WV GIS Technical Center
Dept. of Geology and Geography
West Virginia University

SYSTEM TECHNICAL DOCUMENT

DOT Scanning Project

February 28, 2023

This document describes technical specifications of the DOT Scanning Project

Jim Schindling
GIS Programmer
jpschindling@mix.wvu.edu

Kurt Donaldson
Manager
Kdonalds@wvu.edu

Heather Maxey
Project Lead
hamaxey@mix.wvu.edu

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

OVERVIEW

This document describes technical specifications for the highway plan scanning project sponsored by the WV Department of Transportation (DOT). The purpose of the system is to support the process of scanning DOT Project Plan books and making them available to the public through a web-based interface. The system utilizes Windows based file servers, ESRI ArcGIS Server, and Microsoft SQL Server database software for collecting, storing, and searching a digital collection of plans. Two web applications have been developed using the ArcGIS JavaScript API and PHP, allowing users to search the collection. The first (http://www.mapwv.gov/dotplans) allows users to enter non-spatial search criteria to return lists of matching plans while the second (https://mapwv.gov/DOTPlans/viewer) provides an interactive map interface for performing spatial queries to select and view plan footprints and details.

The digital images are produced by scanning hardcopy highway plan sheets provided by the DOT into TIFF format. The images are then loaded into an ESRI format geodatabase that is used to support spatial searches and web-based access. A second database is used to record additional metadata related to the images, along with a recording of the operational activity involved in the manual scanning process. A Microsoft Access application is used by the operations staff as the interface for recording operational information. Custom developed ArcGIS tools, built using Python, are used to ingest plan footprints into the geodatabase. Figure 1 presents a generalized view of the architecture.

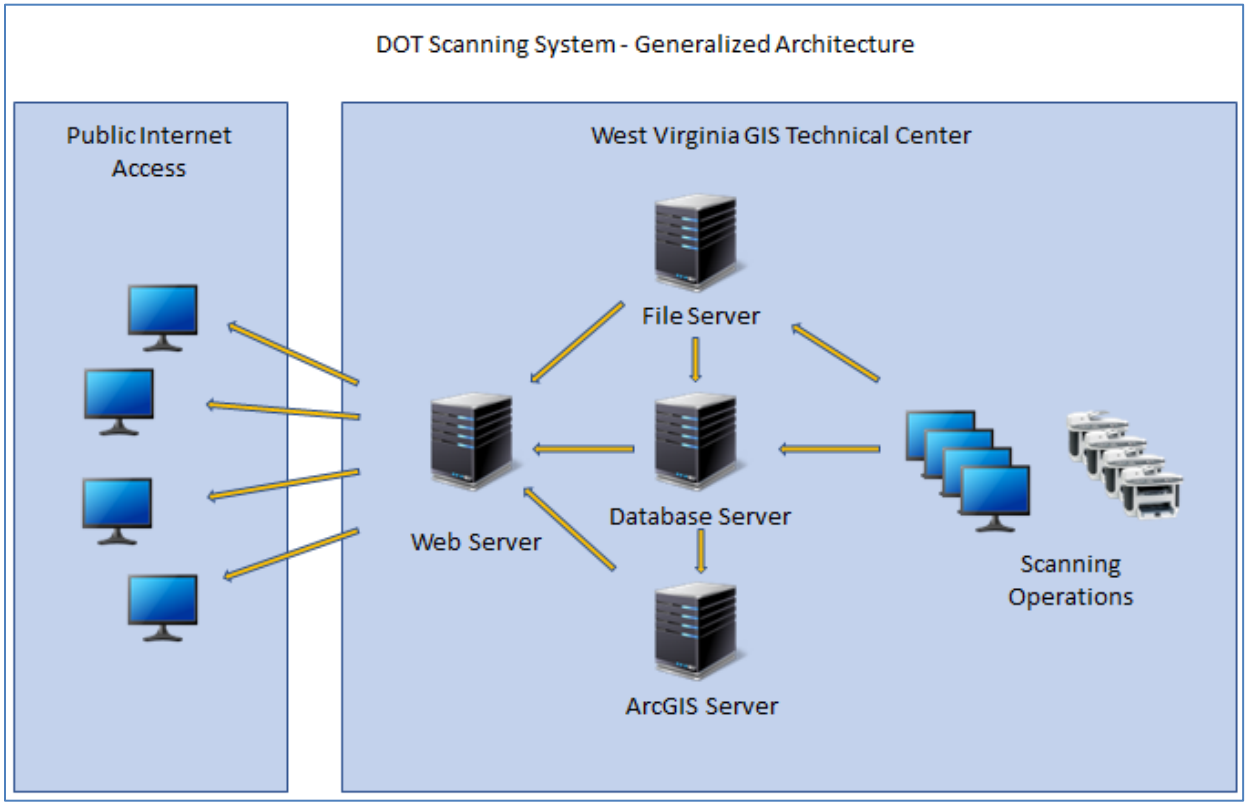

Figure 1: Generalized System Architecture

SERVER HARDWARE CONFIGURATION

The database, web, file storage, and ArcGIS services run on virtual servers that are hosted on physical servers located at the West Virginia GIS Technical Center on the campus of West Virginia University. The servers use Microsoft's Hyper-V virtualization environment, an industry standard virtualization environment. The physical servers are connected via fiber channel to a Storage Area Network (SAN). The SAN uses RAID 10 technology to enhance performance and provide data redundancy that is highly resilient to any hardware failures.

The physical machines themselves run Windows 2022 R2 Server with 2 CPUs each operating at 2.10 GHz. The virtual machine has 284 GB of RAM memory for handling many internal and external users. The database files themselves are stored on a 6 TB data drive located on the SAN. Automated SQL Server processes create full weekly database backups and daily incremental backups.

