OCTOBER 28, 2025

The Power of Connection: A Subscriber's Guide to Data Integrity and Unified Identity with CreditRegistry

THE POWER OF CONNECTION: A SUBSCRIBER'S GUIDE TO DATA INTEGRITY AND UNIFIED IDENTITY WITH CREDITREGISTRY

CREDITREGISTRY
support.@creditregistry.ng
www.creditregistry.ng

Table of Contents

Executive Summary

The Nigerian credit ecosystem is a dynamic and interconnected network, with CreditRegistry at its heart, facilitating trust and enabling informed risk management decisions. The integrity of this ecosystem hinges on the quality and consistency of the data that flows through it. This whitepaper serves as a definitive guide for our valued subscribers to demystify CreditRegistry's data architecture, clarify the principles of entity resolution, and empower each institution to maximize the value derived from our services. It is built on the foundational principle of shared responsibility: the accuracy of the national credit reporting system is a direct reflection of the collective diligence of its participants.

This document addresses common challenges faced by subscribers, such as perceived data fragmentation and query inefficiencies. It reveals that the sophisticated, unified view of a borrower that subscribers seek is already a core component of our system, delivered through a powerful mechanism known as the CorrelationID. This CorrelationID functions as the "Master Entity ID," linking disparate records from across the subscriber network into a single, holistic profile.

By reading this whitepaper, subscribers will gain a comprehensive understanding of:

- Their critical role in maintaining data quality, starting with robust Know Your Customer (KYC) processes.
- The foundational importance of submitting a unique CustomerNumber for each entity.
- The architecture of CreditRegistry's identity framework, distinguishing between the RegistryID and the CorrelationID.
- Practical, step-by-step guidance on how to leverage this system to achieve efficient, accurate, and consolidated credit assessments with minimal effort.

Ultimately, this guide demonstrates that by embracing best practices in data submission and understanding the power of CreditRegistry's entity resolution engine, subscribers can resolve their current challenges, reduce operational friction, and contribute to a more transparent and robust credit market for all Nigerians.

Section 1: The Foundation of a Trusted Credit Ecosystem: A Shared Responsibility

A credit bureau is more than a passive repository of data; it is the central nervous system of a modern economy, facilitating the flow of credit by providing a trusted, impartial mechanism for risk assessment. CreditRegistry, licensed and regulated by the Central Bank of Nigeria (CBN), fulfills this role by aggregating borrower and repayment information from a diverse network of subscribers. This network includes commercial and microfinance banks, telecommunication companies, landlords, government lending institutions, and other businesses that extend credit. The collective data from these partners forms a national asset that, when accurate and comprehensive, lowers the cost of credit, expands financial inclusion, and fuels economic growth.

In a globalized economy, the risks associated with data fragmentation are significant, particularly in emerging markets. Geopolitical and economic pressures can lead to regulatory divergence and increased uncertainty, which in turn can reduce cross-border investment and raise the cost of doing business for all participants. A fragmented domestic credit information system poses similar risks on a national scale. It creates information asymmetry, where lenders operate with an incomplete picture of a borrower's obligations, leading to suboptimal lending decisions, increased defaults, and systemic risk. CreditRegistry's mission is to serve as a bulwark against this fragmentation by creating a unified, reliable source of credit information.

However, the success of this mission is not a unilateral effort. It is a shared responsibility. The accuracy of a credit report, the reliability of a SMARTScore, and the overall health of the credit ecosystem are direct reflections of the quality of data submitted by the entire network of subscribers. Inaccurate, duplicated, incomplete, or outdated data submitted by one institution can have a cascading negative effect, impacting the credit decisions of all other subscribers who access that borrower's file. When a subscriber encounters a fragmented or inconsistent view of a customer, it is often a symptom of a systemic challenge within the data-sharing ecosystem itself. The discrepancies—a slightly different address, a variation in name spelling—are frequently the result of other lenders submitting data based on their own, sometimes less rigorous, onboarding processes.

Therefore, this whitepaper elevates the perspective of our subscribers from that of transactional "data suppliers" to strategic, indispensable partners. Each subscriber is a steward of the data they manage, and their commitment to data quality is a commitment to the integrity and efficiency of the entire Nigerian credit market. By working collaboratively under a common set of standards, we can collectively build a more transparent, efficient, and trusted economic future.

