Regulating Smart Agriculture in India: A Socio-Technical Framework for IoT and Machine Learning Data through Global Comparison

Dr. Tanaya P Kamlakar[1*], Ms. Ruchira Halli[2], Dr. Deepa Dubey[3], Dr. Gaurav Jadhav[4]

Abstract

'Smart Agriculture' is broadly termed as a structured transformation in global agricultural practices which includes architectures of Internet of things (IoT), wireless sensor networks and algorithms of machine learning. While this architype offers noticeable gains in efficiency of resources, optimization of yield and climate resilience, at the same time it generates vast quantity of agricultural data which is sensitive whose governance remains inadequately regulated and poorly understood. This paper examines the challenges in multidimensional aspects of data governance and cybersecurity within smart agricultural system which are IoT enabled, with a particular focus on Indian context and comparative reference to evolving regulatory frameworks globally. The paper also investigates a socio-technical analytical lens as to how diverse architectures of IoT, deployments of edge computing and data flow of multi stakeholder collectively produce vulnerabilities which are structural both technical and institutional in nature that the existing instruments of governance have been ill equipped to address. The paper critically assesses India's Digital Personal Data Protection Act, 2023 (DPDPA) and thereby understanding and identifying important gaps in their suitability and applicability to agricultural data ecosystems in comparison with the newly enacted General Data Protection Regulation (GDPR), AI Act, and Data Governance Act of the European Union. After a synthesis of various literature work and analysis of regulation, the paper proposes an original Socio-Technical Agricultural Data Governance Framework (STADGF) that integrates various aspects of legal, architecture and participation.

While balancing innovation with data sovereignty, cybersecurity resilience and farmer rights, the framework proposes policy recommendations for agricultural stakeholders, policy makers and platform developers.

[1*]Assistant Professor, Department of Law, Specialization in Criminal Law, Maharashtra National Law University Mumbai, India, Email Id: tanayakamlakar25@gmail.com , Orcid Id: 0009-0005-1770-3890

[2]PhD scholar, Department of Law, Specialization in Corporate law, Maharashtra National Law University, Mumbai India, Email Id: ruchira.halli@gmail.com, Orcid Id: 0009-0003-1226-3828

[3]Assistant Professor, Specialization in Constitutional Law & Administrative Law, School of Law, Forensic Justice and Policy Studies, National Forensic Sciences University, Gandhinagar, Gujrat, India, Email Id: deepa.dubey@nfsu.ac.in, Orcid Id: 0000-0003-1355-0143

[4]Assistant Professor, Specialization in Criminal Law, School of Law, Forensic Justice and Policy Studies, National Forensic Sciences University, Gandhinagar, Gujrat, India, Email Id: gaurav.jadhav@nfsu.ac.in, Orcid Id: 0009-0004-7686-0346

Corresponding Author*: Dr. Tanaya P Kamlakar, Assistant Professor, Department of Law, Specialization in Criminal Law, Maharashtra National Law University Mumbai, India, Email Id: tanayakamlakar25@gmail.com ,

Keywords: Smart Agriculture; Internet of Things; Machine Learning; Edge-computing; Data governance; Cybersecurity; Digital Personal Data Protection Act 2023; GDPR; Socio-technical systems; Precision agriculture

1. Introduction

Agriculture continues to underpin India's economic structure and rural sustenance, shaping livelihoods, markets, and food security, The global population is projected to exceed 9.7 billion by the year 2050. It is definitely challenging to feed the population under the conditions of shrinking arable land, accelerating variation in climate and scarcity in increasing freshwater has triggered attention of agricultural system in digitalization [1]. Among the most significant technology solutions to this challenge is implementation of Internet of Things (IoT) infrastructure in agricultural settings, a collection of interconnected linked sensors, actuators, drones, autonomous self-driving vehicles and cloud-based analytics systems that together can monitor, model and control and manage agricultural operations at a level of detail previously unimaginable [2]. These systems can do everything ranging from predictive yield modelling and automatic scheduling of irrigation to real-time soil moisture analysis and detection of diseases in crops, when coupled with artificial intelligence (AI) algorithms and machine learning (ML) [3]. The emerging paradigm, referred to as Agriculture 4.0, or designated as smart farming or precision agriculture, represents a fundamental reorganization of relations of power and knowledge structures that govern food production in addition to a technological advancement [4]. The academic literature as of recently disproportionately concentrated more on the technical performance aspects of smart agriculture systems, such as accuracy of sensor, efficiency of communication protocol, rates of model classification, so on and so forth, while largely neglecting the socio- institutional aspects that ultimately determine whether such systems are adopted equitably and function reliably, despite the obvious promise of transformation [5].

There are three interconnected issues, in particular, which have not received enough systematic attention. *First*, issue of data governance: agriculture which is IoT enabled generates high volume, high velocity, heterogenous data streams relating to microclimatic conditions, characteristics of soil, farm management techniques, pest dynamics, crop phenology and also about the behaviour of the farmer. Under many national regimes, these data streams' ownership, access rights, commercialization potential and retention obligations are still unclear and ambiguous [6]. *Second,* the issue of the cybersecurity: infrastructure of agriculture IoT, which is usually spread across geographically, environments which are resource constrained with a management capacity of limited network, offers a wide and undefended attack surface vulnerable to a variety of threats such as manipulation of firmware, denial of service attacks, unauthorized data interception and ransomware [7]. *Third*, the issue of regulatory imbalance: the transnational nature of agriculture data platforms, where the data generated by the smallholder farmers in rural of Gujarat, Tamil Nadu or Maharashtra may be processed on the servers run and operated by multinational agri-tech corporations in Europe or United States, which creates vital mismatch between the jurisdictions one in which data risk is experienced and the other jurisdictions in which regulatory authority is exercised [8]. These challenges and difficulties are not just hypothetical or theoretical. Several jurisdictions have taken legislative action in response to high profile instances of agricultural data misuse, such as exploitation of unauthorized commercial farm level operational data by manufacturers of equipment and seed [9]. In a nation where about 600 million people rely either directly or indirectly on agriculture, India's enactment of Digital Personal Data Protection Act, 2023 (DPDPA) represents a major milestone in this regulatory landscape by creating for the first time a domestic comprehensive framework for personal data protection [10]. Although, this paper contends that, when applied to the realities of data ecosystem of agricultural IoT, the DPDPA and its rules as formulated currently contain a significant lacuna that neither the text of the Act nor its related policy documents have sufficiently addressed.

Although, the issue in India is not exceptional, it is especially significant because of the size and structural fragility of agriculture in India. According to the most recent Agricultural Census, the average size of landholding in India was 1.08 hectares and more than 86 % of farmers are categorized as small or marginal [11]. These smallholders lack institutional backing, technical literacy and negotiating power to negotiate significant data agreements with major agri-tech platforms. The technological benefits of smart agriculture may disproportionately benefit platform operators and corporate value chains rather than the farmers whose land and labour produce the underlying data due to combination of lax regulatory coverage, asymmetric data power and substantial cybersecurity exposure. [12]. In order to address these convergent issues, this paper develops an integrated socio-technical analysis of cybersecurity and data governance in agriculture which are IoT enabled. Drawing the traditions of Hughes, Bijker and Pinch and further developed through Science and Technology Studies (STS) study, the socio- technical approach maintains that technological systems cannot be sufficiently understood or controlled by their technical specifications [13]. Instead, the social relationships, institutional structures, legal standards, and cultural assumptions that surround them are co-constituted technology. When this perspective is applied to smart agriculture, it becomes evidently clear that the failures of data governance are not only due to the inadequacy of the firewall configuration or insufficiency of the encryption protocols, but also due to the contractual arrangement that concentrate data rights in platform providers, institutional cultures that view farmer-generated data as a raw material rather than a type of intellectual or financial asset, and regulatory frameworks adjusted to various data environments [14].

