

PERSPECTIVES:

The Journey to Data Value:

Navigating the pathway to data consolidation

June 2017

by Alissa Cruz, Lyn McGee, and Dr. Paul Eder



You will never
reach your
destination
if you stop and
throw stones at
every dog that
barks.

Winston S. Churchill

CAN DATA CREATE VALUE?

The fact that data exists does not make it valuable. Data only creates value when it is combined with the way it is used – business value, social value, and the holy grail for government... demonstrable transparency. The starting point for using data, of course, is the ability to find the data and have a common understanding of the data definitions. Thus, we enter the brave new (and rapidly changing) world of data-driven economic value. Standardized, consumable data is the foundation for decisions that increase efficiency and effectiveness. How do we get there and why would anyone want to go on the journey?

Agencies and organizations that are successful in cracking the code on data availability will discover that they require a higher level of skill from employees and leaders so they can effectively use data to drive decisions. The new sought-after talent will be critical to agency sophistication in assessing the possibilities for what can be done against all possible outcomes. Agency program owners will need to consider how data users, stakeholders, and customers are being equipped with the skills needed for this level of sophistication to drive decisions, improve processes, and increase customer satisfaction.

Creating consumable, standardized data that can be either logically or physically consolidated is the starting point for deriving value from agency data. It sounds simple, but anyone who has done it will tell you that the challenges are surprisingly not technical—or mostly not technical. Delivering accessible data that can be used for decision making is all about politics, policies, and people.

Data integration initiatives are complex and risky—many, if not most, agencies end up with static, one-time solutions that can't scale to meet future needs, do not increase the acumen of the people who create and consume data, and are inflexible in the face of new standards and analysis tools.

Planning the journey carefully before you start out will ensure that your agency or program does not end up in unexpected places with unintended consequences impacting critical underlying business processes.

IMAGINE:

- Using agency-wide data to identify duplicative activities, bringing together the program owners, identifying best practices, and creating a more efficient, effective program.
- Simultaneously accessing data from diverse systems across organizational units.
- Discovering and correcting data prior to making decisions.

For a typical Fortune 1000 company, just a 10% increase in data accessibility will result in more than \$65 million additional net income.

Forbes

Government regulations requiring agencies to be more transparent and more accountable to the American people are not going away. To the extent that data consolidation and reporting is not mandated, agencies can make choices about reporting on and creating access to financial and performance data.

There is a real risk for agencies in deciding not to follow a path to data standardization and consolidation. The choice indicates a confidence deficit regarding the quality and usability of agency data assets. Taking it a step further, it is likely that agency management cannot rely on data to inform decision making, limiting the capacity to achieve improved efficiency and effectiveness in delivering the mission.

This clearly impacts the effectiveness of leadership decisions, but may also mean there is a good chance no one is paying attention to data quality for existing agency datasets. Data users notice when data is complete or error free. When agencies decide not to collaborate and adopt data standards, other internal and external groups who have developed their data acumen and sophistication may blind-side them with data issues when they try to use agency data.

THE DISRUPTION OPPORTUNITY

The Center for Organizational Excellence, Inc. (COE) has over 15 years of experience with data standardization, integration, consolidation and reporting for government-wide and agency-wide initiatives like the Office of Personnel Management's Enterprise Human Resources Initiative (EHRI) and DATA Act implementation at the Department of Homeland Security.

Processes to standardize and consolidate data exist. Creating a data-centric culture, however, is foundational for creating sustainable value. It requires a disciplined project management approach and a willingness to create pain in the pursuit of excellence. Culture change can be an incredibly valuable by-product of the data consolidation process:

- **Planning for data integration.** You must identify the data you need, why you need it, what you want to do with it, and how it will integrate seamlessly across systems.
- **Managing the change.** You must consider the needs of internal and external stakeholders and integrate efforts around other concurrent initiatives.
- **Implementing new processes and policies.** You must implement processes that properly feed the data integration system to include accurate ingestion, transformation, validation, and reporting of data.
- **Evaluating and modifying the course.** You must review and analyze your current and future state to determine the most effective ways to align the system and processes through updates and refinements, or system lifecycle refreshment.

