UNITED NATIONS
CENTRE FOR TRADE FACILITATION AND ELECTRONIC BUSINESS (UN/CEFACT)
REGULATORY PROGRAMME DEVELOPMENT AREA (REG-PDA)
E-GOVERNMENT DOMAIN (EGOV-D)

Blockchain White Paper

White Paper on the technical applications of Blockchain to United Nations Centre for Trade Facilitation and Electronic Business (UN/CEFACT) deliverables

SOURCE: Blockchain Project Team
ACTION: Draft for public review

DATE: 30 April 2018
STATUS: Draft v0.1

I. Introduction

1. The international supply chain can be characterised as a set of three flows - of goods, funds and data. Goods flow from exporter to importer in return for funds that flow in the reverse direction. The flow of goods and funds is supported by a bidirectional flow of data such as invoices, shipping notices, bills of lading, certificates of origin and import/export declarations lodged with regulatory authorities. UN/CEFACT standards have played a fundamentally important role in this flow of data since the 1980s, facilitating trade and driving efficiencies in the supply chain.

2. These three flows are supplemented by a layer of trust. Trust, or lack of trust, underlies almost every action and data exchange in international trade, including trust in:

- The provenance and authenticity of goods;
- The stated value of goods for the purposes of insurance, duties, and payment; promises to pay;
- The protection of goods during shipping (i.e. integrity of packaging, vehicle and container conditions, etc.);
- The integrity of information that is used by regulatory authorities for the risk assessments which determine inspections and clearances;
- The traders and service providers involved in a trade transaction.

3. This layer of trust has seen relatively little support from technology and is still heavily supported by paper documents, manual signatures, insurance premiums and escrow and other trusted third-party services.

4. Blockchain, also known as Distributed Ledger Technology (DLT), is a technology that has the potential to deliver significant improvements and automation in this layer of trust.

5. As the focal point in the United Nations framework of the Economic and Social Council, UN/CEFACT needs to ask itself how this new technology impacts its work and whether there are any new technical specifications that it should develop in order to maximise this technology's value to UN/CEFACT's constituency. This paper seeks to answer these questions.

6. Although this paper is primarily focussed on blockchain, it is important to note that blockchain is not alone in its potential to have a disruptive impact on the supply chain. The rise of e-commerce platforms and cloud-hosted solutions are transforming the way organisations do business. The Internet of Things promises a vastly richer flow of granular data for tracking consignments, containers, through conveyances, ports, and warehouses. And other technologies, such as artificial intelligence and InterPlanetary File System (IPFS) as well as technologies under development such as the semantic web offer powerful new ways to understand and access data. Therefore, this paper will also position blockchain within the broader context of other new technologies that have an enormous potential to improve supply chain efficiency and integrity.

7. This analysis has resulted in five specific suggestions for UN/CEFACT work to support these new technologies. These suggestions build upon existing high quality work such as the UN/CEFACT Core Component Library (CCL) and process models.

8. The project team suggests:

- Investigating the development of a reference architecture so that all specifications as well as new technologies can be understood as constituent parts of a consistent whole;

- Reviewing UN/CEFACT process models in order to allow blockchain smart contracts (and other technologies using defined processes) to record key events and resulting changes in the status (state) of an entity such as the approval of an invoice or the release of consignments by a customs authority. This will require process models that are more granular and where the different statuses (states) of key entities are defined. In other words, process models that focus on the state life cycles of key resources in the supply chain such as consignments and containers as well as other key entities such as contracts and payments;

- Performing gap analysis to define what is needed in order to have an inter ledger (i.e. inter-blockchain) interoperability framework for supply chains that establishes cross-ledger trust in the face of the inevitability of a plethora of blockchain solutions;

- Performing gap analysis to define what is needed in order to provide supply chains with a standard way to discover and consume data regardless of which platform hosts information about a resource. It must take into account that cloud-based platforms will be the source of many truths (facts) about supply-chain entities such as parties, consignments and containers; and,

- Relying on a semantic framework that releases new value from existing UN/CEFACT work products such as the CCL. With the UN/CEFACT CCL, supply chains will have tools to process the faster and bigger stream of transactions and granular data that are being generated by platforms, IoT and blockchain. The working group further suggests that UN/CEFACT explore the use of ontologies based on the CCL.

9. As more platforms produce more data that must be understood by more parties, the value of UN/CEFACT semantics will only increase. There are exciting opportunities offered by blockchain and related technologies and looks forward to participating in work within UN/CEFACT to deliver new technical specifications that will release new value by supporting supply chain interoperability, efficiency and integrity.

II. Purpose and scope

10. UN/CEFACT standards such as the UN/EDIFACT directories have successfully supported trade facilitation and supply chain automation since the late 1980’s. As new technologies, such as XML, emerged in the early 2000’s, UN/CEFACT kept pace by releasing new specifications such as the CCL and the Extensible Marked-up Language Naming & Design Rules (XML NDR). However, the last few years have witnessed an unprecedented rate of technological change with the emergence of new technologies such as cloud platforms, the Internet of things, blockchain, advanced cryptography and artificial intelligence.

11. This poses two questions for UN/CEFACT:

- What opportunities do these technologies present for improving e-business, trade facilitation and the international supply chain?

- What is the impact on existing UN/CEFACT standards and what gaps could be usefully addressed by new UN/CEFACT specifications?

12. These questions are being addressed by UN/CEFACT white papers, each focusing on the impact of one key technology. This paper is focussed on blockchain to create a single architectural vision that positions blockchain within a future environment for supply chain automation that makes the best use of each technology.

13. At its heart, blockchain is a cryptographic protocol that allows separate parties to have shared trust in a transaction because the ledger cannot be easily falsified (i.e. once data is written it cannot be changed). This trustworthiness is created by a combination of factors including the cryptography used in a blockchain, its consensus/validation mechanism and its distributed nature.

