Armstrong Wolfe

The Power of Collective Ambition

The International COO Community (**iCOOC**) Industry Paper 2025

Solving the Data Completeness Challenge

Global Economics GROUP

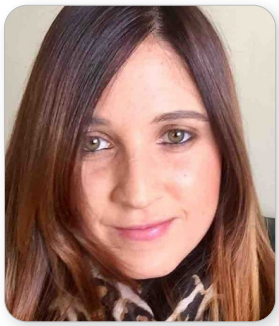
ARMSTRONG WOLFE™

The International COO Community (**iCOOC**) Alliance Partner

Global Economics Group

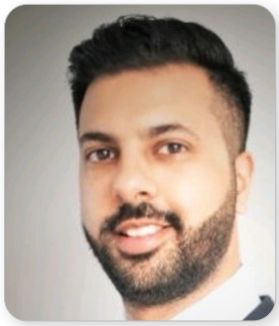
Data Completeness Working Group

Chris Rigg
Principal
Global Economics Group



Lorian Barrett
Data Strategy Delivery Manager
Ninety One

Brian Greenberg
Business Engagement Lead
Enterprise Data Management
BNY



Rishi Larkhani
COO Chief Data Office
DWS

Marc Price
Chief Strategy Officer CMBYND
Former Scotiabank

Rav Swamy
Chief Digital Officer
Pzena Investment Management

Contributing Organisations

BNY

DWS

HSBC

J.P.Morgan
ASSET MANAGEMENT

Ninety One

PZENA Investment Management

The Missing Piece: Solving the Data Completeness Challenge with Effective Governance

Introduction: The Data Completeness Imperative

Data has always been central to delivering financial services. Financial institutions manage risk to achieve returns for their clients and shareholders. Effective risk management depends on timely and accurate data to inform the business' stakeholders about its state. Additionally, the financial services industry has become digitized, materially increasing the pace of activity and requiring management teams to process higher volumes of data.

As financial institutions increasingly integrate AI into decision-making and operations, the demand for complete and well-governed data has become even more critical. AI models rely on high-quality, comprehensive datasets to provide accurate insights, making data governance a foundational element of success in the AI era.

Simultaneously, the volume and frequency of data that financial institutions use to manage their operations are increasing exponentially. The quality of this data is critical; errors can lead to misleading business insights, failed transactions, miscalculated risks, regulatory fines, and negative experiences for both employees and customers. To address these risks and improve the accessibility of core transaction and reference data, institutions have prioritized data governance.

However, introducing new datasets, including unstructured and external data, adds complexity to the landscape. Each new dataset must be carefully assessed for its intended use and evaluated for accuracy and completeness. Stakeholders continually demand additional data to meet regulatory requirements, enhance AI model performance, and deliver increasingly personalized customer experiences.

Recognizing these challenges, Armstrong Wolfe and Global Economics Group convened an industry working group focused on data completeness to explore the data governance challenges created by the financial services sector's ever-growing demand for more data

The group included representatives from buy-side and sell-side institutions across various roles, including data management, technology, compliance, and operations. Throughout three working sessions, we explored:

- The expanding data landscape - the types and domains of data consumed by financial institutions
- Data governance challenges and pain points
- Practical approaches to improving data governance
- The foundational capabilities for effective data governance
- Data governance strategies and success criteria
- Implementation and adoption challenges.

These themes will be discussed in detail below.

Understanding The Expanding Data Landscape

Financial services companies have always relied on core data to manage their businesses effectively. This data spans critical domains, including customer information, products, risks, compliance, sales, and operations. Historically, most of this data has been structured, primarily consisting of numerical values and text. It is often referred to as "core data" because it is central to the institution's core processes. For example, trades are characterized by data such as price, quantity, counterparty, trade date, and settlement date. The quality of this core data directly impacts a trade's economics and is essential for running the business.

In recent years, the demand for data has evolved significantly. Business stakeholders seek access to more diverse datasets, including unstructured data such as text-based research or commentary, images, voice recordings, and videos. This type of data, often referred to as 'extended data,' provides additional context that enhances decision-making. However, the rapid growth of extended data poses challenges for data governance frameworks, which must ensure its quality, accuracy, and completeness amid increasing complexity.

Emerging technologies like artificial intelligence introduce another dimension to the data landscape. While AI thrives on large, varied datasets, its dependence on high-quality and complete data highlights the importance of effective governance. Poorly governed or incomplete data can compromise AI's reliability, leading to operational and reputational risks. Addressing these risks requires ensuring that both core and extended datasets are managed with equal rigor and attention to detail.

