Data Interoperability
An Introduction

Bruce Harold - Esri Inc
Dale Lutz - Safe Software

What is the Data Interoperability Extension?

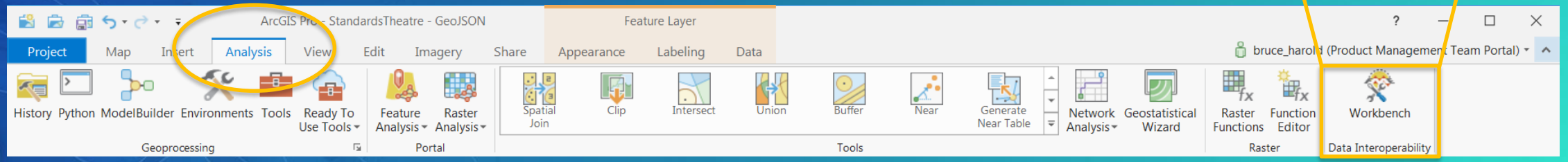
Moving data from app to app – “ETL” in industry jargon

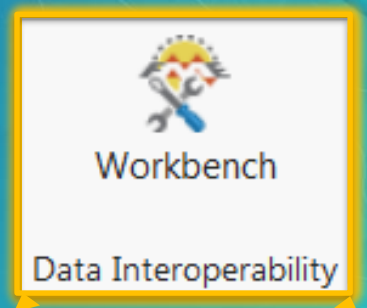
- If you are not the custodian you don’t control the data
  - Format or System of Record
  - Attribute Schema
  - Transport Technology
- Data providers probably don’t care much about your functional requirements
- Data itself certainly doesn’t care about you *at all*
- Data Interoperability empowers the ArcGIS data professional
  - Connectivity & Authentication
  - Data Transformation & Business Rules
  - Workflow Automation
  - Sharing

Where is the Data Interoperability Extension?

ArcGIS Pro, ArcMap, ArcCatalog

- Licensed like any extension
  - Named User (Pro), Concurrent, Single User
  - Evaluation licenses available
- You must *install* the extension
- Immediate authoring from the Analysis ribbon in Pro
  - We'll get to 'authoring' later
- ArcGIS 10.x - enable the extension
  - Tools visible in Pro 2.0





Why Should I use the Data Interoperability Extension?

- Traditionally three workflows – Direct Read, Quick Conversion, Spatial ETL
  - Web GIS needs more than direct format support e.g. feeds, authentication…
- Lets look at it this way

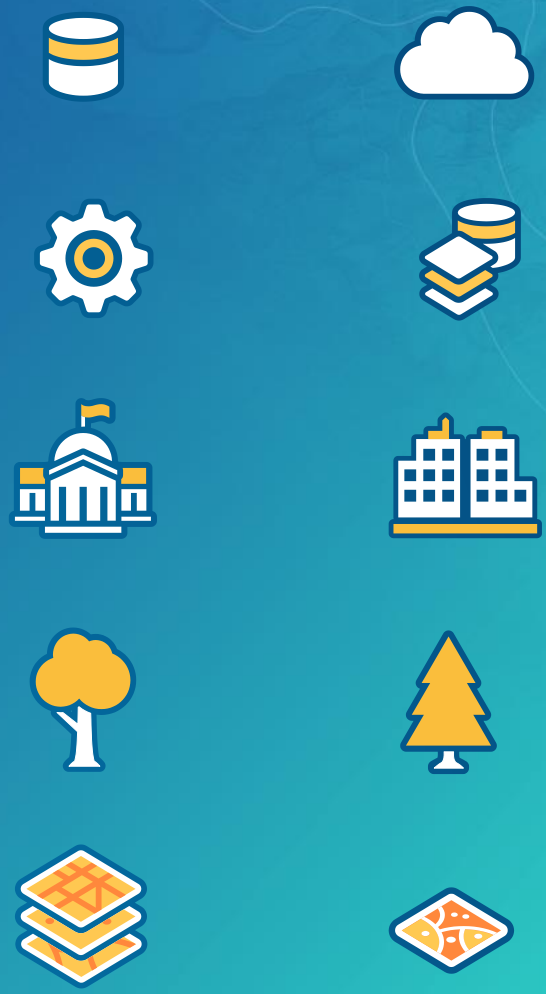
Capability	ArcGIS 10.x	ArcGIS Pro
Direct Read of Formats	✓	✗[1]
Quick Format Conversion	✓	✗[2]
Spatial ETL Tools	✓	✓
FMW Integration	✗	✓

1: Use Workbench; Interoperability Connections Coming 2: Use Workbench

How Do Others use the Data Interoperability Extension?

Top Use Cases

- Migration
  - Adoption of ArcGIS
- Integration
  - Projects with External Parties' Applications
- Publication
  - Formats Accessible to Target Audiences
- Validation
  - Feature Level Business Rule Enforcement
- Detection
  - Cross-Feature/Format/Schema Interaction
- Creation
  - Make Features from Anything (including Nothing)

