



# Global Data Utility - Evolving the Central Counterparty for Data Management (CCDM)

A Research Note by Financial InterGroup



[www.financialintergroup.com](http://www.financialintergroup.com)

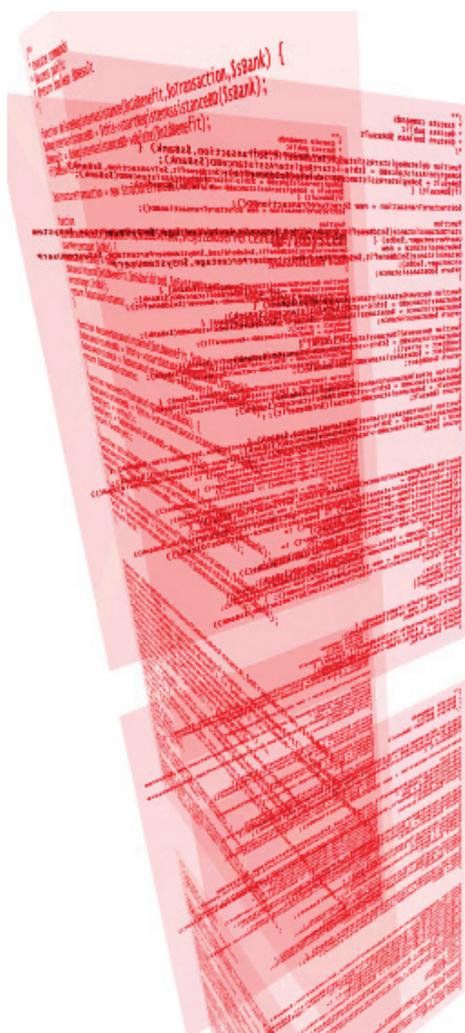
# Contents

The Issue.....	4
Lehman Brothers.....	4
Definition of Reference Data.....	5
Definition of Financial Transaction.....	5
Consequences of Getting it Wrong.....	5
Costs.....	6
History.....	6
BASIC Committee.....	6
Group of Thirty.....	7
Securities Glossary Project.....	8
T+1.....	8
Industry and Government Leadership .....	8
Solution.....	10
Clearing and Settling Reference Data.....	10
Arbitrating Multiple Sources.....	10
Role of Financial Stability Board (FSB).....	10
Role of the Office of Financial Research (OFR).....	11
The Role of the GLEIS.....	11
The U3 Global Identification Coding System (the Barcodes of Finance).....	11
Benefits.....	12
CCDM - A Global Solution.....	12
Multiple Reference Data Utilities.....	12
Input at Source.....	14
Cost Savings.....	16
Getting There From Here.....	16
Data Maturity Progression.....	16
Single Source of Digitized Data and Nomenclature.....	17
Technical Model - a Virtual Database and Federated Network (the Blockchain?).....	18
In Regulators Own words.....	19
Financial Intergroup’s Principals.....	21
Contact.....	22

## Figures

Fig 1 Reference Data Challenge.....	4
Fig 2 The Risk of Getting DataWrong.....	6
Fig 3 Proprietary Data Identities.....	7
Fig 4 The CCDM Utility.....	12
Fig 5 Multiple Reference Data Utilities.....	14
Fig 6 XBRL Message Format.....	15
Fig 7 Data Maturity Progression.....	16
Fig 8 Single Source of Identity and Reference Data.....	19

# Introducing the CCDM, the financial industry's globally shared facility administered by a global standards setting body to assure the completeness, accuracy, security and accessibility of legal, client, instrument, contract and transaction identities and their associated reference data



## Definition of the CCDM

### The CCDM

#### A CENTRAL COUNTERPARTY for DATA MANAGEMENT

*An industry sponsored, government regulated globally shared standards setting body operating a data standards facility. The CCDM maintains a virtualized distributed database connected through an intelligent secure federated network acting as a global repository of standard datasets for use in financial transactions.*

*The CCDM organizes and maintains standard product (Unique Product Identifier – UPI) and legal entity (Legal Entity Identifier – LEI) codes; transaction identification calculation algorithms; and data tags and reference datasets that are used in: reporting financial transactions to regulators; in aggregating financial transactions; and in assuring settlement amongst counterparties.*

*The CCDM vets the underlying transactions' identifiers and reference data for the contracts, instruments, and legal entities at its source, through trusted third parties.*

*The CCDM, by providing identical datasets for assuring the matching of buyer to seller in financial transactions acts as a golden copy of identification and reference data for all manner of unique, unambiguous and universally identified and tagged datasets for purposes of data aggregation of financial transactions and their transfer of value.*

*The CCDM is the global center for managing and overseeing established and new data standards. The resultant datasets are registered and catalogued with the CCDM.*

# The Issue

*“The difficulty lays not so much in developing new ideas as in escaping old ones.”*

John Maynard Keynes

*“We can easily forgive a child who is afraid of the dark; the real tragedy is when men are afraid of the light.”*

Plato

One of the longest standing and intractable impediments to global straight-through-processing in the financial services industry has been the proprietary and non-standard nature of reference data. Regulators were made aware of this while rummaging through the collapsed ‘basement’ of Lehman Brothers and trying to determine who was Lehman, what was its financial exposure and who had been put at risk by Lehman’s bankruptcy.

## Lehman Brothers

Many and varied computer codes were found to be describing the same Lehman Brothers business entity. There were no unique, unambiguous, universal codes to define Lehman the parent or its thousands of legal entities, nor its counterparties, nor the products it traded in. It is now a regulatory priority to fix this ‘plumbing’.

