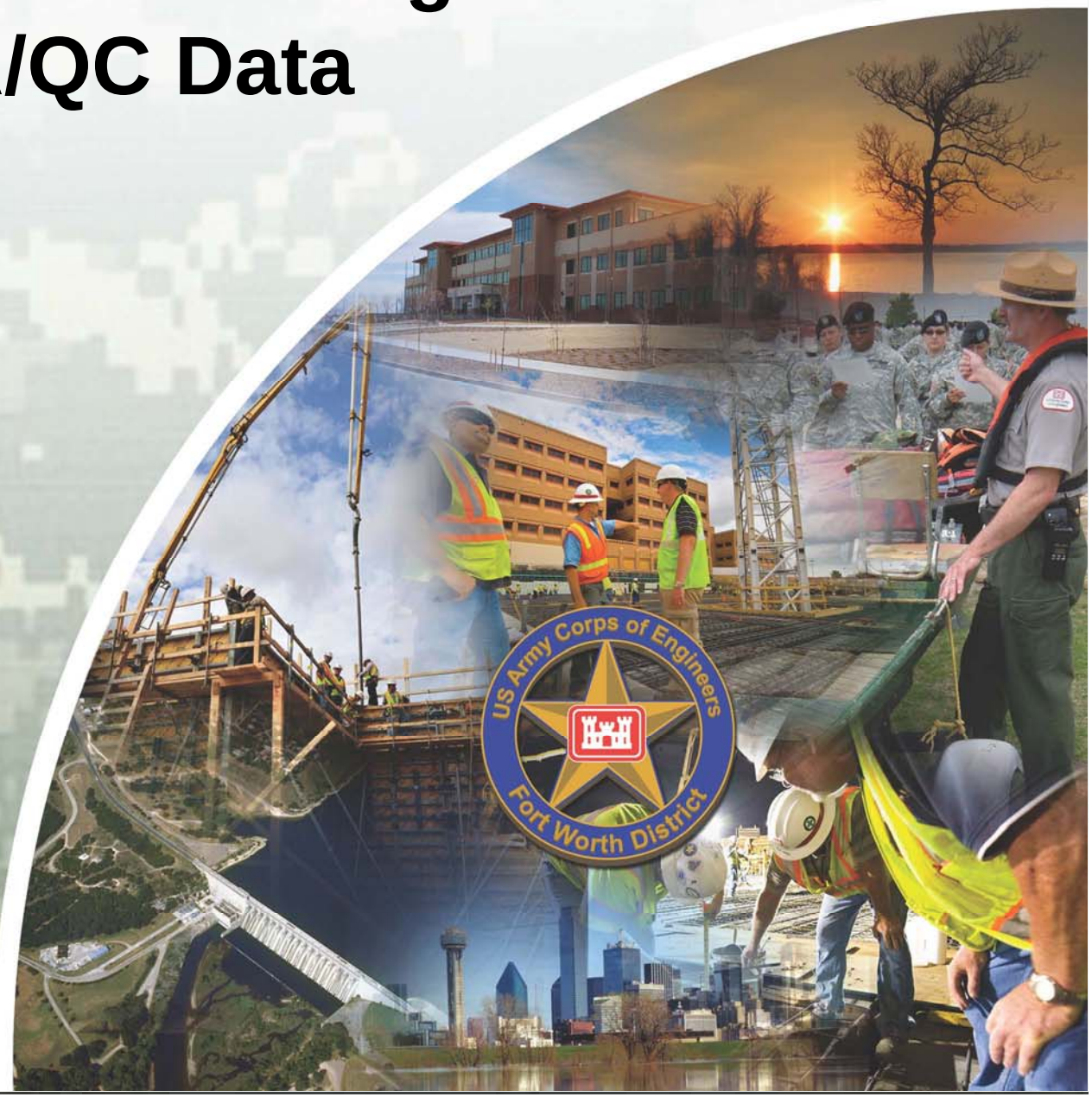
# Python – It's Like Having an Extra Person to QA/QC Data

