Tech Mahindra

issue 1

Changing Landscape of BI & Analytics

Gartner

Executive Summary

The data to insights journey has witnessed a sea change over the last couple of decades. From databases, to warehouses, to Lakes, to Cloud; the appetite for data has increased voraciously along with technology transformations that have reduced the cost of storing data and that which have increased the extracted value by effectively harnessing data.

This voracious appetite for data brings in its wake challenges of a nature hitherto unencountered. What to consider and what to leave behind? There is an ocean to churn but impractical. Today, it's not just about the data that the enterprise generates but data generated by all stakeholders in the business ecosystem. Stakeholders, meaning all those entities that can influence the business. This could be economical, political, social or cultural. B2Bs are slowly turning into B2Cs. Continuous production as we know it may cease to exist. There is discretization in almost all aspects of the value chain.

In such an environment, how do we create differentiated value propositions for storing data and for extracting insights from it? With Artificial intelligence driven by Machine Learning (ML) and Deep Learning (DL) being the toast of the day, how can we create a data ecosystem that is self-learning, self-sustaining, self-replicating and that which continuously adds value?

Industry standard algorithms for predicting weather patterns, actuarial activities, sales forecast, supply chain efficiencies etc. have been in existence for quite a while. If we assume predictions and prescriptions to be of sinusoidal nature where the crests are parameters that we are well aware of and the troughs are areas where we make assumptions; today's scenario provides for reducing the depths of the troughs and increasing the effectiveness of the crests. Keeping the sinusoid as a reference, one could go about looking for data that is relevant to solving a problem based on filling out the troughs. In essence, the problem of churning the ocean gets reduced in scope to churning those Lake waters which could yield meaningful insights.

The ability to utilize mathematical distributions to solve problems of a complex nature have taken a whole different dimension with the availability of data. The ecosystem may change in the time duration between taking decisions and to implementing them rendering the best conceived projections meaningless. There was nary a thing that one could do in the past to address this problem. Mathematical distributions on real-time data from IOT devices help bridge that gap by bringing in an element of variability in the decision-making process. Data scientists look at data for their causality, randomness and utilize time-tested algorithms to take decisions in real-time. Today we are aware that data becomes stale in a matter of hours if not minutes. Taking decisions based on static data is passé. The winning combination seems to be to promulgate an ecosystem where data is curated, made available on demand, tested for its veracity and economized for its potential value.

If Democracy of a nation is where the govt. is for the people, of the people and by the people. Democratization of data is not different. It should for the people in the organizational ecosystem, of the people in the organizational ecosystem and to be built by the people in the organization.

The ensuing pages deal with how highly-available, democratized data powered by AI and ML can be an asset to any enterprise. As a student of data science, it behooves me to quote Robert Frost who couldn't have described the role of a data and analytics person any better.

The woods are lovely dark and deep
But I have promises to keep,
And miles to go before I sleep,
And miles to go before I sleep.

Krishna Thiagarajan
Global Practice Head, Data & Analytics

Source: Tech Mahindra

Traversing BI & Analytics Journey Along With AI and ML

Alongside the focus on generating prescriptive and cognitive insights, Tech Mahindra has been focusing keenly on “rapid data onboarding” triggered by data proliferation.

Leveraging today’s enormous computational capacity and using advanced statistical methods, Tech Mahindra is fostering the use of “Data TO GO” paradigm.

It is engaged with large enterprises on a number of data ingestion projects, which requires both Cloud-based Pipelining as well as creation of Big Data loads.

A host of different architectural patterns such as **API-based architectures and MICROSERVICES-augmented** data ingestion approaches are being increasingly used.

Image recognition market, which is slated to be $77 Billion by 2025, is another area of focus at Tech Mahindra. Advanced methods in Life Sciences to automatically decipher medical images and perform cell counting are being developed at our labs.

Several projects on NLP along with NLG, Conversational Analytics are being pursued and these would be the cornerstone for more advancements in Cognitive Computing.

Enabling enterprises to unleash the power BI & Analytics

A key trend in modern-day Data and Analytics is **“pervasively sourced”** AI and ML algorithms across multiple business functions of an organization, third parties and other vendor products. In many of its recent implementations, Tech Mahindra has been able to enable benefits through its cutting-edge solutions augmented with Advanced Analytics and Machine Learning.

- Enabled **20% reduction in TCO & 10% reduction** in unscheduled maintenance for a US Manufacturing major. One of our key Manufacturing customers in the US, where Tech Mahindra, leveraging Advanced Analytics tools, implemented a solution to monitor functionality & health of system by installed sensors in the process enabling the above business benefits. This involved data emanating from Machine-to-Machine conversations and extensive usage of ML & IOT technologies

An equally disrupting trend is the operationalization of algorithms forcing extensive use of analytics marketplaces and **“monetization”** of algorithms, which has thus far been limited to Data Monetization. Gartner says that;

> “By 2023, 25% of application providers will generate additional value from their customers’ application data, up from 5% today.
>
> By 2021, 25% of organizations with public APIs will have discontinued or rebooted their public API strategy due to security incidents or lack of uptake.”[1]

Tech Mahindra’s in-house solutions around Data Anonymization has enabled our customers in monetization of their data by anonymization and masking to address the data security mandates critical for monetization initiatives.

- An important implementation for enabling Data Monetization has been for a large European TELCO wherein Tech Mahindra’s Anonymization platform was implemented to de-identify the subscriber’s data in compliance with European Privacy and Data Protection Laws. It helped process ~ 250 Million records daily i.e. 500 GB of information, applying complex Data Security algorithms like HMAC, SHA2, K-Anonymity, etc.