Section 2: The Subscriber's Foundational Role: Data Integrity and KYC Compliance

The journey to high-quality credit data begins long before a file is submitted to CreditRegistry. It starts at the very first point of contact with a customer: the Know Your Customer (KYC) and Customer Due Diligence (CDD) process. These are not merely administrative tasks but are regulatory mandates from the Central Bank of Nigeria (CBN) that form the bedrock of financial integrity and risk management. A subscriber's adherence to these mandates directly correlates with the quality of data they contribute to the credit ecosystem and, consequently, the clarity and accuracy of the credit reports they retrieve.

Subsection 2.1: The Regulatory Mandate - Beyond Compliance, Towards Quality

The CBN's regulations on customer due diligence are explicit and comprehensive. Financial institutions are required to identify and verify the identity of all customers, whether they are individuals or corporate entities. This is not a one-time event at onboarding but an ongoing responsibility that requires periodic updates to customer records to ensure information remains current and accurate.

For **individual customers**, the CBN mandates the collection of key information, including but not limited to:

- Legal name and any other names used (e.g., maiden name)
- Permanent physical address
- Date and place of birth
- Bank Verification Number (BVN)
- Tax Identification Number (TIN)
- Nationality and occupation
- A valid, unexpired government-issued identification number (e.g., from a National ID card, International Passport, or Driver's License)

For **corporate customers**, the requirements are equally stringent and include:

- Certificate of Incorporation from the Corporate Affairs Commission (CAC)
- Tax Identification Number (TIN)
- Registered and business addresses
- Information on directors, signatories, and beneficial owners
- Board resolutions authorizing the business relationship

The recent CBN directive to place a "Post no Debit or Credit" freeze on all Tier-1 accounts lacking a BVN or National Identification Number (NIN) from April 2024 underscores the criticality of these national identifiers. This policy elevates the BVN and NIN as the primary, most reliable anchors for individual identity within the Nigerian financial system, reinforcing the logic behind CreditRegistry's validation of these fields in our data submission channels.

Subsection 2.2: From KYC Documents to Bureau Data - Bridging the Gap

The data quality issues that subscribers often perceive as problems originating from the credit bureau are, in fact, lagging indicators of deficiencies in their own internal KYC and data governance processes. The bureau's database acts as a diagnostic mirror, reflecting the collective data health of its contributors. When a subscriber encounters discrepancies, it is an opportunity to identify and rectify weaknesses in their source data collection and verification.

Consider the common pain points:

- **Name Variations:** A record showing "John Emeka OKEKE" from one lender and "Jon Okechukwu EMEKA" from another for the same BVN is not a bureau error. It indicates that at least one of the lenders failed to properly verify the customer's full legal name against a primary government-issued ID during onboarding.
- **Incorrect Dates of Birth:** This issue arises when a customer's stated date of birth is not validated against an official document or is not properly entered into the lender's CRM or banking software system from which data provided to CreditRegistry is extracted. This seemingly small error can have significant impacts on identity matching and credit scoring algorithms.
- **Multiple Addresses:** A customer having three different addresses on file across three lenders is a direct result of inconsistent address verification practices. One lender may have used a utility bill, another the address from the BVN record, and a third may not have updated the record after the customer moved. While this presents a reconciliation challenge, it also provides a valuable risk signal. An individual with a transient data footprint may represent a different risk profile than one with a stable, consistent record across all relationships.

Therefore, the challenges of data reconciliation are not a failure of the bureau's system but a reflection of varying data quality standards across the ecosystem. By interpreting these discrepancies as actionable intelligence, subscribers can improve their internal processes. This proactive approach—treating the bureau's output as a diagnostic tool for their own KYC health—is the most effective way to reduce data-related friction and improve the accuracy of their risk assessments.

Section 3: The Keystone of Data Submission: The Unique CustomerNumber

At the core of the relationship between a subscriber and CreditRegistry is a single, critical data field: the CustomerNumber. This identifier, referred to as CustomerID in the Common Data Template (CDT) documentation, is the keystone that anchors a subscriber's entire data submission process and ensures the integrity of their records within the bureau's database. Understanding and correctly implementing the CustomerNumber is the single most important step a subscriber can take to guarantee data provenance and prevent the very issues of data overwriting and fragmentation that cause concern.

