

# Modernization at the U.S. Census Bureau

William G. Hazard<sup>1</sup>, U.S. Census Bureau\*

## Abstract

The Center for Enterprise Dissemination Services and Consumer Innovation (CEDSCI) is a new paradigm for data and metadata curation, harmonization, and dissemination based on the following core principles:

1. Operationalizing an information-centric model, we can architect our systems for interoperability and openness, modernize our content publication model, and deliver better, device-agnostic digital services at a lower cost
2. Breaking free from the inefficient, costly, and fragmented practices of the past, and building a sound governance structure for digital services
3. Enabling the public, entrepreneurs, and our own government programs to better leverage the wealth of federal data by ensuring that data is open and machine-readable by default

## Why We Are Innovating

By modernizing and centralizing data collection and processing activities, our efforts must keep pace with technological innovations and changing expectations from respondents and data users alike. Deploying shared services via one centralized metadata repository across our diverse programs and datasets not only provides us with efficiencies, it exposes our innovations to all consumers of Census Bureau data by repurposing the data to meet multiple needs.

## CEDSCI goals include:

1. Developing an enabling technology dissemination platform that is flexible and extensible
2. Promoting a set of dissemination tools to meet future dissemination capabilities
3. Adopting a Customer Experience Management (CEM) system
4. Establishing a standardized central metadata repository across all censuses and surveys
5. Designing shared enabling technologies based on business requirements and needs

The challenge is to modernize our technology platforms and leverage existing innovations that enable both internal staff and our data consumers to do more with the massive content we publish year round.

## Intended benefits include:

- Cost savings through elimination of duplicate systems and processes
- Spurring greater innovation
- Systematic quality assurance

---

<sup>1</sup> William G. Hazard, Branch Chief of DataWeb and Applications Staff, U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233, email: [William.G.Hazard@census.gov](mailto:William.G.Hazard@census.gov).

\* Any views expressed are those of the author(s) and not necessarily those of the U.S. Census Bureau.

- Improved customer satisfaction through metadata standardization
- Better utilization of existing tools to meet customer needs
- Greater insights into customer needs

## 1. Vision

*We can use modern tools and technologies to seize the digital opportunity and fundamentally change how the Federal Government serves both its internal and external customers—building a 21st century platform to better serve the American People.*

**Digital Government Strategy**, ([May 23, 2012, p.2](#))

In 2012, President Obama issued a directive entitled *Building a 21st Century Digital Government* ([2012](#)). The directive launched the *Digital Government Strategy*, which set the goal of delivering better digital services to the American people ([2012](#)). The strategy builds on several initiatives, including *Executive Order 13571, Streamlining Service Delivery and Improving Customer Service* ([2011](#)), and *Executive Order 13576, Delivering an Efficient, Effective, and Accountable Government* ([2011](#)).

The strategy identified four overarching principles to guide digital transformation of Federal agencies:

1. An “**Information-Centric**” approach
2. A “**Shared Platform**” approach
3. A “**Customer-Centric**” approach
4. A “**Security and Privacy**” platform

The U.S. Census Bureau had already begun making changes in information dissemination before the release of the Digital Government Strategy, including the following projects:

- **DataWeb** – Store data once and reuse for multiple purposes.
- **Census Data API** – Provide a consistent method to expose data to internal and external users and systems.
- **Digital Transformation** – Use a customer- and data-driven approach to strategically create and enhance digital products and services to benefit diverse user types.

