

# 2010 NSDI CAP Interim Report

Category 2: Framework Data Exchange through Automated Geo-Synchronization

## Coordinating Local, State and National Data Stores

### Framework Data Exchange through Automated Geo-Synchronization

Submitted By:



Presented To:



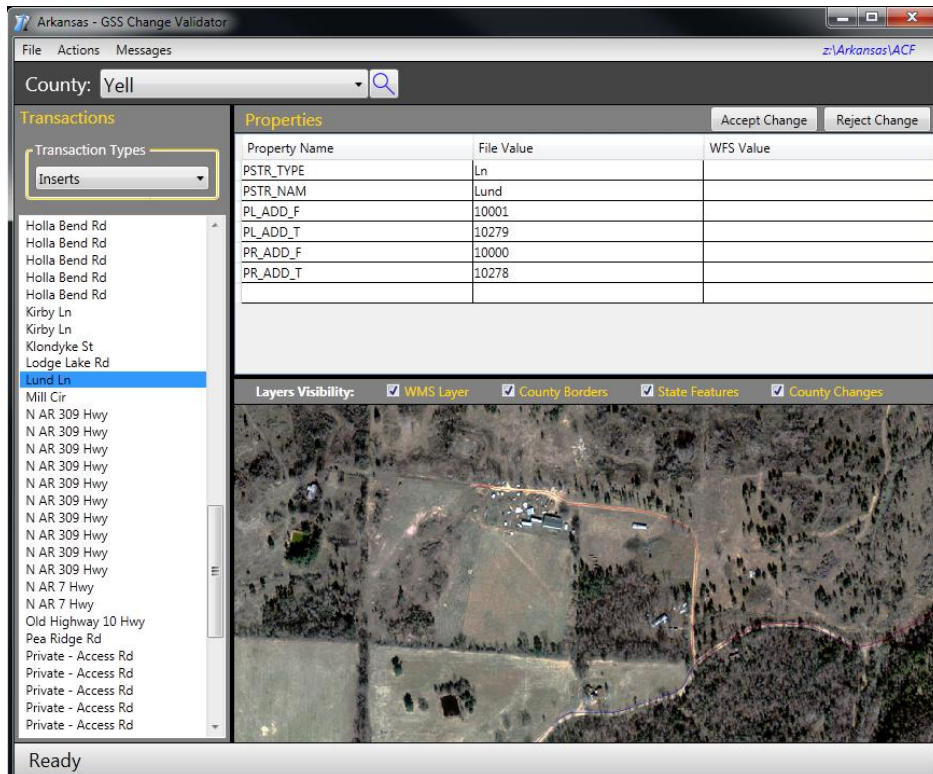
October 22, 2010

Cooperative Agreement Number G10AC00237

## Project Summary

The Carbon Project and the Arkansas Geographic Information Office (AGIO) are pleased to submit this 2010 NSDI Cooperative Agreement Program (CAP) Category 2 Interim Report for the Category 2, “Framework Data Exchange through Automated Geo-Synchronization” effort. Since kickoff in May 2010 the project has made significant progress in developing, deploying and sustaining a federation for exchanging local, state and national framework data using geo-synchronization services and common data models. In particular, the project has developed, tested and started to deploy a unique model for the NSDI that provides methods to ingest county transportation data into the geo-synchronization federation and synchronize state-level databases to federal-level databases. To achieve this first step in a manner not disruptive to current processes, we are now leveraging common folder structures consisting of FTP updates from 73 Arkansas counties. We have also developed polling/transformation capabilities to extract data changes and transmit them to a GeoSynchronization Service (GSS) using ATOM feeds, a new application called the GSS Change Validator (Figure 1) and a prototype Web Feature Server Transactional (WFS-T) on ESRI ArcGIS Server deployed at The Carbon Project site. With this capability our approach provides a bridge between current processes and a new generation of NSDI GeoSynchronization. It will also allow the state to advance data stewardship, and federal systems to automatically keep data sets up-to-date. The solution will also be re-deployable to other sites through open APIs and free tools.

*Progress to date on this project begins to bridge the gap between current file-based production operations and the FGDC vision of developing, deploying and sustaining a federation for exchanging local, state and national framework data using GeoSynchronization services and common data models.*



**Figure 1 – Since the start of this CAP effort we have developed a new capability, called the Arkansas GSS Change Validator, to transform county-based road updates from Shapefiles to geosynchronizable transactions suitable for a federated National Spatial Data Infrastructure (NSDI)**

## Lead Organizations

### **Applicant Organization Information (Submitted on Behalf of Government Partner):**

The Carbon Project / Arkansas Geographic Information Office (AGIO)  
15 New England Executive Park, Burlington, MA 01803  
Internet Address: <http://www.thecarbonproject.com>

### **Government Partner:**

Arkansas Geographic Information Office (AGIO)  
1 Capitol Mall, 2B 900 | Little Rock, AR 72201  
Internet Address: [http://www.gis.state.ar.us/AGIO\\_index.htm](http://www.gis.state.ar.us/AGIO_index.htm)

### **Principal Investigator:**

Learon Dalby, GIS Program Manager  
Phone: 501-682-2929, Email: [learon.dalby@arkansas.gov](mailto:learon.dalby@arkansas.gov)

### **Technical Lead:**

Mark Mattson, Senior Software Engineer, The Carbon Project  
Phone: 781-270-0674, Email: [mmattson@thecarbonproject.com](mailto:mmattson@thecarbonproject.com)

### **Community Coordination Lead:**

Jeff Harrison, CEO, The Carbon Project  
Phone: 703-628-8655, Email: [jharrison@thecarbonproject.com](mailto:jharrison@thecarbonproject.com)

**Applicable Framework Themes:** Transportation (Road Centerline Data) and other themes

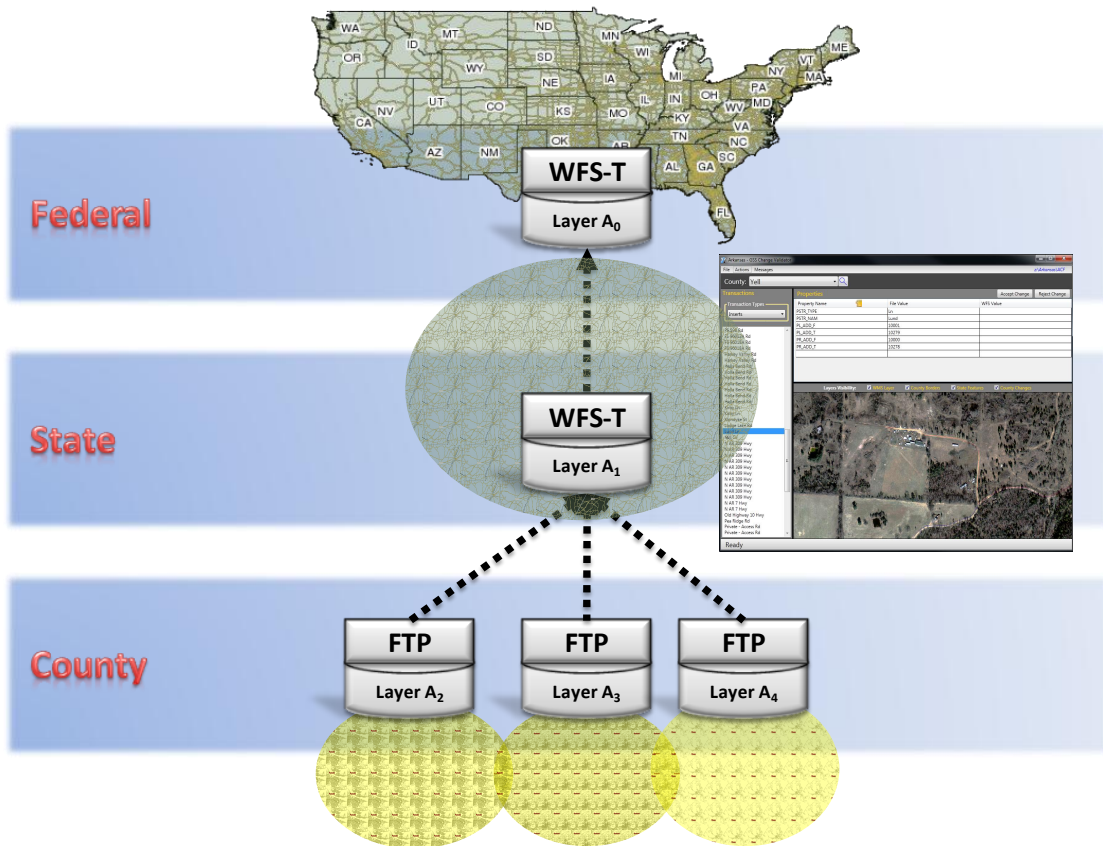
**Geographic Scope or Area:** State of Arkansas and Nationwide

**Previous CAP Participation:** AGIO successfully completed a CAP 2004 project to establish a coordination mechanism for the state's County Assessor Mapping Program. AGIO is also executing a CAP 2009 project to develop a Sustainable Funding Plan for Framework Data. The Carbon Project successfully completed 2006/2007 CAP projects, resulting in the popular, free framework data application, Gaia. The Carbon Project also successfully completed a CAP 2008 project, deploying joint Canadian-US SDI and a 2009 CAP project developing Dashboards for Geodata.gov.

