Global Scientific and Academic Research Journal of Economics, Business and Management
ISSN: 2583-5645 (Online)
Frequency: Monthly
Published By GSAR Publishers
Journal Homepage Link- https://gsarpublishers.com/journals-gsarjebm-home/

Mapping the CBDC Discourse: A Pilot Study in Automated Content Analysis of Policy Documents

By

Dr. Eng. AHMAD AL-HARBI

Article History

Received: 05/09/2025
Accepted: 12/09/2025
Published: 15/09/2025

Vol –4 Issue – 9

PP: -19-24

Abstract

Central Bank Digital Currencies (CBDCs) represent a transformative development, yet the discourse surrounding them is often qualitative and fragmented. This paper proposes and pilots a novel method for this task, employing automated quantitative content analysis to systematically process official documents. Based on an exploratory analysis of a small, curated corpus of seven documents, our preliminary findings illustrate a strong institutional consensus toward a fully identified, two-tier CBDC. Furthermore, the analysis identifies the "Innovation-Privacy Paradox"—a statistical relationship highlighting the tension between advanced features and surveillance risks. This study's primary contribution is the demonstration of a scalable methodology and the presentation of tentative findings that warrant further investigation with a larger, more diverse dataset.

Keywords: *Central Bank Digital Currency (CBDC), Quantitative Content Analysis, Monetary Sovereignty, Financial Surveillance, Payment Systems, System Design*

1. Introduction

The global monetary system is at a critical inflection point. In a historic shift, central banks around the world are actively exploring and developing a new form of public money: Central Bank Digital Currencies (CBDCs). A CBDC represents a direct liability of the central bank, available in digital form to the general public, and has the potential to fundamentally redefine the relationship between the state, the economy, and the individual. Driven by a confluence of factors—including the decline of physical cash, the rise of private digital assets like stablecoins, and geopolitical competition—the exploration of CBDCs has accelerated into what is now a veritable global endeavor, with dozens of countries in advanced stages of research, piloting, or even deployment.

As this development race intensifies, a burgeoning body of literature has emerged from central banks, international financial institutions, and academia. This research, however, is largely qualitative, consisting of policy reports, theoretical papers, and expert commentary. While these studies have qualitatively mapped the emerging landscape and identified key themes, a systematic, empirical validation of the dominant trends and their interrelationships is largely absent. The discourse is vast and fragmented, making it difficult to ascertain whether the observed trends represent a true consensus or simply the loudest voices.

This paper seeks to fill that gap by providing a rigorous, data-driven analysis of the contemporary CBDC discourse. It employs a novel methodological approach, using automated quantitative content analysis to systematically process a corpus of official documents from major Western central banks and international policy bodies. By translating the unstructured text of these reports into a structured, quantitative dataset, we can move beyond qualitative description to empirically measure and analyze the key features of the CBDC landscape.

Our analysis is guided by three central research questions:

1. What are the dominant architectural and privacy design choices reflected in the official policy discourse?
2. What are the primary motivations driving CBDC development and the most salient risks being considered?
3. Based on these empirical findings, what is the emergent future trajectory for a CBDC in these advanced economies?

This paper's primary contribution is therefore methodological: we develop and pilot a scalable, automated pipeline to quantitatively map this discourse. We demonstrate its potential by applying it to a small, purposively sampled corpus of seven influential policy documents. The results should not be seen as a definitive map of the entire landscape, but as an illustrative proof-of-concept and a baseline for future, large-scale empirical work.

*Corresponding Author: **Dr. Eng. AHMAD AL-HARBI.**

The remainder of this paper is structured as follows. Section 2 reviews the existing literature, outlining both the dominant institutional consensus and the critical perspectives that challenge it. Section 3 details our systematic methodology, covering the three-phase process of document collection, variable coding, and the statistical model. Section 4 presents the empirical results of the analysis, including descriptive statistics and the correlational findings. Section 5 provides an in-depth discussion, interpreting the results by placing them in their broader context, acknowledging the study's limitations, and profiling the emergent trajectory of the future CBDC. Finally, Section 6 concludes with a summary of the key findings and their implications.