Tech Mahindra’s in-house, **proven and certified solution offering, PRISM**™ that is a Self-service Predictive Modeling engine and Data Science Workbench, is aligned to address this trend. In addition, this can be tailored for specific Analytics needs of the customers. Collaborative workflows and model governance are gaining increasing importance and hence many of our projects leveraging governance of Machine Learning notebooks through tools such as DataBricks.

- As a part of its PRISM implementation, one of Tech Mahindra’s footprints includes the enablement of 7% increase in Customer visits for a European Retail major; this has led to a commensurate increase in sales.

[1] Gartner Inc., Monetize Your Customer Data, but Do It Without Becoming a Headline, 26 October 2018, G00375565

- Another PRISM implementation for an Energy major in Africa featured an implementation of Revenue Assurance leveraging SAS' Fraud Analytics module enabling prevention of Revenue leakage

In alignment to the modern-day preferences, Tech Mahindra's Analytics practice has a focused CLOUD Analytics competency closely involved in implementation of Analytics on the CLOUD. There is a healthy ecosystem of niche CLOUD analytics partners with whom Tech Mahindra teams up to offer state-of-the-art, accelerated Cloud-based Analytics solutions and value-adds to our customers. This includes Analytics implementation on the CLOUD, migration of ON-PREM solutions to the CLOUD, analysis and proposing strategies for CLOUD Analytics implementations.

- One of the key engagements of Tech Mahindra on CLOUD Analytics has helped us enable a Net Savings of **$35.67 MN** for a Global Healthcare major. It is for one of our key strategic customers in Europe where our footprints included design and build of an Analytics platform on MS AZURE along with Cloud Migration, formulation the Linear programming to provide optimal Spend plan with Maximum Net Profit.

Tech Mahindra as an SI has leveraged the **collaboration across the spaces of Analytics, Artificial Intelligence and IOT**. Ensuring a healthy and enriched partner eco-system has also been a key step in the overall journey to address the modern-day demands of Cognitive technologies, Machine-to-machine conversations and the likes.

- Tech Mahindra's focus on the modern-day MFG Analytics solutions include foray into extensive IOT Analytics. This has helped Tech Mahindra enable reduction of time **from 2 weeks to 2 days** to facilitate flight efficiency improvements for a US Airlines major. For one of the leading Airlines players, an advanced analytics implementation featured the capture of Sensor data from the flights. The solution offered insights and intelligence that helped enabling the business benefits as outlined above

Tech Mahindra has an established Data Management practice comprising of Data Migration, Data Governance, Data Quality and Master Data Management. We have been involved in large multi-year transformation deals in Data Quality, Reconciliation and Governance and have been able to offer significant benefits to our global customers.

- Enabled Revenue-leakage prevention of **$300 MN** over a period of 4 years for a giant UK Telco. One of the TELECOM giants in the UK where Tech Mahindra implemented a Data Quality and Governance program for the TELCO's Enterprise and Corporate customers facilitating a total benefit of ~ 200 Mill GBP over a span of 4 years yielding more than 600& ROI

In house solutions driving the change

PRISM™ on the other hand is Tech Mahindra's Predictive Analytics and Data Science Workbench focusing on democratizing Analytics and is augmented with Data Science and Machine Learning components.

As a solution, PRISM™ too is certified by HORTONWORKS and for its global implementations, we have observed PRISM™ to enable ~ 30% reduction to delivery for analytic services, 20% analytical resource cost savings and a consistent Proven approach and solution minimizing delivery and business risk.

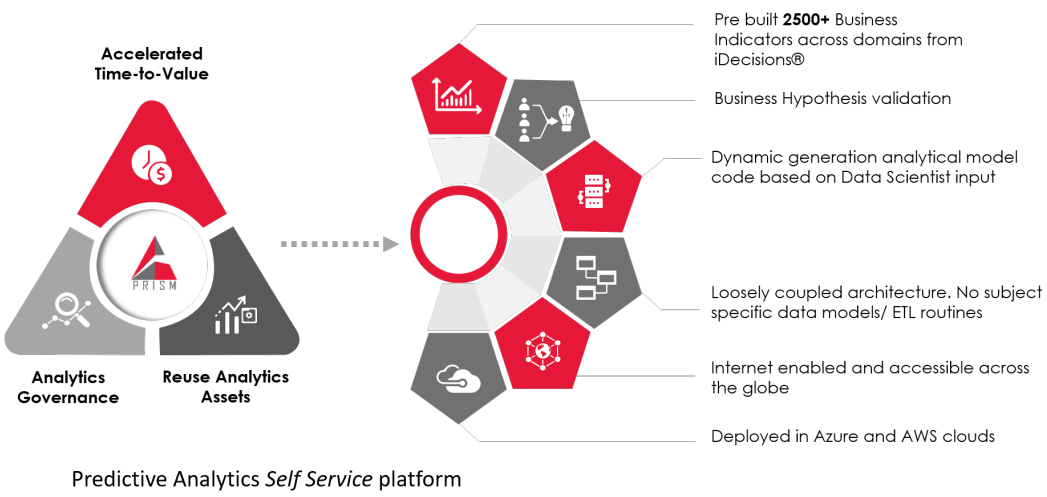
In its recent engagements with few of its strategic customers, Tech Mahindra has focused heavily in enabling DATA-AS-A-SERVICE and ANALYTICS-AS-A-SERVICE as a part of its implementation.

For one of our large customers in Middle East and Africa, Tech Mahindra has implemented DATA-AS-A-SERVICE leveraging one of our partner Big Data Platform solutions. This is a multi-country engagement and we are implementing and enabling DATA-AS-A-SERVICE for their multiple outputs in a phased manner heavily leveraging the comprehensive data-platform built on HADOOP adopting a LAMBDA architecture.

