

CHAPTER 2 - LITERATURE REVIEW

This chapter explores the connection between blockchain technology and privacy concerns, specifically in the agricultural sector data with PDP law. The objective of our research is to analyze and extract valuable information from concepts of theory and previous research.

2.1 Theoretical Perspectives

2.1.1 HARA

HARA is an Indonesian technology company that collects and analyses data specifically for the agricultural sector. (HARA, 2024) utilizes blockchain technology to offer transparent and dependable data exchange solutions. HARA's objective is to gather data from farmers to enhance their access to third-party services like financial services and insurance.

HARA has expanded access to financial services by partnering with multiple financial service providers through an innovative loan origination program. This initiative operates across 98 villages, 29 sub-districts, and four regencies. It covers Banyumas, Bojonegoro, Klaten, and Tuban within Central Java and East Java provinces. HARA's financial services initiatives are focused on paddy and corn commodities. This program has recorded approximately 10,000 hectares of farmland, connecting local farmers to access financial products supporting their planting needs.

Another program currently in the pilot phase involves a partnership with an insurance company focused on the tobacco commodity in one of Java's provinces. This program provides farmers with insurance coverage for crop failure, helping to mitigate financial risks associated with agricultural production.

Another ongoing initiative focuses on completing farmer information to meet EU due diligence requirements for palm oil traceability. This effort supports compliance with the EU Deforestation Regulation (EUDR), which requires that all palm oil products imported into the EU are traceable, deforestation-free, and free from forest degradation (THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION, 2023).

This initiative facilitates data sharing between HARA, its partners, and government bodies, ensuring farmers' needs are met more effectively through streamlined access to financial and insurance services and completing compliance.

The HARA ecosystem includes two technology tools to assist every stakeholder within the system: the HARA acquisition apps and the internal monitoring dashboards. Field agents use the acquisition apps that focus on this thesis to collect general farmer data and additional data specific to the programs in which farmers participate. Meanwhile, HARA utilizes the internal dashboard to monitor and validate the data collected from the field, ensuring its accuracy and reliability. Together, these tools facilitate efficient data management and enhance the overall effectiveness of HARA's operations.

2.1.2 Agriculture Data and HARA's data

Based on (Telangana, 2023), Agricultural data refers to information that is generated, gathered, processed, provided, or utilized in relation to agricultural business, agricultural production, and general agricultural services. The list of each types can be seen in Table 2.1. Agriculture data is crucial for optimizing agricultural production, improving business performance, and enabling effective policy-making (Ngo et al., 2019).

Table 2.1. Types of Agriculture Data.

Agriculture business data	Agriculture production data	General agricultural data
Agricultural credit Insurance Logistics Storage-related services Market-related services Payment services	Farm advisories pest and disease management Water management Quality assurance Soil health	Benefits Land records Research Innovation Education Data principal use.

Agricultural data is being generated and collected at an unprecedented rate from various sources such as IoT sensors, satellites, weather stations, farm machines, and farm management information systems (Lyotos et al., 2020). In fact, the agriculture sector is expected to be one of the largest generators of big data in the coming years (Ariza Ariza et al., 2019). The variety of agriculture data is also increasing with structured and

unstructured data, images, videos, sensor readings, and text. The velocity of agriculture data is also increasing with real-time data streaming from IoT devices and other sources (Ngo et al., 2019).

HARA collects data using input apps. This data collection includes agricultural data and personal data of farmers to verify the authenticity of these individuals. The agricultural data is categorized into several groups: agricultural business data for loan origination, crop monitoring data to track farming activities, and general agricultural data that includes land records to verify the existence and boundaries of the farmland. Table 2.2 shows the list of variables that HARA acquired from the input tools. This data collection ensures that HARA's programs are based on accurate and reliable information, supporting effective service delivery to farmers.

Table 2.2. HARA's Data Acquire by Data Input tools.

Type of Data	Data	Commodities			
		Paddy	Corn	Tobacco	Palm Oil
Farmer's Data	Full Name	✓	✓	✓	✓
	Phone Number	✓	✓	✓	✓
	Citizen Identity Number	✓	✓	✓	✓
	Name of <i>Kelompok Tani</i>	✓	✓	✓	✓
	Regency/City	✓	✓	✓	✓
	District	✓	✓	✓	✓
	Subdistrict/Village	✓	✓	✓	✓
	Residence Address	✓	✓	✓	✓
	E-STDB Document				✓
	<i>Kartu Tani</i> Photo (<i>Kartu Tani ID</i> and full name)	✓	✓		
	Farmer Status	✓	✓	✓	✓

Table 2.3. HARA's Data Acquire by Data Input tools. (cont.)

Type of Data	Data	Commodities			
		Paddy	Corn	Tobacco	Palm Oil
Land Data	Land Name	✓	✓	✓	✓
	Land Area	✓	✓	✓	✓
	Land Coordinate	✓	✓	✓	✓
	Land Polygon	✓	✓	✓	✓
	Province	✓	✓	✓	✓
	Regency/City	✓	✓	✓	✓
	District	✓	✓	✓	✓
	Subdistrict/Village	✓	✓	✓	✓
	Plan Variety	✓	✓	✓	✓
Loan Data	Bank Name	✓	✓	®	
	Number of Saving	✓	✓		
	Estimated Planting Date	✓	✓		
	Photo of Agreement Document	✓	✓		
	Farmer's Selfie with Agreement Document	✓	✓		

2.1.3 Blockchain Technology

The origin of blockchain technology can be linked to the innovative study conducted by (Haber and Stornetta, 1991). They designed an approach for storing documents with timestamps in a chain of blocks. However, it was just after (Nakamoto, 2008) introduction of Bitcoin that blockchain achieved worldwide acknowledgment. Nakamoto merged a decentralized network with a Merkle tree for chronological verification. Although Bitcoin was revolutionary, its early use was limited to documenting cryptocurrency transactions.

