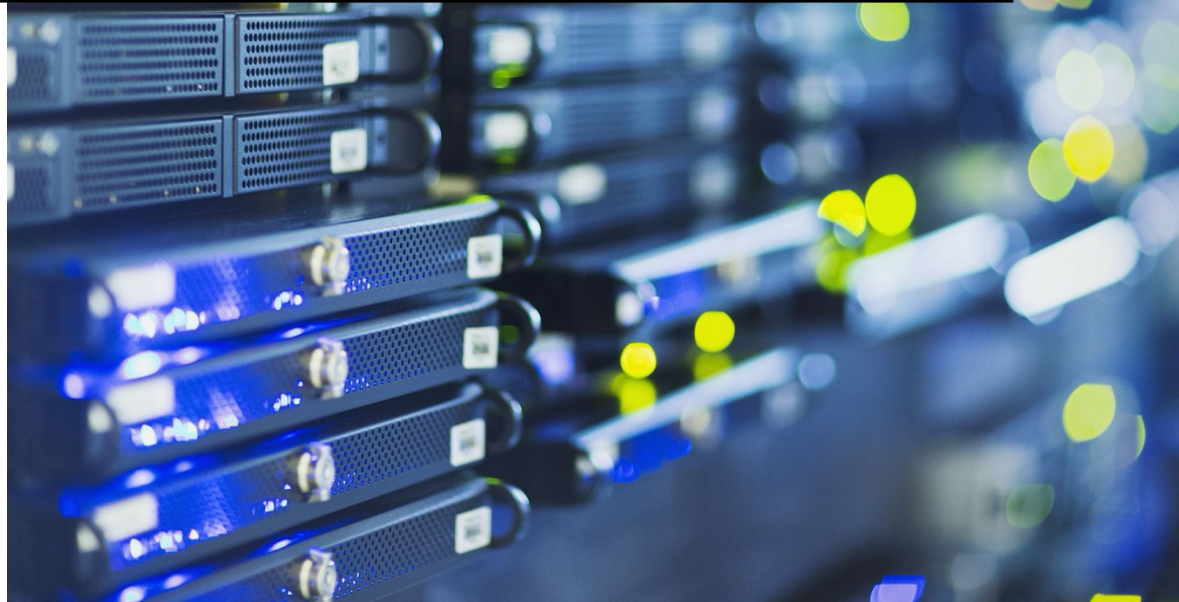


2025

# The Strongbridge Strategic Data Initiative



[www.sb-llc.com](http://www.sb-llc.com)

1/1/2025

# The Strongbridge Strategic Data Initiative

*“Where data meets innovation”*

## Strongbridge Introduction



Strongbridge is a **data-focused** provider of IT solutions to the government. We are an end-to-end *data lifecycle* support company, with 16 years of experience supporting *data-driven decision making*. Our customers in the government have a mandate to collect data on topics such as labor and transportation safety in the United States (U.S.). Our support to such programs is comprehensive and includes lifecycle steps such as: 1) creating a viable data strategy; 2) collecting and transforming the data on a national or worldwide basis; 3) storing, integrating, and processing the data; 4) deriving the mission value from the data; 5) secured sharing of data products; 6) and preservation or destruction of the data in accordance with statutory requirements. Additionally, we help government organizations in maximizing and optimizing *data value* in existing data collections, allowing them to make new connections across data sets for finding new, and perhaps unexpected, data patterns.

The Strongbridge *Strategic Data Initiative* is an effort to leverage a decade of lessons learned, staff skills, and data tools for the benefit of government programs undertaking complex data-focused efforts. The data discussion in this paper is organized around a data lifecycle appropriate for the statutory needs of a government program. Our data initiative brings the following benefits for our customer's data programs:

