



Chief Data Officers Presentation

Ted Kaouk, USDA
Greg Fortelny, Education
Richard Allen, EPA

Advisory Committee on Data
for Evidence Building
February 19, 2021





Chief Data Officer Council

Overview



Authorization

- Established by the Foundations for Evidence-Based Policymaking Act of 2018



Vision

- To lead transformational change that improves the nation's ability to leverage data as a strategic asset



Membership

- All agency-level CDOs are members (including CFO Act and small agencies)



Purpose

- Establish government -wide best practices for the use, protection, dissemination, and generation of data
- Promote and encourage data sharing agreements between agencies
- Identify ways in which agencies can improve upon the production of evidence for use in policymaking;
- Consult with the public and engage with private users of Government data and other stakeholders on how to improve access to data assets of the Federal Government.



Council Goals

- Meet statutory requirements, including the required report to Congress
- Be a community of learning by broadly identifying best practices and resources to facilitate the implementation of Evidence Act
- Provide leadership on the delivery of the Federal Data Strategy (FDS) Action Plan



Components of Evidence

- **Policy Analysis** : Analysis of data, such as general purpose survey or program specific data, to generate and inform policy
- **Performance Measurement** : Ongoing systematic track of information relevant to policies, strategies, programs, projects, goals, and activities
- **Foundational Fact -finding** : Foundational research and analysis such as aggregate indicators, exploratory studies, descriptive statistics, and basic research
- ▶ **Program Evaluation** : Systematic analysis of a program, policy, organization, or component of these to assess effectiveness and efficiency



What is a Chief Data Officer?

CDOs enable data driven decision-making in a variety of ways, including:

- ▶ Developing and leveraging centralized agency analytics capacity
- ▶ Creating tools and platforms that enable self-service across their agencies and for the public
- ▶ Convening and coordinating agency-wide data governance activities
- ▶ Coordinating with any official in the agency responsible for using, protecting, disseminating, and generating data to ensure that the data needs of the agency are met;
- ▶ Engaging agency employees, the public, and contractors in using public data assets, and encouraging collaborative approaches on improving data uses;
- ▶ Supporting the Evaluation and Performance Improvement Officer of the agency in identifying and using data to carry out their functions
- ▶ Ensuring effective lifecycle data management
- ▶ Improving data infrastructure

Council Structure



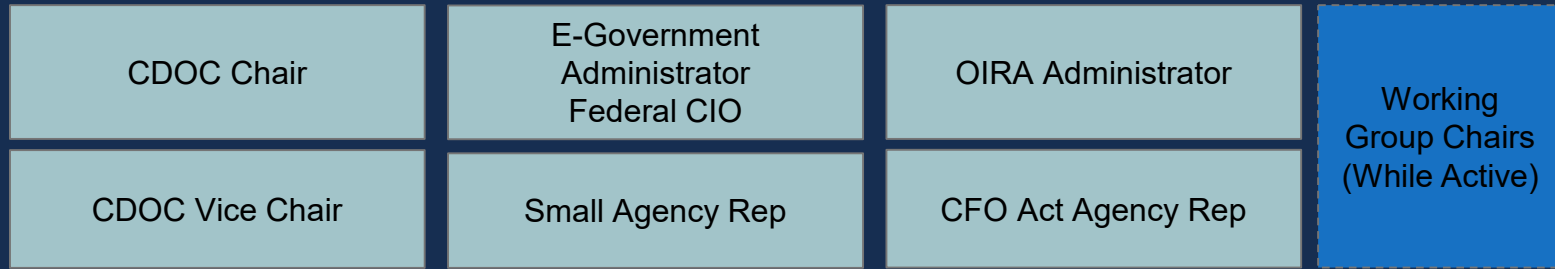
Executive Committee

1. Provide strategic guidance and oversee all activities of the Council, including identifying potential priorities, developing meeting agendas, and promoting collaboration.
2. Provide recommendations to the CDO Council on matters to be considered including the creation or completion of working groups, the approval of reports or other documents for publication.
3. Approve the Ex Officio membership, based on the advice of the appropriate interagency council.

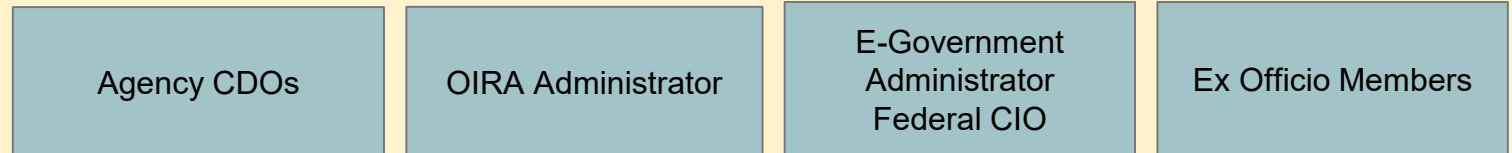
CDO Council Members

1. Attend the regularly scheduled meetings
2. Raise initiatives and issues that might have an impact on the overall Federal CDO community.
3. Review, consider, and provide input and recommendations on CDOC initiatives and be an active participant in the discussions.
4. Cast votes regarding issues presented to the council.
5. Participate in Council priority initiatives, working groups, or projects.
6. Provide input and participate in the development of CDO Council reports.
7. Foster collaboration and shared solutions with Federal, state, local, and international stakeholders.
8. Contribute best practices they are implementing for use by other Agency CDOs.

