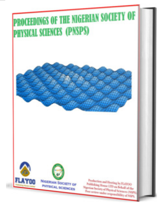
Published by Nigerian Society of Physical Sciences. Hosted by FLAYOO Publishing House LTD

Check for updates

Proceedings of the Nigerian Society of Physical Sciences

Journal Homepage: https://flayoophl.com/journals/index.php/pnspsc



Multisectoral data lakehouse for climate-resilient agriculture and sustainable food security in Nigeria

Adam Omeiza Yusuf*, Taiwo Kolajo, Emeka Ogbuju, Francisca Oladipo

Department of Computer Science, Faculty of Computing, Federal University, Lokoja, Nigeria

ABSTRACT

Nigeria, the most populous country in sub-Saharan Africa, faces multifaceted, overlapping problems of economic instability, security concerns, and an acute energy deficit, further aggravated by the ever-increasing impacts of climate change on its agricultural systems, a primary source of livelihood for millions. These hydra-headed challenges require a scientific, data-driven, informed response. A Multi-Layer Perceptron-based Deep Neural Network (MLP-DNN) prediction model for household-level food insecurity in Nigeria using a socioeconomic dataset was developed. This paper examines how principles from the physical sciences, particularly in multi-sectoral data management, modelling, analytics, and systems integration, can be applied to promote efficient energy resource distribution for climate-resilient agriculture (CRA), thereby boosting the economy and ensuring long-term food security. A mixed-methods approach, combining literature reviews, AI tools, case studies of climate adaptation projects, ontology-driven data harmonisation, and stakeholder consultations, was used in the research. Findings revealed that an integrated National Climate–Water-Agriculture Data Lakehouse can significantly enhance climate-resilient agriculture and energy resource sharing, thereby improving the economy and food security. The initiative becomes a scientific laboratory and a decision-support ecosystem, transforming siloed databases into actionable insights that policymakers, researchers, and practitioners can use to enhance the economy. The paper concludes that a multisectoral ontology-driven Data Lakehouse transforms physical science data into practical tools for climate-resilient agriculture by connecting laboratories to farmlands and datasets to decision-making, thereby enabling evidence-based agricultural governance and smart-agriculture to increase food production and boost the economies of the major and small-holding farmers in the face of emerging climatic chaos compounded by security concerns.

Keywords: Precision Agriculture (PA), Data lakehouse, Energy resource distribution, Ontology-driven data integration, Semantic machine learning.

DOI:10.61298/pnspsc.2026.3.287

1. INTRODUCTION

Nigeria, the most populous country in Sub-Saharan Africa, grapples with numerous challenges, including economic instability, security issues, and severe energy shortages worsened by climate change and poor agricultural data management. Food security re-

*Corresponding Author
e-mail: adam.yusuf-phd@fulokoja.edu.ng (Adam Omeiza Yusuf)

mains one of the pressing human issues affecting many people's livelihoods. According to the United Nations, food security is a situation in which everyone has access, at all times, to enough nutritious and affordable food for an active, healthy life [1]. But such perfect world does not actually exist. An estimated 2.33 billion people were living with moderate to severe food insecurity around the world in 2024, as livelihoods and food systems are increasingly squeezed by climate change, geopolitical tension and ongoing economic volatility. [2]

A combination of climate change impacts, economic volatility, and infrastructure inadequacies is worsening food security in Nigeria. Recent empirical evidence indicates that increasing temperatures, fluctuating rainfall and extreme weather conditions have severely disrupted one of the backbone economies of the country, the agricultural productivity [3]. But in the face of actual scientific data, there is still an enormous rift between the generation and use of such data in agricultural governance.

