Paper 056-2008

Getting Started with the ESRI Map Component in a Unix Environment for 2010 Census Planning

Bob Anderson, Qualex Consulting Services, Inc., Rockville, MD
Annette Quinlan Davis, US Dept. of Commerce, Bureau of the Census, Decennial Management Division, Suitland, MD

ABSTRACT
The basic steps in setting up the ESRI Map Component for geospatial display of data on a SAS BI platform are well documented. As we followed these steps at the Bureau of the Census to get from ESRI map development on Windows to map-displayed OLAP Cube data on our Unix-based BI Platform, we encountered several unexpected roadblocks. This paper reviews how our project evolved, highlighting in detail those experiences and solutions at critical points in the process.

1.0 INTRODUCTION
Looking for a job?

Between now and April 1, 2010, the Bureau of the Census will hire, train and deploy about a million temporary workers in order to conduct and support the field operations of the 2010 Decennial Census of the United States. These address listers, household enumerators, telephone workers, crew leaders, supervisors, office workers, and others, must be selected locally from applicant pools that vary widely in their cultural, socio-economic and demographic composition. This local variability impacts on census planning in at least two ways. First, local areas vary in how difficult they are to recruit from. Second, for similar reasons, local areas vary in how difficult they are to enumerate. In order to ensure a quality count, census planners need to select the right mix of people, methods and materials to optimize the accuracy and efficiency of the count in all the diverse areas of the country and its territories. Targeted marketing areas must be identified and an arsenal of electronics purchased, installed, and supported. All in all, a daunting task, to say the least, but the mandate is iron clad (Article 1, section 2, of the U.S. Constitution) and the deadline is firm. This work will be done, and it will be done on schedule.

To support executive decision makers in their planning, the Decennial Business Intelligence and Analysis (DBIA) team, a unit of the Decennial Management Division (DMD), was chosen to construct a new Planning Database for the 2010 Census (2010 PDB). While it is the Field Division (FLD) which actually conducts the crucial and labor intensive field work of the census, DMD is charged to ensure that the entire process is done on schedule, and within budget. The 2010 PDB is intended to provide not just an MIS, but an information factory to support the planning decisions of FLD and DMD planners for 2010 Census and beyond.

Based largely on its predecessor, the new PDB integrates data from the prior census and from other sources into a comprehensive decision support system on the SAS Business Intelligence (BI) platform with drillable Web-based reports and drillable ESRI maps providing online analytical processing (OLAP) detailed down to the tract, block group, and even to the block level. The geospatial component was required in order to meet the Bureau's need to quickly and accurately identify locations where preparations might be ahead of schedule, and where they are behind. With this tool in hand, the managers responsible for appropriately recruiting and enumerating in every area of the country will be able to perform the required analysis in a geographic context.

In his recent article in ESRI's ArcNews magazine, Jack Dangermond discusses the advantages of integrating maps into decision support systems like ours.

"GIS improves the way we do our work," he writes, "by facilitating better decision making; saving money, time, and resources; and allowing us to more effectively communicate through geospatial visualization." That, in a nutshell, is why our management directed that the 2010 PDB must include maps linked to the data for the reports we designed.

The geography dimension for the OLAP cubes that support the 2010 PDB is required to comprise the following levels: Nation, State, County, Tabulation Tract, Block Group, and Block. As we knew from reading the SAS 9.1.3 Intelligence Platform Administration Guide, the fundamental requirement was to construct an ESRI map with a boundary layer for each of those six levels, and with the attribute tables of those layers each having a column of

values matching one to one with the values of the corresponding levels in our OLAP Cubes' geography dimensions. We were left on our own to decide what other details would make the maps most useful for our customers. We elected to display highways, major roads and streets, important lakes and rivers, plus named places of significant population to help the planners recognize area boundaries quickly, and precisely, and to help them interpret the commonalities and disparities they might find among the various communities on the maps.

The core functionalities required of the application are:

(a) To display pre-summarized (OLAP cube) data in tabulations with drill down capability,
(b) To represent the values of variables in drillable choropleth geographic maps, and
(c) To link those maps to the underlying OLAP data in such a way that drilling up or down through the levels of geography in either the table or the map causes the other to display a corresponding change in the level of its display.

Figure 1 on the following page is a high level roadmap of the tasks you must complete in order for ESRI Maps to display with your Web OLAP Viewer data explorations and (after 9.1.3, SP4) your Web Report Studio reports. The purpose of this paper is to share our approach and the lessons we learned as we traveled this road for our 2010 PDB. Hopefully, knowing where we encountered the most challenges will save you some time along the way.

2.0 THE LONG AND WINDING ROAD

In this paper we group the steps of Figure 1 into five phases. These, and the crucial challenges we faced in them, are as follows:

A: Map construction (Steps 1, 3, and 4)
Challenge: Identifying desired drill paths, and suitable/sufficient sets of feature layers for our customers, and constructing the same in ESRI ArcMap.

B: Server configuration (Step 2)
Challenge: Establishing the connection between our ArcGIS 9 and our SAS BI Platform metadata server so that SMC's Server Manager and Map Service Manager Plug-ins can communicate with ArcGIS.

C: Map migration (Steps 5 through 10)
Challenge: Learning how to migrate our ESRI Maps from their homes in our Windows ArcGIS clients to the ArcGIS 9 server environment in UNIX.

D: OLAP Cube build (Step 11)
Challenge: Define the Geo dimension(s) so that the maps will be in synch with the OLAP cubes, and so that layers and data labels integrate seamlessly to provide users with an intuitive, informative, and easy to use graphic interface to their data.

E: Information Map definition and Report/Data Exploration (Step 12)
Challenge: Not a problem!