## Project Background

The goal of this project is to bridge the gap between the current needs of state data production organizations and the geo-synchronization vision outlined in the 2010 CAP grant. The result will be a practical capability for exchanging transportation framework data between local, state and national data stores through a geo-synchronization service using common data models and services. To achieve this objective the Arkansas AGIO assumed the leadership role for this project and is working with The Carbon Project to deploy a geo-synchronization capability. Deployment of the Arkansas county-to-state geo-synchronization system will be in the form of a simple one-directional tree where the state's data is synchronized against all county layers, thus any operation committed against the county layers will be propagated to the state level (Figure 2). Federal and other interested services can register to be synchronized with the updates. The capability uses geo-synchronization system from The Carbon Project (CarbonCloud Sync) and ESRI's ArcGIS Server. This effort provides a unique capability for NSDI geo-synchronization by:

- 1) Addressing methods to ingest County data into the geo-synchronization federation and;
- 2) Geo-synchronizing a state-level database to a federal-level database.



**Figure 2 – This project deploys and sustains geo-synchronization capability in Arkansas, enabling exchange of framework data between local, state and national data stores through geo-synchronization.**

To achieve the first step we have established a common folder structure consisting of regularly contributed GIS updates via FTP from over 70 Arkansas counties (see Appendix A). A polling/transformation capability is now being used to bring this data (via ATOM feeds) into the geo-synchronization process as outlined by FGDC and OGC. Our project then leverages the Web Feature Server Transactional (WFS-T) capacity of ArcGIS Server 9.3 to promote collaboration at local, state and federal levels. With this capability we will bridge current Arkansas business processes with the “to be” processes outlined by FGDC. Without this recognition of current processes we believe geo-synchronization may not successfully address real-world operations. The solution will be applicable (re-deployable) to other sites and problem domains through an open API, cloud-based services, and free tools so that it may be modified and re-deployed by other Framework data exchange users and providers.

This approach is required because currently the Arkansas state-level GIS receives a drop of Shapefiles features data from counties on approximately a monthly basis. These files are ingested into the state level view, including all provided data changes and updates. Little or no resources are available to deploy WFS-T in each of the 70 counties. The enhanced system can accommodate current process and offer the option for future introduction of WFS-T to counties with minimal disruptions. Furthermore, this system will enable the

Arkansas state-level to be updated with new transactional sources including citizen (or professional) provided updates (Gov 2.0).

This project builds on the initial geo-synchronization capability developed by The Carbon Project during the Canadian Geospatial Data Infrastructure (CGDI) Interoperability Pilot in 2007, as refined for the US Army last year (Figure 3) to provide collaborative geospatial data maintenance and update. To provide a better understanding of the potential of GeoSynchronization the reader is directed to Appendix B and a YouTube video describing the use of GeoSynchronization.<sup>1</sup>

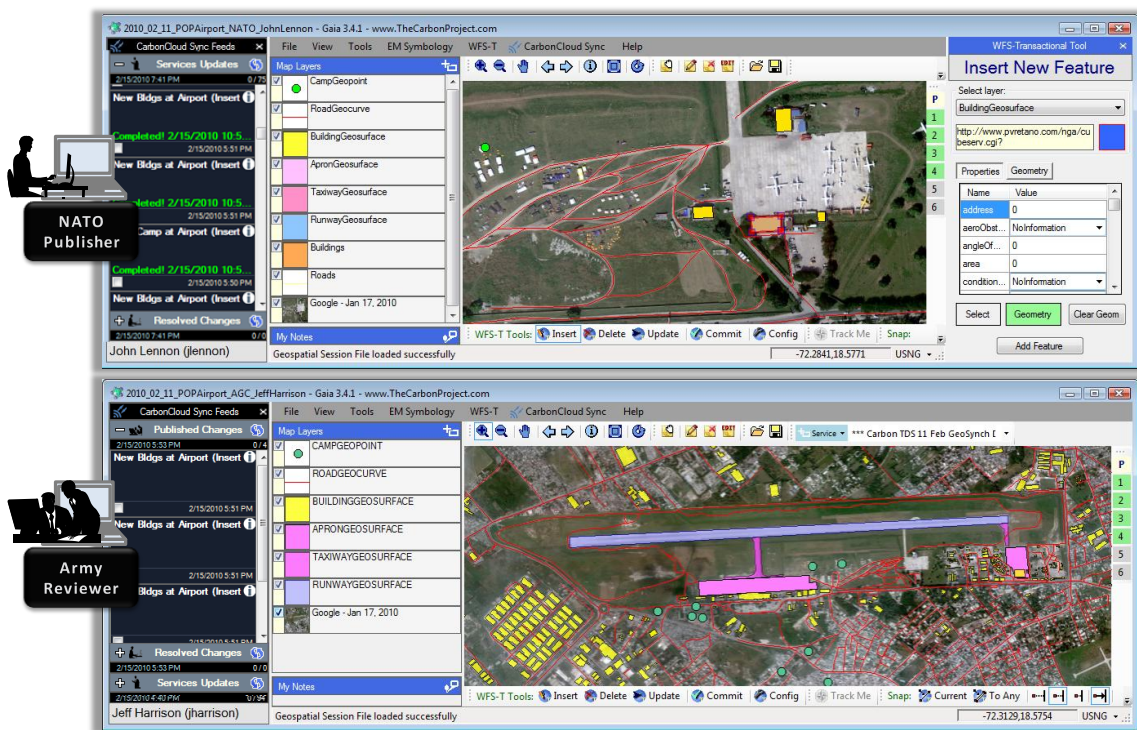


Figure 3 - The Carbon Project's Geo-Synchronization capability was initially developed in the CGDI IP, and then enhanced in 2009 and 2010 with Gaia Extensions and CarbonCloud Sync for the US Army, shown above.

The project is being conducted as a series of collaborative tasks to develop, deploy and sustain the Geo-Synchronization capability in Arkansas *and* engage government stakeholders:

- **Development, Deployment and Sustainment** – These tasks develop, deploy and sustain services and applications in Arkansas to synchronize transportation framework data updates into authoritative State data layers, which are then provided for national-level integration. The initial effort of this project focused on these tasks.
- **Community Engagement** – These tasks define and execute communication and outreach activities to Arkansas/NSDI stakeholders. The next phase of the project will focus on these tasks.

<sup>1</sup> <http://www.youtube.com/thecarbonproject#p/a/u/0/3p4hCHpAFKc>

- **Project Coordination** – This task enables coordination between project staff, participants and the NSDI community. Activities include management, documentation, meetings and maintaining online collaborative environments.

Our approach allows Arkansas to implement geo-synchronization with minimal change to existing technologies and planned development.

## Technical Approach and Progress

Our technical approach is based on the Open Geospatial Consortium (OGC) draft for GeoSynchronization Services (GSS) and the OGC Web Feature Service (WFS). The GSS deployed in this project uses The Carbon Project's CarbonCloud Sync<sup>2</sup> software. This solution will be deployed in Little Rock, Arkansas and designated the Arkansas GeoSynchronization Service (ArGSS). ArGSS will be a powerful capability agnostic to the underlying information infrastructure; it simply interfaces with WFS-T services or FTP common file folder, and includes the ability to analyze transportation Shapefiles data and transform changes into WFS transactions using Geography Markup Language (GML) consistent with Framework Data Content Standards. This approach enables our solution to support current and future NSDI production operations simultaneously.

ArGSS implements OGC GSS and WFS-T and uses the Atom Publishing Protocol to:

- ✓ advertise geospatial database changes
- ✓ create/edit GeorSS feeds for data currency/replication
- ✓ alert subscribers of proposed changes, resolutions, and replication transactions available
- ✓ perform queries on feeds

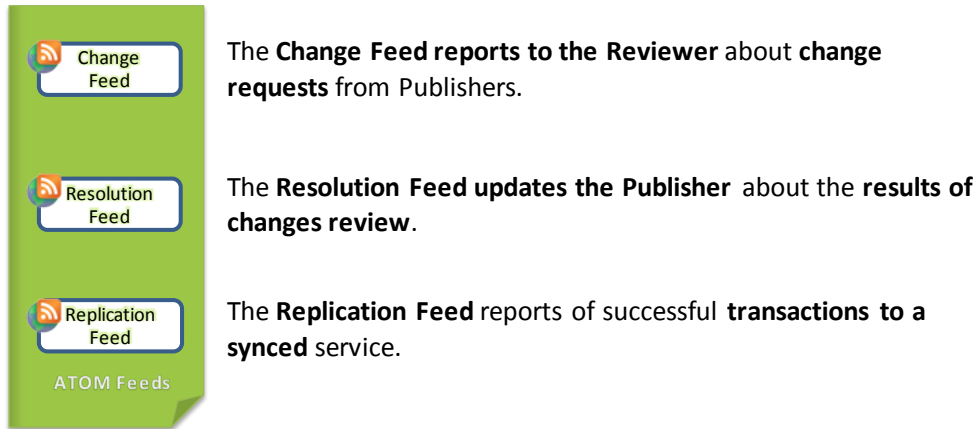
In addition, ArGSS will include capabilities to bridge the gap between the current needs of state data production organizations and the geo-synchronization vision outlined by the FGDC, including:

- ✓ a common folder structure consisting of regular FTP updates from over 70 counties
- ✓ a polling/transformation capability to bring county data input (via ATOM feeds) into the GGS and WFS-T process as outlined by FGDC

The ArGSS, like any GSS, uses three *feeds* (Figure 4), managed by a web service to support data currency and replication in a federated environment.

