ACHIEVING ENTERPRISE SELF-SERVICE DATA:

THE SIX PILLARS OF COMPREHENSIVE CAPABILITIES

Radiant Advisors

Table of Contents

Introduction

A successful enterprise self-service data strategy is dependent on a comprehensive set of core capabilities deployed together to empower business users with data.

Too often, companies embark on enterprise self-service data programs to empower the business user based on a particular need or specific starting point, such as data catalog, only to realize later that the lack of other, related capabilities hinders widespread ability to work effectively with data every day. Successful enterprise self-service data programs are based on a foundation of six pillars, or core capabilities, that comprise a balanced enterprise self-service data platform to support enterprise needs over time. With these six pillars as the foundation, companies can achieve enterprise-class self-service data, enabling business users to work more efficiently and effectively than with standalone tools to solve business challenges.
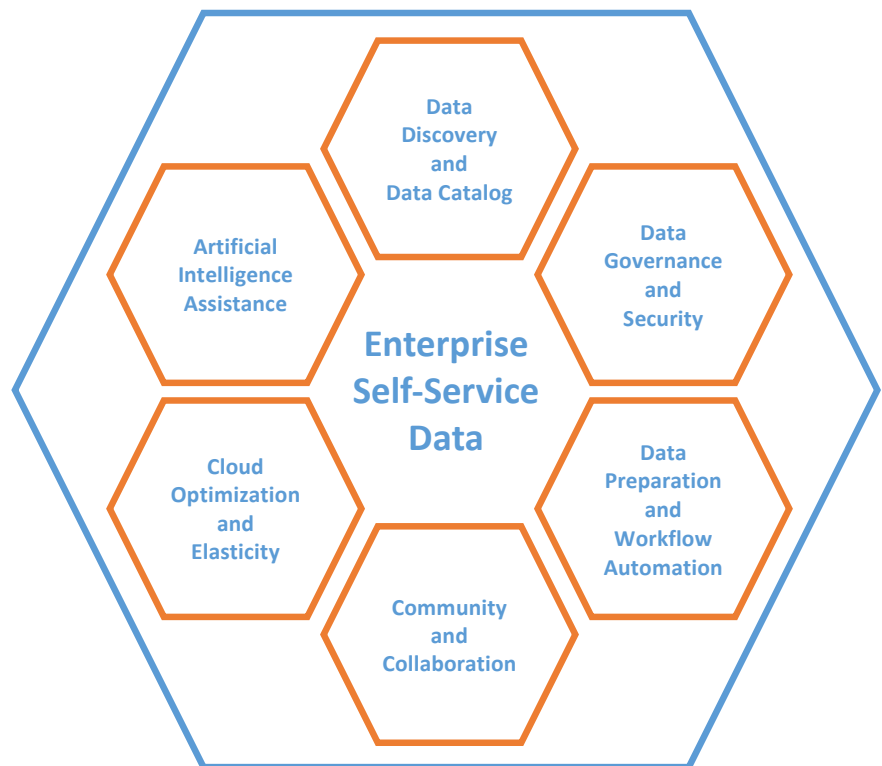
The purpose of a self-service data platform is to supplant the dependency on IT and support the business user's end-to-end process of independently finding, understanding, curating and publishing data for business use. Coupled with a company's cultural shift to establish data proficiency, this is how organizations fully democratize data and become data-centric leaders.

Radiant Advisors, together with support from research sponsor Unifi Software, has defined six pillars of enterprise self-service data based on years of researching how companies develop their enterprise self-service data programs. Companies that understand this related set of core capabilities can avoid the pitfalls that result from use of fragmented solutions.

An enterprise self-service data platform empowers business users to work intuitively with data in an interactive way that quickly applies their unique business domain knowledge to solve complex challenges independently of IT developers and administrators. This concept goes far beyond self-service BI's custom-tailored reports and analyst tools that extract data into another BI server environment for business users to access. Instead, enterprise self-service data follows the end-to-end process that begins with business users searching, identifying and properly accessing data to understand and prepare it to curate valuable enterprise data insights that others can leverage.

These essential six pillars serve the needs of three different groups within the organization: business users, IT development and the enterprise data management culture. Working together transparently, users competently participate in the self-service data process, thus achieving true data democratization.