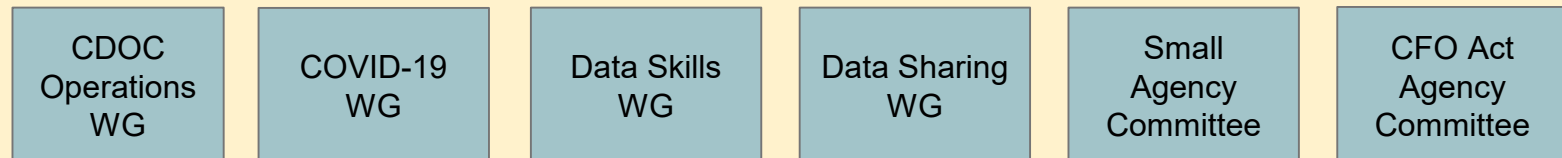
Executive Committee



CDO Council








Working Groups, Standing Committees, Communities of Practice





Focus Areas

-  COVID19 Data
-  Federal Data Strategy
-  Data Skills in the Workforce
-  Data Sharing & Dashboards
-  Data Inventories and Metadata
-  Stakeholder engagement



FY 2021 Activities & Initiatives

01	Regular meetings to enhance use of strategic data	<ul style="list-style-type: none">• CDO Council• Executive Committee• Working groups• Discussion groups
02	Stakeholder engagement	<ul style="list-style-type: none">• Internal• External
03	Strategic linkages	<ul style="list-style-type: none">• Interagency councils (EOC, CIOC, Privacy, Statistical, FOIA Officers Council, FGDC)• Public-private partnerships
04	Member developed projects	<ul style="list-style-type: none">• Data Skills Workforce Development• Governmentwide Dashboards and Infrastructure• Public Comment Analysis Tool• Interagency Wildland Fire Fuels Data Management
05	Federal Data Strategy	<ul style="list-style-type: none">• Provide input in the development of FY21 Action Plans• Provide leadership and support to CDOs as they implement the FDS Action Plans

USDA ENTERPRISE ANALYTICS CASE STUDY



USDA AND ENTERPRISE ANALYTICS

USDA's data and analytics initiative has made progress across two key dimensions:

ANALYTICS

Create analytics products (e.g., dashboards) to...



Support executive leaders perform **data-driven decisions**



Make data available at the 'Click -of-a-Button'



Keep leaders and key stakeholder groups **better informed**

DATA MANAGEMENT

Enable the organization with tools, technology, talent and structure to...



Make data and analytics tools **accessible and integrated** to solve cross-agency questions



Upskill and empower the workforce to perform analytics to address long standing challenges



Achieve **greater cost efficiencies and maximize ROI** on modernization investments

USDA ENTERPRISE ANALYTICS MODERNIZATION FY18-20

CREATE ANALYTICS PRODUCTS (E.G., DASHBOARDS)

500+ dashboards developed ACROSS Mission Area and CXO Offices

10,000+ users across all levels of leadership

40,000+ average views per month

~400,000 thousand in labor hours eliminated from manual tasks (e.g., status of funds)

ENABLE THE ORGANIZATION WITH TOOLS, TECHNOLOGY, TALENT AND STRUCTURE

150+ data sources brought together within department analytics platform (EDAPT) ACDO role established at all Mission Areas

Communities of practice, enterprise working groups, data governance board established

400+ Creator licenses in active use on the Server

Avoided \$10 million in potentially duplicative investment spend

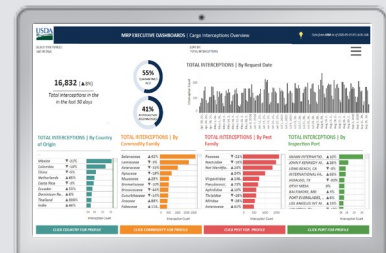
...with variety of analytics, machine learning, & artificial intelligence happening at various levels across USDA...



FOREST SUPERVISOR 360



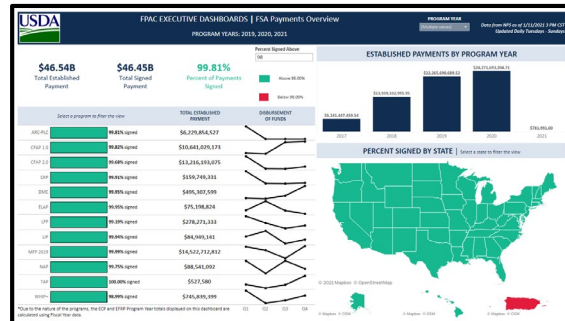
CORONAVIRUS FOOD ASSISTANCE PROGRAM



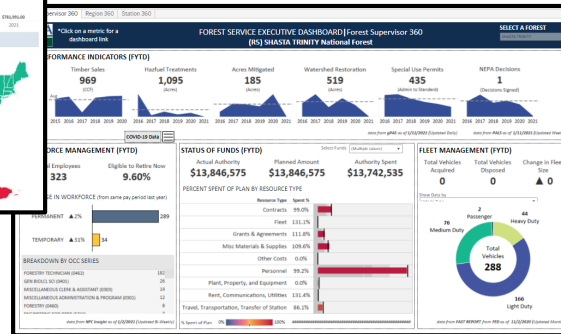
12
PEST INTERCEPTIONS

MISSION AREA DASHBOARDS OVERVIEW

In FY19 the effort expanded beyond Departmental Administration, with a central team assisting to create an initial suite of Mission Area specific program and administrative dashboards and enabling them to push forward independently on their data and analytics journeys.