Four main scholarly contributions are made in this paper. *First,* it provides a systematic mapping of architectures of IoT, flow of data, patterns of ML deployment in smart agriculture, in order to establish the technical substratum for governance analysis depends. *Second,* by differentiating between infrastructure layer, data layer and application layer vulnerabilities, it creates an organized taxonomy of cybersecurity and privacy issues unique to agricultural IoT environments. *Third,* it compares the EU AI Act, EU Data Governance, GDPR and DPDPA, 2023 as they relate to agricultural data contexts, highlighting areas of regulatory overlap, convergence, divergence. *Fourth,* it suggests a novel framework for governance called the Socio-Technical Agricultural Data Governance Framework (STADGF), that combines the technological, legal, and participatory principles upgraded as per the requirements of smallholder agriculture systems in developing economies. The paper is structured as follows. The architecture and deployment environment of IoT systems in smart agriculture are examined in Section 2. The agricultural data ecosystem and multiple data flow it produces are examined in Section 3. An organized evaluation of security and privacy threats is given in Section 4. With a focus on the Indian context, Section 5 discusses the regulatory and governance environment. Section 6 gives a comparative review of global regulatory frameworks. The proposed STADGF is presented in Section 7. Reflections on research limits and future directions round out Section 8.

2. IoT-Enabled Smart Agriculture Systems

A three-layer model that includes *the perception, network and application layers* is typically used to understand the architecture of IoT enabled smart agriculture systems [14]. Heterogeneous sensing devices, such as electrochemical soil sensors, leaf wetness sensors, chlorophyll meters, weather stations, autonomous vehicles with LiDAR autonomous vehicles, and unmanned aerial vehicles (UAVs) installed with multispectral imaging systems, continuously accumulate raw data from the agricultural environment, at the *perception-layer* [12]. These devices function in a wide range of physical conditions, from the open, sparsely connected landscapes of rain-fed smallholder farms to the controlled atmospheres of precision greenhouse.

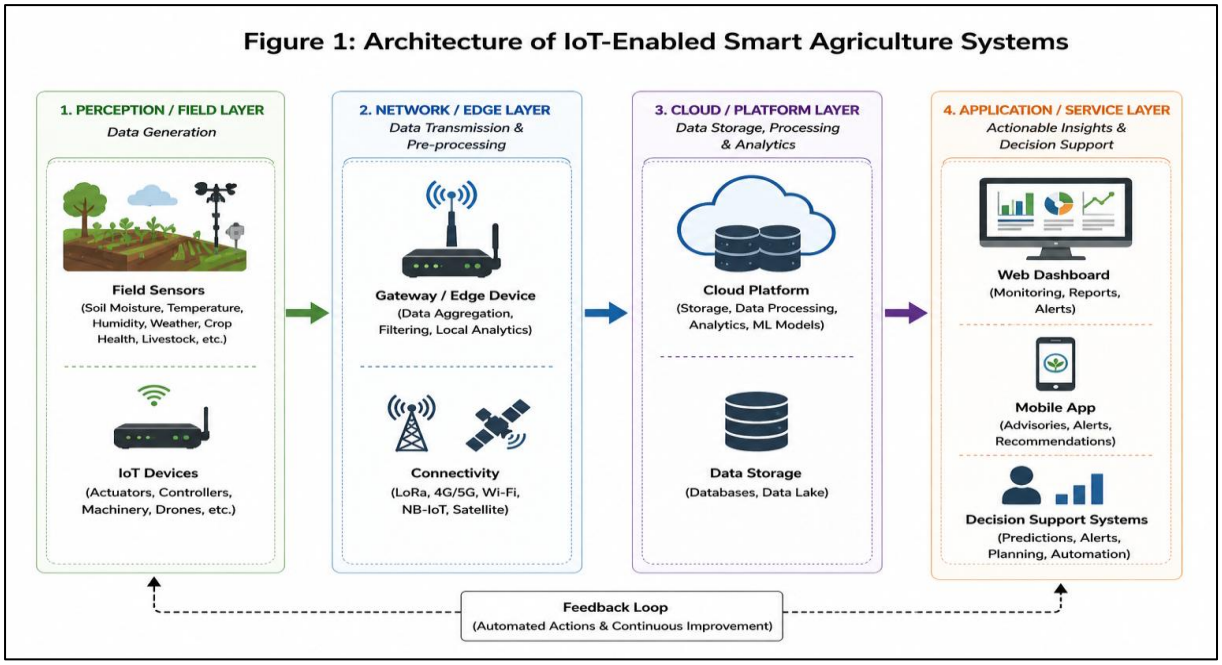

Figure 1. Architecture of IoT-Enabled Smart Agriculture Systems

The *network layer* refers to the communication infrastructure that transmits data collected at the perception layer to processing and storage systems. It is made in the combination of multiple wireless communication protocols: each with different strengths and limitations in terms of data rate, its energy consumption, communication range, and security [15]. In smart agriculture, Low-Power Wide-Area Network (LPWAN) technologies and Narrowband IoT (NB-IoT) are utilized considering less consumption of power. This makes them suitable for precision agriculture, where sensors are often spread across large and remote farming areas [2]. For denser deployments, such as greenhouse and polyhouse installations, where gateway infrastructure can be more widely dispersed, shorter-range protocols, such as Wi-Fi, Bluetooth Low Energy (BLE) and Zigbee, are still very useful. The advent of fifth generation (5G) mobile networks opens up new possibilities for applications which are ultra-low-latency, such as real-time autonomous machinery control but coverage limitations in rural areas continue to constrain near term applicability in many regions. India is at a nascent stage of deployments of 5G infrastructure in rural areas and is also geographically uneven, with coverage concentrated in peri-urban agricultural belts rather than the rain-fed interior [15].

The processing, analysis and decision support tasks that convert unprocessed sensor data into useful agricultural intelligence are included in *application-layer*. Cloud computing platforms, run by major providers like Google Cloud, Amazon Web Services and Microsoft Azure, and Google Cloud, as well as industry specific platforms from Ag Leader, Trimble and Indian companies like Agrostar, Fasal and CropIn, combine data from several farms and use ML models for classification of disease, scheduling irrigation, prediction of yield, and analytics of supply chain [3]. These platforms provide substantial analytical capability and scalability, but they also raise issues relating to commercial data exploitation, data sovereignty and the dangers of centralized data concentration. These issues are covered in more detail in later sections.

In the context of smart agriculture, edge computing refers to the deployment of processing power at or close to the data source, usually in the form of intermediary nodes placed at rural telecommunications towers or edge gateways erected at the farm level [16]. This architecture improves data privacy and lowers the volume transmitted raw data to centralized cloud infrastructure and also enables near-real-time decision-making in low-connectivity environments, by ensuring that the sensitive data of farm level does not leave the perimeter of farm for routine analytical tasks. Classification of crop disease, detection of anomaly in soil moisture and recognition of pest activity all can be done directly on sensor nodes with computational resources with are severely limited, due to the recent work that has shown the feasibility of implementing trained ML inference model on microcontroller-class hardware, therefore, this technique variously referred to as edge AI or Tiny ML [17]. This move towards distributed intelligence at the network edge has significant governance ramifications since it opens up technical possibilities for architectures that allow farmers to maintain local sovereignty over their raw data while only sharing processed analytical outputs with external or other platforms.

