MahaAgX

MahaAgX is a secure, interoperable and consent-based agricultural data exchange platform that enables researchers, innovators and policymakers to access diverse agricultural datasets for smarter governance, accelerating innovation, and building AI-powered solutions.

https://mahaagx.maharashtra.gov.in/

Goverment of Maharashtra

सत्यमेव जयते

Maharashtra Agriculture Data Exchange

Digital Public Infrastructure for Agricultural Data

An Initiative by

Agriculture Department, Government of Maharashtra

Artificial Intelligence and AgriTech Innovation Center

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

Table of Contents

Executive Summary

Agriculture in Maharashtra operates at a scale and complexity that demands timely, reliable, and trusted data. Information related to land records, crop conditions, weather, markets, inputs, and government schemes already exists across multiple public institutions, research bodies, and private entities. However, this data remains fragmented across systems, stored in different formats, and governed by inconsistent access mechanisms. As a result, its potential to improve farmer outcomes, reduce risk, and enable innovation remains significantly under-utilised.

Digital Public Infrastructure (DPI) has transformed sectors such as identity and payments by empowering both public trust systems and private innovation to thrive independently yet work together seamlessly. Agriculture now requires a similar foundation. Without a neutral and trusted data exchange, agricultural innovation remains constrained to isolated pilots, point integrations, and bilateral data sharing arrangements that do not scale equitably. MahaAgX, the Maharashtra Agriculture Data Exchange, serves as the high-speed "digital conveyor belt." Unlike a traditional database that simply stores information, MahaAgX is a decentralized, consent-driven exchange designed to facilitate secure and standardized flow of data between providers (Public organizations, Private organizations, etc.) and consumers (Startups, Innovators, Researchers, Financial Institutions, etc.). In today's agricultural landscape, timely and reliable data plays a critical role in improving decisions, reducing risks, and helping farmers respond better to changing conditions.

The core objective of MahaAgX is to create a trusted environment for data exchange. Data ownership remains with the original data providers, including government departments and authorized private entities. MahaAgX enforces clear rules for discovery, access, usage, and accountability, ensuring that data sharing strengthens trust while complying with applicable legal and regulatory frameworks.

> By building a DPI for agricultural data through MahaAgX, Maharashtra is laying the foundation for a robust agricultural data ecosystem.

This initiative is being implemented by the Department of Agriculture, Government of Maharashtra, in partnership with the Centre of Data for Public Good (CDPG) at IISc Bengaluru and the World Bank.

MahaAgX brings together critical datasets such as land records, weather, soil health, crop health, crop conditions, and market information within a trusted and interoperable data exchange framework. It will enable ecosystem participants to build solutions such as faster and fairer access to credit and insurance, improved market intelligence, more accurate advisory services, and evidence-based policy making. It will also enable responsible use of advanced analytics and artificial intelligence, powered by high-quality, interoperable, and well-governed data shared through the exchange.

MahaAgX represents a long-term public investment in agricultural inclusion, innovation and sustainability. It reflects Maharashtra's commitment to treating agricultural data as strategic public infrastructure, designed to serve farmers while unlocking sustained value across the agricultural ecosystem.

1. Intended Audience and How to Use This Book

This book is intended for a diverse set of stakeholders who play different roles in Maharashtra's agricultural data ecosystem, as well as for individuals and institutions seeking a deeper understanding of Data Exchange DPI.

For Government Leadership and Policymakers, this book provides a strategic view of why agricultural data must be treated as public infrastructure and how a state-led data exchange can improve service delivery, policy design and long-term resilience in agriculture.

For Startups, Industry, and Innovators, this book outlines how MahaAgX lowers the cost and complexity of accessing trusted agricultural data. It highlights how innovators can build solutions for farmers, markets, finance, and sustainability without negotiating individual data-sharing arrangements with multiple agencies.

For Government Departments and Data Providers, this book explains the principles, governance model, and responsibilities involved in participating in MahaAgX. It clarifies how departments can retain control over their data while enabling secure and responsible reuse across the ecosystem.

For Research Institutions and Academia, this book explains how standardized, interoperable, and governed datasets can support evidence-based research, modelling, and evaluation at scale.

For Other States, International Agencies and Countries, this book presents Maharashtra's approach as a reference model for building sector-specific Data Exchange DPIs that balance innovation with trust and public interest.

Readers may choose to engage with this book in different ways. The early sections establish the vision and rationale for DPI-led data exchange. The middle sections explain how MahaAgX works conceptually and institutionally. Later sections describe participation and outcomes. The book does not require sequential reading and is designed to support selective exploration based on the reader's role and interest.

2. Importance of Data Ecosystem to the Agriculture Sector

India's shift from focusing only on agricultural output to building greater value from farming depends heavily on the availability and use of data. Agriculture contributes around 18-19% to the country's Gross Value Added, yet supports around 54-55% of the workforce[1]. For years, the absence of timely and reliable information has resulted in inefficiencies across markets and unpredictable incomes for farmers.

Strategic Insights for Maharashtra: Maharashtra, with its wide range of climatic conditions and diverse farming practices, is well placed to benefit from a more structured approach to data use. The state's strong presence in crops such as soybean, cotton, pulses, paddy and horticultural crops creates significant opportunities to improve planning, decision-making, and overall agricultural outcomes through better access to and use of data.

Value Proposition of the Data Ecosystem:

De-risking Agriculture:

By bringing together weather updates, soil information, and market trends, the sector can move away from guessing outcomes based on the monsoon and instead make informed decisions that help manage risk more confidently.

Market Intelligence:

Instead of planting based on last year's prices (which often leads to gluts and price crashes), real-time sowing data from across the state will help farmers and policy makers balance demand and supply dynamically.

Monetization of Externalities:

A robust data layer will allow farmers to participate in carbon markets. When information on practices like low tillage or reduced use of fertilizers is properly recorded and verified through the Exchange, it can be converted into carbon credits. These credits can then be sold, giving farmers an additional source of income.

By establishing a structured data ecosystem for agriculture, Maharashtra is poised to transform historical information asymmetries into actionable intelligence - enabling precision decision-making, unlocking new income streams for farmers, and building systemic resilience against climate and market volatility across the entire agricultural value chain.

1. https://agriwelfare.gov.in/en/Dept

3. Role of Data Exchange in Agriculture

In recent years, changing climate patterns and growing pressure on food systems have made digital tools in agriculture increasingly important. What was once seen as an optional upgrade is now essential for ensuring stable farm livelihoods and food security. As India advances toward a digital economy, agriculture, which often struggles with uneven access to information and unpredictable weather, needs strong and reliable digital support. In Maharashtra, a leading state in agricultural production yet highly exposed to climate uncertainty, relying on static and fragmented data is no longer sufficient. The state needs a flexible and connected data ecosystem that can help shift from responding to crises after they occur to planning ahead and building long-term resilience.

The Scale of the Challenge and Opportunity:
Maharashtra holds one of the largest agricultural land footprints, with an estimated 16.5 million hectares of net sown area and over 17.11 million farmers[2]. The important crops of the state are soybean, cotton, Sugarcane, Chickpea, Pigeon pea & other pulses, paddy, Sorghum and horticultural crops. The sheer diversity and volume necessitate a scalable, digital infrastructure.

MahaAgX acts as a trust layer to solve three fundamental challenges:

Unifying the Agrarian Digital Footprint

Today, essential information such as digital land records, soil health data, Weather data, satellite data, and market prices is scattered across different government departments and private systems. MahaAgX brings this information together on a common platform, making it easier to find, access, and use. By enabling secure sharing and consistency, MahaAgX supports better decision-making for farmers, businesses, and policymakers across the agriculture ecosystem.