FPAC FSA Payments Overview



NRE Forest Supervisor 360 Dashboard

- Developed and deployed program dashboards for all eight Mission Areas
- Established up Assistant CDO functions and accompanying analytics teams for long term sustainability

USDA FY21-23 DATA STRATEGY

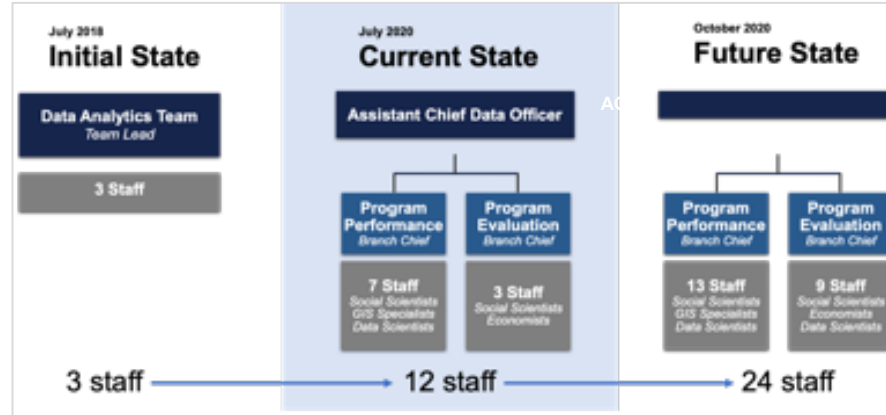
FY21 USDA DATA STRATEGY GOALS	1 Data Governance and Leadership	2 Data and Analytics Workforce	3 Common Data and Analytics Platform	4 Open Data
GOAL DESCRIPTIONS	<p>USDA will strengthen data governance and data leadership to enable a strategic approach to data and analytics development, infrastructure, and tools.</p>	<p>USDA will create a strong, data-driven culture by recruiting, retaining and retraining the workforce to acquire the needed data and analytics skillsets.</p>	<p>USDA will develop and leverage technology, infrastructure, and analytics tools to enable shared access and use of data to achieve our mission and to drive innovation.</p>	<p>USDA will support and promote effective data sharing to provide customers, stakeholders, and the public with deeper insights, value, and transparency.</p>
FOCUS AREAS	<p>Data Governance and Data Leadership</p>	<p>Centralized Analytics Teams Workforce Skills Development Fostering Communities of Sharing and Education</p>	<p>Enterprise Analytics Infrastructure and Tools Streamlining Data Collection Methods and Common Customer Records</p>	<p>Open Data Strategy Open Data Platform Sharing Data with Researchers</p>

