FINAL
DRAFT

TECHNICAL
REPORT

ISO/DTR 6277

ISO/TC **307**

Secretariat: **SA**

Voting begins on:
2023-10-26

Voting terminates on:
2023-12-21

Blockchain and distributed ledger technologies — Data flow models for blockchain and DLT use cases

Reference number
ISO/DTR 6277:2023(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents. ISO shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 307, *Blockchain and distributed ledger technologies.*

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

This document consolidates a set of system-level models from ISO 23257:2022 and ISO/TR 3242:2022 to give a data-flow-centric description framework for blockchain and distributed ledger technology (DLT) use cases. The framework enables a data flow analysis approach for blockchain and DLT use cases which has been defined in ISO 23257:2022, successfully applied across all use cases in ISO/TR 3242:2022 and extended in this document to display more detailed information on data flows.

The robust descriptive capabilities provided by this framework can help to improve blockchain and DLT application design and enhance interoperability. It can be beneficial for:

— clear understanding of data types and data flows in distributed ledger systems that allows for better-designed, fit-for-purpose systems;

— better governance and risk management;

— a sound basis for interoperability modelling for the use cases that require data exchange in hybrid or orchestrated systems environment.

Understanding data flows can be a necessary foundation for DLT users to ensure data privacy and data confidentiality in DLT use cases, or a decision-making basis when implementing technology selection or scheme assessment. From this perspective, data flow analysis is especially essential to scenarios which frequently involve data flows among stakeholders or devices. To illustrate the features of data flows in DLT use cases with above characteristics, this document provides three uses cases which apply the description framework to unfold data flows among devices, data flows along with business process, as well as data flows between physical and virtual spaces. These use cases can also provide an insight into the role of data flow analysis in balancing business value maximization and risk controls.

This document is organized as follows:

— Clause 5 presents an overview of DLT data flows, including data flow categories, data categories, roles/subroles and considerations related to data flow.

— Clause 6 and Clause 7 provide analysis of typical intra-system and inter-system data flows for DLT systems.

— Clause 8 provides three DLT use cases based on a descriptive and visualization template focusing on data flows.

Blockchain and distributed ledger technologies — Data flow models for blockchain and DLT use cases

1 Scope

This document uses a set of models that describe the flows of different types of data between distributed ledger technologies (DLT) and related systems, as well as between different DLT nodes.

It provides a descriptive analysis of data flows in the development of use cases, as well as the basis for understanding the characteristics of DLT data flows, to support DLT application design and system analysis.

The models referenced are in accordance with ISO 23257:2022 and the use case analysis approach provided in ISO/TR 3242:2022.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 22739, *Blockchain and distributed ledger technologies — Vocabulary*

ISO 23257, *Blockchain and distributed ledger technologies — Reference architecture*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 22739, ISO 23257 and the following apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1
cloud computing
paradigm for enabling network access to a scalable and elastic pool of shareable physical or virtual resources with self-service provisioning and administration on-demand

Note 1 to entry: Examples of resources include servers, operating systems, networks, software, applications, and storage equipment.

[SOURCE: ISO/IEC 22123-1:2023, 3.1.1, modified — Note 2 to entry deleted.]

3.2
data category
class of data items that are closely related from a formal or semantic point of view

[SOURCE: ISO 30042:2019, 3.8, modified — Example and Notes to entry deleted.]

3.3
data flow
movement of data through the active parts of a data processing system in the course of the performance of specific work

[SOURCE: ISO/IEC TS 20748-3:2020, 3.5]

3.4
decentralized identifier
DID
identifier that is issued and managed in a decentralized system and designed to be unique within a context

Note 1 to entry: Decentralized identifiers are used in systems that do not rely on central registration authorities.

Note 2 to entry: Decentralized identifiers are often cryptographically verifiable.

[SOURCE: ISO 22739:—[1)], 3.18, modified — Note 2 to entry added.]

3.5
derived data
data created as a result of processing that involves steps other than or in addition to direct retrieval and validation of information from data functions

[SOURCE: ISO/IEC 20926:2009, 3.17]

3.6
DLT account
distributed ledger technology account
representation of an entity participating in a transaction in a DLT system

Note 1 to entry: A smart contract, digital asset, or one or more private keys, for example, can be associated with a DLT account.

[SOURCE: ISO 22739:—, 3.25, modified — Note 1 to entry added.]

