TAC SUB-COMMITTEE ON DATA STANDARDIZATION

UPI and LEI Working Group

Unique Product Identifier (UPI) Approach and Implementation

December 13th 2011

1   Contents

2   Recommendations

ISDA and an inclusive group of market participants including buy-side, CCPs, potential SEFs and SDRs, affirmation platforms and other market facilities have created a proposal for generating UPIs in the marketplace.   TAC members of the Sub-Committee on Data Standardization, UPI and LEI working group support the ISDA UPI proposal and the underlying approach.

The approach is to create a full product representation in a decentralized manner. It avoids the need to build and maintain a central issuing registry for UPIs that would require integration with all market participants and be very costly. The recommended approach is cost effective, leverages existing market standards, scalable and can be used for all asset classes and products.

The approach is detailed in a specification document (UPI Generation specifications) and asset class-specific rulebooks which are currently under construction. The specification describes the procedures used to generate the product representation for any group of price-comparable derivative contracts, with emphasis on the standardized areas of the current OTC market.   F*ull product representation* will include:

- Asset Class/Taxonomy
- Product Description
- UPI Code
- Ticker (to be determined by each asset class where appropriate)

The asset class/taxonomy will allow regulators to classify products for purposes of reporting and monitoring systemic risk. The product description will be human-readable expression of the product, machine readable and will allow for ad-hoc querying or reporting.   The UPI code will be hash generated and effectively unique, opaque identifier that will be machine readable and necessary for IT processing.

For standardized products it is expected the full product representation (shown above) would become available. For exotic or bespoke products where market standard representations do not exist, and where price comparison does not make sense it is expected that they would be represented by the asset class and taxonomy.

The approach is to convert a standard FpML trade representation into a normalized product representation (product info-set) that when input through a hashing algorithm will generate the full product representation including the unintelligible identifier (the UPI code) for swaps and other derivative contracts. If tickers were desired (for the most standardized and liquid markets), these would need to be defined in this process.

Product rulebooks for each asset class are under construction that will detail the process by which a standard trade representation is transformed into a set of discriminating information which, once normalized, can be input into the hashing algorithm to generate the UPI code.

While ISDA will be the publishing entity to make the new approach freely available to all market participants, the product rulebooks will be owned by the ISDA product steering committees.

At this time the approach is being vetted with each of the ISDA product steering committees to ensure the approach is fully understood and supported by the industry. The ISDA product steering committees are looking to understand the costs and ensure there is no impediment to product innovation or trading. In addition there is a discussion around the level of granularity needed for UPIs while still ensuring that there is not a market disruption due to the UPIs being disclosed. There also has been much discussion over the past six months on using a ticker alone to serve as the UPI. The TAC feels that, while the ticker is a useful component of the product representation, that the hashed UPI code will be necessary to maintain the IT advantage of computing readability. The TAC has recommended to the UPI Technology and Data Sub-committee, which is reviewing the technical aspects of UPI, to comeback with a full recommendation on this matter.

There are challenges to full-scale implementation of UPIs within the market place. The exercise that will be undertaken by the asset class rulebook groups to identify all of the price-forming attributes that define the product is a considerable endeavor. The industry-wide regulatory framework will need to be carefully planned, coordinated and tested prior to UPI implementation.

In conclusion, the TAC believes significant work and progress has been made in defining the UPI framework and considerations and would highlight the large change this will create for the industry. The proposal will go through a vetting process over the next couple of months to reach a point of industry consensus and approval driven via the ISDA framework and process. It is expected that UPIs for the most liquid, standardized parts of the market will become available for integration with the overall industry regulatory reporting framework by Q3/Q4 2012. The TAC fully supports the implementation of UPIs into the overall industry regulatory reporting framework.

3   Background

The CFTC has issued proposed rules that require unique product identifiers to increase transparency in the OTC markets. In response to the regulatory requirements, ISDA initiated the development of an approach towards UPI that meets the regulatory requirements and furthers the industry goals of standardization and automation. The constituents who have been involved with developing the approach along with the dealers include: buy-side, potential SEFs and SDRs, CCPs, affirmation platforms and other market infrastructure.

The current approach is being reviewed by the ISDA product steering committees to ensure the industry fully details the benefits/costs and the potential impact (if any) to the business. It is expected that the vetting process will be completed in January 2012.

