

# AISWITCH Ethical AI Manifesto: 12 Principles

**Who should read this: End-user AI leaders/ strategists/ digital business leaders/ service provider client partners/ AI-automation leaders**

Enterprise AI-automation (IT/ business) leaders and end-users/ service providers/ business leaders who are working on AI-automation strategies, governance, policies and adoption initiatives.

## Why is an ethical AI manifesto needed?

2020 surveys on AI-automation adoption shows a consistent trend amongst the end-user leaders: More than 90% of them are eager to scale up the adoption of AI & intelligent automation solutions in their enterprises. But, low level of trust and acceptances are standing in the way of adoption, as the biggest stumbling blocks in form of change resistance. Even in the COVID-triggered wave of fast adoption of certain AI-automation technologies and usecases, e.g. identity verification for remote work and data sharing, this lack of trust and comfort remains a huge concern for leaders.

Visible and demonstrable use of ethical AI practices and policies, in all AI-automation use-cases across business processes and functions, is an elementary first step in alleviating this very valid concern of enterprise leaders and users. Taking the cue, some of the relatively mature, global service provider organizations have been working at the intersections of AI-automation technologies and their usage ethics, for several years now.

For example, among end-user AI stalwarts, companies like Capital One have had ethical and explainable AI practices set up since 2016. Accenture, Deloitte, EY and AI tech leaders like IBM, Google, Microsoft have come up with their technical and applied versions of ethical AI, in terms of responsible and trusted AI, and interpretable, explainable, transparent AI across the entire spectrum from AI-optimized hardware (e.g. INTEL SGX with secure data enclaves) to APIs and algorithms (e.g. Google XAI). Service providers like Capgemini, Infosys and Wipro have been working on AI governance that goes beyond just technical availability and bot-health monitoring to cover not just quantity but the quality of work done by bots. Across all these initiatives, a common minimal ethical AI manifesto can emerge as a stepping stone for any organization looking to productionize AI in a trusted and ethical way.

Ethical AI manifesto must extend beyond algorithms, technology choices, or evolving standards e.g. considerations of data debiasing techniques and XAI algorithms like deconvolution, LRP, LIME etc., or adherence to budding standards like IEEE P7000 series providing ethical governance on variety of AI aspects ranging from student & child data protection to autonomous vehicles. Ethical AI practices need to start from people (employees, users, business process owners, AI solution architects and owners), and end with people (customers, shareholders, regulators).

# What is the Ethical AI manifesto: The 12 key principles

In line with the most popular concepts presented in the **Agile Manifesto**- which is also very relevant in AI & automation, here is a short & simple, basic manifesto- for ethical AI & automation.

**1: Highest priority for AI practitioners is to augment human intelligence applications positively and satisfy the customers and users.**

Use applied AI to design use-cases that can enable the most optimal value, RoI and outcomes from AI and automation capabilities portfolio, with 100% audit-ability, transparency, human ethics and integrity.

**2: Treat the AI and automation requirements of every customer as unique and directly in sync with their business and technology strategies, in an ethical manner.**

One size fits all doesn't work in AI. For the same reasons, AI cannot be delivered straight off the shelf. Don't attempt a Blanket policy on AI use-cases. Some applications [like compliance to regulatory requirements of different verticals] will be common and will also have a direct linkage to ethics. But also, many enterprise applications will be unique and differentiated, within the same ethical frameworks, due to their close alignment to specific strategic priorities of each enterprise. 3 V's- Velocity, Volume and Variety [most important] are the 3 key pillars of learning, for AI.

**3: Build a journey map for every customer, with specific target milestones of efficiency and UX on short timescales of weeks, across the spectrum of AI and automation.**

Think big. Act small. Deliver specific and measurable outcomes in non-linear efficiency & productivity gains, in weekly spurts, rather than waiting for a huge AI project [like most of the Watson programs] to succeed. Start with basics e.g. from RPA, RDA(Robotic Desktop Automation- screen-scraping based, key-stroke mapping based, record-and-replay type bot learning for example), RBA(Run Book Automation), orchestration, SDA (Service Delivery Automation) etc. Then continue climbing up the AI value curve, by further building on cognitive, supervised, unsupervised deep learning and autonomous capabilities, always keeping human in the loop for overall control & governance.

Too large and convoluted project scopes- with loads of inter-dependencies and tightly integrated AI modules, can make any AI project/ program impossible to manage [like any other software project- which is why Agile took over from Waterfalls]. This will eventually lead to an organizational AI fatigue. People will lose their early enthusiasm and sense of wonder with the new tech kid on the block. Top leadership will lose patience and stop funding, without seeing any palpable outcomes on ground.

