

A New Model for Address Point Management

Data interoperability improves customer service in Kentucky county

By Matthew DeMeritt, ESRI Writer



How does a GIS department convert reams of production data into a single data model that can be published and easily consumed by the public?

For Boone County, Kentucky, the solution came through using the ArcGIS Data Interoperability extension, which provides powerful spatial ETL (extract, transform, and load) capabilities. The use of this extension spared Boone County the nightmare of manually consolidating information from many different agencies into a single usable dataset that could be shared by all its clients.

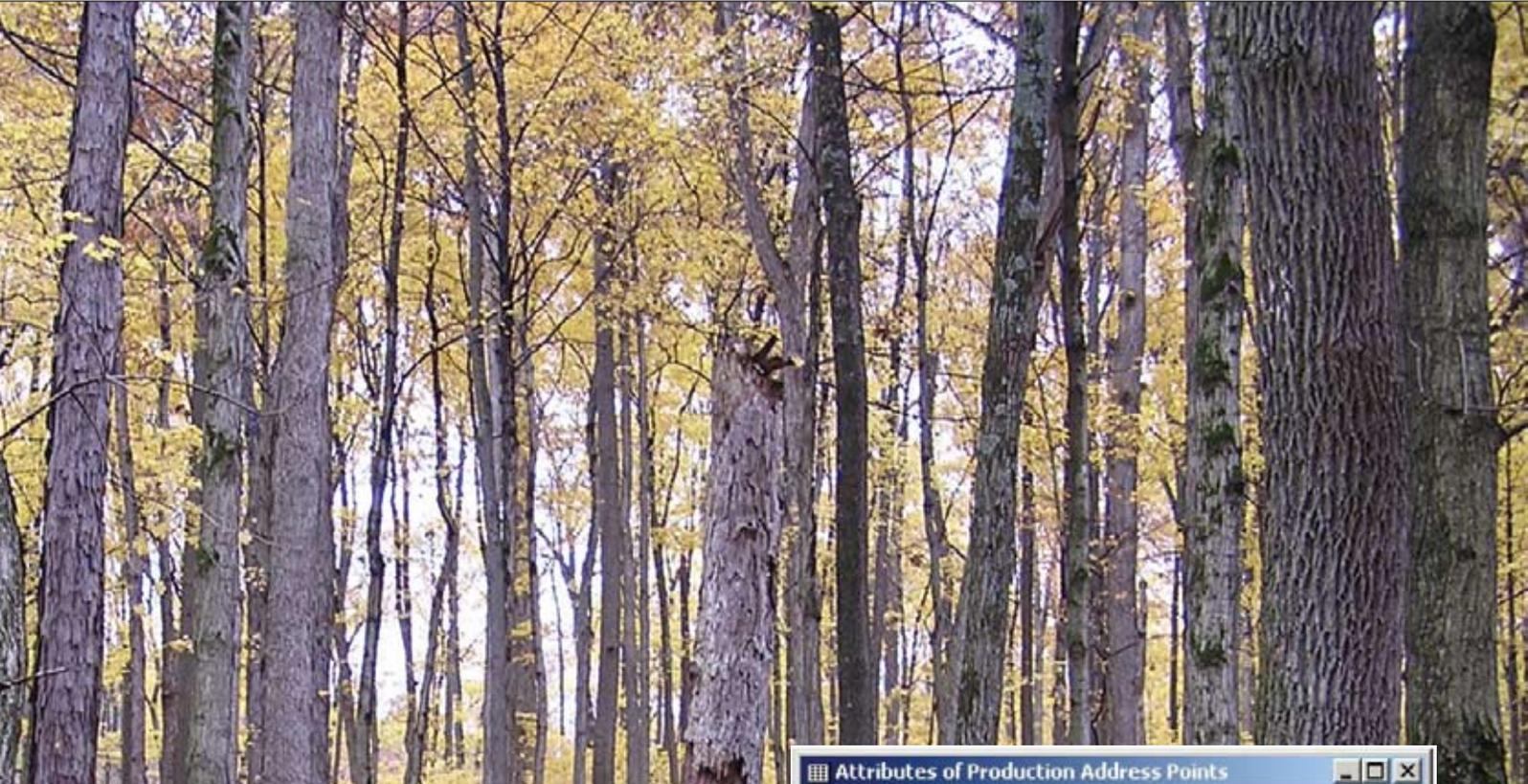
Production and Publication Data

Boone County GIS is a consortium of 30 agencies with more than 700 users in Boone County and northern Kentucky. Like most IT departments with volumes of data stored in disparate formats and data schemas, getting functional data into the hands of end users was an issue.