USDA MISSION AREAS ASSISTANT CHIEF DATA OFFICER ROLE

Centralized analytics team



Building the bench



Dashboard adoption

Enhance User Experience

Make it Fun - Gamification

Regular Touchpoints

Training Opportunities & Resources

Make it Essential

Further dashboard development

Key Performance Indicator (KPI) Trackers

Workforce Activity & Efficiency

COVID-19 Portfolio Risk Analysis

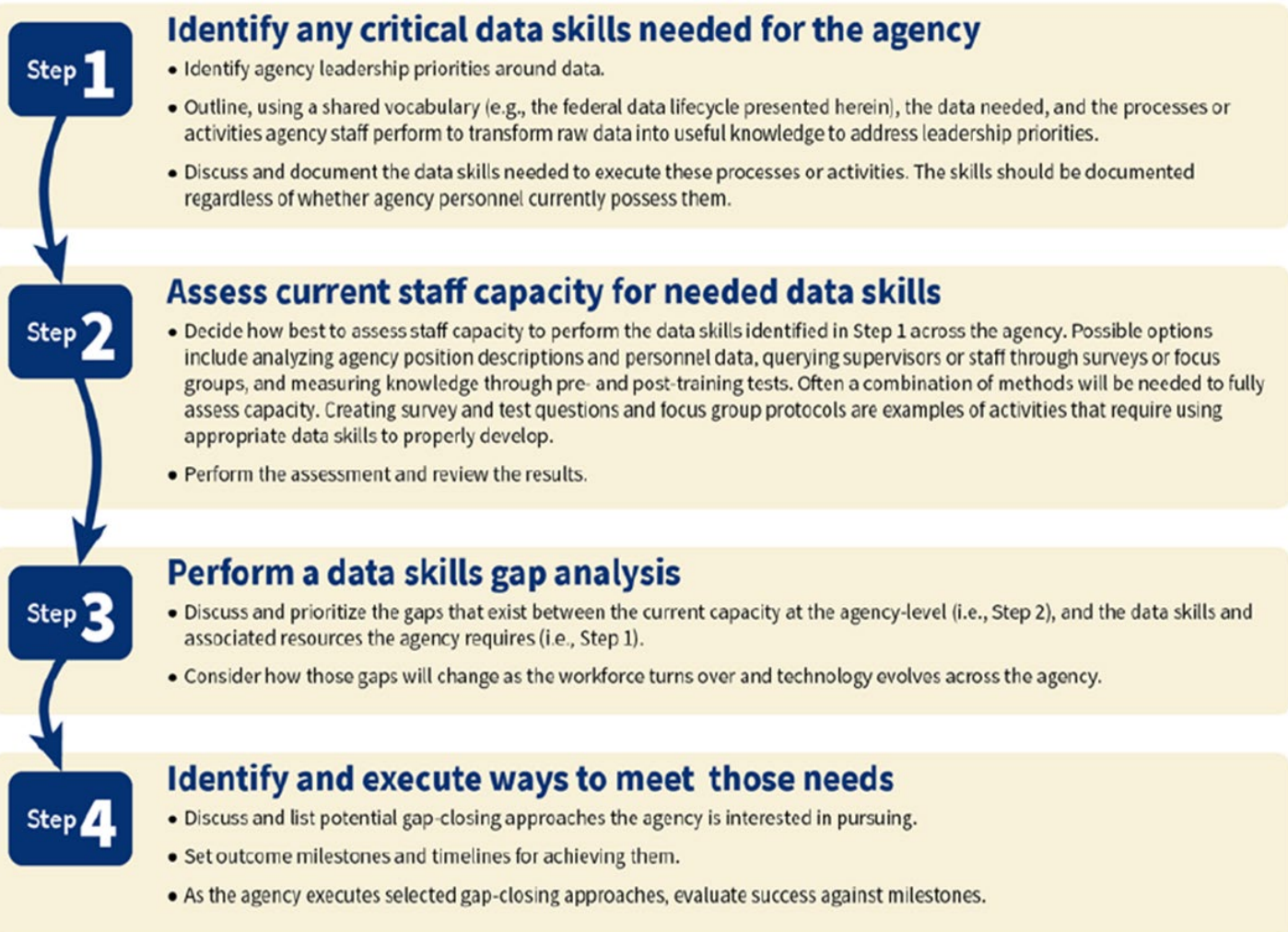
Data Governance Board Structure



USDA Data Skills Workforce Assessment

USDA's Data Skills Workforce Assessment will accomplish Federal Data Strategy **Action 4: Identify Opportunities to Increase Staff Data Skills**

Milestone
Perform an assessment of current staff data literacy and data skills
Conduct a gap analysis between the current staff's skills and the skills the agency requires
Based on the assessment, establish a baseline and develop a performance plan to close the identified data skills and literacy gaps



USDA COMMUNITIES OF PRACTICE

DATA VISUALIZATION COP

The ***purpose*** of the Data Visualization Community of Practice is to *promote the greater adoption of data visualization tools* that would assist all mission areas in achieving the three overarching goals of the federal data strategy: (1) *Building a culture that values data and promotes public use*, (2) *Governing, managing and protecting data* and (3) *Promoting efficient and appropriate data use*.

Objectives:

1. Share success stories and other notable work using Tableau dashboards across as broad a spectrum of USDA staff as possible.
2. Encourage greater utilization of the Tableau dashboards across all mission areas and USDA offices.
3. Encourage new and innovative uses of the Tableau dashboards across mission areas and USDA offices.

ADVANCED ANALYTICS COP

The ***purpose*** of the Advanced Analytics Community of Practice is to *identify, share and foster use of advanced analytic techniques to improve decision-making across USDA and its Mission Areas*. The CoP also aims to *demystify what analytics means* and what it does not, mapping techniques and methodologies to organizational and programmatic objectives and challenges.

Objectives:

1. Unearth and showcase institutional knowledge and experience in relation to analytical approaches so as to share best practices, methodologies and tools;
2. Develop a community of analytic professionals that can provide advice and assistance to others with similar research and/or business questions; and
3. Cultivate a data-driven organization through the development and enablement of the workforce.

PRIORITIES TO MODERNIZE DATA AND ANALYTICS

Technology can be removed as a barrier to the USDA data strategy



Imagery data can be shared more easily within and between agencies



Analysts can quickly and securely leverage data across a variety of analytics tools to meet ever-changing needs



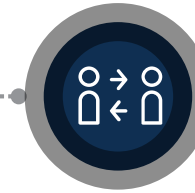
01 Establish single, common data warehouse platform with segregated and secure data storage and access per Mission Area requirements

02 Enable standard data ingestion toolset to bring data together from disparate sources

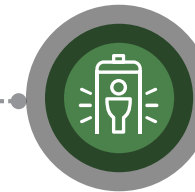
03 Establish standard visualization and data science toolset that includes Tableau, R, Python, SAS and ESRI and is easily accessible through a data science workbench

04 Launch common data cataloguing tool that provides transparency into the full set of data across USDA

05 Expand single, common Open Data Platform to enable efficient publication of USDA data to the public



ERS, ARS, NRCS, NRE can better share data to address important research questions



FSIS, RMA, APHIS can easily access AI & ML tools to better manage fraud compliance and pest outbreaks



USDA can enable Master Customer Data Management between agencies

Data Sharing

GREG FORTELNY, CHIEF DATA OFFICER

U.S. DEPARTMENT OF EDUCATION



Office of the Chief Data Officer

Agenda

1. Overview and Perspective
2. Product Examples
 - College Scorecard
 - Education Stabilization Fund Public Transparency Portal
3. Data Sharing Challenges
4. Intra-Agency Solutions
5. Inter-Agency Solutions



Overview and Perspective



U.S. Department of Education (ED)

- The Department's mission is to promote student achievement and preparation for global competitiveness by fostering educational excellence and ensuring equal access.
- Annual Budget of \$73.5 billion (FY21) and a total FTE of 4,069 (as of 9/30/20) across 16 principal offices.