Data Discovery and Data Catalog

Artificial Intelligence Assistance

Data Governance and Security

Enterprise Self-Service Data

Cloud Optimization and Elasticity

Data Preparation and Workflow Automation

Community and Collaboration

Radiant Advisors

#1. Data Discovery and Data Catalog

For each business question to be answered and hypothesis to be verified, a business user must apply his or her domain knowledge to uncover and explore any data that may be relevant and valuable for the task at hand. The ability for a business user to quickly find and understand available data sources is essential to the self-service data process.

A **data catalog** helps business users search across all data sources for the data that's relevant to their specific needs – to see data characteristics and capture related knowledge. A data catalog and index support the search and discovery process for data from all available data sources (including databases, flat files, unstructured data or SaaS APIs). The ability to use keywords to search and refine search terms iteratively until relevant results are generated saves time and allows users to explore data using familiar methods and syntax. Natural language query further streamlines this interaction and returns better results with semantic context.

The data catalog also helps business users evaluate potential data sources regarding context and applicability. The first aspect of this is the information that can be captured and represented in technical and business glossary definitions. This data is relatively static, such as data type, format and relationships between data sets. A second aspect is the highly valuable information about institutional or tribal knowledge existing with people in the organization relative to the data. Combined, these sets of information quickly allow a business user to determine whether a data source will be useful. To capture this information in the data catalog, business users must be able to easily enrich the data with comments, ratings and reviews that can be preserved for reference and future use.

#2. Data Discovery and Security

Organizations must ensure that access to data and data sources is fully governed and secured so that business users only access what they are allowed to and understand how to use the data appropriately. Establishing and sustaining governance and security procedures that are transparent to users and always present – but do not unnecessarily restrain people while working with data – provides safeguards and ensures that enterprise data compliance and standards exist for all business users.

Data authorization ensures that business users have access only to data that has been granted to them. A structure of roles and groups for users along with permissions for create, read, update and delete (CRUD) data operations must be established in a centralized enterprise function to manage the relationship between data and users. Within the data governance program, data owners are responsible for establishing criteria for who has authorized access to their data along these parameters.

Data compliance needs to be established broadly and deeply from the beginning to prevent inadvertent data breaches, which can be extremely costly. Companies are responsible for complying with their industry, federal and global regulations, such as Health Insurance Portability and Accountability Act (HIPAA), the Payment Card Industry Data Security Standard (PCI DSS), all personally identifiable information (PII) rules, and the upcoming European General Data Protection Regulation (GDPR), and must embed safeguards in data processing to eliminate risks of non-compliance events. The most common starting point is to implement data encryption, masking, tokenizing or other data assignments at the data ingestion step before data is accessible by users.

Data governance is instrumental to fueling enterprise self-service data adoption. When done correctly, governance will incorporate the collaboration and community features of defining and validating self-service data rather than stifle curation with restrictions regarding how to handle various data sets. Governance is especially important to enterprise self-service data, where the goal is to create new derived data sets that must also have definition, context and governance.

Data lineage intuitively shows any business user what data sources were used and the series of data preparation steps in the workflow. This information takes a business user from simply accepting the data output definition to understanding how the data output was derived. The user can then make their own determination of acceptance and validation for their needs. Data lineage also improves the accuracy and speed of impact analysis whenever a change in source data or a transformation is made. This is not always possible when the data workflow has steps outside the data preparation environment. This is especially important when regulated and secure source data is being used to create derived data sets that must also comply with privacy and regulations.

"It's being able to bring in multiple data sets beyond just clinical data that allows me to digest and gain insights that I was never able to do before."

- Dr. Jason Cunningham, West County Health Centers

#3. Data Preparation with Workflow Automation

At the heart of self-service data is the ability for business users to work with data in ways that go beyond what desktop spreadsheets or small databases could do – without requiring a developer to program any data integrations or complex business rules. In the hands of a business user, business knowledge can be seamlessly applied in data discovery, exploration and integration. Within this construct, the business user creates a data workflow that is visually and logically easy to read, design, change and see how data moves through their transformations. The ability for business users to prepare data in an agile process produces the exacting data sets needed by the business for decision-making and complex analytics.

The bulk of **data preparation** workflows centers on three main steps. The first step is to profile data sets to understand the quality of the data and determine what needs parsing, cleansing, filtering or enriching. Next is to analyze multiple data sets for integration while maintaining the correct business context for inner and outer joins across several fields and choosing which fields should be in the final output. Finally, the output data is prepared for publishing with various aggregations and filters. These three steps are highly iterative and powerful in the hands of business users who are able to seamlessly apply business domain knowledge to their data sets during the workflow development.