14. If you are not familiar with blockchain technology yet, the first two pages of Annex I provide the basis.[1] The terminology used in blockchain (and also in this document) as well as related technologies (such as Internet of Things) are explained there.

15. Broadly speaking, blockchain technology can be used for four things (explored further in Annex 1), which are:

- A cryptocurrency platform, the best known of which is Bitcoin;
- A smart-contract platform, leveraging its immutable write-once nature;
- An electronic notary guaranteeing the content and, optionally, the time of issuance of electronically recorded data;
- A decentralised process coordinator, leveraging a combination of attributes, including its addressing techniques (public/private key), smart contracts, and immutability.

16. Since the core business of UN/CEFACT is to develop standards to support trade facilitation and supply chain automation, the focus will be on the smart contract, electronic notary and decentralised process coordination features of blockchain rather than cryptocurrencies. Similarly, although blockchain has wide application in sectors such as digital intellectual property rights, digital voting, digital record keeping, and so on, the focus will remain on its use within supply chains.

17. In this context, there are two types of blockchain implementations (explored further in Annex 1):

- Public blockchain ledgers, such as most cryptocurrency platforms, in which any party can host a complete copy of the ledger and participate in transactions and verifications. The two largest and best known public ledgers are Bitcoin (cryptocurrency) and Ethereum (focussed on smart contracts).
- Private or “permissioned” ledgers, in which a single party or consortium hosts the platform, sets the rules and explicitly grants permissions for other parties to act as nodes and/or perform transactions (performing transactions, which may, depending upon a private ledger’s rules, be open to the public).

18. A useful analogy here is that public ledgers are like the internet while permissioned ledgers are closer to corporate intranets. There are clear value and use cases for each and this paper will discuss both.

19. Given the high interest and potential value of blockchain technology, it is not surprising that there are already a large number of projects focussed on (or impacting in some way) the supply chain. These include shipping information platforms run by carriers, container logistics platforms run by port authorities, goods provenance (traceability) platforms, and many others. Most are permissioned ledger implementations. As with any promising new technology that has a rush of commercial implementations, some will fail and there is likely to be a growth phase followed by some consolidation. Nevertheless, technical limitations as well as commercial and political pressures will ensure that there will never be just one blockchain supporting the entire international

[1] See UN/CEFACT 24th Plenary document ECE/TRADE/C/CEFACT/2018/9.

supply chain. Even a single consignment is likely to touch multiple ledgers during its journey from exporter to importer. Therefore, just as UN/CEFACT has always focused on supporting interoperability between systems, the key technical focus for this paper is on supporting inter-ledger interoperability.

III. Related technologies

A. The rise of platforms

20. A platform-enabled website allows external Application Programming Interface (API) to offer additional functionalities. This means that developers can write applications that run on the platform (located on the cloud), or use services provided from the platform, or both. In pure business terms, a web platform is a business upon which an ecosystem of other businesses can be built. Shared platforms allow for innovation at the platform level, allowing work to be done once while benefiting many. This has allowed business models to emerge that eliminate intermediaries (create disintermediation) and create new efficiencies, disrupting the markets for intermediary services and lowering costs. A classic example of this disintermediation is the market for travel agency services.

21. However, at least as important, is the trend of established businesses such as carriers and couriers to provide APIs that allow their services to be seamlessly plugged into the systems of other businesses. The transition from desktop business applications such as small business accounting packages to cloud hosted platforms is also a notable trend.

22. The rise of e-commerce platforms has some profound impacts on electronic data interchange. Among these impacts are the following:

- The integration paradigm, instead of trying to exchange business-to-business messages between millions of individual businesses, integration is achieved simply by using APIs to connect together a few platform applications.
- Aggregation paradigm, the natural aggregator of businesses is shifting from centralized Electronic Data Interface (EDI) hubs that connect different parties, often on a semi-monopoly basis (because buyers dictate which hub must be used), to platforms where the sellers and buyers use their own platforms and then the platforms exchange data between one another. This means that sellers no longer have to deal with connecting to multiple hubs and it also allows them to take advantage of services on their platform that can analyze/use the data being exchanged.
- Discoverable data, platform APIs offer real time access to the resources (e.g. invoices, consignments, containers, etc.) that they host via simple web Uniform Resource Locators (URLs, i.e. web location). They can also emit events when a resource changes state (e.g. a container becomes “sealed” or “delivered” or an invoice becomes “paid”). What this means is that rather than exchanging large complex data structures as EDI messages, platforms can publish links to their resources and individuals can subscribe for the state changes which they find of interest.

23. There are some business risks with platforms:

- Platform operators may incorporate selected functionalities or services (provided by themselves) into the platform itself which prevents others from innovating in those areas on that platform and creates an incentive to drive innovations off-platform. This is less of an issue with platforms that are decentralized, or are operated in an open way by regulators rather than commercial interests.

- As platform adoption approaches market saturation (meaning most of the market uses the platform), the dysfunctions associated with monopolies (or, when there are just a few firms, oligopolies) come into play with fewer incentives to innovate, improve services and lower costs. In addition, network effects (the value provided to the community of additional users) diminish and zero-sum games become the main economic drivers. This situation naturally drives platforms to exploit asymmetric information advantages (such as surveillance-based business models) and replace their emphasis on innovation and collaboration with an emphasis on cost reduction, even at the expense of customers (a lack of credible alternatives for customers meaning that the platform has less need to be concerned with their satisfaction).

24. In general, the consequence of these kinds of behaviour are new spin-off platforms that attract customers away from more established platforms. To prevent this, platforms sometimes implement lock-in strategies that increase the cost and difficulty of transferring to alternate platforms.

