

WHITEPAPER

Data Fabric: A Case Study

By Shubham Kumar



Contents

- I. Introduction**
- II. Key Components of Data Fabric Solution**
- III. Data Fabric Use Cases**
- IV. Market Analysis**
- V. Patent Trends**
 - 1. Application Year-Wise Patent Trends
 - 2. Country-Wise Patent Trends
 - 3. Assignee-Wise Patent Trends
 - 4. Domain-Wise Patent Trends
- VI. Conclusion**

Introduction

Recently a new technology term has been widely in use broadly related to data and its management. In simple words, data fabric can be explained as an environment that consists of a blend of architecture and services or technologies running on that architecture, which can help organizations manage their data. The main objective of it is to increase the value of users' data to a maximum value and speed up the digital transformation process.

The structure of data fabric can be imagined as similar to a fabric or a weave which is stretched over a large space that connects multiple locations, types, and sources of data, with methods for accessing that data.

The data fabric has an approach very similar to the human brain. It uses a network based architecture to handle information using connections rather than using copies.

This approach enables data fabric so that it can replace point-to-point integration with universal access controls, eliminate data copying, promote collaborative intelligence, end data silos, and create sensible data ownership, probably for the first time since the invention of digitized data. This makes it an important technology for now and the future when policies like the GDPR seek to encrypt the protection of data privacy.

The data can be managed, processed, and stored as it propagates within the data fabric. The data can also be accessed by or shared with internal or external applications for a wide variety of analytical and operation use cases for all organizations, including advanced analytics for forecasting, product development, and sales and marketing optimization.

Objectives

To improve customer engagement through more advanced mobile applications and interactions,

Comply with data regulations, and

Optimize supply chains are a few among them

Key Roles of Data Fabric

Connecting to any data source with the help of pre-packaged connectors and components, thereby removing the requirement of coding.

It is responsible for making the data ingestion and integration capabilities available among sources of data and applications.

It supports batch, real-time and big data use cases.

It is responsible for managing multiple environments like on-premises-cloud, hybrid, and multi-cloud as both data source and data consumer.

It provides built-in data quality, data preparation, and data governance capabilities supported by machine learning and augmented automation to improve the health of data.

It also supports sharing of data with internal and external stakeholders with the help of API support.

As data fabric is a novel terminology, multiple solutions exist that are offered under the name data fabric. However, in reality, there exist only a few solutions that can be considered as true data fabric technology. The seven main components of this technology are explained in section II.

Key Components of a Data Fabric Solution

1. A Network Based Design with Universal Controls Rather than Data Copies

Instead of sharing data copies, permissions can be set for users so that they can access their single source. A data fabric enables one to control these permissions at the data level, which implies that they can set data permissions once instead of on an app-by-app basis. These controls are embedded at that data level; hence it will be there wherever that data appears.

For instance, if an organization allows the Business Development Team so that they can view client email addresses. The organization needs to set this permission only once thereafter; any time a client's email address appears in the form of a dataset, the marketing team can view it. This reduces the hours invested in managing data permissions repeatedly.

2. The Capacity for Autonomous Data

Data fabric enables the process of separating data from the application, creating autonomous data that is not dependent. It can be accessed by multiple applications without requiring point-to-point integration efforts. This solves the previously existing problem of the dependence on copying data and performing expensive integration projects as data has always been tied to the application that created it.

Multiple Uses of Autonomous Data

It allows users to reuse data in a manner similar to API, which allows users to reuse code for new applications.

It enables the easy addition of new features and capabilities to legacy systems.

3. The Presence of Plasticity

Plasticity can be explained as the possibility to re-structure and re-organize the already existing data in a more efficient way. Presently, point-to-point integration implies that data architecture has the maximum amount of connections possible, which may account for a very low IQ score. Data plasticity denotes that these connections may be streamlined to develop actual intelligence for the organization. This has never been properly replicated in machine data earlier.

Builders can make integrations with the help of data contracts or data models so that integrations can be prevented from breaking as data fabric evolves with time. This can be helpful to change the data schema while keeping the internal or external dependencies, including relationships to and from other tables, APIs, or queries, intact. By making the evolution of schema possible, the data model is allowed to evolve without any barriers in the same manner as the human brain continually adapts as it takes on new information.