At the same time, clients are increasingly requesting more granular and transparent data in their reporting. They want to gain a deeper understanding of how the products and services they utilize align with their strategic objectives and values. This rising demand underscores the need for robust governance practices that can adapt to the expanding data landscape and ensure that all data (structured and unstructured) is accurate, complete, and ready for use.

"Emerging technologies, such as artificial intelligence, add another dimension to the data landscape."

What is Data Governance?

Data governance for financial services organizations is a comprehensive framework for managing, controlling, securing, and optimizing data throughout its lifecycle. It encompasses the policies, procedures, and practices that ensure data accuracy, completeness, consistency, timeliness, accessibility, and compliance with internal policies, audits, and legal and regulatory requirements.

This framework includes methods for data collection, storage, processing, analysis, and sharing, enabling financial institutions to make informed decisions based on reliable data while preserving data quality, security, and regulatory compliance.

Data completeness is a crucial element of this framework. It denotes the degree to which all necessary data is present and accessible for use. Completeness guarantees that datasets are thorough and devoid of missing components that could jeopardize their reliability for analysis, decision-making, or regulatory reporting. Incomplete data creates gaps that can result in errors, inefficiencies, and risks, making it a central concern of effective governance.

The key components of data governance include:

- **Data Stewardship:** Designating individuals responsible for overseeing data assets and implementing policies. Data stewards need to come from the parts of the organization that are best positioned to opine on the quality and appropriate use of the data in their charge.
- **Data Quality Management**: Systematic practices to maintain high data quality data by determining the best and most accurate source, validating the data at the attribute level when it is ingested, and creating running routines to cleanse the data.
- **Data Policies:** Detailed guidelines governing data usage and ensuring compliance with internal and regulatory standards
- **Compliance Monitoring:** Continuous oversight to ensure adherence to legal and regulatory requirements
- **Data Security:** Protecting data from unauthorized access

Why Data Governance and Completeness Matter?

Data governance is crucial for financial services organizations because it ensures that data, including its completeness, is fit for purpose. This is essential for:

- Ensuring data accuracy and consistency, which promotes transparency and accountability across the organization
- Maintaining regulatory compliance, reducing the risk of fines and penalties due to inconsistent reporting.
- Improving decision-making processes by providing reliable and accessible data.
- Enhancing data security and protecting sensitive information.
- Fostering customer trust by safeguarding their data and ensuring all necessary information is captured and managed effectively.

By adopting strong data governance practices that prioritize data completeness, financial services organizations can effectively manage their data assets, reduce risks, and leverage data for strategic advantage in an increasingly data-driven industry.

Data Governance Challenges and Pain Points

Effective data governance is inherently challenging. While everyone values quality data, few are willing to engage in the necessary processes to ensure it. In financial services, this challenge is amplified by increasingly complex and distributed data ecosystems. Governance processes can often seem costly and cumbersome, diverting attention from core business objectives like serving customers and driving revenue.

The pressure to meet immediate business goals further exacerbates the issue, leading stakeholders to bypass the officially approved data sources in favor of quicker, more convenient options. Ironically, these same stakeholders often encounter quality issues with their chosen data, ultimately impacting business performance and reinforcing the need for robust governance.

These challenges can be broken down into key pain points that hinder the effective implementation of data governance:

1. **Organization**: Ownership and accountability are two areas where the various data stakeholders disagree. The business, usually represented by the front office, is often seen as the owner of the data because it is closest to the trustworthy sources, including the transactions, markets, and clients whose daily activities generate the data utilized by downstream stakeholders.

   The business frequently prefers to delegate ownership and accountability for data quality to technology, operations, and data office groups that manage the infrastructure and processes responsible for consuming the data and producing outputs based on it. The challenge is that both sides are correct and incorrect.

   Ownership is a complex term that can have multiple interpretations. Is the owner the person who understands the data and how it drives the business, or is it the individual who comprehends how governance works and how it relates to managing data quality? The answer is that both groups share ownership responsibilities and have accountabilities within the governance processes throughout the business lifecycle.

2. **Complexity**: Financial institutions consist of multiple interrelated lines of business that possess significant autonomy in how they serve clients and fulfill obligations. Individual lines of business often share an overlapping client base and must cooperate and collaborate on key concepts that affect the broader organization.