Figure 1: Adding Maps to SAS BI Web Client Reporting with the ESRI Map Component
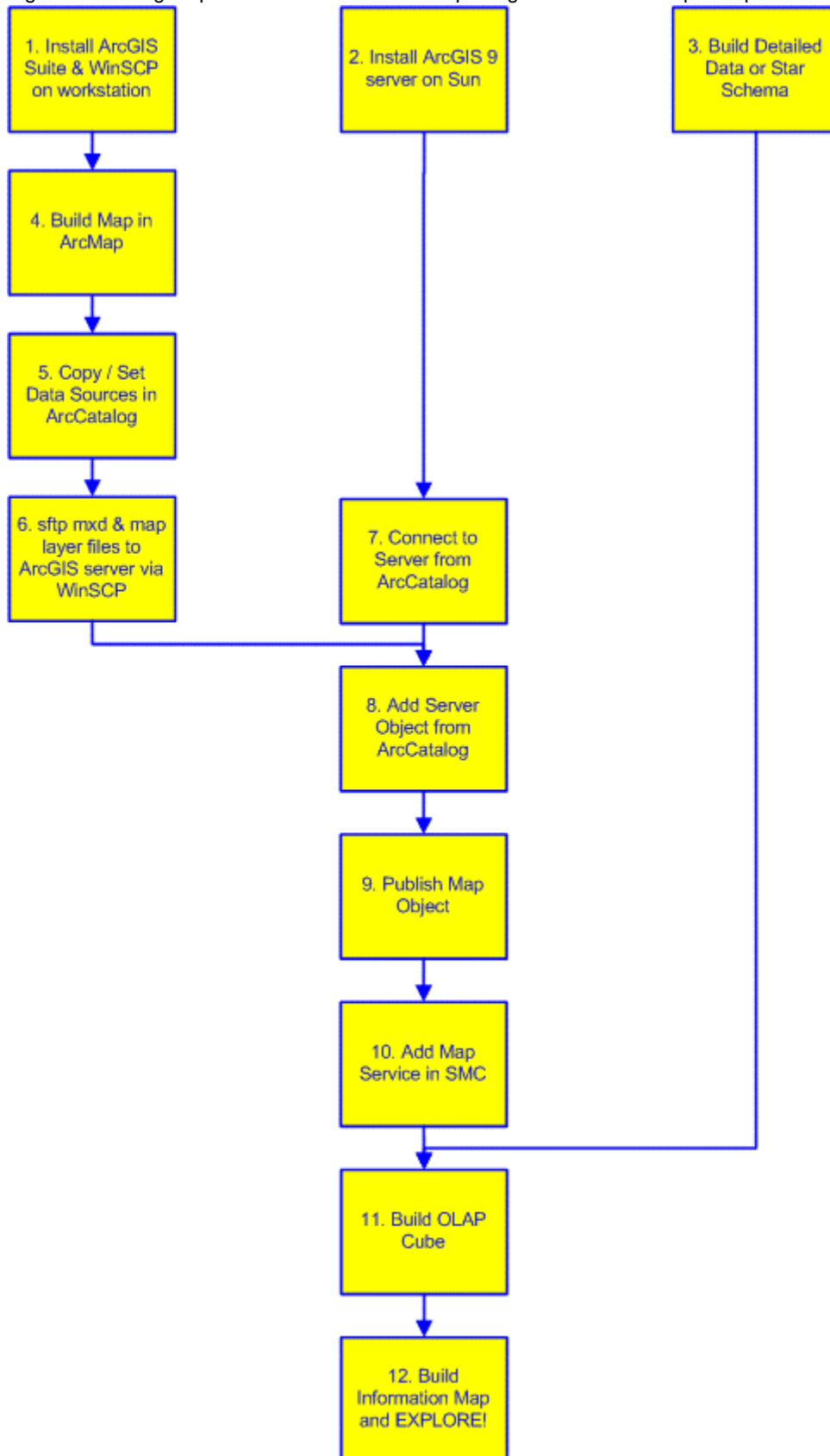
2.1 MAP CONSTRUCTION

Like most SAS professionals, we had limited formal training in geography and our combined professional experience in geospatial analysis was negligible when we began this journey. We had worked with PROG GMAP, of course, building drillable choropleth maps and such, and we were aware of the functionality in SAS/GIS, so we had a good idea of what we were trying to accomplish. But, we had virtually no idea how to integrate ESRI maps into our new BI web environment. So, the first challenge we faced was to figure out what SAS and ESRI products we would need to learn and use to implement the desired results.

LESSON 1: KNOW WHAT SOFTWARE YOU NEED: ESRI Map Component, ArcGIS Server, ArcMap AND ArcCatalog

Having an interest in this paper, you probably already know SAS offers two separate products that integrate SAS with ESRI:

    (1) SAS Bridge for ESRI, and
    (2) ESRI Map Component.

The SAS Bridge for ESRI is an add-on for ArcGIS. If your business is geospatial analysis, and your tools of choice for that work are ESRI's, then installing the SAS Bridge for ESRI on SAS and ESRI clients would greatly enhance your programming environment by supplementing it with the power of SAS Analytics. But if BI and data exploration for decision support is your game, your environment probably doesn't support a staff of ArcGIS application developers and geospatial analysts. For you, as for us, the ESRI Map Component provides all the required functionality.

At present, the 2010 PDB does not want its users in the regional offices or headquarters to apply sophisticated geospatial analysis techniques or to update the information displayed on their maps in real time as a result of such analyses. While they might want that capability in the future, it is not part of the current development. Rather, 2010 PDB only wants its users to have the ability to display OLAP data on an informative map, and to interact with the map in the way the ESRI Map Component allows. That is, as SAS describes it, "a feature that enables you to plot your OLAP data onto an interactive geographic map…" allowing you to "zoom, scroll the map, drill up or down, and expand or collapse regions."

Knowing what we needed to do the job, our next steps were (1) to construct the ESRI maps that would meet the requirements of our application, and (2) to build and export the map objects (.mxd) and underlying data to the ArcGIS server on UNIX. For these steps we needed ESRI's ArcMap and ArcCatalog software installed on our Windows XP Pro desktop workstations. We also installed WinSCP for secure ftp of the associated layer and attribute files.

LESSON 2: TRAIN YOURSELF (…AND/OR FIND AN EXPERT)

Here we met our third challenge: To learn and apply the art of map construction without blowing our production schedule out of the water! The ESRI ArcToolbox provides hundreds of tools for map construction and analysis. We were a bit overwhelmed! Fortunately, there is no shortage of excellent training at the ESRI online campus. (www.training.esri.com)

So, we began our journey by digging into the recommended sequence of courses. "Learning ArcGIS Desktop," number 1 on their list, provided a more than adequate introduction to both ArcMap, and ArcCatalog. Designed with geographers in mind, the course modules do place a premium on understanding projections, coordinate systems, and the like, but grinding through that was a small price to pay for the skills we acquired by following this well-structured hands-on introduction to the many powerful features, options and tools in the software.

Of course, there are always those times when nothing but expert advice will do. Here, again, we were lucky once more, for, in addition to having access to excellent online documentation and tech support both at SAS Institute and at ESRI, finding a live geography expert at the bureau is no problem at all. The place is crawling with them!