For an equally large customer in Western Europe, Tech Mahindra has offered ANALYTICS-AS-A-SERVICE, which is designed and built atop the data platform. The ANALYTICS-AS-A-SERVICE platform offers seamless self-service driven predictive algorithms and integrates with the Campaign Management systems sharing insights and intelligence with it for improved and more effective campaigns.

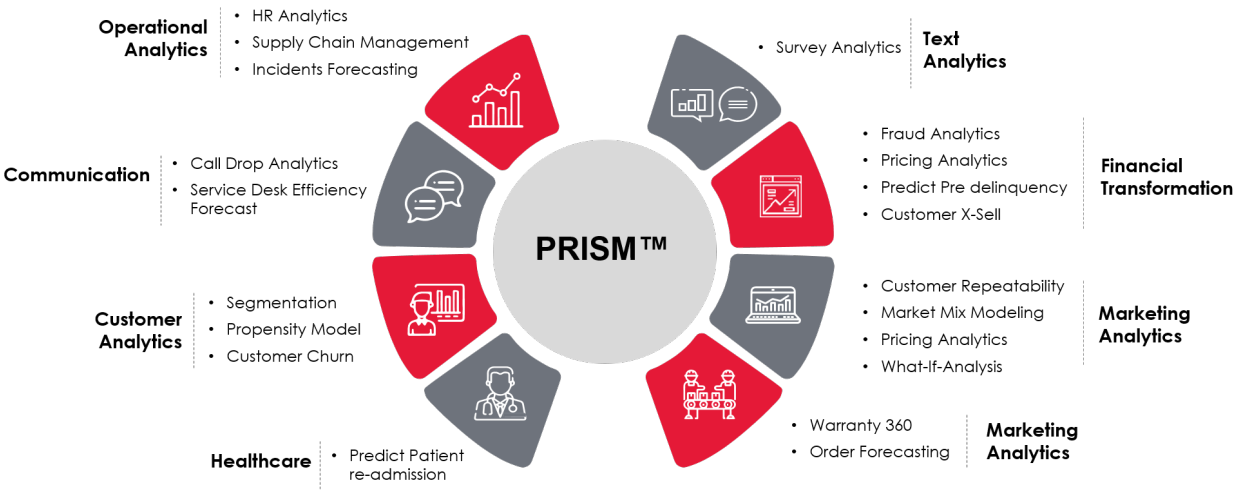
Such initiatives are increasingly becoming a business-as-usual and mundane reality where the Analytics engine communicates and interacts with the E2E systems of the modern enterprise to drive value across the chain. Tech Mahindra has been increasingly involved in such engagements for its global customers which are seeing increasing inclusion of areas around Cognitive Analytics and Neural Networks.

figure 1. PRISMtm - Data science work bench



Source: Tech Mahindra

figure 2. PRISMtm - Pre Built Solution Framework

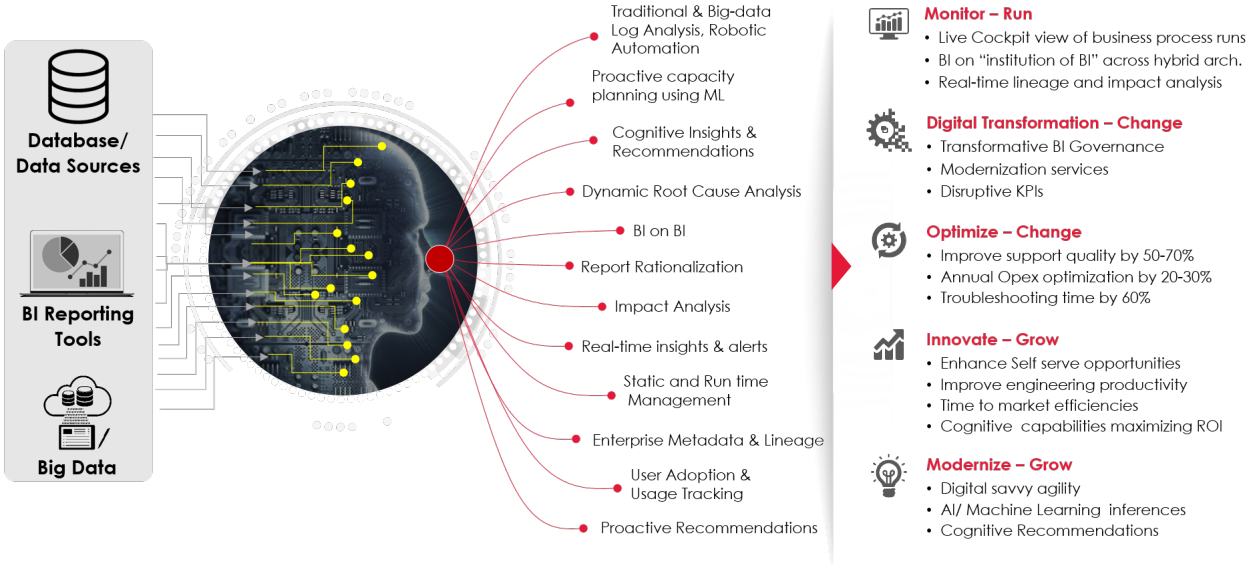


Source: Tech Mahindra

Apart from PRISM™, the other in-house Solution offerings that is aligned to the modern-day customer asks is INFOWISE.

