Journal of the American Society for Information Science and Technology

The Classification of Financial Products

Journal:	*Journal of the American Society for Information Science and Technology*
Manuscript ID:	JASIST-2012-12-0742
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	31-Dec-2012
Complete List of Authors:	Loehrlein, Aaron; University of British Columbia, iSchool Lemieux, Victoria; University of British Columbia, iSchool Bennett, Michael; Enterprise Data Management Council,
Keywords:	faceted classification < classification < organization of information < information operations < (activities and operations), hierarchies < knowledge organization systems < (knowledge and information), finance < (fields and disciplines)

The Classification of Financial Products

Aaron J. Loehrlein[1],Victoria L. Lemieux[1] and Michael Bennett[2]

[1] iSchool, University of British Columbia and [2]Enterprise Data Management Council

Abstract

In the wake of the global financial crisis, the U.S. Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank) was enacted to provide increased transparency in financial markets. In response to the Dodd-Frank act, a series of rules relating to swaps record keeping have been issued and one such rule calls for the creation of a financial products classification system. The manner in which financial products are organized will have a profound effect on data integration and analysis in the financial industry. This article considers various approaches that can be taken in creating hierarchical taxonomies of financial products. It considers the effects that act on these taxonomies, depending on the approaches taken, and how those effects specifically apply to the classification of financial products. The article recommends use of facet analysis in the organization of financial products, as this type of analysis is flexible enough to accommodate multiple viewpoints, but rigorous enough to facilitate inferences that are based on the hierarchical structure. Various use cases are examined that pertain to the organization of financial products. The use cases confirm the practical utility of taxonomies that are designed according to faceted principles.

Introduction and Motivation

The global financial crisis of 2007-2009[1] has highlighted a number of weaknesses in the quality and management of records, information and data leading to an increase in systemic risk. The need for improved data and information was specifically recognized in a Financial Stability Board (FSB) and International Monetary Fund (IMF) report on "The Financial Crisis and Information Gaps" (2009) which noted that, "the recent crisis has reaffirmed an old lesson—good data and good analysis are the lifeblood of effective surveillance and policy responses at both the national and international levels" (FSB/IMF, 2009). Lack of access to information on a timely basis; data quality issues; missing evidence; and poor record keeping have all been implicated in the build up of financial risk and the spread of contagion through global financial markets. Within the international financial community a broad consensus has emerged over the information gaps that need to be filled. Since 2008, international agencies, financial regulators, financial trade associations and market participants have been working to better capture the build-up of risk in the financial sector; to improve data on international financial network connections; to better monitor the vulnerability of domestic economies to shocks; and to improve the communication of official statistics (FSB/IMF, 2009; 2011).

It was with the objective of improving transparency in financial markets that the U.S. Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank), which was signed into law on July 21, 2010, introduced comprehensive regulation to the swaps marketplace under Title VII (Public L. 111-203 H.R. 4173) and amended the Commodity Exchange Act (CEA). In scope of this section of the Dodd-Frank and the CEA are the credit default swaps that were at the center of the global financial crisis of 2007-2009. Among the changes set out in the Act were new requirements to have standardized forms of these financial products trade on open platforms and cleared through central counterparties. Specifically, Section 727 of Dodd-Frank added new section 2(a)(13)(G) to the CEA, which requires all swaps, whether cleared or uncleared, to be reported to swap data repositories (SDRs), which are new registered entities created by section 728 of Dodd-Frank to collect and maintain data related to swap transactions, and to make such data electronically available to regulators.

New section 21(b) of the CEA, added by section 728 of Dodd-Frank, directed the Commodities and Futures Trading Commission (CFTC) to prescribe standards for swap data recordkeeping and reporting. To give effect to Dodd-Frank's provisions, the CFTC engaged in an extensive rulemaking exercise resulting in the Real-Time Public Reporting of Swap Transaction Data Rule (''Final Rule'') which was finalized on January 9, 2012 (77 FR 1182). The Final Rule applies to swap data reporting and recordkeeping for swap data repositories, derivatives clearing organizations, designated contract markets, swap execution facilities, swap dealers, major swap participants, and swap counterparties who are neither swap dealers nor major swap participants (CFTC, 2011b; 2012). In addition, the Final Rules call for use of three unique identifiers in connection with swap data reporting: a unique swap identifier (USI), a legal entity identifier (LEI), and a unique product identifier (UPI). The CFTC proposed the use of these unique identifiers as regulatory tools for linking data together and enabling data aggregation across counterparties, transactions, and asset classes, to fulfill the systemic risk mitigation, market manipulation prevention, and other purposes of Dodd-Frank (CFTC, 2011b; 2012).

This paper focuses on the issue of financial product classification, which is needed to support issuance of a single product identifier (UPI). The purpose of the UPI is to categorize or describe swaps with respect to the underlying products referenced in them, allowing regulators to aggregate, analyze, and report swap transactions by product type, and also to enhance position limit enforcement and real time reporting (CFTC, 2012). Industry experts involved in this initiative and the CFTC's Technical Advisory Committee (TAC) data subcommittee anticipate that it may be possible, once a product classification system is developed, to assign a UPI to approximately 80 to 95 percent of swaps (depending on the asset class involved), while approximately 5 to 20 percent of swaps may be sufficiently bespoke that they can only be described rather than identified by a UPI (CFTC 2011a, p. 101). As noted in the Committee on Payment and Settlement Systems and International Organization of Securities Commissions' report on over-the-counter derivatives data reporting and aggregation requirements, development of a standard product classification system is needed as a first step toward both a system of product identifiers for standardized derivatives products and an internationally-accepted semantic for describing non-standardized instruments (CPSS-IOSCO, 2011). The final rules with respect to UPIs, as published in the Federal Register, are reproduced in Table 1.

Insert Table 1 here

We provide a foundation for discussions about the most appropriate method for developing a product classification system as a foundation for assigning UPIs. It is expected that UPIs for the most liquid, standardized parts of the market will become available for integration within the overall industry regulatory reporting framework in 2013, and so there is considerable pressure to arrive at a consensus on the technical details and operation of the financial products classification system. In spite of these pressures, we believe it is important to take the time to consider all possible options for product classification and to carefully weigh the pros and cons of each approach.

In this paper, we focus on the advantages and disadvantages of various possible approaches to the hierarchical classification of financial products to support risk analysis and management, setting aside any discussion of the mode of operation of a financial products classification system or related technical implementation issues. As such, this paper examines the choices for product classification, predominantly from a knowledge organization-theoretic perspective, drawing on ideas from this field to discuss a range of possible hierarchical classification options before applying them to a set of use cases. Our decision not to focus on the technical mode of operation or means of implementing a financial products classification system means that we also do not focus on the Financial Products Markup Language (FpML)—an XML-based message standard for the OTC Derivatives industry—even though it incorporates a "Product Type" element. FpML is a standard that represents a messaging protocol in financial transactions, but it is not itself a classification system. For similar reasons, we do not focus on Financial Information eXchange Protocol (FIX) and other messaging standards (e.g., ISO 20022), even though these may have common data models. Nor do we focus on the Resource Description Framework/Web Ontology Language (RDF/OWL) specifications. RDF/OWL provides a framework and language for the description of semantic concepts in a domain (i.e. domain ontologies from which classification systems may be derived), but does not in itself specify the content of domain ontologies or classification systems. In other words, our focus is not on standards such as FpML or RDF/OWL; rather, it is on models or principles of organization for a hierarchical classification system itself. Once a particular model is chosen, it is then possible to populate the structure with content, or semantic elements, which may or may not be derived from a domain content standard[2] in order to facilitate grouping of the semantic conceptualizations of financial products for purposes of product valuation and risk analytics.

Related Literature

Systemic Risk

This paper explores approaches to the classification of financial products in support of enhancing financial systemic risk analytic and management capabilities, although we recognize that different audiences will have different requirements for financial product classification. For example, good classifiers should assist in data validation, data integration, pattern detection and risk monitoring at the firm level.

Since the global financial crisis there has been much written about financial systemic risk, but there is still no agreement on a precise definition of the term (Bisias, Flood, Lo, & Valavanis, 2011). Flood, Mendelowitz, and Nichols (2012) point to a recent shift from firm-centric measures of systemic risk to those which focus on the financial system or networks of relationships among financial counterparties. From this perspective, systemic risk may be said to refer to the risk or probability of breakdown (losses) in the individual parts or components of a financial system that affect the operation of the system as a whole, and is evidenced by co-movements (correlation) among most or all parts of the system (Kaufman, 2000). Bisias et al. outlines a range of techniques designed to identify systemic risk, including probability distribution measures; contingent claims and default measures; illiquidity measures; network analysis measures; and macro-economic measures.

Data and Transparency Issues affecting Systemic Risk Analysis and Management

A number of sources, including the FSB/IMF report cited above, have noted the data management issues facing macro-prudential supervisors. Flood, Mendelowitz and Nichols

(2012) observe that financial information has been expanding much faster than traditional technologies can track. The rapid expansion in financial information coupled with shortcomings of existing financial information processing infrastructure has limited the ability of financial regulators and market participants to identify, track and manage systemic risks (Flood, Kyle, & Raschid, 2011; Lemieux, 2012). These issues have precipitated the introduction of global initiatives to enhance data management throughout the global financial system, one of which is the CFTC's Final Rule.

Knowledge Engineering and Financial Product Classification for Systemic Risk Analysis and Management

Efforts to enhance financial data management in support of improved capabilities for financial risk analytics and management increasingly draw upon semantic approaches from the field of knowledge engineering. Knowledge engineering is an engineering discipline that involves integrating knowledge into computer systems in order to solve complex problems normally requiring a high level of human expertise (Feigenbaum & McCorduck, 1983). The relationship between hierarchical classification and knowledge engineering is discussed in Studer, Benjamins, and Fensel (1998) where the role of classification in aggregating instances in support of making inferences is set out in what they refer to as the Problem-Solving Method (PSM). There is a great deal of literature relating to classification in general and its application in a variety of domains, but comparatively little concerning the application of classification theory in the financial domain and less still on its application for use in financial risk analysis and management. Of note is Lemieux and Limonad (2011) who discuss the development of a domain ontology to represent financial crisis using upper level substantive and social ontologies to extend the semantic expressiveness of their modeling language and as a means of identifying the elements required to create ‘good’ representations of domains for the purposes of recordkeeping. Though they discuss classification, it is in the context of articulating the class hierarchies for their domain ontology, and not in respect to the development of a system for the hierarchical classification of instances of financial products which can be used to support the development of inference engines for financial systemic risk analytics, which is the focus of this paper. Similarly, Ye, Yan, Wang, Wang, and Miao’s (2011) paper on knowledge-level modeling for systemic risk management in financial institutions sets out a conceptual model for systemic risk analysis, laying out the static and dynamic constructs for their domain ontology and an implementation in RDF/OWL and SWRL. A related paper in the field of knowledge engineering is Pena, Patino, Palacio, Lochmuller, Ardila, and Villa (2012) which describes a fuzzy model of adaption for operational risk assessment in financial institutions. Their paper, however, focuses on classification of risk events, not financial products.

