

InCLUDE Data Exchange

Integrated **C**adastral and **L**and **U**se **D**ata **E**xchange ■

David Wyatt

Eastern Band Of Cherokee Indians

Project Lead

Julia Harrell, GISP

GIS Coordinator, NC DENR

Julia.harrell@ncdenr.gov



InCLUDE Project Partners

- **NC Department of Environment & Natural Resources:** The State of NC's lead environmental agency
- **Eastern Band of Cherokee Indians:** Has Tribal lands largely in NC
- **NC Property Mapper's Association:** a non governmental professional organization of County Government property mappers
- **NC Division of Land Records:** part of the NC Secretary of State's office
- **NC Statewide Mapping Advisory Committee:**
 - WGSP:** Working Group for Seamless Parcels
 - WGRT:** Working Group for Roads & Transportation

Why use the Exchange Network to share parcel & land use data?

- Many environmental and disaster problems cross multiple jurisdictions. The business of managing and solving these problems has become complex and very information intensive.
- Environmental policy makers and other stakeholders need access to timely, accurate, and consistent data that present a complete picture of the environment. This includes property ownership and land use.
- Exchanging high quality information allows governments and regulated communities to make better planning decisions regarding the environment. Property data is a vital missing piece of the puzzle, not only for planning, but also incident response.

Benefits to NC DENR

- Most regulatory and natural resource programs need current property boundaries, ownership and land use information, which is not consistently available in a common format.
- Some of the 100 NC County systems are mainframe-based and/or extremely proprietary.
- A consistent source of seamless parcel and land use data will aid DENR in: regulatory permitting decisions, mitigation and remediation efforts, land acquisition for parks and protected areas, wildfire response activities, and regional level analysis.

Benefits to Eastern Band of Cherokee

- More easily get copies of county property data and share Trust land boundaries issues digitally with NC County property mapping offices.
- Access NC County land records digitally and allow the Tribal Historic Preservation Office to work more closely with the State Historic Preservation Office.
- Locate culturally important sources of natural resources, such as River Cane and Butternut, and have easy access to property owner information to facilitate permission for harvesting these resources.

Benefits to NC Counties



- Reduce the burden of filling frequent parcel data requests from state and federal agencies.
- Establish a common data exchange format for parcel data without requiring counties to change any of their back-end systems.
- Allows counties to maintain control of transforming their data and determining who should have access to it.
- Will aid in creating regional level economic development and growth plans, as well as mutual aid and incident response compacts.

Major Project Components:

InCLUDE = Integrated Cadastral and Land Use Data Exchange. InCLUDE is the name of both the data flow and the project. Will use a GML “Simple Features” Level 1 schema, following the FGDC Parcel Working Group guidelines as much as is practical.

SPART-NC = Spatial Translator for NC. A secure web application allowing local governments to upload and transform their parcel and land use data to the community schema and submit it to the NC InCLUDE Node.

NCLandGate = a secure web publishing application for InCLUDE data for Government & Tribal users to download desired areas of the transformed parcel & land use data from a map-centric interface.

What has changed in the project ?

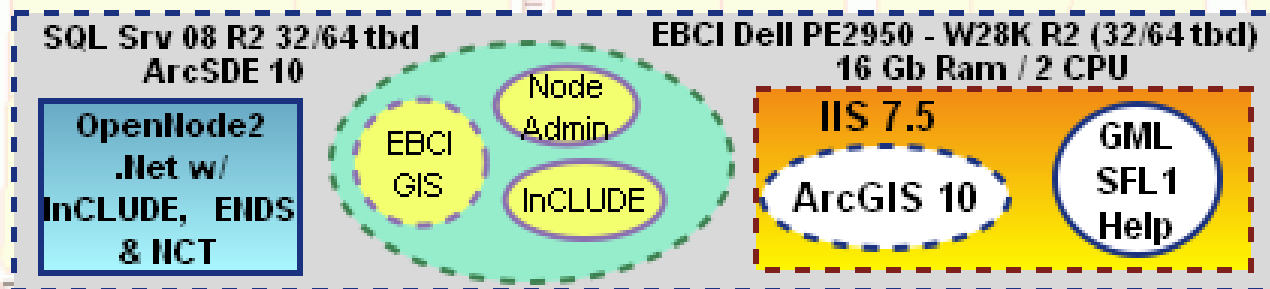
- Cost analysis determined that it would be more economically viable to build the data transformation application as a web-based product rather than as a desktop software application.
- Functionality analysis determined that much of the needed data transformation capabilities already exist in an open source software package designed to provide Spatial ETL (extract, transform & load) services: **GeoKettle**.
- Functionality analysis indicated that most of the needed “on-demand” spatial data requests the InCLUDE data flow must provide are already supported by the OGC’s Web Feature Service (WFS) filter encoding spec, the goal is to integrate that technology rather than develop dedicated spatial selection code for the InCLUDE OpenNode2 plug-in.
- The open source **52North GeoWeb Security** software *may* be used to secure both types of WFS services if the native security proves inadequate.