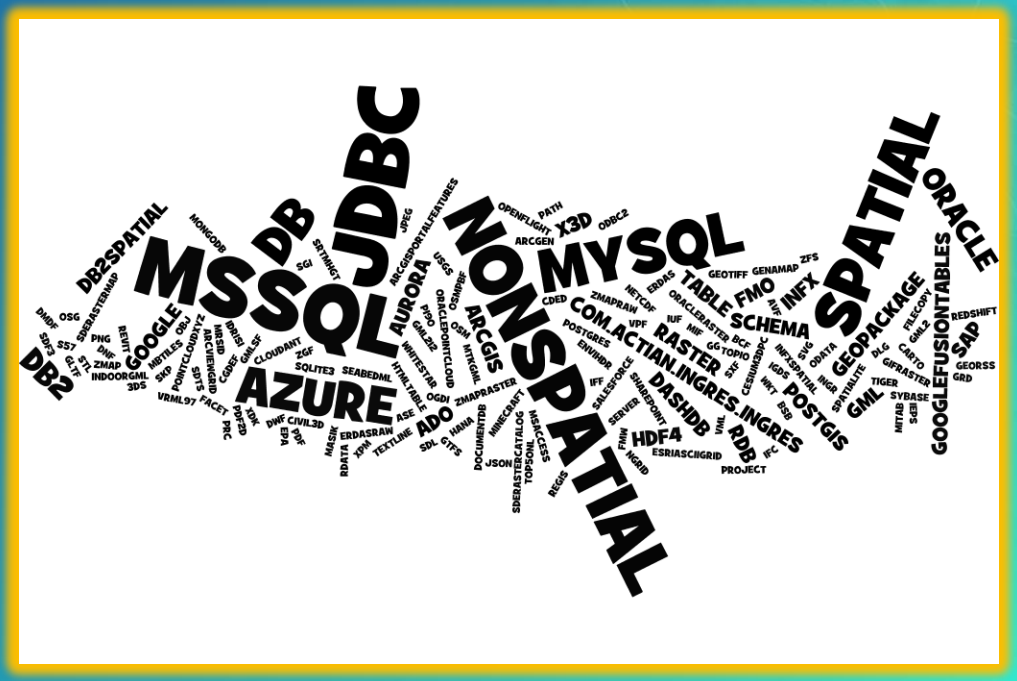


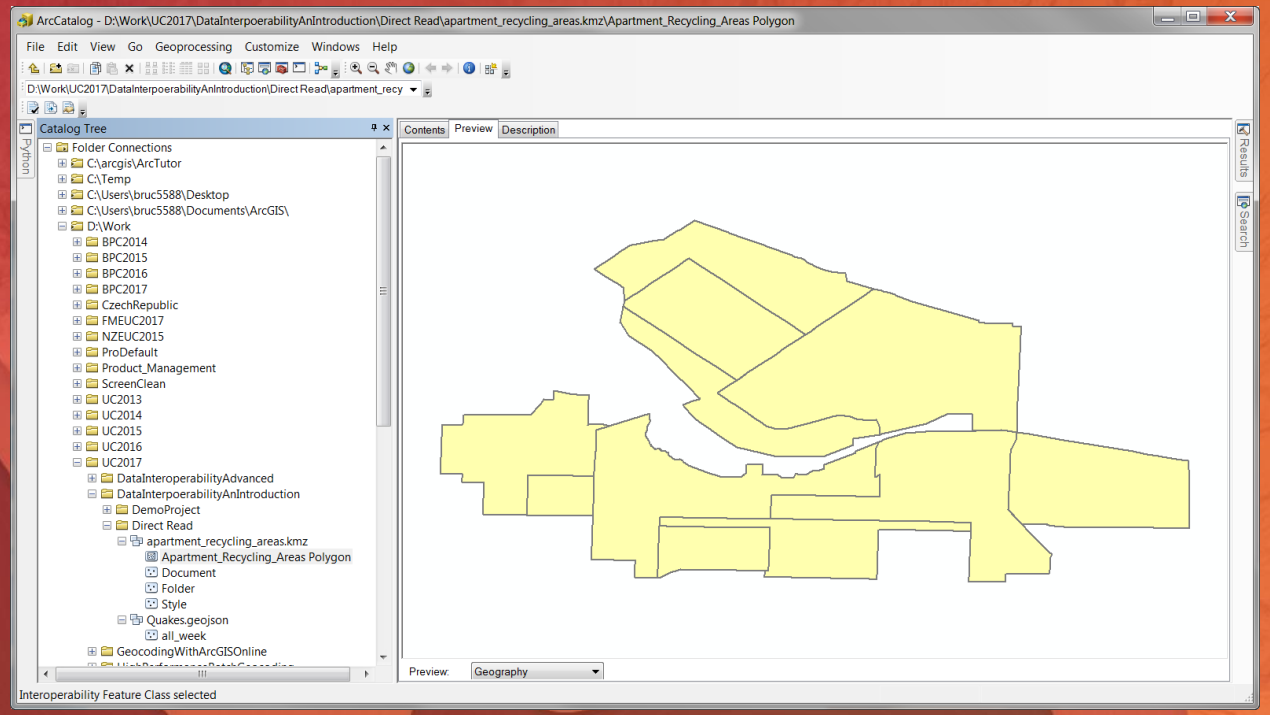
The Classic Use Cases

Direct Read, Quick Conversion, Spatial ETL

Direct Format Read

- ArcGIS is Geo-Relational
  - Feature classes are database tables with an ObjectID and a geometry column
- Some file formats can be automatically interpreted this way
  - Recognized by file extension (.gml, .geojson, .kml, .gpx, .xlsx, .osm, .xyz….)
  - Default behavior assumed
- Some formats can get help from a friend
  - VRT gateway files
- Many data sources are out of scope for direct read
  - Interpretation is required
  - Custom formats are an option (advanced topic)





