TATA CONSULTANCY SERVICES

Building on belief

Ensuring safe, secure, and sustainable AI

Abstract

Given its tremendous scope to empower enterprises, artificial intelligence (AI) has huge expectations and investments riding on it. Therefore, it is important that AI solutions comply with fair norms of operations, deliver trustworthy outcomes wherein contributing factors are managed gainfully, and continue to do these over the long term to become sustainable. In other words, enterprises should strive towards 'safe, secure and sustainable AI' by looking at the larger picture. The AI vision of an enterprise should therefore be drawn upon four key considerations: technology pace, human-machine balance, performance and data.

Considerations for AI vision

Scoping AI maturity

The last few years have witnessed accelerated growth in AI. Today AI is well understood in terms of its maturity beginning with ANI (artificial narrow intelligence), progressing towards AGI (artificial general intelligence), and culminating with the futuristic ASI (artificial super intelligence).

As a first step toward becoming a safe, secure and sustainable AI organization, enterprises should create a strong foundation with today's ANI that can easily mature to tomorrow's AGI and prepare themselves to be future-ready to manage ASI, once it becomes a reality.

Achieving human-AI balance

As enterprises adopt AI, collaborative intelligence should be the goal where AI augments human effort. With machines getting the first right of refusal, enterprises need to find the right balance between machines assisting humans and humans empowering machines. Human oversight will continue to be imperative for careful planning, deploying and sustenance of AI solutions for gainful outcome realization across various technology and business processes.

Defining AI performance

To fully leverage the power of AI, understanding and defining the functional and non-functional performance parameters of AI is very critical. The functional or hard performance indicators include the accuracy and consistency of the AI solution. On the other hand, the non-functional or the soft performance indicators should focus on security, privacy, legality, auditability or explainability, and above all, a fair approach to the AI solution without any social, political, or ethical biases.

Leveraging data

AI has the potential to demystify the data deluge in an enterprise, glean insights, and accelerate decisions. But for AI to deliver the expected competitive advantage, it is important that it gets the minimal required corpus, high-quality data, to train and fine-tune itself. Data security, data privacy, data residency, which fall under the ambit of data sovereignty, and data governance all need to be considered for using data to train machines. Enterprise maturity of best practices around data democratization and governance will go a long way in envisioning a safe, secure and sustainable AI roadmap.

Once the vision for AI journey is in place after careful analysis and planning of the above four considerations, the next step is to establish expectations around trustworthy outcomes based on the six Rs.

Trustworthy AI

For establishing the ‘trust’ and the ‘worth’ of the AI solutions, the following six Rs are the typical considerations:

1. Rational

   Explaining the rationale behind the anticipated output of the AI model helps in its greater understanding and acceptance, paving the way for ongoing enrichment.

2. Righteous

   While planning for an AI solution, enterprises need to consciously ensure the outcome of the AI solution is fair, ethical, and purpose-driven with a larger objective of societal and environmental well-being.

3. Reliable

   The functional and non-functional performance of AI depends largely on the AI solution being reliable in terms of managing:

   Data bias: Training data is not gainfully representing the problem statement to be solved by the AI model

   Data poisoning: Training of the AI model is adversely impacted causing the AI model to behave maliciously

   Model stealing: Technique to learn about the AI model so that unethical acts can be performed for undue advantages

   Adversarial attack: Input to AI model is modified to mislead the model to product incorrect results

   Privacy breach: Information gets accessed without required permission.

4. Revitalize

   AI is all about augmenting human capability and accelerating the growth of enterprises. The right mix of human-machine capabilities to manage business processes is the key to revitalize organizations.

5. Restrictive

   Success of the AI solution is largely dependent on data, and therefore, data privacy is of importance to prevent any undesirable consequences. But enterprises should go one step beyond and ensure the AI solution is technologically secure through the privacy by design approach.

6. RoI and results

   The AI ecosystem in an enterprise must be gainfully managed across different asks spanning multiple stakeholders, technologies and processes. For AI to get a firm foothold in an enterprise, meticulous planning, implementation and change management are critical. AI implementation demands a structured and systematic approach defining the objective (why), the kind of AI solution (what), and when and where AI will be deployed. The approach should also include a RASCI matrix to bring in the right teams into the program and the methodology of how it will be deployed adhering to all compliance norms. Last but not the least, measurement methods and KPIs should be determined. The RASCI matrix must clearly articulate accountability for each aspect of the AI lifecycle. This is very critical to the success of the AI program so that it delivers on the promise by maneuvering the situation and overcoming all possible challenges to maximize the RoI.