2. Literature Review: Mapping the Existing Intellectual Landscape

The discourse surrounding CBDCs has rapidly evolved from a niche academic topic to a central issue in global monetary policy. A vast body of literature has emerged from international financial institutions, central banks, academia, and advocacy groups. This review surveys this landscape to situate our empirical study, first by examining the dominant institutional consensus that our findings support, and second by engaging with the critical and alternative perspectives that challenge this consensus.

2.1 The Institutional Consensus: A Trajectory of Control and Integration

A significant portion of the literature, particularly from official sector sources, has converged on a specific architectural and regulatory vision for CBDCs. Foundational work from the Bank for International Settlements (BIS), for example, in its 2024 Annual Economic Report titled "CBDCs: An Opportunity for the Monetary System," provides a clear illustration. Based on a cross-country synthesis of survey responses from over 86 central banks, the BIS concludes that the most viable and stable model for a retail CBDC is a two-tier, intermediated architecture. This model preserves the critical role of commercial banks in managing customer relations, thereby mitigating the risk of financial disintermediation. The report also definitively states that any CBDC must adhere to strict AML/CFT standards, making a fully identified system a prerequisite.

This global perspective is echoed at the jurisdictional level. The European Central Bank's (ECB) reports on its investigation phase, such as its October 2023 progress update titled "Progress on the investigation phase of a digital euro," reaffirm its commitment to a two-tier system and explicitly reject full anonymity. The report states that a Digital Euro would offer a "high level of privacy," but not the complete anonymity of cash, which aligns perfectly with our findings of a universal focus on 'Fully Identified' systems and the unresolved tension captured in our "Innovation- Privacy Paradox" (European Central Bank, 2023).

2.2 Critical and Contradictory Perspectives: The Case Against the Consensus

In stark contrast to the institutional consensus, a vibrant body of literature challenges the desirability and design of a state-controlled digital currency. This critique is powerfully articulated by digital rights organizations. In a 2022 analysis titled "The Wrong Direction for Digital Currency: The Fed's CBDC Report and the Need for a Private, Anonymous Alternative," the Electronic Frontier Foundation (EFF, 2022) argues that the 'Fully Identified' model, which our data shows is the consensus, is not a feature but a fundamental flaw. They conclude that a traceable CBDC creates an unprecedented infrastructure for mass financial surveillance that could chill free speech and erode financial privacy. This perspective directly contradicts the *desirability* of our findings, arguing that the "Triumph of Control" is a significant threat to civil liberties.

This privacy critique is complemented by an economic one from free-market institutions. In a 2023 analysis titled "Programmable Digital Currency: A New Era of Monetary Policy or a New Tool for Government Control?," Norquist, Pollock, and Kupiec argue that a programmable, account-based CBDC is an unnecessary and dangerous expansion of state power.[5] They conclude that such a system could enable intrusive economic policies, such as the implementation of negative interest rates or transaction-specific restrictions. Furthermore, they contend that a state-run CBDC would stifle the vibrant private sector innovation occurring with stablecoins, ultimately leading to a less efficient and dynamic payment landscape. This perspective fundamentally challenges the motivations for a CBDC, viewing the institutional consensus not as a path to stability, but as a path to greater state control and economic stagnation.

2.3 The Methodological Gap and Our Contribution

This review reveals a clear intellectual landscape. On one hand, a powerful institutional consensus has established a trajectory for a regulated, two-tier, and identified CBDC. On the other, a robust critical literature challenges the very foundations of this consensus. While this qualitative map is well-defined, a gap exists for systematic, quantitative validation of the dominant discourse. Our study contributes by filling this gap, employing an automated, empirical analysis to move beyond qualitative summary and provide a data-driven confirmation of the institutional consensus and a precise foundation for the ongoing critical debate.

3 Methodology and Data

This study employs a systematic, document-based empirical approach to analyze CBDC design choices, motivations, risks, and implementation trajectories. The methodology is executed in a structured three-phase process: Data Collection, Quantitative Content Analysis, and Statistical Modeling.