4. Meaningful Data Ownership

Meaningful data ownership is necessary so that the personal privacy and security of an organization can be protected. It is a basic step for entering into the hyper-intensive data future of AI/ML, IoT, and other emerging technologies.

There has been an increasing motivation from lawmakers to develop and enforce data ownership regulations. But every integration project leads to new data copies, and present organizations already have thousands of data copies to manage. With so many data copies, there is nothing left that can be called “data ownership.”

5. Active Metadata

Traditional metadata is inactive, which already reduces its benefits to a great extent. A data fabric does the job of making this metadata active. This implies that it is updated in real-time and can be questioned, analyzed, and interacted with just like normal traditional data. This is where the true potential of a data fabric comes from.

Active metadata facilitates data management in a natural way. This is the true essence of data fabric technology.

6. Metadata Driven Experiences

A true data fabric has the capacity to replace traditional applications with experiences powered entirely by metadata. The end-user cannot differentiate between these experiences from an API or app but developing them is as easy as working with data in a spreadsheet.

Fully-fledged metadata-driven experiences need a mature data fabric that contains a robust assortment of connected data sources, converting them into a future state technology. But the basics for these experiences must be present in any current technology that call itself a data fabric.

There are multiple advantages of it, from faster build times to easily personalized solutions.

For example, team members of an organization are given the option to create their own custom-tailored solutions for working with their data, even if they have no technical expertise except working in a spreadsheet or SQL. That's exactly what these metadata driven experiences promise to do.

7. The Capacity of Network Effects

Maybe the most promising benefit of true data fabric is the capability for network effects. This is a phenomenon where a network becomes more productive and more effective as multiple nodes are connected.

The more data that is already present on the fabric, the easier it is to move towards new solutions. This is a total opposite from the present model of point-to-point integration, where projects become more complicated and more costly over time.

A true data fabric becomes more efficient if it is more in use.

Data Fabric Use Cases

The main requirement of data fabric comes into picture in order to solve the problem of challenges faced in managing big data.

Key Challenges in Big Data

01

Inability to Handle
Data Growth

03

Delayed System
Response Time

05

Cope Up
with Ai

02

Poor Quality
Data Analysis

04

Expensive
Management

One of the major reasons for a disorganized data management vertical is more than one disconnected team. Data fabric provides a unified architecture to perform automated data management throughout the lifecycle: sourcing, cleaning, orchestrating, preparing, archiving, & analyzing and, in this way, resolving this process inefficiency.

Key Opportunitites in the Data Fabric Market

01

Explosive Rise in
Consumer Mobility

03

Enhanced Adoption
of the Cloud

05

Edge
Computing

02

Data Ops & MLOps
are the Frontier

04

Increased Adoption
of IoT

Some of the data fabric use cases used by organizations are listed below:



Improving Business Understanding

Data Fabric can be thought of as a topological map of anomalies, inflection points and business outcomes across the enterprise his makes it the perfect training and testing set for machine learning and I use to understand the business. This can also make it easier to implement process mining projects that make sense of business processes spanning multiple applications



Simplifying Predictions and Triggered Actions

Data fabrics can also be used to train, configure and deploy simple prediction algorithms and trigger actions that run across various enterprise application endpoints
These types of use cases span everything from security traceability to audit compliance and revenue-generating events like cart abandonment action, ad optimization, customer retention, marketing and even orchestrated selling.



AI Data Collaboration

A data fabric architecture can provide AI engineers with access to broad, integrative data for better-informed decisions.
AI needs broad access to high-integrity data, a data fabric can support the efficient delivery of information to AI applications for quick, well-informed decisions



Creating a Data Marketplace

Enterprises implementing a data fabric architecture can also set up a more accessible data marketplace that makes it easier for citizen developers to weave disparate data sources into new models. A data marketplace allows d ata engineers to set up an infrastructure that can be used across multiple use cases rather than creating fresh infrastructure for each use case individually.