LESSON 3: START WITH EXISTING MAPS (www.census.gov or www.esri.com/data)

Another key to what success we have had to date has been not having to build much of anything from scratch. Neither should you! The Geography Division (GEO) at Census hosts an enormous electronic library of geospatial data, most of which is already warehoused in the industry standard geospatial database formats that ESRI software

expects, and much of which is publicly available. Some of our first maps were built from shape files we downloaded at www.census.gov. ESRI has a download site, too, and also provides their Census-based "Maps and Data Media Kit" free to ArcGIS customers.

The challenge of map construction should not be in finding the boundary and feature files you need, but in assembling them in such a way that (a) they conform to the geography dimensions of your OLAP cubes, (b) they load quickly, and (c) they provide sufficient (not excessive) detail at every drill level.

LESSON 4: AVOID BOUNDARIES THAT DON'T LINE UP. BUILD HIGHER LEVEL BOUNDARIES BY DISSOLVING DETAILED BOUNDARIES.

So, now you are underway. You've learned how to start building a map, and how to add and manage its layers. You've also learned a few techniques for modifying the appearance of the layers and the content of their underlying attribute tables so the map will do your bidding in the Web OLAP Viewer.

Lesson 4 is one of the most important early lessons we learned in map building with ESRI. In our first attempts we had tried to pull layers from different sources. We started with a U.S. and State boundaries map from ESRI, a counties shape file from census.gov, and local area details (tract, block, places, roads and hydrography) in maps we inherited from the parent PDB. The resulting map had fine and precise local boundaries, but where we expected to see layers sharing boundaries, the lines did not coincide. The problem was that, while the layers all employed the same projection, the state, and county map boundaries were designed for display at different scale levels. The states had highly generalized boundaries made from longer line segments that were intended to load more quickly, but were jagged and irregular at lower scales. Not surprisingly, these jagged U.S. and state boundaries did not match up at all well with the more densely segmented county and local files. While the offset at the state level might have been tolerable, the differences between state and local certainly was not. The local area boundaries, made up of a great many, very short segments, had no commonality at all with the increasingly lengthy segments of the state and national.

The solution, unfortunately, was to drop all the layers above tract (county, state, and national) and re-build them from the tract level up. To accomplish this task, we employed the dissolve tool in the ArcToolbox Data Management Tools -- Generalization section. In the following example, we illustrate how to use this tool to generate a boundary for a 9-county region in North Carolina.

Figure 2 shows the map we start with. It is limited to the 9 counties and has a feature layer with the boundaries of each. We want to create a single boundary for the entire area with no interior boundaries.

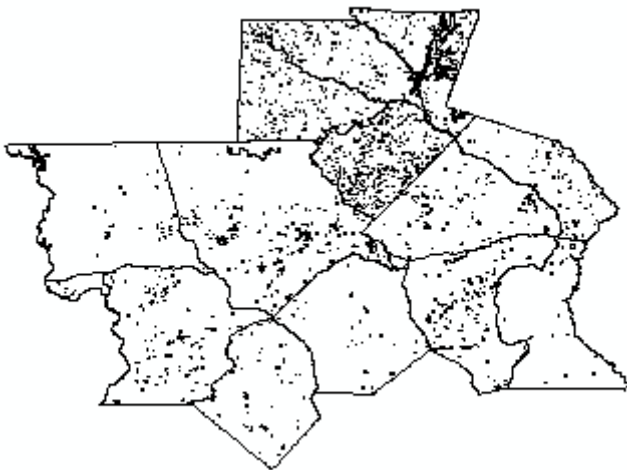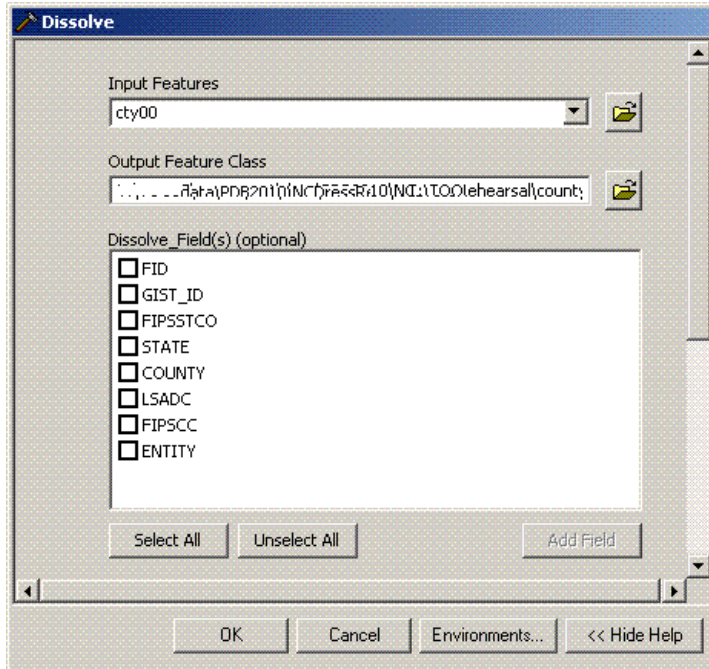Figure 2: Nine North Carolina Counties Before Dissolve.

Figure 3: Dissolve Menu with Options Filled

Double-clicking the Dissolve Tool opens a dialog window called "Dissolve." There are two text boxes and a text area at the top, followed by a text box and table for computing statistics for the new layer. The first list box, "Input Features," has a dropdown that lists the map features that you might want to dissolve into one larger boundary. When you choose one, ArcMap populates the second text box with a default path and file name for the new layer, as shown in figure 3.

To create an LCO layer from the 9 selected counties around Fayetteville, select the county layer (cty00) from the Input Features dropdown list. This will populate both the Input Features item, and the Output Feature Class, and provide a list of available fields in the Attribute Table that you want to include in the output Attributes table.

Since we did not choose a dissolve_Field, a single boundary layer will be output, which is what we want. When you save the map the new layer will be placed in the same directory folder as the layer out of which it was dissolved.

There are no statistics to summarize in this case, but if there were, you would specify them in using the text box and table at the bottom of the dialog (not shown here.) When you click OK, ArcMap creates the new layer. As the work is being done, ArcMap displays a log window so you can see the progress of the new layer being created. When the layer is built, it appears in the map layers tree.

LESSON 5: BUILD SUB-REGIONS FROM A LARGER MAP WITH THE SELECT AND CLIP TOOLS.

Often, you have a map of an entire region and you want to build separate maps of its sub-regions. You can subset the boundary areas you want from the larger map using the Select tool, then Clip the other layers as you overlay them.