Accelerating Financial and Market Resilience

Farmers often face barriers to loans, insurance, and better markets because proving their track record can be costly and slow. MahaAgX solves this by letting authorised organizations securely share verified information, such as land ownership, crop history, and yields, for a limited time. This helps banks and insurers make faster, safer decisions, giving farmers easier access to affordable credit, quicker insurance support, and smoother connections to markets and buyers.

Catalyzing Innovation at the Farm Gate

Farming technology has often been accessible only to large, well-funded farms. MahaAgX changes that by bringing together a wide range of agricultural data, it allows Agri-Tech startups to create practical tools powered by AI that help even small farmers make better decisions, use water wisely, prevent pest problems, and get the most out of their resources.

By addressing these interconnected challenges through a unified data exchange infrastructure, MahaAgX transforms fragmented agricultural information into actionable intelligence - enabling Maharashtra's farming community to move from reactive crisis management to proactive resilience building at scale.

2. Economic Survey of Maharashtra 2024-25

4. MahaAgX - A Digital Public Infrastructure (DPI)

MahaAgX is explicitly structured as a Digital Public Infrastructure (DPI), drawing inspiration from the successful India Stack and aligning with the national strategy to leverage data as a shared public asset. DPIs are non-proprietary, interoperable, and foundational platforms that enable the delivery of critical public and private services at a population scale. MahaAgX is positioned to become the core DPI for the state's agricultural data economy, generating high impact by creating a robust, secure, and equitable foundation.

Key Pillars of MahaAgX as DPI and Strategic Impact:

Foundational Layer of Trust (Separation of Concerns):

Like other DPIs that enable trust in key areas, such as UPI for payments or Aadhaar for identity, MahaAgX provides a secure layer that connects users and data. It manages the systems that handle authentication, permissions, and data organization. This setup keeps the platform neutral, ensures users' control over their data, and protects privacy.

Open Standards for Interoperability

The mandate to use open, widely accepted standards (NGSI-LD, OGC, IETF, JSON-LD) ensures easy interoperability. This is crucial for scaling effectively, as it allows any organization, from large companies to small startups, to create new services without technical restrictions, encouraging a competitive & vibrant market for farmer-focused solutions.

National Scalability and Alignment:

MahaAgX is designed in alignment with the India Digital Ecosystem Architecture (INDEA), the national reference framework for building scalable and interoperable Digital Public Infrastructures. This is the same architectural foundation on which the Intelligent Universal Data Exchange (IUDX) has been built and operationalized across multiple deployments. Adopting INDEA principles ensures that MahaAgX is modular, federable, and future-ready, enabling seamless scalability and potential integration with other state-level and national agricultural data initiatives, while contributing to India's broader digital public infrastructure ecosystem.

Economic and Social Inclusion:

The DPI mandate ensures that the platform is built to be widely accessible and fair. By offering verified information at no or low cost, MahaAgX helps solve long-standing gaps in access to knowledge, making financial services, tailored advice, and support for managing risks available even to small farmers, and helping bridge the digital divide in agriculture.

By establishing MahaAgX as a DPI, Maharashtra positions agricultural data as a public good - creating an open, trusted foundation enabling innovation ensuring farmers, service providers, & institutions participate equitably in the state's digital agricultural transformation.

5. Core Actors and Roles

The MahaAgX ecosystem operates through a well-defined actor framework where each participant has specific responsibilities and operates within explicit boundaries. This role clarity ensures accountability is traceable, trust is systematically verifiable, and data flows remain secure and purpose-bound. The governance model designates five primary actors, each fulfilling distinct functions within the data exchange lifecycle:

Data Principal (Farmer)

The core stakeholder and sovereign owner of their personal data. The farmer provides explicit, digital consent (via the Consent Manager) to share their data (e.g., land records, yield history) for specific purposes and benefits from the innovative services provided by AIUs, driving financial inclusion.

Agricultural Information Providers (AIPs)

The custodians & generators of data (both public and private). AIPs are responsible for the quality, accuracy, security, & hosting of their datasets on their Resource Servers, as well as defining the access policy (Data Permit rules). (e.g., Department of Revenue, Weather Agencies, Agricultural Universities).

Agricultural Information Users (AIUs)

The service providers and innovators who consume data to develop value-added applications, decision support systems, and customized farmer services. AIUs must strictly adhere to the permitted use specified in the Data Permit or Consent. (e.g., Fintechs, Agri-Tech Startups, Insurance Companies).

Platform Administrator

The non-profit or government entity responsible for the operation, maintenance, security, and governance of the MahaAgX Control Plane. Their role includes managing the Catalogue, Authorization Server, and maintaining the platform's core standards and compliance.

Delegate

An agent authorized to act on behalf of any user (AIP or the AIU) in administrative or data-sharing matters. Delegates facilitate efficient participation, particularly for large organizations who wish to manage their data access through a trusted intermediary.

These precisely defined roles prevent ambiguity in ownership, usage rights, and compliance obligations - creating a transparent operational model where every data transaction has identifiable actors, enforceable responsibilities, and auditable outcomes across the entire exchange lifecycle.

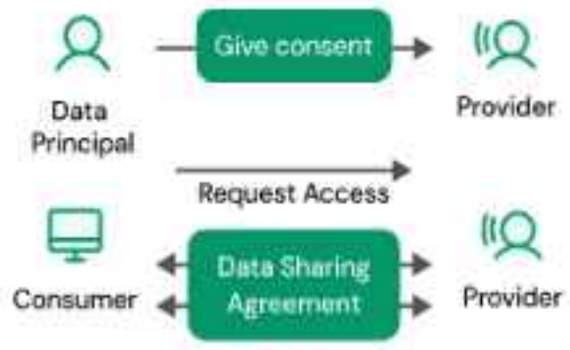
6. Data Governance, Security, and Trust

The MahaAgX governance framework serves two essential purposes: enabling smooth data sharing to drive agricultural innovation, while rigorously protecting farmer privacy and ensuring Agricultural Information Providers retain control over their data.

This section outlines the protective measures built into the system through three interconnected components: 'classification rules' that control who can access what data, 'mutual accountability standards' that ensure all parties use data responsibly, and 'regulatory compliance measures' that protect the rights of farmers and data owners. These safeguards work together as multiple layers of protection, making it significantly harder, both technically and legally, for anyone to misuse data within the ecosystem.

6.1. Data Classification and Access Tiers

All data resources within the MahaAgX ecosystem are strictly classified by the Agricultural Information Provider (AIP) into one of three tiers, which dictate the necessary authentication and authorization requirements for consumption.

Public Data	Restricted Data	Personal Data
		
DEFINITION Non-sensitive, aggregated, or anonymised data for open innovation and free public use.	**DEFINITION** Sensitive commercial data, anonymised granular data, or government datasets that require specific use-case validation.	**DEFINITION** Data linked to a farmer (Data Principal), containing Personally Identifiable Information (PII) like names or Aadhaar ID linkages.
ACCESS Open access for any registered AIU.	**ACCESS** Requires formal approval from the AIP for a specific, limited purpose.	**ACCESS** Requires explicit consent from the Data Principal for a specified purpose.
AUTHORISATION Basic Authentication (AIU registration required).	**AUTHORISATION** Data Permit (AIP approval) and Data Token (issued by Auth Server).	**AUTHORISATION** Individual Consent (via the Consent Manager) and subsequent Data Token.
EXAMPLE (MAHAAGX) Aggregated monthly average market price of Cotton at APMC, District-level Soil Health Status.	**EXAMPLE (MAHAAGX)** Real-time (but anonymized) sensor data from a private weather station	**EXAMPLE (MAHAAGX)** Linking a farmer's name to their 7/12 land record for a loan assessment.