MICROSOFT SQL SERVER DATABASE

ENTERPRISE DATABASES

Both the ESRI enterprise geodatabase and the operational database are hosted using Microsoft's SQL Server 2016 Standard edition. This database supports high performance data access, high data availability, and automated backup process.

BACKUP AND RECOVERY

Full database backups are performed weekly and incremental backups are performed daily. These are performed using scheduled SQL Server Agent jobs that run database maintenance plans. The schedules and plans are authored and managed using Microsoft SQL Server Management Studio. This backup approach allows the data to be restored to its end-of-day state for any selected date. All backups are maintained for one year. Restores to previous versions of the data are performed using built-in SQL Server capabilities. This processing and other database administration is normally performed using the Microsoft SQL Server Management Studio application. Database access permissions are based on the use of Windows Authentication; that is Windows network domain logins.

ARCGIS SERVER CONFIGURATION

ARCGIS SERVER

Geospatial data is managed and served using ESRI's ArcGIS Server 10.3.1 with the data stored in a Microsoft SQL Server 2016 database. Operational staff uses ArcGIS Pro Desktop to perform updates to the geodatabase such as importing plan images and creating map footprints for display on the public website. ArcGIS Pro is also used to compose DOT related feature services and image services. ArcGIS Server publishes the resources, primarily maps and data, as RESTful web map services and/or feature services. The DOT web interfaces support querying and displaying of georeferenced project plans provided by the ArcGIS feature services. The communications between different components can be seen at Figure 2. In the diagram, the ArcGIS SDE block represents the underlying ESRI software component that supports the connection between the desktop application and the database server.

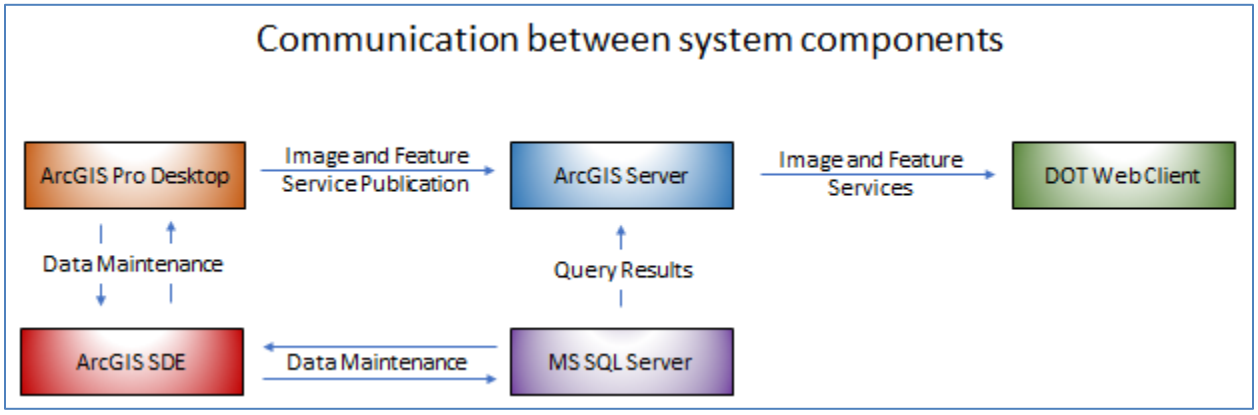

Figure 2:Communications between system components.

ARCGIS SERVICES

Image Services. ArcGIS Image Services make scanned images of DOT Project Plan data available to the Web interface. The DOT Image Service provides a viewing only DOT image layer containing scans of plan documents and outline footprints of each.

Feature Services. ArcGIS Feature services provide accesses to vector data and feature metadata that are stored in the enterprise geodatabase. These data are exposed as REST services by ArcGIS Server for display in the DOT map viewer web application. They are also available for use by external web developers and ArcGIS desktop users to be incorporated into custom map interfaces.

Spatial Reference. The map layers in the SQL Server Database are stored using the UTM Zone 17N projection. The map projection of the feature and images services provided to the web is Web Mercator (Auxiliary Sphere) EPSG: 3857. For positional accuracy, length and area calculations are performed using the UTM Zone 17N projection.

DOT PLAN WEB APPLICATIONS

As mentioned above, two web applications allow users to search the collection. The first (http://www.mapwv.gov/dotplans) allows users to enter non-spatial search criteria to return lists of matching plans while the second (https://mapwv.gov/DOTPlans/viewer) provides an interactive map interface for performing spatial queries to select and view plan footprints and details. Both applications also allow users to download copies of the scanned Project Plans in PDF and TIFF format.

NON-SPATIAL QUERY INTERFACE

The non-spatial DOT Plan Web application allows users to query the database for plan sets based on a variety of metadata attributes. Query results are presented in tabular form with links that allow users to view additional metadata details, as well as access related PDF documents and GeoTIFF images. The interface also provides links to the spatial viewer for quick access to the map-based interface. Non-spatial queries can be performed using either a free-form text search capability that searches for matching information across all metadata or by searching for matching information in specific fields.

The search options available in the non-spatial application are shown in Figure 3. The web application was developed using the ArcGIS JavaScript API and PHP. The PHP application, a server-side HTML embedded scripting language, allows access to the non-spatial scanning data while the ArcGIS JavaScript API provides access to spatial information.

The interface also allows users to view georeferenced plan images overlaid on a map background. Searching can be performed using either a free-form text search capability that searches for matching information across all metadata or by searching for matching information in specific fields. The search interface options are shown in Figure 6.

Figure 3: Non-Spatial Web Search Interface

Query results are presented in tabular form with hyperlinks that allow the user to access additional metadata and plan imagery. Figure 4 shows an example of the query results. The icons on the left of the tabular results provide access to related PDF documents, GeoTIFFs, and georeferenced imagery. The Project Key column provides a hyperlink to additional details related to the selected Project Plan. The Query URL value presented at the top of the results table can be shared with other users as a means of quickly invoking the same query that produced the results.

