Microsoft

A Data-Led Approach to Digital Innovation
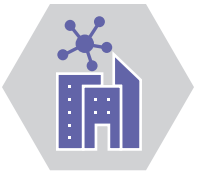
Authors

Meaghan McKeown
Susan Meldahl
Pankaj Srivastava

Key Contributor

Kelsey Radwick

Contents

Microsoft

Introduction

The fourth industrial revolution is upon us, arriving in the form of digital transformation. The leaders – and laggards – of this new era will in large part be determined by a single factor: the effectiveness of their approach to data management. Leveraging data to build better products and higher value services as a point of competitive advantage is nothing new. What *is* new is the volume, velocity, and variety of data that has been enabled by cloud computing and the corresponding need for companies to quickly process and make sense of their data. The world's data is expected to reach 175 zettabytes in 2025 with a 30 percent real-time consumption rate[1]. To put that into perspective, the world's collective internet usage reached just one zettabyte in 2016[2].

The companies that can unlock and activate this vast amount of data will see several topline benefits. According to recent research from McKinsey, data-driven companies are 23 times more likely to acquire a customer, six times more likely to retain that customer, and 19 times more likely to be profitable as a result. To this end, leaders are recognizing data as the new basis of competition and as an underlying force behind innovation, productivity, customer loyalty, and growth.

However, building a '**modern' data estate** is not an easy endeavor and, in some cases, will require a complete overhaul of data and business strategy. As a result, companies are turning to partners for their expertise and guidance on how to make data a central tenet of their cloud journey, opening a range of new and untapped revenue-generating opportunities for partners. Thus, considering the data challenges and opportunities facing customers and partners, this thought piece aims to accomplish three things.

The first is to propose a data-led approach as fundamental to the digital transformation journey. Unlike the application-led approach, which brings data along with applications, thereby leading to piecemealing of data, a data-led approach 'futureproofs' the business by taking a holistic view of data across the organization from source to usage. As such, Section I serves to clearly define what a "data-led" approach looks like and how it differs from other approaches.

Once customers understand the value of a data-led approach the job of the partner is to help facilitate this journey. The second key aim of this thought piece is therefore to underscore the core design principles of a modern data estate as outlined in Section II.

And finally, the third key aim is to outline the opportunity for partners to (a) expand the surface area of their engagements with customers and (b) introduce new models for monetization. With the era of data-powered digitization upon us, the imperative for companies to modernize their data estate and bring advanced capabilities to their businesses has amplified. Partners must move quickly to build rich service offerings that meet this need. Section III describes four categories of managed services in the realm of data, for partners to leverage as a framework for building out their offers.

[1] Seagate: Data Age 2025, The Digitalization of the World
[2] Bernard Marr & Company: How Much Data Is There In The World

Microsoft
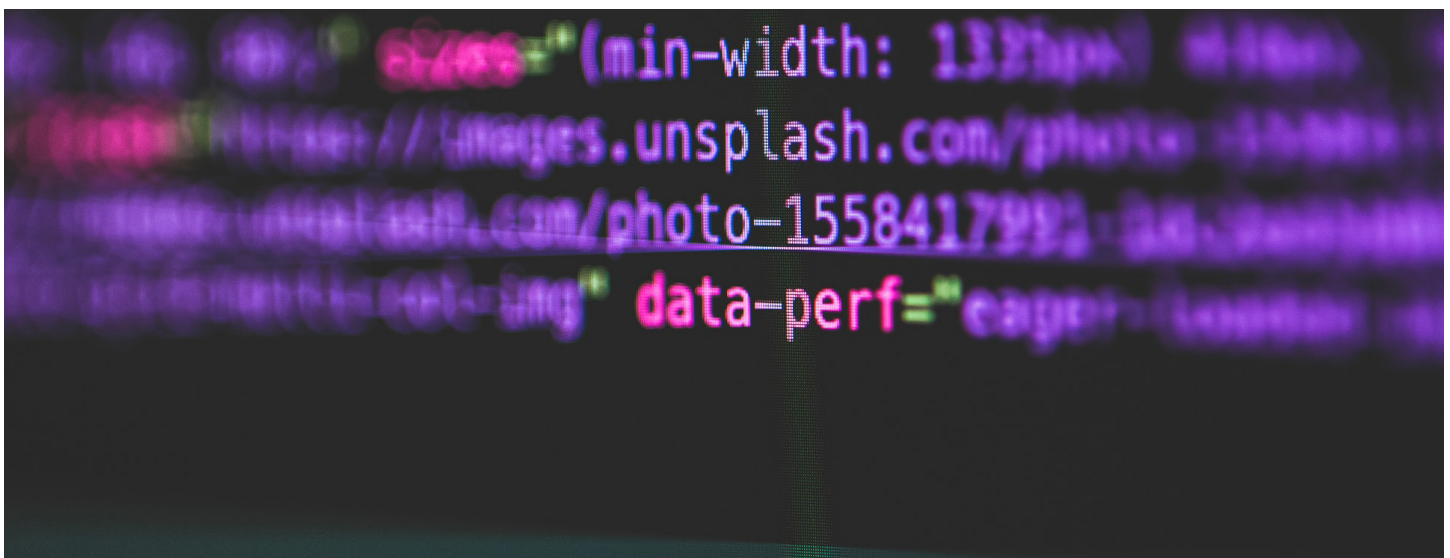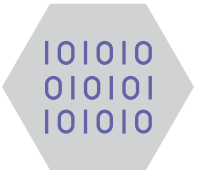
Section I: The Data-Led Approach