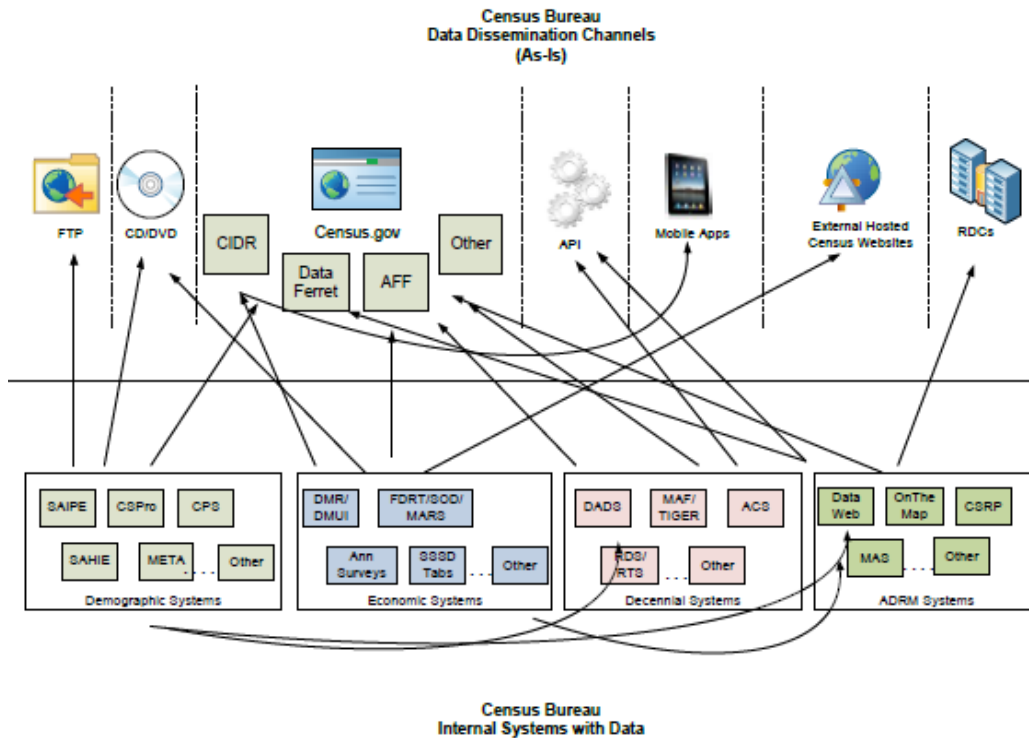
Building on these efforts and aligned with Digital Government Strategy principles, the Census Bureau has recently embarked on an innovative, multiyear effort to transform and expand the dissemination of Census content – the Center for Dissemination Services and Consumer Innovation (CEDSCI).

## 2. CEDSCI Business Challenge: Dissemination System Proliferation

Census Bureau program areas developed over 80 applications to prepare data and over 30 applications to release data, many of which store the data locally, resulting in duplications of data and storage. While the [Census Bureau website](#) (census.gov, including tools residing within the census.gov environment) is the host for most of the data releases, program areas use additional servers and tools to make their data available. Multiple independent data processing and product preparation environments use a variety of hardware platforms, software tools, and middleware, increasing software licensing, operations, and maintenance costs. Each environment replicates needed data, increasing

storage and processing costs. Programs execute survey-specific procedural steps to tabulate data, conduct disclosure checks, and release data.

Similarly, the current architecture for data dissemination involves numerous disparate processes, methodologies, technologies, and applications. During data preparation, Census Bureau programs summarize and aggregate data at different geographies, roll them up at different levels in different dimensions, and store them in multiple different formats. Program areas also define data (via metadata) differently and often tightly couple data to presentation applications. As depicted in Figure 1, for any given data product, the tools used to present the data – mapping, charts, graphs, criteria selection – differ widely across data products. This results in increased costs, complexity for the user to find data, and decreased opportunities to combine datasets for better insights.



**Figure 1:** As-Is Data Systems and Dissemination Channels

In February of 2013, Census Bureau leadership appointed a special task force of Census Bureau professionals from across the agency to look into the future of survey and census work, data dissemination, and emerging technology. Their charge was to define the core capabilities required in the coming years to keep the agency in the forefront of sharing with the whole nation the power of trusted, objective government statistics. In 2014, the task force produced a set of working papers to provide an approach to restructure and enhance dissemination activities across the entire agency to improve customer satisfaction, and grow our audience and customer base.

