**5 Megatrends That Will Shift the Data and Analytics Market**

Over the past year, as I interacted with enterprise customers of all sizes, I have seen five unmistakable megatrends shifting the data and analytics market. We have confirmed these trends by examining the perspectives of investors, research firms, and industry practitioners.

**Trend 1: From Data Center/Cloud to Distributed Cloud**

The shift from data center deployments to the cloud continues apace.

“New data from [Synergy Research Group](https://www.srgresearch.com/articles/2020-the-year-that-cloud-service-revenues-finally-dwarfed-enterprise-spending-on-data-centers) shows that enterprise spending on cloud infrastructure services continued to ramp up aggressively in 2020, growing by **35%** to reach almost $130 billion.

In 2020 COVID-19 helped to further fuel this major shift in worldwide IT operations. Over the decade, the average annual spending growth for a data center was just 2% and for cloud services (IaaS, PaaS and hosted private cloud) was 52%.”

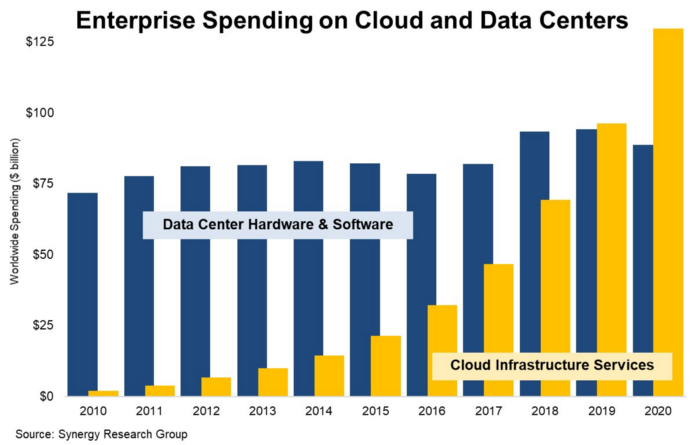


Figure 1. Enterprise Spending on Cloud and Data Centers

On the other hand, the regulation of some countries demands data to reside in a physical data center. This fact has pushed the public cloud services to be available in different physical locations, resulting in various deployment models, such as private cloud, public cloud, and multi-cloud.

**Trend 2: Great Convergence of Analytics and AI/ML**

Large enterprises and tech companies **with sophisticated, complex data needs** usually require their data infrastructure to serve two purposes:

1. help internal users make better decisions through the use of data;

2. build operational AI/ML systems to support data intelligence into customer-facing applications.

Historically, companies had data lakes on one side serving AI use cases and then data warehouses on the other side supporting transactional analytics and business intelligence. In reality, many organizations suffer from the duplicated data pipelines, complex infrastructures, and high costs coming along with this kind of data stack. **As a result, companies are now trying to merge both sides, creating a unified experience for all types of data analytics, including BI and ML.**

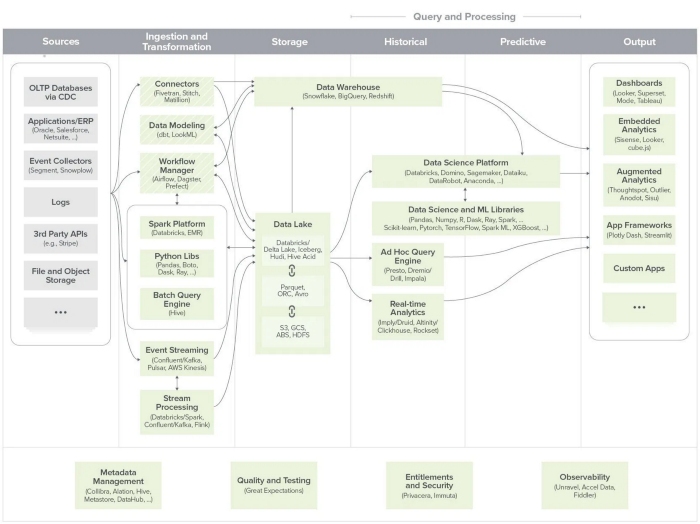


Figure 2.The Emerging Architecture for Modern Data Infrastructure — [Source](https://a16z.com/2020/10/15/the-emerging-architectures-for-modern-data-infrastructure/)

According to [A16z](https://a16z.com/), in an article called “[The Emerging Architecture for Modern Data Infrastructure](https://future.a16z.com/emerging-architectures-modern-data-infrastructure/),” some experts believe data warehouses and data lakes are on a path toward convergence to simplify the technology and vendor landscape for both use cases.

