SETTING UP PREFERENCES

For ArcMap users, setting up project preferences includes setting up the Map Document Properties, environment variables, and the ArcMap Options. In Map Document Properties, fill in the title of your project, create an empty geodatabase into which you will load all of the data you will download and/or create. Typically, you will use a variety of file types and maintaining them in a single geodatabase will reduce your file management overhead and provide you with additional structure for the data. To prevent having to search your entire hard drive or networked drive, specify your default geodatabase so any data you may add or any results you may generate will always be saved here.

Next, check the box indicating that you will store relative path names to data sources; this avoids being tied to hard coded path names. In the Geoprocessing Environment, set your current and scratch workspace to your working folder and database. Under Geoprocessing Options, ensure that you "Overwrite the output of geoprocessing operations" and check "Add results of geoprocessing options to the display" so your results will be added to the table of contents.

Under the Customize menu, under ArcMap Options, visit the Raster tab and Raster Dataset. Be sure that the option to Use the world file to define coordinates of rasters is checked; if you have raster data, it will be registered to real-world coordinates. Under the Metadata tab, select the type of Metadata format you prefer. Under Customize, be sure that any extensions you want to use are enabled.

Lastly, save your map document—and while you are working, save your project often. Create a backup of your map document and your data layers in a different location or storage device. Consider backing up your data in the cloud as well as on physical media.

Finding public domain spatial data

Before the advent of robust search technologies, web pages containing links provided the early one-stops for spatial data by listing sources for a variety of themes. Many of these were assembled by universities and occasionally by nonprofit, government, or commercial organizations. Examples of these include Columbia University Libraries' Digital Social Science Center, Ohio State's Numerical Cartography Lab, and the University of

North Carolina's Spatial Data Online resource. Two of the first were the Alexandria Digital Library at the University of California, Santa Barbara, and the resource established in 2001 by the University of Arkansas, originally named Starting the Hunt. As the technology has developed, a number of purpose-built spatial data portals, such as the Geospatial One-Stop Geodata.gov and the European Commission's INSPIRE portal, were established and financed by government agencies and consortia.

Nowadays, most data users search for public domain spatial data in the same way they search for anything else on the web—using a search engine such as Google and Yahoo!. Although this method can provide quick and relevant results, there are some practical things you can do to maximize the efficiency and effectiveness of Internet searches. First, become familiar with the advanced settings on the search engine you are using; this can save you valuable time. Most have the capability of searching for all words, exact words, one or more specific words and searching within a specific site or domain. Search terms such as "GIS" and "spatial data" can be more effective if used in conjunction with the theme of data you are seeking and the scale or region, such as "global biomes GIS." Including the format of the data, such as the terms raster, vector, shapefile, geodatabase, KML, SHP, MXD, XLS, map package, layer package, SID, and GeoTIFF, in the search string can also help. Knowing how spatial data are stored and made available on the Internet provides more options for searching. For example, including the "allinurl" string will search for a specific path; if the string "allinurl: arcgis/rest/services" is searched, public Esri ArcGIS REST (representational state transfer) services can easily be discovered. These sites are running REST APIs, which provide open web interfaces to GIS services hosted by ArcGIS for Server. As all resources and operations supported via the REST API are accessible through a URL for each published service published, and they use a common path in the URL, they will be located with the "allinurl" search string.

FINDING HIDDEN DATA

Sometimes, data on the web exists but is hidden in historical pages. The Wayback Machine is a digital time capsule created by the nonprofit organization Internet Archive, which archives web pages from 1996 onwards. You may find the data you are seeking there.

With the plethora of web browsers and different versions—some running animations and other functions—if you experience trouble accessing any dataset, your difficulty might be due to the browser, rather than the site. Try another browser, or try an updated version of the same browser, and make sure that whatever plug-ins you need (such as Flash or Silverlight) are installed. There are also tools that allow you to download large numbers of files at one time, such as *wget* in the Ubuntu operating system or Firefox's *download all* add-on (with this tool you can download all the map layers from the national atlas of the United States at once). Utilities like these may be useful for the spatial data libraries you wish to access.

