Data led Digital Transformation in Indian Consumer Banking

Contents

[**Foreword 3**](#_Toc101472716)

[**Executive Summary 4**](#_Toc101472717)

[**Evolution of Consumer Banking in India 5**](#_Toc101472718)

[Technological Developments in the Indian Banking Sector 6](#_Toc101472719)

[Shifting Facets of the Banking Ecosystem 8](#_Toc101472720)

[**Foreseeable Future of Banking 9**](#_Toc101472721)

[What will roll the dice in building “Banks of the Future”? 9](#_Toc101472722)

[Competitive Advantage and Growth in Banking 10](#_Toc101472723)

[**Emergence of Cloud Computing in Banking 13**](#_Toc101472724)

[Data Security 13](#_Toc101472725)

[Realizing the necessity of Securing Data 13](#_Toc101472726)

[Availing Gains by implementing a model for Secure data 13](#_Toc101472727)

[Possible Path to be Followed 13](#_Toc101472728)

[Up Above the Cloud 14](#_Toc101472729)

[**Riding the Cloud Wave with Azure 17**](#_Toc101472730)

[Data Security on Azure 18](#_Toc101472731)

[Multi Data Lake Ecosystems 19](#_Toc101472732)

[**Banking Industry Insights – KPMG Survey 22**](#_Toc101472733)

[Data Vision 22](#_Toc101472734)

[Data Strategy 23](#_Toc101472735)

[Data Technology 23](#_Toc101472736)

**Foreword**

Today’s age of competition and augmentation in every sphere has compelled even the Banking sector to step up, become more agile and cater to the changing demands and needs of customers. This also brings the need to partner and collaborate with start-ups and service providers to further enhance their portfolios. Banks have abundance of data that can be leveraged to gain insights and deliver actionable outcomes. It has become important to devise a digital and data strategy to ensure scalability to process petabytes of data, flexibility to hyper-personalized products and service and, speed to react quickly to the change in demands from the business.

Over the decades, The Reserve Bank of India (RBI) has taken steps to embrace digitalization and IT Transformation in the Indian Banking Sector. With RBI’s regulations and recommendations, Information Technology saw a manifold in innovation with the arrival of Debit/ Credit card in late 1980s and 90s, Electronic Clearing Services (ECS) in late 1990s, Electronic Fund Transfer (EFT) in early 2000s, RTGS in 2004, National Electronic Fund Transfer (NEFT) and the Introduction of ATMs.

With recent technological advancements in the 5G space and the 2021 Government guidelines of EASE 4.0, banking in India is heading towards a revolution marked by state-of-the-art capabilities with focus on 24\*7 availability, analytics, automation, and API based banking. The banking services are growing at a faster pace from loans and deposits to credit cards and co-lending scenarios leading towards seamless and virtual interactions benefiting customers as well as the banks.

This provides a compelling reason for banks to leverage cloud capabilities and ensure seamless scalability in the areas of Big Data Processing, Analytics, Artificial Intelligence and Machine Learning. For CIOs of organizations, it now becomes important to build cloud and data strategy roadmaps that have the potential to meet fluctuating demands and at the same time ensuring data democratization, governance, security, and regulatory compliance.

# Executive Summary

The Banking Industry is at the forefront of the technology revolution. Data & Analytics is rapidly changing the face of the banking industry, as several banks are now turning towards data digitization as well as exploring better opportunities for an enhanced customer experience. With the advent of customized data lake cloud implementation, leading IT product enterprises often in partnership with banks, business owners and consumers facilitate management of financial operations and processes. Although different sectors of these enterprises continue gaining traction today, a large part of it still focuses on the traditional global banking industry.

This publication encapsulates various aspects of Indian Consumer Banks; It provides insights on the evolution of banking sector in India with respect to data and technology. The drivers of key aspects in banking like data led insights, open banking and data security are highlighted and at the same time, key challenges, upsides, and the future trends of the banking sector are also covered. It also highlights the defying challenges for banks to use data to its core capability. Specifically, moving data to the cloud can be a deciding factor for banks from various aspects of data security, cost, volume of the data and infrastructure. It also introduces concepts of data security on the cloud and how the Microsoft Azure suite of products fits a banks’ data and analytics needs.

We at KPMG, India conducted a detailed survey alongside some leading banks of the country to capture on-ground data points and thoughts concerning Data Cloud implementation based on four key themes, mainly –

Vision

Strategy

Technology

KPMG brings to you a holistic picture of insights on said themes, coupled with viewpoints from the most experienced senior leadership.

# Evolution of Consumer Banking in India

Banking eco-system plays a seminal role in the economic development and provides financial resources to an individual, corporation, government, or a sector in need. The health of any economy is directly related to the financial status of its banks.

Over the centuries, numerous changes have taken place within this industry, starting from technology to the diversification of financial services and products. Retail banking faces a more complex environment than perhaps ever before. The social and economic landscape has been radically reshaped while customer needs and expectations continue to evolve continuously.

Since the past few decades, technology and digitalization have been at the forefront for Indian banks to achieve their vision and to reach their milestones. Data is gaining relevance outside the technology domain as the future has significant dependency on data that is being generated across the banking value chain.

The arrival of the Internet in the 1990s lent a plethora of possibilities in worldwide commerce, thereby naturally impacting the very functioning of banks. The internet revolution of the world coincided with the Indian Policy of LPG - Liberalization, Privatization and Globalization - aimed to make the Indian economy the fastest growing economy in the world. It was at the cusp of attracting many foreign banks like HSBC, Citibank, Bank of America etc. to India, thereby opening new markets, new products, and efficient delivery channels for the banking industry. Most of the banks were quick to leverage emerging technologies which led to significant increase in penetration, productivity, and efficiency. The Indian Government encouraged foreign investment, which opened the economy to foreign and private investors, thereby leading to the introduction of mobile banking, internet banking, ATMs, and more.

To stabilize the nationalized public sector banks, the Indian Government formed the Narasimham Committee in 1991 to manage reforms in the banking sector. During this time, the Government approved various private banks. These included the Axis Bank, IndusInd Bank, and ICICI bank.

With the potential to become the 3rd largest banking industry by 2025, India’s banking and financial sector is expanding rapidly. The Indian banking industry is currently worth more than 1 trillion dollars and this significant breakthrough has benefited the citizens on ground in more ways than one by -

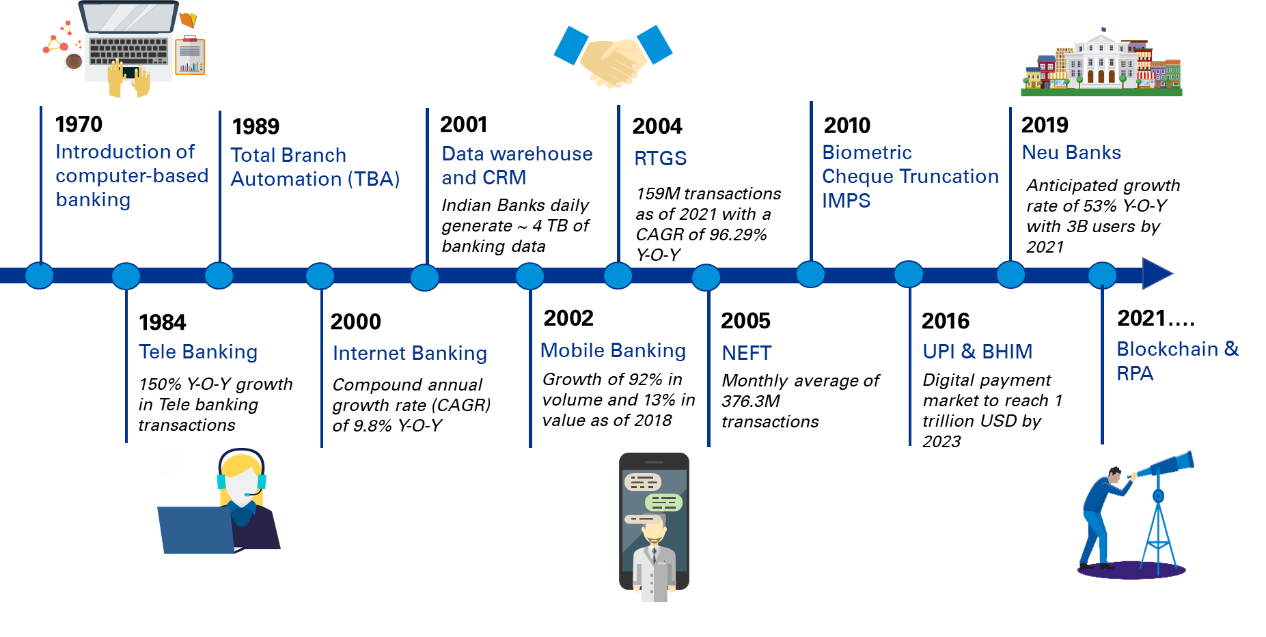
- Driving financial inclusion

- Providing financial assistance to MSMEs

- Enhancing customer experience and transparency

E-Banking has reduced costs drastically and has led to tremendous revenue generation - the cost of a bank transaction on Branch Banking is estimated to be in a range of Rs.70 to Rs.75 while it is around Rs.15 to Rs.16 on ATM, Rs.2 or less on Online Banking and Rs.1 or less on Mobile Banking

