

PoV

A Systematic Approach to Building a Sustainable AI Solution



A Systematic Approach to Building a Sustainable AI Solution

Artificial Intelligence (AI) is no longer the future, but a technology that is knowingly or unknowingly integrated into our everyday life. For example, product recommendations in Amazon, navigation with Google Maps, auto message compose feature in Gmail, Facebook feed suggestions, Netflix production suggestions are common encounters. In the last decade or so, there has been an explosion of AI use cases (especially through methods like machine learning), which has taken automation to the next level.

Based on the examples listed above it's evident, that, a 'well thought-out' AI

strategy can transform a business (in fact, be a game-changer) and is what most organizations are looking for. Corporate giants like Google, Amazon, Meta have already adopted AI on the war footing. They are further improving their existing portfolios or building new products that are AI-driven.

But what about the other organizations with not-so-big budgets and billions of revenues? With limited experience and exposure to what works and what doesn't, and fewer best practices, how can they get into the game of AI to build a successful, yet sustainable solution to transform their business.



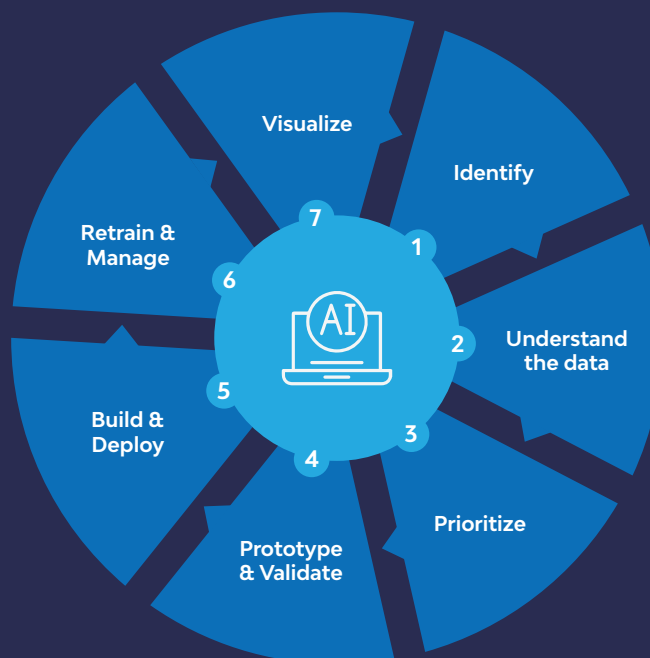
Andrew Ng a well-known figure in the AI world, stated: “Almost all of AI’s recent progress is through one type, in which some input data (A) is used to quickly generate some simple response (B).”

When it sounds so simple how come organizations are battling with its adoption? Well, the simplicity of this technology comes to light only when we understand what it takes to adopt AI-based solutions and the approach towards building one.

Now, there isn’t a golden rule that applies to all organizations and businesses. But a systematic approach is what organizations, both large and small, can focus upon. This article aims to elaborate on this “systematic approach”.

I to V approach to a sustainable AI solution

The I to V approach enables organizations to move systematically with limited risk towards realizing an AI-based solution that can yield business and IT benefits. The approach cuts across **7** stages, initiating with **identifying** a potential opportunity or use case and finally **visualizing** the results of the solution towards business transformation or reduced TCO.



1. Identifying Potential Opportunities

Every innovation/solution starts with an idea. Organizations should ensure that they have a mechanism to generate ideas from inside (employees) as well as outside (customers). Ideas for improvement are generated from all kinds of personas involved in a business process irrespective of them being technical or non-technical. The involvement of different personas helps in identifying opportunities that are both creative and transformational and hence, the opportunity to generate them shouldn't be restricted to a few.

Typically, AI opportunities can be in the form of processes that have the following features:

- Recognition
- Conversation and Human interaction
- Predictive Analytics and decisions support
- Patterns and anomalies detection
- Personalization or recommendations

Once the ideal use case is identified, it's best to start analyzing or discussing with key stakeholders, with simple questions like:

- What are the business problems that can be solved and bring a transformation change?
- Feasibility and ROI?
- What are the data sources for each business problem?
- What could be the potential impact of the solution and its effectiveness over some time?



2. Understand the Data and Its Metadata

Once the use case or opportunity is finalized, the next step is to understand the underlying data and how it can be rightly analyzed for the AI solution. Why rightly? Well, the type of data is the key to the AI solution. More on this in stage 5 i.e. the 'Build' phase.

This stage involves two key activities where most of the time is spent i.e.

- Pre-process, to make raw data more understandable (stemming, normalization, dimensionality reduction, etc.)
- Identify patterns

Pre-process

Understanding the underlying data involves activities such as identifying if the data sources include text, images, video, audio, or any business-sensitive data, then the metadata (metadata – data about data), details i.e. is its employee data, recordings of customer complaints, satellite images of landscapes, etc. Metadata is a must, especially for unstructured data sources, as it makes the data analysis better.