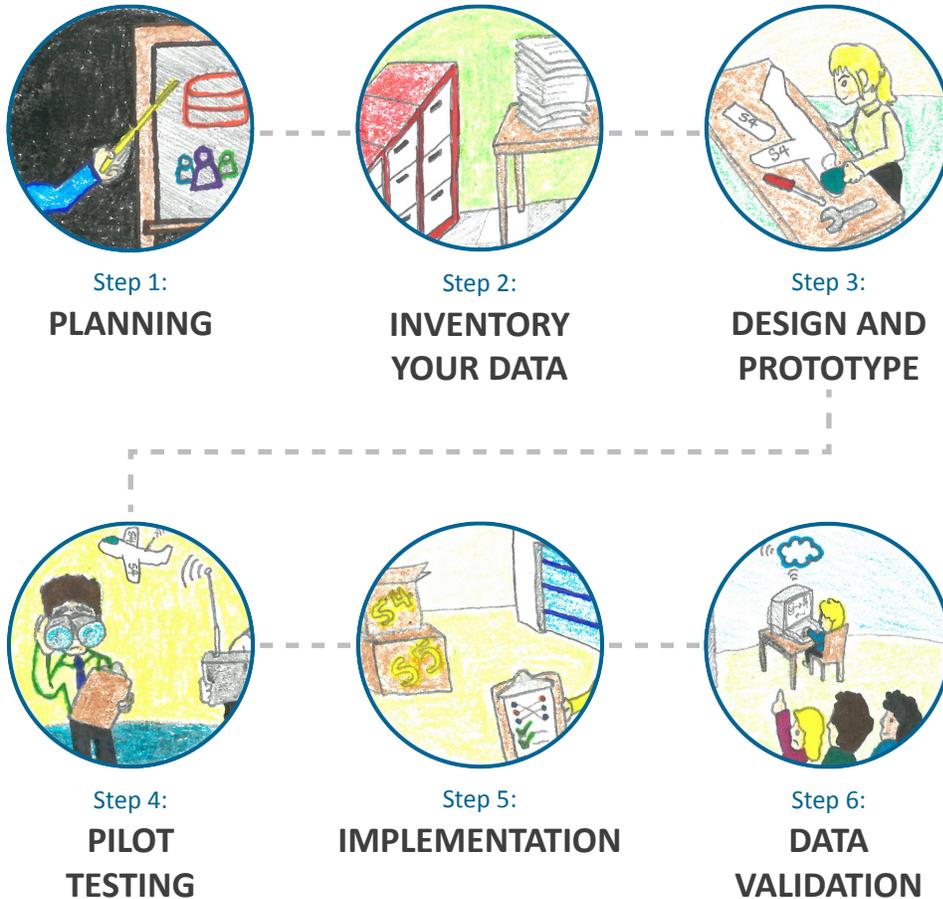


THE APPROACH

A set of common activities is performed in any data consolidation effort. We break them down into six areas, with the caveat that numerous iterative and cyclical events occur within each step.

A successful consolidation effort is not about the specific activities performed or the order in which they are done. It is absolutely about how the processes supporting the overall initiative are executed. The people, policy, and cultural implications within each step are the critical factors making the difference between short-term success and the long-term viability of your data solution.

Done right, this data consolidation approach will pay off for years to come. In the following sections, we share some of our insights into a data consolidation process that builds a sustainable solution, delivering lasting value for our agency clients and COE today and long into the future.



Data has a soul.
It will outlive you,
your project, the systems
you put it in, your agency,
presidents, and countries.
It is absolutely essential
that our data solutions are
built for the future. Solutions
built only for today's needs
are already obsolete.

Marcel Jemio
Chief Data Architect, OPM



Step 1: Plan the Integration

The critical decision that must be made in the planning phase is whether you will centralize the data consolidation effort or build a distributed solution that leverages the data owners for the various data sets that you wish to consolidate. As you identify the business drivers for data consolidation and create a vision for the agency's long-term data integration capabilities, a set of competing imperatives surfaces. Decision parameters like the timeframe in which you need to create the consolidated data set, the degree of influence you wield over impacted stakeholders, data users, and subject matter experts (SMEs), or the resources available for the effort start shape the project plan and direction.

