Conducting a GIS project

Planning a GIS project

4

Beginning with this chapter and through the rest of the book, you will conduct a sample GIS analysis project. The tasks you perform will help you learn the methods for performing your own GIS projects. You'll learn about several specific GIS analysis techniques and, perhaps more important, you'll learn how to plan and carry out a GIS project.

This chapter provides an overview of GIS analysis and presents the steps involved in conducting a GIS project. It then walks you through the first step—planning the project.

The scenario for the project involves finding the best site for a new wastewater treatment plant for the fictitious city of Greenvalley. To find a suitable site you will need to know the site selection criteria. You'll then need to identify the data needed to address these criteria and use the data to find suitable sites for the plant. These are fundamental elements of a GIS analysis project.

What is GIS analysis?

The phrase "GIS analysis" encompasses a wide variety of operations that you can do with a geographic information system. These range from simple display of features to complex, multistep analytical models.

Showing the geographic distribution of data

Perhaps the simplest form of GIS analysis is presenting the geographic distribution of data. This is conceptually the same as sticking pins in a wall map, a simple but powerful method of detecting patterns.
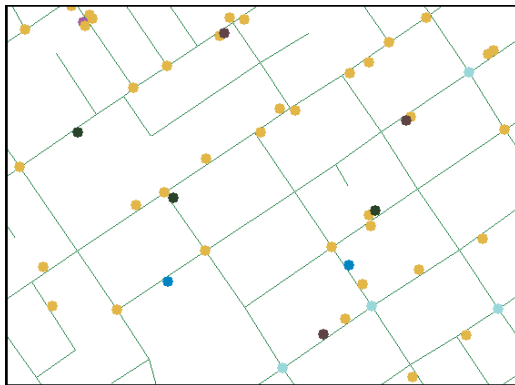

Here, the map is the analysis. A police department might analyze burglary patterns by plotting the addresses of reported break-ins. The department could make the map more informative by displaying the incidents with different symbols to show the time of day, method of entry, or types of valuables reported stolen.

Querying GIS data

Another type of GIS analysis is querying, or selecting from, the database. Queries let you identify and focus on a specific set of features. There are two types of GIS queries, *attribute* and *location* queries.

Attribute queries, or aspatial queries, find features based on their attributes. The police department mentioned above could use an attribute query of its database to obtain a table of crimes that fall into a particular category.
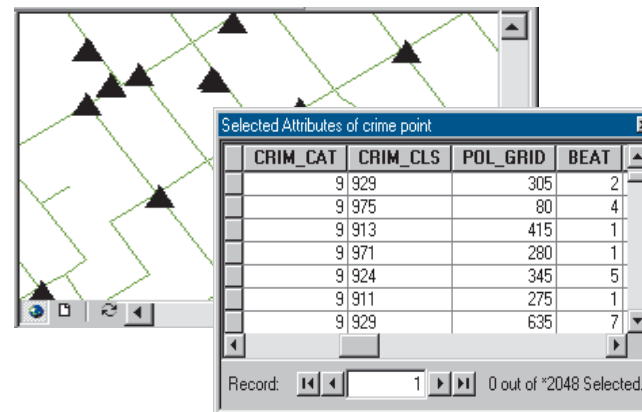
Here are the results from a query on the CRIM_CAT field, showing records where the value in the field is 9. The map shows the results of the query.

Location queries, or spatial queries, find features based on where they are. The police department could use a location query of the database to find crimes that occurred within a given area.
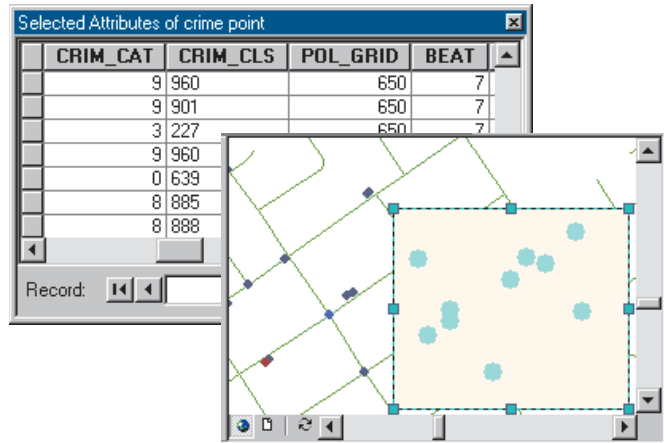

One way to do a location query is by drawing a rectangle on the map. Here, the police department has selected only those crimes that occurred within the rectangle. These crimes could be studied together to determine if any of them are related.