6.2. Core Governance Model (AIP/AIU Accountability)

MahaAgX Core Governance Model establishes a verifiable accountability framework with programmatic enforcement, where every access request, data quality commitment, and usage transaction is auditable in real-time.

MahaAgX implements a bidirectional obligation system: Agricultural Information Providers (AIPs) commit to defined quality standards and access parameters; Agricultural Information Users (AIUs) operate within specified constraints and report data anomalies. Continuous audit mechanisms detect accountability lapses systematically and trigger established escalation procedures.

The following matrix delineates specific responsibilities across critical governance areas:

Governance Area	Agricultural Information Provider (AIP) Responsibility	Agricultural Information User (AIU) Responsibility
Access Control	Defining the usage policy (Public, Restricted) and data retention period. Retains full ownership and revocation rights over datasets.	Requesting access and strictly adhering to the permitted time, volume, and purpose limitations.
Data Quality	Ensuring the published data meets MahaAgX DQA standards and is regularly updated.	Reporting any data anomalies or quality issues discovered during application development.
Usage Audit	Allowing the Platform Administrator to audit access logs against defined policies	Maintaining internal audit logs for data consumption and reporting usage when required.

6.3. Regulatory Compliance: DPDP & Consent Manager

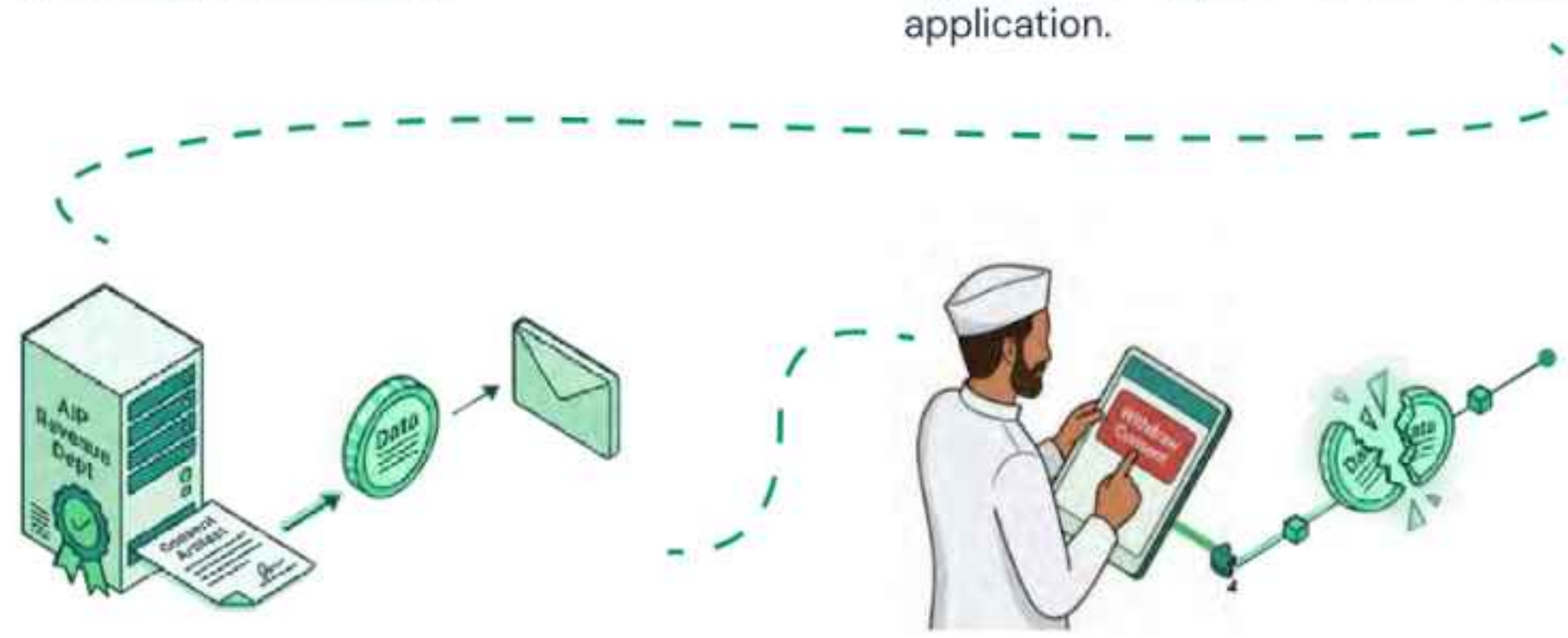
MahaAgX is designed to comply with the Digital Personal Data Protection (DPDP) Act, 2023. This is primarily achieved through a federated Consent Manager architecture.

- **DPDP Alignment:** The entire system enforces the DPDP principles of Purpose Limitation (data must be used only for the consented use) and Data Minimization (only the minimum necessary data should be shared).

- **The Consent Manager:** For any data involving PII (e.g., linking a farmer's name to their land parcel for a loan application), the exchange leverages an integrated Consent Manager (CM).

1. **Consent Request:** An AIU (e.g., a Bank) initiates a request, specifying the exact data fields needed and the purpose ("Lending Assessment").

2. **Farmer Approval:** The request is routed to the farmer (the Data Principal), who reviews and digitally approves or rejects it via a secure application.



3. **Token Release:** Upon approval, a digitally signed Consent Artifact is created. Only upon validation of this artifact does the AIP (e.g., Revenue Dept) release the sensitive data to the AIU.

4. **Revocability:** Farmers maintain the Right to Withdraw Consent at any time, immediately terminating the AIU's access token.

Together, these governance mechanisms establish a structured ecosystem where agricultural innovation advances within clearly defined ethical and legal frameworks - ensuring data serves as an enabler of farmer welfare and sustainable development.

7. Data Lifecycle Management: The Technical Backbone

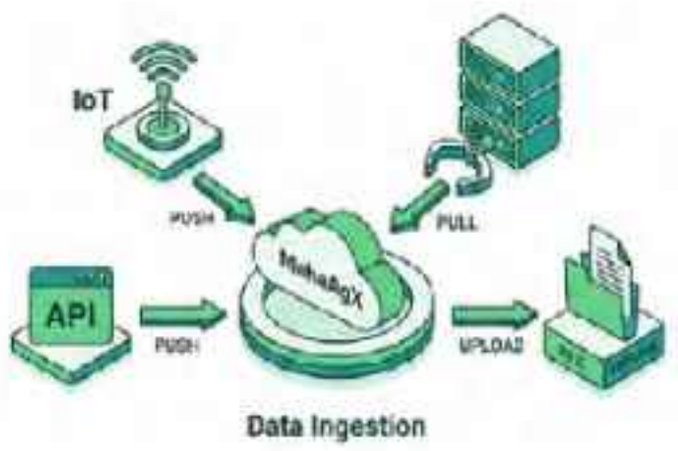
Technical standards require systematic enforcement mechanisms to ensure operational effectiveness. This section defines the protocols that ensure data within the MahaAgX ecosystem maintains interoperability, reliability, and fitness for agricultural applications throughout its entire lifecycle - from ingestion to consumption.