To support this process, **workflow automation** operationalizes the data preparations that business users have created. Most data preparations will not be a one-time use for business users, and saving the preparation work for reuse as new data arrives from operational systems or external sources eliminates rework and saves time and resources. Data preparations fall into one of three categories from which users will select to deploy their workflows for execution:

Radiant Advisors

manual, scheduled or event-driven. In some cases, it's not possible to remove the dependency on the business user, and manual operations are needed for processing. In most cases, refreshing published data sets can be scheduled as appropriate for business needs. For event-driven situations, a form of change data capture (CDC) can be utilized, such as the arrival of a file in a specified location or database activity that initiates the data preparation job.

Overall, workflow automation requires a visual and highly intuitive interface for business users to understand how to operationalize their data preparations. A good workflow automation engine for enterprise self-service data will coordinate all data preparation jobs at scale, manage dependencies and workloads, and provide audit logs for reporting.

#4. Community and Collaboration

Rarely do business users work in a vacuum – nor should they, when independent validations, peer reviews or sharing of data sets with others can enhance the quality of their work. Community-building and collaboration are a core component of a data-centric culture and are instrumental in driving enterprise self-service data programs. Collectively, the overall network effect of leveraging others' work and contributing to the community advances business objectives.

Collaboration capabilities, such as crowdsourced ratings of available data, help users identify which data sources are likely to be best-suited for their project. Users can read comments and pose questions on data preparation workflows to quickly understand data quality, filters and context for why the data was prepared a certain way – as well as determine whether those decisions are applicable to the project at hand.

Many Fortune 500 companies recognize the value of collaboration to fuel a data-centric organizational culture. Without an environment for collaboration, companies struggle to easily share valuable knowledge to those that may benefit from it. Inefficient data sharing slows down data and analytics work, and is counteractive to self-service data programs. In such companies, it is common to hear comments from business analysts such as, "Other analysts and managers frequently stop by my desk and ask me to share snippets of SQL code for their own analysis," as a Radiant Advisors client commented. Platforms that promote crowdsourcing of data and collaboration have a cumulative effect of engaged, active participants working efficiently together across the organization.

One highly valuable aspect of a collaborative and **engaged community** is crowdsourced data cleansing. Individuals closest to the business process of creating and utilizing data are able to make their knowledge and work available for others, assessing and improving the data along the way. This real-time transparency for data quality improvements makes better data available and clarifies its context to avoid the massive amount of time wasted reinventing the wheel for each business user.

The delivery and publishing stage of the self-service data process specifically relies on community and real-time collaborative capabilities. Once a business user is satisfied with the completed data preparation workflow, he or she can share their work for peer reviews and validation to build confidence before publishing the data workflow and data set. The real-time collaborative nature of providing feedback and answering questions will move finished data sets into use faster than traditional processes of coordinating calls, meetings and documentation sessions for review and approval.

Havas Media, a global marketing communication agency, relies on collaboration as one of their value propositions in working with clients. Data discovery and data prep are a part of their creative process, and clients actively participate by uploading their own data into Havas's Unifi Software self-service data platform. The clients are then directly involved with creative teams at Havas to understand campaign performance, explore new ideas and seek insights as a group. "People have their own data and curiosity, and collaboration yields interesting insights from sharing unique perspectives and asking questions of each other," says Sylvain Le Borgne, EVP of Platforms at Havas Media.

Community and collaboration also play an important role in the data governance process by enabling subject matter experts and data stewards to easily collaborate on data definitions and uses. Crowdsourced data governance is effective in allowing business users and data stewards to determine business definitions and context.

#5. Cloud Optimization and Elasticity

As adoption increases with the growth of active users and accumulated data preparation workflows, the infrastructure for executing intensive data workflows will require an increasing amount of computing resources. On-premises deployments face the initial challenge of needing to estimate the capacity of computing resources required for initial usage, peak usage and future growth.

> *"People have their own data and curiosity, and collaboration yields interesting insights from sharing unique perspectives and asking questions of each other."*
> *- Sylvain Le Borgne, Havas Media*

Radiant Advisors

Regardless, scaling infrastructure and service management eventually reach a limitation for most enterprise-class implementations.

Protecting the user experience is critical to ensuring that user adoption is not inhibited. If a business user senses that the environment is sluggish or too slow, not only will this impact the highly iterative thought process of data preparation, but the business user will be inclined to extract and work with the data locally, where isolation of computing resources protects their process. This represents a step backward in an enterprise self-service program.