What is the 'data-led' approach and how does it differ from application-led approach?

Most partners begin customer engagements with a 'cloud envisioning session', a standard methodology for educating customers and helping them make key decisions regarding their move to the cloud. The envisioning session is typically followed by a cloud assessment to gather insights on the customer's application landscape and to identify application dependencies which are later used to design a cloud migration and modernization plan. Typically, these plans are heavily pivoted on lift-and-shift migration, application modernization, and related managed services. Data is often a secondary thought in this approach, usually piecemealed into different application/workload silos without a view of how to ensure trusted, connected data across the entire application portfolio.

At the same time, partners who have invested in building AI capabilities (supported by data science talent) and are executing AI projects at a certain volume (100+/year) are struggling to find repeatability and opportunity for automation across the AI project portfolio. Lack of scalability severely limits operating margin expansion.

By contrast, a data-led approach is grounded in *data* discovery and realizing value from the beginning, very quickly opening the door to broader discussions on data estate modernization. Once this door has been opened, the data journey can be customized to fit the needs and appetite of the customer. For some, the data journey will occur separate from their application journey, accelerating the pace of data estate modernization. For others, the data and application journeys will run in parallel – often converging to build intelligent tools and processes that unleash the possibilities of digital innovation (see Figure I).

When partners implement a data-led approach to customer digital transformation, smaller data projects that would otherwise be executed in siloes materialize into a connected, holistic approach to modernizing the data estate. During the data discovery process, partners may initially start with a relatively small business challenge. Examples could include financial reporting challenges, data quality issues, security incidents, or a business need such as point-of-sale data consolidation, IoT project integration, improved efficiencies, reduced costs, and business insights and visualization. Once value/ROI is proven for the initial business challenge, some customers will be hungry for radical, large scale data estate transformation. Others may take a more phased approach, continuing with smaller use cases and over time, evolving the broader data estate.