Query URL: //mapwv.gov/dotplans/index.php?county=Boone&signSystem=County Route&limitResults=on

20 matches out of 1819 published scans

ID	Project Key	Project Title	Federal Project	State Project	County	Sign System	Year	Route	Mile	Bridge
3	B_03_4_0003_10_000_1994_S00003	Keith Bridge	BRO_0310(009)E	S303-3/10-0.02	Boone	County	1994	3	0	---
184	B_03_4_0001_00_008_2007_S00184	Ashford Truss Bridge	BR-0001(112)E	S303-1-8.28 00	Boone	County	2007	1	8	---
514	P_03_4_0005_00_000_1974_S00514	Coal River Bridge at Seth		B00-5-18-5074	Boone	County	1974	5	0	2973
583	B_03_4_0085_05_000_1986_S00583	Danville Temporary Bridge	BRO-0858(003)5C-1	S303-85/8-0.37(0.2)	Boone	County	1986	85	0	3707
599	B_03_4_0020_00_005_1984_S00599	Boone County Route 20	BRS-0020(45)5	S303-20-4.64	Boone	County	1984	20	5	3372
609	B_03_4_0017_00_000_1985_S00609	Spruce Fork Bridge	BRS-0017(018)	S703-17-8.80	Boone	County	1985	17	0	3432
697	B_03_4_0017_00_001_1979_S00697	State Route 17 over Spruce Fork	BRS-0017(002)	5303-17-0.53	Boone	County	1979	17	1	3247
882	B_03_4_0119_20_000_1982_S00882	Madison Bridge Replacement	BRO-1192(001)	S303-119/20-0.20	Boone	County	1982	119	0	4326
897	B_03_4_0010_00_006_1979_S00897	Route 10 over Mud River		S703-10-6.61	Boone	County	1979	10	6	3251
902	B_03_4_0009_00_000_1979_S00902	Jeffrey Bridge over Spruce Fork	BRS-0009(004)	S303-9-0.01	Boone	County	1979	9	0	3254
905	B_03_4_0119_20_000_1982_S00905	Madison Bridge Replacement	BRO-1192(001)	S303-119/20-0.20	Boone	County	1982	119	0	3258
1059	B_03_4_0000_00_000_1982_S01059	East Madison Bridge			Boone	County	1982	0	0	B105

Figure 4: Web query results example

Clicking on the Project Key hyperlink displays a page containing additional details related to the selected project (Figure 5). The ProjectWise link at the bottom of the form allows DOT staff to directly access other details stored in the ProjectWise information management system. This feature is only available to users who are currently logged in to the DOT internal network.

West Virginia
Department of Transportation

Project: B_17_2_0019_00_026_2012_S00142

Scan ID:	142
Project Title:	Shinnston Bridge
Federal Project:	BR-0019(197)D
State Project:	S317-19-26.19
County:	Harrison
District:	4
Plan Type:	Bridge
Sign System:	U.S. Highway
Route:	19
Mile:	26
Special ID:	N/A
Bridges:	
Fiscal Year:	2012
Supplemental Code:	(20) Construction
Project Status:	Final Design
Book Location:	DOT
Scanning:	Scanned 06/18/2014, 152 Sheets, 152 Scans
PDF:	PDF
GeoTIFF:	GeoTIFF
Map:	Show on Map
Comments:	Sheet #0146 has a different federal project number. Sheets #0143-0145 have no sheet number.
ProjectWise:	pw://DOTB6PWHQ.executive.stateofwv.gov:PW-Primary/Documents/P{09b3cd20-1541-40e5-a71c-3c06dec50c4d}/

Query URL: //mapwv.gov/dotplans/Index.php?scanID=142

Figure 5: Project Plan detail form

SPATIAL QUERY INTERFACE

The spatial DOT Plan Web application allows users to query the database using a map interface. Clicking a point on the map presents the user with information related to all DOT Plans that include the clicked point. Optionally the user can select an area to be queried by drawing a polygon that encompasses an area of interest. Users also have the option of specifying filter criteria to limit the query results to those that are the most relevant. Figure 6 gives an example of the spatial query interface with the filter popup window visible.