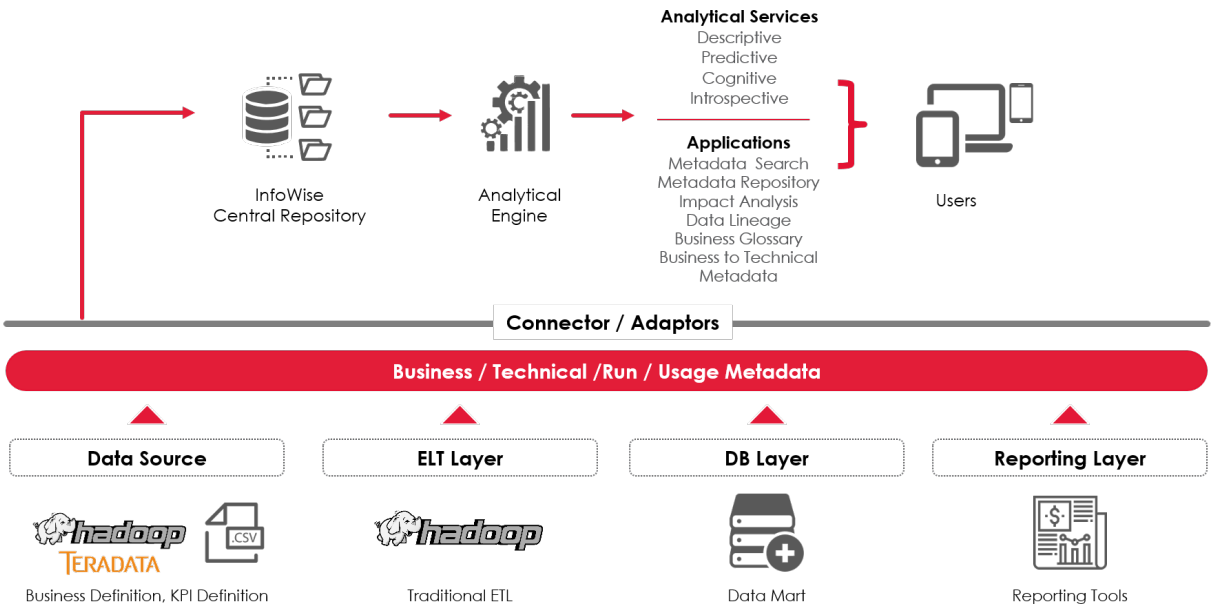
INFOWISE™ is Tech Mahindra's solution for predictive and cognitive BI Run-Ops and Governance. This is a Proven and certified AI-augmented BI Run-Ops solution, which proposes to cover ALL the applications, mentioned. The solution is certified by HORTONWORKS and focusses on the Automation of Root Cause analysis and troubleshooting leveraging a cockpit view of BI process in the form of dashboards. It also offers Pattern identification capabilities which aids in for proactive performance improvements and ensuring predictability to BI Run-Time processes & alert mechanism in conjunction with SLAs.

figure 3. InfoWise™ – Intelligent BIOps WorkBench



Source: Tech Mahindra

figure 4. InfoWise™ – Functional Flow



Source: Tech Mahindra

Implementation of INFOWISE™ for its various global customers has witnessed an improvement in support quality of a next-gen BI estate by 30%, been able to facilitate optimization of OPEX by 15-20% and save time and efforts for analysis and troubleshooting by 40%.

Source: Tech Mahindra

Striding Into the Future With Analytics

Analytics for “Change”

One of the big leaps that BI and Analytics has taken over the years is in their ability to get closer and closer to “next best action” relevant to business processes. Historic analysis has never been as pervasive as today and part of the reason could be attributed to the fact that the “actionable” inside Actionable Intelligence has been rather retrospective and discretionary on whether they are brought back into fine tuning business processes.

Embedded Analytics is fast bridging this gap through inherent integration of analytic content and capabilities within business process applications.

As much as MICROSERVICES driven architectures have brought in loose coupling, enterprises are increasingly looking at analytics automation embedded within their corporate application.

As SI collaborates, we are observing increasing demand from customers to provide additional analytic awareness of transactional systems to enable informed decision making at every step of the process. Consider claim adjudication in the context of Healthcare Insurance, for example.

The ask is not just to adjudicate but also to provide native intelligence within that process to detect fraud and not as an aftermath of the data collected by the process. Even predictive analytics alone cannot drive this kind of agility in a transactional system.

Telecom, BFSI and Healthcare Life Sciences, in our opinion, are the largest beneficiaries in this area. One of the things, which our customers want to achieve through embedded analytics, particularly in these domains, is the ability to generate an auto alerting mechanism for highly regulated processes.

The key is also to leverage data proliferation effectively in generating meaningful actionable intelligence that can be used to autonomously make decisions without human intervention (in some cases).

Tech Mahindra is currently implementing embedded analytics applications in the area of Pharma commercial business processes. These use cases are tailored towards addressing marketing automation and driving more informed offer generation. The key challenge in such initiatives is however to effectively use well rounded datasets and connected applications to incorporate closed loop insight generation and improved learning over time.

At the *EDGE* of network

As devices continue to grow across almost all verticals edge analytics is fast catching up to provide the much needed scalability and automation in the analytics world. In Manufacturing, for instance, Industry 4.0 is propelling the need to respond intelligently and quickly to changing factory floor conditions.

Many of these use cases require ultra-low latency and provide an opportunity to accelerate innovation at a more granular and operational level. Tech Mahindra is investing heavily in this area to yield optimization benefits at different steps of a connected factory.

Medical Devices Manufacturers are also increasingly using edge analytics and considering monetizing valuable health data generated in the process.

EDGE analytics also has the ability to reduce the over reliance on Big Data or large usage of cloud compute for relatively light weight or first level analysis. The very idea of pushing as much analytical workload to the edge can bring in potential benefits through “federation” of insight generation, reducing latency and cost of communication. The analytical heavy lifting can then be performed on a Cloud Big Data lake.