Action Research

The methodology used in the production of this paper is informed by Action Research (AR). AR is a systematic approach to investigation that enables people to find effective solutions to problems they confront in their everyday lives (Stringer, 2007). It works on the assumption that all stakeholders—those whose work is affected by the problem under study—should be engaged in the processes of investigation. Stakeholders participate in a process of rigorous enquiry, acquiring information and reflecting on that information to transform their understanding of the nature of the problem under investigation (Stringer, p. 11). In AR, stakeholders are active co-participants in the process of inquiry, rather than passive research subjects. The goal of this highly interactive and inclusive research approach is to transform the research into practical, reflective, pragmatic action directed toward solving “real-world” problems. For example, in the InterPares 3 Project, researchers comprised of academic researchers, professional archivists, and organizational staff, identified record keeping problems and then worked together through regularly discussion and focussed research activities to identify strategies and solutions to address the problems (InterPARES 3, 2012).

AR is grounded in a qualitative research paradigm. Unlike quantitative research that is based on the precise definition, measurement, and analysis of relationship between a carefully defined set of variables, action research begins with a question, problem, or issue—in this case a question about options for financial product classification—that is broadly defined. Formally, then, AR is phenomenological (focusing on a community's actual lived experience/reality), interpretive (interpreting acts and activities of stakeholders in the community), and hermeneutic (incorporating the meaning stakeholders make of their activities) (Stringer, 2007, p. 20). The literature on action research is quite extensive, indicative of its application to a wide range of professional and organizational contexts. In spite of its wide diffusion in other domains, its use in the financial domain is quite limited. To our knowledge, Waddell (2012) remains the sole example of the application of this method in the area of global finance in spite of its suitability to this domain. In the Waddell study, the project team set out to integrate social, economic and environmental concerns into the "logic" of the global financial system. The researchers engaged in this project did not have extensive expertise in the domain of finance. This is a key differentiator from our study, in which individuals possessing knowledge of the global financial system and of financial products actively participated. This is not to suggest that the researchers who participated in this study are representative of the entire domain; however, they do offer expert if not entirely discursive legitimacy (Huzzard, Ahlberg, & Ekman, 2010).

We also identified as relevant to our purposes the theory of structuration (Giddens, 1984) and, in particular, adaptive structuration theory (Orlikowsky, 1992; Yates & Orlikowsky, 1992). The theory of structuration emphasizes mutual interaction between structures (e.g. financial markets), functions (e.g. over-the-counter trading) and actors (e.g., financial counterparties). Adaptive Structuration Theory (AST) (DeSanctis & Poole, 1994), which studies the interaction of groups and organizations with information technology, is relevant because it draws on the concepts of the theory of structuration to study the interplay existing between social structures, human action and advanced information technologies.

Methodology

In the specific methodology used in development of this paper, stakeholders were involved in co-generation of knowledge. The stakeholders who participated in the research form part of an informal community of practice—the Open Financial Data Group (OFDG)—for which social interaction creates meaning and defines values in respect to data management and information technologies within the global financial system. The OFDG is comprised of financial regulators, market participants, technology developers, academics, and representatives of financial industry associations from a number of different countries. The group meets weekly via conference call, and participation in the weekly calls varies with members joining as they have time and interest in the topic under discussion. The authors acknowledge their place within the OFDG both as researchers and stakeholders.

The ideas presented in this paper took shape over a series of recorded weekly OFDG calls held between 2 March and 13 July, 2012, as well as drawing on discussions on the Group's Basecamp collaboration site. The authors' analysis and synthesis of the weekly call recordings and wiki postings resulted in the production of this paper. Early drafts of the paper were shared with OFDG stakeholders and the authors made revisions based on feedback received. This process continued in an iterative fashion until OFDG stakeholders were satisfied with the final "position" on the options for financial product classification for enhanced systemic risk analytic capabilities.

Financial Product Classification Initiatives

During the period that the issue of financial product classification was under discussion by the OFDG, an industry initiative to create a product classification system to extend classification of financial products to include products primarily based around cash flows was being led by the creators of FpML, in cooperation with experts at FIX (FpML, 2012). The International Swaps Dealers Association (ISDA), and an inclusive group of industry representatives, had developed a proposal and underlying approach to implementing UPIs in the marketplace to support regulatory reporting to swaps data repositories. Participants in the development of this proposal included: dealers, buy-side market participants, central counterparty clearing houses, affirmation platforms, potential swaps exchange facilities, swaps data repositories and other market facilities. The proposal was being vetted with each of the ISDA product steering committees to ensure that the overall approach was fully understood and supported by the industry, including ensuring that

- It presented no impediments to product innovation or trading
- The costs were well understood
- The resulting UPI level of granularity would meet regulatory needs without causing market disruptions

- The creation of a full product representation (as opposed to a ticker-only representation of the product) would result

ISDA's aim was to have the vetting process reach a point of industry consensus on adoption of the framework. As part of this process, a whitepaper on UPI generation and dissemination was produced that would detail options for where UPIs will be generated and the resulting workflows as well as options for how UPIs could be disseminated (e.g. translate a UPI key to a product representation).

At the time, there was a broad industry consensus that the UPI code and system could not be achieved without identifying products that make sense for price comparison. There was also recognition that the UPI code and system would require both human recognizable and computer readable identifiers. It was also thought that the identifier also would have to be extendable beyond vanilla (i.e., standard) products because all products will eventually need a UPI with the exception of exotics (i.e., highly customized products). Further, there was wide recognition that UPIs are best suited for liquid products and least suited for exotic products, with the suggestion that exotics instead be represented via an asset class taxonomy and a generic UPI (e.g., 999999)

The ISDA proposal called for the following elements:

- Asset class and taxonomy
- Product description: A human readable description of the product
- Ticker: For some but not all products where appropriate to be used for price dissemination (ticker tape) and quoting (e.g. by swaps exchange facilities) for human understanding
- UPI key: A computer-friendly, singular, opaque, immutable, fixed length, always unique key

- A ‘rulebook’ for each asset class which would define the price forming attributes which make up the product info-set.

The CFTC’s Technical Advisory Committee Sub-Committee on Data Standardization, UPI and LEI working group supported the ISDA UPI proposal and underlying approach.

Insert Figure 1 here

The methodology which was suggested to implement the proposed ISDA system aimed to:

- Convert a standard trade representation (FpML) into a product info-set by determining the price forming attributes
- Normalize the data by ensuring consistency in the values and sorting etc. in the product info-set
- Generate the UPI key by inputting the product info-set into a hashing algorithm. The resulting output would effectively yield a human readable product description, though the creation of tickers could be used for some standardized and liquid parts of the market.
- Generate the ticker as an additional step after generation of the product description and UPI Key.

As of March, 2012, ISDA had issued its Product Taxonomy Scheme and FpML had stated that this would become the default scheme to determine the FpML “Product Type” element (FpML, 2012). Work to extend international standards (e.g., ISO 10962 on Classification of Financial Instruments) is ongoing (FpML, 2012).

Hierarchical Taxonomies

This paper examines the various approaches that can be taken to the organization of financial products. Emphasis is given to approaches that use taxonomies, especially those that are hierarchical in nature. We use the term *taxonomies* broadly, referring to any system of representation that is used to group, arrange, and describe items according to meaningful principles, and which provides users with an overview of the set of items as a whole (Lambe, 2007). Particularly well known taxonomies include the biological classification first developed by Carl Linnaeus (1735), the periodic table of elements developed by Dmitri Mendeleev (1869), and the Dewey Decimal Classification developed by Melville Dewey (1876). Taxonomies can also take many other forms such as inventories, maps, figures, and diagrams. A grocery list can be considered a taxonomy simply because it groups together items that are needed from a grocery store. Similarly, an anatomical chart of the human body can be considered a taxonomy since it represents how human organs fit together.

We consider ontologies to be types of taxonomies, though they are somewhat atypical of taxonomies. A typical taxonomy might group together categories by using one or two simple relationships (such as *type of*) in a systematic manner. In contrast, an ontology explicitly defines the precise nature of concepts and the relationships between them. For example, Lemieux and Limonad (2011) use an ontology to describe a wide range of concepts that relate to the structure of financial records, including FINANCIAL SYSTEM, FINANCIAL ASSET, and POLICY. These concepts are associated by various types of relationships, including *owns, is part of, secured by,* and *restricts*. These conceptions of taxonomies and ontologies are not universally accepted.

Lambe (2007) maintains that an ontology would be a type of taxonomy. However, Suryanto and Compton (2000), writing from the perspective of knowledge management, conceive of taxonomies as being narrower than and sometimes derived from ontologies. In addition, Guarino (1998) distinguishes between *ontology* as a formal philosophical discipline that concerns itself with developing a system of categories that account for a certain conceptualization of the world and *ontologies* as engineering artifacts constituted by a specific vocabulary that describe a certain domain.