- **Staff Data Expertise, enabled by data tools and training** - We bring expert, hands-on knowledge of the data lifecycle, and staff supported by market-leading/innovative data tools (e.g., for cross-program integration, automation, visualizations, cloud data lakes, AI).
- **Structured Data Engineering Processes** – We apply engineering processes to: 1) unlock organizational data value; 2) find under-leveraged data strategic assets; 3) and to make new high-value data connections/integrations. Our work is requirements-driven and Agile.
- **Organizational Perspective for Data Value** - We implement the goals of the Government Chief Information Officer (CIO)/Chief Data Officer (CDO) in mission systems, applying data services and tools for the agency mission in the public interest. We recognize the organizational value of the CDO Council vision, “to make improvements in the management, use, protection, dissemination, and generation of data in government decision-making and operations.”<sup>1</sup> In doing so, we address today's data challenges (e.g., multi-sourced data interoperability, security, sharing analytics, lowering maintenance costs, meaningful AI use cases and AI bias)
- **Creating and Enabling Data Infrastructure** - We advance the state of the data infrastructure and services in a large, complex organization for users that range in expertise from data scientist to beginning data reporter.

We recognize the complexities and challenges faced by government agencies in managing vast amounts of data. Our team of experts employs state-of-the-art technologies and methodologies to **transform raw data into actionable insights**, ensuring that our customers are

---

<sup>1</sup> <https://www.cdo.gov/index.html>

# The Strongbridge Strategic Data Initiative

*"Where data meets innovation"*

equipped with the knowledge and tools needed to serve the public effectively. Whether it's through advanced data analytics, AI-driven insights, or customized data management strategies, our mission is to empower our customers with data-driven excellence, paving the way for a smarter, more responsive government.

## Strongbridge Data Lifecycle for Government Programs

By applying the data lifecycle shown in the inset figure, Strongbridge specializes in delivering cutting-edge solutions tailored to the unique needs of the Federal Government. Our expertise lies in harnessing the power of **data to drive decision-making**, enhance operational efficiency, and improve service delivery.

The six topics in Strongbridge's Data Initiative, detailed below, establish a methodical approach to leveraging data as a strategic asset. With unprecedented amounts of data stratified across legacy and cloud systems, Strongbridge is a partner that will consider all facets of the data lifecycle, from systems development, data utilization, data visualization, to data consumer experience to ensure iterative value and evidence-based decisions. Understanding the strategic business questions that need answers, we can tailor data solutions for mission value while maximizing the ROI of customer data investments.



The Strongbridge Data Lifecycle outlines the process for thinking about data holistically, from initial strategy to preservation or destruction. Too often systems development focuses on the Systems/Software Development Lifecycle without focus on the underlying data gathered or produced. By utilizing a data lifecycle approach as well, it grounds system development with an understanding that data is not an end of itself but a means to further greater organizational value. Leveraging data from business analysis to more advanced analytical or statistical methodologies recognizes return on system investment through mission value. Leveraging modern system technologies breaks down the silos of legacy business intelligence systems to allow an ecosystem of data sharing that was not available previously. Through modern system technologies and infrastructure, data worth is greater, especially as never before combined data sets are used together in a synergistic way.

The following sections provide detail on each of the six data lifecycle topics.

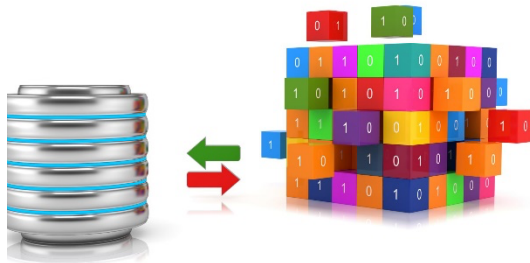


**Data Strategy** – Creating a strategy emerges as a foundational step in the data lifecycle, guiding the overarching mission questions and conceptual framework that steer subsequent data initiatives. In this phase, a comprehensive data strategy is formulated, outlining the organizational vision for leveraging data to achieve mission objectives. Strategic decisions include

# The Strongbridge Strategic Data Initiative

*"Where data meets innovation"*

identifying and assessing data sources, encompassing government databases, commercial datasets, and custom-built repositories. The utilization of data standards is strategically incorporated to ensure consistency, interoperability, and seamless integration across diverse datasets.



