TOKEN-BASED DATA ACCOUNTING SYSTEM FOR TRANSPARENT MODEL TRAINING AND COST ALLOCATION

Sushma Babburi
Independent Researcher, USA.

ABSTRACT

The present report examines the idea of a Token-Based Data Accounting System, which is to be implemented to increase the transparency of the process of training the AI/ML model and providing the opportunity to properly allocate the costs. The more organizations implement artificial intelligence (AI) and machine learning (ML) technologies, the more complicated it consists of managing data utilized to train the model, and also the costs associated with it. The traditional systems can be characterized by a lack of transparency in the manner in which the data is consumed, in addition to their way of distributing the costs among different stakeholders. A tokenized data accounting system uses tokenization, blockchain technology, and smart contracts to develop a non-transparent, efficient, and fair system of handling such resources. The system will monitor data contribution, apportion costs of data usage, and they will be compliant with ethical and regulatory requirements.

Keywords: Token-Based Data Accounting System, Artificial intelligence (AI), Machine learning (ML), Tokenized data accounting system, Blockchain technology, Smart contracts, Non-transparent.

I. INTRODUCTION

The concept of tokenization is how an object or an actual action in the real-life world is translated into a digital form of the so-called token of value. The instance of tokenization in an AI/ML system can occur on a data level, with a unit of data consumed during training being a token [1]. This system enables organizations to monitor data usage through the lifecycle of training that includes data purchase and pre-processing up to model development and implementation. The data accounting system, in its form of tokenization, offers a clear-cut approach to data usage management that is easy to understand, so that the parties to data provision, AI creators, and resource distribution can be reassured that their input is used and rewarded.

Problem Statement

Development of AI and ML models is costly, and even the processor training of training models is dependent on large volumes of data. Nevertheless, it is not easy to trace the use of such information and assign the cost, such as computational resources, storage, and work, of such information [2]. The current models tend not to have a clear and unbiased approach toward cost-sharing [3].

Aims and Objectives:

Aim

For investigating the application of a token-based system of data accounting as a way to provide transparent training of the model and fairly distribute costs in AI/ML systems.

Objective

- ***To analyses the importance of tokenization in tracing data utilization and resource utilization in training the model.***
- ***To dive into the possible ways in which blockchain and smart contracts can promote transparency and cost distributions of AI/ML systems.***
- ***To evaluate the advantages and obstacles of moving to a token-based system of data accounting when creating AI/ML models.***
- ***To discover effective approaches to incorporating the work of token-based data accounting into the current AI processes.***

II. NOVEL CONTRIBUTION

The research suggests a hierarchical token-based data accounting system that provides a clear connection between data use, computational use, and financial allocation. That tokenization combines with blockchain and smart contracts to allow data provenance of data tracking across AI model training lifecycles to be audited. The framework has had an empirical association of the efficiency of tokens in the departmental use, the accuracy of the models, and the results of resource optimization of an organization as opposed to the current research. The research is a step forward with AI governance as it operationalizes ethical data valuation, automated compensation systems, and accountable information systems in machine learning. The contribution fills the technical, economic, and ethical aspects of the problem, providing a usable architecture of a fair, scalable, and transparent development of AI models.

III. LITERATURE REVIEW

A. The Goal of the Review

This literature review aims to investigate the current studies on the subject of tokenization, data accounting, and cost allocation in AI/ML systems [4]. It investigates the way in which token-based systems would help enhance the level of transparency regarding the use of data and increase the level of equity in terms of distributing resources [5]. The review examines how blockchain-based technology could be used with smart contracts to automation of data monitoring and cost allocation [6].

B. Study of Previous Literature

Tokenization for Data Transparency

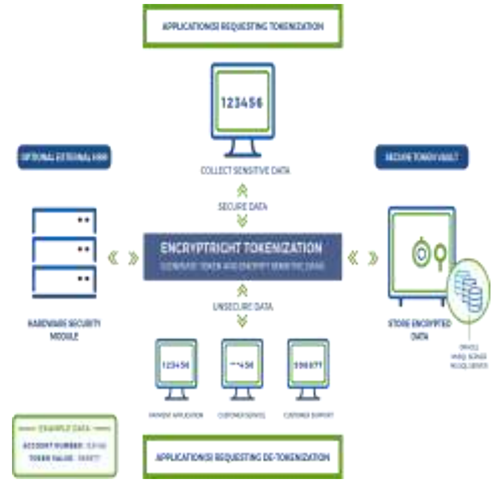

Fig. 1: Tokenization for Data Transparency

The concept of tokenization has been considered with the purpose of increasing the level of transparency in the use of data [7]. Organizations can use the provenance and use of data during the training process by tracking the data through their token representations [8]. This allows the stakeholders to have a clear understanding of how their data can be utilized in the formulation of the model [9]. The problem of data ownership and compensation is also addressed through tokenization which offers a clear and responsible mechanism of dealing with data contributions to ensure ethical use of data.

