Pennsylvania Spatial Data Access
The Public Geospatial Data Clearinghouse for the Commonwealth of Pennsylvania
Final Report
FY Ending June 30, 2005

Submitted
By
Maurie Caitlin Kelly
Penn State Institutes for the Environment
The Pennsylvania State University
August 2005

A. Introduction

Pennsylvania Spatial Data Access (PASDA) is Pennsylvania's official public access geospatial information clearinghouse.

PASDA was developed as a service to the citizens, governments, and businesses of the Commonwealth. PASDA is a cooperative project of the Governor's Office of Administration, Office for Information Technology, Bureau of Geospatial Technologies and Penn State Institutes for the Environment of the Pennsylvania State University. Funding and support is provided by the Pennsylvania Office for Information Technology, Bureau of Geospatial Technologies.

PASDA serves as the Commonwealth's node on the National Spatial Data Infrastructure (NSDI), Geospatial One Stop, and the National Biological Information Infrastructure. PASDA is also a member of the Geography Network

The data made available through PASDA is provided by our data partners to encourage the widespread sharing of geospatial data, eliminate the creation of redundant data sets, and to further build an inventory (through the development and hosting of metadata) of available data relevant to the Commonwealth. PASDA serves as a resource for locating data throughout the Commonwealth through its data storage, interactive mapping/webgis applications, and metadata/documentation efforts.

B. PASDA Services

- **Data Liaison Activities**
  PASDA staff work directly with state agencies to identify, document, and provide access to data.
- **Inventory and Documentation (Metadata)**
  PASDA will create FGDC standard metadata free of charge for any agency, organization, or data stakeholder. The metadata will also be hosted by PASDA for the purposes of developing the state data inventory.
- **Data Storage and Access**
  PASDA works directly with state and federal agencies, local and regional governments, non-profit organizations, and academic institutions to provide access to the widest variety of data possible. There is no fee to store or provide access to data via the PASDA clearinghouse.
- **Metadata and PASDA Training**
  PASDA staff provide free metadata training and training in the use of PASDA and PASDA data for individual organizations and groups. PASDA offers formal training at different locations throughout the Commonwealth.
- **Educational Outreach**
  PASDA works with non-profit organizations such as watershed groups and K-12 schools to promote the knowledge and use of GIS. PASDA develops lessons, tutorials, and guidelines for new GIS users.

- **Presentations & Seminars**
  PASDA staff offer seminars and presentations at meetings, conferences, and schools.
- **User Assistance**
  PASDA is committed to providing timely user assistance with the PASDA site and data. PASDA fields between 10 to 25 requests for information, assistance, and user inquiries per day. PASDA provides full time user assistance via e-mail and assistance via phone Monday thru Friday from 9-5.
- **PASDA Data on CD-ROM**
  All data on PASDA can be obtained on CD-ROM through a cooperative arrangement with the Pennsylvania State Data Center for a nominal charge.

C. PASDA Staff

- **Maurie Caitlin Kelly**, PASDA Coordinator
- **Gary W. Petersen**, Co-Director, Office for Remote Sensing of Earth Resources (ORSER), Environmental Resources Research Institute
- **Rick L. Day**, Assistant Professor of Soil Science and Environmental Information Systems, College of Agricultural Sciences
- **Ryan Baxter**, Information Technology Coordinator
- **Tracey Walrath**, Education Coordinator
- **David Walrath**, Research Assistant & Applications Programmer
- **James Spayd**, Metadata Coordinator
- **Scott Dane**, Data Manager

D. Major Accomplishments

1. User Centered Interface (UCI)

The UCI was undertaken in an effort to address comments and suggestions from users, improve the functionality, navigation, and responsiveness of the website/clearinghouse, and to provide additional data search and retrieval options for users.