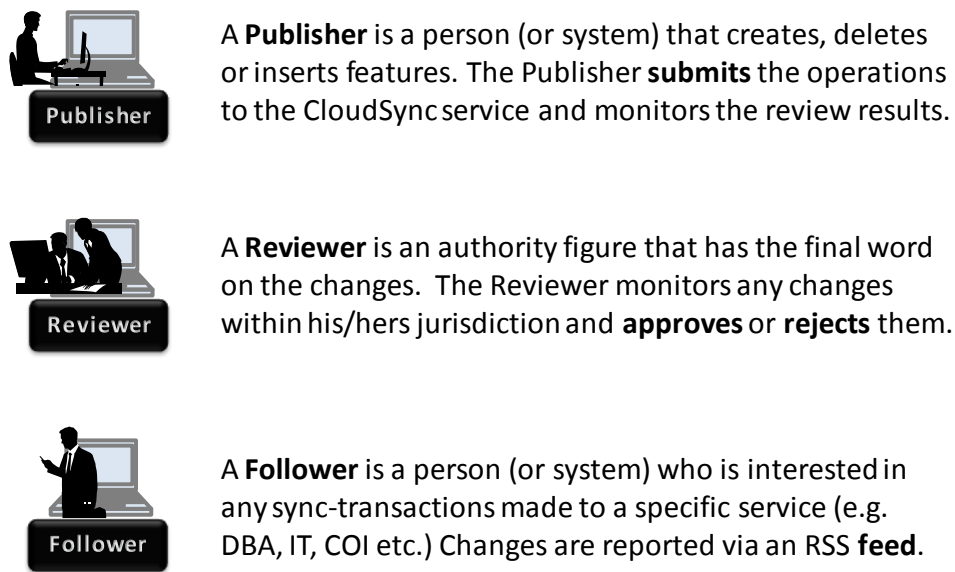
---

<sup>2</sup> <http://www.thecarbonproject.com/Products/cloudsync.php>



**Figure 4 – A GeoSynchronization Service uses 3 types of feeds, managed by a web service**

In the Arkansas operational deployment, three major roles will be supported by the GSS for transportation framework data (Figure 5).

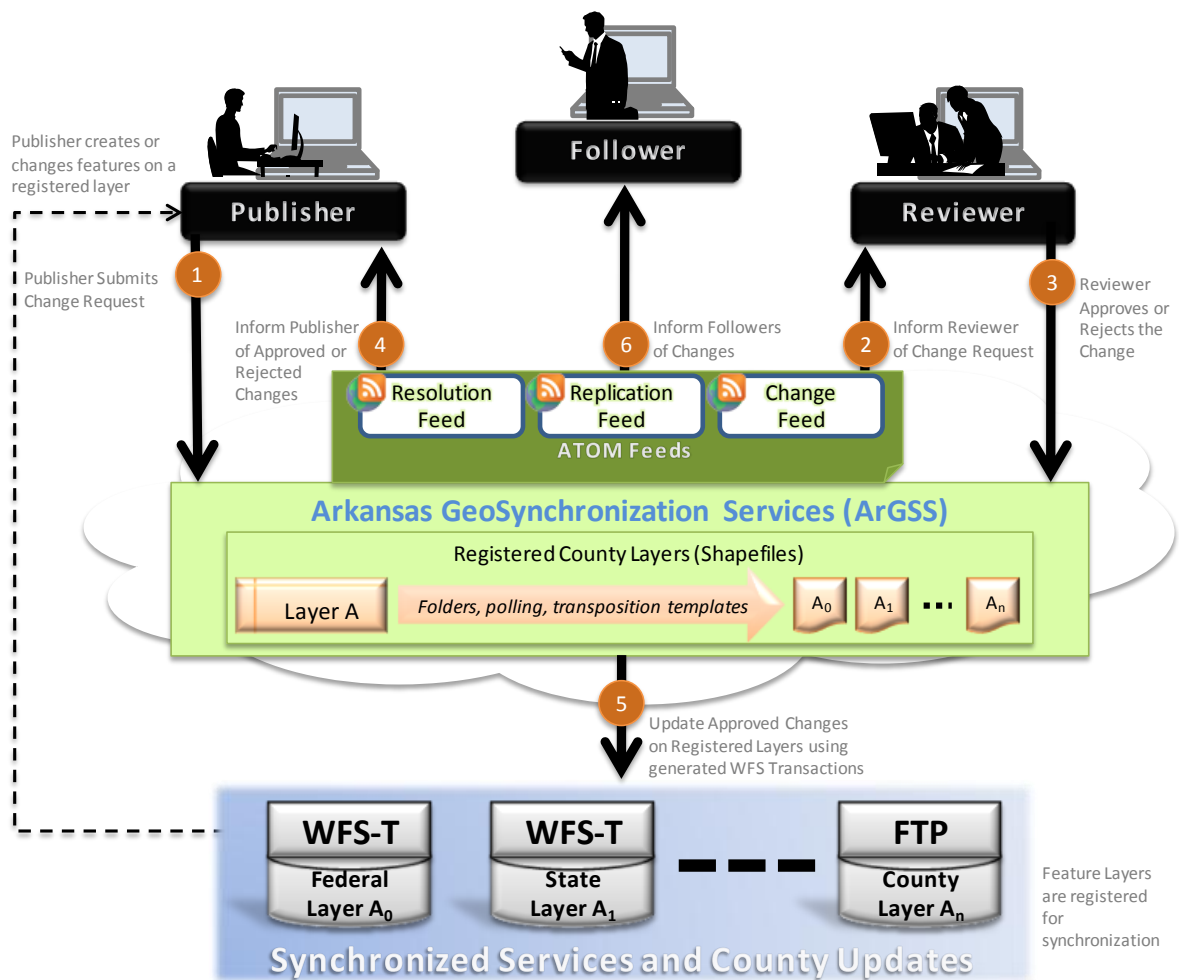


**Figure 5 – Major roles in the ArGSS federation**

Currently Arkansas state GIS receives a drop of Shapefiles based features data from over 70 counties on approximately a monthly basis. These files are ingested into the state-level view, including all provided data changes and updates. Little or no resources are available to deploy WFS-T in each of the 70 counties.

To support this reality we have extended the CarbonCloud Sync services and clients to accommodate current systems and processes with minimal disruptions, while maintaining the powerful synchronization capabilities and benefits the system has to offer to interoperable federated systems. The enhanced system delivered for ArGSS will accommodate the current process and offer the option for future introduction of WFS-T and other technologies to counties with no disruptions. It accomplishes this by reading county Road Centerline Shapefiles in the form currently used (see Appendix A) and then applying the process described below.

There are two key components to the solution. The first component is the CarbonCloud Sync service that disseminates changes throughout a federated deployment of geospatial services, keeping any registered remote services up to date with authorized updates (Figure 6). The second component is a custom standalone application called the GSS Change Validator. The GSS Change Validator application ingests and analyzes the registered county layers (Shapefiles) and manufacture change-based transactions on county-level transportation framework data layers. This two-part approach meets the needs of state data production organizations and the geo-synchronization vision of the CAP grant.

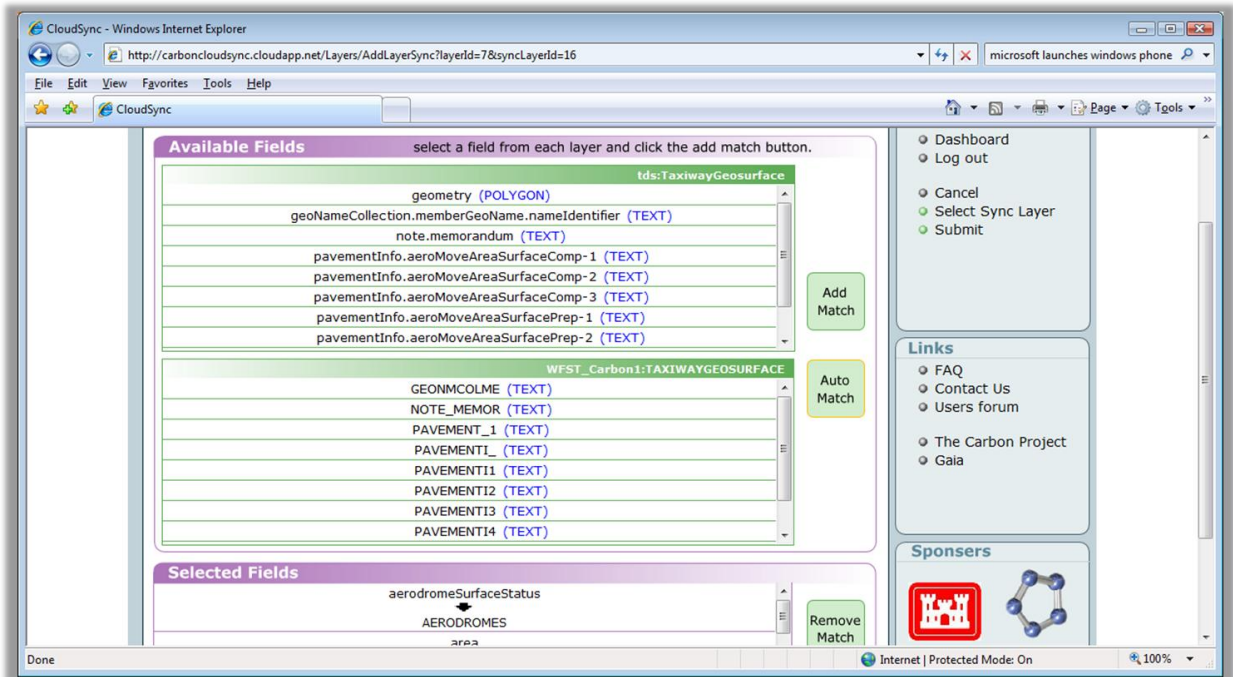


**Figure 6 – The Arkansas Geo-Synchronization capability facilitates maintenance and update of NHD into authoritative State data layers, which can be provided for integration into the NSDI.**

This solution required an initial setup. During this one-time process we created a layer template for the state data (a schema-based description of the data). We will also generate county-level layer templates for each participating county. Once the layers are registered in ArGSS we will register the synchronization links by mapping counties to the state layers, and the federal WFS-T to the state’s WFS-T. This process is performed using the CarbonCloud Sync’s Web-based management client (Figure 7). The process is similar to any



CarbonCloud Sync deployment with the exception of using static Shapefile fields to produce the layer's data templates instead of WFS-provided schema sources.