Direct Read Demo

ArcCatalog

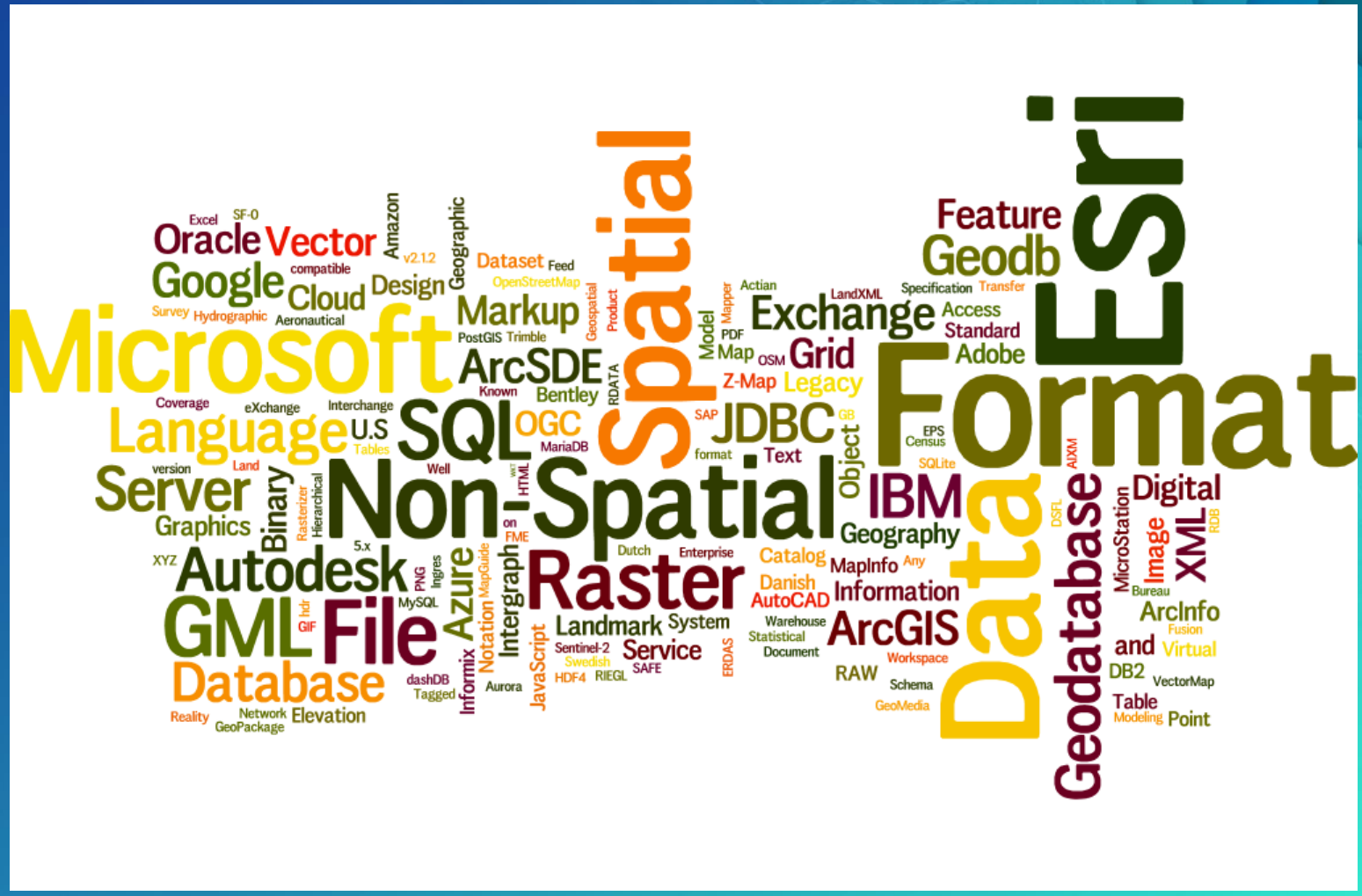
Quick Import and Quick Export

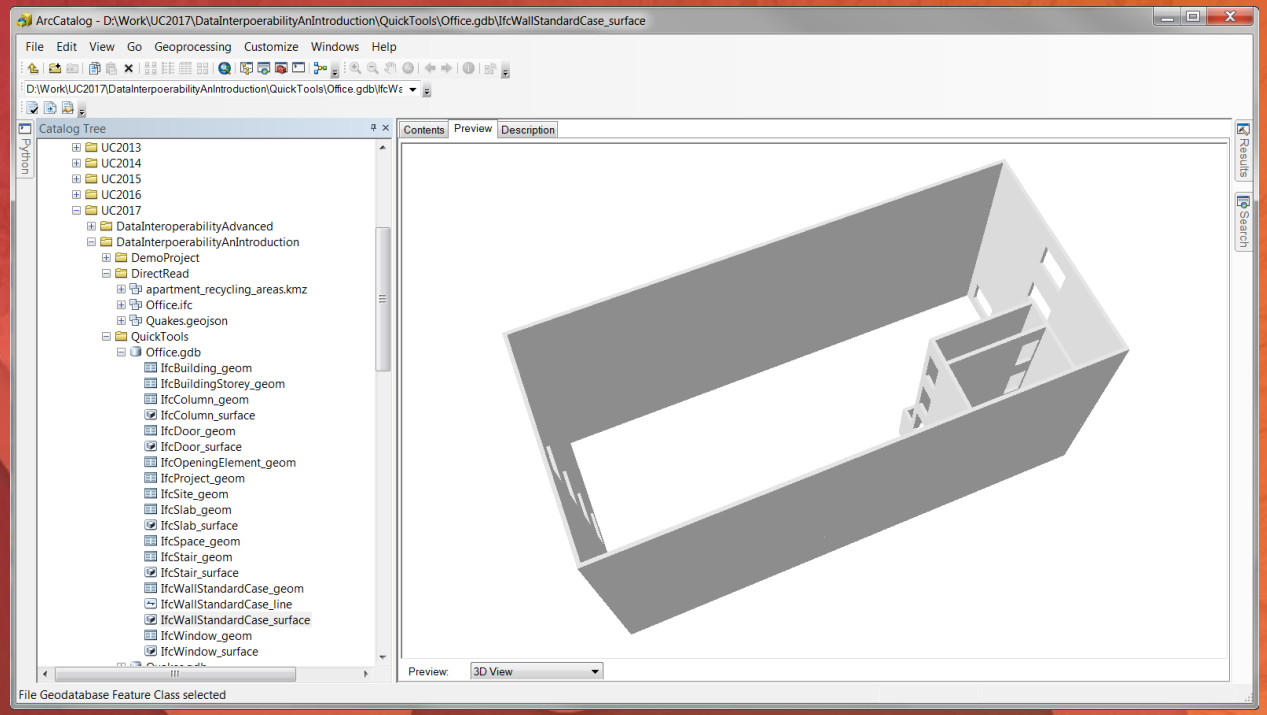
- Easy Button
- When you need data to participate in:
  - Editing
  - Geodatabase behavior
- Quick Import
  - Writes to Geodatabase
  - Create or Overwrite
  - Default Schema
  - Postprocessing often required
    - In Modelbuilder the Select Data tool accesses feature classes
    - Project or Define Projection
- Quick Export
  - Writes Any Format (277 available)
  - Default Schema

The Format Challenge

Dale, I'm sticking with 277

- Pick from 100's
  - On Disk
  - Database
  - Semantic (XML, JSON)
- Web as Filesystem
  - Box
  - DropBox
  - OneDrive
  - AWS S3
  - Google Drive