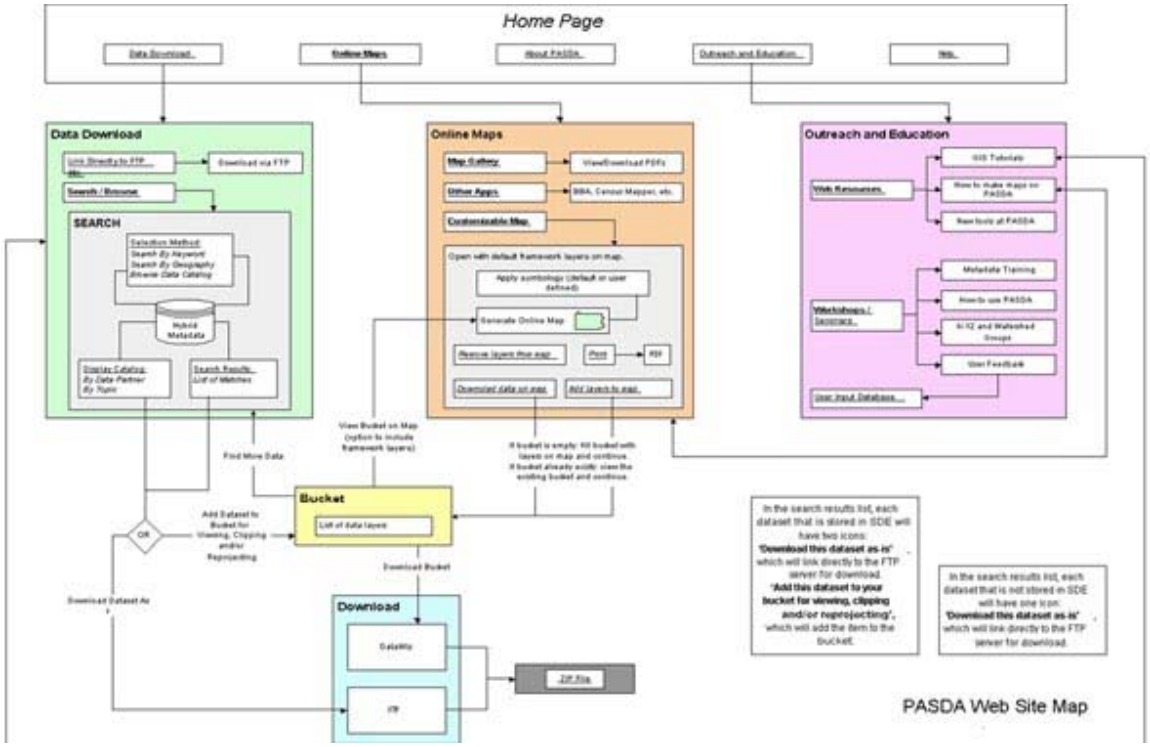


Figure 1. UCI Diagram

The primary goal of this initiative is to streamline and simplify both the interactions of users with the site and data as well as the management of the data/metadata stores, website, and applications.

Figure 2. User Centered Interface Homepage

The UCI is the new focus and core of PASDA's interface architecture. In this architecture, users will be able to add any number of data layers to a shopping cart-like collection prior to viewing or downloading them. Each other component in the new PADSA architecture, i.e., the search tools, the Web mapping tool and the download tool, communicate with the data cart to retrieve the user's data collection. The data cart was implemented by storing a user's session information in a relational database. This information may include an auto-generated user ID, the IDs for each data layer in the bucket, a date stamp and the status of the user's session.