UAVs are a very important and consequential component of contemporary and modern smart agriculture infrastructure. Agricultural drones are fully equipped with thermal, multispectral or hyperspectral cameras that can survey the conditions of crop canopy over hundreds of hectares in one single flight, thereby generating images which are georeferenced and feeds into vegetation index calculations, primarily the Normalized Difference Vegetation Index (NDVI) and its derivatives which are used for yield estimation, mapping of fertilizer prescription and detection of early stress [12]. The amount of data generated or produced by UAV operations are enormous, a single hectare if surveyed at a spatial resolution of five centimeters can generate or produce several gigabytes of raw imagery data, placing heavy burden on processing infrastructure, storage capacity and transmission bandwidth in addition to the governance frameworks required to handle the sensitive operational and locational data and information contained within such datasets.

This three-layer architecture's integration of ML has transformed agricultural decision-making in a number of crucial and key domains. Under controlled conditions, convolutional neural networks (CNNs) are trained with image datasets on large-scale plant disease which have achieved classification accuracies surpassing 95 percent and deployment in resource-constrained environments is becoming more and more possible, all thanks to transfer learning approaches [6]. Architectures like Recurrent neural networks (RNNs) and Long Short-Term Memory (LSTM) have shown the efficacy in crop yield prediction from remote sensing data and multivariate time series of meteorological [3]. Because ensemble methods are resilient on tabular agricultural datasets, they are widely used in prediction of soil properties and optimization of irrigation, especially gradient boosting algorithms like LightGBM and XGBoost [18]. Data governance is at the core of the smart agricultural value offer since all these capabilities are essentially dependent on availability, quality and integrity of the underlying data.

3. Agricultural Data Ecosystem and Data Flows

The ecosystem of agricultural data produced by IoT-enabled smart farming systems is remarkably diverse in multiple dimensions. Data types include high-frequency, low-dimensional sensor streams (e.g. readings of soil moisture at 15-minute intervals) and episodic, high-dimensional imagery datasets produced by weekly drone surveys; structured tabular records (e.g. farm management logs and purchase history input); unstructured text (e.g. pest reports and messages of crop advisory); continuously generated real-time data; and historically accumulated archival records of agronomic practice [19]. Since different data types carry different risk profiles, different commercial values and different sensitivities with regard to competitive and privacy advantages of farmer, this heterogeneity poses a governance challenge in addition to a technical challenge for interoperability and operation of data.

Four main categories of agricultural data generated within ecosystems of smart farming are distinguished by a useful analytical taxonomy [20]. Environmental data captures ambient conditions that define the agroecological context of production such as soil's chemical and physical properties, parameters for quality of water and variables of microclimatic. Plating dates, irrigation events, chemical application records, logs of usage of machinery and other farming operations are all documented in operational data, which creates a comprehensive operational fingerprint of farm management practices that may have significant commercial importance. Direct observations of livestock and crop conditions such as disease incidence records, assessment of growth stage, yield measurements, are included in biological data, which is frequently augmented by remotely sensed spectral data. Lastly, the most direct implications for the security of personal data come from economic and identity data, such as location of farms, records of land ownership, purchase history input and records of sale transaction, which are directly connected to agricultural operations to particular legal persons.

This ecosystem's data flows involve multiple actors, many of which have materially different interests. Through their operations and sensors placed on their land, farmer generate primary data. Operators of agri-tech platform gather, aggregate, and process this data. They then use proprietary ML models to extract insights that are given back to farmers as recommendations for decision support. In order to calibrate product recommendations and target marketing interventions, input suppliers such as fertilizer manufacturers, seed companies and agrochemical corporations are increasingly requesting access to farm-level performance data Agricultural data is used by insurers and financial institutions and insurers to evaluate loan lending risk and underwrite crop insurance products. Government agencies many times reply on aggregated agricultural data for policy planning and food security monitoring. Research institutions use farm-level data for scientific study of agricultural system and emerging trends [19]. This raises multiple questions such as data control and data ownership which are not covered under the existing laws.

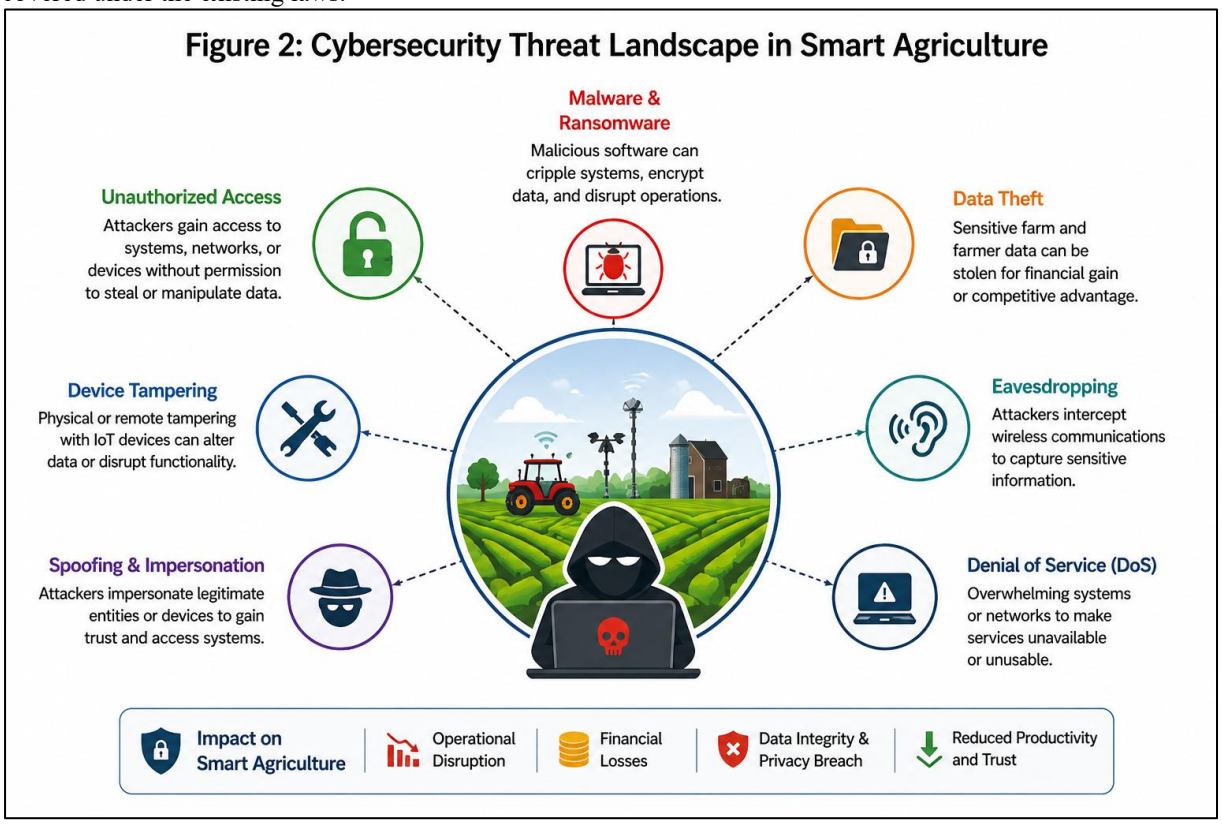

Figure 2. Agricultural Data Ecosystem and Multi-Stakeholder Data Flows

Agricultural data are not fit within the traditional intellectual property categories. At the same time much of this data relates to land, equipment, or crop conditions rather than identifiable individuals, meaning it may also fall outside conventional definition of personal data [8]. Farmers grant data access agri-tech platforms through contractual agreements that have been widely criticized for their structural asymmetry: platform terms of service are often drafted unilaterally, lengthy and technically complex by operators with superior legal resources, and offered without any real chance of discussion and with no meaningful opportunity for negotiation [21]. According to the research conducted by Jakku and his colleagues, even digitally literate Australian farming communities expressed serious and ongoing concerns about who owned their data and what happened to it after it left the farm. These concerns augmented by perceived power disparities between themselves and big corporate platform operators [22]. These power disparities are much noticeable in Indian smallholder environment, where institutional support for farmers in data negotiations is almost non-existent and digital literacy levels are often lower.