**Jennifer M. Holland, GISP**

**28 June 2016**



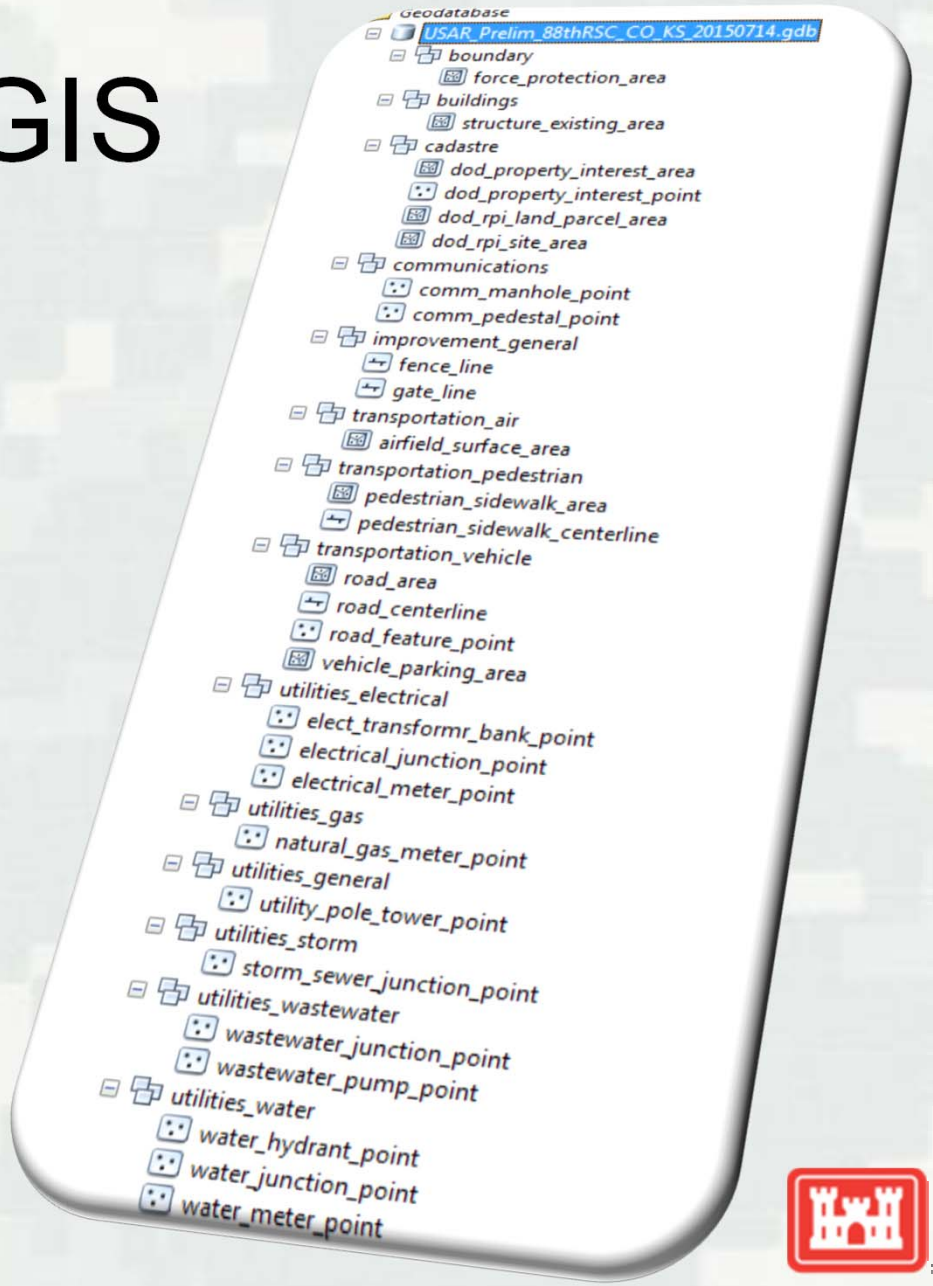

US Army Corps of Engineers
BUILDING STRONG®

# Project Details

- U.S. Army Reserve (USAR)
  - ▶ Army requirement to have complete data
    - Not met
  - ▶ Need completed data to manage their facilities
- Data collection project
  - ▶ 473 sites
    - On site data collection
      - ▷ GIS (map grade GPS)
      - ▷ Space and Facility Utilization

# GIS

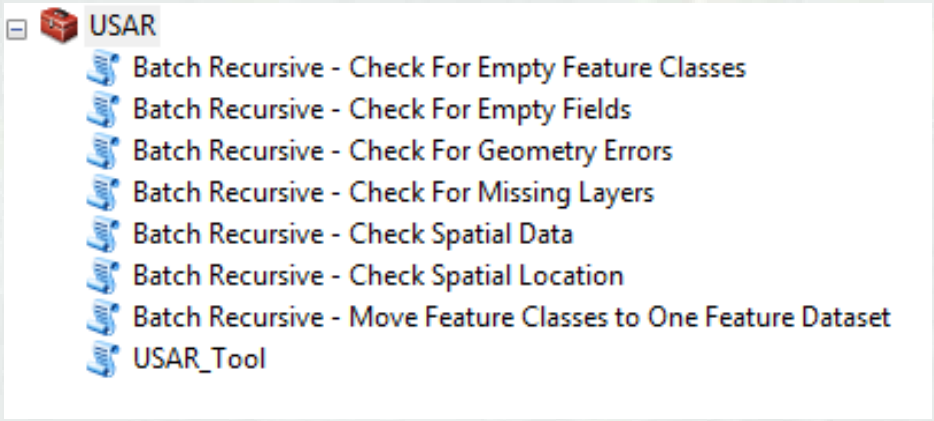
- SDSFIE version 2.6 data schema
  - ► 14 feature datasets
    - 28 feature classes