3.7
edge
boundary between pertinent digital and physical entities delineated by networked sensors and actuators

Note 1 to entry: Pertinent digital entities means that the digital entities which need to be considered can vary depending on the system under consideration and the context in which those entities are used.

[SOURCE: ISO/IEC TR 23188:2020. 3.1.2]

3.8
end user identifiable information
EUII
derived data associated with a user that is captured or generated from the use of the service by that user

Note 1 to entry: Data that is linked to the user but is not DLT user data.

Note 2 to entry: End user identifiable information includes connectivity data, usage data.

[SOURCE: ISO/IEC 19944-1:2020, 3.1.2, modified — Notes 1 and 2 to entry added.]

1) Under preparation. Stage at the time of publication: ISO/FDIS 22739:2023.

3.9
role
set of activities that serves a common purpose

[SOURCE: ISO/IEC 22123-1:2023, 3.1.10]

3.10
peer-to-peer
P2P
relating to, using, or being a network of equal peers that share information and resources with each other directly without relying on a central entity

[SOURCE: ISO 22739:—, 3.70]

3.11
smart contract
computer program stored in a distributed ledger technology system wherein the outcome of any execution of the program is recorded on the distributed ledger

Note 1 to entry: A smart contract can represent terms in a contract in law and create a legally enforceable obligation under the legislation of an applicable jurisdiction.

[SOURCE: ISO 22739:—, 3.88]

3.12
sub-role
subset of the activities of a given role

[SOURCE: ISO/IEC 22123-1:2023, 3.3.11]

3.13
transaction record
record documenting a transaction of any type

Note 1 to entry: Transaction records can be included in, or referred to, in a ledger record.

Note 2 to entry: Transaction records can include the result of a transaction.

[SOURCE: ISO 22739:—, 3.95]

4 Abbreviated terms

API	Application programming interface
DID	Decentralized identifier
DLT	Distributed ledger technology
ICT	Information and communications technology
IoT	Internet of things
n.e.c.	Not elsewhere classified
PII	Personal identifiable information
PoC	Proof of concept
SDG	Sustainable development goal
SME	Small and medium-sized enterprise

VC Verifiable credential

5 Overview of data flow for DLT

5.1 General

The impetus for introducing DLT-specific data flow models is to support technical and business process analysis. The models mentioned in this document are in accordance with ISO 23257:2022 and applied across all use cases in ISO/TR 3242:2022 combined with the behavioural UML models. The focus on the approach in this document is exploring diverse ways of applying it to data flow analysis on DLT use cases. The approach taken in this document derives from architectural approaches in cloud computing and service-oriented design. If a service model is a collection of components that represents a business service, a data flow model that is described across component-based view of a system can help bring clarity to both technical and business objectives in system design.

It can be seen in ISO/TR 3242:2022 and elsewhere that applications and systems deploy blockchain and DLT to provide robust and purposeful system transparency in highly distributed, multi-party systems, on cloud-based execution environments. The data flow models reviewed here can be used in tandem with service-modelling approaches to gain greater insight into system functionality and business process performance.

5.2 Categories of data flows

ISO 23257:2022 and ISO/TR 3242:2022 identify five fundamental data flows relative to DLT systems:

— Data flow N: data flowing within and between the nodes of the DLT system.

— Data flow A: data flowing between separate DLT systems when they interoperate.

— Data flow B: data flowing between a DLT system and non-DLT systems connected to it.

— Data flow C: data flowing between administration applications and a DLT system.

— Data flow D: data flowing between user applications and a DLT system.

NOTE ISO 23257:2022 defines Data flow Z as data flowing among the nodes of the DLT system. ISO/TR 3242:2022 includes data flow within the nodes of the DLT system and defines Data flow Z as data flowing within and between the nodes of the DLT system. This document adopts the definition in ISO/TR 3242:2022 and uses the code N instead of Z to avoid confusion.

5.3 Data categories

This document identifies data categories in the DLT ecosystems, to help understanding DLT data flows and support transparency about DLT data. A data taxonomy is also useful for the conversation about data between different roles/sub-roles. This document provides the following four sets of data categories:

— data categories from data storage perspective;

— data categories from data sources;

— identifier data categories;

— other data categories.

5.3.1 Data categories from data storage perspective

5.3.1.1 General

In order to balance the advantages and performance of DLT systems, off-ledger data storage has increasingly become a common auxiliary way of data storage in DLT system, especially in the DLT applications involving large amount of DLT user data.