Boone County's data environment is divided between production and publication data. Production data is actively edited. Published data has been refined into information that is easily consumed by end users. Boone County's objective was to design a single production data model with a schema that was flexible enough to support the data transformation requirements necessary to pull many different data models—created using various schemas—together. The county's published data is available to the public via its ArcIMS and ArcGIS Server Internet mapping Web site (www.boonecountygis.com) and its own users via a custom built ArcReader/C# application called BooneMap.

SUMMARY

A Kentucky county consolidated information from many different agencies into a single usable dataset for its clients without custom coding and manual processes by using the ArcGIS Data Interoperability extension.



Consolidating Data for End Users

Spatial ETL tools can extract data from the source (production) data model, transform the attributes and geometry of the source data to match the destination (publication) data model, and load the transformed data. Many people think the ArcGIS Data Interoperability extension is limited to converting ESRI formats into other software formats and vice versa, but this is just one of many modifications that can be made during the conversion process. Since Boone County’s data was already being maintained in ESRI formats, conversion between formats was not part of the overall data migration process.

Instead, Boone County wanted a solution that would enable the easy manipulation of production data in a variety of ways to generate value-added published data for end users. “Essentially, we wanted to perform as much of the ‘heavy lifting’ on the back end as possible by preprocessing the data into something a bit more user friendly,” said Steve Gay, the GIS coordinator for Boone County. “Early on, we knew the different ways that we wanted to transform the data, but we thought that in order to achieve this, we would have to invest heavily in a lot of complicated programming. We were desperately seeking a solution that was easy to use and would ultimately enable us to republish our data at a moment’s notice.”

After seeing a presentation by ESRI’s Steve Gris  at the 2005 ESRI User Conference, Gay approached Gris  with his predicament. To his delight, Gris  explained that the ArcGIS Data Interoperability extension would be able to accommodate the data transfer needs of his department

Continued on page 18

Attributes of Production Address Points

SITENUM	SITEDIRPRE	SITESTNAME	SITESTYPE
6073	S	ZIG ZAG	RD
6034	S	ORIENT	ST
615	S	AIRFIELD	DR
6043	S	ZIG ZAG	RD
67	S	MAIN	ST

Record: 0 Show: All electe

Selected Attributes of Published Address Points

Site Address Number	Site Address *
67	67 S MAIN ST
6043	6043 S ZIG ZAG RD
615	615 S AIRFIELD DR
6034	6034 S ORIENT ST
6073	6073 S ZIG ZAG RD

Record: 0 Show: All electe

Address components in the production geodatabase can be concatenated in the published geodatabase. In this case, the address number is available as a separate field as well as concatenated with direction, street name, and street type data in a single field.

A New Model for Address Point Management

Continued from page 17

and manage the transformation of data from the production environment to the publishing environment with no complicated manual programming.

Boone County purchased the ArcGIS Data Interoperability extension and began using FME Workbench, a graphic authoring environment included with the extension that provides more than 200 transformers for manipulating geometric and attribute data. Because the visual interface was easy to use and required no code writing, Boone County staff quickly became proficient in building data transformation workflows. Data transformations authored using the Workbench application are saved as geoprocessing tools or models inside ArcToolbox.

Accommodating Many Schemas

Gay uses the term “atomic attribution” to refer to the data modeling strategy of storing attributes in their lowest possible components, then combining these values together when needed. “It’s much easier to combine separate values together than it is to try and parse a single value into its component parts,” said Gay. This strategy provides flexibility by allowing the data publisher the freedom to choose the field values that need to be concatenated. “Everyone today seems to have their own schema requirements. By employing atomic attribution, we have the capability to provide our data to all of our usual data recipients according to their data formatting requirements,” Gay added.

Boone County currently uses 243 domain tables to support 142 feature classes in its ArcSDE production geodatabase. Conversely, the ArcSDE publishing geodatabase serves 162 feature classes that use no domain tables. Various steps had to be taken during the transformation to accurately map the data sources in the production geodatabase to a single publication data model. The Data Interoperability extension, and specifically the Workbench application, was used to build reusable spatial ETL tools that accurately and efficiently generate publication datasets. These tools are used to work with domain tables, map field attributes, and perform spatial conflation.

Working with Domain Tables and Values

When Boone County publishes its data, any field that uses a domain table gets the domain table’s coded value converted to its corresponding descriptive value. ValueMapper, one of the transformers furnished by the Data Interoperability extension, accomplishes this task.

“Users have an easier time understanding the data when they don’t have to constantly stop and



By providing the address number with and without other address components, end users can easily choose how much detail they want to include in labels.