As customers embark on the data estate modernization journey, there are two best practices partners should leverage in order to maximize value creation for the customer while solidifying themselves as critical allies in the digital innovation process.
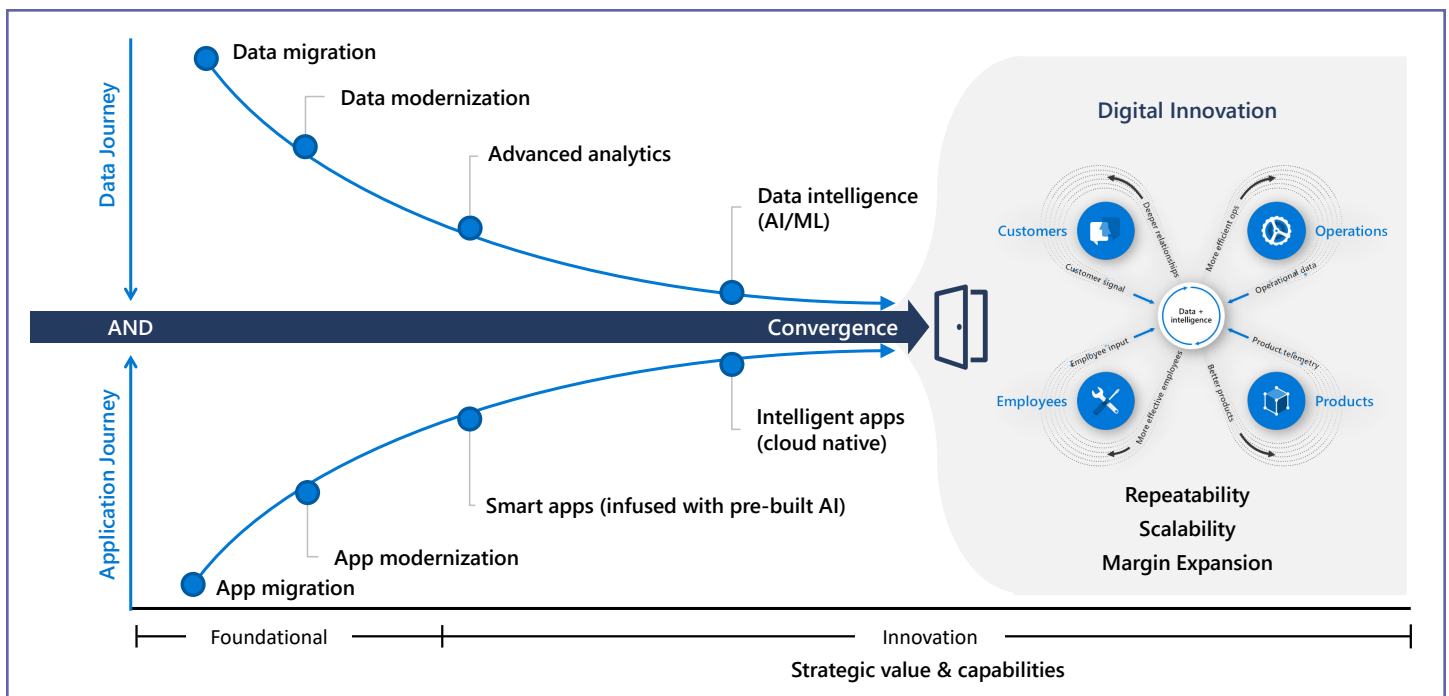

Figure I. Journey to Unlock Digital Innovation

Microsoft

The first is *continuous and iterative design* – Operating in short, one to four-week sprints based on Agile methodology accelerates the discovery process, proves value early, and validates the investment. In the early stages, ROI is best demonstrated by focusing on a single use case and the impact on a set of prioritized success indicators. This helps to achieve customer buy-in early before scaling the customer project to larger parts of the estate. As the process of discovery, agile deployment, and value realization is executed over additional use cases, the value multiplies. The data builds upon itself, creating new value and encouraging new sources which in turn creates more value.

The second best practice is *repeatability and scalability* – Execution across multiple use cases enables partners to catalyze value realization for their customers in the process, establishing structure and governance while realizing operational efficiencies that increase profitability. Execution across multiple use cases enables partners to become increasingly programmatic in their engagements with customers, not only providing valuable structure for customers, but enabling partners to realize operational efficiencies that increase profitability. Three assets in particular are key to building a repeatable and scalable date estate practice. The first is a solution library comprised of algorithms, machine learning/analytical models, and data patterns collected across customer projects to be leveraged for future projects. The second is a continually refreshed centralized knowledge repository containing reference architectures, solution templates, use cases, DevOps process templates, solution runbooks, deployment guides, governance repositories, analytics practice libraries, data connector repositories, and frameworks for assessments, scripts, and security. And finally, the third key asset is a standardized pre-sales methodology to showcase the "art of the possible" to potential customers. This methodology should help customers visualize specific use cases and the business value that can be unlocked by modernizing their data estate. Appendix I is one such visual to help customers image a "before and after" state. Invariably, this conversation leads to discussion and identification of additional use cases.

Microsoft
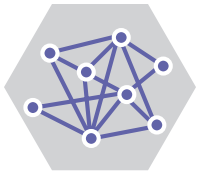
Section II. The Modern Data Estate

What are the core design principles of a 'Modern Data Estate'?

A 'Modern Data Estate' leverages the power of the cloud to ingest, process, store, serve, and visualize structured and unstructured data from multiple different sources. Generally, the litmus test for a Modern Data Estate is whether it delivers on the following four tenets: (1) reduced costs, (2) improved performance, (3) reduced risk, and (4) advanced analytics. These are achieved through adherence to the following fundamental principles:

I. Establish an enterprise-wide data hub consisting of a data warehouse for structured data *and operational analytics* and a data lake *or lakes* for semi-structured and unstructured data. This data hub becomes the single source of truth for data and facilitates implementation of data governance, security and quality.

II. Integrate relational data sources with other unstructured datasets with the use of big data processing technologies.
III. Use semantic modeling and powerful visualization tools for simpler data analysis.