**Technological Developments in the Indian Banking Sector**

****

* **Payment Systems**

Graph 1. Count of payment transactions across the various modes of payment systems over 3 years

*Source : RBI Report – May 2021*

* The payment systems recorded a robust growth of 26.2 per cent in terms of volume during 2020-21 in addition to the expansion of 44.2 per cent in the previous year.
* In terms of value, the contractionary trend which started in the previous year (-1.2 per cent) got further amplified and witnessed a drop of 13.4 per cent, mainly due to lower growth observed in the large value payment system, viz., Real Time Gross Settlement (RTGS) system and decrease in transactions of paper-based instruments. The decline in value of transactions in RTGS is largely attributable to the subdued economic activity.
* The share of digital transactions in the total volume of non-cash retail payments increased to 98.5 per cent during 2020- 21, up from 97.0 per cent in the previous year.
* **Mobile Banking**

Graph 2. Growth of mobile transactions over the last decade

*Reference- RBI*

Looking at these emerging trends, it can be inferred that there has been a significant increase in IT spending over the past few years and similar growth is also forecasted in coming years by many researchers to take advantage of emerging technologies and data trends.

Gartner forecasts IT spending by Indian banks will increase by 9.02% in 2021 and 4.10% in 2022 from their total spending of US$10.36 billion in 2020.

Graph 3. Growth of IT spending by banks in India

*Reference- Gartner*

With the advancement in the digital space, the risk of **cybersecurity** glooms over us. Also, the use of open banking architectures based on improved data exchange increases the chances of cyber frauds. Open APIs can also pose a serious cybersecurity risk because they provide clear access to customer banking data such as transactions and balances stored in the infrastructure. Other potential operational and cybersecurity issues related to the use of APIs including data breaches are misuse, forgery, denial of service attacks, and infrastructure malfunctions. Banking industry has taken several initiatives in cyber security and data protection.

While cyber threats cannot be avoided, it is imperative for the banks to have a secure cyberspace and tackle the cybercrimes with utmost urgency and importance and have a smart and intelligent cybersecurity in place. Thus, implementing an intelligent cross-channel defense mechanism has a huge opportunity for the banking sector to improve the brand value as well as customer trust. The key is to identify just how cyber criminals can cash-out and monetize the access they achieve to our global financial systems, including the payment and clearing systems at the heart of our financial world, introduce a range of measures, consisting of fraud controls, data analytics and other security strategies for each of these key systems and the gateway systems that communicate with them.

**Shifting Facets of the Banking Ecosystem**

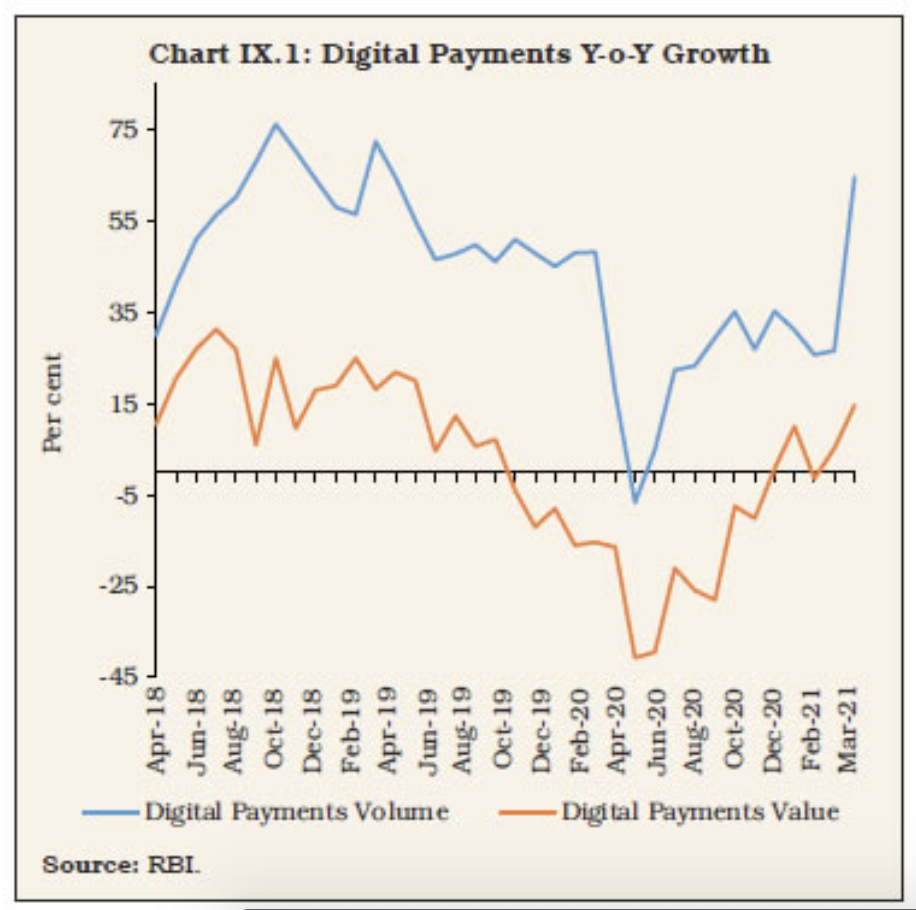
With all the evolution and progress that’s been happening in the banking sector, the eco-system is changing too. Illustration below showcases how the changing facets of a banking ecosystem will need banks to constantly innovate and remain agile.



Fig: Changing facets of the banking ecosystem

# Foreseeable Future of Banking

When in the early 1980s, then deputy governor of RBI, Dr. Rangarajan recommended a phased approach for the “computerization” of the Indian banking sector, little did the country know that decades later, India’s digital banking market would generate a whopping USD 776.7 million in just one financial year (2021).

In the foreseeable future, the market is estimated to grow at a CAGR of 9.8%, earning revenue of around USD 1,485.5 million by the end of 2028.

A combination of next generation technology and IT operating model transformation has led to a 360 degree change in the Banking IT landscape. Banks of present have evolved their operational models into new forms such as:

* **Going Digital:**Banks are focusing on leveraging Digital Retail Banking to make stronger relations with the customer by providing them with customized products and solutions
* **As a service provider –** Cash rich banks are heavily investing in technology and are aiming at introducing this technology induced facilities in the market
* **A massed service delivery model:**New entrants focus on specialization in one of the areas of banking as compared to traditional banks that offer an end-to-end business
* **Reshaping IT operations:** IT infrastructure has drastically changed from the traditional in-house technology ownership with IT operations covering data centers, operating systems and vendor software solutions to a more sophisticated information and communication technology

**What will roll the dice in building “Banks of the Future”?**

The present and future is technology, and as we advance, technology will continue to reshape our lifestyle. Banks will transform alongside the shifts in how we work, live and play. Banking sector will shift their focus to **data, business models, regulation, and evolving technology** to improve the financial service ability which in turn will improve financial wellbeing. Many disruptive technologies will spread across the sector, with primarily focus on connecting people and focusing on rapid customer engagement.