### How Did You Understand Your Exposure to Lehman Brothers and When Did You Know It?

#### A Real World Example of the Reference Data Challenge

What Was Your Exposure to Lehman Brothers?	What Was Lehman Brothers Relation to You?	
Unique Securities Issues Outstanding	Bond Indenture Trustee	General Partner
Lehman Brothers Bank, FS	Commodity Trading Adviser	Investment Adviser
Lehman Brothers Finance SA	Counterparty	Index Vendor
Lehman Brothers Treasury Co.BV	Custodian	Limited Partner
Lehman Brothers Holdings Inc.	Collateral Depot Agent	Market Maker
Neuberger, Inc.	Dealer	Prime Broker
	Depository Agent	Reference Entity
	Escrow Agent	Real Estate Manager
Unique Issuing Identities	Fiduciary	Syndicate Manager
Subsidiary Issuers	Floor Broker	Underwriter
	Futures Commission	
	Merchant	

Fig 1 - Reference Data Challenge

## Definition of Reference Data

Reference data is a broad term understood by operating management, information technology professionals, and risk managers alike. Unfortunately each group understands it differently. To the information professional reference data is “any kind of data used solely to categorize other data found in a database or solely for relating data in a database to information beyond the boundaries of an enterprise”. This concept of reference data is usually referred to as Metadata. To the risk manager it is “...internal and external (third party) data that is used to establish the underlying criteria from which credit risk analysis is performed and credit risk exposure is modeled.” To operating management reference data is information that enables financial transactions to be identified and processed and financial information to be internally and externally reported. To the compliance officer it is data that allows a financial institution to provide assurances to its regulatory overseers that they are complying with documentation requirements for oversight, money laundering and know-your-customer legislation. It is, therefore, no wonder then, owing to basic misunderstandings of this core definition, that risk managers, compliance officers, operating management and information technology professionals have not yet agreed on how to resolve one of the major flaws in the infrastructure of the financial system.

Today’s automated financial markets require data elements of a financial transaction to be accurate throughout the transaction’s lifecycle, from on-boarding to order placement, to reporting of transactions and accumulated positions to clients and regulators. The ability to externally match counterparty information and internally manage financial, business and risk performance depends on its accuracy. The consequence of inaccurate or untimely data manifests itself in significant operational overhead costs, both in individual firms and collectively in the interconnected global financial system. It can be seen in the inability to aggregate risk data and in billions of dollars in regulatory fines for violating money laundering and tax reporting requirements.

## Definition of Financial Transaction

Financial transactions can be thought of as a set of computer encoded data elements that: collectively represent standard reference data, identifying the transaction (the Unique Transaction Identifier - UTI) as a specific instrument or contract (Unique Product Identifier – UPI) bought by a specific business entity (Legal entity Identifier – LEI); variable transaction data such as quantity and amount; and other associated referential information such as price data, credit ratings and other types of fundamental data. Analogous to specific component items of a manufactured product, reference data also defines the products’ changing specifications (periodic or event driven corporate actions such as mergers, acquisitions and spin-offs), occasional changes to sub-components (calendar data, reset dates, credit ratings, historical prices, betas, correlations, volatilities) and seasonal incentives or promotions (dividends, capital distributions and interest payments).

## Consequences of Getting it Wrong

Reference data about a client or product should be consistent across each financial transaction’s life cycle and throughout the financial supply chain. Reference data about changes to either should similarly be consistent across the financial supply chain. However, poor quality and duplication of reference data is pervasive: in large financial enterprises; with data vendors that supply proprietary codes and associated data; amongst financial market utilities and throughout the industry, leading to significantly higher risk and operational costs. When identification codes and reference data that should be identical are not, it causes miscalculated

values, misidentified products and counterparties, and involvement with multiple supply chain partners (trade repositories, custodians, paying agents, et al) to resolve the problem. Inappropriate transactions and individual transaction failures cause monetary loss, higher labor costs, fines and the potential for systemic failure.

## Costs

Standardizing on a common dataset of identifying reference data, variable transaction data, and corporate actions and defining those data elements in a common tagging language would solve some long standing problems for the financial industry: systemic risk caused by mismatched counterparty transaction failures; redundant costs for sourcing, on-boarding and maintaining the fairly static referential data that comprise 70% of a financial transaction; unnecessary costs for reconciling, mapping, transforming and securing this data; and failures from improperly and inconsistently aggregating data for reporting of performance and risk, both internally and for regulatory purposes. In the end, it would save \$2 billion annually for each of the largest financial institutions and mitigate fines that have to date reached as high \$9 billion for a single institution.

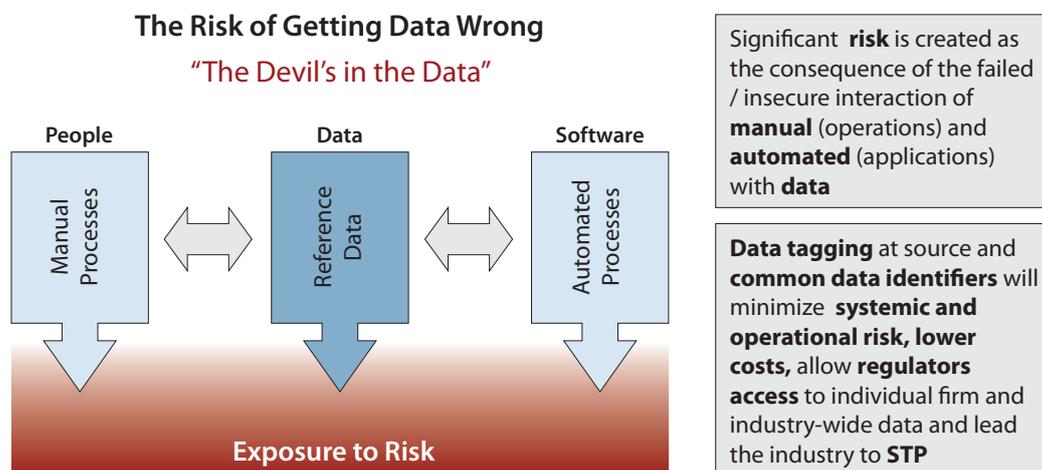


Fig 2 - The Risk of Getting Data Wrong

## History

This issue has been with us since the late 1960s when a working group of executives known as the Banking and Securities Industry Committee (BASIC) was established.

### BASIC Committee

The BASIC Committee was born in the aftermath of the paper crisis that crushed the industry and caused the NYSE to close one day each week to process the accumulated paperwork. This committee championed a single standard for identifying securities for use in computer processing, known as the CUSIP number (Committee on Uniform Securities Identification Procedures); identified uniform numbers for supply chain participants such as transfer agents and brokerage firms, referred to as FINS (Financial Institution Numbering System); and standardized forms used in processing trades from their origination as an order, to their transfer, delivery and payment status. Along the way, in order to facilitate computerized processing, BASIC attempted to standardize the troublesome physical stock certificate by proposing its reincarnation in the form of an engraved punched card, the preferred computer medium of that era.

The punched card concept soon gave way to more efficient methods that immobilized and dematerialized the stock certificate in depositories, thus eliminating the need to make physical delivery but necessitating accurate reference data to assure the electronic completion of a transaction, previously done through physical delivery and site inspection. The establishment of multiple data processing installations throughout the supply chain of banks, brokers, custodians, depositories, investment managers, et al led each to establish their own reference data master files. The multiple representations of this same data led to each firm being burdened with similar costs for acquiring and maintaining data that should be identical but wasn't due to multiple sources of the original textual and/or digital data, clerical input errors, multiple interpretations and/or misinterpretations of the same information, coding errors, and multiple proprietary codes supplied by data vendors. See sample of such codes below:

## How the Computer Knows a Business Entity

There's Got to be a Better Way!