Office of the Chief Data Officer (OCDO)

Responsible for managing and improving the Department's ability to leverage data as a strategic asset.

An analytics division is responsible for maintaining an analytical infrastructure that is responsive to strategic data needs, exploiting traditional and emerging methods to improve decision making.

A governance division oversees lifecycle data management, develops and enforces data governance policies, oversees the agency's information collections approval process, and creates public-facing data products,



Office of the Chief Data Officer

ED Data Sharing Product Examples

COLLEGE SCORECARD

EDUCATION STABILIZATION FUND PUBLIC TRANSPARENCY PORTAL



Office of the Chief Data Officer

College Scorecard

A U.S. Department of Education tool designed to help students make informed decisions about their education options after high school, bringing together information on college costs, graduation rates, student loan debt, post-college earnings, loan repayment rates, and more.

The screenshot shows the College Scorecard interface with the following details:

- Header:** U.S. DEPARTMENT OF EDUCATION College Scorecard. Navigation links: Home, About the Data, Search, Compare.
- Filters (Showing Results For):**
 - Schools (selected) / Fields of Study
 - Search by Name: Type to search
 - Location: State (Alabama)
 - Academic Fields: Start typing to search
 - Degrees/Certificate: Certificate, Associate's Degree, Bachelor's Degree
 - Graduation Rate: ≥ 50 %
 - Average Annual Cost: ≤ \$0 k
- Results (13 Results):**
 - Tuskegee University** (Tuskegee, AL): 2,529 undergrads. Graduation Rate: 76%. Salary After Completing: \$20k-68k. Average Annual Cost: \$35k.
 - Auburn University** (Auburn, AL): 24,147 undergrads. Graduation Rate: 74%. Salary After Completing: \$21k-75k. Average Annual Cost: \$24k.
 - Samford University** (Birmingham, AL): 3,524 undergrads. Graduation Rate: 70%. Salary After Completing: \$24k-55k. Average Annual Cost: \$29k.
 - The University of Alabama** (Tuscaloosa, AL): 32,177 undergrads. Graduation Rate: 67%. Salary After Completing: \$21k-76k. Average Annual Cost: \$22k.
 - Birmingham-Southern College** (Birmingham, AL): 1,265 undergrads.
 - Remington College-Mobile Campus** (Mobile, AL): 313 undergrads.
 - Athens State University** (Athens, AL): 2,732 undergrads.
 - University of Alabama at Birmingham** (Birmingham, AL): 13,486 undergrads.



Office of the Chief Data Officer



>6.5 million visitors



>2,500 API users



>130,000 dataset downloads

Education Stabilization Fund Public Transparency Portal



U.S. DEPARTMENT OF EDUCATION

CARES Act: Education Stabilization Fund



What is the Education Stabilization Fund (ESF)?

Part of the \$2.2 trillion Coronavirus Aid, Relief, and Economic Security (CARES) Act, the Education Stabilization Fund (ESF) allotted \$30.75 billion to the U.S. Department of Education to award grants to States (SEAs), districts (LEAs), and institutions of higher education (IHEs) in response to COVID-19. [More Information](#)

Major Components of the ESF

\$13.2 billion

ESSER

Elementary and Secondary School Emergency Relief

ED awarded ESSER grant funds to State Educational Agencies (SEAs). The SEAs awarded grants to Local Educational Agencies (LEAs) to address the impact of COVID-19 on elementary and secondary schools across the Nation.

[More Information](#)

\$3 billion

GEER

Governor's Emergency Education Relief

ED awarded GEER grant funds to States' Governor's offices. States decided how to award funds within their state to best meet the needs of students, schools, postsecondary institutions, and other educational organizations affected by COVID-19.

[More Information](#)

\$14 billion

HEER

Higher Education Emergency Relief

ED awarded HEER grants to institutions of higher education (IHEs) through several different programs. More than \$6 billion is designated for students as emergency financial aid grants to assist with their expenses arising from the disruption of campus operations due to COVID-19.