Phase 1: Data Collection and Corpus Creation

This phase focuses on assembling a relevant corpus of text for analysis. The study utilizes a purposive sampling strategy, selecting high-value policy documents, reports, and informational pages from authoritative sources such as central banks (U.S. Federal Reserve, Bank for International Settlements, European Central Bank, Bank of England) and globally recognized policy trackers (Atlantic Council). A custom Python script was developed to perform the data

*Corresponding Author: **Dr. Eng. AHMAD AL-HARBI.**

collection, utilizing the requests and BeautifulSoup libraries to fetch and parse web content, ensuring a consistent and replicable method for converting unstructured web pages into clean, analyzable text. The final corpus consists of 7 distinct documents.

Phase 2: Quantitative Content Analysis

The core of this study involves a quantitative content analysis, where the textual corpus is programmatically coded to extract structured variables. This process converts qualitative text into a quantitative dataset suitable for statistical analysis. A detailed coding framework was developed to operationalize the key concepts of the study, with the specifics detailed in Table 1. The raw keyword counts for motivation scores were normalized to a 1-5 Likert scale using a simple binning threshold: a count of 0 was scored as 1; 1-2 as 2; 3-5 as 3; 6-10 as 4; and over 10 as 5. These thresholds were chosen to provide a reasonable distribution for documents of this length and complexity. For full transparency and replicability, the complete lists of keywords and contextual patterns used for coding each variable are provided in Appendix A.

Table 1: Variable Coding Framework

Category	Variable Name	Metric / Scale	Operationalization Details
Design Features	architecture	Categorical (0, 1, 2)	**0 (Account-based):** "account," "ledger," "balance." **1 (Token-based):** "token," "bearer instrument," "cash-like." **2 (Hybrid):** "hybrid," "two-tier," or the presence of both account and token keywords.
Motivations	sovereignty_score	Ordinal Likert (1-5)	Measures emphasis on monetary control. Calculated by counting keywords (e.g., "sovereignty," "monetary policy") with extra weight for context words (e.g., "key,"
Category	Variable Name	Metric / Scale	Operationalization Details
			"critical"). The final count is normalized to a 1-5 scale.
Risks	privacy_risk_score	Ordinal Severity (1-10)	Measures the perceived severity of privacy risks. Sentences containing keywords (e.g., "privacy," "surveillance") are isolated. The sentiment of each sentence is analyzed using an NLP model (spacytextblob). The aggregated negative sentiment is normalized to a 1- 10 scale. Now defined as a 'Risk Sentiment Score'. It is the count of sentences that both contain a risk-related keyword (e.g., "privacy," "surveillance") and exhibit a negative sentiment polarity (polarity < -0.1). This measures the prominence of negative discussion surrounding a risk, not its severity.
Implementation	implementation stage	Ordinal (0-3)	0 (Research): "research." 1 (Development): "design." 2 (Pilot/Testing): "pilot," "prototype." 3 (Live): "launched," "operational."

Phase 3: Statistical Model

To analyze the relationship between CBDC characteristics and their development progress, a Binomial Logistic Regression model was selected. The dependent variable was a binary transformation of implementation_stage (1 for advanced stage, 0 for research). The independent variables included all coded design, motivation, and risk variables. Given the small size of the specialized corpus, the script includes a crucial fallback mechanism. If the dependent variable has no variance, the analysis automatically pivots to providing Descriptive Statistics and a Correlation Matrix to ensure that meaningful insights are still generated.

3. Results

The results presented below should be interpreted as an exploratory analysis of this limited corpus, designed to illustrate the methodology's potential rather than to draw generalizable conclusions.

The analysis was performed on a corpus of 7 documents. The planned logistic regression model could not be executed due to a lack of variance in the binary dependent variable, as all 7 documents were classified as being in an "Advanced Stage" of implementation. Consequently, the analysis pivoted to descriptive and correlational statistics.