The capabilities of blockchain received a significant transformation when (Buterin, 2014a) proposed the concept of smart contracts. This new concept enabled blockchain to expand its functionality beyond the base purpose of recording transactions. Smart contracts operate autonomously based on predetermined principles that are written in code. Buterin's innovation allows blockchain to go extends its initial use in finance and be applied in several industries.

Blockchain is data stored in a decentralized network. A “block” consists of transaction data that is stored in blockchain. “Chain” linked the connection between blocks. It connects the hash of current block with block $i-1$ as seen in Figure 2.2.

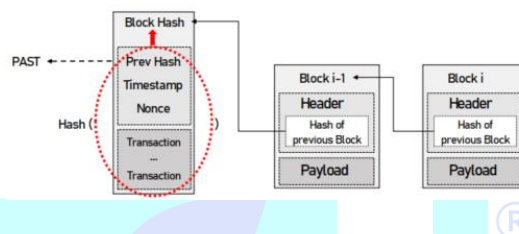


Figure 2.2. The Structure of Blockchain (Ghiro et al., 2021)

Blockchain starts with a genesis block, also known as the initial block, which is the parent of all following blocks. The genesis block is hardcoded into the blockchain protocol, forming the chain's foundation. Each block in the blockchain contains several key components, transaction data, and a block header. Transaction data records all transactions within the block, and a block header summarizes the block's content. The block header consists of the protocol version, a timestamp of the block's creation time, a nonce, a Merkle root representing all transactions in the block, and the previous block's hash, linking it to its previous block. Each block also has its unique cryptographic hash as its fingerprint. The link between blocks is created using the previous block's hash stored in the current block's header. This linkage ensures stability by ordering blocks, immutability by making tampering nearly impossible without affecting the entire chain, and security by maintaining the integrity of the blockchain.

2.1.3.1. Blockchain Network Types

Blockchain networks are categorized according to permission structures, which are permissionless and permissioned models (Yaga et al., 2019). Permissionless, or public,

blockchain networks (Falazi et al., 2019) adopt an open-access principle. It enables anyone to participate by deploying an instance of the corresponding network. Pioneered by Bitcoin (Nakamoto, 2008) this approach is widely adopted by major networks such as Ethereum (Buterin, 2014a), Polygon, and Tezos. Permissioned or private blockchain networks have restrictions on participation. Enterprises usually use it in collaborative processes without relying on third parties (Falazi et al., 2019). Notable examples of permissioned networks include Hyperledger Fabric (Androulaki et al., 2018) and XRP Ledger by Ripple (Chase and MacBrough, 2018). The classification of blockchain networks into two categories highlights their adaptability, allowing for a wide range of applications and user preferences in both public and private domains.

New blocks are created by validator node based on the consensus algorithm implement. **Error! Reference source not found.** shows the difference between three common consensus: Proof-of-Work, Proof-of-Stake, and Proof-of-Authority. Proof-of-Work (PoW) is the most use consensus in public blockchain network. Validator in PoW, known as miner, solved complex mathematical puzzles to validate new blocks. Miner who solved the puzzles is given a reward (Patil and Mitragotri, 2023; Zarrin et al., 2021). Unlike PoW, Proof-of-Stake (PoS) selects validators based on the number of tokens they hold and are willing to stake as collateral. The more tokens a participant holds, the higher their chances of being chosen to validate transactions. Validators are incentivized to act honestly, as they risk losing their staked tokens if they attempt to validate fraudulent transactions (Patil and Mitragotri, 2023). Another consensus Proof-of-Authority mostly use in permissioned blockchain. PoA design the block creation using Byzantine approach (Zarrin et al., 2021). All validators in PoA will have the turn to create new blocks without solving mathematical puzzles.

Table 2.4. Differences between PoW, PoS, and PoA.

Mechanism	Proof-of-Works	Proof-of-Stakes	Proof-of-Authority
Block Creation	The first one who solve complex mathematical puzzles gets to create a block and receives a reward.	The person who has the biggest and longer time of stake has the highest chance to produce a block and receives a reward.	A certain number of actors are pre-determined to be able to produce a block.
Validator Node	All user can become miner	Users stake their tokens to become a validator	Trusted user become sealer
Network Type	Public/Private Network	Public/Private Network	Private Network

When the transaction is submitted, the transaction will be broadcast on the network. A validation run based on the specific consensus. If valid, the transaction will be grouped with other valid transactions into one new block. The new block links to the previous block. Transaction is complete (Ghiro et al., 2021).

2.1.3.2. Blockchain Trilemma Issues

The chosen consensus will lead to the blockchain trilemma (Ghiro et al., 2021). A trade-off between the three properties in Figure 2.3 must be chosen. The blockchain trilemma refers to the challenge of balancing three critical nonfunctional aspects of blockchain technology: scalability, decentralization, and security. It is difficult to achieve optimal performance in all three areas simultaneously, and it often compromises the others (Fu et al., 2024; Quattrocchi et al., 2024). Scalability refers to the ability of a blockchain to process more transactions than a single regular node to ensure that the system can handle increased usage without significant performance degradation. Decentralization highlights that the blockchain operates without relying on a centralized entity. It is generally understood that participation should not depend on trust. Security ensures the blockchain can handle attacks from a significant percentage of participating nodes, ideally up to 50%, maintaining the integrity and reliability of the network even under malicious conditions.(Quattrocchi et al., 2024)

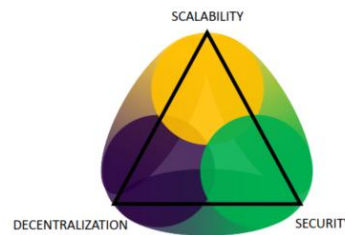


Figure 2.3. Blockchain Trilemma

(Kim et al., 2018) examined many solutions to tackle the critical issue of one of the trilemmas in scalability. These strategies may be categorized into five unique approaches: On-chain, Off-chain, Sidechain, Child-chain, and Inter-Chain. On-chain solutions refer to direct alterations made to the primary blockchain protocol with the goal of enhancing both its capacity and throughput. Off-chain solutions, on the other hand, reduce network congestion by moving certain transactions away from the main blockchain. Side-chain solutions facilitate the creation of separate blockchains that are connected to the main chain. Child-chain solutions adhere to a comparable premise but function with their own distinct set of regulations and protocols, guaranteeing a certain level of independence.