BERKSHIRE HATHAWAY INC	
Data Identity:	
Service (partial list)	Identifier
Cusip Issuer	084670
S&P Rating	100264
Compustat Issuer	002176
Dun & Bradstreet	001024314
Edgar Online	0001067983
Red/Clip CDS	08CAD7
Fitch Rating	80090742
Telekurs	20823

Fig 3 - Proprietary Data Identities

## Group of Thirty

These issues and regulators' responses were taken up by the Group of Thirty (G30) in their study of the 1987 global market disruption. That study resulted in the recognition of an interconnected global financial system, the need to shorten the trade date to payment date cycle, and a need for global standards of identification and associated reference data. In 2006, after twenty years of monitoring, the G30 study concluded that the issue had not been resolved. The financial crisis began in 2007 and reached its catalytic event in 2008 with the Lehman failure. It was at that point that regulators recognized that they could not wait for the industry to fix its own plumbing. Regulatory compulsion was applauded and an integrated identification system was requested, one which satisfied the industry's desire for straight-through-processing and regulators' need to observe accumulating systemic risks.

*"The implementation of reference data standards has proven difficult. With no global owner of reference data and friction between the needs of the domestic and cross-border market users, progress has been slow. Future progress will require greater efforts by market infrastructure operators and international institutions with global reach."*

Group of Thirty Final Monitoring Report  
Global Clearing & Settlement Committee  
May 22, 2006

## Securities Glossary Project

In the intervening years the industry further globalized and expanded beyond equities and fixed income securities to a myriad of financial products, distinct from each other, collections of one and another, and derived from each other. In 1993 the solution to the reference data problem was presented as the Securities Glossary project, a vision to standardize and centralize reference data in a central Securities Master file. The impetus for this was the high cost that each firm was enduring in acquiring and maintaining the same set of data. Also important was the bankruptcy in 1990 of brokerage firm Drexel Burnham Lambert, which exposed firms to having valued differently the same securities issued by Drexel, each having sourced and used different close-of-day valuation prices.

## T+1

Then in 2000 the industry again recognized the importance of reference data in realizing the T+1 vision. Many technology companies (Accenture, Asset Control, Capco, Counterparty Link, Golden Source, IBM and SmartStream amongst others) attempted to develop Reference Data Utilities, primarily around securities data. All these attempts were part of implementing the mantra of "straight-through-processing", the financial industry's still unrealized vision of a seamless electronically connected realtime infrastructure where risk is minimized and operational efficiency maximized. With few exceptions these efforts as originally conceived failed to attract a critical mass of clients primarily because the data vendors were reluctant to have their data aggregated with each other so that they can be compared on quality. They were also reluctant to discount their data costs to a shared utility that would compromise their one-off sales to individual firms and to the silos within those firms.

## Industry and Government Leadership

The calls for solving this problem persist to this day, with thought leaders and regulators proposing solutions, both to resolve product reference data and, more recently, business entity reference data.

*"..now is the time to begin a serious dialogue on a collaborative solution to the securities industry's reference data problem... The cost to move forward will be significant, but with the securities business growing evermore complex, the cost of inactivity would be far greater."*

Tower Group  
Dec, 2007

*"The Governing Council of the European Central Bank (ECB) is today issuing a proposal to all European central securities depositories (CSDs) to join the TARGET2-Securities (T2S) initiative. ...T2S offers a single securities accounts reference data model for all connected CSDs."*

European Central Bank  
"CSDs invited to join the  
TARGET2-Securities initiative"  
Press Release May 23, 2008

*“CCDM has suggested an index of data in the cloud, providing a mechanism for regulators to access the data linking to the cloud and allowing them to query and search data they want to look at, at the time they want to look at it”*

Financial InterGroup  
Waters Magazine  
“CCDM Proposes Utility in the Cloud”  
Dec 7, 2010

---

*“Together, we have created a Central Data Utility (CDU) in order to offer a service that mutualizes the cleansing of reference data”*

Euroclear and Smart Stream  
Brochure 2012

---

*“Our ultimate aim is to support the industry’s call for a comprehensive, centralized platform to effectively manage virtually all client reference data”*

DTCC  
“DTCC to partner with global banks to develop comprehensive reference data service”  
Press Release  
Sept 30, 2013

---

*“I can’t tell you how many times over my long banking career I’ve had a client say, ‘We already provided that information’, or ‘I went to another bank and provided it one place and then they turned around and asked it somewhere else [in the same institution]’ ”*

Stephanie Wolf,  
Managing Director, global transaction services business,  
Bank of America  
American Banker: “The Race to Build a Know-Your-Customer Registry”  
Dec 18, 2014

---

*“J.P. Morgan Chase & Co., Goldman Sachs Group and Morgan Stanley are working to create a company that will pull together and clean reams of reference data at a lower cost than what they would spend individually, according to people familiar with the matter. The new entity, which will create a stream of consistent data that banks use to help determine pricing and transaction costs, is the latest example of increasingly cost-conscious banks coming together to save on head count, expenses and time.”*

*“The initiative is currently dubbed “SPReD”, which stands for Securities Product Reference Data, and is likely to be launched as a new entity in the next six to 12 months, the people said. Each founding bank is investing “seven figures” for the entity, the people said. The company will work specifically with reference data on financial instruments, including identifiers like names, codes and symbols that each institution already buys. It will start with listed derivatives and equity data, with fixed income-related data added later.”*

Wall Street Journal  
“J.P. Morgan, Goldman Sachs,  
Morgan Stanley to Form Data Company”  
Aug. 19, 2015

---

## Solution

To solve this long standing problem the first step is to vet identifiers and their associated reference data at the earliest point in the data life cycle, at the source of its creation and, thereafter, at the on-boarding stages.

### Clearing and Settling Reference Data

Two-sided matching concepts that traditionally have been applied to validate valued transaction data in settlement systems and central counterparties should be applied to the matching, "clearing" and "settlement" of other reference data (prices, currency rates, codes for product categories, market codes, etc.) at the pre-trade financial transaction assembly point. By doing so, at the immediate frontend of a financial transaction's lifecycle, the post trade payment and regulatory reporting environment can be improved significantly.