Quick Tools Demo

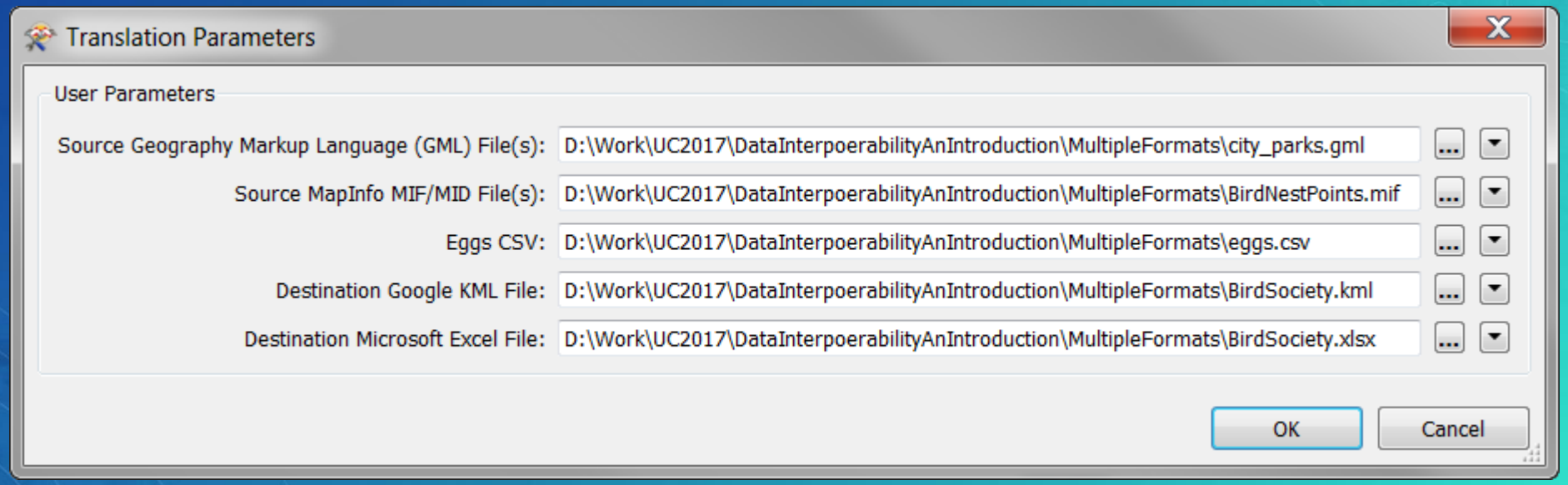
ArcCatalog

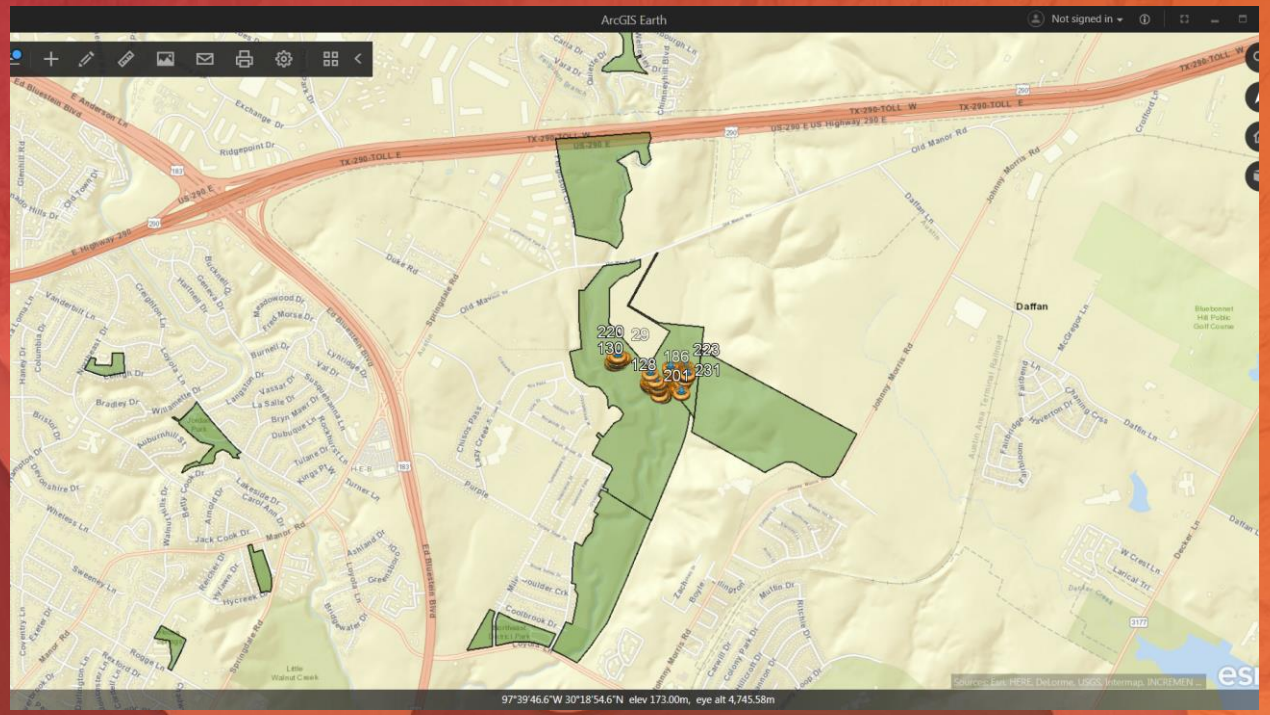
Multiple Formats

Combining is OK

- Read Many
- Join Many
- Spatial Interact Many
- Write Many

GML
MIF
CSV
KML
XLSX





Multiple Formats Demo

ArcCatalog
ArcGIS Earth

Spatial ETL – Moving Data With Transformation

Extract Transform Load

- Doesn't ArcGIS Already Read Most Useful Formats?
  - Arguably, but the details really matter
- Spreadsheets
- Character Separated Values
- Shapefile