(Gangwal et al., 2023) further elaborates on scalability solutions by categorizing them into Layer-1 (L1) and Layer-2 (L2) solutions. Layer-1 solutions involve core changes to the blockchain design itself, such as block data modification, the introduction of new consensus mechanisms, network sharding, or the adoption of Directed Acyclic Graph (DAG) technologies. These changes directly alter the fundamental structure of the blockchain to improve scalability.

The Layer-2 solutions do not modify the core blockchain but instead operate on top of it, providing scalability enhancements without altering the existing blockchain network. Layer-2 solutions include channels (such as state and payment channels), side/child chains, cross-chains, and hybrid solutions. These approaches facilitate faster transactions, greater flexibility, and enhanced scalability by offloading some of the transaction processing to external systems or layers, reducing the strain on the primary blockchain.

2.1.3.3. Ethereum

Ethereum is a decentralized, open-source blockchain platform that enables developers to build and deploy smart contracts and decentralized applications (dApps). It was proposed in late 2013 by Vitalik Buterin and development began in early 2014. (Buterin, 2014) introduce four key points of Ethereum from the whitepaper consisting smart contracts, decentralized applications (dApps), Ether (ETH), and Ethereum Virtual Machine (EVM).

- Smart Contracts: These are self-executing contracts with the terms of the agreement directly written into code. They automatically enforce and execute the terms when conditions are met.
- Decentralized Applications (dApps): These are applications that run on the Ethereum blockchain, leveraging its decentralized nature to avoid downtime, fraud, and interference from third parties.
- Ether (ETH): This is the native cryptocurrency of the Ethereum platform. It is used to pay for transaction fees and computational services on the Ethereum network.
- Ethereum Virtual Machine (EVM): This is the runtime environment for smart contracts in Ethereum. It allows anyone to run any program, regardless of the programming language, given enough time and memory.

In 2022, Ethereum transitioned its consensus mechanism from Proof-of-Work (PoW) to Proof-of-Stake (PoS) (Smith, 2024). PoS is significantly more energy-efficient, eliminating the need for intensive computations. It also lowers entry barriers by reducing hardware requirements, making participation more accessible. Additionally, PoS enhances decentralization by encouraging more nodes to secure the network. With lower energy consumption, less ETH issuance is required to incentivize participation. Economic penalties for misbehavior make 51% attacks more costly, and in extreme cases, the community can socially recover an honest chain if such an attack occurs. After implementing PoS, the block creation time was reduced to twelve seconds, as shown in Figure 2.4.

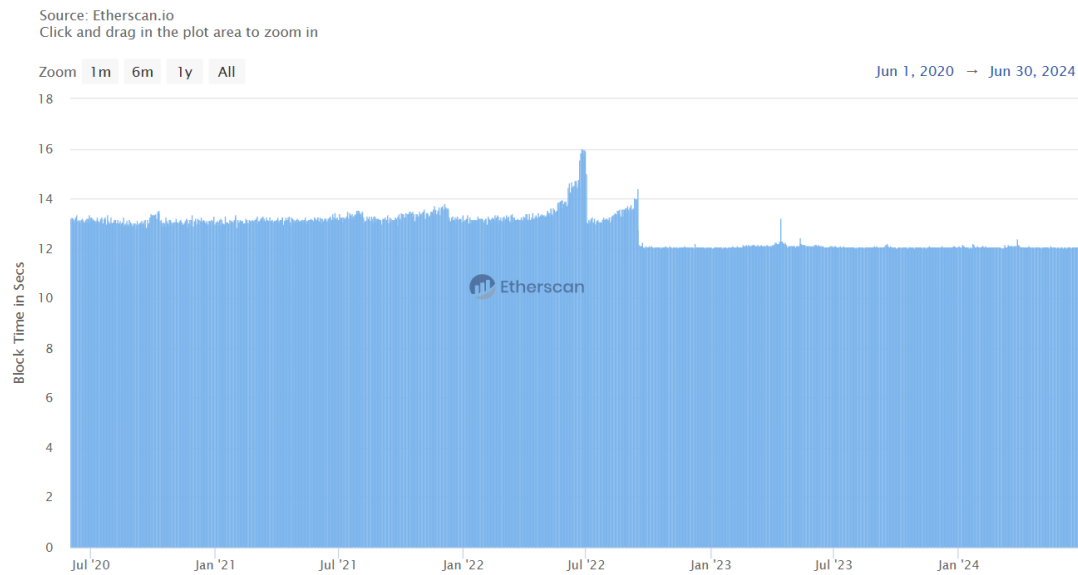


Figure 2.4. Ethereum Block Time Creation from January 1st, 2020, to June 30th 2024. (Etherscan, 2025)

Based on (Popper and Gibson, 2021) in stack overflow survey, Ethereum is the most widely used blockchain technology among respondents working with blockchain as seen in Figure 2.5 .This has resulted in a large and active community surrounding Ethereum-based technologies. As blockchain is still a relatively new technology, the growth of this community is highly beneficial for addressing any challenges that arise.