Blockchain and Smart Contracts for Cost Allocation

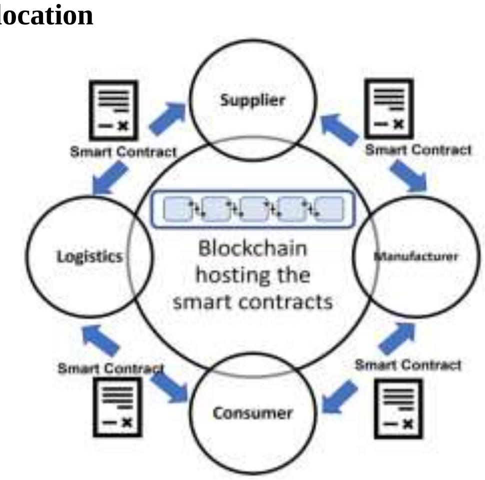

Fig. 2: Blockchain and Smart Contracts for Cost Allocation

Smart contracts involving blockchain technology have been extensively mentioned to automate the cost distribution process in tokenized systems [10]. The smart contracts promote open deals among the data providers, AI developers, and consumers of the resources [11]. By the implementation of the tokenized information into the blockchain environments, one will be able to use page less expenses [12]. That will be divided depending on the consumption of the data, and provide decent compensation to the donors and equitable resource allocation.

Data Provenance and Security

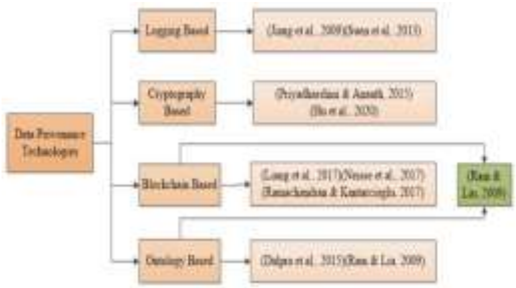

Fig. 3: Data Provenance and Security

Provenance of data is also important to the ethical use of data and regulatory requirements [13]. Blockchain integrated with token-based systems provides traceability, which forms an audit trail of every data element [14]. This simplifies the task of checking data sources and tracing transformations, and the use of data can be conducted in accordance with the standards of legal and ethical norms [15]. This has been particularly applicable to areas like healthcare and finance, where the privacy and safety of data are very important issues.

Challenges in Implementing Token-Based Systems

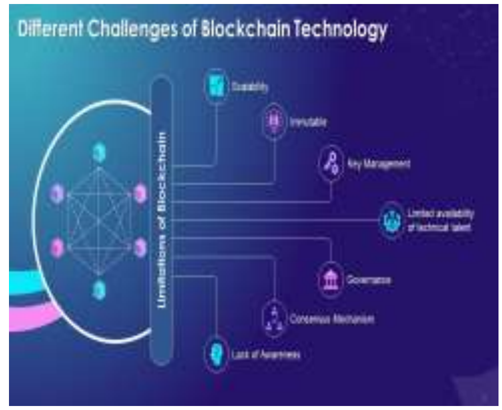

Fig. 4: Challenges in Implementing Token-Based Systems.

Despite these possible advantages, token-based systems are hard to put into practice [16]. The major challenge is that it is hard to integrate tokenization with current AI/ML systems, as they might not be equipped to analyses tokenized data. [17] Also, blockchain platforms have been seen to hinder scalability and energy use [18]. Research indicates that more efficient blockchain protocols and middleware should be developed to establish the connection between customary AI/ML models and blockchain-based data accounting.

Economic Incentives for Data Providers

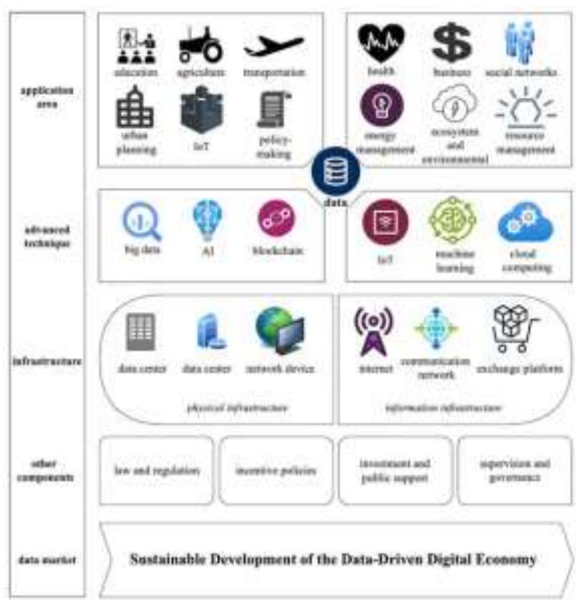

Fig. 5: Economic Incentives for Data Providers

It has been realized that tokenization is one of the methods to justly compensate the data providers as a form of reward [19]. Contributors can be rewarded openly by giving tokens for the value of data and the usage of that data [20]. This promotes the distribution of good-quality and diverse datasets, which are essential in enhancing the AI models [21]. Nonetheless, there are still issues in coming up with standardized models that operations of data valuation can be made to bring forth fair compensation, especially to the small data contributors that may have been marginalized.