In order to understand the challenges, precedents and next steps towards the implementation detailed in this paper, we first outline the current approach to generating UPIs.

In April 2011, ISDA released a whitepaper entitled "Product Representation for Standardized Derivatives,"[1] which offered an initial approach for creating unique product identifiers for OTC derivatives.

---

[1] http://www2.isda.org/attachment/MzQyMA==/ISDA%20UPI%20Implementation%20Plan%20Final.pdf

The white paper release was followed by a comment period. Additionally, ISDA put a governance structure in place to manage the further development and implementation of UPI for OTC derivatives. The ISDA Data Working Group and the FpML Standards Committee have overall ownership of the UPI work. A UPI Steering Committee was formed to oversee the project and the work in the various working groups on a day to day basis[2] and to provide direction and guidance for the project and the asset class specific groups that were created.

Changes in the approach based on the comments received have been further developed in the working groups and are codified into the "UPI Generation specification", produced by the tech and data subcommittee. While applicable to all OTC asset classes, each asset class will need to produce a rulebook, work that just has started, that details the set of trade attributes which are price forming, and specify how these attributes will be grouped into products.

The creation of the Asset Class-specific rulebooks will be a significant undertaking. Participants will be tasked with the review of each attribute to determine if the attribute is price-forming. The product rulebooks will be available to all market participants and will be approved by each of the ISDA asset class Steering Committees.

4   Centralized vs. Decentralized Issuing Registry

The UPI Steering committee determined that the approach described in the April 2011 whitepaper, calling for a central, issuing UPI registry would not be the right approach to implementing UPIs in the marketplace. A central issuing UPI registry would in effect become an industry-wide point of failure. In addition, the time to build and complexities around the full integration with all industry participants were drawbacks to this approach.

The current approach towards generating UPIs, as detailed below, decentralizes the creation of UPIs.

5   Approach to Generation of Unique Product Identifiers

The current UPI approach is detailed in the UPI Generation specification[3], which describes the different types of identifiers and the process to generate them from a given set of data. This document is accompanied by the asset class-specific rulebooks which are currently under development.

The approach avoids the dependency on a central computerized registry or service to be involved at run-time in the generation of a UPI for inclusion in a trade price report. Instead, the procedures required to generate UPIs are made available to all market participants.

The output of the UPI code generation process is an alphanumeric code that represents the effectively unique identifier. In addition, a product description will be produced. A ticker can also be developed for sections of the market where they are warranted.

The *full product representation* can be thought of as (shown for illustrative purposes):

---

[2] ISDA taxonomy and UPI implementation plan:
http://www.fpml.org/news/ISDA%20UPI%20Implementation%20Plan%20Final.pdf

- Taxonomy- Credit/Standard North American Corporate
- UPI Code- e23d223c4293e032421dc78ba32fc2ef
- Product description- Reference Entity FORD MOTOR COMPANY-RED 3H98A-1.00%-Following-2016_Sept_20-Standard North American Corporate_Modified Restructuring_Senior_USD
- Ticker- FORD_CDS_1_Sept2016_MR_Sen_USD

The taxonomy, which includes the asset class, will allow for regulators and market participants to report in a more systematic way, allowing for better data aggregation, leveraging the ISDA OTC Taxonomy work.

The UPI as an alphanumeric code will enable machine processing, including post-trade price comparisons for the standardized parts of the market. The UPI code also will provide a unique identifier which could eventually be leveraged by market participants in a RFQ process, trade workflow and post-trade processing.

The use of a hashing algorithm guarantees that with very high certainty that each UPI code will be effectively unique. In addition, the approach will lead to identifiers that will remain immutable. The trade-off required to produce a unique, immutable identifier is a deficiency in human readability or ability of a market expert to consume and understand the UPI code.

The product description generated as part of the process provides the details of the data that details the unique product. It is also machine and human readable and will allow regulators the ability to drill-down into the data in a custom manner (see also section 6 on UPI dissemination).

It is anticipated that tickers will be created for the most standardized and liquid parts of the market. This is viewed as an important innovation that will occur but is not central to the implementation of UPIs as detailed in the UPI Specifications and product rulebooks.