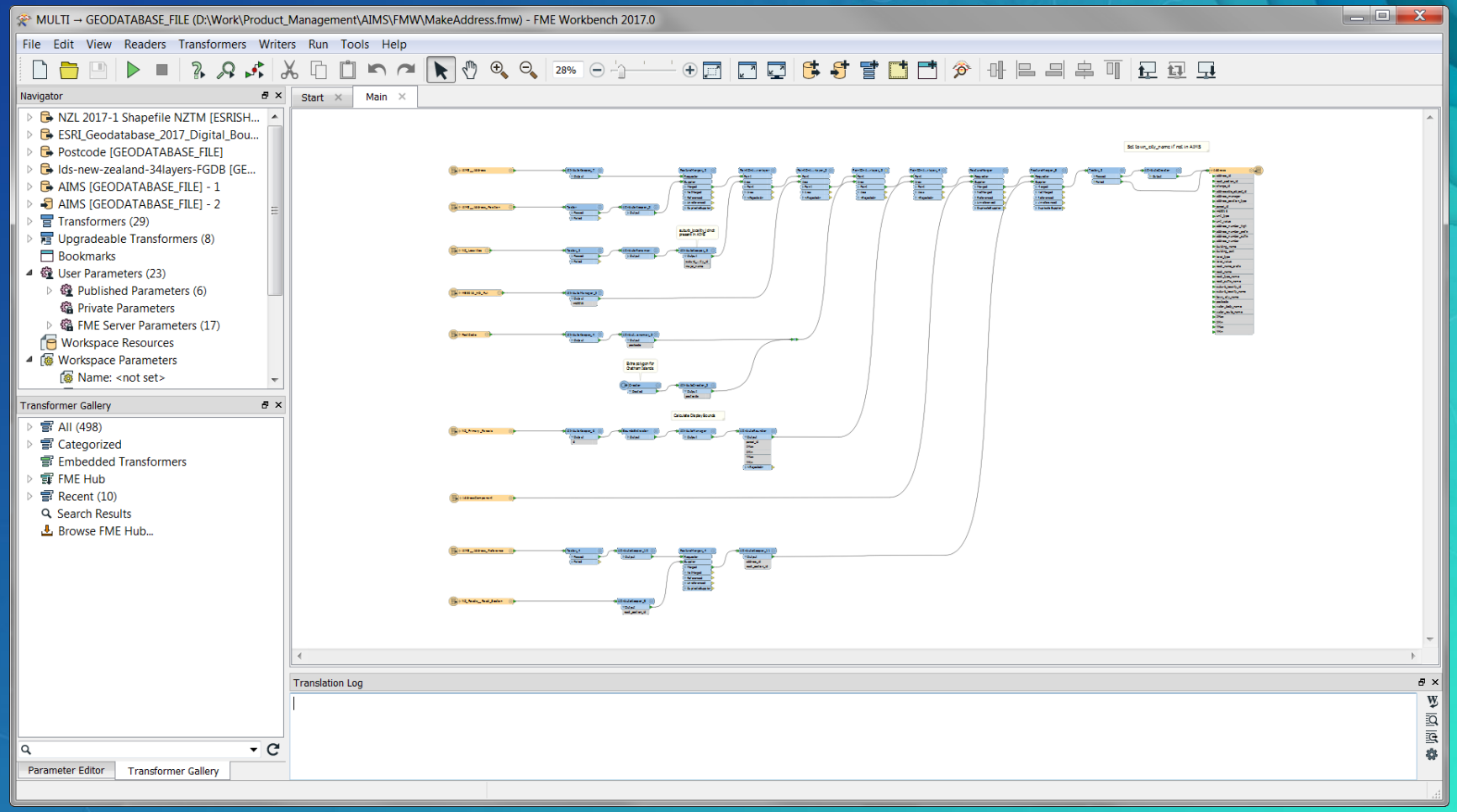
Issues

- Did you notice the non-spatial formats above?

Workbench

Graphical Programming Application

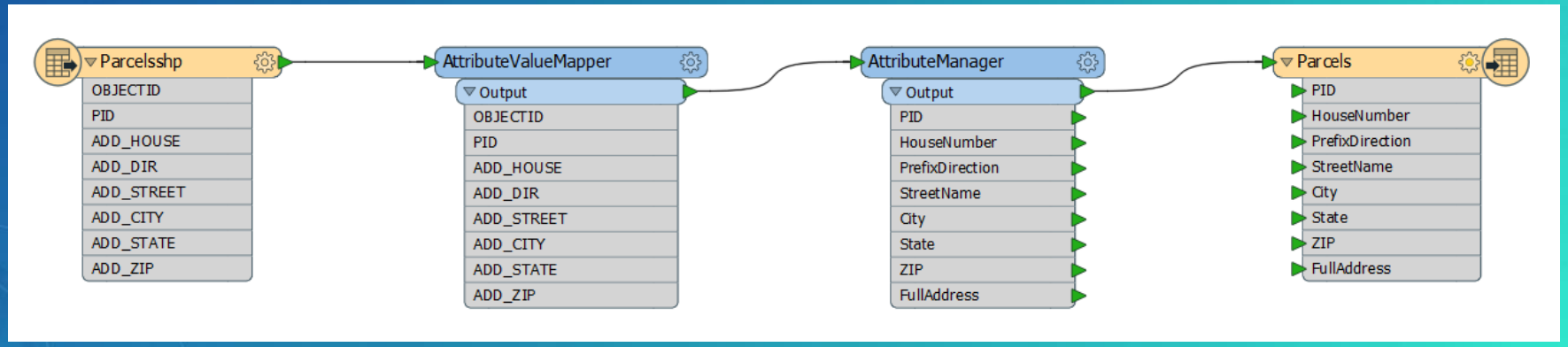
- Separate Process
- Menus, Controls
- Canvas
  - Readers
  - Transformers
  - Writers
- Navigator
- Transformer Gallery
- Log



What's a Transformer?

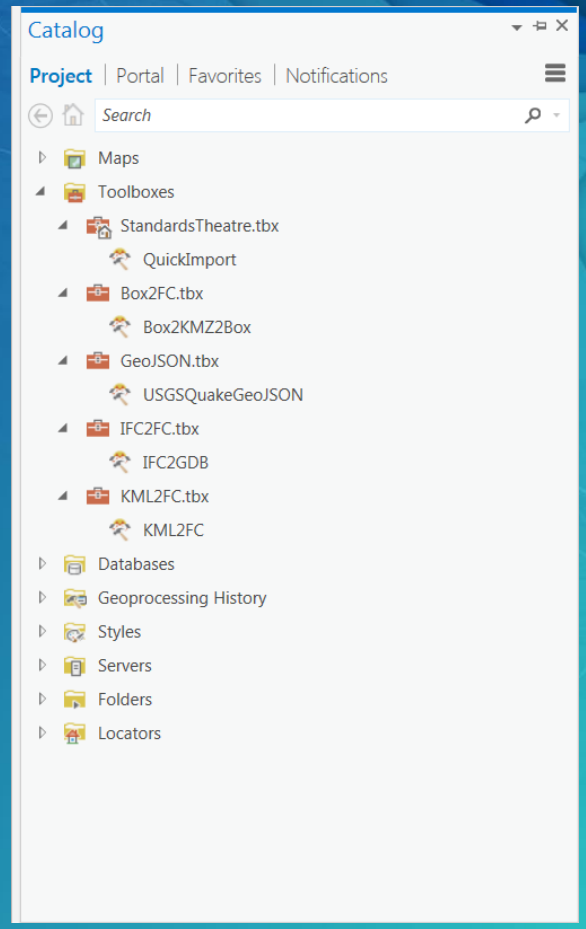
Extract Transform Load

- Transformers Are Analogous to a Geoprocessing Tool
- Have Input and Output Port(s)
- Typically Operate Per Feature
- May Operate on Attributes and/or Geometry
- Optionally Operate Per Group of Attribute Value
- May Operate With Parallel Multiprocessing
- Show Me