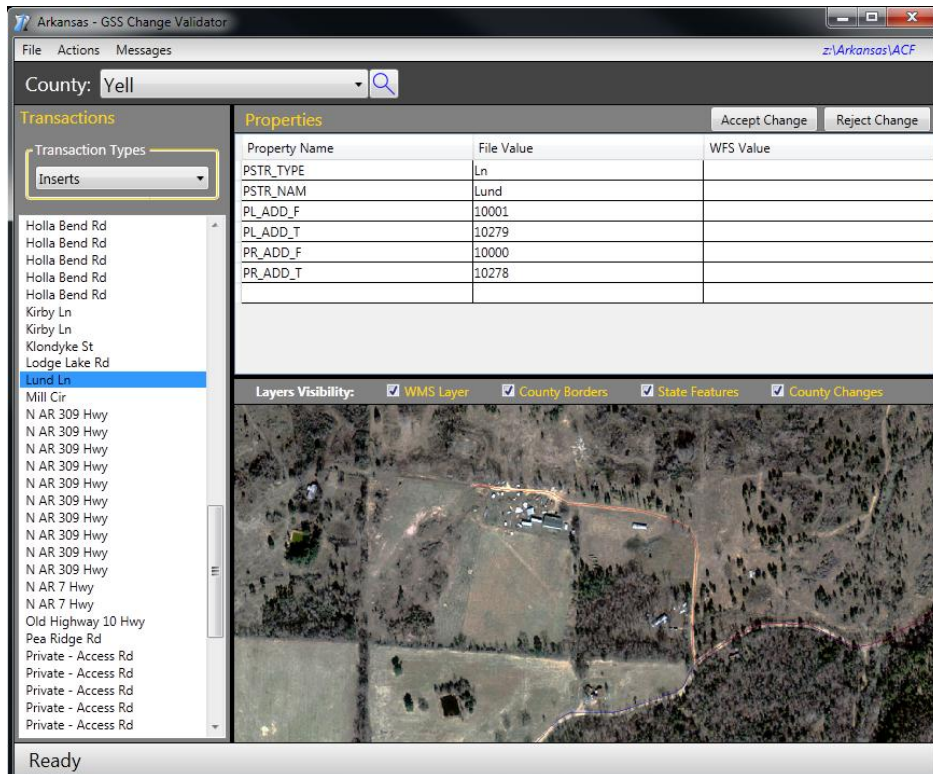
**Figure 7 – ArGSS can register synchronization links by mapping county-to-state layers, and the federal WFS-T to the state's WFS-T – a process performed using Web-based management clients.**

Deployment of the Arkansas county-to-state system is in the form of a simple one-directional tree where the state level is synchronized against all county layers, thus any operation committed against the county layers will be propagated to the state level, using the appropriate data transpositions (traversing from one schema to another). USGS and other interested and appropriate WFS-T can register with ArGSS to be synchronized with the updates (see Figure 2 for reference).

### GSS Change Validator

The GSS Change Validator is a new standalone application for the NSDI community (Figure 8). The application is presented to the user as a simple, form-driven tool that digests and analyzes county level data and automatically determines changes. Its primary purpose is to take data compiled by each of Arkansas' counties, compare that data against existing statewide data, determine what has changed and make the changes by submitting them to a GeoSynchronization Service (GSS) called CarbonCloud Sync. The application is also able to access any to use any OGC compliant data source such as Web Map Services (WMS) and Web Feature Services (WFS) endorsed by the US [Federal Geographic Data Committee](http://www.fgdc.gov/) (FGDC) Steering Committee (along with other standards developed externally to FGDC).<sup>3</sup>

<sup>3</sup> <http://carboncloud.blogspot.com/2010/10/ogc-standards-officially-endorsed-by-us.html>



**Figure 8 - The GSS Change Validator is a new standalone application for the NSDI community. The application takes data compiled by each of Arkansas’ counties, compares that data against existing statewide data, determines what has changed and makes the changes by submitting them to a GeoSynchronization Service (GSS) called CarbonCloud Sync.**

### Data Importing/Schema Transposition Process

County data comes in a variety of different information models. Our challenge was to convert the county schemas into the standard ACF schema that is used at the state level. The majority of the fields mapped one to one from county to state. These relationships are expressed in an XML file that can be updated at any time to reflect any changes in schemas. However, certain fields contained a complex arrangement of data. In those cases, it was required that we move the processing into the code level and mark that field with a generic FUNCTION attribute in the XML. Here is an example of data that required this type of processing:



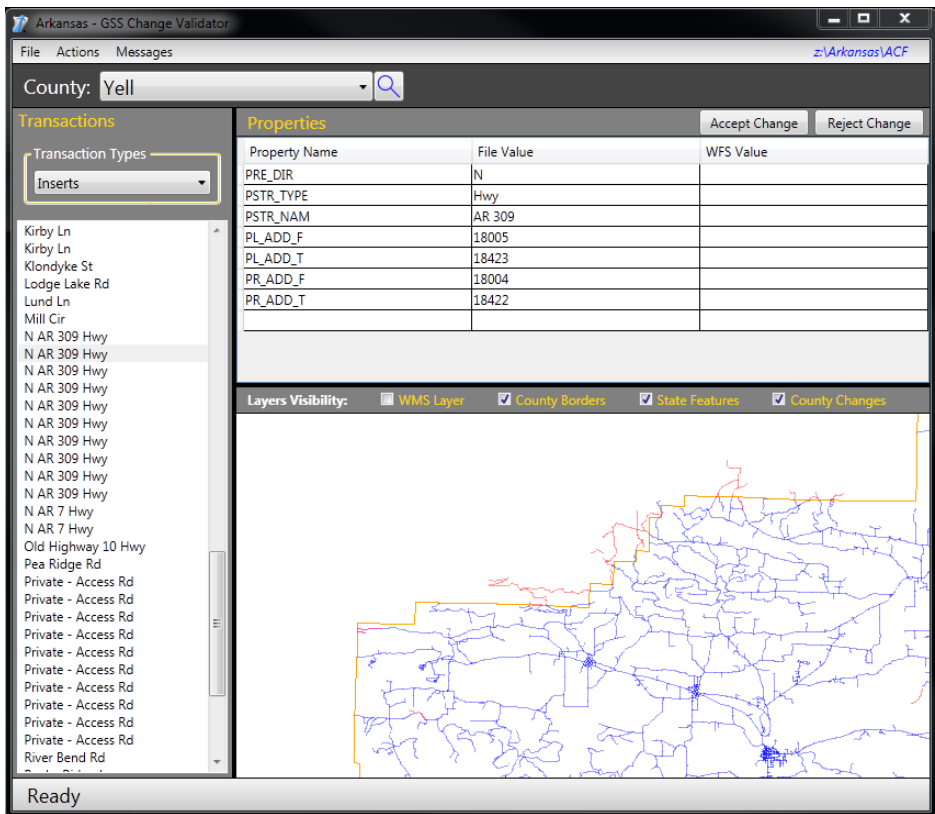
Our current processing algorithms contain logic for prefix/suffix directions, all possible PSTR\_TYPE values (Rd, St, Dr, etc.), some specialize highway rules and several other elements. After the schema is transposed into the proper format, the data is validated against the ACF rules and is modified (if necessary) to comply. Things like case and max length are checked here.

### WFS Data Retrieval

When a county is selected for processing, we need to compare it against the existing WFS features. To do that, we must first query the server for all features from a particular county. We do this by sending a WFS Filter Encoding (FE) query to the server. However, because the county geometries are often quite complex, it's often not possible to send it as a Filter Encoding. So to augment the process we send the Bounding Box of the extent of the entire county to the server. The resulting data set is then checked to see which features actually belong with the county and those that do not are removed. The result is all the features that are contained or intersect with the current county geometry.

### Comparison

Once we have the county data in the proper format and the WFS data from the server, we can initiate our comparison algorithms. We compare both the geometry and the properties from both the county and the state files. The property comparison is done based on the fields in the county file. The geometry comparison is done with a small variance built in (usually under a meter and you can adjust this variance in the settings if you need to.) The geometry comparison also takes into account the line direction, so it will match the geometry even if the lines are drawn in the opposite direction. Any features that are matched are immediately removed from both data sets. If either the geometry or the properties completely match, then we mark this as an update and remove it from both data sets. Features that remain on the county side are considered inserts. Features left on the WFS side are considered deletes. If an update has occurred which changes both features and geometry, the system will be unable to match it with the state data and will result in both an insert and a delete. The resulting data (Inserts/Updates/Deletes) are presented to the user for inspection. It's up to the user to determine if the changes are valid and if they should be acted on.



**Figure 9 – The GSS Change Validator with inserts are shown on the map in red**

In the preceding image, the inserts are being shown on the map in red. The current WFS features are displayed in blue. The user must select a feature in the left list, determine if it is valid and then click the accept change button to indicate that it's a valid change. We build in the ability to export changed as GML in the event it might be easier to use that change data.