To date mutualized risk sharing within matching, clearance and settlement systems has only been applied to the value portion of transactions (principally quantities, transaction prices and currency values). These same techniques, however, can be applied to the reference data components of these transactions. While reference data is not value-bearing, computer matching algorithms do not differentiate reference data from other data and mismatches and failed transactions occur regardless of the data elements' business intent. Acquiring, maintaining and managing such data is costly with faulty data being at the core of significant components of operational losses. It can be argued that one of the core processes of financial market utilities is to match disparate datasets to allow interoperability across the interconnected financial system, an activity that could be eliminated if a single golden copy of data was available.

### Arbitrating Multiple Sources

A Central Counterparty for Data Management (CCDM) would organize and maintain identification codes and their associated reference data, match multiple incoming original sources of other reference data, "clear" this data through best-of-breed computer analysis, and "settle" (distribute) industry accepted, CCDM assured datasets to primary participants and, in turn, to their downstream correspondents. This would entail an industry-wide effort not dissimilar to the clearing entities, netting systems and central depositories that emerged as industry-wide solutions to past industry-wide problems. Leading this effort could well be the largest of financial enterprises, now identified by regulators as systemically important financial institutions (SIFIs) given to first satisfying their own collective needs as they currently absorb the most cost and risk and are required to set aside additional capital under new capital guidelines (Basel III).

### Role of Financial Stability Board (FSB)

This approach is made simpler now that there is a global institution, the G20's Financial Stability Board that is directing infrastructure projects through member states' regulatory institutions. The first such project, the Global Legal Entity Identifier (LEI) System (GLEIS) will provide a unique identifier and 'business card' reference data for counterparties and other financial market participants in the swaps supply chain. Thereafter, all financial market participants that can enter into financial contracts for any financial product are to register for a LEI. Other projects currently underway are the UPI (Unique Product Identifier) and UTI (Unique Transaction Identifier – UTI). Again, initially for swaps contracts and transactions, but eventually for all financial products and all financial transactions.

## Role of the Office of Financial Research (OFR)

Since the financial crisis regulators came to realize they too needed such financial entity and financial product 'catalogues' (the Dodd-Frank legislation requires they be maintained by US Treasury's Office of Financial Research). A 'business entity' catalogue was left out of the legislation but came later as a private sector initiative in cooperation with regulators. The GLEIS is the outcome of that cooperation.

Beyond the labeling of legal entities in the swaps markets, the codes would also have to be useful for data aggregation and systemic risk. Here multiple codes pertaining to a single business entity (i.e. an ultimate parent or ultimate controlling entity) would need to be rolled up. It is yet to be demonstrated that this is doable within real-time time tolerances and without creating huge mapping tables that introduce time delays, reconciliation issues, synchronization issues and additional costs.

## The Role of the GLEIS

Technically, the GLEIS is to be a federated intelligent network, federating multiple LEI registries domiciled in the sovereign domain of each participating country. The intention of the CCDM is to utilize the federated network of the GLEIS to embed existing proprietary business entity code maps into the networks' servers and routers to permit in-network code mapping to and from the LEI identifier code. However, the LEI code in its current form will introduce another layer of mapping – firms will have 100's, 1000's, some over 10,000 legal entities. This outcome is the opposite of the objectives for the LEI – eliminating mapping over time.

## The U3 Global Identification Coding System (the Barcodes of Finance)

To solve this problem FIG has developed the U3 (unique, universal and unambiguous) identification code so that the pre-LEI can be transitioned to a more efficient code construction for global data aggregation and internal use. The U3 codes have been developed as a common code standard that is extensible for product, transaction and corporate event identification. This code construction is the subject of continuing discussions with regulators and industry members. Research on the code construction for the LEI is available - see "Gateway to the Barcodes of Finance - The Global Legal Entity Identifier (LEI) Implementation" at <http://sales.financialintergroup.com>.

# Benefits

## CCDM - A Global Solution

The benefit of a single CCDM - a global 'central counterparty' for setting data standards, for example for the LEI and UPI and maintaining their reference data (i.e. one "golden copy") is transformational for both industry and regulators. The cost savings of a single virtual database distributed as nodes across a peer-to-peer network vs. multiple golden copies, one centralized data warehouse each for each firm (the Enterprise Data Management or EDM model), or multiple ones shared by multiple firms/facilities in multiple outsourced facilities is obvious – we have estimated those savings at \$50-\$100 billion annually for the 100 largest Global and Domestic SIFIs. The need to standardize the calculation method of an unambiguous and universal transaction identifier (UTI) should, likewise, be obvious.

Both the common code standard and the CCDM can be organized as a government/private sector initiative at the global level. With modifications the CCDM is not unlike the combination of the GLEIS, a free open standard and data utility for all and Omgeo's regulatory structure, which operates in the post trade equity markets as an exempt clearing company, now owned by DTCC under US regulation but, in turn, is owned by its largest contributing industry members.

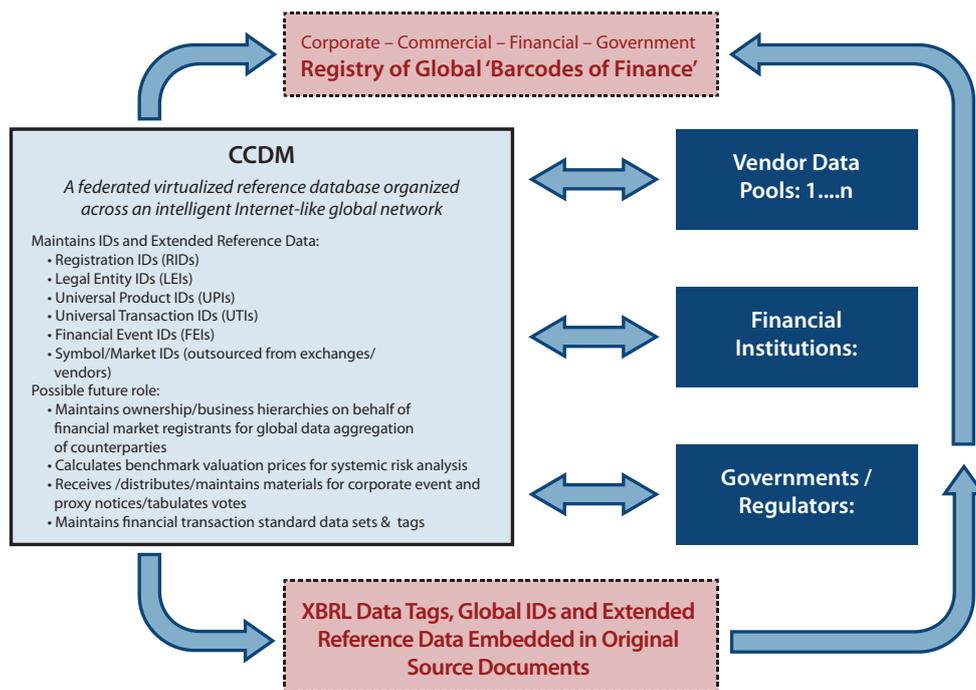


Fig 4 - The CCDM Utility

## Multiple Reference Data Utilities

Multiply sourced, multiple copies of these golden copies cannot solve the straight-through-processing problem, even when all are using the same transmission standards, standard data tags or content standards or when everyone has one of their own golden copies in their own firms or in each central securities depository or clearing facility, or in collective facilities that serve multiple firms.