The product rulebooks are currently being developed for each of the asset classes, with an initial focus on Credit and Rates. The product rulebooks will detail the attributes that represent the price forming attributes that form the product representation. The UPI Generation specification will detail how to convert the standard trade representation into the product representation (including all components described above).

The following figure illustrates the procedure[3]:

_____
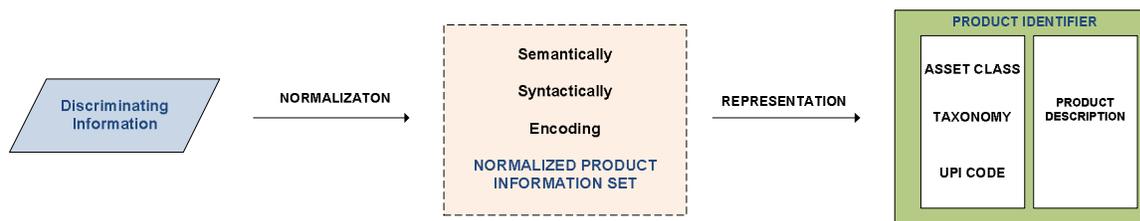
[3] Unique Product Identifier (UPI) Generation DRAFT Version 0.2, 25 September 2011

5.1 UPI Generation Specification

The Unique Product Identifier (UPI) Generation specification contains the procedures used to generate a UPI code for any group or class of price-comparable derivative contracts, with an initial emphasis on centrally cleared derivatives.

The UPI Specifications define a common framework or approach, irrespective of asset-class, while taking account of the specific characteristics of different asset classes when it comes to the application of that approach.

They build upon pre-existing market standards and conventions. For example they rely on pre-existing public industry standards for data representation by using FpML. The Specification - for the generation of the alphanumeric UPI code -demands the use of a public standard cryptographic hash algorithm that is widely available in commonly-used technology suites.

UPI generation is a 5 step process[4]:

1. Convert a standard trade representation to a product describing "Reference fragment"
2. Standardize the format of data associated with each tag (e.g. "5%" vs. "0.05")
3. Sort the data attributes into a standard order
4. Convert standard representation into list of names and values
5. Generate the UPI code by inputting the product description into an algorithm.

5.2 Asset Class Rulebooks

The starting point for generating the UPI code for all asset class rulebooks will be a full trade representation expressed in FpML. The rulebooks will define the price-forming attributes that, after normalization, will represent the normalized product information set for a particular product.

The first step is to transform the full trade representation into a fragment that captures the product present in that trade. This is done by reviewing each data attribute and characterizing it into one of three categories:

- Not included in product description (always)
- Included in product description (always)
- Included in product description (conditionally) - For each conditional attribute, rules must be defined which determine when an attribute is retained and when it is eliminated.

5.2.1 Rulebook guidance[5]

The Guidance for Asset-Class Working Groups was produced to detail the objectives, success criteria, scope and approach for the product rulebook working groups. While each of the asset classes needs to create their own rulebook, the guidance document ensures a consistent approach across asset classes by defining the tasks required of each asset class rules development group, and provides guidance on common issues.

---

[4] UPI Development: Guidance For Asset-Class Working Groups Version 0.2, 9 November 2011
[5] UPI Development: Guidance For Asset-Class Working Groups Version 0.2, 9 November 2011

An example of this specific guidance is shown here:

Conditional Inclusion

When an attribute is conditionally included, objective rules must be developed to evaluate its inclusion in each specific instance.

- Rules should be written in an objective "IF (condition), THEN include, ELSE, exclude" or the inverse "IF (condition), THEN exclude, ELSE, include"
- Rules may reference the presence or absence of other data attributes
- Rules may reference the value associated the data attribute
- Rules may **not** reference the following:
  - Taxonomy
  - Trading venue
  - Expected clearing / settlement arrangements

5.2.2   Creation of Normalized Product Information Set

The following diagram describes the process for the creation of the normalized product information set that underpins each distinct product:

Process for determining price forming attributes:
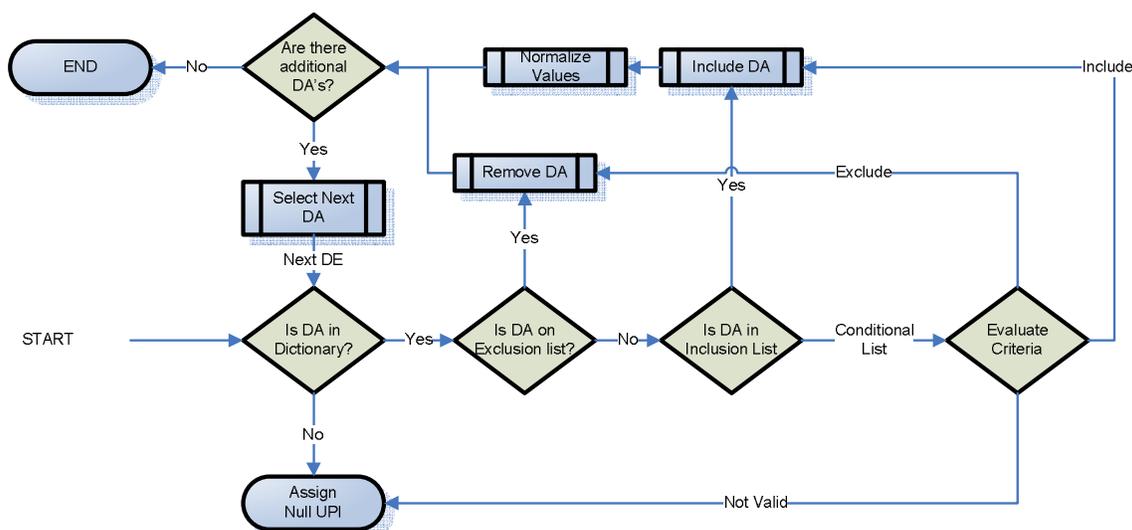
"DA" = data Attribute

Description:

- Determine the starting representation in the form of an FpML representation.
- Identify data attributes which will be covered by the initial version of the Rule Book.
  - Note – This includes both included and excluded attributes.
- Document the vanilla products types covered by the selected data attributes.
  - Note: These are the minimum products which the specification will cover. Variations of the "in scope" data attributes will likely also cover simple variations of these standard products.

- Identify data attributes which are always removed.
- Identify data attributes which are always retained.
- For remaining data attributes, draft conditional rule.
- For each retained or conditional attribute, define the format for the data associated with each data attribute.
  - E.g. are Rates expressed as %, bps, or decimals? Are leading 0's present? Are trailing 0's present?
- Identify a methodology for specifying tenors, if applicable.
- Identify a methodology for specifying price, if applicable.
- Identify any "infinitely variable terms", if any.
- Identify any terms which are product determinative, but whose specific values are not.

6   UPI Generation and Dissemination

The UPI Tech and Data Sub-committee have commissioned a whitepaper to analyze the options for the generation and dissemination of UPIs.   Dissemination refers to a necessary service to provide end-users the ability to look up the product representation if given a UPI code. The whitepaper will detail potential options and pros and cons of potential workflows. Below is a summary of the main points we expect to address on the final paper on UPI generation and dissemination.

6.1   UPI Generation Options

The UPI codes could be generated/coined upstream in the workflow by the SEFs or DCOs (first touch) or downstream by the SDRs. The workflow and dissemination whitepaper will consider how and where trades are executed and the resulting workflows and provide recommendations for UPI generation.

6.1.1   First-touch UPI Generation

First touch UPI code generation is clearly the preferred approach for SEF executed trades. The SEFs will support standardized trades, whose UPIs could be generated prior to execution which would greatly reduce any latency for regulatory reporting. SEFs would also be perfectly placed to assign tickers to the most liquid, standardized and centrally cleared trades. SEFs could also communicate the UPI code early in the trade workflow to all participants and pass them to all downstream participants.

For non-electronically executed trades the reporting party would need to apply the algorithm or utilize an internal security master to determine the UPI code and provide the UPI code to the counterparty and SDRs.

It is expected that first touch UPI generation is a strategic goal that the industry will move towards over time.

6.1.2   SDR UPI Generation

Downstream UPI code generation is another option where the SDRs would generate the UPI code. The SDRs will be able to generate the UPI if given all of the discriminating information necessary to normalize the data for input into the hashing algorithm. SDRs could potentially receive multiple formats to identify and describe the products.