Artificial intelligence (AI) elements like machine learning, deep learning and the ability to process historical and real-time data using today's formidable computing power means AI has become capable of accurately predicting what is coming next with suggestions in some cases. Machine Learning and Deep Learning models are trained on large datasets. The model then encodes finer patterns and decreases overfitting, which brings more approximately better generalisation and predictions/planning. Jameel Francis, a Forbes Councils Member, cited Gartner, saying that poor data quality and relevance cause 85% of all AI models to fail [4]. Predictions of AI work as well as the pertinence and quality of the data on which such decisions are made. A "fragmented" data picture exists across the MDAs as it stands. The climatic data, as observed and recorded for International Water Management Institute (IWMI) study area is the responsibility of Nigerian Meteorological Agency under the Federal Ministry of Aviation [5]; hydrological data on flood forecast and predictions are domiciled with the Nigeria Hydrological Services Agency (NIHSA) while Dam water level to facilitate dry season irrigation farming is domiciled in River Basin Development Authorities - all under the Federal Ministry of Water Resources and Sanitation (FMWRS). The geospatial data on vegetation and arable land from earth observation stations and meteorological sensors for crop health assessment are domiciled with the National Space Research and Development Agency (NASRDA), which is under the Federal Ministry of Science and Technology [6]. Despite these agencies producing and housing reams of essential agriculture-related data, from when the rains start to when it stops, weather patterns and satellite imagery to flood predictions, drought-resistant crops to pest and disease, there's no neat way for these "data silos" to talk so that you can model complex agricultural systems as a whole. This unique challenge creates an obstacle for us to grapple with, as we try to solve issues of food security and prosperity for millions of Nigerians who ply their trade in farming through informed decision-making.

To mitigate the evidence of scattered data, an ontology-based Multisectoral Data Lakehouse for Climate-Resilient Agriculture and Sustainable Food Security in Nigeria is proposed. This new architectural model is an on-balance approach between the flexibility of data lakes and the data management features of data warehouses [7]. This system seeks to ingest data from 40+ departments and agencies under the four Federal Ministries for Machine Learning (ML) models, which can drive climate-smart agriculture and predict food security outcomes in a more accurate fashion. The multisectoral ontology-driven National Climate-Water-Agriculture Data Lakehouse transforms physical science data into practical tools for climate-resilient agriculture by connecting laboratories to farmlands and datasets to decision-making, thereby enabling evidence-based agricultural governance and smart-agriculture to increase food production and boost the economies of the major and small-holding farmers in the face of emerging climatic chaos compounded by security concerns.

2. RELATED WORK ON MULTI-SECTOR DATA INTEGRATION

Climate variability is a major underlying cause of food insecurity in Nigeria, affecting the economies of big and smallholder farmers. Efficient data governance is key to informed decision-making. Various researchers have proposed various approaches to making data available to decision makers.

The NASA Hydrological Forecast and Analysis System (NHyFAS) was reviewed, which provides forecasts of hydrological and agricultural drought [8]. The system is designed to provide multi-model seasonal forecasts that align with predictive analytics for food security. We are improving this by using localised data for national food security predictions, as opposed to using NASA data.

The use of AI-driven yield forecasting in European hydroponic systems and resource optimisation in Southeast Asian aquaponics was proposed [9], showcasing localised efficiency gains. The study adopts a mixed-methods approach, including systematic literature analysis and regional case studies and concludes by outlining future directions, emphasising context-specific AI implementations, the need for public–private collaboration, and policy interventions to enhance scalability and adoption in food security contexts. The conclusion of the paper aligns directly with our design of a multisectoral semantic-based Data Lakehouse for sustainable food security in Nigeria, as it provides for easy access to curated food security data.

The development of an ontology for Climate Smart Agriculture (CSA) [10] was proposed, which aims to provide a formal representation of relevant information and solutions, automating tasks and facilitating access to practices that mitigate climate change effects. The CSA ontology adds the element of automation to a task that would otherwise be daunting and extremely labour-intensive. It focuses on CSA practices and techniques that will help reduce the effects of climate change. The paper aligns with our concept of a Semantic-based Data Lakehouse that would organise and make accessible relevant data for sustainable food security.

An ontology development was proposed, a formal representation of CSA information and solutions [10]. It highlights the role of the ontology in automating tasks and providing quick access to suitable practices for climate change mitigation, aligning with the concept of a data Lakehouse that would organise and make relevant data accessible for sustainable food security planning and prediction.