Our understanding of taxonomies is informed by various theories and standards, including those taken from knowledge organization theory (e.g., Vickery, 1960; Ranganathan, 1962; Kwasnik, 1999; Dextre Clarke, 2001; Jacob, 2004; Broughton, 2006), professional indexing (Aitchison, Gilchrist, & Bawden, 2000; ANSI/NISO, 2005), taxonomists (Lambe, 2007; Hedden, 2010), information architecture (Farkas & Farkas, 2002; Morville & Rosenfeld, 2006) and ontologists (Sowa, 2000; W3C, 2004a; 2004b; Gruber, 2008). Some of these models are standards that explicitly constrain the manner in which a taxonomy is to be structured. For example, the Resource Description Framework Schema (RDFS) is constrained such that all instances of a class must also be instances of the class's superordinate class. Similarly, if one wishes to adhere to the recommendations of the ANSI/NISO Z39.19 standard regarding monolingual controlled vocabularies, only certain types of relationships are allowed between categories.

Some models are not standards, but theories that seek to describe taxonomies that are already in existence. For example, when Dextre Clarke (2001) discusses thesauri, her purpose is not to provide constraints as to what is allowable in thesauri, but to describe the types of relationships that are commonly observed in thesauri and to provide explanations for inconsistencies that are observed between thesauri. These approaches derive from observation of taxonomies "in the wild" (Glushko, Maglio, Matlock, & Barsalou, 2008) in the same manner as a botanist might observe specimens of plants.

Many taxonomies are easy to create and use, but prove difficult to describe, since they are more complex than they appear to be. For example, when making a taxonomy of animals, it may be easy and natural to create the categories DOGS, CATS, and HAMSTERS and to group them under PETS. However, this structure can be fundamentally misunderstood. What about dogs that are not pets? Because the category DOGS is nested under the category PETS, it is possible to interpret the structure above such that only dogs that are also pets should be included in DOGS. It is also possible that the category DOGS is meant to include all dogs and that PETS is simply the most convenient section of the taxonomy in which to place DOGS. Upon further investigation, the user of the taxonomy may find that DOGS is nested under both PETS and CANINES. In that case, he or she may infer that HAMSTERS is nested under both PETS and RODENTS, only to find that there is no category called RODENTS in the taxonomy. There may also be difference in opinion as to whether it is valid to add the category PET CARE under PETS. Some people may argue that, since all the other categories under PETS are types of pets, adding PET CARE violates the principles of how PETS should be organized. Other people may argue that any principles for the organization of PETS are merely inferred and that PET CARE is a valid subcategory. Further complications may ensue if some people use the taxonomy to organize animals, while other people use the taxonomy to organize books that are about animals. When the taxonomy is used to organize animals, a pet dog may be considered to be a member of both the categories DOGS and PETS. However, when the taxonomy is used to organize books, a book that is about dogs (or even pet dogs) might not be considered a valid member of the category PETS, since the book is not about pets in general. As these examples illustrate, even a very simple taxonomy can be subject to a wide range of interpretations and misunderstandings.

Because taxonomies can be deceptively complex, it is perhaps not surprising that no one theory of taxonomies is universally agreed upon. Also, different theories of taxonomies employ slightly different terminologies. Kwashnik (1999) uses the term *classification* to refer to a system that employs a "meaningful clustering" of items (p. 24), while Jacob (2004) uses the same term to refer to the "orderly and systematic arrangement" of items into a "system of mutually exclusive and nonoverlapping classes" (p. 522). Some theorists use the term *hierarchy* to refer to any system of subdivision (e.g., Morville & Rosenfeld, 2006), while others use *hierarchy* only to refer to subdivisions into general types (e.g., Kwashnik; Lambe, 2007). Many professional guidebooks do not provide formal definitions or descriptions when discussing taxonomies, preferring informal language and illustrative examples (e.g., Morville & Rosenfeld). For these reasons, the terminology that we use in the discussion of taxonomies is not necessarily consistent with the terminology used in other theories and standards.

The remainder of the discussion of taxonomies focuses on hierarchical taxonomies, in which concepts are systematically divided and subdivided (Morville & Rosenfeld, 2006). Hierarchies are useful because they are fairly simple to construct and easy to navigate. Hierarchies tend to be modular, making it relatively easy to identify subsections within the overall hierarchical structure. In addition, it is possible to make inferences about a concept based solely on its position in a hierarchy. For example, if a hierarchy of military ranks is organized such that Generals have power over Colonels, and Colonels have power over Majors, it can be inferred that Generals have power over both Colonels and Majors.

Although it can be easy to identify a structure as being a hierarchy, it can be surprisingly difficult to identify the qualities of hierarchies in general. For example, Lambe (2007) claims that a hierarchy is, among other things, a tree structure, which is a series of nested one-to-many relationships. Kwashnik (1999) describes tree structures as structures that "divides and subdivides its classes based on specific rules for distinction," such as military ranks (p. 30). A taxonomy of military ranks has a tree structure if each soldier is linked to multiple soldiers over which he or she has power. A General would be linked to the Colonels over which he or she has power. A taxonomy of physical regions has a tree structure if a region such as Canada is linked to its constituent parts.

These distinctions are useful, but the presence of one-to-many relationships is not necessary or sufficient in determining whether a taxonomy is a hierarchy. In some cases, a category might have multiple categories nested underneath it in a one-to-many relationship that are not hierarchical in nature. For example, a person might be linked to his or her various phone numbers. In addition, a hierarchical structure might include relationships other than one-to-one relationships. A hierarchical taxonomy might divide the United States into its various states. It might further divide each state into its congressional districts. Some states, such as Wyoming, have only one congressional district. Therefore, a hierarchical structure that was based on US states and congressional districts would include a few one-to-one relationships between a state and a congressional district. In addition, some hierarchies have no relationships at all between categories. For example, in a hierarchy of income brackets, the wealthiest 1% of the population might occupy the top bracket while the poorest 20% might occupy the bottom bracket. This hierarchy might not include any relationships between these income brackets aside from their position on a continuum. If hierarchies are thought of only as implementations of one-to-many relationships, then structures without relationships would not be considered hierarchies.

Nevertheless, the presence of one-to-many relationships in a taxonomy is highly characteristic of a hierarchical structure. Many of the advantages of hierarchical structures, such as modularity and the potential for transitive inferences, are facilitated by one-to-many relationships. Accordingly, hierarchical structures tend to be fertile ground for the development of one-to-many relationships. We consider a hierarchical structure to have two characteristics. Firstly, it consists of categories that are organized along an identifiable continuum, such as *specificity* or *size*. Secondly, more items tend to be placed at one end of the continuum than at the other. A taxonomy of military ranks is a hierarchy if it groups people according to military rank, it sequences those ranks according to their relative power, and if the lower ranks tend to have higher populations than the higher ranks. If the continuum is transitive in nature, then it is possible to make inferences based on the hierarchy. People in military hierarchies do not simply have different degrees of power. Instead, some people command others. For these reasons, it is fairly straightforward to represent this hierarchy with a transitive "commands" relationships. Because the taxonomy tends to be lop-sided, most of the relationships will be one-to-many in nature. However, there may be a few cases where a person commands only one other person, or where a person answers to multiple commanding officers. Because hierarchies need not consist entirely of one-to-many relationships, the presence of other types of relationships does not disqualify a taxonomy of military ranks as being a hierarchy.

Types of Hierarchies

Having laid the theoretical groundwork for the essential characteristics of a hierarchy, we now examine different types of hierarchies. Some hierarchies are structured so that it is not possible for a category to have more than one superordinate category. We refer to these hierarchies as *strict* hierarchies. Other types of hierarchies allow categories to have multiple superordinates and are known as *polyhierarchies* (Kwashnik, 1999). Well-known examples of strict hierarchies include the biological classification and the Dewey Decimal Classification.

Each of the proposed ISDA OTC Derivatives Taxonomies is also a strict hierarchy, since no category in those taxonomies can have more than one superordinate. For example, the ISDA Commodities Taxonomy includes many categories that are labeled CASH. Like all categories in the ISDA Taxonomy, each CASH category has only one superordinate. Therefore, CASH as it is nested under METALS/PRECIOUS/SWAP is a different category from CASH as it is nested under METALS/PRECIOUS/OPTION. Strict hierarchies are often used to organize physical objects, such as books on a bookshelf, which can exist in only one location at a time. Further, they are often organized according to a single relationship type. Often, it is easier to interpret the organizational principles of a strict hierarchy than a polyhierarchy. However, a strict hierarchy tends to be useful only to people who have an interest in the specific set of principles upon which the hierarchy is based. For example, if the category CONVERTIBLE BONDS is nested under BONDS in a strict hierarchy, it cannot also be grouped under EQUITIES. In that case, people who are interested in types of equities would find that the taxonomy does not suit their needs.

Insert Figure 2 here

Insert Figure 3 here

In a polyhierarchy, however, CONVERTIBLE BONDS could be grouped under both BONDS and EQUITIES. However, in most polyhierarchies, there is no explicit distinction that is made between different types of hierarchical relationships, each of which may support different types of inferences. For example, if DOGS was nested under both CANINES and PETS, a person might infer that, because all dogs are canines, then all dogs must also be pets. Similarly, if CONVERTIBLE BONDS is grouped under both BONDS and EQUITIES, a person with incomplete knowledge of financial instruments might infer that convertible bonds have the same relationship to bonds that they have to equities, when in reality a convertible bond ceases to be a bond once it is converted to equity. When using a polyhierarchy, there is usually no way to distinguish between different types of hierarchical relationships, or the different implications that a relationship carries. For that reason, making inferences is often more fraught when using a polyhierarchy than when using a strict hierarchy (Morville & Rosenfeld, 2006; Lambe, 2007).

Hierarchies can also be distinguished by the type of relationships that they use. Hierarchies could make use of virtually any kind of transitive relationship. As seen earlier, a hierarchy of military ranks might use a *commands* relationship, while a hierarchy of regions may use a *partitive* relationship in which each category is part of its superordinate. It is even possible to organize a group of partygoers into a hierarchy that is based on the *spoke to* relationship in order to track how an idea at the party spreads from person to person. Each of these relationships is *recursively transitive*, such that if A has a relationship with B, and B has the same relationship with C, it can be inferred that A also has that relationship with C (Kwashnik, 1999). However, sometimes exceptions occur. It may be that a corporate executive supervises a departmental manager, and that the manager supervises a technician, but that the executive has no form of power over the technician. Additionally, the category CHAIR may in some cases include non-traditional chairs such as chairlifts on which it is possible to sit. CHAIR, in turn, is typically included in the category FURNITURE. However, it does not follow that chairlifts should be included in category FURNITURE, however comfortable they may happen to be (Hampton, 1982).

