BLOCK CHAIN BASED INCOME TRACEABILITY SYSTEM FOR EQUITABLE WELFARE DISTRIBUTION

A Project Report

submitted to the APJ Abdul Kalam Technological University

in partial fulfillment of the requirements for the award of degree of

Bachelor of Technology

in

Computer Science and Engineering

by

DEVATHMAJ A KALIYATHAN (STM22CS021)

MUHAMMED JINAS T P (STM22CS038)

NEVIN R PRADEEP (STM22CS043)

NIRANJ C N (STM22CS044)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

ST. THOMAS COLLEGE OF ENGINEERING AND TECHNOLOGY

MATTANNUR

October 2025

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

ST. THOMAS COLLEGE OF ENGINEERING AND TECHNOLOGY

MATTANNUR

2025 - 2026

CERTIFICATE

This is to certify that the report entitled **BLOCK CHAIN BASED INCOME TRACEABILITY SYSTEM FOR EQUITABLE WELFARE DISTRIBUTION** submitted by **DEVATHMAJ A KALIYATHAN** (STM22CS021), **MUHAMMED JINAS T P** (STM22CS038), **NEVIN R PRADEEP** (STM22CS043), **NIRANJ C N** (STM22CS044) to the APJ Abdul Kalam Technological University in partial fulfillment of the B.Tech. degree in Computer Science and Engineering is a bonafide record of the project work carried out by them under our guidance and supervision. This report in any form has not been submitted to any other university or institute for any purpose.

Mr. Jithin S
Project Guide
Assistant Professor
Department of Computer Science
and Engineering
St. Thomas College of Engineering
and Technology
Mattannur

Mr. Jithin S
Project Coordinator
Assistant Professor
Department of Computer Science
and Engineering
St. Thomas College of Engineering
and Technology
Mattannur

Dr. Sreekesh Namboodiri T
Head of Department
Dept. of CSE
St. Thomas College of Engineering and Technology
Mattannur

DECLARATION

We hereby declare that the project report **BLOCK CHAIN BASED INCOME TRACEABILITY SYSTEM FOR EQUITABLE WELFARE DISTRIBUTION**, submitted for partial fulfillment of the requirements for the award of degree of Bachelor of Technology of the APJ Abdul Kalam Technological University, Kerala is a bonafide work done by us under supervision of **Mr. Jithin S**

This submission represents our ideas in our own words and where ideas or words of others have been included, we have adequately and accurately cited and referenced the original sources.

We also declare that we have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in our submission. We understand that any violation of the above will be a cause for disciplinary action by the institute and/or the university and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree, diploma or similar title of any other university.

DEVATHMAJ A KALIYATHAN
MUHAMMED JINAS T P
NEVIN R PRADEEP
NIRANJ C N

Mattannur
21-10-2025

ACKNOWLEDGEMENT

We take this opportunity to express my deepest sense of gratitude and sincere thanks to everyone who helped us to complete this work successfully. We are extremely thankful to our Principal **Dr. SHINU MATHEW JOHN** for giving us his consent for this project.

We express our sincere thanks to **Dr. SREEKESH NAMBOOTHIRI T**, Head of the Department, Department of Computer Science and Engineering for providing us with all the necessary facilities and support.

We would like to express our sincere gratitude to the Project coordinator, **Mr. JITHIN S**, Asst. Professor, Department of Computer Science and Engineering for the support and co-operation.

We would like to place on record our sincere gratitude to our project guide **Mr. JITHIN S**, Asst. Professor, Department of Computer Science and Engineering for the guidance and mentorship throughout this work.

Finally we thank our family, and friends who contributed to the succesful fulfillment of this project work.

DEVATHMAJ A KALIYATHAN
MUHAMMED JINAS T P
NEVIN R PRADEEP
NIRANJ C N

Abstract

Ensuring equitable welfare distribution and accurate income assessment for daily wage and informal sector workers in India remains a persistent challenge. Existing methods such as income tax filings, government surveys like the Periodic Labour Force Survey (PLFS) and the Socio-Economic Caste Census (SECC), ration card declarations, and banking data are either infrequent, fragmented, or exclude the vast informal workforce, which constitutes over 80ily manipulated, or self-declared income records. Consequently, many individuals falsely claim Below Poverty Line (BPL) status, inflating the BPL population and causing misal- location of welfare resources. This deprives genuinely deserving populations of adequate benefits, contributing to social and economic inequities. To address these challenges, we propose an Income Traceability System powered by blockchain and AI/ML technolo- gies. The system immutably records every digital wage payment on a private blockchain, creating a tamper-proof, verifiable income history. AI models accurately estimate in- come bands, dynamically classify individuals for welfare eligibility such as BPL or Above Poverty Line (APL). Additionally, the system offers a government-ready, real-time mon- itoring dashboard for policymakers, ensuring privacy-preserving data handling, context- aware income profiling, and automated welfare eligibility flagging. Compared to tradi- tional methods, this system offers greater accuracy, real-time insights, and adaptability to the fluctuating nature of informal sector incomes. By integrating blockchain-verified wage records with AI-driven analytics, the framework enhances the fairness, integrity, and efficiency of welfare delivery, ensuring benefits reach the truly deserving. Combined with strong governmental reforms promoting digital and cashless transactions, the system can also help curb illegal activities like money laundering, tax evasion and unaccounted financial flows, contributing to a more transparent, inclusive, and economically resilient society.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

In India, many welfare schemes have been implemented to aid and uplift different sections of society who are financially incapable of becoming empowered on their own. While these schemes provide incentives based on the current financial status of citizens through classifications such as APL and BPL, the eligibility criteria mainly depend on their income pattern throughout a financial year or their annual income. The methods employed for gathering this income data, such as manual surveys, have often been found to produce inaccurate and inconsistent records. The failure to devise appropriate strategies that could overcome the shortcomings of the existing approaches in an effective and economical manner has been the root cause of the continued use of these outdated methods even in the current scenario.

While existing systems can accurately record the annual income of formal sector workers, the lack of means to verify the credibility of self-proclaimed incomes of informal sector workers has resulted in major inaccuracies. Moreover, many citizens who are aware of the potential benefits of the incentivized categories and the liabilities of the category they truly belong to may provide false information to enjoy undeserved benefits. When examining actual census records of annual income, the statistics for such misclassification are alarmingly high. The consequences include financially weaker sections being deprived of their rightful incentives, undeserving individuals receiving tax exemptions, and reduced opportunities for the government to design better and more targeted welfare schemes for the truly deserving population.

We have identified the need for an approach that integrates blockchain technology to record financial transactions made throughout the year in combination with the Unified

Payments Interface (UPI). This can create an immutable and traceable storage system that can be used for accurate income data collection. Through this paper, we present our proposed system, which aims to address these challenges while effectively eliminating the need for tedious tasks such as manual data collection through traditional census methods and facilitating an equitable welfare distribution.

1.1 Motivation

The methods currently in practice for annual income data collection for both informal and formal sectors of workers have been found to be insufficient and inaccurate. While the credibility of records can be ensured to a certain degree for income data from the formal sector, the same cannot be said for the informal sector, where a large portion of the data is self-proclaimed. This leaves room for both unintentional and intentional inaccuracies. Moreover, the data being stored in centralized systems makes it vulnerable to tampering.

The direct consequences of these inaccuracies are reflected in the welfare policies introduced by the government to aid financially backward sections. Undeserving individuals who are misclassified into these categories increase the overall count of beneficiaries, thereby reducing the share of funds available to each deserving person. Furthermore, a significant portion of the funds allocated for genuine welfare schemes is misused in this way. Such misclassified individuals are also exempted from taxes, leading to a reduction in government tax revenue, which in turn forces the government to increase the tax burden on genuine taxpayers.

As elaborated above, these inaccuracies stunt the economy's growth and, consequently, the development of the country as well as the standard of living of its citizens. Realizing these issues encouraged us to explore a blockchain-based, AI-integrated approach to address the shortcomings of the current system. Blockchain, when used to record wage transactions, ensures immutability and traceability, while AI can be used to estimate income patterns of individuals throughout the year. This enables accurate annual income calculation even for workers in the informal sector.

By implementing this pilot project, we aim to explore its potential for adaptation and integration into the economy. With proper governmental policies and UPI integration, we can achieve a more equitable welfare system and a much more transparent economy.

1.2 Problem definition

Determination of eligibility for welfare schemes and other governmental benefits, as well as the imposition of various taxes in India, follows the classification of citizens into different categories based on their annual income. The current methods in practice for obtaining annual income data, such as manual census, have been found to produce inaccurate results. This inaccuracy is highly significant in the informal sector, where no verifiable wage records exist and data collection is heavily reliant on self-declarations. As a result, individuals may exploit loopholes in the system to get falsely classified as beneficiaries.

The issue this generates is the large-scale misallocation of welfare funds to undeserving individuals, while those truly in need are left with minimal aid owing to the inflated beneficiary population. Such individuals are also exempted from their liabilities, such as income taxes, further reducing the government's revenue that could be channeled into additional welfare schemes and developmental strategies. This not only prevents economically backward classes from improving their circumstances but also negatively impacts the overall economic and social standards of the country.

Another issue we address here is that the centralized storage of personal annual income data is vulnerable to tampering, and the lack of transparency and traceability further makes the current system inadequate. We devised a blockchain-based, AI-integrated pilot project named Tracient to solve these identified issues and verify its performance in a simulated environment when integrated with the UPI system of India.

Chapter 2

Literature Review

A literature survey is an essential component of any research or project work that involves studying and analyzing existing research papers, publications, and related systems in the chosen domain. It provides an overview of the progress made in the field, highlights existing challenges, and helps in identifying the research gaps that the current project aims to address. By reviewing previous works, researchers gain valuable insights into the methodologies, technologies, and design strategies adopted by others, allowing them to build upon proven concepts while avoiding known limitations.

The literature survey also establishes the theoretical foundation and credibility of the project. It ensures that the proposed system is not developed in isolation but rather extends and improves upon existing knowledge. Through a proper review, one can evaluate the strengths and weaknesses of prior approaches, understand the feasibility of integrating specific technologies, and identify opportunities for innovation within the selected field.

Although we were unable to find any previous implementations directly addressing income traceability and welfare distribution using blockchain, the literature papers we referred to during our survey offered essential insights that guided the development of our proposed system.

The paper “Blockchain-Based KYC and Access Verification for Financial Institutions” provided valuable understanding of improving data security and ensuring privacy through decentralized identity verification. The work “Blockchain-Integrated Digital Payment Systems in the Public Sector: Innovation, Transparency, and Economic Efficiency in Emerging Markets” offered clarity on the perks and challenges of integrating blockchain technology with India’s Unified Payments Interface (UPI), highlighting both its feasibility and potential societal benefits. Similarly, the paper “On the Integration of Artificial Intelligence and Blockchain Technology: A Perspective About Security” provided guidance on the challenges and best practices for properly integrating AI mechanisms with blockchain frameworks to enhance decision-making and data reliability.

Additionally, our base paper, “An Explainable Federated Blockchain Framework with Privacy-Preserving AI Optimization for Securing Healthcare Data,” significantly influenced the technical direction of our project. This work demonstrates an effective integration of blockchain and AI within a federated learning environment to ensure data privacy, transparency, and model optimization. Many of the core principles and architectural concepts presented in this study, such as decentralized validation, explainable AI, and privacy-preserving mechanisms, form the foundation for several technical aspects implemented in our system as well.

Together, these studies provided the conceptual and technical groundwork necessary to design our blockchain-based, AI-integrated income traceability system, ensuring it is secure, transparent, scalable, and adaptable for equitable welfare distribution.