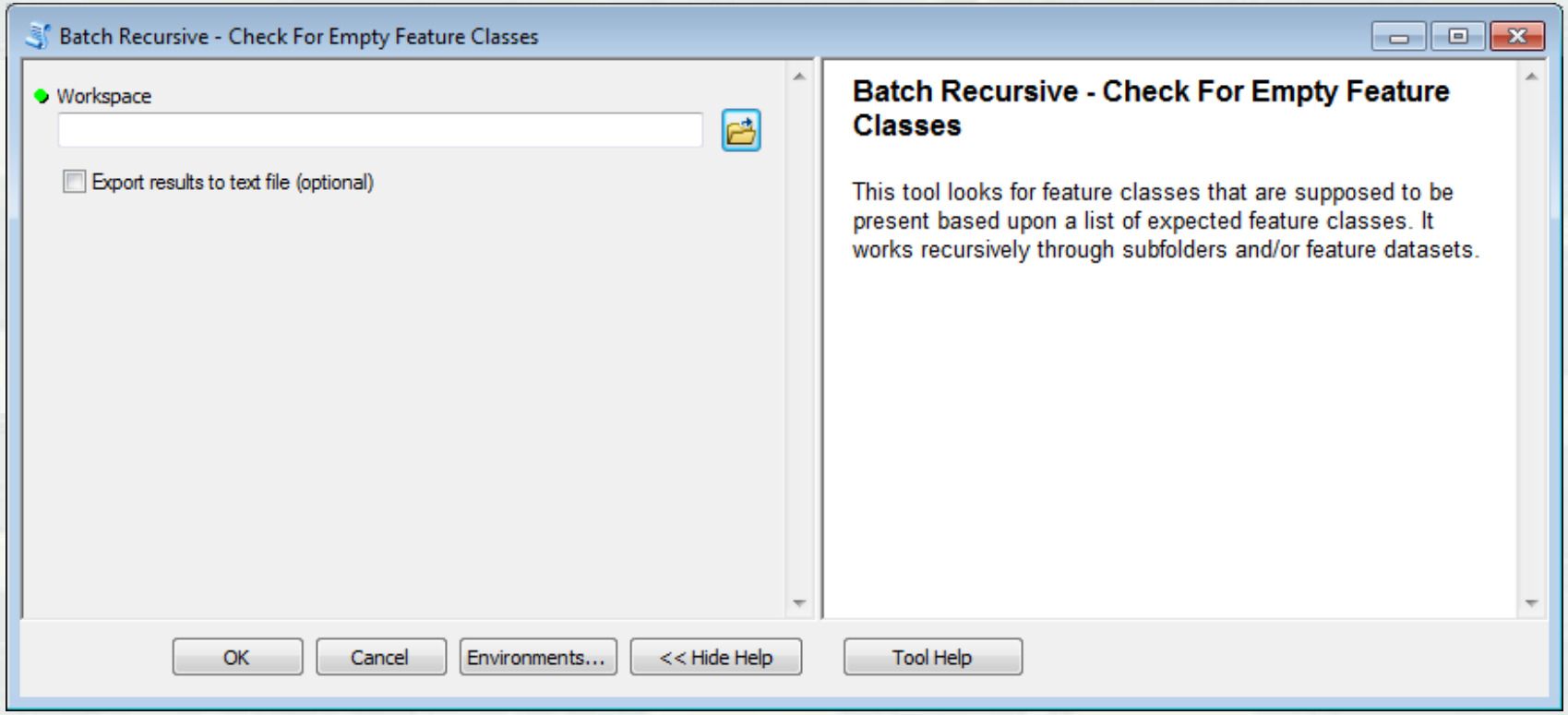
# Fort Bliss

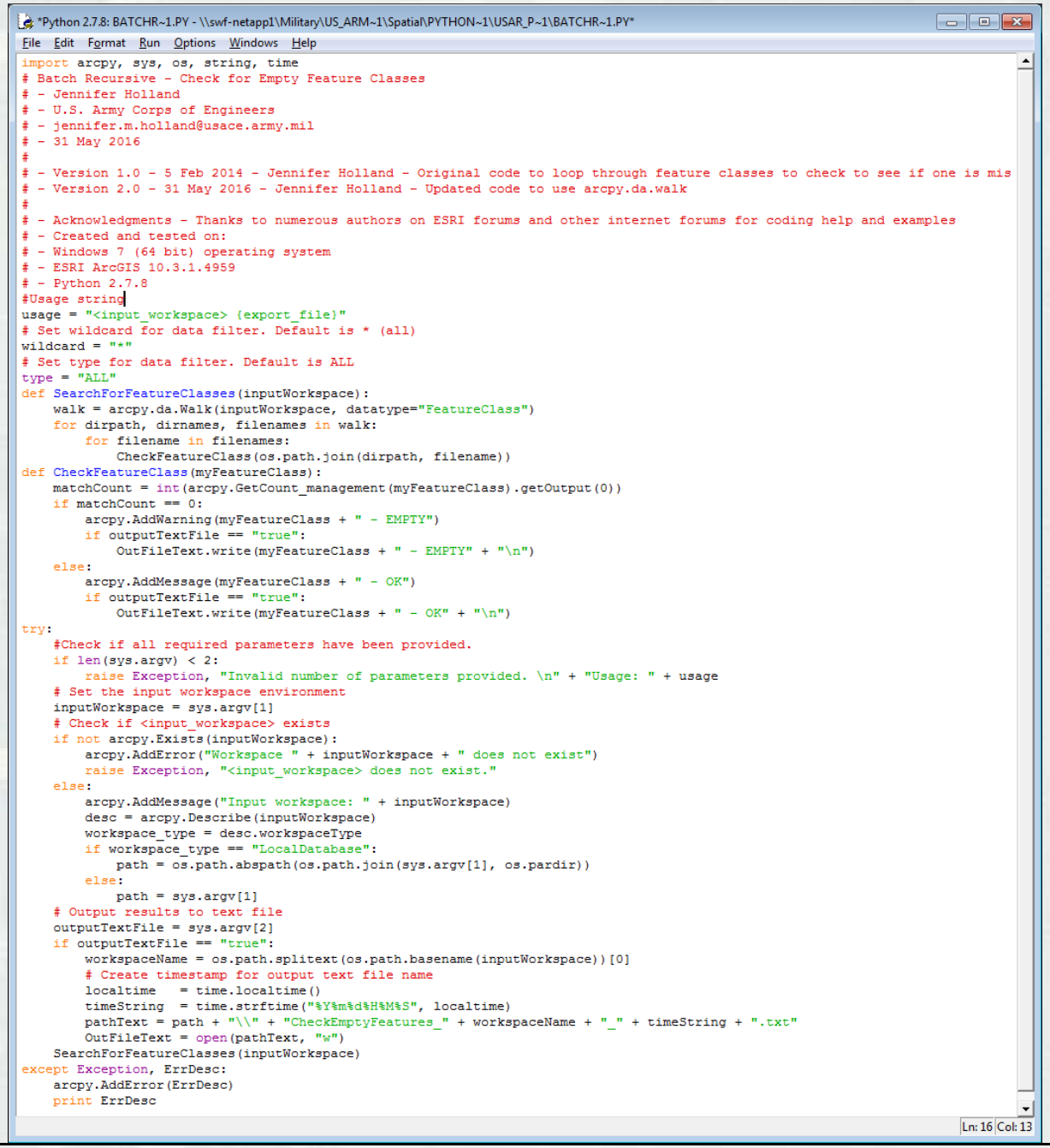
# QA/QC Process

- Requirements
  - ▶ Must QA/QC at least 20% of the GIS data
    - Spatial
    - Attributes
- Interested in automating the process as much as possible
  - ▶ Python