In the case of derivatives, we may note, somewhat similarly, that spots, forwards, options and swaps are all kinds of over the counter contracts but spots are not generally considered to be derivatives. Similarly there is debate as to whether a credit default swap is in fact a swap.

One of the most commonly used relationship types in a hierarchy is the *generic* relationship, which is sometimes referred to as the *type of* relationship or *is-a* relationship. This relationship organizes items according to type. The generic relationship exists between DOG and CANINE, since dogs are types of canines, but does not exist between COLONEL and GENERAL, since Colonels are not types of Generals. Different theories provide slightly different accounts of the nature of generic relationships. Some models describe generic relationships as applying in cases where all members of a category are included in another category (e.g., Kwashnik, 1999; Dextre Clarke, 2001; Farkas & Farkas, 2002; W3C, 2004b). Other approaches describe generic relationships as *parent/child* relationships (e.g., Morville & Rosenfeld, 2006), *broader term/narrower term* relationships (e.g., Hedden, 2010), or relationships between a category and a *species* or *member* of that category (e.g., Kwashnik, 1999; Aitchison et al, 2000; Jacob, 2004; ANSI/NISO, 2005). In all likelihood, each of these models conceives of the generic relationship in a broadly similar manner. However, generic hierarchies can still differ from each other in important ways, such as whether they organize items or topics, or with respect to the implicit understanding of the concept *species* outside of its traditional usage in biology.

Types of Generic Relationships in Hierarchies

A generic relationship can be thought of as a series of inclusion sets in which items in a category are also included in that category's superordinate category. However, many generic relationships form inclusion sets that are not precise. In those cases, most, but not all, of the items in a category are also members of the category's superordinate. A generic relationship between DOGS and PETS is imprecise because a dog is not necessarily a pet. Imprecise inclusion sets often form when a generic hierarchy is designed to be easy to use rather than rationally precise. For example, a generic relationship might exist between RATS and VERMIN simply because the people who use the taxonomy find it convenient to place these two categories near each other and no other type of relationship is possible. Similarly, CONVERTIBLE BOND could be nested under EQUITIES, even though not all convertible bonds are traded by the equities desk of a firm. If the taxonomy only uses the generic relationship, then in order for the categories CONVERTIBLE BOND and EQUITIES to be associated with each other, one must be nested under the other. Obviously, an imprecise generic relationship could result in an incorrect inference, such as a convertible bond being inappropriately treated as though it was traded by an equities desk. Nevertheless, given the time constraints under which many taxonomies are created, it is not always practical to create a generic hierarchy that consists only of precise inclusion sets. Even for categories that are not necessarily inherently fuzzy, certain instances of the category may be highly exceptional and therefore inferences drawn based on placing those instances within a particular category may prove false or misleading.

Furthermore, some concepts are inherently fuzzy in nature. As Rosch and Mervis (1975) demonstrated, it is exceedingly difficult, or perhaps impossible, to accurately and precisely stipulate the essential characteristics of the concept FURNITURE. As a result, some people may feel that all seats count as furniture, while others discount atypical seats such as car seats or miniature seats that are found in doll houses. Many naturally occurring categories are notoriously difficult to define precisely. For example, metallurgists do not agree on the essential characteristics of metals (Murphy, 2004). The essential nature of biological species such as TIGER has also proven elusive.[3] Similarly, the CFTC's definition of a swap, though intended to precisely establish the scope of rules relating to swaps, still raises questions about whether certain types of contracts that were not intended to be treated as swaps might fall within the scope of the definition (Brush, 2010). If a hierarchy includes categories that are inherently fuzzy, some of the inclusion sets in the hierarchy will not be strictly precise.

Some standards, such as ANSI/NISO Z39.19, discourage the use of imprecise inclusion sets in taxonomies. In addition, many theorists hold that inclusion sets must be precise in order for a relationship to be truly generic, or even hierarchical (e.g., Cheti & Paradisi, 2008). In this paper, we do not take a position on whether the use of precise inclusion sets is recommended, or legal, but instead observe the type of relationships that are commonly used. We also do not interpret standards such as ANSI/NISO Z39.19 as prescriptive definitions of taxonomies, such that any structure that violated the standard could not be considered a taxonomy. Rather, we view ANSI/NISO Z39.19 simply as a set of recommendations.

As stated earlier, generic relationships have been described as links between a category and a *species* of that category. An understanding of the generic relationship that is based on the concept of *species* is fraught for the reasons described above. That is, a species of a superordinate category may include members that do not belong to that superordinate category. In addition, even if the hierarchy is rigorously constructed so that each category falls entirely within its superordinate, it may be unclear just what a species is. Intriguingly, although many theorists and standards use the term *species* to describe generic relationships (Vickery, 1960; Kwashnik, 1999; Aitchison et al., 2000; Jacob, 2004; ANSI/NISO, 2005), it is difficult to find a definition of *species* in the context of taxonomies. It is possible that theorists use *species* synonymously with *inclusion set*, so that a species is simply any set of items or concepts that is included within a larger set. However, it is also possible that species is only used to refer to sets whose identities are well established within a field or discipline. For example, SQUARE is generally considered to be a species of QUADRILATERAL because squares have all of the properties of quadrilateral in addition to having properties that are not necessarily associated with quadrilaterals (Smith & Medin, 1981). In addition, the category RED SQUARE is clearly included within the general category SQUARE. However, it is not clear that RED SQUARE is a species of SQUARE, since RED SQUARE is not an established type, but is rather an ad hoc combination of the well-established concepts RED and SQUARE. In the domain of finance, EQUITY BASKET OPTION might be considered a type of OPTION without being considered a species of option. Vickery (1960) seems to subscribe to this interpretation of *species* when he suggests that, when a taxonomy combines terms in compound subjects, it "breaks free from the restriction of traditional classification to the hierarchical, genus-species relation" (p. 13). Here, it is important to keep in mind the kind of system for which the classification scheme is designed and its purpose (e.g. in a political classification system one would expect to find the concept RED SQUARE as in a city square in Moscow).

It could be argued that to contrast *type* with *species* is to make a distinction without a difference. However, the distinction speaks to the fundamental purpose of the taxonomy. If the purpose of a hierarchical taxonomy is to organize items or concepts according to consistent principles, it may not be necessary to make a distinction between well-established concepts and ad hoc intersections of concepts. However, if the purpose of the taxonomy is to organize established concepts, then the term *species* might be used to exclude marginal, novel, or ad hoc concepts. For example, in the Emtree Thesaurus, ARTIODACTYLA (a type of hoofed mammal) is nested under PLACENTAL MAMMALS, which in turn is nested under MAMMALS. In the CAB Thesaurus, ARTIODACTYLA is nested directly under MAMMALS (Dextre Clarke, 2001). It is unclear whether the Emtree Thesaurus includes PLACENTAL MAMMALS because its users may wish to refer to that category, or because PLACENTAL MAMMALS is simply a useful means for grouping together a set of subordinate categories.

Some hierarchical taxonomies include relationships that do not represent the hierarchical structure. For example, many thesauri identify all non-hierarchical relationships collectively as *associative relationships*. These relationships are generally used to link together categories that have an important relationship, but which exist in different sections of the hierarchy (Dextre Clark, 2001; ANSI/NISO, 2005). An associative relationship might be used to link together BEES (which may be nested under INSECTS) and HONEY (which may be nested under CONDIMENTS). Associative relationships can be very useful in helping a user navigate a taxonomy. However, the unspecified nature of associative relationships implies that associative relationships are based on a variety of principles that cannot be discerned from the relationship itself. For that reason, it is often not possible to make inferences based on associative relationships. For example, a taxonomy may use an associative relationship to connect the categories BEES and HONEY and a second associative relationship to connect HONEY and TOAST. However, the associative relationships cannot be used to support inferences between BEES and TOAST (apart from the fact that both categories are related to HONEY in some manner).

Taxonomies are typically thought of as organizing items according to principles that apply to the items themselves. If an animal has wings, it is likely to belong in the category BIRDS, which tends to include animals that have wings. However, taxonomies that organize documents tend to organize those documents according to their topics. These hierarchies are often structured slightly differently than other hierarchies (Vickery 2008, Loehrlein, 2011). The differences are partly due to the fact that documents represent topics at particular levels of specificity, and partly due to the fact that collections of topics often include concepts that are of fundamentally different types. When a document represents a topic at a particular level of specificity, it not only represents the characteristics of the topic, but it attributes the characteristics specifically to that topic. For example, we could say that all robins have wings, but we cannot tell from examining a robin whether *has wings* is a characteristic of robins, birds, or animals. A book about robins not only represents the characteristic *has wings*, it attributes that characteristic to a particular level of specificity via statements such as "like all birds, robins have wings." In contrast, a book about birds in general is likely to attribute *has wings* to birds without specifically mentioning whether robins have also wings. For that reason, categories of documents do not form inclusion sets. Books that are about robins would not be included among books that are about birds in general. The two sets of books are likely to be disjoint, apart from a few books that happen to be devoted to birds in general and to robins in particular. Nevertheless, taxonomies of topics are often organized as though they were taxonomies of items. When organizing books by topic, the topic ROBINS might be nested under the topic BIRDS, not because books on the former topic are included among books on the latter topic, but because robins are included among birds.

However, hierarchies that are organized by topic are particularly susceptible to imprecise nesting, since these hierarchies tend to bring together concepts that are of fundamentally different types. In a hierarchy of living things, all the items are living organisms, but in a hierarchy of topics, almost any type of concept can be included. For example, FRENCH GRAMMAR is not a type of FRENCH, but a book on French Grammar might be considered a type of book on the French language. Similarly, a hierarchy that organizes scientific journal articles might nest the class ABSORPTION within the class CHEMISTRY even though absorption is a process and chemistry is a discipline. The concepts are nested perhaps because most studies of absorption are thought to be undertaken by chemists. It is not possible to evaluate the appropriateness of the hierarchical relationship based on the members of each category, since the journal articles that are about absorption are disjoint from the journal articles that are about chemistry in general. For that reason, if hierarchies of topics are to be constructed in a precise manner, great care must be taken to ensure that the categories are organized according to consistent principles.