B. The Internet of Things

25. The Internet of Things (IoT) describes a network of sensors or smart devices that are connected to the Internet and generate a stream of data. Many blockchain trade applications use data generated from the IoT for processing by smart contracts. For example, sensors in containers and in ships, ports and railway infrastructure might be used to track container movements and then this information could trigger actions based on previously agreed smart contracts.

26. IoT data feeds are generally owned by infrastructure operators, value-added service providers, or specific platforms, and their availability is already being used as a source of differentiation and competitive advantage between platforms. This data is often made available through platform APIs or using message-based approaches. The impact on international trade and blockchain applications will be a significant increase in the volume and timeliness of supply chain data.

IV. Risks and opportunities

A. A plethora of ledgers

27. An increasing number of individual corporations, government agencies, and industry consortia are recognizing the value of blockchain technology (beyond cryptocurrencies) and are building platforms that intersect in some way with the international supply chain. Some are focussed on transport logistics, others on trade financing, others on goods provenance (traceability). Some are international and some are local or regional. As with any new technology there is likely to be a surge of initiatives followed by some market consolidation. Nevertheless, the eventual landscape will be characterised by a plethora of different ledgers, with different characteristics including trust. Furthermore, data about a single consignment is likely to be provided to or obtained from several different blockchain ledgers.

28. Possible examples of related data being recorded on different blockchain ledgers include:

- The commercial invoice may be recorded on financial industry ledgers focussed on trade financing and insurance;

- Consignment and shipping data may be recorded on ledgers run by freight forwarders and couriers;
- Container logistics information and bills of lading may be recorded on a ledger run by carriers and/or port authorities;
- Permits and declarations may be recorded on ledgers run by national regulators.

29. Blockchain technology does not solve the interoperability problem that UN/CEFACT standards have always supported. Also, different blockchains are far from equal in terms of the level of trust that participants should place in them. A permissioned ledger run by a single corporate entity with very or relatively few nodes will have much less resistance against hacker attacks than a public ledger such as Bitcoin, a permissioned ledger with thousands of nodes, or a large multi-party permissioned inter-ledger operated by multiple entities.

30. At the same time, the implementation of blockchains, together with other technologies such as the IoT and cloud platforms, is creating more and more electronic data that needs to be shared across supply-chain participants.

31. The opportunities for UN/CEFACT are:

1) To ensure that its semantic and business process modelling standards are fit for purpose in blockchain environments, and
2) To identify what needs to be done in order to ensure the most efficient and effective use of blockchain technology by supply chains and all their participants, including government authorities.

B. A profusion of platforms

32. There is likely to be some overlap between the scope of a platform and the scope of a blockchain ledger. In some cases there could be a 1:1 relationship where a given platform is also the host of a single permissioned ledger. Some platforms won't use blockchain at all, others will interact with multiple blockchain ledgers and still others may share a blockchain ledger. A potential use case could be a national platform hosting approved certificates of origin and participates in a multi-country blockchain ledger created through multilateral arrangements and in which multiple national platforms handling certificates of origin each host a node.

33. In any case, while blockchain ledgers are intended to provide a certain level of trust, platforms support the flow of data. As discussed in the previous section on the rise of platforms, they can provide data, which in some cases is authoritative, about a resource such as a consignment or a container. In a few rare cases, a single platform might hold all the authoritative data about a single consignment and its related data (commercial and logistical). In that case, the problem of discovering all related information about a consignment would be simply a case of querying the single platform. However, this is most likely to be the exception rather than the rule. Therefore, the interoperability challenge includes a discovery problem - given an identifier of an entity (e.g. a container or consignment number), how to locate the detailed information about it?

34. There is an opportunity for UN/CEFACT to identify what needs to be done in order to ensure that all supply-chain participants can locate the data that they need (and that they are entitled to access) about a given transaction, even if the data is scattered across different platforms and blockchains. Such a resource discovery protocol, allowing supply chain participants to discover the detailed data about a resource given its identifier, would allow a profusion of platforms to work like a virtual single global platform.

C. A torrent of data

35. While traditional structured document exchanges (of invoices, bills of lading, declarations, etc.) will remain a critical part of the data landscape, the rise of platforms and IoT will bring an additional stream of more granular data such as the events in the lifecycle of a consignment or container or conveyance. This granular data might be discovered by following a link in a blockchain, or by following the identifier of a resource in a document. Whatever the discovery mechanism, there remains a challenge to actually making sense of the transaction or data stream if different platforms, different blockchain networks and different IoT applications present the same information (semantic concept) differently.

36. There is an opportunity for UN/CEFACT to leverage its existing semantic standards such as the CCL.

V. Putting it all in context

37. Technologies such as blockchain, IoT and platforms can each, independently, contribute to increased supply chain efficiency. At the same time, when working together within a standards-based framework, the sum can be much greater than the parts. In this context, it could be very useful to develop a conceptual model of the international supply chain that shows the role of each technology within the broader map of stakeholders, services, and standards. Such a model would work equally well for the domestic supply chain, which is just a simpler subset of the international supply chain.

A. A context model for trade technologies

38. The diagram in Figure 1 shows a draft conceptual model of the international supply chain with relevant technologies. Importers and exporters often facilitate the flow of goods, funds and data, as well as the relevant trust by using a variety of service providers and third parties. Overlaying blockchain and other emerging technologies on the model can show the relationship with the new UN/CEFACT specifications suggested later in this paper. Some other observations related to this diagram are:

- All parties in this example use one or more platforms to conduct their business. This may be a single organisation-level internal platform, e.g. a corporate Enterprise Resource Planning (ERP) system, but increasingly will be cloud-hosted web platforms for most participants.
- Platforms may use IoT data sources and APIs to improve the information flow.
- Platforms may use private blockchain ledgers to improve trust by recording immutable and auditable transactions.
- An inter-ledger framework, eventually prepared by UN/CEFACT, could provide trust between platforms.
- A resource discovery framework, eventually prepared by UN/CEFACT, could provide a means to locate the authoritative data source for a resource based on its identifier.
- UN/CEFACT trade data work such as the CCL provides semantic anchors to facilitate data exchange.