Additional levels of governance complexity are introduced by cross-border data flows. Data generated by framers in Uttar Pradesh or Andra Pradesh may be managed under the legal regimes far removed from regulatory authority of Indian law, or in some cases in jurisdictions with no meaningful data protection regime. This is because many agri-tech platforms that serve Indian farmers operate cloud infrastructure as seen and experimented in the Europe, US or Singapore. The modern agricultural data economy depicts a growing discord about where the data risks actually arise and where the legal power exists parallelly. This gap should be addressed while making the policies. Despite creating a government-notified list of acceptable recipient nations, the DPDPA's cross-border data transfer requirements currently lack the granularity and sectoral specificity necessary to effectively handle this issue in the agricultural context [10].

4. Security and Privacy Risks in Smart Agriculture

The cyber security risks associated with smart agricultural IoT systems are often overlooked by the users [7]. An attack surface that is both vast and structurally challenging to defend is created by unique features of agricultural IoT deployments including a lack of technical expertise on site, geographical dispersion across vast rural areas, device hardware with limited resources, a dependence on commercial off-the-shelf components, and extensive operational lifecycles measured in years or decades.

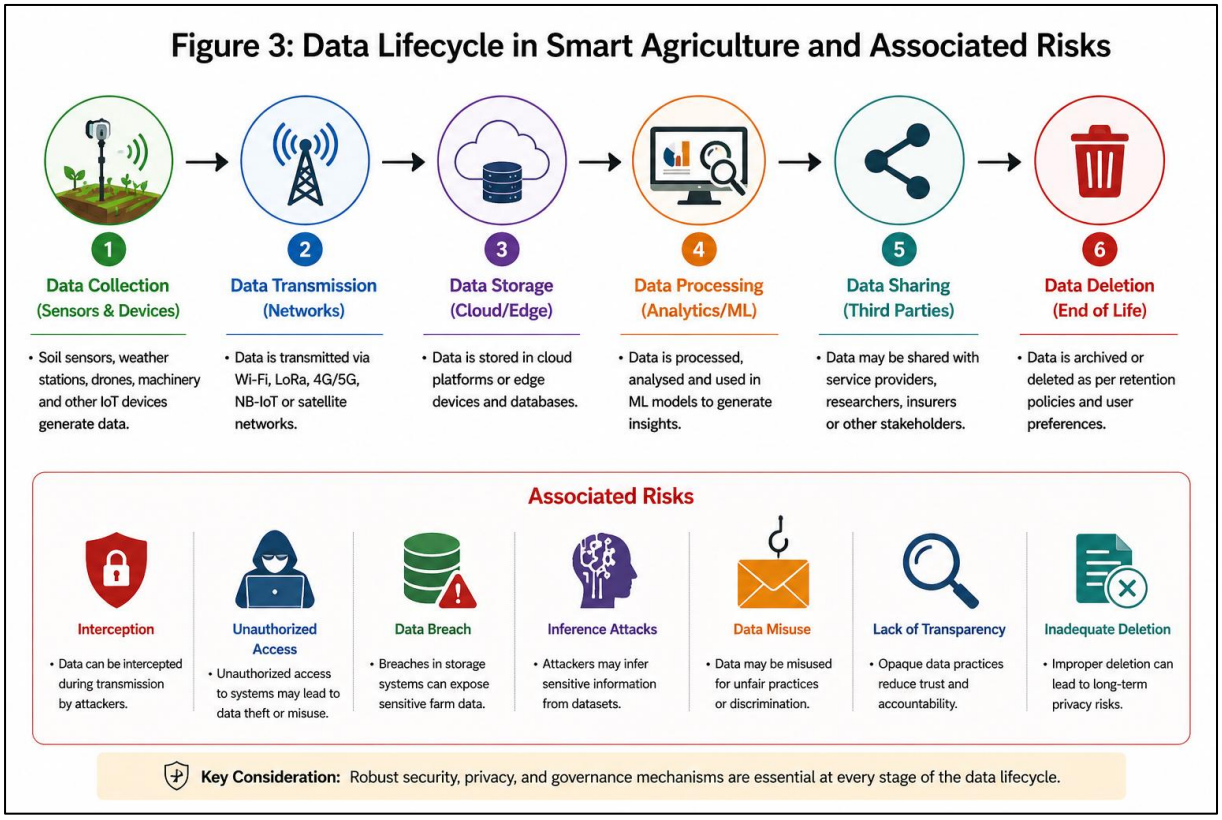

Figure 3. Cybersecurity Threat Landscape in Smart Agriculture

The most common vulnerabilities at the infrastructure layer are caused by the extensive use of IoT devices with unencrypted firmware, weak or default credentials and infrequent nonexistent or infrequent patch management procedures. Farmers and agronomists who lack expertise or knowledge in specialized cybersecurity often install agricultural sensor nodes and edge gateways; they might not be aware of the necessity to update device firmware or change default passwords provided by the manufacturer [23]. Once compromised, these devices can be used as entry points for lateral movement into infrastructure of larger farm network, employed into botnets for distributed denial-of-service (DDoS) attacks, or manipulated to generate fabricated and forged sensor readings that cause incorrect automated responses, such as an autonomous irrigation system that over-irrigate or under-irrigate a crop based on spoofed soil moisture data [9]. The latter type of attack, sometimes referred to as an actuator manipulation attack or false data injection attack, is especially sneaky because its effects can take weeks to materialize and show up as crop failure or yield loss that is hard to tell apart from typical agronomic mishandling as opposed to a planned security incident.

At the communication layer- the wireless transmission channels used in agriculture IoT systems are very prone to the interception, spoofing and replay attacks. The LoRaWAN's, specification includes message integrity checking and AES-128 encryption, but this also has implementation flaws like poor cryptographic nonce management that can result in key stream reuse [24]. UAV-based systems and precision guidance equipment are particularly vulnerable to GPS spoofing, which could lead to autonomous vehicles operating in inappropriate field locations, applying inputs to inappropriate areas, or colliding with physical structure. Adversaries may be able to intercept or secretly alter data streams without being discovered by either endpoint using man-in-middle attacks that target the communication link between edge gateways and cloud platforms.

The privacy risks and hazards related to smart agricultural systems are complex and interrelated at the data and application layer. When combined, the operational data produced by interconnected farming systems create a very sensitive economic profile of the farm business. Crop selection patterns, the use of chemical inputs, the amount of machinery used, and the timing of markets can all be used to identify competitive farm management tactics whose unauthorized disclosure could seriously jeopardize the financial interest of farmers. Aggregated agricultural data at larger scale may disclose national and regional production trends whose early or selective disclosure could be used by those with privileged data access to speculate commodities market [25]. When

financial and identity data are integrated into agri-tech platform ecosystem, there is a risk of data linkage, which is the merging of seemingly unrelated datasets that together allow for identification or inference.

Concerns over the use of farm level data for discriminatory profiling without farmer consent and knowledge are also becoming more widespread and well documented. Lenders may utilize remotely sensed land quality evaluation or farm performance scores derived from IoT data to calibrate lending access in ways that harm farmers in marginalized locations, or insurers may use them to refuse or price-discriminate in crop insurance provision [21]. Without regulatory transparency standards, it would be almost impossible to detect and contest such applications because they might be algorithmically mediated and opaque to the impacted farmers.