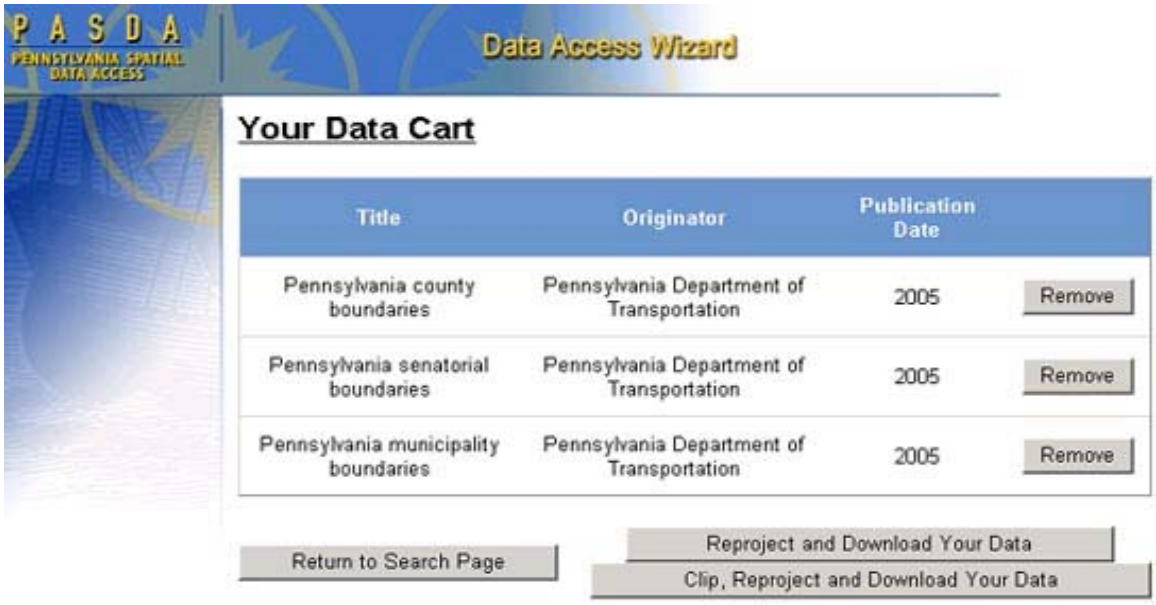


Figure 3. User Centered Interface Data Cart.

The data cart has a simple interface where users can view the list of data layers currently in their collection, add new data layers by returning to the search tools, or delete unwanted layers.

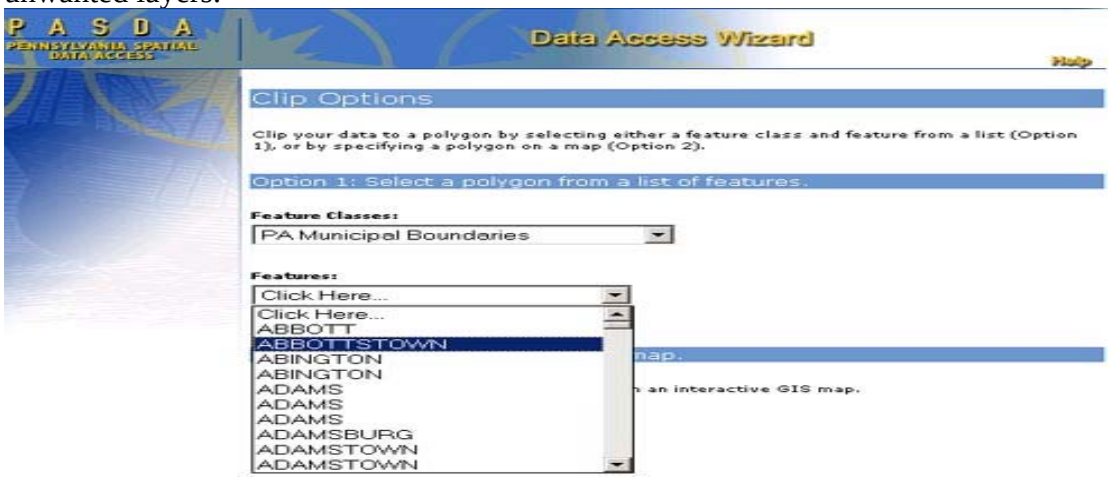


Figure 3. Data Wizard/Clip & Reproject Interface

In addition, this page provides links to the Web mapping tool, where the data layers can be viewed, and to the download tool where data can be processed (i.e., clipped and reprojected) and downloaded.