## Submitting Changes

After a feature change has been accepted, it must be sent to the GSS for processing. To do that, users simply select the *Actions->Submit Accepted Changes* menu item and the changes will be sent to the GSS service. The transactions will show up in a GSS aware client like Gaia for approval.

## Use of WMS and WFS for Reference

Because we are using CarbonTools PRO, a geospatial interoperability toolkit, behind the scenes, we are able to use any OGC compliant data source.<sup>4</sup> In this case, we have an Arkansas GeoSTOR 1 meter imagery WMS that we are overlaying the change on top of and Arkansas GeoSTOR WFS with Road centerline data for reference.

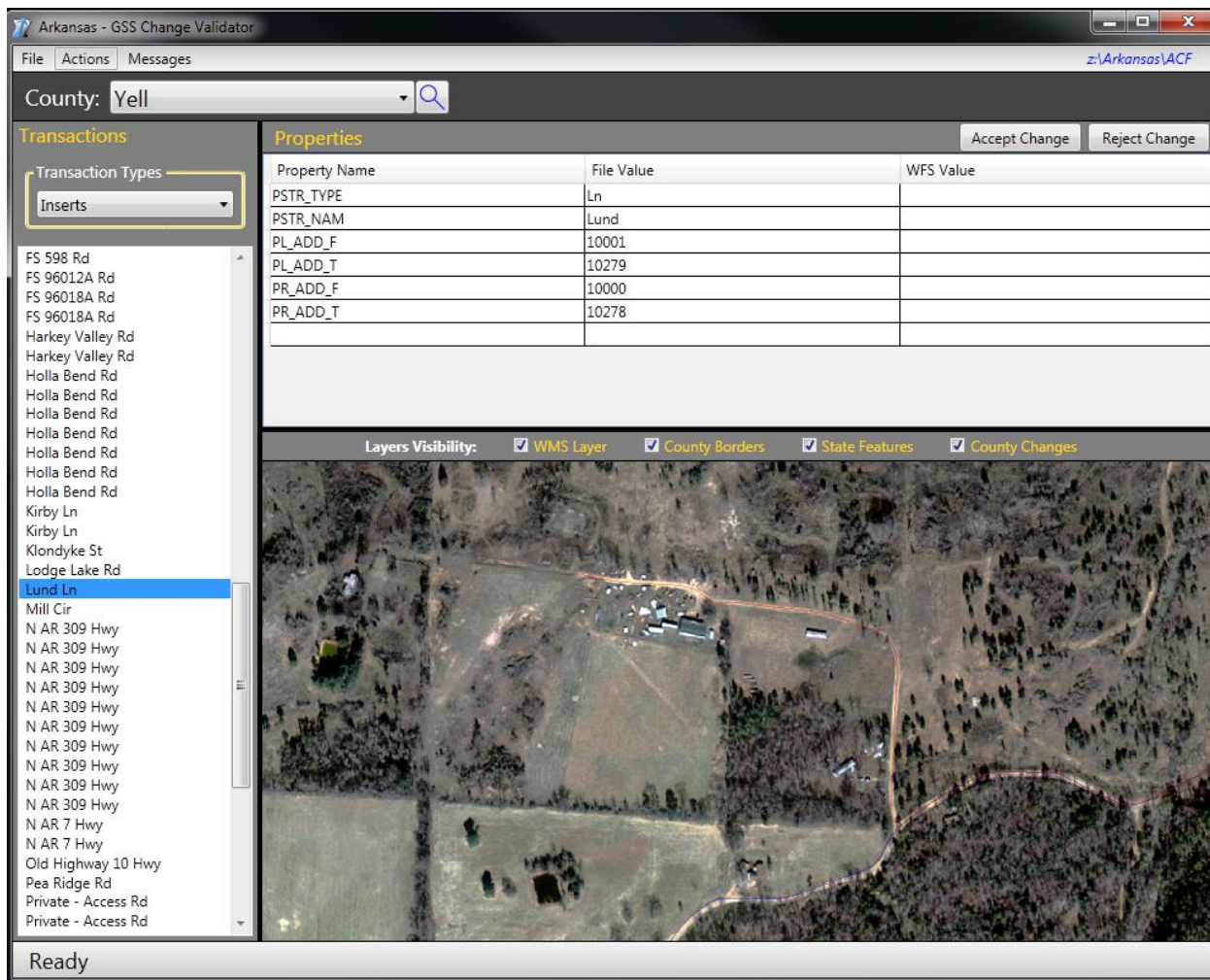


Figure 10 – The GSS Change Validator is able to access Arkansas GeoSTOR WMS and WFS for reference

This process bridges the gap of the current system’s handling of transactions originated as static files and WFS-T operations in automated geo-synchronization. The system uses advanced processing and comparison

<sup>4</sup> [www.CarbonTools.com](http://www.CarbonTools.com)

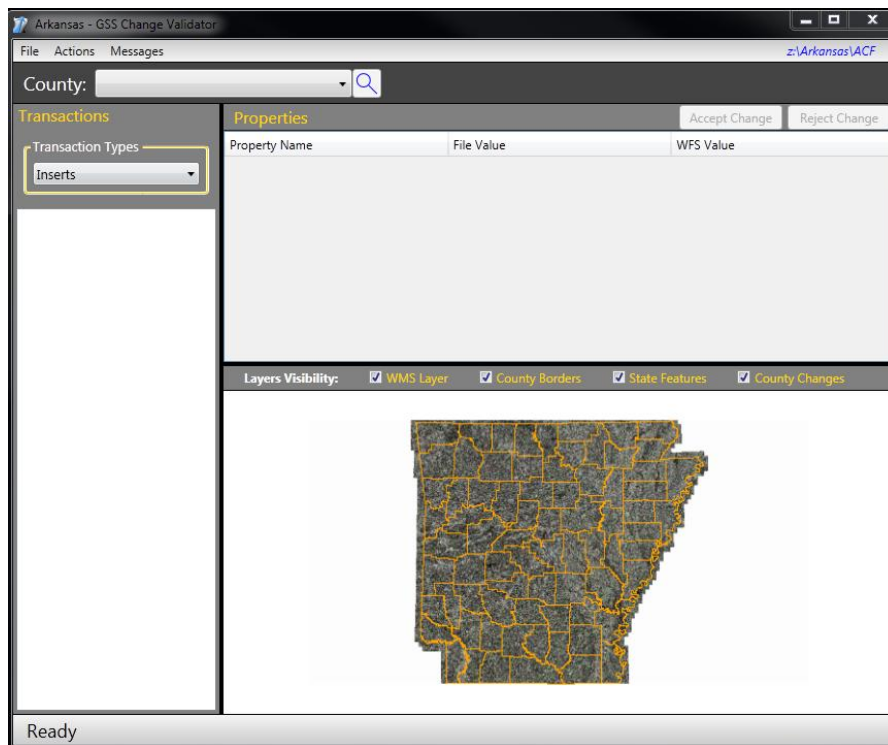
techniques to generate the WFS-T like operations.<sup>5</sup> From there, the normal CarbonCloud Sync processing will take care of propagating the changes down the deployment chain. This process allows the state's current Shapefile and FTP driven exchange process to evolve to a geo-synchronized federation.

Finally, in the next stages of this project the system will deploy an ArcGIS Server WFS-T at the state-level that will enable the Arkansas transportation framework data to be updated with government or citizen-provided changes. This system will use a suite of Gaia WFS-T<sup>6</sup> and geo-synchronization Extenders for Gov 2.0 operations (Figure 3 and Appendix A).

## Sample Use Case

To help understand the process of using the GSS Change Validator a sample operational Use Case is provided below.

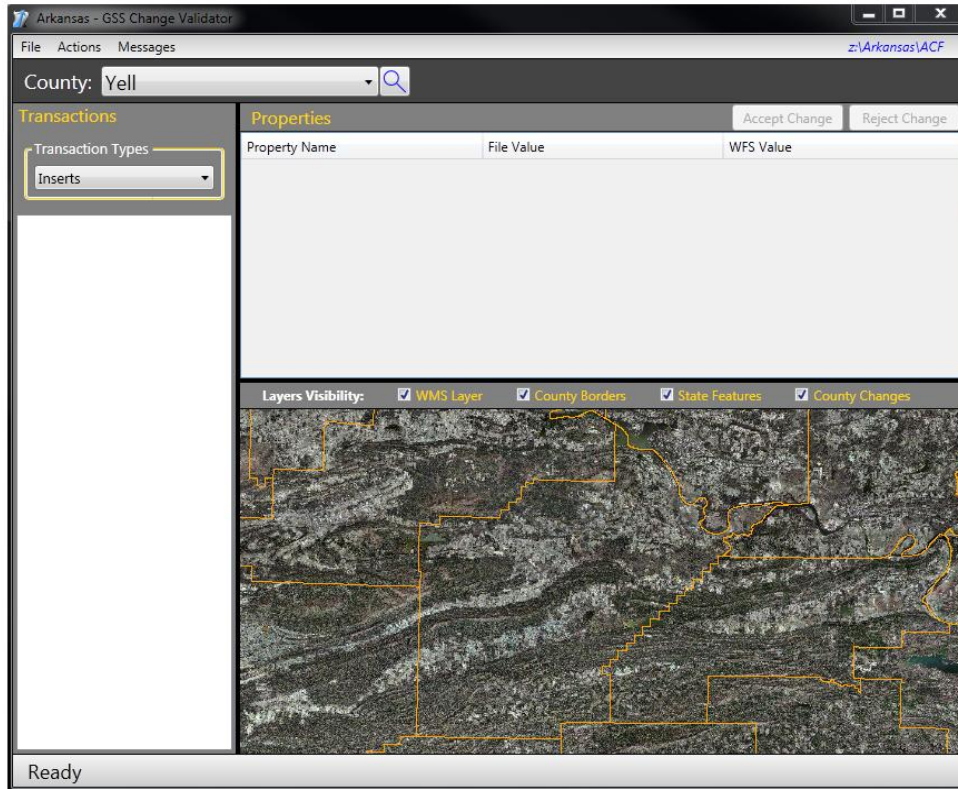
### *Step 1 – Open the GSS Change Validator Application*



