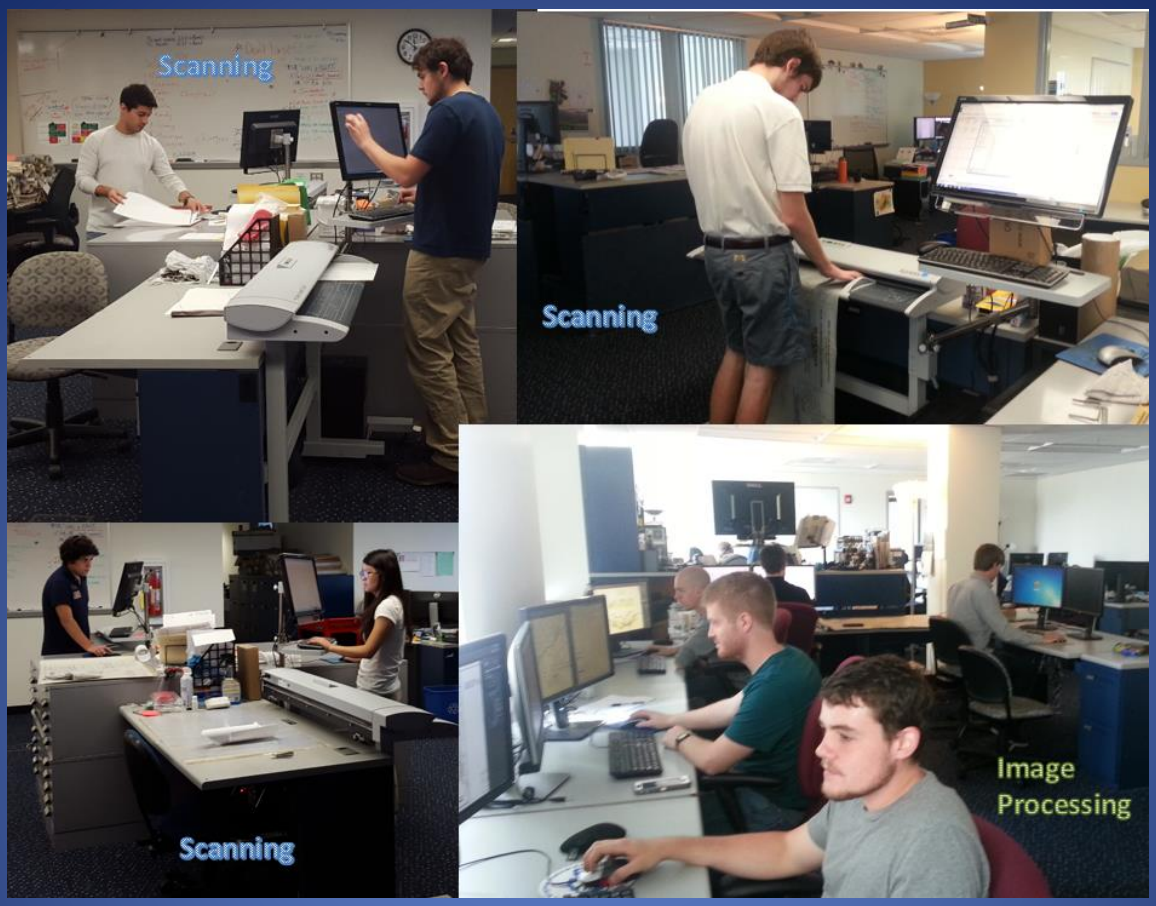

TECHNICAL SPECIFICATIONS
FOR
SCANNING WV DOT HIGHWAY PLANS

PREPARED BY

WV GIS Technical Center

West Virginia University

12/2/2016

TECHNICAL SPECIFICATIONS

SECTION 1. OVERVIEW

The West Virginia GIS Technical Center is assisting the West Virginia Department of Transportation (DOT) to scan large format highway plan sheets into digital raster images. The WV DOT Plan Scanning Project is fulfilled through seven major work tasks that are described throughout this documentation and in more detail in the Scan Procedural Manual. All the highway plan sheets are scanned, image enhanced, quality checked, and named according to prescribed specifications. For every highway plan book, the index map is spatially referenced and a single Adobe PDF file of all the scanned sheets is created. Upon completion, the digital highway plans are published to the Highway Plans Locator web application (www.mapwv.gov/dotplans) to allow users to search, identify, view and download digital highway plan sets.