2. Web GIS Applications

PennCat

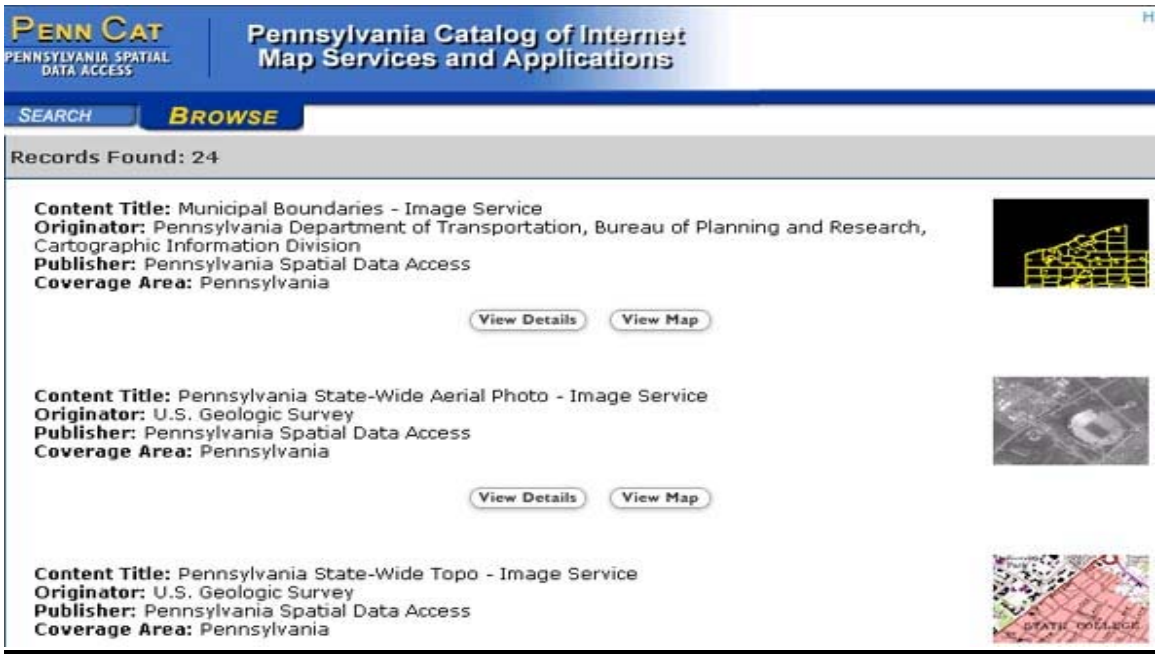