The main concern for on-ledger data and off-ledger data is the difference of scope of data processing and use. In most cases, processing and use of off-ledger data have no much difference with data in other types of IT systems, however, it is usually impossible to delete on-ledger data, which makes it more crucial to provide specific ways of ensuring transparency and privacy protection.

5.3.1.2 On-ledger data

On-ledger data are data stored inside a DLT system, which includes small amounts of data from DLT users. Due to the size-limit of DLT ledger, on-ledger data can also include hashes related to off-ledger data from DLT users.

5.3.1.3 Off-ledger data

For large amounts of data from DLT users, or personal identifiable information (PII) which might need to be deleted or updated by the PII principal, it is common practice to store the data on a cloud on the user's infrastructure or a public storage provided by a third party or network. Off-ledger data also includes data from the DLT provider data and data derived when a DLT application is used.

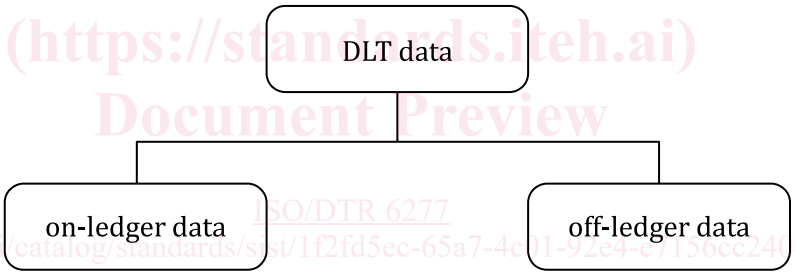

Figure 1 — Data categories from storage perspective

5.3.2 Data categories from data sources perspective

5.3.2.1 General

ISO 23257:2022 defines six DLT roles, including DLT developers, DLT administrators, DLT users, DLT providers, DLT governors and DLT auditors. Among these DLT roles, DLT administrators, DLT users and DLT providers are most closely related to DLT data. In order to support different stakeholders to carry out data-related activities, or to formulate data-related policies, this document identifies seven data categories which are associated to different DLT roles (see Figure 2).

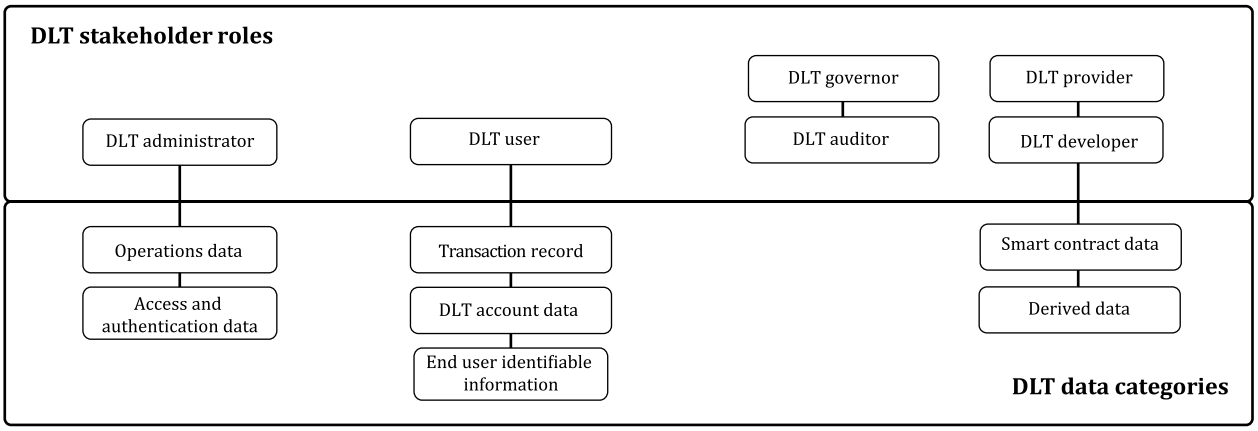

Figure 2 — DLT stakeholder roles and related DLT data categories

5.3.2.1.1 Transaction record

Transaction record can be financial transaction data, or generalized transaction data such as genome data, voter record, transport data, product production data. Transaction records might be on-ledger data or off-ledger data with related hashes stored on ledger. Transaction record is generated when a DLT user uses a DLT system to record a transaction.

5.3.2.1.2 DLT account data

A smart contract or digital asset, for example, can be associated with a DLT account. Corresponding DLT account data are data representing an entity whose data are recorded on a DLT system. DLT account data are generated or updated when a DLT user creates a DLT account or uses the account.