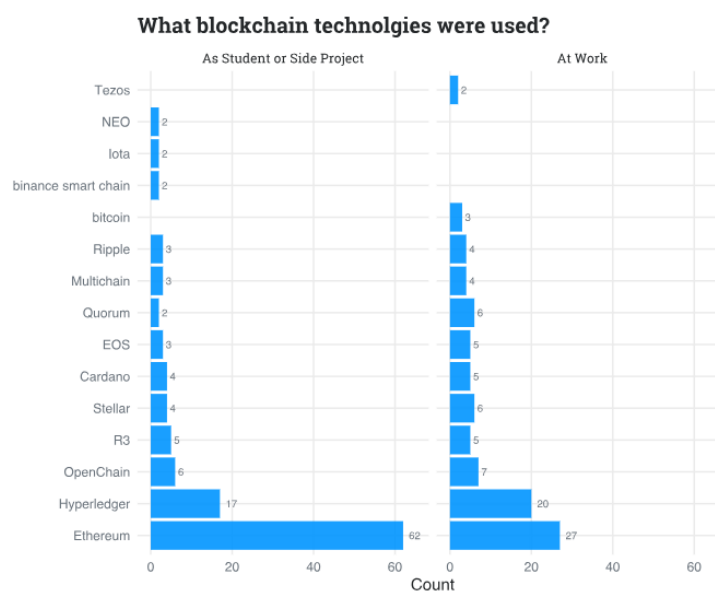


Figure 2.5. What blockchain technology were used? Survey of Stack Overflow. (Popper and Gibson, 2021)

As Ethereum was the first to implement the EVM-based blockchain, many popular public networks have adopted EVM compatibility. Binance Smart Chain, Polygon, Avalanche, and Ripple either support or use sidechains compatible with the EVM. This enables developers to create multi-supported platforms and fosters interoperability between multiple networks.

With the rapid growth of the Ethereum community and its widespread compatibility with many popular blockchain networks, HARA has chosen to build its platform on top of the Ethereum Virtual Machine (EVM). By leveraging the EVM's robust and proven infrastructure, HARA ensures it remains interoperable with other EVM-compatible networks such as Binance Smart Chain, Polygon, and Avalanche. To meet the unique requirements of its business model, HARA has implemented custom configurations and optimizations on the EVM, allowing it to tailor its blockchain solutions more effectively. This approach not only enhances scalability and performance but also positions HARA to seamlessly integrate into the broader blockchain ecosystem, offering flexibility and innovation to its users.

2.1.3.4. Blockchain Data Management

Effective data management is essential for ensuring the security, integrity, and efficiency of blockchain systems. It plays a crucial role in leveraging the full potential of blockchain technology. (Wilson et al., 2024) introduces a detailed data management taxonomy specifically designed for blockchain technology. This taxonomy, in Figure 2.6, provides a structured framework to categorize and address various aspects of data management within blockchain networks. This includes:

- **Data Flow:** Understanding how data moves within the blockchain network is crucial for maintaining system efficiency and security.
- **Data Storage:** Efficient storage solutions are necessary to handle the large volumes of data generated by blockchain transactions.
- **Access Control:** Implementing robust access control mechanisms is vital to protect sensitive data and ensure that only authorized users can access it.
- **Querying:** Effective querying methods are needed to retrieve data efficiently from the blockchain

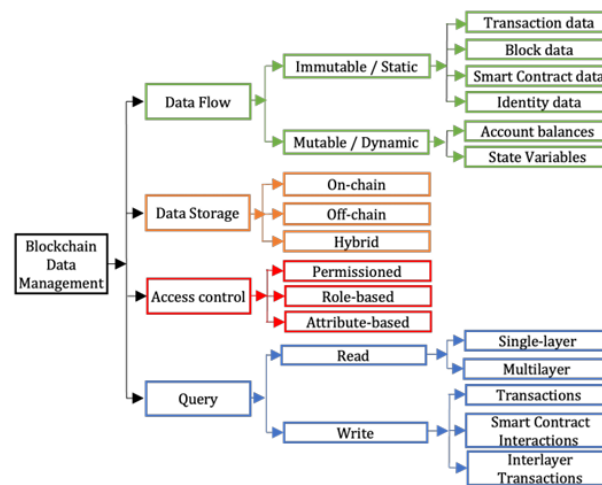


Figure 2.6. Blockchain Data Taxonomy by (Wilson et al., 2024)

2.1.3.5. HARA Blockchain Implementation

HARA utilizes blockchain technology to provide data integrity, transparency, and security. Blockchain guarantees that every data is unalterable and traceable. The data's signature is documented to preserve hash integrity, ensuring that whenever the data is accessed, it is verified against the recorded signature to confirm its integrity. Blockchain offers a transparent ledger that allows all participants to view transactions and data exchanges, ensuring that any stakeholder can access the history and provenance of the data, so rendering the information reliable and verifiable. The decentralized architecture of blockchain strengthens security by dispersing data over numerous nodes, rendering it exceedingly challenging for any singular actor to manipulate the data.

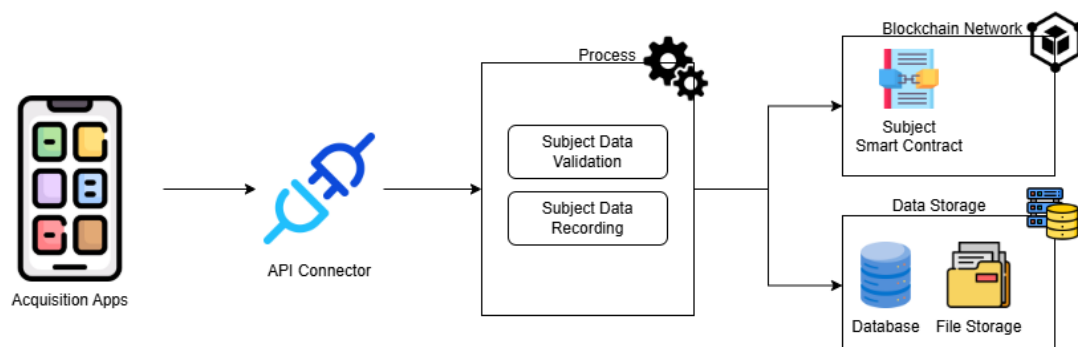


Figure 2.7. Flow of data from acquisition recorded to blockchain and data storage.

