Designing and Running the 2020 Census and
Building the Census Business Ecosystem

JITSWCF

Michael Thieme, US Census Bureau, Senior Advisor for IT and Operations

September 19, 2024

Census Data Stewardship – Ensuring Your Data Is Safe And Secure
## Our Culture Values Data Security

Data Stewardship is the formal process the Census Bureau uses to care for respondent information—from the beginning, when a respondent answers, to the end, when the statistical data products are released.

Data Stewardship is a comprehensive framework designed to protect information over the course of the information lifecycle, from collection to dissemination, and it starts with creating a culture of confidentiality that is based on the law and designed to maintain public trust.

Information Security

Data Stewardship

Confidentiality

Privacy
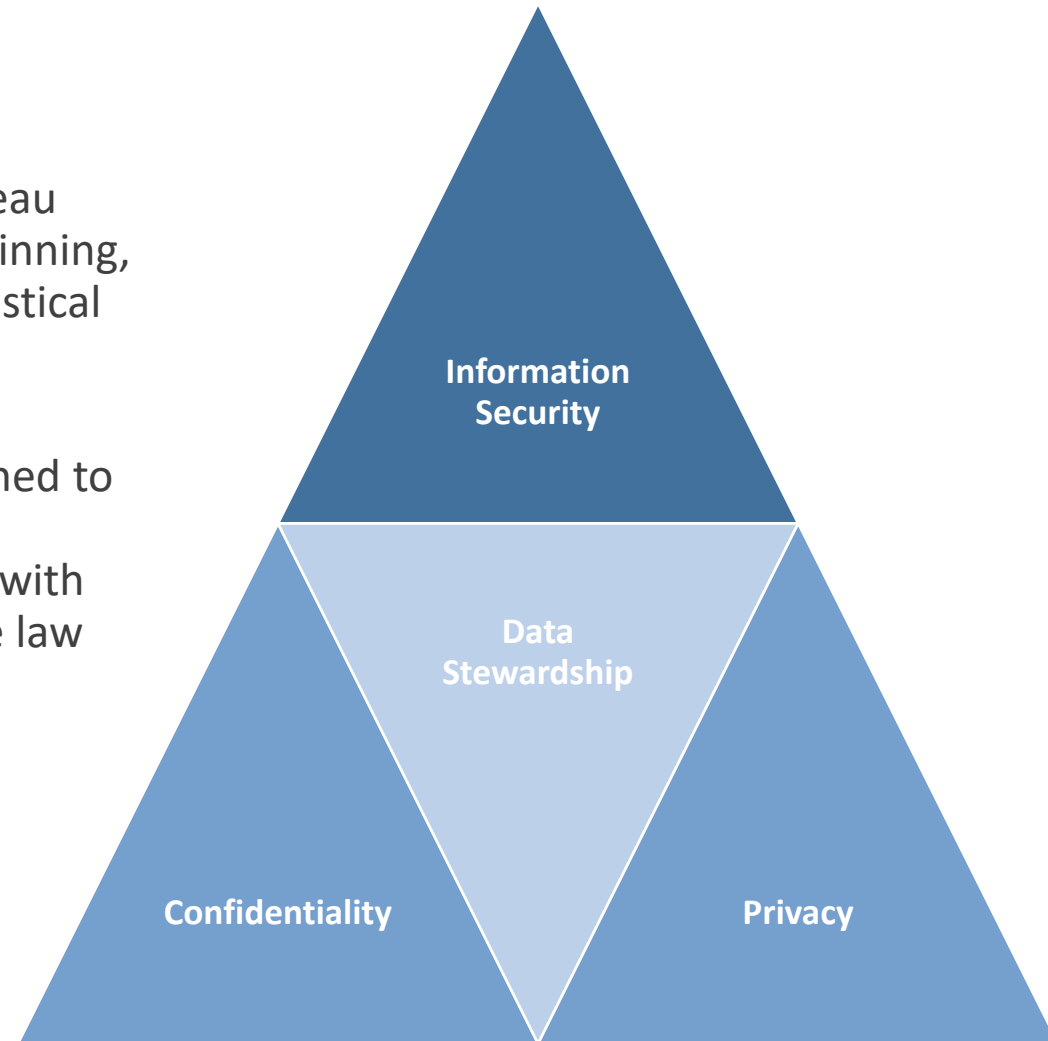
Secure System Design
Designed to Contain, Sustain, and Maintain Public Trust

