CII Conclave

Big Data & Analytics

Confederation of Indian Industry (CII) organized a one day conclave on Big Data & Analytics with the theme "Data Renaissance: Sculpting Tomorrow with Insights Today" in New Delhi. The conclave facilitated the knowledge sharing, industry networking and strategic discussions among senior leaders, fostering data-driven innovation and excellence.

Talking about the event, Abhishek Singh, additional secretary, MeitY said, "From catalyzing innovation, fostering fruitful collaborations and serving as a source of inspiration for advancement in the field, I am sure that this conclave has played a pivotal role in shaping New Horizons in Big Data analytics."

TelecomLive brings you the key deliberations of the event.

Big Data Analytic will not only drive job market but boost Indian economy in coming years

Vinod Sood, *Chairman, CII Regional Committee (NR) on Digital Transformation (DX) and Co-Founder & MD, Hughes Systique*

The recent development at the global level shows growth in cloud based data warehousing, real time analysis and Big Data platforms largely influencing business and intensifying the competition in this space. Together data analytics is the single most powerful catalytic for the change in the enterprises. The effective management of corporate data has grown in importance as businesses are subject to increasing number of compliance regulation. Companies are moving from business intelligence, standardization to experimentation. Data quality concerns are easing a bit and cloud based real time Big Data platforms are on the rise. This has led to an exponential increase in the number of people using these technologies and their interest in R&D activities.

Big Data has become much talked phenomenon in the business in last few years. It's changing the way decisions are taken routinely. India is witnessing high growth in Big Data analytics industry. With more and more smart devices in the market coupled with the increasing usage of internet, the amount of data churned out will only increase in the coming years. Already, India ranks second highest in the internet usage across the world. This means that with the expected growth analytics will be very exciting career for the Indians to work on. The increasing usage of Big Data is creating many jobs in the field of data analytic, not only for freshers but experienced

professionals are also signing up for online courses to improve their career prospect.

The Indian government has also understood the value of Big Data and formulated a Big Data management policy which was launched by the Comptroller and Auditor General of India (CAG). The objective of the policy is to manage the tremendous volume of data that are generated by different states and central government. For example, the government is working on the historical and current power consumption pattern in India to analyze and draw insights from it. This will help in better governance in the power sector. There are lot of high risk industry that have relevant use cases for Big Data in their daily operations.

If you see the example of banking and the financial sector, the Big Data analytic is a useful way to minimize the losses and avoid the risk in the insurance sector. When a consumer approaches insurance company for the insurance or customer approaches bank for the loan, they can check his/her profile and take an informed decision. Big Data analytics also have several use cases in the fields of research, customer service, sales and marketing, health care, public welfare agriculture in fact across all the industry. It will help in solving business problems using advance technology for data analysis as it makes it possible to seek, store and understand data sets which were earlier deemed to be unmanageable.

Startups and enterprise across the board are already working on making devices that can handle critical data 24x7 from the customers and environment around. This in turn will improve the services provided to their customers.

Big Data is also proofing itself in the field of marketing. Marketers are using data available from the digital marketers and the industry to better understand the behavior and the needs of the customers. This analytic is improving the planning of marketing strategy, thus, improving business performance. It is also improving the time taken to deliver the product to the market and the end customer through automated systems. All the major corporations and enterprises are using advance Big Data method tools to work on their data and gather relevant customer insights. They are also using it to enhance the user experience on their website and application. The unmanageable data of the past can now be effectively managed and understood. Thanks to the Big Data and technologies like AI and ML.

Now is the time to mobilize the synergy of human expertise, data analytic and AI to optimize business value and outcomes by making the best informed decisions. Big Data Analytic is one of the top five catalysts that will not only drive the job market but will also boost the Indian economy in the coming years.

Need to ensure our Data is secure and governed by the use of latest technology

Himani Agrawal, Microsoft India

We are at the cusp of technological revolution, a transformation where AI is poised to revolutionize everything whether it is healthcare, transportation, education, communication, manufacturing, financial services to everything that we know today. Live, work and interact; all of that is going through significant changes and that will bring a lot of opportunities in front of us and lot of challenges which we will need to keep tackling and solving as we move forward together.

Today, Generative AI is the buzzword on every table, boardroom and even household that's changing the world significantly, which is pretty much that a leader is thinking about. How to leverage generative AI and enable my organization to change it, transform it significantly and be ahead of the competition, market and also to empower our employees the way they work or empower our customers the way they interact with our products and services.

How do I create more growth opportunities with that? How do I make process simpler and intelligent so that we know that how do we manage the topline and the bottom-line? That's what is on top of every leader's mind. But AI is as good as the data it uses to produce that intelligence. Data is at the center of the AI and in order to move from Business Intelligence (BI) to AI enabled insights, to walk the journey from what happened to why did it happen, to what will happen, to how can we make it happen. This requires a lot of clean and rich data and very strong tight analytical systems that needs to provide that information which can actually convert it into AI driven insights.