Cloud optimization and elasticity offer greater efficiency for enterprise self-service data when the workload is unpredictable and the overall growth and adoption rate is largely unknown. AWS offers the advantage of affordable storage scalability, support for specific functions such as S3 bucket encryptions and data access, and can also include optimizations for per-second usage billing. Similarly, cloud-based Hadoop environments leverage cost-effective data storage solutions along with distributed processing execution engines such as Spark or Hive.

Cloud optimization can also provide the additional unique feature of business department charge-backs, which some companies find valuable for accounting purposes. Further, data preparation engines should utilize cost-based optimizers to ensure efficient utilization of cloud computing and manage associated charges. Enterprise self-service data preparation must have the ability to elastically add and release computing resources as needed for executing data preparation workflows. Optimizing for elasticity in a cloud self-service data environment prevents the "sticker shock" that most companies are concerned about when considering cloud options.

#6. Artificial Intelligence Assistance

Artificial intelligence (AI) has a very important role as an embedded intelligent assistant for making recommendations in real-time as business users work with data. AI recommendations guide business users through their process to solve complex challenges with sophisticated algorithms that can leverage and learn from other business users' activities to make highly relevant and probabilistic predictions.

In the self-service process, AI can improve search results within the data catalog by going beyond the index-based keyword searches to incorporate other relevant user searches based on which data sources were determined by users to be most useful. Another proven scenario is **recommendations** that

use AI-assisted data integration to infer which data fields are likely intended to be joined based on statistics from the data fields and from other business users' work. Ultimately, an underlying AI engine will tackle the many micro-facets of self-service data to predict what the user is trying to do, minimize wasted time and improve the confidence of decisions.

At Havas Media, creative team members may not always instinctively know how to join data sets to further explore campaign performance. In these instances, AI makes recommendations that remove these obstacles. "We have English majors on their first job being asked to be clever for their clients. Recommendations show them related data when they didn't know where to start," says Le Borgne.

Further, AI is exceedingly valuable in scenarios that involve data security and compliance. Monitoring a constantly growing set of data for business users without having to manually run validation checks ensures continuous security without gaps. For example, AI recognizes data patterns, formats and relationships to make recommendations for likely personally identifiable data that needs to be masked or secured. To be conservative, self-service platforms may allow identified data at risk to be automatically secured or quarantined pending further governance reviews. As a digital assistant for business users, AI-enabled data profiling supports GDPR, PII and other compliance mandates by automatically recognizing potential data security impacts for a user.

Interweaving each of the six pillars together establishes the foundation necessary to work with data independently and efficiently as an enterprise. While each of the six has its value and merit in the self-service data process, the capabilities are interdependent and need to be present at every stage of the process for all of them to work optimally. This combination of the six pillars makes a well-balanced foundation, and when any of the components are missing, substandard or hampered technologically, enterprise self-service data cannot reach its full potential.

All six pillars working together creates a complete self-service data capability. Data catalogs and data prep functions supplant the business user's need to request information about available data sources or the need for IT to program basic data curation. Community and collaboration eliminate the challenge of finding other people in the organization for domain knowledge or validation when using and creating new data assets. Data governance and security as a part of the foundation allows IT to carry out its responsibility to ensure that the business is securely and properly utilizing enterprise data as intended. AI assists business users with intelligent recommendations while cloud optimizations assist IT with managing an elastically scalable computing environment.

Many companies that piece capabilities together by adopting different tools over time don't realize the full potential of having all six pillars as their foundation. Some companies become stuck after one or two years of adoption and will question whether they selected the best tool or need additional tools to achieve their enterprise self-service data goals. In some cases, companies realize limited success in small groups or departments when one or two key capabilities are provided, and they rely on their group knowledge and relationship with IT processes for limited self-service data. Even when companies believe their self-service data is working sufficiently, they don't realize the additional burden of multiple tool connections that are being maintained for the same data sources, the additional efforts in moving and synchronizing data from one application to another or the opportunity to strike a more efficient balance between self-service data prep and other modern programming-oriented environments, such as Apache Spark. In these cases, companies experience improvement from their previous non-self-service environments but are not yet aware of the full potential with a robust platform that encapsulates the six pillars.

"Many companies that piece capabilities together by adopting different tools over time don't realize the full potential of having all six pillars as their foundation."

