

DC GIS Steering Committee Meeting June 7, 2007

Barney Krucoff GIS Director Office of The Chief Technology Officer Barney.Krucoff@dc.gov 202-727-9307



Agenda

- Introductions Barney Krucoff
- DC GIS News Barney Krucoff
- System update Zhen Lo
- Data Report Mario Field
- Training Report Tim Abdella
- Income Data from OCFO Kelly Dinkins and Fitzroy Lee
- DC WASA Asset Identification Louis Desjardins



News Items

- ESRI Enterprise License (Back at the table)
- United States Geological Survey Planning Grant
 not this year
- Regional GIS News
 - COG data exchange HUB (DOH)
- DC GIS showcase with WDCEP in Vegas
- ESRI User Conference June 18 22

– Any items to address with ESRI or community?



New Administrative Projects

- Mayor's Transparency Web Site
- Department of Environment
- MPD Reporting Forms
- Nuisance Property Task Force
- Office of Property Management web site
- Economic Development Cluster Focus on housing data
- WASA & FEMS Fire Hydrants
- DDOT
 - Asset Management
 - Various Permitting Applications
 - Vending
- DCRA
 - Permitting
 - Vending



DC GIS Systems Report

Zhen Lo



9.2 Upgrade Status

- ArcGIS 9.2 Upgrade Proceeding
 - ArcGIS Desktop on Citrix Complete
 - ArcSDE for DCGISCENTRAL 7/09/07
 - ArcGIS Server on going (development)
 Includes ArcIMS
- DC Atlas All-In-One
 - Delayed until ArcIMS 9.2 deployed
- DC Guide



Things to Consider

- DCGISCentral in ArcSDE 9.2
 - only access with clients in the 9.2 family
- Need ArcGIS 9.2 software for local install?
 - \\10.128.100.32\Public\arcgis92\Desktop\Desktop
- Network Deployment tips
 - Copy package from share to local network
 - Install using active directory to deploy software
 - Cmd: Msiexec /I (location of the arcgis92 setup.msi file) /qb SOFTWARE_CLASS=Viewer ADDLOCAL=ALL



DC GIS Data Report

Mario Field Dr. Data



Updated DC GIS Data

- ABRA Licensee
- Assessment Neighborhood
- Assessment Sub-neighborhood
- DC Charter School
- DC Property Point
- Fire Station
- Historic District
- Metro Entrance
- Metro Line
- Metro Station
- MAR
- Neighborhood
- Street Centerline
- Sub-watershed
- Watershed
- Zoning



Current DC GIS Data Update

- Advisory Neighborhood Commission 2002
- Address Point
- Address Alias
- Assessment Neighborhood
- Collaborative Area
- Historic District
- Military Area
- Parking Zones???
- Regional Evacuation Route
- Zoning
 - June 2006 publication



2008 Planimetric Data Update

#	Deliverable	Format
1	Digital color orthophoto imagery – 3" pixel	GeoTiff and ERDAS IMAGINE
2	Photogrammetric mapping of select planimetric features	ArcGIS 9.x GeoDatabase
3	Detailed Digital Elevation Model with Breaklines	ArcGIS 9.x Geodatabase
Options		
4	1 meter First Return LIDAR	ERDAS IMAGINE
5	Bare Earth Raster	ERDAS IMAGINE
6	Topographic Data (2 foot contour)	ArcGIS 9.x Geodatabase
7	3-Dimensional Building Model Update	TBD
8	Oblique Imagery	ERDAS IMAGINE or GeoTiff



2008 Planimetric Data Update

Tentative GIS Layer Update List:

- Bollard*
- Building
- Bridge and Tunnel
- Curb
- Digital Terrain Model
- Geodetic Control
- Grate*
- Guardrail

- Inlet*
- Manhole
- Obscured Area
- Planter (above ground)*
- Digital Terrain Model
- Railroad
- Road
- Sidewalk
- Wooded Area

*Depending on capture rate



DC GIS Training Report

Tim Abdella GeoSpatial Education Director



CWD Courses

T	he Courses: (In suggested order of completion)	Number Day	
•	Overview of DC GIS Services and Applications	#232	1
•	Introduction to DC GIS using ArcGIS 1	#230	3
•	Working with address & point based data	#233	1
•	GIS for ESF (Emergency Service Functions)	#234	2
•	GIS for Professional Analysts	#235	2

