Spatial Data Discovery: Assessing Enterprise Data Standards for ArcGIS Pro

LORI ODEGARD & JENNIFER STILL
INTEGRATED INFORMATICS INC
ABOUT US

• Integrated Informatics was founded in 2002 as an Engineering Company
• Specializing in Geographic Information Systems
  • Enterprise Geographic Information System (GIS)
  • Spatial Workflow Process and Analyses
  • Decision Support and Application Development
  • Spatial Data Management
  • Education and Support Services
• Three core offices: Houston, Calgary, and St. John’s
• Company works across multiple industries and fields – including Oil and Gas, Electric Utilities, Environmental, etc.
• Esri Business Partner (Silver Tier International) since 2006
1. The New MXD: Key Challenges
2. Case Study: The Migration Process
3. Going Pro with Our Data Management Standards
THE NEW MXD

KEY CHALLENGES
ARCGIS 101

ARCMAP

• ArcGIS Map Document (.mxd)

• Licensing is serviced through ArcGIS Administrator.

• Spatial data is handled in a 2D environment.

• A single MXD can contain multiple data frames – but not intended for housing multiple layouts.
ARCGIS 101

ARCGIS PRO

• ArcGIS Pro Map Project (.aprx)
• Licensing is handled through a user’s ArcGIS Online account.
• Allows users to work in both 2D and 3D environments.
• A single APRX file can contain multiple maps and layouts.
• ArcGIS Pro’s version of the **MXD** is the **APRX**.
• Similar concept, but different approach...
  • ArcGIS Pro can create an APRX file based on ArcGIS Map Documents (.mxd), scenes (.sxd), and globes (.3DD).
    • The new file cannot be opened within ArcMap, and it cannot be converted back to its original format.
• Once you are in Pro, you are **full** Pro.
Creating a new ArcGIS Pro Map Project (.aprx) results in two things ArcMap users are not always aware of...

1. An APRX file must be saved prior to opening – That is, no more Untitled.mxd files.
2. An empty file geodatabase is created alongside each new APRX file by default.*

*This can be changed within ArcGIS Pro’s Options > Application settings once you get going.

The impact of this is that unused APRX files and file geodatabases may live on the network unless there is a plan of attack upfront – or the Default Geodatabase settings are altered.

Both features are slated to be modified in the ArcGIS Pro Product Plan. No word on when.
Our biggest hurdles?
- **Absolute/Relative Paths**
- **Metadata**

The way ArcGIS Pro works with **path settings**...
- Connections on the **same drive** as the APRX file are stored as **relative paths**.
- Connections on **different drives** are stored as **absolute paths**.
- There **no way to toggle** between these path settings like in ArcMap.

The way ArcGIS Pro works with **metadata**...
- Metadata functionality within this platform is not as robust as ArcMap.
- **Example** – Importing, exporting, or updating (i.e., update FGDC metadata to ArcGIS metadata) metadata is currently a no-go. This must be done beforehand within ArcMap.

Why does this matter during migration?
- Know **what** data belongs **where**.
- Know **what we have** and **what we need**... in this case, as far as metadata goes.
- Modify path settings and/or metadata **prior** to conversion.
THE MIGRATION PROCESS

CASE STUDY
Migrating to a new platform will essentially revolve around **Spatial Data Discovery**, or *the art of knowing what we have and how to use it.*

It is a time of *trial and error* as well as the perfect chance to better understand **what we are doing** and **what we should be doing instead**.
THE MIGRATION PROCESS

DISCOVER DATA.
Identify the spatial data on the network – including its health (i.e., age, brokenness).

CLEAN UP.
Tidy up data – especially unhealthy files. Resolve any data issues between platforms.

CONVERT.
Move ArcGIS Map Documents (.mxd) to their ArcGIS Pro Map Project (.aprx) equivalents.
PHASE 1 | DISCOVER DATA

WHAT TO LOOK FOR?

• Key areas evaluated include:
  1. LOCATION
  2. OWNERSHIP
  3. RELEVANCY
  4. STATUS

BUT WHY?

• Inventorying enterprise data allows us to answer questions like...
  • What spatial data is stored on the network?
  • Who owns it?
  • What is it connected to?
  • Are there any broken layers?
  • Is it potentially outdated material?
  • Who can we place blame on?
LOCATION, LOCATION LOCATION

• One of the most important pieces of information = Knowing *where to find data*.

• The **full path** of the file is essential to finding, sharing, and evaluating the data at hand – and determining its **health**.
  • This applies to Layer Files, File Geodatabase, ArcGIS Map Documents, etc.

• **Location** also helps us to determine the data type, discovering broken links, fixing said broken links, and moving data across the network – not just between Esri applications.
This information also allowed us to discern hot spots on the network to help determine which areas to focus on first – aka directories with poor location health.

By analyzing location, we determined that C:\testdata – especially C:\testdata\rnr\input – held a high number of broken Feature Layers and Raster Datasets.