The Select tool is found in the Toolkit's "Analysis Tools – Extract" section. It allows you to create a new feature by sub-setting the source with a query against its attribute table. In this example, we create a map of the Census Bureau's West South Central Division by selecting U.S. from a map of the 50 states and major U.S. Rivers, shown in Figure 4.
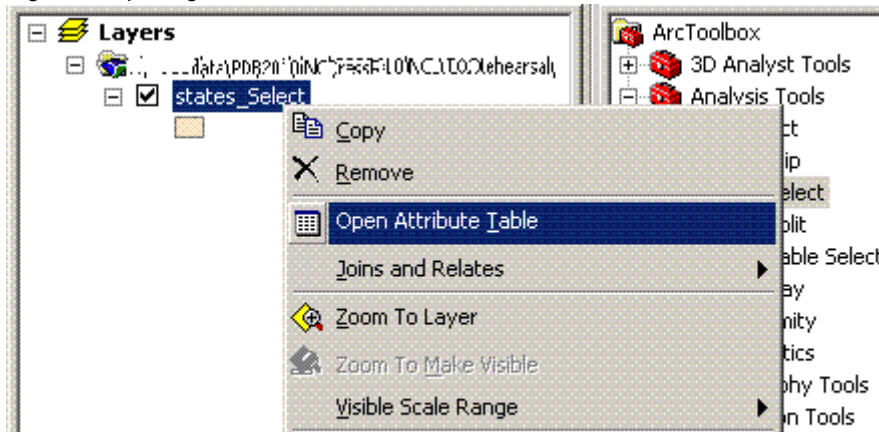
Figure 4: U.S. Map with Rivers Layer.
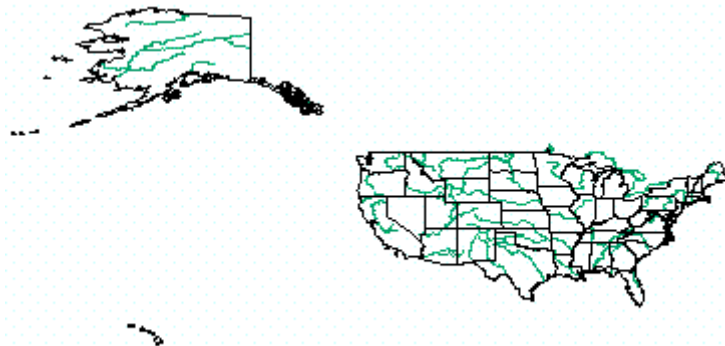
Examine the attributes table of the layer you want to sub-set to determine the criterion you will use in your query. (See Figure 5.)

Figure 5: Opening an Attribute Table

Looking at the table in Figure 6, you can see that the query (where clause) you will need for this sub-set will be "SUB_REGION='W S Cen'"

Figure 6: Attribute Table for States

Now start the Select tool by double-clicking its name. This opens a dialog titled "Select" which has three text boxes. Like the Dissolve tool, the first box has a dropdown list of optional layers to select from. In Figure 7 we have selected the states layer, and again, ArcMap has put the path and name of the physical layer file to be created in the "Output Feature Class" text box.

Now you just need to tell ArcMap what criteria to use to select the sub-set of areas to select. This where clause expression, which we discovered above, goes in the text box labeled "Expression (optional)." Here, you have three choices. You may (1) leave the box empty, in which case all areas will be selected; (2) type in the expression; or (3) build the expressioin using ArcMap's Query Builder. We like the Query Builder because we don't have to type as much, so we click the 'SQL' icon.

Figure 7: Select Dialog with Feature Chosen
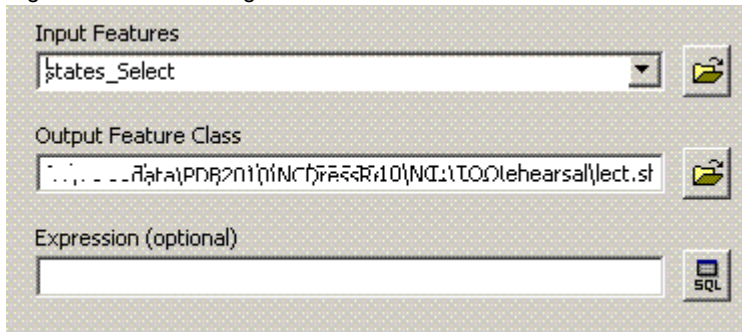
Figure 8 shows the Query Builder dialog, and illustrates how we used it to select our West South Central states.

Figure 8: Query Builder
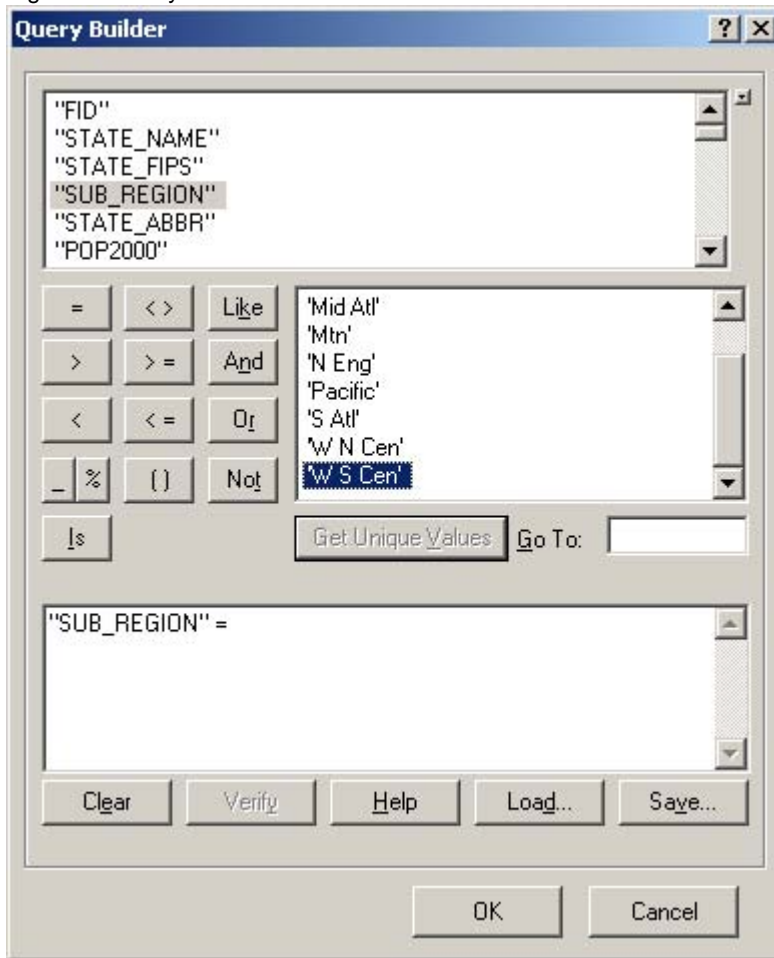
We no longer need the original states layer, so right-click on its name in the layers tree and "Remove". Now the map only shows the S W Cen state boundaries. However, it still shows the rivers of the entire country. We fix that by building a new rivers layer for just the new sub-region. To do this, we use the Clip tool, also in the Analysis Tools – Extract section of the ArcToolbox.