A well-defined quality plan is integral to the data strategy, outlining methodologies for data validation, cleansing, and verification to uphold data integrity. Legal considerations, including Intellectual Property (IP) issues, are carefully addressed to navigate potential challenges related to data ownership and usage rights. Contracting arrangements are established to formalize

relationships with data providers and vendors, ensuring compliance with legal and ethical standards. Ultimately, "Data Strategy" sets the stage for a systematic and purposeful approach to managing, utilizing, and deriving value from data throughout its lifecycle, aligning with the broader mission and objectives of the organization.

The strategy will address topics such as:

- What are the data sources flowing into or created by the system?
- What are the business/mission questions this system and/or data is supposed to answer?
- Who is the audience for this data and in what format do they need outputs?
- What tools are needed from a data perspective to interact with the data after it is created?
- What are the security protocols around the data and what is the plan to ensure they are followed from ingestion through downstream utilization?
- How can you minimize people performing data entry and move them to analyzing outputs?
- What staff skills are needed, data engineers, data scientists, statisticians, etc. to move data from its point of origin to business analysts and consumers for right on time analysis?

Thinking through and documenting the answers to these questions is a solid starting point in ensuring data is not an afterthought in systems projects.

**Experience Example:** Strongbridge Chief Innovation Officer (CIO), Geoff Raines, organized the data strategy for the CIO of the Department of Defense (DOD), one of the largest and most complex data-focused organizations in the world.

## 2) Data Collection, Generation, Acquisition

**Data Collection, Generation, Acquisition** - Thinking through the data strategy first allows for a more efficient and complete understanding of the underlying needs of the data and what needs to be collected. Moving past the forest view of the data strategy, the need begins to think about the trees of implementation. It is also at this step that quality should be assessed to ensure proper upstream functionality and downstream needs. This is where a quality plan can be leveraged not only from legal standpoint, but from the data provenance

## The Strongbridge Strategic Data Initiative

*"Where data meets innovation"*

perspective as well. The type of data should be understood (structured vs unstructured) as well to beginning to think about the proper design of the storage solutions. Mapping of data flows at this point can help with understanding frequency of updates and data lineage as it moves through an ecosystem from creation to utilization in analysis or business decision making.



Here the input mechanism for data should be considered.

The source of incoming data, whether leveraging data from an existing IT system or database internal to an organization, external data being brought in, and/or human input data can influence the level of control over the incoming data. For internal systems, data anomalies may be managed at the source vs in the new environment, or you can tailor your data model to internal norms. For external sources, a vetting and transformation process may be necessary to

**Challenge - Setting Data Exchange Standards Across Government and Commercial Providers** – Creating a community standard for data exchange in a government program requires collaboration, participation and knowledge sharing, combined with legal and regulatory understanding of the mission. We have led the creation of data exchange standards, with input from all parties.

ensure incoming information aligns to the existing data model. For human input information, assessment of the input mode, from forms to web pages, is necessary to ensure they are intuitive to the average user. Managing the mode and limiting responses to the universe of acceptable answers in a best practice. The completeness of a form should be considered based on the data usage, eliminating future data gaps with comprehensive data from the start.

**Experience Example:** Strongbridge supports large data programs for the National Highway Traffic Safety Administration's (NHTSA) Crash Data Acquisition Network (CDAN). The NHTSA CDAN is a pivotal nationwide data collection system that plays an essential role in enhancing vehicle safety and improving road traffic conditions across the United States. This program gathers comprehensive crash data from a variety of sources and the data collected is meticulously analyzed to identify patterns, assess risk factors, and develop strategies to prevent accidents and mitigate their consequences. Strongbridge recognizes the importance of the CDAN as a tool for informing policy decisions, guiding research on vehicle safety, and fostering innovations in automotive technology. Our approach to managing the CDAN will ensure the integrity, accuracy, and security of the data, while also making it accessible to stakeholders who rely on this information to make informed decisions.