The CustomerNumber is defined as the subscriber's unique internal identifier for a specific customer. Its primary function is to allow CreditRegistry to issue and maintain a stable, unique RegistryID for every distinct subscriber-customer relationship. This architectural choice is deliberate and crucial. It prevents the "nightmarish situation" described by subscribers where incoming data from one institution could potentially overwrite or corrupt the valid data submitted by another. By creating a unique RegistryID for each CustomerNumber, CreditRegistry preserves the integrity of every subscriber's sovereign view of their customer.

This approach effectively creates a "namespace" for each subscriber. Even if two different institutions report on the same individual (identified by the same BVN), their records are initially treated as distinct and separate entities within the bureau. For example, if Subscriber A reports on a customer with their internal ID 'CUST-123', and Subscriber B reports on the same person with their internal ID 'ACCT-XYZ', CreditRegistry will create two separate RegistryIDs. This prevents a scenario where a data submission from Subscriber B, perhaps with a different spelling of the customer's middle name, could alter the record that Subscriber A relies upon. This preservation of the original, unaltered record from each source is a cornerstone of robust data governance and is essential for dispute resolution and auditing.

Recognizing that not all subscribers have sophisticated IT systems for generating unique internal IDs, CreditRegistry provides clear and unambiguous guidance:

- **For subscribers without unique internal identifiers, it is recommended to use the customer's BVN (for individuals) or their Corporate Affairs Commission (CAC) Registration Number (RC Number) or TaxID (for businesses) in the CustomerNumber field.**

This practical instruction ensures that every subscriber, regardless of their technical capacity, can meet this fundamental requirement for data submission.

Data can be submitted to CreditRegistry through two primary channels:

1. **The AutoCred8 Data Submission API[i]:** This is the modern, recommended method. It enforces data quality at the point of entry by performing real-time, mandatory validation of key fields. For individual records, the BVN is a mandatory and validated field. For business records, at least one of the Registration Number or TaxID must be provided and validated.
2. **Partner Connect Portal [ii]:** Subscribers can also upload data files using the Common Data Template (CDT) format via the secure CreditRegistry Partner Connect Portal. For historical reasons, dating back to a period before the BVN was widely adopted, mandatory verification of these key national identifiers is not enforced on this channel.

To ensure the highest level of data integrity and to prevent submission errors, CreditRegistry strongly encourages all subscribers, particularly new ones, to utilize the AutoCred8 API *(or later)* for their data reporting. This channel provides immediate feedback and helps maintain the quality of the data from the outset.

Section 4: Unlocking a Unified View: Understanding CreditRegistry's Entity Resolution Engine

A common and valid concern among subscribers is the need for a single, consolidated view of a borrower's credit history. Many have expressed a desire for an "Entity Resolution system" that can de-duplicate records and provide a "canonical Master Entity ID" that maps to all underlying records for a single legal entity. This is an excellent and precise description of the core architecture that CreditRegistry has successfully operated for years. The "Master Entity ID" that subscribers propose is precisely what our system refers to as the CorrelationID.

Subsection 4.1: The Subscriber's Proposal and Our Implemented Solution

Entity Resolution is a globally recognized best practice in data management and data science. It involves the complex process of identifying, linking, and merging records that correspond to the same real-world entity across different data sources. In financial services, this technology is indispensable for creating 360-degree customer views, detecting sophisticated fraud rings, and ensuring regulatory compliance with standards like KYC and Anti-Money Laundering (AML).

CreditRegistry's platform is built upon a sophisticated entity resolution engine designed specifically to address the challenges of the Nigerian market. The system is engineered to resolve discrepancies and link records, providing the very unified view that subscribers require. The key to unlocking this capability lies in understanding the two fundamental components of our identity framework: the RegistryID and the CorrelationID.

Subsection 4.2: Deconstructing the Identity Framework: RegistryID vs. CorrelationID

To eliminate confusion arising from "ID multiplicity," it is essential to distinguish between these two identifiers and their specific functions.