The acquisition tool captured the farmer and land data as the source of the commodity. It will process and record both blockchain and data storage. On the on-chain side, a

“subject” smart contract, shown in Figure 2.7 on blockchain network section, records the entities data including farmer and land. Meanwhile, off-chain data storage utilizes a managed database and cloud file storage to store off-chain data.

The entire HARA ecosystem which integrates blockchain technology, is illustrated in Figure 2.8. The blockchain component, referred to as the on-chain layer, operates on a permissioned network called the HARA Ledger, utilizing a Proof-of-Authority (PoA) consensus mechanism. The sealers responsible for securing the HARA Ledger include both HARA and its partners, ensuring a balanced level of decentralization.

The current HARA Ledger addresses two aspects of the blockchain trilemma: scalability and decentralization. Proof-of-Authority (PoA) enhances scalability by setting a block creation time of five seconds in the HARA Ledger, along with a higher gas limit compared to public networks. For context, Ethereum’s average block creation time is twelve seconds. For comparison, Ethereum's average block creation time is twelve seconds that is mentioned in section 2.1.3.3. However, the trade-off in security is not yet fully addressed due to the limited number of external sealers from HARA’s partners. This thesis does not further explore the security trade-offs of the HARA Ledger, but this topic can be examined in future research.

The blockchain component, referred to as the on-chain layer, operates on a permissioned network called the HARA Ledger, utilizing a Proof-of-Authority (PoA) consensus mechanism. The sealers responsible for securing the HARA Ledger include both HARA and its partners, ensuring a balanced level of decentralization. Within current HARA ledger, it covers the scalability and decentralization of blockchain trilemma. PoA by defining new blocks every 5 second and more gas limit than the public network make more scalable than public network (i.e. Ethereum average block creation is 12 seconds). The trade off in security is not cover yet because the minimum number of sealer from external HARA partner that join. This topic is not dicuses further in this thesis and can be further research.

On the other hand, the off-chain layer interacts with the on-chain layer via the HARA Connection Node (connode) service and the Blockchain Core Service. Data collected from acquisition applications is stored in off-chain storage (e.g., a database) and then recorded on a smart contract through the Blockchain Core Service. All transactions recorded on the blockchain can be monitored using a block explorer.

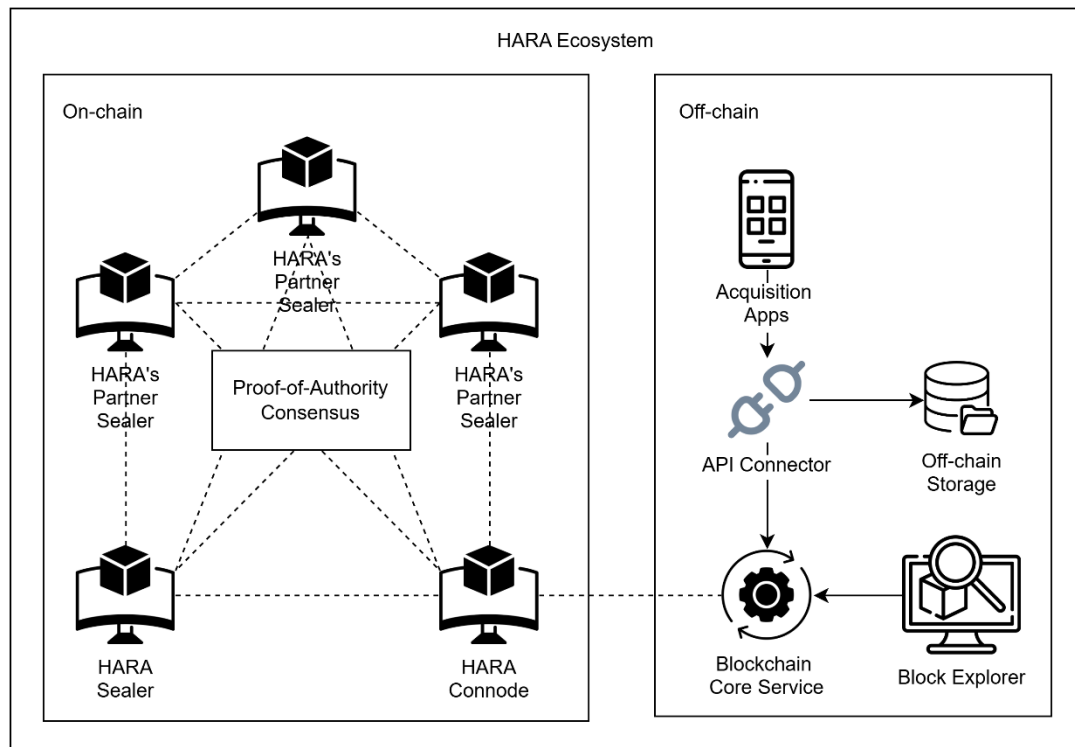


Figure 2.8. HARA Blockchain Ecosystem.

2.1.4 Data Privacy Regulation and PDP LAW

The concept of privacy has had a profound evolution, shifting from the straightforwardness of rural communities to the complexities of contemporary cities, where issues about safeguarding data and legal entitlements to privacy have become of utmost importance in the era of digitalization. During the Elizabeth era, there was a significant prevalence of espionage and intrigue, whereas the Victorians introduced identifying techniques such as fingerprints, pictures, and the creation of criminal records (Marshall and Thomas, 2017). The Middle Ages, a period largely neglected in the study of privacy history, were characterized by small rural communities where direct interpersonal communication was prevalent, fostering a sense of familiarity among individuals and constraining the notion of personal boundaries.

Many nations understand the necessity to regulate privacy to protect their citizens' rights and well-being. The rise of digital platforms and technical advances have driven governments worldwide to create comprehensive privacy frameworks and laws. The

regulations tried to balance technical progress and personal data protection. Governments are progressively adjusting legislative frameworks to the dynamic challenges of the digital era to safeguard privacy and data protection in a changing global landscape.