When considering financial products, derivatives could be thought of as a description of their underlying asset. For example, a foreign exchange (FX) swap can be considered to be a specialized description of the currencies to be exchanged in the swap. For that reason, a financial taxonomy might nest derivatives under their underlying asset, even though derivatives and assets are fundamentally different types of concepts. While this approach is likely to confound inferences about the concepts organized, it is the same general approach that is used in schemes that organize documents by topic.

Faceted Schemes—A Tool for Creating a Generic Hierarchy

When designing a generic hierarchical structure, there are often a variety of approaches that can be taken when dividing and subdividing categories. A taxonomy of vehicles could divide the set of vehicles into well-established types: cars, trucks, trains, airplanes, boats, etc. These general types could be subdivided according to different principles. That is, cars could be subdivided according to their types of chassis, producing the subcategories sedans, compact cars, and subcompact car. In contrast, airplanes could subdivided according to types of engines, producing the subcategories jet planes and propeller planes. When people use this taxonomy, they may find it difficult to predict how a given category will be subdivided. For that reason, some generic hierarchies are structured so that categories are subdivided according to a consistent set of principles. In order to ensure that principles of division are consistently applied, hierarchies may use a schedule of principles of division and the order in which to apply them. These schedules are referred to as *faceted schemes*. A faceted scheme takes a bottom-up approach. That is, it does not organize items, but instead identifies the characteristics of items and organizes them into facets. At that point, the characteristics from multiple facets can be combined to represent an item. For that reason, each item is represented as a specific intersection of characteristics (Broughton, 2006; Cheti & Paradisi, 2008).

Many theorists refer to facets as *categories* or *fundamental categories* (e.g., Kwashnik, 1999; Hedden, 2010). However, because the term *category* can refer to any type of grouping, more precise language is necessary to understand how facets differ from other types of categories. According to Taylor (2006), facets are “clearly defined, mutually exclusive, and collectively exhaustive aspects, properties, or characteristics of a class or specific subject” (p. 394). Ranganathan (1962) describes facets as “trains of characteristics” (p. 71), which can be understood to be sets of characteristics that are ordered according to a single principle. When that situation occurs, the values are *alignable* (e.g., Gentner & Markman, 1994). For example, the characteristics TWO-SIDED, THREE-SIDED, and FOUR-SIDED could be grouped together into a facet that is organized by the principle NUMBER OF SIDES. Ideally, the characteristics within a facet should be identical to each other, or nearly identical, except with respect to a single principle. However, facets could also be sets of values that are abstract or subjective in nature, such as different degrees of comfort or entertainment value (Johnson, 1984). For example, consider the following faceted scheme of financial derivatives:

Facet A: Contract structure

Option

Forward

Swap

Swaption

Contract for Difference

(etc.)

Facet B: Notional maturity

1 month contract

6 month contract

1 year contract

(etc.)

Facet C: Type of Underlying

Foreign Exchange

Rate Based

Index

Interest Rate

Asset

Debt

Equity

Commodity

Credit

The characteristics in the NOTIONAL MATURITY facet are identical except for a single principle: the amount of time to the contract's notional maturity. However, the characteristics in the TYPE OF UNDERLYING facet can be differentiated according a wide set of principles. For example, there are many differences between a contract wherein the underlying is foreign exchange versus that in which the underlying is a commodity.

Once characteristics of items have been organized into facets, the facets can be used to divide and subdivide the items in a hierarchy. One facet is chosen to make the initial division of items, while the other facets are used to make a series of subdivisions. For example, the set of all derivatives can be divided first according to the CONTRACT STRUCTURE facet, including options, forwards and swaps. Each of these groups could be subdivided by the values in the NOTIONAL MATURITY facet, so that options, forwards, etc., are divided according to the length of time to notional maturity. The TYPE OF UNDERLYING facet could be used to divide the categories a third time. As a result, the following hierarchy emerges:

FINANCIAL DERIVATIVES

- OPTIONS
  - OPTIONS, 1 MONTH CONTRACT
    - OPTIONS, 1 MONTH CONTRACT, FOREIGN EXCHANGE
    - OPTIONS, 1 MONTH CONTRACT, INDEX
    - OPTIONS, 1 MONTH CONTRACT, INTEREST RATE
  - OPTIONS, 6 MONTH CONTRACT
    - OPTIONS, 6 MONTH CONTRACT, FOREIGN EXCHANGE
    - OPTIONS, 6 MONTH CONTRACT, INDEX

OPTIONS, 6 MONTH CONTRACT, INTEREST RATE

(etc.)

One of the great advantages of the faceted structure is that once the structure is in place, it is relatively easy to alter the order in which the facets apply. This order is commonly referred to as the scheme's *citation order*. For example, a person who is primarily interested in types of underlying could apply the TYPES OF UNDERLYING facet first, followed by the CONTRACT STRUCTURE facet. In that case, a different hierarchy emerges:

FINANCIAL DERIVATIVES

FOREIGN EXCHANGE

FOREIGN EXCHANGE, OPTIONS

FOREIGN EXCHANGE, OPTIONS, 1 MONTH CONTRACT

FOREIGN EXCHANGE, OPTIONS, 6 MONTH CONTRACT

FOREIGN EXCHANGE, OPTIONS, 1 YEAR CONTRACT

FOREIGN EXCHANGE, FORWARD

FOREIGN EXCHANGE, FORWARD, 1 MONTH CONTRACT

FOREIGN EXCHANGE, FORWARD, 6 MONTH CONTRACT

FOREIGN EXCHANGE, FORWARD, 1 YEAR CONTRACT

(etc.)

Both hierarchies are strict hierarchies. In both cases, a single principle of division is employed at each hierarchical level. The user of the system has control over which principles of division to apply to the hierarchy and the order in which they can be applied. As a result, people with different priorities can use the same faceted scheme to arrange the same set of items according to different sets of principles.

It is important to remember that a faceted approach alone is not guaranteed to produce hierarchies that are organized according to consistent principles. If a facet itself is organized in an imprecise manner, any hierarchy that is organized according to that facet will also be imprecise. For example, a facet of derivatives could contain the values COMMODITIES, METALS, CASH, and CREDIT DEFAULT. However, these values are not alignable by any identifiable principle. If six month contracts were subdivided according to this facet, different types of relationships would form, since the relationship between "six month contracts" and "six month credit default contracts" is different from the relationship between "six month contracts" and "six month metals contracts." In addition, facets can be constructed such that an item fits two or more of its values. Certain contracts are both credit default contracts and commodities contracts. In a hierarchy that is based on the above facet, those contracts would appear in more than one location.

The categories that result from faceted schemes tend to be ad hoc intersections of characteristics rather than well-established concepts. This phenomenon occurs because the process of dividing categories by facet naturally results in categories that are combinations of values from each facet. A faceted structure could easily be devised that would divide squares into red squares, yellow squares, and blue squares. However, it would be more difficult to devise a faceted structure that divided a well-established concept such as QUADRILATERAL into another well-established concept such as SQUARE (Smith & Medin, 1981). A faceted scheme that attempted to do so may look something like this:

Facet A: Number of sides

3 sides

4 sides

(etc.)

Facet B: Equality of sides

All sides equal in length

All sides not equal in length

Facet C: Equality of angles

All angles equal in size

All angles not equal in size

In that manner, four-sided shapes (quadrilaterals) could be broken down into four-sided shapes whose sides are equal in length (rhomboids) and further broken down into four-sided shapes whose sides are equal in length and whose angles are equal in size (squares). However, this approach is cumbersome and of questionable utility. How often will a user be interested in a set of shapes that is defined only by having sides that are equal in length? In addition, this approach often produces null sets, such as triangles with sides that are all equal in length but angles that are not equal in size. Similarly in classifying derivatives products, many intersections may be null, such as credit default futures and so on. For that reason, people classifying financial instruments may opt for a looser set of organizing principles.

In general, users who are primarily interested in hierarchies of well-established concepts will often be frustrated with the categories that can be formed out of faceted schemes. However, it is equally true that users may be interested in certain combinations of properties that they deal with on a routine basis, but which do not have well-established names. This is certainly the case in the area of risk analytics as applied to financial products. In that case, a faceted scheme could have a great deal of utility by helping analysts to identify and combine the ad hoc categories of products that are of particular interest to them.

Because faceted schemes produce categories that are based on intersections of characteristics, hierarchies that are based on facets cannot form inclusion sets that are imprecise. Because sub-sections can be only formed by taking a category and dividing it, all members of a category must also be members of the category's superordinate. For example, it is not possible to nest the category RATS under the category VERMIN when using a faceted scheme. Instead, RATS and VERMIN would each be treated as characteristics that could be combined in various ways to form different categories. A faceted scheme that includes these characteristics may take this form:

FACET A: ROLE

PET

VERMIN

FACET B: SPECIES

MOUSE

RABBIT

RAT

The facets could be combined to form either this hierarchy:

PET

PET MICE

PET RABBITS

PET RATS

VERMIN

MICE THAT ARE VERMIN

RABBITS THAT ARE VERMIN

RATS THAT ARE VERMIN

Or this hierarchy:

MOUSE

PET MICE

MICE THAT ARE VERMIN

RABBIT

PET RABBITS

RABBITS THAT ARE VERMIN

RAT

PET RATS

RATS THAT ARE VERMIN

Both hierarchies consist only of precise inclusion sets, and in neither case could RATS be nested under VERMIN. Similarly, when using a faceted scheme to organize financial products, it would not be possible to nest CONVERTIBLE BONDS under EQUITIES. Instead, the hierarchy would nest CONVERTIBLE BONDS under BONDS.

Facets are also a useful means for separating intrinsic and extrinsic characteristics of items. A bond may be intrinsically convertible, but might be traded on the equities desk in some circumstances and the fixed income desk in other circumstances. A faceted scheme tends to naturally group together intrinsic characteristics into certain facets and extrinsic characteristics into other facets.

Finally, it should be noted that it is possible to use to faceted scheme without necessarily applying the scheme to an entire taxonomy. For example, in the hierarchy below, it would be possible for FINANCIAL PRODUCTS to be organized according to a faceted scheme, but for the categories nested under CLOTHES, ELECTRONICS, and FOOD to be organized into non-faceted hierarchies.