Enhancing security

A data fabric can also improve security applications by tying together data and applications from across physical and IT systems



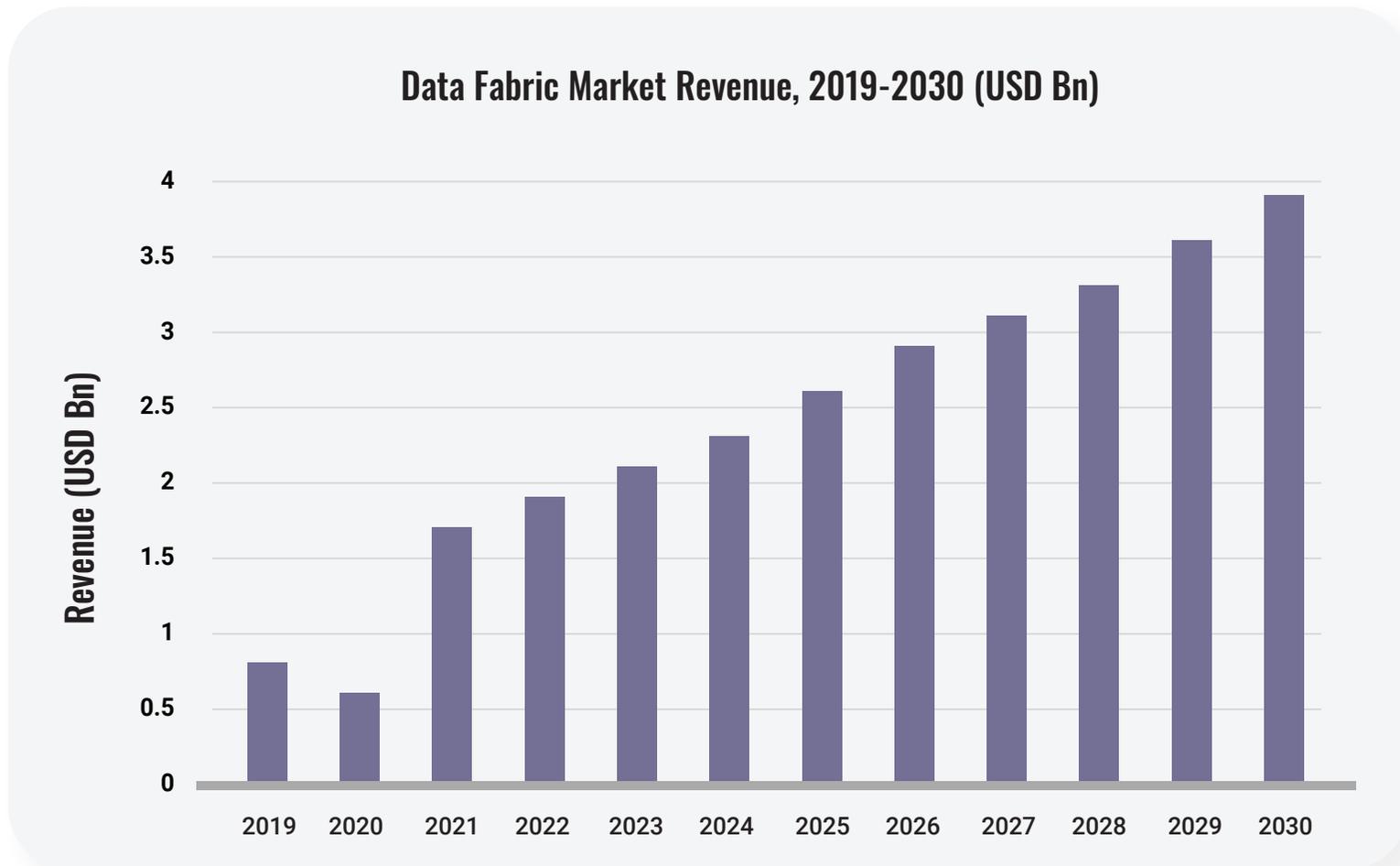
Creating a Holistic Customer View

Organizations can also use a data fabric architecture to weave together data from a customer's activities along with the various roles that interact with them to get a more holistic view.
This could incorporate real-time data of various sales activities, potential revenue realization, customer onboarding time and customer satisfaction metrics.