Figure II illustrates the interaction of these principles. Data from a variety of unstructured, structured, and semi-structured sources is published to a virtual lake that users centrally access to generate reports, APIs, and dashboards. To eliminate duplication and optimize efficiency, compute is brought to the data lake as opposed to data being extracted, copied locally, and then compute applied. A layer of shared managed services and analysis tools are also made available to maintain the integrity of the estate and help users. Finally, the entire system is contained within a single security context to ensure data is protected and secure.
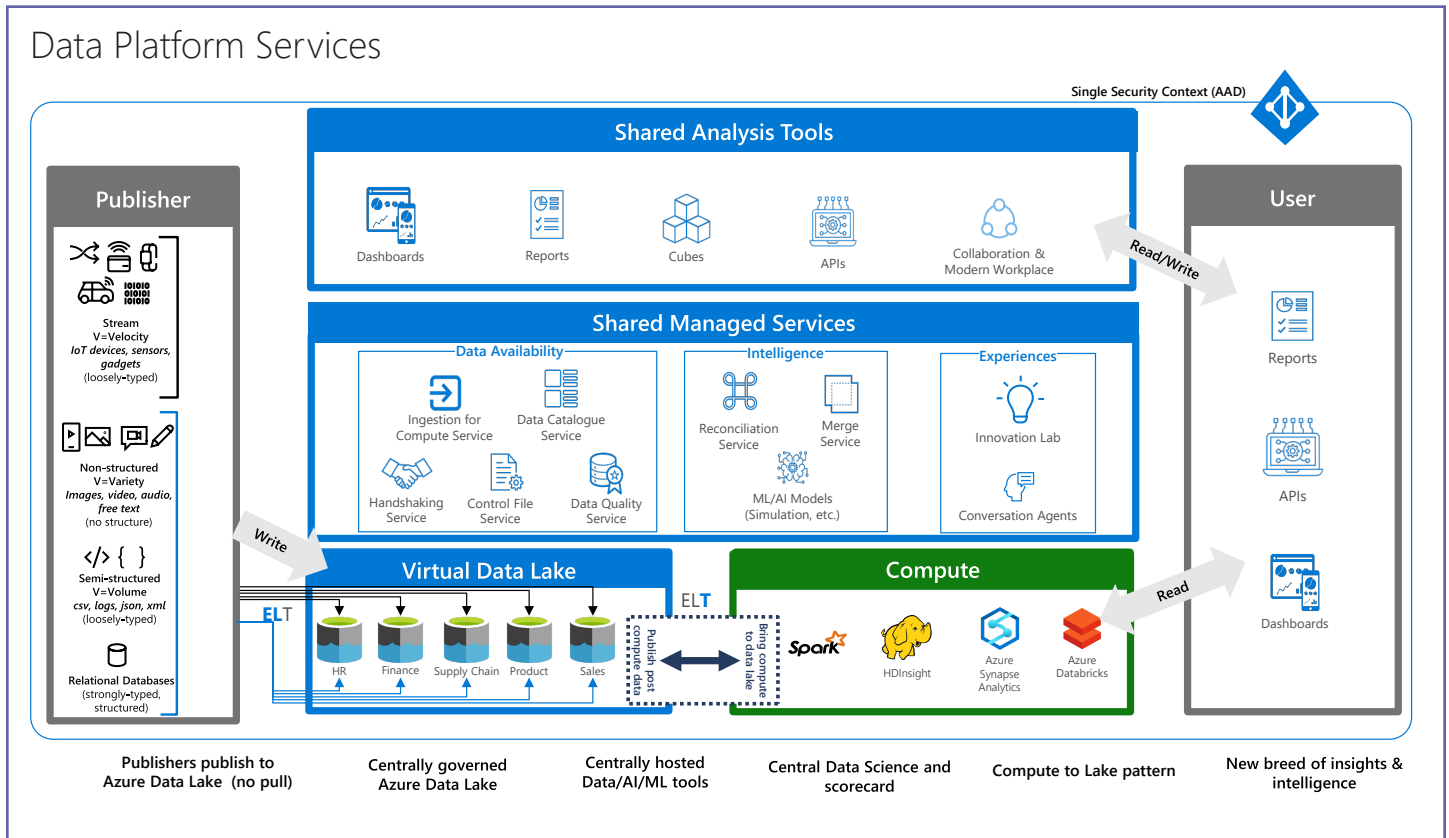