**4: Link all AI and automation outputs to ethical business and experience outcomes, including applications of AI and automation in IT.**

What we mean by "ethical business and experience outcomes" is that- these outcomes should only effect human life positively- e.g. measurably augmenting the human productivity and Quality of Work [e.g. by improving process TAT's, process accuracy- reducing human errors & rework, enabling 100% audit-ability etc.], measurably improving our Quality of Life [e.g. by freeing our brain-time up from boring, mundane, routine tasks, so that we can participate in more socially relevant, value-adding activities and creative pursuits: These are measurable by Emp-Sat, Cust-Sat scores, NPS for an employer brand or for customers or both].

## **5: Build appropriate rewards and recognition systems to promote the right behaviour to motivate others to change and adopt the right approach & use-cases.**

The choice to apply and use, is always with the humans. AI is just yet another tool, albeit a very powerful one. Hence, the 'right behaviour' towards AI will be manifested in leveraging AI & automation in a value-systems-based, socially conscious manner. Build appropriate rewards and recognition systems to promote the right behaviour to motivate others to change. Given that all AI programs require significant change in the ways people perceive their work & their value, in a 100+ years' old system, getting them to share knowledge, to collaborate with others including bots, and to train the bots, will require a paradigm shift in staff and user mindset.

Make the self-motivated individuals and early adopters 'AI and automation evangelists'. Give them top leadership visibility and support, and the right environment, lifting all possible constraints.

## **6: Build and utilize persona-based story-boards for right messaging, answering the most critical 'WIIFM' question for each persona, across layers up-down and across.**

Build and utilize persona-based story-boards for targeted and value-based messaging. Given AI will impact people's work the most, every major stakeholder will have this question: "What's in it for me [WIIFM]". Story-boarding around these target persona's helps the affected people to internalize and institutionalize adoption of the new ways of work, enabled by AI & automation. To design the most effective communication to drive this change, use a "PUDA" approach, e.g. "Persona's - Up Down and Across", to cover all the strata of users/ staff. For example, 1- how an AI program will impact a CXO [up], 2- how will it impact a BU manager or line manager [across- peers], 3- how will it affect the shop-floor workforce/ operations staff [down].

Also be mindful of the fact that **Buying Centre persona's [customers] of AI have changed dramatically in recent years and it's NOT a tech challenge targeted towards the CIO BUT a Business Challenge targeted towards Digital CXO's.**

## **7: Recognize that there is no such thing as "perfect AI".**

Recognize that there is no such thing as "perfect AI", just as there is no such thing as perfect human intelligence. Whatever algorithms and techniques across the AI & automation spectrum work best for the customer at a given point in time, are the most apt and best value solutions for that time-frame. Don't wait to release an AI & automation program only after ALL the features are done, instead, release bundles of features at regular intervals and let the system capabilities increase in a spiral.

Therefore, expect slow and steady improvement in model accuracy, with more quality and quantity of data [Ref 3 V's]. Time to accuracy is one of the most visible races that all AI-ML algorithms are currently running, along with the "optimized performance" [read 'compute requirements'] race. **Time is going to become the most critical AI currency, going forward [i.e. who/ which algorithm can reach what accuracy levels within what time-frame, using how much resource].**

## **8: Establish and follow robust governance and design principles.**

Establish and follow robust governance and design principles, from day 1. AI & automation are long-term and strategic capabilities hence risk profiling and mitigation plans must be part of every such program.

**9: First simplify, then automate.**

First simplify, then automate. Automating a sub-optimal process that has not been made lean, will inherently introduce leakages around performance and efficiency and will not let us reach the target levels for our customers.

**10: Let the AI & automation solution teams self-organize, based on customers' requirements, in a time-bound manner.**

Let the AI & automation solutions teams self-organize, based on customers' requirements, in a time-bound manner. Self-organizing teams can help leaders avoid loads of controversies, conflicts and confusions, and blaming and counter-blaming culture.

**11: Build an experimentation-focused organization where failures are seen as opportunities to innovate.**

Build an experimentation-focused organization where failures are seen as opportunities to innovate, and hence are not looked down upon.

**12: Adopt agile/ DevOps as The Default mode to deliver AI use-cases/ features.**

Last but not the least, basis the above principles especially 3 and 7 etc., it is imperative that AI use-cases are built and delivered [and life-cycle managed] through the DevOps mode. Use Data-driven DevOps [ref. Gartner archived research - Cameron Haight & Tapati Bandopadhyay] for defining measurable agility and value-of-speed, TTV, TTM Metrics for AI use-cases delivered in DevOps mode.