But what does organizations have today, all the analytics system that exist in most of the organizations are the maze of a very special-

Vinod Sood
Chairman, CII Regional Committee (NR) on Digital Transformation (DX) and Co-Founder & MD Hughes Systique

Himani Agrawal
Country Head - Azure Microsoft India

Karthik Neelakandan
VP & Head India Business Infosys

Alok Gupta
Founder & CEO Pyramid Cyber

Gaurav Malik
Chief Strategy Officer Successive Digital

ized and disconnected services with lot of issues. It could be proprietary platforms that are logged in or costly integration and upkeep time and maintenance which is theirs.

This scale issue comes along with the infrastructure when you have to grow and off course the data and security risk, governance issues etc. All of these become much bigger issues because the velocity with which the data is growing and rising and increasing every day. The amount of data that will be created probably in next 2-3 years won't have been created in last 30 years. So, the scale and volume of the data is so different and high, that you would need much better cohesive approach to address these needs. So, there is good news and bad news.

The good news is there are tons of innovation that has happened in the data and AI in the last 5-10 years. The bad news is that it has added lot of complexities.

Today, the companies need open, trusted and secure data platforms so that you can focus on your resources on making those decisions and not integrating those silos of information. Don't make your Chief Data Officer as Chief Integration Officer. They are crying from the roof top for the help to bring the silos information and spend the time in integrating those systems. I want you to focus on what this data can do for the organization. That's why we have introduced a unified SaaS product that takes care of everything from a data lake to a business user. You need to make the decision whichever industry you are on, as it addresses everything.

It also delivers AI copilot which brings in the power of the Generative AI to create the natural language that interface like ChatGPT interface which is required for the business user to interact with data that you are providing to them. We announced this as a preview few months back. In the last few months more than 25,000 customers have already started journey on it. Almost 67 per cent of the fortune 500 companies are using it because it is that simple.

We are living really in the exciting times to ensure that AI works for all of us. We need to focus on our data and for that we need to focus on tools and tech which can make our team's life simpler right from a person who is working on the data to the person who is consuming that data by ensuring that it is secure and governed by the use of latest technology like confidential computing etc.

Industry 4.0 initiative will not be successful without Big Data and Analytics

Karthik Neelakandan, *Infosys*

The industry 4.0 is the integration of all digital technologies to improve manufacturing and industrial processes. There are plethora of technologies that are being used to enable this outcome like IoT, robotics and automation. But without Big Data and analytics, no Industry 4.0 initiative will be successful. Industry 4.0 enables the manufacturing companies increase their topline by creating next generation products and services and also, improve their bottom-line by driving efficiencies.

The volume of data that is used across the world has significantly increased research such that over 10 Exabyte of data is being generated on daily basis. Now, you have to make sense of this data. Variety of data is also huge. There is structured data, semi-structured data and unstructured data. You have to handle these varieties of data and need technologies to handle the data. The velocity impacts the decisions that you are making. Recalibration of machines based on the previous data can improve efficiency and effectiveness. In other domains like market

like stock exchanges millions of billions of dollars can me made or lost because of the way we handle velocity of data.

Now, two more V's has been added recently from the Big Data perspective. One is about the "Veracity" which talks about quality and accuracy of the data. If there are missing pieces than we might not get the insights we want. The final 'V' is about "Value". Ultimately, if you are not able to get the value out of this data, then all this exercise of collecting data, analyzing and visualizing them is of very little use.

One of our customers has process bottlenecks due to lack of formal scheduling processes. They didn't have shift by shift planning and no predictability. There was large backlog of order and visibility information utilization. This led to inability to ramp up the production, there was no visibility to availability and performance of especially old machine. There was a broken energy and waste management system. We used the Big Data and analytics to provide them the solution which created real time dashboard with the visibility of OEE (operating equipment effectiveness) performance and quality monitoring. We provided them intelligent manufacturing schedule and product optimization through smart planning algorithms which is now becoming more and more rampant in the manufacturing industry.

We are using data sensors, power meters which are installed for tracking liquid and solid waste and integrating with the building management system and waste data with Manufacturing Execution System (MES) for real time production correlation as a result of all these solutions component. Ultimately, the client was able to increase the production possibility from two days to 21 days and OEE which is increased to 70 per cent and decrease in solid liquid waste from machine above 10 per cent. This is the classic case, how we have used data structured and semi-structured, un-structured data from machines, from sensors, from integration with the rest of the ecosystem to deliver meaningful value to the customer. There are similar stories about how we enabled digitization manufacturing operations for customers across the world to deliver the kinds of outcomes like optimize resource utilization, productivity gains improved agility, flexible operations and business growth and innovation.

A very reliable industry estimate says that 80 per cent of the project doesn't see light of the day. The key reasons are objective is not very clear. Without clear goal it would be challenging to determine what data to collect and how to use it effectively.

