

# Organizing a Century of Data

*Spatial data infrastructure created by world's leading gold company*

By Matthew DeMeritt, ESRI Writer

As the largest gold mining company in the world, much of Barrick's stock market value depends on its reserve base (i.e., gold proven to be minable but not yet mined). With annual gold production of nearly 8 million ounces, Barrick must add at least that much gold to its reserve base each year or its stock market value will decline. Consequently, Barrick's geological data directly affects the company's bottom line.

## Impressive Shelf Life

Geologic data has an unusually long life cycle. Because ore-forming processes act over millions or tens of millions of years, data collected decades ago is still relevant to any mining company. Barrick has more than a quarter century of self-collected and purchased data as well as more than 100 years' worth of data inherited from acquired companies. Effectively managing this rapidly growing collection of geology, geochemistry, geophysics, and remotely sensed data is crucial. To make best use of its impressive data archives, Barrick recently began an initiative to better organize and serve its data throughout the company, and a spatial data infrastructure (SDI) was born.

Within a mining company, different departments often have data that other groups can use. At Barrick, when surveyors collect as-built data for a mine site, that data can be of interest to the Mine Geology, Security, Health and Safety, and Exploration groups. An SDI simplifies the way spatial data is disseminated and accessed throughout an organization. "Although our vast store of exploration data was the catalyst for the initiative, the project grew to include Environment, Land, Security, and Health and Safety [groups]," said Iain Allen, global spatial data systems coordinator for Barrick. "Making every group's data available to everyone else eliminates the likelihood of duplicate data collection and/or purchasing."

## The Migraine of Fine Grain

The crucial first step in creating an easily searched data catalog is generating metadata for each dataset. However, as anyone who has created metadata knows, the devil is in the details. "Everything depends on metadata, but no one likes to do metadata," laughed Allen. "We use a custom metadata editor, MetaTools 5, by Peter Barrs of Data Arterial in Stanthorpe, Australia, which runs in ArcMap as well as ArcCatalog." Barrick standardized on 14 mandatory metadata elements and four optional ones. "We do not have dedicated data managers, so we depend on the end user for metadata. If it's too tedious, they won't do it. We tried to find a balance between everything we would like to know about a dataset and what we could realistically expect people to do."

The Barrick metadata editor uses pick lists for every metadata attribute. The pick lists automatically sync with a server-based Master Keyword Database every time the editor opens, ensuring everyone always has the most current set of keywords. Adding new keywords is done through a Web form, which adds them to the master database. The editor also facilitates the copying of metadata. Working in ArcMap, where the datasets typically have many common metadata attributes, can greatly speed metadata creation. The user completes all the common metadata elements for one dataset, copies them to all the other datasets, then completes the one or two remaining variable elements individually.

## Publish or Perish

Detailed metadata is worthless if the data can't be queried and accessed. To publish the data, Barrick uses MXD Publisher, created to Barrick specifications by NGIS in Australia. MXD Publisher, which will not allow data

to be published if it does not have complete metadata, creates a file geodatabase at a specified location on the network, also capturing all the layer files and creating a copy of the MXD that references the newly created file geodatabase. "This is our way of overcoming 'C-drive syndrome.' We find that a lot of data exists only on a particular user's laptop, so no one else even knows it exists," said Allen. "Through the MXD Publisher, all the datasets used in the published MXD are copied to the network drive. Data previously available only on their laptop is now also available, with complete metadata, on the network."

The final step in the Barrick workflow is SDE Publisher, another custom NGIS tool, that publishes the file geodatabase on the network to the enterprise geodatabase in

ArcGIS Server. Each night a metadata harvester runs and updates the data catalog with the metadata for any new datasets (the data catalog uses the ESRI GIS Portal Toolkit [now known as the *ArcGIS Server Geoportal extension*] for presentation and ArcGIS Server with SQL Server for storage). Dennis Geasan from GIS Technologies in Anchorage, Alaska, was instrumental in setting up

**"Everything depends on metadata, but no one likes to do metadata."**

—Iain Allen

the back-end processes.

Search results include links to the individual datasets and to the source MXD. The results are presented with selected metadata, and buttons allow users to either stream the dataset or complete MXD directly from the server to their desktop client or download the data directly. FME Server is used for the ETL (extract, transform, and load) function, giving Barrick the ability to deliver search results in any format desired. This is important because Barrick supports two desktop GIS packages.

#### Data Security

"People tend to be very protective of their data, so we have two 'data security' related metadata attributes," Allen said. One controls metadata visibility, the other controls data