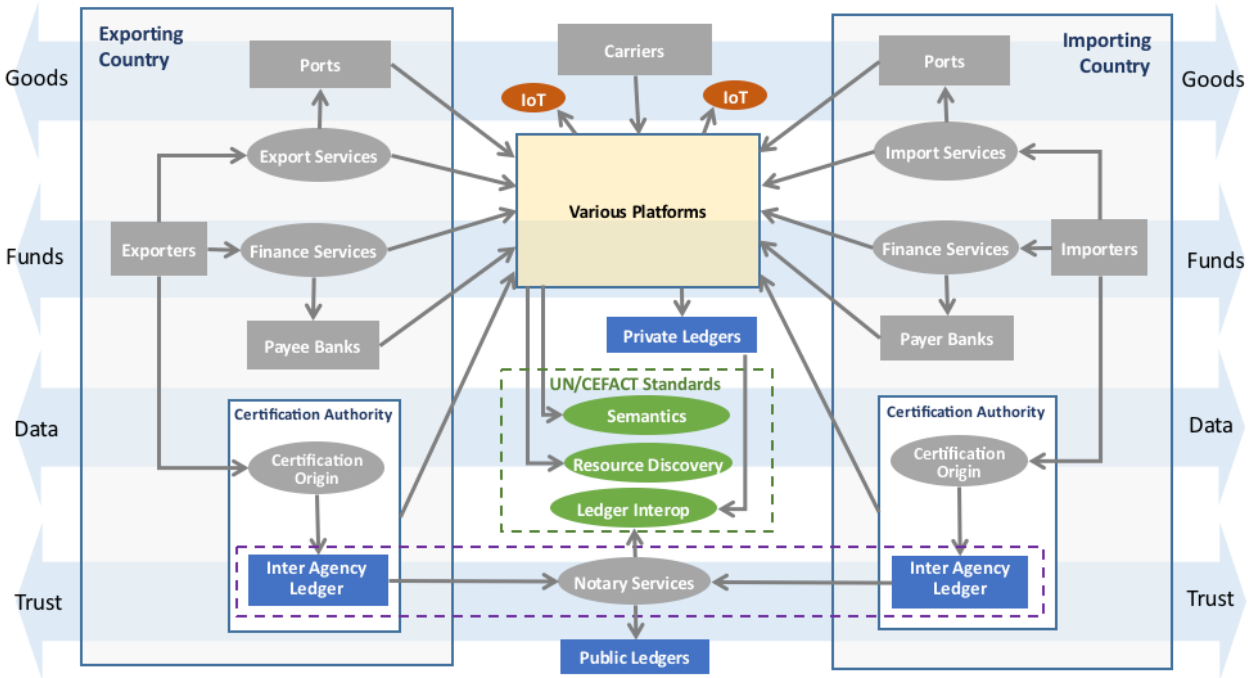

Figure 1 - Draft Context Model for ICT Trade Technologies

39. Arrows between boxes/ovals in the diagram represent dependency relationships so should be read as "uses" or "depends on". They do not represent flows of information which are between various platforms and ledgers.

40. Multiple platforms exist to address different needs in the trade and transport sectors, and will continue to evolve through innovation in IoT, AI and other emerging technologies.

41. National regulators play a special role in the network as they provide a unique point of convergence for data in each jurisdiction.

- Data is often being integrated data from multiple sources ranging from traditional document-based data sources to more detailed digital data entries and can come in higher volumes and can be delivered in real-time. The same can be true for key transport hubs such as sea, air and dry ports.
- Authorities are unlikely to surrender control of their information and processes by conducting regulatory business on a shared platform outside their jurisdiction. They will, undoubtedly, maintain independent systems, but find new ways to verify and appropriately share data with other countries.

42. All of the above underlines the growing complexity and multiplication of systems and data that traders and authorities will need to deal with in the near and long-term future.

43. Standards-based semantic models could facilitate this widening network of data exchange around trade transactions and support traders as they look for flexible integration across a diversity of platforms (including diverse blockchain-based applications).

44. A complete example of a possible blockchain trade scenario is presented in Annex 2.

VI. Suggested way forward for UN/CEFACT

45. Based on the opportunities identified and positioned above, there are some clear gaps that UN/CEFACT is uniquely positioned to fill. The project team suggests that UN/CEFACT work with national delegations and its experts to establish working groups to progress the following new technical specifications.

A. A UN/CEFACT Architecture Reference Model

46. Just as UN/CEFACT semantics standards are mapped to UN/EDIFACT and XML through technical specifications like the XML NDR, so it must be shown how UN/CEFACT's semantics can be mapped to newer technologies such as blockchain, big data, and web platform APIs. Also, as data flows become more granular, it will be increasingly important to model the detailed semantics of processes as well as data.

47. All of these drivers will lead to a number of new technical specifications and related semantic work. In order to have these specifications understood as parts of a consistent bigger picture, it is suggested that a reference architecture specification be developed that shows how all technical specifications work together. This work could use the context model presented in this document as a starting point.

B. Process modelling in support of smart contracts

48. Significant economic commitments between agents may be associated with specific events in the lifecycle of a resource. Possible examples include:

- An invoice transition from "received" to "approved" may trigger the release of low cost trade financing for small suppliers.
- A consignment transition from "landed" to "cleared" represents the release of goods by a regulatory authority.
- A shipment resource that transitions from being in the possession of agent X to agent Y when containers are sealed and loaded under a bill of lading.

49. If these events can be notarized as smart contracts in a trusted blockchain ledger, then there is a unique opportunity to improve and automate this trust in the supply chain. But only if there is a clear shared understanding of the meaning of each state transition (including the triggering conditions).