Zoning District	C-2
Water Provider	1
Sewer Provider	3
Stormwater Provider	3
Fire District	HFD
Fire Emerg. Service Zone	1020
Law Enforcement Provider	2
Police Beat	B41
Voting Precinct	A109
Polling Location	29
School District	1
Elementary School	GE
Middle School	CM
High School	CH

Zoning District	COMMERCIAL TWO
Water Provider	BOONE COUNTY WATER DISTRICT
Sewer Provider	SANITATION DISTRICT #1
Stormwater Provider	SANITATION DISTRICT #1
Fire District	HEBRON FIRE DISTRICT
Fire Emerg. Service Zone	1020
Law Enforcement Provider	BOONE COUNTY SHERIFF DEPARTMENT
Police Beat	BOONE 41
Voting Precinct	HEBRON 2
Polling Location	HEBRON FIREHOUSE #1
School District	BOONE COUNTY UNIFIED
Elementary School	GOODRIDGE ELEMENTARY
Middle School	CONNER MIDDLE
High School	CONNER HIGH

When Boone County publishes its data, fields that use a domain table have the domain table’s coded value converted to its corresponding descriptive value.

determine what a coded value represents,” said Gay. “We make it explicitly clear by providing them with the textual description rather than some obscure code that may or may not mean something to the end user.” In some cases, the domain table’s coded values and description are exported into two separate fields. “Even though most users prefer the spelled out description, there may be an additional need to preserve the coded value in the published data for labeling purposes. The Data Interoperability extension allows us to export whatever we want—the code, the description, or both if necessary,” Gay continued.

Attribute Field Mapping

A simple example of field mapping combines multiple fields from the source data into a single field in the destination geodatabase. Using several different transformers provided by the Data Interoperability extension, GIS staff at

Boone County can easily concatenate field values. Boone County publishes the address number by itself in its own field as well as in another field combined with all other address components in data published by its own users. “Doing this allows the end user to decide if they want to label a location with its full address or just the address number if they’re trying to make the map less busy,” explained Gay. Because the address number for both fields originates from the same value in the production data, there is never a risk that the two published fields will contain conflicting address numbers.

Spatial Conflation

Publication data can also be generated by populating new datasets with information derived from spatial relationships stored in the production data. Boone County had a somewhat unique need to transfer some address information from the Address point layer onto



Photographs courtesy of Steve Gay

two polygon layers—Buildings and Parcels. The Data Interoperability Point-On-Area Overlay transformer performs a spatial join between two layers. Since Boone County’s business rules dictated that (1) address points should be located on top of the associated building polygon feature and (2) all parcels must contain an address point, the Point-On-Area Overlay transformer could be used to enter address attributes from the production address points into published building and parcel polygons by spatially joining the layers.

The ArcGIS Data Interoperability extension does more than convert data between ESRI formats and other formats. It enhances the manipulation of data in the production environment, increases the quality of published information, and improves customer service.

The same Point-On-Area Overlay transformer was also used to compare every address point’s location to 19 different polygon layers. To help visualize this, Gay used this analogy, “Think of overlaying all of your large-scale administrative boundary polygon layers, such as fire district, voting precincts, and census

blocks, and then hammering a nail through them all at the same location.” This process transfers information relevant to each address from the intersected polygon features to the published address points. Boone County adds this spatially related information to its published address points to enable end users to focus much of their query efforts on the address points. “This methodology empowers our users by providing them with the ability to perform multicriteria attribute queries on one layer and generate an address list tailored to their needs,” added Gay.

This process also explains the discrepancy in the number of feature classes between Boone County’s production geodatabase (142 feature classes) and its publishing geodatabase (162 feature classes). The additional published feature classes contain features that do not have to be actively maintained (and therefore don’t exist in the production geodatabase). Instead, these features are derived by running geoprocessing operations on features from other feature classes. Boone County has automated this creation process with the Data Interoperability models. These models facilitate periodic updates of feature classes.

A perfect example of this are the polygon layers created from buffering certain features. “We publish several polygon layers that are essentially buffers of other point or polygon features,” said Gay. “As these source features change, the associated buffers may need to be updated as well.” Instead of maintaining these



buffer areas in their own layer, Boone County re-creates the layer’s geometry and attributes using Data Interoperability models. Gay points out that this not only frees up staff time but also provides an easily repeatable and documentable method of creating layers.

Conclusion

As Boone County demonstrates, the ArcGIS Data Interoperability extension is more than a way to convert data between ESRI formats and other formats. It greatly facilitates the manipulation of data in the production environment, increases the quality of published information, and improves customer service. In the process, Boone County set a precedent by creating a new model for managing address points—one that speeds the production-to-publication process without complicated programming or additional resources.