The challenge of the 80% ahead of analytics

It is well known that bulk of the effort spent in a typical analytics or machine learning project can be attributed to data engineering and data preparation. Apart from the effort that goes in, this is also amongst one of the reasons why businesses are still not completely empowered and analytics is still not as pervasive across all business processes. We believe Augmented Analytics will bridge this gap in the coming years.

Traditional BI platform vendors are under pressure to remain relevant in a market where data discovery is driving most net new buying. Smart Data Discovery is the fundamental building block, which can make analytics truly agile. A number of use cases, which we are implementing with our customers, are data led rather than incepted with a firm business context in mind. These then mature on to full-blown projects with a complete business buy in on their tangible outcomes. To drive more of such use cases what we really need is a mechanism to identify patterns, perform sophisticated relationship analysis across data elements.

Much of this, if automated within data discovery tools, will accelerate the data and insights journey.

While data discovery tools have existed since long, Tech Mahindra wants to differentiate by bringing in the power of hypothesizing to business stakeholders. We would like to automate data preparation and manipulation leveraging Machine Learning and Cognitive Intelligence. Furthermore, we would like to make it accessible within Tech Mahindra frameworks such as PRISM and this is an area where we are increasingly focussing on. Accessible Augmented Analytics will assist Citizen Data Scientists and improve accountability and empowerment.

Covering the last mile!

The Transactional and the Analytical. Two worlds, which featured distinctly different types of systems with different data operations not so long ago. The Transactional systems featuring primarily data ‘manipulation’ operations whilst the Analytics systems mostly data ‘select’ operations.

Most often, the Analytical systems were insulated from the Transactional ones lest their operations adversely affect each other. A dividing line or an iron curtain hence prevailed between the Transactional and the Analytical world.

This iron curtain is vanishing gradually and the dividing lines are being rapidly diluted. The interactions between the Analytical and Transactional worlds are on the rise. Whilst the Transactional systems have matured consistently, the role of the Analytical systems have undergone a radical transformation. A metamorphosis in the least from a mere “eyes and ears” of an enterprise to its “intelligence and insight”.

Hence the emerging importance of covering the “LAST MILE”!

The step where the Analytical engine offers the insight to the Transactional systems for facilitating improved data-enabled transactions. The step where the Transactional systems like Campaign Management or Service Delivery Assurance are able to reap the benefits from the underlying Data and insights which the Analytics engine enables and equips them with. Very advanced and niche aspects of Analytics like Natural Language Generation (NLG) are used to good effect for crossing the last step.

This, in essence, is the LAST MILE which is yet to be a step to be fully conquered as the Analytical systems the world over are in a fit of struggle to seamlessly share and inject insights and intelligence into the Transactional stack.

This step is, in essence, where the benefits of an Analytics engine is realized. A step, which many of the large Customers, Corporates and System Integrator alike are heavily scaling up the path.

Source: Tech Mahindra

Conclusion

The modern-day analytical tools being more powerful and less expensive than what were previously available offers an advantage to bring forth enhanced value-for-money to the customer's table.

Empowering customers to **store and analyze data in significantly larger chunks and far wider varieties and types**. The number crunching, the analyses and associated recommendations enabled are far quicker, at a much lower latency and in most instances, real time.

With the growing importance of Analytics especially across the last decade, the large corporates are increasingly attributing their **decisions and drivers to informed, intuitive facts and data;** dependency on Prediction and Forecasting has increased manifolds, especially so in the past 5 years. The journey has not being restricted to mere **'Prediction'** and the expectations have evolved into the usage of Analytics in a **'Prescriptive'** capacity as well.

We are at a juncture where Analytics' role is increasingly considered as a conduit to infuse insights and intelligence into the organization's **transactional and operational eco-system**. Into its core operations and system orchestration.

As a prominent differentiator and influencer to the three key objectives of large enterprises, **TCO reduction, Revenue enhancement and improvement of Customer experience**.

Looking back where Data & Analytics stood as a decade and a half ago, this is indeed a transformation of royal proportions

A complete metamorphosis of its role, its expectations and its perceived value in the eco-system of an enterprise.

The gradual transition of Analytics from a 'diagnostics' to a 'predictive' to a 'prescriptive' to an 'intuitive' engine is evident. A path where it not only traverses across its milestones of maturity but also is at its cusp of its transformational journey from "BI to AI".

A juncture where Pattern recognition, Reasoning and Interaction, Human intelligence, Conversational and Cognitive Analytics gain increasing prominence and areas where Analytics seems to be having increasing expectations as a domain.

As the world gravitates towards connectivity and convergence, Analytics' role as well becomes increasingly pervasive, stepping into the areas of Artificial Intelligence, Augmented Reality, Machine-to-Machine conversations.

We have embarked upon a journey constituting the Man, the Machine and a host of algorithms on the CLOUD to predict, preempt, recognize, learn and interface. A cognitive and connected journey with Analytics at its focal point, assuming the pivotal role.

And for certain, the journey is only to continue and evolve much further!

"Emerging technologies have always been at the core of Tech Mahindra's business service offerings. We harness the power of emerging technologies like BI & Analytics to devise business and technology roadmap of our customers. BI & Analytics have emerged as DNA of business transformation across the Enterprises. We at Tech Mahindra are enabling this transformation by designing business strategy, analytics & AI strategy, technology architecture and executing those initiatives whether it is revenue assurance, cost optimization, customer experience redesign, reduction in time to market, or ideation of new revenue models."