### 3) Storage, Integration, and Processing

**Storage, Integration and Processing** – These steps are foundational for optimizing data management and facilitating downstream analyses. The process begins with transformations to enhance the quality and relevance of raw data, followed by meticulous data cleaning to ensure accuracy. Data mapping techniques are then employed to establish meaningful relationships between various datasets, fostering a cohesive and interconnected data



# The Strongbridge Strategic Data Initiative

*"Where data meets innovation"*

landscape. Cross-source integrations play a crucial role in merging data from disparate origins, promoting a comprehensive understanding.

As part of this step, data aggregations are implemented, including the creation of data lakes and warehouses, providing centralized repositories for efficient storage and retrieval. Live data services and APIs are leveraged to enable real-time access to dynamic datasets, ensuring that organizations work with the most current information. Downstream data integrations

complete the process by seamlessly connecting processed data to various departments or systems within an organization, facilitating cohesive workflows and analysis.

---

**Example Tools:** *Cloud Data Lakes (e.g., Data Factory Pipelines, Azure Data Lake Gen 2), APIs and microservices (e.g., AWS Lambda, MS Functions)*

---

In essence, "Storage Integration and Processing" optimizes the data lifecycle by incorporating transformations, cleaning, mapping, integrations, and aggregations. This strategic approach not only ensures the integrity of data but also sets the stage for subsequent phases, providing a solid foundation for organizations to extract valuable insights from their data resources.

**Experience Example:** Strongbridge supports the FRA's Business Intelligence Services (BIS). BIS is an integrated platform of business intelligence, data and predictive analytics tools in a cloud-based environment that will ingest, integrate and standardize structured and unstructured enterprise data using business rules established by the agency. This project unites capabilities for enterprise data warehousing, data integration, visualizations, and analysis, and supports highly scalable data collections of any size in a data lake in the Azure cloud ecosystem. For analysts, researchers and data scientists the BIS environment we provide supports rail and safety data research of any scale, with both structured and unstructured data sets. For FRA projects, BIS supports the integration of embedded analytics for end users (e.g., Power BI), live data services and reporting into rail project applications.

## 4) Analysis Driven Mission Value and Insights

**Analysis Driven Mission Value and Insights** – This phase represents a critical stage where data is transformed into actionable intelligence. Visualizations play a pivotal role during this step, translating complex datasets into comprehensible representations, enabling stakeholders to glean insights effectively. The process involves dynamic analysis, where the data is

continually scrutinized and interpreted to extract meaningful patterns. Shared pre-built analytics further enhance efficiency, allowing for the rapid deployment of proven analytical models and tools. Leveraging data science techniques, such as statistical analysis, contributes to a deeper understanding of trends and correlations within the datasets.

---

**Example Tools:** *Tableau, SAS, Microsoft Power BI, AI deep learning and predictive models, Cloud NLP, sentiment analysis*

---

Reporting becomes a central element, encompassing leadership briefings, internal reviews, and public disclosures. The insights generated through this analysis-driven phase provide crucial decision support for mission-critical activities. Additionally, a confidence assessment is

# The Strongbridge Strategic Data Initiative

*"Where data meets innovation"*

conducted, ensuring the reliability and accuracy of the derived insights. This multifaceted approach not only empowers internal stakeholders with valuable intelligence but also fosters transparency and accountability when shared with the public. Ultimately, "Analysis Driven Mission Value and Insights" elevates data from raw information to a strategic asset, guiding mission-critical decisions and contributing to organizational effectiveness and success.

AI services can play a crucial role in data analysis by enabling organizations to extract valuable insights, make data-driven

decisions, automate repetitive tasks, and

uncover patterns that may not be apparent through traditional methods. For example, we recently applied cloud-based AI Natural Language Processing (NLP), sentiment analysis, and machine learning (ML) to the analysis of contractor past performance data in the procurement cycle. AI algorithms, including machine learning and deep learning models, excel at identifying patterns and trends in large datasets. These algorithms can uncover complex relationships and correlations that may not be immediately apparent to human analysts.