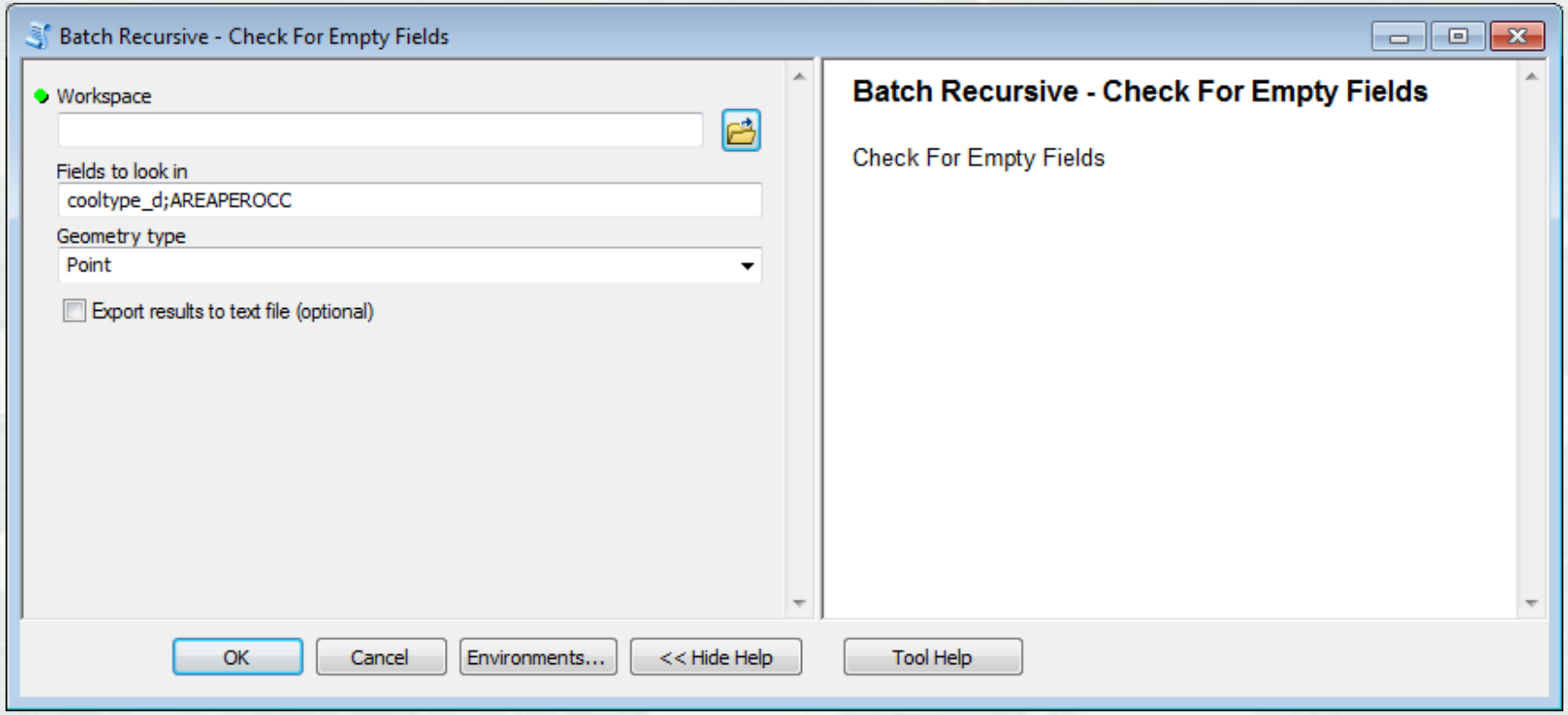

USAR
Batch Recursive - Check For Empty Feature Classes
Batch Recursive - Check For Empty Fields
Batch Recursive - Check For Geometry Errors
Batch Recursive - Check For Missing Layers
Batch Recursive - Check Spatial Data
Batch Recursive - Check Spatial Location
Batch Recursive - Move Feature Classes to One Feature Dataset
USAR_Tool