[More Information](#)

Funds Awarded

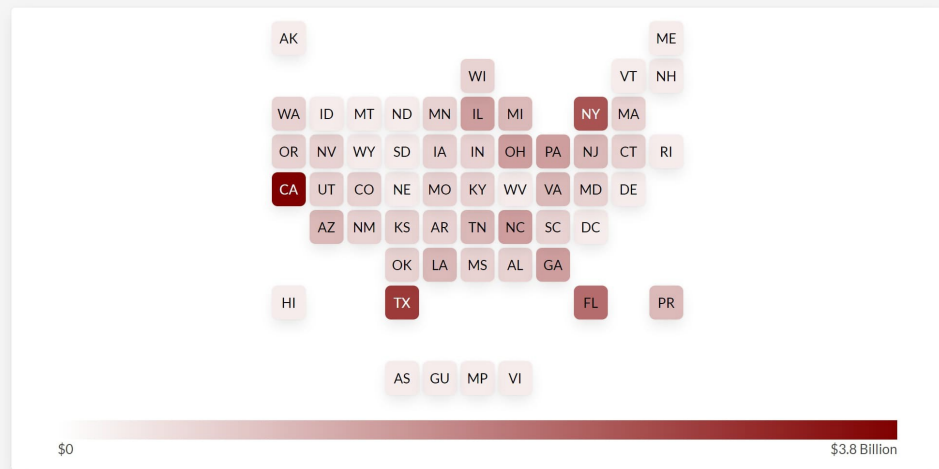
Data reported through September 30, 2020

Awarded Spent % Spent

View Funds by Program

Total ESSER GEER HEER

Includes Outlying Area ESF-SEA and ESF-Governor funds



Click on a state/outlying area or select below for more information:

A U.S. Department of Education application providing public transparency and accountability to the \$30.75 billion provided through the Coronavirus Aid, Relief, and Economic Security (CARES) Act.



Office of the Chief Data Officer

Data Sharing Challenges



Office of the Chief Data Officer

Data Sharing Challenges



Significant attention has already been placed on technological solutions, especially those catering to the secure transfer or integration of data across organizational boundaries.



Less attention seems to be placed on people and processes. Until these components are resolved, data sharing will continue to be an ad hoc and costly exercise inhibiting participation.

Individual agencies (intra-agency) must take steps to ensure their data management practices lend itself to secure sharing. Government (inter-agency) must take proactive steps to reduce costs and increase perceived value of secure data sharing.



Data Sharing Solutions

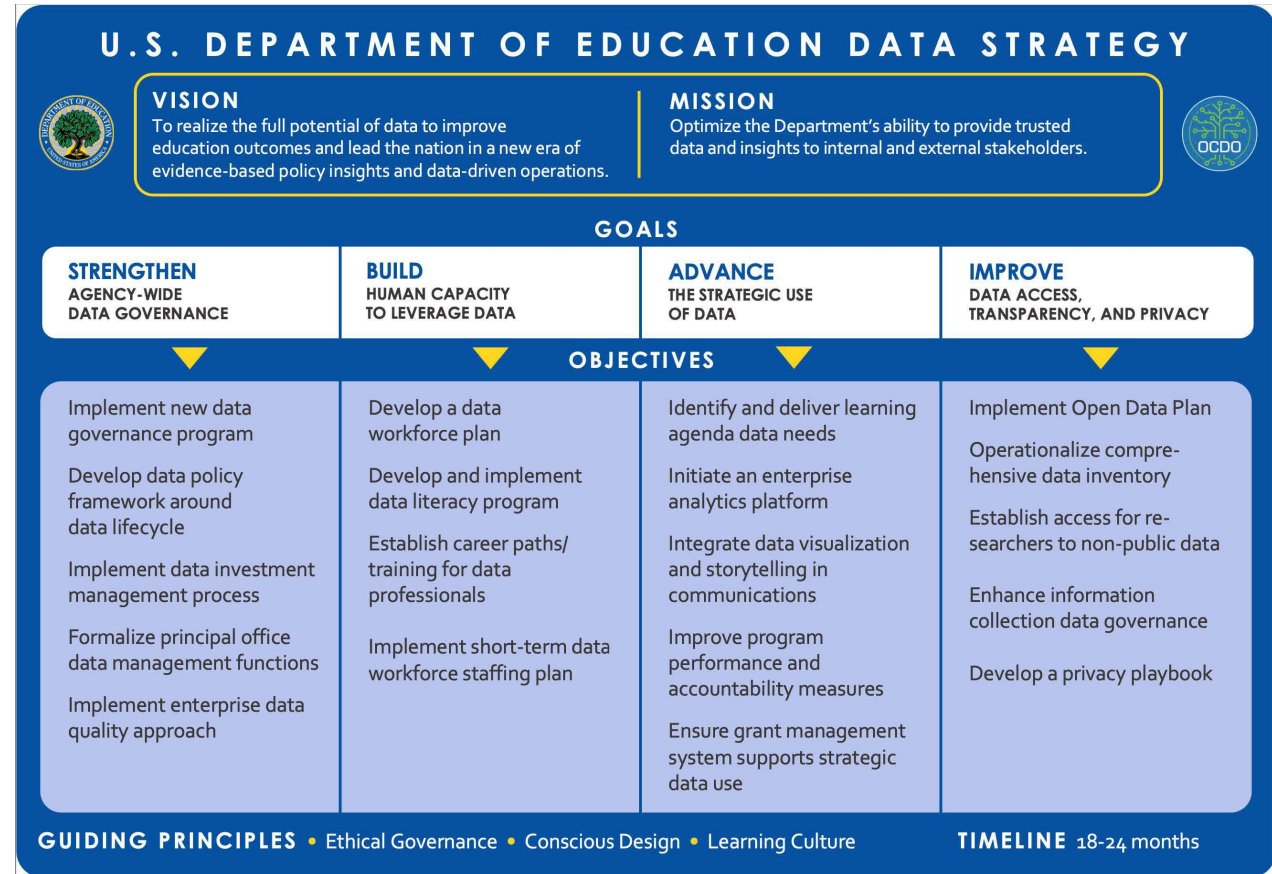


Office of the Chief Data Officer

Intra-Agency Efforts:

Department of Education Data Strategy

- The U.S. Department of Education published its first-ever Data Strategy in December 2020.
- It recognizes that a holistic point of view of its data assets is increasingly necessary to effectively meet mission objectives and establishing funding priorities.
- It also seeks to ensure the agency's data assets lend themselves to "FAIR" principles.