*Continued on page 22*

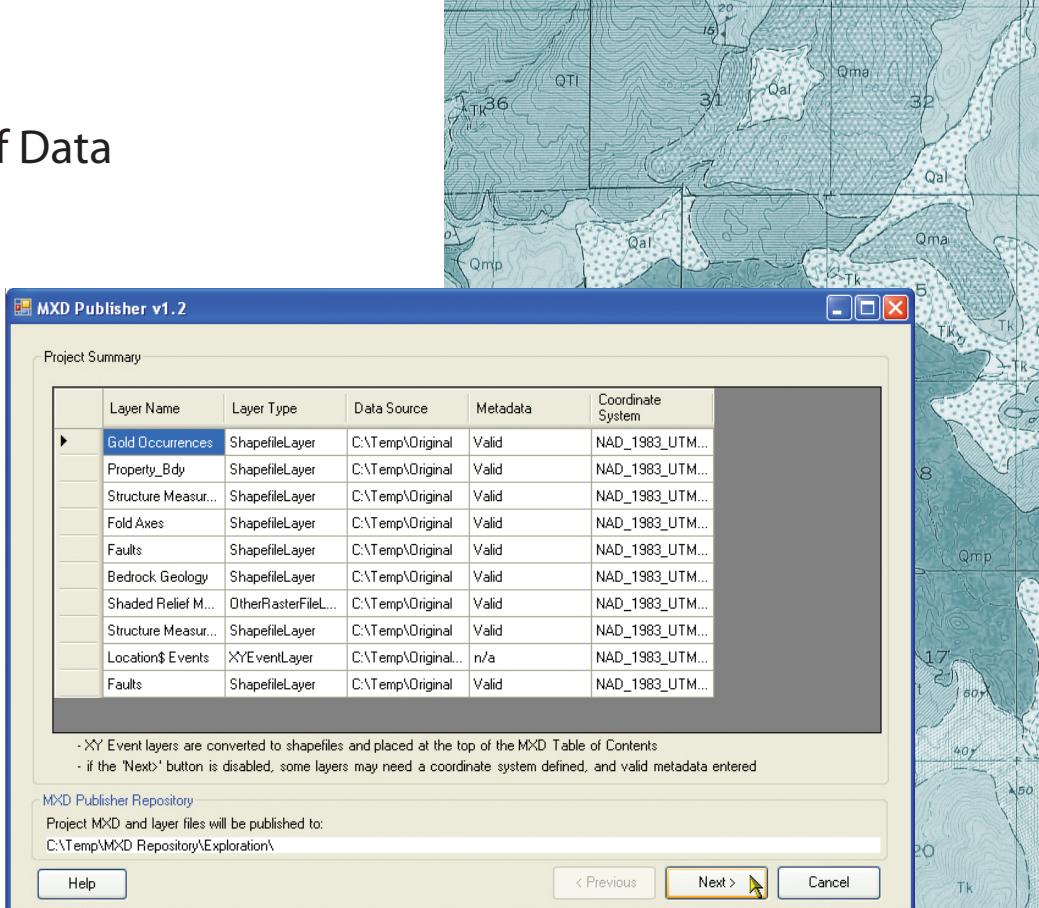
# Organizing A Century of Data

Continued from page 21

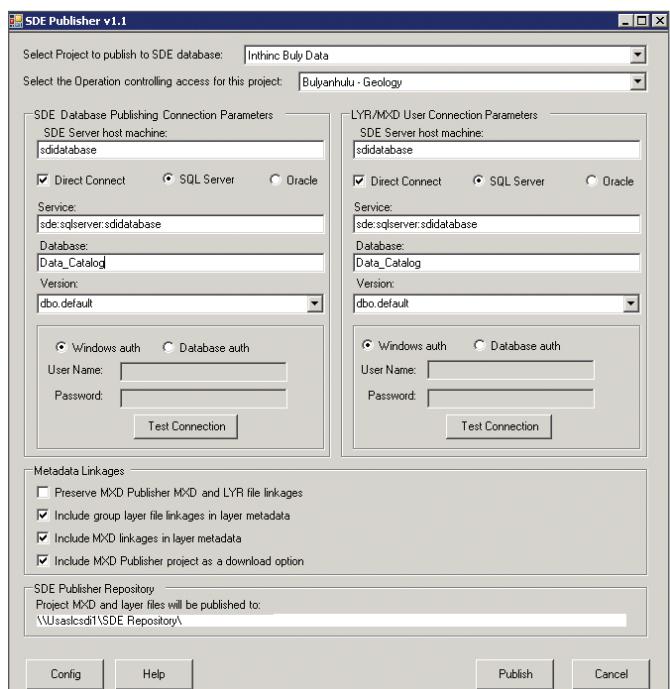
access. This gives Barrick the flexibility to let everyone know a dataset exists. When the metadata is added to the catalog, it will show up in search results. However, it restricts data access to the group most likely to need it. If others want access, they must contact the persons listed under the Responsible Party metadata attribute and make their case. For example, Barrick's land managers want people to know whenever there is land data available for their search area, but they do not want people to have access to that data without consulting a land professional. This procedure ensures that anyone using the data is made aware of any confidentiality or area of interest agreements associated with the land polygon.

## Future

By tagging its huge store of geologic information with metadata, Barrick has taken a significant first step in increasing the value of this data. Future work will include incorporating spatial data stored in other repositories, such as acQuire from acQuire Technology Solutions for geochemical and drilling data and EQuIS from ESRI business partner EarthSoft for environmental data. Barrick will also serve data from the SDI to other applications. For example, for vehicle tracking applications, Web services from the SDI data provide context for vehicle movement.



*In the first step in publishing data with MXD Publisher, all datasets are listed and checked for valid and complete metadata.*



*SDE Publisher publishes the file geodatabase from the MXD Repository to the SDE Repository.*