It starts with a discussion with the concerned stakeholders about their understanding and what they believe can be achieved from the data sources involved and the document.

For instance, discussion with the HR leader can help to understand employee data and to what extent the HR believes the data can predict the employee's engagement levels.

The above activities require specific skill sets to ensure it is done right. One of the critical skillsets is that of a data scientist. Data scientists come with vast expertise and experience in handling the volume of data and with great attention to detail.

Identify patterns

The patterns in data reveal a lot of insights. But looking for patterns within a huge volume of data (especially unstructured) becomes very tedious for humans. There will be multiple patterns within the data sources, and once identified it creates the path for a lot of AI solutions. For example, if we have identified an opportunity for an AI solution that helps in making recommendations to new users, looking at the comments by the customers on a product page will not only reveal the sentiment but also an opportunity to understand the buying patterns, potential features, issues from user experience, competitors features, etc.

3. Prioritize the Opportunities

With the list of ideas and feasibility and the associated data sources nailed down; it is time to prioritize the ideas to get into the next step.

For prioritization, consider the below questions:

- What will be the impact (magnitude of transformation or competitive advantage) of the AI solution and how will the success be measured?
- Who will be key stakeholders benefitting from this AI solution?
- Are there any legal or ethical issues for building AI solutions (especially on data privacy)?
- Do we have the skillset to build the AI solution?
- What does it take to sustain the AI solution once deployed?

By asking questions like the ones listed above and analyzing the responses, organizations need to pick one or two “Quick wins” from the list to begin the journey. Those not shortlisted need not be discarded but can be kept as backlogs and can be later assessed and on acceptance adopt the below procedure thereafter.

Prioritization needs to be revisited at least once a year or whenever there is a change in landscape/leadership/business, etc.



4. Prototype and Validate

It's best said, 'Before jumping into the ocean, let's get to know the depth of it and our skills to survive'. Keeping this in mind, the shortlisted opportunities should be converted to a simple prototype. Building a prototype for a potential solution will offer multiple benefits. From a technical standpoint, it will provide the opportunity to try out various ML algorithms (classification, regression, clustering, etc.) that best suits the problem and the data at hand. ML models are the driving force behind AI solutions. More importantly, it will help to build confidence in terms of its impact, usefulness to business, and the overall benefit it can bring once it is productized.

From a business standpoint, it will help to see the immediate impact of a small investment and justify the need for more, based on the results.

When it comes to the selection of the best ML algorithms to solve a given problem, one should know the major categories. They are Supervised, Unsupervised, and Reinforcement learning. Based on the nature of the problem, the algorithm to be chosen and in turn nature of the data is also very important. Below is a quick snapshot of which category of machine learning should be adopted for which type of data.

Prototype	Type of problem/outcome	Nature of data can be used	Common algorithms
Supervised	Predict discrete values (0,1) (True, False) (Spam, Not spam).	Labeled dataset for training	Classification: XGBoost, Naive Bayes, Random Forest, Logistic Regression, etc
	Predict continuous values like the weather forecast.	Labeled dataset for training.	Regression: Linear Regression, Support Vector.
Unsupervised	Pattern detection / Segmentation.	Unlabeled dataset.	Clustering: K-Means.
Reinforcement learning	Learn on the go, like self-driving cars, chess games, etc.	Real-time data to learn on the go.	Markov Decision Process, Q Learning.

Now, a point to note is that one algorithm will not fit to every problem and every dataset. Hence, it requires a lot of research and trying out multiple algorithms before finalizing the solution.

Prototyping will also reveal, the skill gaps, resources needed, effort required, timeframe, and the potential risks associated when building a full-blown solution later.

Building prototypes should be time-bound with a defined goal, scope, and acceptance criteria. The incremental approach best suits an agile approach of sprints (not more than 4 weeks) and having small cross-function teams working closely with business stakeholders involved.

During prototyping, organizations or relevant teams should invest a lot of time and effort in research, understanding data, processing the data, trying out multiple algorithms, and measure/comparing the outcomes with different approaches.

The presence of skillsets like data scientists helps given their expertise in the art of testing the ML models, building a validation framework to ensure the model works best even on unseen data. There are a lot of parameters to measure the outcome quantitatively (like accuracy, precision, recall, F-Measure, the sum of squared error, etc.), but the challenging part is qualitative.

Taking feedback from key stakeholders as we build the prototype helps to understand their needs and satisfaction.

Human-in-the-loop AI takes effect in this stage to mainly validate the outcome and provide feedback. This enables to continue training with the targeted feedback and make the model more accurate.

Once the ML models are validated, next step is identifying the key stakeholders to validate whether it adds value to their line of business.

It's necessary to document the findings during the entire journey of the prototype. For any reason, if any prototype doesn't meet the business expectations, is technically not feasible, has data quality/volume issues, skillset not available, move on to the next idea from the backlog and start the process again.

Nowadays, prototyping can also be done by non-technical personas through **AutoML** - Automated Machine Learning tools available in the market. These tools will help to quickly model and test the solutions. Examples include Google **AutoML**, H2O **AutoML**, Auto-sklearn, etc.