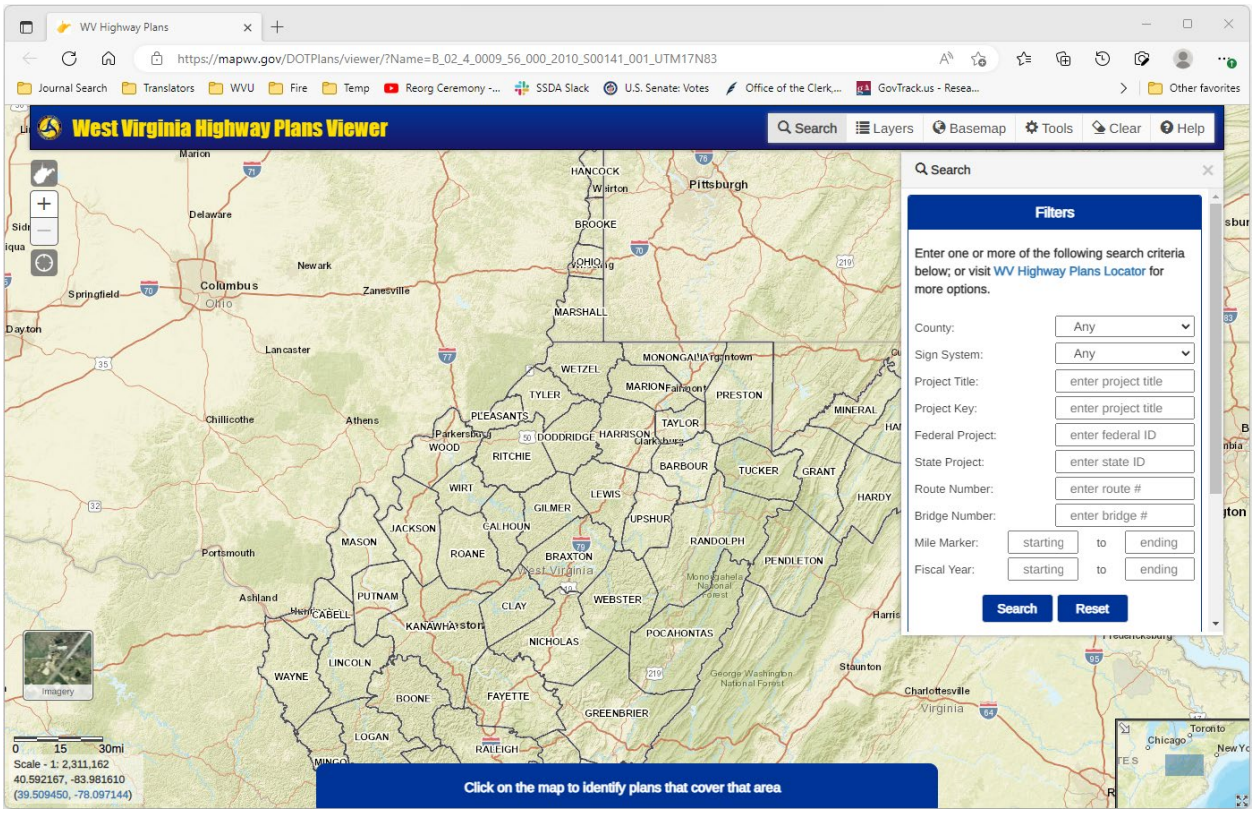

Figure 6: Spatial Web Search Interface showing filter options.

Figure 7 gives an example of a georeferenced image of a project plan overlaying the map interface along with a pop-up containing relevant metadata. Links are available in the pop-up window that allow quick access to additional metadata related to the project plan as well as downloadable versions of the scanned images.

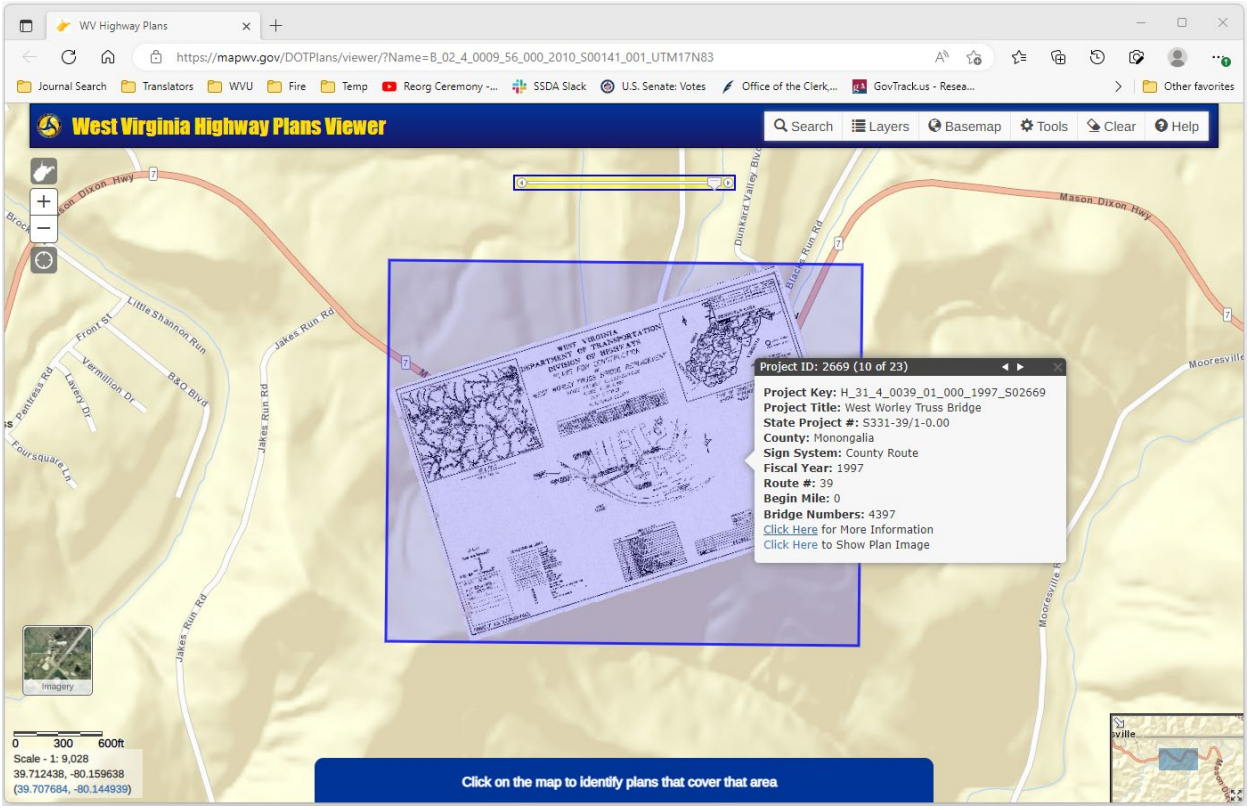

Figure 7: Spatial Web Search Interface showing a scanned project plan overlaying the map interface.

DOT PROJECT TRACKING DATA SPECIFICATIONS

PROJECT TRACKING DATABASE

The model used to manage DOT scanning data is designed to store all attributes required by DOT while also enforcing data integrity and consistency. The database also records the activities of the scanning operators for performance planning purposes.