Security

User Experience

Data Security

Performance
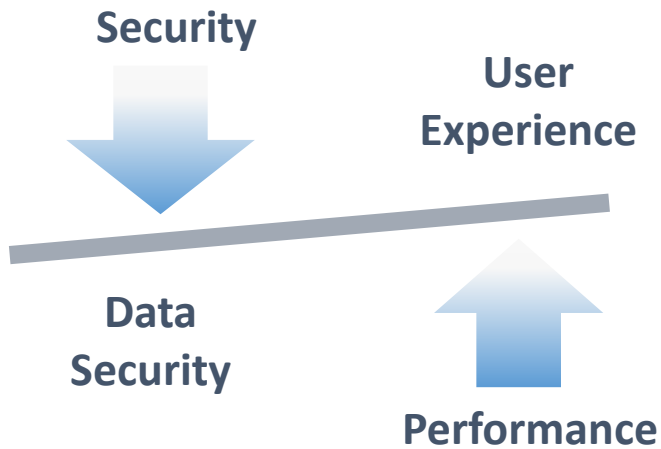
Census design is focused 1st on data security to protect respondents' data and 2nd on user experience so that respondents may confidently respond to the 2020 Census.
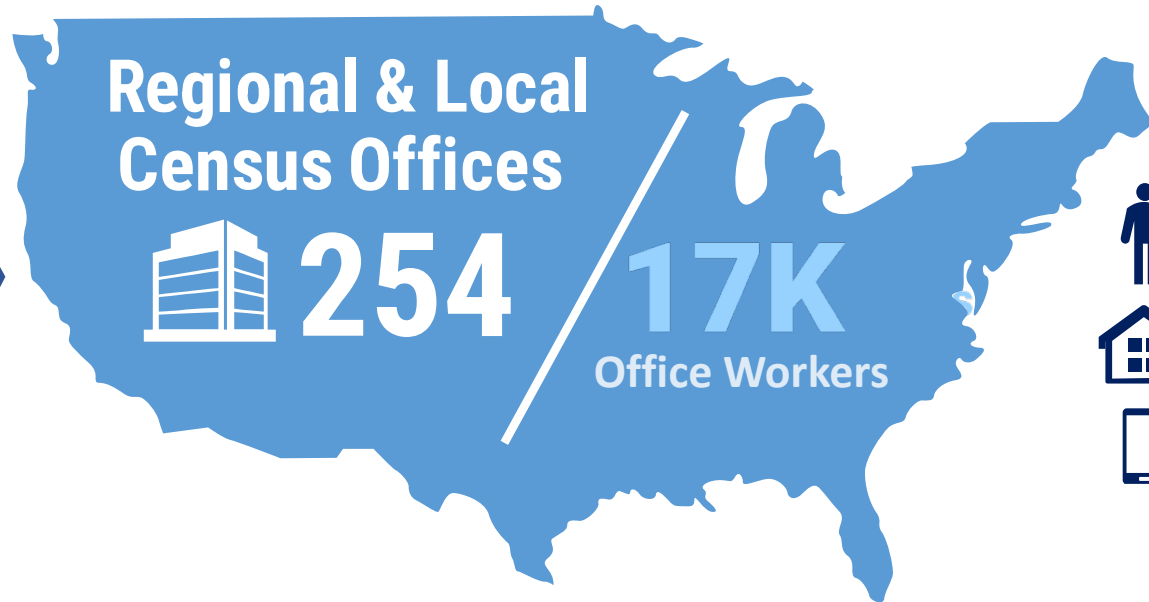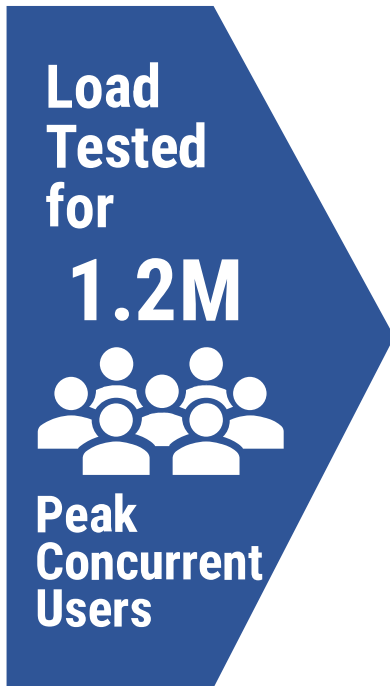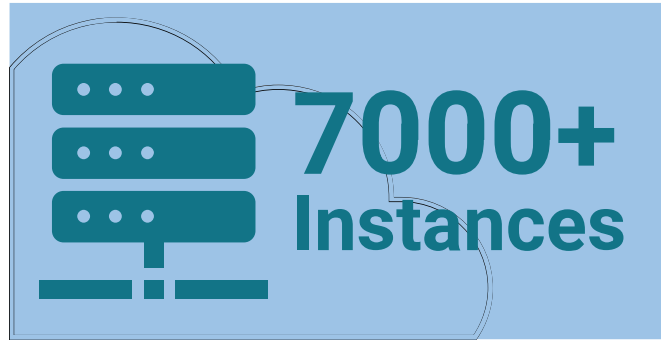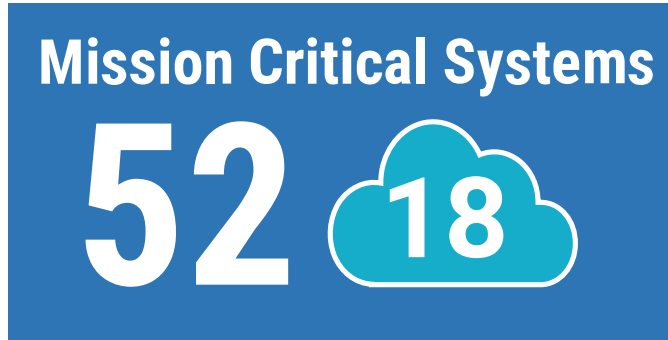
Contain Issues + Sustain Services = Maintain Public Trust