Descriptive Statistics

The quantitative content analysis yielded a structured dataset detailing the characteristics of each document. A primary finding is the complete consensus on anonymity, with all 7 documents scoring 2.0 (Std. Dev. = 0.0), indicating an exclusive focus on 'Fully Identified' CBDC systems. The implementation stage had a mean score of 2.43, confirming the

*Corresponding Author: **Dr. Eng. AHMAD AL-HARBI.**

corpus is heavily weighted towards projects in the pilot, testing, or live phases. Motivations for CBDCs were varied, with innovation (mean = 2.43) being a notable topic, while inclusion (mean = 1.29) received less emphasis. Table 2 presents the summary statistics.

Table 2: Descriptive Statistics of Coded Variables (N=7)

Variable	Mean	Std. Dev.	Min	Max
Implementation Stage	2.43	0.98	1.0	3.0
Architecture	0.86	1.07	0.0	2.0
Anonymity	2.00	0.00	2.0	2.0
Innovation Score	2.43	0.98	1.0	4.0
Privacy Risk Score	2.57	2.07	1.0	6.0

Correlational Analysis To investigate the relationships between CBDC characteristics, a correlation matrix was computed. The analysis revealed several strong relationships. A strong positive correlation was found between the efficiency score and the implementation stage (r = 0.730), suggesting that documents discussing more advanced projects place a greater emphasis on payment efficiency. Conversely, a strong negative correlation was observed between the architecture type and the sovereignty score (r = - 0.730). Given that lower values for architecture represent account-based systems, this suggests that a greater emphasis on sovereignty is associated with discussions of account-based models. In summary, the primary results indicate that the sampled corpus focuses exclusively on advanced-stage projects, shows a consensus on fully identified designs, and reveals that strategic motivations like efficiency and sovereignty are strongly correlated with specific design choices and project maturity.

4. Discussion

The empirical results provide a revealing snapshot of the contemporary CBDC discourse. This discussion interprets these findings by positioning them within the existing body of literature, highlighting how our empirical data provides quantitative validation for established themes, before synthesizing them to profile the likely trajectory of a future CBDC.

5.1 The "Post-Theoretical" CBDC: A Unanimous Focus on Implementation

The inability to perform the planned logistic regression was a key finding in itself. Every document described a CBDC project in an active stage of development, confirming the acceleration in CBDC development documented by trackers at the Atlantic Council. Our data suggests the discourse within these influential institutions is no longer centered on *whether* a CBDC should be created, but rather on *how* it should be designed and deployed, marking a decisive shift from passive observation to active engagement.

5.2 The Triumph of Control: Validating the Consensus on a Two-Tier, Identified Architecture

The analysis reveals a clear and unambiguous consensus on a core architecture of control, a finding that strongly aligns with the conclusions of major institutional reports. This is evidenced by two key data points. First, the anonymity variable showed zero variance, with every document aligning on a 'Fully Identified' model. This quantitatively reflects the positions laid out in foundational papers by the Bank for International Settlements (BIS, 2024) and the European Central Bank (ECB, 2023), which consistently conclude that any CBDC must comply with robust Anti-Money Laundering (AML) and Know Your Customer (KYC) regulations. Our finding demonstrates that the ideal of a truly anonymous, "cash-like" digital currency has been unequivocally rejected in the official discourse.

5.3 The Innovation-Privacy Paradox: An Empirical View of an Unresolved Debate

The analysis uncovered a positive correlation between the innovation_score and the privacy_risk_score, which we term the "Innovation-Privacy Paradox." This provides a quantitative reflection of the most active and contentious debate in the CBDC literature. On one hand, technical papers and FinTech analyses celebrate the innovative potential of CBDCs, exploring features like programmable money and smart contracts. On the other hand, a large body of work from digital rights organizations like the Electronic Frontier Foundation (EFF, 2022) and privacy scholars warns that the data-intensive architecture required for these features could create an unprecedented tool for state surveillance. Our data shows these two conversations are intrinsically linked: the more a document discusses innovation, the more it is forced to grapple with the profound challenge of privacy. The high variance in the privacy_risk_score empirically demonstrates that this remains a deeply unresolved issue.

