ArcGIS Enterprise: Data Storage Strategies

Adam Ziegler
Agenda

- What is a data strategy and why would I need one?
- Storage options in ArcGIS Enterprise
- Technical architecture of data in ArcGIS Enterprise
- Example data strategies
- Related topics

Note as the software evolves, so does this information!
What is your role?

- DBA?
- System Architect?
- Analyst?
- Developer?
- Executive?
- All of the above?
ArcGIS Enterprise: Data Storage Strategies

Your organization’s plan for achieving its goals.
What is a data strategy?

A comprehensive plan for how your organization will *store, access, and manage* your GIS data.

A data strategy is *feasible, economical, tailored to your workflows and users, and evolves as necessary.*
What is a data strategy?

A data strategy can take on many forms, from prescriptive manuals, to checklists, to general user guidelines and handbooks.

Whatever form it takes, a data strategy as a integral part of GIS.
Why is it important to have a strategy for data?
You are working with more data than ever before:

- Imagery
- 3D
- Raster
- Real-time
- Big data
- Urban
- Indoor
- Field
- Demographic
- Third party
- Living Atlas
- Unstructured
- Vector & tabular
- Utility networks
- Drone
- Cloud

ArcGIS Enterprise supports your data workflows
A data strategy makes the best use out of your data.

Flexible
Gives you room as data, workflows, and your user base changes and grows.

Accessible
Enables your users to access data when and where they need it (mobile, desktop, web / via direct connections, services).

Functional
Provides the right capabilities and functionality to successfully execute your workflows across your organization.

Interoperable
Integrates with other parts of the ArcGIS platform and technology as needed.
It also gives you a strong foundation to bring on new challenges, workflows, and innovation.
Options for data have evolved throughout the years…

- **1990s**: Coverages
- **1990s**: Shapefiles
- **2014**: SDE (enterprise geodatabase)
- **2014**: Personal geodatabase, file geodatabase
- **2014**: ArcGIS Data Store
### Available options in ArcGIS Enterprise today

<table>
<thead>
<tr>
<th>Geodatabases</th>
<th>Folders &amp; files</th>
<th>ArcGIS Data Store</th>
<th>Cloud storage</th>
<th>Big data storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Enable on top of commercial RDBMS</td>
<td>- Local or network, E.g. file geodatabase</td>
<td>- Storage included with ArcGIS Enterprise</td>
<td>- Amazon, Azure integrations</td>
<td>- Hadoop, Hive</td>
</tr>
<tr>
<td>- Large scale, multi-user, authoritative data</td>
<td>- Storage for different file formats (csv, shp, etc)</td>
<td>- 3 different flavors</td>
<td>- Store map and image caches</td>
<td>- Input and output of big data analysis (vector, tabular)</td>
</tr>
<tr>
<td>- Spatial and attribute integrity across datasets</td>
<td>- Powers hosted data</td>
<td>- Feature data, 3D scenes, high volume real-time and big data</td>
<td>- Optional output of raster analytics</td>
<td></td>
</tr>
<tr>
<td>- Versioning, archiving</td>
<td>- Feature data, 3D scenes, high volume real-time and big data</td>
<td>- Input and output of big data analysis (vector, tabular)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
With all of these options, it can seem daunting at first.

(That’s probably why you are here!)
We can break down a lot of these concepts using two terms:

*User managed*

*ArcGIS managed*
What does this mean?

**User managed:**
- Data storage that you manage independently
- You provision, scale, tune, the underlying database
- You make it accessible to ArcGIS Enterprise by registering it

**ArcGIS managed:**
- Data storage included with ArcGIS Enterprise
- You install the software component as part of your ArcGIS Enterprise deployment

Direct management of the underlying storage

Management through ArcGIS interfaces and APIs
Relationship to data

- **User managed** (enterprise geodatabase)
  - The system references the data in place

- **ArcGIS managed** (hosted, ArcGIS Data Store)
  - The system hosts (stores) the data for you

This defines the term "hosted."
Storage types

**User managed**
(enterprise geodatabase)

- Enterprise geodatabase
- Cloud storage
- File shares
- Big data storage

**ArcGIS managed**
(hosted, ArcGIS Data Store)

- ArcGIS Data Store
  - Relational
  - Tile Cache
  - Spatiotemporal
Publishing and access

**User managed**
(enterprise geodatabase)

- Data doesn’t move: referencing in place
- Accessed through database connections, REST services and items in portal
- Delete the service, data remains

**ArcGIS managed**
(hosted, ArcGIS Data Store)

- Copy data or publish directly in your portal
- Accessed through REST and items in portal
- The data *is* the service
Use cases

**User managed**  
(enterprise geodatabase)

- Authoritative system of record
- Utility networks and parcels
- Need strict spatial and attribute quality
- Support for multi-user versioning workflows
- Comprehensive, relational database

**ArcGIS managed**  
(hosted, ArcGIS Data Store)

- Often used for self-service portal workflows
- Good alternative for storing file-based data
- Some advanced options (domains, views, etc)
- Relatively isolated, standalone datasets
Whitepaper: Data in ArcGIS