50. Therefore, a review is suggested of the existing UN/CEFACT Modelling Methodologies and standards (Business Requirement Specifications and Requirement Specification Mappings) to identify what modifications would be needed to support blockchain and smart-contract based applications.

C. Inter-Ledger interoperability framework

51. As more and more applications anchor their transactions into various private and public blockchain ledgers, there will be an increasing need for a means to discover and integrate transactions across blockchains.

52. As discussed earlier, each transaction on the chain contains only the hash of the actual data and a minimal amount of metadata about the document or transition state. With clear semantics in the metadata, parties can discover data of interest in other ledgers by observing linked-data anchors and traversing them to obtain appropriate access.

53. Also, as discussed earlier, each node on a blockchain has a complete copy of the ledger. Specific ledgers (and the nodes that verify transactions) will typically exist for a specific geographic or industry segment. But if a specific international consignment touches a dozen different ledgers, it is impractical for a party that wishes to verify the transactions to host a dozen different nodes. A common inter-ledger notary protocol would allow authorized parties to verify transactions irrespective of which ledger they are created on.

54. Therefore, the project team suggests the establishment of a technical working group to review existing work by standards organizations in order to identify if there is a need to collaborate with them on a possible framework for inter ledger interoperability specifications that would define:

- Standards for on-chain metadata;
- Standards for inter-ledger notarization.

55. This specification will most likely build upon (and not duplicate) existing specifications such as Hyperledger chain code, Ethereum solidity code, and multi-hash.

D. Resource discovery framework

56. Resources, such as invoices, consignments, certificates of origin, containers, etc., will be increasingly hosted on web platforms. This means that the source of truth about supply chain entities will be online and discoverable, vastly increasing supply chain transparency. At the same time, even for a single international consignment, these truths (information resources) will exist on many different platforms. It is impractical to expect every authorized party to be a registered member or customer of every platform that holds some relevant data. However, it could be possible, given the identifier of a resource, to develop a consistent means to discover where it is hosted and be granted access to appropriate data. If this were done, then the disparate web of platforms could work as one.

57. As a result, it is suggested that UN/CEFACT develop a specification that bridges independent platforms to discover resource data independent of where it is stored. Basic requirements for the specification would include the ability to:

- Resolve the identity of parties, platforms and other agents participating in trade-related activities, using identity providers from all jurisdictions and sectors.
- Access current and authoritative information about the public keys of participants, to enable secure direct interaction and communications.
- Support a diversity of entity types (e.g. businesses, jurisdictions, platforms, containers) including high volume entity types (e.g. consignments).

58. This specification should build upon (and not duplicate) existing, relevant technical elements from existing specifications.

E. Trade data semantics framework

59. After all the technological wizardry, organizations in the supply chain still must be able to make sense of the data that is discovered / exchanged by various platforms, ledgers, or even network connected sensors. However, as described in the chapter on the rise of platforms, the landscape is changing from EDI hub-centric models to peer-to-peer exchange where platforms are the natural aggregators. The traditional document centric

transaction is complemented / enriched by a fast moving stream of events about all the resources in the supply chain.

60. In this context, there is an opportunity to increase the value of UN/CEFACT semantic standards through a technology where:

- UN/CEFACT explores the use of ontologies based on the CCL and if this approach may be better adapted to the use of blockchain technologies.
- Communities of interest (e.g. fast moving consumer goods in a country) can overlay the core UN/CEFACT semantics with an industry / geography specific framework that effectively says "this is how we use the UN/CEFACT standards in our context".
- Platform operators can release semantic frameworks that map their interfaces to UN/CEFACT standards.

61. As a result of the above, runtime tools (called inferencing technology) for a particular business in an industry sector that uses a specific platform could overlay all three semantic frameworks to consistently use and create UN/CEFACT standard data from any platform that meets their industry / geography specific needs.

F. Blockchain application data needs

62. There is an immediate need to work with blockchain application developers to identify data that requires definition and is not covered by current UN/CEFACT standards (specifically, the CCL) and to develop related Business Requirement Specifications and core components in order to cover that gap. In particular, there is a requirement, from within a business document or transaction, to reference data (one or many) located in a particular blockchain (out of many).

63. This review should also look at any new needs created by off the chain data used in blockchain applications. Most data will not be kept on a blockchain, rather it will be referenced (pointed to) together with a hash for data verification and perhaps a time stamp. There may also be a requirement to describe various cryptographic primitives for the purpose of referencing them from business documents. For example, hashing algorithms, key distribution, cryptographic signatures and encryption schemes.

64. At the same time, this blockchain capacity will result in an exponential growth in systems that reference data which has been generated by diverse sources - resulting in either high costs for harmonization or high error rates as data is used that is based on different definitions. In conclusion, there is an urgent need to look at not just blockchain data but, perhaps even more importantly, the data used by blockchain-based applications especially in areas like trade that are horizontal and use data from almost all sectors of economic activity. As a result, it is suggested that UN/CEFACT consult and engage with technical standard bodies and review existing technical standards to see what might be relevant for developing trade facilitation applications using blockchain.

Annex 1 – Blockchain; How it works

I. Blockchain - How it works

1. At its heart, blockchain is a cryptographic protocol that allows separate parties to have shared trust in a transaction because the ledger cannot be easily falsified (i.e. once data is written it cannot be changed). This security is due to a combination of factors including the cryptography used in a blockchain, its consensus/validation mechanism and its distributed nature.

2. This annex does not aim to provide an in-depth review of blockchain technology - there are plenty of web resources to help readers achieve that goal. Rather, it will cover the core concepts which are needed to understand the potential application of blockchain in international supply chains.

