IS/DM Part 1

1. Industry Overview (Aung)

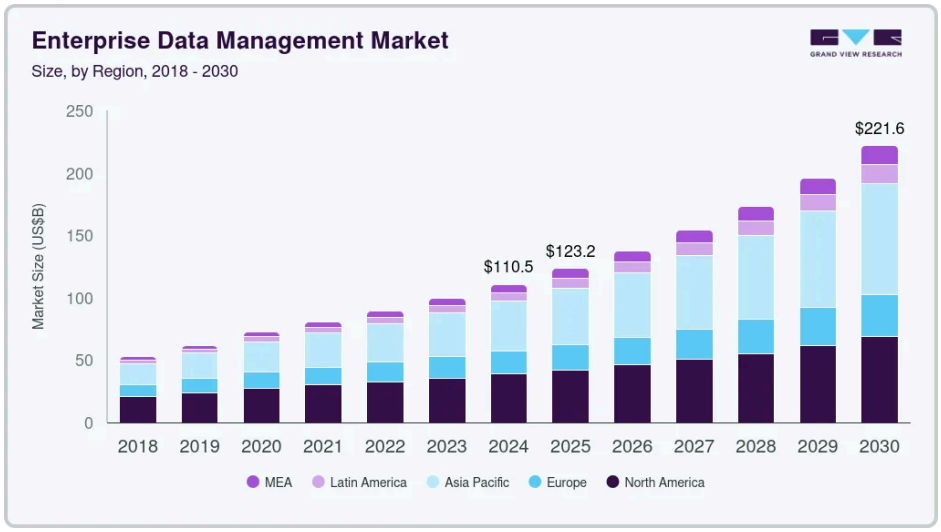
1. **What is Information System/Data Management (IS/DM)?**

   A set of processes that involves using systems to collect, store, and organize data and use these data and information to strategically support business operations, decision making, and management.

   - How does IS/DM collect, store, and organize data?
     - Stores structured (tables), unstructured (text, images, video), and semi-structured data in databases, data warehouses, data lakes, file systems, etc…
     - Once data is stored, the system organizes the data (through steps like modelling, integration, and transformation) to make the data accessible and meaningful for analysis and decision making.
   - Why does this matter?
     - With the significant increase in the volume and variety of data, organizations that manage data properly will be in a stronger position.
2. **What are some related sectors**
   a. Data Analytics/Business Intelligence (BI)
      i. Difference: assumes the data is already there. IS/DM is more foundational
   b. Data/Information Integration
      i. Subset of IS/DM, focus on combining data from multiple sources into a unified, usable form
3. **IS/DM Companies**
   a. Snowflake
      i. Offers a cloud-native platform for storing, processing, sharing, and analyzing structured, semi-structured, and unstructured data across multiple clouds.
   b. Databricks
      i. Provides an AI platform that supports data engineering, streaming, analytics, and gen-AI workloads in just one system.
   c. Informatica
      i. Delivers an "Intelligent Data Cloud Management" that includes data integration, data quality/observability, master data management (MDM), and governance across hybrid/multi-cloud environments.
4. **Current Market Size & Major Players**

   Figure 1

   Enterprise Data Management Market (2018 - 2030)



Source: (Grand View Research)