1. European Union (EU): The EU has implemented the General Data Protection Regulation (GDPR), a stringent framework that guarantees thorough protection of personal data. The GDPR places a strong emphasis on individual rights, granting individuals the power to access, correct, and delete their personal data. Organizations that deal with personal data have certain responsibilities, such as getting consent, establishing security measures, and quickly informing individuals of any data breaches (Dauden-Esmel et al., 2021; Martins Gonçalves et al., 2023; Ramadhan and Wijaya, 2022).
2. Singapore: Singapore has been following the Personal Data Protection Act (PDPA) since 2012. This act sets out the rules and regulations for how businesses can acquire, use, and share personal data. The PDPA establishes regulations for the acquisition of consent, ensuring the accuracy of data, and providing individuals with the right to access and rectify their personal information. Furthermore, the PDPA requires the designation of data protection officers and sets down procedures for reporting data breaches (Ramadhan and Wijaya, 2022).
3. Malaysia: The Personal Data Protection Act 2010 is the governing law that regulates the handling of personal data by business enterprises in Malaysia. This Act requires enterprises to obtain consent, offer clear notification of data gathering, and implement strong security measures for personal data. The Act also appoints a commissioner to supervise adherence and enforce the provisions of the law (Ramadhan and Wijaya, 2022).
4. Japan: The Act on the Protection of Personal Information (APPI), Act No. 57 of 2003, is Japan's primary privacy law for corporations and governments. Entities must seek consent and express explicit data collecting

intentions under fairness, legality, and purpose limitation laws. To prevent unauthorized access, sensitive information requires special security. The APPI controls cross-border data transfers and allows persons to request data disclosure, correction, and deletion. The Personal Information Protection Commission enforces laws strengthened by amendments, and non-compliance results in sanctions(The Government of Japan, 2003).

In 2022, Indonesia implemented the UU *Perlindungan Data Pribadi* under UU No. 27, 2022. These moves posed challenges for industries and organizations. Based on interviews with various companies, it became evident that many of them lacked documentation for their data, including both the data they possessed and its flow. Closing this gap to ensure strict compliance with the constitution is expected to be a time-consuming process. Another challenge is meeting the requirement for cross-border data transfers. This requires getting data subjects' approval, confirming that the recipient nation maintains a sufficient level of data protection, and informing the PDP Authority about the transfer. Establishing appropriate policies and procedures for cross-border data transfers is crucial to solve this, as is making sure that third-party service providers follow these rules as well (HukumOnline, 2022).

2.1.4.1. Current State of HARA's Data and Compliance in PDP Law

Based on the data captured through HARA's data acquisition process, each variable has been evaluated in alignment with the provisions outlined in Indonesia's current Personal Data Protection (PDP) Law. This assessment ensures that the categorization adheres to legal standards for identifying and handling Personally Identifiable Information (PII). For a clearer understanding, the classification of these variables, highlighting which data points fall under the scope of PII, is detailed comprehensively in Table 2.5.

Table 2.5. Classification of Personal Data (PII) Under Indonesia's PDP Law.

Type of Data	Data	PII under PDP Law
Farmer's Data	Full Name	Yes
	Phone Number	Yes
	Identity Number	Yes
	Name of <i>Kelompok Tani</i>	No (unless it identifies individuals)
	Regency/City	No (unless combined with PII)
	District	No (unless combined with PII)
	Subdistrict/Village	No (unless combined with PII)
	Residence Address	Yes
	E-STDB Document	Yes
	<i>Kartu Tani</i> Photo (<i>Kartu Tani ID</i> and full name)	Yes
	Farmer Status	No
Land Data	Land Name	No
	Land Area	No
	Land Coordinate	Yes (if tied to personal ownership)
	Land Polygon	Yes (if tied to personal ownership)
	Province	No
	Regency/City	No
	District	No
	Subdistrict/Village	No
	Plan Variety	No
Loan Data	Bank Name	No
	Number of Saving	Yes
	Estimated Planting Date	No
	Photo of Agreement Document	Yes (if it includes personal identifiers)
	Farmer's Selfie with Agreement Document	Yes

2.1.4.2. Transparency, Deletion, and Extermination in the PDP Law

Article 27	Article 43	Article 44
<i>Pengendali Data Pribadi wajib melakukan pemrosesan Data Pribadi secara terbatas dan spesifik, sah secara hukum, dan transparan.</i>	(1) <i>Pengendali Data Pribadi wajib menghapus Data Pribadi dalam hal:</i> <i>a. Data Pribadi tidak lagi diperlukan untuk pencapaian tujuan pemrosesan Data Pribadi;</i> <i>b. Subjek Data Pribadi telah melakukan penarikan kembali persetujuan pemrosesan Data Pribadi;</i> <i>c. terdapat permintaan dari Subjek Data Pribadi; atau</i> <i>d. Data Pribadi diperoleh dan/ atau diproses dengan cara melawan hukum.</i> (2) <i>Data Pribadi sebagaimana dimaksud pada ayat (1) dilakukan sesuai dengan ketentuan peraturan perundang-undangan.</i>	(1) <i>Pengendali Data Pribadi wajib memusnahkan Data Pribadi dalam hal:</i> <i>a. telah habis masa retensinya dan berketerangan dimusnahkan berdasarkan jadwal retensi arsip;</i> <i>b. terdapat permintaan dari Subjek Data Pribadi;</i> <i>c. tidak berkaitan dengan penyelesaian proses hukum suatu perkara; dan/ atau</i> <i>d. Data Pribadi diperoleh dan/ atau diproses dengan cara melawan hukum.</i> (2) <i>Pemusnahan Data Pribadi sebagaimana dimaksud pada ayat (1) dilakukan sesuai dengan ketentuan peraturan perundang-undangan.</i>

Figure 2.9. PDP Law Article 27, article 43, and article 44.