The scanning project consists of seven major work tasks (Table 1) which are briefly described below. For a more detailed explanation refer to the **Scan Procedural Manual**.

- ftp://ftp.wvgis.wvu.edu/pub/temp/DOH/scan/Scan_Procedural_Manual_20161202.pdf

Seven Major Work Task

SECTION 2. SPECIFICATIONS BY WORK TASK

(1) Book Preparation and Data Management

Once the highway plan project books from the DOT are received, the book preparation and data management tasks are executed. First, the technician selects a project book and then reviews the completed transmittal sheet (Appendix B) which is used to populate the metadata of each project book in the Scan Tracking Database. Each DOT project book gets an entry in this database table, which is filled out at the time the project book is scanned.

This database consists of key transportation fields for identifying the highway book, unique scan order and project key identifiers, scan information for tracking progress, index map coordinates, and useful online links for the PDF Book, geographic location, ProjectWise location, etc. The highway plan tracking fields include the following: ScanID, ProjectTitle, FederalProject, StateProject, PlanType, CountyID, County2ID, SignsystemID, RouteNumber, SubRoute, BeginMile, FiscalYear, District, SpecialCodeID, ProjectStatus, Project Status, BridgeNumbers, BookLocation, PublicationStatusID, ScanDate, ScanCount, URN, and Comments. Table 2 describes these fields.

Table 1: Specifications by Work Task

Work Task	Specifications	Notes	Deliverables
(1) Book Preparation and Data Mgmt.	• Populate Tracking Database • Create Scan Order ID • Generate unique Project Key for Plan Set	All fields in the Scan Database are completed and assigned a Scan Order ID which is annotated on transmittal sheet. A Project Key is generated from Scan Tracking Database.	Plan Set information entered accurately into Scan Tracking Database including Scan Order ID.
(2) Sheet Scanning	• 300 dpi Resolution • 8-Bit 256 Colors • TIFF.LZW file type • Named(Scan ID) folder containing TIFF of each scanned sheet • S0XXXX-PageNo (Scan ID and Page Number file format)	File Folders are created for work tasks (1) Scanned TIFFS and image enhancements, (2) PDFs, and (3) GeoTIFFs. The TIFF folder contains all the scanned image files that *correspond* to the correct sheet number of plan set. See sheet naming standards.	300 dpi, 8-bit color scanned sheets in TIFF LZW file format without scanning image defects. Each file named to corresponding page number and sequential order of plan sheets.
(3) Image Processing	• Auto Contrast • Crop	All scanned sheets in the TIFF folder undergo Auto Contrast image enhancements and Crop extraneous white space at margins	Each TIFF scanned image auto contrast enhanced and cropped to neat line
(4) PDF Creation	• OCR Text Recognition • Reduced File Size	Image processed sheets combined into a single PDF and placed into the PDF folder. The PDF is to be reduced in file size and with OCR text search capabilities.	Single PDF of each plan set with *OCR Text Recognition* and saved as *Reduced File Size*
(5) Geo-Referencing	• Projection UTM Zone 17, NAD83 • Set RGB Values to 0,0,0 • Rectification: 4 control points, No Data = 0 • Name the file S0XXXX-indexpagenumber_UTM17N83.tiff	The index location map of each project book is georeferenced using four control points. Follow geo-referencing and file naming specifications.	Index map of each plan set spatially referenced to UTZM Zone 17 NAD83 projection, Esri GeoTIFF format, compressed to a zip file.
(6) Quality Control & File Renaming	• Image Processing, PDF, Georeferencing quality checks • Scan Tracking Database fields validated • File Names renamed to full Project Key name	All work tasks and entries of Scan Tracking Database are quality checked. Scan Database validated. The File Renaming Script renames all files with Scan Order ID to the full 30-digit Project Key Identifier.	Quality checks completed for highway plan set and tracking database. All scan order file names renamed to full Project Key name.
(7) Publish	• Publish files to Web Application • X,Y geographic extent boundary coordinates added • Update completion status of Scan Tracking Database	Digital plan sets are published to the Highway Plan Locator application and completion status in Scan Tracking Database updated.	Digital plan set published to Highway Plan Locater for users to view or download plans.