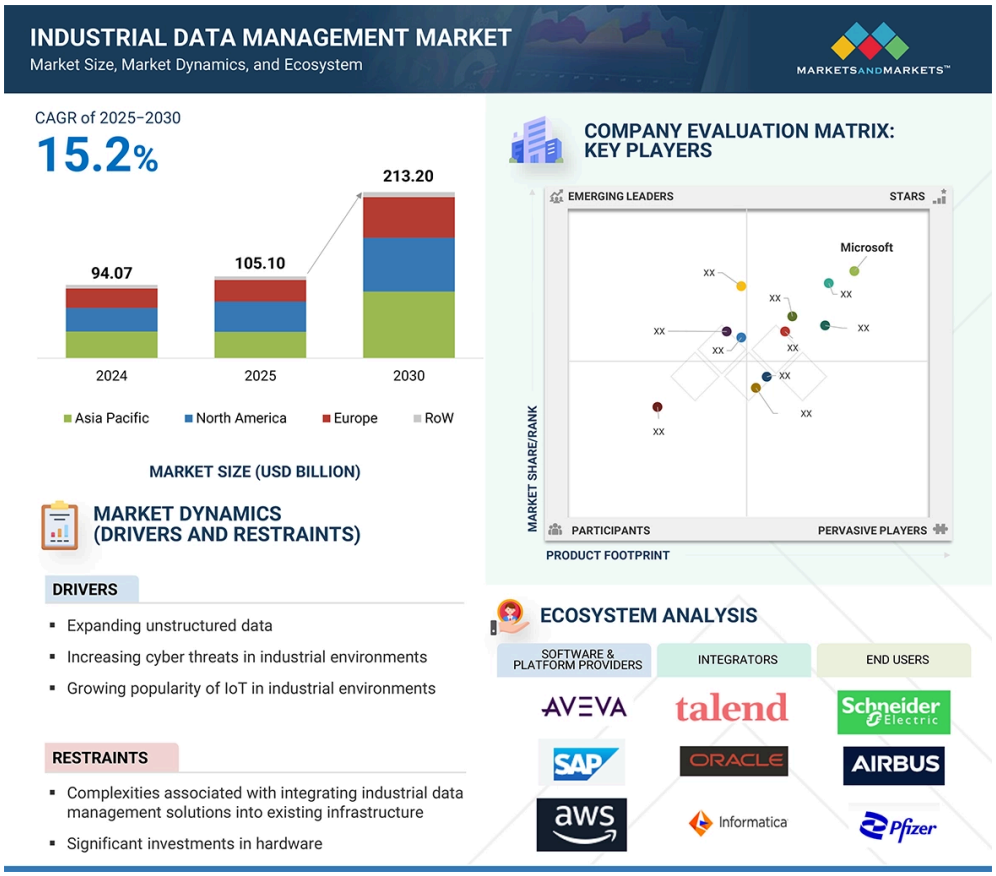
a. As shown in Figure 1, in 2024, the global enterprise data management market size was estimated at USD 110.53 billion.
    i. Anticipated to reach USD 221.58 billion by 2030, growing at a CAGR of 12.4%.
b. Major players in this industry include IBM, Oracle, SAP SE, Informatica, etc…

2. Market Growth & CAGR (Rohan)

Globally, the Information Systems/Data Management (IS/DM) sector is experiencing explosive growth due to increasing digital transformation, cloud migration and exponential enterprise data growth. Grand View Research (2024) using the title Enterprise Data Management Market: The global enterprise data management market size is anticipated to reach USD 221.58 billion by 2030, rising from USD 110.53 billion in 2024, according to a new report by Grand View Research, Inc., expanding at a CAGR of 12.4% over the forecast period seems appropriate. In the meantime, IMARC Group (2024) forecasts an only slightly lower development with the market value projected to increase from USD 85.5 billion in 2024 to USD 200.6 billion by 2033, at a CAGR of approximately 9.45%. Within this broader ecosystem, some sub-sectors are growing even faster. The Master Data Management (MDM) market, for instance, is projected to grow from $19.9 billion in 2023 to $60.7 billion by 2030, with a strong 17.4% CAGR, while the Data Management Platform (DMP) market is anticipated to expand at 13.5% CAGR through 2035.

Figure 2

Industrial Data Management Market projected to reach USD 213.2 billion by 2030, growing at a CAGR of 15.2% (2025–2030)



Source: MarketsandMarkets

While this GRAPH is focused on the Data Management segment, it represents similar underlying technologies and growth propellants that are layering, with increasing sophistication, across the IS/DM market - cloud penetration, IoT interplay and increasing demand to manage unstructured data at scale.

The aforementioned momentum is driven by the interplay of structural and technological factors:

- With the flood of both structured and unstructured data pouring into organizations today, companies need scalable IS/DM systems to house, tame and obtain value from their dataspace.
- The evolution of local to-the-cloud and multi-cloud storage solutions has successfully allowed businesses to efficiently manage massive, diversified datasets while enhancing access, security, and cost savings.
- The incorporation of AI, ML and automation tech: AI and ML is infiltrating data management workflows to automate cleansing, integration and analytics - by speeding up insight generation and predictive capabilities that directly contribute to strategic decision making.
- There is a growing shift in focus towards data privacy, governance and compliance. With the arrival of global data protection legislations (LIKE CCPA or GDPR), organizations have had to substantially re-invest in IS/DM solutions, which guarantee good stewardship, lineage trackability and audit trail.
- Cross-enterprise digital transformation projects: Industries like finance, healthcare, manufacturing, and retail are all investing in real-time analytics, business intelligence and system integration which drives IS/DM demand.
- The arrival of new storage and integration technologies: The emergence of data lakes, data lakehouses, and distributed databases (Snowflake, Databricks) has changed how businesses handle complex data at scale and unify analysis.