5. Build the Ultimate Solution

Now it is time to move the validated prototypes to a full-fledged solution. After prototyping, we will have more knowledge about how to solve the business problem, what algorithm best fits the data, what type of resources we need, and what will be the effort needed to build the solution.

The business owner (sponsor) who will be funding this AI solution, is a key stakeholder whose buy-in needs are to be secured for the given business case and anticipated ROI.

Next, a cross-function team with data scientists, data analysts, data engineers, SMEs, and an agile coach needs to be set up. The scrum master/agile coach needs to prepare a sprint plan with clear deliverables. The team must showcase incremental demos to all the stakeholders and take feedback as early as possible.

Once completed and deployed, there must be a period where parallel executions should happen (wherever possible) where both AI and humans perform the same actions to produce the outcome. Since most of the algorithms are Black-Box type of models, even the developer will not be able to explain the logic behind the decisions/predictions from the AI solution. Hence, it is important to test the solution on real scenarios and compare its outcome

with the outcome of a human performing the same action with the same dataset. The differences in the outcome need to be analyzed and the necessary remediate actions like retraining the model, fine-tuning the parameters, or more datasets must be taken up.

Explainable AI (XAI)

Recently, there are developments in the field of AI for white box type of models called "Explainable AI (XAI)". This will help to bring transparency on how an algorithm works, how decisions/predictions are made and help in debugging as well. This is invaluable for the users to provide not only the "What", but also the "Why" for every outcome from the model.



6. Retrain and Manage the Solution

Once the model is deployed to production, the process doesn't end here. There must be effective change management in place, to make sure the solution stays relevant, retrained, and effective in the future as well. The impact and benefit of AI solutions also need to be measured continuously.

The deployed model must be retrained to accommodate the changes in data over the period. When the data changes are significant compared to the original dataset

used for training, then retraining becomes very critical for AI solutions to work efficiently. Once the model is retrained and tested, it must be redeployed.

Elimination of bias is crucial in governing the deployed model. Building a test suite to test the model every time it gets retrained and ensure there is no bias for a particular feature/dimension/set of data. Explainable AI can also be used to understand the model's outcome on various test suites.

7. Visualization

Finally, "Visualization" i.e. a set of dashboards that provides the information and insights to the various persona with ease will help to identify the solution impact. The data analyst's role is to create these visualizations by making use of any platforms such as Tableau, Power BI, etc. Visualizations

should be built with the target audience in mind and they should be interactive. The output reveals the value provided to users through AI solutions. Visualization should be interactive to get the users involved and make use of the dashboards.



Conclusion

Every organization and business should start looking (if not done already) at building AI capabilities, frameworks, solutions which will bring the digital transformation at a large scale, sustain, and improve the user experience. Intelligent solutions will not only transform the business or organization but also make humans as superhumans. To all these, there needs to be a method to the madness.

While the overall approach discussed looks simple, generic, and straightforward; organizations rarely follow a systematic approach to such transformation solutions. In reality, most AI solutions don't go through these stages and are developed based on trial and error concepts. Not just the approach, even the skillset, and experience are overlooked. The role of data scientists, data engineers, and data analysts is paramount in the overall journey. However, adopting such solutions arbitrarily can lead to more costs and issues rather than benefits.

Looking at the recent developments in the field of AI like **Creative AI, Language Modeling, Nowcasting, XAI, AutoML**, etc., shows us the speed at which transformations are going around and the need to accelerate to gain advantage, but again there should be a systematic approach as described in this document in which organizations should adopt.



About the Author



Karthikeyan Gangatharan

Associate Principal – Consulting, NWOW Automation, LTIMindtree

Karthikeyan is leading Custom Applications Operations [CAO] Automation Team. He has more than 15 years of experience in the field of IT services with specialization in IT Process Automation (ITPA) and Automation CoE activities for the past 7 years. He was instrumental in creating the automation assessment framework and other processes/standards that are needed for successful automation delivery. He works with various delivery units/accounts within LTIMindtree to assess their automation potential and qualify the automation opportunities/use cases for further solutioning and implementation. He has successfully setup Automation Delivery engine in multiple delivery units of LTIMindtree using BOT Model. He also focuses on developing reusable and end to end AIOPS use cases.

LTIMindtree is a global technology consulting and digital solutions company that enables enterprises across industries to reimagine business models, accelerate innovation, and maximize growth by harnessing digital technologies. As a digital transformation partner to more than 750 clients, LTIMindtree brings extensive domain and technology expertise to help drive superior competitive differentiation, customer experiences, and business outcomes in a converging world. Powered by nearly 90,000 talented and entrepreneurial professionals across more than 30 countries, LTIMindtree- a Larsen & Toubro Group company- combines the industry-acclaimed strengths of erstwhile Larsen and Toubro Infotech and Mindtree in solving the most complex business challenges and delivering transformation at scale. For more information, please visit www.ltimindtree.com.