ESRI describes this tool as a "cookie cutter". You tell it what boundary you want to use as the "cutter" in one text box ("Clip Features"), and what Input Features you want to cut (i.e., Clip) in the next. The result is a new layer which contains only the portion of the original feature which fits inside the boundaries of the Clip Feature. Again, after you have created the clipped feature layer, "Remove" the original, unclipped, layer, and expand the map to its Full Extent by clicking on the globe icon in ArcMap's toolbar.

Figures 9 and 10 show the map before and after clipping the rivers to contain only the parts that flow through the West South Central states.

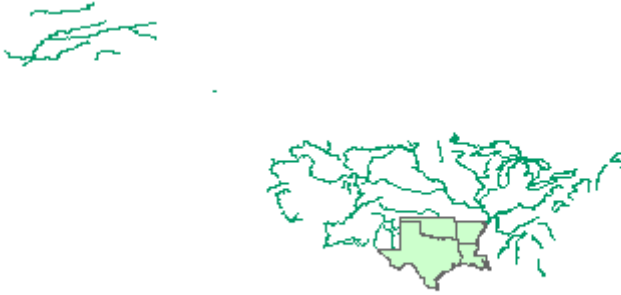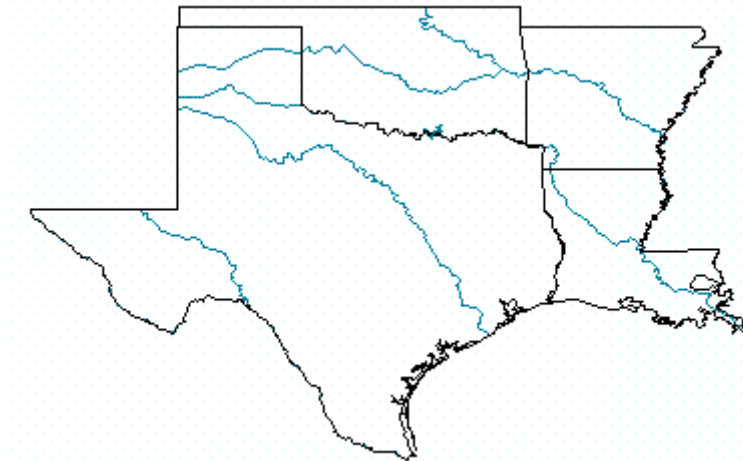Figure 9: Sub-Region Selected but Rivers Not Clipped

Figure 10: Sub-Region Selected and Rivers Clipped

LESSON 6: BUILD FROM THE BOTTOM UP, AND DON'T COLOR-FILL POLYGON FEATURES.

An early lesson in map building with ArcMap is that you order the layers from bottom to top. In this way, the most local features (e.g., streets and rivers) do not prevent higher level boundaries from being displayed when multiple layers appear in the map viewer simultaneously, and the Web OLAP Viewer will know the boundaries that correspond to the current hierarchy level.

While color choices are very important when it comes to boundaries, it does no good that we know to color fill your polygon layers in ArcMap. Leave them transparent (Fill = "No Color"). First of all, you know that the Web OLAP viewer will be assigning colors to the areas according to the values of the selected variable being displayed, so you wouldn't want to confuse the picture with other colors, even if the Viewer would display them. Furthermore, if you color fill areas in ArcMap while building the map, you will not be able to see how underlyinig features will look in the the final product.

LESSON 7: DIFFERENT MAP AREAS AT THE SAME DRILL LEVEL DISPLAY AT DIFFERENT SCALES. YOU MAY NOT BE ABLE TO FORCE FEATURES TO DISPLAY AS YOU WANT THEM TO IN ALL AREAS WHEN YOU DRILL DOWN IN THE WEB OLAP VIEWER.

To "zoom in" or "zoom out" on most internet map applications you simply select from a finite set of points on a scale. Google maps have 20 levels. Mapquest offers 14. (Never mind Google Earth!) At each level you only expect to see boundary and feature details (roads, lakes, rivers, etc.) appropriate to the scale. As a practical matter, for example, there just isn't enough screen real estate to display the names of all the neighborhood streets in Dallas on a map whose scale encompasses all of Texas. Would we really want to identify every small town in the Southwest on a map that displays the entire region? Probably not.

Fortunately, ESRI Maps will automatically prevent excessive labeling based on certain rules about font size and overlapping labels. By default, it will not overlay feature labels or put them at an incorrect place on a map, given its scale. While these automatic restrictions keep maps from getting cluttered, they often result in some, but not all feature labels appearing at any given zoom level. In hopes of making our features and boundaries appear at once and be labeled at once, we chose to take control over when our features and their labels would appear. Here we learned through a lot of trial and error.

ESRI has designed intuitively simple controls into ArcMap that allow you to set scale ranges with ease, once you know the ranges you want. Simply select the Out beyond: and "In beyond" scale for features and for labels in the properties of each layer. Figure 11 shows the Scale Range section that appears on the General tab of our states layer properties window. You can set these ranges for the feature here.

Figure 11: Scale Range for Layer
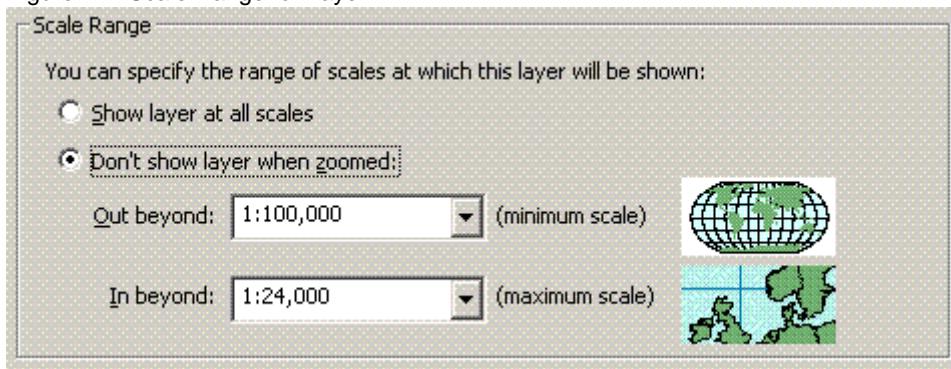
To set the ranges for the labels, click the "Scale Ranges" button on the Labels tab.

