县界，或者也可成为自然区域的天然屏障，如，野生动物栖息地。.



河流可以形成边界.

面状要素

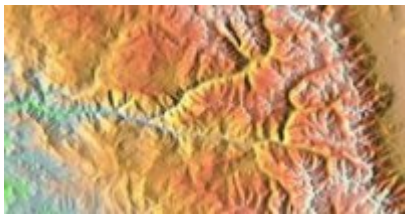
河流可以表示为一个面状的要素，可以精确地表示河岸、溪流、及可供航行的河道。



河流的面状表达

在表面模型中是断槽(Trough)

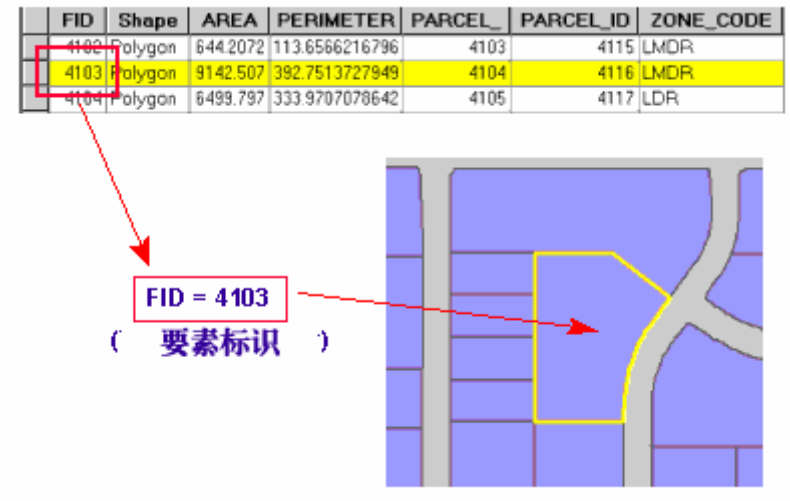
在表面模型中，错综复杂的一条河流形成断槽。沿着河流走向，河流穿过地形表面，你可以计算其剖面、下降速率、划分流域、计算指定降水量下的淹水潜势。



在基于地形的表面中，河流是断槽

关联要素及其属性

每个单独的要素被赋以一个惟一的数字标识,其特点是地理空间上的惟一位置及相应的特征被记录在一个要素属性表中。 理解要素、标识、及属性记录间的一对一映射关系是十分重要的。

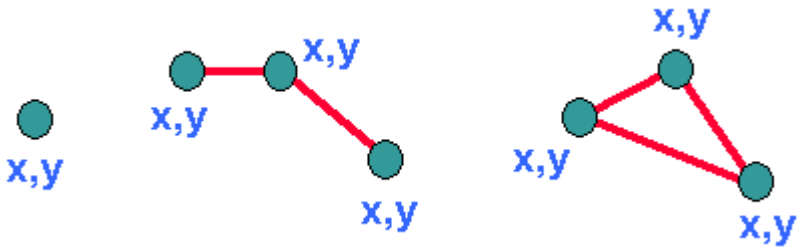


每一个要素是数据表中的一行。一个惟一的标识将要素与其属性相关联。

存储抽象的对象

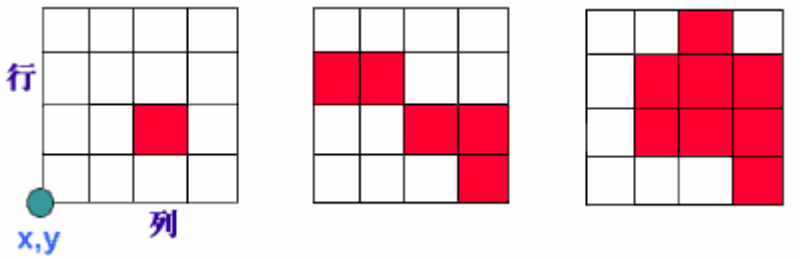
有两种基本的模型实现地理数据的存储：矢量数据模型及栅格数据模型。

矢量存储方式要求点、线、多边形要素具有明确位置坐标，比如 X,Y 笛卡尔坐标系统。每个点有一个 X,y 坐标对定位其位置。 线段可以被存储为一系列的(x, y)坐标，这些坐标对定义了构成线段的直线线段。线段也可以存储为两个终点坐标和一个曲线公式。多边形由封闭的线段构成。



矢量数据模型存储每几何要素的空间坐标。

栅格数据存储方式使用一个等大的方格格网实现点、线、或多边形的存储。具有地理参考的栅格数据格式不是通过明确的(x,y)坐标，而是通过坐标原点(通常是在数据图层的左下角)、行和列、像元大小来定义各种形状。有了这四个参数，栅格数据集中的要素就能根据它们的具体位置进行赋值。

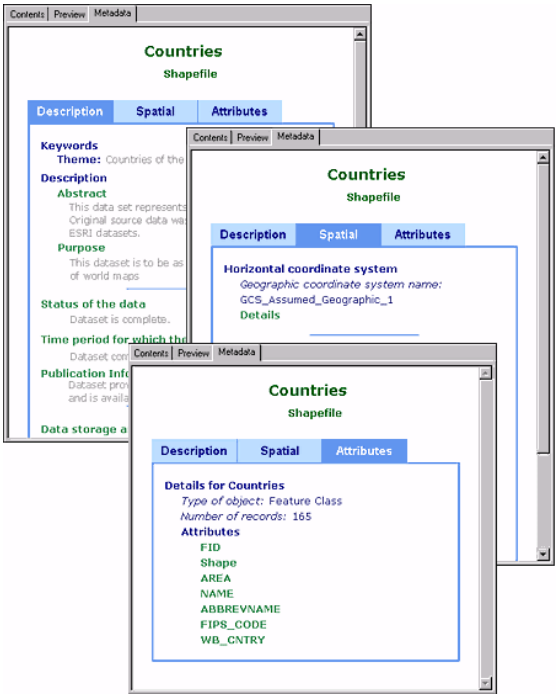


栅格数据模型使用正方形的单元栅格表示现实世界中的实体。

在如何关联要素与属性方面，矢量和栅格格式之间有细微的不同。前面所述的要素和属性表模型主要适用于矢量数据。

介绍元数据

元数据的通常定义是：“数据的数据”。但是这个定义不是准确的，它也不是那么有意义。更加正式的应该是：元数据是提供有关数据的描述信息。



在 ArcGIS,有三类元数据：描述（Description），空间(Spatial)，和属性（Attributes）. 描述信息包含数据集的基本信息，包含数据源，组织，日期，用途与限制。空间信息包含坐标系统信息及数据集的地理区域。属性信息包括：字段、属性值域及相关信息。

空间数据的元数据包括描述信息，比如日期、创建人、地图数据范围、坐标系统及属性值域区间。无数据使得空间数据更可信，而且，在很多情况下，如果没有元数据，空间数据可能无法被解析和使用。维护元数据的标准是很重要的，这样它才可以比较容易地分发和解析。

正如你所了解的，使用 ArcGIS ，你可以根据给定的数据集，生成、编辑、导入、导出元数据。元数据存储在称为扩展标识语言(XML)的数据文件中。这意味着你可以在其它可识别 XML 文档的软件中使用元数据，包括浏览器。