Prior to migrating, these needed to be cleaned up (i.e., correct paths found, broken sources fixed, old data removed) and noted for addition to a new maintenance plan.
OWNERSHIP

• Finding users who created and/or maintained data helps to grasp what actions to take for the **now** and the **future**.

• **Ownership**, or *knowing the user account with which data is associated*, is useful for a number of reasons...
  
  • **Hints at Maintenance** – Data attributed to a specific account (such as a retired employee or one who’s since changed roles) provides insights into whether its contents have been properly maintained.
    
    • For example...Has Gerta been gone for a while? The files she maintained – or did not maintain – may need refreshing.

  • **Point of Contact** – In instances where more information is needed, reach out to this user or their team.
OWNERSHIP

• Further analysis of spatial data based on ownership cross-referenced with location health shows those users responsible for broken datasets.

• How is this information used outside of determining maintenance and contacts?
  • Evaluate User Performance – Identifying these accounts allows management to better discern an employee’s use of the network and decide if they may benefit from more defined data management guidelines.
The whens of change allow us to see how well files have been maintained and if they could be out of date.

Here, relevancy refers to temporal-based details for each dataset.

- When was it created?
- When was it last modified?
- When was it last accessed?

Being aware of relevancy provides insight into how regularly data is used, how frequently it is maintained, or if it is altogether irrelevant.
• Analyzing this information can also help answer...
  • Of files with the same name and content – what has been used recently?
  • How often do we need to perform a scan on this section of the network?
  • What location on the network houses the most unused data (i.e., not touched within five years) and just how many files are there after all?

*
STATUS – SELF AWARENESS

• Seeing when we last checked on data allowed us to get a better idea of what could be lurking around the corner... Where be dragons?

• Our information about location, ownership, and relevancy are just a snapshot in time. These things can and do change.

• As important as it is to know when data was created, it is equally as important to know the status, or know when we last checked – therefore, when we need to run a refresh.
  - This information alone helps us to better understand our own data practices and evaluate current/future standards – for both ArcMap and ArcGIS Pro.

• What are reasons to perform a data check-up?
  - Timeframe – The date of the last status is older than the check-in date or timeframe requested by IT, manager, big boss, etc.
  - Activity – We know that data has been added, modified, or removed – or heaven forbid, crashed – since the last scan.
PHASE 2 | CLEAN UP

• What were the most important pieces of wisdom gained from inventory?
  • Did we find broken layers?
  • Did we come across files that have aged out?
  • Too many Untitled.mxd files floating around?

• Migration Wisdom: Keep in mind what you found and what you want when cleaning up any size network.

• This is the perfect opportunity to begin evaluating current enterprise data standards – both official and unofficial – and think toward the future.
PHASE 2 | CLEAN UP

- Take this time to address issues that could compromise spatial data or negatively impact conversion...
  
  1. Nix duplicate data.
  2. Fix broken data sources in ArcGIS Map Documents.
  3. Decide if those vaguely named files are worth keeping around.
  4. Determine if certain files – or whole drives – should be archived.
  5. Ensure naming schemes and formats for all files are up to date with standards.
  6. Watch out for metadata updates that require ArcMap to resolve – such as importing/exporting or updating FGDC metadata to ArcGIS metadata for use in ArcGIS Pro.
  7. Toggle relative and/or absolute paths to heart’s content.
PHASE 3 | CONVERT

- ArcGIS Pro’s **Import Map** tool can be used to convert an ArcGIS Map Document (.mxd) to an ArcGIS Pro Map Project (.aprx), keeping in mind...
  - **Import Map** only converts one file at a time.
  - Be mindful of where you save these new files. *Are they being saved to the default location or one of your choosing?*
  - Data frames within an MXD are saved as maps within ArcGIS Pro. **Multiple data frames = multiple maps.**
- **Are you moving forward?** Educate yourself and your team on both the shiny new features – and limitations – awaiting you in this new platform.
DEVELOP NEW* STANDARDS

*BONUS IF THEY’RE BETTER
NEW STANDARDS FOR FUTURE YOU

• Once we have made it through a successful migration, take this time to reevaluate data standards.
  • This refers to everything from where and how data is stored, who has access to what, how often the system is scanned for any issues with data and/or to record details like status, etc.

• The Starter Kit for Developing New Standards
  1. Make note of the problems encountered during the Migration Process – even better, include how they were fixed.
  2. Schedule regular data check-ups to examine inventory, brokenness, and usage.
  3. Keep track of what has been migrated and what has not.
For more info…

LORI ODEGARD
INTEGRATED INFORMATICS INC.
ljodegard@integrated-informatics.com

JENNIFER STILL
INTEGRATED INFORMATICS INC.
jlstill@integrated-informatics.com

www.integrated-informatics.com
www.marcostud.io/solutions-arcgis-migration