2.1 Blockchain-Integrated Digital Payment Systems in the Public Sector: Innovation, Transparency, and Economic Efficiency in Emerging Markets

The 2025 paper by Vani Tyagi and Dr. Yogita Beri, titled ”Blockchain-Integrated Digital Payment Systems in the Public Sector: Innovation, Transparency, and Economic Efficiency in Emerging Markets,” offers a detailed look at how blockchain can improve government-backed payment systems in places like India. It doesn’t just talk about theory; it uses India’s Unified Payments Interface, or UPI, as a real-world example to build its case. The main issue the paper tackles is that even with successful systems like UPI, public finance still struggles with things like corruption, a lack of full transparency, and high costs for certain transactions. The authors suggest that instead of replacing UPI, we should enhance it. Their proposed solution is a hybrid model. In this setup, the super-fast and scalable UPI system would keep handling all the everyday, small-value payments we make. At the same time, a new blockchain layer would be added to handle specific, high-stakes transactions, like government subsidy payments or large public procurement contracts, where having a secure and unchangeable record is most important.

To show why this is a good idea, the paper spends a lot of time on the success of UPI. It details its massive growth from its launch in 2016 to becoming a system that processes billions of transactions every month across hundreds of banks. This part of the analysis is important because it proves that a large-scale, government-led digital finance project can actually work. To illustrate this, our report should include Figure 2 and Figure 3 from their paper, which show the growth in UPI's transaction volume and the number of participating banks. The paper also brings in user survey data showing that even though people use UPI a lot, a good number of them are still worried about security and whether the system can be tampered with. This is where the authors argue that blockchain's core feature of creating a permanent, tamper-proof ledger can help build more trust.

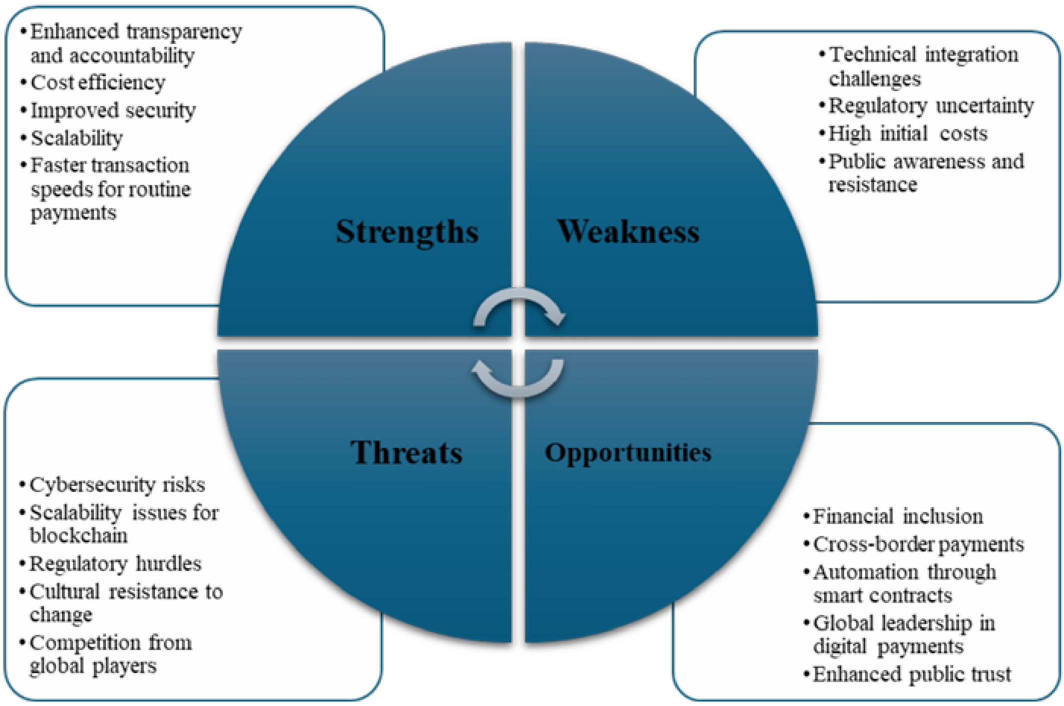

Figure 2.1: SWOT Analysis:Challenges in a Hybrid System

The most useful part of the paper for understanding the whole picture is its SWOT analysis of this proposed hybrid system. It gives a really balanced view of the pros and cons. The strengths are pretty clear: you get better transparency and accountability from blockchain, while keeping the speed and low cost of UPI for normal payments. The weaknesses are also significant, mainly the technical difficulty of connecting a centralized system with a decentralized one, the lack of clear regulations for blockchain, and the high initial cost to build it. The opportunities are exciting, like improving financial inclusion for people in rural areas and using smart contracts to automate payments. Finally, the threats are what you might expect, such as cybersecurity risks and the challenge of getting people and institutions to adopt a new way of doing things.

The paper is significant because it shifts the discussion about blockchain from a ”what if” scenario to a more practical ”how to” guide. By suggesting we integrate blockchain with a system that already works, instead of trying to replace it, the authors present a much more believable path forward. The idea feels grounded and achievable because it's built on the proven success of UPI. The paper also looks at what other countries like Singapore and Estonia have done, which adds a global perspective and offers lessons on what works. In terms of practicality, the paper is very realistic. It acknowledges that just flipping a switch to a blockchain-based system is not going to happen overnight. It even points to Figure 1 from its research, which shows that globally, blockchain adoption is still very low, with a majority of companies having not even considered it. Because of this, the paper's recommendation for a phased approach is very practical. It suggests starting with small pilot projects in a controlled environment, often called a ”regulatory sandbox.” This would let the government test the technology on a smaller scale, like for a specific welfare scheme, to see how it performs and fix any issues before thinking about a nationwide rollout.

Relevance

- Validates Your Problem Statement:The paper's core argument—that public finance systems in emerging markets need enhanced transparency and efficiency to reduce corruption and better serve citizens—directly aligns with your project's mission to create an ”equitable welfare distribution” system. It provides a strong academic backing for why a system like TRACIENT is necessary.

- Supports Your Hybrid Architecture: Your project architecture, which includes an Integration Subsystem to connect with UPI and Aadhaar, is a practical implementation of the hybrid model this paper advocates for. The paper explains why this hybrid approach is superior to a pure blockchain solution, giving you a solid rationale for your design choice. You can argue that TRACIENT is not trying to replace UPI but is instead providing a specialized, secure ledger for a specific public finance function (income traceability), which can then leverage UPI for the actual payment disbursal.

- Highlights the Challenges You Will Face: The SWOT analysis in the paper essentially provides a pre-made risk assessment for your project. The weaknesses and threats it identifies—technical integration challenges, regulatory hurdles, public awareness, and scalability issues—are precisely the challenges your TRACIENT system will encounter. You can use this section of the paper to build out the ”Challenges and Limitations” chapter of your own report and discuss how your project plans to mitigate these risks.
- The paper’s policy recommendations offer a plausible adoption pathway for TRACIENT. You can position your project not as a final, nationwide solution but as a proposed pilot project that aligns with the paper’s recommendation. This makes your proposal seem more practical and grounded, suggesting it could be tested within a single district or for a specific welfare scheme first, allowing the government to gather data and assess its effectiveness in a controlled manner before scaling up.

2.2 Blockchain-Based KYC and Access Verification for Financial Institutions

The paper ”Blockchain-Based KYC and Access Verification for Financial Institutions” by Beauden John, set for June 2025, looks at a common but frustrating problem in banking: the Know Your Customer, or KYC, process. The author points out that the traditional way of doing KYC is slow, expensive, and annoying for customers because they have to repeat the same verification process for every new financial service they use. More importantly, it’s not very secure. Banks store tons of sensitive personal data in centralized databases, which makes them a prime target for hackers, and data breaches are a huge risk. The solution proposed in the paper is to build a KYC system on a blockchain. The idea is to create a shared, decentralized ledger that all participating financial institutions can trust.

The proposed system is designed as a permissioned blockchain, meaning only authorized banks and financial entities can join and validate information. Customer identities would be stored securely on this network, and smart contracts would be used to automate the whole verification process. A key part of the design is how it handles data storage. Instead of putting bulky, sensitive documents like ID scans directly on the blockchain, which is inefficient and risky for privacy, the system stores them in a secure off-chain location. Only a cryptographic hash, which is like a unique digital fingerprint of the document, is recorded on the blockchain. This way, you get the security and immutability of the blockchain to prove the document's integrity without actually exposing the personal data on the ledger. The framework also uses something called Decentralized Identifiers, which allows customers to have control over their own digital identity and decide which pieces of information to share with which institution. The significance of this research lies in its potential to completely reshape how identity is managed in the financial world. It offers a solution to a problem that costs the industry billions of dollars and is a major source of friction for customers. By creating a single, verifiable, and user-controlled identity, it could eliminate the need for repetitive KYC checks, speed up customer onboarding, and significantly reduce administrative costs for banks. This would make the entire financial ecosystem more efficient and secure. It represents a shift from a system where institutions own and control customer data to one where customers have more sovereignty over their own information, which is a big deal for privacy.

In terms of practicality, the paper is quite realistic and highlights some major hurdles. The biggest challenge is interoperability. For a shared KYC system to work, all the different banks, with their different legacy IT systems, would need to agree on a common standard, which is a massive coordination problem. Another practical issue is the regulatory uncertainty. Laws around data privacy, like GDPR in Europe, have strict rules, including the "right to be forgotten," which directly conflicts with the immutable, permanent nature of blockchain. It's still unclear how a blockchain-based identity would be treated legally. The paper also acknowledges that blockchain technology itself can have performance limitations, like low transaction speeds, which could be a problem for a system that needs to handle real-time verifications at a large scale.

Relevance

This paper is highly relevant to our TRACIENT project for several key reasons. First, the core of TRACIENT is about creating a trusted and verifiable record for an individual, in our case, a worker's income history. The KYC paper provides a solid framework for the identity management part of our system. We can adopt its model of using a permissioned blockchain and Decentralized Identifiers to securely register and manage the identities of workers, employers, and government authorities in our system.

Second, the paper's approach to data privacy and storage is directly applicable to TRACIENT. We will be handling sensitive financial data about workers' wages. The hybrid on-chain and off-chain storage model is a perfect solution for us. We can store the unchangeable proof of a wage transaction on the blockchain, like the amount, date, and hashes of the employer's and worker's identities, while keeping the detailed, personally identifiable information in a secure off-chain database. This would give us the transparency and auditability we need without compromising worker privacy.

Finally, the KYC paper's discussion on using smart contracts for access control is a great blueprint for our policy dashboard. TRACIENT needs to ensure that workers can only see their own data, employers can only manage their own wage payments, and government officials can only access aggregated, anonymized data for policymaking. We can use smart contracts to automatically enforce these access rules, creating a system that is transparent and secure by design. The challenges it outlines, especially around regulation and user adoption, are also the same ones we will face, which helps us anticipate and plan for them in our own project.

2.3 On the Integration of Artificial Intelligence and Blockchain Technology: A Perspective About Security

On the Integration of Artificial Intelligence and Blockchain Technology: A Perspective About Security offers a focused literature review on the convergence of Artificial Intelligence (AI) and Blockchain Technology (BCT), highlighting how their integration can strengthen security and trust in digital systems.

1. **Principles of Integration: Decentralized AI**
   The research addresses security challenges inherent in conventional computing systems. While AI relies on machine learning (ML) and deep learning (DL) processes, BCT offers a decentralized, immutable ledger for transaction records.

Centralized AI models are identified as vulnerable to hacking and manipulation, as all data is concentrated in a single location. The integration of AI and BCT establishes a decentralized AI model, which leverages blockchain to enable processing and decision-making based on trusted, digitally signed, and secure shared data, as illustrated in Figure 2.2. This decentralized approach facilitates collective decision-making, with outcomes verified by network nodes to ensure traceability and integrity.