Conclusively, combined with these trends, IS/DM has risen to emerge as a core component in the current enterprise application stack. The ongoing momentum in the market reflects not only the growing strategic importance of data as a corporate asset but also the necessity for an end-to-end solution to ensure compliance, innovation, and competitive advantage. Not only do analysts anticipate double-digit growth, but they anticipate such growth through the next 10 years as companies like IBM, Oracle, SAP, etc. battle it out for dominance in global enterprise data management.

3. Key Trends (Atharv)

The Information Systems and Data Management (IS/DM) industry is at the center of a digital revolution driven by artificial intelligence, modernizations, and the massive demand for trustworthy data. The following five trends capture how technology, regulation, and user behavior are collectively reshaping the industry's direction.

Trend 1: AI's Integration into Everyday Life

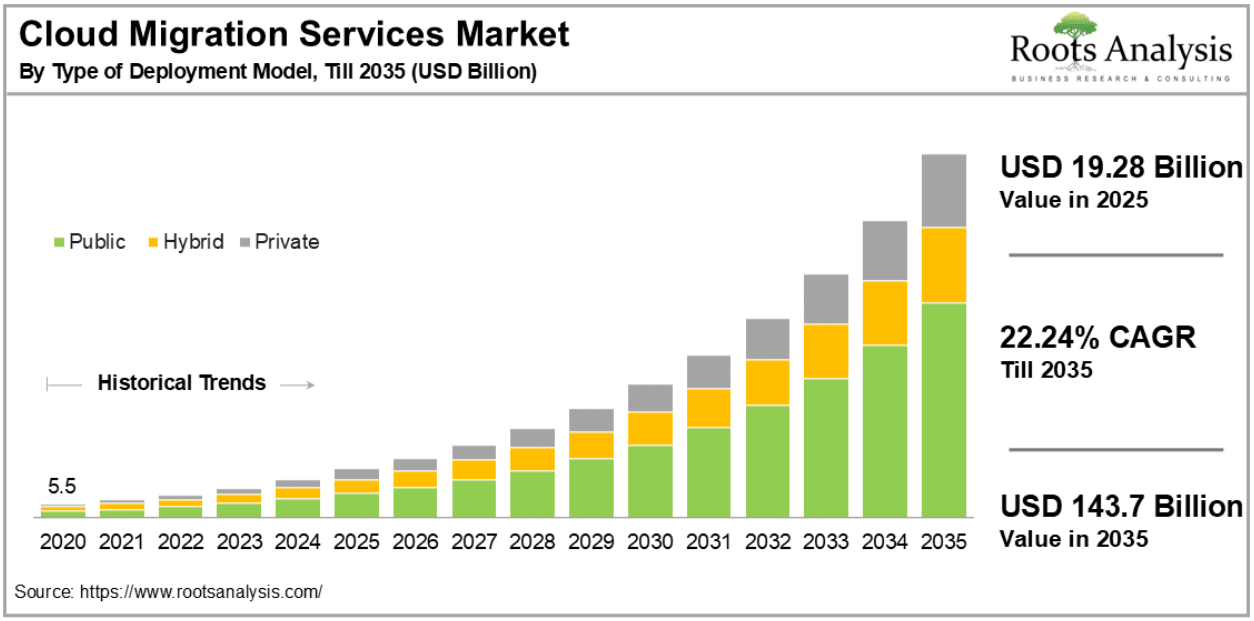
Artificial intelligence has moved far beyond research labs; it is now woven into the daily routines of billions of people. From AI copilots drafting emails and generating content, to recommendation engines shaping what we buy and watch, to autonomous driving systems interpreting real-world environments, AI is transforming from a piece of technology into public utility, and in time, as foundational as electricity or the internet. This surge in everyday AI usage forces a massive transformation in IS/DM. To keep these systems accurate and reliable, companies must train and manage enormous, high-quality data sets. Firms like OpenAI, Anthropic, and Google have effectively become data companies as much as AI firms, because data management is now the key determinant of AI capability.