The poor data quality has improved but it is still a concern. Another issue is insufficient storage capacity. Working with Big Data sets, it is important to have enough storage to accommodate all kinds of data and making sure to have enough space on your servers or cloud or considering using compression technique to reduce the size of your data set. Data security also remains a challenge that leads to data theft and identity theft. Governance and clear ownership, texting, lack of documentation many of these are problems in multiple Big Data and analytics project and handling them well is important for critical success of your project.

Network forensic can protect both edge as well as network from any cyber attacks

Alok Gupta, *Pyramid Cyber*

The kind of data that we are generating every day is close to 230 million terabytes. Big Data analytics refines the data. With AI and ML, more refinement is possible.

Generative AI is transformative enough and brings in the data augmentation. It means you are trying to mimic the real data which actually can be used for training model further. If these machine models are trained further then accuracy of these models is going to be far better.

Cyber security is one of the biggest use cases for Big Data because there is about 2.3 million ransomware attack that is happening. Imagine ransomware attack can create a havoc like what happened in AIIMS or to any large corporation. The attacks happen because all these threats actor use AI to generate attacks. These attacks are automated. The whole thing can be prevented if we do threat analytics using the Big Data which is available with organizations. Because every lock, device there generates a log and if you are able to carefully collect that data and do analysis on that then we know what could be the pattern behavior which the attacker might be using to attack you.

The one way to prevent yourself from ransomware is threat analytics. There are ways available both in the security management systems which are able to predict this is the kind of attack which may happen which then can be used for fortifying your security infrastructure further. Another aspect is ethics. I think ethics is being talked about in every AI start-up, any venture capitalist and private equity fund which is trying to fund startups in AI. They are talking about ethical use of AI. We have been hearing about fake videos, fake news all that is detrimental to the growth of AI and that's where ethics comes into the picture.

Now, in the light of new Digital Personal Data Protection Act (DPDP, 2023) Bill, it is very important to classify data because most of the organization generate huge amount of data. But they don't know which data is sensitive, which is personally identifiable information (PII) and classified as something which can be protected because you don't need to protect every kind of data. You only need to protect data which is sensitive in nature where your penalization can happen on DPDP Act. So, it is important to do classification and

also deploy technologies like information right management, digital right management which is then able to protect data at the data level. You don't need to look at the application at the database level and you can protect them.

Another area is forensic and that's where I find there is lot of data being generated by the devices and network. Now, if you are able to analyze the data that can help in network forensic, you can protect both your edge as well as network from any cyber-attacks. So, network forensic plays a major role in actually helping you in analyzing what kind of attacks have happened.

Predictive and prescriptive analysis of data is going to change the world

Gaurav Malik, *Successive Digital*

Big Data is the key catalyst for the organization growth enabling predictive capability and providing strategic edge for staying ahead in today's competitive landscape. The important thing is to get crux out of the data. How do we refine the data? The importance is around the quality of the data. For example: how do you measure data when you are evaluating data, growth. Airbnb just uses one data set for their growth, the data set is number of nights sold. In order to get that data, imagine the number crunching that happens in the background, the quality of data that is required because Airbnb is worldwide.

Top five Key Performance Indicators (KPIs) are measuring for your growth and business. First is knowing what your KPIs for the growth are, second is how I get those KPIs, third is how to implement them. It all starts with the quality of data, how do we rinse the data, what are the systems that you implement to get the insights and how do we create a nice little dashboard. Then comes the challenge of real time analytics. The data is being given but most of the reports are probably three months old. We work on quarterly, monthly and how about the work on the daily data.

These are the things that are changing in the industry. Predictive and prescriptive analysis is one of the big themes that is going to change the world. Predictive analysis uses historical data to predict the future outcomes and while prescriptive analytics goes a step further to suggest and take optimal result. Both of them likely to become more prevalent as organizations look for competitive edge. We all live in competition and doing similar things. The difference that can be is of the predictive edge of the work that we do. We do predictive analysis in our work in the first step you might be 30 per cent correct, 50 per cent in step 2. But when you have tried this 100 times, you are 90 per cent correct.

So, the organizations are going to win. It is not easy but possible. That's where the power of AI comes in. The AI and ML is going to help us in predications. We need to know how to use the algorithms.

Now, lot of buzzword started with Generative AI. But the biggest challenge is, people don't know how to use AI. One of the right way to use ChatGPT is knowing the right prompt. We have to train our version of ChatGPT to tell who you are. The more models know about you, better outcomes you will predict for you.

The key takeaways are that data is your oil. Hence, make friends with data. Have KPI in terms of what you are going to measure, have predictive analysis for anything that you are doing, keep data as part of your design of your steps, of your initiative that you are doing and make AI your friend and tell AI what exactly you are looking for and outcomes would be phenomenal.