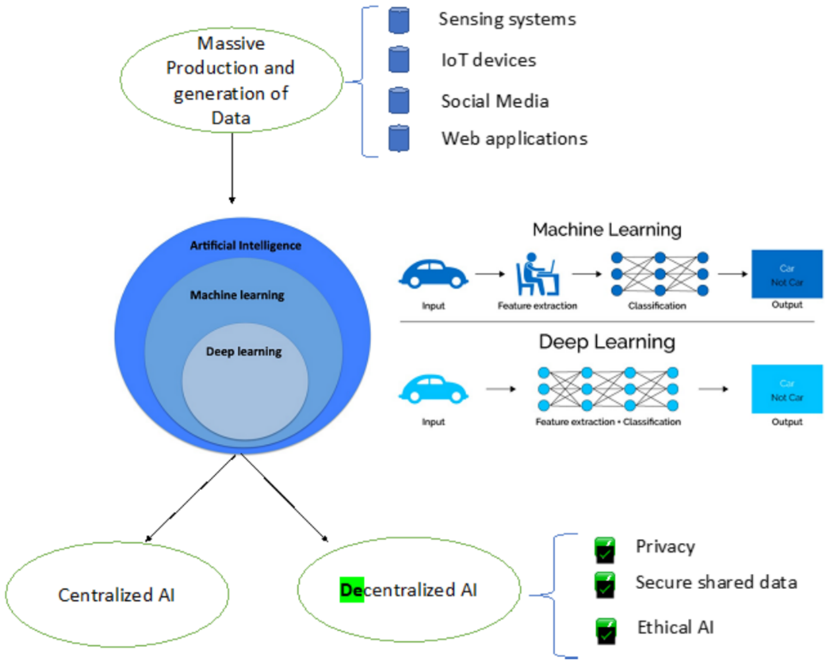

Figure 2.2: Blockchain and AI: centralized and decentralized AI

2. **Benefits of Integration**

   Integrating BCT and AI generates several advantages with direct applicability to secure system design:

   - **Enhanced Data Security and Trust:** The decentralized and immutable nature of blockchain ensures that AI algorithms operate on credible, verified data.
   - **Transparency and Explainability:** Storing AI decisions on the blockchain produces a secure, transparent record, addressing the "black box" issue often associated with deep learning models.
   - **Efficiency and Scalability:** AI can optimize blockchain network operations by enhancing consensus mechanisms and accelerating validation, mitigating

scalability challenges.

- **Model Security:** Blockchain-based Model Version Control allows every iteration of an AI model to be immutably recorded, maintaining a transparent history of model development.

3. **Integrated Security Challenges and Mitigation**

   Despite the benefits, combined AI-blockchain systems are subject to security challenges inherent to both technologies (Figure 2.3). The vulnerabilities and associated mitigation strategies are summarized as follows:

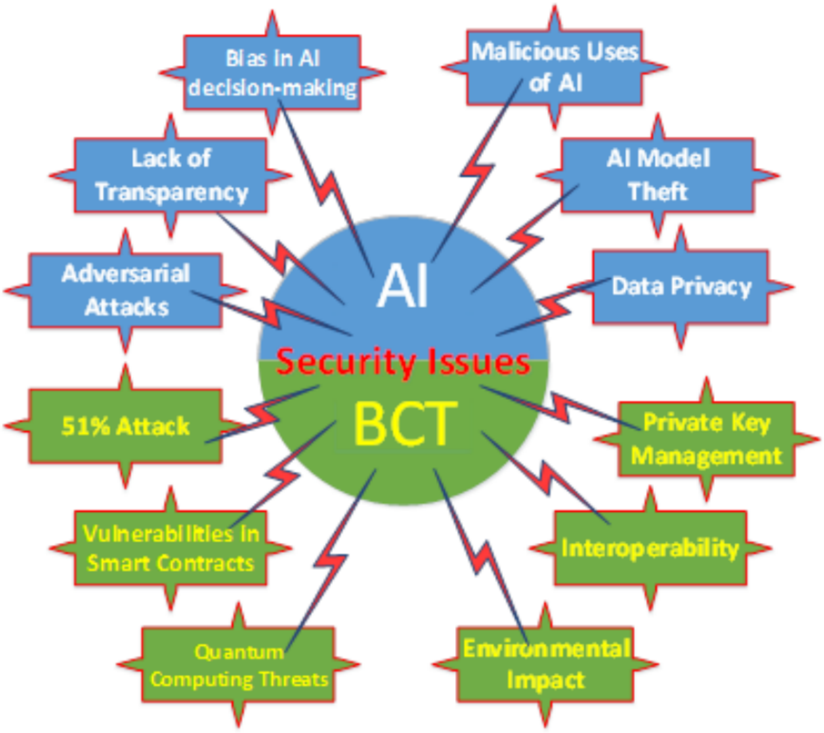

Figure 2.3: Existing security issues associated with Blockchain and AI

- **Artificial Intelligence Security Issues:** AI systems may face adversarial attacks that manipulate input data, bias in decision-making due to training datasets, and privacy concerns. Mitigation strategies include robust model design, privacy-preserving techniques such as federated learning or differential privacy, and transparent AI models to detect and correct biases (Table 2.1).

- **Blockchain Technology Security Issues:** Despite its decentralized architecture, blockchain remains susceptible to 51% attacks, smart contract vulnerabilities, and private key management risks. Mitigation strategies include

Table 2.1: Overview of existing security issues associated with AI

Security Issue	Elaboration
Adversarial Attacks	Input data can be manipulated to deceive AI models
Data Privacy	AI often requires large amounts of data, raising concerns about data protection
Lack of Transparency	The decision-making process of AI systems can be opaque and difficult to interpret
Bias in AI decision-making	Biased data can lead to skewed or unfair AI decisions
AI Model Theft	Theft of valuable AI models, particularly in cloud-based machine learning solutions
Malicious Uses of AI	Misuse of AI technology for malicious purposes, such as deepfakes

adopting improved consensus mechanisms (e.g., Proof-of-Stake) and using smart contract verification tools prior to deployment (Table 2.2).

Table 2.2: Overview of existing security issues associated with Blockchain

Security Issue	Elaboration
51% Attack	If a single entity gains control of the majority of the network's mining hashrate, they can disrupt the functioning of the Blockchain
Private Key Management	Loss or theft of a private key can lead to loss or misuse of Blockchain assets
Vulnerabilities in Smart Contracts	Smart contracts can contain bugs or vulnerabilities, which if exploited can lead to significant losses
Interoperability	Ensuring secure interaction between different Blockchain systems is a significant challenge
Quantum Computing Threats	Quantum computers could potentially break the cryptographic algorithms that secure Blockchain
Environmental Impact	High energy consumption of some Blockchains, particularly proof-of-work Blockchains

- **Regulatory and Trust Implications**

  The study emphasizes that effective regulation must be flexible and technology-neutral. Promoting transparency and accountability is essential for establishing public trust, particularly in applications involving critical welfare distribution systems.

Relevance

The work aligns closely with the major project, Blockchain-Based Income Traceability System for Equitable Welfare Distribution, by supporting its core objective of utilizing blockchain security alongside AI/ML techniques for accurate income analysis.

- **Validation of Decentralized AI (Figure 2.2):** The project employs a private blockchain to maintain verifiable digital wage histories, with AI/ML models estimating income and determining dynamic welfare eligibility. The concept of decentralized AI validates this hybrid approach, providing increased security and trust through digitally signed, verifiable data.
- **Addressing AI Model Risks (Table 2.1):** Potential risks such as adversarial attacks and bias in AI/ML income estimation are critical for the project's success. The paper provides a framework for anticipating and mitigating these threats.
- **Securing the Blockchain Foundation (Table 2.2):** Understanding blockchain vulnerabilities, including smart contract issues, is necessary to maintain integrity and reliability of digital wage records.
- **Mitigation Strategies:** Recommended approaches, including robust model design and privacy-preserving techniques like federated learning or differential privacy, guide the development of a secure, transparent, and equitable system.

2.4 An explainable federated blockchain framework with privacy-preserving AI optimization for securing healthcare data

The base paper for our project is "An explainable federated blockchain framework with privacy-preserving AI optimization for securing healthcare data" by Tanisha Bhardwaj and K. Sumangali. This paper introduces a really advanced framework they call PPFBXAIO, which stands for Privacy Preserving Federated Blockchain Explainable Artificial Intelligence Optimization. The core problem it tries to solve is a big one in healthcare: how to use machine learning to analyze sensitive patient

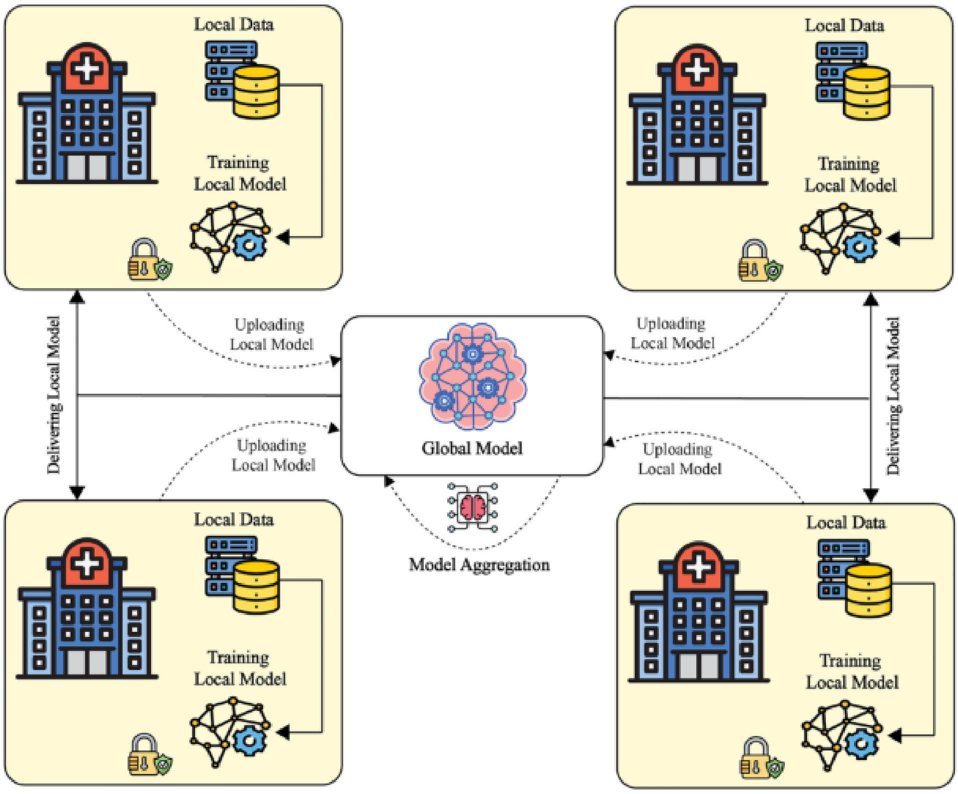

Figure 2.4: A FL Framework

data from different hospitals without actually moving or centralizing that data, all while making sure the process is secure, private, and easy to understand for doctors.

Their solution is pretty complex and brings together a few different technologies. First, they use Federated Learning, which is a way to train a shared AI model across multiple locations without the raw data ever leaving the original hospital's server. This is great for privacy.

Second, they use a blockchain to create a secure, tamper-proof log of all the model updates. This means every change to the shared AI model is recorded permanently, which adds a layer of trust and auditability. They also use smart contracts to automatically validate the quality of these updates before they are accepted.

The third component is Explainable AI, or XAI. A common problem with complex AI models is that they are like a ”black box”—they give you an answer, but you don't know how they got it. In medicine, that's a huge issue because doctors need to trust and understand the AI's reasoning. The PPFBXAIO framework uses XAI tools like SHAP to make the model's predictions interpretable.