After all the required information is entered in the Scan Tracking Database, the technician then records a 6-digit Scan Order Number (e.g., S01287) on the transmittal sheet and proceeds to scanning (Task 2). Later during the Quality Control and File Renaming work task (Task 6) the Scan Order ID is converted to a 30-digit Project Key that provides a fuller description of the unique highway plan set. The Project Key is generated from the Scan Tracking Database and is composed of eight elements: Plan Type, County Code, Sign System Code, Route, Sub-Route, Begin Mile Marker, Scan Order Number, and Fiscal Year. The unique Scan Order and Project Key identifiers follow prescribed naming conventions.

See Table 2 for Data Fields of dbo_Scanning tracking database or Appendix D the Procedural Manual for entering information of a new Highway Plan Book into the Scan Tracking Database

Table 2: Scanning Database Attributes (*dbo_Scanning*). **Boldfaced** elements form 30-digit ProjectKey.

Field Name	Description	Notes
ScanID	Unique 6-digit scan order ID	1-digit location (S, D, X) character plus 5-digit number
ProjectTitle	Name of the project that the plan represents.	If no title can be found, name the project after something relevant, such as the county it is in and the route it is on.
FederalProject	Federal Project number	Not listed on the transmittal sheet. Must be found on the physical project book (generally in the top right corner).
StateProject	State Project number	Not listed on the transmittal sheet. Must be found on the physical project book (generally in the top right corner).
PlanType	Must be a letter. Currently permitted values: R = Right of Way, P = Construction Plan, B = Bridge, S = Shop Drawings, H = 12X18 size book	Corresponds to "Type" field on transmittal sheet.
CountyID	Name of county where project is located	Referred to by County ID number on transmittal sheet. Check physical book for county name (generally in the top right corner).
County2ID	Name of second county, if project crosses county borders	Not listed on the transmittal sheet. Check physical book for county name (generally in the top right corner).
SignSystemID	Denotes sign system. Code specified by WV DOT. Example: 1 = Interstate, 2 = US Route, 3 = State, 4 = County	Numbers on transmittal sheet correspond to a dropdown box in access. If no ID can be found, mark as N/A.
RouteNumber	Denotes the project primary route number	Corresponds to transmittal sheet. If no route can be found mark as 0.
SubRoute	Denotes the primary sub-route number	Corresponds to transmittal sheet. If no sub route can be found mark as 0.
BeginMile	Denotes the project's beginning milepost, as stated on the project cover/index sheet	Corresponds to the "Begin Milepost" field on the transmittal sheet. If no begin milepost can be found, mark as 0.
FiscalYear	Year project was commissioned	If no year can be found mark as 9999.
District	District in which the county is located	Corresponds to transmittal sheet.
SpecialCodeID	Denotes a special code. This code is intended to differentiate between multiple contracts or projects that fall within the same route and milepost.	Corresponds to "Special Code ID" on transmittal sheet. If no ID is listed mark as 00.

Field Name	Description	Notes
	Code to be specified by WV DOT. Example: 01 = contract 1, AA = archived set	
ProjectStatus	Denotes Project book Status ID.	Corresponds to the “Status ID” field on the transmittal sheet.1 = initial design, 2 = preliminary design, 3 = final design
BridgeNumbers	Denotes relevant bridge numbers for the project.	List all bridge numbers for the project book, found in the “Bridge No.” section of the transmittal sheet.
BookLocation	Location book was scanned	
PublicationStatusID	Indicates if book has been published on website	Drop down menu
ScanDate	Date scanned	Leave blank
Scan Count	Number of Sheets in book	Leave blank
Comments	Comments on project book	Add if necessary.
SupplementalCodeID	Denotes a supplementary code identifier which refers to a key provided by WVDOT	Mark as “Not Applicable”
DirectionID	Denotes the directional heading of the roads depicted on the map	Mark as “Unspecified”
GeoTIFFSheet	The page number used to find the georeference.	Leave blank
XMax	The latitude of the top corner used to hold the GeoTiff in correct location on map.	Leave blank
XMin	The latitude of the bottom corner used to hold the GeoTiff in correct location on map.	Leave blank
YMax	The longitude of the top corner used to hold the GeoTiff in correct location on map.	Leave blank
YMin	The longitude of the bottom corner used to hold the GeoTiff in correct location on map.	Leave blank
URN	Uniform Resource Name (URN)	Web link to WV DOT ProjectWise Content Management System

Table 3 illustrates the file naming convention of the Project Key, a fixed-length alpha-numeric name of 30 digits (23 digits plus 7 underscores) that describes the primary elements of the highway plan set. Folder structures and scanned files shall adopt this convention for file organization and standardization. Leading zeroes and null values are incorporated when necessary so that all Project Key Numbers have the same fixed length format. Data field elements of the Project Key file name are separated by underscores to improve readability.