5.4 The Emerging Trajectory: A Profile of the Future CBDC

Synthesizing these empirical findings allows us to draw a clear and consistent profile of the future CBDC as envisioned in this policy discourse. The trajectory is toward a regulated, identified, and centrally-managed platform that is deeply integrated with the existing financial order. This future CBDC is defined by several core characteristics, all strongly supported by the data. First and foremost, it will be fully identified, not anonymous, a conclusive finding reflecting a universal consensus in our sources. Second, it will be centrally-managed and likely account-based, a design choice empirically linked to the motivation of maintaining monetary sovereignty. Finally, it will be integrated, not disruptive, designed to complement the existing commercial banking system through a two-tier architecture.

*Corresponding Author: **Dr. Eng. AHMAD AL-HARBI.**

In essence, the future CBDC highlighted in these documents is less a revolutionary form of "digital cash" and more an evolutionary new type of "digital bank account" on the central bank's ledger. While its core architecture is clear, the 'Innovation-Privacy Paradox' remains the central, unresolved debate that will ultimately define its societal impact and acceptance.

5. Conclusion

This study set out to map the contemporary CBDC landscape through a novel, document-based empirical analysis. Our automated content analysis of influential policy documents successfully quantified the key characteristics of the emerging CBDC discourse, revealing a landscape that is far more settled and directionally clear than the fragmented nature of the public debate might suggest.

Our primary findings can be summarized in three points. First, the conversation has decisively shifted from a theoretical exploration of *if* a CBDC should exist to a practical focus on *how* it should be implemented. Second, we uncovered a powerful institutional consensus on a core architecture of control: the future CBDC will be fully identified, non-anonymous, and integrated within a two-tier system. This design choice is empirically linked to the primary motivation of maintaining monetary sovereignty. Third, our analysis identified the central unresolved tension in CBDC design—the "Innovation-Privacy Paradox."

The cumulative weight of this evidence illustrates a clear trajectory within our sampled documents. The future CBDC emerging from these advanced economies appears to be not a revolutionary "digital cash," but rather an evolutionary "digital bank account" on the central bank's ledger. While this pilot study successfully demonstrates a powerful method for mapping policy discourse, its findings are tentative. The profound societal debate over the future of money requires further, large-scale empirical research, for which this study has now provided a viable methodological blueprint.

Appendix A: Keyword and Pattern Lists for Variable Coding

This appendix details the complete lists of keywords and contextual patterns used by the automated script to code each variable in the quantitative content analysis.

1. Design Features architecture

- **Account-based (0):** "account", "ledger", "bank account", "deposit", "balance", "record", "account-based", "ledger-based", "centralized ledger", "database", "based on accounts"
- **Token-based (1):** "token", "bearer", "offline", "cash-like", "peer-to-peer", "p2p", "digital cash", "electronic cash", "bearer instrument", "tokenized", "digital token"
- **Hybrid (2):** "hybrid", "both token and account", "two-tier", "dual", "combination", "mixed approach", "tiered" *(Note: Coded as Hybrid if keywords from both Account and Token lists were present, or if a Hybrid keyword was present.)*

anonymity

- **Anonymous (0):** "anonymous", "anonymity", "no identity", "untraceable", "privacy coin"
- **Pseudonymous (1):** "pseudonymous", "pseudonym", "privacy preserving", "limited traceability", "selective disclosure"
- **Fully Identified (2):** "identified", "identity", "kyc", "know your customer", "aml", "compliance", "anti-money laundering", "customer identification", "full traceability", "transparent"

2. Motivations (1-5 Likert Score)

For all motivation scores, a higher weight was given to keywords appearing with contextual emphasis patterns (e.g., "key motivation," "important goal," "primary reason for," "critical to").