Benefits of an Enterprise Platform Approach

In research and client work with Fortune 500 companies across the Radiant Network, Radiant Advisors has watched the self-service data industry rise in recent years. We have concentrated a significant portion of our research in the area to understand the approaches and best practices surrounding the implementation of the six pillars.

Over the years, we have seen impressive best-of-breed products in each of the six pillar categories, and we began our hypothesis with "achieving enterprise self-service data requires an ecosystem of partnerships and tightly integrated products." From our self-service assessments done over time, we have observed companies adopt a single product or focus on specific department's needs and question why enterprise self-service data is not being realized. When single tools are adopted, the challenge is ensuring that analysts across a company with access to a tool also have the complete set of six pillars for the necessary shared capabilities at an enterprise level. Additionally, we found that companies struggled to maintain multiple integrated product updates, data integrations and user training across the ecosystem. Where there are a few capabilities and one or two products that can deliver them, the integrations are manageable.

However, with the six pillars required to be the foundation of enterprise self-service data, a single platform that incorporates all six is the ideal approach. A single platform preserves end-to-end governance, provides a consistent user experience and simplifies technology management and scalability.

End-to-End Governance

Data governance rises in importance along with the demand for enterprise self-service data; with more business users accessing data, governance is more necessary. It is this balance between self-service data and governance that allows data usage to become pervasive in the enterprise.

Two common governance failures cause enterprise self-service data roadblocks. Companies either face the risks associated with allowing business users to work with data without controls, or they give access to a select few business power users – who then become the new bottleneck for working with data in the business. Without coupling tightly with data governance, self-service data programs at companies are then limited to only those power users and data stewards with data access and context knowledge.

A self-service data platform provides end-to-end lineage and integration that is necessary for data governance to be present in every aspect of the self-service data process. It is a best practice to apply data governance at data ingestion into the self-service data platform. This ensures that data privacy is always protected and proper context for the data is established and preserved for all subsequent derived data work and publishing. For data discovery and data catalogs, data governance must properly define the data and its references, owners and data stewards for communication. As data preparation and automated workflows are created and deployed, the newly created data assets must also go through the data governance process before making them available for others. Without having data governance as part of the platform, the organization is at risk of individuals inadvertently introducing non-compliant derived data sets into the business environment. Collaborative data governance and community engagement help ensure that data is validated and governed at the end of the data prep process for end-to-end visibility of data lineage of the many-branched data workflows used through the enterprise.

Consistent User Experience

The single-platform approach to self-service data fosters enterprise adoption and broad use as it lends itself to a common user experience. Working through the data process to gain business insights is a highly iterative process, and when users must traverse multiple products within their self-service data environment (e.g., between data discovery, catalog and data preparation), context switching causes significant disruption to the user experience.

A user's thought process is interrupted every time they have to stop their work, log into another application (remember another user name/password) and switch to another application user interface and working paradigm. While users may find intuitive user interfaces, layouts and visual design in products that support self-service data, the differences in application design and terminology further inhibit their ability to work smoothly. These tools also require multiple user training programs. Users prefer to stay on one platform and focus on the task at hand for efficiency and continuity.

The business user's work is accelerated when all the capabilities exist transparently within a single platform. Switching capabilities between the six pillars should be as easy as accessing a menu or alternate tab. When performing data discovery, a business user should have data catalog and collaborative access embedded or have AI recommendations occurring at the point in the process when they are most useful.

Technology Management and Enterprise Scalability

Technology management is the most common adoption benefit of platform solutions over architecting an ecosystem, and its impact cannot be understated. Maintaining integrated environments is very challenging for IT, as each vendor has their release schedules and technology stacks, not to mention technology vendor management. Coordinating multiple data connections and regression testing integrations will require IT time and resources to execute and verify. With a platform, all the integrations are internal and, therefore, tested by the vendors before releasing upgrades or patches. Furthermore, when issues do arise, IT benefits from a single vendor that will take ownership for problem resolution and deliver any corresponding patches. Despite best intentions from multiple vendors, debugging integrated applications is difficult, as is determining where the problem resolution should ideally occur.

Achieving enterprise-wide scalability to handle thousands of users' activity and data-intensive workflows requires that each component in the self-service data environment is scalable. With an ecosystem of products, any component that does not scale prevents the ecosystem from being fully scalable. With a platform, this is greatly simplified because the components scale together.