Krishna Thiagarajan
Global Practice Head, Data & Analytics

Source: Gartner Research, G00375565, John Santoro,Eric Hunter, Bart Willemsen, 26 October 2018

Monetize Your Customer Data, but Do It Without Becoming a Headline

Customers' application data offers a source of innovation for your products and a source of value for your business. When tech CEOs monetize this data, they must minimize the associated legal and reputational dangers.

Key Challenges

- Opportunities to monetize SaaS application data can be appealing, but providers must do so without betraying customer agreements and privacy expectations.
- Providers must offer valuable APIs to encourage ecosystem participation, but participants could abuse the APIs and harm the provider's reputation.
- Privacy scandals and maturing privacy legislation, such as the GDPR, increase both customer expectations and legal requirements for data protection.

Recommendations

To strengthen their product strategy, tech CEOs should monetize customer data by doing the following:

- Deliver customer peer group comparisons or machine learning-enabled predictive insights by generating anonymized, pseudonymized or aggregated application data.
- Do not rely on an assumption of only "good actors" using APIs. Protect against worst-case scenarios by controlling data and API access.
- Provide flexibility and adaptability for emerging legislation and consumer rights surrounding personal data by engineering solutions with data suppression capabilities.

Strategic Planning Assumptions

By 2023, 25% of application providers will generate additional value from their customers' application data, up from 5% today.

By 2021, 25% of organizations with public APIs will have discontinued or rebooted their public API strategy due to security incidents or lack of uptake.

By 2021, organizations that bought privacy risk and are caught lacking will pay 100% more in compliance cost than best-practice-adhering competitors.

Introduction

This research focuses on how to ethically and safely take advantage of the opportunity provided by customer data. To complicate these efforts, however, Gartner clients around the world are facing maturing privacy and data protection requirements.[1] This document is adapted from "Beyond GDPR: Don't Let Privacy Concerns Prevent You From Ethically Monetizing Your Customer Data," which contains technical details better suited to technology product managers.

Although many technology providers have successfully migrated on-premises applications to SaaS, many still do not fully leverage the capabilities made possible by SaaS architectures. For multitenant applications, a pool of customers' application data offers an appealing resource that technology providers can use to the benefit of all parties. For example, providers can use aggregate data to train machine learning models more accurately than if a customer relied solely on its own data. Another capability that customers may value is the ability to compare results against those of their peer groups or competitors.

Gartner's research on "infonomics" can help vendors better understand how to value the data collected.

Technology providers can leverage this data by adding features to their application and by creating APIs that expose the data to partners or a paying public. A 2018 Gartner survey[2] found that 19% of responding end-user organizations currently monetize APIs directly or indirectly, with a further 28% planning to monetize APIs in the future. It should be noted that not all API monetization involves data monetization through data-as-a-service APIs. It could involve selling usage of an algorithm through an API, for example. However, most API monetization takes the form of data monetization, since many APIs are built for access to data. In fact, the same survey found that data accessibility is one of the top three business goals or objectives organizations address with APIs.

Our survey also found that API capability matters for API monetization. Organizations rating their API capability level as intermediate/advanced or expert are more likely to monetize (32%) or plan to monetize APIs (30%) compared with organizations with lower API capability levels (8% currently monetize and 24% plan to monetize APIs).

Analysis

Deliver Anonymized, Pseudonymized or Aggregated Customer Data

For many applications, customers may be curious to know how their organization's performance compares with the competition. Direct access to this competitor data generally is not feasible and would not only erode customer trust, but also potentially violate legal restrictions. However, through anonymization, pseudonymization3 or the aggregation of customer data, application providers can offer customers benchmarks without jeopardizing privacy agreements and violating laws.

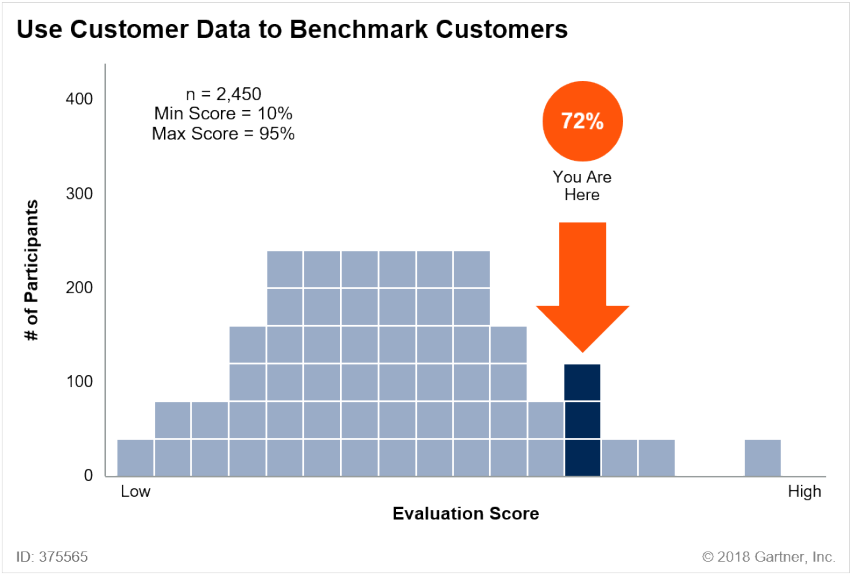
For example, an application that manages courses and students could allow customers to compare with similar customers their catalog size, percentage of courses filled and repeat customers (see Figure 1 for an example). Some financial technology (fintech) companies offer aggregated data that allows lenders to build risk profiles. Since the data required to provide this information does not reveal who the other customers are or the particulars of their data, technology providers would not violate their privacy.