Employ "Physical Security" Techniques
- "Layer" physical entry with the appropriate level of security (doors, walls)
- "Isolate" separate areas within layers to efficiently handle interactions (lines, guards)
- "Lock down" valuables behind closed doors (vaults, safes)

United States® Census Bureau

Application Layer
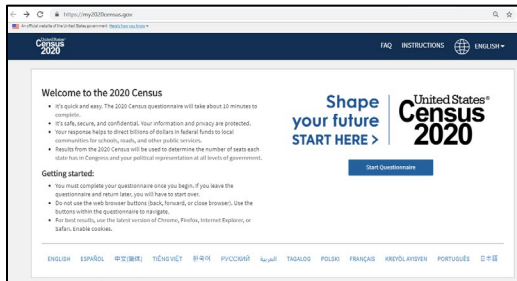
Service Concept	Design	Preparation	Operation

High availability service to US public

- Minimal impact to respondent during failure

- Multiple language support
- Responsive design
- 508 Compliant

Resiliency

1. **Multiple Software Apps** (different tech) for Internet Self Response (ISR)
2. **Redundancy:** AWS Cloud Regions; multiple Availability Zones
3. **Data Replication across Regions** (Census ID, Session Information)
4. **Native AWS resiliency features** (DynamoDB, Transit Gateway, Lambda, Multi-AZ RDS)