Table 3: Scanning Project Key (1-23 digits) and Sheet Number/Suffix (24-28 digits)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	**23**	24	25	26	27	28
–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	S	–	–	–	–	–	–	–	–	–	–
Type	County ID	Sign System	Route Number					Sub Route	Begin Milepost				Fiscal Year				**Unique Scan ID**						Sheet Number				Suffix

PRIMARY FIELDS: Plan Type, County Code, Sign System Code, Route, Sub-Route, Begin Mile Marker, Scan Order Number, Fiscal Year, ScanID, Sheet Number, and Suffix. For example, the plan set **B_23_3_0010_00_013_2012_*S00153*_157B** the part "B_23_3_0010_00_013_2012_S00153" represents the Project Key and "157B" denotes Sheet Number 157, Suffix B.

PROJECT KEY NUMBER: The unique ID number generated for each project book. The first 8 fields are concatenated and separated with underscores to create this Project Key Number which is also the file folder name.

EXAMPLE FILES NAMES: Boldfaced text represents the Project Key Number (30 characters with underscores) composed of eight fields including the alphanumeric Scan Order ID (6 characters). For examples, the Project Key number **B_06_1_0064_00_011_2012_*S00149*** includes the Scan Order ID ***S00149***. Addendum information may include file type, sheet number/sheet suffix (5 digits), and projection/datum information.

1 2 **3** 4 5 6

12345678901234567890123456789012345678901234567890123456789012

B_06_1_0064_00_011_2012_*S00149*	(Project Key Folder name)
B_02_3_0081_08_000_2008_*S00150*.pdf	(PDF highway plan set)
P_47_2_0048_00_073_2005_*S00152*_002.tif	(TIF file scanned sheet number 2)
B_52_4_0036_00_002_2010_*S00151*_010A1.tif	(TIF file scanned sheet 10, suffix A1)
B_23_3_0010_00_013_2012_*S00153*_157B.tif	(TIF file scanned sheet 157, suffix B)
R_17_1_0079_00_132_1971_*S00156*_001_UTM17N83.tif.zip	(Compressed Geo-Referenced Tiff) .

(2) Sheet Scanning

Scanner Settings: After all data is recorded and the Scan ID determined, the sheet scanning can commence. Project books are unbound and each sheet individually scanned, following best scanning practices to achieve high quality scanned images void of defects. Scanning specifications (Table 4) are 300 dpi resolution, 256 Color Mode (8-bit indexed color), and file type TIFF LZW. The scanned file name matches the page or sheet number of the particular plan set and if applicable includes a sheet number suffix to differentiate duplicate page numbers.

Table 4: Scanning Parameters

Resolution	300 dpi
Color Mode	256 Colors with auto palette (8-bit indexed color)
File Type	TIFF LZW
File Name Format	S0XXXX-XXXXX.tif (A hyphen separates Scan Order ID and Page/Sheet Number)
Quality	No scanning image defects

FOLDER, FILE AND PAGE NAMING

Folder Name: On the scanning computer create a folder named after the Scan Order ID for the project book being scanned.

Scan File Name: The naming convention for individual TIFF digital scans is the Scan Order ID followed by the page or sheet number. Example: Sheet number 25 from project book S01287 would have a file name of S01287-025.tif. For duplicate pages, alpha numeric suffixes may be added after the sheet number as appropriate (e.g., S01287-025A, S01287-025A1, S01287-157B) to maintain consistency between project book page numbers and TIFF scan page numbers. Importantly, sheet numbers are to follow the correct sequential order of the plans sets so that a correctly referenced PDF Book can be created in Task 4.

Sheet/Page Naming: Use the "auto sheet naming" function of the scanning software to name the sheets (or pages) in sequential order automatically. A hyphen separates the Scan ID from the Page Number. When sheets are out of order turn off the auto naming feature and manually name the sheets appropriately.

Important Notes: First, ensure the sheet number of the digital TIFF image scan corresponds to the correct page number of the highway plan books. Second, the auto naming formats of sheet numbers may vary with different scanning software so the Project Leader should be contacted before production scanning commences.