The police department could also do more complex spatial queries using polygon features, such as census tracts, selected from another layer. One of the most useful features of a GIS is that you can see the results of both spatial and aspatial queries on the map.

Identifying what's nearby

A third type of GIS analysis is finding what is near a feature. One way to find what is near a feature is by creating a buffer around the feature.

A city planning commission could identify the area within 1,000 meters of a proposed airport by buffering the airport feature. The buffer could be used with other layers of data to show which schools or hospitals would be near the new airport.
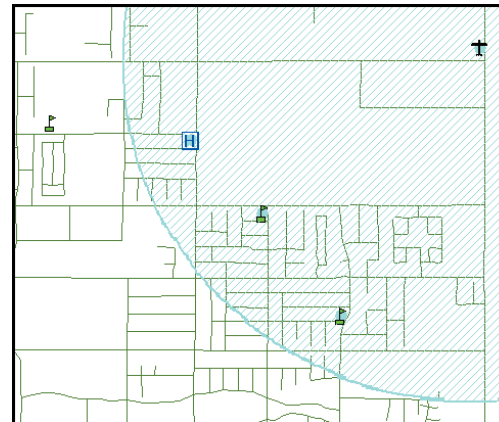

A powerful function of GIS analysis is that the output of one procedure can be used in another. Here, the buffer around the airport is used in a location query. Two schools and a hospital that are within the buffer have been selected. The school that is outside the buffer was not selected.

Overlaying different layers

A fourth type of GIS analysis is overlaying different layers of features. You can create new information when you overlay one set of features with another. There are several types of overlay operations, but all involve joining two existing sets of features into a single new set of features.

For example, a farmer wants to find how much land can be planted with a new crop. The crop can't be planted on hillsides and needs soils that are highly permeable.

The farmer combines, in a *union* overlay, two existing layers of data about the farm: polygons of the ground surface classified by slope and polygons of soil permeability. The farmer can now select the new polygons that have low slopes and high permeability.
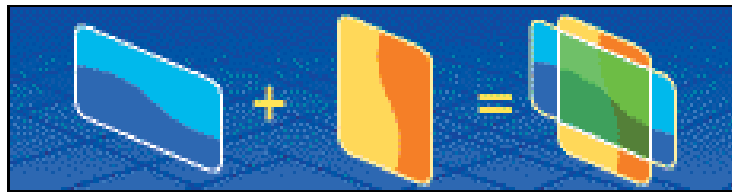
Slope   Permeability   Slope and
                                     permeability

There are several different spatial overlay and spatial manipulation operations that you can use on layers including union, intersect, merge, dissolve, and clip.

Doing a complex analysis

You can combine all of these techniques and many others in a complex GIS analysis. With a GIS you can create detailed models of the world to solve complicated problems. Because a GIS can perform these operations rapidly, it is possible to repeat an analysis using slightly different parameters each time and compare the results. This can allow you to refine your analysis techniques.

This section has provided a brief overview of some of the common types of GIS analysis. For more information on these and other kinds of analysis you can perform using GIS, see *The ESRI Guide to GIS Analysis*.

In the next section you'll learn about the steps in a typical GIS analysis project.

The steps in a GIS project

In a typical GIS analysis project, you identify the objectives of the project, create a project database containing the data you need to solve the problem, use GIS functions to create an analytical model to solve the problem, and present the results of the analysis.

Step 1: Identify your objectives

The first step of the process is to identify the objective of the analysis. You should consider the following questions when you are identifying your objectives:

- What is the problem to solve? How is it solved now? Are there alternate ways to solve it using a GIS?

- What are the final products of the project—reports, working maps, presentation-quality maps?

- Who is the intended audience of these products—the public, technicians, planners, officials?

- Will the data be used for other purposes? What are the requirements for these?

This step is important because the answers to these questions determine the scope of the project as well as how you implement the analysis.

Step 2: Create a project database

The second step is to create a project database. Creating the project database is a three-step process. The steps are designing the database, automating and gathering data for the database, and managing the database.