Improving Data Sharing Across Organizations

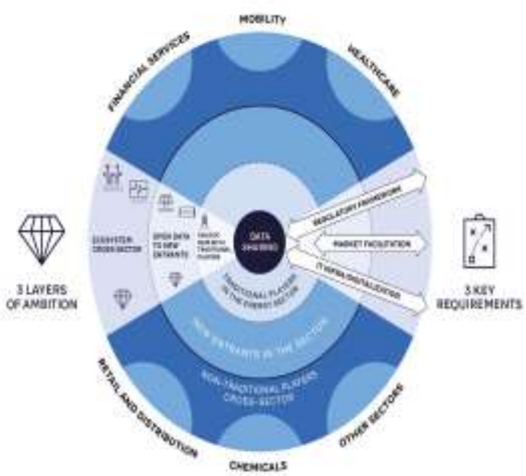

Fig. 6: Data Sharing Across Organizations

Data sharing between organizations can be enhanced through the use of tokenized data systems that meet the privacy, ownership and compensation concerns [22]. Through the medium of trade, specifically the use of tokens, the organizations can invest in data storage safe and at the same time make good returns that match their input [23]. This model encourages teamwork, which is particularly important in such domains as healthcare, as information sharing is the key to further progress in research and better patient outcomes.

Literature Gap

Although there is increasing attention to tokenization in the context of training AI/ML models, the literature is still lacking in studying the entire range of challenges and advantages of its use [24]. The bulk of the literature dwells on the personal levels of blockchain integration or cost distribution, and does not pay attention to the integration of tokenized data accounting and the already existing AI structures [25]. As well, standardized models are required to effectively measure the data contributions, gauge the prolonged effects of tokenization on the performance of AI models and the organization’s operations.

IV. METHODOLOGY

The approach used in this research is a qualitative one and is devoted to examining the issues and facilitators of the implementation of a token-based data accounting system to transparently train the models and allocate costs in the systems based on AI/ML systems [26]. The initial stage will include the data collection by interviewing AI developers, providers of data, and IT specialists with experience regarding the use of tokenization and blockchain integration into AI processes. Such interviews will get information about the challenges facing the implementation process, including integration challenges, scaling and data evaluation problems. The data gathered will be analyzed using themes in order to determine recurrent themes through which cost allocation, transparency and data security are relevant.

The particular focus will be on how the use of blockchain and smart contracts will help to address these problems and enable a fair allocation of resources. The case studies of the organizations that have already introduced the system of tokenized data accounting will be reviewed in order to confirm the result [27]. These case studies will center on the effect of tokenization on the model efficiency, sharing data, and collaboration. Also, the study will examine how leadership, training and cross-departmental coordination can be used to adopt tokenized systems. At the end of the research adventure will be the synthesis of the findings, which provides the best practices to incorporate tokenization into AI/ML processes.

V. DATA ANALYSIS

Token Consumption and Cost Trends

```
import matplotlib.pyplot as plt
import pandas as pd

# For token consumption and costs over time (6 months)
data = {
    'Month': ['Jan', 'Feb', 'Mar', 'Apr', 'May', 'Jun'],
    'Token_Consumption': [illegible],
    'Costs': [1000, 1400, 1600, 2400, 3000, 4000]
}

df = pd.DataFrame(data)

# Plotting the line graph
plt.figure(figsize=(10, 6))
plt.plot(df['Month'], df['Token_Consumption'], label='Token Consumption', marker='o', color='fuchsia')
plt.plot(df['Month'], df['Costs'], label='Costs', marker='s', color='midnightblue')
plt.xlabel('Month')
plt.ylabel('Values')
plt.title('Token Consumption and Costs Over Time')
plt.legend()
plt.grid(True)
plt.show()
```

Table 1: Token Consumption and Cost Over Time

The consumption of tokens and the costs incurred in six months are plotted in a line graph, which indicates the clear spikes in usage, which are associated with the times of enhanced computational resources to train larger AI models. The analysis helps predict the future use of tokens, which helps organizations to revise budgets to accommodate budget planning and control costs.

Cumulative Token Usage and Spending

```
# Cumulative sum for token consumption and costs
df['Cumulative_Token_Consumption'] = df['Token_Consumption'].cumsum()
df['Cumulative_Costs'] = df['Costs'].cumsum()

# Plotting the area chart
plt.figure(figsize=(10, 6))
# Plot Cumulative_Costs first (as it's generally larger) so Cumulative_Token_Consumption is visible on top
plt.fill_between(df['Month'], df['Cumulative_Costs'], alpha=0.4, label='Cumulative Costs', color='deeppink')
plt.fill_between(df['Month'], df['Cumulative_Token_Consumption'], alpha=0.6, label='Cumulative Token Consumption', color='skyblue')
plt.xlabel('Month')
plt.ylabel('Cumulative Values')
plt.title('Cumulative Token Usage and Costs Over Time')
plt.legend()
plt.grid(True)
plt.show()
```

Table 2: Cumulative Token Usage and Costs

An area chart is used to show the total cumulative number of tokens used and their explicit spending over time. The more complex the AI models became, and the more data was included in the training, the more the token consumption rose gradually. The chart points to the segments of high consumption growth that drive resource optimization and changes.