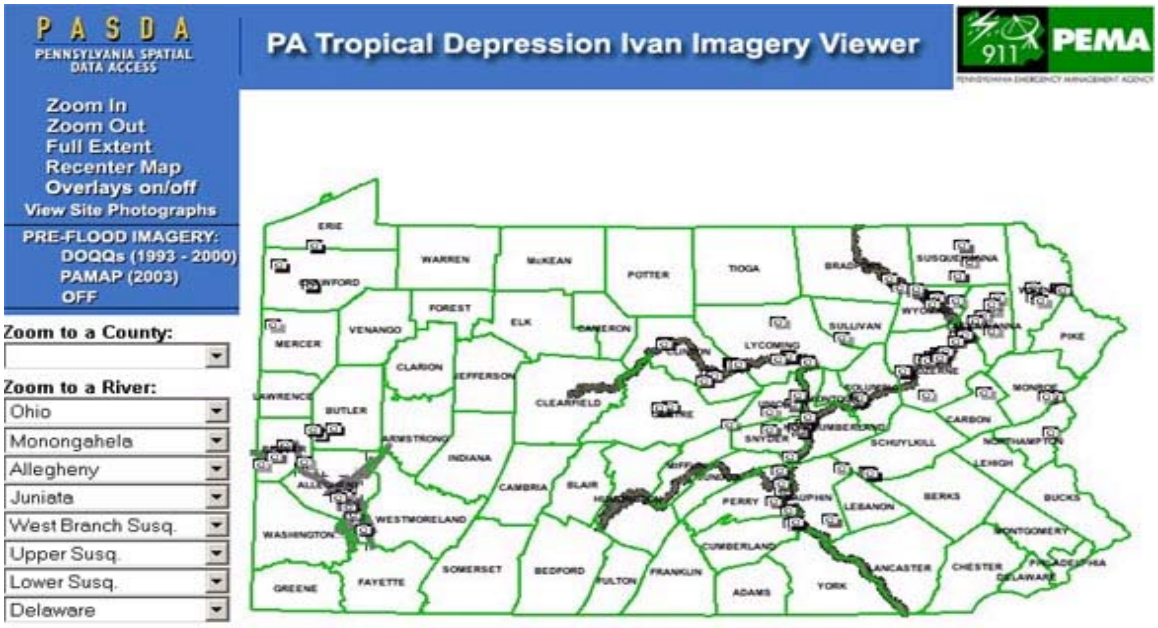
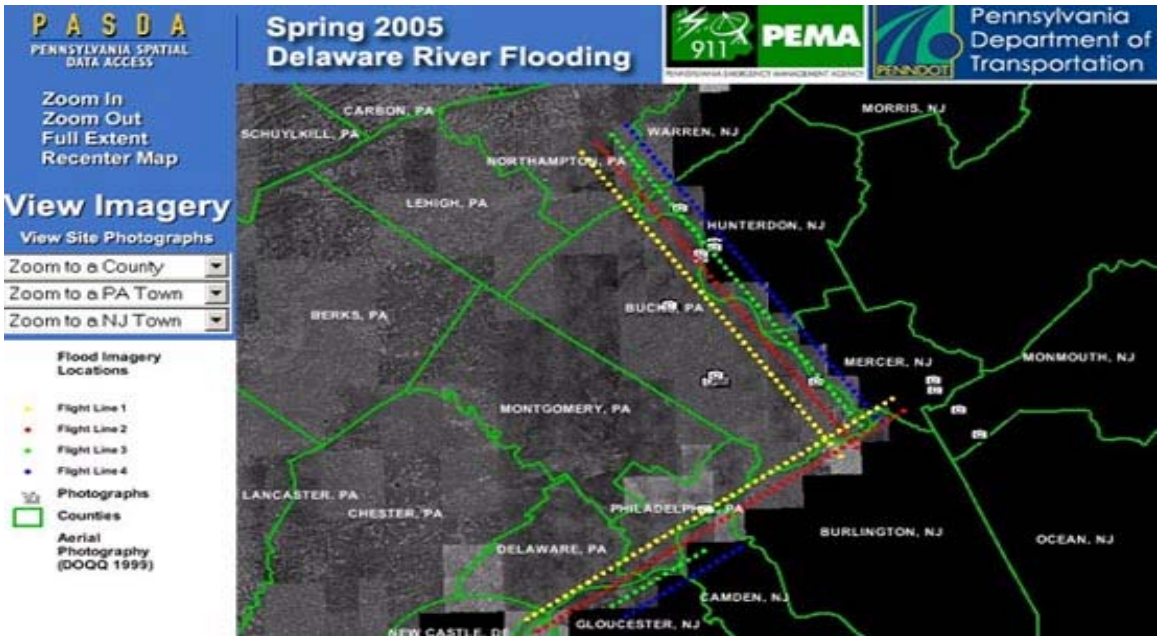