- **RegistryID:** This is a unique identifier generated by CreditRegistry for a single entity record as reported by a *single subscriber*. It is permanently linked to the CustomerNumber provided by that subscriber. The RegistryID represents that subscriber's specific, sovereign view of their customer. It ensures that the data as submitted by the lender is preserved with complete integrity and is never altered by information from another source.
- **CorrelationID:** This is the master identifier, or "Master Entity ID," that links multiple RegistryIDs—often from different subscribers—that CreditRegistry's engine has determined belong to the same real-world individual or business. The CorrelationID is the key to the unified, 360-degree view.

The process of linking these records is known as correlation. It is a multi-layered process that employs both deterministic and probabilistic matching techniques to achieve a high degree of accuracy.

1. **Primary Anchor (Deterministic Matching):** The process begins with a deterministic, or rules-based, match on a unique national identifier. This is the strongest signal for a potential match.
   - For **individuals**, the **Bank Verification Number (BVN)** is the primary anchor.
   - For **businesses**, the **CAC Registration Number (RC Number)** or **Tax Identification Number (TIN)** serves as the primary anchor.
2. **Secondary Heuristics (Probabilistic Matching):** Once records are grouped by a matching primary anchor, the engine applies a set of probabilistic, or fuzzy-matching, algorithms to confirm the link with a high degree of confidence. These heuristics analyze other key data points to calculate the likelihood that the records refer to the same entity. These include:
   - **Name Similarity:** The system looks for a match of at least two tokens within the entity's name (e.g., "Adebayo" and "Williams" in "Adebayo Tunde Williams").
   - **Phone Number Match:** The engine compares reported phone numbers.
   - **Date of Birth Match:** Where the date of birth is deemed reliable and consistent

across records, it is used as a strong corroborating factor.

If the primary anchor matches and the secondary heuristics meet a predefined confidence threshold, the system assigns the same CorrelationID to all the relevant RegistryIDs, effectively merging them into a single, comprehensive entity profile.

The relationship between these key identifiers is summarized in the table below.

Table 1: The CreditRegistry Identity Framework

Identifier Name	Source / Owner	Scope	Core Function
CustomerNumber	Subscriber	Unique per customer *within that subscriber's portfolio*.	**Anchor:** The subscriber's primary key. Links a subscriber's internal record to a unique RegistryID.
BVN / RC Number / TIN	Federal Government (via NIBSS, CAC, FIRS)	Unique per individual/business *nationally*.	**Primary Correlation Key:** The main government-issued ID used to initiate the matching process across different subscribers.
RegistryID	CreditRegistry	Unique per subscriber's CustomerNumber.	**Sovereign Record:** Represents a single, un-altered view of an entity from one specific subscriber, ensuring data provenance.
CorrelationID	CreditRegistry	Links multiple RegistryIDs across all subscribers.	**Master Entity ID:** The "single source of truth" identifier.

			Groups all related records to form a complete, 360-degree view of an entity.

Section 5: Practical Application: Maximizing Efficiency and Accuracy in Credit Enquiries

Understanding the technical architecture of the RegistryID and CorrelationID framework is the first step. The second, more critical step is applying this knowledge in daily operations to conduct efficient and accurate credit assessments. By following a simple set of rules, subscribers can eliminate redundant queries, simplify their internal processes, and gain a complete view of a borrower's creditworthiness.

The Golden Rule of Correlation

The design of the correlation system leads to a straightforward and powerful principle for all credit enquiries:

If a search (e.g., by BVN) returns multiple records that share the same CorrelationID, a subscriber only needs to pull a full credit report or SMARTScore using *any one* of the associated RegistryIDs. The resulting product will automatically consolidate all loan records, repayment histories, and demographic data from all correlated RegistryIDs.

This single principle directly resolves the most common subscriber complaints regarding inefficiency and complex reconciliation. There is no need to make multiple API calls for each RegistryID within a correlated group. There is no need to build complex internal logic to stitch together data fragments. CreditRegistry's system performs this aggregation automatically on the backend. When a report is requested for any RegistryID that is part of a correlated set, the system uses the shared CorrelationID to gather all associated information and present it in a single, unified report. This is the intended and most efficient method for using the service.

Addressing the Exception: Un-correlated Records with the Same BVN

In rare instances, a search by BVN may return two or more records that have *different* CorrelationIDs. This is not a system failure; it is a deliberate and critical feature designed to preserve data integrity in the face of significant data quality anomalies.