Figure II. Data Platform Services

Figure III brings this framework to life with an example scenario of a Modern Data Estate architecture in Microsoft Azure. The solution architecture demonstrates the ability of a Modern Data Estate to ingest a variety of complex data from multiple sources, including high velocity data from IoT devices, sensors, and other gadgets as well as high variety and high volume data from images, video, audio, csv, xml, log, and more. This data is ingested through streaming or batch processing and then processed and stored in a centralized, raw storage location before being transformed from its source format to the format required for analysis. Processes will vary, such as for streaming data, through which near real time insights can be unlocked, this data can be transformed in the streaming pipelines themselves, as opposed to being stored first being transformed. Finally, the data is ready to 'model and serve', ultimately powering data-driven applications and supplying rich data visualizations to users.
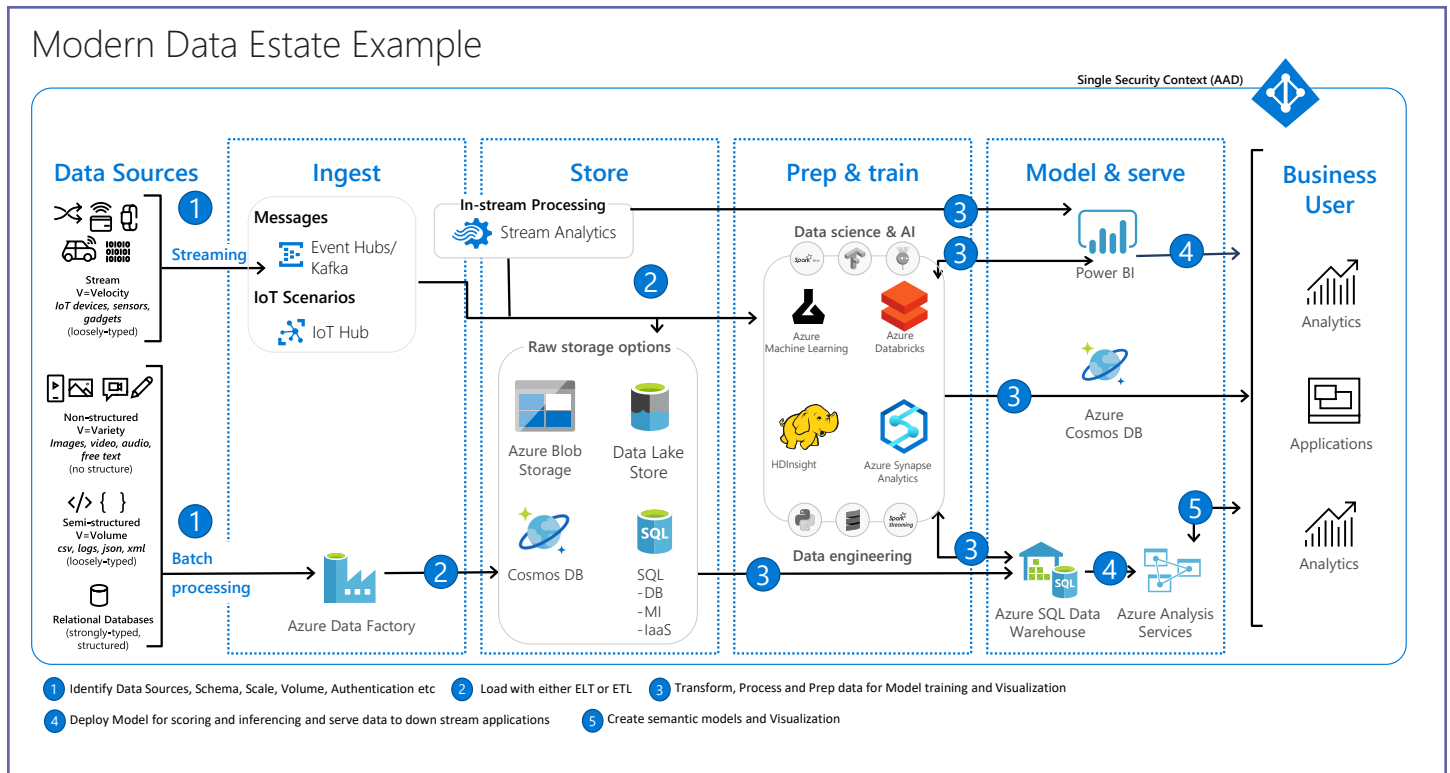

Figure III. Modern Data Estate Sample Scenario

Section III. Opportunity for Partners

A data-led approach to digital transformation presents two opportunities for partners: (a) increased surface area for professional and managed services and (b) new approaches to monetization for improved profitability.