PRODUCTS

- CLOTHES
- ELECTRONICS
- FINANCIAL PRODUCTS
- FOOD

Hierarchical Taxonomies—Conclusions

Hierarchical taxonomies have the following characteristics.

They consist of items or categories that are organized on a continuum. One side of the continuum is likely to be more densely populated with categories than the other side. This environment is particularly conducive to the identification of **one-to-many relationships** between categories.

They can be **strictly hierarchical**, where categories are permitted to have no more than one superordinate, or **polyhierarchical**, where categories are permitted to have multiple superordinates. Strict hierarchies tend to be organized according to a smaller set of principles than polyhierarchies.

The relationships used in a hierarchy can taken any form, as long as it is transitive in nature. Perhaps most common type of relationship is the **generic relationship**, which nests categories under general types. Another common transitive relationship is the **partitive relationship**, which nests parts under wholes. Other transitive relationships are also possible, such as the *commands* relationship, the *sold to* relationship, etc. Transitive relationships are often recursive, but exceptions can occur. For example, it is possible for a rock to be considered a chair, and for chairs to be considered furniture, without the rock being considered furniture.

Hierarchies can organize **items** or **topics**. Hierarchies that organize topics are used in the organization of books and other documents. If a hierarchy of topics is based on the generic relationship, each document that is organized in the hierarchy is likely to occupy a single hierarchical level. Hierarchies of topics are more likely to include concepts that are of fundamentally different types than are hierarchies of items.

Hierarchies based on the generic relationship have the following characteristics.

A category is considered to be included in its superordinate category or categories. However, a category may or may not be **included precisely** in its superordinate. RATS might be nested under VERMIN even if not all rats are vermin. Imprecise inclusion sets are more likely to occur when the categories involved are inherently fuzzy (e.g., VERMIN is a fuzzier category than RATS), since membership in the category is uncertain.

A subordinate category might be a **well-established species** of its superordinate, or it might represent an **ad hoc combination of characteristics** that are used to divide the superordinate. For example, subordinates of QUADRILATERAL might include SQUARE (a well-established concept), or RED QUADRILATERAL (an ad hoc combination of QUADRILATERAL and RED THING). Some users may prefer that a taxonomy consist only of well-established concepts, while other users appreciate categories that are plausible for the domain, but which have not been blessed with a well-known linguistic label.

Generic hierarchies can be made with the benefit of a **faceted scheme**. Faceted schemes are very useful in ensuring that categories in a hierarchy are organized according to a consistent set of principles. They are also useful in distinguishing between principles that are intrinsic to the item and principles that are based on an item's role or function. By their very nature, hierarchies that are based on faceted schemes are always strictly hierarchical and never include imprecise inclusion sets. However, the resulting categories tend to be ad hoc combinations of characteristics rather than well-established concepts. By changing the order in which the facets apply, it is possible for different users to derive different hierarchies from the same faceted scheme.

Identifiers for Items and Classes

In many information systems, an item must receive a unique identifier. In the case of financial products falling with the CFTC's Final Rule, it will be the Unique Product Identifier (UPI). Therefore, it may seem reasonable to infer that an item can appear in only one location, as per a strict hierarchy. However, in addition to a number that is used to uniquely identify the item, the item may also be given a separate set of identifiers that refer to its position in an organizational scheme. There is no theoretical limit to the number of location identifiers that an item can receive. For example, in the Medical Subject Headings (MeSH), the heading for MOLECULAR COMPUTERS has the unique identifier D039301, plus two tree numbers: L01.224.230.260.315 (COMPUTERS) and L01.224.300 (COMPUTING METHODOLOGIES) (NLM/NIH, 2012). The tree numbers indicate the two positions in the polyhierarchical structure in which MOLECULAR COMPUTERS appears. This situation is roughly analogous to a person who maintains multiple addresses: a home address, a work address, etc. The person is unique, but can be found at multiple locations.

Use Case and Implications

In this section, we discuss several use cases to illustrate the operational implications of choosing one classification model versus another. We posit nine specific use cases related to macroprudential supervision, though we again emphasize that these are not the only use cases for financial product classification. We draw from the questions raised in the Canadian Securities Administrators CSA Consultation Paper 91-402 on Derivatives Trade Repositories (Canadian Securities Administrators, 2011):

1. Aggregate notional data for all contracts traded or settled in Canadian dollars, including a breakdown by reference entity and/or sector.
2. A list of the top counterparties trading Canadian dollar denominated contracts with each counterparty's aggregate notional position and aggregate position by contract type.
3. A list of the top counterparty positions for each of the largest financial groups in Canada.

4. Aggregate notional data for contracts written on Canadian-domiciled corporations (reference entities), including a list of the top aggregate notional counterparty positions for contracts written on each firm.
5. A list of the top counterparties' aggregate notional positions where the contract references the debt of the government of Canada.
6. A list of top counterparties' aggregate notional positions where the contract references a specific commodity.
7. A list of the top counterparties' aggregate notional positions where contracts reference the debt of one of the ten largest Canadian financial groups.
8. Data on the overall level of activity of each of the Canadian banks in each asset class.
9. Each of the Canadian bank's overall positions in specific products within an asset class.

We focus our attention on use cases two, four, five, six, seven, eight, and nine as all of these cases rely upon a system of classification for financial contracts.

Beginning with use case two, we note that we require a system of classification that permits identification and grouping of all Canadian dollar denominated contracts. This case also requires that we further segregate these contracts by contract type. It is easy to see that a strict hierarchical system, such as that proposed by ISDA, will fall short on meeting this requirement. First of all, the issue of the currency type in which the contract is denominated is likely to be of subordinate interest, thus producing a classification scheme where contract types is the superordinate and currency type of contract is the subordinate. In such a scheme the grouping together of contracts by the subordinate feature—currency type—can be cumbersome (and depending on the technology of implementation, practically impossible). In this case, a polyhierarchical structure may work better. We could then expect to see CURRENCY TYPE having a number of different superordinates:

By contract type: FUTURES, OPTIONS, SWAPS, etc,

By underlying type: ASSET, FX, COMMODITY, CREDIT, INTEREST RATE, etc.[4]

This arrangement, however, still does not resolve the difficulty of grouping by currency type first, followed by contract type, followed by underlying asset type. For this to work most effectively, it would be better to use a faceted scheme wherein currency types, contract types and asset types exist as faceted categories which can be combined together in order of preference depending on the analytic task—in this case, with currency taking the first position in the hierarchy.

In the fourth use case (which seemingly implies but does not state that the contracts in question are credit default swaps, since only these are generally described as being "written on" some company—most derivatives are written by some company, not on it), the classification system must support grouping of contracts by the country in which the reference entity is domiciled (e.g., Canada). The challenges in this use case are similar to those delineated in the discussion of use case two, with a similar result: a faceted classification scheme provides the required flexibility and analytic capabilities.

Use case five is challenging not only from the perspective of classification, but also because it is semantically underspecified. For example, does "contract references the debt of the government of Canada" (GOC) refer to the underlying or to collateral? Both could be relevant for the purposes of systemic risk analysis. Moreover, where the contract relates to an index, it would be necessary to have sight of the asset composition of the index in order to meet the requirement. Assuming that the requirement is to prepare a list of contracts referencing GOC debt as both the underlying and as collateral, and that we are also interested in indices that include GOC debt, we can now consider various classification options. A strict hierarchy by contract type is not going to be of much assistance, unless the type of collateral is a level in the hierarchy. Even so, it is highly unlikely that a hierarchy that is organized by contract type would include the heading GOC debt, which is subdivided by various collateral types that also reference GOC debt. Since we are more interested in grouping types of contracts that reference GOC debt whether as underlying or collateral, we need a mechanism to bring these contracts together in the most efficient manner. A strict hierarchical system would require a significant amount of data wrangling before the desired list could be achieved. The next option is a polyhierarchical approach. We could have a system where GOC debt is subordinate to both contract type and collateral type, thus we must still decide how to handle index composition.[5] It would not be possible to make GOC Debt Index a part of a class called Index—Debt Index—GOC Debt Index that also contained sub-headings Equities Index, and so on, because a single index portfolio may hold all of these asset types. We would, therefore, be unable to classify or group the index portfolio under the single class Index—Debt Index—GOC Debt Index as it would improperly characterize the index portfolio and render opaque and unanalyzable the other features of that portfolio. The latter problem does not disappear with a faceted scheme, but the added flexibility of a faceted approach means that one could more likely create an "on-the-fly" analytic classification classes to address the need to segregate indices containing GOC Debt from those that do not. For example:

ASSET TYPE

FX

EQUITY

DEBT

GOC DEBT

INDEX

[CLASSIFIED BY INDEX STRUCTURE]

CONTAINS DEBT

CONTAINS GOC DEBT

$500,000

$1M

$1.5M

CONTAINS EQUITY

(etc.)

COLLATERAL

CASH

EQUITIES

DEBT

GOC DEBT

Use case 6 presents essentially the same challenges, but in reference to commodities instead of GOC Bonds.

Based on the analysis of the Canadian use cases, we conclude that a faceted scheme has the requisite flexibility to meet the Canadian requirements. We also wish to emphasize that the utility of this approach is extended when one considers that analytic requirements, especially in stressed conditions or unprecedented market disruptions, may be unpredictable and emergent. Under such conditions, it is especially important that the structure of the financial product classification scheme be capable of supporting new analytic requirements. This would not be the case with a strict hierarchical model, but would be achievable with a faceted scheme.

Another feature to tease out here is that for some analytics, and in particular for risk aspects, what is of interest may not be features intrinsic to the thing itself (facts which may be established with reference to the instrument contract) but extrinsic features such as whether the holder also holds some related asset (“naked”—no collateral—versus “covered”—backed by collateral—credit default swaps), whether the issuer (or a subsidiary or parent) has some unrelated but systemically important position with respect to the holder and so on.

Use case eight is interesting as it calls for grouping by asset class as opposed to contract type. Clearly then, any strict hierarchical scheme designed on a principle of division by contract type is going to be unsuited to the task at hand. A strict hierarchy divided by underlying asset types would work (assuming this is what is meant by “asset class” in this context); however, it would not work for the other use cases. Thus, we would be forced to adopt different classification systems for different analytic tasks, which is not an attractive prospect given the added complexity. In this case, it would make most sense to have a strict hierarchical faceted scheme that permitted us to combine classes divided by contract type with those divided by asset type.