This scenario occurs when the primary anchor (the BVN) matches, but the secondary heuristics fail to meet the required confidence threshold for an automatic link. For example, consider a single BVN associated with two records:

- Record 1 Name: 'Adebayo Tunde Williams'
- Record 2 Name: 'Funke Grace Adebayo'

While they share a BVN (likely due to a data entry error by a financial institution) and one name token ('Adebayo'), the other names are completely different. In this case, the entity resolution engine will correctly refuse to automatically correlate these records. To merge them would be a high-risk assumption that could pollute the valid credit file of one individual with the potentially erroneous or fraudulent data of another.

This "failure" to correlate is, in fact, the system succeeding in its primary mission: preventing incorrect data fusion. The "extra work" required of the subscriber—pulling two separate reports to investigate—is a necessary due diligence step triggered by a high-risk data anomaly. The system is effectively flagging a severe data quality issue that requires human review.

The ultimate solution to this situation is not for CreditRegistry to loosen its matching rules and risk data corruption. The solution lies with the subscribers who submitted the conflicting data. When this scenario is encountered, it should trigger an internal data quality review process. The recommended actions are:

1. Acknowledge the red flag raised by the system.
2. Review both credit reports to understand the nature of the discrepancy.
3. Contact the customer to verify their correct and complete details against official identification documents.
4. Submit a corrected data file to CreditRegistry during the next reporting cycle.

Once the source data is corrected by the responsible institution(s), the entity resolution engine will be able to correctly link the records on a subsequent run, resolving the issue for all future queries. This collaborative approach to resolving data anomalies strengthens the entire credit ecosystem.

Section 6: Addressing Subscriber Concerns: A Practical Q&A

This section directly addresses the four primary problems outlined by subscribers. By applying the concepts of shared responsibility, data quality, and the CorrelationID framework, it becomes clear that these challenges are not system limitations but are opportunities for process improvement and a deeper understanding of the service.

"Problem 1: Inefficient and Costly Querying"

- **The Issue:** "The customer cannot retrieve a complete credit report with a single API call. If their initial search...returns three different Registry IDs for the same individual, their system must run three separate report requests."
- **Our Answer:** This inefficiency stems from a misunderstanding of the CorrelationID system. As explained by the "Golden Rule of Correlation," if the three RegistryIDs share the same CorrelationID, only one API call is necessary. Requesting a report with just one of those RegistryIDs will yield a complete, consolidated report containing the data from all three. The system is specifically designed to prevent the need for multiple calls, thereby saving API usage, reducing network traffic, and accelerating the decision-making process.

"Problem 2: Complex and Risky Data Reconciliation"

- **The Issue:** "The customer has to build complex internal logic to stitch together the multiple report fragments...They must then decide how to aggregate the debts, facilities, and scores."
- **Our Answer:** CreditRegistry performs this reconciliation on behalf of the subscriber. The CorrelationID engine *is* the logic for stitching fragments together. The consolidated report delivered by our API is the final, aggregated view. Subscribers should not be building their own reconciliation logic, as this duplicates the core function of the credit bureau and introduces the risk of error. The bureau's aggregated report and the single, comprehensive SMARTScore calculated from all correlated data are designed to be the definitive source for a lending decision.

"Problem 3: Poor Customer Experience and High Support Cost"

- **The Issue:** "The multiplicity creates confusion...If a customer searches and is shown three separate records to choose from, they don't know which is the 'right' one...damaging confidence in the API's authority."
- **Our Answer:** The presentation of multiple records should be interpreted as a valuable data signal, not a system flaw. It indicates that different lenders in the ecosystem hold slightly different information for the same individual. The presence of a common CorrelationID across these records should instill confidence, not damage it. It is proof that our system has successfully navigated these minor inconsistencies to identify a single, unique entity. The "right" record for pulling a full report is any one of them, as they all point to the same consolidated profile.