The above mentioned four primary areas will enhance the financial services’ ability to deliver improved financial wellbeing and below are the details:

|  |  |
| --- | --- |
| **Data** | **Regulation** |
| Increasing transaction data will connect with other information to support new digital service development. Banks will need to compete through their data assets, which will likely result in material acquisitions of data capabilities for them to retain relevance in the digital economy. With the exponentially increasing importance of data, banks will have to build high security. | The major shifts in banking industry will require governance and regulating bodies to come up with completely new ways to identify and manage risks, regulate the activities being undertaken by a broader range of participants and judged on the outcomes they deliver to the customers. |
| **Business Model** | **Evolving Technology** |
| The widespread variety of data will fuel new entrants and top banking platforms will help in building a business model. A futuristic model can be designed to connect the front, middle and back offices. It can focus every process, function, and relationship of the organization on meeting customer expectations, creating business value and driving sustainable growth. | Technology is and can be an enabler and driver of change, and it is already seen over last 10 years how it has impacted every industry. Cloud computing, IOT, AI, Quantum, block chain are the game changers in the emerging world. Like any other industry, banking has also redefined its face. |

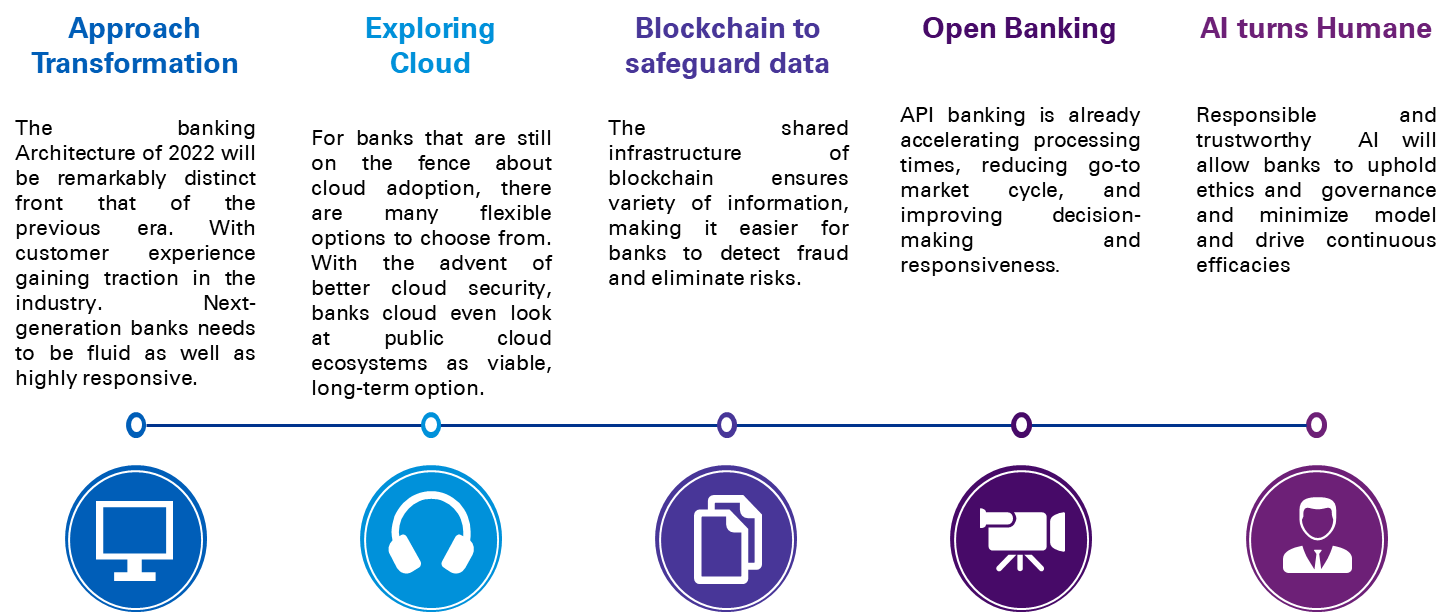
The banking industry of the future will be completely different to what we see today. Banks will no longer be limited to a financial institution but a trusted interface for life, embedded within the needs and lifestyles of consumers. Banks will have to be much more adaptive in nature to accommodate the ever-increasing customer needs. Rapid innovation in technology would compel the industry to be adaptive and customer centric to sustain the ever-rising competition.

## Competitive Advantage and Growth in Banking

In the wake of the pandemic, the change that was expected to happen over the next 5 years is now taking place in 24-36 months. To keep pace, banks needed to innovate through the lens of changing lifestyles, evolving business processes, and emerging smart technologies, rather than taking a traditional product approach. New skills are needed in the organization, with fintech and entrepreneurial mind-sets at a premium, and a ‘fail fast’ culture to shorten time to market. Advanced data analytical capabilities are critical, including cloud-based AI and predictive modeling techniques combining internal with external data to give a truly robust view. To provide real value, data must connect to inputs from outside the bank. Better data availability and quality enhanced analytical capabilities informed by AI and machine learning, and faster reporting – these must be the hallmarks of a risk function of the future. Using dynamic, predictive models to better understand customers and the associated risks could also drive opportunities to create competitive advantage and growth.

* *Open banking giving platform for greater customer value* - Open banking involves many moving parts. To keep customer data safe and minimize risk, financial institutions need to carefully consider each phase of the open banking process. This essentially facilitates collaboration between traditional banks and the new fintech-based business model. The reality of open banking is that it uses a focused omni-channel approach. Open banking enables open account information, this means that financial institutions can provide their clients with a complete picture of their financial transaction history. Access to this information gives clients centralized insights that can help them improve their financial behavior and decision making. Digital agility refers to the institution's ability to adapt quickly to change and implement solutions. The high level of agility depends on many factors, including the digital process and software used.
* *Rise of alternate data and how it’s changing finance* - Alternative(non-traditional) data that is culled from the offset of the traditional data set which succor to provide the indication of future performance of a bank outside of traditional sources. Digital innovations such as increased research and adoption of AI / ML techniques, cutting-edge breakthroughs in natural language processing (NLP) and natural language generation (NLG) for textual data is also contributing to exponential rise of alternative data use. Alternative data goes beyond the traditional 5 C’s of credit i.e., character, capacity, capital, conditions, and collateral. By enriching traditional financial information and datasets (bank statements, tax returns, credit reports, etc.) with new sets of data related to customer’s behavior and their perception of products and services, the combination can bring numerous advantages. Very soon, Banks will regard these non-traditional data sets as primary rather than so-called ‘Alternate’ data.
* *Building a next generation experience for customers* – Be it any industry, consumer priorities have shifted to top of the list towards heightened convenience, reliability, authenticity, and transparency. So the rule remains the same in consumer banking as well. Customers are the flag bearers of a brand as their trust defines the market value of the organization. Leading banks are investing at a large scale in their front office to enhance the end-to-end customer journey.
* *Hyper-personalized offerings and seamless omni-channel* – Today’s customers are more well-versed, connected and demanding than ever. And while they have come to expect the highest standards of personalization, choice, speed, satisfaction and security in every digital interaction, industries have served to intensify the customer experience 2x better. Banks understanding the customer behavior, with the benefit of technology in hand, have started offering their customer a rationalized approach that delivers the fast, convenient, secure, personalized, consumer-centric services that are revolutionizing every corner. A customer can experience all these offerings from any front of banking operations. Other than offering the same banking operations either through a website or mobile application, a call center, a bank’s branch, or any other available channels. Omnichannel banking platform also allows **real-time data synchronization** between different channels. That means any customer can start on-boarding process with one channel and finish it with another without the need to provide same data repeatedly. Omnichannel platform can improve marketing performance, simplify onboarding processes, boost customer retention rates, and much more.
* *Transforming credit risk operating model to forecast bank’s revenue* - Credit risk refers to the prospect of loss due to failure in making payments. With the help of correct strategy and understanding, a customer at portfolio levels banks is trying to mitigate the losses that can be incurred. Banks traditionally use traditional credit risk models to predict categorical, continuous, or binary outcome variables (default/non default), as ML models are difficult to interpret and are not easily verifiable for regulatory purposes.AI based decision tree techniques can result in easily traceable and logical decision rules despite having non-linear characters. Unsupervised learning techniques can be used to explore the data for traditional credit risk modelling while classification methods such as support vector machines can predict key credit risk characteristics.
* *Virtual assistants and chatbots swapping humans & technology in identifying frauds and threats –* To meet an increased demand for fulfilling the customer’s prescriptions either over internet or live at an organization, bots are placed to assist. AI is embedded in banking business models to transform the potential in banking. With bots replacing humans at work, the time and effort of the customer in getting a generic query answered is minimal. Development and adoption are growing fast and are likely to accelerate further as firms intensify their investments. Banking sector’s mobility in the new era of technology also offers identifying the fraudulent transactions using AL/ML techniques. The predictive models identify and notify the authenticity of a transaction, redeeming both customers and banks from any fraud.