Services Opportunity

As customers grow to understand the value of modernizing their data estate, the surface area for partners to deploy professional and managed services also expands. Not only is the partner able to build and sell net new services, but feedback from partners indicates that data estate modernization projects are "stickier", quicker to market, and have faster time to value realization compared to heavy enterprise line of business applications that may take years to modernize and move to the cloud.

Project services are an important point of entry for partners. However, supporting customers through the initial transformation of their data estate is only part of the opportunity for partners. Rather, managed services offer a clear path to long-term profitability through higher margins and increased customer lifetime value. Further, the current lack of managed services for data estate management means partners who move quickly to grow their portfolio will see first-mover advantages.

As partners develop their portfolio of offers, there are four categories of data managed services to consider: foundational services, data availability services, intelligence services, and experience services. The first, foundational services, are those that transcend the underlying data estate to maintain overall health and integrity while the remaining three categories build upon each other to enhance the estate. Data availability is the first of these building blocks, and helps to ensure customer data is the quality needed discoverable, and fit for the intended use case. Intelligence builds on data availability and helps the users derive deep and meaningful insights from their data. And finally, the third tier of data services offer users personalized, interactive, and automated experiences via advanced analytics and powerful BI reporting. Real-time data is consumed seamlessly across devices and mediums, empowering users to find hidden stories within their data and make smart, proactive decisions to move their business forward. Appendix II provides just a small sampling of relevant managed services across each of these categories.

[3] Processing methods will vary based on the type of data, use case and if using ELT or ETL process.

Microsoft

Monetization Opportunity

When leveraging an application-centric approach, most partners monetize their services by charging a certain percentage or tiered fixed fee on top of the customer's cloud consumption. Margins are healthy, often exceeding 50 percent for applications and infrastructure. A data-led approach neither replaces nor diminishes this existing revenue/margin. Rather, a data-led approach expands the surface area for customer engagement (as discussed above) while introducing new monetization opportunities akin to IP provider models. This means that rather than a percentage or tiered fixed fee approach, partners can monetize their data services using royalties or a fee-per-user/fee-per-instance licensing model.

Data managed services – especially data availability, intelligence, and experiences – are essentially IP and can be monetized as such. For example, a managed service such as Data Quality Services or a Merge Service can be executed in a customer's own subscription and licensed to customers for a fee or executed in a partners own subscription in Azure and accessed by customer users/publishers for an access fee. The first model provides customers with

total control over their own environment. If executed in the partner subscription, the customer will not control their own environment, but can tap into insights and learnings from a multitude of "like" use cases spanning various customers and industries. By adopting the architectural considerations of a Modern Data Estate outlined in the paper and accelerating the customer's journey to rich data intelligence through managed services, partners will be well-positioned to grow the number of subscriptions per customer. By investing in both models outlined above, partners will be well positioned to support customers with diverse needs and use cases.

Because this is a relatively novel monetization approach for managed service providers, they may need to retool their commerce and Cloud Management Platforms to effectively invoice under these models. However, the margin realization on these services is directly proportional to the value delivered to the customer and far exceeds that of infrastructure and application projects/services. For example, IP services that have been enriched with automation and machine learning can lead to margins exceeding 70 percent.

Microsoft

Conclusion

The pivot to becoming a data-driven organization is fundamental for businesses and thus represents an enormous opportunity for partners to become catalysts for digital innovation. To fully capture this opportunity, partners must rethink the way they engage customers on their cloud journey by shifting from an application only approach to an application and data-led approach. This means conducting early discovery efforts and assessing the data estate as a whole, apart from any singular application, to ensure repeatability and scalability across current and future customer use cases that impact business outcomes. For partners looking to adopt a data-led approach and thus capitalize on a proposition not yet offered at scale in the market, key tenants are:

I.   A data-led approach should be designed around two best practices: *continuous and iterative design* to prove value early and *repeatability and scalability* to catalyze value realization for customers and increase partner profitability.

II.   When defining success with the customer, always refer back to the Modern Data Estate litmus test (i.e. whether the investment delivers on the following core tenants: *reduced costs, improved performance, reduced risk,* and *advanced analytics).*

III.   To maximize customer value while also expanding partner margins, build and attach managed service offers across the four data services catagories: *foundational services, data availability services, intelligence services,* and *experience services.*

These tenets offer a roadmap for partners looking to become the "go to" trusted advisor for navigating customers through their data journey. In so doing, they become the linchpin for companies in a time where data may very well determine the winners and losers of the fourth industrial age.

Data managed services fall into one of four categories: foundational, data availability, intelligence, and experiences.