MahaAgX implements data lifecycle management as a continuous quality assurance system where each stage - onboarding (ingestion and transformation), assessment, evaluation, description, and exchange - functions both as an independent quality gate and an integrated component of the broader trust infrastructure. Data quality, security, and compliance are embedded into operational processes through automated monitoring and audit mechanisms.

7.1. Data Onboarding Interactions

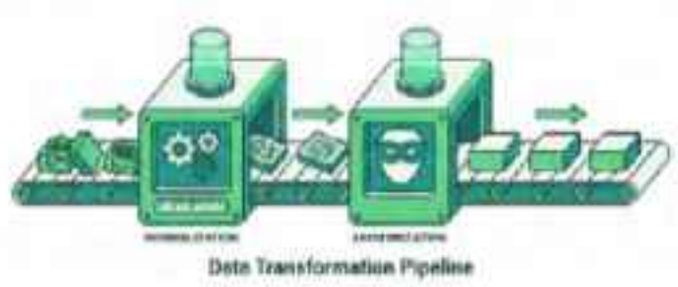
MahaAgX simplifies the transition from "siloed data" to "exchange-ready data" through flexible ingestion pathways.

1. **Standardized Metadata:** Every resource must carry a non-negotiable metadata "passport." This includes Resource IDs, update cadences, clear access classifications (Public vs. Restricted) and geospatial footprints (Polygons/Points).



Data Ingestion

2. **Data Ingestion:** AIPs can either utilize Push-based mechanisms (real-time data streams from IoT sensors or APIs) or the MahaAgX can employ Pull-based adaptors to fetch data from legacy government systems (e.g., departmental databases). If the data is unstructured and a static one-time push, the AIPs can utilize MahaAgX file services to upload the datasets.


Data Transformation Pipeline

3. **Data Transformation Pipeline:** Before onboarding the dataset into the platform, it passes through a mandatory transformation pipeline to ensure compliance with MahaAgX standards:
   - **Normalization:** Ensuring units, formats (e.g., date/time), and coding schemes are consistent across the exchange.
   - **Anonymization/Masking:** Implementing mandatory masking, aggregation, or pseudonymization of Personally Identifiable Information (PII) at the source or the ingress point. Unless explicit consent is obtained via the Consent Manager, no PII is permitted in the exchange.

7.2. Data Assessment

MahaAgX treats data quality as a dynamic metric, not as a one-time check. Every dataset receives a Data Readiness Score and a report generated using a hybrid model that uses sophisticated, custom algorithmic methods as well as Large Language Model inputs with safeguards that provide semantic context to evaluate the following metrics:

Completeness: Automated validation of mandatory fields (e.g., Gat numbers, soil health parameters).

Consistency: Logic-gate checks (e.g., ensuring "harvest date" is after "sowing date" and yield values sit within historical biological bounds).

Timeliness: Real-time monitoring of latency - crucial for "perishable" data like weather alerts or commodity prices.

Granularity Verification: Ensuring the data resolution (Village vs. District) matches the metadata claim.

7.3. Data Evaluations

Data Evaluations encompass the ongoing monitoring, auditing, and continuous improvement mechanisms necessary to maintain trust and relevance in the Exchange.

Continuous Evaluation

MahaAgX platform continuously monitors data usage patterns, token request volumes, and AIP uptime to ensure the reliability of the platform.

Usage Audit

MahaAgX servers maintain immutable logs of all data transactions, detailing who accessed the data, when, for what purpose, and the volume consumed. These logs are auditable by the Platform Administrator.

Actionable Feedback

If an AIP's data frequently fails DQA checks or generates user complaints, the platform provides actionable feedback to the AIP for remediation, maintaining a publicly visible data quality score in the Catalogue to guide AIU decision-making.

7.4. Data Description

Data is only useful if it is understood. MahaAgX platform mandates a uniform language for describing all resources to ensure interoperability.

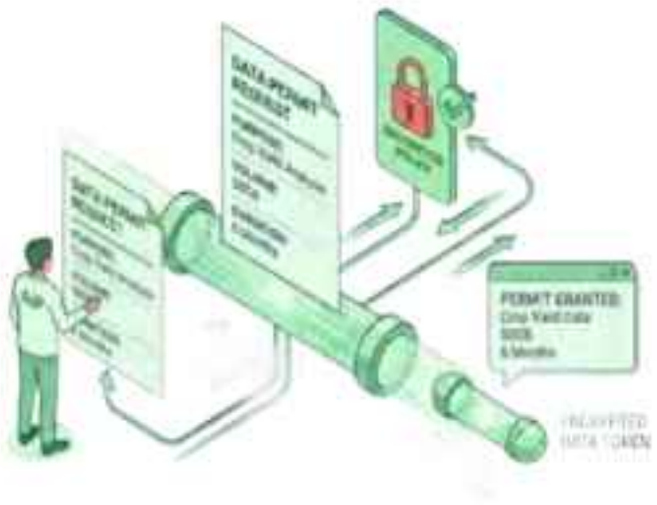
- **Standardization Data Descriptions:** All data resources must be described using JSON-LD (Linked Data). This ensures that every piece of data is contextualized (e.g., it's not just a number, but **soilMoisture** measured at **Location X** at **Time Y**).
- **Metadata:** Rich, non-negotiable metadata must accompany every resource, including the resource ID, geospatial footprint (Lat/Long or Polygon), update frequency, license type, and the AIP's details.

7.5. Data Exchange Mechanism & Permits

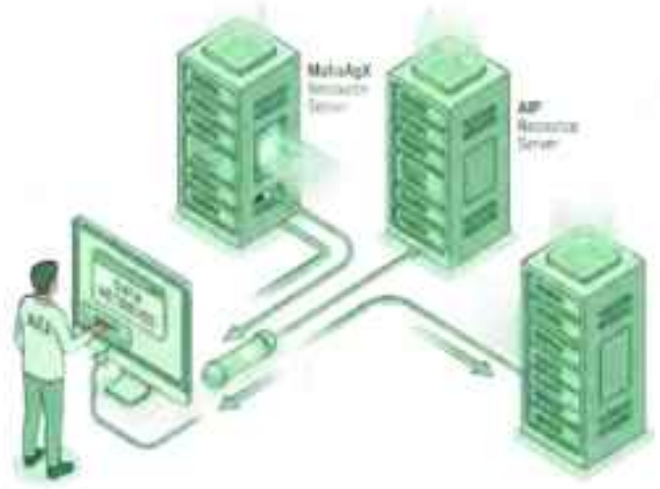
The exchange uses a zero-trust model, meaning access to data resources is not automatic, but granted on a per-request basis and tightly controlled by the AIP. This mechanism ensures data sovereignty and adherence to the defined usage policy.

1. **Discovery:** An AIU searches the MahaAgX portal for a dataset, viewing its metadata and the AIP's access policy (Open, Restricted).

2. **Permit Request:** If the data is restricted, the AIU sends a formal request (a Data Permit) to the AIP, outlining the specific use-case (Purpose), the volume needed, and the required duration.



3. **Authorization:** Upon AIP approval, the MahaAgX Authorization Server issues a short-lived, encrypted Data Token, which encapsulates the granted permissions (the Permit).



4. **Data Flow:** The AIU uses this token to retrieve the data directly from the MahaAgX Resource Server or from the AIPs Resource server.

Collectively, these data lifecycle protocols ensure that data entering the MahaAgX ecosystem is not merely technically accessible, but institutionally trustworthy, legally compliant, and operationally valuable - transforming fragmented agricultural information into a strategic asset for sustainable innovation.