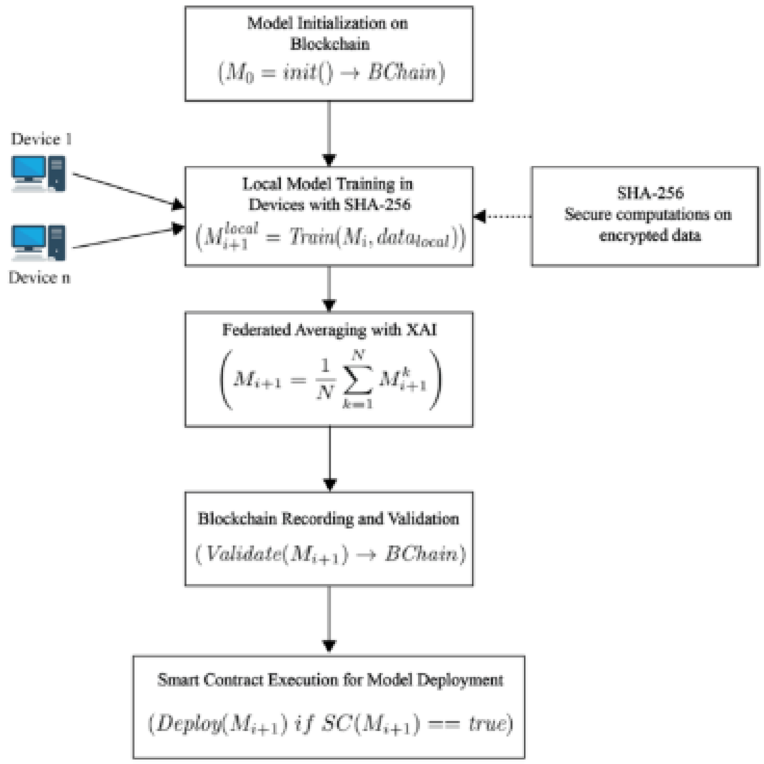

Figure 2.5: Workflow of Proposed Framework for PPFBXAIO

Finally, they use an optimization algorithm called the Levy Grasshopper Optimization Algorithm, or LGOA, to automatically select the most important features from the data, which helps make the AI model more accurate and efficient. They tested their framework using two well-known medical datasets, the Heart Disease dataset and the Wisconsin Breast Cancer dataset, and their results showed very high accuracy and better performance compared to other standard federated learning methods.

The significance of this paper is that it provides a complete, end-to-end architecture for building trustworthy AI in a high-stakes field like healthcare. It's not just about making the AI accurate; it's about making it secure, private, and explainable all at the same time. This holistic approach is what makes it a strong foundation for our own work. It shows that it's possible to design a system that respects data privacy while still allowing for collaborative machine learning.

From a practical standpoint, the paper is ambitious. Integrating all these different technologies is technically very challenging. The authors themselves point out in their study that adding blockchain and XAI increases the computational overhead, meaning it requires more processing power and time compared to a simpler federated learning setup. For example, running XAI calculations and validating transactions on a blockchain adds latency to the system. The framework, as presented, is also quite theoretical and tested in a simulated environment. Getting multiple real-world hospitals to adopt and integrate such a complex system would face enormous logistical and financial challenges. So, while it's a very powerful design, its real-world implementation would require a lot of resources and careful planning.

Relevance

This paper is the direct inspiration for our TRACIENT system, and its relevance is fundamental. We are essentially taking the core architectural concepts of the PPFBXAIO framework and adapting them from the healthcare domain to the socio-economic domain of public welfare. The way the base paper combines blockchain for security and AI for intelligent analysis is the exact same principle we are using. We can point to its success in securing sensitive medical data as a strong justification for why a similar approach can work for securing sensitive income data.

More specifically, we are borrowing the idea of using a private blockchain to create a tamper-proof audit trail. In their case, it's for AI model updates; in our case, it's for wage transactions. This proves that blockchain is a solid choice for creating the verifiable wage history that is central to TRACIENT. Also, their use of advanced AI for classification and optimization inspires our own use of AI/ML models for income estimation and eligibility classification. The LGOA algorithm they use for feature selection is a powerful technique that we could consider for optimizing our own models to ensure they are as accurate as possible.

However, our project also makes some important advancements. The base paper's framework is designed for a very specific, institutional use case and doesn't consider integration with large-scale public infrastructure. A major part of TRACIENT's novelty is its design to connect with real-world systems like Aadhaar for identity verification and UPI for payments, making it much more practical for a nationwide public service. Furthermore, their AI models are for static medical diagnoses, whereas our AI models need to be dynamic and adaptive to handle the constantly changing nature of income and employment in the informal sector. So, we are not just copying their framework; we are evolving it to solve a different, more dynamic problem in a public-facing context.

Table 2.3: Literature Review

Method	Description	Author Details
Blockchain-Integrated Digital Payment Systems in the Public Sector: Innovation, Transparency, and Economic Efficiency in Emerging Markets	This paper proposes a hybrid model combining India's centralized UPI system with a decentralized blockchain. The objective is to improve transparency and security in public finance by using blockchain for high-value government transactions while UPI handles routine payments. It also highlights the practical challenges of integration and scalability.	Vani Tyagi, Dr. Yogita Beri (2025)
Blockchain-Based KYC and Access Verification for Financial Institutions	This research details a decentralized system to replace slow and insecure bank KYC processes. It uses a permissioned blockchain and smart contracts to create a secure, automated identity verification framework where users can control their own data. The design stores sensitive documents off-chain to maintain privacy while ensuring data integrity with on-chain hashes.	Beauden John (2025)
On the Integration of Artificial Intelligence and Blockchain Technology: A Perspective About Security	This survey analyzes how integrating AI and blockchain can enhance security and trust by offsetting each technology's weaknesses. It describes how blockchain can provide secure, trusted data for AI models, enabling decentralized decision-making. The paper also identifies key security threats for both technologies, such as adversarial AI attacks and 51% blockchain attacks, and suggests mitigation strategies.	Oleksandr Kuznetsov, Paolo Sernani, Luca Romeo, Emanuele Frontoni, Adriano Mancini (2024)
An Explainable Federated Blockchain Framework with Privacy-Preserving AI Optimization for Securing Healthcare Data (Base Paper)	This paper introduces the PPFBXAIO framework, which combines three key technologies to securely analyze sensitive healthcare data. It uses Federated Learning (FL) to train AI models without centralizing patient data, a blockchain for a tamper-proof audit trail of model updates, and Explainable AI (XAI) to make the model's decisions understandable to doctors.	Tanisha Bhardwaj, K. Sumangali (2025)

Chapter 3

Proposed System

3.1 Proposed System Description

The proposed system, Tracient, is a Blockchain-Based Income Traceability System for Equitable Welfare Distribution. It aims to revolutionize the welfare distribution process in India by introducing transparency, accountability, and data-driven accuracy in income assessment, especially among the informal and unorganized workforce. The system bridges the credibility gap present in traditional data collection methods and ensures equitable allocation of welfare benefits.

Tracient integrates the strengths of Blockchain, Artificial Intelligence (AI), and Machine Learning (ML) into a unified architecture. It creates a secure, decentralized ecosystem where income records are immutable and verifiable, while AI-driven analytics support intelligent policy decisions.

System Layers Overview

The proposed system is designed as a multi-layered architecture, each component performing a distinct role in ensuring transparency, accuracy, and security. The layers are as follows:

(a) Blockchain Layer – Handles immutable and transparent recording of verified income transactions.

(b) AI/ML Layer – Performs intelligent analysis, income estimation, and dynamic categorization.

(c) Dashboard and Decision Support Layer – Provides a real-time, visual interface for policymakers to monitor, analyze, and take informed decisions.

Each of these layers interacts seamlessly to ensure that the system maintains trust, fairness, and adaptability in income evaluation and welfare distribution.

3.1.1 Blockchain Layer

The blockchain layer serves as the foundation of trust in the proposed system. A private blockchain network is implemented to record verifiable income transactions for both formal and informal sector workers. Each transaction, whether it involves wage payment, self-declared income, or verified employment data, is hashed and stored in the blockchain ledger, ensuring immutability and traceability. Smart contracts are used to automate validation processes such as verifying the authenticity of wage submissions and ensuring compliance with government-approved formats. By decentralizing record storage, the system prevents data tampering and provides full auditability across all welfare-related financial records.

3.1.2 AI/ML Layer

The Artificial Intelligence and Machine Learning layer enhances the system's analytical capability by processing the income data recorded on the blockchain. It estimates probable income for individuals based on employment patterns, transaction trends, and historical data. The models dynamically classify citizens into Below Poverty Line (BPL) or Above Poverty Line (APL) categories using adaptive algorithms, and they also detect anomalies such as underreported or exaggerated incomes. These intelligent processes ensure the credibility of income assessments while maintaining data privacy and improving the overall accuracy, adaptability, and fairness of welfare decision-making.

3.1.3 Dashboard and Decision Support Layer

The system incorporates a privacy-preserving administrative dashboard that provides real-time insights for policymakers and welfare officers. It visualizes aggregated and anonymized information such as verified income histories, welfare eligibility classifications, and regional or demographic income patterns. By offering clear, data-driven visualizations and analytics, the dashboard assists policymakers in identifying income disparities and optimizing the allocation of welfare resources. This ensures that assistance reaches those who need it most, promoting a more transparent and efficient welfare system.

3.1.4 Benefits and Key Advantages

The proposed system introduces several transformative benefits:

- Transparency: Every income record is permanently traceable through blockchain.
- Security: Immutable ledgers and cryptographic verification prevent manipulation.
- Fairness: AI-based dynamic categorization ensures equitable welfare allocation.
- Accountability: Tamper-proof digital records promote honest income reporting.
- Scalability: The modular design allows integration with national systems like Aadhaar and UPI in the future.

3.1.5 Comparative Advantage

Unlike traditional survey-based or centralized income data collection systems, Tracient ensures accuracy, adaptability, and integrity. It overcomes the limitations of static, self-reported data and provides a real-time, evolving view of the economic landscape. By using blockchain as the record base and AI for analysis, the system minimizes human bias, reduces operational inefficiency, and establishes a robust digital framework for fair and accountable income verification.

3.1.6 Socio-Technical Impact

Beyond its technical framework, Tracient addresses a significant social challenge by ensuring that welfare benefits reach the truly deserving sections of society. By combining technological innovation with ethical governance principles, it promotes data transparency, decentralization, and equity. The system has the potential to transform welfare governance in developing economies, setting a global precedent for blockchain-integrated public sector reforms that enhance both social justice and administrative efficiency.

3.2 Requirement Analysis

Requirements analysis for the Tracient system was a crucial process for defining its core functionalities and operational constraints. The goal was to understand and specify what the system must do to create a robust, transparent, and accurate income verification framework. This framework is intended to serve as a single source of truth for determining citizen eligibility across a wide spectrum of government services and benefits, including welfare schemes (BPL/APL classification), educational grants, and income-based reservations for examinations, thereby addressing the limitations of existing fragmented methods. This analysis was divided into two key areas: functional and non-functional requirements.

3.2.1 Functional Requirement

One of the key functional requirements was the ability to immutably record every digital wage payment on a private blockchain, creating a tamper-proof and verifiable income history. This necessitated the use of Hyperledger Fabric and the development of smart contracts (chaincode) in Go with functions to record wages (recordWage), query wage histories (queryWageHistory), and update eligibility flags (updateBPLStatus). Another critical function was the integration of an AI/ML engine to intelligently process this on-chain data. This required the AI to perform several tasks: estimate a worker's annual income from fragmented payment data using models like Linear Regression , dynamically classify their eligibility status for various schemes based on a flexible set of policy rules and income thresholds using models like Random Forest , and detect anomalies that could indicate fraudulent activity using Isolation Forest. Furthermore, the system needed to provide distinct user interfaces: a simple interface for employers to submit wage payments, a portal for workers to view their own income history and status, and a real-time policy monitoring dashboard for government authorities to view aggregated analytics and eligibility distributions.

3.2.2 Non Functional Requirement

The non-functional requirements were equally important to ensure the system's security, trustworthiness, and widespread adoption.Security and Privacy were prioritized. The system must operate on a private, permissioned blockchain to ensure only authorized entities like government bodies and employers can participate. All sensitive data, such as PAN or Aadhaar numbers, must be anonymized before being recorded on-chain, for instance, by using a secure hash like SHA256. The architecture must be privacy-preserving, with provisions for features like Zero-Knowledge Proofs (ZKPs) to verify income status without revealing the underlying data. Usability and Accessibility were also critical. To cater to a user base with varying levels of digital literacy, the system must support multiple modes of interaction, including simple web/mobile apps, as well as SMS and IVR-based interfaces in regional languages. Scalability is essential for a system with national-level ambition. The architecture must be designed to handle a high volume of transactions and be deployable on cloud infrastructure using container orchestration technologies like Docker and Kubernetes

3.2.3 Other Requirement

Hardware requirements

The hardware infrastructure for the Tracient system is primarily cloud-native, meaning the production environment runs on platforms like AWS or Azure, which provide the underlying physical servers as a managed service. This involves utilizing virtualized resources such as scalable compute instances (VMs) to host the blockchain nodes and backend services, managed Kubernetes for container orchestration, and high-performance cloud storage for ledger data. For the development phase, a local workstation with a multi-core processor, a recommended 16 GB of RAM, and an SSD is necessary to efficiently run the entire software stack within a Docker environment. The hardware requirements for end-users are minimal, necessitating only a basic smartphone for the mobile application or any standard phone for interaction via the planned SMS and IVR interfaces.