In the Indian agricultural IoT context, supply chain security is an additional and understudied risk dimension. Agricultural IoT systems' hardware and software are acquired from a globalized supply chain that may incorporate corrupted or counterfeit components. Additionally, software dependencies may have undisclosed vulnerabilities that remain unnoticed across product generations. In the context of ongoing geopolitical tensions, India's heavy reliance on imported IoT hardware, primarily from China, has been identified as a strategic vulnerability, raising documented concerns about the possibility of hardware-embedded backdoors in devices deployed across crucial rural and agricultural digital infrastructure [10].

These risks' magnitude is real. One of the biggest meat processing companies in the world, JBS Foods, was the target of a ransomware attack in 2021 that disrupted the production of beef and pork in the US, Canada, and Australia. This illustrates how cyberattacks have the potential to significantly disrupt supply chains in the food and agriculture sectors [26].

5. Governance and Regulatory Challenges

In August 2023, the Digital Personal Data Protection Act, 2023 (DPDPA) was given Presidential assent and thereby concluding India’s regulatory answer to the challenges of data governance of the digital economy [10]. DPDPA is not only significant achievement in terms of legislation but also first comprehensive piece of legislation in personal data protection. Purpose limitation, data minimization, storage limitation and enforceable rights of data principals to access, correct and erase their personal data, such various internationally recognized principles are incorporated in DPDPA. Data Protection Board of India is also established under DPDPA, which acts a primary enforcement authority. It also has received financial sanctions of up to INR 250 crore (approximately USD 30 million) for certain categories of breach and also has tiered penalty regime.

Although, when we examine the DPDPA in the context of data governance in smart agriculture, it discloses substantial restrictions. “Any data about an individual who is identifiable by or in relation to such data”, this definition of personal data in DPDPA largely bring into line with international standards. Thus, the application of this definition to agricultural data still remains uncertain, due to the fact that much of the farm level operational data does not directly count as an individual data in the most conservative form. Meanwhile such data might be linked to land records, connected to specific farm operator, or combined with other data sets to enable re-identification [10]. For example, there are few concepts which the Act fails to clarify whether they fall under the definitions of personal data or not, such as records of crop phenology, telemetry of soil sensory or co0ordinates of field boundary. The rights and obligations of the farmers and digital operators are not understood clearly due to legal uncertainty in the inclusiveness of agricultural concepts and also complete absence of any guidance in the Act’s text or policy document.

In an agricultural setting, the DPDPA’s consent framework creates further challenges. DPDPA defines consent as “free, specific, informed, unconditional and unambiguous” consent for processing personal data. In contrast, obtaining meaningful consent as required under DPDPA might be very difficult and far-fetched where awareness of data rights and digital literacy in farmers often remains limited in rural communities [28]. It raises a serious question, as to how a truly informed consent is obtained from a farmer on a smartphone that too in English or any other non-native language, which include lengthy and technically complex platform service terms. DPDPA’s provision of ‘deemed consent’, which permits and allows the platform operators to take advantage and infer consent from the context of legitimate activity to justify broad data processing without genuinely informed farmer agreement. In such circumstances, where digital literacy limitations make meaningful engagement with consent mechanism nearly impossible.

A question of agricultural data ownership, though distant, but largely related to personal data protection is not addressed in DPDPA. Whether or not farm-generated data qualifies as personal data under the law, ownership refers to who has the financial and legal rights to regulate its use and commercialization. Extensive rights of farmer generated operational data is often given to the platform terms of service, which may allow companies to use such data to train models of ML, to develop benchmarking tools or to share information with corporate partners [21]. In practice these provisions are often hidden in lengthy and technical agreements that most farmers are unlikely to read or fully understand.

Institutional fragmentation also causes further complication in framework of agricultural data governance in India. Ministry of Agriculture and Farmers Welfare, the Ministry of Electronics and Information Technology, the Ministry of Commerce and Industry and various state-level bodies are responsible for digital infrastructure, agricultural policy, consumer protection and data standards. Despite this, there is no single coordinating mechanism or dedicated governance framework for agricultural data. India’s AgriStack initiative highlights both the opportunities and risks of digital agricultural governance. It seeks to create an agricultural public infrastructure

which is digital through a national farmer's registry, records of land which are geo-referenced and interface which is unified farmer service. However, civil society organizations have raised concerns about weak mechanism of consent, surveillance over data and data concentration. These issues remain largely unaddressed by specific regulatory safeguards [29].

6. Comparative Global Frameworks

The European Union has developed an interlocking suite of legislative instruments that, when combined, offer a significantly more comprehensive governance architecture than any single national statute has yet achieved, in response to the governance challenges of agricultural IoT data. Comparative analysis of these tools reveals both their inherent limitations and relative strengths, and it provides valuable insights for the creation of regulations in India and around the world.

Strong safeguards for the processing of personal data are established by the EU's General Data Protection Regulation (GDPR), which has been in effect since May 2018 and applies to all economic sectors, including agriculture [30]. Compared to the DPDPA's existing formulation, the GDPR's broad definition of personal data, which includes any information pertaining to an identified or identifiable natural person, is more easily applicable to farm-level data connected to specific farmer identities. Data protection standards have been raised throughout the EU agricultural sector as a result of the GDPR's requirements for data protection by design and by default, its provisions for data portability, the creation of supervisory authorities with substantial enforcement powers, and its extra-territorial scope, which applies to processors handling the data of EU residents regardless of their own location. However, the GDPR was created as a horizontal tool and does not address sector-specific agricultural data governance issues that call for more focused regulatory action, such as ownership of non-personal operational data and the governance of developing agricultural data spaces.

A more focused and creative solution to these issues is the EU Data Governance Act (DGA), which went into effect in September 2023 [31]. A significant amount of agricultural IoT data, such as soil sensor telemetry, weather station records, and crop monitoring imagery, falls under the DGA's framework for the governance of non-personal data and mixed datasets with both personal and non-personal components. In agricultural context, the provisions relating to certified data intermediary services and recognized data altruism organizations are particularly relevant. They provide a potential legal framework for farmer-governed data cooperatives that are able to aggregate and handle farm level data on behalf of their members. Expanding on this concept, the Common European Data Spaces program of the European Commission has suggested a specific agricultural data space. The main objective is to establish a reliable cross-border infrastructure for data exchange in agricultural contexts, while protecting and preserving farmer data sovereignty and facilitating useful data sharing for policy creation, research and adaption of climate.

In June 2024, the European Union Artificial Intelligence Act (AI Act) was adopted officially, it introduces regulatory framework in smart agriculture which is risk based for AI systems with direct implications for ML applications [32]. There are four types of categories of AI systems such as unacceptable risk, high risk, limited risk and minimal risk in the AI Act. However, the classification of many agricultural AI applications remains uncertain. For illustration, categories of limited or minimal risk can or cannot include detection system for routine crop disease or precision irrigation tools. Whereas, high risk category can include farm level data used in credit scoring processes and any AI systems used in insurance underwriting or management of infrastructure. Such systems would be stricter obligations, including assessments of conformity, requirement of technical documentation, obligatory registration and monitoring post market. Startups and small agri-tech developers may be burdened by these compliance requirements, which could skew market structures in favor of major incumbents.

The US has for agricultural data governance mostly relied on voluntary self-regulation rather than enacting legally enforceable federal legislation. The most well-known industry-led governance initiative is the American Farm Bureau Federation's Privacy and Security Principles for Farm Data, which were published in 2014 and later revised. However, agricultural economists and legal scholars have criticized the document for lacking sufficient enforceability and for failing to establish the substantive data rights that statutory protection would grant farmers, particularly portability and non-discrimination rights [33]. There is now a sizable regulatory gap in the largest agricultural technology sector in the world since proposed federal legislation addressing agricultural data rights has not yet produced fully adopted statutes.