5.3.2.1.3 End user identifiable information

End user identifiable information is linkage to the user but is not directly created by DLT users. End user identifiable information includes connectivity data, usage data.

5.3.2.1.4 Operations data

Operations data are data which is used for supporting the operation of DLT system, which includes service logs, configuration data.

5.3.2.1.5 Access and authentication data

Access and authentication data are data used within DLT system to manage access DLT capabilities, DLT data or smart contract, which includes passwords, cryptographic keys, security certificates. Access and authentication data are controlled by DLT administrator and are critical to its administrative activities.

5.3.2.1.6 Smart contract data

Smart contract data includes not only executable codes of program but also execution results. Smart contract data can be generated when a DLT developer creates or maintains the smart contract, or a DLT operator runs the smart contract.

5.3.2.1.7 Derived data

Derived data include data describing the connections of the DLT system, data describing the usage of the DLT services, etc. Derived data also include end user identifiable information.

5.3.3 Identifier data categories

ISO/TR 6039 specifies the identifiers data including:

— Subject identifiers: natural persons and legal entities used by administrations of countries, for specific (international) government functions and in some industries. Subjects are entities with rights and obligations.

— Object identifiers: object identifiers used for government purposes and in several industries. Objects are entities without right and obligations.

5.3.4 Other data categories

5.3.4.1 Overview

There are many other types of data and data types of selected examples that are important in financial use cases are presented in 5.3.4.2 to 5.3.4.4.

5.3.4.2 Market price data

Market prices are available for many objects. These include, but are not limited to:

— stock prices of listed companies at exchanges;

— currency rates of Forex markets;

— commodity prices for commodity markets;

— derivatives rates;

— interest rates (Central Banks reference rates and market rates);

— prices of retailers and web-retailers for goods and services.

5.3.4.3 Accountancy data categories

— Stock data: Accountants use specific terminology included in the IFRS standard which uses a balance sheet that can be presented in an annual report. These stock data include: the value of the goods, debtor position, creditors position presented on the balance sheet. The stock data covers the aggregation of data at a certain point in time.

— Flow data: Accountants use terminology included in the IFRS standard for flow data such as a) profit and loss account and b) cashflow statement. Flow data includes the flow of values, as they change over time, during a predefined time period.

5.3.4.4 Message data in networks

— Messages: Data in messages can be structured or unstructured. The (industry) networks used for Business to Government (B2G) or for Business to Business (B2B) purposes include mostly instructions on the structure of the data messages used in the network involved.

5.4 Roles from the perspective of data flow

5.4.1 DLT stakeholder roles and stakeholder data

Data flows are triggered by the data-related operations of stakeholders, between system components that belong to or are associated with them. Stakeholders achieve their aims by means of role-based interactions with the DLT system.

ISO 23257:2022, 9.2 describes a set of roles/sub-roles which address the main activities associated with DLT systems and gives overall descriptions for activities of these roles/sub-roles. When discussing data flow and data taxonomy, the detailed data-related activities of these roles/sub-roles are necessary information.

5.4.2 Roles/sub-roles and their activities related to data flow

5.4.2.1 Data-related activities of DLT users

A DLT user often uses a DLT system on their devices, by using a DLT application or an application that interacts with a DLT API. Examples of its activities include:

— providing data to DLT systems;

— using data obtained from DLT systems on their devices;

— installing applications on devices.

5.4.2.2 Data-related activities of DLT administrators

A DLT administrator performs administrative activities which might be data-oriented or produce data. Examples of its activities include:

— developing plans to ensure blockchain data backup and recovery, as well as possible data replication and failover;

— ensuring compliance of data storage and processing;

— discovery, classification and protection of data;

— managing access and authentication data;

— managing application configuration data;

— managing cross-ledger data exchange;

— managing long-term preservation of data.

5.4.2.3 Data-related activities of DLT providers

As defined in ISO 23257:2022, DLT provider is the business stakeholder owning and operating one or more DLT nodes. DLT providers can have data interaction with DLT users and are also responsible for the operation of data.

DLT providers include three sub-roles:

— DLT system operators;

— DLT node operators;

— DLT application operators.

5.4.2.3.1 Data-related activities of DLT system operators

Examples of DLT system operators' activities include:

— providing audit data activities according to audit's requests;

— maintaining operation data.