A closer look at these imperatives, however, likely reveals short-term issues that can create risk for a viable long-term data capacity within your agency. For example, project leaders may be tempted to meet urgent requirements by building a centralized IT solution with the capacity to apply data standards and store data from a broad range of systems, which gives them control and removes many external dependencies. The alternative is collaborating with data owners to implement new data standards within the source systems, performing validation locally, which takes time, training, patience, and is subject to the limitations within legacy systems.

Placing responsibility for data standardization, quality, and validity in the hands of data owners and source systems is a key step forward in creating a data-centric agency culture. Data owners develop competency around inventorying their data, defining and implementing new standards within their systems, developing and efficiently executing robust test plans, building remediation processes to address errors, and a host of other skills that result in expanded capacity and data sophistication for your agency.



- Gain leadership support
- Establish a PMO, project charter, governance plan
- Develop a change management plan
- Create cross-functional working groups

Step 1:

PLANNING





Step 2: Inventory the Data Environment

It goes without saying that before you can standardize and consolidate data, you must have a complete picture of the data within your agency, where it lives, the processes that create it and consume it, how and where it interconnects, what the current data definitions are, and who the stakeholders are for any given system.

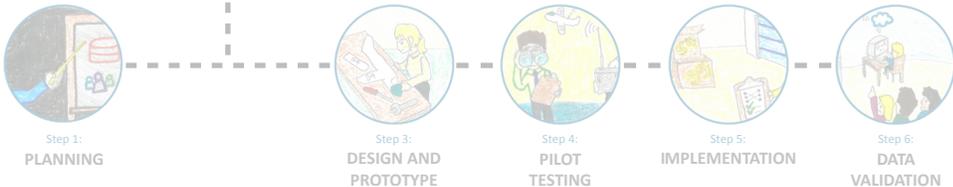
One way to build your data map and increase data acumen across the agency is to engage stakeholders in creating a standardized, templated view of existing data across your agency. The more common approach is to do a data call and then normalize the input to create an apples to apples picture of the agency’s data assets. There are some unintended consequences with this approach, however, not the least of which is that the central group that grooms the data call results will have to do it every time there is a new data integration requirement.

Building a set of tools and templates that educates system owners on data definitions and the ways that the consolidated data will be used, along with training, brown bags, and working group sessions builds agency capacity to adopt data standards. Data standards are the foundation for actionable data—you know what it means, you can trend it, compare it to results from disparate processes and external sources, etc. This requires collaboration among stakeholders and creates relationships that will increase efficiency as the integration effort continues.



- Identify data stewards for each data set and system
- Create an AS-IS data map including interconnections
- Document existing data standards
- Identify data gaps and feasibility of acquiring missing data
- Assess data consolidation impacts on current processes

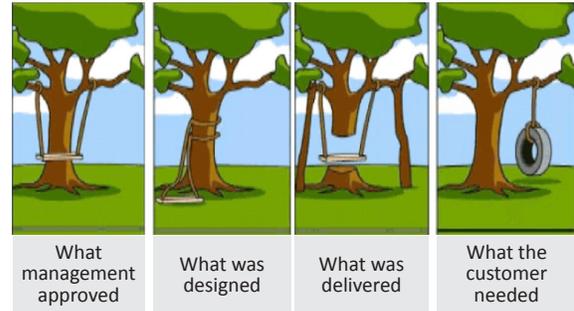
Step 2:
**INVENTORY
YOUR DATA**





Step 3: Design and Prototype Process, Policy, and System Solutions

Capture requirements. Five syllables that can make or break any IT initiative. And the fallacy of making sure all requirements are captured up front is amply demonstrated in the iconic project management “tire swing” cartoon. It is also the definition of infinity.