Office of the Chief Data Officer

Inter-Agency Efforts:

CDO Council Data Sharing Working Group

FOUNDATION

- The Foundations for Evidence Based Policymaking Act recognized that the full value of data can only be realized when it is securely accessible to others including the public. It mandated requirements, but challenges persist in practical implementation.
- The Commission on Evidence-Based Policymaking identified several legal barriers and bureaucratic processes that inhibit data sharing. Those hurdles persist today and regularly require leadership intervention to overcome.

LOOKING FORWARD

- In its simplest form: An agency's perceived value gained from a data sharing relationship with others must exceed the agency's perceived costs.
- The CDOC Data Sharing Working Group aims to identify challenges—legal, policy, technical, and measurement—that inhibit data sharing.
- It will also recommend, develop, and execute solutions that make sharing easier with strong privacy protections.



Questions?

GREGORY.FORTELVNY@ED.GOV



Office of the Chief Data Officer



Enhancing a Data Driven Culture



Capture Hearts and Minds

We must capture the imaginations of people who manage and use data, at all levels, if we are going to achieve transformative, meaningful change in our collective use of data for evidence building.

Working Across EPA to Address Data Culture

- Talk to Everyone to Build Trust
 - Had data focused discussions with every component
 - Shared summary results
- Invite Everyone to Participate
 - Data governance body is a mob
 - >60 people from every sub-unit
 - Work with peers and senior executive class
- Listen to What the People Say
 - Vet priorities
 - Bring people on to help



Focus on What the People Need

- **Data Standards:** Comprehensive approach for data standards
- **Data Quality:** Comprehensive approach for better data quality
- **Data Governance:** Policies, guidance, training
- **Data Management Resources:** People and money
- **Awareness of Data:** Better communication of data resources
- **Accessing Data:** Permissions, technology, formats and standardization
- **Integration of Data:** Gaps in connecting data (programmatic & technical)
- **Coordination/Collaboration:** Challenges across regions and program offices, with states, municipalities, and academics
- **Data Stewardship:** Improved stewardship community
- **Data Skills:** Training in collection, storage, interpretation, analysis
- **Data Security:** Need for data management guidance on handling, protecting and transporting data

Foster Common Understanding

Standards – are highly domain and technology specific.



Format



Definition



Structure



Metadata



Tagging



Transmission



Use



Management

Quality – has many dimensions importance of each varies with intended data use.


- Relevance
- Accessibility
- Timelines
- Punctuality
- Granularity
- Accuracy and reliability
- Coherence
- Scientific integrity
- Credibility
- Computer and physical security
- Confidentiality

To Share or Not to Share: Why is the Question

Obvious reasons to share – Make our argument

Benefit for themselves, their group, the public, the environment...

Reasons not to share – Build better counter arguments

- Data are classified or otherwise controlled
 - Data quality are not where we want them to be
 - Data quality are without reproach for one use but not for all
 - Data are complex and would require significant explanation (context)
- 

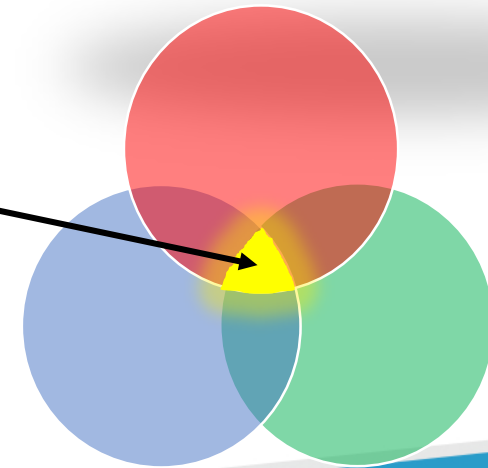
Data Management Focus: Facilities

- **Project Goals**

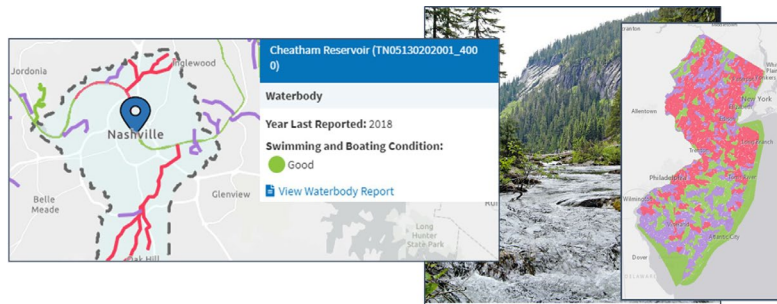
- Provide an entirely fresh look at facility data in EPA
- Gather requirements for business needs (current and future)
- Identify business processes that rely on facility data
- **Redesign EPA's facility management processes in order to better meet business needs**