Figure 4. A Few of the Image Services from PennCAT

Tropical Depression Ivan Flood Mapping Application



Delaware River Spring 2005 Flood Mapping Application



3. PAMAP

PASDA worked with PA DCNR, BGT, and the PAMAP program to provide access to PAMAP program data. The most prominent acquisition this year was the south central imagery.

PAMAP South Central Imagery

4. FTP

During the last few months of the 2004/2005 contract, the PASDA staff reviewed all data on the FTP site and prepared a list of data to be either updated or eliminated. The data to be eliminated consisted mostly of data that had been reprojected into several different projections or data that had been updated throughout the year. All data, with the exception of river conservation programs, are available jointly through the ftp site and UCI.

5. Metadata

Reorganization/Hybrid Database Design

This task refers to the effort to migrate from the current file-based system in use on the PASDA site wherein the source XML files are indexed using the open-source ISITE package in use by many of the NSDI nodes for Z39.50 compatibility. The reorganization refers to a restructuring of the documentation scheme to provide a collection level documentation along with the file-level that currently exists. The search hierarchy will also need to be modified to allow users to search at their choice of collection or file level, i.e. search for DOQQ's as a collection or for a specific DOQQ within that collection.

The Hybrid Database will allow us to migrate from the ISITE file-based structure to a database driven search for the site and search services while maintaining the integrity of contributed metadata files. This involves parsing the metadata out to indexed database tables for rapid search and reporting while still allowing full access to the source XML document as a binary in the database table. Also incorporated into this is the integration of the different paths to access that a single dataset may have, i.e. built into an application, map/feature service, interactive download or raw downloadable data.

Training

PASDA provided several metadata training sessions/informational sessions throughout the year—most significantly to Berks County in 5/05. In addition, PASDA wrote and updated the PGDSS standard for metadata for PA.

Development

Some highlights of metadata development—please note that all new data on PASDA has metadata developed for it as it is acquired or updated:

- Provided input to BGT on metadata standards
- SRTM
- City of Philadelphia
- Soils
- Juniata County
- PennDot
- Centre County
- PAMAP
- Berks County
- Western Pennsylvania Conservancy
- PEMA
- PASDA data inventory was reviewed for categorization of data into ISO metadata standard theme/keyword thesaurus. Each data set will have one (or more) theme/keywords assigned to it in the database that will reflect the standard categories accepted by ISO. These standard theme keywords will be used in the User Centered Interface (UCI) design and database as the basis for keyword searching and to display the catalog by topic. These standards are accepted at the national level for initiatives such as geospatial one stop.

6. Data

There were several major data acquisitions during the year, these include:

PAMAP-- Orthoimagery for south-central Pennsylvania captured in April of 2003. These images are MrSid compressed natural color orthoimages that were produced at 2-feet pixel resolution. Each tile represents a portion of a ten county area located in south-central Pennsylvania. The following ten counties are covered within the project area: Fulton County, Franklin County, Adams County, York County, Huntingdon County, Cumberland County, Dauphin County, Mifflin County, Juniata County, Perry County. A shapefile index to the tile scheme is also available.

Tropical Depression Ivan-- A mirror of the PEMA FTP site containing MrSID mosaics of aerial imagery captured September 19-20, 2004 depicting the flooding caused by Tropical Depression Ivan in the previous days. To document the damages and impact of flooding along the major river corridors of the Commonwealth, PEMA authorized aerial photography missions for the 3 Rivers in Pittsburgh as well as Lake Raystown and the Juniata River, and the complete Susquehanna River basin.

NAIP-- This data set contains imagery from the National Agriculture Imagery Program (NAIP). Funding for upgrading the resolution and scope of the NAIP data in Pennsylvania was provided by DCNR Bureau of Forestry and OA Bureau of Geospatial Technologies. This imagery is 1-meter resolution color infra-red (CIR) imagery captured during leaf-on conditions primarily between June and August of 2004 though conditions necessitated recapture of some tiles through the fall of 2004.

NASA Shuttle Radar Topography Mission-- The Shuttle Radar Topography Mission (SRTM) aboard the Space Shuttle Endeavour, launched on Feb. 11, 2000. SRTM used the same radar instrument that comprised the Spaceborne Imaging Radar-C/X-Band Synthetic Aperture Radar (SIR-C/X-SAR) that flew twice on the Space Shuttle Endeavour in 1994. SRTM was assigned to collect 3-D measurements of the Earth's surface. To collect the 3-D data, engineers added a 60-meter (approximately 200-foot) mast, installed additional C-band and X-band antennas, and improved tracking and navigation devices. The mission is a cooperative project between NASA, the National Imagery and Mapping Agency (NIMA) of the U.S. Department of Defense and the German and Italian space agencies. It is managed by NASA's Jet Propulsion Laboratory (JPL), Pasadena, CA, for NASA's Earth Science Enterprise, Washington, D.C. Data for Pennsylvania is available via the PASDA site in both GeoTIFF and USGS DEM file formats.