It is essential that there is a vision for the data consolidation effort. It is not essential to know everything about what data will be included, what the standards will be for the data, how it will be used, what reporting tools will need to access it, etc. Because all of those conditions are going to continue evolving as your mission evolves and data skills within your agency evolve. Any solution you identify must be capable of rapid adaptation to new requirements. It must be agile, modular, open source, tool agnostic, and responsive. The key to achieving this is building understanding and agreement that the solution will be continuously evolving with all agency stakeholders, data owners, and users. There is no one and done. This is the new normal—new regulatory requirements, new data elements, new data definitions, and more data becoming available from internal and external sources—and the solution has to be able to adapt.

- Finalize TO BE data map, including new standards
- Collect input from functional stakeholders
- Capture business requirements and develop technical requirements
- Analyze technical solutions, options, and feasibility



- Document required policy and process changes
- Determine central storage requirements for validation or reporting
- Draft data architecture solution
- Design initial technical capability

Step 3:

DESIGN AND PROTOTYPE





Step 4: Pilot the Solution

The pilot phase is a chance to assess not just the data validity and the performance of the data ingest or interface mechanisms, but also the processes, governance, and policies that are impacted by the new solution. The pilot project is not about software bugs, it is about the efficiency and effectiveness of the end to end process —Are there unexpected manual processes that need to be executed? Scripts that have to be written to correct problems? What happens when there is an error? What are the tolerance thresholds for errors and what happens when they are exceeded? This is an iterative process that drives modifications to processes, documentation, and the technology solution. It is also a chance to observe the efficacy of the change management and governance process that approve modifications.

Testing provides your agency with significant lessons learned as they relate not only to your future data solution, but related processes associated with the mapping, ingesting, transforming, and uploading/posting/submitting of your agency’s data. The lessons learned after piloting shed light on potential compromises that need to be made when implementing a data consolidation solution. Ultimately, these lessons learned will help your agency course correct by identifying potential inefficiencies in your processes and new perspectives on opportunities for process improvement. Piloting also allows testers, data users, and solution implementers to share a common understanding of the final data standards.

- Develop testing plans
- Identify pilot data sets and users
- Evaluate the pilot results



- Modify technical solution as needed
- Repeat

Step 4:
PILOT TESTING





Step 5: Implement Data Consolidation

With a successful pilot phase under your belt, it is tempting to assume that full implementation will be likewise successful. The significant difference is the expanded user population and number of systems and datasets that are all trying to get along for the first time. It is important to maintain engagement with testers, data users, and solution implementers throughout this step. Tap into their expertise and lessons learned from piloting to conduct “mock” production runs of your agency’s data.

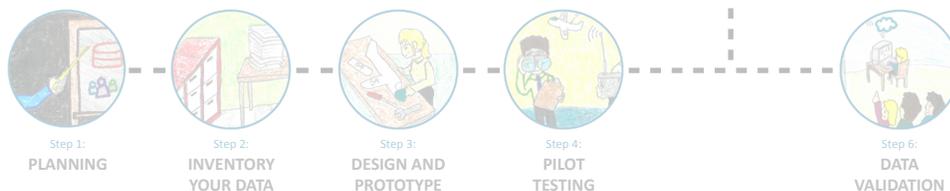
Here, you want to ensure that the processes tested during piloting are now scalable to your agency’s broader stakeholder group, end users, and consumers of your data. The focus remains on building a data-centric agency culture, which takes time. Identify data stewards and super users that can become change agents in their business unit or agency component. Create a mechanism for capturing user feedback and respond to it. Test, test, and test again. Implementation does not mean error-free completion. It means that your agency can start using the data. It does not mean that the data is perfect, which is important for agency leaders to understand. Perfecting the data happens only when it becomes available for use.

- Implement a v1.0 data dictionary
- Finalize the data model, definitions, and data formats
- Create an effective governance framework
- Evaluate source system data processing platforms
- Develop data ingestion and transformation activities for each source
- Assess performance and modify solution to optimize query capabilities



Step 5:

IMPLEMENTATION





Step 6: Confirm Data Accuracy and Quality; Conduct Reporting

Achieving data quality is not a one-and-done effort. For large data integration efforts, it is a constant struggle informed by the feedback of the people using the data for analysis and decision making.

Data consolidation solutions generally encompass both automated and manual checks for accuracy, completeness, and logical consistency. We have found that systems that rely solely on automated quality checks often miss the mark. Why? They allow valid entries to slip through that don't make sense contextually. Having valid numbers doesn't mean the numbers are correct, and often human eyes are needed to discern practicality—even on a small sample of entries.

