

# What is Open Data & Open Source?



Presented by Bill Savino  
Branch Chief ADSD COTS Integration  
& Census Liaison to ICSP/SCOPE

# Open Data: Executive Order's and Memorandums

- **Executive Order 13642** – 5/9/13 – Making Open and Machine Readable the New Default for Government Information
- **M-13-13** Open Data Policy-Managing Information as an Asset

“Open data means taking data that is sitting in the vaults of the government, that the tax payers have already payed for, and jujustu-ing into the public domain as machine-readable fuel for entrepreneurship and innovation.”

Todd Park, U.S. Chief Technology Officer

# Open Data Metadata

- Data Set Discovery/Descriptive Information
  - Title, Description, Quality, Sensitivity/Access Level, Value, System of Record #, Detail Metadata Schema
- Detail/Variable Level
  - Defined by local or International standards? (Geo-JSON, JSON-stat, SDMX, XBRL, DDI, etc...)
- Dissemination/Format
  - Indents, decimal places, fonts, order of magnitude, etc.

# Open Data Publishing

- Enterprise Data Inventory (EDI)
  - All Data Sets, Governance/Metadata through the Data Management System (DMS), and Data stays in Silos
- Public Data List (PDL)
  - All Published Data Sets, Discovery/Descriptive Metadata Published through *data.gov*, and Data through Bureau APIs
- Information Asset Inventory (IAI)
  - Unstructured, Informational, Content Management of 'ALL' Assets, Adobe Experience Manager or SharePoint

# Open Data at Census

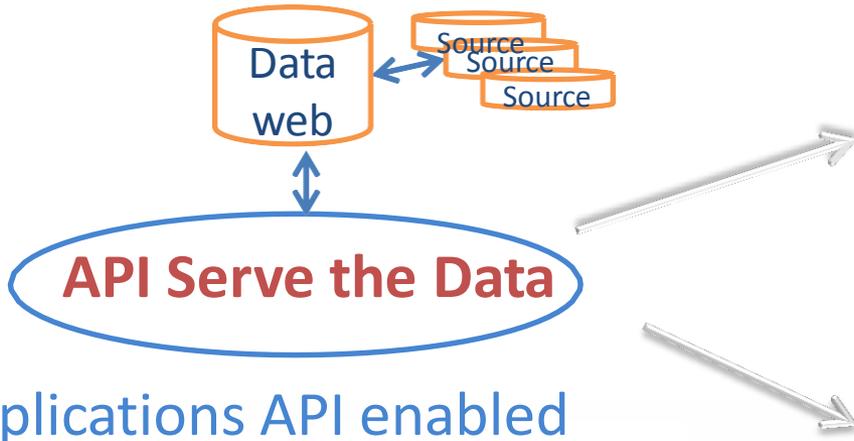
- PCO/Open Government, Policy, and Data Stewardship
- ITD/CTO - Enterprise Data Strategy & Data Model
- ITD/ADSD - Data Management System
- CEDCaP & CEDSCI
- ADRM - Data Web & Application Programming Interface (API)
- CNMP - Digital Strategy, Customer Experience

# Open Data Government Wide

- EOP/Office of Science and Technology Policy
- GSA/Office of Citizen Services and Innovative Technologies (Digital Services Innovation Center & 18F) (FedRAMP, and Data.gov)
- EOP/OMB (E-Government Office, and Office of Scientific and Statistical Policy)
- ICSP/SCOPE (Metadata and Privacy Sub-Teams)
- National Archives and Records Administration

# Digital Strategy

APIs provide new methods of data access and dissemination



## Advanced Search

The image shows two screenshots of the Census Bureau's search interface. The top screenshot shows a search for 'population' with 327,000 results. It lists three main results: World (7,004,110,246), United States (308,745,538), and California (37,253,956). A 'Data Tools' sidebar is visible on the right. The bottom screenshot shows a search for 'income' with a result for 'United States | Median Household Income, 2006-2010' valued at \$51,914. Below this, search suggestions include 'income distribution', 'income inequality', and 'income'.

## Web Applications API enabled

The image shows the 'Easy Stats' web application interface. It features a navigation menu with three steps: '1 Select Geography' (with a map and 'Select a state' dropdown), '2 Choose a topic:' (with icons for Financial, Jobs, Housing, People, and Education), and '3 Get Results' (with a 'Clear Choices' button). The main content area is titled 'About Easy Stats' and describes the tool's purpose. It is powered by 'The American Community Survey and Census API'.

## Mobile Applications

The image shows two mobile application screens. The left screen, titled 'Result' for 'New York, NY', displays a map and a list of statistics: Total Population (8,128,980), Median Housing Value (\$514,900), Age 20 to 34 yrs. old (25%), Married (43%), Education: Bachelor's Degree (19%), and Commute: Carpool (5%). The right screen, titled 'Sitegeist', shows 'AVERAGE AGE 49.8' and a 'RELATIVE AGE DISTRIBUTION' pyramid chart.

# Open Source: Laws and Memorandums

- **OMB Memorandum M-04-08**, Maximizing Use of SmartBuy and Avoiding Duplication of Agency Activities with the President's 24 E-Gov Initiatives (February 25, 2004)
- **OMB Memorandum M-04-16**, Software Acquisition (July 1, 2004)
- **Federal Information Technology Acquisition Reform Act (FITARA)**, Required that each agency have a software asset and management plan.
- **OMB Memorandum M-16-21**, Federal Source Code Policy: Achieving Efficiency, Transparency, and Innovation through Reusable and Open Source Software

# The Open Source Definition

- Free Redistribution
- Source Code
- Derived Works
- Integrity of the Author's Source Code
- No Discrimination Against Person or Groups

# The Open Source Definition

- No Discrimination Against Fields of Endeavor
- Distribution of License
- License Must Not Be Specific to a Product
- License Must Not Restrict Other Software
- License Must Be Technology-Neutral

# Open Source Actions

## ➤ Procuring

- Acquiring for use: Plug-Ins, Additional Modules, Applications (R), Languages (Python), Libraries, Operating Systems (Linux), Researcher's Utilities.

## ➤ Developing

- Writing for use: Collaboration in the wild, Git-Hub

## ➤ Publishing

- Writing for others to use: CitySDK

# Open Source Risks

## ➤ Procuring

- License, Patches, Upgrades, Security, Documentation, Help Desk, Training

## ➤ Developing

- Skill Set, Code Standards, Code Security Review, Public Scrutiny

## ➤ Publishing

- License, Patches, Upgrades, Security, Documentation, Help Desk, Training

# Open Source at Census

- PCO/Open Government, Policy
- ITD/OIS
- ITD/CTO - Engineering Office - Coding Best Practices
- ITD/ADSD - Open Initiatives, Web Applications
- CEDCaP & CEDSCI
- CNMP - Digital Strategy, Customer Experience

# Open Source Government Wide

- EOP/Office of Science and Technology Policy
- GSA/Office of Citizen Services and Innovative Technologies (Digital Services Innovation Center & 18F) (FedRAMP & Data.gov)
- EOP/OMB (E-Government Office, Office of Scientific and Statistical Policy, and Office of Federal Procurement Policy)
- ICSP/SCOPE (Cybersecurity and Metadata Sub-Teams)

# What is Open Data & Open Source?

## Questions?