Market Analysis

1. Data Fabric Market by Year

By analyzing the opportunities discussed in the above section, the data fabric market is analyzed as shown in the graph below:



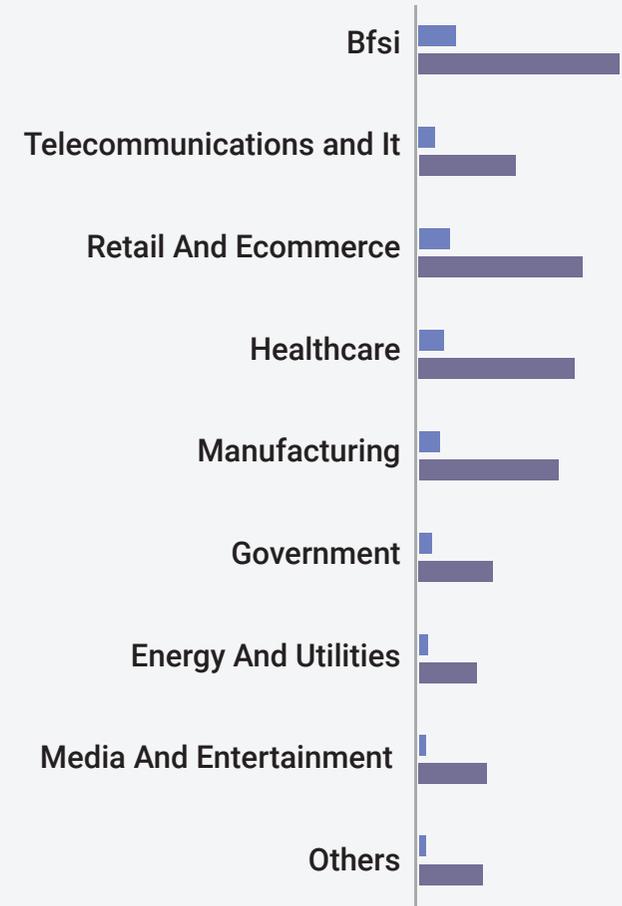
- During the forecast period of 2020–2026, It is predicted that the global data fabric market size may grow from USD 1.0 billion to USD 4.2 billion in 2026 at a Compound Annual Growth Rate of 26.3%.
- Factors influencing the rising demand of data fabric services include: enhanced volume and variety of personal as well as business data which will push the need for streaming analytics in real-time.
- The Global Data Fabric Market size is approximated to be USD 0.82 billion in 2019 and may reach USD 3.88 billion by 2030 with a CAGR of 15.6% from 2020–2030.

2. Data Fabric Market by Sector

BFSI, i.e., Banking, financial services, and insurance domain, owns the largest amount of data, and hence, this sector is the major driver in the growth of data fabric services as it needs a large amount of data to be managed effectively.

It is followed by the telecommunications and IT sector, retail and commerce, and healthcare sectors.