Software requirements

Go (Golang) is utilized as the primary programming language for developing the on-chain business logic, specifically the smart contracts (chaincode) that run on the Hyperledger Fabric network. These contracts manage the immutable recording of wage data and the updating of eligibility statuses. Hyperledger Fabric itself is the core blockchain framework, chosen for its private, permissioned nature which is suitable for a government-collaborated system. Python, along with libraries such as Scikit-learn, is used to develop the AI/ML models responsible for income estimation, BPL/APL classification, and anomaly detection. A Node.js server using the Express.js framework acts as the backend API layer, handling requests from the frontend, interacting with the blockchain network, and invoking the AI engine for analysis. The frontend dashboard and user interfaces are developed using React. The entire multi-component architecture is containerized using Docker and Docker Compose, which ensures consistency across development and production environments and simplifies deployment. Finally, Git is used for version control throughout the project lifecycle.

3.3 Feasibility Study

A feasibility study was conducted to rigorously evaluate the viability of the Tracient system across four key dimensions: technical, operational, economic, and performance. This analysis ensures that the proposed framework for a unified income traceability system is not only conceptually sound but also practically implementable, financially justifiable, and capable of meeting its performance objectives.

3.3.1 Operational Feasibility

Operational feasibility evaluates how well the system can be integrated into existing government workflows and adopted by its intended users. The system is strategically designed to align with major national initiatives like Digital India, eShram, and the India Stack (Aadhaar, UPI), ensuring it complements rather than disrupts the existing digital infrastructure. A significant operational challenge is the varying digital literacy among informal sector workers, which is addressed by providing multiple user-friendly interfaces, including simple mobile apps, SMS-based interactions, and IVR (Interactive Voice Response) systems in regional languages. To further ease adoption, the rollout plan suggests partnering with Common Service Centres (CSCs) and NGOs to assist workers with onboarding and training. By replacing the repetitive, manual submission of income proofs with a single verifiable digital record, the system streamlines the current bureaucratic process for both citizens and officials, with a phased pilot deployment ensuring that any operational issues are resolved before a national-scale rollout.

3.3.2 Economic Feasibility

Economic feasibility focuses on the financial viability of the project by weighing the initial and ongoing costs against the substantial long-term benefits. The initial costs are associated with the development effort and setting up the cloud infrastructure, while ongoing costs would involve the maintenance and scaling of these cloud resources based on user and transaction volume. However, the economic benefits are extensive and transformative; the primary return on investment comes from the significant reduction in financial leakage caused by fraudulent claims for welfare schemes, educational grants, income-based reservations, and other subsidies. By providing an accurate and tamper-proof income history, the system ensures that benefits are targeted precisely to the deserving population. Additionally, it can improve tax compliance within the informal sector and help curb unaccounted financial flows, contributing to a more transparent economy. The long-term savings from reduced fraud and increased efficiency are projected to far outweigh the project's costs, making it a highly valuable investment.

3.3.3 Performance Feasibility

Performance feasibility assesses the system's ability to meet its objectives for accuracy, speed, and scalability, which are critical for its real-world application. The core value of the system depends on the accuracy of its AI models in estimating annual income and classifying eligibility, as inaccurate predictions would undermine the system's purpose of fair distribution. The performance of the blockchain itself is also crucial; the system must handle a high volume of wage transactions quickly, a requirement that Hyperledger Fabric is designed to meet with its capacity for higher throughput compared to public blockchains. Finally, the end-to-end system must exhibit low latency, ensuring that data from wage submission to blockchain confirmation, AI analysis, and dashboard updates occurs in near real-time. This rapid response is essential to support dynamic monitoring and timely decision-making by policymakers using the government dashboard.

Chapter 4

System Design and Schedule

The Blockchain and AI Enabled Income Traceability System is designed as a secure, multilayered architecture that ensures fairness, transparency, and real time accountability in welfare distribution. It follows a client–server model, where the client side consists of mobile and web applications that act as the main interaction point for workers, employers, and policymakers. These interfaces allow users to submit digital wage data, verify their identity through Aadhaar or PAN (optionally anonymized for privacy), and access income records or welfare eligibility information.

At the heart of the system is a private blockchain network implemented in Go and structured around Hyperledger Fabric principles. This network records every verified wage transaction as an immutable entry in a distributed ledger. Each transaction contains essential metadata such as payment mode, timestamp, and job type, creating a transparent and tamper proof wage history. Smart contract logic manages record validation, integrity checks, and controlled data sharing between authorized entities such as government departments and NGOs.

Working alongside the blockchain is an independent AI analytics engine that functions outside the chain. This engine analyzes blockchain verified wage data to estimate income, detect anomalies, and classify individuals into categories such as Below Poverty Line or Above Poverty Line using statistical and machine learning models. The insights produced by this engine, including predicted annual income, anomaly alerts, and dynamic eligibility scores, are delivered to a real time policy dashboard that provides decision makers with a comprehensive and privacy preserving view of income patterns and welfare distributions.

By combining blockchain based immutability with AI driven intelligence, the system creates a transparent and adaptive digital framework for income traceability. Its modular design supports seamless integration with India's existing digital public infrastructure such as UPI, Aadhaar, and eShram, paving the way for accurate welfare targeting and a more equitable economic ecosystem.

4.1 System Architecture

The proposed system architecture is a sophisticated, multi-layered platform engineered to facilitate complex interactions between workers, employers, and policy administrators. It is built upon a modern microservices paradigm, creating a clear separation of concerns that integrates advanced capabilities like artificial intelligence, blockchain, and critical third-party services. This design ensures the platform is not only robust and scalable but also secure and intelligent in its operations. The primary entry point for all users is through a set of specialized front-end applications: the Worker App, the Employer App, and the Policy Dashboard. Each application is tailored to the specific needs of its user group—for instance, workers might use their app to apply for benefits and track payments, employers to verify employment and manage compliance, and administrators to monitor the ecosystem's health and enforce policies via the dashboard.

All traffic from these applications is securely managed and directed by an API Gateway. This component serves as a unified interface to the backend, handling crucial cross-cutting concerns such as request routing, load balancing, authentication, and rate limiting. By abstracting the backend complexity, the gateway provides a single, secure entry point, protecting the internal services from direct exposure. At the heart of the system lies the central App Service, which functions as the primary orchestrator. Upon receiving a validated request from the API Gateway, this service implements the core business logic, coordinating tasks and delegating responsibilities to the various specialized subsystems.

A key feature of this architecture is its data-driven decision-making capability, powered by the Intelligence Subsystem. This subsystem's core is the AI Engine, which is invoked by the App Service to perform advanced analytical tasks. These tasks include Income Estimation to predict a worker's income for credit or subsidy applications; Anomaly Detection to monitor for fraud or system abuse; and Eligibility Classification to automate the process of determining if an individual qualifies for certain policies. The intelligence of this engine is not static; it is continuously refined through an offline Training Phase. This represents a complete Machine Learning Operations (MLOps) pipeline, including Data Sourcing, Data Preparation, Model Training, Model Validation, and Model Saving, ensuring the models are always up-to-date and accurate.

For processes that demand the highest levels of integrity and auditability, the system leverages an On-Chain Subsystem. This private blockchain network provides a tamper-proof distributed ledger for recording critical transactions. The App Service communicates with the network's Peers, which host the ledger and execute business logic encapsulated in ChainCode (smart contracts). The network's integrity is maintained by an Orderer, which sequences transactions chronologically, and a Certificate Authority (CA), which manages the digital identities of all participants to ensure only authorized entities can access the network. To function within a real-world context, the architecture relies on an Integration Subsystem to connect with essential external services like the Aadhaar Service for identity verification and the UPI Service for instant financial transactions. Finally, to ensure performance and resilience, a Persistence Subsystem stores Snapshots of system state or frequently accessed data, reducing latency and providing a mechanism for state recovery.

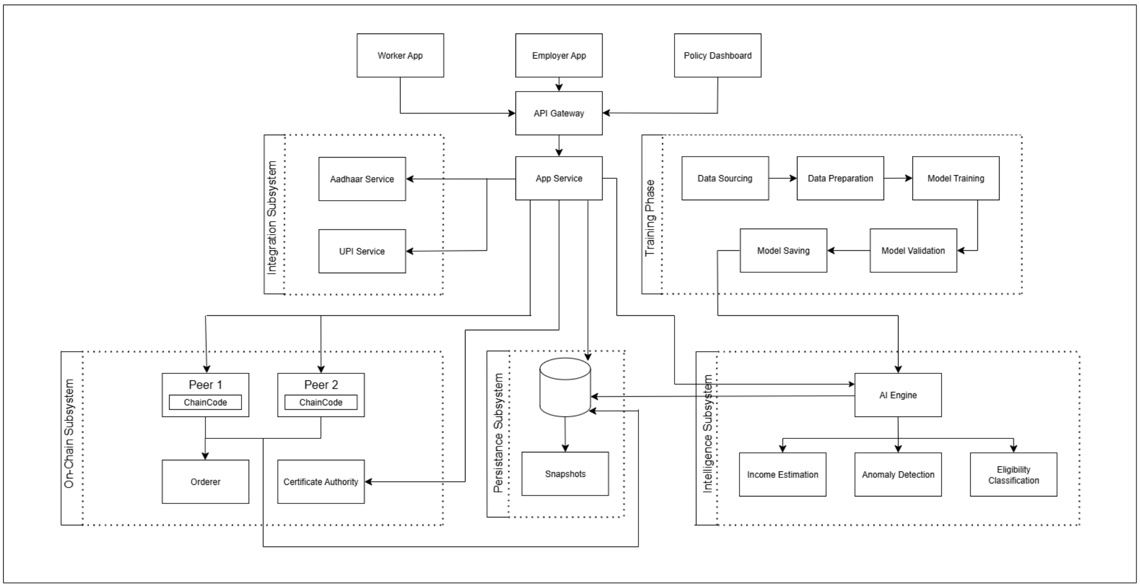

Figure 4.1: Architecture Diagram

4.2 Use Case Diagram

The use case diagram for the Tracient system illustrates the interaction flows between the primary actors—the System Administrator, Employer, Worker, and Government Authority—and the core functionalities of the platform. It defines the specific roles and permissions each actor has within this blockchain-based income traceability ecosystem.

For the System Administrator, the responsibilities are purely technical and foundational. This actor is in charge of the initial setup and ongoing maintenance of the system's infrastructure. Their key actions include managing the blockchain network, deploying and updating the AI models that perform the income analysis, setting the initial policy rules in the system, and managing overall system access for other users.

The Employer's role is central to the data input process within the formal employment flow. Their primary function is to manage their roster of workers and, most importantly, submit digital wage payments directly into the system. Each submission creates a verifiable, immutable transaction on the blockchain, forming the basis of a worker's income history.