Additionally, the ecosystem approach presents further challenges with managing multiple vendors' ability to scale their applications and their integrations across open standards, whereas a platform conducts these communications internally. Cloud platforms make scalability achievable by leveraging elastic computing and storage. Further, a self-service data platform will have a lower total cost of ownership when considering the technology management and enterprise scalability aspects of an ecosystem of products.

Radiant Advisors

Several important initiatives clearly benefit from self-service data's key value propositions. Customer 360 programs unlock speed and agility with the domain knowledge of business users. Compliance programs such as GDPR ensure that data is protected and business users enabled with self-service data work within those standards. ERP consolidation and systems migration projects leverage self-service to benefit from business users' specific business domain knowledge for discovering complex data mappings and exception handling. The comprehensive capabilities of the six pillars within a self-service data platform allow these leading business initiatives to be more agile and higher quality, ultimately delivering value more quickly.

Customer 360 and Customer Intelligence Initiatives

Customer 360 programs yield deep customer insights and understanding by analyzing all customer interactions for a single view of the customer to optimize marketing and sales efforts, increase total customer lifetime value, and improve customer satisfaction and retention. A primary challenge lies in requiring the knowledge to appropriately and usefully integrate data from a variety of disparate sources across many business functional areas. Only business users possess the knowledge to understand and curate the data they need and translate it into valuable data sets of customer interactions. The self-service data platform approach uniquely solves the challenge of connecting business users' domain knowledge with their data to discover, profile and make sense of data integrations and share data across other business groups needing a complete view of the customer.

By far, the biggest benefit of self-service data for Customer 360 is derived from data discovery and data prep.

Marketing departments rely on data discovery and data prep for targeted marketing and micro-segmentation to capture and make use of customer attributes at a finer grain. Self-service data prep, coupled with AI recommendations, empower these marketing teams to integrate data sets from sales customer profiles, external data on companies and customer service interactions to develop micro-segments of customers based on data for accurate targeted marketing. Routinely, marketing will receive results from the latest marketing campaigns and A/B testing that need to be quickly prepared, understood and integrated to create the next, improved campaign.

Rapid, ongoing data prep and integration of many internal and external data sets provides insights for sales, marketing and support teams to work from

"Rapid, ongoing data prep and integration of many internal and external data sets provides insights for sales, marketing and support teams to work from and collaborate across the customer journey."

and collaborate across the customer journey, from targeted marketing to sales nurturing to customer satisfaction and support.

Customer support and marketing teams leverage sentiment analysis from multiple forms of customer interactions – ranging from direct, indirect and anonymous or audience sentiment – to improve satisfaction, retain customers and increase customer lifetime value. While data from sales and CRM systems are more traditional data types, other data from marketing campaigns and digital engagement, such as websites and mobile apps, are typically non-standard or unstructured. Integrating social sentiment with these systems and other third-party data, such as survey and market data, allows customer support to integrate sentiment analysis results to correlate key sentiment categories to customer groups, employees and products for improvements. Self-service data is the most effective way for teams to quickly discover and integrate all of these varying data sources, relying on an enterprise data catalog, governance and AI to aid and enhance the process.

Throughout all Customer 360 efforts, careful attention must be paid to protecting the privacy of customers. A single self-service data platform that preserves end-to-end governance across all data work ensures that everyone fulfills their responsibility to secure and protect the privacy of customer data with data masking and data lineage for other departments to leverage that data.

GDPR and Compliance

Following a two-year transition period, in May 2018 the European Union's General Data Protection Regulation (GDPR) will go into effect and join the list of data regulations that companies are mandated to follow or incur substantial fines. Companies that hold any data about people in the EU will be required to demonstrate that every piece of personal data is managed and accounted for thoroughly. The intent of GDPR is to ensure a person's right to having their personal data kept private and secure as well as permanently erased if requested.

"GDPR is coming soon and it will have a big impact," says Le Borgne at Havas Media. To Le Borgne, compliance and security are imperative to not only meet standards and mitigate risk but also to reassure clients their data is safe and protected.

Self-service data platforms must follow a proactive approach to ensure that end-to-end data governance and security are incorporated to identify and manage critical elements of personal data throughout the enterprise. While

most companies are focused on the personal data related to customers, any personal data related to employees, suppliers and partners is also covered in the GDPR mandate. A self-service data platform that is designed to account for the protection and governance of data end-to-end across systems and processes is critical to ensure compliance while enabling business users to work with data. Once established data governance and security procedures are in place, the self-service data platform can provide a record of processing activities through detailed audit trail reports.