Figures 8 and 9 show high-level data model diagrams of the DOT database. Figure 8 shows the tables used to record and manage the project book scanning metadata. Figure 9 shows the tables used to record and monitory the operational activity related to the scanning process. The individual tables and fields are defined below in Tables 1 through 15.

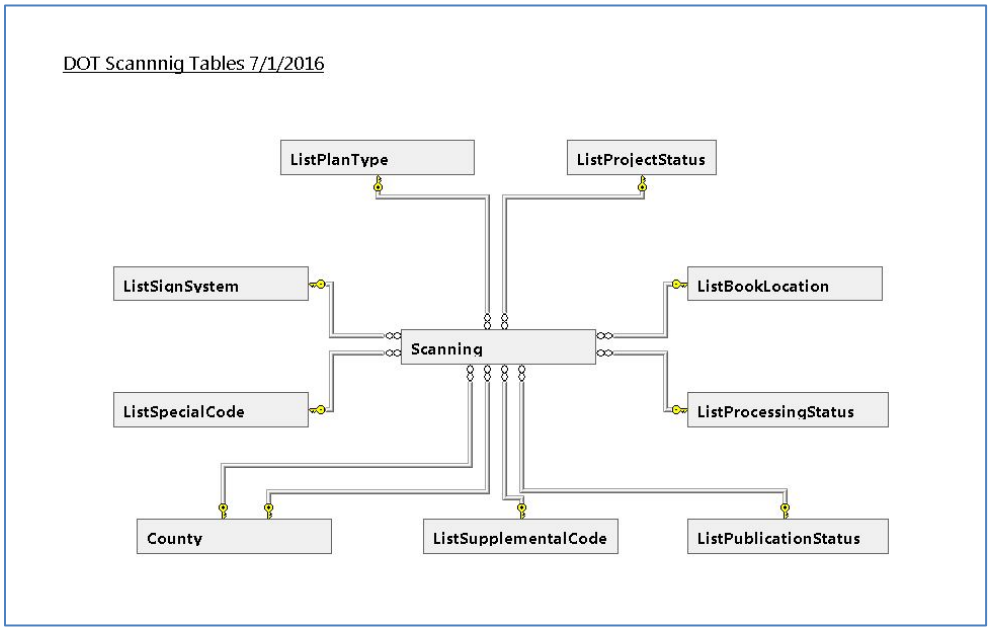

Figure 8: DOT Database tables used for storing Project Book metadata.

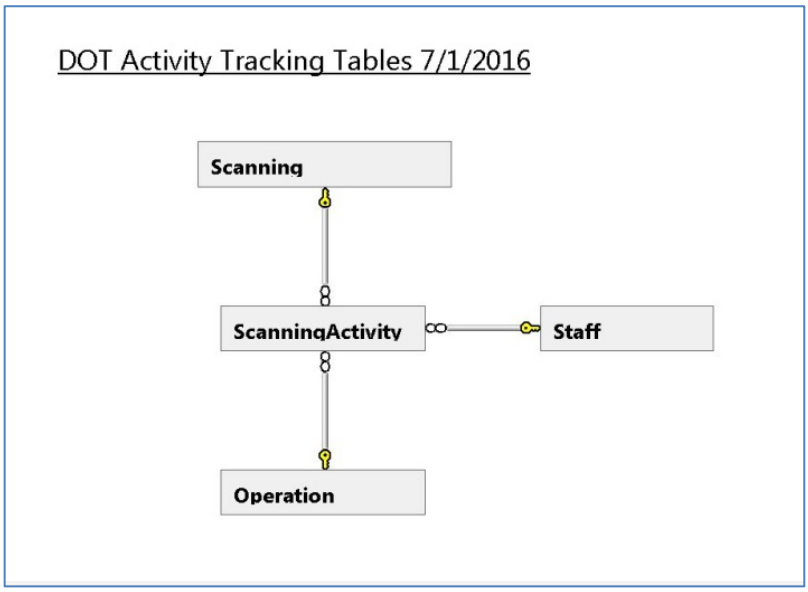

Figure 9: DOT Database tables used for tracking operational scanning activity.

DOT DATA DICTIONARY

County - Table of West Virginia Counties							
Column Name	**Description**	**Primary Key**	**Foreign Key**	**Data Type**	**Length**	**Allow Nulls**	**Default Value**
ID	Primary Key - Unique Identifier	Yes	--	int	4	--	--
Name	The name of the County	--	--	varchar	50	--	--
Code	Two-digit numerical identifier for the County provided by CAMA data source	--	--	varchar	2	Yes	--
FIPSCode	Five-digit combination of State and County FIPS codes	--	--	int	4	Yes	--

Table 1: County table definition

ListBookLocation - Lookup table containing a list of possible locations for hardcopy DOT project books							
Column Name	**Description**	**Primary Key**	**Foreign Key**	**Data Type**	**Length**	**Allow Nulls**	**Default Value**
ID	Primary Key - Unique Identifier	Yes	--	int	4	--	--
Name	Name of the location	--	--	varchar	50	--	--
SortOrder	Value used for sorting on reports and UI	--	--	int	4	--	--

Table 2: ListBookLocation table definition

ListDirection - Lookup table containing the list of possible road directions along with corresponding coded values							
Column Name	**Description**	**Primary Key**	**Foreign Key**	**Data Type**	**Length**	**Allow Nulls**	**Default Value**
ID	Primary Key - Unique Identifier	Yes	--	int	4	--	--
Name	Name of the road direction	--	--	varchar	50	--	--
Abbreviation	Abbreviation of the road direction	--	--	varchar	5	--	--
Notes	Note describing the record	--	--	varchar	500	Yes	--
SortOrder	Value used for sorting on reports and UI	--	--	int	4	--	((1))