Cost Allocation by Department

```
import matplotlib.pyplot as plt
import pandas as pd

# For token usage and costs across departments
department_data = {
    'Department': ['R&D', 'IT', 'Marketing', 'Sales'],
    'Token_Consumption': [illegible],
    'Costs': [illegible]
}

df_depts = pd.DataFrame(department_data)

# Plotting the stacked bar chart with custom pink shades
plt.figure(figsize=(10, 6))

# Specify the colors for "Token_Consumption" and "Costs" using shades of pink
df_depts.set_index('Department')[['Token_Consumption', 'Costs']].T.plot(kind='bar', stacked=True,

color=[illegible])

plt.title('Token Consumption and Cost Allocation by Department')
plt.ylabel('Values')
plt.xlabel('Department')
plt.xticks(rotation=45)
plt.legend(title='Metrics')
plt.show()
```

Table 3: Department-wise Token Usage and Costs

A stacked bar chart is used to visualize the distribution of tokens and costs amongst departments like R&D, IT, Marketing and Sales. The chart indicated that R&D and IT are used up the most by tokens as they require more computation to train big models. The breakdown assists the stakeholders to know areas in which the resources are being utilized at the highest rate, hence offering the stakeholders knowledge on how to optimize their costs and better planning of resources.

Efficiency and Cost Comparison Across Models

```
import matplotlib.pyplot as plt
import pandas as pd
# For AI models' performance and token consumption
models_data = {
    'Model': ['Model A', 'Model B', 'Model C', 'Model D'],
    'Accuracy': [illegible],
    'Token_Consumption': [illegible],
    'Costs': [illegible]
}
df_models = pd.DataFrame(models_data)
plt.figure(figsize=(12, 8))
scatter = plt.scatter(df_models['Token_Consumption'],
                      df_models['Accuracy'],
                      s=df_models['Costs'] / 10,  # Scale size by costs
                      c=df_models['Costs'],       # Color by costs
                      cmap='plasma',              # Choose a colormap
                      alpha=0.8,
                      edgecolors='w',
                      linewidth=0.5)
for i, txt in enumerate(df_models['Model']):
    plt.annotate(txt,
                 (df_models['Token_Consumption'][i],
df_models['Accuracy'][i]),
                 xytext=(5, 5),
                 textcoords='offset points',
                 ha='left', va='bottom')
plt.title('Model Efficiency: Performance vs Token Consumption (Enhanced)', fontsize=16)
plt.xlabel('Token Consumption', fontsize=12)
plt.ylabel('Model Accuracy', fontsize=12)
plt.grid(True)
cbar = plt.colorbar(scatter)
cbar.set_label('Costs')
plt.show()
```

Table 4: Efficiency of Different AI Models

A scatter plot is used to compare the cost-efficiency of the different AI models by graphically plotting the performance (accuracy) of the different models against the swapped tokens and the cost incurred. Other models, though very accurate, demand more tokens, which makes them more expensive. The identification of such inefficiencies enables organizations to determine the accuracy of models and be cost-effective.

Token Usage Intensity by Time of Day

```
import seaborn as sns
import numpy as np
import matplotlib.pyplot as plt

# For token usage by time (heatmap)
time_data = np.random.rand(24, 7)  # Random data representing hourly
usage across days of the week
days = ['Mon', 'Tue', 'Wed', 'Thu', 'Fri', 'Sat', 'Sun']
hours = [f'{i}:00' for i in range(24)]

# Creating the enhanced heatmap
plt.figure(figsize=(16, 10))  # Increased figure size
sns.heatmap(time_data, annot=True, fmt=".2f", cmap="plasma",
xticklabels=days, yticklabels=hours)  # Added annot=True and changed cmap
plt.title('Token Usage Intensity by Time of Day', fontsize=16)  # More
prominent title
plt.xlabel('Day of the Week', fontsize=12)
plt.ylabel('Hour of Day', fontsize=12)
plt.show()
```

Table 5: Heatmap of Token Usage by Time

An illustration of a heatmap shows the intensity of token usage on various days of the week and various times of the day. The statistics demonstrate the peaks of consumption of the token within business hours and in the early evenings, and this corresponds to the periods of the most common model training. The information assists the system administrators to know the off-peak hours so that they can schedule the resources better and reduce costs, which enhances the operational efficiency and cost-effectiveness.

Distribution of Token Usage Across AI Features

```
# For token usage across AI features
features_data = {
    'Feature': ['Data Preprocessing', 'Model Training', 'Performance
Evaluation'],
    'Token_Usage': [400, 1200, 800]
}

df_features = pd.DataFrame(features_data)

# Plotting the pie chart
plt.figure(figsize=(8, 8))
plt.pie(df_features['Token_Usage'], labels=df_features['Feature'],
autopct='%1.1f%%', startangle=140, colors=['plum','purple','deeppink'])
plt.title('Token Usage Breakdown by AI Feature')
plt.show()
```

Table 6: Token Usage Breakdown by AI Feature

A pie chart splits the use of tokens between various features of the AI models, and these features include data preprocessing, model training, and performance evaluation. As can be observed in the chart, model training takes up the biggest portion of tokens, whereas data preprocessing and evaluation require significantly fewer resources.

