CMAP web environment

November 2010

Overview

CMAP is committed to serving as the authoritative source of information about the Chicago region; our online data program is the core expression of that commitment. CMAP's online data program is built around these guiding principles:

- Comprehensiveness
- Currency (data updates)
- Flexibility of outputs
- Connectivity with partners

Comprehensiveness. To achieve comprehensiveness CMAP seeks out all data that is relevant for planning and analysis of the policy issues described in CMAP's GoTo2040 regional plan, including transportation, land use, housing and the economy.

Currency. We acquire updates on a regular basis so that CMAP's data assets are kept current. This resource-intensive practice ensures that CMAP and its constituents can recognize and react to change as it occurs in our region and beyond.

Flexibility. CMAP makes its data available in a wide range of formats to meet a wide variety of needs. Outputs include raw data, maps, charts, reports, feeds and web services. Spatial and non-spatial data are equally important in this mix.

Connectivity. Collaboration with partners is an important aspect of CMAP's data program. Connectivity includes data sharing, specifically focusing on real-time data sharing via machine-readable formats. Maintaining these partnerships is our best hope for achieving comprehensiveness and currency of the data.

Workflow

CMAP's data program depends on a full-spectrum workflow, with tasks ranging from raw data processing to data visualizations and modeling. This workflow follows the contours of our data architecture, focusing on four stages:

- Raw data processing (ETL)
- Data storage
- Data retrieval
- User interface

Raw data. CMAP's ETL (extract ,transform, load) process converts raw data into standardized database assets suitable for web applications. Raw data is acquired from many sources and in many formats.

Sources include the U.S. Census Bureau, other government agencies, commercial providers, surveys and partner organizations. Data formats include HTML, XML, PDF, DBF, Excel, delimited text and spatial formats. Our most commonly-used ETL tools are:

- Microsoft Visual Foxpro
- Microsoft Sql Server
- SAS
- Python

Data storage. CMAP's online data systems are based on a Microsoft Sql Server database. Tables adhere to a standard CMAP structure that is optimized for the matrix of geographies, data categories and time periods that characterize CMAP's data.

Data retrieval. Data retrieval is managed primarily through Microsoft's web platform including:

- IIS (Internet Information Server)
- ASP.NET (Active Server Pages .net)
- WCF (Windows Communication Framework)

Spatial data is retrieved primarily through ESRI map services in combination with .NET web services. Almost all data retrieval is handled by a small number of .NET data access components. Most data retrieval for online systems is handled by an API (application programming interface). The API consists of web services that return XML (extensible Markup Language) in response to data requests submitted by online applications. The API includes services in the following formats:

- AMF (Action Message Format), managed by WebOrb
- JSON (Javascript Object Notation)
- POX (plain old xml)
- REST (representation state transfer)
- SOAP (simple object access protocol)

For more information about CMAP's API, see http://www.data.cmap.illinois.gov/API.aspx.

User interface. Any web-based user interface tool that can work with standard APIs could display CMAP data. But CMAP focuses on a small number of user interface tools:

- Adobe Flex (general UI, charts and maps)
- ASP.NET (general UI)
- ESRI ArcGis Server (maps)
- WebFocus by Information Builders (charts and reports)

Connectivity. CMAP values collaboration and coordination among government agencies, universities, non-profits, foundations and other partners. Data sharing is an important aspect of such collaboration. Data can be shared using traditional media such as disks or electronic file transfer. But data that changes very frequently is best exchanged via server-to-server linkages—particularly web services.

Conclusion

CMAP's online data program is motivated by our belief in the importance of good data for planning, development and efficiency in government. Our workflow—from raw data to finished products such as data visualizations—is designed to bring disparate information resources into an accessible pool for our partners and the public.