Developing an ontology for Climate Smart Agriculture (CSA) to fill a gap in existing research by creating a knowledge base ontology for CSA [10]. The methodology involves gathering information from secondary sources and using Description Logics and Web Ontology Language to formalise the ontology, thereby indicating a structured approach to ontology development and resulting in a machine-readable ontology that can be used for data storage and sharing. The result aligns with our goal of designing a Semantic-Based climate-resilient-agriculture Data Lakehouse for sustainable food security and prediction.

The importance of sustainable agriculture and the role of data integration in supporting it [11]. The paper reviews current semantic web technology applications, the construction and maintenance of ontologies in the agricultural domain, while focusing on integrating heterogeneous data using semantic technologies and linked data, aligning with the design of a semantic-based data Lakehouse knowledge base that can support farmers and researchers. The development of a semantic data management platform using ontology models for automotive data lakes [12]. The concept of designing an ontology-driven Data Lake involves using ontology models for data management. Developing a knowledge graph (BDAKG) to integrate agricultural data in Bangladesh, [13] focusing on data integration and making it more accessible and interoperable, which aligns with the concept of a semantic-based Data Lakehouse.

An ontology development for structuring digital databases in the context of grain production was proposed [14]. The ontology is designed to operate in a cloud environment and uses Data Lake storage solutions and semantic treatment of big data. The objective is focused on decision-making and the integration of heterogeneous data, which aligns with the goal of a Semantic-based Data Lakehouse for sustainable food security.

The challenges faced by modern agriculture due to climate change and resource constraints was discussed [15] and further highlight the need for data innovation to improve productivity and sustainability. The paper further identifies the challenges posed by data heterogeneity and the need for a new data management infrastructure that follows FAIR principles, which aligns with our design of a Multisectoral Data Lakehouse for climate-resilient agriculture and sustainable food security.

Developing a spatio-temporal semantic data model for precision agriculture was proposed [16], which involves managing large volumes of sensor and device data. They validated the model using real-world data from livestock and crop scenarios, indicating practical applications of integrated multisectoral data. The use of InfluxDB for evaluation suggests a focus on efficient data management, which is one of the benefits of the Multisectoral Data Lakehouse.

The NiMet historical data (1993 - 2023) Climate view report has shown warming trends and temporal variability in rainfall, resulting in decreased yields of rice and groundnut [17]. In addition, the use of Random Forest models on historical NiMet data (1988 - 2018) has shown that climate change for planning purposes can be accurately predicted using machine learning [18]. But climate data without a hydrologic context is not enough. NIHSA's flood information is crucial for forecasting disasters in flood-prone agrarian areas, but it is rarely considered alongside real-time crop performance records.

The applications of remote sensing to precision agriculture are one example of new frontiers in agricultural data governance. We have also adopted an enhancement index on the top public register, based on satellite imagery data obtained from NASRDA, as a foundational layer for both land-use change, soil moisture, and crop health analysis. Coupled with farm profile data in the FMAFS, specific locations at risk of yield failure can be identified from this geospatial data. The difficulty here again arises in the semantic mismatch between the satellite-pixel data and the organised record structure of government ministries in RDBMS or Data Warehouse. Similarly, NiMet temperature and humidity data are directly associated with the presence of certain pests and plant diseases that impinge on crop yield. However, these records are not related. Strategic plan to ramp up food production for sustainable food security development needs to pool all these data together as input for deep learning to guide appropriate planning, monitoring and predictions toward targeted results.

An experiment with large volume of datasets was conducted and compared with other test with single dataset for prediction [19]. The result revealed that large number of observations and high-dimensional agricultural data are the prerequisites for the effective application of Deep Learning in Precision Agriculture. This highlights the necessity of integrating multiple-sector food security datasets to improve planning and forecasting.