Out of the 76 articles in 16 chapters of the PDP Law, this thesis focuses on three key articles that shown in Figure 2.9: articles 27, 43, and 44. Article 27 highlights the fundamental principles for personal data controllers when processing personal data must emphasizes limited and specific, legally valid, and transparent. Transparent means ensuring that the data subject is aware of the personal data being processed and how it is processed. Additionally, all related information and communication must be easily accessible, understandable, and conveyed in clear language. Article 43 addresses data deletion, which must be carried out when personal data is no longer needed for processing purposes, subject to consent withdrawal, requested for deletion by the Data Subject, or obtained or processed unlawfully. Article 44 discusses data extermination, which refers to the act of removing, eliminating, or permanently destroying personal data so that it can no longer be used to identify the Data Subject. This means that all links between the data and the Data Subject must be completely erased.

2.1.5 Other Law Related

2.1.5.1. POJK Number 14/POJK.05/2015 and POJK Number 39 /POJK.05/2020

The Otoritas Jasa Keuangan (OJK) Regulation No. 14/POJK.05/2015 outlines the requirements for self-retention and domestic reinsurance support in the Indonesian insurance industry (Otoritas Jasa Keuangan, 2015). The regulation mandates that insurance and reinsurance companies must optimize the use of domestic reinsurance capacity before seeking overseas reinsurance. The regulation includes guidelines for determining self-retention limits based on risk exposure, financial capacity, and reinsurance strategies.

Regarding the data retention period under this regulation, Article 17, Paragraph (1) states that exceptions to automatic reinsurance support may apply if a company does not obtain or does not require such support. Article 16 mentions a retention period related to reinsurance rejection, specifying that if domestic reinsurers refuse coverage, the company may seek overseas support within a maximum of one year from the rejection date. However, for broader compliance with Personal Data Protection (PDP) Law, the retention period should align with data processing and storage regulations applicable to insurance and financial services.

To enhance the regulatory framework established by POJK No. 14/POJK.05/2015, a new regulation, POJK No. 39/POJK.05/2020, was introduced in 2020 (Otoritas Jasa Keuangan, 2020). This regulation provides greater flexibility for insurance industry players and allows for more effective risk distribution through reinsurance support programs. The objective is to enhance economic efficiency while maintaining best practices in risk management. Key updates in this regulation include:

1. Gradual removal of mandatory domestic reinsurance support for simple-risk insurance.
2. Adjustments to minimum placement requirements in domestic reinsurance for non-simple risk coverage.
3. New eligibility criteria for obtaining reinsurance support from foreign reinsurers.
4. Strengthened enforcement through revised sanction provisions for non-compliance.

In relation to the Personal Data Protection (PDP) Law, these regulatory changes reinforce the importance of secure and compliant data management in the insurance sector. As insurance companies increasingly rely on reinsurance mechanisms, ensuring that personal data shared across entities aligns with PDP Law principles

2.1.5.2. Undang-Undang No. 8 Tahun 1997

Undang-Undang Nomor 8 Tahun 1997 tentang Dokumen Perusahaan is an Indonesian law that regulates the management, storage, and disposal of corporate documents (Presiden Republik Indonesia, 1997). It aims to ensure that businesses maintain proper records for accountability, legal compliance, and financial transparency. The law defines various types of corporate documents, outlines retention periods, and provides guidelines for document archiving and destruction.

The retention framework outlined in Article 11 mandates that primary business documents, including financial and accounting records, must be preserved for at least 10 years from the end of the relevant fiscal year. Supporting financial administration documents may be retained based on operational needs, while other corporate records should be maintained according to their functional value. Article 19 permits the destruction of records following the expiration of their retention period, a practice that is critical in mitigating risks related to data breaches, unauthorized access, and prolonged data processing beyond its intended purpose. By harmonizing corporate document retention with personal data protection regulations, businesses can enhance compliance, uphold individuals' rights, and foster a secure data management environment.

2.1.6 Qualitative Research

Qualitative research is an exploratory approach used to understand complex social, technological, and regulatory phenomena (Patton, 2015). Unlike quantitative research, which focuses on numerical measurement, qualitative research emphasizes context, meaning, and process. In the context of blockchain and PDP Law in HARA data acquisition product, qualitative research provides the necessary flexibility to assess how regulatory frameworks interact with real-world data management practices.

Key characteristics of qualitative research that align with this thesis include:

- Contextual understanding: Captures how HARA's blockchain-based data acquisition product function under Indonesia's PDP Law.
- Exploratory analysis: Identifies compliance gaps and proposes solutions.

A case study approach is used to analyse HARA's data acquisition product, following (Yin, 2008) case study research methodology. A case study is appropriate when:

- The research investigates a contemporary issue (blockchain data governance under evolving PDP Law).
- The research requires in-depth analysis of a single entity (HARA) to identify challenges and best practices.
- The research environment cannot be controlled, as it is an ongoing, real-world system.

Types of Case Studies and Justification (Yin, 2008) identifies different case study designs:

- Exploratory case study: Used to investigate new or emerging phenomena.
- Explanatory case study: Used to understand causal relationships.
- Descriptive case study: Used to provide a detailed account of an event/system.

Since this study seeks to analyse HARA's blockchain-based data practices and propose a compliance-aligned framework, it follows an exploratory case study approach. This approach enables the identification of compliance gaps and the development of regulatory-aligned recommendations.

2.2 Previous Studies

2.2.1. Challenges and Solutions in Aligning Blockchain with Privacy Regulations

Numerous scholarly studies have examined the convergence of blockchain technology and privacy regulations, focusing on countries including PDP law in Malaysia and GDPR in the European Union (Baskaran et al., 2020). The immutability inherent in blockchain represents a critical feature that challenges data privacy. It ensures that users cannot easily remove any recorded information in the blockchain, presenting a conflict

with the 'Right to be Forgotten' principle in the General Data Protection Regulation (GDPR) highlighted by (Godyn et al., 2022). This raises concerns about how HARA's blockchain-based solution can comply with such requirement as the basis in RO1.