The Worker is a primary beneficiary and user of the system with several key interactions. They can view their complete wage history as recorded on the

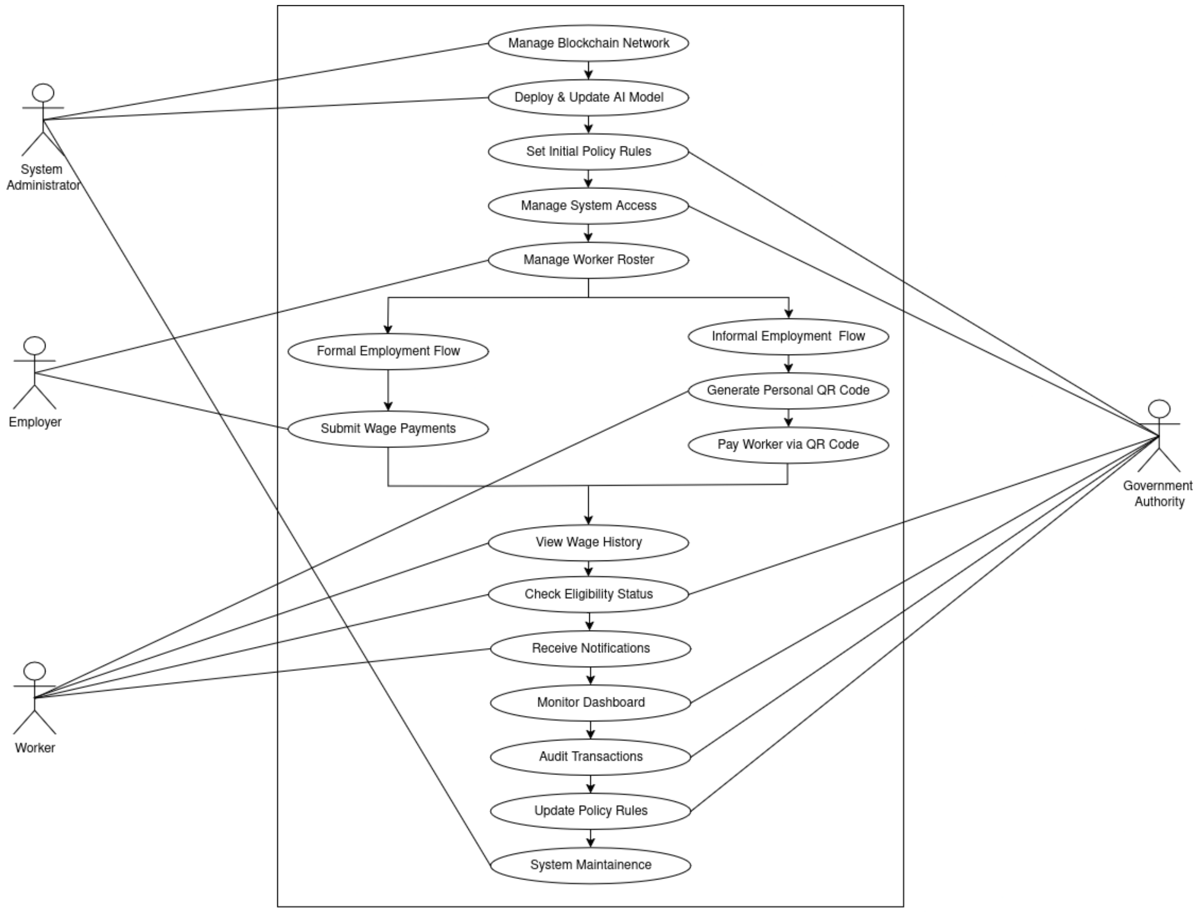

Figure 4.2: Use Case Diagram

blockchain, check their real-time welfare eligibility status as determined by the AI engine, and receive notifications from the system. The diagram also outlines an ”Informal Employment Flow” where a worker can generate a personal QR code, allowing them to receive payments from various sources in a digitally traceable manner, which is crucial for those without a single, formal employer.

Finally, the Government Authority acts as the primary consumer and overseer of the system’s data. They can monitor aggregated, anonymized income trends and eligibility distributions through a dedicated dashboard. They have the capability to audit transactions to ensure transparency and accountability. Crucially, they can also update the policy rules (e.g., income thresholds for BPL status) that the AI engine uses for its classifications, ensuring the system remains aligned with current government policies.

Overall, this workflow clearly defines how each actor contributes to and benefits from the system, creating a closed-loop process for transparent wage logging, intelligent income analysis, and dynamic policy management to ensure equitable welfare distribution.

4.3 Data Flow Diagram

A Data Flow Diagram (DFD) is a graphical representation that depicts how data moves through a system, showing the relationship between processes, data stores, and external entities. In the TRACIENT – Income Traceability System, the DFD provides a clear visualization of how wage-related information is collected, verified, processed, and analyzed across different stakeholders such as Workers, Employers, Banking Systems, and Government Officials.

4.3.1 Level 0 Data Flow Diagram

The Level 0 Data Flow Diagram for the TRACIENT – Blockchain-Based Income Traceability System for Equitable Welfare Distribution illustrates the system as a single, high-level process interacting with multiple external entities. This top-level representation highlights how data enters and leaves the system, focusing on the major information exchanges between users, organizations, and government bodies.

The worker interacts with the system by registering and submitting personal and income details. Their wage data is verified and stored securely on the blockchain. The worker can also view verified income records, check their eligibility status, and receive welfare related updates based on their verified income category.

The employer provides wage and work details of the employees. These records are verified and uploaded into the system to ensure transparency in wage payments. Each transaction entered by the employer is authenticated and converted into a digital proof, preventing manipulation or false reporting of income.

The banking system plays the role of verifying the payment transactions made by the employer. It confirms that the wages transferred to workers are genuine and completed successfully. This verified data is then stored on the blockchain through the TRACIENT system, ensuring all wage records are legitimate and immutable.

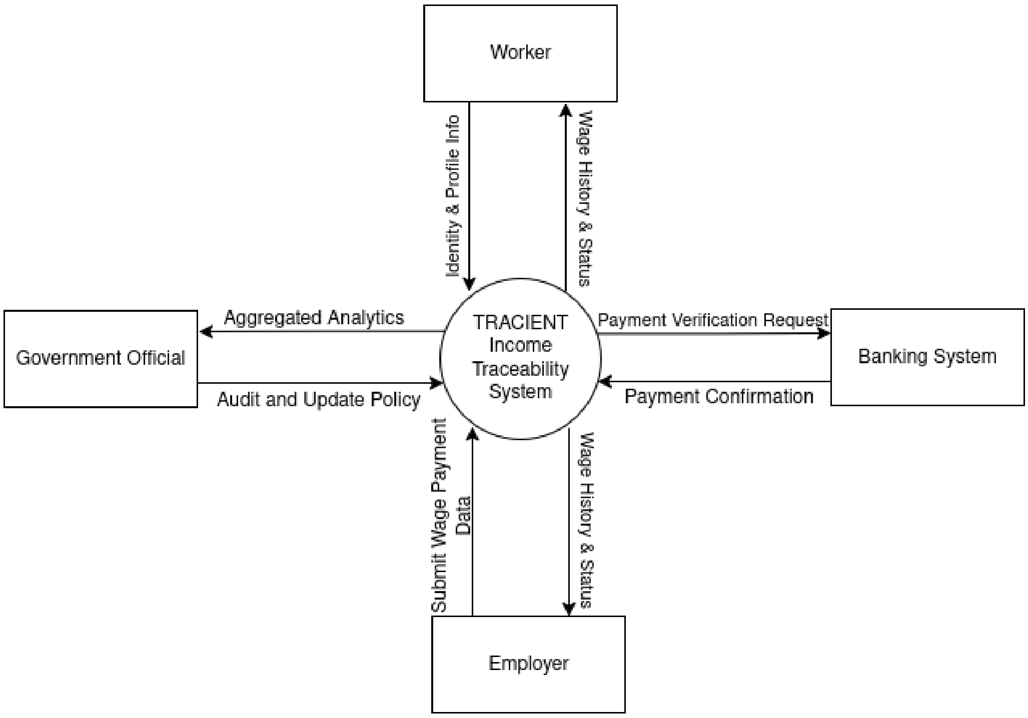

Figure 4.3: Level 0 Data Flow Diagram

The government authority accesses the verified data to analyze income patterns and classify citizens based on income levels such as APL or BPL. The system provides summarized reports generated using AI analysis, which help in fair and accurate welfare distribution and policy decisions.

4.3.2 Level 1 Data Flow Diagram

The Level 1 Data Flow Diagram (DFD) provides the system's highest-level functional overview. It focuses on the entire TRACIENT system as a single, central process (Process 1.0). This diagram clearly establishes the system's boundaries by identifying the three major external entities (Worker, Employer, and Government Authority), the two key data stores (D1 and D2), and the essential data flows that convert raw income reports into verifiable, categorized data for equitable welfare distribution.

The Daily Wage Worker (DWW) is the first external entity, acting as the primary source of input for the system. They send their raw income and transaction data into the system, initiating the process of tracing their financial activity to establish a verifiable record.

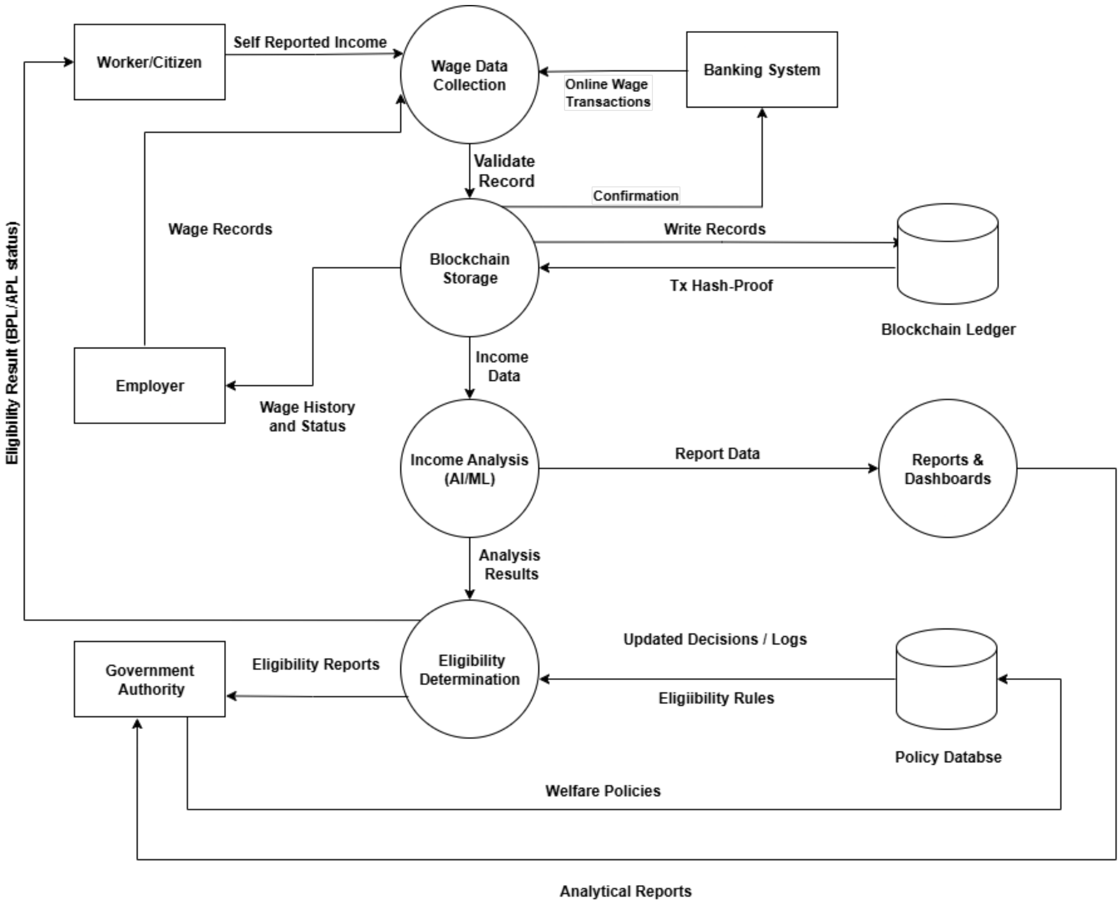

Figure 4.4: Level 1 Data Flow Diagram

The Employer is the second critical external entity whose role is validation. They interact with the system to verify the accuracy of the employment and transaction data provided by the DWW. This check is crucial for ensuring the integrity and trustworthiness of the initial income records.

The Government Authority is the main consumer entity. This authority retrieves the finalized data output from the system, specifically the verified and categorized DWW financial profiles, which are then used to make informed and targeted decisions about the allocation of welfare resources.