The plan called for a multiyear, innovative effort to transform and expand the dissemination of Census Bureau content. To manage this effort, the Census Bureau chartered CEDSCI with representation from across multiple directorates. The CEDSCI vision is: *To enable the public to make better decisions using data through a continuously*

*adaptive, customer-centric, open and accessible dissemination environment that sets the standard in the federal community and spurs innovation. Specific goals include:*

1. Developing a flexible and extensible Census Bureau dissemination platform
2. Promoting a set of dissemination tools to meet future dissemination capabilities
3. Adopting a Customer Experience Management (CEM) system
4. Establishing a standardized central metadata repository across all censuses and surveys

The CEDSCI team is an experienced set of Census Bureau experts with extensive knowledge of internal dissemination systems and external customer needs. Its diverse membership is composed of technologists, scientists, and analysts from Decennial, Economic, Demographic, Communications, Research and Methodology, and other offices. Their challenge is to modernize Census Bureau technology platforms and leverage existing innovations to provide a set of shared dissemination services that enable both Census Bureau staff and our data consumers to do more with the massive content we publish year round. Their process will be to rapidly prototype, test, and deploy the transition steps necessary to move us from the current architecture to a new permanent model.

### **3. CEDSCI Dissemination Modernization Effort**

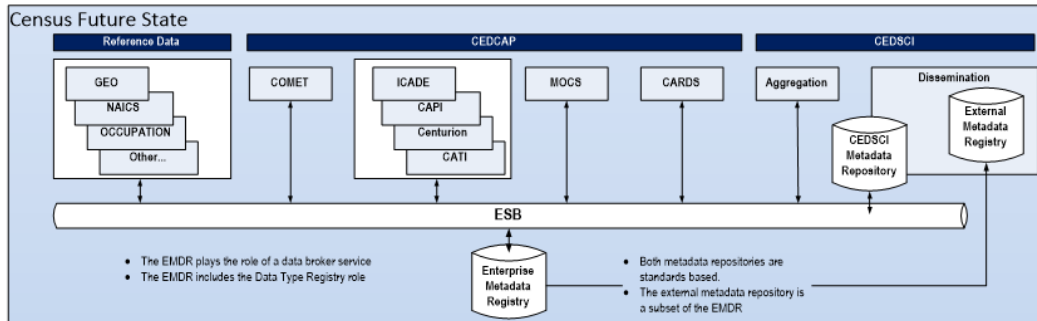
Herein, we describe how the CEDSCI dissemination modernization effort is realizing Digital Government Strategy principles.

#### **3.1 Information-Centric**

*An information-centric approach moves us from managing “documents” to managing discrete pieces of open data and content which can be tagged, shared, secured, mashed up, and presented in the way that is most useful for the consumer of that information. (2012, p. 5)*

Managing “discreet pieces of open data and content” requires standardized metadata. CEDSCI will establish a standardized enterprise metadata repository for all disseminated information. This will include all structured (statistical data) and unstructured (reports, visualizations, tools, etc.) content, which will allow users to find the information they need and build relationships between the content, allowing them to find related information easily. Customers will use metadata to find information, including standardized searchable topics, faceted filtering, a unified search, and geographic equivalents.

An enterprise metadata model creates the foundation to allow for standardization across data products, geography, and survey life cycle phases, as depicted in Figure 2.

