Modern data architecture as a strategic lever in the competitive landscape

Data has become the life blood of businesses and properly managing that data to gain the most value is becoming ever more important as businesses seek to remain competitive. This Insights paper will address the importance of investing in the processes, practices, and technologies to maximize the value of data in an enterprise.

Why is an investment in enterprise data architecture so important?

Many organizations are still struggling today to manage and organize the massive amounts of data they have generated or acquired through third party interaction. It is a situation that prevents business from extracting the most value from their data which turns data into more of a liability than an asset. Although there are many challenges associated with managing and orchestrating data, organizations need to bring order to the chaos by building a modern data architecture, which will enable leverage of this data at scale in an organized and more purposeful way.

Business reasons for pursuing a modern data architecture?

The analysis and interpretation of data has become one of the competitive differentiators that businesses can use to improve operations, increase efficiencies, can use to improve operations, increase efficiencies, and garner a clearer picture of various markets. Businesses looking to remain competitive can no longer ignore the importance of data and should now be investing in data architecture and operations as a key lever to achieve their strategic goals.

Enterprises across multiple industries generate large amounts of data as part of their operations and interactions with external parties. Well organized and clear understanding of this data can lead to key benefits like improving product development, enhancing finance operations, managing risk and driving data driven actionable events. These examples take form in our day to day lives when we get those fraud alerts that notify us that our account could be compromised or the proactive marketing that we get for a new car loan because we are close to needing a new vehicle. With data that is organized, analytics whether predictive or reactive is possible and separates leaders from the laggers! With that, let's discuss what modern architectures look like and how you can be a leader in your industry.

Creating and managing a high value trusted data environment to benefit the various functions of the company is more an art than a science and is unlike application development, which has very specific and predetermined paths for how data is to be used. The continuous need for data discovery with a highly iterative design can make managing data challenging. In addition, the need to support the highly changing data environment through a mature operating model is critical.

Enterprise data applications need to support various consumption patterns that can't always be pre-determined. However, the need to protect data from inappropriate use is a fundamental requirement of any data solution, especially on a modern data stack which entails a key dependency on vendor partners to ensure a secure and trusted infrastructure. Data governance and data tagging is also a key dependency when it comes to data protection thus maturity in data governance capabilities is more so imperative.

Investing in data as a strategic pillar drives efficiency and is mindful of data consumers and key stakeholder needs. Investing in technical architectures proves to be strategic and building a modern data architecture executes on strategy that is critical for realizing the value offered by data.

Board members and C-suite executives view "Inability to utilize data analytics and 'big data' to achieve market intelligence and increase productivity and efficiency" among the Top risk issues for their organizations over the next decade.

Executive Perspectives on Top Risks for 2023 and 2032, Protiviti and NC State University ERM Initiative, December 2022

Building a successful modern data architecture requires combining an organizational model with an operational model and that can be a significant challenge for many organizations.

Key points on building a modern data architecture:

There are several best practices that help to define how to establish a modern data architecture. In

our experience, we recommend focusing on the following areas:

- Defining a target state conceptual architecture with prescribed design patterns
- Discovering the risks of not establishing a foundational architecture and common set of capabilities
- Ensuring organizational models guide process and establishing key roles
- Putting the puzzle pieces together with well-organized data operations