Short-term impact: Explosive demand for AI-ready infrastructure and synthetic-data pipelines to support training at scale.
Long-term impact: Data management and AI infrastructure fully merge — every modern IS/DM system will natively support model training, validation, and governance.

Trend 2: Legacy Modernization and the Cloud-First Transformation

As companies phase out outdated systems, they are embracing cloud-native, modular platforms capable of handling real-time analytics, scalability, and integration — a process known as legacy modernization. This modernization wave underpins the ongoing global digital transformation, driving organizations toward greater efficiency, flexibility, and innovation.

Figure 3: Cloud Migration Services Market by Deployment Model (2020–2035)



Source: Roots Analysis (2024). Forecasted CAGR 22.24%, USD 19.28 B in 2025 → USD 143.7 B by 2035.

Cloud-first IS/DM solutions such as Snowflake, Databricks, and Informatica have become the operational backbone of enterprises, enabling seamless data flow across hybrid and multi-cloud environments. The scale of this transition is striking: as shown in Figure 4, the Cloud Migration Services Market is projected to grow from USD 19.28 billion in 2025 to USD 143.7 billion by 2035, reflecting an extraordinary 22.24 percent compound annual growth rate (CAGR). This surge demonstrates how modernization has evolved into one of the defining growth engines of the global IS/DM industry.

Short-term impact: Heavy investment in infrastructure migration, workforce retraining, and system integration as enterprises transition to cloud environments.

Long-term impact: Consolidation around intelligent, cloud-native IS/DM platforms that are secure, compliant, and designed for continuous AI integration.

Trend 3: Rising Data Governance and the Era of Digital Accountability

At the cusp of massive digital transformation, data privacy has shifted from being a choice to a crucial pillar of the IS/DM industry. With expanding global regulations, companies are being forced to redesign their data frameworks around transparency, consent and traceability.

These global regulations include the General Data Protection Regulation (GDPR) in Europe, which grants users full control over how their personal data is collected and processed; the

California Consumer Privacy Act (CGPA), which allows consumers to opt out of data sales and demand disclosure of data usage; and new U.S. state laws such as those in Maryland, Minnesota, and Nebraska, which extend these principles by requiring companies maintain stricter data security. Together, these laws mark the rise of a “privacy-by-design” era: where compliance, auditing, and ethical data handling are built directly into the infrastructure rather than treated as add-ons.

Short-term impact: Increased spending on compliance infrastructure, data lineage systems, and region-specific storage setups.

Long-term impact: Development of globally standardized governance frameworks and the emergence of trust-centric data platforms that treat transparency as a competitive differentiator.