# Check for Empty Feature Classes

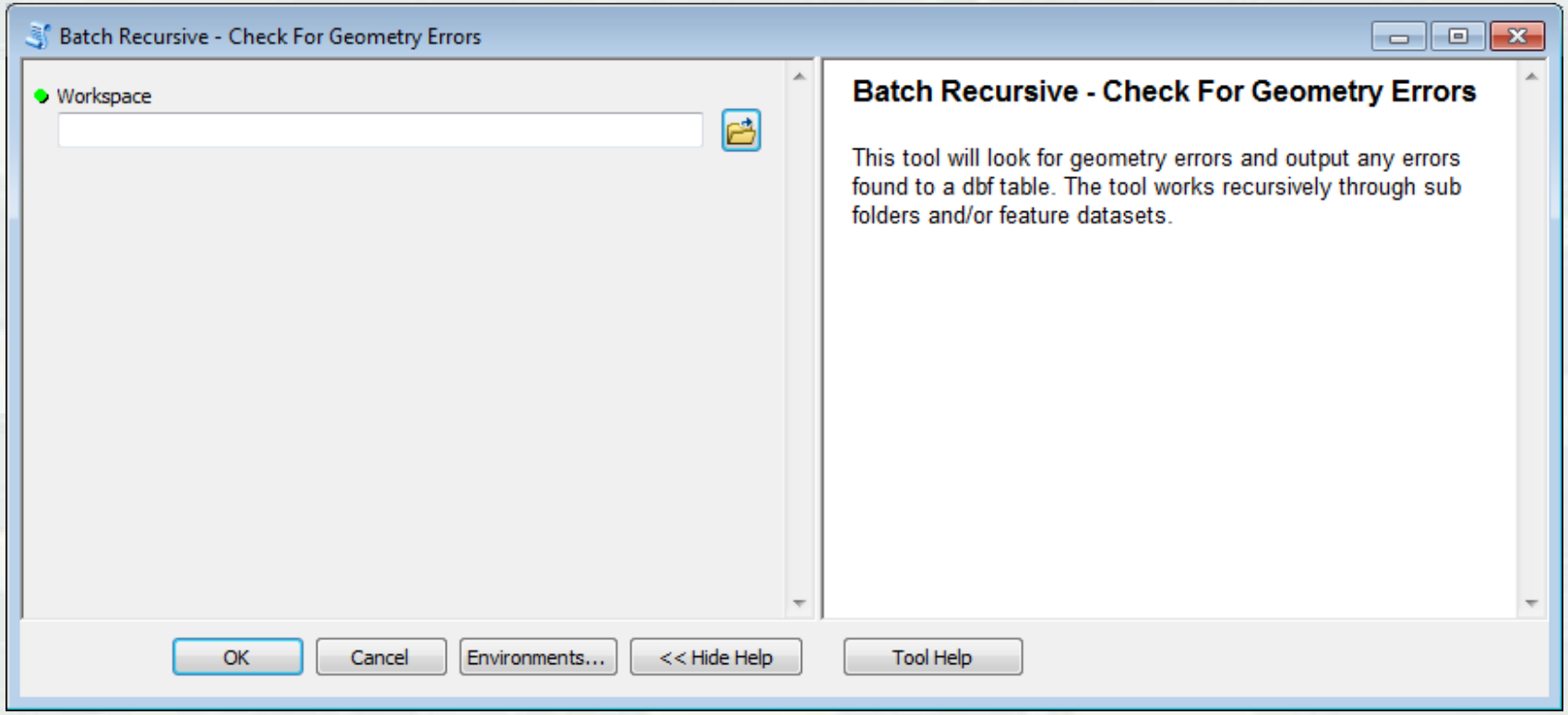




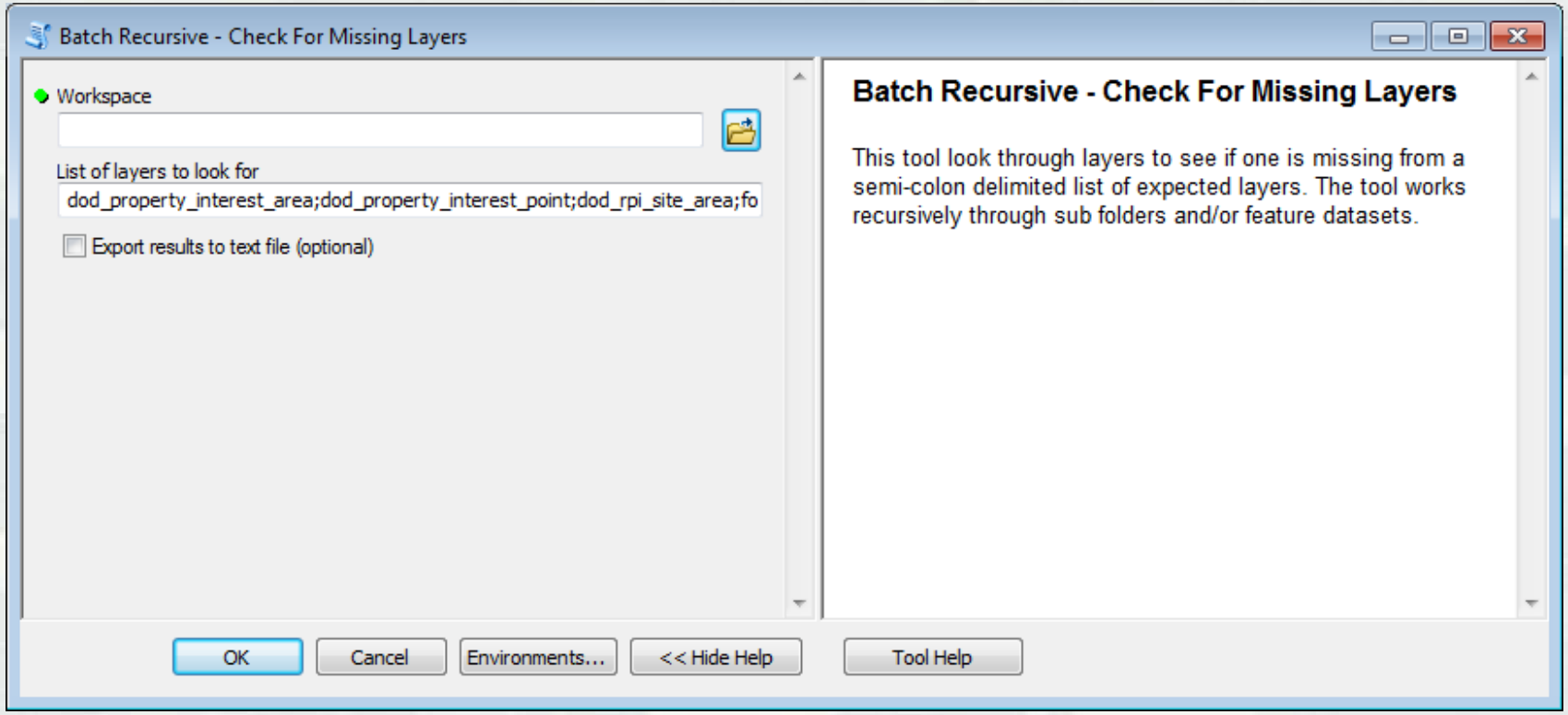
# Check for Empty Fields



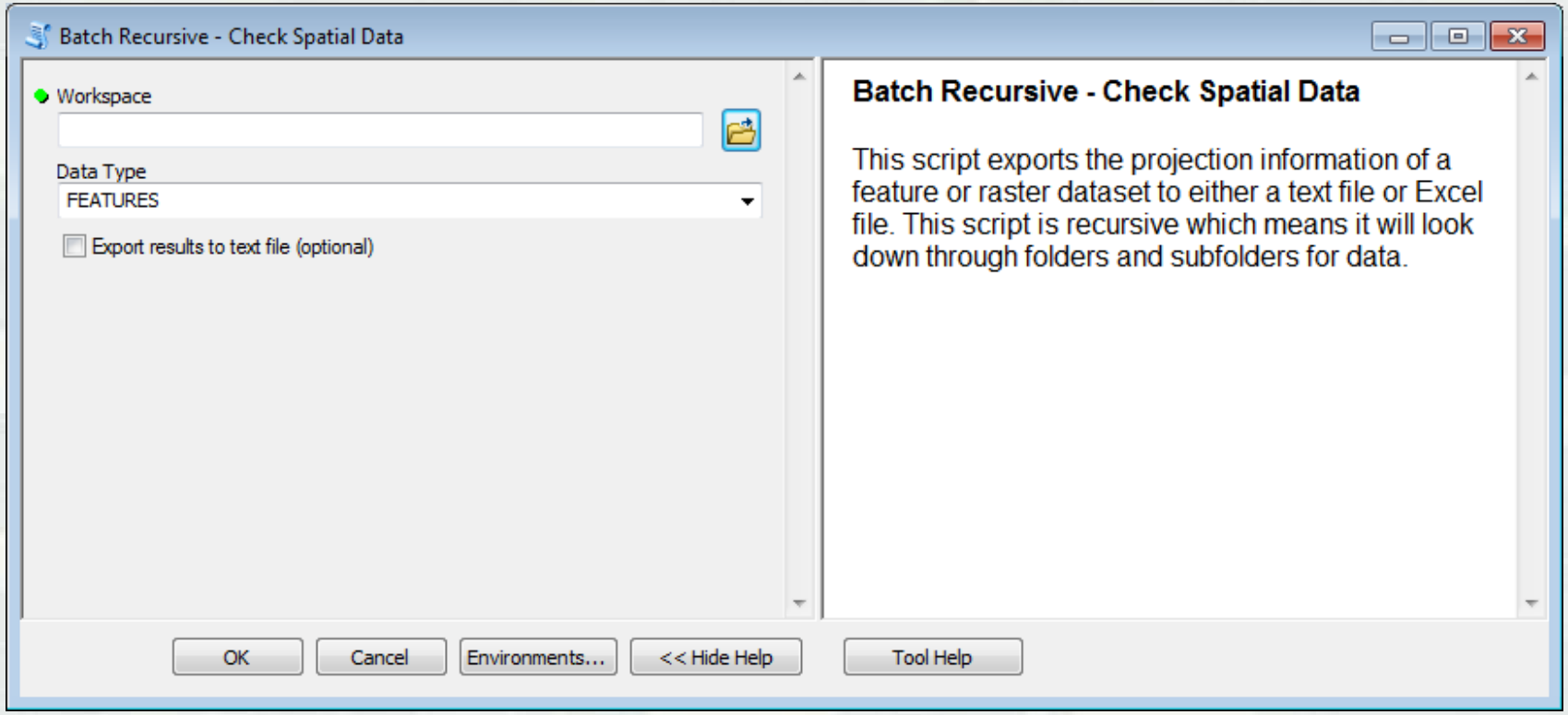
# Check for Geometry Errors



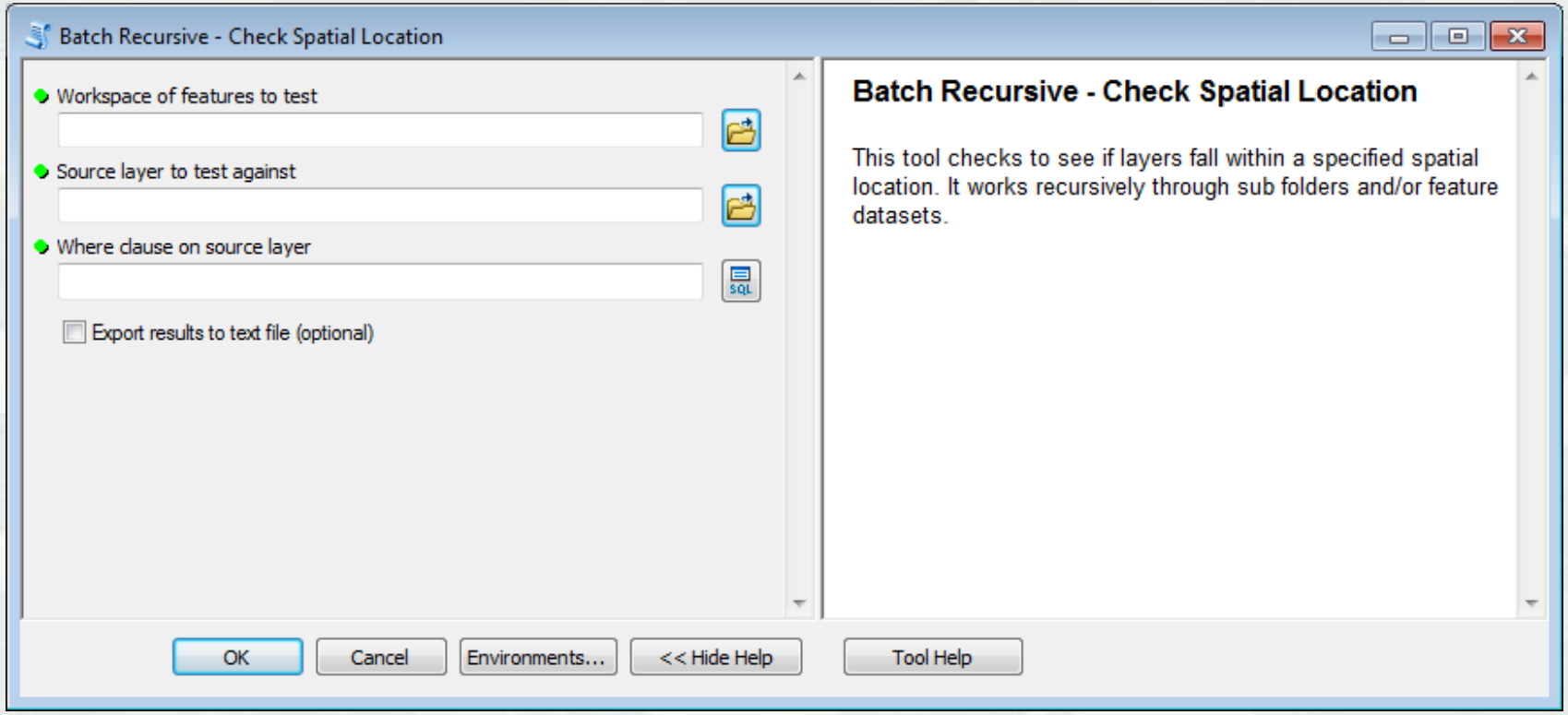
# Check for Missing Layers



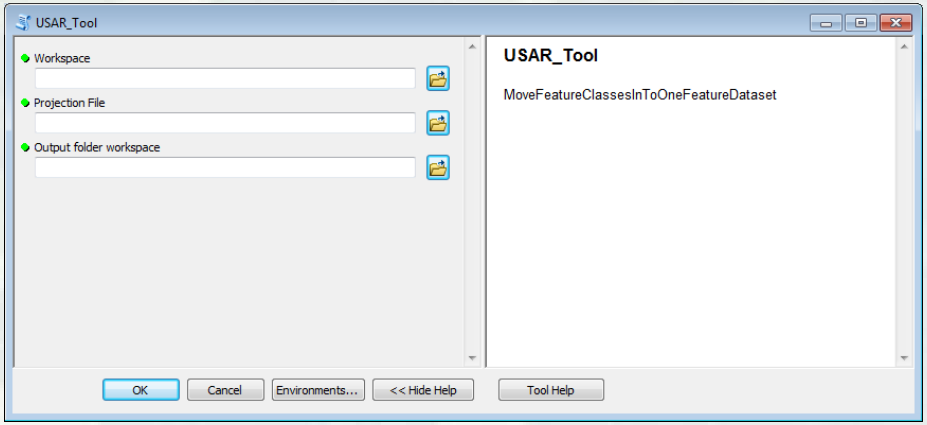
# Check Spatial Data



# Check Spatial Location



# USAR Tool




- Create FGDB
- Create feature dataset
- Create topology
- Add feature classes
- Load topology rules
- Validate topology

# Totals

- ~60,000 features
  - Manual review
    - 20% review
    - 333 hours
    - 8.5 weeks
  - Automated review
    - 90% review
    - 400 hours
    - 10 weeks

# Lessons Learned

- Data Reviewer
  - Has *most* of the functionality needed out of the box
  - Has some additional functionality that was not needed for this project
- USAR Python Tools
  - Topology
  - Simple
  - Can be targeted to *exactly* what is needed with customization

# Conclusion

- Automating the QA/QC process saved
  - Time
  - Money
  - Data Quality (most important)
    - More data reviewed (+/- 90%)
    - More time spent visually inspecting data
- Great alternative to Data Reviewer
- Tools available for other projects
  - "Toolified"

# Questions

## Python!

It's like having an extra person to QA/QC data

Jennifer M. Holland, GISP
Geographer
U.S. Army Corps of Engineers
Fort Worth District
jennifer.m.Holland@usace.army.mil