

ANALYTICS INDIA MAGAZINE & TAPMI

TOP DATA SCIENCE & AI TRENDS FOR 2022

Key trends in AI and Data Science that would define the industry
in the coming year.

FOREWORD



Dear Reader!

Greetings of the new year from TAPMI!

It's a great pleasure to collaborate with Analytics India Magazine and present the exciting Data Science and AI trends report.

AI has ceased to be just a buzzword, as organizations are increasingly understanding the merits of using analytics and Artificial Intelligence for automated and augmented decision-making. Hence the adoption and use of analytics and AI has become pervasive across industries. The pandemic has acted as a catalyst for rapid adoption and use of digital technology and process automation, bringing the use of AI and , analytics and decision sciences to the forefront.

This report provides the reader with rich insights about industry trends in AI, analytics and Decision Sciences. Based on inputs from leading practitioners across diverse industry verticals, the report provides a holistic view of industry-wide adoption and trends in Decisions Science and AI. In a nutshell, the report captures the voice of the industry.

Realising the potential and the future pervasiveness of AI and data science in the industry, TAPMI has included AI, decision sciences and analytics in the curriculum for several years now. Students of our flagship MBA programs can major or minor in Analytics in the second year of their curriculum. We have also tied up with Mu Sigma, one of the largest pure-play analytics companies in the world, for training the next generation of decision scientists.

TAPMI has created a cutting-edge analytics and data science lab to help us integrate and leverage analytics in our curriculum. We are also looking to foray into the undergrad education space and nurture high quality talent in the AI, decision sciences and analytics domain.

We thank you for your interest in the report. Happy reading!

Regards,
Madhu Veeraraghavan
Director - TAPMI

EXECUTIVE SUMMARY

After a pandemic-driven start to 2020, enterprises across industries realised the significance of AI and Data Sciences. Its adoption accelerated, and leaders correctly predicted growth in the industry in all aspects. Overall, organisations invested more in Data Science, and there was an upswing in the Data Science jobs. While the median salaries of analytics professionals saw a slight decline at the start of the year, a rising trend was witnessed again in the recent months, which will continue to be the case in the coming year. The inefficiencies of Data Science teams from development to deployment in the real world were observed before but they became even more evident due to the pandemic. The operationalisation and scaling of Machine Learning models through structured frameworks was the talk of 2021. These processes will start getting streamlined in the coming years.

The Data Science industry also realised the breadth of roles needed for these deployments. While generalists will continue to be in demand, niche roles will play an important role going forward, especially Data Engineers. Subsequently, the role of education will also evolve. It will become further formalised with more specialisation courses introduced. First, they will be introduced as certification courses and then as undergraduate or postgraduate programmes. Data Engineers will also play an important role in establishing data management architectures as companies look to democratise data access and establish efficient pipelines. Organisations will redefine their data strategies according to these processes. Large language models will become huge, and new age algorithms will be used even by smaller companies.

Finally, this year was marred with controversies surrounding Big Tech and the use of biased or unethical algorithms. The impact of AI/ML algorithms on society and individuals is becoming apparent, and the accountability of organisations building them is increasing. While organisations in the Western world have realised this and started taking steps, the field of Ethical or Responsible AI is still at a very early stage in India. This will change in 2022. Companies will actively hire AI Ethicists, and third-party auditing will become a part of the modelling process. Finally, leaders have also realised the importance of localising AI/ML. Training models on local data will not only give better business results but also provide better accuracy. This will play an important role in improving customer engagement.

ABOUT TAPMI

Established in 1980, T.A. Pai Management Institute (TAPMI) is one among the 7 institutes in India and 5% of the top business schools worldwide to have been accredited with the illustrious Association to Advance Collegiate Schools of Business (AACSB) accreditation, the oldest accreditation in the world of business education. It is also the 5th institution in India to have both the AACSB as well as Association of MBAs (AMBA) accreditation. TAPMI is not just a B-school; it is a centre of business excellence with over 30 years of experience in academics, research and Executive Education. A few years ago, TAPMI and Mu Sigma introduced one of the most exciting programs in the area of analytics and decision sciences. The goal of this program is to create exceptional leaders in the digital age, thereby developing a new approach to decision-making called the new Art of Problem-Solving.