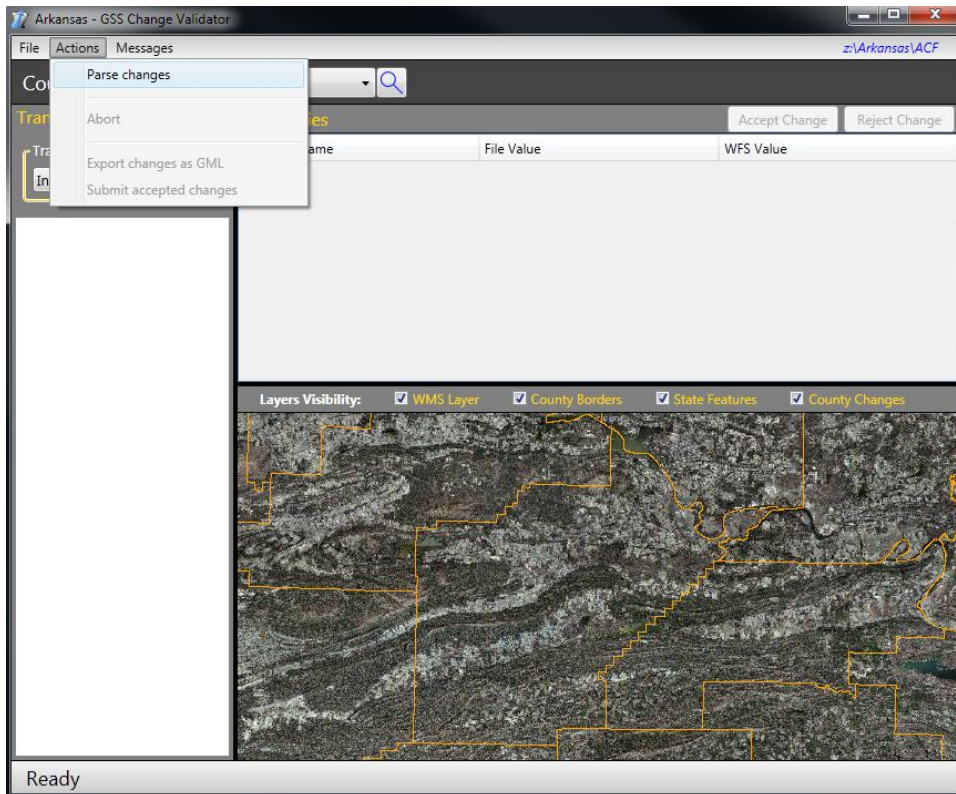
### *Step 2 – Select a county*

<sup>5</sup> Made possible by deploying CarbonTools PRO ([www.carbontools.com](http://www.carbontools.com)) on the server

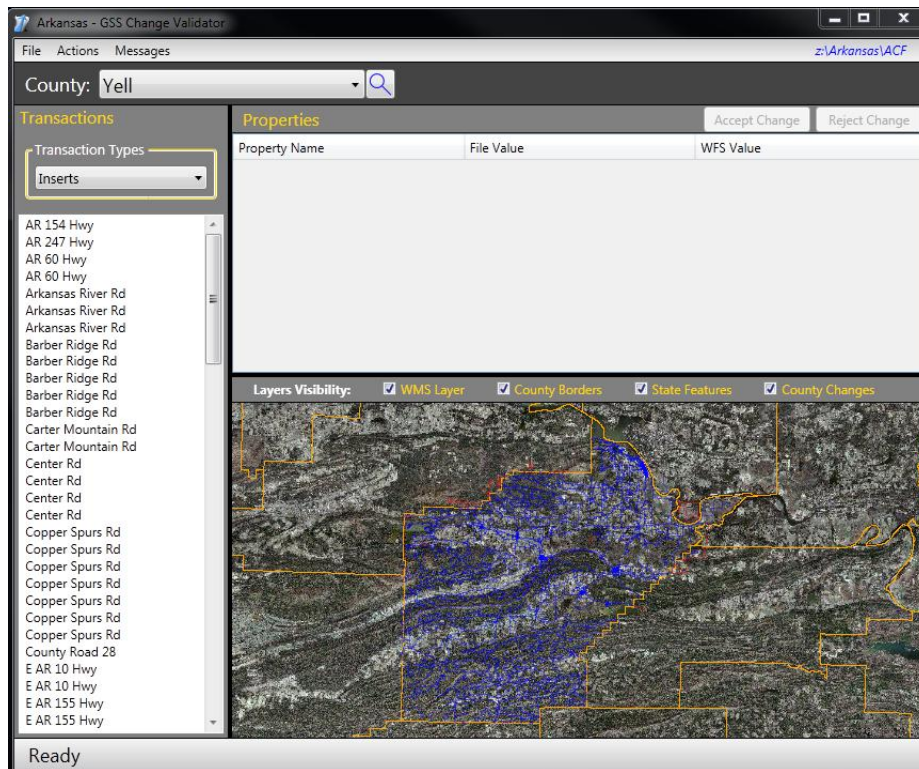
<sup>6</sup> [http://www.thecarbonproject.com/Products/gaia\\_wfst.php](http://www.thecarbonproject.com/Products/gaia_wfst.php)



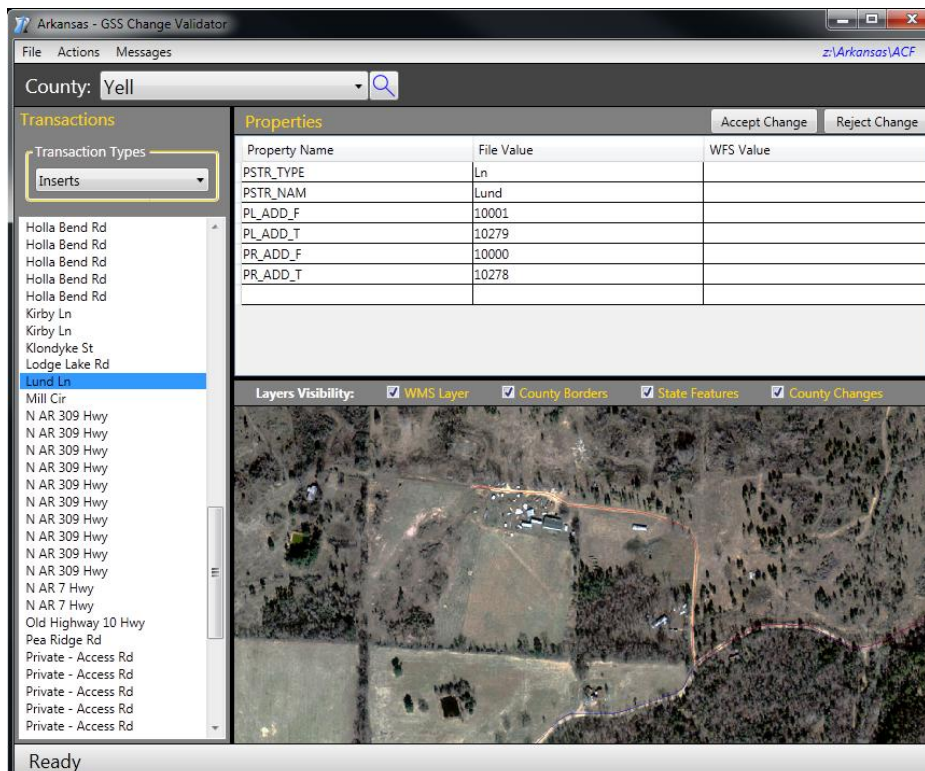
**Step 3 – Parse changes**



The results of the parsing are shown in the left panel. The dropdown selection box determines which type of changes are being shown (inserts/updates or deletes).