What has not changed in the project?

- InCLUDE Plug-in will be built for both the Java & .Net flavors of OpenNode2. DENR will use Java & EBCI will use .Net
- The OpenNode2 administration database will be ported to run on the PostgreSQL + PostGIS platform for use with the NC DENR InCLUDE Node. This will be reusable with other instances of ON2. EBCI will use an existing instance of SQL Server + ArcSDE as their InCLUDE spatial database.
- DENR will use the open source GeoServer as its companion WFS server and EBCI will use ArcGIS Server 10. AGS will require a “assist module” to provide WFS support for GML SF L1.
- A secure web-based data publishing application (NCLandGate) will be available to Government & Tribal staff for ease in obtaining parcel data in shapefile format rather than InCLUDE data flow GML.

InCLUDE System Components Diagram

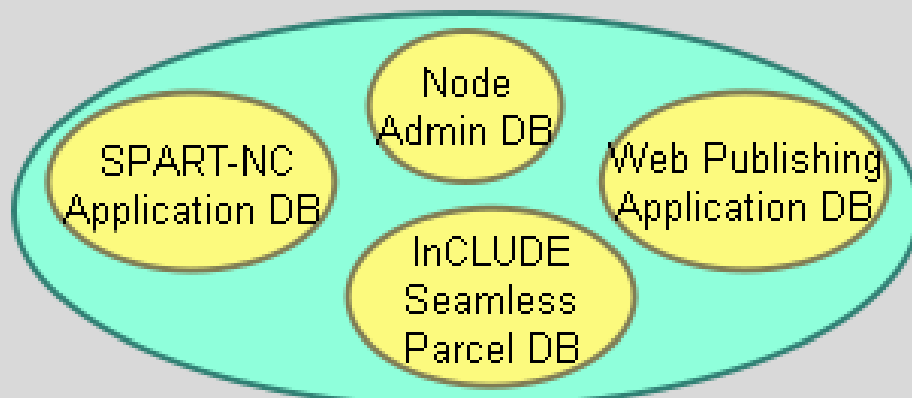
NC InCLUE System
Proposed Components
(dashed = existing)



DENR VMWare 4.5 VM, OS = W28k Server R2 (32/64 tbd) 16 GB Ram / 2 CPU

GeoKettle 3.2 (or latest)

SAN Storage - File system for uploaded data, metadata, transform recipes, etc.



PostgreSQL 9.0.3 + PostGIS 1.5.2
(or latest stable release when dev. starts)