TOP DATA SCIENCE & AI TRENDS FOR 2022



1. MLOPS WILL ACQUIRE WEIGHT IN OPERATIONALISING AND SCALING AI/ML



2. RESPONSIBLE AI WILL GAIN SIGNIFICANCE IN INDIA



3. DEARTH OF DATA ENGINEERS WILL BE FELT MORE THAN DATA SCIENTISTS



4. LARGE LANGUAGE MODELS WILL BECOME LARGER

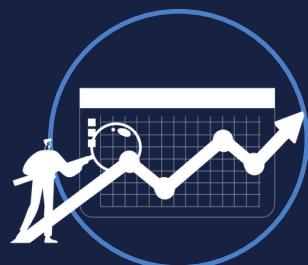


5. FURTHER FORMALISATION OF DATA SCIENCE EDUCATION WILL LEAD TO SPECIALISED COURSES

TOP DATA SCIENCE & AI TRENDS FOR 2022



6. HOLISTIC DATA FABRICS WILL BE AT THE CENTRE OF REDEFINING DATA STRATEGY



7. DATA SCIENCE INDUSTRY WILL SEE A CONSOLIDATION OF MORE IPOS OR ACQUISITIONS



8. ANALYTICS PROFESSIONALS WILL COMMAND HIGHER SALARIES



9. NEW AGE ALGORITHMS WILL SEE HIGHER UTILISATION IN INDUSTRIES



10. EFFORTS IN AI/ML LOCALISATION WILL IMPROVE



1. MLOPS WILL ACQUIRE WEIGHT IN OPERATIONALISING AND SCALING AI/ML



Around 72% of organisations that began AI pilots couldn't deploy even a single application in production, according to a 2019 Capgemini [report](#). Similarly, a 2020 [survey](#) showed that around 55% of the companies actively engaging in Machine Learning had not deployed a single model. While many data scientists build machine learning models, they lack engineering knowledge leaving a glaring gap between development and deployment. MLOps brings Data Scientists and IT Engineers/developers together to deploy ML models faster and at scale.

IT Engineers have been using DevOps for years now. DevOps is a set of practices and tools that increase the organisation's ability to deploy applications in the real world. MLOps builds on DevOps ideas to facilitate the automated development and deployment of machine learning models and applications. MLOps observes growing importance among Data Science or AI leaders, who have realised the limitations of data scientists to be good programmers.

While MLOps as a term was [coined](#) in 2015, it has gained traction in recent years.

The rising significance of Data Science and the need for automation (especially after the pandemic) got leaders talking about the topic more. AI leaders predict a growing focus on MLOps in 2022. Its use will be streamlined and decide investment in ML models. Furthermore, the gap between Data Science and Engineering will narrow.



We will increasingly see maturity in systems where AI and Engineering elements are built to allow easily scalable, automated and “low-touch” operations. The field of MLOps will allow for easier and faster scalability to AI/ML operations, and we will see increasing use of “AutoML” tools that will make the process of building AI models easier.



SURAJ AMONKAR, VP - AI@SCALE AT FRACTAL ANALYTICS



AI/ML Ops will gain a lot of traction due to demand for faster experimentation and execution of AI/ML models at scale owing to the phase where a lot of prototypes envisage production. Along with it, AI/ML on cloud and low code AI will become more prominent.



RUBLE JOSEPH, VP - DATA SCIENCE AND ANALYTICS AT ECLERX



"The problem of moving from experimentation to production is a known challenge, and we have seen reports stating a large percentage of AI/ML experiments don't make it to production. Post the covid crisis, when we see a significant increase in the adoption of Data Science to augment human intelligence drive business decisions and outcomes, both topline and bottom line, we will see an increased focus on MLOps, especially in the non-digitally native organisations. The thought process behind setting up experiments will also be a lot wider, and the focus will be around the end-to-end execution rather than a quick fix experiment."



PARIKSHIT NAG, CLUSTER LEAD, DATA SCIENCE AT HUL



AI/ML Ops will gain a lot of traction due to demand for faster experimentation and execution of AI/ML models at scale owing to the phase where a lot of prototypes envisage production. Along with it, AI/ML on cloud, low code AI will become more prominent.



RUBLE JOSEPH, VP - DATA SCIENCE AND ANALYTICS AT ECLERX



According to Gartner, the performance, scalability, interpretability, and reliability of AI models need robust AI engineering. The three pillars of AI engineering are DataOps, ModelOps, and DevOps. DevOps deals mainly with high-speed code changes, but AI projects experience dynamic changes in code, models, and data, and all must be improved. Therefore, organisations must apply DevOps principles across the data pipeline and the machine learning model pipeline.



ANISH AGARWAL, DIRECTOR - DATA & ANALYTICS AT NATWEST GROUP



Value of ML for businesses has now been well proven, and large organisations have started using ML across hundreds of use cases (big & small). This is leading to a challenge of managing all these models - many times thousands of them getting as frequently as every minute. ML as an engineering discipline will come of age.



PRADEEP GULIPALLI, CO-FOUNDER AT TIGER ANALYTICS



2. RESPONSIBLE AI WILL GAIN SIGNIFICANCE IN INDIA



Analytics India Magazine [conducted](#) a survey in April to analyse the state of Responsible AI in India. It was observed that while some Indian enterprises are making an effort in adopting guidelines or frameworks, they are still behind when it comes to conducting third-party audits or impact assessments.

Around **a third (33%) of the companies** do not have any internal risk evaluation or auditing frameworks, and only 7% of firms had adopted third-party auditors.

In addition, only a quarter of them (28%) have bias detection frameworks.

Leaders predict that Ethical AI frameworks will play a significant role in 2022, with audits becoming a part of modelling cycles. Considering how Big Tech has been held accountable for their biased AI algorithms, we will see an increased focus on the responsible and ethical development of AI/ML.

Beyond the ML model's decision-making, the ethicality tests will also extend to privacy. These will be subject to the data privacy bills introduced by the Indian government. Companies will start standardising processes for the same, which will play a significant role in the organisation's data strategy.

“

With great power comes great responsibilities. While we look at the advances made in AIML over the past decade, it's hard to ignore the controversies which marred the likes of Google and Facebook (Meta) around biased algorithms. With a stark increase in AIML adoption, Ethical AI Frameworks will play a significant role moving forward, and model audits may even become a part of the modelling lifecycle.



PARIKSHIT NAG, HEAD OF DATA AT INDUS OS (OS LABS)

“

The areas of privacy-preserving algorithms and other ethical AI paradigms will see higher traction as we will see the development of algorithms that can inherently factor for fairness, bias, privacy, and transparency. We will see systems that are built with “ethical” considerations baked-in and will help provide a better than before ability in these areas.