Table 3: ListDirection table definition

ListPlanType- Lookup table containing the list of possible plan types along with corresponding coded values							
Column Name	**Description**	**Primary Key**	**Foreign Key**	**Data Type**	**Length**	**Allow Nulls**	**Default Value**
ID	Primary Key - Unique Identifier	Yes	--	int	4	--	--
Code	Code for the Plan Type	--	--	varchar	1	--	--
Name	The name of the Plan Type	--	--	varchar	50	--	--
SortOrder	Value used for sorting on reports and UI	--	--	int	4	--	--

Table 4: ListPlanType table definition

ListProcessingStatus - Lookup table containing possible operational status values for in-process project books							
Column Name	**Description**	**Primary Key**	**Foreign Key**	**Data Type**	**Length**	**Allow Nulls**	**Default Value**
ID	Primary Key - Unique Identifier	Yes	--	int	4	--	--
Name	The name of the Processing status	--	--	varchar	50	--	--
SortOrder	Value used for sorting on reports and UI	--	--	int	4	--	((1))

Table 5: ListProcessingStatus table definition

ListProjectStatus - Lookup table containing the list of possible DOT defined project book statuses							
Column Name	**Description**	**Primary Key**	**Foreign Key**	**Data Type**	**Length**	**Allow Nulls**	**Default Value**
ID	Primary Key - Unique Identifier	Yes	--	int	4	--	--
Name	The name of the Project Status	--	--	varchar	50	--	--
SortOrder	Value used for sorting on reports and UI	--	--	int	4	--	--

Table 6: ListProjectStatus table definition

ListPublicationStatus - Lookup table containing list of possible publication statuses							
Column Name	**Description**	**Primary Key**	**Foreign Key**	**Data Type**	**Length**	**Allow Nulls**	**Default Value**
ID	Primary Key - Unique Identifier	Yes	--	int	4	--	--
Name	The name of the Publication Status - Indicates the visibility of the Project Book data	--	--	varchar	50	--	--
SortOrder	Value used for sorting on reports and UI	--	--	int	4	--	((1))

Table 7: ListPublicationStatus table definition

ListSignSystem - Lookup table containing DOT defined Sign System values along with corresponding coded values							
Column Name	**Description**	**Primary Key**	**Foreign Key**	**Data Type**	**Length**	**Allow Nulls**	**Default Value**
ID	Primary Key - Unique Identifier	Yes	--	int	4	--	--
Code	Code for the Sign System	--	--	varchar	1	Yes	--
Name	Name of the Sign System	--	--	varchar	50	--	--
ShortName	Shorter version of the name - Used for compact display and reporting	--	--	varchar	50	--	--
SortOrder	Value used for sorting on reports and UI	--	--	int	4	--	--

Table 8: ListSignSystem table definition

ListSpecialCode - Lookup table containing DOT defined 'Special' codes used to distinguish project types							
Column Name	**Description**	**Primary Key**	**Foreign Key**	**Data Type**	**Length**	**Allow Nulls**	**Default Value**
ID	Primary Key - Unique Identifier	Yes	--	int	4	--	--
Code	Code for the Special Code	--	--	varchar	2	Yes	--
Name	Name of the Special Code	--	--	varchar	50	--	--
SortOrder	Value used for sorting on reports and UI	--	--	int	4	--	--

Table 9: ListSpecialCode table definition

ListSupplementalCode - Lookup table containing DOT defined Supplemental Codes							
Column Name	**Description**	**Primary Key**	**Foreign Key**	**Data Type**	**Length**	**Allow Nulls**	**Default Value**
ID	Primary Key - Unique Identifier	Yes	--	int	4	--	--
Name	Name of the Supplemental Code	--	--	varchar	50	--	--
SortOrder	Value used for sorting on reports and UI	--	--	int	4	--	--

Table 10: ListSupplementalCode table definition

Operation - List of individual operations tasks involved in scanning and loading DOT project books							
Column Name	**Description**	**Primary Key**	**Foreign Key**	**Data Type**	**Length**	**Allow Nulls**	**Default Value**
ID	Primary Key - Unique Identifier	Yes	--	int	4	--	--
Name	Name of a specific processing operation	--	--	varchar	50	--	--
SupervisorOperation	Boolean flag indicating if the operation can only be performed by a supervisor	--	--	bit	1	--	--
SortOrder	Value used for sorting on reports and UI	--	--	int	4	--	((1))

Table 11: Operation table definition

Scanning - Primary table used to store Project Book metadata							
Column Name	**Description**	**Primary Key**	**Foreign Key**	**Data Type**	**Length**	**Allow Nulls**	**Default Value**
ID	Primary Key - Unique Identifier	Yes	--	int	4	--	--
ProcessingStatusID	The ID of the corresponding Processing Status	--	Yes	int	4	--	((1))
PublicationStatusID	The ID of the corresponding Publication Status	--	Yes	int	4	--	((1))
ProjectTitle	The DOT project title	--	--	varchar	100	Yes	--
FederalProject	DOT supplied Federal Project Title	--	--	varchar	50	Yes	--
StateProject	DOT provided State Project Title	--	--	varchar	50	Yes	--
CountyID	The ID of the corresponding County record	--	Yes	int	4	Yes	--
County2ID	The ID of a secondary County related to the project	--	Yes	int	4	Yes	--
District	DOT supplied District name	--	--	int	4	Yes	--
SignSystemID	ID of the corresponding Sign System	--	Yes	int	4	Yes	--
RouteNumber	DOT supplied Route Number	--	--	int	4	Yes	--
SubRoute	DOT supplied Sub-Route Number	--	--	int	4	Yes	--
ProjectLength	Length in miles of the section of road affected by the project	--	--	varchar	255	Yes	--
ProjectDate	DOT Supplied date of the project	--	--	Date time	8	Yes	--
FiscalYear	Fiscal year that the project was budgeted to	--	--	int	4	Yes	--
SheetCount	Number of sheets in the project book	--	--	int	4	Yes	--
ScanCount	Number of sheets actually scanned	--	--	int	4	Yes	--
PlanTypeID	ID of the corresponding Plan Type	--	Yes	int	4	Yes	--
ScanDate	The date that the scanning of the project book was completed	--	--	Datetime	8	Yes	--
BookLocationID	ID of the corresponding Book Location	--	Yes	int	4	Yes	--