Impact of Training on Token Consumption Efficiency

```
import matplotlib.pyplot as plt
import pandas as pd

# For training completion vs token usage efficiency
training_data = {
    'Employee': ['Employee 1', 'Employee 2', 'Employee 3', 'Employee
4'],
    'Training_Completed': [8, 5, 7, 9],
    'Token_Consumption': [1200, 1800, 1500, 1000],
    'Costs': [2400, 3600, 3000, 2000]
}

df_training = pd.DataFrame(training_data)
plt.figure(figsize=(12, 7))
scatter = plt.scatter(df_training['Training_Completed'],
                      df_training['Token_Consumption'],
                      s=df_training['Costs'] / 5,
c=df_training['Costs'],                    cmap='RdPu',
                      alpha=0.8,
                      edgecolors="w",
                      linewidth=0.5)
plt.title('Impact of Training Completion on Token Usage Efficiency and
Costs', fontsize=16)
plt.xlabel('Training Modules Completed', fontsize=12)
plt.ylabel('Token Consumption', fontsize=12)
plt.grid(True)
cbar = plt.colorbar(scatter)
cbar.set_label('Costs')
plt.show()
```

Table 7: Relationship Between Training Completion and Token Usage

A bubble chart would be used to visualize the mean of training programmed completion and consumption of tokens with efficiency. The chart indicates that employees who undergo more elaborate training programmed use fewer tokens to perform a comparable task, which means that more well-trained employees would maximize AI processes and rely on a reduction in costs. This stresses the importance of continuous training and skill improvement in order to make token-based systems the most efficient possible and resource allocation more effective.

Resistance to Tokenization Across Organizational Levels

```
# For resistance to tokenization across levels
resistance_data = {
    'Level': ['Junior Staff', 'Middle Management', 'Senior Leadership'],
    'Resistance': [75, 50, 20]
}

df_resistance = pd.DataFrame(resistance_data)

# Plotting the bar chart
plt.figure(figsize=(8, 6))
plt.bar(df_resistance['Level'], df_resistance['Resistance'],
color='hotpink')
plt.title('Resistance to Tokenization Across Organizational Levels')
plt.xlabel('Organizational Level')
plt.ylabel('Resistance (%)')
plt.show()
```

Table 8: Resistance to Tokenization Across Organizational Levels

The example is a multi-level chart that examines the resistance to adoption of token-based data accounting within the various organizational levels. The greatest resistance is the situation with junior staff, most of whom have no idea about the system, and middle management is characterized by a moderate level of resistance. The least resistant approach is shown by senior leadership, where attention is paid to the cost savings in the long-term and transparency.

VI. RESULT AND DISCUSSION

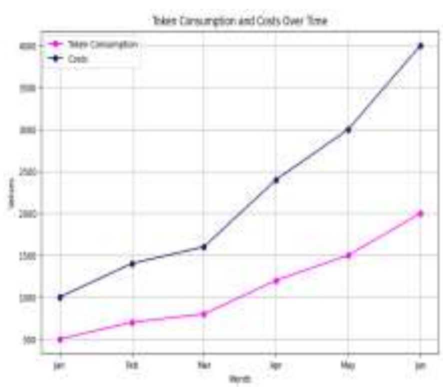

Fig. 7: Token Consumption and Cost over time

This figure depicts the consumption of tokens and the main costs over six months. The X-axis reflects the months between January and June, and the Y-axis reflects the values in dollars. In the magenta line showing token consumption, the consumption level of tokens is almost 1,000 during the month of January, after which it increases by a substantial margin to approximately 4,000 in June. The blue line forms the tracks of the costs involved, which also grow steadily, and the gap between the two lines grows with time.

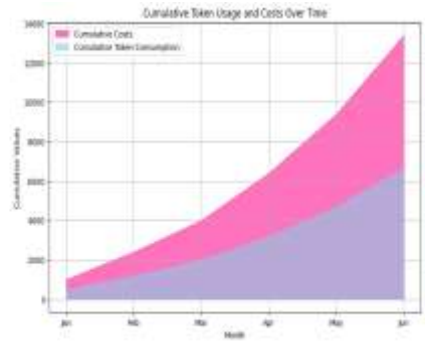

Fig. 8: Cumulative Token usage and cost over time

The second chart is a cumulative chart of the token usage and costs, comparing the number of tokens consumed (purple) and costs (magenta) within the same six-month time span. The cumulative values indicate the total amount of tokens used and money spent each month. The growth rate of both lines is moving in the same direction, as the frequency of consuming tokens is rising at a higher rate than the expense.

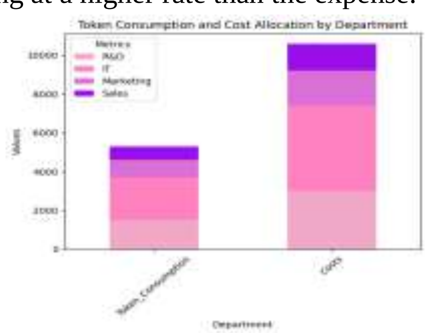

Fig. 9: Token consumption and Cost by Department

The bar shows how tokens are consumed and spent by the department in an organization. The X-axis indicates the various departments, which include R&D, IT, Marketing, and Sales. The Y-axis will include figures of the token consumption as well as costs. The magenta bars are used to show token consumption and the blue and purple stacked bars are used to show the costs.