In conclusion, we would say that banking is embarking on an exciting journey with technology in general and cloud in specific backbone of its innovation and growth. If there are top 5 trends that banks should look out for in the coming years, it would be these:



# Emergence of Cloud Computing in Banking

## Data Security

### Realizing the necessity of Securing Data

Digital banking makes banking more convenient and easier for customers to manage their finances at any time and from anywhere. While consumers prefer online banking over traditional approaches, transfer of personal data and financial information through digital channels raises an important concern, whether data is protected and kept private. According to a 2020 study published by KPMG, 87% of consumers say data privacy is a basic human right. Yet 68% say they don’t trust companies to ethically sell their personal data. Banks and other financial institutions manage a large volume of sensitive information about their customers, and the breach of such data can have dire consequences.

There has been a conscious effort from the Central Bank to emphasize the need for information security by means of providing frameworks and guidelines. In addition, the IT (Amendment) Act, 2008 has laid the foundation for strengthening cyber security and data protection in India. This will have implications on the existing regulatory landscape of the banking industry especially with introduction of section 43A that mandates corporates to implement ‘reasonable security practices’ for protecting ‘sensitive personal information’.

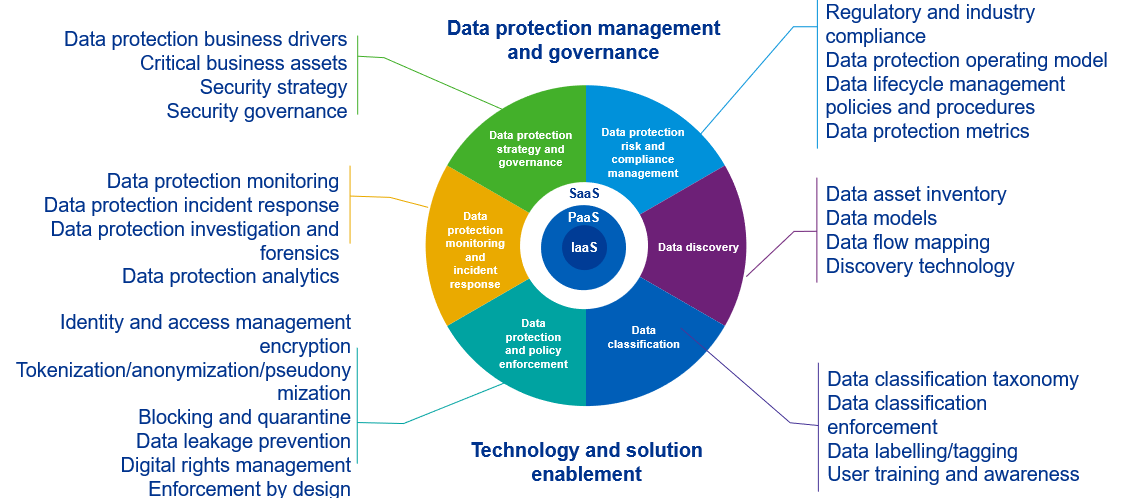
### Availing Gains by implementing a model for Secure data

With the help of a strong data security plan, organizations gain the trust of customers, and this also helps retain their brand value. Organizations are levied penalties in cases of non-compliance. Having compliance standard stimulates banks to focus on cybersecurity. When a bank stays compliant, it ensures that it is meeting consensus security and protecting customer data. When a bank faces a data breach, consumers lose time, money and trust in the bank. A bank may recover fraudulently spent money fully or partially, but counter actions are painstakingly long and stressful.

### Possible Path to be Followed

As technology and cloud solutions become more sophisticated, so do the efforts of hackers endlessly crafting creative new ways to access sensitive data. With businesses turning to the cloud, now is the time to ensure these services are governed and monitored by corporate IT, risk and cyber security professionals who understand today’s emerging threats and regulatory requirements. With better strategy, banks will be able to win their customer’s trust and retain the brand value of their organization. These days banks have deployed the two-factor authentication solution to prevent data breaches.

One Possible forward-looking approach can be **Cloud Capability Architecture**. The data protection reference architecture consists of six domains to safeguard and maintain data security. The below domains are part of the data reference architecture, with capabilities for each domain.

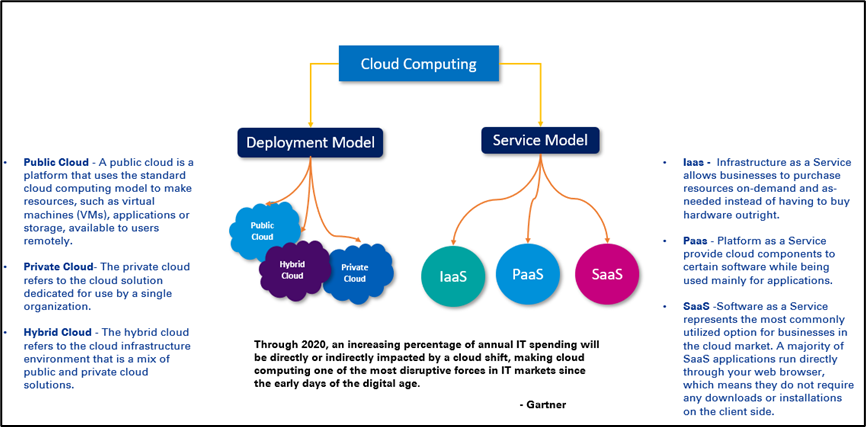


## Up Above the Cloud

While cloud has gone mainstream and, as one of the crucibles of the new digital economy, innovative cloud services, platforms and infrastructure are helping to deliver high levels of scalability, flexibility, and resilience, one of the major hurdles for implementing cloud data lakes in banking has been ensuring data security. The thought is that data is safer in a physical on-premises data center than in a private cloud. 90% of all data ever produced by humans has been made in the last two years. The volume of information is constantly growing, and organizations are finding it hard to manage their data and control where and how sensitive information is actually stored and transmitted.

Cloud solutions are helping to unlock leading capabilities for banks pursuing workforce productivity gains, enhanced efficiency, and new ways to meet rapidly evolving consumer expectations. With this, data protection for cloud environments is becoming critical as more businesses transition from building and managing their own data centers to storing applications and data in the cloud.

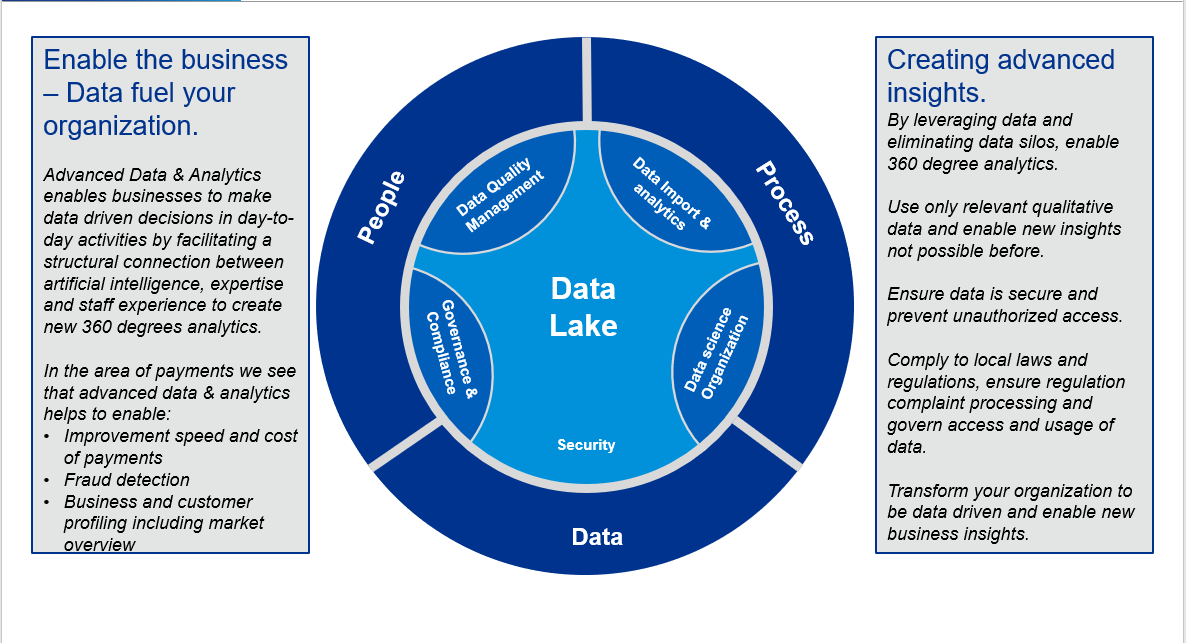
Many organizations, banks included, are still in the early stages of their migration to cloud Infrastructure as a Service (IaaS), grappling with issues that include obstinate legacy architecture, data privacy compliance and the role of cloud providers versus the organization. Others may be more advanced in their adoption of increasingly popular Platform as a Service (PaaS). Meanwhile, almost every organization today relies on some form of cloud Software as a Service (SaaS) for standard office productivity tools, online training, enterprise-wide HR management platforms and more. To build in or migrate their on-premises structure to cloud, banks can opt for services available as below:



While most banks today are already hosted or using several services of cloud at some level, several other banks are still reluctant to transform their ecosystem. Below are some myths associated why banks hesitate to build their infrastructure in cloud:

|  |  |
| --- | --- |
| Myths | Fact |
| Cloud is always about money | Banks can reduce the cost and increase capacity at the same time by optimizing the operating expenses |
| Insecurity/ Privacy issue | Cloud provides strong security to ensure the compliance. The data hosted is fully secured. |
| Data breaches | API based cloud security prevents these mishaps. |
| Banks lose control over data | Banks maintain the rights to who gets access to what data |
| Regulatory/Compliance | Cloud compliance refers to the vast set of regulations and **principles that organizations must follow when using systems delivered through the cloud** |

Moving to cloud is no longer an option for banks, its rather a necessity for prospering in the new era of modern banking. Banks have set up their data warehouses, data lakes and multiple other applications in different cloud system.



Some of the key advantages of cloud implementations are:

* **High Availability:** Irrespective of location, data is available at other geographies during natural disasters.
* **Low Infrastructure Cost for Decommissioning**: It eliminates end-of-life concerns for physical infrastructure. For decommissioning any on-premises servers there is a laborious process & huge cost involved in decommissioning of infrastructure.
* **Quick Adoption** of new technologies without investing in physical infrastructure.
* **Numerous Subscription Models** offers flexibility of services, leading to a huge reduction of burden on Internal IT.
* **Data Collaboration** **and Integration** with external parties and third-party consumers.
* **Faster Implementation** and adoption of newer tools and technologies
* A **strategic shift** by moving from CapEx to OpEx.
* **Reduces TCO** (Total cost of ownership) with software licensing and assurance handled by the cloud service provider to large extend.

**Riding the Cloud Wave with Azure**

Compared to other industries, Banking sector is held up with legacy and traditional technologies and not inclined to adopting emerging products till the Digital Security teams are confident on the product’s maturity and stability. Data security and Data privacy are key for any Bank to protect PII data. Having secured data storage and encryption with high performance is vital for operational efficiencies, application performance and regulatory compliance.

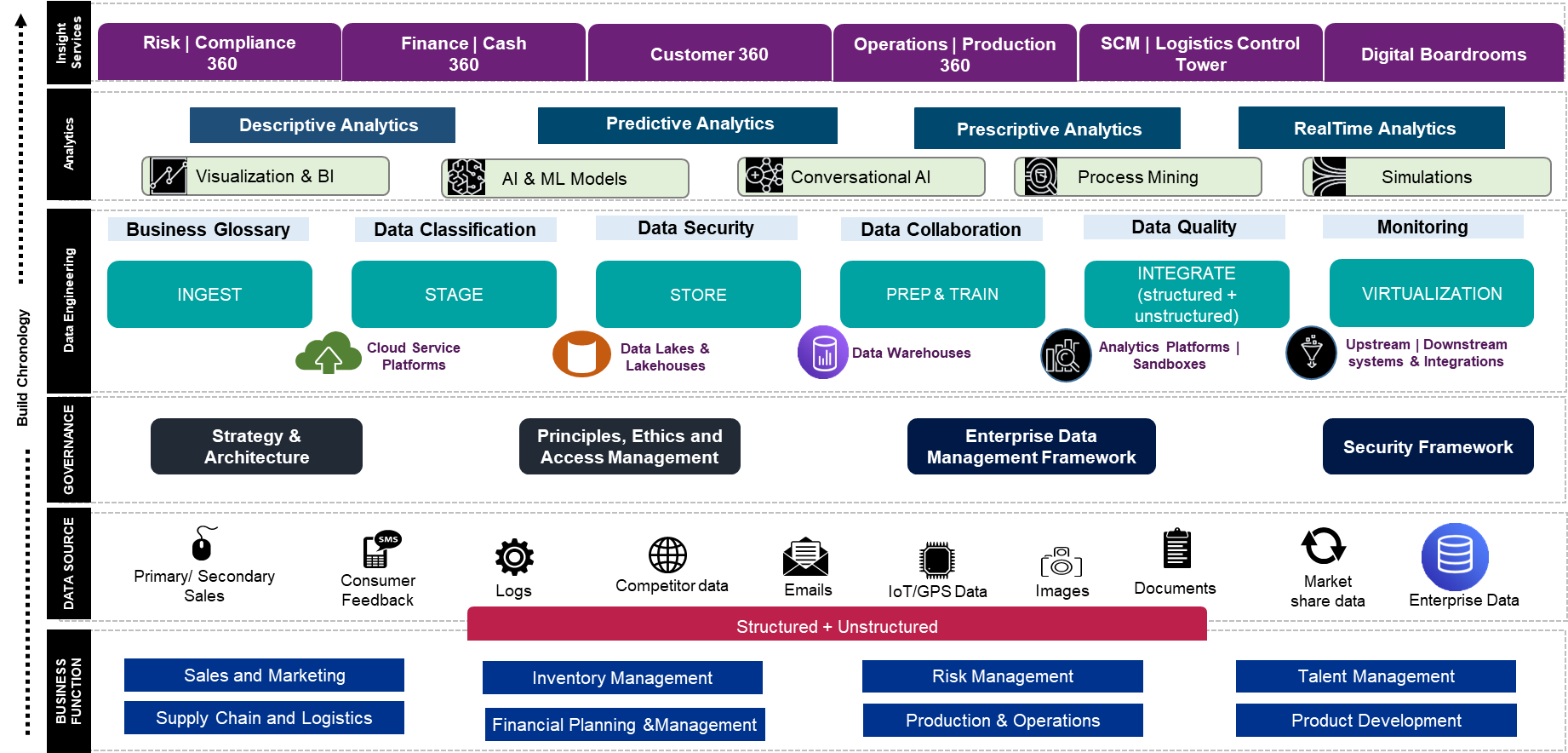
Banks hold valuable large datasets with high eminence that are sourced from multiple data sources for analytics. Banks are limited by the IT Infrastructure and IT team to provide resources for Data Scientists and Analysts for performing these analytics. With entire world stepping towards identifying themselves as data-driven organizations, banks are well inclined to perform advanced analytics for key decision making. Banks would need to embrace cloud Computing environments that are highly secured, encrypted and protected for enabling themselves as data-driven organizations. Azure is one of the leading cloud solution providers that provides a secured, restricted, and highly available environment for banks to leverage and adopt.

Banks can extend their on-premises data centers with Azure by building a Virtual Network (VNET) on Azure within proximity and having an encrypted & secured connectivity established between the two using Express Route connectivity. Large Datasets can be transferred over the secured network to store them on azure data lake storage/s that are enabled with private links to make sure data is and only accessible from within the Bank’s network. Storages with hierarchical namespaces can be integrated with Azure Active Directory (AAD) and enabled with RBAC and ACLs to conform on the data protection rules. Analytics and Machine Learning engines can access the data only when enabled with private links to the virtual network.