8. MahaAgX Platform

The MahaAgX Platform represents a comprehensive cloud-based solution designed to revolutionize agricultural data management and enable data-driven decision-making across the agricultural ecosystem. MahaAgX provides a secure, scalable, and interoperable platform for aggregating, processing, and disseminating agricultural data for public good.

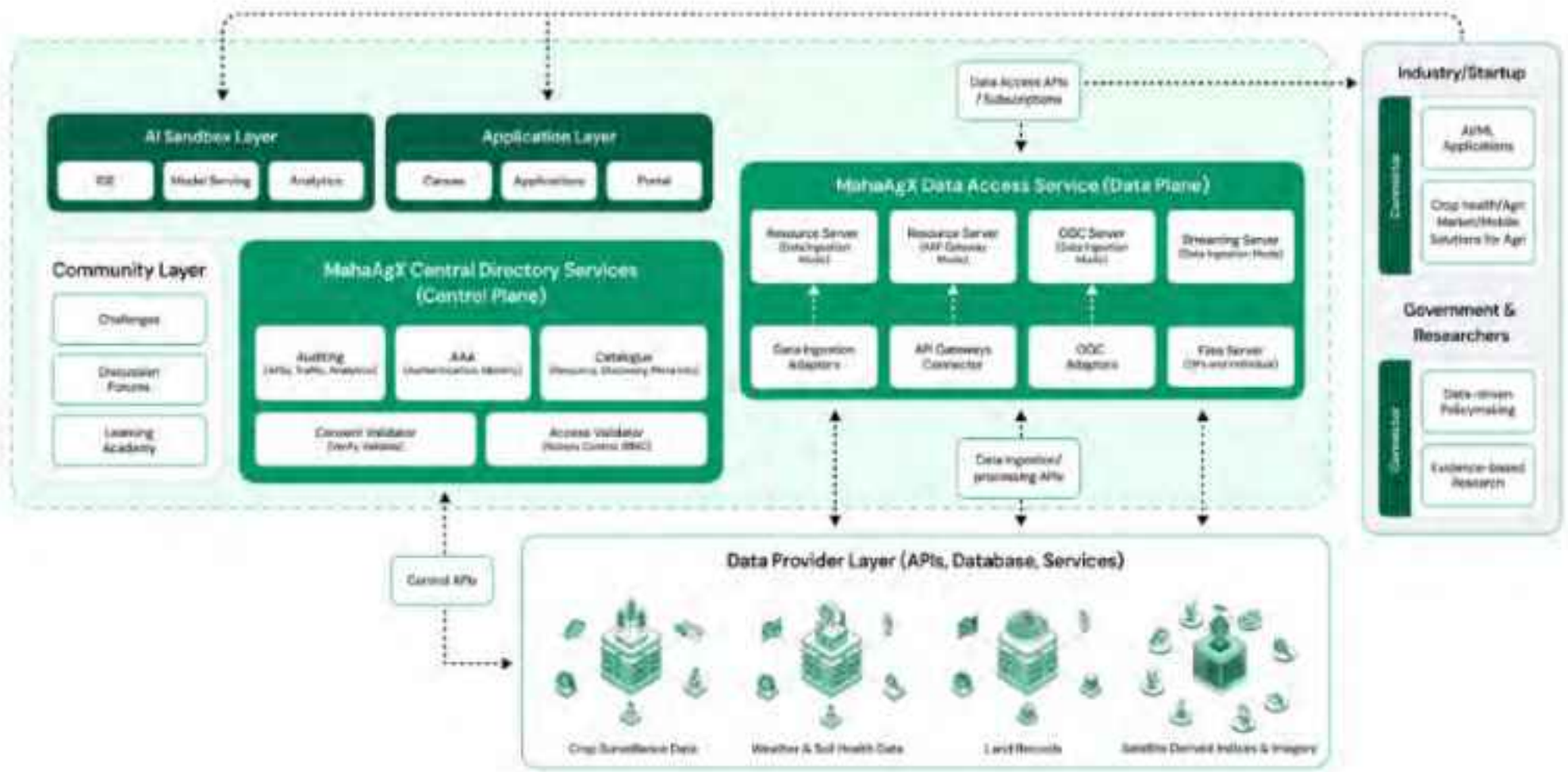


Figure 2: Components and modules of MahaAgX Platform

MahaAgX is architected as a multi-layered cloud platform that facilitates seamless data exchange between diverse stakeholders. The platform addresses critical challenges in agricultural data management including data silos, lack of standardization, access control complexity, and the need for real-time processing capabilities.

MahaAgX operates as a six-layer integrated system. The 'Data Provider Layer' aggregates agricultural datasets from government departments, universities, and field systems. The 'Data Plane' standardizes this heterogeneous information using international protocols (NGSI-LD, OGC).

The 'Control Plane' manages authentication, authorization, and auditing to enforce governance. The 'AI Sandbox' provides GPU-accelerated computing resources for innovation. The 'Application Layer' offers portals and visualization tools that convert data into actionable insights. The 'Community Layer' accelerates solution development through competitions and collaborative forums. These layers work together to enable secure, standards-based data exchange that supports evidence-based decision-making across Maharashtra's agricultural ecosystem. Each component is detailed below:

8.1. Data Provider Layer

At the foundation of the MahaAgX platform lies a robust Data Provider Layer that consists of diverse agricultural datasets, including:

AgriStack	Foundational farmer registry and land records integration
Farmer Field School (FFS)	Crop condition, soil moisture, pest infestation etc. data captured during technology demonstrations to farmers at field-level.
CROPSAP	Disease and pest surveillance data from 16,000+ locations, collected digitally on a weekly basis through field machinery
Pest & Disease Registry	Crop wise pest and disease registry, diagnostic symptoms, treatment protocols, and geographic distribution.
Seed Registry	Notified varieties with yield and agronomic traits, zone-specific recommendations, and disease resistance profiles
Pesticide datasets	Registry of approved and banned substances, laboratories, manufacturers, importers, exporters, dealers, pesticide licenses, Certificates of Registration, and manufacturing facilities
Soil Health datasets	Soil Nutrient profile, pH levels, organic carbon content.
Satellite-derived indices & imagery	Spectral datasets supporting vegetation indices, water stress indicators, crop classification, and phenological monitoring
Weather data	Live and historical meteorological parameters from 2,300+ Automatic Weather Stations (AWS)
Crop Cutting Experiment (CCE) data	Sample field-harvest plot data used to estimate average crop yield.
Market intelligence	Live and historical mandi prices and arrival volumes from 300+ Agricultural Produce Market Committees (APMCs)
Scale of Finance (SOF)	Standardized credit norms and per hectare financing values for crops, animal husbandry, fisheries, and beekeeping operations
Infrastructure data	Information from 200+ warehouse centers, Krishi Vigyan Kendras (KVKs), cooperative banks, and agricultural service centers

> Also, for the first time, research data of four state agricultural universities is now made accessible for the innovators.

All this heterogeneous data is exposed through standardized APIs and control mechanisms, enabling uniform access patterns regardless of the underlying data source characteristics.

8.2. Data Plane

MahaAgX data plane represents the core processing infrastructure of MahaAgX, similar to a "data highway", which is used for implementing multiple "data route" based operational modes to accommodate diverse data consumption patterns using multiple international standards. These multiple "operational modes" implemented to achieve the "data highway" are as follows:

Resource Server (Data Ingestion Mode)

Like a massive grain silo, resource server collects huge batches of data (e.g., annual census records), cleans them, and stores them. To achieve this, MahaAgX Resource server implements European Telecommunications Standards Institute (ETSI), Next Generation Service Interfaces for linked data (NGSI-LD) - an API specification originally developed for context information management standard which handles high-volume batch data ingestion from multiple providers, performing validation, transformation, and persistence operations. Supports scheduled data pipelines and ETL workflows.