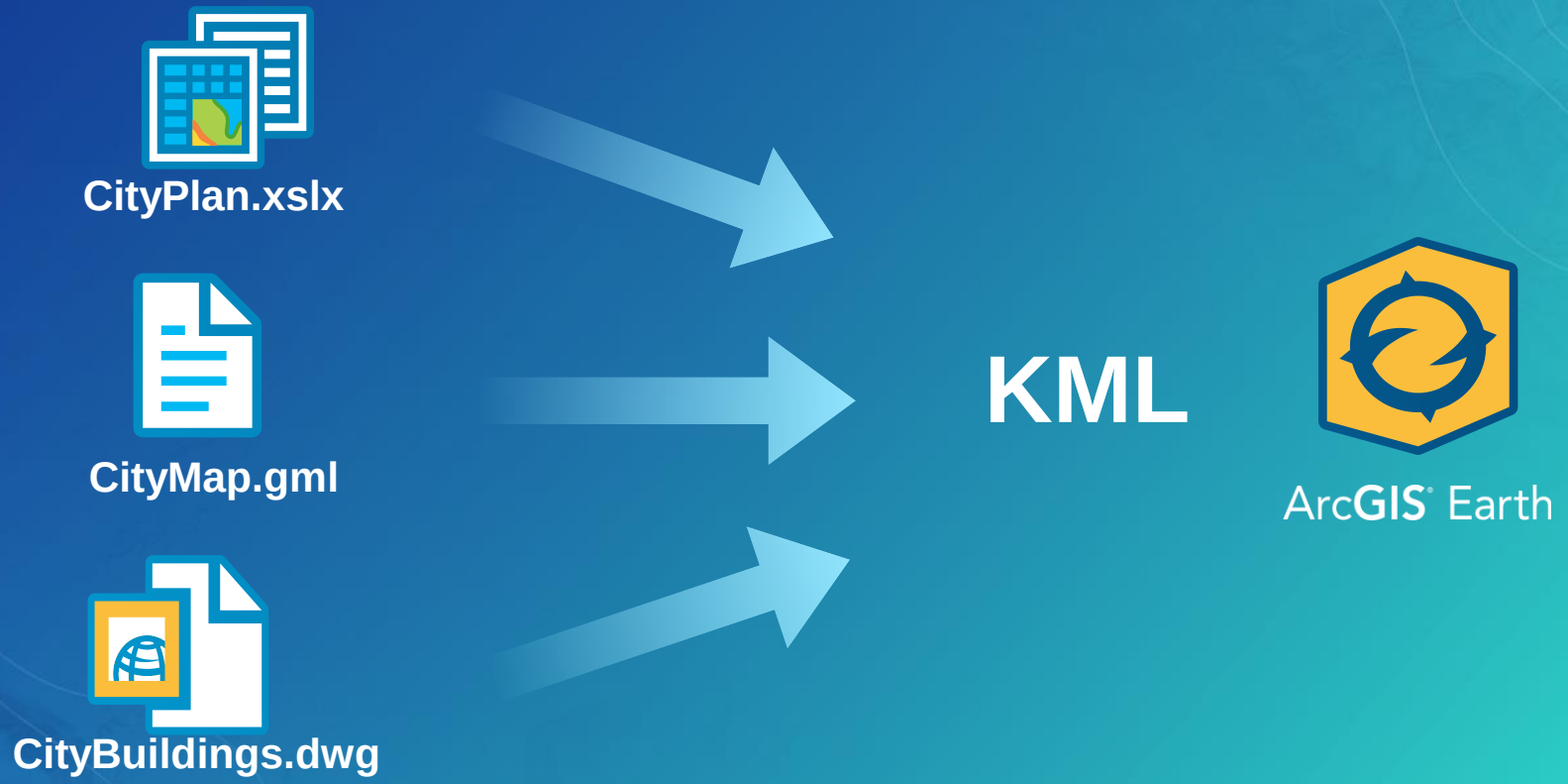
Common Workflow #1

- **Transformation Automation**
- Spatial ETL Tools ***are*** ArcGIS Geoprocessing Tools
  - Use in Modelbuilder
  - Use Spatial ETL Tools in standalone Python scripts
  - Use ArcPy inside Workbench
  - Publish to ArcGIS Enterprise (Windows, 10.x)
- ArcGIS Pro 2.0 interoperates with FME
  - Read and Write .fmw files
  - FME 2017 Engine
- Publish maps that use non-native formats copied to a Data Store folder