Banks are not limited for creating a single VNET on azure but can built multiple VNET’s that are network peered and connected to on-premises datacenter via the hub virtual network using Hub and Spoke network topology. Azure Data services like Synapse, Azure Data Factory, Azure Databricks that can be provisioned within the VNET or connected with private link can be consumed for performing advanced analytics and executing of machine learning models.

Modern technology and analytics play a prominent role in transforming the IT as well as data landscape of the banking industry. With Microsoft’s azure as a service, we can build a modern data landscape where the data can be quickly integrated from heterogenous sources and stored in enterprise data lake to perform analytics Additionally, data analytics paves way for banks to engage in better decision making and how big data can facilitate in extracting information quickly and easily. The benefits arising out of analytics help in driving sales, boosting retention, improving service, and identifying needs, so the right offers can be served at the right time

Below is a semantic view of the data architecture starting with functions and sources at the bottom and building upwards into the governance, data platform and analytics layers.



## Data Security on Azure

Banks can benefit from multi-layered security provided by Microsoft across physical datacenters, infrastructure, and operations in Azure. They can gain from the state-of-art security delivered in Azure data centers globally. It is best practice to rely on a cloud that is built with customized hardware, has security controls integrated into the hardware and firmware components, and added protections against threats such as DDoS.

Enterprise data governance and Digital Security regulations look for high access controls on the data lakes like GDPR for data governance and Principle of least privilege for digital security. Azure provides multiple mechanisms to protect the environment and data; below are some of the key features -

**Secured Network** – Azure virtual networks can be protected from external intrusions without any data being compromised. Network appliances can be deployed to the VNET along with network security groups (NSG) and user defined routing (UDR) to provide certain measure of network security. To enable security at high levels of stack, third party appliances can be deployed like Palo Alto, etc. Establishment of express route connectivity to on prem applications with encryption assures that data in transit is secured, protected and encrypted.

**Manage Identity and Control Access** – there are many best practices that can be implemented to have the best identity management policies. A few of them are:

* Integrate on-premises AD with Azure AD for a hybrid AD.
* Synchronize on premise AD users, groups, and other AD objects with AAD using password-hash synchronization, with an exception for high privileged accounts like Global Admin, etc.
* Enable single sign-on, MFA, and follow the principle of least privilege using RBAC controls.
* Use trusted identity model with managed identities and service principals for azure resources and application authentication.

**Safeguard data –**

* Data that is stored in ADLS Gen2 is encrypted and decrypted using AES 256-bit encryption. Data is encrypted using TLS 1.2 whilst in transit. The storage can be enabled with azure private links to have a secured connectivity with PaaS services. Disk encryption can be performed for disks attached to the VM’s.
* Multiple authorizations mechanisms are provided by Azure like access keys, SAS tokens, RBAC, ACLs. Considering storage to be DFS compliant with hierarchical namespace enabled, fine-grained access controls can be provided both at role-based access controls at azure resource level and POSIX style permissions on files and folders with ACLs.
* Data in the relational or MPP databases can be encrypted and masked. For fine-grained permission, Microsoft supports both row-level security (RLS) and column-level security (CLS). SQL and AAD authentication are supported along with option to enable MFA for studio access.

**Key, secrets, and certificate management** – Keys and certificates can be stored in the Key Vaults. It is designed to support passwords, credentials, certificates, and API keys with RBAC and Key Vault access policies for access model.

**Get centralized visibility and prevent attacks** – Azure provides security center that allows teams to check for misconfigurations, identify and remediate risks. Azure sentinel provides single solution for threat visibility, threat response, proactive hunting and alerting. All the Azure services are enabled with auditing and logging to azure monitor along with alerting and notification mechanisms. Dashboards can be built on log analytics workspace and can be further analyzed with azure data explorer.

## Multi Data Lake Ecosystems

Banks can build and extend their data lakes on Azure. They can either build one or more data lakes with minimal effort and time. Each data lake can be deployed on independent spokes to provide a clear separation of data domains. Having one centralized data lake for an organization like that of a bank which have multiple domains and diverse set of data available from multiple data sources is not a recommended approach.

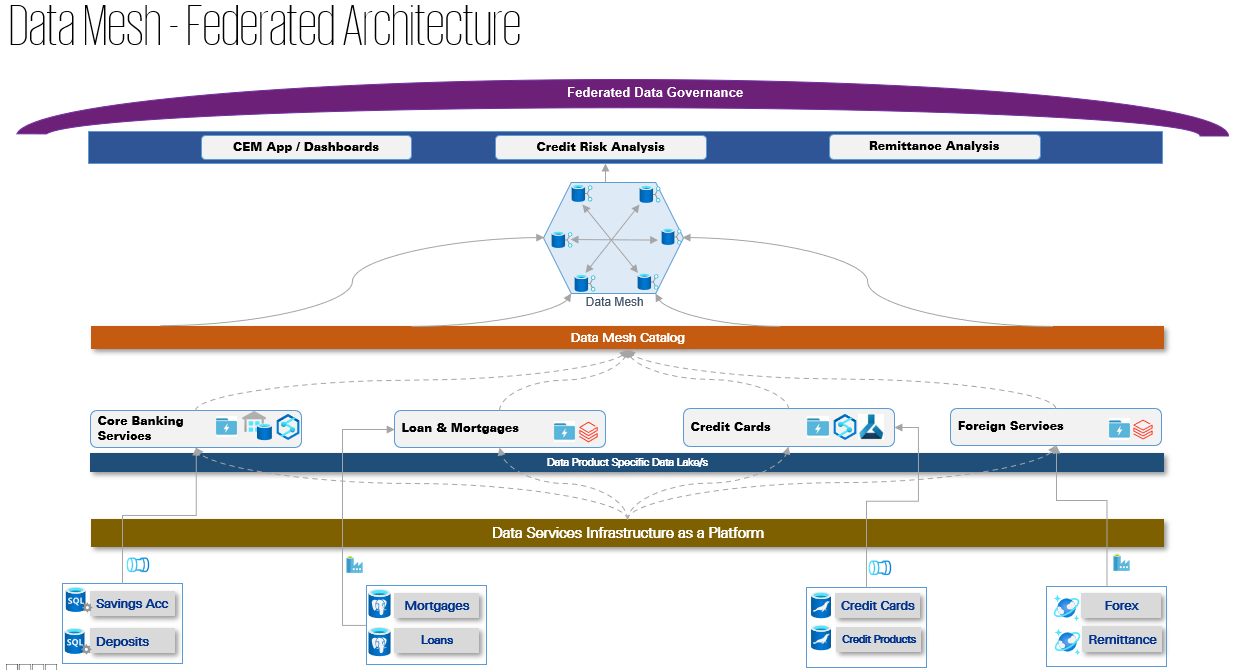
Domain experts are aware of the value that can be brought to the organization by providing best customer experience based on the data and hyper-personalization. Employees have the zeal to incorporate new methodologies by building various trend analysis with BI platforms.

Data for a specific domain would be owned and managed by specific domain managers or teams. Moving towards data domains would eliminate redundancy and improve the quality of data stored. Data storage has become more affordable, and data being stored in raw formats, might not require detailed architecture. Low-cost storages can be used for machine learning and analytical purposes.

This can be achieved by building domain expertise and domain specific data platforms. Multiple data lakes reduce data redundancy and scaling issues for analytics that impede the monolithic data lakes. Domain experts can provide discovery and cataloging of their data which reduces the effort of analysts to only extract missing data.