3. First, some nomenclature:

- **Node**: System that hosts a full copy of the blockchain ledger.
- **On-chain transaction**: Automated procedure that creates or updates the state of an address in the blockchain database by appending new data to the ledger. Examples include digital asset exchange, or execution of an automated business process.
- **Validation**: Work performed by all nodes in parallel, that verifies transactions using a consensus algorithm. Different networks may use different consensus algorithms. When mutual validation results in a consensus, then the nodes all commit (record) the transaction onto their blockchain.
- **Block**: Data that is appended to the ledger by consensus. Once a block is written to the chain, it cannot be changed or deleted (without replacing all subsequent blocks).
- **Hash**: Fixed size, unique cryptographic fingerprint of data. A hash is a one-way function; this means that given the data, one can easily verify that the hash is the correct one for that data. However, it is not possible to reverse-engineer the hash, so you cannot use it to re-create the data. This is a key feature because it allows users to confirm that no changes have been made. For example, even an additional space or empty line in a text would change its hash.

4. An important characteristic of blockchain systems is the way consensus allows users to trust that transactions have been executed and trust information about those transactions (for example, their date and content). As a result, blockchain systems can be used as an independent umpire in processes that might otherwise expose participants to the risk of one party not living up to its contractual obligations (counterparty risk) and third party guarantors are reluctant to intervene and assume part of that risk. In the case of public blockchains, the umpire is the society of all nodes that choose to participate in the consensus. In the case of private blockchains, the umpire is the consortium of nodes trusted to (given permission to) create consensus on the network.

A. It's a distributed ledger

5. Ledgers are a kind of database, kept digitally or with paper records, where transactions are recorded once and not subsequently updated (also known as a journal database). Each record can be read many times but written only once. The term ledger comes from

accounting where entries, once written into a ledger (accounting journal), cannot be changed.

6. A blockchain is described as distributed because there are multiple copies which are kept on different nodes. The multiple copies are updated in a coordinated way that ensures they remain consistent, using a consensus algorithm (of which there are many). Specifically, the consensus algorithm decides (by mutual agreement between the nodes) which block is added to the chain next. In essence, a blockchain database is a sequence of data blocks that have been added in a specific order, by consensus of the network operators, to each of multiple copies of the ledger and where each block contains a fingerprint (hash) that can be used to verify the content of all the previous blocks.

B. It writes transactions

7. Each block of data written to the ledger contains at least one or many records of transactions. A familiar example of a transaction would be “debit one coin from account A, and credit one coin to account B”, although many other kinds of transactions are possible. Some blockchains support a limited sub-set of transactions (operations or algorithms), such as this simple double-entry bookkeeping operation. Some blockchains support a much wider set of transactions covering any solvable algorithm (i.e. a Turing-complete computer programming language[2]). These types of transactions are variously called smart-contracts, chain-code, transaction families, or other, equivalent terms. In summary, all blockchains support a variety of data operations on their chains, but not all blockchains support Turing-complete transaction languages.

C. To a cryptographically signed block

8. Blockchains implement two kinds of cryptographic technology: hash functions and public/private key cryptography. Hash functions are used to construct the fundamental proof that links each block to the rest of the chain before it. Hashes, in a different context, can also be used to provide proof of validity for data that is referenced by blocks. Public/private key cryptography is used for identifying transactors and controlling access to data. An analogy is e-mail where the public key is your email address (which others can use for sending messages to you) and the private key is your password which gives access to the private material which is your messages. So, on a blockchain, a public key can be used, for example, to implement a transaction that sends a document or a payment to a party, but only the party with the private key can access those documents or payments after they are sent.

D. That independent nodes must verify

9. There are various consensus algorithms used by different blockchain systems. For example, Bitcoin uses proof of work algorithms which allow miners to recover the cost of computationally expensive work in exchange for transaction fees. Permissioned ledgers use a consortium of collectively trusted (but not necessarily individually trusted) nodes to agree on the output of a consensus process, which are generally cheaper and faster than Bitcoin’s proof of work. All consensus processes require a mechanism to settle

[2] Turing complete programming language is capable of solving any mathematical problem computationally (if you know how to program it). In general, this means it must be able to implement a conditional repetition or conditional jump (while, for, if and goto) and include a way to read and write to some storage mechanism (variables).

disputes, or uncertainty, about which block should be written next. Most of these mechanisms are based upon using the block which is agreed upon by more than 50% of the nodes.

10. The nature of the consensus mechanism determines some key characteristics of a blockchain system. For example, Bitcoin has deliberately made mining (the creation of blocks) expensive. This protects the blockchain by making the cost of capturing more than 50% of the nodes (the number needed to approve a block, and thus to manipulate the blockchain) prohibitively expensive. To compensate for this cost, miners are rewarded both an amount of Bitcoin for each block they create and fees for each transaction written to the blockchain[3]. Each block has a size limit and transaction costs are determined on a free market basis, so the more transactions are requested, the more the price increases for each transaction. This is necessary for the Bitcoin economic operating model, which seeks to obtain an honest consensus in an unregulated market of potentially anonymous and economically rational operators (i.e. operators who might, being anonymous, and having no costs for doing so, steal assets). As an additional incentive, if a node/miner does not accept the block voted on by over 50% of the other nodes, it is, effectively, kicked off the blockchain, thus losing the possibility of earning future Bitcoins and transaction fees. As a consequence, Bitcoin has extremely low bandwidth (due to the cost of generating blocks) with transactions taking more than 10 minutes to be confirmed. In addition, its very large number of nodes and users (generating large amounts of data), together with its block size limits, makes storing data on the Bitcoin blockchain expensive as well as being inefficient (given the duplication of information across all nodes, it is generally inefficient to store significant amounts of data on any public blockchain). Bitcoin still supports many billions of US dollars worth of Bitcoin and other high-value transactions, but its speed and volume limitations make this blockchain unsuitable for many enterprise applications.