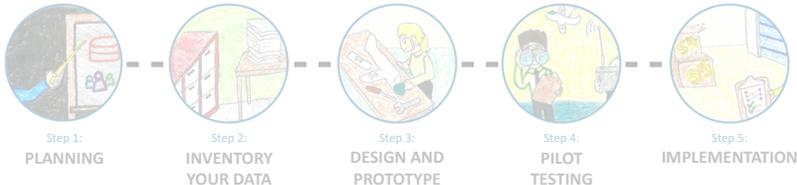
Agencies need to establish tools and processes to manage data sets, data quality monitoring, data validation, and security/quality controls. Formally documenting processes for examining validity, robustness, and defensibility drives accountability with data owners, which in turn, increases adoption and engagement with the new solution.

Vetting data and implementing feedback loops for suppliers to cleanse inputs as needed is an inherent requirement for usable data. After data quality reaches acceptable levels, agencies are then able to start realizing the value of the data by processing and reporting on the collective data set in complex and useful ways.

- Data quality monitoring, data validation, and security/quality controls
- Test data submissions for validity, robustness, and defensibility
- Implement feedback loops for cleansing/re-entering data
- Provide reporting functionality



Step 6:
**DATA
VALIDATION**





The ultimate outcome of this step is data usability. Is the data in a consumable format, i.e., open source? Is data structured in a way that all stakeholders can use? Can it be used to drive effective decision-making in your agency? Does your data help your agency fulfill its mission? The goal is to be able to answer “yes” to any of these questions.

How much is data worth? Facebook is now worth about \$200 billion. United Airlines, a company that actually owns things like airplanes and has licenses to carry passengers on routes all over the world, is worth \$34 billion.

Federal agencies are not in the money-making business. The value of data for government is conveyed by adding value to people’s lives. Agencies can better serve their constituents when they understand them, when they can assess their operations, and when they can measure their impact. Data, combined with the way that it is used, delivers lasting value that makes a difference. How will your agency get there?

ABOUT THE AUTHORS

Alissa J. Cruz is a Senior Consultant and Communications Strategist at The Center for Organizational Excellence, Inc. (COE) with over 13 years of experience helping clients execute their internal and external outreach initiatives. Ms. Cruz functions as a subject matter expert for business-focused strategic communications. In addition, she has led efforts in program management, change management, organizational assessment, competency modeling, and data analysis. She is also a certified Project Management Professional, Human Capital Strategist, and Change Management Facilitator.

Lyn McGee is a Vice President at The Center for Organizational Excellence, Inc. (COE) with over 20 years of experience developing and delivering data and internet solutions. She leads the Client Solutions for COE and driving the company to delivering management and technology solutions that measurably improve the efficiency and effectiveness of federal clients. She is also a Certified Public Accountant (CPA) and a member of the American Marketing Association and Women in Technology.

Paul Eder, PhD is a Lead Consultant at The Center for Organizational Excellence, Inc. (COE) with more than 16 years of experience helping clients achieve strategic management solutions involving data innovation. Dr. Eder is a certified Project Management Professional (PMP®) and Lean Six Sigma Black Belt. He has led consulting initiatives across the government and private sectors, supporting such organizations as the U.S. Department of Defense, the U.S. Department of Homeland Security, and the U.S. Securities and Exchange Commission. She is also a certified Project Management Professional, Human Capital Strategist, and Change Management Facilitator.

ABOUT COE

For nearly 35 years, COE has been a trusted partner in helping organizations and programs transform to more efficiently and effectively accomplish their mission. We help our clients achieve meaningful and measurable outcomes by designing and delivering consulting solutions in the areas of strategic planning, organizational effectiveness, human capital, training, development, technology and data management.



The Center for Organizational Excellence, Inc.

15204 Omega Drive, Suite 300

Rockville, MD 20850

www.center4oe.com

For more information, please contact:

Alissa Cruz, Senior Consultant

(240) 361-9262 | alissa.cruz@center4oe.com