2018
2026

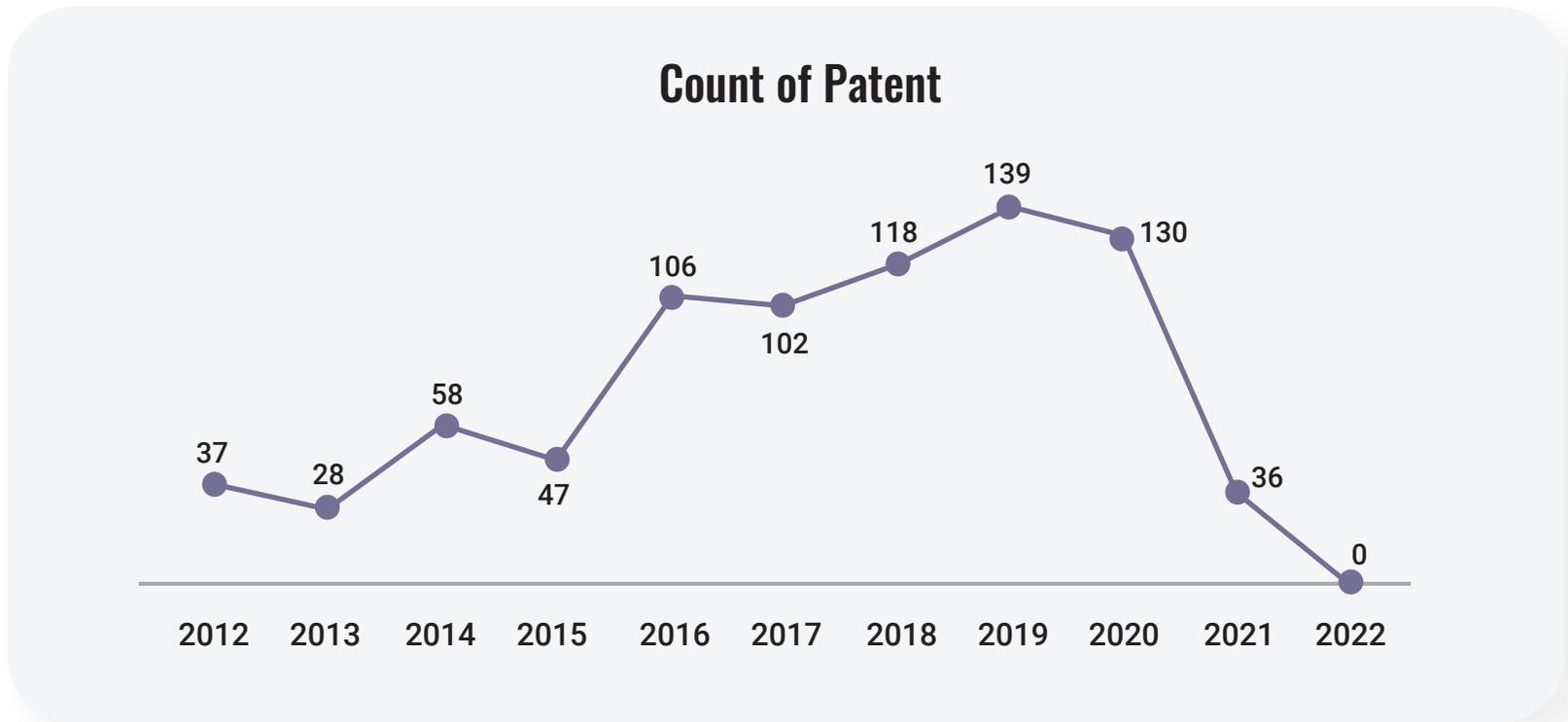


Patent Trends

1. Application Year-Wise Patent Trends

The term 'data fabric' was first given by data management systems provider NetApp. The company explained the challenges of managing data in hybrid cloud environments in a whitepaper in the year 2016. The patent filing in the data fabric technologies originally started from 1971 with one patent filed, and proper filing started from the year 1988 when 21 patents were filed.

A number of patents have been filed related to data fabric thereafter. We can view the patent trends of last 10 years from the graph below:



2. Country-Wise Patent Trend

Superpowers of the big data economy

According to a market report, the expected top countries which will adopt and use data fabric technologies in the future can be:



United Kingdom

Adoption of big data in the UK is found to be growing at a rapid phase. At present, big data is widely adopted by telecom industries. Other business verticals are taking efforts towards adopting big data in their business operations as well.



United States

The collection of massive amounts of data from various sources is the key factor that drives the adoption of big data in the US markets.



China

In the Chinese market, policy support and technology integration are the two most important factors that contribute towards the growing adoption of big data.



Japan

The increased usage of social media to understand the end-users/customers is one of the most important factors that contribute to the adoption of big data in the Japanese market.



South Korea

The South Korean government has decided to expand its big data market to 10 trillion won by 2022.



India

The big data market in India has been gaining momentum in recent years with the involvement of many small and large big data and AI-based companies. This could potentially make India to be the largest big data market in the long run.