SCANNING PROCEDURES FOR 12x34 AND 12x18 SHEETS

After the scanner and file formats are set and Scan ID determined, the sheet scanning can commence. Each plan set is wiped off with the cleaning products to protect the scanners from dirt and any potentially damaging debris. If the technician finds that the project book is especially dirty or is in bad condition, then document protectors should be used.

Select a named plan book with a Scan ID and take it to the scanner staging area. Remove anything from the sheet such as clips, tape, or staples. Rearrange pages in proper sequence and add page number suffixes if necessary for duplicate pages. Verify that scanner settings are correct. On the computer operating the scanner, create a folder named using the Scan Order ID from the transmittal sheet. Then wipe clean the sheets to protect the scanner and improve the image quality of the scan and proceed to scan the sheet according to the proper specifications. Be sure to check the digital image quality of each scan, making sure its orientation is straight and free of any streaks or artifacts made by the scanner. Repeat this process until all the sheets have been scanned. After the entire book is scanned, the TIFF files are moved to the assigned project folder of the central storage drive.

SCANNING PROCEDURES FOR 8.5x11 REPORTS

Sometimes 8.5 x 11-inch letter sized reports are included within certain 12 x 18 inch books. Since these 8.5 x 11 inch sheets are of standard letter sizes, a photocopier can be used in order to speed up the scanning process. These sheets will be scanned and sent via email automatically, in which the technician will then combine and append the 8.5 x 11 sheets to the end of the 12 x 18 project book PDF.

SCANNING QUALITY

Scanned images should be inspected for signs of a dirty lens and/or rollers at the time of scanning. This could include long horizontal streaks on the image, or dots that appear repeatedly along a horizontal line, or any number of other markings that appear on the scan but are not visible on the physical sheet. Any scans with excessive visible marks as a result of dirty lens/rollers should be rescanned after the scanner is thoroughly cleaned. If the scanner pickup rollers are unable to feed a crumpled or frayed sheet, then either flip the sheet before scanning and rotate it with the scanning software, or place it in a document protector and rescan. Lastly, the scanners should be properly calibrated so that scanning artifacts do not appear. *Refer to the Sheet Scanner Section of the User Procedural Manual for more information.*

(3) Image Processing

To improve image quality, enhancements are performed on each scanned image with software like Adobe Photoshop. Image processing tasks include the Auto Contrast and Crop functions. Auto Contrast adjusts the overall contrast of an image without affecting its color. Auto Contrast maps the lightest and darkest pixels in the image to white and black, which makes highlights appear lighter and shadows appear darker. It is important when making image enhancement adjustments to ensure no data is lost during the process. Scans are cropped to remove excess white space in the margins and to reduce file size. Technicians should be careful to not crop out any information or markings visible on the scan. *Refer to Image Processing Section of the Procedural Manual.*

(4) PDF Creation

After image processing has been completed, the edited TIFFs are compiled into a PDF document. For each WVDOT project book a single PDF document is created that contains every scan in that project book and follows the correct page order. When creating this PDF, the OCR recognized text and file size reduction tools are used. *Refer to PDF Book Section of the Procedural Manual.*

(5) Geo-Referencing Highway Index Map

The technician uses Photoshop, Google Maps, and ArcGIS to locate and geo-reference each project book's index map to UTM Zone 17N projection, NAD83 Datum. The target index map to be georeferenced is typically found on the first page of the highway plan set, but it can be found in various places throughout the project book. The page that contains the map to be georeferenced must be copied from the TIFF folder and then pasted into the GeoTiff folder.

The index map of the highway plan book is spatially referenced to control points identified on ortho-rectified imagery and other suitable reference layers. Each index map is rectified to UTM Zone 17 NAD83 using detailed specifications listed in the Procedural Manual. The GeoTIFF is named according to prescribed file naming specifications (e.g., S0XXXX-indexpagenumber_UTM17N83.tiff). *Refer to Georeferencing Section of the Procedural Manual.*

(6) Quality Control & File Renaming

Due to the complexity in the several previous processes, a special emphasis is placed on quality control and assurance to confirm that the best possible product is achieved. The quality control technician verifies that all the plan book sheets are scanned, all pages face the correct direction, no pages are askew, and there are no streaks or marks from the scanner. Other quality procedures involve checking the completeness and accurateness of all work tasks to include Scan Tracking Database entries, file names, image enhancements, PDFs, GeoTiffs, etc.

