Distributed Ledger WG Call Notes

20 Feb 2018

# Attendees

* Mike Bennett
* Anthony Coates
* Mike Bennett
* Rob Nehmer
* Nick Stavros
* David Saul
* Pete Rivett
* Dan Webb

# Headline Points

Last week we considered the need for complex data structures to be posted to a distributed ledger, including terms with other qualifying terms (currency etc.), relationships among data posted at different times, and pedigree and provenance information.

This week we simplified this problem: given that XML is simply text and can be posted to any Block in any architecture, we don’t really need to worry about other datatypes (Boolean, integer, float etc.) or complex datatypes, since this can all be handled within any XML textual structure.

Further, XML includes RDF/XML which means that RDF graphs can be posted to a block, and can make reference to a formal ontology.

This means there are two immediate simple opportunities that don’t require any complex data structures support or any knowledge of or convergence between architectures:

* For any business process already defined in ISO 20022, the precise scope of what information needs to be posted to the block can simply be the corresponding ISO 20022 message definition (XML schema)
* Better, any information can be posted that is RDF data framed with reference to FIBO and other ontologies.

This also opens up the possibilities of being able to support reasoning over such data.

Consequently:

* Our PoC will assume posting of RDF/OWL structures to the block
  + Analysis of what needs to go alongside each datum needs to be analyzed as the next step in the PoC
  + We can make reference to ISO 20022 to help with these where applicable
* The DIDO Reference Architecture scope was also expanded to cover not just the Distributed Immutable Data Records architectures but also their use alongside RDF graphs and ontologies.

# Exploration

## Order Concepts

We started by exploring a link provided by Nick Stavros on kinds of ‘Order’ in the securities trading sense. Mike provided some background on how FIBO and ontologies derived from it make use of the formal REA Ontology for transactions, which covers these concepts more formally.

See link:

<https://www.investopedia.com/university/intro-to-order-types/duration.asp>

### Discussion

This is about Orders. Missing in Blockchain, where nearly all txns are immediate. This post describes the different of orders there can be. for example, how long is a given number valid for. Would apply to e.g. an exchange rate.

Mike Bennett This relates to last week's discussion on ancillary data.

Nick This can also apply to commodities like onions, potatoes, etc.

Nick I'm selling onions and they are only valid for 1 week, after they go off

### Link3: Market Orders

<https://www.investopedia.com/university/intro-to-order-types/market-orders.asp>

MB this all relates to the REA Transactions semantics we use in FIBO - any txn is an exchange of commitments, with different timings.

### REA Explanation

Each Commitment (each being one side of a transaction) has a beginning and an end, ie a.ife, which may involve several events or activities (it is incurred, it is discharged, potentially parts of the commitment are paid down at different times etc.). Simple retail transactions are the simplest as both sides, as commitments, are discharged right away. A typical securities transaction may have a T+3 or other settlement time (by market convention), both fo the delivery and the paymnt settlement. A simple swap e.g. a basis swap may have one leg as a short lived payment commitment (T+n) and the other side as an agreed set of cashflows over time as in a basis swap or bond return swap. For an IR Swap, like our example, both sides are of long duration (typically both are as long as the contract) and consist of a single commitment to pay different cashflows at predetermined times and under defined terms.

### Link 7 Duration Types

This is specifically the order duration (T+3 and that kind of thing; IR Swaps 3 month cycles etc.)

Similarly Margin - posted and updated daily. Not good beyond that day.

These txn kinds (actually txn side kinds): Duration is per side, whereas some of the terms listed on this page, like FOK, would describe the whole txn i.e. the two sides. So we are talking about the record of (the Side / the Txn) in various cases.

## Data Requirements for the Above

So we are talking about the record of (the Side / the Txn) in various cases.

There will be Ts and Cs on the overall txn and on each side of the txn e.g. a commitment to deliver onions is subject to ongoing variation in the quality of the onions.

Q: So is the Distributed Ledger environment a marketplace or something that reflects what happens in the marketplace?

A: The answer to 'is it this or ' is it that' is Yes. Can surely do either?

### What is Posted on the Block?

DS: If I buy a bond there would be multiple entries as interest posts every month. I can't wait until the end of its life (redemption) to post something. I need to post something every time cash moves. So saying there is an equivalence to a ledger post and a transaction is not suportabble.

NS: point was that the ancillary record that support the txn needs more information, e.g. how long a given exchange rate is good for.

DS: The rate is good for the particular transfer that takes place.

So this is the opposite - the ledger entry only represents the immediate action but the txn can run over months or years.

NS: the thing needs additional information - possibly on a separate oracle (MB: in some architectures).

DS: if I do an IR Swap and I need to cancel that I need all the information to do the cancellation. Includes the ccy rate.

MB see also bonds dirty price v clean price. The dirty price for a bond has certain accompanying information without which it has no meaning (Dirty Price = Clean Price + Accrued Interest; this changes daily)

Nick Lifecycle of commitment

Exploring further:

## Multi-chain

Nick Multi-chain is a good candidtae.

See link: <https://www.multichain.com/>

### Data streams

See note on Data-streams on this page – this names a few standard complex data problems, which makes us wonder if it supports the more general problem of linking data.

#### Data streams

Create multiple key-value, time series or identity databases on a blockchain. Ideal for data sharing, timestamping and encrypted archiving.

### Observations

This seems to answer our question in the negative: if they are identifying these specific kinds of complex data structure (key value pairs, time series etc.) that seems to suggest that the broader notion of being able to link any datum to any other datum, is not covered; or is hidden away somewhere.

