ATARC

Developing an Effective Data Management Strategy

Summary of Roundtable, hosted by ATARC in October 2022

WHITE PAPER

Effective data management has become increasingly important as government organizations continue along the path towards digital transformation. As operations steadily move into the cloud, managing vast amounts of data is critical to not only agency mission success, but also national security.

In a recent roundtable discussion hosted by the Advanced Technology Academic Research Center (ATARC), topic experts from various Federal agencies exchanged ideas and approaches to data management, and the hurdles agencies must overcome to manage and share data effectively.

> "Often the sense of security from the status quo is a false sense of security."
>
> *Roundtable Participant*

Approaches to Data Management

Data management is one aspect of digitalization that agencies are paying particularly close attention to. As agencies reorganize to accommodate Chief Data Officers and the expansion of data collection, data management is at the forefront of most digital strategies. Several agencies oversee multiple missions that collect and use data for vastly different purposes. Part of a successful digital strategy is connecting these disparate data sources and converting enormous amounts of data into meaningful information in a secure manner.

When thinking about data management, agencies should consider data ownership, collection methods, data authority and usability. Understanding data bias and collection methods is imperative to improving data quality and usability. Creating governance structures allows agencies to effectively use the data as an enterprise resource, rather than for a narrow, programmatic scope.

Data Management and Innovation

Strong data management and governance practices enable agency innovation and modernization. Roundtable participants note that accelerated AI and modern software adoption is dependent on greater access to quality data. Some agencies are implementing Zero Trust practices in their data management strategies. By doing so, agencies can use machine learning to inventory and classify all data in order to build out Zero Trust architecture. Agencies are then better able to manage and secure sensitive information and give stakeholders proper access to the right data, at the right time, and in the right way. Machine learning can also be used to better understand and identify fraud. By understanding the people accessing networks, agencies can apply Zero Trust rules to emails to determine risk levels if certain data were exposed.

> "Data can be a four letter word if you don't know what you're doing with it."
>
> *Roundtable Participant*

Data is core to an effective Zero Trust architecture and strategy. Understanding data patterns through an organization, and who should be accessing the information is critical to maintaining a strong security posture. Automation tools can help agencies not only manage data more efficiently, but also monitor for any discrepancies, data loss, or unauthorized access of data. The benefits of effective data management and open data practices are many, but a culture that's hesitant to share and exchange data threatens further progress.

Agencies are also making a concerted effort to separate analytics management from data management, although there is still a tendency for the two to intermingle. To do this properly, agencies have started to set up executive data boards and to place Chief Data Officers outside of IT departments. Although connected in certain ways, data governance and IT governance are separate business practices and should function independently from one another. Adopting clean and agile data management practices will make it easier for development teams and stakeholders to understand the value of data and encourage better collaboration.

Data and Culture

Roundtable participants point to a dominant culture hesitant and unwilling to share data as a primary challenge to effective data management. Many agencies are working to promote a culture that strategically manages data as an asset to optimize business value and application performance. Helping stakeholders to realize the shortcomings of not taking advantage of the insights provided by open data is key to gaining traction.

"Culture remains a huge barrier to anything we do at the department and programmatic level. We have a sound strategy. We have a data governance process. We have representation. We have a core lead from the Chief Data Officer. But we still struggle from a cultural standpoint."

Roundtable Participant

Some of the hesitancy surrounding data sharing relates to security risks. Generally, data is collected for a particular use case, which makes data usability challenging for other use cases unless stakeholders are trained to work with data properly. Users need to understand data bias, data collection methods, and data modeling to properly use data. Upskilling the workforce in data analytics and data management best practices is necessary to encourage safe data sharing and proper usage.

"Separation and isolation are insufficient in the modern era. We have to learn from an entire global industry that sees their information as their most valued assets."

Roundtable Participant

Ultimately, the paradigm of risk needs to change in order for data to effectively solve enterprise scale problems. Data scientists must accept using imperfect data to deploy tools before they're perfected in order to make progress with modernization. By integrating new capabilities that are coming from the commercial sector, agencies can learn by taking an experimental, agile approach. To find traction, agencies should be willing to take on a certain level of risk in order to share and use imperfect data that is perfectly capable of driving decisions. Technology, talent and culture must be in sync in order for agencies to take full advantage of the new capabilities that technology enables.

Public Data and Zero Trust

Agencies should evaluate the risks inherent to sharing data, and whether the risk of leakage would offset its value. The American public has paid for this data and allowed its creation, so agencies should consider whether they have a responsibility to operationalize the value of that data.

Knowing what data should be secured and how to secure it is challenging for many agencies. Understanding how data can be used and interpreted individually and in aggregate is a big undertaking, but roundtable participants point to Zero Trust to make this process more efficient. Agencies are still determining how to implement Zero Trust on systems that are open to the public, as it is much easier to understand internal user behavior than the behavior of public end users. Security will remain in tension with data utilization, so agencies must consider the value and utility of data once shared with the public.

Contact us today to learn more and get involved in ATARC's AI & Data Working Groups and other ATARC events!