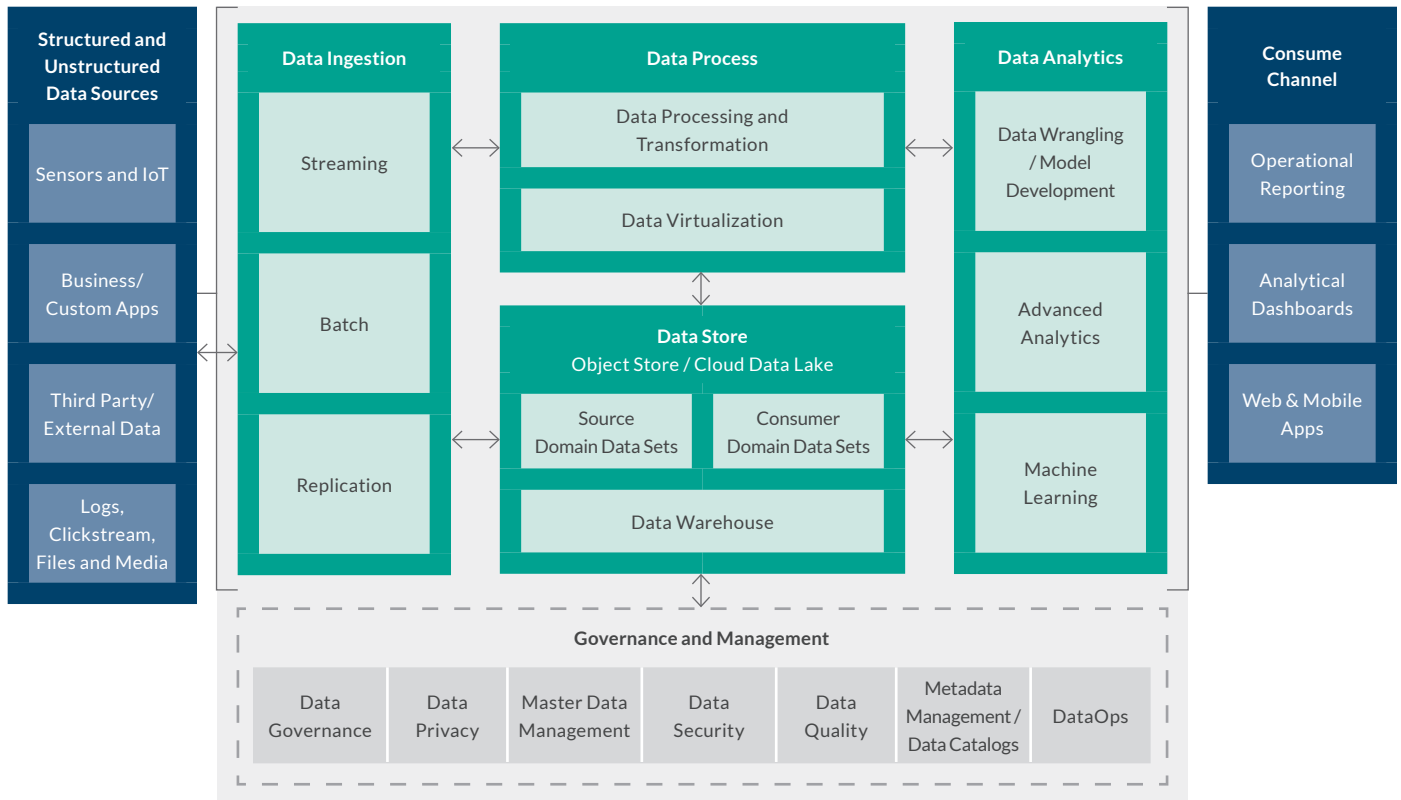### Defining a target state conceptual architecture with prescribed design patterns:

It is critical to understand the organization's current data providers or producers and its data consumers when defining a target state conceptual architecture that features prescribed design patterns. A good architecture will account for functional and non-functional needs for the platform.

Functional needs include data accessibility methods, data quality and timeliness standards, data privacy, access controls, change management and stakeholder specific data requirements. While, non-functional needs include availability, scalability, performance and other infrastructure capabilities like backup/recovery and disaster recovery.

A conceptual architecture describes the various high-level components of an enterprise data eco-system, which takes into account the business model, goals and domain design. There are times when those components correspond to specific capabilities, such as a data warehouse or analytic application.

However, a conceptual design is chiefly concerned with mapping the purpose and logic of how the system functions in support of business goals. Common constructions or relations of these components is termed a design pattern and helps architects compare or attempt to replicate the setup of systems known to have certain advantages.

Historical constraints in understanding and working with underlying data infrastructure once inhibited business leaders and key stakeholders from playing a more interactive role in the data system design process. However, modern architectures make it possible for applied requirements to be gathered first.