SURAJ AMONKAR, VP - AI@SCALE AT FRACTAL ANALYTICS



There will be an emphasis on data governance and user consent. More and more countries are likely to apply restrictions on the usage of user data. India will likely see significant progress in the Data Governance Bill in 2022. This trend is likely to gain prominence in other G-20 nations outside of those that have frameworks for these already in place. Organisations need to get ready to adhere to these governance frameworks. This will involve investment, education and expertise to come together within the organisations to ensure adherence.

SUBRAMANIAN M S, HEAD OF CATEGORY MARKETING AND ANALYTICS AT BIGBASKET



As organisations scale up on their AI adoption journey and data privacy issues gain prominence, Data Governance will become the foundational pillar of an organisation's data strategy. In simplest terms, data governance is about managing data as a strategic asset and democratising data responsibly.



ROHINI SRIVATHSA, NATIONAL TECHNOLOGY OFFICER AT MICROSOFT



The irresponsible or unregulated use of AI/ML within business operations can pose serious risks to an organisation's credibility and reliability. To minimise this risk in the year ahead, enterprise leaders should establish a responsible AI strategy and framework.

By doing this, organisations will build trust in their product and services through business metrics evaluation, legal compliance, data interpretability and explainability, data-pattern change mitigation, added reliability, privacy, and autonomy for customers and employees.

SREEKANTH MENON, VP - DATA SCIENCE AT GENPACT





3. DEARTH OF DATA ENGINEERS WILL BE FELT MORE THAN DATA SCIENTISTS



The annual [salary study](#) conducted by AIM Research in June 2021 showed that **Data Engineers commanded a median salary greater than Big Data Scientists** or AI Engineers. This is indicative of the increase in demand for these professionals but a lack in supply. Data Engineers are analytics professionals responsible for generating, cleaning, processing, and storing data in a way that is ready for analysis.

Data Engineers lay the foundation for Data Scientists or AI/ML professionals to do their jobs.

With the increase in digital transformation after the pandemic and the ability to collect data from various sources and formats, data engineering will play an important role. The industry already faces challenges in hiring Data Science talent. In addition, the dearth of data engineers will be felt even more in 2022.



The pace of digital transformation, post-pandemic, has increased the demand for data engineering capabilities manifold. Most newer projects focus on creating a robust data layer before applying sophisticated machine and deep learning algorithms. In addition, the explosion of data itself, the need for more quality and comprehensive data, as well as increasing cloud capabilities are adding fuel to the growing traction of data engineering skillset, which is now being rightly termed as the sexiest job of the current times.



SWATI JAIN, VP ANALYTICS AT EXL



With an increase in the volume and sources of data and the continuous evolution of data processing platforms like the cloud, the task of Data Engineering is becoming crucial & challenging by the day. With Data Engineering quickly becoming the nerve centre of Digital Strategy, organisations are finding it challenging to align themselves and build teams of Data Engineering Talent. As of 2021, LinkedIn is showing more than 29K job opportunities in data engineering as organisations still face a significant shortage with not enough data engineering talent in the market. As the range of skills entailing Data Engineering spans 15-25 technologies, hence building an integrated multi-disciplinary team is the key to success.



SRIRAM NARASIMHAN, HEAD OF DATA, ANALYTICS AND AI AT COGNIZANT

“

As I see it, the industry is already facing a lack of data engineers, and it's only going to become amplified as we go into 2022. Companies need to build robust in-house training and invest in public certifications to equip engineers to get technically good at data engineering skillsets. Also, as a business, one should start looking at available tools and DIY platforms that can be used to supplement some of the skills data engineers bring to the table.



SURESH CHINTADA, CTO AT SUBEX



”

During yesteryears, talent scarcity was primarily around "Data Scientists", where the need for Quality data was kind of overlooked. But, now we realise that the availability of Quality Data is crucial to solving Problem using Data Science - and, the Quality of the Data will directly influence the Quality of the Solution. Data Engineers are solely responsible and play a pivotal role in generating this data (from the relevant source(s) using the right tools & techniques) so that Data Scientists can curate and consume it to find a Machine Learning model that can effectively solve the problem in hand.



PADMAKUMAR NAMBIAR, DIRECTOR OF PRODUCT DEVELOPMENT AT ORACLE





Data science is an advanced level of data analysis driven by computer science and machine learning. Data engineers tend to have a far superior grasp of this skill while data scientists are much better at data analytics. Data engineers have three critical tasks to perform – design, build and arrange data pipelines. In contrast, data scientists analyse, test, aggregate and optimise data. Behind every one data scientist, 2-3 data engineers are required. There is exponential demand growth for data scientists and the Bureau of Labor Statistics predict growth of over 30% every year in the next 10 years. With this calculation, the requirement of Data engineers will grow by approximately 50%, which is huge by any standards.

PRADEEP MISHRA, SR. VICE
PRESIDENT AT VECV





4. LARGE LANGUAGE MODELS WILL BECOME LARGER



GPT-3 made a lot of fuss in 2020 when it published an [article](#) in The Guardian. However, its methodology made the model's shortcomings very apparent, and it was written off by many leaders in the industry. In all, experts have divided opinions on the subject. While some have expressed surprise over the rapid progress various language models have made, others see significant limitations.