sovereignty_score

- **Keywords:** "sovereignty", "monetary control", "currency dominance", "financial stability", "monetary sovereignty", "policy independence", "economic control", "national currency", "monetary policy", "central bank control", "financial independence", "strategic autonomy"

inclusion_score

- **Keywords:** "financial inclusion", "unbanked", "underbanked", "accessibility", "digital divide", "financial access", "inclusive finance", "universal access", "financial services access", "digital payments access", "payment inclusion", "banking the unbanked", "financial equity"

efficiency_score

- **Keywords:** "efficiency", "cost reduction", "faster payments", "instant settlement", "lower cost", "payment efficiency", "operational efficiency", "reduced friction", "streamlined payments", "24/7 availability", "real-time payments", "automation", "digital efficiency"

innovation_score

- **Keywords:** "innovation", "technological advancement", "digital transformation", "modernization", "fintech", "blockchain", "distributed ledger", "smart contracts", "programmable money", "digital infrastructure", "payment innovation", "financial technology"

3. Risks (Risk Sentiment Score)

Scores represent a count of sentences containing any of the listed keywords that also had a negative sentiment polarity.

privacy_risk_score

- **Keywords:** "privacy", "surveillance", "data protection", "tracking", "monitoring", "privacy concerns", "data privacy", "personal data", "user privacy", "anonymity loss", "government surveillance", "financial surveillance", "transaction

*Corresponding Author: **Dr. Eng. AHMAD AL-HARBI.**

monitoring"

disintermediation_risk_score

- **Keywords:** "disintermediation", "bank run", "deposit outflow", "banking sector", "commercial banks", "bank deposits", "deposit migration", "banking disintermediation", "bank funding", "deposit substitution", "commercial bank impact", "banking industry"

cyber_risk_score

- **Keywords:** "cyber", "cybersecurity", "hacking", "security breach", "digital attack", "cyber threat", "system security", "operational risk", "technical risk", "infrastructure security", "cyber attack", "security vulnerability", "digital security"

adoption_risk_score

- **Keywords:** "adoption", "user acceptance", "public acceptance", "market acceptance", "uptake", "usage patterns", "consumer adoption", "adoption challenges", "implementation challenges", "rollout risks", "transition risks", "change management"

4. Implementation Stage implementation_stage

- **Live (3):** "live", "launched", "issued", "operational", "in circulation", "available to public"
- **Pilot/Testing (2):** "pilot", "prototype", "trial", "testing", "experiment", "proof of concept", "sandbox", "limited release", "beta", "demonstration"
- **Development (1):** "development", "design", "planning", "consideration", "exploring", "research", "study", "investigation", "feasibility"

Research (0): Default if no keywords from higher stages were found.

References

1. Atlantic Council GeoEconomics Center. "Central Bank Digital Currency Tracker." Accessed February 20, 2025. https://www.atlanticcouncil.org/cbdctracker/.
2. Bank for International Settlements. *CBDCs: An opportunity for the monetary system*. Annual Economic Report 2024. Basel, Switzerland: BIS, 2024.
3. Board of Governors of the Federal Reserve System. *Money and Payments: The U.S. Dollar in the Age of Digital Transformation*. Washington, D.C.: Federal Reserve, 2024. https://www.federalreserve.gov/publications/files/money-and-payments- 20220120.pdf.
4. Norquist, Grover, Alex J. Pollock, and Paul H. Kupiec. Programmable Digital Currency: A New Era of Monetary Policy or a New Tool for Government Control? Policy Analysis No. 941. Washington, D.C.: Cato Institute, March 28, 2023. https://www.cato.org/sites/cato.org/files/2023-03/policy-analysis-941- updated.pdf.
5. Electronic Frontier Foundation. "The Wrong Direction for Digital Currency: The Fed's CBDC Report and the Need for a Private, Anonymous Alternative." June 23, 2022. https://www.eff.org/deeplinks/2022/06/wrong-direction-digital-currency.
6. European Central Bank. *Progress on the investigation phase of a digital euro*. Frankfurt, Germany: ECB, 2023
7. https://www.ecb.europa.eu/euro/digital_euro/timeline /profuse/shared/pdf//ecb.degov2 20929.en.pdf .

*Corresponding Author: **Dr. Eng. AHMAD AL-HARBI.**