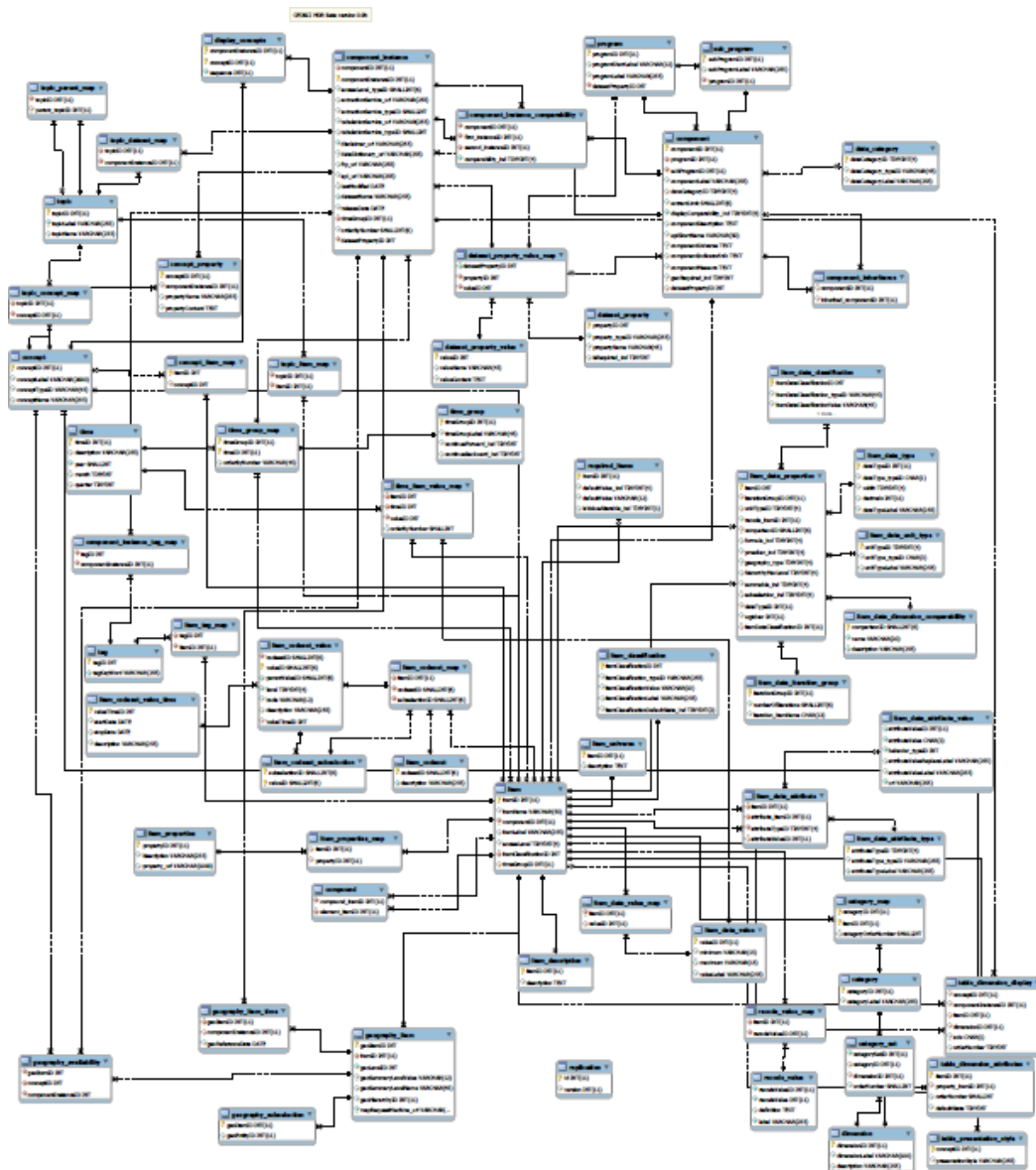


**Figure 2:** Metadata standardization and consolidation in an Enterprise Metadata Registry (EMDR) supports applications in all phases of the survey lifecycle

Commonly defined geographies and code sets allow joining information from different datasets. By using consistent topics, code sets, and labels (e.g., geographic names) users will be able to find information more easily, and better understand similarities and differences across data sets and their contents. As part of the standardization process, we are working closely with stakeholders across all program areas in the Census Bureau. Governance processes will ensure metadata remains consistent over time and across the enterprise.

Standardized metadata enables advanced analytics. For example, metadata for data items stores only differences from one time period to the next, thus defining comparability over time for that item. Data consumers can use this information to easily create time series, whether pre-defined by an analyst or created on the fly by a user. Metadata standardization prevents users from incorrectly defining a time series using an item that is not comparable over that time span. In addition to the benefits to the end user, this model also reduces the amount of work necessary by data providers; if metadata remains consistent for a new time period, there is no need to produce or deliver metadata – simply extend the time the data items are available.

Figure 3 illustrates the complexity of the Census Metadata Database schema, which provides a consistent definition of the data to all tools.