4. Headwinds (Evan)

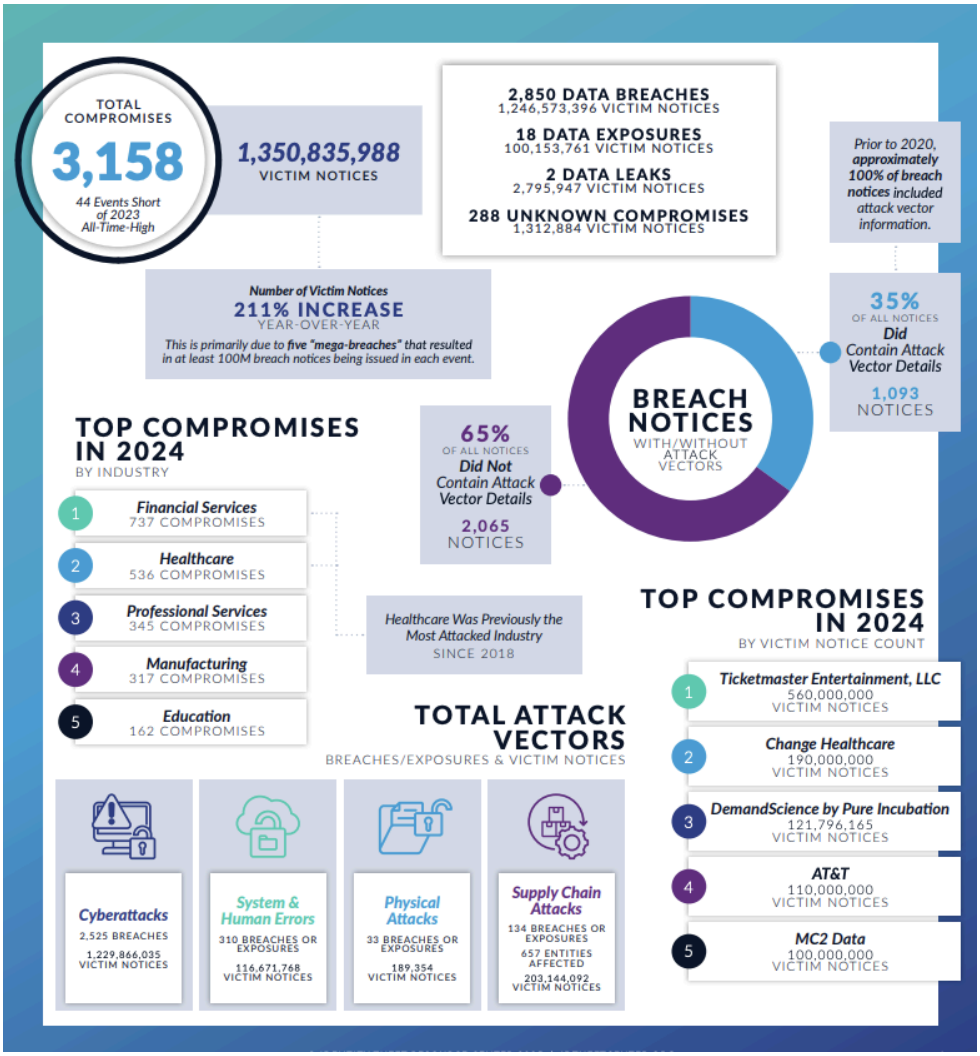
Increase in Data Compromises Quality

The more data an organization collects, the harder it is to ensure that the data is accurate, consistent, and usable. For instance, if you're collecting billions of records (IoT devices, user logs, sensors, etc.) even a tiny error rate (say 0.1%) can create millions of bad entries. Additionally, differences in how devices record or transmit data (such as units, timestamps, or calibration errors) can further reduce reliability of data. Thus, without rigorous and strong data governance, the massive influx of data can overwhelm systems, which results in inaccurate or unusable datasets. The loss in quality of data sets is bad because as data quality deteriorates, firms will make flawed decisions based on inaccurate/misleading datasets. This could lead to incredible financial losses or stunted efficiency. (Source 1, Source 2, Source 3).

Security Exposure and Breach Costs

As the IS/DM sector grows, cybercriminals have even more incentive to target data. They breach systems to commit fraud, demand ransoms, or even commit espionage. Attackers are also getting more sophisticated: instead of relying on simple encryption, many now use multi-pronged extortion and data theft, stealing sensitive information and threatening to leak it even if a company can restore from backups. While the global average breach cost is roughly **$4.44M**, United States breaches now average **$10.22M.** Additionally, **97%** of organizations have reported an AI-related security incident and lacked proper AI access controls. Additionally, **49%** of victims paid the ransom to get their data back. Additionally, this problem is only growing. The Identity Theft Resource Center counted **3,158** U.S. compromises in 2024 (second-highest year on record) as well as **1732** breaches in just the first half of 2025, which is slightly ahead of 2024's pace.(Source 1, Source 2, Source 3, Source 4).

Figure 4



Source: [Identity Theft Resource Center]

a. As shown in Figure 3, there were 3158 compromises in 2024, the second-highest ever in history.
b. Victim Notices (notifs. Company sends to people whose data may have been exposed to a breach) up 211% from the previous year
c. Financial services industry most affected.
d. Big-name companies like AT&T and Ticketmaster Entertainment were top compromises in 2024.