11. Permissioned ledgers strike a different balance between bandwidth, capacity and trust. For example, because they have more control over who participates, permissioned ledgers can use other consensus mechanisms, even if some of them are somewhat less robust than the proof of work used by Bitcoin. For examples, there are consensus mechanisms based on the amount a node has invested in a network (called proof of stake), or where a consensus by a subset of nodes is verified by a larger group. In addition, there is a great deal of research going on to identify and test a range of other consensus mechanisms. Using these alternative consensus mechanisms, some permissioned ledgers can support hundreds or even thousands of transactions per second (rather than an average of one new block per 10 minutes, as with Bitcoin) and petabyte scale databases.

E. The block is written to the ledger after it is verified

12. When consensus is reached (which includes agreeing that a block contains legitimate data, and that it is the block that should be written next), each node adds the agreed block to their local copy of the ledger. In this way, all nodes maintain an identical copy of the ledger each time a block is written. This is guaranteed (proven) by the next block to be written, because it will contain a hash of the block before it.

---

[3] Bitcoin is designed so that, over time, mining rewards are reduced with the objective of eventually having all mining rewards come from transaction fees.

F. The new block is linked to previous blocks - creating immutability

13. Recall that a hash is a one-way function that produces a unique fingerprint of some data. Also note that a hash function produces a fixed-size fingerprint regardless of the amount of data being hashed. For example, there is no way to know from looking at the hash if the data was a single small document or a database holding many billions of records.

14. Each block in a blockchain contains some transaction data, plus the hash of the previous block (which is always the same size, no matter how much data it represents). Given a consensus that this new block forms part of the chain, it is possible to verify the previous block from its hash. And from the previous block, the block before it, and so on all the way to the first (or genesis) block in the chain. The hash of the previous block is said to be anchored in the subsequent block.

15. Tampering with the contents of any block in the chain will change the hash of that block, which will change the hash of the block after it, and so on for every subsequent block in the chain. If this occurs then the tampering is easily detectable by any node, and the consensus algorithms will prevent new blocks from being written to a chain because the hashes don't match.

16. This characteristic is the origin of the word "chain" in "blockchain" because each block is anchored to the previous block and proves the existence of all the data it references going back to the first "block" of data in the "chain".

II. Blockchain - Types

A. Public Ledgers

17. Public ledgers can be read by anyone. They are also permissionless in the sense that anyone can participate and utilise consensus mechanisms without depending on a regulator to enforce acceptable behavior. Bitcoin, Ethereum and more than 10 other cryptocurrencies with market capitalization over USD 1B operate this way, allowing any transaction that is logically valid even between anonymous parties.

18. One of the fears about blockchain technology is that, if a malevolent actor were to control a majority of the nodes, then they could decide to reach a consensus in contradiction of the interests of other stakeholders. This threat is described as a Sybil attack in the cryptographic literature. A successful Sybil attack on a public blockchain cryptocurrency could result in a catastrophic redistribution of assets or double spending. Public ledgers are designed to operate according to rules that do not require governance or regulatory mechanisms to intervene in order to prevent antisocial transactions, because those mechanisms might themselves be exploited for antisocial outcomes, for example, if they were to be hacked by a third party or abused by the trusted regulators. These systems operate with absolute trust in their algorithms and are designed to avoid any need to trust any counterparties. This is why (public) blockchains are sometimes referred to as being trust-less.

19. Public ledgers typically compromise other aspects of performance in order to achieve strong resistance to Sybil attacks. They also rely on the transparency of the public ledger, and also on the transparency of the open source software involved.

B. Permissioned/Private ledgers

20. Like conventional (operational/analytic) databases, the contents of a private blockchain ledger may be a guarded secret that is only available to selected users (and node operators) through a role-based access control mechanism. Unlike a traditional database, a private blockchain ledger is immutable (cannot be updated) and transactions are verified by a consensus mechanism that is established by the network operators.

21. Private ledger technology is typically applied in enterprise use-cases where immutable transactions are required, that can be verified by a closed community of nodes. These nodes may be independent of parties to the transactions on the blockchain and may be subject to oversight and governance that is not possible (or considered desirable) in a permission-less blockchain system.

22. Permissioned ledgers operate with a different threat model to the public ledgers. The operators of permissioned ledgers are not anonymous, they are subject to some kind of governance controls and are collectively trusted by the users. Antisocial behaviour of a node or participant could result in that party being evicted from the network, and their transactions blocked or even rolled-back from the blockchain by consensus of the remaining operators. The expectation of users of a permissioned ledger is that the operators will intervene in antisocial behaviour but not commit antisocial behaviour themselves.

23. On permissioned ledgers, the level of security, and so the confidence users can have in the immutability of the data, varies depending upon the rules established for that permissioned ledger (including its consensus mechanism). Permissioned ledgers can also create a false sense of security because only trusted participants are allowed to maintain nodes and participate in verification. At the same time, even trusted participants can become untrustworthy upon being hacked; permissioned ledgers with single points of failure are vulnerable should anything happen to that single point, and poorly tested smart contracts can create bad consequences for participants – even if no harm was originally intended, and especially if the blockchain network does not have adequate controls in place.

C. Interledger: implementing transactions across blockchains

24. Today, many different blockchains exist and, in the future, there will be even more. Already, a supply chain transaction, from beginning to end, could involve writing or reading data from multiple blockchains. In addition, it is easy to foresee an increasing need for the exchange of information and the implementation of transactions across blockchains (i.e. interledger).

25. As mentioned earlier, blockchains can reference data outside of that blockchain. This includes data in other blockchains as well as non-blockchain systems. There are two broad categories of external data references that can occur in a blockchain system: linked data and blockchain-spanning transactions.