One thing is sure that as language models grow, their capabilities change in unexpected ways. For example, the **GPT-3 model had 175-billion parameters**, 100 times more than GPT-2. These parameters have now reached trillions, with Google [releasing](#) GLaM trained on 1.2 trillion parameters and DeepMind released Gopher trained on 280 billion dollars in 2021.

Along with the number of parameters, newer models also are seeing improvement in computation efficiency and training text.

The importance of language models is growing significantly. This is especially true in the post-pandemic world, where conversational AI plays a big role in customer engagement. With more accuracy, the models will achieve several tasks like writing articles, synthesising reports, searching, and code generation. Leaders believe that we will see many enterprises in the industry next year focusing on solving pertinent challenges in the area.



The current arms race in NLP to achieve accuracy on multiple tasks using a single transformer model will heat up in the coming year. This is quite pertinent to the performance of conversational AI (a manifestation of NLP technology) convincingly beating the Turing test. The challenge, of course, will be to move from generalised tasks to domain-specific tasks so that transformer models can perform when given a short prior context. Our DS groups experiments with GPT-3 models have shown that this is where the industry needs to focus next.



ASHWIN SWARUP ADURTHI, VP - DATA SCIENCE AT DIGITÉ



The newly launched Language models like Gopher from DeepMind and Generalist Language Model (GLaM) from Google are certainly helpful in improving the overall accuracy of NLP (NLU & NLG). However, we still need to have a contextual language model for the use case being implemented. Try saying What's up in Google Assistant, Alexa, PAi (in NPCI), DISHA (IRCTC) - you would see the difference in the understanding and the answers! Better and more the contextual language data model, better the accuracy! In the coming years, organisations will work towards introducing these contexts in language models.



ANKUSH SABHARWAL, FOUNDER & CEO AT COROVER

“

PIM

Large language models (LLM) will fuel the next wave of automation. Arguably LLMs could be very close to the AGI or artificial general intelligence. There are many organisational tasks, business processes that still rely on the human ability of language and technical writing skills. Even in the media industry, for content repurposing there is great demand for automation of tasks for subtitle generation, creating storylines, publishing in multiple languages simultaneous. LLMs can bring more than 90% of automation in each case. The LLMs are not just about language alone. It learns from image, video and thus in the true sense bridge the context gap which is usually one of the major drawbacks of other types of models. This multimodality is one of the key improvements and will be found generally useful for the usage of LLMs.



BISWAJIT BISWAS, CHIEF DATA SCIENTIST
AT TATA ELXSI

“

2021 observed a growing increase in the use of natural language for routine analysis to detect trends, multilingual language construct and sentiment in data. These have downstream use cases in providing seamless multilingual experiences in chat, search, coding, media, even literature and art. This is in part fuelled by the ever-expanding penetration of technology worldwide. Both these trends will continue to intensify in 2022 with increased requirements for multilingual processing. An example of how this technology can evolve is Open AI's GPT-3. While garnering mixed opinions and still a long way to go, it has shown capabilities in creating human-like language in code, conversations etc. GPT-4 is expected to incorporate text and visual patterns to improve this.



VANITHA D'SILVA, DIRECTOR DATA SCIENCE
AT SKORUZ TECHNOLOGIES



5. FURTHER FORMALISATION OF DATA SCIENCE EDUCATION WILL LEAD TO SPECIALISED COURSES



As Data Science becomes more ubiquitous across industries, the demand for data science talent will grow further. Today, the lack of data science talent is probably the most pressing concern the industry faces in India.

While the subject needs continuous upskilling, formal education will play a big role in addressing this talent gap.

Private and public institutions have realised this and have introduced newer courses. These courses go beyond introducing specialisation modules in an existing course and moving to full-fledged undergraduate and postgraduate programmes. Just this year, AIM's academic rankings got nominations of more than ten different postgraduate and undergraduate programmes that are running their first batches.

These courses will also get more specialised. Industry leaders realise the breadth of data science roles and the importance of fulfilling the positions through the analytics pipeline. As the demand for specialised professionals like data engineers, NLP engineers, Computer Vision, etc., increases, more courses specific to the subject areas will be introduced.



Data Science education has become more formalised, and it will continue to do so. When I started work, there were no courses in data science. People from statistics, econometrics, computer science or even business got together to create the discipline. Now that has changed, and as the discipline gets more structured, we will see even more structured courses, both broader 'end-to-end' ones and new niche areas. I also think for niche courses it will be a demand-pull process where corporate clients would be the first ones to reach out to institutions to ask for very specific areas to upskill their talent. I have seen this earlier for HR analytics, supply chain, healthcare analytics, etc. Eventually, as the broader mandate picks up, these courses become a mass market. I predict that will be happening next in Cloud-Based AI, IoT and Edge, AI/ML in Cybersecurity or Fintech among others.



DIPYAMAN SANYAL, ACADEMIC HEAD AT HERO VIRED



Three things are likely to happen: 1. Data Science education will become mainstream in undergraduate engineering and science colleges in the form of BSc/ B.Tech/ BE in Data Science. 2. As the area matures, institutes will offer sharper specialisations ranging from Data Management and Business Intelligence to Data Engineering, Machine Learning Engineering and Data Visualisation (and several others). 3. Domain-specific Data Science will gain currency, especially for mid-level managers, and there will be industry-academia partnerships to develop Data Science courses customised for domains like supply-chain, retail, health, agriculture, finance, marketing, telecom, manufacturing etc.