The six dimensions of data quality in the AI era are accuracy, completeness, consistency, timeliness, relevance and validity. If an organisation neglects the quality of data, it will have very serious and multi-channel effects. It has been proven that poor-quality data results in quantifiable damage to financial value, decision making, operational performance, reputation and missing strategic possibilities [19]. The Space industry has, in recent past, experienced one of the most dramatic impacts of poor-quality data. NASA's Mars Climate Orbiter was obliterated because a piece of software that engineers had written to measure thrust forces, which were measured in pounds, but the push-back generated by the oven-cleaning fan was recorded as metric, which resulted in a $125 million failure of consistency. It is a classic example of how poor data validation and ineffective governance can result in disastrous losses. The failure highlights the fact that data quality is not just a technical issue; it has become an organisationally significant problem and one that needs strong verification through multiple levels [20]. Now, data needs to be shaped not just for human consumption but also for machine consumption and the training of machine learning models. This double-duty nature of data quality goes beyond the traditional dimensions.

Beyond physical sciences, data on the socio-economic indicators (market prices of staple foods and frequency of conflicts in a region), monitored by NBS and NGOs, are essential for a comprehensive estimation of food security. In addition, food safety data, which are often neglected, also have a significant impact: in recent studies, Neural Networks were effective in the prediction of food safety and contamination risks, obtaining good accuracy [21]. It is the establishment of such a generative model that necessitates a strong architectural framework for the integration of these heterogeneous datasets.

Relational Database Management Systems (RDBMS) are the backbone of data management in agriculture. Online Transaction processing (OLTP) type RDBMS enforces a rigid schema and the Atomicity, Consistency, Isolation, Durability (ACID) transaction property for their persistence to guarantee that the data is correct, consistent, and recoverable. The Federal Ministry of Agriculture and Food Security (FMAFS), among others, in the Nigerian food security ecosystem, uses an RDBMS to manage structured records; this includes farmer registries, subsidy beneficiaries, fertiliser distribution logging, etc. Their core competency is administrative semantics: transaction records are accurate, auditable, and failure-proof.

For food security prediction and early warning, however, conventional RDBMS hit their limits easily. Predictive analytics relies on combining data across agencies and fields, but public-sector data governance is typically siloed. Siloed database was characterised as a radical mess [22], reflecting that many institutions gather disjoint and incomparable data sources, leading to persistent silos of data. Because of this, important indicators like spikes in local food prices or displacement due to conflict are stuck in siloed systems and are impossible for policymakers to analyse together. RDBMS are great at handling with transactional workloads, but they do not deal well with the type of heterogeneous cross-sector data to successfully model complex food security crises prediction.

The problem of Unstructured Data has also been an issue facing data warehouses since their inception. In order to provide a unified access point to that integrated, historical data located in various transactional systems and made available for analysis - Data Warehouses (DW) were proposed. From an architectural point of view, DWs are based on a schema-on-write paradigm that implies data cleaning, transformation and structuring before insertion. They have been widely applied in agriculture to assess historical yields, production patterns, and the impacts of policies over long-term periods [23]. Their biggest merit is that they help build a single source of truth that helps maintain consistent reporting and long-term planning across government departments.

However, classical data warehouses do not scale and are too slow to handle the variety of food security data. Modern forecasting is more and more being characterised by the usage of unstructured, high-velocity data such as satellite images for drought monitoring or real-time hydrological sensor streams for detecting floods [24]. These types of data are hard to fit into strict warehouse schemas without lots of preprocessing. By the time unstructured data is shaped into tabular formats that any DW can receive, most of its real-time value is already lost, and rapid crisis response or dynamic prediction becomes impossible.

Because of the data warehouse's disadvantages, we have the Data Lake to the rescue, in which huge volumes of raw data can be stored using its native formats based on schema-on-read. This flexibility is what's particularly alluring in agricultural and climate research, where organisations such as NiMet, NASRDA produce loads of data from climatic records to high-resolution pictures taken from satellites. Data Lakes defer any data modelling, allowing the raw features to be stored at low cost and can be inspected granularly before putting it into a desired system (e.g., machine learning or advanced analytics) [7].

And yet that flexibility doesn't come cheap. In the absence of good governance, metadata or semantic structure, expect these to degrade into "data swamps" filled with poorly documented "garbage long tail" of disconnected datasets. It was pointed out that, in the realm of climate adaptation planning, this piecemeal and uncurated information is more likely to be a barrier than a facilitator for decision-making. For food security prediction (e.g., when defining terms like drought or crop failure, it is crucial that there are clear and correct definitions), the lack of semantic constraints may lead to unreliable models and spurious insights.