The trick, of course, is in knowing the scale where you want each feature and label to appear, and the scale where you want them to disappear (if ever) as you drill in and out on your maps.

At the national level, we wanted to reveal state boundaries, but nothing more. At the State level, we wanted all the county boundaries within the state to show, but not the tracts. Not until the user drilled to the county level did we want detailed features such as streets to appear. The problem, as you might suspect, is that counties are geographically diverse, ranging in size from 26 to 87,860 square miles, with the smallest 10% being less than 350 and the largest 10% being over 1,270.   To have roads display in every county, you would have to set the display ranges for the scale that will display the largest county. At that scale, local streets would be cluttering the maps of many states! If you set the display range so they appear only when you drill the scale of a typical county, SAS BI users will have to use the zoom tool on large counties to see these layers. Simply stated, the right scale in some regions will not be the right scale in others.

As we saw above, ArcMap allows you to choose the scale at which boundary and feature layers will appear. In the SAS Web OLAP Viewer, the scale of the map depends on the size of the targeted area. When you drill down into geography levels in SAS BI, the Web OLAP Viewer displays the selected area at the lowest scale that still allows the entire selected region to be shown in the Viewer's available map space.

We found the best solution to this problem was to resign ourselves to allowing some boundaries to appear before they are large enough to label, and to train our users to use the zoom tool instead of drilling down, when they want to see

more detailed information. (When you use the zoom tool to take a closer look at a smaller area within a region, you do not change the cube information displayed, but you will surface other layers of ESRI information at the new scale.) We also trained our users to use the "View details about a region" icon on the map's menu bar in the Web OLAP Viewer to see the details of specific regions on the map.

The ArcGIS map server can serve partitioned maps, if the calling application can accept them, and that might lead to a better solution to the problem. OLAP Cube Studio expects a single map object with all layers arranged hierarchically in parallel with the GEO dimension's hierarchy. We have not learned how to use partitioned layers in our map construction, yet.

LESSON 8: TAKE ADVANTAGE OF ATTRIBUTE TABLES.

Most often, the physical source behind your ESRI map layer is a "shape" file. Shape files are actually composed of multiple files. For example, if you look into the source directory structure of our 2000 census tracts layer (trt00) you will see trt00.shp, trt00.shx, and trt00.dbf, corresponding to the geography, the index of the geography, and the associated attributes. The .dbf, or dbase file, is what ArcMap refers to as the Attributes Table. Among the fields in these tables are the items which must match with the geo dimension levels for your maps to work in the Viewer. There are many times when you will need to add or modify information in the .dbfs.

Figure 12: Options on the Attributes Table Editor.
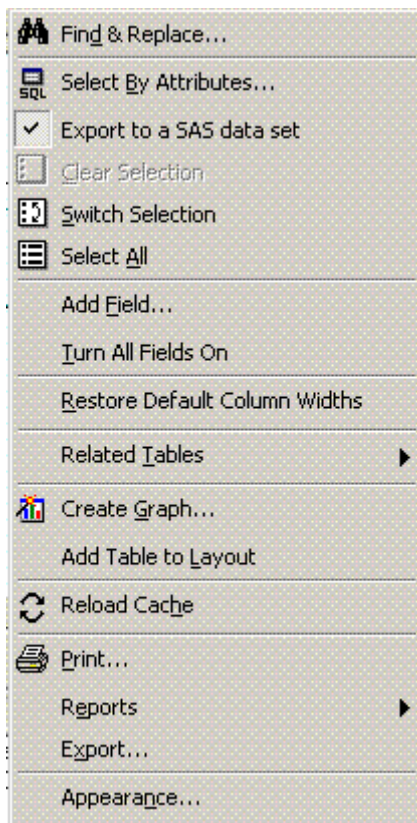
One way to modify the columns is by right-clicking on a column name and then selecting "Field Calculator" to add an expression. You can also open the dbf in Excel, modify it, and save it a .dbf. However, if you want to add a column using this method, you must first add the new column to the attribute table using the "Add Field" tool (in ArcToolkit section "Data Management Tools – Fields"). You can also access this tool directly from the attributes table by clicking the "Options" button at the bottom of the screen, revealing all the options shown in Figure12:

2.2 SERVER CONFIGURATION

We can't say enough good things about our great partners in the server support area, Decennial Systems and Processing Office (DSPO), whose server administrators build and support our SAS BI and ArcGIS servers.   While we were busy having fun learning to build maps in ArcGIS, DSPO moved forward with the installation, configuration, and support of our many UNIX system servers.

LESSON 9: UNIX SOLARIS VERSION MATTERS.

SAS and ESRI share many two great strengths: both evolve rapidly as customer needs evolve, and both have a long history of supporting multiple operating environments. In practical terms this means that whenever a supported operating system plans a new release, SAS and ESRI are likely to release patches, too. If the marriage between them is strong, such changes will be compatible, and will release at about the same time. From the customer's point of view, server admins have their hands tied with regards to upgrading the OS until all supporting software applications are up to speed. So it was in Spring of 2006 in DSPO. We were still running our first BI prototype on Tru64 servers. Solaris was still at version 8 on some machines, and at version 9 on others. Everyone was anxious to get off Tru64 and on to Solaris 10.

The first attempts to install ArcGIS 9.1 under Solaris 10 led to a lot of troubleshooting complicated by the fact that ArcGIS 9 would not support containers (zones) that are a standard feature of Solaris 10 installs. After a lot of back and forth between tech support at ESRI and at SAS, the machine was abandoned in favor of a Solaris 9 server, which ArcGIS 9.1 supported, for our early prototype. Ultimately, we did we did move forward to Solaris 10 *without* employing zones, but only after getting this configuration certified for support by ESRI.

LESSON 10: UNDERSTAND HOW THE COMPONENTS OF YOUR ArcGIS SERVER COMMUNICATE

We've already mentioned that our ArcCatalog client is on Windows, and our ArcGIS Servers are on UNIX. To do its job of adding and managing services on GIS for your maps, ArcCatalog must establish communication with ArcGIS. To do this always involves authentication so your systems are secure. Under ArcGIS 9.1 on UNIX, this involves three processing components: MainWin, ArcGIS Server, and the ArcGIS IdentityServer. Our system administrator describes the setup: "The MainWin process listens for requests for ArcGIS Services. It passes the login information to the IdentityServer for authentication, and finally communicates with the actual ArcGIS server." As developers, we must authenticate via the Identity Server to add our map services. This is where things broke down. What we saw in ArcCatalog was that we could not connect to the ArcGIS Server. The reason was that the user's network password, not the local host password, had to be set in the IdentityServer. The solution was for developers to synch up network and server passwords. That has since been remedied.