Structured and Unstructured Data Sources	Data Ingestion	Data Process	Data Analytics	Consume Channel
Sensors and IoT	Streaming	Data Processing and Transformation	Data Wrangling / Model Development	Operational Reporting
Business/ Custom Apps	Batch	Data Virtualization	Advanced Analytics	Analytical Dashboards
Third Party/ External Data	Replication	**Data Store** Object Store / Cloud Data Lake — Source Domain Data Sets — Consumer Domain Data Sets — Data Warehouse	Machine Learning	Web & Mobile Apps
Logs, Clickstream, Files and Media				

Governance and Management

Data Governance	Data Privacy	Master Data Management	Data Security	Data Quality	Metadata Management / Data Catalogs	DataOps

It is critical to have buy-in across the organization. Without feedback from senior stakeholders down to everyday, front-line users, it becomes impossible for data-driven programs to realize their full strategic potential. It is crucial to have input and feedback from all these voices informing the context and objectives of an enterprise data solution.

Traditional architecture patterns were focused mainly on an IT centric view. The move to modern data architecture patterns are growing in adoption across the industry. This growth can be attributed to the emergence of architecture models that support higher levels of federation like the Data Mesh.

In this domain-based data architecture model, domain-managed and governed data products versus raw data sets are the primary inputs into the process. These data products need to be ready to use when they are published on the data platform. The responsibility lies with the domain owners to ensure that the product is fully governed and of optimal quality.

To create consistency across all the business data domains, an overarching central data governance process drives how the data products are published through data standards, catalogs, lineage, policies, common and shared infrastructure, and tools.

Automation is a critical component of modern data architecture, enabling agile onboarding of new data sets and publishing data products. Getting data products to market faster using an automated factory model, results in quicker monetization or speedier compliance with the regulatory risks that are increasing every day. Metadata-driven approaches to moving and transforming data through the architecture are best practices to accomplish optimal automation and agility in releasing data products.

DevOps automation is also a key factor in both architecture agility and nimbleness in supporting business needs. Traditional operating models where data engineering and data operations worked in silos to create data products are now obsolete; the modern data stack operates much like a software application development stack, providing the data engineer greater ability to develop, test and deploy data pipelines in an agile manner. Modern data architecture must allow for integration of native or third-party tools and technologies that enable continuous DataOps to ensure faster delivery of data products and analytics.

With global, country specific, local and state regulations, such as GDPR, CCPA/CPRA, CPA, VCDPA (Virginia), UCPA (Utah) and CTDPA (Connecticut) (State Level Comparison Charts of Data Privacy Laws in the U.S. | Bloomberg Law), data privacy and security should be a major focus for organizations. The modern data architecture needs to handle the myriad of data security and privacy requirements. On the data security front, the architecture should be able to adhere to various audit and compliance requirements and legislations like SOC1, SOC2, HITRUST, FedRAMP, HIPAA and others.

The ability to model data obfuscation rules around PII and PHI is a critical requirement to protect sensitive data. For public companies, it is also imperative to comply with SOX requirements like access control/information security, change management, incident management, physical and logical security, and backup/recovery. Establishing data clean rooms in an organization enables data sharing and monetization of data products in a compliant manner.

Modern data architecture powers a myriad of business use cases today such as fraud alerts on personal bank account activity and IOT (telematics) based transportation improvements.

These are some real world examples of how well thought out modern data solutions are used to develop and deploy business-driven data products that positively impact businesses and communities.

DataOps is an emerging process based on agile software engineering and DevOps that encapsulates many data management best practices and helps generate better quality data and larger quantities of data analytics products. DataOps can enable companies to deliver data products faster and stay ahead of their competition.

The DataOps Manifesto - Read The 18 DataOps Principles).

Discovering the risks of not establishing a foundational architecture and common set of capabilities:

To reduce risk, organizations should establish key foundational principles and embedded control frameworks. Risk can come in many forms, such

Links to modern data architecture examples by the major cloud vendors:

- **AWS:** A new era of data: a deep look at how JPMorgan Chase runs a data mesh on the AWS cloud — SiliconANGLE

- **AWS:** Design a data mesh architecture using AWS Lake Formation and AWS Glue | AWS Big Data Blog (amazon.com)

- **GCP:** Data Mesh on the Google Cloud — A Technical Architecture Sketch | by Sven Balnojan | Towards Data Science

- **Azure:** Cloud-scale analytics — Microsoft Cloud Adoption Framework for Azure — Cloud Adoption Framework | Microsoft Learn

as security breaches, ineffective controls, and data proliferation. It is important to note that the federation of responsibilities in the delivery model does not mean federation of the architecture, software and technical design.