The final use case, number nine, requires that we classify or group our objects of analysis first by asset type and then by product type. A strict hierarchical scheme would run us directly into the problem we encountered earlier with convertible bonds. A polyhierarchy would resolve this issue, but a faceted scheme would allow for maximum flexibility as we could combine lists of asset types with lists of product types to produce a combination that would cover assets that are both bonds and those that are bonds traded as equities and that are of the product type "Convertible Bond."

Recommendations

This paper has discussed the theory of classification and the strengths and limitations of various approaches to classification as a contribution to the issue of establishing and assigning UPIs to categorize or describe swaps with respect to the underlying products referenced in them, allowing regulators to aggregate, analyze, and report swap transactions by product type, and also to enhance position limit enforcement and real time reporting as a means of identifying and managing financial systemic risk. Rather than definitively recommending one approach to financial product classification, we prefer to acknowledge the limitations and applicability of approaches in different contexts. The facets used may depend on the context in which the instrument is considered. For example, the 'available for sale and held to maturity' (AFS/HTM) distinction may be important in an accounting context, but irrelevant to a risk manager. Similarly, if the object of analysis is to discover how the behavior of a contract might contribute to systemic risk under certain market conditions, it is more likely that one would need to infer this information from combinations of facets. A purely faceted approach is: 1) very flexible and clean, and 2) tends to exclude certain concepts. Separation of the registration or I.D. number for a product instance from the classification number is recommended, however, regardless of the other choices.

It is worth noting that a formal ontology provides the means to capture meaningful concepts across as many facets as are considered to be worth naming and defining in the domain of discourse. Other technologies, when used to try to support classification, identification and so on, will unnecessarily constrain the number of available options to those that that specific technology supports—including typically the imposition of a strictly hierarchical taxonomy on the subject matter or the imposition of one and only one order of precedence of the different facets by which the subject matter may be classified. This approach may thus prove too brittle for effective risk analytics.

Here we have laid out the foundations for an approach to classification of financial products, with emphasis on financial systemic risk as a use case. We suggest that the core principles set out here may be applied more broadly than shown here, to embrace other facts of relevance such as the combinations of instruments in portfolios, effects of market behavior and so on. We also do not attempt to address issues of implementation and operation of the financial products classification system, however readers may be concerned about the cost of supporting a faceted approach to product classification given the innumerable categories that could presumably be created and creatively combined to support product valuation and risk analytics. While cost is an issue, we think that a move toward the use of semantic technologies by market participants will alleviate much of the cost burden because it is technically possible to automatically tag (using FIBO, for example) elements within a financial contract in order to classify them according to a faceted scheme. The details of how a financial products classification system based on a faceted approach could be developed and operated, we must leave to a follow-up paper.

It is worth noting that many important stakeholders in the discussion about financial products classification have not participated in the work that resulted in this paper. It is our hope that through dissemination of this work, other stakeholders will be able to engage in this discussion as well.

References

Aitchison, J., Gilchrist, A., & Bawden, D. (2000). Thesaurus Construction and Use: A Practical Manual, 4th ed. New York: Europa Publications.

ANSI/NISO (2005). Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies ANSI/NISO Z39.19. Retrieved December 5, 2012, from http://www.niso.org/kst/reports/standards?step=2&gid=&project_key=7cc9b583cb5a62e8c15d3099e0bb46bbae9cf38a

Bisias, D., Flood, M., Lo, A., & Valavanis, S. (2012). A survey of systemic risk analytics. U.S. Department of Treasury, Office of Financial Research No. 0001. Retrieved on December 5, 2012, from http://ssrn.com/abstract=1983602

Broughton, V. (2006). The need for a faceted classification as the basis of all methods of information retrieval. Aslib Proceedings, 58(1/2), 49–72.

Brush, S. (2010). CFTC approves swaps definition triggering Dodd-Frank rules, Bloomberg. Retrieved November 19, 2010, from http://www.bloomberg.com/news/2012-07-10/cftc-votes-4-1-to-approve-swap-definition-starting-overhaul-1-.html

Canadian Securities Administrators, Derivatives Committee (2011). CSA Consultation Paper 91-402, Derivatives: Trade Repositories. Retrieved November 18, 2012, from http://www.osc.gov.on.ca/documents/en/Securities-Category9/csa_20110623_91-402_trade-repositories.pdf

Cheti, A., & Paradisi, F. (2008). Facet analysis in the development of a general controlled vocabulary. Axiomathes, 18(2), 223–241.

Committee on Payment and Settlements System and Technical Committee of the International Organization of Securities Commissions (2011). Report on OTC derivatives data reporting and aggregation requirements. Retrieved November 18, 2012, from http://www.bis.org/publ/cpss96.pdf

Commodities and Futures Trading Commission (2011a). Swap data recordkeeping and reporting, 17 CFR Part 45 RIN 3038-AD 19. Retrieved December 5, 2012, from http://www.cftc.gov/ucm/groups/public/@newsroom/documents/file/federalregister122011b.pdf

Commodities and Futures Trading Commission (2011b). Final rule on swap data: Recordkeeping and reporting requirements. Retrieved November 18, 2012, from http://www.cftc.gov/ucm/groups/public/@newsroom/documents/file/sdrr_factsheet_final.pdf

Commodities and Futures Trading Commission (2012). Real-time public reporting of swap transaction data. Federal Register, 77(5), 1182–1266. Retrieved December 19, 2012 from http://www.cftc.gov/ucm/groups/public/@lrfederalregister/documents/file/2011-33173a.pdf

DeSanctis, G. & Poole, S. M. (1994). Capturing the complexity in advanced technology use: Adaptive structuration theory. Organization Science, 5(2), 121–147.

Dewey, M. (1876). A classification and subject index for cataloguing and arranging the books and pamphlets of a library [Dewey Decimal Classification] [Electronic version]. Retrieved December 5, 2012, from http://www.gutenberg.org/files/12513/12513-h/12513-h.htm

Dextre Clarke, S. G. (2001). Thesaural relationships. In C. A. Bean & R. Green (Eds.), Relationships in the Organization of Knowledge (pp. 37–52). Boston: Kluwer Academic Publishers.

Dmitrii M. (1869). On the relationship of the properties of the elements to their atomic weights. Zhurnal Russkoe Fiziko-Khimicheskoe Obshchestvo 1, 60–77.

Farkas, D. K., & Farkas, J. B. (2002). Principles of Web Design. New York: Longman.

Feigenbaum, E. A., McCorduck, P. (1983), The Fifth Generation, Reading, MA: Addison-Wesley.

Financial Products Markup Language (FpML) Organization (2012). OTC taxonomies. Retrieved November, 18 2012, from http://www.fpml.org/samples/taxonomiesandupi.html

Financial Stability Board and the International Monetary Fund (2009). The financial crisis and information gaps. Retrieved December 5, 2012, from http://www.imf.org/external/np/g20/pdf/102909.pdf

Financial Stability Board and the International Monetary Fund (2011). The financial crisis and information gaps: Implementation progress report. Retrieved December 5, 2012, from http://www.imf.org/external/np/g20/pdf/063011.pdf

Flood, M., Mendelowitz, A., & Nichols, W. (2012). Monitoring financial stability in a complex world. In V. Lemieux (Ed.), Financial Analysis and Risk Management: Data Governance, Analytics and Life Cycle Management (pp.15-40). Berlin: Springer Verlag.

Gentner, D., & Markman, A. B. (1994). Structural alignment in comparison: No difference without similarity. Psychological Science, 5(3), 152–158.

Giddens, A. (1984). The Constitution of Society: Outline of the Theory of Structuration. University of California Press, Berkeley and Los Angeles.

Glushko, R. J., Maglio, P. P., Matlock, T., & Barsalou, L. W. (2008). Categorization in the wild. Trends in Cognitive Sciences, 12(4), 129–135.

Gruber, T. R. (2008). Ontology. In L. Liu and M. T. Ozsu (Eds.), Encyclopedia of Database Systems. Springer-Verlag. Retrieved April 6, 2012, from http://www-ksl.stanford.edu/kst/what-is-an-ontology.html

Guarino, N. (1998). Formal ontology and information systems. In Proceedings of International Conference on Formal Ontology in Information Systems (FOIS'98) (pp. 3–8). Amsterdam: IOS Press.

Hampton, J. A. (1982). A demonstration of intransitivity in natural categories. Cognition, 12(2), 151–164.

Hedden, H. (2010). The Accidental Taxonomist. Medford: NJ: Information Today.

Huzzard, T., Ahlberg, B.M., & Ekman, M. (2010). Constructing interorganizational collaboration. Action Research 8(3), 293–314.

InterPARES 3 Project (2012). Retrieved December 20, 2012 from http://www.interpares.org/ip3/ip3_index.cfm

Jacob, E. K. (2004). Classification and categorization: A difference that makes a difference. Library Trends, 52(3), 515–540.

Johnson, M. D. (1984). Consumer choice strategies for comparing noncomparable alternatives. Journal of Consumer Research, 11(3), 741–753.

Kaufman, G. (2000). Banking and currency crises and systemic risk: Lessons from recent events. Economic Perspectives, 9–28.

Kwasnik, B. H. (1999). The role of classification in knowledge representation and discovery. Library Trends, 48(1), 22–47.

Lambe, P. (2007). Organising Knowledge: Taxonomies, Knowledge and Organizational Effectiveness. Oxford: Chandos Publishing.

Lemieux, V., & Limonad, L. (2011). What ‘good’ looks like: Understanding records ontologically in the context of the global financial crisis. Journal of Information Science, 37(1), 29–39.

Linnæus, C. (1735). Systema Naturæ, Sive Regna Tria Naturæ Systematice Proposita Per Classes, Ordines, Genera, & Species. Lugduni Batavorum: Haak.

Loehrlein, A. (2011). An examination of interdisciplinary theory between cognitive categorization and knowledge organization. In Proceedings of the Third North American Symposium on Knowledge Organization, Toronto, June 16-17, 2011.

Morville, P. & Rosenfeld, L. (2006). Information Architecture for the World Wide Web, 3rd ed. Cambridge: O’Reilly Media, Inc.