Regulatory Fragmentation and Data Residency

Privacy and data rules vary widely by region, so a company may need one configuration for Europe, where the GDPR (information regulation law) governs how personal data is collected, used, and shared, and a different setup for other jurisdictions with their own requirements. The

EU's GDPR also carries harsh penalties (up to 20 million Euros) for violations. For the IS/DM industry, this turns into a lot more work: keeping detailed data maps and tags(who's data, where it lives, why you have it), building region-specific pipelines and storage, and enforcing consent and retention rules inside data platforms, encrypting and limiting access by role, auditing activity, and managing vendors and cross-border transfer. Essentially, all this extra work raises costs, causes slow rollouts, and limits how teams can combine datasets or train AI models. (Source 1, Source 2, Source 3)

5. Tailwinds (Mason)

Huge Growth in Data Creation

Every day, individuals and companies alike create massive amounts of data. Because there's a plethora of information, businesses need better ways to store, organize, and understand it. Global data creation is expected to hit 181 zettabytes in 2025, which is more than 1.45 trillion gigabytes created daily, with loT devices generating about half of the new data. Beyond data storage, there has been a massive push for the monetization of data. Companies like Amazon and Netflix have been selling behavioral and mobility data to third-party analytics, fueling the demand for growth and management within the industry. For example, in Netflix/Amazon/Uber all are selling anonymized data about user behavior or travel patterns to other firms, which are cutting huge checks and reaping lots of profit.

Use of Artificial Intelligence (AI) and Machine Learning (ML)

AI and ML assist computers in the analysis of faster data and the ability to find useful trends in predicting customer behavior or improving company efficiency to optimize operations. This, in turn, makes companies want to invest more in better systems that target consumers more effectively. AI and ML are helping to pave a new path for pattern recognition amongst users and fraud detection/ predictive analytics. These analytics provide structured data and optimize performance in advanced systems to track analytics.

Focus on Data Privacy and Security

New laws are requiring companies to protect people's information in an age of digital identity theft. In turn, this pushes corporations to incorporate better security systems that track and secure user and company data. For example, in Nebraska, Minnesota, and Maryland, several new U.S. privacy laws have been put into place in order to protect user data. This is pushing companies like Google, Meta, and Apple to continually upgrade their platforms in compliance with new laws and regulations

1. Children's Online Privacy Protection Act (COPPA) Amendments (Effective June 2025): The FTC expanded requirements for operators collecting data from children under 13, giving parents more control over how children's data is used and shared. Organizations must comply by April 2026.
2. Executive Order 14117 and Data Security Program (April 2025): This program restricts access by foreign countries of concern to Americans' bulk sensitive personal data (e.g., financial, health, biometric, geolocation, and government-related data).

3. Protecting Americans Data from Foreign Adversaries Act (PADFAA): This new federal law restricts data brokers from transferring sensitive personal information to foreign adversaries.

Faster Computing

Companies in areas such as finance and healthcare now need to access information instantly to stay competitive. Quicker decision-making, as well as optimization of digitized tasks, is needed for these companies, so there is a great need in the startup space for these new IS companies. As companies require faster and more reliable data pipelines, IS/DM solutions help optimize these processes. By the end of 2025, 75% of enterprise-generated data will be processed outside of data centers, as projected by Gartner. IS and DM companies, therefore, have been optimizing their data pipelines, which is extremely important for growth within the industry.

Increased Demand for Business Intelligence

Leaders and managers across the industry want to ensure smarter choices based on customer analytics. Tools are becoming more accessible to help turn raw data into charts and insights, making it easier to plan and compete. Therefore, the demand for Business Intelligence is increasing, and IS/DM products and services are being forced into combination with current workflows..

New improved ways to store and use data

New technologies such as data lakes and distributed databases make it cheaper and easier for companies to store and analyze huge amounts of data. This helps businesses take advantage of and utilize advanced data tools in order to save on data usage. Data lakes and cloud databases are two pieces reshaping how the industry manages massive amounts of data. They are software architectures that incur a lower cost, improving cross-domain analytics, and creating central building blocks for firms and enterprises alike. Adoption rate for these types of companies also continues to climb as businesses are prioritizing the integration of AI databases. One example of this is how data lake solutions is holding over 60% which reflects a desire in the market for integrated platforms, analytics, and storage.

atlan

https://atlan.com/data-lake-use-cases/

9 Essential Data Lake Use Cases You Must Know - Atlan

Explore data lake use cases from advanced analytics to IoT. Dive into their benefits, challenges, and real-world applications. Stay informed.

Toward data lakes as central building blocks for data management ...

