Using Python to Gather Information about Data in SDE

Replication Requirements

David Watkins
Pacific Southwest Region GIS Coordinator
USDA Forest Service
Background

- The USDA Forest Service manages data in SDE and utilizes replication as a workflow to edit data.
- This Presentation illustrates using Python scripts to determine whether datasets meet the requirements for participation in replication.
Replication Requirements

• In Order to participate in replication a dataset must meet the following:
  – It must be Registered As Versioned
  – It must be stored with High Precision
  – Each Feature Class must have a GlobalID field
  – It must have ‘Public’ view access granted
Forest Service Scenario

• Data is stored in Oracle database
• ArcSDE 10.1 is installed
• All data has been converted to High Precision
• Each forest has a GIS Coordinator that uses a Proxy account to connect to the data
  – This account allows them to manage the data with elevated privileges
The Issue

• It can be time consuming to find out if all the data meets the requirements for replication if you have many feature datasets.

• The status changes frequently
The Solution

• Create a Python script to gather the necessary information for you

• All you need is a valid Proxy connection file to the database
The Script

- Two Parameters
  - SDE Workspace
  - Output CSV file
- Uses `arcpy.da.walk` module to walk through the datasets in the geodatabase
- Uses `Describe` function to get info on Precision and whether the dataset is registered as Versioned
- Checks field names for presence of ‘GlobalID’
- Uses SQL statement to determine ‘Public’ access
- Uses `csvwriter` to write the output file
Script Setup

- Import required modules
- Get the parameters and set variables
- Owner is [S_R05] – need to drop []s
- Setup the SQL connection
Setup the CSV file for writing
Set the da.walk parameters
Check to see if data is in a feature dataset
Check to see if data exists
Describe Object

• Create a Describe object
• Check the ‘IsVersioned’ property
• Get the Spatial Reference and check the ‘ishighPrecision’ property
GlobalID field

- Get the field list using ListFields
- Check for presence of GlobalID field

```python
if datatype in ['FeatureClass', 'Table']:
    fieldlist = arcpy.ListFields(os.path.join(dirpath, filename))
    fieldnames = []
    for f in fieldlist:
        fieldnames.append(f.name)
    if 'GLOBALID' in fieldnames or 'GUID' in fieldnames:
        hasGUID = 'True'
    else:
        hasGUID = 'False'
else:
    hasGUID = 'NA'
```
Public Access

- Create SQL statement
  - Check all_tab_privs table
  - Check for ‘Public’ access granted
Wrapping Up

- Write the CSV file
- Clear workspace cache
Questions?