Streaming Server:

Instead of waiting for a daily report, applications can process high-frequency telemetry data with low latency using MahaAgX streaming services. This allows developers to build "Instant-Alert" systems—for example, automatically turning on an irrigation pump the moment soil moisture drops too low or sending a frost warning to a smartphone as the temperature hits a critical point. MahaAgX manages real-time data streams from IoT devices and sensors, processing high-frequency telemetry data with low latency for time-critical agricultural monitoring using low latency databrokers.

Resource Server (Gateway Mode):

Like a drive-thru window, resource server in gateway mode allows an app to ask for a specific piece of info (e.g., "Is Farmer X verified?") and gets an instant, live answer directly from the provider's database. In order to achieve this, MahaAgX Resource Server provides synchronous API endpoints for real-time querying, allowing Data consumers to directly integrate their databases and applications with the MahaAgX Data Plane.

OGC Server:

Just as a GPS app layers traffic, terrain, and street views on top of each other, the OGC Server takes complex geographic data from different sources—like government land boundaries and private satellite imagery—and aligns them perfectly so an app can "see" exactly which crop is growing in which specific field. MahaAgX implements Open Geospatial Consortium standards for geospatial data ingestion, enabling interoperability with GIS systems and spatial analysis tools in order to achieve this.

8.3. Control Plane

MahaAgX control plane serves as the nerve center of the platform, orchestrating all platform operations through sophisticated service components:

Authentication, Authorization, and Accounting (AAA): Implements role-based access control (RBAC) and attribute-based access control (ABAC) mechanisms. Manages user identities, validates credentials through OAuth 2.0 and OpenID Connect protocols, and maintains comprehensive audit logs of all data access activities.

Catalogue (Resource Discovery Meta Info): Provides comprehensive metadata management and discovery capabilities. Maintains searchable indices of all datasets, APIs, and resources with rich descriptive information including data schemas, quality metrics, update frequencies, and provenance information.

Consent Validator (Policy Validator): Enforces granular consent policies defined by data providers. Validates data access requests against user permissions, usage constraints, and temporal restrictions. Ensures compliance with data protection regulations and organizational policies.

Auditing (APIs, Traffic, Analytics): Implements comprehensive monitoring and auditing infrastructure. Tracks all API calls, data transactions, and system events. Generates analytics on platform usage patterns, popular datasets, and system performance metrics.

Access Validator (Access Control RBAC): Implements fine-grained authorization checks at the resource level. Evaluates user roles, resource ownership, and contextual factors to determine access permissions in real-time.

8.4. AI Sandbox Layer

MahaAgX offers a dedicated innovation environment provided through a secure data virtualization layer that democratizes access to AI/ML capabilities for agricultural applications by providing isolated development environments with enterprise-grade GPU infrastructure and pre-configured tools. Features include GPU-accelerated Jupyter notebooks (IDE) for deep learning workloads, ML model serving infrastructure (Model Serving) for deployment and inference, and advanced analytics tools for exploratory data analysis.

MahaAgX implements a credit-based usage model that democratizes access to expensive computational resources. Startups and ecosystem participants receive allocated credits enabling immediate access to GPU clusters without upfront infrastructure investment. Transparent cost tracking via OpenCost integration automatically deducts credits based on actual GPU hours, CPU time, and storage consumed, ensuring fair resource allocation and eliminating financial barriers to AI innovation.

8.5. Application Layer

MahaAgX application layer provides user-facing interfaces and analytical capabilities using inbuilt portals, canvas based visualisation and data driven applications:

MahaAgX Portal

The centralized portal interface provides comprehensive platform capabilities:

- **Catalogue:** Searchable repository of all available datasets and APIs with advanced filtering and discovery features
- **Admin Portal:** Platform administration interface for user management, resource provisioning, and system configuration
- **Notebook IDE:** Integrated development environment for data scientists and analysts to perform interactive analysis

MahaAgX Canvas

Interactive visualization interfaces with map and temporal plot visual interfaces for real-time plotting of datasets such as agricultural metrics, crop health indicators, environmental parameters, etc., within the platform.

Applications

Domain-specific applications for use cases such as irrigation optimization, Disease and pest management, yield prediction, Price Forecasting etc. can be built on standardized APIs for consistent data access patterns. These applications can also be developed by start-ups and ecosystem participants to enable data driven decision making.

8.6. Community Layer

MahaAgX creates a vibrant innovation ecosystem through a Kaggle-inspired competitive platform. Organizations and government bodies can announce agricultural challenges (e.g., crop disease detection, yield prediction, Price forecasting) with defined problem statements, evaluation metrics, and prize pools. Participants, ranging from university researchers to agri-tech startups, submit solutions that are evaluated against holdout test datasets, with real-time leaderboards fostering healthy competition. This challenge-driven approach accelerates solution development by crowdsourcing innovation from diverse perspectives, transforming complex agricultural problems into solved use cases speedily.

Knowledge Collaboration & Community Discourse:

The discussion forums enable deep technical discourse where participants collaborate on feature engineering strategies, debate model architectures, and troubleshoot implementation challenges, creating a knowledge repository that benefits the entire agricultural AI community.

From Competition to Production:

Successful solutions transition from competition winners to production-ready applications, directly impacting thousands of farmers. This comprehensive community infrastructure establishes MahaAgX as not just a data platform but an innovation catalyst that transforms agricultural research from isolated academic exercises into collaborative problem-solving ecosystems delivering measurable field impact.

8.7. Summary and the Road Ahead

Overall, MahaAgX represents a paradigm shift in agricultural data management, transforming fragmented data silos into a unified, accessible, and actionable data ecosystem. By providing standardized APIs, robust access controls, and powerful analytical capabilities, the platform enables stakeholders across the agricultural value chain to make data-driven decisions that improve productivity, sustainability, and farmer livelihoods.

The platform's sandbox environment democratizes access to advanced AI/ML capabilities, fostering innovation by startups and researchers who would otherwise lack the infrastructure to develop agricultural intelligence applications. This creates a vibrant ecosystem where novel solutions can be rapidly prototyped, validated, and scaled.

MahaAgX platform is designed to be compliant with the DPDP Act requirements. As we continue to evolve, proposed enhancements include confidential clean rooms, secure multi-party computing for federated learning capabilities and privacy-preserving model training, consent management, integrated with interactive data visualization tools, no-code and low-code environments and sophisticated tools for data providers to integrate with the platform. These developments will further strengthen the platform's role as the foundational infrastructure for digital agriculture in Maharashtra and serve as a replicable model for agricultural data exchanges globally.

9. Potential Use-Cases that arise out of MahaAgX

Infrastructure value manifests through the applications it enables. MahaAgX's structured data exchange unlocks transformative applications across the agricultural value chain, converting fragmented data into actionable intelligence that addresses critical challenges for farmers, administrators, and enterprises.

The following potential use-cases demonstrate outcomes: reducing input costs through precision agriculture, mitigating risks through predictive analytics, improving market access through information transparency, and ensuring program integrity through verification mechanisms. Each application leverages multiple interoperable data sources, creating value that wouldn't have been possible in fragmented data environments:

Precision in Irrigation Scheduling

DATASETS REQUIRED

Digital Crop Survey/ e-Pik Pahni, Weather Forecasts, Soil Moisture, KC Values / Crop coefficients.