Pmc.ncbi.nlm.nih

2025 State Privacy Laws: What Businesses Need to Know for ...
whitecase
150 Data Privacy Statistics For 2025 You Need To Know About
Usercentrics
How AI/ML is Revolutionizing Data Management and Integration
Leadingedgeonly
How Much Data Is Created Every Day: Key Insights & Trends
designrush
Amount of Data Created Daily (2025) - Exploding Topics
Explodingtopics
The power of distributed data management for edge computing ...
datasciencecentral

IS/DM Part 2

Product Overview (2 People) - Evan, Aung

The main product that is sold by IS/DM companies is a cloud-based set of tools known as the Modern Data Stack. The Modern Data Stack is a set of products that allows organizations to collect, store, organize, secure, and analyze data. Essentially, it acts as a highly-automated, digital database that enables companies to use messy information to make calculated decisions. The Modern Data Stack can be separated into four main categories: data storage and processing, data integration, governance and security, and analytics or AI platforms.

Data storage and processing platforms are primary repositories for enterprise data. Basically, these include modern data warehouses such as Snowflake, Databricks, Google BigQuery, Amazon Redshift, and Azure Synapse. The main function of these is to store massive amounts of data (think of these like giant digital warehouses where all the company info is stored). Essentially, every time a business collects data, it needs somewhere to keep it safe and easy to find. Data storage platforms are usually built on the cloud, which means that they don't live on physical computers in one building. Rather, they exist on the internet and they can grow or shrink depending on how much data the company has. Additionally, data storage platforms can quickly organize data into different shapes and formats, allowing them to be accessed efficiently.

Next is data integration and pipeline tools. These are essentially like delivery trucks and sorting machines. Data does not just appear in one place. For instance, data can come from apps, websites, stores, etc. Data integration and pipeline tools collect all the data that comes from these places, cleans it up so that it is consistent, and then moves it to the data storage platforms where it can be stored. Some tools can make sure that this process happens regularly (every hour or even every few seconds) to make sure that information that companies use is always up to data and relevant.

Next comes the data governance and security tools. These are essentially tools that make sure that all company data remains safe, organized, and trustworthy. As an analogy to help understand this concept better, you can think of data governance and security tools as librarians who label every book, record who borrowed it, and make sure that nobody reads stuff they're not allowed to. In the same way, data governance tells companies where their data came from, who is allowed to use it, and how it is being used. It also checks for mistakes or missing pieces to data. Additionally, data governance entails making sure companies follow privacy laws that will protect customer information. If you have ever filled out an online form and checked a box that says that a company can store your data safely, data governance is the kind of software that helps companies keep that promise. It helps to build trust and reliability for companies so that people are willing to let companies use their data.

Finally is the analytics/AI platforms. These turn all the messy data that lies in data storage into readable and useful models in which people can understand. For instance, raw data that lies in the data storage can come in the form of millions of numbers, records, and transactions that don’t really mean much on their own. Analytics and AI platforms can turn this info into graphs and charts, dashboards, etc. These models can help companies analyze trends better, enabling them to make predictions and other important decisions about the company.

Ecosystem & Supply Chain (2 People) - Atharv

The Information Systems and Data Management (IS/DM) industry operates as a deeply interconnected system that transforms raw data into meaningful insights for decision-making. At its core, the industry depends on a complex hierarchy of technology layers, infrastructure providers, and strategic partnerships that allow information to move seamlessly from collection to application.

The foundation of the IS/DM industry lies in its physical and digital infrastructure. Cloud providers such as Amazon Web Services, Microsoft Azure, and Google Cloud deliver the computing power, storage, and scalability that make modern data management possible. Beneath them, hardware manufacturers and data center operators like NVIDIA, AMD, and Equinix supply the chips, servers, and facilities that physically host and process data. Without these unseen components, the massive data operations behind everyday digital systems—whether enterprise software, AI applications, or analytics platforms—would not exist.

On top of this infrastructure sits the data management layer. This includes systems that store, organize, and process information, such as databases, cloud data warehouses, and distributed computing platforms. Tools like Snowflake, Databricks, and Apache Hadoop clean, structure, and integrate data coming from multiple sources—transactions, sensors, user behavior, or enterprise systems—so it can be used effectively. Once data is processed, machine learning models and analytical engines turn it into predictions, visualizations, and strategic insights. Business intelligence platforms such as Tableau or Power BI then help organizations interpret these insights and make data-driven decisions.