Four findings stand out as particularly clear when applying comparative lessons to the Indian context. *First,* to address the unique features of agricultural data ecosystems, sector-specific agricultural data governance tools are required, as opposed to solely depending on horizontal personal data protection laws. *Second,* a potentially revolutionary tool for rebalancing data power between individual smallholder farmers and giant agri-tech platforms is provided by data intermediary organizations, such as farmer data cooperatives, agricultural data trusts, or government-run fiduciary data platforms. *Third,* since rights that cannot be realistically exercised because of proprietary data formats are merely rights in name, technological standards for data portability and interoperability in agricultural IoT systems are a crucial prerequisite for genuine data rights. *Fourth*, in order to guarantee that governance tools reach and are comprehended by farming communities, data governance standards must be incorporated into agricultural policy frameworks rather than being treated as a unique digital economy regulatory issue.

7. Proposed Socio-Technical Agricultural Data Governance Framework (STADGF)

The preceding sections' analytical work shows that technical or legislative solutions alone will not be sufficient to overcome the cybersecurity and data governance issues of IoT-enabled smart agriculture. Legal frameworks without technical implementation mechanisms remain abstract and practically unenforceable; technical security measures without supportive legal frameworks deprive farmers of enforceable rights; and both legal and technical interventions will fail to accomplish their goals if they do not meaningfully engage with the institutional and social realities of agricultural communities. By utilizing the best features of comparable international frameworks, **the Socio-Technical Agricultural Data Governance Framework** (STADGF) put forth here aims to incorporate these dimensions into a logical, workable governance architecture tailored to the unique circumstances of smart agriculture in India.

The STADGF is built around four interconnected pillars as demonstrated below.

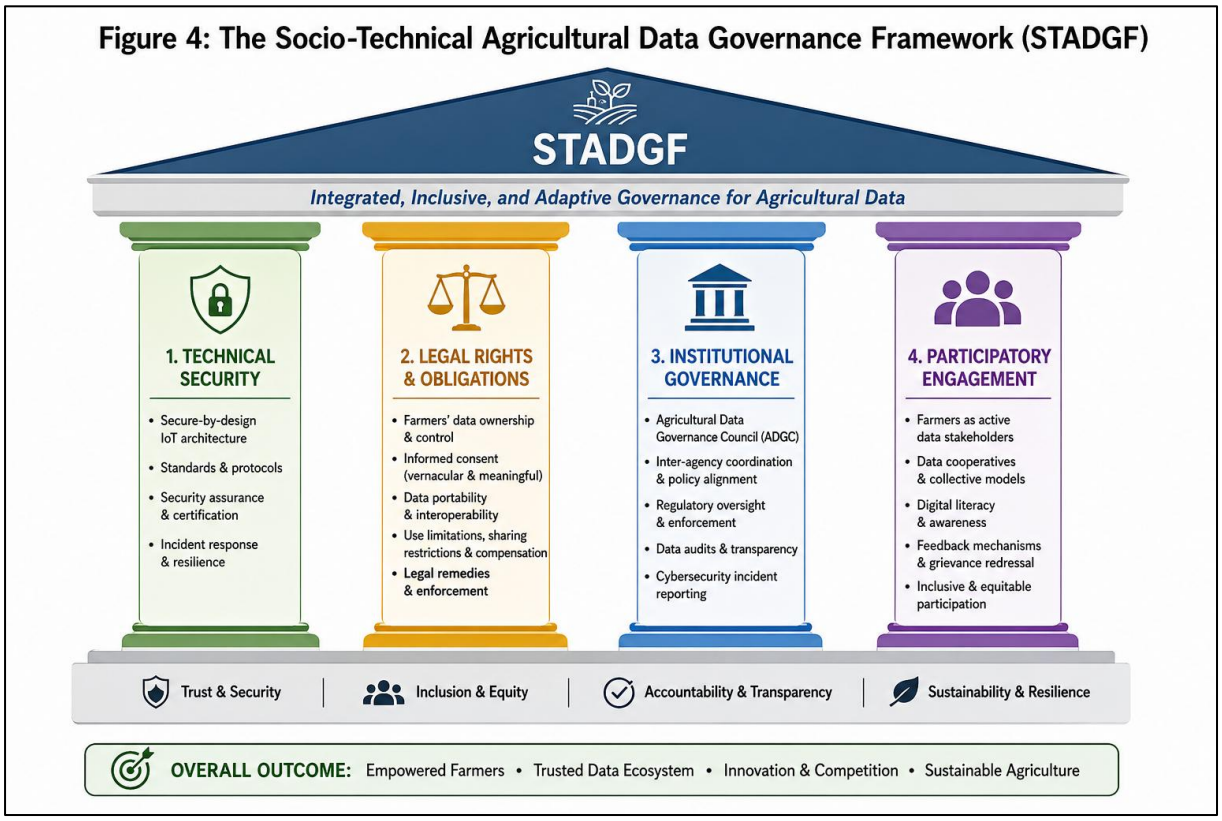

Figure 4. Socio-Technical Agricultural Data Governance Framework (STADGF)

7.1. Pillar of Technical Security

The Technical Security Pillar proposes a set of design requirements, technical standards and security assurance mechanism for agricultural IoT systems deployed in India. The adoption of a Security by Design principle, which is required by IoT product regulations applicable to devices marketed for agricultural use, is essential to this pillar. Before products can be sold or imported, manufacturers must implement unique per-device credentials at the point of manufacture, hardware-enforced secure boot processes, encrypted over-the-air firmware update mechanisms with cryptographic integrity verification, and documented vulnerability disclosure and patching policies [34]. This legislative strategy is similar to the EU Cyber Resilience Act already in effect and the UK's Product Security and Telecommunications Infrastructure Act 2022, both of which set mandatory baseline security measures for connected products as a condition of market access.

The STADGF advises the installation of agricultural edge data vaults at the edge computing layer. These are tamper-resistant hardware security modules that are shared at the village hub level or installed at the farm level. They carry out selective data processing, integrity verification, and local data encryption prior to any transmission to external cloud platforms. Unless individual farmers specifically grant extended data access permissions through an authenticated and revocable consent mechanism, these devices would implement hardware-enforced access controls that guarantee platform operators can only receive processed analytical outputs, such as irrigation volume recommendations or pest risk alerts, rather than raw sensor data streams. In addition to improving privacy and lowering the attack surface related to data transmission over public networks, this compute-at-source design facilitates the application of the DPDPA's data minimization principles [16].

In addition to standardized agricultural IoT data format specifications created through a participatory multi-stakeholder standards process involving farmer organizations, technology providers, government agencies, and research institutions, the STADGF recommends mandatory adoption of LoRaWAN 1.1 or equivalent protocol

versions that incorporate appropriate cryptographic key management and replay attack protection for communication security. Interoperability standards are crucial for making data portability effective. It enables and allows the farmers to move their data from one platform provider to other. It is therefore a valuable tool for promoting competition thereby preventing exploitative practices.

7.2. Pillar of Legal Right and Obligations

The Legal Rights and Obligations Pillar recommend legislative and regulatory reforms to proposes sector-specific legal protections for agricultural data thereby strengthening the data governance framework under the (DPDPA). An important recommendation is the enactment of Agricultural Data Rights Act or a dedicated chapter within a future Digital India Act. Such legislation could establish several important rights and obligations. *First,* farmers should have recognized ownership rights over operational and environmental data generated on their land, regardless of whether such data qualifies as personal data under the DPDPA. *Second,* agri-tech platforms operating above a defined scale threshold should be required to provide data sharing agreements in plain language and in relevant Indian vernacular language. *Third,* farmers should have a right to data portability, allowing them to export their data in standardized, machine- readable formats within a specified timeframe.

The law should also prohibit the use of farm operational data for purposes unrelated to agricultural service delivery unless separate and specific consent is obtained. In addition, the sale, licensing or secondary disclosure of farm level data to commodity traders, insurers or credit providers should require explicit informed consent and fair compensation mechanism linked to the commercial value of the data.