Beyond peer group benchmarking, customer data can help with machine learning. Supervised machine learning relies on large amounts of training data to refine the models the process produces. Data from a single isolated company or consumer may be insufficient to produce an accurate model. However, by using customers' anonymized data, technology providers can produce more relevant models (anonymous, because the same privacy rules apply to test data management).

This process does require careful consideration and planning. Product teams must ensure they account for adequate time in both the early design and delivery phases. A Carnegie Mellon University study demonstrated with a high level of success that even data thought to be anonymized could be used to determine a person's identity with the use of a five-digit ZIP Code, gender and date of birth. Beyond this example, there are specific considerations that should be applied relative to characteristics of a specific population for which data is compared. With this in mind, ensure the appropriate review by legal counsel of processing techniques and controls.

Some customers may not want to contribute their data for these use cases, as they may fear that competitors will learn about their successes or failures in the market. To encourage contributions and discourage the exchange of something for nothing, require customers to "give to get" — in other words, they can benefit from the capability only by contributing their data. There are a variety of mechanisms that can further encourage participation, such as gamification, monetary or participation incentives, and exclusive product and insight offerings.

figure 1. Example Use of Customer Data to Benchmark Customers



Source: Gartner (October 2018)

Control Data and API Access

The privacy scandal with Facebook and Cambridge Analytica arguably hinged on two API-related problems:

- **Problem 1:** Facebook's APIs provided more information than was necessary for the stated purpose — Why would a researcher need actual usernames or the names of the users' friends?

- **Problem 2:** Facebook granted access to the APIs for research, not for targeting political messages. However, once it granted access, Facebook had no way to prevent abuse through indirect uses or subsequent sharing of the data.

In many areas of law and policy, organizations rely on the assumption that people and users will be "good actors," meaning individuals will operate within the established guidelines and intended constraints of a system.

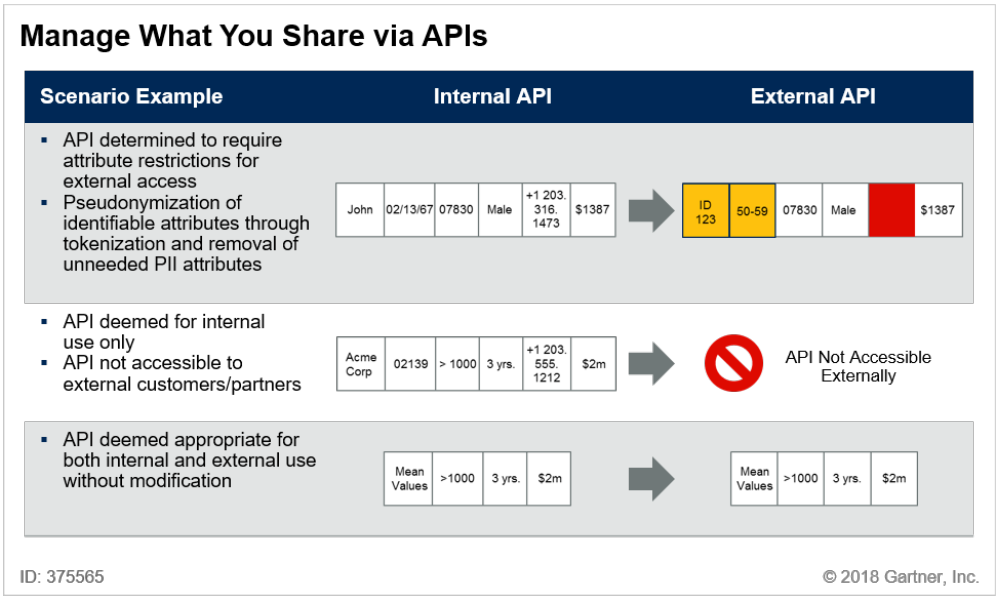
Try to encourage "good actors" by providing a review process and certification for those who use external APIs. Although the fear of detection and termination of access may deter some, it likely will not be enough to prevent willful abuse. As a result, try to prevent customers' misuse of data access by limiting access to it in the first place.

When deciding whether and how to limit data access via APIs, try to imagine the worst possible misuse of data. If the worst-case failure mode outcomes would not sound good if made public, consider confining those APIs to internal developers. Figure 2 shows how to modify or prohibit access to internal APIs for customers and partners.

LinkedIn provides an example of how to learn the lessons from Facebook. LinkedIn will provide access to its customers' application data, but only in aggregated, anonymized form. In addition, the company will vet all proposals for access to ensure that any use of data meets its ethical and legal standards.

However, technology providers will need to closely collaborate with their legal counsel and security experts to ensure appropriate contractual and logical controls preventing reidentification in subsequent processing of such data. Enriching initially considered anonymous data or deploying advanced technology (such as a deep neural network) might — given enough time — eventually lead to anonymous data becoming pseudonymous again, or even directly identifiable. Mutual responsibilities of both providers and users of such APIs must be contractually established.

figure 2. Limit What You Share via APIs for Customers and Partners



PII = personally identifiable information
Source: Gartner (October 2018)

API management allows the cataloging, managing and monitoring of internal and external APIs from a single location. API management allows companies to better control access to their data and protect customers' privacy by selecting which APIs to expose externally, controlling and tracking who and how many times they call APIs, and securing APIs to reduce hacking risk.

Organizations should also prepare for the scenario where their API may be provided through a third-party API marketplace. Ensure that appropriate security controls are in place, and that your participation in a third-party API marketplace is aligned with your goals.

Engineer Data Suppression Capabilities

Define architectures supporting API monetization to meet present-day and future consumer privacy legislation demands. Governments are drafting many other forms of legislation, so the upfront design for consumer privacy controls is a must. General Data Protection Regulation (GDPR) provides a reasonable framework against which to assume specific design use cases, regardless of the impact of specific legislation on a company. Moreover, GDPR is not the only reason. It has kick-started a global tidal wave that is leading to similar legislation's adoption by other jurisdictions.