Russia

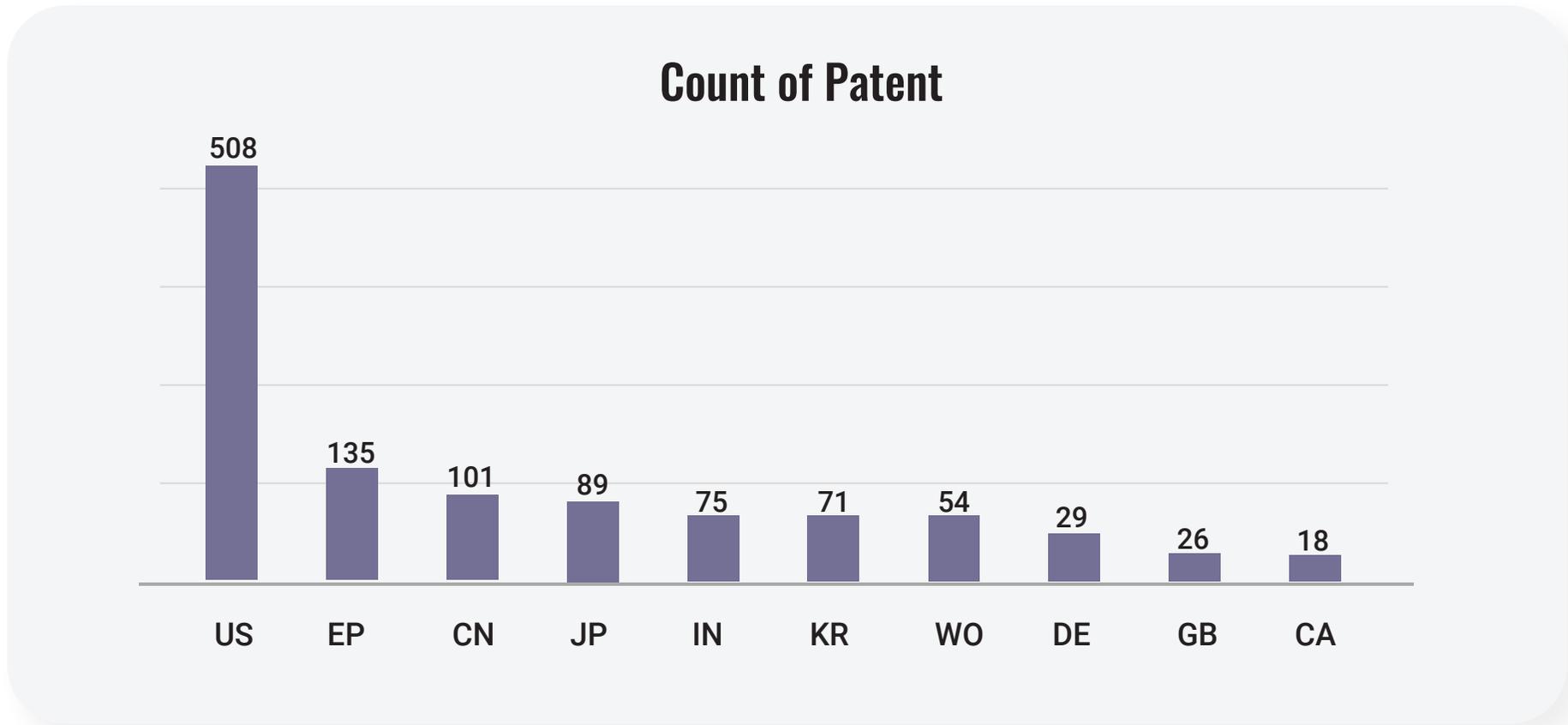
In Russia, the three major verticals to adopt big data are finance, telecom and retail.