Scaling

1. **Primary ISR:** AWS Autoscaling
2. **Secondary ISR:** Cluster based scaling
3. **Both Apps:** "Contain & Sustain" strategy

Resiliency

1. **Execution / practice** of failover procedures
2. **Production Validation** conducted before go-live
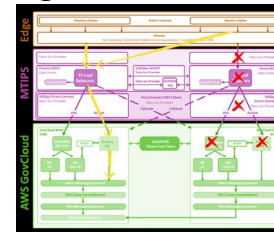
Scaling

1. **P&S tested to support 1.2M** concurrent users in each region (East & West)
2. **Determined breaking point** by simulating 2.4M concurrent users
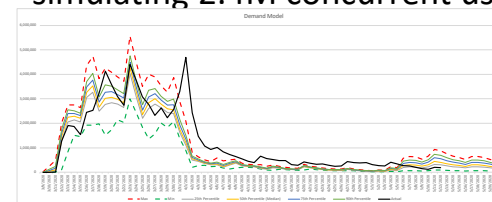
1. **Monitoring each service component**
2. **Splunk** as primary aggregator of system metrics
3. **Advanced monitoring dashboards** developed w/key stakeholders
   - Dashboard dry run during production validation
   - Highly relevant/well understood KPIs:
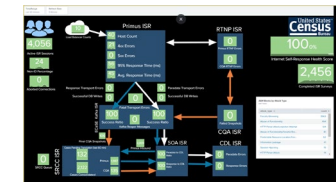
4. **Examples of resiliency/ scaling during ISR operations**
   - Network Outage
   - Patching
   - Failover for Call Centers

2020 Census
Cyber Threats

Cyber Threats

External Threats
Beyond Census Bureau Control

Compromising Respondent Device
Viruses on personal mobile devices, tablets, laptops, or PCs

Impersonation of the U.S. Census Bureau
Rogue websites, phishing

Compromised External Network Access
Website routing, internet providers

Invalid Response
Improper use of Census Bureau ID or Non ID ISR path

Internal Threats
Within Census Bureau Control

Disruptions to the Internet Self-Response Web Site
Denial of service, Computer-generated traffic

Data Breaches
Brute force, malware, phishing, insider threats

Compromised Employee Devices
Mobile devices, talets, laptops, or PCs

2020 Census
Major Contracts

Technical Integrator

Ensured 2020 Census System of Systems (SoS) integrated, performed, scaled, was secure, and met business objectives

Provided 2020 Census Cloud and On Premise Infrastructure

Provided systems integration and testing

AWS GovCloud (US)

2020 MPLS Network
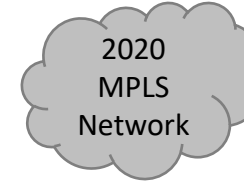
2020 Census Data Center

Census Questionnaire Assistance (CQA)

Provided call center capability for self response

Assisted respondents with responding to and completing census questionnaires

Provided Interactive Voice Response (IVR)

C-SHaRPS – Recruiting & Selection

Provided recruiting and selection COTS for up to 400K Schedule A Field Staff

C-SHaRPS Fingerprinting

Provided fingerprinting services for field staff

Decennial Device as a Service (dDaaS)

Provided Mobile Devices for Field Data Collection Operations

Software Integration, Cellular, Logistics, Shipping, Asset Management, Break/Fix and Disposition

Field IT Deployment (FITd)

Provided the IT infrastructure for 2020 Census field sites

IT & Telecommunications Hardware, Image Integration, Logistics, Shipping, Installation, Break/Fix and Disposition