Designing the database includes identifying the spatial data you will need based on the requirements of the analysis, determining the required feature attributes, setting the study area boundary, and choosing the coordinate system to use.

Automating the data involves digitizing or converting data from other systems and formats into a usable format, as well as verifying the data and correcting errors.

Managing the database involves verifying coordinate systems and joining adjacent layers.

Creating the project database is a critical and time-consuming part of the project. The completeness and accuracy of the data you use in your analysis determines the accuracy of the results.

Step 3: Analyze the data

The third step is to analyze the data. As you've seen, analyzing data in a GIS ranges from simple mapping to creating complex spatial models. A model is a representation of reality used to simulate a process, predict an outcome, or analyze a problem.

A spatial model involves applying one to three categories of GIS functionality to some spatial data. These functions are:

- Geometric modeling functions—calculating distances, generating buffers, and calculating areas and perimeters
- Coincidence modeling functions—overlaying datasets to find places where values coincide
- Adjacency modeling functions—allocating, pathfinding, and redistricting

With a GIS you can quickly perform analyses that would be impossible or extremely time-consuming if done by hand. You create alternative scenarios by changing your methods or parameters and running the analysis again.

Step 4: Present the results

The fourth step is to present the results of your analysis. Your final product should effectively communicate your findings to your audience. In most cases, the results of a GIS analysis can best be shown on a map.

Charts and reports of selected data are two other ways of presenting your results. You can print charts and reports separately, embed them in documents created by other applications, or place them on your map.

What's next?

Now that you have reviewed the steps in a GIS project, you are ready to begin planning your own project. The next section presents an overview of the steps for the Greenvalley wastewater treatment plant project. The first step—identifying the project objectives—is covered in this chapter. The rest of the steps comprise the remaining chapters in this book.

Planning your project

Planning is a critical step in any GIS project and can save you time and effort once you get to the database creation, analysis, and mapping steps. During the planning phase you identify the project objectives, define the criteria for the analysis, and identify the data required to support the analysis. You should also consider the approach you'll use for the analysis and what the final products of the project will be. Once you've done this, you can proceed to create the project database.

Throughout the rest of this book you will be working on a small GIS analysis project. In the process, you'll learn how to plan a GIS project and how to use ArcMap and ArcCatalog together to carry it out. While you'll be performing a specific type of analysis—finding a site for a new facility—the steps you'll follow in the project, and many of the specific tasks, will be applicable to a range of GIS projects. The scenario for the project is to find a suitable site for a new wastewater treatment plant.

The city of Greenvalley is growing. To support this growth, the city is building a new wastewater treatment and recycling plant. The city plans to use conservation and wastewater recycling to help meet its expected water needs.

The diagram to the right outlines the steps in a GIS project and shows where each step is covered in the remaining chapters in this book.

In this chapter, you'll carry out step 1—identifying the project objectives. You'll also do some planning for the remaining steps.
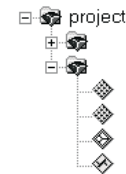
Steps in a GIS project

Step 1: Identify the project objectives—Chapter 4

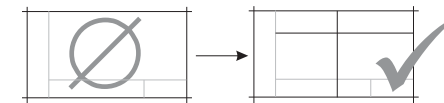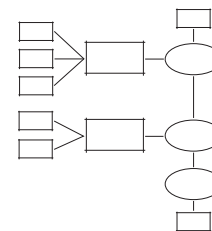| CRITERIA	DATA SET	ATTRIBUTES

Step 2: Create the project database
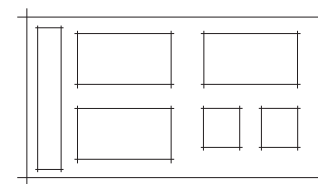
Assemble the data— Chapter 5

Prepare the data for analysis—Chapter 6

Step 3: Analyze the data—Chapter 7

Step 4: Present the results—Chapter 8

Step 1: Identify the project objectives

The objective of this GIS analysis is to find a suitable site for the city's new wastewater treatment plant. The city has never used a GIS model to site a wastewater treatment plant. The existing plant was sited many years ago using a quadrangle map, acetate overlays, and the city council's knowledge of the area in consultation with the city engineer. This approach was adequate but time-consuming, and the public was not involved in the process.