CHARANPREET SINGH, FOUNDER & DIRECTOR AT PRAXIS BUSINESS SCHOOL FOUNDATION



Data Science (DS) education has come a long way in the last decade. Initially, the number of marquee courses was very limited. Candidates looking to make their careers in DS would have either statistics or computer science background and gain experience and knowledge primarily on the job. Then the number of courses in the MOOCs platforms like Coursera increased, and candidates began to take advantage of such courses to make a career shift.

Now we see full-fledged courses (undergraduate, postgraduate, and certificate) rolled out by almost all educational institutes. Hence the education has become more formalised. I see DS education to be more formalised in the near future. With the evolution of the field, new skills will be needed, and it may be very difficult to have such specific skills incorporated into the generic courses rapidly.



SAYAN MUKHERJEE, PROGRAM CHAIR – LEADERSHIP THROUGH ANALYTICS AND DECISION SCIENCES (L.E.A.D), TAPMI



In my view, data science will become more and more formalised over the years. Because an increasing number of disciplines benefit from data science—domain knowledge in these areas will be incorporated. Some specialised data science courses could be: Data modelling with unsupervised learning, Financial engineering modelling, Data science project management, Data governance and law, Natural language processing models, AI and deep learning.



JAIDEEP GHOSH, PROFESSOR AT SHIV NADAR UNIVERSITY

“

While the increased reliance on Data Science to improve business decision making, assess market trends, decrease losses, and boost profitability is fueling innovation, it is also pushing global recruiters to look out for talent that has the in-depth subject knowledge and addresses the growing demand for niche job roles.

Few profiles like data engineers, data architects, data analysts, Machine Learning engineers, business intelligence developers, etc., are getting popular and require the understanding of programming languages and coding tools which never existed a few years back.

We can expect Data Science as a subject and a domain to get formalised for good as we are already in an era where every day we have innovations taking place, and there's a constant demand for a skilled workforce who are quick to adopt the market evolution and accelerate profitability for the company.

MAYANK KUMAR, CO-FOUNDER & MD,
UPGRAD





6. HOLISTIC DATA FABRICS WILL BE AT THE CENTRE OF REDEFINING DATA STRATEGY



IBM defines data fabrics as a data management architecture that can optimise access to distributed data and intelligently curate and orchestrate it for self-service delivery to data consumers. The idea is to ensure data access to all the right stakeholders irrespective of where it is generated or stored. It is a powerful architecture that standardises data management practices across cloud, on-premises, and edge devices.

This can help in effective and sustainable digital transformation and improve the value of data within the organisation while also reducing costs.

Along with data access and control, it also addresses concerns related to data governance and security.

The increasing ability of AI applications to tightly integrate with the hardware and operate intelligently on their own will facilitate the formation of robust data fabrics. Intelligent edge devices will play an important role in data fabrics that will help save bandwidth, reduce latency, and further improve privacy and security.



Data fabrics have emerged as the key element to designing a successful enterprise data strategy. They serve integrated layers of data connecting processes and distributing valuable insights across operations, users, and platforms. In addition, Artificial Intelligence technologies within the data fabric will dramatically improve enterprises' return on investment, while significantly reducing operational costs.

SREEKANTH MENON, VP - DATA SCIENCE
AT GENPACT



We will increasingly see AI applications that will be tightly integrated with hardware and will see AI applications that live outside "cloud" or "server" based systems. Self-driving cars, robots, A.I. based cameras and similar applications will see increasing maturity and movement towards production applications. AI algorithms that learn quickly: The traditionally "slower learning" and data-hungry algorithms will see movement towards more sophisticated and agile algorithms that learn in 'few-shot', 'self-supervised' or 'active-learning' based paradigms. As these approaches mature and gain traction for accuracy and efficiency, we will see an easier path to build algorithms in areas that were low resourced for data.



SURAJ AMONKAR, VP - AI@SCALE AT
FRACTAL ANALYTICS



“

Intelligent edge devices like IoT and large-scale AI models that are becoming more like platforms will create Ambient Intelligence - digital environments that are aware, contextual and responsive.



ROHINI SRIVATHSA, NATIONAL TECHNOLOGY OFFICER AT MICROSOFT.

“

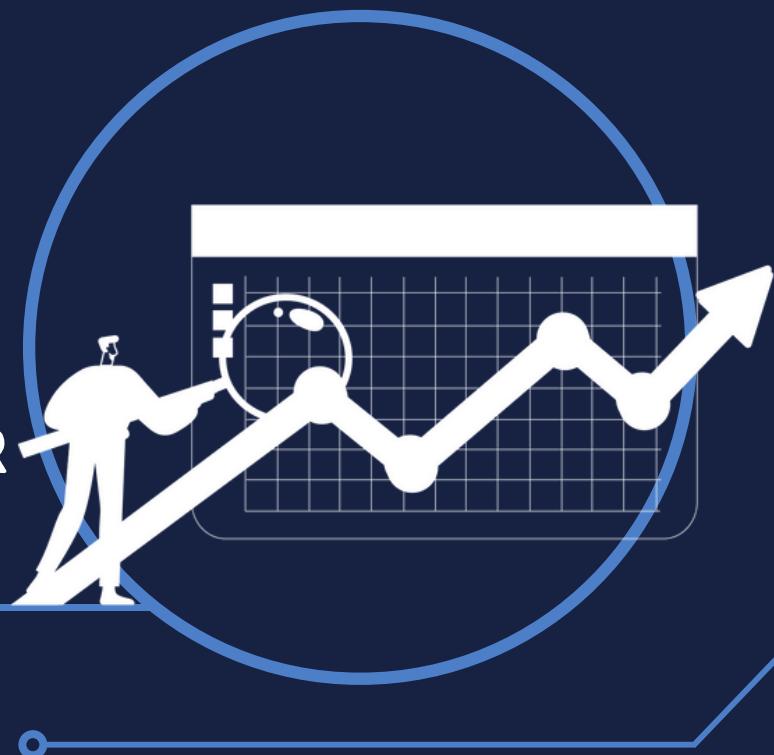
One of the biggest game-changer will be 'Edge Computing'. This will enable companies to store, access and retrieve AI-based data storage to remain local rather than keep it remote in the cloud. This will enable quick and faster decision making. Thus AI will help companies make decisions, take actions and switch strategies in real-time.