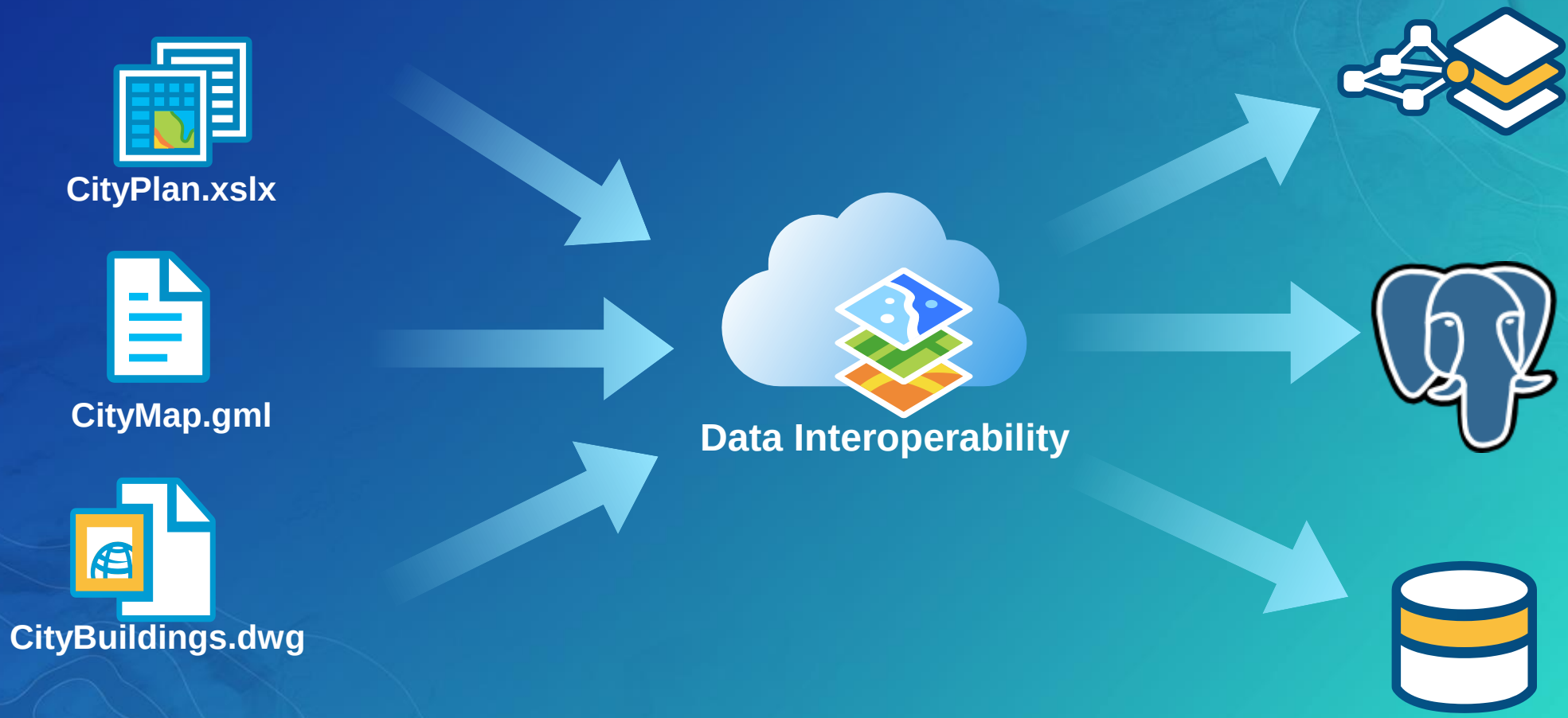
Common Workflow #2

- Integrate Multiple Formats



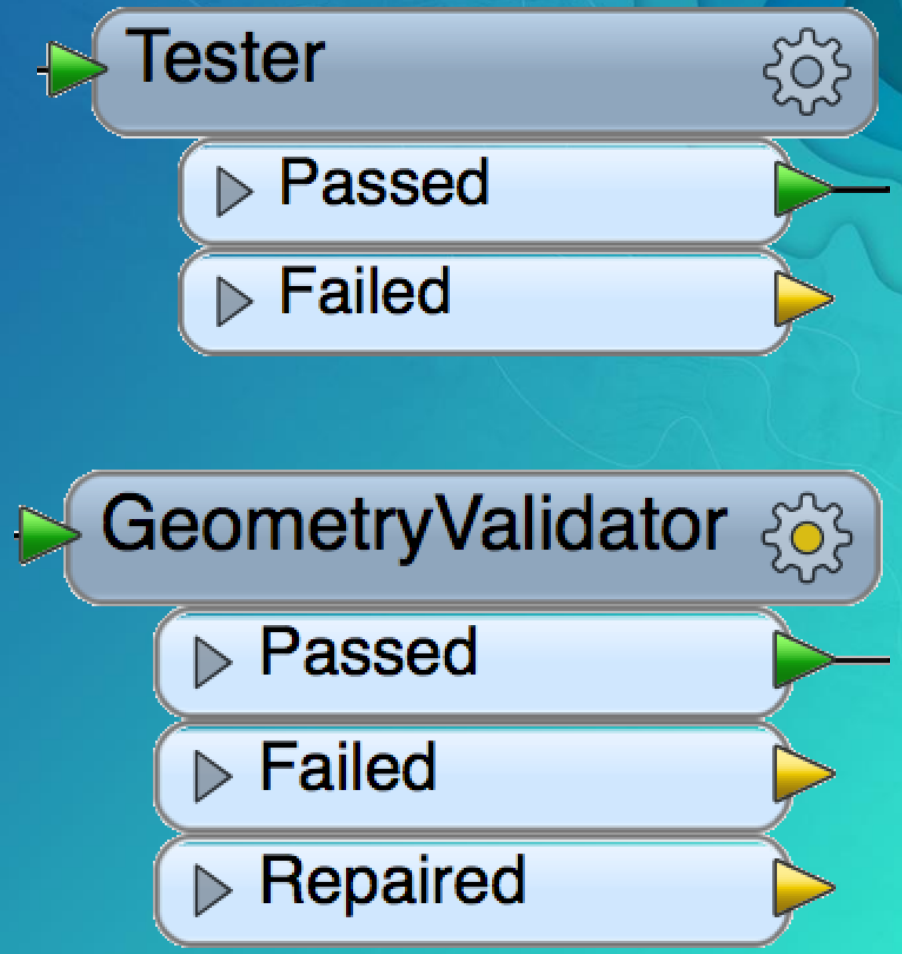
Common Workflow #3

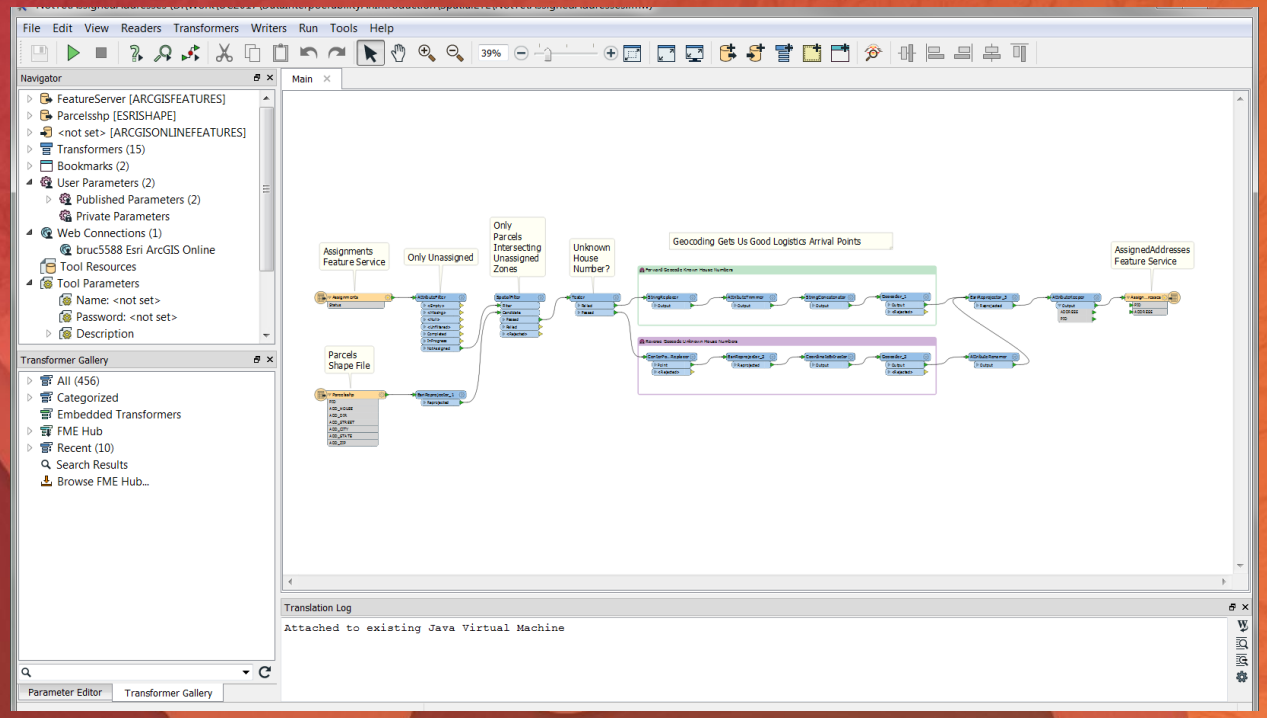
- Distribute/Migrate Data Between Systems



Common Workflow #4

- Validate Data
- Examples:
  - Consider the *previous* feature in some custom sort order
  - Build expressions using multiple attributes
  - Use statistical analysis to drive tests
  - Reach out to web services for real time checks
  - Cast geometry to another type temporarily
- Create reports of issues found





Spatial ETL Demos

ArcGIS Pro 2.0

Try it, Learn it

- 60-day Trial:
- http://www.esri.com/software/arcgis/extensions/datainteroperability/evaluate