Apache 2.2.17 & mod_cas

SPART-NC Web App

InCLUE Map Web App

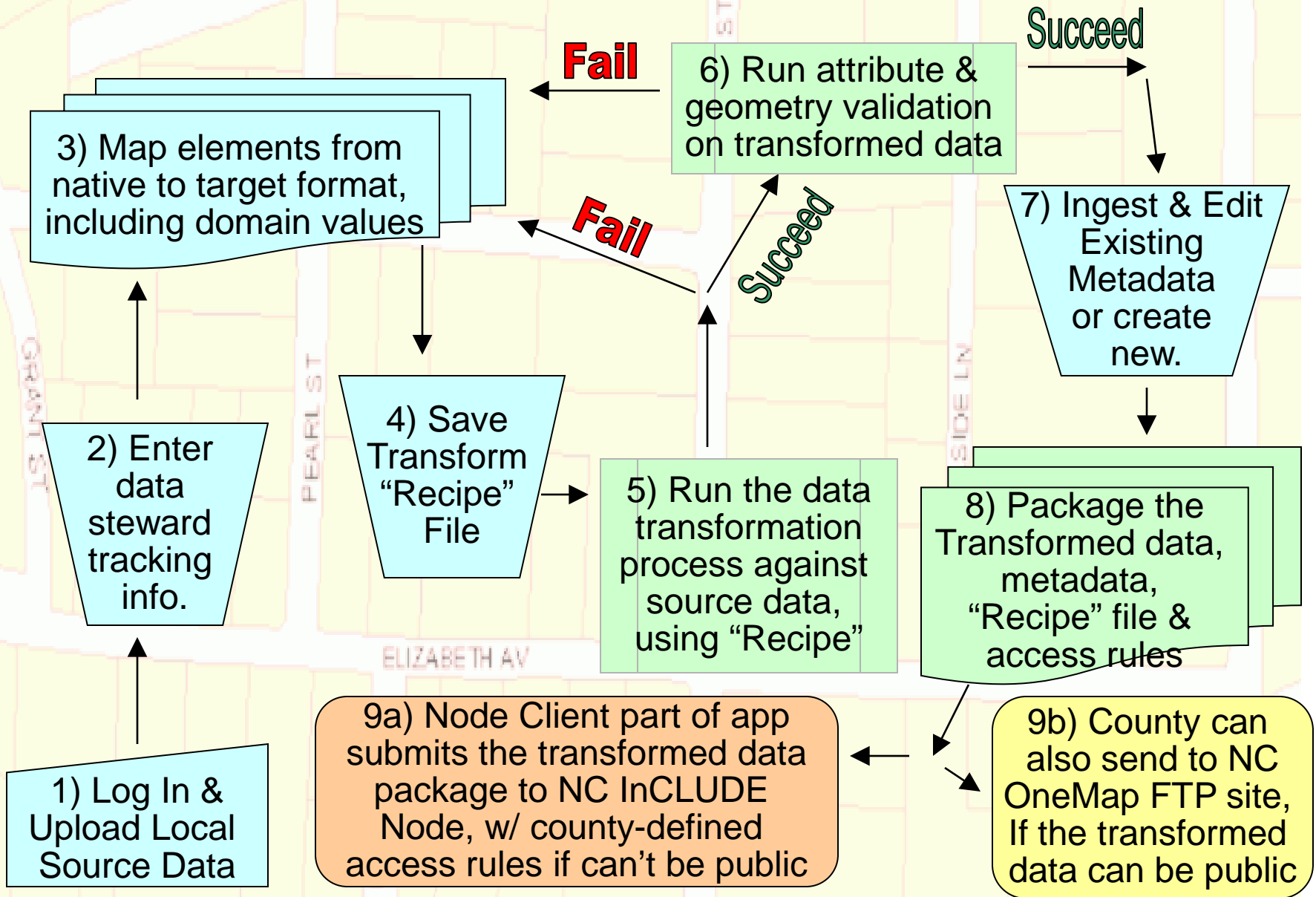
ApacheTomcat 6.0.18

52North WAS/WSS /WSC.Web

Geoserver 2.1 & AppSchema Ext.

OpenNode2 Java w/
InCLUE, EHDS & NCT

SPART-NC: Flow of User Operations



SPART-NC: behind the scenes

- Uploaded source shapefile & DB file ingested into NC InCLUDE's PostgreSQL + PostGIS spatial database.
- SPART-NC reads source element structure out of NC InCLUDE spatial DB and transformed data structure out of InCLUDE GML schema.
- User defines which source elements correspond with transformed destination elements, maps source domain values to destination domain values in InCLUDE schema. App has Javascript string/number functions and simple conditional logic to "slice & dice" the source data.
- App saves user defined data mappings in GeoKettle's existing xml transformation 'recipe' file format. App sends 'recipe' file to GeoKettle's Carte web service interface for transformation.
- Validation & reporting tools provide user-friendly error reports. Underlying GeoTools/JTS library used for geometry validation.
- User ingests & edits existing metadata, or creates new metadata.
- Validated data, metadata, recipe file, and access permissions sent to NC InCLUDE Node & are loaded into the spatial database.

SPART-NC: Data Mapping Process

The screenshot illustrates the data mapping process in SPART-NC. It features a central mapping table, a domain value matching interface, and a transformation recipe builder.