   One business may develop custom processes and infrastructure to manage its core data but may not recognize that this data needs to be consistent throughout the organization. A tension exists between centralized functions, such as risk, compliance, and audit, and decentralized business units. Central groups often lack familiarity with the specifics of how the business utilizes core and extended data, leading to conflicts during governance processes' execution.

   Additionally, different functions and groups within the business maintain varying standards for data quality. A risk function may choose to use lower-quality data earlier in the lifecycle because it might contain valuable signal information. In contrast, other functions, like compliance, demand the most accurate data possible to fulfill their obligations to stakeholders.

"The pressure to meet immediate business goals further exacerbates the issue..."

3. **Legacy Infrastructure:** The age and complexity of the institution's technology infrastructure also challenge data governance because the techniques, tools, and processes currently employed to manage data quality did not exist when many of these technologies and applications were first implemented.

   Adapting this legacy infrastructure to newer approaches is considered expensive and crowds out investments that could grow revenue, increase efficiency, or enhance the customer experience. Business owners are often reluctant to fund changes to business processes and technology because they do not comprehend the risks to the organization posed by ineffective data governance.

   Regulators are focused on data quality and have taken enforcement actions, including fines, which highlight the need for change and compel management to make the necessary investments.

4. **Culture:** Financial institutions aim to meet customers' expectations by delivering products and services. While data is central to that process, culturally, individuals may be encouraged to act quickly and undertake actions that boost revenue at the expense of good governance.

   There is often a lack of understanding regarding the importance of data quality for the organization's success, leading to attitudes that see concepts like stewardship as someone else's responsibility. Leadership must elevate the significance of data quality and support initiatives that foster accountability and ownership throughout the organization.

"Ownership is a complex term that can have multiple interpretations."

Practical Approaches to Data Governance

Data governance as a function within financial services institutions is not new. As the number and types of regulations applied to the industry have increased, institutions have made significant investments in the data supply chain necessary to source the data needed to meet their regulatory obligations. This includes governance to ensure the quality of the data and demonstrating to regulators that institutions understand the lineage of the data in their reports. While data governance processes and tools are present in most institutions, they often struggle to achieve their objectives at scale, and some organizations are further along the maturity curve than others. The working group discussed several approaches that have proven successful, including:

1. **Policy-Driven:** Many institutions have developed data policies that, when properly implemented, fulfill their obligations under essential regulations such as BCBS 239 and GDPR. The successful implementation of this approach involves several key capabilities:
    a. **Roles and Responsibilities:** Defining the roles within the organization for those responsible for providing, consuming, managing, and governing data. Implementing new roles necessitates establishing new organizational structures to accommodate specific roles and obtaining consensus to allocate additional responsibilities to existing roles across the front, middle, and back-office organizations.
    b. **Data Management Infrastructure:** Clearly defining which components (applications and databases) of the institution's technical infrastructure serve as **authoritative data sources**. Authoritative sources are responsible for adhering to critical elements of data policies to ensure the quality and integrity of the data. The data policies mandate that data consumers must utilize authoritative sources to generate outputs for customer communication, key management reporting and decisions, and regulatory reporting.
    c. **Policy Enforcement:** Requiring all stakeholders to utilize the authoritative infrastructure or face the consequences while conducting audits and leveraging technology-driven capabilities to monitor compliance.
2. **Technology-Driven:** This approach leverages an integrated data management environment to automate much of the data governance process, including:
    a. Implementing automated data quality management tools for real-time identification and resolution of data inconsistencies.
    b. Utilizing data cataloging and metadata management systems to centralize data assets and provide context.
    c. Employing data lineage and provenance tracking to maintain transparency and regulatory compliance
    d. Leveraging AI and machine learning for enhanced data governance processes.
3. **Culture and Organization-Driven:** Emphasizing the significance of data governance throughout the organization and facilitating interactions between data producers and data consumers within the governance processes and compliance with policies, including:
    a. Taking a people-centric approach by rebranding roles and gamifying data management tasks to engage employees.
    b. Prioritizing collaboration across departments to promote better data sharing and decision-making.
    c. Focusing on incremental integration of governance practices rather than a top-down, heavy-handed approach.

There is a natural tension between a policy-driven approach, which can establish the conditions for effective data governance, and conflict and disruption if the organization lacks maturity at either the centralized governance layer or the business unit level. Implementing policies that lead to significant non-compliance from the outset can do more harm than good. Regardless of which approach is preferred, technology plays a crucial role in providing the necessary capabilities for data governance.