City of Philadelphia—Refresh of all Philadelphia data including:

Buildings, Census Block Groups 1990 and 2000, Census Blocks 1990 and 2000, Census Tracts 1990 and 2000, City Plan Boundaries, City Council Districts, 1990 and 2000, Curb

Edges and other Transportation Features, Empowerment Zones, Enterprise Zones, Fairmount Park, Fire Stations Points, Generalized Rail, Geographic Market Areas, Highway District Boundaries, Hydrology, Impervious Surfaces, OrthoPhotos, Parcels, Planning Analysis Sections, Police Districts, Police Divisions, Political Divisions, Political Wards, Recycling Day Boundaries, Rubbish Day Boundaries, Sanitation Areas, Sanitation Districts, Street Centerlines, Topographic Countours 10 foot and 2 foot, Zip Codes, Zoning 2000.

PA Game Commission—New data for state gamelands and wildlife management units.

Soils—Interim data for the remaining counties in PA were acquired as well as 40 counties of final data.

Historic Markers—Over 400 marker points were gathered and added to the historic marker dataset.

PennDOT-- Revised update from the Pennsylvania Dept. of Transportation of the state maintained roadway layers extracted from the Roadways Management System (RMS). This update corrects some attribute problems with the earlier 2004 roadway files. PennDot 2004 state roadway files downloaded from PASDA previous to 11/18/2004 should be replaced with this version.

Congressional District Boundaries-- Updated Congressional district boundaries representing the delineation of the boundaries set forth in Pa. Act 34 of 2004. These boundaries supercede the Act 1, 2002 Congressional District boundaries which were available through the PASDA site previous to October 10, 2004. These are the judicially approved results of the redistricting process following the 2000 U.S. Census count. The shape files was compiled by PENNDOT's Cartographic Information Division from information provided by the Reapportionment Commission for the Bureau of Commissions, Elections and Legislation of the Pennsylvania Dept. of State.

Alliance for Aquatic Resources Monitoring-- Water quality sampling results from the Alliance for Aquatic Resource Monitoring volunteer sampling program. Contains 39,500 sampling records collected since 1986 at 734 sampling points throughout Pennsylvania.

Western Pennsylvania Conservancy—125 data layers from WPC were acquired and metadata created.

Mason Dixon-- Received data from Mason Dixon line. Reviewed data points and images. Data requires metadata development.

National Wetlands Inventory-- Reviewed existing National Wetlands Inventory data on PASDA ftp site. Began acquiring new National Wetlands Inventory data. This data has been updated by NWI project and new updates must replace existing files on PASDA.

National Hydrography Dataset—Acquired NHD for PA. Also provided project description for the development of an NHD web feature and mapping service to BGT.

Heritage Conservancy-- Reviewed data from Heritage Conservancy after discussion about providing updates. Current data offerings are complex and difficult for users to fully utilize. Heritage Conservancy will be providing streamlined updates and a new conservation place--Lower Neshaminy.

MapTech--Completed putting together the 350 MapTech 15-minute historic quad images. Will now begin the process of registering and rectifying the images in UTM.

EPA--Envirofacts data to be used on the PASDA website. Have edited Excel files for CERCLIS data (585 PA facilities, 97 on National Priority List). Now working on TRI data (2618 facilities) for Pennsylvania.

Agriculture--Acquired Census of Agriculture 2002 data.

DEMs--Inventoried all 10 meter dems for completeness. PASDA set is up to date with all available 10 meter files. Reported out to Thom Whitfield. PAGS was not regularly receiving these data from USGS.

GDT--Copied GDT data and shipped back to BGT.

7. Education, Meetings, & Outreach

User Guides & Documentation—User guides and documentation were prepared for PASDA applications and submitted to BGT 12/04.