Source Data	Parcel Schema
First Name	Parcel_ID
Date Recorded	Parcel_ID_Alt
Deed Book	Owner_FirstName1
PID	Owner_LastName1
Deed Page	Source_Reference
Last name	Source_Date
PIN	Site_St_Number
Physical Address	Site_St_Name
	Site_St_Type

Coded Domain Values Matching Lists

Input Domain	Output Domain
Code Value 1	Code Value 1
Code Value 2	Code Value 2
Code Value 3	Code Value 3

Save Transformation Recipe to Local XML configuration file

Browse to Location & Enter Filename

Javascript and/or Python expression builder and scripting interface to allow the "slicing and dicing" of source data during the transformation process

OK Cancel Undo Help

+ - / * & = > < <> And Or Not Like () Paste Help

- Forms
- Reports
- Functions
 - Built-In Functions
 - GeoKettle Functions
- Constants
- Operators
- Common Expressions

Conversion	Asc
Database	CBool
Date/Time	CByte
DDE/OLE	CCur
Domain Aggregate	CDate
Error Handling	Cdbl
Financial	CDec
General	Chr
Input/Output	Chr\$
Inspection	CTr

SPART-NC web GUI will be simple AJAX, much simpler than GeoKettle desktop GUI, but all underlying functions can still be included in the recipe file.

NC LandGate: User Operations

- Authorized Government or Tribal user logs into secure NCLandGate app using NCID or NAAS account.
- Selects a County or Tribal Area Subdivision Name from a pick-list, or zooms on the map interface to the location where they want to get InCLUDE parcel data by one of these methods:
 - Lat/Lon coordinate
 - NC State Plane coordinate
 - US National Grid coordinate
 - Street Address (*Mapstraction* API w/ Google / Yahoo / OSM)
- Methods for selecting InCLUDE parcel data to download:
 - By County or Tribal Area Subdivision Name
 - By User-Defined Bounding Box (within reason)
 - By User-Defined Center Point & Buffer Radius (within reason)
 - By 10-digit Watershed Boundary ID
- User defines desired output file format (from a short list) and coordinate system (with a valid EPSG code)