PRADEEP MISHRA, SR. VICE PRESIDENT AT VECV



7. DATA SCIENCE INDUSTRY WILL SEE A CONSOLIDATION WITH MORE IPOS OR ACQUISITIONS



India has enlisted 55 companies in 2021 (until Dec 15), raising over 1.2 lakh crores. This is 4.5 times more than the number of companies listed and 3.2 times more than the money issued through IPOs in 2020. In addition, several more companies are in the pipeline for the coming year. Many of these IPOs are new-age companies that are predominantly tech-based and leverage AI and Data Science. Some of the names include Zomato, PayTM, and PolicyBazaar. While several factors contribute to the increase of the overall IPOs, the Private Equity/Venture Capitalist investment cycle of Data Science or Tech companies gives it a further boost.

This is reflected in the over-subscription of LatentView IPO. India saw the first pure-play analytics company listed in the year, paving the way for others.

The number of acquisitions in India also increased significantly. In August, India completed 155 acquisitions - this is more than the total number of acquisitions in 2020. Large companies are looking to buy startups to make way for themselves in the digital economy. A significant investment in the tech startups, along with the agile model that they are set up in, makes it easy for bigger companies to acquire them. Niche AI/ML companies are solving complex problems with great accuracy. Bigger companies see this as an opportunity to improve their technical capabilities, while smaller companies see this as an avenue to improve their market penetration.



The data science and analytics industry has been experiencing tremendous growth. This growth is just the early stages of a deep, long-term change that was set into motion by the shift to digital, and further accelerated by the pandemic. Analytics has proven to be a 'must-have' capability for businesses to succeed and thrive. So, it comes as no surprise that the past two years have ushered in a number of acquisitions. More recently, the successful launch of India's first analytics IPO is another positive indicator of the tremendous potential of the analytics industry ahead. I believe we can expect to see many such strategic investments in the years to come and look forward to accelerating our momentum.



SUNIL MIRANI, CO-FOUNDER & CHIEF EXECUTIVE OFFICER AT UGAM



The demand for Artificial Intelligence in the world today is like never before. Especially after Covid, the adoption of AI and Digital is exponentially growing across all industries. Companies are building their internal capabilities but are also looking for external partners to educate them on the possibilities of AI and reimagining their businesses with AI inside. Next 5 years will see phenomenal growth of AI solution providers leading to more IPOs, strategic partnerships and acquisitions across the industries and geographies. Companies with the ability to scale talent and capabilities with the right focus on innovation will dominate the supply side of the equation.



PRITHVIJIT ROY, CEO AND CO-FOUNDER AT BRIDGEi2i

“

IPOs and acquisitions are force multipliers if done rightly. Think of it as a virtuous cycle. It can have a huge impact across the enterprise value, including employees, clients, suppliers, investors, and partners. As IPO markets gain momentum, data science companies are enjoying favourable market sentiment, high liquidity in the financial system and rising investments in analytics. On one hand, large enterprises are acquiring mature analytics practices to attain speed for innovation. On the other, analytics vendors themselves are racing to differentiate by consolidating capabilities. Companies focusing on building and acquiring customer-centric capabilities will survive to enjoy market dominance in the next stage.



SHASHANK DUBEY, CO-FOUNDER &
CHIEF REVENUE OFFICER AT
TREDENCE





8. ANALYTICS PROFESSIONALS WILL COMMAND HIGHER SALARIES



The scale of growth of any domain is evinced by the salaries drawn by experienced professionals or the growth in the proportion of experienced professionals drawing higher salaries. In addition, the pandemic accelerated the need for data-driven decision-making and intelligent automation, increasing the demand for data scientists.

According to the [salary report](#) published by AIM Research in June 2021, the median salary slightly declined compared to 2021 from 14.4 lakhs to 13.4 lakhs (still higher than the 2019 median of 12.6 lakhs).

However, a recent analysis in AIM showed that the salaries are again seeing an upward trend with the median salary for data scientists at 13. 6 lakhs in August 2021. This trend will continue in the coming year.

The supply-demand gap and need for niche technical skills will lead to analytics professionals commanding more salaries in the coming year.