The DPDPA's consent framework should also be adapted for agricultural contexts. One possible approach is a system of vernacular consent certification. Under such a system, a public or independent third-party body would assess whether platform consent mechanism are understandable, specific and accessible in relevant Indian languages and for users with varying literacy levels.

Platforms that obtain certification could benefit from a regulatory safe harbour for certain categories of data processing. Platforms that fail to obtain certification may face enhanced audit obligations, increased liability, or restrictions on market access in regions with large smallholder farming populations.

7.3. Pillar of Institutional Governance

The Institutional Governance Pillar addresses the institutional fragmentation discussed in Section 5. It proposes structural reforms to improve inter-agency coordination and strengthen enforcement capacity. The STADGF recommends the creation of an Agricultural Data Governance Council (ADGC) as an inter-ministerial coordination body. It should include representatives from the Ministry of Agriculture and Farmers Welfare, Ministry of Electronics and Information Technology, the Data Protection Board of India, the Competition Commission of India and farmer representative organizations such as national agricultural cooperatives federations. The ADGC could be tasked with developing sector- specific data governance guidelines, overseeing the design and implementation oof AgriStack, coordinating cybersecurity incident response in the agricultural sector and advising o international data governance negotiations affecting Indian agriculture.

The framework also recommends the appointment of an Agricultural Data Auditor, either within or formally linked to the ADGC. This body would have statutory authority to conduct periodic and ad hoc audits of agri-tech platforms. These audits should examine data processing practices and cybersecurity readiness. Similar to the requirements under the EU's Network and Information Security Directive (NIS2), reporting obligations should be triggered when incidents are likely to affect the data or operational continuity of significant number of farmers.

7.4. Pillar of Participatory Engagement

The social and community aspects of agricultural data governance that are beyond the scope of simply technological or legal solutions are addressed by the Participatory Engagement Pillar. In order to collect farm-level data on behalf of their members, negotiate data agreements with agri-tech platforms collectively rather than individually, monitor compliance with agreed data use terms, and distribute the financial benefits of data commercialization to participating farmers through transparent and auditable dividend mechanisms, the STADGF recommends significant investment in Farmer Data Cooperatives, which are farmer-owned and farmer-governed collective institutions supported by appropriate legal structures similar to India's established agricultural credit cooperative framework [35]. This type of data cooperative has already been tested in several European countries, including the Netherlands and Ireland. These initiatives demonstrate a new model for enabling the agricultural data economy to flourish and generate the rightful data in favour of farmers.

With the global reliance on technology, digital literacy is vital for participatory data governance. It is essential to understand the value and risks associated with digital deals in order to give free consent, and to bargain fairly with the platform operators, Farmers who do not understand the value, risks and use of their data cannot give meaningful consent, negotiate fairly with platform operators. It is equally important for making the informed choices between digital services.

The STADGF therefore recommends integrating agricultural data literacy modules into existing agricultural extension systems. These modules could cover topics such as data rights under DPDPA, interpretation of platform agreements, practical consent processes and basic cybersecurity practices. Existing outreach networks such as

Krishi Vigyan Kendras (KVKs) and State Agricultural Universities can play an important role because of their rural presences and the trust they already enjoy among farming communities.

The STADGF is intended to be a dynamic governance framework rather than a fixed blueprint. It should be reviews and refined periodically as technology, cybersecurity threats and regulatory standards evolve. The framework is also meant to complement India' Digital Public Infrastructure (DPI) initiatives such as India Stack and proposed AgriStack.

Further, the framework could guide India's participation in global data governance debates, such as G20 Digital Economy Working Group and its bilateral digital cooperation agreements with the European Union, ASEAN Countries and other key partners in agricultural technology. It has four pillars mutually reinforcing, each is dependent on the others.

8. Conclusion

This paper has explored the data governance and cybersecurity issues in IoT driven smart agriculture, with a focus on India and global regulatory developments. Our analysis reveals that theses challenges are interdisciplinary. They cannot be solved by technical security measures alone, legal frameworks alone or institutional constraints alone. The challenges need to be tackled through a combination of approaches. These include the technical security of agricultural IoT systems, legal ambiguity over ownership and consent, institutional accountability and social factors that impact the ability of farmers to exercise effective control over the data produced from their land and labor.

The Socio-Technical Agricultural Data Governance Framework (STADGF) is one possible solution. It draws on the IoT security, comparative regulatory and socio-technical studies pf digital agriculture and farmer-platform relations. The four elements of its governanace framework: technical security, legal rights and obligations, institutional governance and participatory engagement, capture the complexity of the challenge. A one-dimensional approach is unlikely to lead to effective and sustainable governance.

This study has certain limitations. It draws on published academic literature, regulatory and policy documents. It is not based on primary empirical data from farmers or agri-tech providers. As a result, the lived experiences of these stakeholders are not fully reflected in the analysis. In order to ground-truth the STADGF's recommendations against the practical realities of smart agricultural deployment in various regional and socioeconomic contexts across India, future research should use qualitative and participatory methods, such as farmer interviews, platform operator surveys, regulatory impact assessments, and participatory action research with farmer cooperative initiatives.

Future research on smart agriculture data governance will face some important issues that this article has not been able to fully address. As a possible addition to the governance architecture suggested here, the use of distributed ledger technologies and blockchain-based smart contracts for automated data licensing and agricultural data provenance monitoring merits thorough technical and legal assessment. Large foundation models and generative AI systems have implications for agricultural decision-making that are not yet sufficiently covered by current frameworks. These implications include new methods of data extraction, model opacity, and potential bias amplification that affects already marginalized farming communities. Furthermore, new concerns regarding data integrity, third-party verification, and the allocation of benefits between data-generating farmers and carbon market intermediaries are raised by the intersection of agricultural data governance with voluntary and compliance-driven carbon credit markets, which heavily rely on verifiable farm-level emissions and sequestration data. The IoT engineering, data science, legal, and agricultural policy research communities must give these issues consistent multidisciplinary attention since they will significantly shape the agricultural data governance research agenda for the upcoming ten years.

References

1. FAO, *The Future of Food and Agriculture: Alternative Pathways to 2050*, Food and Agriculture Organization of the United Nations, Rome, 2018.
2. O. Elijah, T.A. Rahman, I. Orikumhi, C.Y. Leow, M.N. Hindia, *An overview of Internet of Things (IoT) and data analytics in agriculture: benefits and challenges,* IEEE Internet of Things Journal 5(5) (2018) 3758 3773. https://doi.org/10.1109/JIOT.2018.2844296
3. K.G. Liakos, P. Busato, D. Moshou, S. Pearson, D. Bochtis, *Machine learning in agriculture: a review*, Sensors 18(8) (2018) 2674. https://doi.org/10.3390/s18082674
4. S. Wolfert, L. Ge, C. Verdouw, M.J. Bogaardt, *Big data in smart farming — a review*, Agricultural Systems 153 (2017) 69–80. https://doi.org/10.1016/j.agsy.2017.01.023
5. D. Pivoto, P.D. de Waquil, E. Talamini, C.P. Spanhol Finocchio, V.F. Dalla Corte, G. Martinelli, *Scientific development of smart farming technologies and their application in Brazil*, Information Processing in Agriculture 5(4) (2018) 519–531. https://doi.org/10.1016/j.inpa.2018.08.002

A. Kamilaris, F.X. Prenafeta-Boldú, *Deep learning in agriculture: a survey*, Computers and Electronics in Agriculture 147 (2018) 70–90. https://doi.org/10.1016/j.compag.2018.02.016