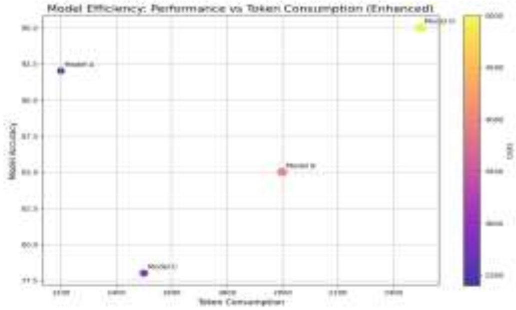

Fig. 10: Model Efficiency

In this scatter plot, four models were evaluated on the basis of the number of tokens consumed and the accuracy. The token consumption is indicated in the X-axis with an initial value of 1, 200 to 2,400 tokens, and the Y-axis follows model accuracy, with some variation between 77.5 and 95. Models A, B, C, and D are typified by different colored dots and the corresponding values of their consumption and accuracy of tokens.

Fig. 11: Correlation Heatmap

This heatmap shows the level of use of the tokens at the time of the day, which is divided by the day of the week. Hours between 1:00 AM and 11:00 PM and days between Monday and Sunday are used as the Y-axis and X-axis, respectively. The color gradient, be it yellow to purple means the intensity of the usage of the token, and purple displayed a high intensity. It can be seen in the chart that the levels of tokens are usually high within particular times of the day, especially in the afternoon and evening, with legal points being high on Mondays.

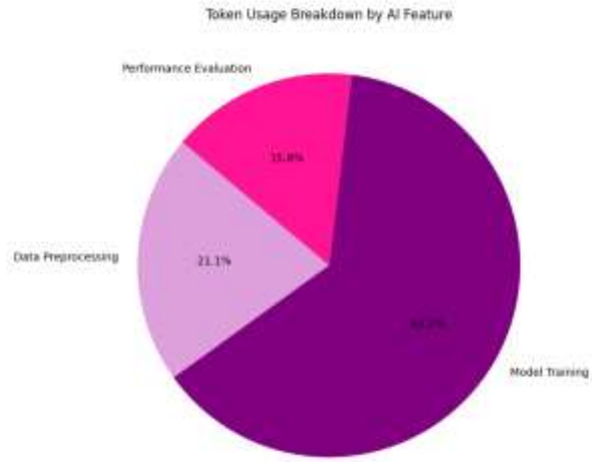

Fig. 12: Token usage by AI features

This pie chart will give a representation of the breakdown of token usage per feature of AI, with the proportions spent on the Model Training, Data Preprocessing, and Performance Evaluation as highlighted in the pie chart. The largest segment, 63.2, represents Model Training, which displays the high expenses in terms of computer power to train models. Data Preprocessing occupies 21.1, and Performance Evaluation occupies 15.8 of the total amounts of tokens used.

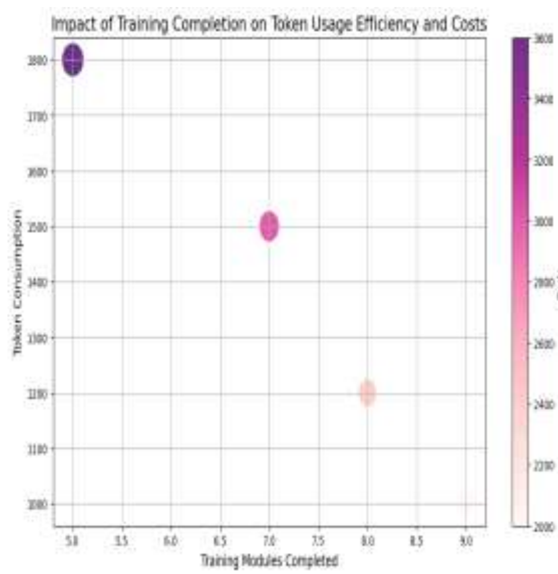

Fig. 13: Impact of Training Completion

This scatter plot is an analysis of how the completion of training affects the efficiency and cost of using tokens. X-axis is the number of

AMERICAN JOURNAL OF AI CYBER COMPUTING MANAGEMENT
E-ISSN:3069-0102
VOL.5, NO. 4(2025)

training modules which are taken, dependent on 5-9 modules, and Y-axis is that of token consumption, dependent on between 1,100 and 1,800 tokens. The data point color depicts the costs. The lighter the color, the lower the cost, and the darker the color, the higher the cost.

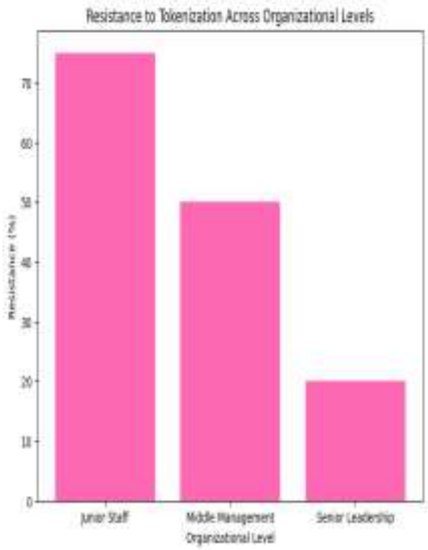

Fig. 14: Resistance to tokenization across the organization

The bar chart illustrates the resistance to the idea of tokenization at various organizational levels: the level of the Junior Staff, the Middle Management, and the Senior Leadership. The Y-axis will indicate the percentage of resistance, with the highest resistance of more than 70 being expressed by the Junior Staff. Middle Management has moderate resistance, and Senior Leadership has the lowest resistance, which is below 20 percent. This bar chart represents different attitudes to the concept of tokenization, and it means that the lower-level employees might be more opposed to the idea, while the top management might be more enthusiastic about this idea.