Canada

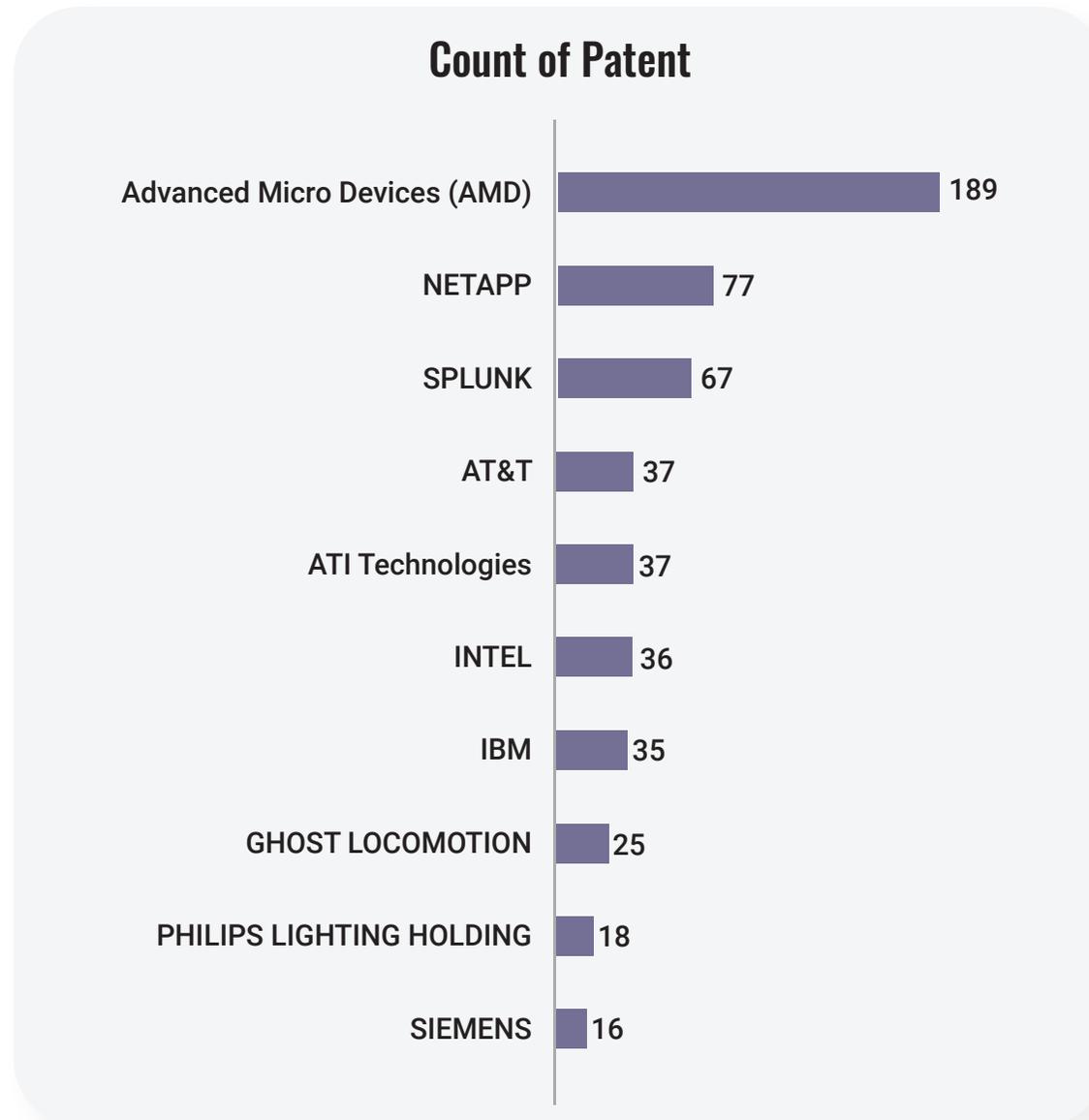
the Canadian market for big data and analytics was found to be \$1,866.6 million in 2017. By the end of 2022, this value is expected to grow with a CAGR of 9.4% between 2017 and 2022. The development of the software industry is believed to be the key factor that drives the adoption of big data in the Canadian market.

However, country-wise distribution of patents in data fabric technologies can be seen from the graph below:



3. Assignee-Wise Patent Trend

Top 10 players who have filed patents in the data fabric technology are shown in the graph below:



Apart from these some of the topmost providers of data fabric technologies are:



IBM

IBM's Cloud Pak solution offers access to the right data at the right time across any cloud and on premise platform. This solution of IBM provides intelligent data fabrication for faster and trusted data outcomes



HP

HPE's Ezmeral data fabric is mainly developed to simplify data management on a global scale with enterprise-grade reliability, flexibility and performance. This data fabric platform is built using MapR technology to deliver a unified platform where the data are collected, stored, managed and applied in various formats from various sources.



K2View

K2view has a single data fabric solution that can integrate, transform, enrich, prepare and deliver data in a single platform. K2View's Data Fabric organizes data from various sources based on data such as customer, location, device, etc.



SAP

SAP's data fabric solution consists of the combined capabilities of SAP data intelligence and SAP HANA. SAP Data Intelligence transforms the collected data into valuable information which can be accessed at the right time using the right context.



Oracle

Oracle is well known for its data management systems. Oracle's data fabric solution includes industry-leading ELT, data preparation, replication and can effectively work together. This data fabric solution can transform data without having an impact on the systems, cleanse and repair data to make it trustworthy and reliable, provide data with 0 downtime for operational purposes, replicate and recover data in case of any failures and finally, use algorithms to streamline data pipelines.