The problem has become more difficult as the area has become more developed and environmental and public health regulations more stringent. The council has chosen to use a GIS model in order to speed the process and to ensure that the necessary regulations are complied with.

Because the council recognizes that siting such a plant can be controversial, it wants the analysis to identify all of the parcels that could be used for the plant site, then identify the highly suitable parcels, based on specific criteria. The possible sites will be discussed at a public meeting. The map you create for the meeting should make clear which parcels are highly suitable, which are less suitable, and which are unsuitable.

The city has provided you with a list of the criteria for a suitable site. The parcels chosen for the site must be:

- Below 365 meters in elevation, to minimize pumping costs
- Outside the floodplain, to avoid spillage during storms
- Within 1,000 meters of the river, to minimize pipeline construction for treated water that is discharged
- At least 150 meters from residential property and parks, to minimize the impact on the city's residents

- On vacant land that can be developed, to minimize land acquisition and construction costs

In addition, to further reduce construction costs, the city would prefer that the site be:

- Within 1,000 meters of the main wastewater junction (within 500 meters is considered even more suitable)
- Within 50 meters of an existing road

The plant will also require a total of at least 150,000 square meters in area.

A preliminary review of existing paper maps has shown that the most likely location for the plant is in the northwest corner of the city, near the river, and in a low-lying area. This will be the study area for the project. The GIS analysis will allow you to combine the criteria to identify specific parcels that are suitable sites.

Step 2: Create the project database

Creating the database for this project will be a two-step process. First you'll assemble the existing data and review it. Then you'll prepare the data for analysis. Some of the data will be usable as is; other layers will need additional processing. You may also need to automate some data. You'll assemble the data in Chapter 5, 'Assembling the database', and prepare the data for analysis in Chapter 6, 'Preparing data for analysis'. At this point, though, you can plan for those tasks by identifying the data layers you'll need and the sources of the data.

Assemble the project data

Several city of Greenvalley departments maintain GIS data and have working arrangements to share data for city projects. Some of this data is already stored in the city's

GreenvalleyDB database. The city also has data sharing agreements with several regional and state agencies.

Because a database containing much of the data you need already exists, you will not need to spend as much time on designing and automating your project database as you would otherwise. However, you will still need to do some database design work for your project database. You'll need to identify the dataset and any attributes required for each criteria. You'll then research the available data to see which layers will meet your needs.

Each of the city's criteria will require a layer of data for the analysis. Here is a list of the criteria and the corresponding datasets and attributes.

CRITERIA	DATASET	ATTRIBUTES
LESS THAN 365 METERS ELEVATION	ELEVATION	ELEVATION IN METERS
OUTSIDE THE FLOODPLAIN	FLOODPLAIN	N/A
WITHIN 1,000 METERS OF THE RIVER	RIVER	N/A
AT LEAST 150 METERS FROM RESIDENTIAL PROPERTY	PARCELS	LAND USE
AT LEAST 150 METERS FROM PARKS	PARKS	N/A
ON VACANT LAND	PARCELS	LAND USE
WITHIN 1,000 METERS OF THE WASTEWATER JUNCTION	WASTEWATER JUNCTION	N/A
WITHIN 50 METERS OF A ROAD	ROADS	N/A
AT LEAST 150,000 SQ. METERS	PARCELS	AREA IN SQUARE METERS

Note that the parcels dataset will be used for several criteria.

You can now take inventory of the data that you have and see which layers correspond to the required datasets. You can also identify other layers that you need to obtain or create.

To find areas below 365 meters elevation, you need a source of elevation data. A colleague at the State Department of Transportation (DOT) has provided an elevation grid. Because you simply need to know whether or not a parcel is below 365 meters, you will use a polygon of areas below 365 meters, which your colleague at DOT has created from the grid. This data is in a shapefile format.

To find parcels outside of the floodplain, you will use the City Planning Department's digital flood zone layer, stored as a feature class in the city's GreenvalleyDB geodatabase.

To identify areas within 1,000 meters of the river, you will first need a layer of the river. The County Water Resources Department has a shapefile of the river.

You will need a dataset of the parcels in your study area. The city tax assessor has a tiled database of parcels stored as shapefiles. Two of these tiles cover your study area. The parcel database includes a land use attribute that you will use to identify residential parcels—so you can buffer them to 150 meters—and vacant parcels. You will use the area attribute of the parcel shapefile to identify possible sites with an area of at least 150,000 square meters.