3. THE PROPOSED SEMANTIC-DRIVEN MULTISECTORAL DATA LAKEHOUSE

Our proposed solution, the Multisectoral Data Lakehouse, combines the flexibility and scalability of a data lake with the structured performance of a data warehouse, enabling both raw data storage and optimised analytical querying within a unified architecture. Unlike legacy data warehouses that rely on rigid predefined schemas and ETL processes, which often compromise data integrity by overwriting or discarding historical values, and unlike traditional data lakes that lack efficient support for structured queries, the Data Lakehouse preserves raw data fidelity while delivering high-performance, schema-aware analytics.

We use the term Lakehouse to refer to a new data management paradigm that combines the best of both architectures: the simple and scalable architecture of a Data Lake, and the ACID (Atomicity, Consistency, Isolation, Durability) transactions, fine-grained security, and performance optimisations of an analytical Data Warehouse. Lakehouse, therefore, allows the advantages offered by data lakes and data warehouses to be combined with access to low-cost storage in open formats using a range of systems from the former, alongside rich management and optimisation interfaces from the latter [7].

For unifying the heterogeneous data coming from NiMet, NASRDA, NIHSA, RBDA and other NGOs and so on, a semantic harmonisation and an abstract representation of knowledge concepts along with their inter-relationship is employed by the system [25]. For example, an ontology can semantically link the metric "precipitation" of NiMet to FMAFS "Rainfall", NIHSA's "river discharge" levels and Dam "water volume" from RBDA, producing an integrated knowledge graph. This semantic layer allows the system to "interpret" that a decrease in rainfall at a given coordinate (NASRDA) is associated with a risk of drought on the farm profile, say (FMAFS), while it also interprets that an increase or reduction in temperature or humidity reading value (from NiMet), are expected to account for incidence of certain crop pests and disease considered harmful to crop yield and food security outcome.

The curated data in the Lakehouse is the advanced analytic training ground for:

1. Predictive Modelling: Machine Learning models trained with curated multisectoral historical climate, water and yield data can predict food availability months in advance more accurately.
2. Risk Assessment: ML classifiers can help classify local government areas (LGAs) according to the risk level of experiencing food insecurity through analytics that use real-time inputs, including flood alerts and fluctuations in market prices.
3. Forecasting: Artificial intelligence-assisted models can be extended to forecasting safety outbreaks in food, guaranteeing that food loss is not only quantity-based but also quality-based [21]. Such prediction allow decision makers to take