Schedule on-line through September 2007





CWD Registration

Course Registration:

- Select Classes
- Get Supervisor Approval
- Complete the Training Form
- E-mail to Training Coordinator for his/her review and signature
- Confirmation of enrollment in the course will be emailed to you
- Notify CWD if you can not attend within 3 days of the class
- For more information about CWD, call (202) 727-1523
- All Information is on dcgis.in.dc.gov \rightarrow Training Website

http://www.dcop.in.dc.gov

Click on:

Training and Development





3D Data update

• Still receiving small shipments from Cyber City





Demonstration

- "Live" data feeds Geolocation Table
 - Use by joining on "DCSTATADDRESSKEY"

- Buffer tool
 - New DC GIS Tool on Tool bar

Buffer	
Step 1: Specify a Buffer Method Buffer By Address	 Search By Location Search By SSL 615 3rd st ne Find It
Step 2: Enter a Distance	Use Lot Boundary
200 Feet	615 3RD STREET NE 100
Step 3: Select Layer(s) to Report On DOH Abatements CDW GeoLocation Table Tax Lots Record Lots	
District Boundary	View Image

Join Data 🛛 🗶
Join lets you append additional data to this layer's attribute table so you can, for example, symbolize the layer's features using this data.
What do you want to join to this layer?
Join attributes from a table
1. Choose the field in this layer that the join will be based on:
AVAPP.DOHABATEMENTS.DCSTATADDRESSKEY_A
 Choose the table to join to this layer, or load the table from disk: CDW GeoLocation Table Image: Comparison of the stable of the stable form
✓ Show the attribute tables of layers in this list
3. Choose the field in the table to base the join on:
DCSTATADDRESSKEY
Advanced
About Joining Data OK Cancel





Timothy L Abdella GeoSpatial Education Director Office of the Chief Technology Officer 441 4th Street, NW Washington, DC 20001 tim.abdella@dc.gov 202-727-4946 Office 716-308-0000 Cell

Thank you



Agency Presentations

- OCFO ORA
 - Income tax Data
 - Kelly Dinkins and Fitzroy Lee

- DC WASA
 - Asset Identification
 - Louis Desjardins



Geocoding the Individual Income Tax Data

Kelly Dinkins, Data Manager Office of Revenue Analysis Office of the Chief Financial Officer June 7, 2007



What is ORA ?

- The Office of Revenue Analysis:
 - Forecasts Revenue
 - Develops Fiscal Impact Statements
 - Performs Tax Expenditure Analysis
 - Special Research Projects



Income Tax Return Structure

- The Office of Tax and Revenue (OTR) maintains a computerized tax database as part of the Integrated Tax System (ITS)
- Information in the ITS system is accessed through the SAND querying system



• ITS has over 200 tables

- Taxpayer demographics
- Financial transaction details for each tax form filed

• ID_Internal is a key field

- Given to each individual or business taxpayer
- Represents the entity in all ITS tables
- Allows the tables to be joined



- More than one entity can have the same ID_Internal
- For example, two persons filing a joint return will share the same ID_Internal





- One entity can have multiple ID_Internal's
- For example, if an individual's marital status changes, he or she will be issued a new ID_Internal based on their new filing status



- ID_Internal does not uniquely identify a taxpayer
- ID_Internal does uniquely identify a taxpayer return
- ID_Internal is used to link the individual tax return table to other tables such as SSN, name and address



- The ITS System is not optimized for querying
- SAND system is a web-based system optimized for general querying by OTR staff and ORA
- Performing queries in the SAND system can be complicated

Data Refresh

Difficult to create "snapshot" of data

Multiple tables

 Income return data separate from name, address, SSN

Knowledge of Key Fields

 Must join the ID_Internal field with account type (Individual Income Tax, Corporate Franchise tax, etc.), and the beginning date of the account period