The current talent scarcity will contribute to leaps in salary benchmarks for analytics professionals. However, there are two factors complicating this phenomenon: a growing scarcity of skilled professionals, despite the high demand for analytics professionals, as well as an emerging trend where professionals are waiting and making conscious choices to join growing unicorns rather than established organisations. Given these various factors, nurturing talent rather than relying on existing talent becomes essential for organisations to ensure a balanced and diverse workforce. Investing significantly in training/upskilling programs, and even infusing professional L&D programs right down into academic curricula themselves, will help encourage a new generation of skilled, resourceful, and future-ready experts.



SAYANDEB BANERJEE, CO-FOUNDER AND CEO AT THEMATHCOMPANY



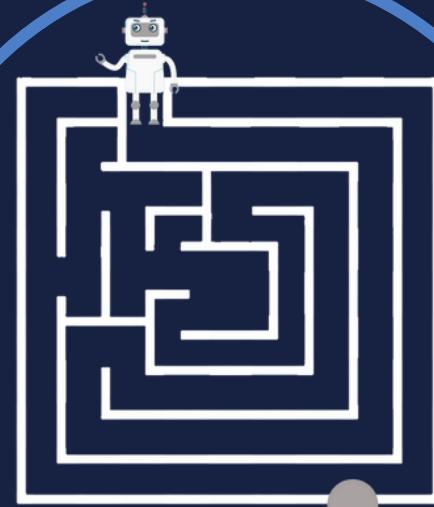
With multiple use cases across industries, the need for quality data analysis is ever-increasing. Professionals in the data science domain leverage their skills for complex functions, which help businesses and other organisations make informed decisions for better economic and social outcomes. Combining the current demand-supply dynamics, where there is a genuine need for analysts and data scientists, and the aforementioned need for technical skills, one can clearly foresee the rise in the remuneration of this sort of talent in India.



SHASHANK RANDEV, FOUNDER VC AT 100X.VC



9. NEW AGE ALGORITHMS WILL SEE HIGHER UTILISATION IN INDUSTRIES



Federated machine learning is an ML technique that can train an algorithm across multiple decentralised edge devices. While the importance of intelligent edge devices has already been established, the significant rise in data breaches coupled with stringent data privacy laws will see a rising demand for federated learning.

Reinforcement Learning techniques have significantly outperformed previous ML algorithms, but the implementation is resource-intensive and costly.

Also, they are extremely sensitive to hyper-parameters. However, its implementation on the cloud has proven to be beneficial in several ways. Firstly, the pay-as-you-go model makes it significantly cheaper to train the models and secondly, you also benefit from better control on hyper-parameters.

Until now, the likes of FL and RL were only implemented by the big tech companies. Going forward, we will see even smaller data science organisations using them. The lowered cost and the privacy demands, along with the overall advancement in the research of these techniques, will drive this adoption.



As lockdowns became the new normal, businesses and consumers increasingly ‘went digital’, providing and purchasing more goods and services online. With changing trends of customer behaviour, the business and models could no longer rely only on historical data. Hence RL algorithms are becoming increasingly famous to build dynamic systems with adjustments for uncertainties. With the increase in cloud-based frameworks and lower technological costs, smaller and bigger firms alike are rushing to use the power of RL. Several reports by McKinsey, HBR and IDC state that RL is the next big thing, and by 2022 – one in six customer experience applications will use RL. On the other hand, with the exponential increase in consumer data, the risks associated with data privacy also increased manifold. Several data breaches over the past years have nudged developers towards techniques like FL that ensures data privacy along with collaborative learning, especially in the post-pandemic digital world.

ANIRBAN NANDI, VICE PRESIDENT, DATA SCIENCES & ANALYTICS AT RAKUTEN



We have seen the transformer architectures disrupt the NLP areas within AI and seeing some initially good results of using similar approaches for machine-vision in the form of vision-transformers. We will see increasing use of similar and other non-CNN based architectures that may provide newer architectures for building deep-learning models. Neuroscience will help the rapid growth of both these fields of knowledge.



SURAJ AMONKAR, VP - AI@SCALE AT FRACTAL ANALYTICS



Annotation is a tedious and time-consuming task, specifically in domains like Pharma and Medical, where the task cannot be outsourced to uninformed individuals and needs expertise and knowledge of the domain even to annotate. While active learning has been applied to tasks like image classification and text classification of late, it is still a work in progress when it comes to identifying named entities recognition as this involves training a deep learning model with sequences at every iteration of the cycle. Privacy-preserving algorithms and Federated Learning is also picking up pace with Hi-Tech customers especially in countries where data privacy is crucial.



MUTHUMARI S, HEAD OF DATA SCIENCE
AT BRILLIO



Deepfakes are great examples of synthetic data. The use of deepfakes for deceptive purposes (political, religious) has caused some disrepute to the field of deep learning. More tools will become available for businesses and entities to identify and cull out deceptive use of deep fakes. In parallel, the field of synthetic data (images, voices, data) will see increased investments given the ability of well crafted synthetic to help in training models that have otherwise been languishing due to lack of good data.



SUBRAMANIAN M S, HEAD OF
CATEGORY MARKETING AND ANALYTICS
AT BIGBASKET



10. EFFORTS IN AI/ML LOCALISATION WILL IMPROVE



AI has primarily grown in English-speaking countries, and ML models have been trained on data in those countries. This makes algorithms less accurate when they function in other countries, especially when it comes to language models. On the other hand, localising content delivers substantial business benefits and improved customer engagement. Indian leaders have realised the importance of adopting AI to local parameters.

Through AI localisation, people train AI engines with hyperlocal content and in-market user experiences-generated data.