<table>
<thead>
<tr>
<th>Feature</th>
<th>User managed (enterprise geodatabase)</th>
<th>ArcGIS managed (hosted data in the ArcGIS Data Store)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Database Technology</td>
<td>Option of RDBMS (Oracle, SQL Server, PostgreSQL, Informix, DB2, SAP HANA, etc.)</td>
<td>The database is chosen by ArcGIS; not a bring-your-own-database nor a general-purpose database.</td>
</tr>
<tr>
<td>User Access</td>
<td>Multiuser, with both direct connection from desktop applications and via KSS-based map and feature services.</td>
<td>Multiuser, solely via REST-based map and feature services.</td>
</tr>
<tr>
<td>Rendering</td>
<td>Map services can render the output server-side using map image layers. Both map and feature services can render client-side via feature layers. Map image layers support advanced renderers and cartographic options not available in feature layers.</td>
<td>Hosted feature layers only support client-side rendering with out-of-the-box symbol set and cartographic options.</td>
</tr>
<tr>
<td>Versioning and archiving</td>
<td>Supports both traditional versioning and branch versioning. Archiving historical snapshots is supported.</td>
<td>Versioned editing is not supported; the last edit submitted is stored for the feature layer. No archiving.</td>
</tr>
<tr>
<td>Topology</td>
<td>Topology rules can be created and enforced.</td>
<td>Topology rules not supported.</td>
</tr>
<tr>
<td>Scalability</td>
<td>The underlying RDBMS can be scaled to support a large number of users, editors, and data using native RDBMS features and scaling of the server hosting the database. Individual map and feature services that reference user-managed data sources can have a substantial memory footprint on the server. Typical configurations can support in the range of hundreds of services.</td>
<td>The relational data store type of ArcGIS Data Store can be scaled vertically, adding more capacity &amp; resources to a single machine. A secondary machine can be added as passive backup and a failover host. Individual hosted feature layers have very low memory footprint on the server. Hundreds to thousands of hosted feature layers can be supported in a standard configuration.</td>
</tr>
<tr>
<td>Use as managed database for hosting server</td>
<td>Not supported, however the geodatabase can be registered as an additional data store, or the managed database on a non-hosting server.</td>
<td>Supported – only the relational ArcGIS Data Store can be registered as the managed database for the hosting server site.</td>
</tr>
</tbody>
</table>
The ArcGIS Data Store is not intended to replace your enterprise geodatabase.

It is a complement to your existing storage options and can be used in conjunction with them.

As part of your data strategy, consider what workflows your organization needs and where best to store and access your data.
Architecture
ArcGIS Enterprise architecture

Enterprise geodatabases, folders, cloud storage

ArcGIS Enterprise portal

ArcGIS Server

ArcGIS Data Store
ArcGIS Managed

User managed
Workflow: Publish by reference from ArcGIS Pro

ArcGIS Enterprise portal
*Feature layer (item) created*

ArcGIS Server
*Feature service created*

ArcGIS Data Store

Enterprise geodatabases, folders
*Data remains here*
Workflow: Publish as a copy from ArcGIS Pro

- **ArcGIS Enterprise portal**
  - Feature layer (item) created

- **ArcGIS Server**
  - Feature service created

- **ArcGIS Data Store**
  - Copy of data stored here

- **Enterprise geodatabases, folders**
  - Data copied from here
Workflow: Directly upload a CSV to your portal

ArcGIS Enterprise portal
*Feature layer (item) created*

ArcGIS Server
*Feature service created*

ArcGIS Data Store
*Data stored here*
Example strategy
A data strategy can take on many forms

**ArcGIS Online**
- Public content
- Open data
- Non-employees (volunteers, contractors)
- Collaborated data from ArcGIS Enterprise for field operations

**ArcGIS Enterprise**

Enterprise geodatabase:
- Continuous, multi-user datasets

Hosted data:
- Innovation, projects, PoC, learning
- Replacement for file geodatabases
A data strategy can take on many forms

<table>
<thead>
<tr>
<th>Business Objectives</th>
<th>Personnel</th>
<th>Metrics</th>
<th>Applications and Tools</th>
</tr>
</thead>
</table>
| Improve quality of data captured in the field            | Field scientists, Geologists, GIS Professionals | • Reduce space for taking maps offline  
• Validate data at the time of collection  
• Ensure that collected data is punctually provided to QA-tier users | • ArcGIS Enterprise  
• ArcGIS Pro  
• Collector  
• Offline map areas  
• Domains                                                     |
Where do I start?
Think about your end goal first
And then work backwards

Start here:
“\textit{I want to make a cake.}”

Then put the pieces in place to get there:
\textit{Butter, flour, a mixing bowl, an oven, ...}
Working from your end goal backwards
What do you want to do?

I want to maintain an accurate inventory of parcels in my city.

I need to be able to have many editors working at once and to track changes.

My users have ArcGIS Pro licenses and we’re using ArcGIS Enterprise.

I’ll use an enterprise geodatabase and use branch versioning off of web services.
Example questions to jumpstart your strategy:

**Collected**
- How will we capture your data?

**Edited**
- Who will need to make changes?

**Kept accurate**
- What type of quality assurance is needed?

**Accessed**
- Who needs to be able to find and use it?

**Scaled**
- Will our data grow? Will our userbase grow?

**Used**
- What is the function of our data? Where will it be used?
Related topics
Related topics

- **Distributed collaboration**
  - *Sharing data between ArcGIS Enterprise environments and with ArcGIS Online*

- **ArcGIS Enterprise sites**
  - *Tailored landing pages for your users to discover and interact with your GIS*

- **Bulk publishing**
  - *A new option for publishing all of your enterprise geodatabase data as web services*
Please Share Your Feedback in the App

Download the Esri Events app and find your event

Select the session you attended

Scroll down to “Survey”

Log in to access the survey

Complete the survey and select “Submit”
Thank you!

Questions? Comments?