Scanning - Primary table used to store Project Book metadata							
Column Name	**Description**	**Primary Key**	**Foreign Key**	**Data Type**	**Length**	**Allow Nulls**	**Default Value**
Comments	General comments regarding the scanning activity for the project	--	--	varchar	255	Yes	--
SupplementalCodeID	ID of the corresponding Supplemental Code	--	Yes	int	4	Yes	--
DirectionID	ID of the corresponding Direction record	--	--	int	4	Yes	--
BeginMile	The beginning mile marker for the project	--	--	int	4	Yes	--
ProjectStatusID	ID of the corresponding Project Status record	--	Yes	int	4	Yes	--
SpecialCodeID	ID of the corresponding Special Code record	--	Yes	int	4	Yes	--
SpecialID	Not Used	--	--	varchar	5	Yes	--
BridgeNumbers	Comma delimited list of DOT supplied bridge numbers related to the project	--	--	varchar	100	Yes	--
OldProjectKey	Not used	--	--	varchar	50	Yes	--
OldDOTKey	Not used	--	--	varchar	20	Yes	--
GeoTIFFSheet	The number of the sheet that the GeoTIFF was created from	--	--	varchar	3	Yes	--
XMax	Bounding rectangle Maximum X coordinate	--	--	numeric	9	Yes	--
XMin	Bounding rectangle Minimum X coordinate	--	--	numeric	9	Yes	--
YMax	Bounding rectangle Maximum Y coordinate	--	--	numeric	9	Yes	--
YMin	Bounding rectangle Mimimun Y coordinate	--	--	numeric	9	Yes	--
URN	DOT internal pathname to the ProjectWise record for the project	--	--	varchar	500	Yes	--

Table 12: Scanning table definition

ScanningActivity - Table used to record the work involved in scanning and loading DOT project books							
Column Name	**Description**	**Primary Key**	**Foreign Key**	**Data Type**	**Length**	**Allow Nulls**	**Default Value**
ID	Primary Key - Unique Identifier	1	--	int	4	--	--
StaffID	ID of the Staff member.	--	Yes	int	4	--	--
ScanningID	ID of the Scanning record that this activity is related to	--	Yes	int	4	--	--
OperationID	ID of the corresponding Operation type	--	Yes	int	4	--	--
ActivityDate	The date and time that the activity completed	--	--	date	3	Yes	GetDate
Minutes	The number of minutes that the activity took	--	--	int	4	Yes	((0))
Quantity	The number of items processed	--	--	int	4	Yes	((0))
Notes	Note related to the activity	--	--	varchar	1000	Yes	

Table 13: ScanningActivity table definition

ScanningCountyUsage - Used to link Counties to DOT Projects - CURRENTLY NOT USED							
Column Name	**Description**	**Primary Key**	**Foreign Key**	**Data Type**	**Length**	**Allow Nulls**	**Default Value**
ID	Unique identifier - Primary key value	1	--	int	4	--	--
ScanningID	Foreign key to the related Scanning record	--	Yes	int	4	--	--
CountyID	Foreign key to the related County record	--	Yes	int	4	--	--
IsPrimary	Boolean flag indicating if the referenced County is the primary county related to the project	--	--	bit	1	--	--

Table 14: ScanningCountyUsage table definition

Staff - Table containing a list of WV GIS Tech Center workers involved in the project book scanning project							
Column Name	**Description**	**Primary Key**	**Foreign Key**	**Data Type**	**Length**	**Allow Nulls**	**Default Value**
ID	Value used for sorting on reports and UI	1	--	int	4	--	--
FirstName	The persons first name	--	--	varchar	20	--	--
LastName	The persons last name	--	--	varchar	20	--	--
IsActive	Boolean flag indicating if the person is an active employee. Only active employees can perform processing operations.	--	--	bit	1	--	((1))
Notes	General notes related to the person	--	--	varchar	500	Yes	

Table 15: Staff table definition

APPENDIX A – Estimated Sizes of Scanned Products

Files Size Estimates for Each File Format. (Yellow highlight indicates required file type for each project plan)

Product	DPI	Unit	Est. Size	X 100,000	Notes
ORIGINAL SCANS					
(1) Original TIFF 8-bit indexed color; not gray-scale because files need to be homogeneous (all color or gray-scale) for web raster mosaic.	300	Sheet	75 MB	7.5 TB	Best format for archival purposes and image quality.
PDF PLAN BOOK					
(2) PDF Sheet Saved as "Reduced Sized PDF"	300	Sheet	1 MB 55 MB avg. project book size	0.1 TB	75 to 1 compression ratio or about 1 MB per page; allows for OCR recognition. Combines all pages of plan into single file. Some image quality lost due to compression but readable.
GEO-REFERENCED					
(3) GeoRef TIFFs 8-bit indexed color (UTM projection)	300	Sheet	70 MB Zip 25 MB	7 TB 2.5 TB	Varies by crop size. Assume only index map needs to be geo-referenced.
COMPRESSED					
(4) ZIP Lossless Compression of GeoTIFFs	300	Sheet	25 MB	2.5 TB	3 to 1 compression of **GeoTIFFs**
WEB RASTER MOSAIC					
(7) Raster Mosaic for Web Map Index (Web Mercator Projection; GeoTIFF)	300	Mosaic	Files stored on server	500 GB to 7 TB	An extra processing step using Photoshop is also required to remove color value artifacts (Appendix G)