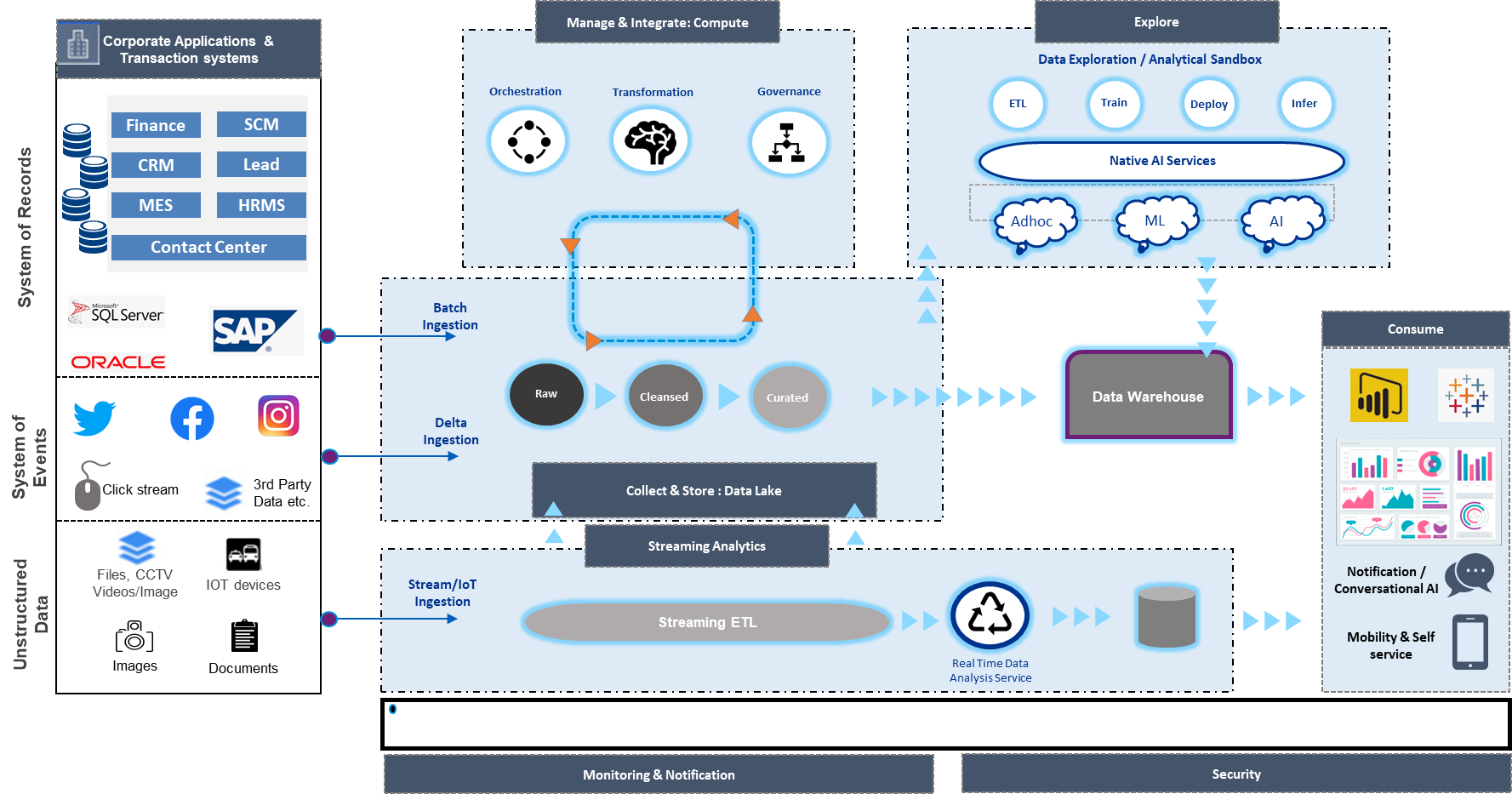
Existing monolithic data platforms lack feasibility to leverage innovative offerings to gain control of data and understand its value towards organizational growth, which indeed is critical for banking. Most of the organizations are now inclined towards data democratization to become data-driven organizations. The monolithic architecture increases complexity for individual business to segregate domain specific data, to structure their specific business purpose models for analysis. This disparity en-routes to the lack of proper data governance, metadata management and leading to compromised data quality.

Enter Data Mesh architecture and these challenges can be addressed easily. This architecture provides a strategic shift from centralized to federated ownership. Moving to a data mesh architecture can be a significant shift for business and IT. In a strategic and cost-effective manner, a data mesh architecture can help bridge the gap between people, process and the system as well as establish federated ownership and governance of data.  This can be architected, designed, and implemented leveraging Azure cloud technologies with ease.



With the recommendation towards federated data architecture, Banks need to be cognizant that each data lake has data specific to their domain and if there are analytics use cases that require cross-functional / cross-domain data – it would surely have data integration requirement which will lead to data gravity and its implications (along with increased data movement costs). Would have a spike in infrastructure cost, which should be managed by having the Data service Infrastructure teams to adopt centralized resource provisioning using Azure DevOps and Infrastructure as Code (IaC) with Bicep or ARM templates.

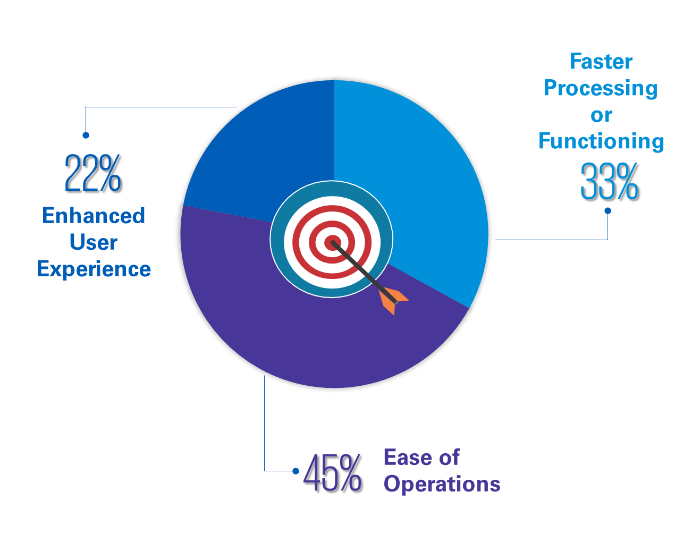
Enterprises must be vigilant that even after data movement the ownership of the data still holds with the source data teams or data managers. Colocation of data, proper writeback mechanisms, maintenance of data integrity, all of these can be managed with proper data management and governance policies. Azure services like Purview can aid to build the right data management and governance policies. With increasing amount of data revolving around an enterprise to perform analytics in data-intensive platforms like artificial intelligence, internet of things and data analytics, data gravity has become a concern in the recent past. Data cataloging is key to expose the valuable information by using data gravity index. This can help us understand how the data is growing, number of data sets being exposed and the number of applications consuming the data. Data gravity requires a connected community approach between enterprises, connectivity, cloud, and content providers at centers of data exchange to remove barriers and unlock new capabilities.



Above is a representative architecture of the Azure Data ecosystem. This is an illustrative architecture of what makes up an individual data lake within the data mesh architecture represented above

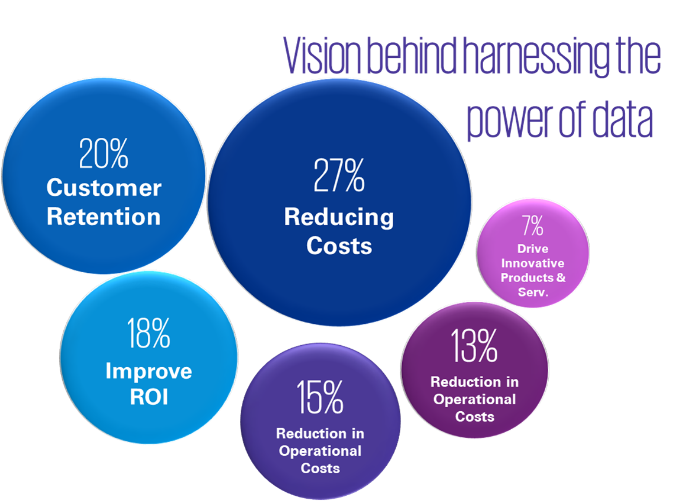
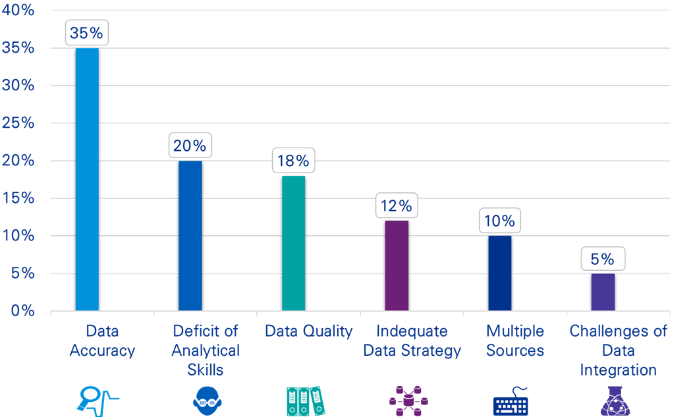
Having established the future of banking, drivers for cloud and Azure nuances, you would now be able to apply these principles to your requirements and your organization. We would also like to throw some light on how other practioners in this industry have been going through their data journey. Next section gives you a glimpse of the same through a survey we conducted.