**Regional Census Centers
Area Census Offices
Island Area Offices
Paper Data Capture Centers**

Decennial Service Center

Provided help desk services for 2020 Census Operations

2020 Census Contracts

Project/Program	Contract Type
Census Schedule A Human Resources and a Recruiting, Payroll System (C-SHaRPS)	Firm Fixed Price and Time & Materials
Decennial Device as a Service (dDaaS)	Firm Fixed Price and Time & Materials
Decennial Service Center (DSC)	Time & Materials
Field IT Deployment (FITd)	Firm Fixed Price and Time & Materials
Fingerprinting	Firm Fixed Price and Time and Materials
Questionnaire Assistance (CQA)	Cost Plus Award Fee
Technical Integrator (TI)	Time & Materials

2020 Decennial Census
Key Accomplishments for Successful Delivery

1. **COVID Resilience** - Collected nearly 80% of self-responses via the internet during COVID-19, helping meet the critical mission of the US Census of counting everyone

2. **Relentless Automation** - Modeled workloads, leveraged on-demand capacity with automated cost management, to establish a budget and trigger investigation when exceeded.

3. **DevSecOps** - Used automation and DevSecOps approach to maintain operations

4. **ACD Integration** - Leveraged Cloud native services integrated with an experienced active cyber defense solution (using Threat Modeling) to ensure a secure movement of mission critical systems to the cloud

5. **No Downtime** – Response systems did not have a single minute of downtime

United States® Census Bureau

Looking Forward

The Census Bureau's Business Ecosystem

Census Business Ecosystem
Background and Overview

Traditional Survey Methods Face Challenges

- Increased demand for timely statistics
- Reduced response rates and concerns about respondent burden
- Rising costs to conduct surveys

Business Ecosystem Video

Census Business Ecosystem Initiatives

The Frames Program (Frames): This initiative links the foundational datasets of addresses, businesses, jobs, and people that will ensure greater alignment across data sources, facilitate advanced research while simplifying data analysis, reduce duplication in effort, and ultimately enable the production of statistical products – possibly without fielding a survey at all.

Data Ingest and Collection for the Enterprise (DICE): This program provides a common suite of applications for data ingest and survey data collection that enable streamlined survey design, integrated response data across multiple collection modes, and near real-time collection metrics. The program delivers improved internal and external user experiences with more engaging and user-friendly interfaces.

Enterprise Data Lake (EDL): This is the cloud-based hub of the ecosystem designed to meet the evolving needs of data storage, management, and processing. It provides modern data access, streamlined security, multiple development environments, advanced data analytics, greater efficiencies in data processing, and centralized support for innovative data products.

Census Enterprise Dissemination Services and Consumer Innovation (CEDSCI): This program provides state-of-the-art tools and an improved user experience to facilitate discovery, visualization, and dissemination of Census data to the public.