A well-architected solution embeds controls and governance to allow proper oversight and shared investments that are key to avoiding rogue processes and data proliferation. Data security at a minimum requires an operational model that ensures governance and central oversight.

Ensuring organizational models guide process and establish key roles:

As mentioned above, there is an element of art when designing the elements of a model. A good example comes in the form of evaluating a particular situation. Take for example, a consumer bank. Many banks are organized by lines of business that have their own organizations to support their respective operations, systems and business goals. Given that, we would want to understand how the conceptual and domain design should align with that organization as a data producer and data consumer.

As data producers, the bank has responsibilities for providing data back to enterprise functions such as risk, finance, compliance, and others. Due to this federated responsibility, a collaborative model should agree on the architecture based on each organization's skillsets,

roles and overall data management operational model and principles. Investing in leadership to create the optimal organizational model with clear roles and responsibilities is necessary.

When the chief data officer (CDO) role first came about, it was a key IT role and was accountable for architecture and infrastructure. Yet, over the past decade or so, it has evolved into a more business-aligned role that is responsible for creating and governing the roadmap of data needs and goals.

With that, it is also important to have an equally senior and accountable role in technology to ensure IT standards, controls, common investments, and such are enabled. We challenge companies to consider this key role and ensure it has appropriate weight in the organization to enable these foundational capabilities. The concept of a "CTO of data" has evolved and is gaining some traction across organizations.

Putting the puzzle pieces together with well-organized data operations

Once an architecture with design patterns is defined, the organizational model reviewed with each line of business or data producers, the data usage needs understood, and roles and responsibilities agreed upon in the overall model, it is time to establish the operating model.
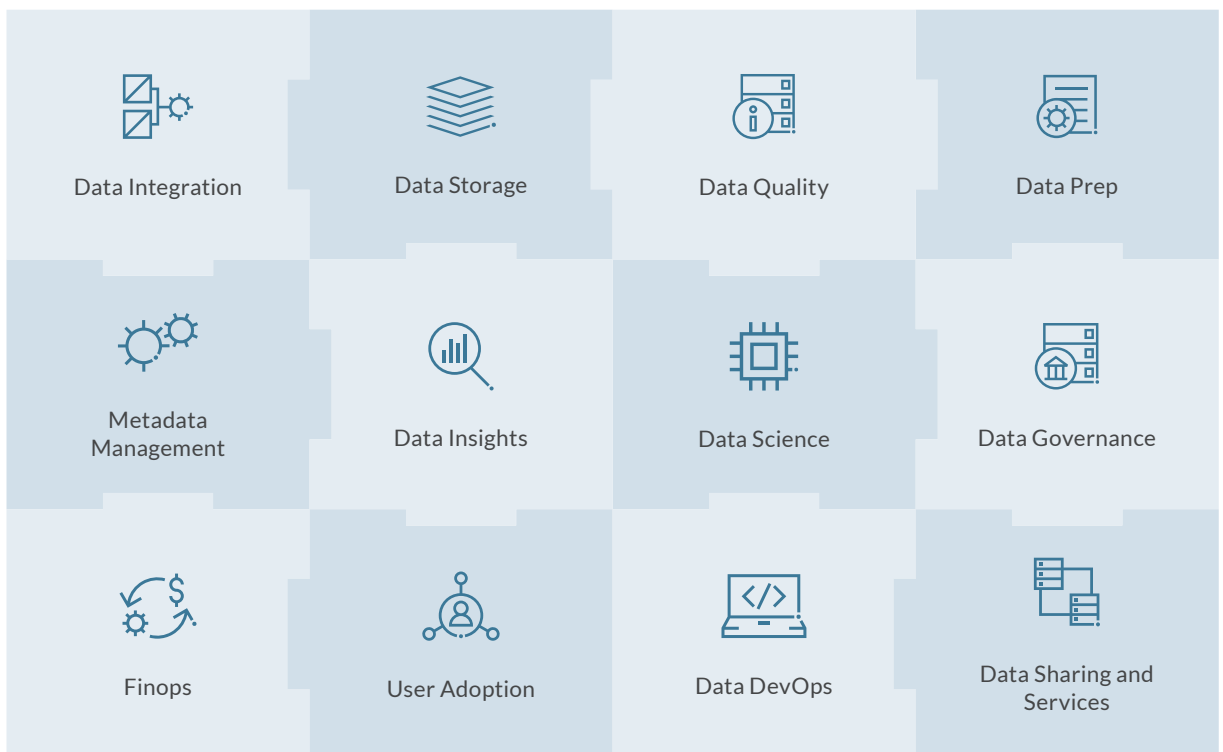
The key tenet of an operating model is ensuring all key stakeholders have a way to interact (typically though stakeholder committees, data governance forums and/or executive committees). Defining the goals and outcomes of those interactions is key, as well as decision-making protocols.