Other little problems came up, too, many of which could have been avoided by giving the workforce a little more prep time. (See The Enterprise Melting Pot (Part II), Paper 055-2008 elsewhere in these proceedings.)

Sometimes, server problems seem to be broken, when version control in BI is really the solution.   When we first attempted to identify our map server in the SAS Management Console (SMC) New Server Wizard, "ESRI Map Server" was not in the list of content servers as it should have been. Here, the solution was simply to upgrade the SMC and Server Manager Plug-in to 9.1.3 SP4.

The configuration we run today is Solaris 10 without zones, WebSphere 5.1.1.5 (very backlevel, but suport is certified for SAS BI 9.1.3), and ArcGIS 9.1 (presently going to 9.2). ArcGIS 9.2 has simplified and streamlined the installation procedures and avoids many of the pitfalls of 9.1. One little detail to be careful about, though, is how you setup the owner account for the install. (Lets call it agssrv). This user is to be the owner of all 9.2 libraries and processes. If he cannot issue xvfb (x-windows virtual frame buffer) commands, you may have to modify his .profile.

More recently, a patch to Solaris 10 had to be backed out because it broke the ArcGIS server. Again, our server support folks advise that you beware of recent Solaris patches that impact on the xvfb.

Even after we were sure we had the all the configuration issues worked out, there were times when we would see map objects in the Map Service Manager plug-in on SMC that had no layers to choose. We've run into this problem more than once, and the cause has always had something to do with the underlying shape files, not the configuration. Either the layers they were not named correctly (see section 2.3, below), or they were simply placed in the wrong

subdirectory in the ArcGIS Server. In any case, the problems were weeded out one by one, and in the end, our layers did begin to appear.

2.3 MAP MIGRATION
As mentioned in the previous section, it was not a trivial exercise to create ArcGIS map objects (.mxd) from the maps we created on the WinXP platform. Its a tricky process and its not particularly well documented anywhere we could find, so we offer it in its own section here.

Having created our initial map document in ArcMap and configured our server appropriately, we can now move the map document over to the UNIX machine. Once all of the files are moved, we are one step closer to creating a map object that the metadata server will recognize.

The first time we moved our map document and all the associated layer files over to the UNIX machine, we naively moved all files as they existed in the windows environment. We soon discovered that this method was not going to give us what we needed to create a map object on the ESRI server.

LESSON 11: DO NOT USE SPACES IN WINDOWS DIRECTORIES OR FILE NAMES WHENEVER POSSIBLE.

It may be fine to use spaces in the directories and file names when you are creating your maps; however, UNIX does not always recognize these spaces. You need to remove all spaces from the directory paths and file names for the map document and all associated layer files. You will use ArcCatalog to complete this file preparation step.

In ArcCatalog, copy all layer files and the map document to a work directory you've created in Windows. This ensures that even after you change all file names and references, you will still have a working map in your Windows environment. In the work directory, you want to review all layer file names and remove any spaces.

We now need to update the map document to recognize any new layer names and define the paths the map document will use to locate the layer files once it's moved to the UNIX machine. We will do this using the "Set Data Sources" option of the map document file from ArcCatalog.

LESSON 12: DO NOT USE UPPERCASE CHARACTERS IN UNIX PATHS

When you right-click on the copy of the map document in your Windows work directory, select "Set Data Sources" to change all path references. Before doing this, you will need to determine the paths where you will store your layers and map document file on the UNIX machine in order to assign them through Set Data Sources. These paths must correspond to the paths that were defined in the Server Configuration. In the set data sources window, by default, your Output File will automatically be saved to the current location with "Copy_of" preceding the map document file name. This creates a new copy of the map document that will contain the new paths. You will be moving this new map document file to the UNIX server.

After you enter a new file name, you will focus on updating the layer files. In the Data Frame section of the window, select <All> to display all data frames. You should now see all data frames and layers within your map document. You want to update the "New data source" column to reflect the directory location where your layer files will be stored on your UNIX server. In some cases, you can use the "select all" and "replace all" options to modify parts of or entire path names. The UNIX path names cannot contain uppercase characters! At this point you will also want to verify that any layer names you modified earlier to remove spaces are also updated here.

You are ready to start moving all of these files to the UNIX machine. We used a windows-based secure ftp session for this process; this allowed us to drag and drop all of the necessary files to the new locations. We moved all layer files and the map document file to the directory paths we identified in our server configuration and in the set data sources dialog box.

Back in ArcCatalog, we will build a Map Object on our GIS Server using the new map document file we just moved to the UNIX server. The Map Object is what you will see surfaced through the metadata server in SAS Management Console when you go to create the map service. The map service is what you will need when you identify the map in your OLAP Cube.

Connect to the GIS Server you created during the server configuration and select "Add Server Object" to open the New Server Object wizard. You will need to identify a name for you map object, this is the name that will appear in

SAS Management Console when you create the map service. You will then need to identify the path and file name of the map document that you just moved to your UNIX server. For our purposes, we were able to accept all other defaults in the New Server Object wizard. If you do not start the object when you create it, don't forget to "start" the map object from ArcCatalog. You will only be able to see map objects that are started on the GIS Server. With ArcGIS 9.2 you can also manage your map services in the ArcGIS Manager web application once you have added the object to the server from ArcCatalog.

The last step in our map migration process is to define the map service through SAS Management Console (SMC). This is clearly documented in the SAS 9.1.3 Intelligence Platform: Web Application Administration Guide , but we will summarize here. In the navigation tree in SAS Management Console, right-click the Map Service Manager and select "New Map Service" to open the New Map Service Wizard. You will enter the name of your new map service; this map service name is what you will need when assigning your map to your OLAP Cube. The Configuration you want to select is the map object you just finished creating in ArcCatalog. Next, you will select all of the layers you want to include in your map service to display with your OLAP Cube data. You must select a layer to correspond to each level in your OLAP Cube. Keep in mind you can also choose to display additional layers on the map that are not associated with any levels in your OLAP data. After selecting the layers, you need to define the field for each layer whose data corresponds with a field in your OLAP Cube data. For additional layers, you can select any field from the map data.