One of the problems plaguing most spatial data collections and portals is that they become difficult to maintain and keep up-to-date. A further complication is the continually changing nature of the web itself. For example, one of oldest and largest collections, Oddens, is no longer maintained; visitors to the site are advised that "This site is no longer maintained. Use 'as is' is still possible." Still, such collections should be considered when you are searching for data, because some of the individual entries in the collections are often missed by search engines. Collections like these are also particularly useful if they are maintained by GIS professionals rather than a general web developer. Data collections developed and supported by the GIS industry—such as the GIS Data Depot maintained by the GeoCommunity, Dundas Data Visualization Inc., and Esri's ArcGIS Online—generally stand a better chance of enduring as a long-term spatial data resource than grant-funded or ad-hoc academic projects, although some exceptions to the rule do exist.

GIS is an analytical tool used in many research studies throughout the world. It may well be worth the effort to investigate relevant research literature in your search for data, including journals, edited volumes, books, and even newsletters from major research organizations and professional societies. Your search in professional society literature should include the geospatial field, of course, such as the Association for Geographic Information, the Geospatial Information & Technology Association, the Urban and Regional Information Systems Association, the American Congress on Surveying and Mapping, and the American Society for Photogrammetry and Remote Sensing, to name a few. You should also examine those related to geography, such as the International Geographical Union, the Royal Geographical Society, and the Association of American Geographers. Many organizations and GIS software companies host annual conferences, including the Esri International User Conference, GeoInt (GeoIntelligence), and GeoWeb. Many of the proceedings from those conferences are available online.

Since GIS is also applied in many disciplines, consider the professional societies and literature tied to these disciplines, such as human health, business, engineering, transportation, energy, public safety, law enforcement, environmental studies, hydrology, climate, and biology. Your search should include major publishers of GIS-related textbooks and journals, such as Taylor & Francis, Wiley, and Esri Press, and also online indexing services, such as university-hosted bibliographies and Google Scholar. These resources may highlight a study that is investigating the same issues that you are addressing and you may find that the author(s) of the study is willing to share data with you or point you to a relevant data source.

Remember, even in the twenty-first century, despite all the online portals, clearinghouses, and server farms, not all geospatial data are online. A vast amount still resides on individual workstations, local intranets, and even CD-ROMs and other physical media. In many cases, the organizations or individuals have neither the staff nor the time to place the data online or cannot upload it because the data are restricted in some way. An often overlooked but effective way of obtaining data is the more traditional direct approach of contacting the data provider directly via e-mail or telephone. It may save you precious project hours or even days.

Social media resources are often more up-to-date than web directories. For example, blogs provide an excellent way of keeping in touch with new datasets posted online. Some bloggers, such as the OpenData Free GIS Data Tips, Publications and Open Resources blog and the Free GIS, Remote Sensing, Spatial, and Hydrology Data blog, focus on free spatial data. It may also be worthwhile following blog authors on Twitter, Facebook, Google +, or LinkedIn to stay up-to-date with the very latest information and technology updates.

Some organizations have Facebook or LinkedIn pages that provide regular updates on new data resources. In contrast to the often inadequately monitored contact facilities offered by many websites, social media tools have the added benefit of allowing you to send a message to the page administrators, or even the authors directly, to inquire about a particular posting or spatial data resource. As these pages are updated frequently, you have a much greater chance of getting a response to your inquiry.

Other considerations

Like many other digital technologies, GIS developed rapidly and became established before concerns were raised about the ethical implications of its use. As a user of GIS,

you are responsible for the ethical use of the information and the conclusions you derive:

Are you propagating any errors in the data?
Are you protecting people's privacy in the data you are using?
Are you creating maps that are clear and unambiguous and will not mislead your readers?
Are you disclosing your data sources?
Are you being fair in your assessment of the problem?
Are you respecting people's lives, rights, and property?

In chapter 4, we noted that spatial data are a means to an end, not an end in themselves, in that data provide the basis for analysis and generate the information for decision-making. If the data you choose are inappropriate or inaccurate, the value of anything you derive from those data will always be tarnished.

FOR FURTHER READING

To learn more about planning GIS projects, an excellent place to start is the book: *Thinking about GIS: Geographic Information System Planning for Managers* (Esri Press 2011) by Roger Tomlinson, a key founder of GIS. A companion book, *Building a GIS: System Architecture Design Strategies for Managers* (Esri Press 2012) includes a Capacity Planning Tool that guides the process of deploying and management of a GIS. The book *Mapping Global Cities: GIS Methods in Urban Analysis* (Esri Press 2006) illustrates how GIS can be used to plan and sustain urban areas. *The GIS Management Handbook* (Kessey Dewitt Publications in association with URISA 2009) from the Urban and Regional Information Systems Association provides information on how to develop, implement, and manage GIS programs and projects, aimed at managers.