The entire functionality of the project is encompassed within Process 1.0: Income Tracing and Welfare Management. This is the core engine that receives input from both the DWW and the Employer, applies the Blockchain security layer, analyzes the data, and transforms it into the essential financial categorization used for welfare management.

D1 Income Records is the first data store, acting as the system's permanent ledger.

It securely houses all the verified, raw transaction and income data received from the external entities. The data here forms the immutable historical record, secured by the underlying Blockchain technology.

D2 DWW Profile is the second data store, holding the processed and summarized information. It contains the personal and final financial categorization of each Daily Wage Worker, which is the direct output of Process 1.0. This is the store the Government Authority queries to retrieve actionable welfare decision data.

4.3.3 Level 2 Data Flow Diagram

The DFD Level 2 provides the crucial expansion of the system's core function by performing a decomposition of the single Process 1.0 (Income Tracing and Welfare Management). This detailed view breaks down the overall function into specific, manageable sub-processes, revealing the internal mechanics of the TRACIENT system. It explicitly illustrates how verified income input is securely recorded via the blockchain, subsequently analyzed through intelligent models, and ultimately transformed into the categorized profile data that the Government Authority requires.

The Daily Wage Worker (DWW) entity is the origin point, sending their claimed employment and transaction data to the system. This entity's interaction initiates the entire income traceability sequence, making the DWW the primary data subject whose financial activities are being authenticated and recorded.

The Employer entity plays the role of the necessary external validator. They interact with the system to confirm and verify the records submitted by the worker. This cross-verification step is vital for establishing the foundational trustworthiness of the income data before it is permanently stored or processed.

The module for Data Submission and Validation acts as the system's intake and scrubbing area. It receives the raw data from the DWW and the confirmation from the Employer. Its function is to perform initial integrity checks, ensure the data is properly formatted, and package the fully verified and clean income information.

The Blockchain Record Creation module is the security core of the system. It receives the validated data and executes the cryptographic and consensus

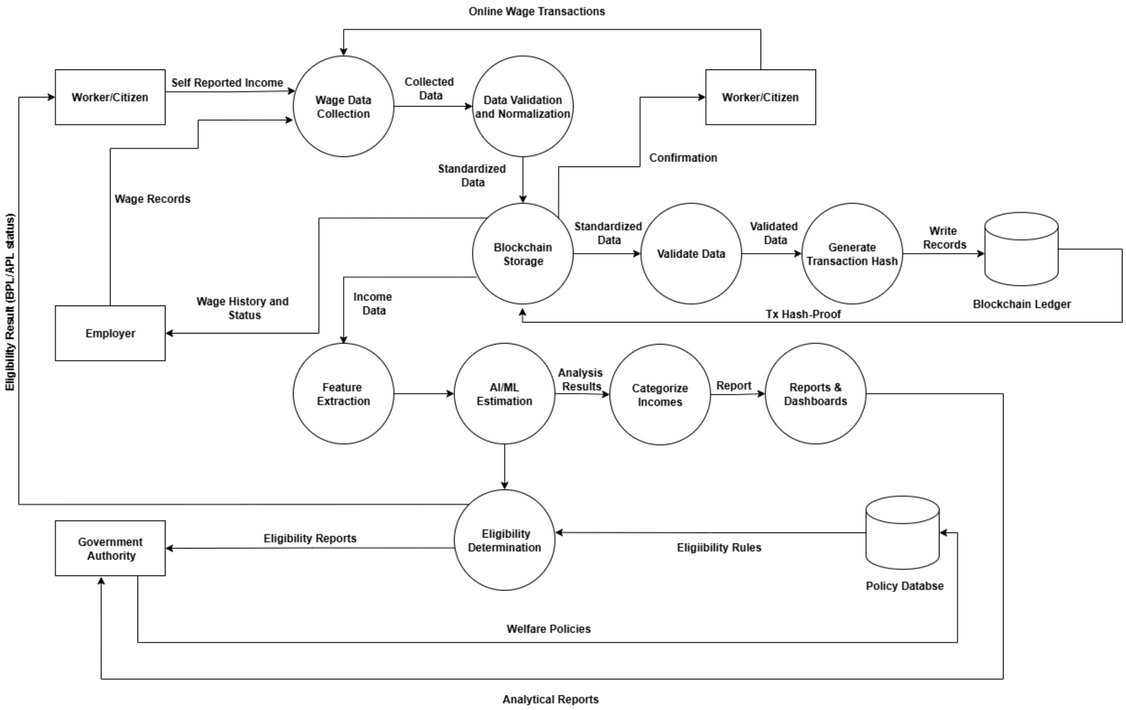

Figure 4.5: Level 2 Data Flow Diagram

mechanisms required to securely and immutably write the transaction record onto the distributed ledger. This action is responsible for updating the D1 Income Records data store, guaranteeing a tamper-proof history.

The Income Profile Analysis module represents the system's intelligence. It retrieves the secured transaction history from the D1 Income Records and applies sophisticated AI/ML logic to assess transaction patterns, calculate average income, and determine the worker's official financial categorization (e.g., eligibility criteria).

The final module, Welfare Data Finalization and Delivery, handles the preparation and delivery of the final output. It takes the categorized status from the analysis module, uses it to update the strategic D2 DWW Profile data store, and manages the secure communication link to the Government Authority for data retrieval.

D1 Income Records is the foundational, secure data store. It maintains the complete, chronological, and cryptographically verified history of every income transaction within the TRACIENT system. This store serves as the single source of immutable truth for the downstream analytical modules.

D2 DWW Profile is the strategic data store, containing the summary-level information. It holds the finalized financial status and categorization for each Daily Wage

Worker. This store is specifically designed to provide the Government Authority with clean, immediate, and actionable data needed for welfare resource allocation.

The Government Authority is the recipient entity, concluding the data flow by interacting with the final module to securely obtain the Verified DWW Categorization. This final interaction confirms that the system has successfully translated raw input into its defined goal: reliable data for equitable welfare distribution.

4.4 ER Diagram

This Entity-Relationship (ER) diagram models a system for managing worker wages, employment history, and regulatory oversight, potentially leveraging blockchain technology. The diagram outlines the relationships among seven key entities: Employer, Worker, Transaction, Wage History, Policy Rule, Anomaly, and Government.

The core of the system revolves around the interaction between an Employer and a Worker. An Employer, identified by `employer_id` and containing details such as company name and GSTIN/PAN, can employ multiple Workers. Each Worker is uniquely identified by `worker_id` and has attributes including `identity_info` and `UPI_id` for receiving payments.

The financial exchange is captured by the **Transaction** entity. An Employer can initiate multiple Transactions, each received by a Worker. Each Transaction includes attributes such as `transaction_id`, `wage_amount`, `payment_method`, and an `on_chain_hash` to indicate that records may be stored immutably on a blockchain. Each Transaction is recorded in the **Wage History** entity, which maintains a Worker's `total_income` and `eligibility_status`. The diagram specifies a one-to-one relationship between a Worker and their Wage History.

Regulatory and compliance aspects are handled by the **Policy Rule** and **Government** entities. Policy Rules, which define specific regulations with `start_date` and `end_date`, can be applied to multiple Workers, and each Worker can be subject to multiple Policy Rules, establishing a many-to-many relationship. The Government entity oversees multiple Wage Histories and is linked to the **Anomaly** entity. If a

Transaction is flagged for an anomaly (e.g., for security or compliance reasons), the Government is notified, ensuring a clear channel for monitoring and fraud detection.

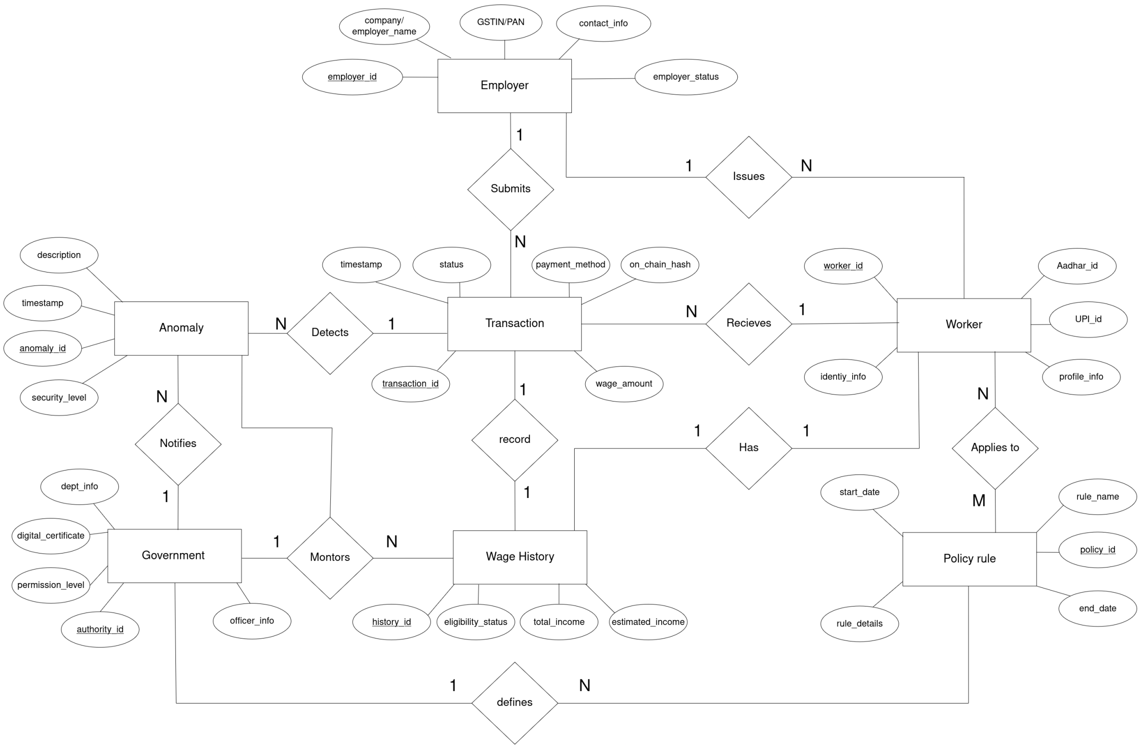

Figure 4.6: ER Diagram

4.5 Gantt Chart

The overall project timeline is systematically managed using a Gantt Chart spanning a fourteen-week period, from July to October. This chart visually represents the project's phased approach, ensuring that all foundational and design activities are completed in a structured sequence before the final documentation phase begins. The total duration allocated for the documentation and design tasks establishes a clear schedule for the deliverables.

The initial phase of the project focused on establishing the groundwork. This began with the crucial Topic Selection, which was completed over 10 days. This was immediately followed by the extensive Literature Survey, which was allocated 21 days to ensure a thorough understanding of the existing research landscape, protocols, and technologies relevant to the TRACIENT system. These tasks lay the theoretical and informational foundation for the entire project.

The core system design phase was then undertaken, focusing on high-level architectural and functional mapping. This phase started with the Architecture Diagram (14 days) and the Usecase Diagram (7 days). This was followed by a sequential documentation of the system's data flow, beginning with DFD Level 0 (3 days) and progressing to the more detailed DFD Level 1 (7 days) and DFD Level 2 (7 days). Simultaneously, the logical structure of the database was finalized through the ER Diagram (6 days), ensuring data integrity was planned early in the process.

Finally, the project shifted focus toward implementation planning and formal documentation. The Module Division task was completed over 5 days to clearly define the coding structure and assignment of sub-systems. The entire project then culminated in the single longest activity: Report Generation, which spans 18 days. This final phase includes all necessary writing, editing, reviewing, and formatting to produce the final comprehensive project report, thereby concluding the scheduled work breakdown structure in October.