Budget planning and aligned roadmaps are a critical exercise and cadence with the teams to ensure consistent alignment to those goals is needed. Also, aligning financial management and transparency with agreed upon cost allocations and such are more of a focus in a modern architecture, to ensure proper design to accurately report and allocate those costs are predesigned and planned.

Lastly, ensuring involvement of key stakeholders from compliance functions and cyber security early on is an absolute to ensure alignment to policies and requirements, as well as prevent risks.

Data operations also include the technical aspects of change management, access management and various other controls which can be automated as part of data devops frameworks. That model/design should be established to ensure the IT/code operational model with proper roles and responsibilities.

• • • **Data operations**

Data Integration	Data Storage	Data Quality	Data Prep
Metadata Management	Data Insights	Data Science	Data Governance
Finops	User Adoption	Data DevOps	Data Sharing and Services

How Protiviti Can Help

Organizations are harnessing the power of data to improve processes, drive new business opportunities and increase competitive advantage. We provide services to design, source, transform and analyze data to empower your business by modernizing your enterprise data architecture. Using our combination of strategic vision, proven expertise and practical experience, we will collaborate with you to enable the development of a cutting edge and pragmatic data architecture.

Our capabilities to enable a modern data architecture include:

- Developing a data strategy and roadmap tailored to your organization's specific needs and growth objectives inclusive of architecture, organizational planning and data operations

- Establishing a data governance framework with aligned data management policies, tooling and operations

- Creating best-practices-based, streaming or batch ETL/ELT frameworks on a variety of cloud platforms to ensure your data is flowing properly

- Providing high-performing storage designs and implementations for data lakes and data warehouses supporting both operational and analytical data workloads

- Establishing policy authoring and design standards that deliver high-performing design standards and implementations for data lakes and data warehouses

- Launching a data security and privacy program that incorporates the appropriate data backup and recovery testing strategies, methodologies and testing models

- Designing a master data management strategy that will carry your organization well into the future

- Delivering analytics and reporting capabilities to enable self-service reporting, real time events, data discovery and more

Contacts

Wendy Luebbe
Managing Director —
Technology Consulting,
Enterprise Data & Analytics
+1.980.939.3323
wendy.luebbe@protiviti.com

Venu Nair
Director — Technology
Consulting, Enterprise Data
& Analytics
+1.312.399.4606
venu.nair@protiviti.com

Protiviti (www.protiviti.com) is a global consulting firm that delivers deep expertise, objective insights, a tailored approach, and unparalleled collaboration to help leaders confidently face the future. Protiviti and its independent and locally owned Member Firms provide clients with consulting and managed solutions in finance, technology, operations, data, digital, legal, governance, risk and internal audit through its network of more than 85 offices in over 25 countries.

Named to the 2022 Fortune 100 Best Companies to Work For® list, Protiviti has served more than 80 percent of *Fortune* 100 and nearly 80 percent of *Fortune* 500 companies. The firm also works with smaller, growing companies, including those looking to go public, as well as with government agencies. Protiviti is a wholly owned subsidiary of Robert Half (NYSE: RHI). Founded in 1948, Robert Half is a member of the S&P 500 index.

protiviti®