6.2   UPI Dissemination Options

The dissemination of the product identifier (UPI Code, asset class, taxonomy and product description) could be delivered via a number of methods. A dissemination (or 'look-up') service will enable users to gain the Product Identifier (including asset class, taxonomy and product description) by inputting the unintelligible UPI code.

Regulators have stated that the UPI and related data will need to be freely available and accessible by all market participants. There should be no barriers for the public to access the information. For this initial phase, the URL of the SDRs' website could be disseminated with the UPI to give market participants the access to look-up the UPI code and unveil the product definition, taxonomy and asset class.

Because of the federated approach towards generating UPIs, it is desirable to implement a look-up service to consolidate the information and serve as a hub for product identifiers and UPIs. This central service could consolidate data from multiple SDRs. The role of Digital Object Identifiers (DOIs) could serve as a potential candidate for a single directory to lookup UPIs given the distributed nature of generating them.

6.2.1   Dissemination key consideration

There is concern amongst market participants around the dissemination of UPIs for non-flow products. For example, taking physical delivery of a commodity is dependent on the delivery point since the price of transport makes a difference. Therefore the delivery point is a key trade attribute that is used in discriminating pricing and should therefore be part of the UPI. However if the delivery point is given for some trades you can very reasonably determine the party taking delivery and therefore the price they paid for the commodity.

A related concern is for an illiquid product, observers of real time data would know from the price on which side of the trade a dealer sits and as a result that there is now a dealer sitting on a certain position they will be looking to hedge potentially providing an opportunity for predatory behavior.

These concerns have been expressed to regulators in numerous forums and venues. Filtering UPIs from real-time reporting, so as to not destabilize the market or infringe upon the anonymity of market participants, will likely be necessary.

7   UPI Implementation

Considerable efforts are underway for the industry to develop the reporting as required by the rulemaking. Below are sequences of steps that will need to be taken together with implementation challenges that must be met for the industry to present a full product representation as defined previously.

7.1   Inter-operability testing

A successful inter-operability test is a milestone which is an important indication and verification of the specification document (consisting of not only the UPI Generation document but the asset class rulebook(s) used in the inter-operability testing).

7.2 Industry-wide testing

After successfully completion of the inter-operability testing the UPI Generation and Workflow will need to be tested within the industry by a number of market participants. This is a critical step as a seemingly small difference by one party (i.e. 1% as opposed to 100 bps) will lead to different UPIs for the same trade. End-to-end testing environments will need to be leveraged. Participants may include:

- Dealers and buy-side
- Swap execution facilities
- Middleware
- CCPs
- SDRs

Each of these participants will need to enhance and test their systems to either produce the UPI or ensure the smooth passage of the UPI into/out of their infrastructure.

7.3 Integration with existing systems, workflows and market participants

There is a strong dependency on the final rules when considering potential workflows for UPIs. As discussed in the UPI Generation Options section above, there are several possibilities that exist as to who will be 'coining' the UPI. In addition, there could be separate workflows depending on if the trade is a) electronically executed and cleared, b) electronically executed and bi-lateral or c) done on paper.

Market participants will need to determine where the UPI generation will occur, how to incorporate the UPI and other product descriptors into their security master. It is expected that there will most likely be many execution platforms that will need to enhanced and tested prior to going live with sending UPIs to the SDRs or generating it themselves so that they can check the values being produced by the SDRs.

7.4 Integration with industry reporting infrastructure and timing

Once there is more clarity around the final rules industry participants will need to determine how UPI fits into the workflow. Once inter-operational testing and industry-wide testing is completed successfully UPI generation will need to be incorporated into the industry's transparency program. There are considerable moving parts (UPI, LEI and USI); ensuring that the identifiers are incorporated into the infrastructure in a logical and coherent manner will be a priority.

8 Timeline

The overall implementation of the UPIs is interlocked and dependent upon the industry's broader market adoption for Dodd-Frank compliant OTC derivatives reporting. The timelines are currently being refined and will include a phased approach.

Each asset class has 3 phases:

1. Design/Build/Test Phase- Includes design work, interoperability testing and beta testing with market facilities including SDRs.