After all the quality checks are preformed the technician executes a file renaming program to rename all the Scan ID file names to the full 30-digit Project Key root name (e.g., S00152-002.tif to P_47_2_0048_00_073_2005_*S00152_002.tif*) for each project book. It is recommended that technicians do not quality control their own work. *Refer to Quality Control & File Renaming Section of the Procedural Manual.*

(7) Publishing

The final process to the WV DOT Scanning Project consists of publishing the TIFFs, PDF, and GeoTiff to the website. First, execute the PublishGeoTiff.pyt script using ArcPython copy and replace GeoTiff's into a new folder that will be used to add scan rasters to the mosaic. Once relocated into a new folder and added to the mosaic, the technician executes the DOTFootprints.pyt tool to create footprints of the georeferenced image (corners of the image). Next, the footprint extents are copied into the *dbo_Scanning* table. This is done with the UpdateFootprintextents tool in ArcMap and automatically updated the X and Y Min/Max geographic extent boundary fields in the dbo_scanning table. After this tool is finished, the technician must update the PublicationStatusID column in *dbo_Scanning*, which publishes the plan set to the web application (www.mapwv.gov/DOTplans). *Refer to Publishing Section of the Procedural Manual.*

APPENDIX A: SPECIFICATIONS OUTLINE

Task 1: Book Preparation and Data Management

- Populate the project tracking database to record all designated information.
    - The Scan Order ID is created and recorded on the project book transmittal sheet.
- A project key will be generated by the populated fields and stored in the tracking database.

Task 2: Sheet Scanning

- Each sheet in the project book are to be scanned in the following format:
    - The image file format is to be scanned as a *Tagged Image File Format*- TIFF.LZW.
    - The image file Resolution is to be scanned at a *300 dpi* resolution.
    - The scanner is to be set to scan with *8-bit 256 colors*.
    - The image name for each corresponds to the page number of plan set.
        - Name: "S0XXXX-PageNumber"
- Project Books files are generated with a name that corresponds to the Scan Order ID and contains the following folders:
    - TIFF folder- This folder stores the scanned sheets in the TIFF.LZW format. It is to be populated during the Sheet Scanning work task and edited during the Image Processing work task.
    - PDF folder- This folder stores the PDF and will be populated during the PDF work task.
    - GeoTIFF folder- This folder stores the index sheet tiff files that are used in georeferenceing and is populated during the Georeferencing work task.

Task 3: Image Processing

- Each scanned sheet is to be image enhanced and cropped. Edits will be saved in the TIFF folder.
    - Auto Contrast will be performed as the image enhancement.
    - The image will be cropped to the sheet's neat line. If information is found outside the neat line the image will be cropped to contain all relevant information.

Task 4: PDF Creation

- A PDF is to be created for each project book that will contain all image processed scanned sheets and is to be saved in the PDF folder.
- Each PDF will be reduced in file size and will recognize text within the document.

Task 5: Georeferencing

- The sheet containing spatial information within the project book will be used to georeference the sheet. The georeference is to be saved the GeoTIFF folder and will meet the following criteria.
    - Four control points are to be used in Georeferencing.
    - The image will be Rectified:
        - No Data: 0
        - Output Location: local data storage file (GeoTIFF folder)
        - Compression Type: None
        - Name: S0XXXX-indexpagenumber_UTM17N83.tiff

Task 6: Quality Control & File Renaming

- The project book's folders and files are to be checked for completion and quality.
- Files within the GeoTIFF folder are to be compressed and prepared for file renaming.
- The File Renaming Tool is to be used to name the files in a format compatible with website operations.
    - Example: S00152-002.tif to P_47_2_0048_00_073_2005_*S00152_002.tif*

Task 7: Publish

- Completed projects are to be uploaded and maintained on the Highway Plan Locator application.
- The Project database is to be updated with the completion status and the georeference coordinates.

APPENDIX B: TRANSMITTAL SHEET

Project Key	Project Title:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17		18	19	20	21	22	23
Type	County ID		Sign System	Route Number				Sub Route		Begin Milepost			Fiscal Year					*Scan Order Number* **S-** Scanned at WVGIS TC **D**- Digitally Received by District X- Letter for Other Entity *(To be completed by scan tech.)*						

Supplemental

1	4	5	6	7	8
_	_	_	_	_	_
District #	Special ID		Status ID	Supplemental Code	

9														
Bar #														
_	_	_	_	_	_	_	_	_	_	_	_	_	_	_
Bridge #														
_	_	_	_	_	_	.	_							

Notes:	ProjectWise: Y / N