The industry's operation also relies heavily on partnerships and integration. System integrators like Accenture, Deloitte, and Infosys connect technologies across different vendors and ensure they align with each client's business goals. Managed service providers maintain ongoing operations, monitor performance, and ensure compliance with privacy and security regulations. Independent software vendors create specialized tools that extend the capabilities of larger systems, while data brokers and API marketplaces provide external data that enhances analysis and forecasting. This interconnected ecosystem allows each participant to focus on a specific layer of value creation while depending on others for complementary capabilities.

What sustains the IS/DM industry is the balance between technological innovation and collaboration. Data exchange is built on shared standards, secure protocols, and regulatory compliance frameworks such as GDPR, HIPAA, and SOC-2. Open-source foundations also play a key role by maintaining common tools and promoting interoperability, preventing vendor lock-in and encouraging innovation.

Overall, the IS/DM industry is a collaborative web of hardware, software, and human expertise. From raw data generation to actionable insight, every layer depends on another—cloud infrastructure supports data platforms, which in turn support analytics tools and business applications. Together, these components form a continuous feedback loop in which data drives decisions, those decisions create new data, and the cycle repeats—constantly refining how organizations understand and interact with the world.

Customers (1 person) - Mason

IS and DM tools have been becoming more and more valuable and popular amongst professionals across industries. These tools help provide fast and easy automation, enable better decision-making based on stored data, and provide an increased sense of precision across complex interconnected industrial environments. These tools are used by almost all corners of industry, from creating genetic workflows, to security, to helping people across the board adapt to the new age style of work.

In the finance industry these tools are utilized greatly through looking for trends in markets, for calculating risk on certain investment opportunities, and for helping to automate vast amounts of tedious work. Recently in sales and trading and investment banking algorithms from information systems have been developing to replace tedious transaction reporting, detecting fraud, and creating faster turn around for data-informed decision-making (Koufax and Fluix are examples of this). Applications like Bloomberg terminal and pitch book are helping these individuals find high quality information as well. Essentially the financial industry is going through a huge revamp with the intergration of new data analytics tools.

Healthcare has been evolving rapidly with the same types of analytical tools being implemented to keep patient data longterm and secure. This has been integrated into tools like the EHR(Electronic health records), through making check-ins at hospitals and clinics quicker and more efficient, and through utilizing AI to search through medical claims to identify fraudulent medical claims.

The retail industry has been evolving through utilizing IS and DM tools too.

The retail industry has been evolving with the use of IS and DM tools, especially to keep track of inventory, track trends in products being sold, and figure out how to market/personalize products to different individuals through the use of tools like Bi tools and Stitch Fix. These tools calculate the trends in user data and can create meaningful predictions and marketing tactics to optimize and efficiently move products.

Manufacturing has been an industry that has also been equipping the use of such tools but not as mainstream as of yet. The tools being used so far are maintenance predictions, making sure machinery is running efficiently, and leading to improved interdisciplinary planning, which has been accelerated by these different tools using modeling based on manufacturing processes that have helped combine with human insights for optimization.

With consumer expectations growing and companies constantly looking to become more competitive through operating at cheaper costs, improving efficiency, and deciphering analytics

to improve customer quality, IS and DM tools have become increasingly more valuable as AI continues to raise the bar. The IS and DM industry can do an insane amount of backend work in such a short amount of time from security to creating models and these tools are being utilized and equipped into varying industries to complement their needs.

IS/DM Part 3

Subsectors to focus on: data infrastructure (storage/processing), data analytics and visualization, data governance

IS/DM Industry Coverage Report — Comps table

Trading comps & precedent (Aung and Mason)

- Balance out companies from different sub-sectors

Competitive analysis (Rohan Kapoor, Evan Lin)

Market share (Atharv Gupta)

IS/DM Part 4

Case Study (FiveTran)

- Due diligence questions (1 per person)
- Financials analysis (Mason Schultz , Aung Myint)
- Risks of investment (Atharv Gupta)
- Reasons to invest **(Rohan Kapoor, Evan Lin)**

Typical Due Diligence Questions (5 total - 1 per person)

Potential Red Flags (5 total - 1 per person)

Thursday Night - Finish everything on the doc
Friday Night - Finish Presentation
Saturday - Practice Run through
Sunday - Presentation