

Proposal

As more organizations jump on the digital transformation bandwagon using AI and ML, there has been an emergent need to manage the plethora of ML assets generated along their maturity journey. The optimized management of these assets drives ML excellence within an organization (or function), offering actionable business insights and key competitive advantages in an ever evolving technology landscape. In this proposal, the authors present a working framework towards effective ML asset management. This work draws inspiration and builds on the top of the [Gartner Maturity Model for Data and Analytics \(ITScore for Data and Analytics, October 23, 2017\)](#) by introducing a novel concept in the form of **AIML Service Kits**** to help organizations **start, accelerate, scale, synergise** and **transform** in their journey towards ML excellence. As a part of this talk/panel discussion, the authors propose the following:

1. Define ML assets
2. The need of ML asset management
3. Establish the ML asset management features
4. Introduce **AIML Service Kits** using the famous PPT (People, Process and Technology) management framework
5. Contemplate the evolution and future of ML assets management

The authors strongly believe that the proposed AIML Service Kits, explores the solution space corresponding to ML asset management, which is one among the key topics of interest at WAMLM 2022. Using the aforementioned solution, any AI business leader (key participants in the conference workshop) can map their current state, identify the existing gaps and leverage the service kits across people, process and technology to transition into higher ML maturity levels, either incrementally or transformatively.

*****A brief snapshot of the service kits is available at Appendix A of this document***

A short Bio of Ashish Balla

Data Science professional with over 5+ years of industry experience in developing and managing AI/ML solutions at scale. Currently, an AI Consultant with the Global Delivery Center at Google and aiding some of Google Cloud's biggest customers with their digital transformation using Machine Learning, Deep Learning and Natural Language Processing. Passionate to talk about AI ethics, product management & cooking!

Appendix A

Organization ML Maturity*		L-I or Basic	L-II or Opportunistic	L-III or Systematic	L-IV or Differentiating	L-V or Transformational
Definition		Leadership is yet to explore and leverage ML for business decision making.	Leadership is ML aware and leverages it inconsistently for business decision making.	Leadership is ML astute and leverages it consistently, for business decision making, across key functions.	Leadership is ML empowered and leverages it consistently for business decision making, across the enterprise.	Leadership is ML transformative and makes strategic investments towards advancement of AI/ML
ML Challenges	Professional Skills	The data science talent stack consists of data engineers only.	Data scientists, besides L-I professional skills.	ML engineers, besides L-II professional skills.	Product managers and DevOps, besides L-III professional skills.	ML researchers and external AI advisors, besides L-IV professional skills.
	ML Infrastructure	Lack of sufficient storage and compute resources.	Non-optimised utilization of both storage and compute resources.	Optimized utilization of storage resources only.	Optimized utilization of both storage and compute resources.	Use novel Big Data frameworks and ML hardwares (GPUs, TPUs, quantum computers, etc)
	Data	Bronze Standard: Data exists in silos and no single SoT	Silver Standards: Siloed data and project-specific schemas are created.	Gold Standards: Data exists in a datalake with well defined schemas and tables.	Platinum Standards: Data exists on a hybrid cloud, and readily available.	Diamond Standards: Unified data on a hybrid/multi cloud, available via self service.
Driving ML Asset Management						
Kits		AIML Starter Kit	AIML Accelerator Kit	AIML Scaler Kit	AIML Synergizer Kit	AIML Transformer Kit
People		Resource talents like annotation professionals and data scientists.	Add ML engineers to the resources in the Starter kit.	Add PMs to the resources in the Accelerator kit.	Add ML researchers and external AI advisors to the Scaler kit.	Add AI Thought Leaders and advanced researchers to Synergizer kit
Process		<ul style="list-style-type: none"> - Define general data management processes. Document low effort and high value business use case fit for ML application. - Establish Data Science practices and train key staff members. 	<ul style="list-style-type: none"> - Establish functional-specific processes for data procurement and management. - Identify similar ML opportunities by setting up synergies across functions 	<ul style="list-style-type: none"> - Establish enterprise wide data consolidation and management processes. - Document enterprise wide process improvement use cases having high business impact. 	<ul style="list-style-type: none"> - Deploy enterprise-wide reusable and scalable data management practices. - Budgetary provisions for AI research. - Establish AI leadership capability across the enterprise. 	<ul style="list-style-type: none"> - Continuously define and refine data management, enhancement and enrichment processes to shape industry best practices, making organizations efficient and compliant.
Technology		<ul style="list-style-type: none"> - Create project-specific datasets and experimentation notebooks. - Document usecase, business impact and technical architecture. - Exclusively provision compute and storage resources for ML projects. 	<ul style="list-style-type: none"> In addition to the ML assets in the Starter kit, - Publish reusable and curated datasets - Implement feature engineering pipelines - Develop ML pipelines with version control - Document process flow diagrams - Create basic tutorials - Provision cluster computing platform with fault tolerance 	<ul style="list-style-type: none"> In addition to the ML assets in the Acceleration kit, - Create a data marketplace - Establish feature store and centralized repos for shared strategies and implementations - Implement reusable data pipelines - Create ML workflows notebooks - Create advanced tutorials - Move to cloud - Invest in ML project collaboration platforms 	<ul style="list-style-type: none"> In addition to the ML assets in the Scaler kit, - Use cloud AI platforms. - Create advanced notebooks for ML orchestration(eg. Kubeflow notebooks). - Manage version controlled ML environments. - Document E2E process flows and integration architecture. - Create highly interactive training materials and tutorials. 	<ul style="list-style-type: none"> In addition to the ML assets in the Synergy kit, - Publish notebooks/pipelines/workflows for external use - Contribute towards open source community