Microsoft
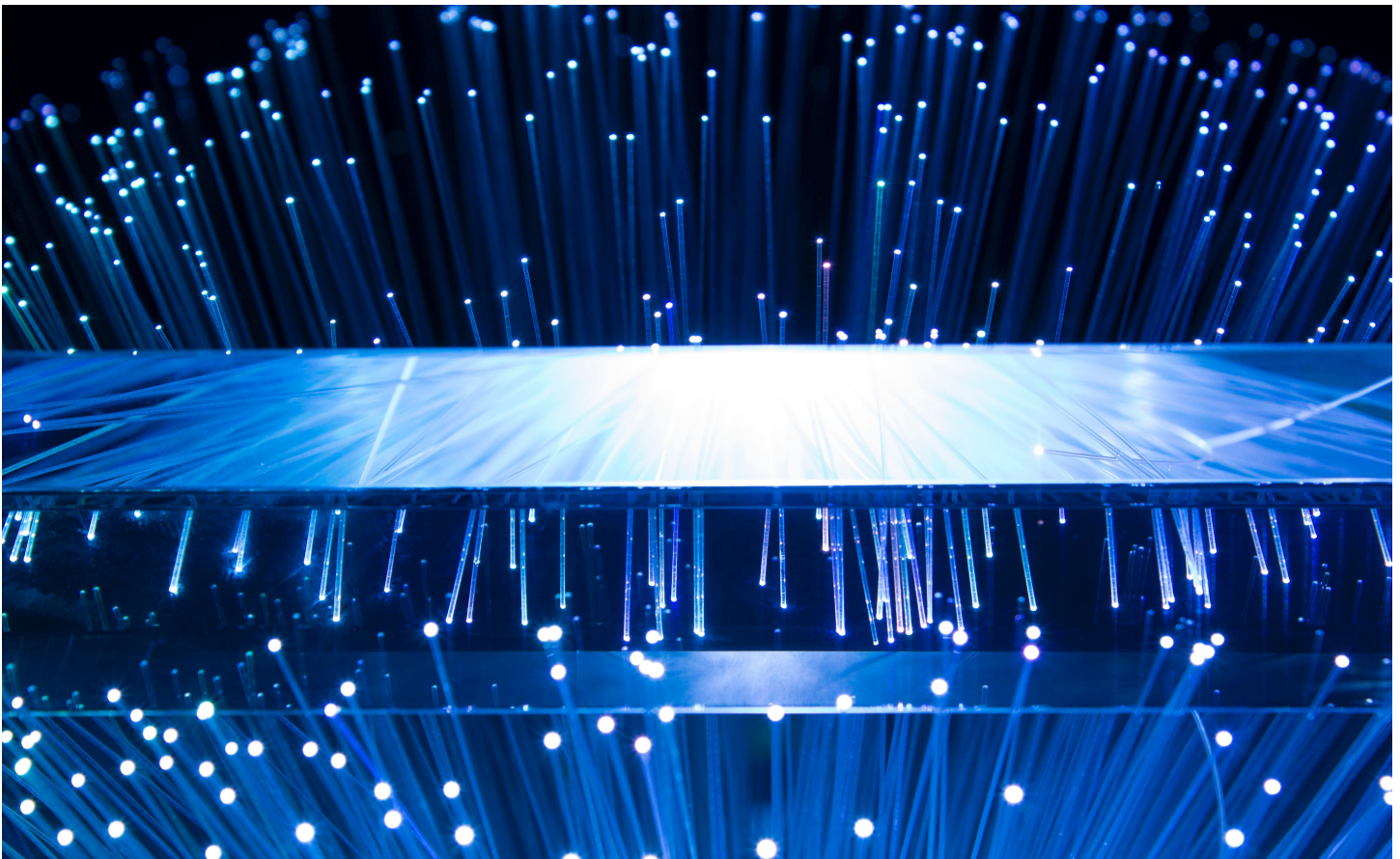
Appendix I: Visualizing the Modern Data Estate
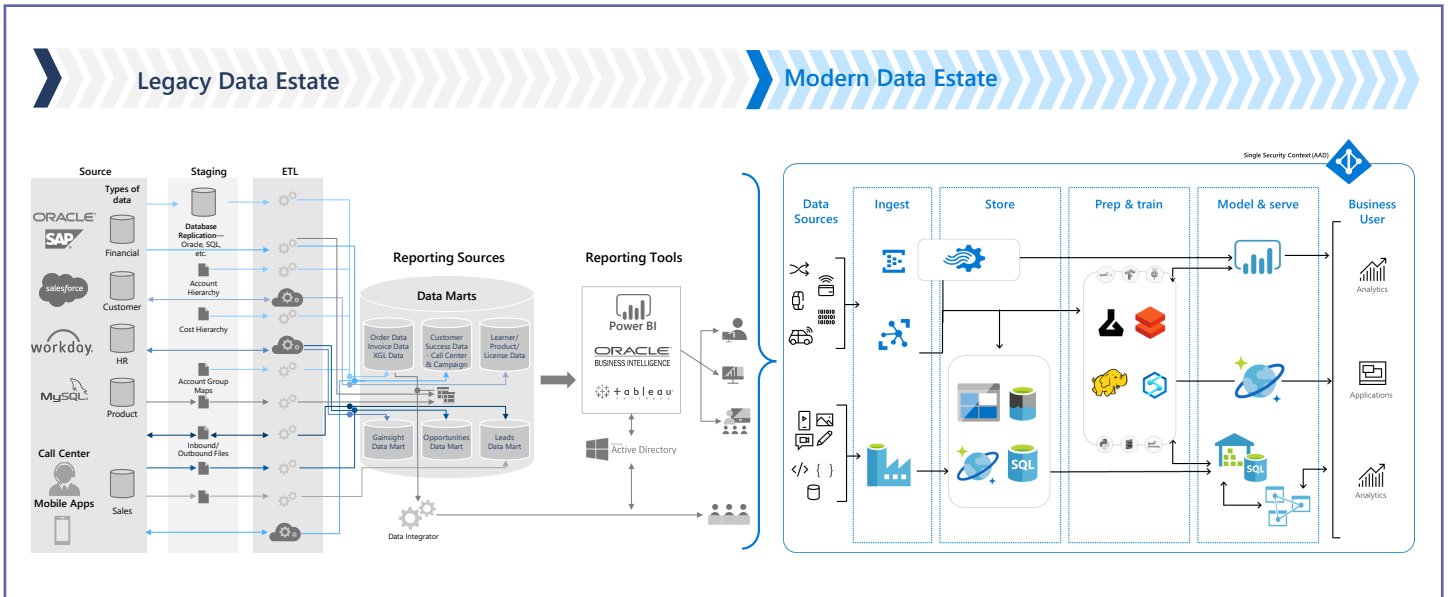
Appendix II: Managed Services Categories