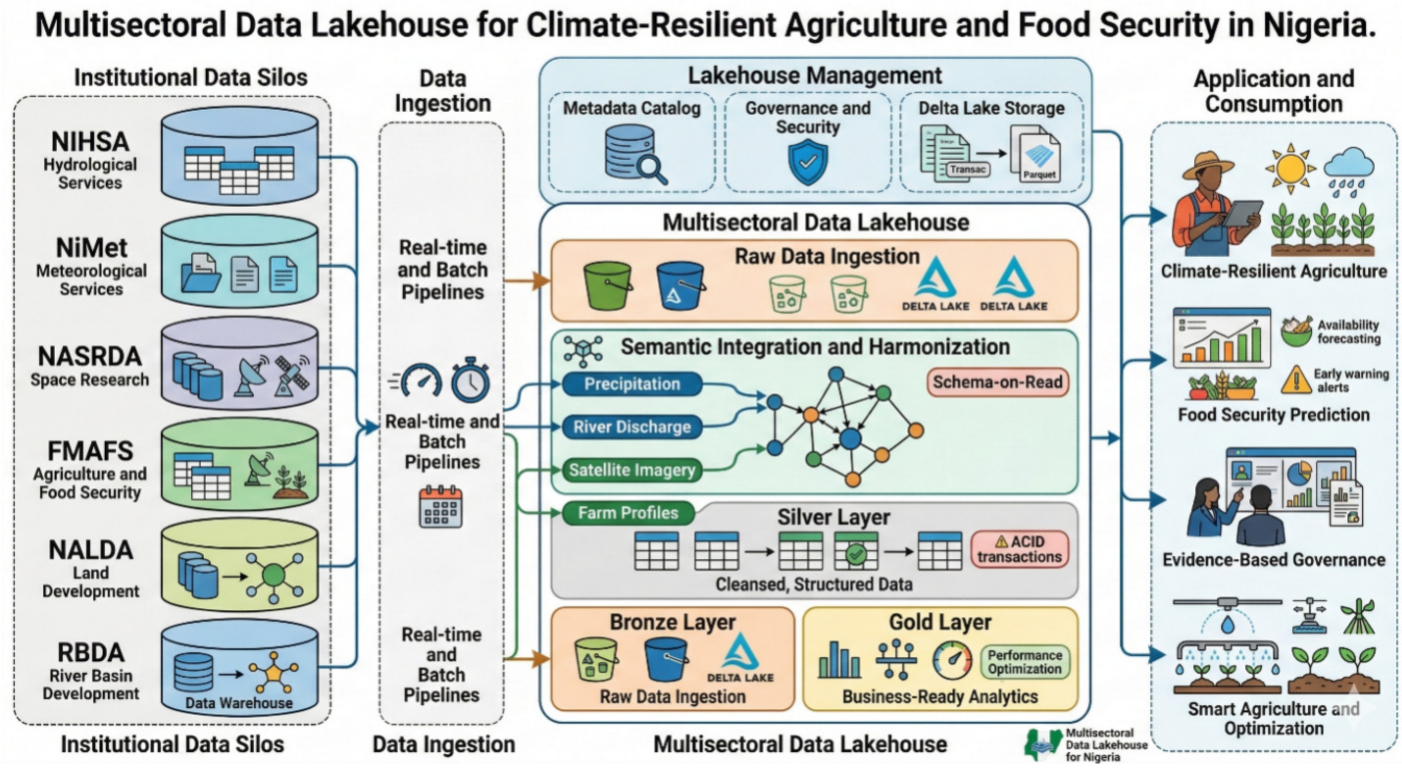


Figure 1. Schematic diagram of the multisectoral data lakehouse.

proactive measures to prevent or mitigate the negative consequences associate with the coming anomalies.

4. Cropping Calendar Planning: Data from NiMet on the onset and cessation of rains provides guidance for the farmers whose activities are weather-dependent, to ensure maximum yield. Temperature from NiMeT and surface water inventory records from NASRDA will assist the irrigation planner in ascertaining the command area adequate for the optimum utilisation of available water in the Dams or reservoirs, as a factor in food security.

4. EXPECTED RESULT

Transitioning to a Data Lakehouse architecture represents a strategic updating of Nigeria's agricultural data governance infrastructure. Centralising access to data from more than 40 agencies and departments, the Multisectoral Semantic-Based Data Lakehouse will eradicate duplication along with delays as a result of manual exchanges of data, as well as the vagaries of data swamps. The adoption of ontologies also has the advantage of scalability, as new agencies or NGOs come on board, the data can be more easily mapped into the existing knowledge graph without having to kick over the architecture.

Finally, the integration of socioeconomic data from NBS and NGOs brings in the human aspect of food security. Data on food prices or conflict-driven displacement can also be combined with drought data to offer precise policy suggestions. This is consistent with reports that highlight the importance of essential measures to reduce the impact of climate change and secure affordable, available food [3].

5. DISCUSSION

The results of this study mark a major shift from traditional agricultural prediction models. Historically, these models have relied on isolated, single-domain datasets. However, as previous research has repeatedly shown, relying on just one type of data naturally limits how accurate our predictions can be, especially when trying to understand something as complex and multi-layered as climate-induced food insecurity. Time and again, studies have pointed to data fragmentation as the main roadblock. When key national agencies like NiMet, NIHSA, and the FMAFS keep their data in separate silos, we lose the ability to harness the massive, high-velocity information they generate. This fragmentation has been the single biggest hurdle holding back the country's transition toward smart agriculture and effective, national early warning systems.