There will be an increased importance to the localisation for language translation and curating accurate ML predictions in the Indian context. Beyond data, artificial intelligence and human intelligence will have to work closely together to observe improved results.



Vernacularisation & localisation need to proliferate more. Speaking in the local language and the 'Bharat' segment is another big step. AI has been primarily grown in the US market with the English language. Adopting AI to recognise vernacular languages and the vernacular settings in developing market is essential for us to move further.



MATHANGI SRI, VP DATA SCIENCE AT GOJEK



Switching between languages involves more than just exact translation. There is a need to understand the context and specific language differences to provide a properly adapted chatbot version. The foundation to achieve such seamless translation lies in how strong the knowledge base is. A strong chatbot program is built on a knowledge base that consists of primary data, facts, assumptions, and the rules of the system available to solve a problem. The chatbot's ability to connect and interact with the customer is dependent on how well-built and expansive this knowledge base is.



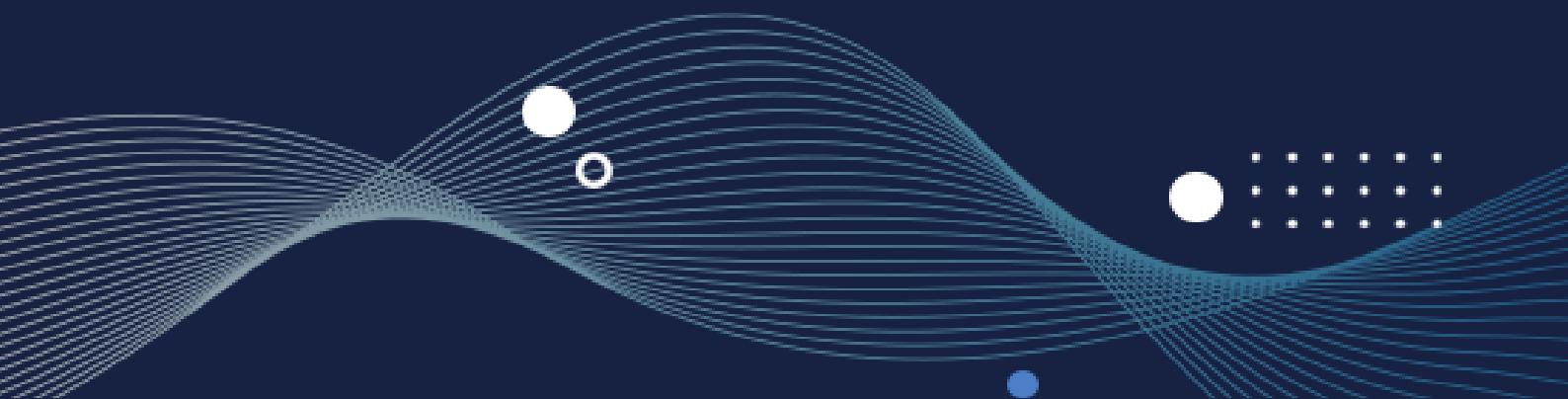
ANKUSH SABHARWAL, FOUNDER & CEO AT COROVER

“

Data science and AI will spread to niche areas previously eluding the purview due to lack of data, limited capabilities, minimal use cases and value. This will give rise to specific but more optimised solutions like better language models, AI for healthcare, AI for manufacturing and supply chain optimisation etc. This will warrant the need for closer integration of artificial intelligence with human intelligence to be able to create niche solution with domain and process knowledge embedded.



RUBLE JOSEPH, VP - DATA SCIENCE AND ANALYTICS AT ECLERX





About Analytics India Magazine

Analytics India Magazine was founded in 2012 and has since been dedicated to passionately championing and promoting the analytics ecosystem in India. It has been a pre-eminent source of news, information and analysis for the Indian analytics ecosystem. It extensively covers opinions, analysis and insights on the key breakthroughs and developments in the field. It engages in the promotion and discussion of ideas with smart, ardent, action-oriented individuals who want to change the world.

Email: info@analyticsindiamag.com

Website: www.analyticsindiamag.com



DISCLAIMER

This document is the result of research carried out by AIMResearch. Permission may be required from AIMResearch for the reproduction of the information in this report. Reasonable efforts have been made to source and present data that is believed to be reliable but makes no representations or warranty, express or implied, as to their accuracy or completeness or correctness. All rights reserved with the aforementioned parties.

© 2020 Analytics India Magazine Pvt Ltd and its affiliates. All rights reserved. Images or text from this publication may not be reproduced or distributed in any form without prior written permission from Analytics India Magazine. The information contained in this publication has been obtained from sources believed to be reliable. Analytics India Magazine disclaims all warranties as to the accuracy, completeness or adequacy of such information and shall have no liability for errors, omissions or inadequacies in such information.

This publication consists of the opinions of Analytics India Magazine and should not be construed as statements of fact. The opinions expressed herein are subject to change without notice.



POSTAL ADDRESS

#280, 2nd floor, 5th Main,
15 A cross, Sector 6, HSR layout
Bengaluru, Karnataka 560102

WEBSITE

www.analyticsindiamag.com

EMAIL ADDRESS

info@analyticsindiamag.com



POSTAL ADDRESS

P.B. No. 9, Manipal – 576 104,
Karnataka,
India

WEBSITE

www.tapmi.edu.in

EMAIL ADDRESS

tapmi@tapmi.edu.in