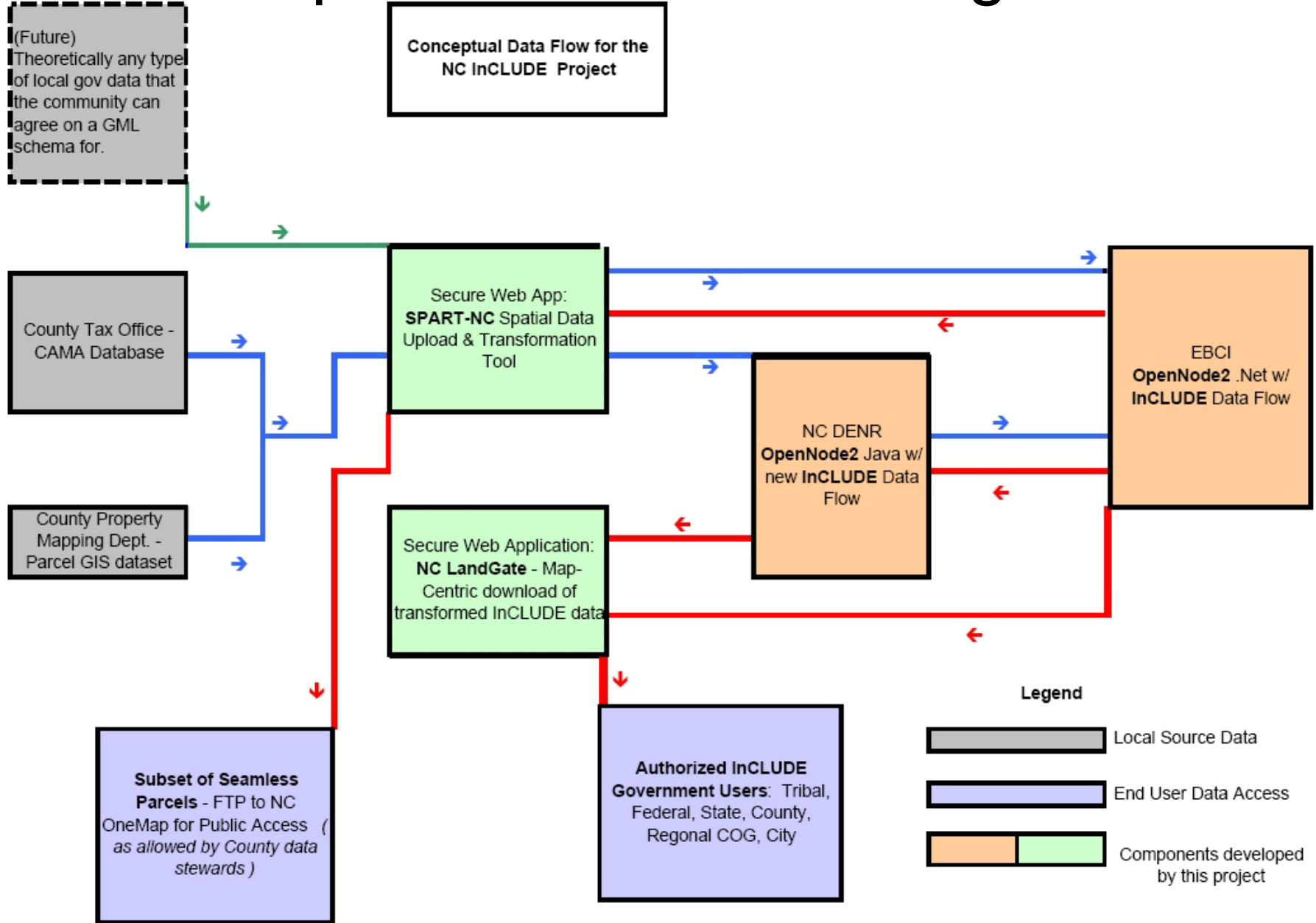
WFS Filter Request Methods

- InCLUDE data flow will support the following “on-request” data provider methods:
 - Get Parcel polygons by County or Tribal Area Subdivision Name
 - Get Parcel polygons by user-specified rectangle
 - Get Parcel polygons within x distance from a user-specified point
 - Get Parcel polygons that overlap a user-specified 10-digit Watershed Boundary polygon
- Will not code these methods directly into the InCLUDE plugin, but plugin will modify the EN data flow request XML to the corresponding WFS filter query syntax and pass on to the InCLUDE WFS to handle the spatial filtering. Keep it generic.
- Corresponding WFS Filter Query requests:
 - equalTo or is Like
 - BBOX (extents within reason)
 - DWithin (center point & buffer radius, within reason)
 - Overlaps
 - Note: *ArcGIS Server does not support DWithin yet, may have to generate our own “circle” polygon, and use it w/ the Overlaps filter*

NCLandGate: Behind the Scenes

- User specifies desired data output format (GML, shapefile, MapInfo, DXF, Personal Geodatabase) and Coord. System.
- Web App takes user-defined info and formulates appropriate request, sends it to InCLUDE Node along w/ user credentials.
- InCLUDE logic checks if user has permission to access parcel data for the areas their current data request covers.
- If OK, InCLUDE plugin sends request to secured WFS service (Geoserver or ArcGIS Server), which serves the transformed “seamless parcels” data in DENR or EBCI InCLUDE spatial DBs.
- If user selected GML, the WFS returns GML in the InCLUDE schema. (Geoserver uses App Schema extension, AGS will need custom “assist” code to support SF Level 1).
- If user selected Shapefile or other format that does not support multi-valued properties (ex: a parcel with > 1 owner names) the multiple values are saved in a “related” tabular file that can be joined to the output file by the user in their GIS system of choice.

Proposed Data Flow Diagram



(Future)
Theoretically any type of local gov data that the community can agree on a GML schema for.

Conceptual Data Flow for the NC InCLUDE Project

County Tax Office - CAMA Database

Secure Web App: SPART-NC Spatial Data Upload & Transformation Tool

EBCI OpenNode2 .Net w/ InCLUDE Data Flow

County Property Mapping Dept. - Parcel GIS dataset

Secure Web Application: NC LandGate - Map-Centric download of transformed InCLUDE data

NC DENR OpenNode2 Java w/ new InCLUDE Data Flow

Subset of Seamless Parcels - FTP to NC OneMap for Public Access (as allowed by County data stewards)

Authorized InCLUDE Government Users: Tribal, Federal, State, County, Regional COG, City

Legend

- Local Source Data
- End User Data Access
- Components developed by this project