The Foundation of a Data Governance Capability

A successful data governance program will likely integrate elements of the aforementioned approaches.

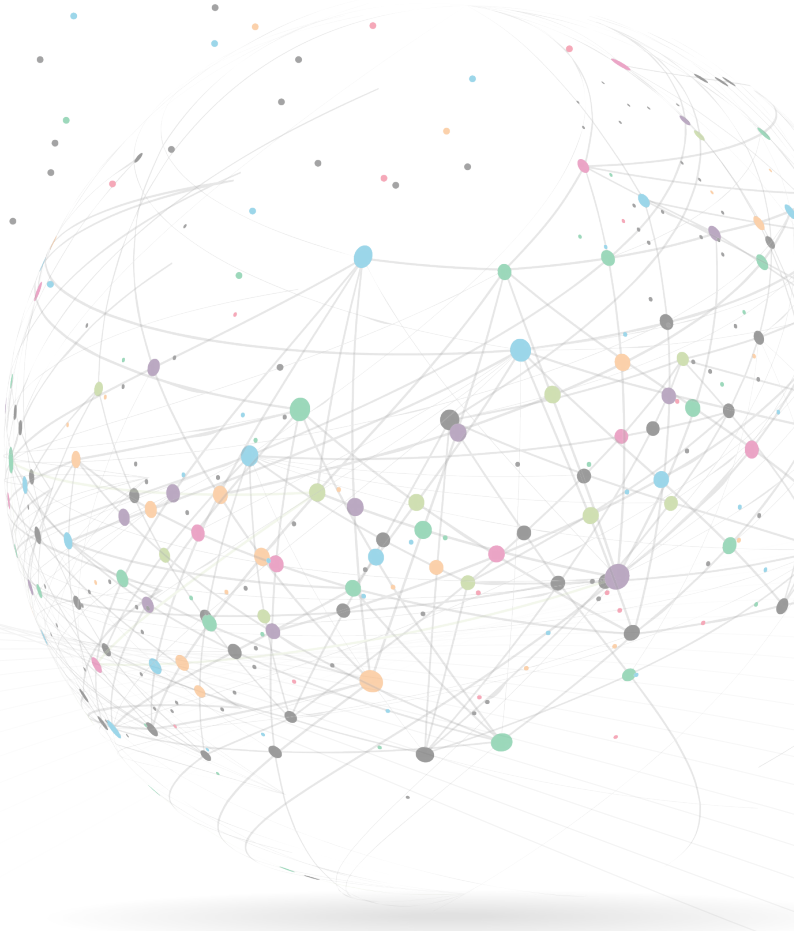
Effective data governance in an environment characterized by rapid data growth and a regulatory obligation to maintain high quality and lineage necessitates a set of technical capabilities that can enhance the people and processes involved. The guiding principle for a data governance capability is to concentrate on the features that provide various stakeholders with the information needed to make sound decisions and the capabilities to implement and monitor those decisions effectively.

- **Data Asset Inventory:** You can't govern what you don't know exists. Data assets encompass more than just raw data; they include policies, data quality rules, applications, reports, models, metrics, and metadata. A comprehensive data asset inventory features both automatic detection capabilities, such as database crawling, and manual entry options that allow data owners to record and publish the data assets and products they are responsible for.
- **Context-Aware Search:** Both data consumers and producers should be able to search for and discover data assets. Facilitating easy access and usage is one of the most effective ways to ensure stakeholders utilize the appropriate data sources and embrace proper quality management practices.
- **Data Quality Measurement:** Automated tools for managing data quality, enabling timely identification and resolution of inconsistencies.
- **Data Lineage and Provenance:** Ensuring transparency and traceability for compliance and operational efficiency.
- **Change Detection:** Data assets are continually evolving, making it essential for any data governance capability to recognize changes.
- **Active Workflow Management:** Data quality issues should be assigned to the appropriate individuals and monitored to ensure their resolution.
- **Self-Service:** The capacity for data consumers and data producers to independently address data quality issues without help from a technology or operations team.
- **Policy-Driven:** Clearly articulated policies that outline the acceptable and appropriate use of data are essential for ensuring the adoption of data governance processes.
- **Measurement-Driven:** You can't manage what you can't measure. Establishing suitable metrics that evaluate and inform stakeholders about the program's health is essential for achieving the transparency necessary for effective governance.
- **Continuous Improvement:** The standard for data governance should continuously be elevated across all program dimensions, including both outcomes and inputs. Over time, governance processes should become more efficient, effective, and cost-effective as their core principles are seamlessly integrated into the organization's daily operations and culture. This ongoing evolution ensures that the governance framework remains adaptable and consistently delivers value in an ever-changing data landscape.
- **Sustainability:** Data management is a continuous responsibility, not a one-off Task.