Financial transactions would not match more than occasionally, within swaps data repositories and within the global payment, clearance and settlement system; collateral would still be

valued differently at times; and customers and traders would from time-to-time be improperly notified of corporate events, if at all, with monies received or positions adjusted incorrectly. Mappings and data transformations would still be necessary and the straight-through-processing vision would remain unrealized.

Regulatory reporting would be inconsistent with different hierarchical constructions of legal entities caused by different constructions subject to financial firm and data vendor interpretations.

Mistiming of changes to product and legal entity data would result in different product or legal entity codes and differences in reference data, resulting in deterioration of data quality over time. Systemic risk and excessive cost would still be built into the industry's infrastructure due to the still unmitigated risk and duplicated costs.

To summarize, multiple versions of identification and reference datasets, whether in central data warehouses of each financial firm, or available from multiple outsourced facilities will be ineffective because of:

- ▶ the limited availability of budgets to source data from multiple vendors;
- ▶ different vendors chosen for each firm or existing infrastructure facility thus imbedding a variance in the datasets maintained by each firm and each outsourced facility;
- ▶ each firm/facility with different rules for accepting "best-of-breed" data;
- ▶ duplicated activities and costs for each firm/facility essentially trying to do the same thing;
- ▶ regulators and firms still dealing with faulty definitions of aggregated risk for a counterparty whose hierarchies and definitions of business entities are determined separately by each firm/vendor;
- ▶ firms still only finding out data faults when they try to send a transaction through its settlement process and it fails to complete;
- ▶ the industry still lacking the ability to accommodate STP in any time frame approximating trade date settlement, let alone real-time settlement;
- ▶ regulators still rejecting electronically filed regulatory reports because they couldn't match incoming data sent electronically from firms to regulators' databases; and
- ▶ regulators accepting electronically filed reports because they did match incoming data from firms, but the regulators' databases had different meanings (descriptions of business entities, instrument identities, data attributes, etc.) for the matched data elements.

Notwithstanding this, multiple industry led initiatives are being pursued, many focused on the data and documentation required for legal entities: under various money laundering and know your customer regulations; under new and active due diligence reporting in taxing jurisdictions; and in legal entity identification requirements under new derivatives regulations. The clients of financial institutions submit a single set of identity documents to either the financial firms themselves or directly to any of the shared registries and can then make them available in digitized form to all of their authorized counterparties. Other industry initiatives are focused on issued trading instruments and their price and corporate event data. Below is a list of a number of shared data utility initiatives already announced:

Utility	Description
<b>SPReD</b>	A shared utility for securities and listed futures reference data organized by J.P. Morgan Chase & Co., Goldman Sachs Group and Morgan Stanley. Facilities operator is SmartStream
<b>Euroclear and SmartStream</b>	A central utility for securities reference data
<b>ClarientGlobal LLC</b>	A DTCC company that will offer centralized know-your-customer and reference data for clients and legal entities founded with BNY Mellon, Barclays, Credit Suisse, Goldman Sachs, JPMorgan Chase and State Street
<b>Know Your Customer (KYC) Registry</b>	A centralized repository sponsored by SWIFT with Barclays, Deutsche Bank, Erste Group Bank AG, HSBC, ING, Raiffeisen Bank International AG, Bank of America, Merrill Lynch, Citi, Commerzbank, J.P.Morgan, Societe Generale and Standard Chartered
<b>Accelus Org ID</b>	A KYC register sponsored by Thomson Reuters with TradeWeb and FXAll
<b>Markit joint venture with Genpact</b>	...with Citi, Morgan Stanley, Deutsche Bank and HSBC
<b>KYC Net Exchange AG</b>	...with DZ Bank, Commerzbank, Société Générale and Standard Chartered
<b>KYCme</b>	A Jersey (Channel Island) based shared on boarding documentation vault
<b>Central Counterparty for Data Management (CCDM)</b>	An ID standards body and reference data utility to be owned by all industry members

Fig 5 - Multiple Reference Data Utilities

## Input at Source

The industry is already embarked on rationalizing content standards such as instrument and contract codes (the UPI) and business entity identifiers (the LEI) and implementing EDM and outsourcing models. These current efforts will make it much easier to evolve to the final phase, the CCDM model of applying the central counterparty concept to data management globally. Here, the focus is both on cost efficiencies and risk mitigation, as well as the recognition that the financial industry has become global, transcending sovereign state regulations and, even, regional regulatory compacts.

The problem for regulators is similar as they too have to properly identify the business entities, the products traded in, the corporate events being reported. They are dependent on corporations and financial institutions to present information in identical form, at least for the referential portion of the data. Given that each financial firm has sourced, interpreted and applied reference data differently, the aggregation of this information to report to each regulator, as well as in reporting within one's own financial institution can be different.

Regulators throughout the world are now requiring electronic transmission of issuer financial statements, underwriter's prospectuses, and other submissions, and, along with this, the further innovation of "at source data tagging" (essentially surrounding data with computer readable tags). Already the SEC has mandated such tagging activities and other regulators as well.