**Challenge – Increasing Bias in Generative AI** – Generative and predictive AI output is ultimately a result of training data. As government adopts AI services, the potential for output bias should be monitored. Strongbridge provides mission focused bias testing.

**Experience Example:** Strongbridge's *Director of Data Strategy*, Sven Carlson Jr, led teams of data analysts across the General Services Administration (GSA), in the Federal Acquisition Service (FAS), Public Buildings Service (PBS), and for the GSA Administrator, delivering mission value through innovative data collection analyses, data process improvements, and through organizational data democratization.

## 5) Data Sharing

**Data Sharing** - In the continuum of the data lifecycle, the stage of "Data Sharing" emerges as a pivotal step where the emphasis is on making data accessible and valuable to a wider audience. This involves intentional data publishing and dissemination strategies to ensure that datasets are shared efficiently and securely. Adherence to Community Data Standards becomes paramount during this phase, fostering consistency and interoperability across shared datasets. Creating Data.gov excerpts plays a significant role, providing a curated and centralized platform for the public to access pertinent datasets, contributing to transparency and collaborative knowledge building.

## The Strongbridge Strategic Data Initiative

*"Where data meets innovation"*

The significance of "Data Sharing" extends beyond mere accessibility; it actively supports research studies by providing researchers with the necessary resources for analysis and exploration. Embracing public downstream data users is integral to this step, as it encourages diverse applications and interpretations of the data, contributing to a vibrant ecosystem of insights and innovations. To govern this exchange responsibly, organizations establish clear data sharing agreements, defining the terms of usage and ensuring ethical and compliant practices. In essence, "Data Sharing" transforms data from a static asset to a dynamic resource, fostering collaboration, research, and informed decision-making within the broader community.



To manage and govern this exchange, establishing clear data sharing agreements is essential. These agreements outline the terms and conditions for data usage, ensuring responsible and ethical handling while safeguarding against misuse. In essence, the "Analysis Driven Mission Value and Insights" step serves as a bridge between data accumulation and its practical utilization, fostering a collaborative environment where data becomes a powerful tool for innovation, research, and informed decision-making.

**Experience Example:** Our work for the Department of Labor's Wage and Hour Division (DOL WHD) involved the development and maintenance of a comprehensive data management system that was instrumental in enforcing federal labor laws regarding minimum wage, overtime pay, recordkeeping, and youth employment. We were responsible for the design and execution of sophisticated data analysis protocols, which allowed DOL WHD to identify non-compliance issues, understand labor market trends, and make informed policy decisions. Furthermore, Strongbridge's commitment to data security and privacy was paramount in our work with DOL WHD. We ensured that all sensitive information was handled in compliance with federal standards and regulations, protecting the privacy of workers and employers alike.



**Data Preservation or Destruction** - In the concluding phase of the data lifecycle, attention to data preservation and destruction becomes paramount. For long-term preservation, organizations must implement robust archive solutions, considering appropriate formats that ensure data integrity and accessibility over extended periods. Tiered storage cost models provide a

strategic approach, allowing for the classification of data based on importance and usage frequency, optimizing storage costs. Addressing the transition from proprietary to sustainable formats is crucial, as it mitigates the risk of data obsolescence. Conformance with National Archives and Records Administration (NARA) record schedules ensures that data preservation aligns with regulatory requirements and compliance standards.