The National Climate–Water-Agriculture Data Lakehouse, when deployed, would offer a direct, practical solution to this deep-rooted problem. The semantic harmonisation process introduced right at the data ingestion stage by our framework actively breaks down these institutional walls. The semantic layer acts as a universal translator, taking disparate and previously incompatible datasets and seamlessly weaving them together into a clear, machine-readable knowledge graph. As a result, this framework does more than just bypass the issue of fragmented data; it creates the deep interoperability necessary for highly accurate, high-fidelity AI predictions. Ultimately, this integrated approach provides the solid, multi-agency foundation required to turn raw data into actionable smart agriculture, helping to sustainably secure the nation's food supply.

6. CONCLUSION

This paper presents a pathway to modernising agricultural decision-making in Nigeria, based on the Multisectoral Semantic-Based Data Lakehouse for climate-resilient agriculture and sustainable food security, by combining climate information from NiMet, satellite data from NASRDA, flood survey results of NIHSA, water availability/irrigation command area from RBDAs, and farm profile data by FMAFS populated into a semantic ecosystem in Nigeria. This curated data repository stored in the semantic-based Data Lakehouse can be leveraged for agricultural planning purposes and to train ML models to mitigate potential threats to food security. This scientific process turns raw, isolated data silos into actionable intelligence, enabling a climate-smart agricultural sector to feed the growing nation. The Semantic-based Data Lakehouse transforms physical science data into practical tools for climate-resilient agriculture by connecting laboratories to farmlands and datasets to decision-making, thereby enabling evidence-based agricultural governance and smart agriculture to increase food production and boost the economies of major and small-holding farmers in the face of emerging climatic chaos compounded by security concerns. This aligns with the theme of this international conference, "From Lab to Live".

DATA AVAILABILITY

We do not have any research data outside the submitted manuscript file.

ACKNOWLEDGMENT

I wish to acknowledge the valuable contributions from my able supervisors and conferees at ZAZZAU2026.