# Banking Industry Insights – KPMG Survey

With KPMG’s depth of experience in the financial sector, we were able reach out to 80+ banking professionals from Indian and International banks operating in India to gather their thoughts and ideas around 3 major areas - Data Vision, Data Strategy and Data Technology. The survey was taken by professionals who were data owners, data curators and data consumers; and some interesting insights were observed which have been captured below.

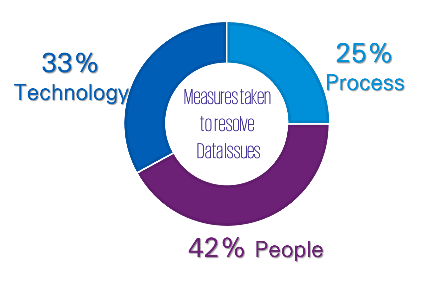
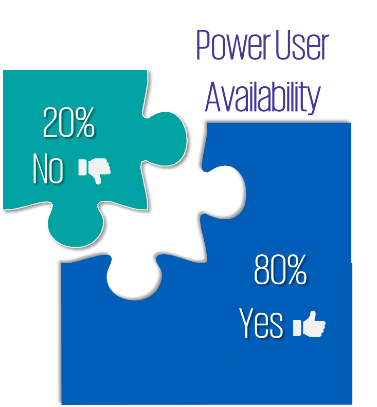
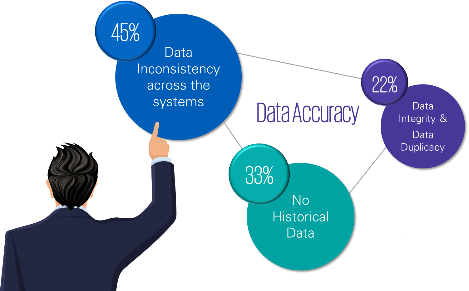
Firstly, when asked about their organizational goal with the data as the backbone i.e., how they would expect data to assist the organization’s goals, this is what we heard.

**Data Vision**

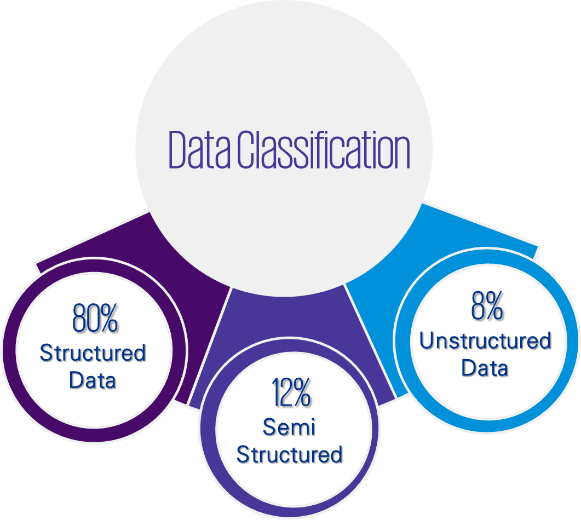
Banks are often of the belief that data can play a dynamic role in driving innovations in products and services. Data also can reduce operational costs and indeed remains a power asset in driving business performance. These graphs depict the on-ground point of view and vision of an Indian leading bank to harness the power of Data

However, there are multiple bottlenecks which are hampering banks in effectively harnessing power of data. Data accuracy stood out to be biggest bottle neck with Deficit of analytical skills within the organization being a close second.

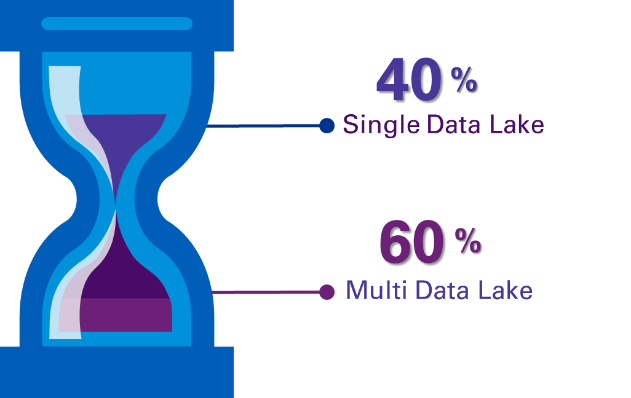
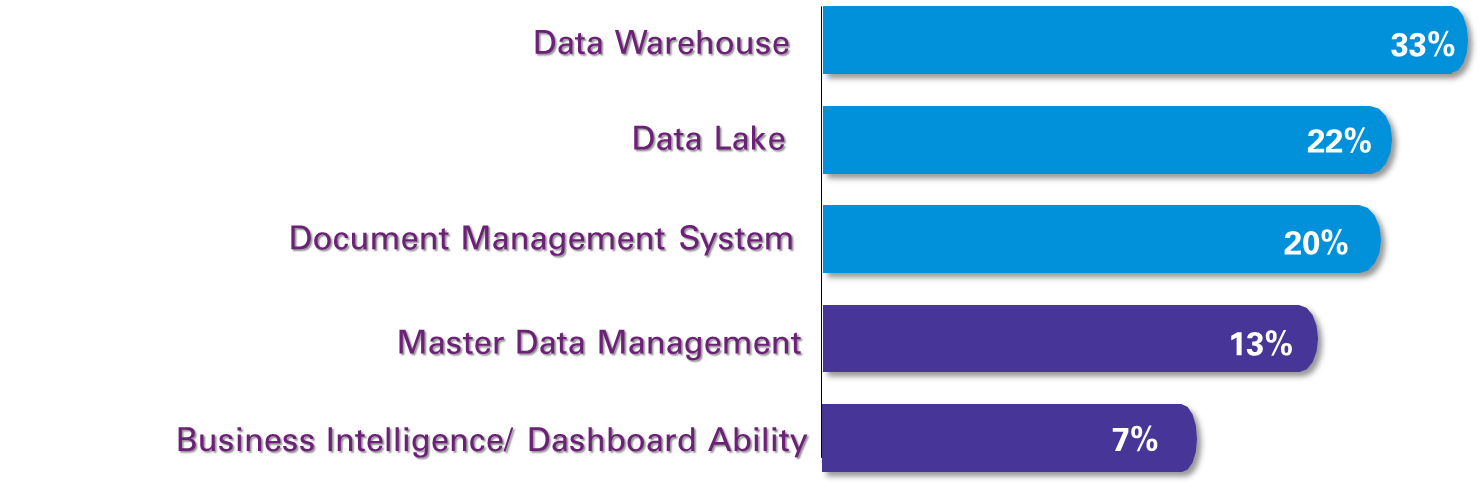
## Data Strategy

Most leading banks are confident about their growth prospects which many a times are driven by established data strategy practices and proper utilization of data. The survey features insights which emphasize on having an appropriate Data Strategy and processes as a key area in driving business growth in the next five years. These stakeholders responded that to uplift the power of the organization exponentially, focus should be given to the mining of the right data along with appropriate technology and skilled resources (data scientists and analysts).

## Data Technology

Data architecture forms the core of an organization’s digital platform as it provides the key capabilities to enable customer proposition and translates directly to the bottom-line enabling organization to be driven by data and insights. The study identified that the data architecture of the banks is either available on enterprise data lake or on data warehouse and is majorly only structured.

Study also shows that there is scope to improve things and therefore a long-term Data Strategy wrapped in an effective data governance framework will enable the appropriate access, management and sharing of data and drive the right data culture within the organization. Through the existing data assets, we noted that, mostly descriptive and predictive insights are being generated. The CDO’s (Chief Development Officer) team understands the use cases, provides infrastructure and platform for data digital transformation implementations.

Also, when questioned about data lakes and applications to be implemented in their organizations, the responses were as follows. 40% of the organizations having a data lake were on a single lake while the other 60% had seen the need for multiple lakes. And the graph on the right showcases the applications that organizations would like to implement as their next IT investment.