HOW IT CREATES VALUE

Calculates crop water needs based on soil type, growth stage, and rainfall predictions to generate optimal irrigation schedules that conserve water while maintaining yields

PRIMARY BENEFICIARIES

Farmers, FPOs, Extension Officers,

Harvest Timing Optimization

DATASETS REQUIRED

Digital Crop Survey, Weather Forecasts, APMC Prices, Warehouse Data

HOW IT CREATES VALUE

Analyzes crop areas, weather and advises harvest timing across districts, recommending early / late harvest or Predicts market gluts and suggests to store the produce in a warehouse to avoid distress sale.

PRIMARY BENEFICIARIES

Farmers, FPOs, Warehouse Operators

Seed Variety Performance Tracking

DATASETS REQUIRED

Seed Registry, Digital Crop Survey, CropSAP, Soil Health Card, Weather Data

HOW IT CREATES VALUE

Aggregates real-world yield data across thousands of farms to rank seed varieties by performance in specific soil and climate conditions, replacing marketing claims with evidence

PRIMARY BENEFICIARIES

Farmers, FPOs, Seed Companies, Agricultural Universities

Pest Outbreak Prediction

DATASETS REQUIRED

Weather Forecasts, CropSAP Data, Pest Registry

HOW IT CREATES VALUE

Analyzes pest surveillance data, weather patterns, and crop stages to predict outbreaks days in advance, enabling timely preventive spraying instead of reactive damage control

PRIMARY BENEFICIARIES

Farmers, FPOs, Extension Officers,

Pest and Disease Identification & Control Measures

DATASETS REQUIRED

Pest & Disease Registry, Crop Registry, Pesticide Registry, SAU Research Database, Weather Data

HOW IT CREATES VALUE

Analyzes farmer-uploaded crop photos using AI image recognition to identify pest/disease within seconds, matches symptoms against diagnostic database, recommends appropriate treatment options based on economic thresholds and weather conditions, provides bilingual advisories with nearby pesticide dealer locations and application instructions.

PRIMARY BENEFICIARIES

Farmers, Extension Officers, Pesticide Retailers

Custom Fertilizer Blending

DATASETS REQUIRED

Seed Registry, Digital Crop Survey, CropSAP, Soil Health Card, Weather Data

HOW IT CREATES VALUE

Aggregates real-world yield data across thousands of farms to rank seed varieties by performance in specific soil and climate conditions, replacing marketing claims with evidence

PRIMARY BENEFICIARIES

Farmers, FPOs, Seed Companies, Agricultural Universities

Maturity-Based Harvest Optimization for Sugarcane and Horticultural Crops

DATASETS REQUIRED

Digital Crop Survey/e-Pik Pahni, Remote Sensing Data, Weather Forecasts, Crop Age/Planting Date Registry, Field Observations (Brix levels for sugarcane, color/firmness for fruits), Historical Yield Data, Market Price Intelligence, Transportation/Cold Storage Availability

HOW IT CREATES VALUE

Predicts optimal harvest windows by analyzing crop maturity indicators (sucrose content in sugarcane, ripeness markers in fruits/vegetables) against weather patterns and market demand. Generates harvest schedules that maximize sugar recovery rates in sugarcane, minimize post-harvest losses in perishables, and align harvesting with peak market prices and logistics availability. Reduces premature or delayed harvesting that degrades quality and profitability.

PRIMARY BENEFICIARIES

Farmers, FPOs, Sugar Mills, Food Processing Units, Cold Chain Operators, Market Aggregators

Farm Gate Logistics Optimization

DATASETS REQUIRED

Digital Crop Survey, GPS Network, APMC Prices, Weather Forecasts

HOW IT CREATES VALUE

Clusters farmers by location and harvest timing, generates optimal collection routes, and coordinates shared transport to reduce costs and post-harvest losses

PRIMARY BENEFICIARIES

Farmers, FPOs, Cold Chain Operators

Farm Machinery Custom Hiring Service availability

DATASETS REQUIRED

MahaDBT, GPS Locations, Weather Forecasts

HOW IT CREATES VALUE

Connects farmers needing equipment with nearby available custom hiring centers and other service providers to rent out the machines, optimizes usage routes,timely availability, Price discovery of service and makes mechanization affordable.

PRIMARY BENEFICIARIES

Farmers, FPOs providing Custom Hiring Center services

Fodder Demand-Supply shortage mitigation

DATASETS REQUIRED

Livestock Census, Weather Forecasts, DCS

HOW IT CREATES VALUE

Forecasts fodder shortages in advance by analyzing drought predictions and livestock populations, triggers imports and matches surplus regions with deficit areas to prevent price spikes and distress livestock sales.

PRIMARY BENEFICIARIES

Cattle Rearers, Dairy industry and Cooperatives, Policy Makers, Drought Relief Programs

Smart Crop Loan

DATASETS REQUIRED

Digital Crop Survey, Scale of Finance of crops

HOW IT CREATES VALUE

Calculates farmer credit score, using yield history, and potential income, enabling faster loan approvals

PRIMARY BENEFICIARIES

Farmers, FPOs, Banks, NBFCs

Farm-to-Fork Traceability

DATASETS REQUIRED

Digital Crop Survey, Soil Health Card, Pesticide Records, APMC Transactions, Cold Chain IoT

HOW IT CREATES VALUE

Creates blockchain-verified supply chain records from farm to consumer, enabling instant recall during contamination events and commanding premium prices for traceable organic/export produce

PRIMARY BENEFICIARIES

Farmers, FPOs, Exporters, Food Safety Authorities, Consumers

Relief and Rehabilitation

DATASETS REQUIRED

Weather Data, Satellite Imagery, Crop Spectral Signatures, Digital Crop Survey

HOW IT CREATES VALUE

Monitors weather triggers and verifies crop losses automatically using satellite data, enabling quicker relief measures decision making.

PRIMARY BENEFICIARIES

Farmers, Government

These use-cases reveal how MahaAgX operationalizes data as a strategic asset - transforming previously inaccessible agricultural information into practical tools that enhance farm profitability, expand financial inclusion, and enable evidence-based governance across Maharashtra's agricultural ecosystem.

10. Benefits of MahaAgX

The value of infrastructure lies in the tangible outcomes it creates for its users. MahaAgX is designed to deliver strategic benefits across the agricultural value chain by systematically reducing friction, risk, and operational costs for every stakeholder while unlocking opportunities that have historically remained inaccessible.

10.1. Benefits for the Farmer (Data Principal)

Financial Inclusion & Access to Credit:

By sharing verified and approved information such as land details and farming history, the farmer can overcome gaps in knowledge, making it easier and quicker to get loans from banks and financial services at better terms.

Risk Mitigation & Insurance:

Verified data helps improve crop insurance by setting fair premiums, speeding up claims, and offering better protection against weather and climate risks.

Precision and Efficiency:

Farmers receive hyper-local, actionable advisories (e.g., precise fertilizer prescriptions, irrigation schedules) that minimize input wastage and optimize yields, improving profitability.

Data Sovereignty:

The Consent Manager ensures control over who uses the personal data, for what purpose, and for how long, guaranteeing privacy and trust.

10.2. Benefits for the Government (AIPs and Policy Makers)

Evidence-Based Policy:
Real-time, standardized data on crop yields, market transactions & weather conditions allows for accurate, timely decision-making regarding procurement, and disaster relief.