To find areas more than 150 meters from parks, you'll need a parks layer. The City Parks and Recreation Department has a feature class of existing parks, which is stored in the GreenvalleyDB geodatabase.

There is also a recently discovered historic site in the project study area. The city plans to develop a park around the site, but the proposed park boundary has not been placed into the park feature class yet. You will get this information into your project database by digitizing from a scanned image of the draft park boundary map.

To find parcels within 1,000 meters of the main wastewater junction, you will need a layer that includes the junction. The City Utility Department has a coverage of the wastewater mains and the junction.

To identify parcels that are within 50 meters of a road, you will use the existing streets feature class from the GreenvalleyDB geodatabase.

The table below lists the layers you'll assemble for the project database, based on the available data. The source and the format of each layer are also listed.

Layer	Source	Format
Elevation	State dept. of transportation	Grid
Elevation < 365 m	State dept. of transportation	Shapefile
Floodplain	City planning dept.	Geodatabase
River	County water resources dept.	Shapefile
Parcels	City tax assessor	Shapefiles (tiled)
Parks	City parks and recreation dept.	Geodatabase
Historic park	City parks and recreation dept.	Scanned image
Wastewater junction	City utilities dept.	Coverage
Streets	City streets dept.	Geodatabase

The database will also include the scanned image of the historic park, which you'll use to digitize the new park. You'll also include the elevation grid, as you may want to display it on your final map.

In Chapter 5, 'Assembling the database', you'll assemble the data and organize it so it's easily accessible within a single project folder. You'll then review the data to see which layers will require additional processing.

Prepare the data for analysis

Based on your review of the data, you'll determine which layers are currently usable and which require additional processing for use in the analysis. Some of the common tasks involved in preparing data for analysis include:

- Checking data quality—making sure the data is accurate and up-to-date
- Converting data between formats
- Automating data by digitizing, scanning, converting, or geolocating
- Defining coordinate systems
- Projecting layers to a new coordinate system
- Merging adjacent layers

You'll need to perform some of these tasks for your project database. You already know, for instance, that the boundary of the proposed park surrounding the historic site will need to be digitized. You have a scanned map of the proposed boundary that you will register to the city's geodatabase and digitize using the parcels layer as a backdrop. The new park feature will be added to the existing park feature class in the GreenvalleyDB geodatabase.

You will also need to merge the two parcel tiles for your study area to make the analysis easier to perform.

Once you've reviewed the existing data, in Chapter 5, 'Assembling the database', you'll be able to see which other layers require additional processing.

Most of the data for the project is already in coverage, shapefile, geodatabase, or raster format, all of which ArcGIS can use, simultaneously. There may be cases, though, where you'll need to convert data to a different format—for example, if converting from vector to raster format or from shapefile to geodatabase feature class, for storage in an existing geodatabase.

ArcGIS can display and overlay layers in different coordinate systems as long as the coordinate system for each layer is defined. You'll need to check this, especially for data you've obtained from other sources.

You'll perform the necessary data processing tasks in Chapter 6, 'Preparing data for analysis'.

Step 3: Analyze the data

During the planning stage of the project, you'll want to consider the analysis methodology and list the major steps in the process. That way you can make sure you are aware of all the datasets you will need and can include them as you create the project database. You may want to create a schematic diagram of the process as a guide.