References

[1] Fao, "The state of food security and nutrition in the world 2022", Food and Agriculture Organisation of the United Nations (2022). https://doi.org/10.4060/cc0639en.
[2] Nhsjs, "From data to action: using machine learning to combat food insecurity", National High School Journal of Science (2025). https://nhsjs.com/2025/from-data-to-action-using-machine-learning-to-combat-food-insecurity/.
[3] F. O. Oderinde, O. I. Akano, F. A. Adesina & A. O. Omotayo, "Trends in climate, socioeconomic indices and food security in Nigeria", Frontiers in Sustainable Food Systems **6** (2022). https://doi.org/10.3389/fsufs.2022.940858.
[4] J. Francis, "Why 85% of your AI models may fail", Forbes.com (2024). https://www.forbes.com/councils/forbestechcouncil/2024/11/15/why-85-of-your-ai-models-may-fail/.
[5] Nigerian Meteorological Agency (NIMET), "2025 seasonal climate prediction (SCP)" (2025). https://nimet.gov.ng/publication_detail?id=42.
[6] Nasrda, "National space research and development agency". https://central.nasrda.gov.ng/nasrda-set-to-revolutionize-agriculture-with-satellite-technology/.
[7] M. Armbrust, A. Ghodsi, R. Xin & M. Zaharia, "Lakehouse: a new generation of open platforms that unify data warehousing and advanced analytics", Proceedings of CIDR 2021 (2021). https://www.cidrdb.org/cidr2021/papers/cidr2021_paper17.pdf.
[8] K. R. Arsenault et al., "The NASA hydrological forecast system for food and water security applications", Bull. Am. Meteorol. Soc. **101** (2020) E1007. https://doi.org/10.1175/BAMS-D-18-0264.1.
[9] S. B. Dhal & D. Kar, "Transforming agricultural productivity with AI-driven forecasting: innovations in food security and supply chain optimization", Forecasting **6** (2024) 925. https://doi.org/10.3390/forecast6040046.
[10] N. Naidoo, S. Lawton, M. Ramnanan, J. V. Fonou-Dombeu & R. Gowda, "Modelling climate-smart agriculture with ontology", International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD) (2021). https://doi.org/10.1109/icABCD51485.2021.9519380.
[11] S. Jiang, R. Angarita, R. Chiky, S. Cormier & F. Rousseaux, "Towards the integration of agricultural data from heterogeneous sources: perspectives for the French agricultural context using semantic technologies", Lect. Notes Bus. Inf. Process. **382** (2020) 89. https://doi.org/10.1007/978-3-030-49165-9_8.
[12] S. Schmid, C. Henson & T. Tran, "Using knowledge graphs to search an enterprise data lake" (2019) 262. https://link.springer.com/chapter/10.1007/978-3-030-32327-1_46.
[13] R. Pratap Deb Nath, T. Rani Das, T. Chandro Das & S. M. Shafkat Raihan, "Knowledge graph generation and enabling multidimensional analytics on bangladesh agricultural data", IEEE Access **12** (2024) 87512. https://doi.org/10.1109/ACCESS.2024.3416388.
[14] R. A. Neves & P. E. Cruvinel, "Ontology for structuring a digital databases for decision making in grain production", in 2021 IEEE 15th International Conference on Semantic Computing (ICSC), IEEE (2021) 386. https://doi.org/10.1109/ICSC50631.2021.00071.
[15] Y. Pan et al., "Transforming agriculture with intelligent data management and insights", in 2023 IEEE International Conference on Big Data (BigData), IEEE (2023) 3489. https://doi.org/10.1109/BigData59044.2023.10386589.
[16] M. San Emeterio de la Parte, J.-F. Martínez-Ortega, P. Castillejo & N. Lucas-Martínez, "Spatio-temporal semantic data management systems for IoT in agriculture 5.0: challenges and future directions", Internet of Things **25** (2024) 101030. https://doi.org/10.1016/j.iot.2023.101030.
[17] N. N. Ubachukwu, T. N. Nzelibe, I. U. Nzelibe & O. O. John, "Climate change adaptive strategies for enhancing food security in Biu-Nigeria", **8** (2025) 2657. https://doi.org/10.48346/IMIST.PRSM/ajlp-gs.v8i10.59550.
[18] C. O. Egbunu, M. T. Ogedengbe, T. S. Yange, M. A. Rufai & H. I. Muhammed, "Towards food security: the prediction of climatic factors in Nigeria using random forest approach", Journal of Computer Science and Information Technology **7** (2021) 70. https://doi.org/10.35134/jcsitech.v7i4.15.
[19] M. Guillen-Aguinaga, E. Aguinaga-Ontoso, L. Guillen-Aguinaga, F. Guillen-Grima & I. Aguinaga-Ontoso, "Data quality in the age of AI: a review of governance, ethics, and the FAIR principles", Group **2882** (2025) 930X. https://doi.org/10.3390/data10120201.
[20] N. Davidson, "The cost of poor data quality on business operations" (2026). https://lakefs.io/blog/poor-data-quality-business-costs.
[21] A. O. Kehinde, M. A. Onafowokan & O. O. Onalaja, "Leveraging machine learning techniques for the prediction and enhancement of food safety standards in Nigeria", FUDMA Journal of Sciences **9** (2025) 130. https://doi.org/10.33003/fjs-2025-0904-3562.
[22] M. F. E. Anupam & S. A. Mustafa, "Integrated data governance and predictive analytics for enhancing public service management in bangladesh: a framework for evidence-based decision-making", Bangladesh Journal of Administration and Management **38** (2025) 1. https://doi.org/10.56379/bjam.V38i1.100.
[23] A. A. Harby & F. Zulkernine, "From data warehouse to lakehouse: a comparative study", 2022 IEEE International Conference on Big Data (Big Data) (2022) 389. https://doi.org/10.1109/BigData55660.2022.10020719.
[24] A. A. Effiong, B. E. Okon & I. J. Udo, "Smart irrigation in Nigeria: a panacea for curbing hunger and food insecurity", Asian Journal of Advanced Research and Reports **19** (2025) 53. https://doi.org/10.9734/ajarr/2025/v19i61041.
[25] M. Horvat, A. Krtalić, A. Akagić & I. Mekterović, "Ontology-based data observatory for formal knowledge representation of UXO using advanced semantic web technologies", Electronics **13** (2024) 814. https://doi.org/10.3390/electronics13050814.