- **Overall Goals**

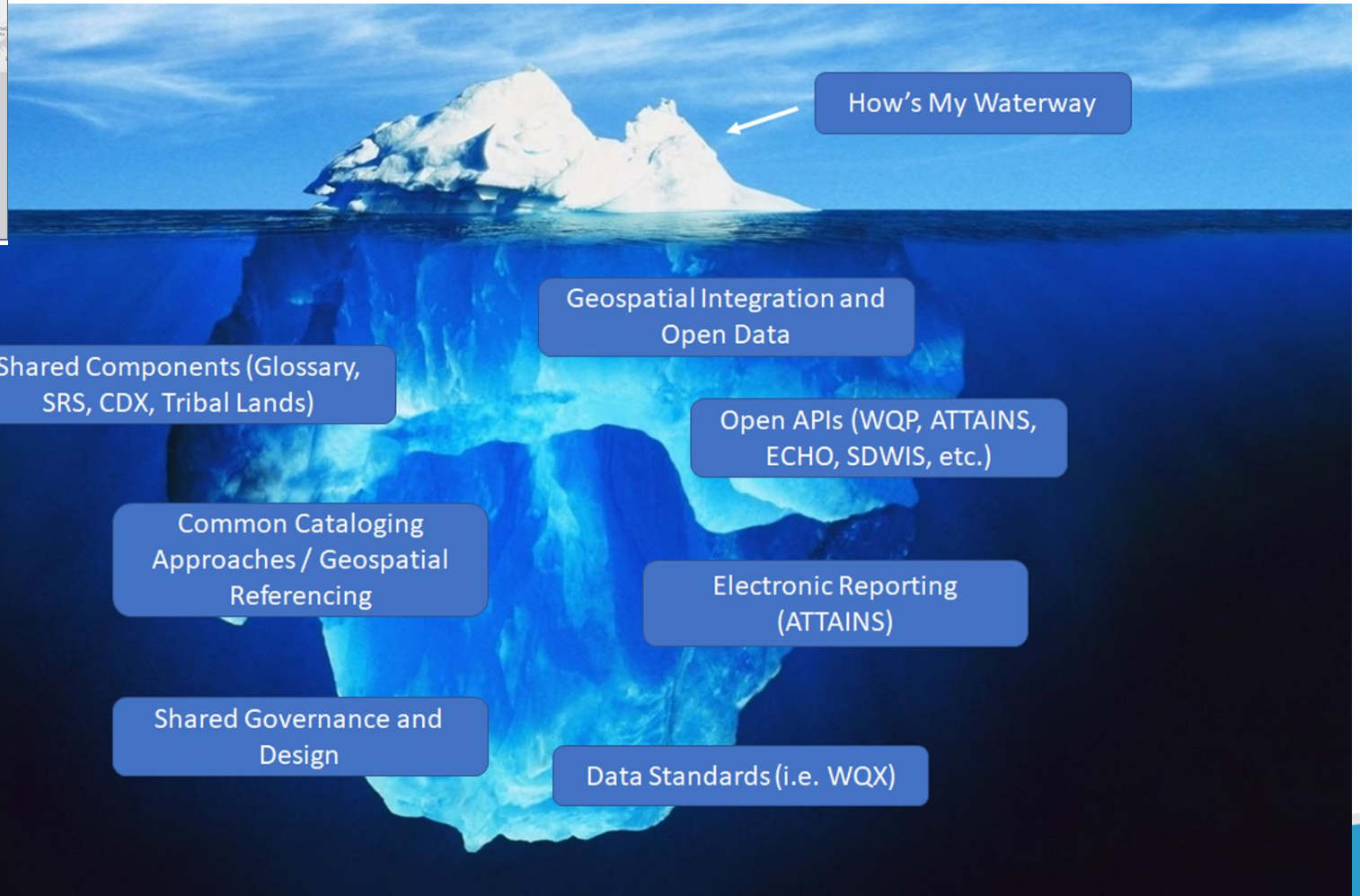
- Build process to improve priority data management
- Identify key activities that rely on priority data
- Iteratively improve those data assets



How's My Waterway: Built on Data Sharing




- All information is based on a service (API first)
- Data standards matter
- Integration of data across multiple systems
- Shared ownership between states / tribes / EPA
- User centered design
- Code is reusable and publically available on GitHub



Respect the Data but Drive at Purpose

- Too often we become slaves to the data we manage
- Too often we spend our max efforts on controlling the data
- Look above the data to the people and their needs
- These needs will inform us about data and its fitness for purpose
 - Quality
 - Standards
 - Utility
 - Other features specific to the intersection

Points of Focus

- Communicate, communicate, communicate
 - Invite participation
 - Build a common understanding
 - Promote data sharing
 - Focus on programmatic needs not specifically data needs
 - Manage for data services and integration
 - Focus on people and build community
- 

Culture change takes time.

Capturing imagination and building coalitions around new ways of thinking can be the lightning that serves as a catalyst for cultural change.