DC GI

ORA has simplified the process and developed its own data model





- The model executes SAS codes which extracts data from 12 tables in the ITS system and combines, modifies, and manipulates the data into 9 SAS datasets:
 - (1) SAS data set corresponding to each of the (7) tax forms
 - > (1) SAS data set which holds taxpayer information
 - > (1) SAS data set which holds account information



Entity Address Country



Geocoding Process

- Reasons for geocoding were twofold:
 - 1) Policymakers oftentimes make requests that require spatial analysis
 - ITS system has numerous inconsistencies, discrepancies and errors in the data especially with respect to names and addresses



ID_Internal	2002	2003	2004
1234567	Jane Do	Doe, Jane	J. Doe

ID_Internal	2002	2003	2004
1234567	112 2 nd Street, NE	11 22 nd Street	112 2 nd St,
	Washington DC	NE, Washington,	Washington, DC
	20000	DC 20000	20000

Geocoding Process cont...

²2)

DC GIS

Geocoding allowed us to verify and correct the address data

Total Records:	646,856
Records in DC:	418,884
Total Records Geocoded:	399,736 (95.4%)
Records needing additional research: (MAR Team)	19,148 (4.6%)



The second secon

- ORA imported the MAR file (138,133) and All DC Records (418,884) file into SAS
- Joined *by Address ID* to get all geospatial data
- Joined resulting file with 2005 Individual Income Tax Return <u>by ID_INTERNAL</u> to get income tax data with geospatial data



Data Aggregation Process

- Spatial Join between Census tract and income tax data with the following fields:
 - ➤ Taxes
 - Income
 - Filing Status
 - DC Earned Income Tax Credit (DCEITC)
 - Federal Earned Income Tax Credit (FEITC)

Data Aggregation Process cont...

Attributes were summarized by:

➤ Minimum

DC GIS

Maximum

➤ Sum

➢ Average



Data Dictionary

ld_Int	(ID_Internal) The internal identification is given to each taxpaying entity by the Office of Tax and Revenue
Wages	(Wages, Salaries, tips, unemployment compensation, etc) As reported on Line 3 of the 2005 D40 Individual Income Tax Return
FAGI	(Federal Adjusted Gross Income) -reported from federal 1040 tax form
DCAGI	(DC Adjusted Gross Income) Derived from subtractions and additions to the federal adjusted gross income (FAGI)
No_Dep	(Number of Dependents) Number of Dependents claimed by a taxpayer
Depend	(Dependent) The taxpayer is claimed as a dependent on another individual's return for the tax year, and as a result the taxpayer cannot claim an exemption for him/her
FEIT	(FEITC) Federal Earned Income Tax Credit - a refundable credit given to low-income working head of household filers with dependents.
DCEI	(DCEITC) DC Earned Income Tax Credit is a special tax break (based on the federal EITC), which is designed specifically for low- and moderate-income workers. Individuals who qualify for the EITC will pay less in taxes or receive a refund. District taxpayers who are allowed the credit in filing their federal individual income tax return and do not claim the District Low Income Tax Credit are eligible for D.C. EITC. For tax years starting 2005, those District taxpayers who are allowed the credit in filing their federal individual income tax return and did not claim the District Low Income Tax Credit are allowed a D.C. EITC equal to 35 percent of the amount allowed by the Internal Revenue Service



Data Dictionary cont...

HeadHH	(Head of Household) A taxpayer who is unmarried or legally separated as of December 31 of the tax year and paid over half of the cost of keeping a home for a qualifying person, such as a child or parent. Certain married people who lived apart from their spouse for the last 6 months of the tax year may also be able to use this filing status
Joint	The taxpayer is married and both spouses were D.C. residents as of December 31 of the tax year, or the spouse died in the tax year and the taxpayer did not remarry in the tax year. If legally separated, this filing status can't be used
MFCS	(Married Filing Combined Separately) The taxpayer and spouse must combine their separate income amounts so that they will receive one refund or make one payment. The taxpayer may also claim a credit for child and dependent care expenses, which are not allowed if the taxpayer files on separate returns. If the taxpayer and spouse were part-year residents during different periods of the tax year, then the tax payer cannot file separately on the same return; separate returns must be filed
MFS	(Married Filing Separately) The tax payer is married and both spouses had income. Each would report only their income, exemptions, deductions and credits, as well as report one half of the income from any securities bank accounts, real estate, etc that are registered or titled in both names
Sum	Sum of all values
Avg	Average of all values
Max	Maximum of total values
Min	Minimum of total values