26. Linked data uses hashes and may also use digital identifiers and public key cryptography (as long as it is used consistently across the blockchain and whichever system the linked data is stored on). This implies that the more standardized the use of public key cryptography, the easier and less expensive it will be to link data – and the same can be said for the semantics defining the data.

27. Extrinsic blockchain references (also known as anchors) can be used to prove the existence or unchanged nature of the data pointed to. This is different from a hyperlink or Uniform Resource Locator (URL) on the Internet where the information at an address

may change depending on the time it is accessed. For example, if you click on a link on a television news website, which changes on a regular basis as it is updated, what you find tomorrow may be different than what you find today. With a blockchain anchor data link, the information in the blockchain is a guarantee (proof of existence) that the data being pointed to has not been changed.

28. As well as linking data between two blockchain systems (cross-chain references) and pointing to data that may be used by a smart contract (for example a test certificate), linked data can also be used to incorporate off-chain big data into a space-constrained blockchain system. Supplementary data can either be in public/open distributed data systems such as InterPlanetary File System (IPFS – an open, content-addressable memory that uses standard internet protocols), or it may reference data in private databases that are selectively available to permissioned ledger users. With private off-chain or cross-chain references, it is possible for network operators to know that some data exists, but to have their access limited by additional controls. This can be very interesting from a privacy standpoint as it is possible to access data in order to know that, for example, someone is over 21, without giving their age, or that they live in London, without giving their address.

29. Inter-ledger (blockchain-spanning) transactions use cross-chain references and components (e.g. smart-contracts) on both blockchains that interact in a coordinated way. This is an emerging field, however there are mechanisms that already exist and are in use. These are primarily focussed on exchanging value (digital assets) between ledgers, for example Ripple interledger and the Lightning Network.

Annex 2 - Making it real with a hypothetical working example

1. As an aid to understanding the context model and the positioning of new technologies and UN/CEFACT standards, below is a hypothetical end to end story of a consignment of wine from an Australian exporter to a Chinese importer. Entity names are fictional and not intended to represent any real organisations:

- Wine producer Perfect Pinot Ltd. is a registered business on the Australian national business register at abr.gov.au with Australian Business Number (ABN) 111222 and is located in New South Wales (NSW).
- Perfect Pinot Ltd. produced and bottled 100,000 bottles of its 2016 vintage. Each bottle has a unique serial number identified by a signed Quick Reference code (QR code) on each bottle using a system from Smart Tags Inc.
- Smart Tags Inc. writes the batch of QR codes to an Ethereum blockchain anchored goods provenance system that they run on behalf of wine producers.
- Wine exporter Fine Reds (ABN 222333) negotiates an export deal with Chinese wine importer Hunan Wines which is registered on the China National Enterprise Credit Information system with an Administration for Industry and Commerce number (AIC number) 444555.
- Hunan Wines places an order for 1,000 bottles of Perfect Pinot Ltd. with Fine Reds. Using a resource discovery framework, Fine Reds' platform looks up the Hunan Wines platform and e-invoicing internet address and sends the commercial invoice directly to the target platform in accordance with UN/CEFACT semantic standards.
- Because Fine Reds and Hunan Wines are on different platforms and because the commercial invoice is one of the foundations of trust, the invoice is also notarized/registered on a public blockchain using an inter ledger notary framework. Hunan Wines indicates their acceptance of the invoice (also notarized).
- Fine Reds grants permission to access the notarized invoice to their bank which provides lower cost trade finance when transactions are notarized.
- The conditions of carriage require that the wine remains under 25 degrees and above 5 degrees centigrade during the shipment, so Fine Reds engages the services of Cool Shippers for freight forwarding. Cool Shippers have instrumented containers with IoT temperature sensors and Global Positioning System (GPS) tracking.
- Cool Shippers provides Fine Reds with the container ID and Fine Reds uses a resource discovery framework to find the container web internet address and subscribe to the container data feed.
- Cool Shippers provides the signed and notarized invoice and the smart tags blockchain reference to the NSW chamber of commerce which verifies the data and issues an automated and signed certificate of origin which is registered on a blockchain.
- Cool Shippers creates a consignment reference using their logistics platform and provides the consignment ID to Australian customs via an authenticated session established by the single window API. Australian customs uses the resource

discovery framework to locate the consignment data and subscribes to data feeds about the consignment.

- The consignment data includes a reference to the notarized invoice, the container ID, the carrier ID, and the certificate of origin ID. So Australian customs can discover full data about each entity, verify integrity, and create an approved export declaration. The export declaration (with links to supporting data) is recorded as a smart contract on an inter-organization ledger.
- The importer clicks a button to review and approve all export & shipping documentation and submit the import declaration.
- China Hunan province customs authority observes a new import declaration. China customs verifies the trade documents and confirms that Fine Reds and Hunan Wines have a sufficient history of high integrity trading. The consignment is pre-cleared by Hunan customs.
- On arrival in Dadukou Port, the container data feed indicates that the cargo has landed and un-packed. The temperature history is notarized and confirms that temperature has remained below 25 and above 5 degrees centigrade for the duration of the journey.
- When the pallet of wine is scanned into Hunan Wines warehouse, the consignment resource IoT device emits the “received” event. This, together with other notarized transactions is sufficient information for Fine Wines’ bank to release an invoice finance payment at very reasonable terms.
- Hunan Wines releases the Perfect Pinot Ltd. wine to a number of retail outlets in Hunan province. A customer buys a bottle and scans the QR code on the bottle. The smart tags platform confirms the authenticity of the wine and records the scanning event against the specific bottle serial number.

2. This example is, of course, fictional but nevertheless entirely feasible. The key difference between this future state vision and current state reality is that each authorized party has direct access to the single source of truth about each entity (party, invoice, consignment, container, etc.) and that all key data is notarized in a blockchain ledger aiming at high levels of trust and so is independently verifiable.