"Problem 4: Failure to Provide a Single Source of Truth"

- **The Issue:** "A single primary identifier, such as a BVN, may return multiple reports and multiple scores...This happens when the same BVN is linked to several distinct, un-merged RegistryIDs."
- **Our Answer:** This is the most critical point of clarification. The CorrelationID *is* the single source of truth at the entity level. The scenario described—a BVN search yielding two records with two different CorrelationIDs and two different SMARTScores—is an intentional and vital feature of our risk management design. It is a red flag indicating a severe data quality conflict. The system is correctly and responsibly refusing to merge these records because the secondary data (e.g., names, DOB) are too dissimilar to do so with high confidence. Presenting two distinct profiles is not a failure of data integrity; it is the *preservation* of it. The correct action for the subscriber is to treat this as a high-risk case, investigate both records to understand the discrepancy, and conduct enhanced due diligence before making a lending decision. The system is protecting the subscriber from the risk of acting on potentially corrupted, incorrectly merged data.

The following table provides a clear summary of how a subscriber's actions directly influence their experience with the system.

Table 2: Mapping Subscriber Responsibility to System Outcomes

Subscriber Action / Input	CreditRegistry Process	Subscriber Outcome / Experience
GOOD: Submits clean, verified data with a	Records are easily matched with others for	**Efficient & Accurate:** A single API call on any

consistent CustomerNumber for an entity.	the same entity. All are linked under one CorrelationID.	RegistryID returns a complete, 360-degree report. High confidence in the single SMARTScore.
POOR: Submits data with name variations, typos, or an incorrect DOB.	Correlation heuristics may fail. The record may not be linked to others, resulting in a separate CorrelationID.	**Inefficient & Risky:** A BVN search returns multiple un-correlated records. Subscriber must pull multiple reports, manually reconcile, and deal with conflicting SMARTScores. This is the "pain point" scenario.

This demonstrates that the path to an efficient, accurate, and seamless experience is through a commitment to high-quality data submission at the source.

Section 7: Conclusion: A Partnership for a Stronger, More Transparent Credit Market

The integrity, efficiency, and fairness of Nigeria's credit market are built on a foundation of high-quality data. This whitepaper has sought to illuminate the sophisticated yet understandable architecture that CreditRegistry employs to manage this data, and to clarify the indispensable role that each subscriber plays in this vital ecosystem. The principle of shared responsibility is not merely a concept; it is the operational reality of our interconnected financial landscape.

The key to unlocking a unified, 360-degree view of any borrower lies in understanding and leveraging the CorrelationID. This powerful feature serves as the master key, linking disparate records into a single, coherent profile and providing the consolidated view that subscribers require for effective decision-making. The challenges of data fragmentation and query inefficiency are not inherent limitations of the system but are symptoms that can be resolved through adherence to best practices in data submission and a correct interpretation of search results.

Our system is robustly designed to handle the real-world imperfections of data. However, its effectiveness is magnified exponentially when it is fed with clean, accurate, and verified information from our subscriber partners. The path forward is a collaborative one, where the bureau provides powerful tools for data aggregation and analysis, and subscribers commit to the highest standards of data stewardship.

To that end, CreditRegistry provides the following actionable recommendations to all subscribers:

- **Review and Enhance Internal Processes:** Conduct a thorough review of internal KYC and data governance policies to ensure they align with CBN regulations and data quality best practices.
- **Standardize the CustomerNumber:** Ensure that a unique and persistent CustomerNumber is assigned to every customer and is consistently used in all data submissions. For institutions without internal ID systems, adopt the recommended practice of using the BVN or RC Number/TIN.
- **Train Relevant Staff:** Educate credit officers, data managers, and IT personnel on how to correctly interpret CreditRegistry search results, specifically on how to leverage the CorrelationID to avoid redundant queries.
- **Prioritize API-Based Submission:** Transition to or prioritize the use of the AutoCred8 Data Submission API *(or later)* for all new and updated records to benefit from real-time validation and improve data quality at the point of entry.

By working together in this partnership, CreditRegistry and its subscribers can continue to build a more inclusive, efficient, and robust financial ecosystem. This collaboration will reduce credit risk, improve the speed and accuracy of lending decisions, and ultimately contribute to the sustainable economic growth of Nigeria.

CREDITREGISTRY
SUPPORT.@CREDITREGISTRY.NG
www.creditregistry.ng

[i] **AutoCred8:** https://www.creditregistry.ng/how-to-use-autocred8/
[ii] **CreditRegistry Partner Connect:** https://connect.creditregistry.ng/