Software planned for use with InCLUDE

Function	Open Source Software	Commercial Software
Presentation / Web Client	OpenLayers, MapFish, and Mapstraction API (all Javascript based) http://openlayers.org/ http://mapfish.org/ http://www.mapstraction.com/	none
Exchange Network	OpenNode2 + NAAS + ENDS http://code.google.com/p/opennode2/ http://www.exchangenetwork.net/	none
GIS Application Server	GeoServer + 52North GeoDRM services (WSS/WAS) http://www.geoserver.org/ http://52north.org/communities/security/index.html	ESRI's ArcGIS Server http://www.esri.com/software/arcgis/arcgisserver
Spatial Data ETL	GeoKettle http://www.spatialytics.org/geokettle	none
Database	PostgreSQL RDBMS + PostGIS spatial extender http://www.postgresql.org http://www.postgis.org	Microsoft SQL Server RDBMS + ArcSDE spatial extender http://www.microsoft.com/sqlserver http://www.esri.com/software/arcgis/arcade

What are OpenLayers and MapFish ?

- **OpenLayers:**

- Was created in 2005 by MetaCarta and presented at the *Where 2.0* Conference as an open alternative to the Google Maps API.
- Is an AJAX toolkit for freely combining geographic data layers *from virtually any source* in a web mapping client.
- Is a pure JavaScript API with no server-side dependencies.
- Is an OSGEO (Open Source Geospatial Foundation) project.

- **MapFish:**

- Is a flexible framework for building rich web-mapping applications. Has both a client and a server component. We are interested only in the client portion.
- Provides a complete RIA-oriented JavaScript toolbox, testing environment, and tools for compressing JavaScript code. The JavaScript toolbox is composed of the *ExtJS*, *OpenLayers*, *GeoExt* JavaScript toolkits.
- Currently undergoing incubation to become an OSGEO project.

What is the Mapstraction API ?

- Is an open source Javascript library that provides a common API to a number of “Mass Market” online mapping providers (Google Maps, Yahoo Maps, MapQuest, Cloudmade, Bing Maps, etc.)
- Is integrated with the OpenLayers library, and can also access information from OpenStreetMap.
- Allows web mapping applications to use maps and services from many different providers without having to write a separate block of code for each one.
- This project will use Mapstraction primarily to access geocoding web services from more than one mapping service provider in the NC Landgate application.

What are GeoServer & 52North geoDRM?

- **GeoServer:**

- The open source reference implementation of the Open Geospatial Consortium's Web Feature Server (WFS) specification.
- A Java-based (GeoTools) GIS application server that supports both publishing and editing of geospatial data via the internet.
- Supports many different types of geospatial data sources.
- Is an OSGEO project

- **52North geoDRM & Web Security Services:**

- Open Source java servlet-based software that provides security for OGC web services (WMS, WFS, WPS, SOS).
- Supports WSS (Web Security Service) Protocol: Allows you to restrict access to your OGC Web Services.
- Supports WAS (Web Authentication Service) Protocol: Allows applications to authenticate users and return a SAML ticket.

What is GeoKettle?

- A spatially-enabled version of Pentaho's Data Integration Suite
- A metadata-driven spatial ETL tool with direct execution of transformations (think Safe Software's FME, only open source)
- Support for most RDBMS formats and over 30 vector GIS formats (via OGR & GeoTools)
- Processing of Geometry objects (via GeoTools & Java Topology Suite)
- Has a Javascript API for accessing Geometry Objects
- Supports topological predicates, and aggregation operators (intersect, envelope, union, etc.)
- Supports Coordinate Transformations (via GeoTools & Proj)
- Has cartographic preview and can be deployed via web services

What are PostgreSQL and PostGIS?

- **PostgreSQL:**

- An enterprise-class, object-relational database system (15+ yrs in development), very comparable to Oracle.
- Runs on all major operating systems (Linux/Unix, MacOS Windows).
- Is fully ACID compliant, includes most ANSI SQL:2008 data types, also supports storage of binary large objects.
- Extremely scalable - there are active PostgreSQL systems in production that manage > 4 terabytes of data.

- **PostGIS:**

- An open source “spatial extender” for PostgreSQL that allows it to manage geographic features in the RDBMS.
- Follows the OGC “Simple Features Specification for SQL” spec.
- Offers support for “geography” data types for managing geodetic (lat/lon) data, support for rasters, many spatial predicate operations (intersect, union, dwithin, etc.), geometry processing functions (buffer, convex hull, polygonize, etc.), linear referencing, and much more.
- Is an OSGEO project