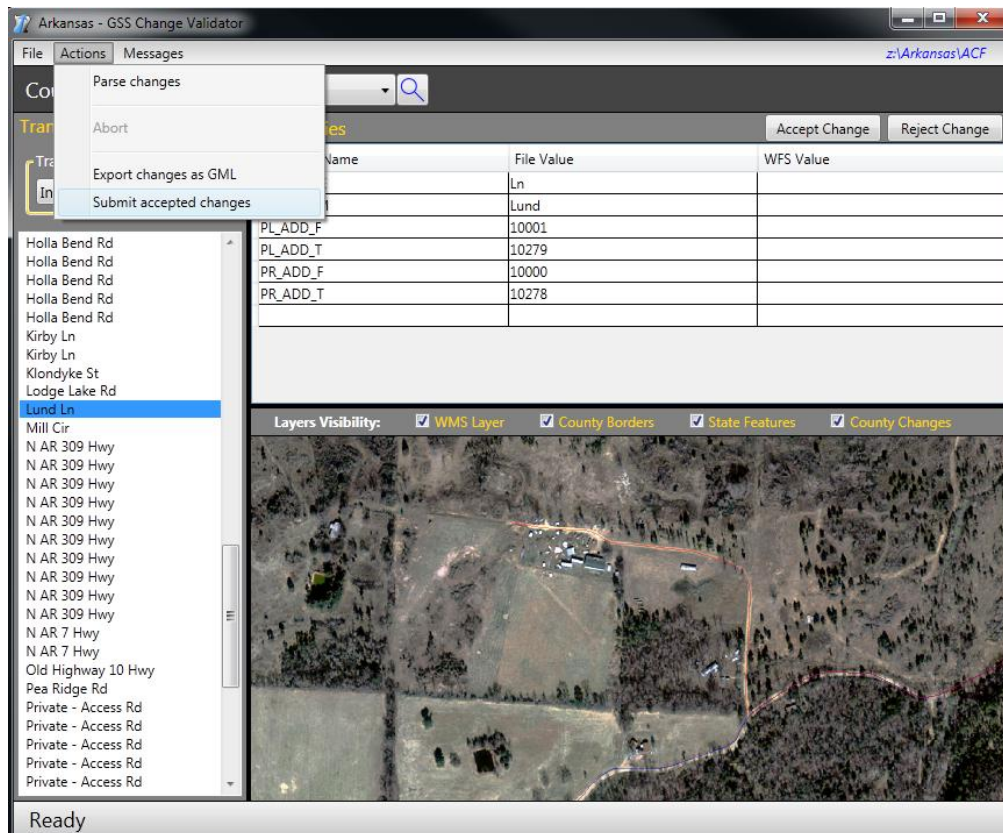
**Step 4 – Select a change.** The properties are shown in the properties panel and the map zooms to the selected feature.



The user has the ability to accept/reject the change. The user can also change or add properties to the property grid if he/she chooses. The user will accept the change by selecting the Accept change button, which will turn the text green. Rejecting the change will remove it from the list.

**Step 5 – Submit changes to the GSS server.**

Once the user has some accepted changes, he/she must then submit those changes to the GSS server. This is done by choosing the Submit changes menu option:



At this point the proposed update are added to the ArcGIS Server WFS-T.

**Step 6 (optional) – NSDI GeoSynchronization**

The process also allows for NSDI users to “follow” along with county-state changes. This is accomplished by first setting up a “sync”

## Community Outreach

This project has already delivered distributed services and applications that can be widely used for GeoSynchronization. The project team has successfully avoided practices that would inhibit the use of the distributed data and deliver short term gains. This project has focused on interoperable SDI solutions and created a GeoSynchronization environment that will entice additional participation from NSDI community of practices.



As part of this project upcoming phases we will engage in a community outreach program to reach users in Arkansas. This community engagement will supplement our attendance at GeoSynchronization community online meetings and the collaborative project workspace established by USGS and FGDC. In addition, we announced the project to US online audiences via press releases<sup>7</sup> and blog posts.


































As our Community Outreach advances we will seek community venues and conferences to brief the results of the effort and engage additional participation. We will also engage additional potential participants using social media such as our blog, as well as Twitter and YouTube. These techniques are vital to engaging a broad spectrum of potential users.

***In summary, progress to date on this project begins to bridge the gap between current file-based production operations and the FGDC vision of developing, deploying and sustaining a federation for exchanging local, state and national framework data using GeoSynchronization services and common data models.***

---

<sup>7</sup> [http://www.thecarbonproject.com/news\\_arkansas.php](http://www.thecarbonproject.com/news_arkansas.php)

## Appendix A – Sample View of County Road Shapefiles

 Arkansas.dbf	5/26/2010 1:26 PM	DBF File	3,411 KB
 Arkansas.prj	5/26/2010 1:26 PM	PRJ File	1 KB
 Arkansas.sbn	5/26/2010 1:26 PM	SBN File	43 KB
 Arkansas.sbx	5/26/2010 1:26 PM	SBX File	4 KB
 Arkansas.shp	5/26/2010 1:26 PM	SHP File	642 KB
 Arkansas.shp.xml	5/26/2010 1:26 PM	XML Document	46 KB
 Arkansas.shx	5/26/2010 1:26 PM	SHX File	33 KB
 Ashley.dbf	5/26/2010 1:26 PM	DBF File	3,128 KB
 Ashley.prj	5/26/2010 1:26 PM	PRJ File	1 KB
 Ashley.sbn	5/26/2010 1:26 PM	SBN File	33 KB
 Ashley.sbx	5/26/2010 1:26 PM	SBX File	3 KB
 Ashley.shp	5/26/2010 1:26 PM	SHP File	1,287 KB
 Ashley.shp.xml	5/26/2010 1:26 PM	XML Document	5 KB
 Ashley.shx	5/26/2010 1:26 PM	SHX File	26 KB
 Ashley.txt	7/9/2010 9:58 AM	Text Document	1,047 KB
 Baxter.dbf	5/26/2010 1:26 PM	DBF File	11,149 KB
 Baxter.prj	5/26/2010 1:26 PM	PRJ File	1 KB
 Baxter.sbn	5/26/2010 1:26 PM	SBN File	100 KB
 Baxter.sbx	5/26/2010 1:26 PM	SBX File	7 KB
 Baxter.shp	5/26/2010 1:26 PM	SHP File	2,164 KB
 Baxter.shp.xml	5/26/2010 1:26 PM	XML Document	44 KB
 Baxter.shx	5/26/2010 1:26 PM	SHX File	78 KB
 Benton.dbf	5/26/2010 1:26 PM	DBF File	15,503 KB
 Benton.prj	5/26/2010 1:26 PM	PRJ File	1 KB
 Benton.sbn	5/26/2010 1:26 PM	SBN File	220 KB
 Benton.sbx	5/26/2010 1:26 PM	SBX File	13 KB
 Benton.shp	5/26/2010 1:26 PM	SHP File	4,471 KB
 Benton.shp.xml	5/26/2010 1:26 PM	XML Document	26 KB
 Benton.shx	5/26/2010 1:26 PM	SHX File	176 KB
 Boone.dbf	5/26/2010 1:26 PM	DBF File	4,961 KB
 Boone.prj	5/26/2010 1:26 PM	PRJ File	1 KB
 Boone.sbn	5/26/2010 1:26 PM	SBN File	53 KB
 Boone.sbx	5/26/2010 1:26 PM	SBX File	5 KB

## Appendix B – Understanding GeoSynchronization – Haiti Scenario

A GSS service and Client GSS client from The Carbon Project were demonstrated in a simulated response scenario over Haiti in early 2010. This scenario was selected in consultation with government representatives and designed to represent a simulation of synchronizing geographic updates in real-time from many sources to any geospatial database using open standard web services - bridging a variety of databases, GIS vendors, and data schemas.

In this scenario, emerging CONOPs and Political, Military, Economical, Social, Infrastructure, and Information (PMESII) and Area, Structures, Capabilities, Organizations, People and Events (ASCOPE) data types were geosynchronized to ESRI and Oracle databases (simulating NATO and Army users). In this scenario, it was assumed that vector data related to PMESII-ASCOPE is missing i.e., Port Au Prince airfield apron, existing camp locations, condition of damaged buildings (impacted areas), hospital locations, buildings, power generation networks, lines of communication and roads (operational corridors). This data was added, updated and deleted on multiple WFS-T enabled services in near-real time simulating operations ‘internal’ to the Army geospatial enterprise as well as ‘external’ operations, with simulated users performing the role of Geospatial Engineering Teams.

Results of this test indicated GeoSynchronization via open standard web services can provide a key capability is to integrate geospatial information and analysis capabilities across collaborative environments. The scenario highlights how with modern geospatial interoperability tools and standards updates can be ‘published’ digitally by field users, reviewed by other users located at more centralized sites, and uploaded to multiple data stores whenever the updates are ‘accepted’.