Tagged Documents and Data Examples: *More work for Computers - Less for People*

XML Message Format*	XBRL Message Format*
<pre> &lt;BusinessEntity&gt; &lt;Type&gt;Corporation&lt;/Type&gt; &lt;Name&gt;"Berkshire Hathaway"&lt;/Name&gt; &lt;Identificationnumber&gt; 999-99-9999 &lt;/Identificationnumber&gt; &lt;Addressline1&gt;1313 Terrace oad&lt;/Addressline1&gt; &lt;Cityormunicipality&gt;Omaha&lt;/Cityormunicipality&gt; &lt;Stateorprovince&gt;NE&lt;/Stateorprovince&gt; &lt;Ziporpostalcode&gt;12345&lt;/Ziporpostalcode&gt; &lt;/BusinessEntity&gt;  *FIXML Trade Transaction format                     </pre>	<pre> &lt;/element&gt; &lt;element id="usfr-seccert_PrincipalOfficerName" name="PrincipalOfficerName" type="xbri:stringItemType" substitutionGroup="xbri:item" xbri:periodType="duration" nillable="true" /&gt; &lt;element id="usfr-seccert_PrincipalOfficerSignature" name="PrincipalOfficerSignature" type="xbri:stringItemType" substitutionGroup="xbri:item" xbri:periodType="duration" nillable="true" /&gt; &lt;element id="usfr-seccert_PrincipalOfficerTitlePosition" name="PrincipalOfficerTitlePosition" type="xbri:stringItemType" substitutionGroup="xbri:item" xbri:periodType="duration" nillable="true" /&gt; &lt;element id="usfr-seccert_ReportAccounting Changes" name="ReportAccountingChanges" type="xbri:stringItemType" substitutionGroup="xbri:item" xbri:periodType="duration" nillable="true" /&gt; &lt;/schema&gt;  *From SEC Filing Format                     </pre>

**Fig 6 - XBRL Message Format**

The SEC’s initiative, the Interactive Data Electronic Application computerized “tags,” are similar in function to barcodes, which identify individual items in a company’s financial disclosures. With every number on an income statement or balance sheet, and every data item in mutual fund filings and offering prospectuses individually labeled, information can be easily accessed on the Internet, downloaded into spreadsheets, reorganized in databases, and put to any number of other comparative and analytical uses by investors, analysts, journalists, and financial intermediaries. (The SEC’s final rules on Interactive Data reporting went into effect October 31, 2014.)

Importantly, by requiring such tags to surround the business identity and security instrument or contract codes, and prescribing that other information normally described in text-only press releases and word documents, such as corporate event

notifications, be presented in standard format, we begin to set the stage for retooling the infrastructure of the financial services industry. Setting standards where proprietary and conflicting identification codes now exist across the entire range of referential data, including such fundamental identifiers as symbols for corporate issuers and their issues; symbols used in contract markets such as options and derivatives; and numbering conventions for securities, supply chain business entities, and counterparty identifiers, would be a transforming event.

**“The Securities and Exchange Commission voted unanimously Wednesday to propose a rule requiring companies - by as early as next year - to file financial statements in an “interactive data” format. The proposed schedule is a landmark moment for interactive data-tagging, using the system known as XBRL, for extensible business reporting language. Christopher Cox, the SEC chairman, called the development something that would “significantly transform the SEC’s business model,” and compared XBRL’s importance to that of the first personal computers and the requirement that financial statements be published online in the Edgar database.”**

SEC Maps Interactive Data-filing Mandate  
CFO Magazine May 14, 2008

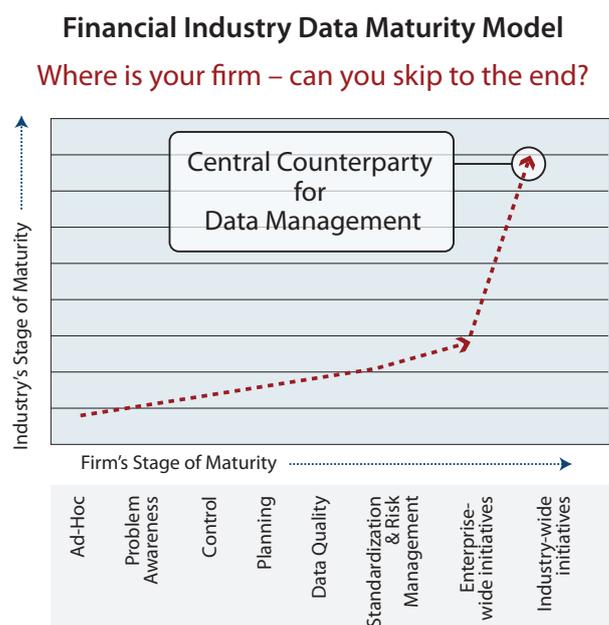
The next step beyond the initial phase of the GLEIS is easy to fathom, even easy to implement, but needs corporate and commercial users' cooperation. First, voluntary corporate events such as mergers and acquisitions, and other reportable items such as dividends that are declared in board meetings, transcribed and typed in text and disseminated in press releases need to be standardized and registered at source. Standard formats and data tags within a finite number of corporate event types have already been codified by a working industry trade group, XBRL International. Standard templates would be filled out by corporate filers and uploaded to the SEC's IDEA database as well as the GLEIS and CCDM for all to retrieve and use in their own ways for their own purposes. Many financial intermediary activities such as retrieving, interpreting, coding, inputting, and transforming this data into proprietary computer formats would be eliminated.

The final step is corporate and commercial entities' commitment to identify and codify corporate event data. The LEI project manager, the Regulatory Oversight Committee (ROC) has a stake in this. Without a standard way of reporting corporate events that change external LEIs the GLEIS will quickly become out of synch with internal codes that are changing to reflect mergers, acquisitions, spin-offs, etc. FIG has developed a financial event identifier (FEI) for identifying at-source corporate events.

## Cost Savings

Most importantly the fulfillment of the straight-through-processing vision would diminish significantly the risk now associated with misidentified items, unreported events, and transposition, transformation and mapping errors. Ultimately these errors find their way into incorrectly updated securities positions, unreported income, failures to adjust traded quantities, and improperly reported performance and risk information to regulators. Just in this simple but transforming way upwards of \$10 billion in annual trading losses could be eliminated toward savings of nearly \$100 billion in annual expenses that financial institutions spend unnecessarily.

## Getting There From Here



## Data Maturity Progression

Over time, probably measured in a decade or two, data vendors, the CCDM and others will be accessing data directly from its originating sources in a completely electronic, standard form, thus eliminating the majority of data sourcing errors the industry now deals with at considerable cost and with significant embedded systemic risk.