Defining Success: Metrics and Outcomes

While governing data is an evergreen process, it is essential to establish key criteria and metrics that indicate success to ensure that the investments made to establish and support these processes produce a measurable and appropriate return. Key success criteria established by the team include:

1. **Data Quality Standards:**
    a. Thorough identification of business and critical data elements with clear definitions and organized processes.
    b. Establishing scorecards to measure compliance and progress, focusing on data usage compliance from both producer and consumer perspectives.
2. **Success Metrics:**
    a. Reduction in incidents and exceptions related to data quality or data breaches.
    b. Faster responses to compliance and regulatory inquiries.
    c. Improved AI model performance.
    d. Increased adoption and trust in the authoritative data sources.
    e. Demonstrating ROI through cost savings, risk reduction, and revenue generation.
    f. Highlighting tangible commercial use cases, such as reduced reconciliation efforts and efficiency gains in market data spending.
    g. Identification and addressing of root causes to eliminate repeat issues.
    h. Development of new products and services based on accurate insights from data, improvements in client satisfaction scores, or an increase in the number of new clients gained.
3. **Culture:**
    a. Feedback mechanisms like surveys to capture organizational satisfaction with the governance processes, including auditors, regulators, risk, compliance, and business stakeholders.
    b. Transitioning from operational focus (e.g., settling trades) to understanding and treating data as a valuable asset.
    c. Embedding a mindset that connects data governance to regulatory, operational, and commercial goals.
    d. Securing agreement on governance priorities and ensuring alignment and accountability across all organizational levels.
    e. Moving from a culture of reactive localized symptom control to proactive mitigation of root causes to eliminate issues instead of treating them.
    f. Achieve data management at scale, i.e., control once and use many times. It is not commercially viable to add control on top of control across all levels of the organization.



Barriers to Adoption of Effective Data Governance Processes

Establishing an effective data governance program at a large financial institution requires agreement from many different parts of the organization to change how they approach data consumption and production. Change is inherently problematic in an environment pressured to satisfy customer demands, deliver returns for investors and stakeholders, and meet ever-increasing regulatory obligations.

1. **Cultural and Behavioral Resistance:** As the saying goes, "Old habits die hard." Many of the processes and tools organizations use for data management and governance are years, if not decades, old and were developed before data governance became a critical issue. Key functions such as risk and operations source data from legacy systems, and tools like Excel are employed to manage and manipulate it. These habits are deeply entrenched, and key stakeholders believe that the cost of migrating to a more controlled and governed process outweighs any perceived benefits..
2. **Policy Challenges:** Creating and implementing new policies can be challenging if the organization is unprepared to comply. A policy that immediately places a large percentage of the organization in a non-compliant state is impractical, introduces risks, and is likely to generate significant pushback, jeopardizing the overall program. Policies must achieve a balance: they should be enforceable yet not overly stringent from the outset. The consequences of non-compliance (e.g., internal audit scrutiny) need to be clearly defined, and concerns regarding scalability and the timeline for implementing new policies must be addressed.
3. **Education and Awareness:** Individuals often do not understand or are unaware of the necessity for effective data governance and its critical role in the company's success. Leadership messaging sometimes lacks clarity or is not emphasized effectively, resulting in the misconception that data governance is a blocker to progress rather than an enabler.
4. **Operational and Legal Hurdles:** Financial institutions confront a complex array of global regulations that impose restrictions on data, including who can access it, where it can be stored, what consents are required from stakeholders, how and with whom it can be shared, and what measures must be in place to protect it from unauthorized access. Additionally, some legacy systems contain internal data and workflows that are inconsistent with the new data governance processes.