# The Strongbridge Strategic Data Initiative

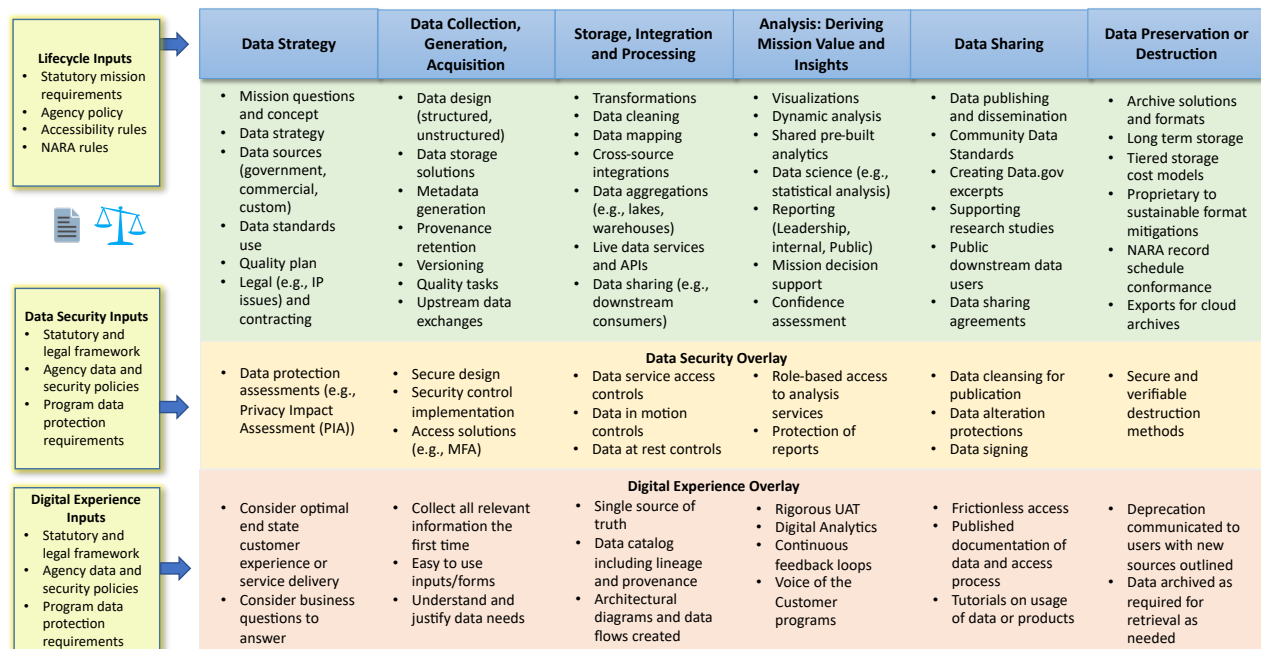
## "Where data meets innovation"

Simultaneously, data destruction is a critical aspect of managing the end of the data lifecycle. Organizations need to employ secure and compliant methods for data destruction, ensuring that sensitive information is permanently and irreversibly removed. Implementing exports for cloud archives facilitates a seamless transition for preserving valuable data in cloud environments, allowing for efficient retrieval and safeguarding against data loss. Striking a balance between data preservation and destruction is essential for maintaining information governance, regulatory compliance, and efficient data management throughout the complete data lifecycle.



**Experience Example:** Strongbridge is supporting DOT's FRA with data digitization and records management tasks in full compliance with NARA's records schedule. We support the FRA Federal Records officer (FRO) to ensure that all Controlled Unclassified Information (CUI) records schedules are properly observed, the records eLibrary functions properly, and transfers to NARA are conducted effectively and efficiently.

### Exploring the Data Lifecycle in a Government Program



**Figure 1: Strongbridge Data Lifecycle – End-to-end program support**

Strongbridge has captured the data lifecycle for government programs in Figure 1 above, but also overlaid the critical considerations for Data Security and Digital Experience. The data lifecycle comprises a strategic progression of stages, each playing a vital role in the holistic management and utilization of information. The lifecycle begins with setting a Data Strategy, where organizations define their mission questions, conceptual framework, and overall vision for leveraging data. This sets the foundation for subsequent phases. "Data Collection,

## The Strongbridge Strategic Data Initiative

*"Where data meets innovation"*

Generation, and Acquisition" follow, involving the systematic gathering of data from various sources, including government databases, commercial datasets, and custom-built repositories.

The next phase, "Data Storage, Integration, and Processing," focuses on optimizing the storage of collected data, employing transformations, data cleaning, and integration techniques to ensure efficiency and accessibility. "Data Analysis, Deriving Mission Value and Insights" marks a critical step, where raw data is transformed into actionable intelligence through visualizations, dynamic analysis, and shared pre-built analytics.