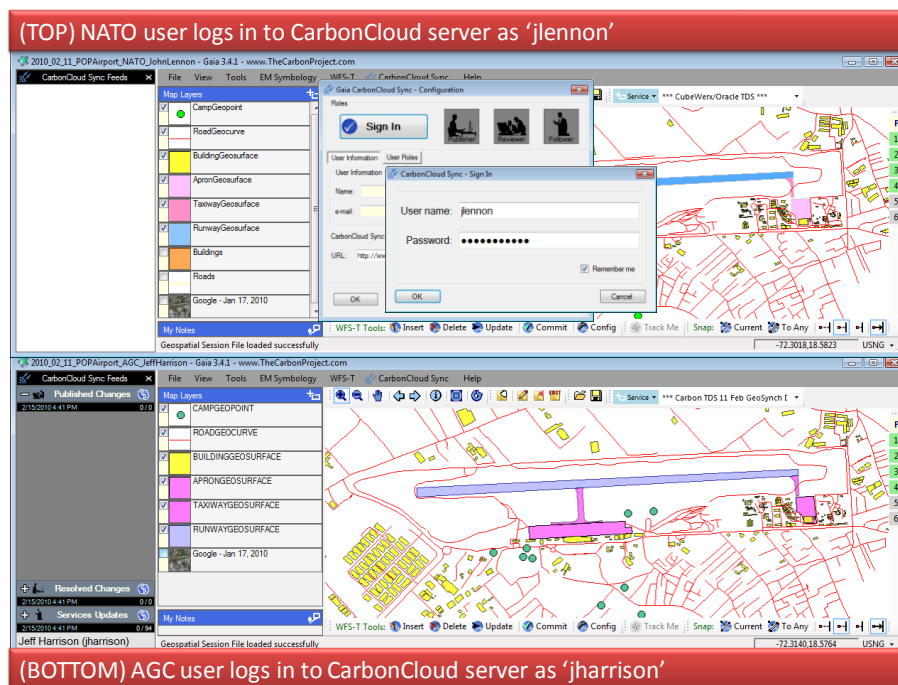
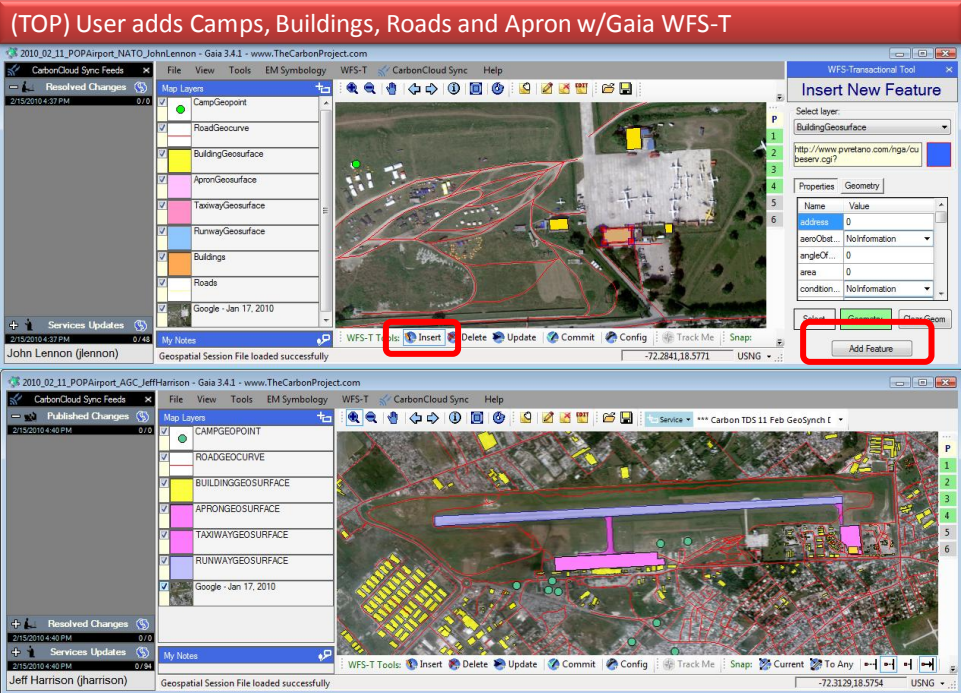
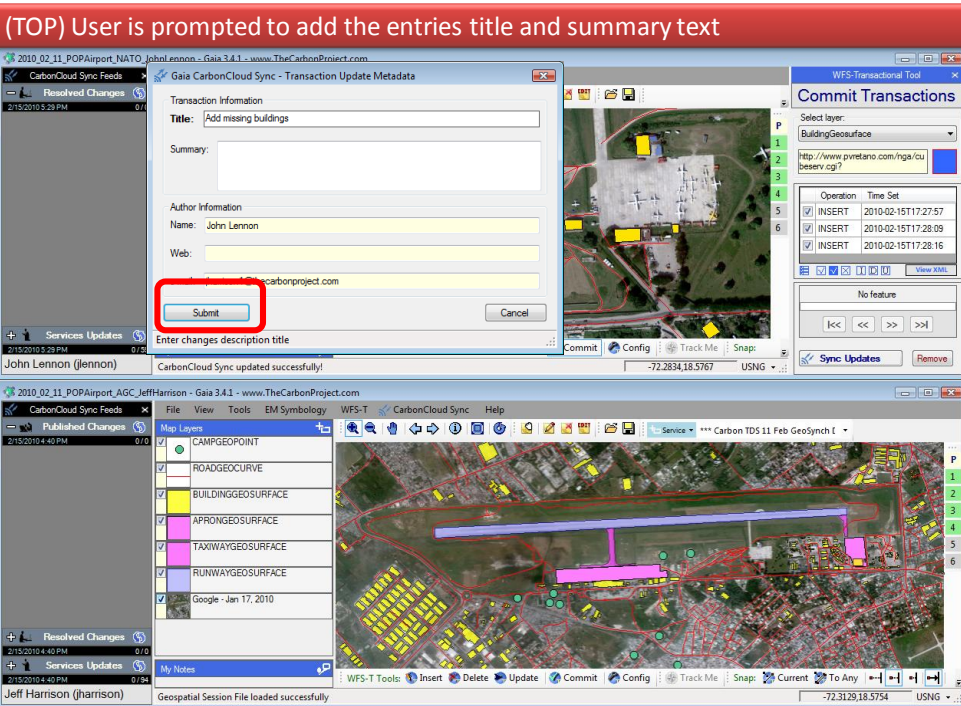


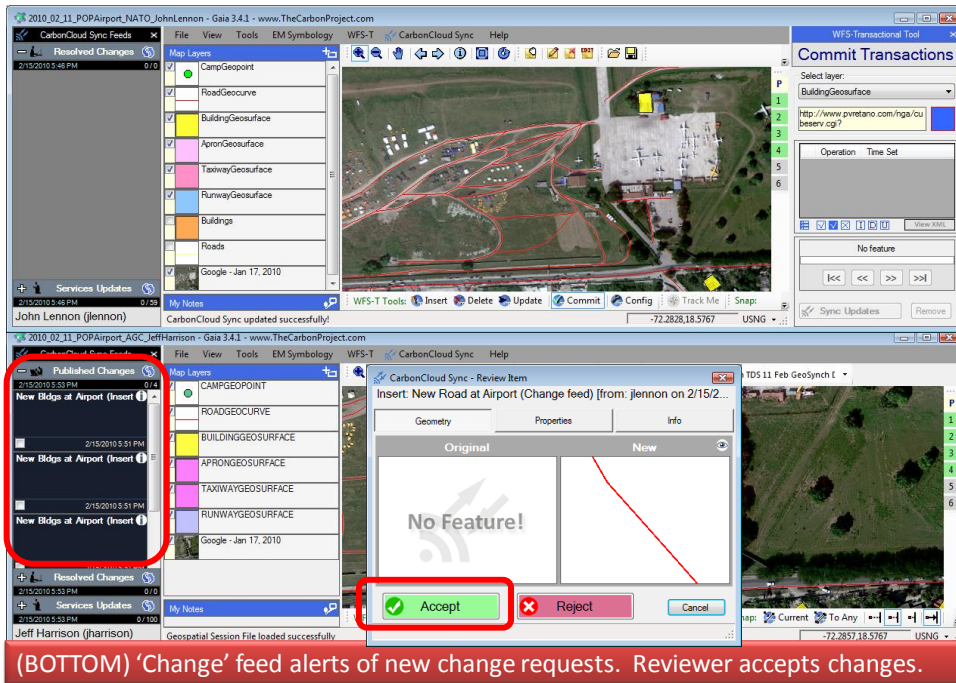
Figure – Publisher log-in to GSS



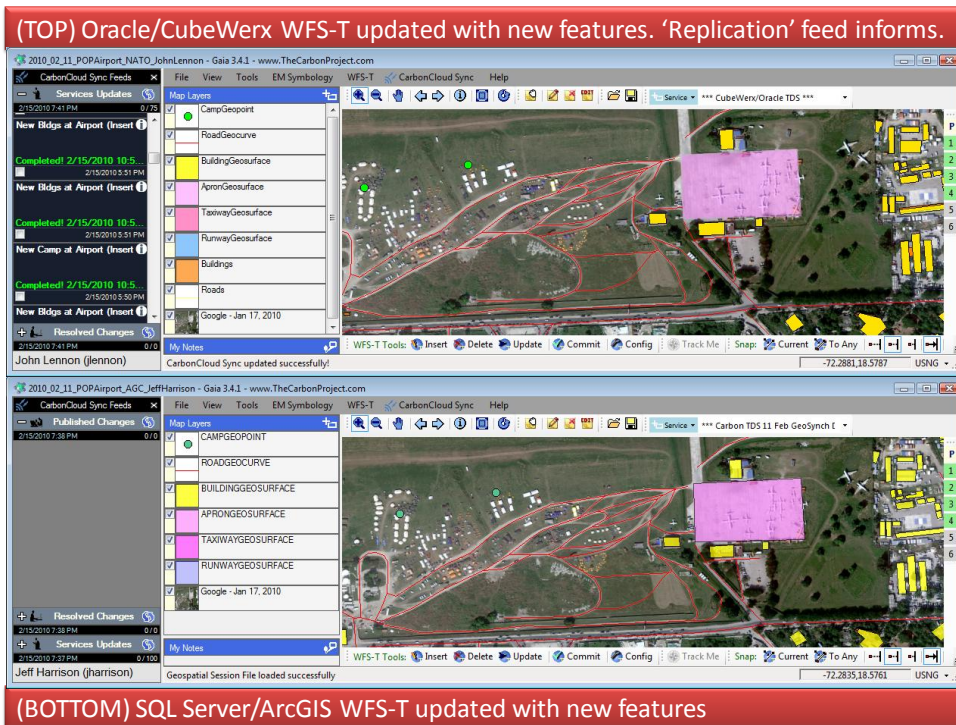
**Figure – Geospatial updates using WFS-T editing tools**



**Figure – Publisher submission of proposed updates to GSS**



**Figure – Reviewer validation of proposed geospatial updates**



**Figure – Gaia Extenders working with the CarbonCloud Sync Geo-Synchronization Service (GSS) to collaboratively update geospatial data in SOA of ESRI and Oracle web services**