Compliance and privacy requirements are also extended to derived data, which is an area of unique concern for enterprise self-service data. This is where data lineage plays a critical end-to-end role for data processing. Business users need to understand where data sources come from before they can consider using the data in their work. Likewise, business users who are authorized to work with personal data must be careful when integrating it with other data sources because the resultant data sets must also fully adhere to compliance standards. A single self-service platform will establish the compliance and governance standards within the platform, thus protecting sensitive data even when a business user may not be aware their actions could result in non-compliance. This reduces pressure on business users and enables them to work safely and confidently with data.

With the huge volumes of data in the enterprise, companies realize the need to support, sustain and enhance compliance efforts by relying on forms of AI to proactively inspect data at ingestion, at rest and at access. AI routines can be trained to detect personal data patterns and inferences and recommend that a data element be masked or marked accordingly for users with proper data access. AI detection routines intercept data at the point of ingestion into the enterprise. The same AI routines can be embedded in the self-service environment during data preparation to detect the creation of personally identifiable information when combining different data fields. AI is becoming the digital assistant that helps companies with the challenges of maintaining data privacy and quality.

ERP Consolidations and Migration

Self-service data plays a critical role in major ERP consolidations and system migrations to accelerate projects and reduce risks. At the heart of these major initiatives is understanding the state of current data, mapping this data into new business data and process structures and validating data quality during parallel processing. This is mission critical in the consolidation of financial systems from company acquisitions and mergers. However, companies also migrate other

"A self-service data platform that is designed to account for the protection and governance of data end-to-end across systems and processes is critical to ensure compliance while enabling business users to work with data."

Radiant Advisors

major operational systems, such as Sales CRM and business applications, to different business applications and cloud software-as-a-service platforms for modernization efforts and to meet evolving business needs.

Initially, companies need to understand and validate the data in the current systems for quality based on current and legacy business processes. Typically, these systems and their data existed while the business processes and rules evolved over time. Consequently, only the business domain experts can conduct this assessment by working first-hand with the data in a highly iterative and collaborative manner. Data discovery empowers these business users with collaboration, data governance and a data catalog to capture the formal and tribal knowledge for the new system. This exploratory phase of the process enables teams of business users to perform self-service data analysis in order to make better decisions for mapping and migrating the data to the new ERP or business system.

Major system consolidations and migrations are a one-time custom project; any development and testing of data migrations will likely not be used again. Therefore, it is inefficient for financial analysts to define mappings and transformation rules for IT to then develop and test code prior to the migration. Self-service data speeds the project time because financial analysts can conduct the exploration, document their data preparation workflows and validate individual exceptions based on their domain knowledge.

An enterprise self-service data platform is ideal for the process of accessing ERP data, profiling and verifying the numerous business rules and then properly mapping its context into a new ERP or business system. Empowered with all the capabilities of the six pillars, financial analysts can build workflows with the assistance of AI recommendations for data quality and integration and then collaborate with other domain experts in the community to verify and increase accuracy. Workflow automations allow for parallel processing of ERP systems to assure quality and validate reporting. This is also an opportunity for financial analysts to capture information in the central data catalog for future use.

Conclusion

The promise of enterprise self-service data is attainable when all six pillars are present. There is no quick route to becoming a data-centric organization; people, processes and technology need to come together for this purpose. The six pillars provide the foundation for companies to navigate and assess their own journey to democratize data across the enterprise efficiently and safely.

Collectively, the pillars support business users with the freedoms of a self-service data process and the protection of data governance and security in place as they use data to address business needs. Recognizing and accounting for these six pillars will lead to a more complete definition of an enterprise self-service data strategy and the capabilities required to make it a reality in your organization.

About the Author

John O'Brien is Principal Advisor and CEO of Radiant Advisors. A recognized thought leader in data strategy and analytics, John provides research, strategic advisory services and mentoring that guide companies in data strategy, architecture, analytics and emerging technologies.

This research report sponsored by:

unifi

To learn more about Unifi Software, visit www.unifisoftware.com

About Radiant Advisors

Radiant Advisors is an independent research and advisory firm that delivers innovative, cutting-edge research and thought-leadership to transform today's organizations into tomorrow's data-centric industry leaders.

Radiant Advisors
Boulder, CO USA
Email: info@radiantadvisors.com

To learn more, visit www.radiantadvisors.com