Overcoming Barriers to Adoption

It is essential to recognize that many organizations have struggled to adopt new policies and processes across a wide range of subjects, as the volume of regulatory-driven change in the financial services industry has been significantly higher since the global financial crisis. Some of these changes have not succeeded, creating the perception that the next new policy is merely another "down from the top" mandate that does not adequately consider the impact on the organization's ability to meet its core objectives. Overcoming institutional resistance to change requires a thoughtful and deliberate approach, including:

1. **Demonstrating Value:** Sell the outcome, not just the remedy. Emphasize use cases demonstrating tangible commercial advantages (e.g., AI readiness) to promote adoption and ensure that governance-driven efficiencies are evident and measurable.
2. **Training and Communication:** Educate teams on governance principles, showcase successful applications of high-quality data, and shift the focus from perfection to actionable improvements.
3. **Building Scalable Frameworks:** Create processes that ensure compliance while being practical for businesses to adopt. Embed data governance into existing processes, such as new product approvals, client communications, and regulatory reporting.
4. **Leveraging AI and Tools:** Position AI as a catalyst for proper data usage while addressing the risks of misusing it as a shortcut for inadequate data. Foster adoption through tools that seamlessly integrate governance into workflows. Utilize industry solutions or tools already available to accelerate the time-to-market.
5. **Drive Accountability:** Integrate data management responsibility at every level of the organization.

Where do we go From Here?

We began this exercise by discussing the challenges posed by incomplete data, which ultimately highlights the need for strong and effective data governance. Governance is an interconnected system of processes, functions, structures, rules, and norms designed to:

1. Establish boundaries of acceptable conduct and practices
2. Control decision-making through rules, policies, and enforcement mechanisms
3. Manage and allocate resources efficiently
4. Provide strategic direction to address collective organizational challenges.

The level of formality and rigor within a governance framework depends on the criticality of the data and the organization's risk appetite. Data has evolved into one of the most critical assets driving performance and innovation in the financial services industry. This growing reliance on data necessitates more formal and rigorous governance processes - an essential evolution that, while beneficial, may initially seem restrictive and face resistance.

Ultimately, each institution must decide how to integrate the tools, techniques, and strategies discussed by the working group to enhance its data governance processes and outcomes.

A Practical Path to Success

It is important to recognize that we are not starting from scratch. Every organization governs its data to some extent, and many existing activities, processes, and tools provide value and lay a foundation for the future state. Transitioning to a more effective data governance process should focus on incremental improvements that build upon what works while introducing innovative concepts and practices.

The business needs to see and feel the results and not concern itself with the what and the why. Trying to sell the idea of governance as a standalone virtue is not an effective way to motivate the rank and file to embrace the new approach.

The key is to demonstrate tangible results. Start small by identifying a specific group or function that is struggling with the current approach. Empower them with new tools and processes to address their pain points, creating visible success stories.

These early wins will build momentum and encourage broader adoption. Over time, as more business units embrace the new approach, the organization's culture will begin to shift toward prioritizing data governance and completeness.

Finally, keep in mind that governance is a continuous journey. Organizations are dynamic and constantly evolving to address new challenges and demands for more data to drive growth, manage risks, and satisfy stakeholder expectations. By embracing governance as an ongoing process, organizations can adapt and prosper in an increasingly data-driven world.

The Global Economics Group

The Global Economics Group is a leading advisory firm that provides customized solutions for today's financial services industry, focusing on risk management, regulatory compliance, and operational consulting.

Using deep expertise and advanced technology, they offer a variety of innovative products. Their flagship platform, RiskPulse, uses AI agents to automate and enhance financial crimes compliance, KYC onboarding, sanctions screening, and transaction investigations—making workflows more efficient while ensuring strong policy enforcement and regulatory compliance.

FactorIQ, a tailored analytics platform driven by AI/ML and Microsoft cloud technologies, provides real-time, centralized insights into financial and operational performance across portfolios, supporting better decision-making and resource allocation for private companies and their investors.

The RCSA Accelerator applies a digital twin approach to transform Risk and Control Self-Assessment (RCSA) programs, allowing financial institutions to simulate scenarios, uncover hidden vulnerabilities, automate documentation, and proactively manage risk to improve business performance.

This comprehensive portfolio highlights Global Economics Group's dedication to delivering actionable insights, efficient compliance, and robust, technology-driven risk solutions to its clients.

www.globaleconomicsgroup.com

Global Economics GROUP

Contact

Maurice Evlyn-Bufton
CEO Armstrong Wolfe
maurice.evlyn-bufton@armstrongwolfe.com

Chris Rigg
Armstrong Wolfe Alliance Partner
crigg@globaleconomicsgroup.com

Find us on LinkedIn: Armstrong Wolfe

Find us on LinkedIn: Women in the COO Community

Global Economics GROUP

ARMSTRONG WOLFE™