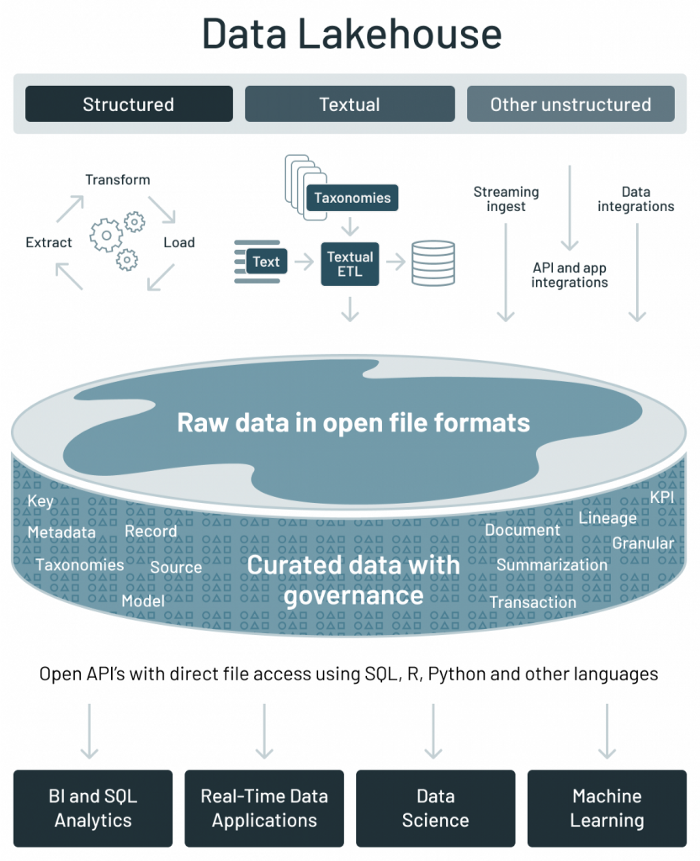


Figure 3. Data Lakehouse Diagram — [Source](https://databricks.com/blog/2021/05/19/evolution-to-the-data-lakehouse.html)

Furthermore, with the debut of the [lakehouse](https://databricks.com/blog/2020/01/30/what-is-a-data-lakehouse.html) concept, Databricks also believes that the data infrastructure of the future belongs to the one that can support both BI and ML use cases.

**Trend 3: From IT-Centric Data Workflows to Self-Service Analytics**

Diverging from the on-demand manual delivery method that relied on IT in the past, business users will now meet their analysis requirements through data assets provided by a self-served analytics platform. This trend echoes the [citizen data scientist](https://blogs.gartner.com/carlie-idoine/2018/05/13/citizen-data-scientists-and-why-they-matter/) concept raised by Gartner.

A self-service analytics platform can satisfy 80% of the needs of business users based on the existing datasets. In comparison, the remaining 20% of new requirements can be quickly defined based on the current data assets. Behind the scenes, the datasets are still **well governed by IT** to prevent wrong usage and high cost. Therefore, this new workflow of delivering data is well optimized with the right balance between efficiency and cost.

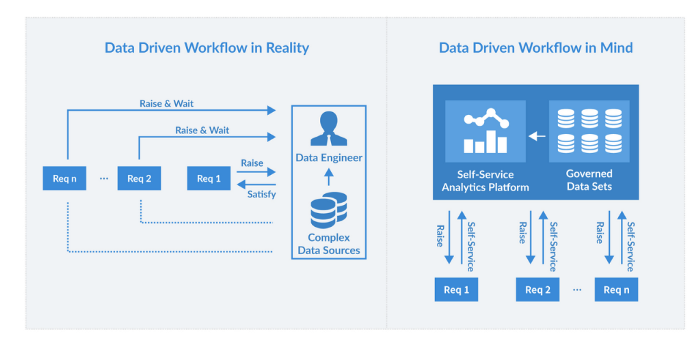


Figure 4. Data-Driven Workflows

**Trend 4: Manual Analytics & Insight to Augmented Analytics**

The traditional approach uses visual-based data discovery to enable business analysts and business users to uncover essential insights. However, when the data is complex, large, and highly dimensional, users must either focus on exploring their hypothesis or manually explore all possible combinations of their data to get an accurate conclusion. This approach is very time-consuming and prone to bias.

Gartner believes that Augmented analytics, a new paradigm of data and analytics, has emerged. Augmented analytics uses machine-learning-based automation to assist human users in uncovering insights and providing direct actions from data. The beauty of Augmented analytics is that it applies a range of algorithms to the data in parallel and reduces **the risk of missing essential insights in data.**

**Trend 5: Convergence of Delayed Insights and Real-time Analytics**

In the past, the refresh of the data always comes in batches. However, more timeliness of **business decision-making** has called for a lower latency of data. The ability to take immediate action at the first indication of an incident is a crucial factor in the success of business operations, especially in use cases such as fraud detection and campaign monitoring.

Consumers are also demanding a more **engaging and seamless user experience** that requires applications and businesses to provide response and feedback at the speed of thought. When it comes to business decision-making, there is no way to identify problems and anomalies within the current situation without knowledge of what has happened in the past. This need has driven the trend of convergence of batch data and real-time data.

In summary, the end goal of this overall trend is the democratization of data analytics to empower all levels of a business to get faster, more timely and better-informed insights.