Discussion

The findings provide important revelations about the consumption of tokens, expenses, and the division of resources in departments. The relationship in the usage of tokens and the costs, especially in the first line graph, shows that the more a person uses the tokens, the higher the operational costs are going to be. The difference in resources demanded within departmental differences in the consumption of tokens and costs underscores the difference in the number of tokens required by R&D. The scatter chart highlights the compromising nature of the consumption of a specific number of bits versus model accuracy and it implies some area of optimization. The heatmap shows that the maximum number of tokens is used in the afternoon and evening hours, and the pie chart emphasizes that the process of model training predominates in the resources of the highest use, which can be exploited.

Research Limitations

The information on token consumption and expenditure was founded on the past trends, which might not capture future changes in AI model's usage, system structure, and the optimization of resource technologies [28].

VII. FUTURE DIRECTION

The potential avenue of future research is to investigate the scalability of token-based data accounting systems to work in environments of various sectors and compare the findings of different sectors [29]. Moreover, the research into the combination of tokenization with newer AI designs, more energy-efficient blockchain solutions, and real-time optimizations tools would also increase the performance of the system and make it cost-effective.

VIII. INDUSTRY IMPACT

The suggested system allows the organizations to track AI training expenses correctly and enhance budgeting clarity and strategic decision-making processes in data-heavy operations. Using tokens to distribute costs in a proportional manner, enterprises can trace inefficiencies and optimize computational means, as well as prove the investment in the development of advanced AI projects [30]. Industries that operate with sensitive data enjoy increased traceability, regulatory adherence and auditable data administration ensured by blockchain accountability frameworks. The incentives could be token-based to motivate equitable compensation to the data contributors to ensure an increase in data quality, partnerships, and sustainable data-sharing environment. However, the framework promotes the scalable adoption of AI, efficiency, and innovation that are based on trust in the areas of healthcare, finances, and other controlled environments.

IX. CONCLUSION

The study shows the importance of using tokens to account for data in terms of increasing transparency and cost optimization when training AI models. The findings reveal an obvious dependence between the consumption of tokens and operational expenses, where there is a clear difference in the usage of resources in different departments. Improvements in the allocation of resources of departments can be identified with gaps in the field of R&D and IT. Another finding of the study is that model accuracy and resource consumption have efficiency trade-offs and therefore, it is important to optimize the costs. Nevertheless, there are some obstacles like system integration, scalability and energy usage.

X. REFERENCES

[1] PATCHIPALA, S., 2023. Data Anonymisation in AI and ML Engineering: Balancing Privacy and Model Performance Using Presidio. *IRE J*, *7*(6), pp.1-10.

[2] Khan, A.A., Laghari, A.A., Li, P., Dootio, M.A. and Karim, S., 2023. The collaborative role of blockchain, artificial intelligence, and the industrial Internet of things in the digitalisation of small and medium-sized enterprises. *Scientific Reports*, *13*(1), p.1656.

[3] Fusco, N., Sils, B., Graff, J.S., Kistler, K. and Ruiz, K., 2023. Cost-sharing and adherence, clinical outcomes, health care utilisation, and costs: a systematic literature review. *Journal of Managed Care & Speciality Pharmacy*, *29*(1), pp.4-16.

[4] Ranta, M., Ylinen, M. and Järvenpää, M., 2023. Machine learning in management accounting research: Literature review and pathways for the future. *European Accounting Review*, *32*(3), pp.607-636.

[5] Harish, A.R., Wu, W., Li, M. and Huang, G.Q., 2023. Blockchain-enabled digital asset tokenisation for crowdsensing in environmental, social, and governance disclosure. *Computers & Industrial Engineering*, *185*, p.109664.

[6] Srinivasa Kalyan Immadi, "Harnessing Artificial Intelligence In Oracle Hcm: Revolutionising Workforce Management With Automation And Predictive Analytics," International Journal of Data Science and IoT Management System, vol. 4, no. 4, pp. 7–13, Oct. 2025, doi: 10.64751/ijdim.2025.v4.n4.pp7-13.

[7] Gagliardelli, L., Zecchini, L., Ferretti, L., Beneventano, D., Simonini, G., Bergamaschi, S., Orsini, M., Magnotta, L., Mescoli, E., Livaldi, A. and Gessa, N., 2023. A big data platform exploiting auditable tokenisation to promote good practices inside local energy communities. *Future Generation Computer Systems*, *141*, pp.595-610.

[8] Sai Maneesh Kumar Prodduturi, "Efficient Debugging Methods And Tools For Ios Applications Using Xcode," International Journal Of Data Science And Iot Management System, Vol. 4, No. 4, Pp. 1–6, Oct. 2025, Doi: 10.64751/Ijdim.2025.V4.N4.Pp1-6.