Require your corporate legal department to review the formal architecture, design and development artifacts throughout the planning and development process. With legal review, they can ensure compliance from concept to delivery with organizational standards and international regulatory guidelines.

Layers of levels of anonymity can be implemented to ease the subsequent compliance of processing, as there must be a separation between the directly identifiable customer information and external environments. Fully anonymous reports contain zero personal data, and therefore are no longer subject to regulatory requirements or even retention periods.

Directly and indirectly identifiable personal data is governed by various regulatory privacy requirements, where anonymous data is not.

Nonetheless, the ultimate anonymous data is "deleted" data. Organizations automatically inherit the lasting responsibility to continuously assess the risk of reidentification of data. Reidentification is aided by time, the amount of data and the inevitable progress of technological advance.

Conclusion — Take Advantage of the Opportunity in a Safe and Ethical Manner

The loss of good faith with customers will far outweigh the gain of providing these dditional capabilities.

Successfully using customers' application data relies as much on digital ethics as it does on APIs. Besides complying with privacy legislation, clearly communicate to customers what and how data will be used.

Consider sharing an "acknowledgment video" to replace lengthy, complex (and unread) contracts. In the video, demonstrate which data you plan to use and the benefit that using it will provide to customers. Doing this reduces the likelihood that a customer will complain about being surprised by the implementation, and it also serves as a bit of marketing to make customers enthusiastic about the provided value.

Do not rely on "fine print" in the contracts when seeking permission. Instead, present a transparent opt-in versus overly verbose opt-out strategy with customers. The loss of good faith with customers will far outweigh the gain of providing these additional capabilities. With that in mind, an appropriately structured and incentivized model can provide mutual benefits to your business, end-user customers and customer peers through shared insights.

Acronym Key and Glossary Terms

GDPR	General Data Protection Regulation
PII	personally identifiable information

Evidence

[1] EU's General Data Protection Regulation [GDPR] and California's AB 375, to pending bills in India and Brazil

- "Brazilian Internet Law"
- "White Paper of the Committee of Experts on a Data Protection Framework for India"

- "AB 375 Privacy: Personal Information: Businesses. (2017-2018)," California State Legislature.

[2] Gartner Research Circle Survey: "API Usage and Its Role in Digital Platform Growth."

- This research was conducted via an online survey in March 2018, among Gartner Research Circle members — a Gartner-managed panel composed of IT or IT-business professionals.
- In total, 129 members participated.
- Qualified participants included business end users with either an IT or IT-business focus as a primary role. Respondents' organizations currently use API (73%); currently implement APIs (17%) or plan to implement API in the next 12 months (10%), and respondents are involved in their organization's API strategy in some capacity.
- The survey was developed collaboratively by a team of Gartner analysts, and was reviewed, tested, and administered by Gartner's Research Data and Analytics team.

[3] Anonymization renders data no longer identifiable, directly or indirectly, and is therefore no longer subject to privacy regulations. However, not in all analytics cases do such changes maintain sufficient information value. Given the options available — for example, with consent provided — privacy protection by default includes at least the option to pseudonymize that data. Pseudonyms might still be subject to legal requirements, though usage of it is less invasive and should be considered where anonymization is not possible. See, for example, regulatory guidance by the European Commission's Article 29 Working Party.

About Tech Mahindra

Having initiated its practice around the start of the millennium, Tech Mahindra has stepped successfully through the different phases of evolution of BI and Analytics. A global spread spanning various verticals, its core offerings include end-to-end services in Data & Analytics across **13 Verticals, Telecom, BFSI** and **Healthcare & Life Sciences** being amongst the TOP 3. Equipped with **7500+** Data & Analytics consultants with an even mix of Technical, Process and thought-leadership it has several BI Strategy and consulting engagements to its credit across multiple global customers.

The practice includes footprints on **60+ Technologies** with dedicated Data & Analytics Centers of Excellences for end-to-end BI landscape and competencies across all the key technologies. Credited with its impressive credentials of **circa 160 customers**, Tech Mahindra currently serves an active customer base of **135 customers**. With **35+ successful, global BI & Analytics engagements**, Tech Mahindra has consistently been enabling its clients leverage "**enterprise data**" as an asset for growth.

Adopting a differentiated approach, Tech Mahindra has a host of **proven, in-house and certified** BI Solutions that have stood the test of time on addressing the evolving demands of customers in this space. These include offerings across the **core BI & DW, Data Management, Data Science Workbench and AI-Augmented BI Run-Ops**.

Tech Mahindra

Tech Mahindra is positioned as a Niche Player in the 2019 Gartner Magic Quadrant for Data and Analytics Service Providers, Worldwide.

Tech Mahindra represents the connected world, offering innovative and customer-centric information technology experiences, enabling Enterprises, Associates and the Society to Rise™. We are a USD 4.9 billion company with 121,840+ professionals across 90 countries, helping over 935 global customers including Fortune 500 companies. Our convergent, digital, design experiences, innovation platforms and reusable assets connect across a number of technologies to deliver tangible business value and experiences to our stakeholders. Tech Mahindra is the highest ranked Non-U.S. company in the Forbes Global Digital 100 list (2018) and in the Forbes Fab 50 companies in Asia (2018). We are part of the USD 21 billion Mahindra Group that employs more than 200,000 people in over 100 countries. The Group operates in the key industries that drive economic growth, enjoying a leadership position in tractors, utility vehicles, after-market, information technology and vacation ownership.

Contact Us:

For more information contact us at: www.techmahindra.com