Tutorials—Numerous tutorials were written during this contract year. Each one addressed either an application, tool, or specific software need—see below.

PASDA's GIS Tutorials

As part of its mission to support and serve the GIS community, PASDA has created a series of tutorials designed to assist PASDA users with data, software, and applications available through the site. While some of them are associated with specific GIS software, they are based upon principles and techniques common to most GIS products.

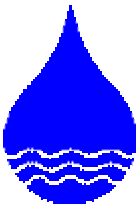
To start any of the tutorials, simply click on its title or its graphic.

The ***"GIS Basics"*** tutorial is a brief introduction to the basic concepts and principles behind Geographic Information Systems. If GIS is new to you, or you're looking for a simple refresher, this tutorial is for you.

The ***"Learning ArcExplorer"*** tutorial guides users through the basic principles and techniques of a simple, and free, GIS software package. It is ideal for users interested in getting their feet wet with GIS software.

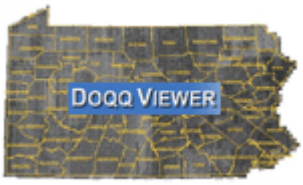
The ***"Learning ArcView"*** tutorial features a more robust and fully featured GIS program, and is better suited for users at least somewhat familiar with GIS who seek to extend their professional skills.

The ***"Map Your Watershed"*** tutorial was created to provide step by step instructions for making a watershed based map using tools and data available through PASDA.



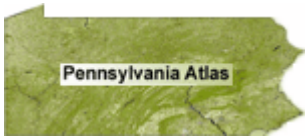
The ***"Learning ArcGIS"*** tutorial features a more robust and fully featured GIS program, and is better suited for users at least somewhat familiar with ESRI's ArcGIS 8.0 and higher versions who seek to extend their professional skills.

The ***"Downloading Data from PASDA"*** tutorial guides users through the basic principles and techniques of searching for and downloading data from PASDA.



The *"How to use the DOQQ Viewer"* tutorial guides users through the application, explaining how to view and download aerial photographs for the state of Pennsylvania.

The ***"How to use the Pennsylvania Atlas"*** tutorial was created to provide step by step instructions for viewing and downloading many base data layers for the state of Pennsylvania.



The ***"How to use PennCat"*** tutorial provides instructions on PASDA's catalog of internet map services and applications.

The ***"How to use the Breeding Bird Locator"*** tutorial was created to provide instructions on how to buffer an area around a street address and retrieve a list of all the breeding birds in that area. This is based on the 1990 Breeding Bird Atlas.

The ***"How to use the Census Mapper"*** tutorial provides instructions on PASDA's catalog of internet map services and applications.

The ***"How to use the Data Reprojector/Clipper"*** tutorial

In addition, PASDA was the first winner of the GIS Service Aware from the Western PA GIS Conferences in September 2005. PASDA participated in many meetings, conferences, and sessions during the year. These include:

- Wissahickon Video conference 5/05
- Western PA GIS Conference GIS Service Award 9/04
- PA GIS Conference 5/05
- GIS Day/GIS Day Committee 11/04
- GIS Day at Penn State 11/04
- Monthly GTC meetings
- Geospatial Communities of Practice meeting 6/05
- Rural GIS conference 10/04
- CCAP presentation 10/04
- York Advanced Skills Center Training 10/04
- DEP Metadata Project—Continued to assist with this effort until September

PASDA staff also continued to meet with data providers and users throughout the year. These included meetings with:

- City of Philadelphia
- PEMA
- DCNR
- Western Pennsylvania Conservancy
- State Soil Scientist
- Juniata County
- Centre County
- USGS
- PAGS
- DEP
- Submitted one page synopsis of educational opportunities for GIS in PA
- Provided special request of PAMAP imagery to Union County

8. Cooperative Efforts

Geospatial One Stop—Worked throughout the end of 2004 to integrate PASDA with GeoSpatial One Stop.

NBII—Provided access to metadata search and retrieval for the National Biological Information Infrastructure.

Agencies—Continued to work with agencies to identify data and metadata needs.