[9] Siva Sankar Das. (2025). Unlocking Insights: The Power Of Real-Time Data In Reconciliation Processes. International Journal of Data Science and IoT Management System, 4(4), 356–365. https://doi.org/10.64751/ijdim.2025.v4.n4.pp356-365.

[10] Kopp, A. and Orlovskyi, D., 2022. Towards the Tokenization of Business Process Models using the Blockchain Technology and Smart Contracts. CMIS, 3137, pp.274-287.

[11] Omar, I.A., Hasan, H.R., Jayaraman, R., Salah, K. and Omar, M., 2021. Implementing

decentralised auctions using blockchain smart contracts. *Technological Forecasting and Social Change*, *168*, p.120786.

[12] Rani, P., Sachan, R.K. and Kukreja, S., 2023. Academic payment tokenisation: an online payment system for academia utilising non-fungible tokens and permissionless blockchain. *Procedia Computer Science*, *230*, pp.347-356.

[13] S. R. Nelluri and K. M. Poluri, "Real-Time Marketing Optimization through Scalable Telemetry Data Engineering: A Framework for Enhanced Engagement and ROI," International Journal of Computer Trends and Technology, vol. 73, no. 1, pp. 26–31, Jan. 2025, doi: 10.14445/22312803/ijctt-v73i1p103.

[14] S. Maneesh Kumar Prodduturi, "Leveraging Big Data And Business Intelligence To Revolutionise Corporate Strategy," International Journal for Research Trends and Innovation, vol. 8, no. 7, 2023, doi: 10.56975/ijrti.v8i7.207667.

[15] Hakimi, L., Eynon, R. and Murphy, V.A., 2021. The ethics of using digital trace data in education: A thematic review of the research landscape. *Review of educational research*, *91*(5), pp.671-717.

[16] Todupunuri, A. 2023. The Role of Artificial Intelligence in Enhancing Cybersecurity Measures in Online Banking Using AI. *International Journal of Enhanced Research in Management & Computer Applications*, 12(1), pp.103–108.

[17] Kovvuri, V.K.R. 2023. Serverless Analytics for Multi-Objective Optimisation in Last-Mile Delivery: Balancing Cost, Environmental Impact, and Customer Satisfaction. International Journal of Creative Research Thoughts (IJCRT), 11(2), pp.614–620. ISSN 2320-2882.

[18] "Scalable Suspicious Activity Detection Using Teradata Parallel Analytics And Tableau Visual Exploration," Inaternational Journal of Communication Networks and Information Security, vol. 8, no. 3, Jun. 2025, doi: 10.48047/ijcnis.8.3.236.

[19] Banda, S. 2023. Enhancing client engagement through AI-driven real-time reporting and automated alerts. *International Journal of Enhanced Research in Science, Technology & Engineering (IJERSTE)*, 12(11), pp.111–113.

[20] Assiri, F., 2020. Methods for assessing, predicting, and improving data veracity: A survey. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, *9*(4), p.5.

[21] Kotte, G. 2023. Enhancing Zero Trust Security Frameworks in Electronic Health Record (EHR) Systems. *SSRN Electronic Journal*, Paper No. 5283668.

[22] Yao, Y., Ghai, T., Ravi, S. and Szekely, P., 2021, October. Amppere: A universal abstract machine for privacy-preserving entity resolution evaluation. In Proceedings of the 30th ACM International Conference on Information & Knowledge Management (pp. 2394-2403).

[23] Ra, G., Kim, D., Seo, D. and Lee, I., 2021. A federated framework for fine-grained cloud access control for intelligent big data analytics by service providers. *IEEE Access*, *9*, pp.47084-47095.

[24] Naga Charan Nandigama, "A Data Engineering And Data Science Approach To Strengthening Cloud Security Through Ml-Based Mfa And Dynamic Cryptography," American Journal of AI Cyber Computing Management, vol. 5, no. 4(2), pp. 76–81, Nov. 2025, doi: 10.64751/ajaccm.2025.v5.n4(2).pp76-81.

[25] Colombo, G. and Giulivo, R., 2021. Financing and stakeholders engagement in infrastructure projects through blockchain tokens: facts or fantasy?.

[26] Das, S.S. 2020.Optimising Employee Performance through Data-Driven Management Practices. *European Journal of Advances in Engineering and Technology (EJAET)*, 7(1), pp.76–81.

[27] Zhao, L., Sen Gupta, S., Khan, A. and Luo, R., 2021, April. Temporal analysis of the entire Ethereum blockchain network. In Proceedings of the Web Conference 2021 (pp. 2258-2269).

[28] Yadav, S.K., Sharma, K., Kumar, C. and Arora, A., 2022. Blockchain-based synergistic solution to current cybersecurity frameworks. Multimedia Tools and Applications, 81(25), pp.36623-36644.

[29] Paruchuri, V.B. 2021. Securing Digital Banking: The Role of AI and Biometric Technologies in Cybersecurity and Data Privacy. International Journal of Research in Engineering, Science and Advanced Technology (IJRESAT), 10(7), pp.128–133. ISSN 2456–5083.

[30] Habib, G., Sharma, S., Ibrahim, S., Ahmad, I., Qureshi, S. and Ishfaq, M., 2022. Blockchain technology: benefits, challenges, applications, and integration of blockchain technology with cloud computing. *Future Internet*, *14*(11), p.341.