Fig 7- Data Maturity Progression

## Single Source of Digitized Data and Nomenclature

In the identification space of the Legal Entity Identifier (LEI), the Unique Product Identifier (UPI) and the Financial Event Identifier (FEI) these efforts all have a single goal of transforming legally drafted definitions of products, business entities, contractual relationships, notices of corporate reorganizations, etc. from paper or word-processed documents into digital form. The originating source of this information is documents – offering memoranda, prospectuses, corporate resolutions, master agreements, collateral agreements, trust agreements, articles of incorporation, word processed documents, etc. It would, therefore, seem reasonable that the preferred method to transform this information into computer readable form is to use the standard of XBRL, the eXtensible Markup Language (XML) for Reports for this transformation. Three quarters of the globe’s regulators already use XBRL to transform other regulatory information reported to them in this way.

In the transactional space data is not created from paper documents. Data is simply typed into or retrieved from a computer in an existing data format. Information such as a price or notional, or a buy or sell indicator, or a reset date, tenor or interest rate, and many other codes and input items are placed into existing computer generated templates. Here such standards as FpML (Financial Product Markup Language) and FixML (Financial Information Exchange Markup Language) are in broad use in the financial industry. Each can be incorporated into XBRL as well as stand apart, depending upon the application.

The biggest challenge is to conform to a common nomenclature, a set of nouns that describe in the smallest number of characters possible what industry members conclude is the best description of the data element the tags describe. This is a task yet to be carried out. It would seem logical to do so under FSB oversight and, where necessary, regulatory mandate to assure conformity. To this end the FIBO (Financial Industry Business Object) language, the most recent attempt at standard tagging nomenclature has shown promise. These efforts should form the basis for a Working Group under FSB oversight to bring finality to a harmonized tagging nomenclature.

However, for these tags to be useful a series of steps must be taken: first communicating seamlessly across business silos within financial enterprises through standard data tags, and business entity and product codes; then between financial businesses and regulators bilaterally; and, finally, communicating within multi-lateral payment, clearing and settlement systems seamlessly in a straight-through-processing environment. The ultimate goal is simultaneous, near real-time trade assembly, execution and finality of payment, settlement and cash flow transfers. Along with this will come a new generation of automated business applications built within each financial firm or in each separate component of the payment, clearance, settlement and regulatory reporting system with the potential of accessing the CCDM network for all manner of standardized referential datasets.

***“Results showed that there remain some significant common challenges to full compliance with the Principles:***

- Banks’ dependence on manual processes;***
- The need to develop common data dictionaries and data taxonomies; and***
- The inability to create accurate and timely risk data reports during stressed or crisis situations.”***

Basel Committee on Banking Supervision (BCBS)  
“Progress in adopting the principles for effective  
risk data aggregation and risk reporting (BCB239)”  
January, 2015

## Technical Model - a Virtual Database and Federated Network (the Blockchain?)

The technical model for the CCDM is an intelligent federated network, a secure virtual private network (VPN) overlaid on the Internet, federating sovereign databases as a single virtual view, geographically distributed, organized across individual firms, regulator sponsored facilities and regional compacts of either or both. The CCDM is to be formed initially by the largest financial institutions as an industry sponsored, government regulated and mutually shared risk mitigating central counterparty facility. This counterparty mechanism has precedent as the industry's proven way of providing assurances to each participant that the use of the datasets from such a facility will be accepted as a faultless standard, both from a regulator's perspective and within the global payment, clearance and settlement mechanism. The responsibility for systemic failures in such a system will be shifted from the tax payers to the private sector.

The private sector, initially financial institutions, will benefit through stripping its own infrastructure of the technology, people and data costs of duplicate reference data and multiple mappings of identifiers. Instead financial institutions will be able to access 'component parts' in the external data management layer of the shared 'parts and supply chain participant catalogues' and build business applications on top of them, rather than incorporating such catalogues in each business application. See diagram in Figure 8 on the following page. This should foster a return to the basics, an industry not anticipating regulators forcing their breakup because of their size or complexity nor forcing their nationalization of our financial institutions. Rather an industry simply getting cooperation from regulators to allow the industry to solve this long standing problem permanently and with its own resources, getting on with reengineering their financial institutions.

This technical model has similarities to the most recent manifestation of technical prowess, the much touted immutable distributed database ledger technology of the Blockchain. All commentators and collaborators, and there are many now in financial circles, are supporting experiments in Blockchain technology. While a diverse set of objectives for first implementations are being considered, they all have one thing in common, a recognition of the needed prerequisite of a universal set of financial product and financial supply chain participant identification standards and associated reference data.

However, these Blockchain visionaries and collaborators are not placing the needed priority on global identifiers and are in denial of the existing mature technologies as described above that can already support their visions. That vision is the displacement of financial infrastructure such as post trade clearing, settlement and payment mechanisms and real-time finality of financial transactions from order placement to posting to digital ledgers.

To this end, a first industry collaboration is needed around the current efforts of the G20's Financial Stability Board to bring unique, universal and unambiguous identification standards into existence. This effort is now bogged down in the one market it is being tested in, the global swaps market. The true test, thereafter, is the global financial industry's willingness to cooperate further around the promise of distributed database technology. This distributed capability exists and was in use long before the Blockchain incorporated such techniques, albeit not in finance.

This technology, in whatever form, can be used to establish the one missing global utility to make all the Blockchain global visions practicable. That utility is the universal product and participant catalogue, what has been described as a golden copy of global identifiers and associated reference data. Along with standard data tags and common datasets that describe financial transactions a distributed ledger utility can be created – in this case as the

Central Counterparty for Data Management - to underpin all subsequent legacy systems and infrastructure reengineering promised by Blockchain visionaries. Without it no consequential global industry transformation can take place as is the collective vision of Blockchain enthusiasts.

### Reengineering the Financial Corporation Around the Externalised Global Identification System and Reference Data Platform