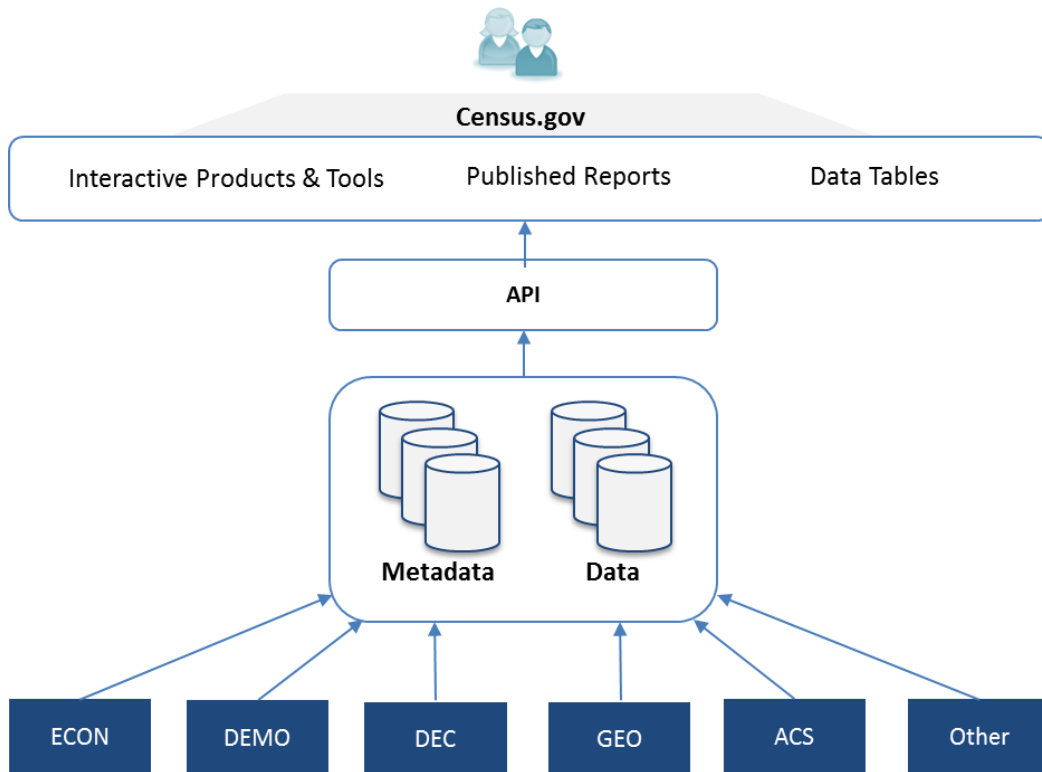
**Figure 3:** A complex metadata database supports harmonized data across Census programs

The Census Bureau will expose a metadata API that will make the same metadata available to external tool developers. It is also possible to develop utilities that generate standard metadata formats (e.g., Data Documentation Initiative [DDI]) for organizations and users that may desire those formats.

### 3.2 Shared Platform

*A shared platform approach helps federal enterprises work together, both within and across agencies, to reduce costs, streamline development, apply consistent standards, and ensure consistency in how we create and deliver information. (2012 p. 5)*

The first goal of CEDSCI is to develop an enabling technology platform that is flexible and extensible. CEDSCI is transforming the thicket of Census Bureau dissemination pathways depicted earlier in Figure 1 into the streamlined platform depicted in Figure 4.



**Figure 4:** A dissemination platform shared by Census program areas

We are developing the platform in a componentized way ensuring that the platform can easily adapt as technology advances. Key components include the discovery (index) service, data service, metadata service, and mapping service.

CEDSCI will develop and promote a set of dissemination tools that will meet future dissemination needs. All datasets that we publicly disseminate will provide for use basic dissemination services, such as those for displaying and creating tables, graphs, and maps. This will provide a consistent experience for users, and reduce the costs associated with the siloed creation and maintenance of similar tools. Because we build these tools on a set of shared, componentized services, we can modify and extend them as needs and technology change.

The Census Data API is the foundation for provisioning the data for all of the data tools. This is the same API available to public developers, who can create tools using the same data and access method as Census Bureau developers.

### 3.3 Customer-Centric

*A customer-centric approach influences how we create, manage, and present data through websites, mobile applications, raw data sets, and other modes of delivery, and allows customers to shape, share and consume information, whenever and however they want it. (2012, p. 5)*

In addition to chartering CEDSCI, the Data Dissemination Task Force called for a greater investment, as well as enterprise management of the data tools and services deployed on census.gov. The Census Bureau is in the midst of a multiyear set of improvements to the website, beginning with improved search and navigation features and evolving to enhanced and more robust integration of datasets into our data tools. A collaborative effort manages features such as mobile apps, Quick Facts, Economic Indicators Dashboard, and the Pop Clock as an integrated system, delivering greater consumer utility and enabled by the new technology platform. The Web Transformation effort has grown into a wider “Digital Transformation” that is customer-centric.