Foundational

Underlying services to maintain the health and integrity of the estate, including:

I.   *Governance* – Set standards and policies that define data availability, usability, integrity, and security.
II. *Security* – Minimize risk to business operations and protect data at both data lake and consumer layers.
III. *Lineage* – Track data's origins and movement over time.
IV. *Alert/Monitoring* – Monitor the implementation of each data lake to ensure compatibility with standards.
V.   *Data Catalog* – Make data easy to find by automatically populating to a central catalogue.
VI. *Common Data Model Service* – Provide a shared data language for business and analytics applications.

Data Availability

Services to ensure customer data operates at the highest level of quality and efficiency in line with best practices.

I.   *Data Lake as a Service* – Ensure lake consistently conforms to standards such as schema, data availability schedule, type of file format available (csv, parquet, json), and access protocols (REST, NFS, HDFS).
I.   *Ingest for Compute Service* – Support publication to data lake across multiple fabrics.
II. *Handshaking Service* – Understand what data is published to the data lake and when.
III. *Control File Service* - Enables users' 'compute' to read data files.
IV. *Data Catalog Service* – Enables discoverability of what data is available in the data lake.
V.   *Data Quality Service* - Ensure data is fit for intended usage in operations, decision making, and planning.

Microsoft

Intelligence

Services to support user ability to derive deep and meaningful business intelligence.

I. *Merge Service* – Merge data based on user(s) configuration.
II. *Reconciliation Service (Recon)* – Compare transformed data against original source data to reduce duplication. and avoid risk of differing results across users leveraging the same data set.
III. *AI/ML Models* – Provide insights to users based on most commonly needed information and queries.

Experiences

Services to provide users personalized, interactive, and automated experiences.

I. *AI/ML* – Leverage automation to streamline processes and improve data driven decision making.
II. *PowerBI as a Service* – Make customer data actionable through visual analytics.
III. *Azure Virtual Assistant* – Deliver a conversational assistant personalized to the user.

Special Thanks To Our Reviewers

Moritz Berger, Anil Dwarakanath, Andy Gore, Phil Harvey, Barnaby Jeans, Robert Venable

Microsoft