2.4 OLAP CUBE BUILD

At long last, we can link our OLAP Cube with our ESRI map! We will accomplish this by identifying the Geographic dimensions and hierarchies that match the ESRI maps. You can either add the ESRI information to an existing cube or create a new one. To configure your OLAP Cube for ESRI integration, open OLAP Cube Studio or Data Integration Studio, whichever tool your organization has decided to use to build OLAP Cube data. This procedure is well documented in the SAS 9.1.3 Intelligence Platform: Web Application Administration Guide and we will again summarize here as to, hopefully, provide one continuous document which you can use as a reference to integrate geographic mapping capabilities into your SAS BI Platform.

In DI Studio or OLAP Cube Studio, navigate in your existing cube or create a new one to the point where you define the dimensions in the Cube Designer. If you are defining a new cube, this is where you create your dimensions and hierarchies, including the dimension that contains the geographic hierarchies. When creating a new cube, in the Dimension Designer wizard for the geographic dimension, select GEO from the Type drop-down list. If you are editing an existing cube structure, the dimension that contains the geographic data should already be created. Select the dimension that contains the geographic data and click Modify to open the Dimension Designer wizard; from here select GEO from the Type drop-down list, if it has not already been selected. Follow all other steps you would normally take to create your dimensions in your cube.

After assigning the GEO type to the appropriate dimension and completing the steps to create or modify the dimension, select the dimension that you just modified, or created, and click Specify Map. In the Specify Map dialog box, select the Map Server you created in the server configuration section (2.2) and the Map Service you create in the map migration section (2.3). This will automatically populate the Levels selection box.

For each dimension level in the Levels selection box, you must identify the Map Layer and the Map Field Id from the Map Service; and the Field Id Column from the Cube Input Table. All levels in your dimension should be linked to a layer in your map service. The Map Layer is the ESRI map layer that corresponds to the selected OLAP dimension level. The Map Field Id is the field in your ESRI map attributes table that uniquely identifies the regions in that layer. The Field Id Column from the Cube input table is the field in your OLAP data that uniquely identifies the regions in that level. The Map Field Id and the Field Id Column must contain identical values for each level even though the field names do not need to match.

LESSON 13: THE LEVEL DOES NOT NEED TO MATCH THE FIELD ID COLUMN IN THE SPECIFY MAP DIALOG BOX.

This lesson took us a little while to discover. Your levels can be independent of the Field Id Column – only the Map Field Id needs to be identical to the Field Id Column. For example, you have layers in your cube for states and counties, but for the cube display, you only want to see the individual state abbreviation and county name. In your map, your county layer combines multiple states and in order to have unique county codes you must concatenate your state and county values to create the Map Field Id. You layers can still show individual states and counties as long as you have an additional field in your cube that concatenates the state and county codes. It is this combined field that you will use as your Field Id Column to link to your Map Field Id.

Continue creating your cube as desired once you have completed specifying the map information.

LESSON 14: DOCUMENT, DOCUMENT, DOCUMENT

Many of these steps may seem very straightforward, and they are, as long as you keep very good notes along the way. There are countless times where you will need to backtrack if you don't know: the map document name, path, map object name, map service name, the map layers, the unique identifier in the map layer that you will need in the cube, the fields in the cube that uniquely identify the levels, etc.

Finally, as we've mentioned several times, you will doubtless accumulate a huge list of links to technical doc and training materials from SAS, ESRI, and elsewhere. The experiences we've shared here barely scratch the surface, to be sure.

2.5 INFORMATION MAP DEFINITION AND REPORT/DATA EXPLORATION
You are ready to create an information map and a report or data exploration using your new cube.

There are no additional or modified steps in creating your information map, its business as usual.

In Web Report Studio and in a Data Exploration (using the Visual Data Explorer in the SAS Information Delivery Portal) you will now have the option to add a geographic map object. Look for the green and blue globe icon in the toolbar.

Have fun exploring your new map!

CONCLUSION
The ESRI Map Component in SAS BI 9.1.3 (SP4) with ESRI maps served by ArcGIS 9.2 under Solaris 10 will add an important and informative geospatial dimension to support decision making for planning the 2010 Decennial Census. The process of implementing this functionality has been full of the kind of pitfalls one would expect with new software being implemented on platforms where only a few have installed them software before. We suggest that considerable effort be dedicated to learning the many powerful tools and techniques available for map building in ArcMap. Further, we advise paying exceptionally close attention to software configuration and version control details in both the WinXP and Solaris environments. Careful selection of file names and target locations for map objects and their underlying files is essential.

The maps and OLAP Cubes we have used to date have not been volume tested with a full-blown map of every census block in the U.S. and its territories. Doubtless, we will learn (and unlearn) many lessons along the way.

We hope this paper will prove useful as a guide to others who want to go down this long and winding road. We'll share more as we learn more. Please do the same. Stay tuned!

REFERENCES

Dangermond, Jack. "A Framework for Understanding, Managing, and Improving Our World: GIS – The Geographic Approach," *ArcNews* vol. 29, No. 3, ESRI, Fall 2007.

SAS Institute Inc. (2006), "Overview of the ESRI Map Component", in *SAS 9.1.3 Intelligence Platform Administration Guide*, Cary, NC.

National Association of Counties. "County Square Mileage," http://www.naco.org/Template.cfm ?Section=Data_and_Demographics&Template=/cffiles/counties/ sqmil_srch.cfm, 2005, (2 Feb. 2008).

SAS Institute Inc. (2007). "Configuring the ESRI Map Component", in *SAS 9.1.3 Intelligence Platform: Web Application Administration Guide*, Cary, NC.

Davidson, Gail, and Mills, Guinevere. 2008, "The Enterprise Melting Pot (Part II), SAS BI – The Plan versus Reality Or Don't Try This at Home, Kids," Paper 055-2008 in *Proceedings of the 2008 SAS Global Forum*, San Antonio, TX.

ACKNOWLEDGMENTS
Thanks to ESRI and SAS Technical Support staffs for many hours of crisis intervention.
Thanks to Guin Mills and the rest of the DBIA Team for picking up some slack while we figured this stuff out.
Special thanks to Albert Hotton in DSPO at the Bureau of the Census for his dogged perseverence in resolving every problem on the server side.

CONTACT INFORMATION
Your comments and questions are valued and encouraged. Contact the authors at:

Bob Anderson	Annette Quinlan Davis
Qualex Consulting Services	U.S. Department of Commerce,
301-763-9359	Bureau Of The Census
bob@QLX.com	301-763-8884
robert.m.anderson@census.gov	annette.quinlan.davis@census.gov

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.
Other brand and product names are trademarks of their respective companies.