Digital Transformation uses a customer and data-driven approach to strategically create and enhance digital products and services to benefit diverse user types. It operates under the goals of increasing customer satisfaction, raising awareness around agency statistics and programs, and growing our audience of consumers. CEDSCI will advance this strategy and establish the foundation for these enterprise capabilities, focusing first on a common, shared, enabling technology platform and the harmonizing of our datasets and other content. The foundation will benefit Digital Transformation, as well as other internal programs, other government agencies, our customers, private entrepreneurs, and nonprofits that develop products dependent on our API and dissemination services.

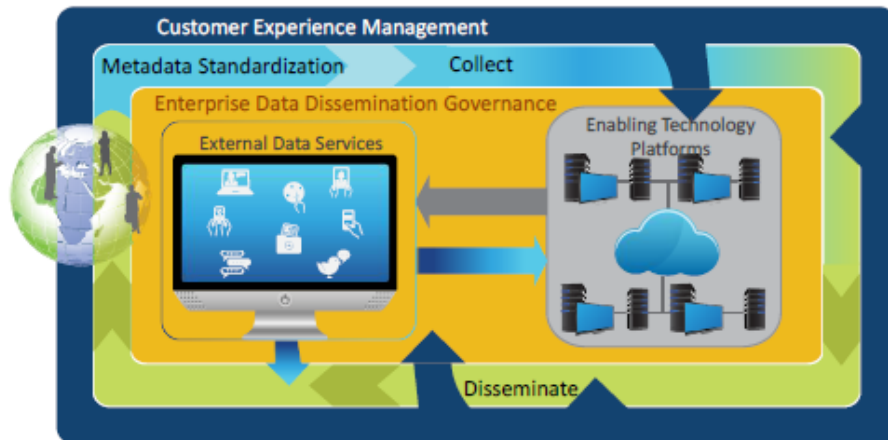
As CEDSCI matures during the transition, Digital Transformation evolves in parallel, delivering on the customer-driven Enterprise Information Services (EIS) called for by the Data Dissemination Task Force. EIS will leverage tools, analytics, and lessons from Digital Transformation, Customer Experience Management, and CEDSCI to continuously improve the customer experience with our content dissemination.

### *3.3.1 Customer Experience Management*

Insufficient standardized, routine consideration of consolidated customer interaction data exists. The Census Bureau captures customer contacts in disparate systems, such as call center logs, email case management, or paper logs, but there is little integration of the data and no standardized system to measure, compare, and contrast cases across offices or programs. As a result, the Census Bureau is currently unable to answer questions as simple as how many customers access data directly from the Census Bureau. We cannot reliably report how often customers typically use our data, how satisfied they are with finding, accessing, and using our data, or whether the Census Bureau is getting better at serving them.

Over the last several years, the Census Bureau launched an effort to begin an enterprise approach to CEM. The solution integrates some of the key customer engagement platforms in use at the Census Bureau, conducts analysis of the data from these platforms, and shares the data and the findings with management.





**Figure 5:** Customer-centric dissemination model

To deliver on a customer-driven dissemination vision, the Census Bureau will change its concept of customer relations, begin to think of “customers” as diverse users with differing needs, centralize the measurement and study of customer behavior and sentiment, and broadly share the analytical findings from this enterprise view of customers. The CEDSCI team will rely on continuous feedback through CEM to provide continuous enhancements to the transition.

CEDCSI will help guide future investments toward integrated tools and systems that capture, code, compare, and mine customer interactions, which will build an enterprise view of all customer behavior and customer sentiment. We will share the data from these systems, and more importantly, analytical insights developed from measuring and tracking the findings across the agency as widely as practical.

As a result, we will move to a dissemination model with continuous feedback and improvement. We will compare and contrast customer insights across our surveys and programs to achieve more complete data integration, and we will enhance data tools that will empower others to put our data to work for them.