- Courses – search ‘ETL’ in the Training Catalog:
  - Transforming Data Using Extract, Transform, and Load Processes
  - Controlling Data Translations Using Extract, Transform, and Load Processes

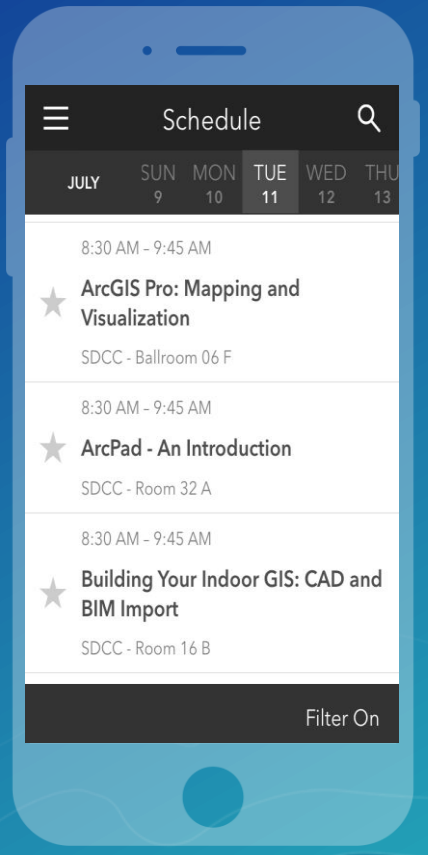
Question Time

- Bruce Harold bharold@esri.com
- Dale Lutz https://www.safe.com/live-chat/

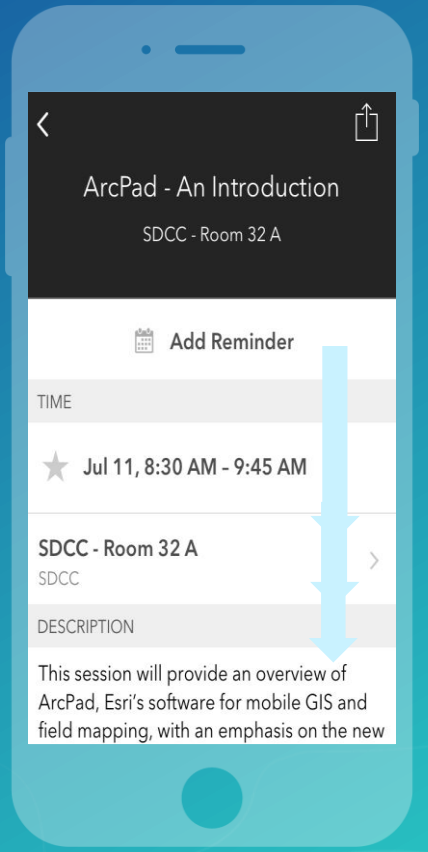
Please Take Our Survey on the **Esri Events App!**

Download the Esri Events app and find your event

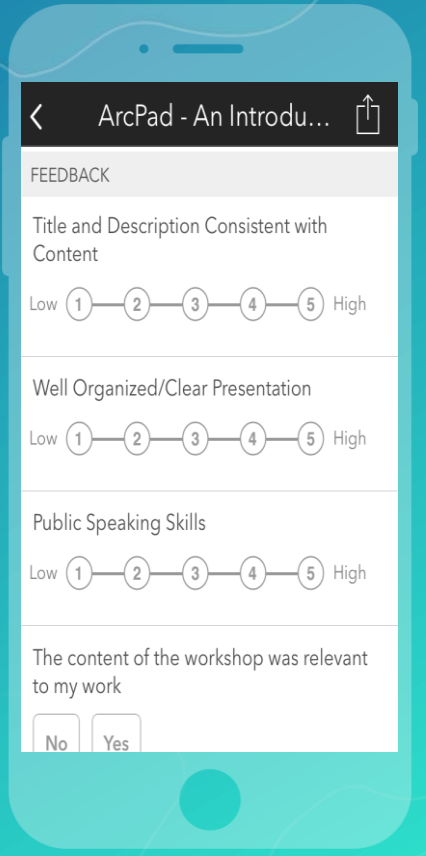
Select the session you attended



Scroll down to find the survey



Complete Answers and Select "Submit"



esri
THE
SCIENCE
OF
WHERE