To do this, we need link data…

DS without some means of linking data together, how are we to track where the data goes together, e.g. date, country, currency. What if someone creates 5 txns at the same time? Need some kind of universal product identifier that each of these figures is related to.

**Conclusion:** We need to understand the lifecycle of a given commitment, the data that needs to be present with each record in order to understand that record and the relationships from that record to other records they relate to (trades to reverse; interest on a given commitment) and / or to the original contract (terms sheet) to which that posted figure applies.

### Identifiers

NS: DIDO cites the need for identifiers

Mike Bennett We know that, but the question remains, where does it go? How do you post a thing to the Block that makes reference to this? It's al very well having a FIGI but how do I know, when I post a figure of 3.25%, that it is the accrued amount for that FIGI on that date?

NS believes none of them have taken this question head on.

DS agrees. Tech folks think they can represent an entire txn - they are not looking at long-running txns. But this problem has existed for a long time and is a problem that can be solved, e.g. via messaging system (MT series, ISO 20022 etc.) which represent incomplete parts of txns. Still need some way of linking these together.

e.g. ISO 20022 and ASN.1 (syntax for MT series)

# Proposal: Post XML as Text

Mike Bennett Can you simply take the whole of a given MT or ISO 20022 message and post it as a single piece of txt, to the Block?

Nick Perhaps it shows the need for "negotiation"

Mike Bennett Yes - e.g. in SWIFT you can tell when a message is associated with a given ongoing transaction.

MB trivially if you post all this as XML or ASN.1 text, you can parse it out at the other end and identify the relevant data elements. Just the same as with any message interchange medium.

Because XML Schema has its own defined datatypes, there is no longer a need to think about data posted to a block or ledger as decimal, boolean, float or anything else and no need to think about combinations of data of different types being posted together somehow. It’s just text to the block and the XML Schema layer and XML parsers deal with the rest.

DS: as long as everyone recognizes the parts of the message and the syntax, and adheres to it, as they do today, it can work.

### Extending to RDF

Mike Bennett Another possibility is to post RDF graphs as text.

RDF is also XML. So everything we have said about posting text for XML or for ASN.1 (SWIFT MT messages) can also be said of RDF graphs.

### Complexity, Rules and Reasoning

Nick The transactions of Bitcoin are instantaneously done. The idea that a transaction goes over time and needs to be negotiated is lacking.

At present, this complexity seems not to be known in the DLT development community (really?) so where there are things that require rules, these need reasoning engines.

David Saul I agree.

MB there are things you need parsing for, in order to consume the data from the block.

Further to that, there is also a role for reasoning engines and higher order logic, to implement rules. These are not the same thing. We were talking about being able to post and consume complex data structures by posting XML as single text records to the block. Now, separately, we are talking about how when this XML text happens to be RDF, there is a potential se for OWL based reasoning. These are not the same thing.

At present these are already done in existing systems, which can be as complex as you like. There are opportunities to improve on that using the Semantic Web. So RDF brings more opportunities to the table, but this has nothing to do with Blockchain. The use of RDF fits well with the application contexts in which we would be using DLT in the first place i.e. DLT + Triple store architectures can be a powerful combination.

# Conclusions

For our PoC we will not concern ourselves with whether or not there are other ways to read or write complex datatypes to a block. Instead we will focus on what you can do by posting RDF graphs (as RDF/XML i.e. text; presumably also TTL?).

# Next item: Name for DIDO (and redefined scope)

Previously we brainstormed some possible names or things that needed to be in names for what is currently called DIDO. DIDO = Distributed Immutable Data Objects.

We had words like Block, Web, Data web etc. that we felt needed to be in.

Meanwhile Distributed Immutable Data Architecture makes sense.

### Scope

NS wants to include use of reasoning engines within the scope of architectures covered by DIDO-RA.

Questions:

* Does DIDO now want to include use of reasoning engines?
* Exclusively, or as a further options?

NS: the thinking is that ontologies can be part of the distribution. Thereby supporting reasoning.

NS believes that this combination of reasoning support and distributed objects will be a powerful combination.

David Saul I really like including the ontology connection.

**Conclusion:** So DIDO intends to incorporate Bitcoin blockchains, distributed ledgers, tangles, hashgraphs and the rest. DIDO also talks to distributed computing, distributed objects and intelligence. Thereby embracing SmartContracts.

### DIDO Scope: So this is 3 layers right there:

1. Distributed data objects
2. Distributed smart objects that describe behaviors i.e. Smart Contracts
3. Distributed ontological representations of things i.e. ontologies.

DS agrees.

# Proof of Concept

To the IR Swaps Process diagram:

Description: MB has replaced the swimlane for ‘The Block’ with a swimlane for variable data. Static data (terms sheet, per Contract and Leg) are added as data objects in the swimlane for overall process. This is now therefore a conceptual model fo the information structures and process flows, without reference to what goes on a block a separate Oracle or anything else, as these would veary between architectures and application use cases.

Now we identify the kinds of data that vary. We just need to think about them like messages i.e. what piece of information gets posted along with each (similar to CD Dirty Price in a bond trade)

These are the ones that would now be posted as complex graphs containing the necessary complex data structures, qualifying terms, relations to other records, relations to the original contract, identifiers and so on.

All agree?

Yes this seems to be the way forward.

Dan Webb Agreed

Mike Bennett We have what we needed.

**Next Step:** We can start to logically design what information accompanies each of the figures we have identified.

[MB adds in edit: for processes that are defined in ISO 20022 these can be a direct copy of the ISO 20022 message schema for that process workflow.]

# AoB / Next Call

No call next week (27 Feb).

Next call Tuesday 6 March