### **3.4 Security and Privacy**

*A platform of security and privacy ensures this innovation happens in a way that ensures the safe and secure delivery and use of digital services to protect information and privacy. (2012, p. 5)*

The U.S. Census Bureau collects its survey and census data under the U.S. Code’s Title 13, which promises confidentiality to its respondents. Other laws, including Title 26 and the Confidential Information Protection and Statistical Efficiency Act also govern the collection and dissemination of our data. The Census Bureau also has the responsibility of releasing high quality data to the public. The Center for Disclosure Avoidance Research develops and improves disclosure avoidance methods to ensure that the Census Bureau effectively disseminates the maximum amount of high quality data about the nation’s people, housing, and economy while fully meeting the Census Bureau’s legal and ethical obligation to protect the confidentiality of respondents and the information

they provide. Program areas and the Disclosure Review Board review all data released to the public to ensure their appropriate protection from disclosure.

Additionally, extensive security plans cover all IT systems included in CEDSCI, which undergo rigorous security testing following a continuous monitoring model.

#### 4. Why Are We Innovating?

Keeping pace with the demands of an information-driven economy is important to the Census Bureau's core business – data collection. Strong Census Bureau brand awareness and acceptance drives higher response rates. The larger the share of households and businesses that value and use Census Bureau data, the higher the likelihood they will cooperate with requests to participate in one of our surveys. In many ways, CEDSCI is the perfectly paired mate to the consolidation of survey data collection processes. Both transformation efforts are connected and interdependent in the sense that they will share a technology platform, will deliver efficiency and quality improvements, and ultimately, will enhance the customer's experience when interacting with the Census Bureau – and grow our customer base. CEDSCI benefits include:

- **Lower costs** through elimination of duplicate systems and processes
- **Greater innovation** via standardized metadata and data access
- **Improved customer satisfaction and insight** into customer needs
- **Systematic quality assurance** across dissemination systems

As we modernize and centralize our data collection and processing activities, our dissemination efforts must similarly keep pace with technological innovations and changing expectations from respondents and data users alike. Deploying shared services across our diverse programs and datasets not only provides us with efficiencies, it exposes our innovations to more of the American public on whom we rely to produce the content we return to them. We must demonstrate to the public that the Census Bureau can innovate and adapt, lessen the burden of response, and meet our budgets and deadlines if we are to preserve their trust in our work.

#### 5. CEDSCI Status

To construct the new dissemination platform, the Census Bureau is employing an agile development methodology that emphasizes iterative development and integrated customer feedback. Our 2-week sprints aggressively incorporate datasets from Census Bureau programs (e.g., Decennial Census and American Community Survey) and features, including the standardized metadata model and state-of-the-art faceted search. Our 90-day release cycles regularly deliver alpha releases to internal and selected external users. Beta releases available to the public will begin in late 2016 and enable extended feedback from a broad spectrum of customers prior to the 2017 release of the platform as the Census Bureau dissemination system of record. Throughout CEDSCI alpha, beta, and production deployments, we will sustain the agile rhythm of frequent releases, elicitation of user feedback, and integration of feedback into upcoming versions.

## References

- Exec. Order No. 13571, 76 Fed. Reg. 24339. (2011). Retrieved from: <https://www.whitehouse.gov/the-press-office/2011/04/27/executive-order-13571-streamlining-service-delivery-and-improving-custom>
- Exec. Order No. 13576, 76 Fed. Reg. 35295. (2011). Retrieved from: <https://www.whitehouse.gov/the-press-office/2011/06/13/executive-order-13576-delivering-efficient-effective-and-accountable-gov>
- Federal Information Office (2012, May 23). *Digital Government: Building a 21<sup>st</sup> Century Platform to Better Serve the American People*. Retrieved from: <https://www.whitehouse.gov/sites/default/files/omb/egov/digital-government/digital-government-strategy.pdf>
- White House, Office of the Press Secretary (2012, May 23). *Memorandum for the Heads of Executive Departments and Agencies: Building a 21<sup>st</sup> Century Digital Government*. Retrieved from: [https://www.whitehouse.gov/sites/default/files/uploads/2012digital\\_mem\\_rel.pdf](https://www.whitehouse.gov/sites/default/files/uploads/2012digital_mem_rel.pdf)