Murphy, G. L. (2004). The Big Book of Concepts. Cambridge, Mass.: MIT Press.

Orlikowski, W. J. (1992). The duality of technology: Rethinking the concept of technology in organizations. Organization Science, 3(3), 398–427.

National Library of Medicine and National Institutes of Health (2012). Medical subject headings. Retrieved December 6, 2012, from http://www.nlm.nih.gov/mesh/

Pena, P. A., Patino, A., Palacio, C., Lochmuller, C., Ardila, H., & Villa, S. (2012). A fuzzy model of adaption for Operational Risk assessment in companies of the financial sector: Operational Risk. In 2012 7th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1 –6).

Ranganathan, S. R. (1962). Elements of Library Classification. Bombay: Asia Publishing House.

Reinhardt, C. M., & Rogoff, K. (2009). This Time Is Different: Eight Centuries of Financial Folly. Princeton University Press, Princeton.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. Cognitive Psychology, 7(4), 573–605.

Smith, E. E., & Medin, D. L. (1981). Categories and Concepts. Cambridge, Mass.: Harvard University Press.

Sowa, J. F. (2000). Knowledge Representation: Logical, Philosophical and Computational Foundations. Pacific Grove, CA: Brooks/Cole.

Studer, R., Benjamins, V. R., & Fensel, D. (1998). Knowledge engineering: Principles and methods. Data & Knowledge Engineering, 25(1), 161–197.

Suryanto, H., & Compton, P. (2000). Learning classification taxonomies from a classification knowledge based system. In S. Staab, A. Maedche, C. Nedellec, and P. M. Wiemer-Hastings (Eds.), ECAI'2000 Workshop on Ontology Learning, Proceedings of the First Workshop on Ontology Learning OL'2000, Berlin, Germany, August 25, 2000. Held in conjunction with the 14th European Conference on Artificial Intelligence ECAI'2000, Berlin, Germany. CEUR Workshop Proceedings 31 CEUR-WS.org 2000.

Taylor, A. G. (2006). Introduction to Cataloging and Classification, 10th ed. Westport, CT: Libraries Unlimited.

U.S. Congress (2010). Dodd-Frank Wall Street Reform and Consumer Protection Act. Public Law 111-203 H.R. 4173, 124 Stat. 1376. US Government Printing Office, Washington, D.C.. Retrieved November 18, 2012, from http://www.gpo.gov/fdsys/pkg/PLAW-111publ203/html/PLAW-111publ203.htm

Vickery, B. C. (1960). Faceted Classification: A Guide to Construction and Use of Special Schemes. London: Aslib.

Vickery, B. (2008). Faceted Classification for the Web. Axiomathes, 18(2), 145–160.

W3C (1994a). OWL Web Ontology Language: Reference. Retrieved December 5, 2012, from http://www.w3.org/TR/2004/REC-owl-ref-20040210

W3C (1994b). RDF Vocabulary Description Language 1.0: RDF Schema. Retrieved December 5, 2012, from http://www.w3.org/TR/2004/REC-rdf-schema-20040210

Waddell, S. (2012). Global finance as an action research domain: Testing the boundaries. Action Research, 10(1), 40–60.

Yates, J., & Orlikowski, W. J. (1992). Genres of organizational communication: A structurational approach to studying communication and media. Academy of Management Review, 299–326.

Ye, K., Yan, J., Wang, S., Wang, H., Miao, B. (2011). Knowledge level modeling for systemic risk management in financial institutions. Expert Systems with Applications 38(4), 3528–3538.

Footnotes

[1]The dates of the so-called global financial crisis are open to debate; for example, Reinhardt and Rogoff (2009) refer to the crisis as the financial crisis of the late 2000s. We have chosen to standardize the dates to 2007-2009.

[2]For example, Enterprise Data Management Council's Financial Industry Business Ontology (FIBO) <http://www.edmcouncil.org/financialbusiness>

[3]For example, if we conclude that all tigers possess an essential tiger-ness, then we must conclude that at some point in the evolutionary process, a non-tiger must have given birth to a tiger.

[4]Note that this assumes that when the referenced document says "contract" it means any kind of contract, and not for example "any kind of swap" or "any kind of interest rate derivative". Sometimes the top level classification is implied by context and not made explicit.

[5]Particularly since "Index" may also be a type of underlying—seemingly hidden away alongside interest rates in the ISDA taxonomy.

Tables and Figures

Table 1. CFTC's final rules concerning unique product identifiers in swaps data reporting and recordkeeping.

Figure 1. Data model of the proposed ISDA UPI scheme.

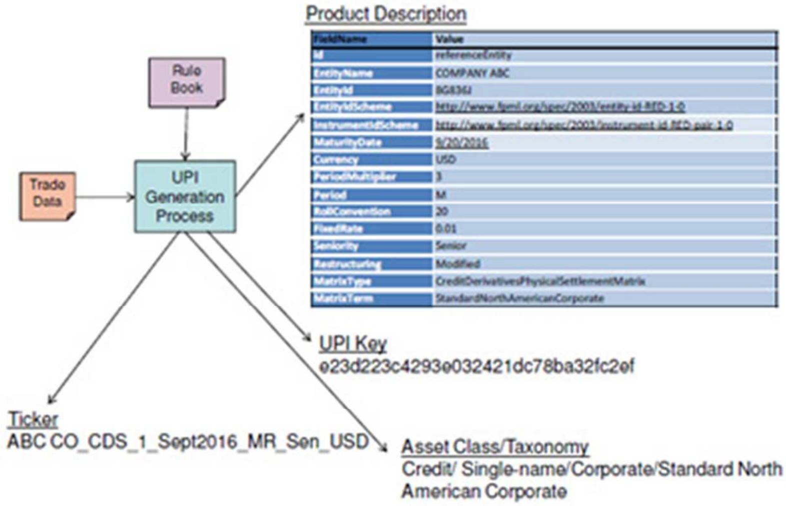
Figure 2. The ISDA product taxonomy showing commodities. The taxonomy also covers credit derivatives, interest rate, foreign exchange and equity-based products.

Figure 3. Graphic illustrating the strict hierarchical structure of the ISDA classification scheme.

3. Final Rule: § 45.7

After considering the comments and input received concerning the UPI and product classification system, the Commission has determined that, as called for in the Notice of Proposed Rulemaking, the final rule provides that each swap subject to the Commission's jurisdiction must be identified in recordkeeping and swap data reporting pursuant to this part by means of a unique product identifier and product classification system acceptable to the Commission, when such an identifier and classification system are designated by the Commission for this purpose. The unique product identifier and product classification system will be required to identify and describe the swap asset class and the sub-type within that asset class to which the swap belongs, and the underlying product for the swap, with sufficient distinctiveness and specificity to enable the Commission and other financial regulators to fulfill their regulatory responsibilities.

The final rule provides that the Commission will determine when a unique product identifier and product classification acceptable to the Commission and satisfying these requirements is available, and when it so determines will designate the unique product identifier and product classification system for use in compliance with this part, making this designation in a Commission order. The final rule requires registered entities and swap counterparties subject to the Commission's jurisdiction to use the unique product identifier and product classification system in compliance with this part when this designation is made. Prior to this designation, each registered entity and swap counterparty must use the internal product identifier or product description used by the SDR in all recordkeeping and swap data reporting pursuant to this part.

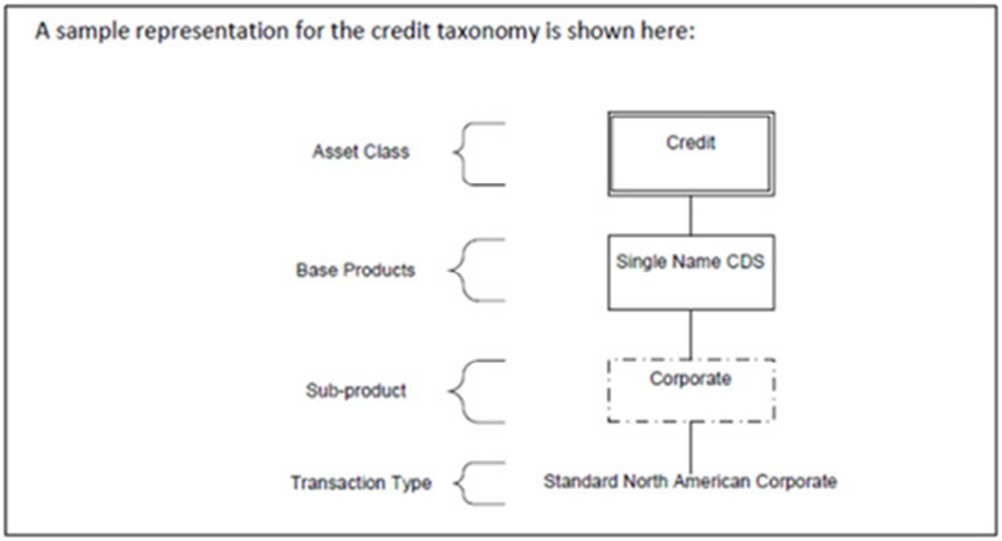


152x90mm (72 x 72 DPI)

Commodity Taxonomy

#	Asset Class	Base Product	Sub-Product	Transaction Type	Settlement Type
1	Commodity	Metals	Precious	Spot Fwd	Physical
2	Commodity	Metals	Precious	Swap	Cash
3	Commodity	Metals	Precious	Option	Cash
4	Commodity	Metals	Precious	Option	Physical
5	Commodity	Metals	Precious	Loan Lease	Cash
6	Commodity	Metals	Precious	Loan Lease	Physical
7	Commodity	Metals	Precious	Exotic	
8	Commodity	Metals	Non Precious	Spot Fwd	Physical
9	Commodity	Metals	Non Precious	Swap	Cash
10	Commodity	Metals	Non Precious	Option	Cash
11	Commodity	Metals	Non Precious	Option	Physical
12	Commodity	Metals	Non Precious	Loan Lease	Cash
13	Commodity	Metals	Non Precious	Loan Lease	Physical
14	Commodity	Metals	Non Precious	Exotic	

208x71mm (72 x 72 DPI)



189x101mm (72 x 72 DPI)