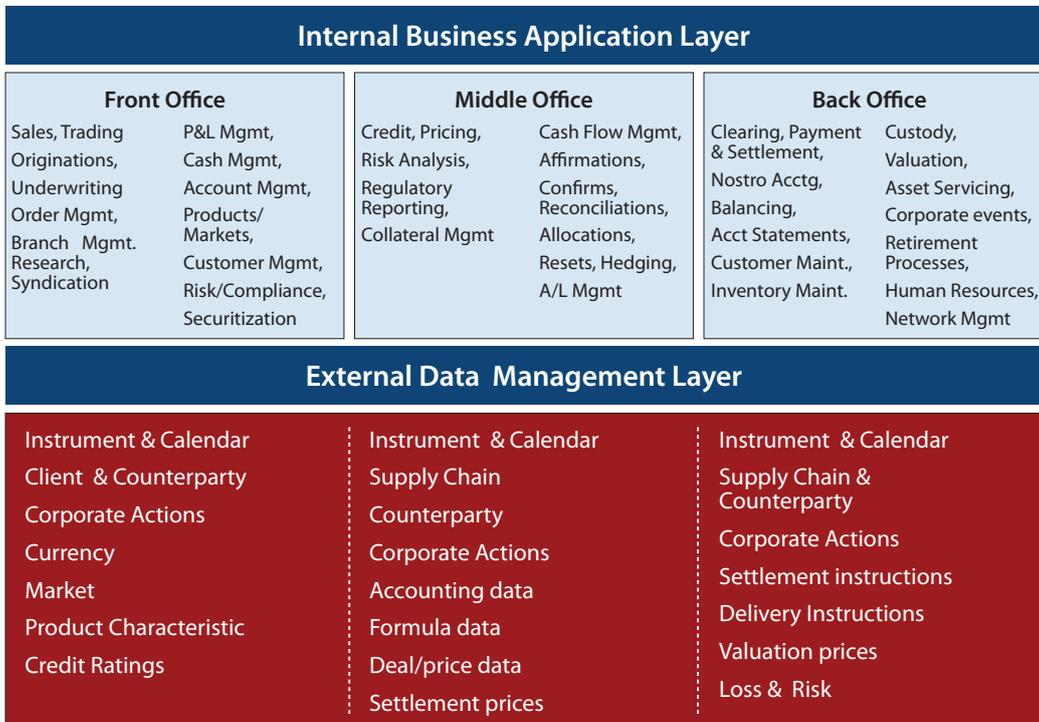


Fig 8 - Single Source of Identity and Reference Data

### In Regulators Own words

The benefit to regulators is obvious, achieving true transparency by being able to finally “see” into the standard data used by all, thus fulfilling their oversight role each are mandated by law to perform on behalf of the public. Here, we believe the governing body for the Central Counterparty for Data Management can play a significant role in coordinating these efforts providing, for the first time, a global center for all data management initiatives.

*“Currently, because the underlying data in firms’ risk-management systems are incomplete or are maintained in non-standardized proprietary formats, compiling industry-wide data on counterparty credit risk or common exposures is a challenge for both firms and supervisors. Further, institutions and investors cannot easily construct fairly basic measures of common risks across firms because they may not disclose sufficient information.”*

Testimony before the Banking Subcommittee  
on Security and International Trade and Finance  
Daniel K. Tarullo,  
Member, Board of Governors,  
Federal Reserve System  
February 12, 2010

***“There is widespread agreement among the public authorities and financial industry participants on the merits of establishing a uniform global system for legal entity identification.”***

A Global Legal Entity Identifier for Financial Markets:  
FSB Report to the G20  
June 8, 2012

---

***“In protecting the broad public interest, the objectives of the ROC are to ensure that the Global LEI System meets broad public and private sector requirements, including ensuring the uniqueness, consistency, exclusivity, accuracy, reliability, timeliness of access, portability, and persistence of the LEI code and reference data”***

Charter of the ROC for the Global LEI System  
November 5, 2012

---

***“Many banks lacked the ability to aggregate risk exposures and identify concentrations quickly and accurately at the bank group level, across business lines and between legal entities. Some banks were unable to manage their risks properly because of weak risk data aggregation capabilities and risk reporting practices. This had severe consequences to the banks themselves and to the stability of the financial system as a whole.”***

Basel Committee on Banking Supervision  
“Principles for Effective Risk Data Aggregation and Risk Reporting”  
January 10, 2013

---

***“Data standards are not an issue to set the pulse racing. And the technical challenges involved in arriving at a global common financial language are nothing short of daunting. But if the financial crisis taught us anything, it is that the prize could not be larger.”***

“Knowledge needed to prevent Lehman repeat-  
Global financial language essential to close dangerous data gaps”  
Andrew Haldane, Chief Economist, Bank of England  
Aurel Schubert, Director-General Statistics, European Central Bank  
Richard Berner, Director, US Office of Financial Research  
Financial Times  
January 14, 2015

---

# Financial InterGroup's Principals



## Allan D. Grody

Allan is the founder of the Financial InterGroup companies. He has been active in the financial industry for nearly five decades and has had hands-on experience in multiple sectors of the financial industry. He advises on domestic (USA) and international issues related to financial institutions' global strategies, restructuring and acquisition needs, information systems, communications infrastructures and risk management systems.

In an earlier career, he was the founder and Partner-in-Charge of Coopers & Lybrand's Financial Services Consulting Practice, which was subsequently merged with Price Waterhouse and eventually sold to IBM. Professor Grody founded and taught the only graduate level Risk Management Systems course at NYU's Stern Graduate School of Business. He is a former founding Board member of the Technology Committee of the Futures Industry Association. He is currently an editorial board member of the Journal of Risk Management in Financial Institutions and a board member of the Blue Ribbon Panel of the Professional Risk Managers International Association. He writes, speaks and advises on issues where data management, risk management and technology converge. He has participated in expert panels sponsored by local and global regulators on these subjects.



## Peter J. Hughes

Peter is a Principal of Financial InterGroup and Managing Director of its UK based company. He is a former country/regional executive with JPMorgan Chase, Fellow of the Institute of Chartered Accountants in England & Wales, a member of the advisory board of Durham University Business School's Banking, Risk & Intermediation research group and a visiting research fellow at the Leeds University Business School.

At Financial InterGroup he leads consulting projects and provides advisory and training services to some of the globe's leading banks, global IT and consulting firms, trade associations and banking institutes with particular emphasis on cross-enterprise risks, operational risk, Basel II & III, capital management (including the Internal Capital Adequacy Assessment Process - ICAAP), finance transformation, accounting (including IFRS), data management, risk measurement and management systems and risk based auditing.

# Contact Us

For more information on Risk Accounting please contact:

▲ **Allan D. Grody**  
**Financial InterGroup - USA**

169 East 69th Street - 18th floor  
New York,  
New York 10021  
USA

Mobile +1 917 414 3608  
Email [agrody@financialintergroup.com](mailto:agrody@financialintergroup.com)

▲ **Peter J. Hughes**  
**Financial InterGroup - UK**

Lambs Green Lane  
Wimborne,  
BH21 3DN  
UK

Tele +44 (0) 1202 842087  
Mobile +44 (0) 7766 916541  
Email [peter.hughes@financialintergroup.com](mailto:peter.hughes@financialintergroup.com)

**Note:**

The CCDM and the U3 Identification System are trademarks of Financial InterGroup (FIG) Holdings Ltd. FIG holds patents and pending patents on the CCDM and U3 methods and systems.