The diagram to the right shows the process for the wastewater treatment plant siting analysis.
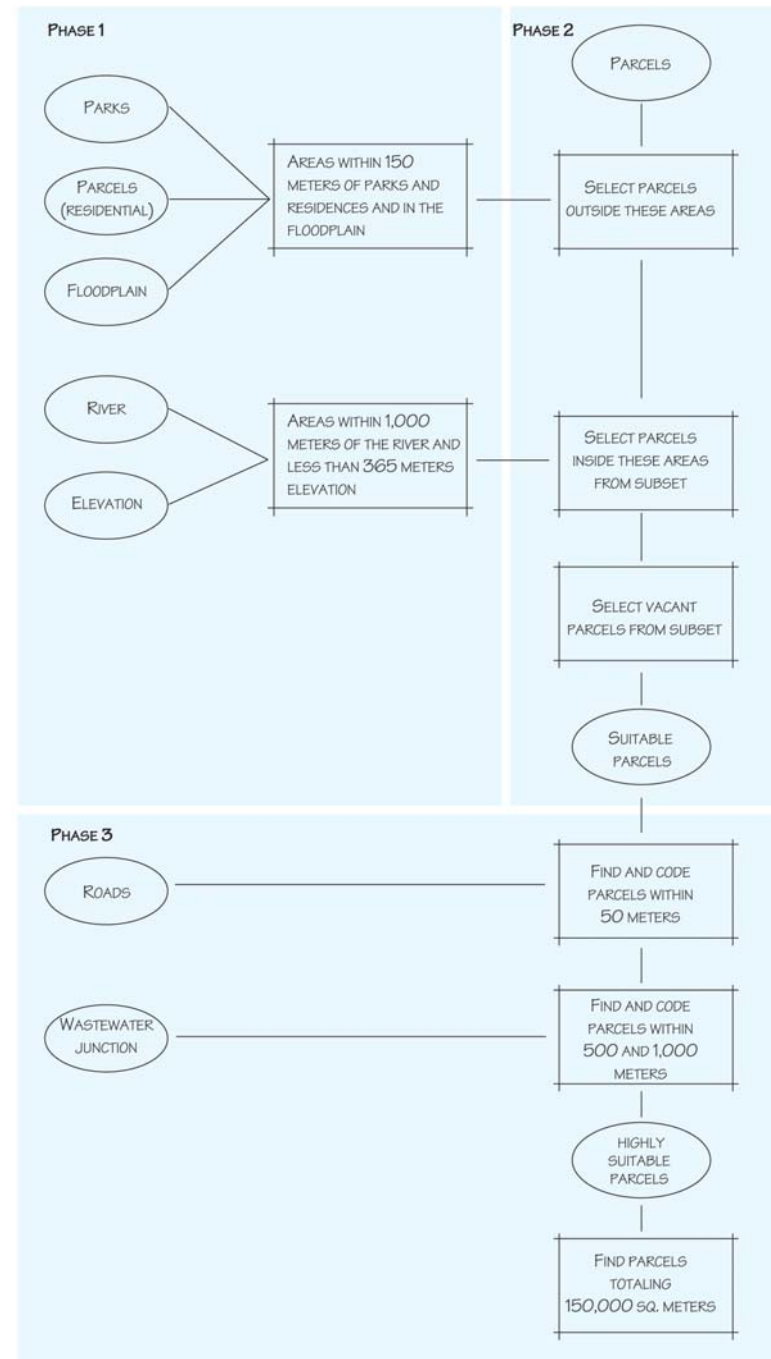
The analysis consists of three phases.

In the first phase, you'll create a layer of the areas the plant should be outside of and another layer of the areas the plant should be within.

In phase two, you'll use these layers to select a subset of parcels that are in a suitable location. You'll then select the subset of these that are vacant to create a layer of suitable parcels.

In the third phase, you'll consider the city's additional criteria that define the highly suitable parcels. You'll find the suitable parcels within 50 meters of a road and those within 500 and 1,000 meters of the wastewater junction, then tag them with the appropriate codes so they can be identified on the map. You'll also check to see which parcels are large enough for the construction of the plant.

While the schematic shows the major steps in the process, there are a number of interim steps you'll complete in each phase. You'll develop the detailed methodology and perform the analysis in Chapter 7, 'Performing the analysis'.

Step 4: Present the results

During project planning you should consider the purpose and audience for your final products. For this project, you'll present the results of the analysis on a presentation-quality map that shows the parcels that are suitable and highly suitable sites. The map will be presented to a general audience at a public meeting.

While you don't need to design the map layout at this point, you'll want to consider the layers that will be shown on the map. In addition to the analysis layers, you may want to include other layers that will provide context for the analysis results and make the map easier to read and understand.

For this project, in addition to the layers you'll use and create during the analysis, you'll want to show the elevation grid as a backdrop so map readers can see the areas of lower and higher elevation in the city, as elevation has a major impact on the location of the wastewater treatment plant.

You'll design and create the presentation map in Chapter 8, 'Presenting the results'.

What's next?

Now that you've reviewed the steps involved in a GIS project, identified the project objectives, and engaged in some project planning, it's time to get started.

You'll assemble the data for the project and review the data in the next chapter.