Efficient Scheme Delivery:
Government programs (e.g., direct benefit transfers, farmer registration) can be targeted more effectively using verified data linkages, reducing leakage and administrative overhead.

Resource Management:
Data integration helps in macro-level planning for water resource allocation, storage capacity, and infrastructure development based on accurate forecasting.

10.3. Benefits for the Private Sector (AIUs and Innovators)

Reduced Data Acquisition Costs:
AIUs gain access to a standardized, unified catalogue of validated data (from diverse AIPs) without having to build bilateral agreements or complex integration layers with every data source.

Accelerated Innovation:
Standardization and clear access mechanisms allow startups and large corporations to focus on developing applications rather than cleaning and standardizing raw data.

Creation of New Markets:
The reliable, secure flow of data enables the development of entirely new financial products (e.g., parametric insurance, supply chain traceability tools) previously infeasible due to lack of trusted data.

Collectively, these benefits establish MahaAgX as a strategic enabler of Maharashtra's agricultural ecosystem - where infrastructure investments translate into sustained improvements in farmer incomes, government efficiency, and private sector innovation across the entire value chain.

11. Guide for Departments to Build Data Exchanges

MahaAgX has been designed not only as a solution for agriculture, but as a reference model for how domain-specific Data Exchanges can be developed across departments. The approach, principles, and governance mechanisms described in this book are reusable and adaptable to other states and sectors such as health, transport, water, urban development, education, and social welfare, etc. Departments seeking to establish their own Data Exchange can draw from the following guiding principles:

Start with Data Stewardship, Not Technology
A successful Data Exchange begins by identifying high-value datasets, understanding their ownership, sensitivity, and usage potential, and defining clear rules for sharing. Technology should serve these decisions, not drive them.

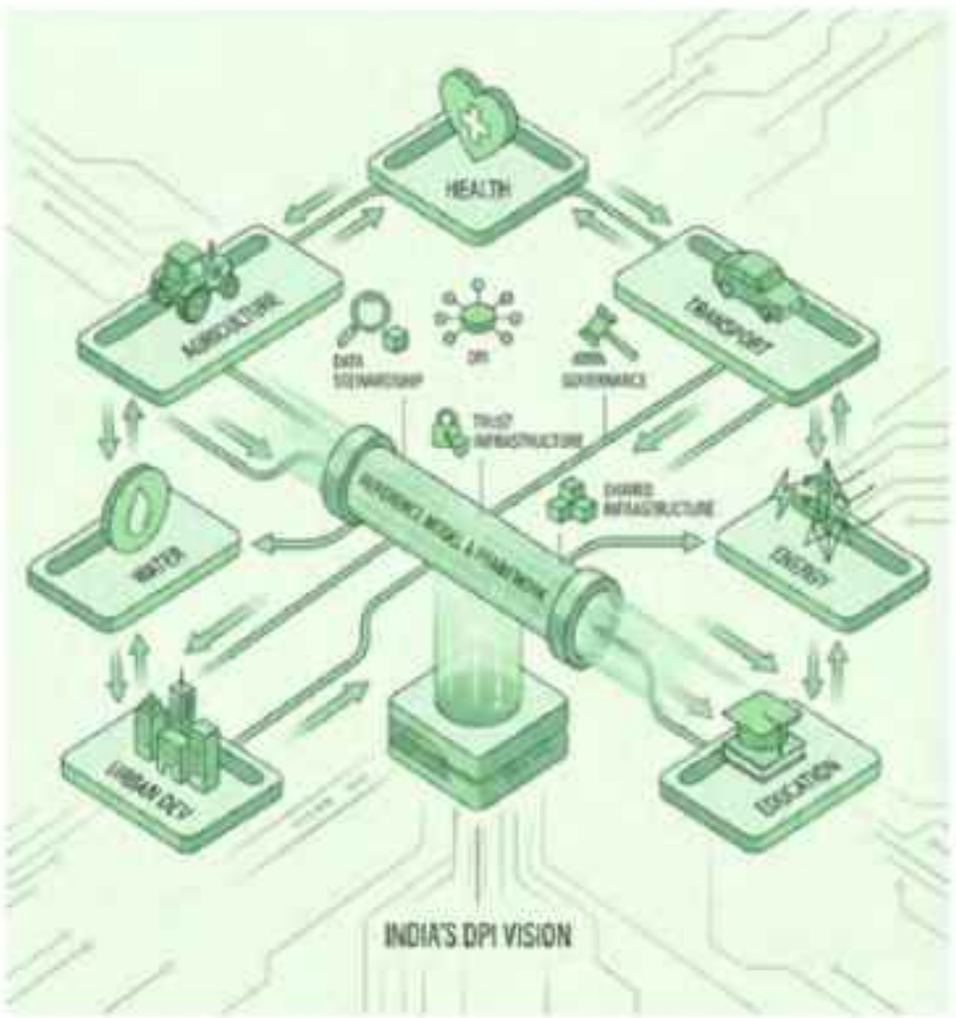
Adopt DPI Principles from the Outset
A Data Exchange should function as a digital public infrastructure rather than as a standalone application. This means prioritizing openness, interoperability, federated data ownership, and ecosystem participation, while avoiding centralized data hoarding.

Separate Trust Infrastructure from Applications
The Data Exchange should provide discovery, access control, consent, and audit mechanisms. Applications, analytics, and dashboards should be built by departments or ecosystem partners on top of the exchange.

Embed Governance and Consent by Design
Clear roles, accountability, and access rules must be defined upfront. Where personal or sensitive data is involved, consent-based access and purpose limitation should be enforced as core features, not add-ons.

Leverage Shared State Infrastructure & Standards
Wherever possible, departments should reuse common standards, identity systems, consent mechanisms, and exchange components developed at the state level. This reduces duplication, accelerates deployment, and ensures interoperability across sectors.

By following these principles, departments can build Data Exchange platforms that are aligned with India's broader DPI vision, interoperable with MahaAgX and other exchanges, and capable of supporting innovation while safeguarding public interest.



12. Summary: Strategic Imperative

The Maharashtra Agriculture Data Exchange (MahaAgX) represents a fundamental shift in how the state views and manages its most critical sector. By enabling seamless data exchange between government, industry, research institutions, and farming communities, MahaAgX bridges the digital divide and ensures that the benefits of agricultural technology reach even smallholder farmers, contributing directly to the vision of sustainable and inclusive agricultural development. MahaAgX will:

Boost Farmer Income

By enabling AIUs to build personalized, low-cost services for credit, insurance, and market linkages, MahaAgX will contribute to reducing risks and improving price discovery for the Data Principal (the farmer).

Ensure Policy Effectiveness

The government gains access to real-time, high-quality data (e.g., Market Prices, Weather, etc.) leading to evidence-based policymaking, better resource allocation, & highly targeted scheme delivery.

Accelerate Digital Innovation

By standardizing data and simplifying access via the Exchange mechanism, MahaAgX drastically lowers the barrier to entry for Agri-Tech startups, fostering a vibrant ecosystem of innovation within the state.

Guarantee Data Sovereignty

Through the strict enforcement of DPDP and the Consent Manager, the farmer retains control over their personal data, ensuring that the digital revolution in agriculture is inclusive, ethical, and equitable.

Thus, MahaAgX serves as both technical infrastructure and institutional framework – enabling Maharashtra to advance agricultural digital transformation while ensuring that economic growth and farmer welfare progress as complementary outcomes of well-designed public systems.