6. M.A. Ferrag, L. Shu, X. Yang, A. Derhab, L. Maglaras, *Security and privacy for green IoT-based agriculture: review, blockchain solutions, and challenges,* IEEE Access 8 (2020) 32031–32053.

https://doi.org/10.1109/ACCESS.2020.2973178
7. N. Kshetri, *The economics of Internet of Things (IoT) data governance,* IT Professional 21(3) (2019) 32–38. https://doi.org/10.1109/MITP.2019.2896944
8. S. Sontowski, M. Gupta, S.S.L. Chukkapalli, M. Abdelsalam, S. Mittal, A. Joshi, R. Sandhu, *Cyber attacks on smart farming infrastructure*, in: Proceedings of the 2020 IEEE 6th International Conference on Collaboration and Internet Computing (CIC), Atlanta, GA, 2020, pp. 135–143. https://doi.org/10.1109/CIC50333.2020.00025
9. Ministry of Electronics and Information Technology, *The Digital Personal Data Protection Act, 2023,* No. 22 of 2023, Government of India, New Delhi, 2023. Available at: https://www.meity.gov.in
10. Ministry of Agriculture and Farmers Welfare, *Agriculture Census 2015–16: All India Report on Number and Area of Operational Holdings,* Government of India, New Delhi, 2019.
11. A.D. Boursianis, M.S. Papadopoulou, P. Diamantoulakis, A. Liopa-Tsakalidi, P. Barouchas, G. Salahas, G. Karagiannidis, S. Wan, S.K. Goudos, *Internet of Things (IoT) and agricultural unmanned aerial vehicles (UAVs) in smart farming: a comprehensive review,* Internet of Things 18 (2022) 100187. https://doi.org/10.1016/j.iot.2021.100187
12. W.E. Bijker, T.P. Hughes, T. Pinch (Eds.), *The Social Construction of Technological Systems: New Directions* in the Sociology and History of Technology, MIT Press, Cambridge, MA, 1987.
13. [14] M.S. Farooq, S. Riaz, A. Abid, K. Abid, M.A. Naeem, *A survey on the role of IoT in agriculture for the implementation of smart farming*, IEEE Access 7 (2019) 156237–156271. https://doi.org/10.1109/ACCESS.2019.2949703
14. H.M. Jawad, R. Nordin, S.K. Gharghan, A.M. Jawad, M. Ismail, *Energy-efficient wireless sensor networks for precision agriculture: a review*, Sensors 17(8) (2017) 1781. https://doi.org/10.3390/s17081781
15. W. Shi, J. Cao, Q. Zhang, Y. Li, L. Xu, *Edge computing: vision and challenges,* IEEE Internet of Things Journal 3(5) (2016) 637–646. https://doi.org/10.1109/JIOT.2016.2579198
16. P.P. Ray, *Internet of Things for smart agriculture: technologies, practices and future direction*, Journal of Ambient Intelligence and Smart Environments 9(4) (2017) 395–420. https://doi.org/10.3233/AIS-170440

A. Tzounis, N. Katsoulas, T. Bartzanas, C. Kittas, *Internet of Things in agriculture, recent advances and future challenges*, Biosystems Engineering 164 (2017) 31–48. https://doi.org/10.1016/j.biosystemseng.2017.09.007

17. L. Klerkx, E. Jakku, P. Labarthe, *A review of social science on digital agriculture, smart farming and agriculture 4.0: new contributions and a future research agenda*, NJAS – Wageningen Journal of Life Sciences 90–91 (2019) 100315. https://doi.org/10.1016/j.njas.2019.100315
18. S. van der Burg, M. Bogaardt, S. Wolfert*, Ethics of smart farming: current questions and directions for responsible innovation towards the future,* NJAS – Wageningen Journal of Life Sciences 90–91 (2019) 100289. https://doi.org/10.1016/j.njas.2018.12.001

A. Lajoie-O'Malley, K. Bronson, S. van der Burg, L. Klerkx, *The future(s) of digital agriculture and what they mean for agricultural sustainability: an analysis of high-level policy documents*, Futures 115 (2020) 102490. https://doi.org/10.1016/j.futures.2019.102490

19. E. Jakku, B. Taylor, A. Fleming, C. Mason, S. Fielke, C. Sounness, P. Thorburn, *If they don't tell us what they do with it, why would we trust them? Trust, transparency and benefit-sharing in smart farming*, NJAS – Wageningen Journal of Life Sciences 90–91 (2019) 100285. https://doi.org/10.1016/j.njas.2018.11.002
20. [23] S.S.L. Chukkapalli, M.B. Seal, M. Gupta, N. Piplai, S. Mittal, A. Joshi, *Ontologies and artificial intelligence systems for the cooperative smart farming ecosystem*, IEEE Access 8 (2020) 112243–112258. https://doi.org/10.1109/ACCESS.2020.3002162
21. F. Adelantado, X. Vilajosana, P. Tuset-Peiro, B. Martinez, J. Melia-Segui, T. Watteyne, *Understanding the limits of LoRaWAN*, IEEE Communications Magazine 55(9) (2017) 34–40. https://doi.org/10.1109/MCOM.2017.1600613
22. K. Bronson, I. Knezevic, *Big data in food and agriculture*, Geoforum 96 (2019) 1–3. https://doi.org/10.1016/j.geoforum.2018.07.024
23. Federal Bureau of Investigation, *Cyber Division, Ransomware Actors Use Significant Financial Events and Stock Valuation to Facilitate Targeting and Extortion Against Victims*, Private Industry Notification PIN 20210901-001, FBI, Washington D.C., 2021.
24. Federal Bureau of Investigation, *Food and Agriculture Sector Increasingly Targeted by Ransomware Actors*, TLP:WHITE, PIN 20220420-001, FBI, Washington D.C., 2022.
25. V. Saiz-Rubio, F. Rovira-Más, *From smart farming towards agriculture 5.0: a review* on crop data management, Agronomy 10(2) (2020) 207. https://doi.org/10.3390/agronomy10020207
26. Ministry of Agriculture and Farmers Welfare, *India Digital Ecosystem of Agriculture (IDEA): Conceptual Framework*, Government of India, New Delhi, 2021.
27. European Parliament and Council, *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation),* Official Journal of the European Union L 119 (2016) 1–88.
28. European Parliament and Council, *Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance (Data Governance Act),* Official Journal of the European Union L 152 (2022) 1–44.

29. European Parliament and Council, *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act),* Official Journal of the European Union L (2024).
30. American Farm Bureau Federation, *Privacy and Security Principles for Farm Data*, American Farm Bureau Federation, Washington D.C., 2014 (revised 2016).
31. M. Frustaci, P. Pace, G. Aloi, G. Fortino, *Evaluating critical security issues of the IoT world: present and future challenges,* IEEE Internet of Things Journal 5(4) (2018) 2483–2495. https://doi.org/10.1109/JIOT.2017.2767291
32. Verdouw, H. Sundmaeker, B. Meyer-Aurich, J. Wolfert, E. Verhoosel, *Digital twins in smart farming,* Agricultural Systems 189 (2021) 103046. https://doi.org/10.1016/j.agsy.2020.103046
33. N.N. Misra, Y. Dixit, A. Al-Mallahi, M.S. Bhullar, R. Upadhyay, A. Martynenko, *IoT, big data, and artificial intelligence in agriculture and food industry*, IEEE Internet of Things Journal 9(9) (2022) 6305–6324. https://doi.org/10.1109/JIOT.2020.2998584
34. M.A. Ferrag, L. Maglaras, S. Moschoyiannis, H. Janicke, *Deep learning for cyber security intrusion detection: approaches, datasets, and comparative study*, Journal of Information Security and Applications 50 (2020) 102419. https://doi.org/10.1016/j.jisa.2019.102419