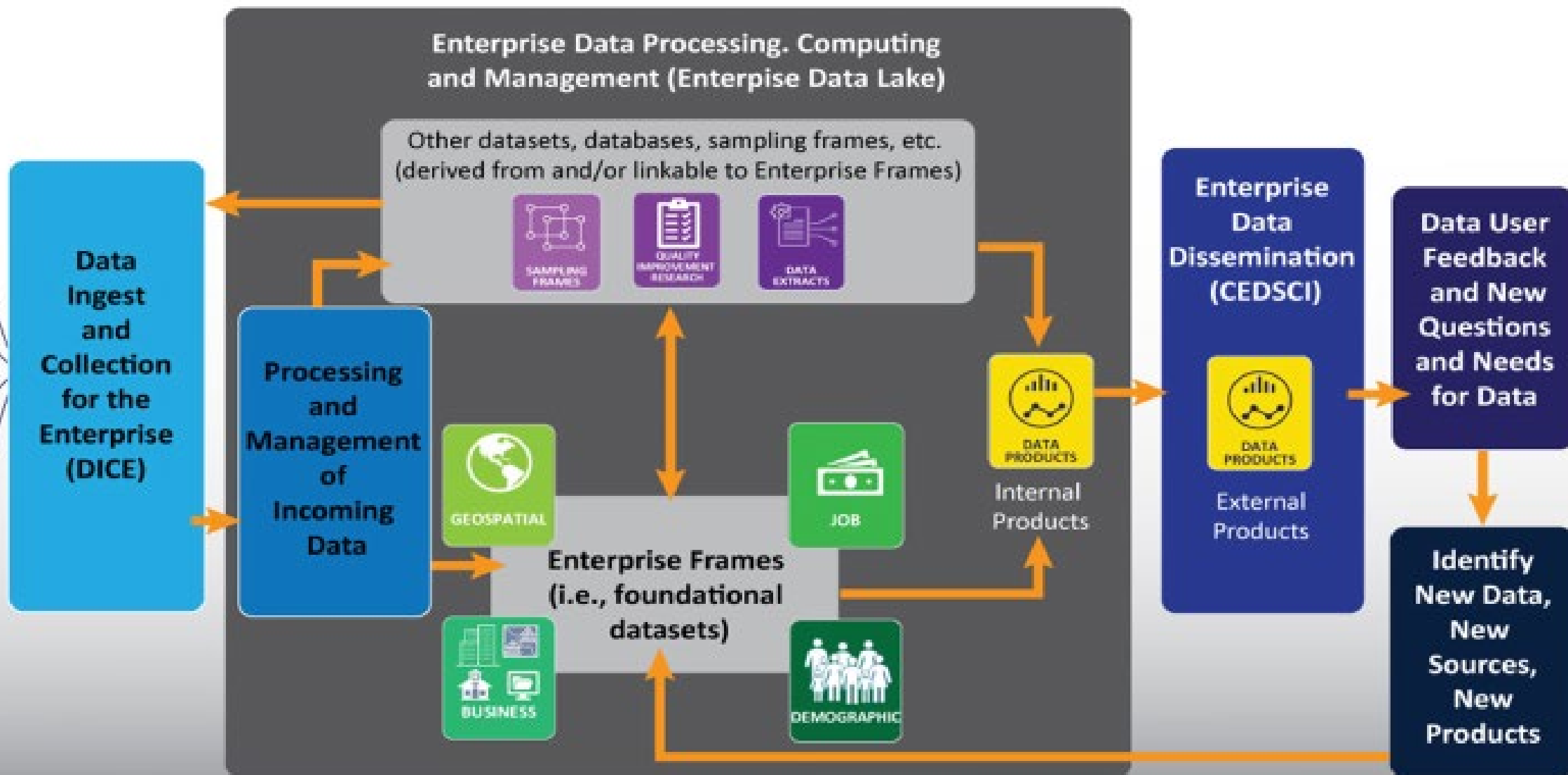Creating a Business Ecosystem to Modernize the Census Bureau's Statistical Foundation

Census Business Ecosystem
Where is the BE Now

Guiding Documents

- True North (November 2021)
- Phase One Plan (April 2022)
- Phase Two Plan (July 2023)

Focus

- Statistical Product First Mindset
- Onboard Surveys to the BE
- Move All the Bureau's Data Assets to the EDL (cloud)

Census Business Ecosystem Activity

AIES – Currently in full production surveying 375,000 businesses nationwide (launched 3/15/24)

Special Census – Currently in full production in Pingree Grove, IL. (launched 3/29/24). Ten Special Censuses booked

Current Population Survey (CPS) – In development, CPS-ISR Test (6/4/24 Start) and CPS-Field Test (3/1/25 Start)

Enterprise Data Lake (EDL) – Supporting production ops (above), continuing to develop and enhance secure, foundational data lake capabilities

Frames – Prototype Demographic Frame established in EDL. Work begun on standardizing data for all four frames in a common and shared set of schemas with common (meta)data definitions and structures

Communications – Biweekly "BE Insights" forum averaging attended by more than 5000 participants. Quarterly Program Briefings open to all, articles and blogs posted periodically

United States® Census Bureau

Questions

Michael T. Thieme,
Senior Advisor to the Deputy Director for IT and Operations,
U.S. Census Bureau

michael.t.thieme@census.gov