More in the summary, (Zanchi, 2022) has highlighted blockchain as a potential threat to user privacy. To address this conflict, (Freund et al., 2021) proposed various approaches to align with GDPR requirements, including the establishment of a central authority, the concept of shared responsibility, and the implementation of pseudonymization. (Stach et al., 2022a, 2022b) have delved deeper into this concern, offering more nuanced and technically oriented studies. These scholars concur on the challenge posed by blockchain's inherent characteristics, such as immutability and tamper-proofing, which can limit privacy. In response, (Baskaran et al., 2020) and (Liang et al., 2023) proposed a solution that involves incorporating permissioned blockchain, off-chain storage, and stealth addresses to mitigate privacy concerns associated with blockchain. (Bansod and Ragha, 2022) contributed to advancing our comprehension of privacy preservation mechanisms within the context of blockchain-based platforms. Their studies have provided valuable insights into the intricate interplay between blockchain technology and privacy mechanisms, thereby enhancing our understanding of both elements. By delving into the intricacies of how blockchain can be leveraged for privacy preservation, their work serves as a foundational resource for scholars and practitioners seeking a deeper understanding of the synergies between blockchain and existing data privacy frameworks. These solutions provide a foundation for this research's RO2, which aims to develop a compliance framework for HARA's data acquisition system, ensuring PDP law compliance while maintaining data integrity and operational efficiency.

Despite these proposed solutions, scholars generally agree that blockchain, in its current state, remains difficult to fully align with privacy-centric regulations. This ongoing challenge supports the H1 of this research, particularly in areas such as data collection, processing, retention, and third-party data management. Furthermore, the research conducted by (Baskaran et al., 2020) and (Liang et al., 2023) supports H2. By leveraging these prior studies, this research will provide a practical framework for integrating regulatory compliance with blockchain-based agricultural data systems,

contributing to both academic discussions and real-world applications in Indonesia's agricultural sector.

2.2.2. On-chain and Off-chain Approach in Blockchain Recording Data

Blockchain systems often combine on-chain and off-chain mechanisms to balance transparency, scalability, and privacy. On-chain data recording focuses on storing metadata, such as hashes of private data, consent, and transaction details, ensuring data integrity and traceability without exposing sensitive information (Bai et al., 2022; Martins Gonçalves et al., 2023). Off-chain storage addresses blockchain's limitations by securely managing sensitive data, such as Personally Identifiable Information (PII) and Personal Health Information (PHI), in systems like IPFS or AWS Cloud. Encryption techniques, such as Paillier Homomorphic Encryption, ensure confidentiality (Liang et al., 2023). This hybrid model enables efficient data management while meeting privacy regulations like GDPR.

Consent management and access control mechanisms are commonly implemented through smart contracts to automate compliance and regulate data usage (Dauden-Esmel et al., 2021; Liang et al., 2023). Under PDP law, data subjects must have control over how their personal data is processed, stored, and shared, aligning with Article 27, Article 43, and Article 44. these requirements raise RQ1 and RQ2. The prior studies provide in Table 2.6 provide strong evidence that HARA's blockchain-based data management can achieve compliance by adopting a hybrid storage model, leveraging smart contracts for hash integrity record and using encryption for sensitive data storage.

Table 2.6. On-chain and Off-chain Paper Summary.

Title	On-chain Recording	Off-chain Recording
GDPR Compliant Data Storage and Sharing in Smart Healthcare System: A Blockchain-Based Solution (Bai et al., 2022)	Public information. For Private Information, the hash of data.	Private Information Personally Identifiable Information (PII) and Personal Health Information (PHI) stored using the InterPlanetary File System (IPFS) in an encrypted form.
Implementing GDPR-Compliant Surveys Using Blockchain (Martins Gonçalves et al., 2023)	Hash Data, User ID, Timestamp, Consent.	Personal information is stored off-chain using the InterPlanetary File System (IPFS)
A Privacy-Preserving Blockchain Platform for a Data Marketplace (Klaine et al., 2023)	Transactions Between Parties: This includes fixed-term contracts that must be completed by all parties and a deposit mechanism linked to smart contracts for marketplace purchases. Performance and rating of buyer and seller.	Traded [®] data stored in external storage, AWS cloud.

Table 2.7. On-chain and Off-chain Paper Summary (cont.).

Title	On-chain Recording	Off-chain Recording
Lightweight Blockchain-based Platform for GDPR-Compliant Personal Data Management (Dauden-Esmel et al., 2021)	Consent Smart Contract: Defines the categories of personal data that can be collected and specifies the duration of consent. Purpose Smart Contract: Grants permission for the use of specific data within a limited time frame.	Personal Data.
PDPChain: A Consortium Blockchain-Based Privacy Protection Scheme for Personal Data (Liang et al., 2023)	Hash of the encrypted personal data and access control.	Personal Data stored in distributed private cluster. Personal data encrypted with Paillier Homomorphic Encryption.

Despite its advantages, the on-chain and off-chain approach faces challenges such as ensuring off-chain data availability and balancing decentralization with external storage (Klaine et al., 2023). These challenges support H1 that implementing the PDP Law will require significant changes to HARA's data acquisition product, particularly in data retention, processing, and subject rights management. Additionally, the success of hybrid blockchain frameworks in GDPR-compliant data management (Bai et al., 2022; Liang et al., 2023) reinforces the H2 that a combination of on-chain and off-chain model will be the most effective strategy for ensuring.

By leveraging these insights from previous research, this study will propose a practical framework for integrating privacy-compliant blockchain storage into HARA's

agricultural data system, ensuring alignment with Indonesia's PDP Law while preserving the security and usability of blockchain-based data management.