Responsible AI

Responsible AI is typically governed by the **five C tenets**. The first C of **C**omprehensible or explainable AI and interpretable machine learning models is similar to the rational consideration of trustworthy AI. **C**lean AI with a fair output aligns with the righteous consideration. The other three Cs include **C**ertain or reliable AI model that's robust and safe by design; **C**omplementary AI ably augmenting human capability as in the case of the revitalize parameter; and conformed AI remaining **C**ompliant throughout its life cycle and drawing parallels with the restrictive and the results components of trustworthy AI.

AI governance

Gainful deployment of a trustworthy and responsible AI solution, which adheres to the six R guiding principles and the five C tenets, is necessary to protect the brand from reputational, regulatory and legal risks. AI governance should have both top-down and bottom-up approaches. While the top-down approach leverages a cross-functional team and an enterprise-level framework to define policies and set expectations around trustworthy and responsible AI, the bottom-up approach will take care of project/product team empowerment for contextual adoption of AI.

AI governance needs to address AI at three levels: content, models and operations. At the content level, there should be a balance between the theoretical approach and what practically works for the business. AI modeling includes preliminary steps such as building a business case, data preparation, model development and evaluation. On the other hand, AI operations include model deployment, execution and management. Governance teams will need to be vigilant about risks such as violation of social/ethical norms, scope change, model behavior guarantee and model malfunctioning. A well-defined ETVX (Entry, Tasks, Validations, Exit) matrix with agreed-upon accountability and responsibility can help ensure proper AI governance.

AI governance can be successful if it follows the three-D approach as outlined below:

1. Depository

   Data governance should exist at a permanent knowledge repository level, where the repository holds content used by AI programs such as context, reference data, business rules, actions, and results.

2. Discipline

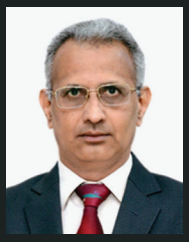
   It is about the controls, which are required to make governance transparent and address core values — both business and social or political.

3. Dissemination

   Information dissemination or communication requirement stems from both awareness and execution perspectives. At the awareness level, it includes organizational change management, wherein proven techniques like communication, training and incentives can be leveraged. From an execution perspective, it is about ascertaining thresholds and overall accountability to individual and group actions, spanning areas such as AI-to-AI/machine, audits, human readability, immutability and non-disclosure agreements.

   Achieving safe, secure, sustainable AI is not purely a technology initiative and cannot be addressed in isolation. It requires a blend of technology, process, and cultural change in an enterprise. Similar to data becoming a board-level agenda in the past, AI must also become a board-level agenda.

About the author



Mahesh Kshirsagar,
CTO – Analytics & Insights, TCS

Mahesh Kshirsagar is the Chief Technology Officer, Analytics and Insights, at TCS. In this role, Mahesh is responsible for shaping innovative and purpose-led solutions that accelerate the growth and transformation agenda of enterprises.

Mahesh is an engineering graduate, who started his career in TCS in 1990, and has IT expertise of 30+ years spanning technology domains, industry verticals and software processes. Over the years, he has incubated several high-impact business-aligned IT solutions/services having high revenue potential, focusing on thought leadership and innovation to enable growth and transformation. His IT expertise stems from his experience in system integrator companies, end-user organizations, IT products and BPO organizations.

Conceptualizing, architecting and delivering state-of-the-art business IT solutions are his strengths. He has applied for more than 30 patents for his solutions, of which around 10 have been granted. His solutions have also won several prestigious industry awards.

Contact

Visit the Analytics and Insights page on https://www.tcs.com

Email: BusinessAndTechnologyServices@tcs.com

About Tata Consultancy Services Ltd (TCS)

Tata Consultancy Services is a purpose-led transformation partner to many of the world's largest businesses. For more than 50 years, it has been collaborating with clients and communities to build a greater future through innovation and collective knowledge. TCS offers an integrated portfolio of cognitive powered business, technology, and engineering services and solutions. The company's 469,000 consultants in 46 countries help empower individuals, enterprises, and societies to build on belief.

Visit **www.tcs.com** and follow TCS news **@TCS_News.**

Corporate Marketing | Design Services | M | 06 | 21