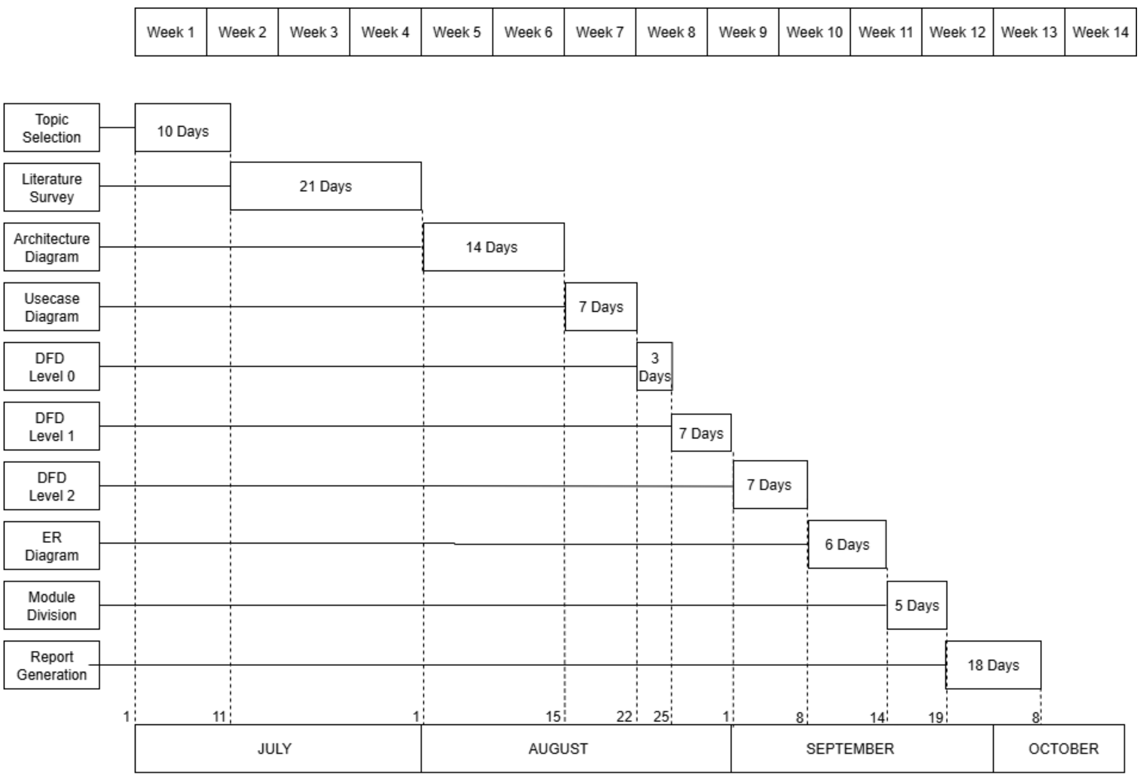

Figure 4.7: Gantt Chart

4.6 Module Division

The project is divided into four specialized modules, each assigned to a dedicated team member responsible for its development, implementation, and quality assurance. This modular approach ensures focused expertise, parallel development, and efficient integration of the Tracient system's components. Below is a detailed explanation of each module.

4.6.1 Blockchain Development and Writing Smart Contracts

This module focuses on establishing the core infrastructure of the Tracient system using Hyperledger Fabric, a private blockchain framework ideal for enterprise and government applications requiring privacy, scalability, and permissioned access. The primary objective is to create a secure, immutable ledger for recording digital wage payments and maintaining verifiable income histories for informal sector workers.

Key Responsibilities:

- Design and implement a multi-organization Hyperledger Fabric network with appropriate channel configurations, consensus mechanisms, and access control policies.
- Develop custom chaincode (smart contracts) in Go programming language to handle core business logic including wage recording, income history queries, and eligibility flag updates.
- Implement cryptographic functions for data privacy, such as SHA-256 hashing for anonymizing sensitive identifiers like Aadhaar or PAN numbers.
- Ensure blockchain scalability to handle high transaction volumes typical of India's informal sector workforce.
- Integrate with Certificate Authorities (CA) for identity management and role-based access control (RBAC) to distinguish between workers, employers, and government entities.

Deliverables:

- Deployed Hyperledger Fabric network with custom chaincode.
- Comprehensive chaincode functions: `recordWage()`, `queryWageHistory()`, `updateEligibilityStatus()`.
- Blockchain network configuration files and deployment scripts.
- Unit tests and integration tests for chaincode functionality.

Integration Points: This module provides the foundational data layer that feeds into the AI backend for income analysis and serves as the data source for frontend dashboards and government monitoring interfaces.

4.6.2 AI & Backend Development (Training, Validation and Testing)

This module encompasses the intelligent core of the Tracient system, combining artificial intelligence and machine learning with robust backend services to process blockchain data, estimate incomes, and determine welfare eligibility. The focus is on developing accurate predictive models while ensuring reliable API endpoints for system integration.

Key Responsibilities:

- Develop and train AI/ML models using Python frameworks (scikit-learn, TensorFlow/PyTorch) for income estimation and BPL/APL classification.
- Implement algorithms including Linear Regression for annual income prediction, Random Forest for classification tasks, and Isolation Forest for anomaly detection to identify fraudulent wage reporting.
- Create a comprehensive dataset pipeline, initially using simulated data and later incorporating real wage transaction data from the blockchain.
- Build a backend API layer using Node.js/Express or Python Flask/FastAPI to expose AI model predictions and handle data processing requests.
- Implement model validation and testing protocols, including cross-validation, performance metrics (accuracy, precision, recall, F1-score), and continuous model retraining capabilities.

- Ensure AI model interpretability and explainability to maintain transparency in welfare eligibility decisions.

Deliverables:

- Trained and validated AI/ML models with performance benchmarks.
- Backend API endpoints for income estimation, eligibility classification, and fraud detection.
- Model training scripts, validation reports, and deployment pipelines.
- Integration with blockchain data queries and frontend data visualization.

Integration Points: This module consumes raw wage data from the blockchain module, processes it through AI models, and provides processed insights to both the frontend user interfaces and the integration module for system-wide deployment.

4.6.3 Frontend & UI/UX Development

This module addresses the user-facing aspects of the Tracient system, creating intuitive and accessible interfaces for diverse stakeholders including informal workers, employers, and government policymakers. The emphasis is on designing a responsive, inclusive, and user-friendly experience that accommodates India's diverse user base and technological landscape.

Key Responsibilities:

- Design and develop responsive web applications using React.js or similar frameworks for multi-role dashboards (worker, employer, government).
- Create user interfaces for wage recording, income history visualization, eligibility status display, and policy monitoring.
- Implement UI/UX best practices ensuring accessibility, including support for regional languages, screen readers, and mobile-responsive design.
- Develop mobile-first interfaces considering the needs of informal workers who may use feature phones or have limited literacy.
- Design data visualization components for income trends, eligibility distributions, and policy insights using libraries like D3.js or Chart.js.

- Conduct user testing and iterate on designs based on feedback from target user groups.

Deliverables:

- Complete web application with role-based dashboards and responsive design.
- Wireframes, mockups, and UI prototypes.
- User testing reports and accessibility compliance documentation.
- Integration with backend APIs for real-time data display.

Integration Points: The frontend consumes data and services from both the AI backend and blockchain modules, providing the primary interface through which users interact with the system's core functionality.

4.6.4 Integration, Version Control, Deployment and Documentation

This module ensures the cohesive functioning of all system components, manages the development lifecycle, and prepares the system for deployment and maintenance. It acts as the glue that binds the specialized modules into a unified, deployable solution.

Key Responsibilities:

- Establish and maintain version control using Git, managing repositories, branches, and merge workflows across all modules.
- Implement continuous integration/continuous deployment (CI/CD) pipelines for automated testing, building, and deployment.
- Coordinate integration testing across modules, ensuring seamless data flow between blockchain, AI backend, and frontend components.
- Develop deployment strategies for cloud platforms (AWS/GCP) using containerization (Docker) and orchestration (Kubernetes).
- Create comprehensive documentation including API references, deployment guides, user manuals, and system architecture documentation.

- Manage project milestones, conduct code reviews, and ensure quality assurance across all deliverables.
- Plan and execute pilot deployments, including environment setup and monitoring.

Deliverables:

- Integrated system deployment on cloud infrastructure.
- Complete documentation suite including technical specifications, user guides, and API documentation.
- CI/CD pipeline configuration and deployment scripts.
- Integration test suites and deployment checklists.

Integration Points: This module orchestrates the outputs from all other modules, ensuring they work together as a cohesive system and are properly deployed for end-users and stakeholders.

Chapter 5

Conclusion

5.1 Conclusion

The Tracient project has successfully demonstrated the transformative potential of integrating a private blockchain with artificial intelligence to create a robust, transparent, and equitable framework for income assessment in India's informal sector. By leveraging the immutability of Hyperledger Fabric, the system establishes a verifiable, tamper-proof wage history, while its AI/ML engine provides the intelligence for dynamic income estimation and eligibility classification. This synergistic approach directly addresses the critical societal challenge of misallocated welfare resources that stems from unreliable and fragmented income data.

The project's key achievement lies in its multifaceted benefits. Technically, it delivers transparency, decentralization, real-time tracking, and scalability. Socially, it enables targeted welfare distribution, empowers informal workers with verifiable income records, and significantly reduces the potential for corruption. Ethically, the system promotes equity, fairness, and data privacy, building trust between citizens and the government. In essence, Tracient provides a comprehensive blueprint for a more efficient, accountable, and just welfare distribution mechanism, ensuring that benefits reach the truly deserving.

References

[1] O. Kuznetsov, E. Frontoni, P. Sernani, L. Romeo, and A. Mancini, "On the integration of artificial intelligence and blockchain technology: a perspective about security," unpublished, 2025.

[2] V. Tyagi and Y. Beri, "Blockchain-integrated digital payment systems in the public sector: innovation, transparency, and economic efficiency in emerging markets," *Indian Council of Social Science Research, Vasanta College for Women (BHU)*, Tech. Rep., 2025.

[3] B. John, "Blockchain-based KYC and access verification for financial institutions," unpublished, Jun. 2025.

[4] T. Bhardwaj and K. Sumangali, "An explainable federated blockchain framework with privacy-preserving AI optimization for securing healthcare data," unpublished, 2025.

[5] R. Mahipal, "Blockchain technology as a catalyst for sustainable development: exploring economic, social, and environmental synergies," *Academic Journal of Interdisciplinary Studies*, vol. 13, no. 2, p. 151, 2024.

[6] O. O. Adefila and O. K. Onasanya, "The economic impact of financial inclusion: a pathway to reducing poverty in the digital economy," *IOSR Journal of Economics and Finance (IOSR-JEF)*, vol. 15, no. 6, pp. 58-67, 2024.

[7] Stanford Graduate School of Business, Consensys, and World Economic Forum, "Study blockchain for social impact: moving beyond the hype," Stanford, CA, 2018.

[8] A. N. Al Hassan, et al., "A survey on blockchain technology: evolution, architecture and security," *IEEE Access*, vol. 9, pp. 47766-47792, 2021.

[9] Bank for International Settlements (BIS), "The organisation of digital payments in India - lessons from the unified payments interface (UPI)," *BIS Paper no. 152*, 2024.

[10] A. P. Singh and M. P. Singh, "A hyperledger fabric-based framework for secure and transparent public distribution systems," *IEEE Transactions on Engineering Management*, vol. 69, no. 4, pp. 1234-1248, Aug. 2022.

[11] S. Gupta and R. Sharma, "AI-driven predictive analytics on blockchain data for enhancing financial inclusion in emerging economies," in *2023 IEEE international conference on big data (Big Data)*, Sorrento, Italy, 2023, pp. 567-576.

[12] N. Kumar and V. Kumar, "A self-sovereign identity model using hyperledger fabric and zero-knowledge proofs for aadhaar verification," *Computers & Security*, vol. 125, p. 102987, Feb. 2023.

[13] D. Shah and P. Patel, "Enhancing transparency and reducing leakage in direct benefit transfer (DBT) schemes using a permissioned blockchain," *Government Information Quarterly*, vol. 40, no. 1, p. 101789, Jan. 2023.

[14] M. Roy and A. Sen, "Digital interventions for income assessment in India's informal sector: challenges and a proposed blockchain-based solution," *Information Technology for Development*, vol. 30, no. 1, pp. 45-67, 2024.