Moving forward, "Data Sharing" facilitates the dissemination of valuable insights to a broader audience, supporting research studies and fostering collaboration. Finally, the data lifecycle concludes with "Data Preservation or Destruction," a crucial phase where data is either archived for long-term retention or securely destroyed in alignment with regulatory requirements. This comprehensive approach to the data lifecycle ensures that organizations not only collect and analyze data but also derive meaningful insights, share knowledge, and responsibly manage information across the entire lifecycle.

**Security Overlay** - Throughout the entire data lifecycle, data security is a paramount consideration, and it is systematically addressed in each phase to safeguard sensitive information and maintain the integrity of the data. As shown in the security overlay in the bottom half of the figure, security protocols are established, defining access controls, encryption standards, and outlining policies that guide the handling of data. This sets the groundwork for a secure foundation for subsequent phases.

Security measures focus on ensuring the confidentiality and integrity of data. Robust encryption during data transfer, secure APIs, and compliance with data protection regulations contribute to a secure collection process. Data storage involves implementing stringent access controls, encryption at rest, and secure data storage solutions. Metadata is managed carefully to avoid unintentional exposure of sensitive information, and provenance tracking is employed to monitor data



changes and ensure data integrity. Access to analytical tools is controlled, and encryption is maintained during data processing. Anomalies and potential security threats are actively monitored to prevent unauthorized access. In the Data Sharing phase, security considerations extend to data sharing agreements, outlining terms to protect data integrity and limit usage to authorized entities. Secure data exchange mechanisms, like encrypted APIs, are employed to ensure the safe dissemination of insights. Finally, in "Data Preservation or Destruction," security protocols dictate the secure archiving of data for long-term preservation or the secure deletion of data according to regulatory requirements. Data destruction processes are carefully executed to prevent any potential breaches.

Throughout the entire data lifecycle, a comprehensive and evolving approach to data security is essential. This includes regular security assessments, updates to security protocols, and adherence to industry and regulatory standards to protect data at every stage of its journey.

## The Strongbridge Strategic Data Initiative

*“Where data meets innovation”*

**Digital Experience Overlay** – To fully leverage data’s value, consideration for how the data will be accessed, consumed, and/or utilized has to be baked into the data lifecycle. The voice of consumers will help hone the data lifecycle and inform design decisions to optimize the user experience.

The user experience is paramount to the adoption of data centric tools or processes. Human centered design techniques foster research into users’ feelings leading to downstream products that provide actionable intelligence easily and on time. This can help with user interface and design decisions to meet users where they are to provide frictionless access to information. In the Data Strategy phase, outlining the users’ wants and needs is crucial to ensuring that the strategy is in alignment. In the Data Collection phase, ensuring user inputs or forms are clear and concise is paramount. Collecting all the relevant data up front avoids re-work and/or data gaps. In the Analysis and Sharing phases, a best practice is to set up a Voice of the Customer program with continuous feedback loops to iteratively improve products or functionalities. Incorporating this feedback should be a routine function. Another quality indicator of successful data products is to monitor usage. This can indicate which items are most utilized and can help with rationalization decisions. Combining these methods with rigorous User Acceptance Testing will lead to better data outcomes. In the Preservation or Destruction phase, communication with users is key to ensuring there are no data gaps upon deprecation of one product or source of data.

### About Strongbridge

Strongbridge is a forward-thinking data-focused company that specializes in delivering comprehensive system solutions tailored to meet the unique needs of various agencies within the US government. With a deep understanding of the intricacies and challenges faced by government entities in their data programs, Strongbridge leverages its industry expertise to develop, implement, and manage sophisticated solutions that enhance operational efficiency, security, and service delivery. Our team of dedicated professionals is committed to providing exceptional service and innovative solutions that address the critical needs of our clients.

Website: [www.sb-llc.com](http://www.sb-llc.com)

#### Contact:

Grafton DeButts, *Director of Strategic Growth*

Phone: 703.431.8443

Email: [gdebutts@sb-llc.com](mailto:gdebutts@sb-llc.com)