2. <u>Initial Implementation Phase</u>- UPIs generation begins for the most liquid and standardized products within the asset class. UPIs become available for integration with existing market infrastructure.
3. <u>Full integration with Infrastructure</u>- UPIs are integrated with the regulatory reporting infrastructure and reporting can commence for the first phase of products in each asset class.

Each phase is predicated on the previous phase. Thus, a slippage in the design/build/test phase for implementing Credit UPIs could have a knock-on effect impacting all subsequent phases.

It is expected that UPIs will become available for integration with the overall industry regulatory reporting framework by Q3/Q4 2012. The initial focus will be on the most liquid, standardized parts of the market.

9   Next Steps

9.1   Product Rulebooks

The priority is to develop the Credit and Rates asset class rulebooks which will define the price-forming attributes that make up the trade infosets representing the products. The product rulebook writing will then be broadened to include the rulebook working groups from the other asset classes (Foreign Exchange, Commodities and Equities).

Within each of the product rulebook working groups the starting point will be the most commonly traded/liquid and highly standardized products. These products will serve as the products that will be used in the interoperability and industry-wide testing noted above.

9.2   UPI Specification version 0.3

The next version of the Unique Product Identifier (UPI) Generation specification document version 0.3 is expected to be released in mid-December. The expectation is that this spec will be used to begin the inter-operability testing between market participants.

9.3   Algorithm selection criteria and considerations

a) Publicly available & non-proprietary
b) Ability to generate a unique product ID with a consistent character length
c) Format of output- Numeric vs. alpha-numeric or binary vs. octal
d) Robust / Widely used - Tried and tested
e) Highly collision-resistant
f) Interoperability- Separate versions run by different market participants need to generate the same UPI given same set of inputs
g) NIST or ISO certified

10 Dependencies

10.1 Final rule uncertainties

There is currently uncertainty around final rules and the impact to the approach, generation, workflow and dissemination of UPIs.

10.2 Broader Industry Regulatory Reporting Framework

As discussed in the timeline section the UPI implementation is intertwined with the overall industry OTC reporting framework.

11 Appendix 3: Criteria for evaluating UPI proposals

The UPI Steering Committee produced a set of criteria and considerations for evaluating UPI proposals.

1. Uniqueness- Approach will need to have an output (UPI code) that is unique and immutable.
2. Asset-class consistency– Each asset class should follow a consistent approach.
3. Granularity –UPIs must be granular enough to allow true, meaningful price comparison.
4. Time to market– The time to deliver a workable solution is an important consideration.
5. Scalability- Approach should be scalable to implement across products and asset classes.
6. Market Standards – UPI proposals should leverage open standards.
7. Universality– Proposals should be consistent globally, and designed to meet relevant known and reasonably anticipated regulatory requirements.
8. Rapid Implementation – UPI proposals should not deter or delay trading in and reporting of existing or new products.
9. Workflow- UPI proposals should be able to be incorporated into existing trade workflows.
10. Reversibility– A way of distributing the input attribution (product attributes) to the UPI, which could be a look-up service that is freely available to market participants.
11. Interoperable- Approach will need to generate the same UPI based on the same inputs for different implementations (or participants).

12 Appendix 4: Algorithm issues and proposals

1. Hashing the product information set is the way in which the UPI will be generated
2. SHA-256 is a widely available, NIST-certified edition that is part of the NIST application guidelines recommended algos for message digests. SHA-1 and MD5 are not recommended but are available widely, NIST certified etc.
3. SHA-256 can be truncated to any number of bits with a commensurate loss in collision-resistance – this is a permitted activity under NIST application guidelines but they warn that careful specification of how this is done is required to achieve safe interoperation.
4. UUID allows use of SHA-1 and MD5, doesn't seem a good option from that angle, but from every other angle it's a very attractive option (tooling, familiarity/precedent).
5. SHA-256 output is long.
6. Output is longer if encoded in Base32, to ensure case-insensitivity.
7. Truncate SHA-2 to half-length, and add an unofficial method 6) to UUID?
8. Write an RFC for a double-UUID, and then become the first use of it?
9. Another issue: endianism. Consider the output as a bit string. Do we chew it from the left or from the right? If we could view it as a Java or C# big integer, which endianism would be appropriate? Do Base64 and Base32 have word-endianism assumptions – or do they view the input as an arbitrary-length bit stream/string?