Total Size Estimates for 100,000 Scanned Images

File Type	X 100,000 scans
TIFFs/GeoTIFFs	7.5 TB
Zipped TIFFs/GeoTIFFs	3.8 TB
PDF Individual Sheets	0.1 TB

Zip File Compression Ratios of Different File Types

File Name	Type	compression ratio	Original Size (KB)	.zip Size (KB)
00003_1_wmA84.tif	GeoTIFF	2.8 to 1	17,019	6,085
00003_4_wmA84.tif	GeoTIFF	2.8 to 1	17,966	6,394
00003.pdf	PDF	1.1 to 1	16,156	15,139

FILE TYPE STANDARD

- Sheets are original scanned at 300 dpi as TIFFs. Geo-referenced sheets are in GeoTIFF format, zip compressed, and in a UTM 17N map projection.
- The PDF book format is used because it combines all map sheets of a particular construction plan set into a highly compressed single file for viewing purposes.

APPENDIX B – Advantages and Disadvantages of File Types

TIFF

TIFF (Tagged Image File Format) is recognized by the extensions .tif or .tiff. It is used especially for text and black and white images. Though it is not widely supported by web browsers, it remains the standard format for printing, scanned documents and Optical Character Recognition. The advantage of TIFF format is that it no data loss due to compression.

Purpose: Best format for original scans.

GeoTIFF

GeoTIFF is a public domain metadata standard which allows georeferencing information to be embedded within a TIFF file. Other information may include map projection, coordinate systems, ellipsoids, datums, and other information necessary to establish the exact spatial reference for the file. The GeoTIFF format is fully compliant with TIFF 6.0, therefore software incapable of reading and interpreting the specialized metadata will be able to open a GeoTIFF format file.

Purpose: Best format for spatially referenced maps using Esri GIS software.

Adobe PDF

Invented by Adobe Systems and enhanced over 20 years, Portable Document Format (PDF) is now an open standard for electronic document exchange maintained by the International Organization for Standardization (ISO). Scanned text can be converted using optical character recognition (OCR) technology to make text searchable. Files are read using free Adobe Reader software.

Purpose: Best format for viewing an entire set of project sheets at a very highly reduced file size or compression ratio.

JPEG

JPEG is a commonly used image format. Its name derives from the name of the people who developed the JPEG compression technique: the Joint Photographic Experts Group. Common file extensions associated with this format are .jpg, .jpeg. JPEG is a good format for imagery served over the web because it offers a reasonable compromise between picture size and picture quality. However, as with most image compression techniques, images can lose their quality based on the amount of compression applied and when images repeatedly edited and saved with additional compression applied. JPEG format is not recommended for scanned documents or text that is to be used with OCR software because Loss of content created by the compression logic can reduce image quality. This can affect the ability of OCR software to accurately interpret image content.

Purpose: Best format for viewing compressed TIFF scans.

Zip Compressed

ZIP is an archive file format that supports lossless data compression. A ZIP file may contain one or more files or folders that may have been compressed. Lossless data compression is a class of data compression algorithms that allows the original data to be perfectly reconstructed from the compressed data.

Purpose: A suitable compression format to compress TIFFs and GeoTIFFs. It is also useful for combining multiple files, like geo-referenced TIFFs, into a single file allowing for easier download.

APPENDIX C – Example Project Plan Links

Users can share online links to specific Project Books. Some examples are listed below:

(1) Keith Bridge:
http://mapwv.gov/DOTPlans/ImageViewer.html?Name=B_03_4_0003_10_000_1994_S00003_001_UTM17N83

(2) Trus Joist MacMillan Access Road:
http://mapwv.gov/DOTPlans/ImageViewer.html?Name=P_49_4_0013_00_001_1995_S00002_001_UTM17N83

(3) Elk Two Mile Watershed:
http://mapwv.gov/DOTPlans/ImageViewer.html?Name=P_20_4_0046_05_000_1995_S00004_001_UTM17N83

(4) St Albans Nitro Bridge:
http://mapwv.gov/DOTPlans/ImageViewer.html?Name=R_20_3_0025_00_000_1977_S00010_001_UTM17N83

(5&6) McMechen Slide Corrections 1 & 2: (two projects overlap)
http://mapwv.gov/DOTPlans/ImageViewer.html?Name=P_26_3_0002_00_000_1979_S00005_001_UTM17N83
http://mapwv.gov/DOTPlans/ImageViewer.html?Name=P_26_3_0002_00_000_1983_S00006_001_UTM17N83

(7) Grantsville Bridge:
http://mapwv.gov/DOTPlans/ImageViewer.html?Name=P_07_3_0016_00_000_1959_S00001_001_UTM17N83

(8) Guyandot and Gideon District:
http://mapwv.gov/DOTPlans/ImageViewer.html?Name=R_06_1_0064_00_010_1959_S00011_001_UTM17N83

(9) Washington and Scott Districts and Danville Corporation:
http://mapwv.gov/DOTPlans/ImageViewer.html?Name=R_03_2_0119_00_000_1970_S00007_001_UTM17N83

(10) Logan Relocation:
http://mapwv.gov/DOTPlans/ImageViewer.html?Name=P_23_2_0119_00_000_1950_S00008_001_UTM17N83

(11) Oak Hill Expressway Contract No. 6:
http://mapwv.gov/DOTPlans/ImageViewer.html?Name=R_10_2_0021_00_000_1961_S00009_001_UTM17N83

Full Extent:
http://www.arcgis.com/home/webmap/viewer.html?webmap=8cb52a4578cd47e3bb2189b3207a27a1&extent=-84.367,36.7083,-76.2536,40.3803