NetApp

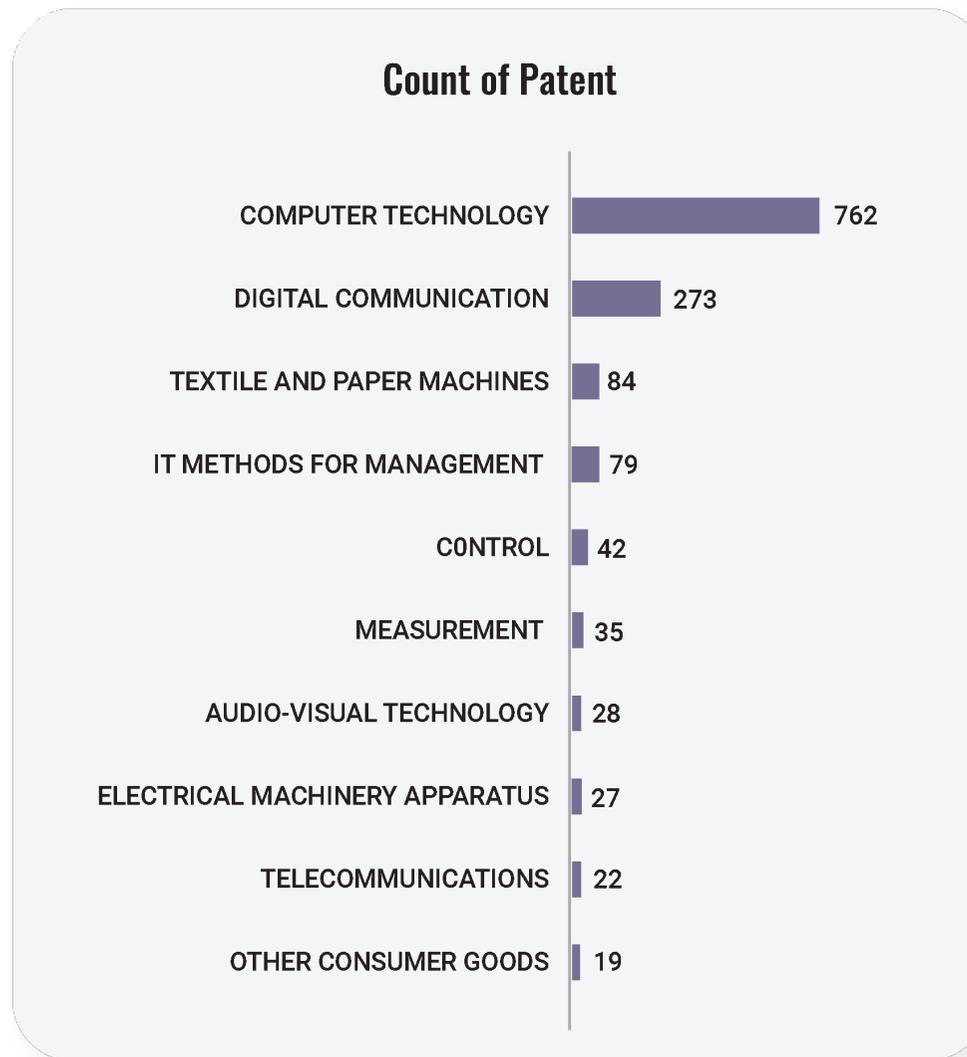
NetApp's data fabric solutions are mainly focused on providing simplicity and agility. NetApp's Fabric Orchestrator increases operational efficiency seamlessly in on-premises as well as cloud environments by providing the necessary features to build and manage the data fabric.



VMWare

VMWare, a global leader on Cloud infrastructure, has developed a solution to make accurate business decisions based on event-driven information. VMWare's approach aims at reducing costs when compared to traditional database architectures. VMWare's data fabric solution can offer dynamic scalability without increased costs, real-time data and event-driven communication, Zero-latency, transparency in transactions, faster and cost-effective processing and simplified database provisioning and governance.

4. Domain Wise Filing Trend



As can be seen from the graph, the highest number of patents disclosing data fabric technologies are filed in computer technology, followed by digital communication.

Conclusion

- » The demand for data fabric is increasing in all the fields with the enlargement in the size of data.
- » Data fabric enables frictionless access and sharing of data in a distributed data environment and eliminates the need to copy data.
- » It enables the organizations to optimize the data for faster insights. It also helps to minimize the inconsistency of data and improves its quality.
- » Organizations need to develop and adopt data fabric technologies to organize their data in a better and more efficient manner.

Get in Touch

India Office

207-208 Welldone TechPark, Sohna Road
Sector 48, Gurugram, Haryana 122018, India
+91 124 429 4218

US Office

4 Heinrick Way Bridgewater,
New Jersey 08807, USA
+1 347 480 2054
+1 202 697 9162



info@iebrain.com