Geocoded and aggregated average D.C. individual income tax liability displayed by U.S. census tracts









Income Tax	2007			
Data	Q1	Q2	Q3	Q4
2006				Х
2005		X		
2004			X	
2003			Х	
2002			Х	
2001			Х	



Other Applications

- Examples
 - EITC can be used as proxy for low-income working families to target service delivery
 - Combine with crime, health, education and other statistics for planning purposes or general research



Presentation of

The use of the US National Grid to Uniquely Identify Asset

By

Louis Desjardins

DC WASA.



Naming Assets

Uniquely and unequivocally name assets that will all be stored in the same information container

Historically, these assets were located based on an address or simply were not being differenciated.

The possible solutions:

Non-sensical key (OBJECT ID, UUID, etc)

Pros: Easy to implement

Cons: There is no logical order to the identifiers when you look at a map, can be tough to search for. Sort order is not well defined

Location based key

Pros: Features are in the same geographical location

Cons: PK might change if the initial location was wrong.

Other naming convention:

???



What is the US National Grid?

The U.S. National Grid System is an alpha-numeric reference system that overlays the UTM coordinate system. It is a "Federal Geographic Data Committee" (FGDC) standard developed to improve public safety, commerce, as well as aid the casual GPS user. The USNG provides can easy to use geoaddress system for identifying and determining locations with the help of a USNG gridded map and/or a USNG enabled GPS system.

The objective of the U.S. National Grid standard is to create a more interoperable environment for developing location-based services within the United States and to increase the interoperability of location services appliances with printed map products by establishing a nationally consistent grid reference system as the preferred grid for NSDI applications. The U.S. National Grid is based on universally-defined coordinate and grid systems and can, therefore, be easily extended for use world-wide as a universal grid reference system.





USNG values have three components.

- A Grid ZoneDesignation (GSD)

First, the U.S. geographic area shall be divided into 6-degree longitudinal zones (UTM Zones) designated by a number and 8-degree latitudinal bands designated by a letter. Thus each area is given a unique alphanumeric Grid Zone Designator (GZD) Zones 10 - 19 cover the conterminous US.

- 100,000-m Square Identification

Each GZD 6x8 degree area shall be covered by a specific scheme of 100,000-meter squares where a two-letter pair identifies each square

- Grid Coordinates:

A point position within the 100,000-meter square shall be given by the UTM grid coordinates in terms of its Easting (E) and Northing (N). For specific requirements or applications, the number of digits will depend on the precision desired in position referencing.

An example:

The Jefferson Pier USNG: 18S UJ 23371 06519. UTM: 323371E, 4306519N



The US National Grid

Users determine the required precision. These values represent a point position (southwest corner) for an area of refinement.

Four digits:	23 06	Locating a point within a 1,000-m square.
Six digits:	233 065	Locating a point within a 100-m square (football field size).
Eight digits:	2337 0651	Locating a point within a 10-m square (modest size home).
Ten digits:	23371 06519	Locating a point within a 1-m square (parking space size).

Applied to our area



Applied to our area





How accurate does it need to be?

Based on the mapping accuracy requirements defined by the FGDC¹:

	Target	Feature Positi	on Tolerance	Contour Interval SI/IP
	Map Scale SI/IP	Horizontal SI/IP	Vertical SI/IP	
Surface/subsurface Utility Detail Design Plans Elec, Mech, Sewer, Storm, etc	1:500 40 ft/in	100 mm 0.2-0.5 ft	50 mm 0.1-0.2 ft	N/A
		0.2 ft : so	= 0.061m 0.1m	

(1) - Federal Geographic Data Committee, Part 4., Standards for Architecture, Engineering, Construction (A/E/C) and Facility Management, Geospatial Positioning Accuracy Standards, FGDC-STD-007.4-2002: Washington, D.C., 2002.



Applied to our Assets





Thank you

Next DC GIS Steering Committee Meeting:

Thursday, September 6, 2007 at 1:00 p.m. Conference Room TBD 441 4th Street NW