INTERNATIONAL STANDARD

ISO 20614

First edition
2017-11

Information and documentation — Data exchange protocol for interoperability and preservation

Information et documentation — Protocole d'échange de données pour l'interopérabilité et la préservation

Reference number
ISO 20614:2017(E)

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

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 documents 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).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

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

For an explanation on 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 the following URL: www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 46, *Information and documentation*, Subcommittee SC 4, *Technical interoperability*.

Introduction

The Data Exchange Protocol for Interoperability and Preservation (DEPIP) aims at facilitating interoperability between a digital archive and information systems of its partners: producers who have created the documents themselves (Originating Agency), intermediaries who are acting on behalf of producers and are not responsible for the intellectual content per se (Transferring Agency), consumers (Consumer) and control authorities (Control Authority). This document provides a framework for data exchange between systems. It is based on the OAIS Reference model. It is generic and may be adapted to all types of information, whether printed or in a born-digital format.

DEPIP is intended for:

— commercial software vendors, in order to complete and/or improve their applications;

— archives, in order to standardize ingest of data destined for preservation; to provide access to archived data; and to facilitate the exchange of data between archives;

— application programmers, in order to enable interoperable data transactions between Archives and the information systems they are developing;

— third parties responsible for transferring documents to Archives;

— data storage service suppliers.

The parties to data exchange may rely on this document to:

— define their archiving/preservation processes or harmonize their existing processes with the best practices;

— organize the management of the archiving/preservation processes; and

— control the creation and management of metadata, whatever descriptive models are used.

Note that DEPIP may be implemented either in its entirety or only partially. However, it is impossible to foresee what the implementations will be like in different domains such as libraries, archives or museums. Domain-specific or generic International Standardized Profiles specifying different levels of interoperability may be developed in the future to support the implementers of this document. The minimum implementation, or level 0, is: Transfer and Delivery, including at least mandatory metadata. Note that in level 0, existing metadata is not redefined and that the initial request and the final reply are required.

Note that DEPIP is a conceptual standard to be considered as a data dictionary. The model defined in DEPIP is independent of implementation issues. In different implementations, DEPIP can be supplemented by relevant technical protocols (like HTTP) allowing implementers to handle technical exchanges between systems.

Information and documentation — Data exchange protocol for interoperability and preservation

1 Scope

DEPIP specifies a standardized framework for the various data (including both data and related metadata) exchange transactions between an archive and its producers and consumers. Interchanges between archives (including archives integrated in organizations, public archives, storage service suppliers) are also considered. This document defines five transactions (Transfer, Deliver, Dispose, Modify and Restitute), which the partners can use to exchange Data Objects. It also specifies the syntax and semantics of the messages that are exchanged during these transactions.

Internal organization of the information systems of the partners is excluded. Information received in conformance with the data exchange model is intended to be handled by various software components. These applications, however, are not the object of this document. The impacts of major risks (for instance, disappearance or incapacitation of the producer of the data) are also excluded.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

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

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

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

3.1
Access
transmission of information by an *Archive* (3.2) to a *Consumer* (3.5), with the authorization, if required, of the *Originating Agency* (3.11) and of the relevant *Control Authority* (3.6)

3.2
Archive
organization that intends to preserve information for *Access* (3.1) and use by a designated community in accordance with the current legal, regulatory or contractual conditions

Note 1 to entry: The DEPIP Archive is not fully equivalent to the OAIS Archive as the latter does not check the compliance with the contractual conditions.

Note 2 to entry: The DEPIP Archive is not necessarily the same as an archive in a traditional sense in which legal custody is permanently transferred. It may temporarily offer preservation services for content, in accordance with a legal agreement that has a set time frame. At the end of that time frame, or in accordance with some other stipulation of the legal agreement that would permit it, the *Data Objects* (3.7) that had been transferred to the archive could be returned to the *Originating Agency* (3.11) or a third party appointed by it.

3.3
Binary Data Object
digital item which contains information, for instance, an electronic file, i.e. a named and ordered sequence of bytes that the file system of an operating system may handle as a unit

3.4
Business Identifier
identifier used to identify the *Archive* (3.2) and its partners, messages, *Submission Agreements* (3.18), etc.

3.5
Consumer
individual or organization wishing to consult information kept by the *Archive* (3.2) in accordance with the current legal, regulatory or contractual conditions

3.6
Control Authority
internal or external individual or organization that, if applicable, may authorize the delivery or the disposal of information held by an *Archive* (3.2)

Note 1 to entry: Control Authority is partially equivalent to the OAIS management (like management, it may be external to the archive), but it has narrower role than management in that Control Authority is only interested in legal or regulatory compliance.

3.7
Data Object
either a digital (sequence of bits) or a physical object which is to be preserved, and technical metadata (representation information, integrity information and identification information)

3.8
Disposal notification
notification by an *Archive* (3.2) to an *Originating Agency* (3.11) of information disposal

3.9
Disposition rule
information required to manage the data lifecycle to indicate a retention period, beyond which *Data Objects* (3.7) shall be disposed or preserved

3.10
Modification notification
notification by an *Archive* (3.2) to an *Originating Agency* (3.11) of the modifications made to the submitted and/or archived *Data Objects* (3.7)

Note 1 to entry: These modifications may be necessary during the ingest or later to ensure proper storage of information (e.g. changing the file format, or adding, correcting or updating representation information or preservation description information).

3.11
Originating Agency
individual or organization that made or received the information within the context of its activities

Note 1 to entry: It is often the producer in the OAIS model.

Note 2 to entry: The Originating Agency may act as a *Transferring Agency* (3.20) or may use a Transferring Agency as an intermediary for sending data to the *Archive* (3.2).

3.12
Physical Data Object
physical item which contains information, for instance, a file, a box, a CD-ROM, etc.

3.13
Preservation
combination of policies, strategies and actions developed by the *Archive* (3.2) to ensure that digital information of continuing value remains accessible and usable

3.14
Preservation profile
adjustment of the descriptive model based on the types of exchanged *Data Objects* (3.7)

3.15
Restitution
transfer of information by an *Archive* (3.2) to an *Originating Agency* (3.11) in order to shift back to the Originating Agency the responsibility for data retention

Note 1 to entry: This transaction is partially equivalent to the OAIS functional entities archival storage and access.

3.16
Rights metadata
metadata concerned with the limitations and restrictions regarding the access to and use of data commonly including elements such as copyright status, use restrictions and information about licensing agreements

3.17
Service Level
quality of the services provided by the *Archive* (3.2) to its partners and planned by the *Submission Agreement* (3.17), including secure preservation, guarantee of the integrity of the stored data, availability rate, etc.

3.18
Submission Agreement
agreement or regulation used as a framework for the relationships between the *Archive* (3.2) and its partners

Note 1 to entry: In order to facilitate DEPIP implementation, an agreement should describe at least the following:

— international standardized profile(s) supported (if any);

— the types of transactions (transfer, deliver, modify, dispose and restitute) supported, specifying, when necessary, whether a preliminary authorization of the Control Authority is required;

— the list of individuals or organizations involved, their roles and responsibilities in these transactions;

— the selected code lists and models to be used during these transactions;

— Preservation profiles (i.e. rules for creating descriptive metadata according to the type of documents or applications being preserved) including: Service Levels, access and Disposition rules and information about how the terms in the original agreement may have evolved.

Note 2 to entry: Details concerning data transactions may be incorporated in the Submission Agreement. It is also possible to create a separate agreement (complementing the Submission Agreement), which provides the required technical information. In DEPIP, it is assumed that everything is included in a Submission Agreement.

3.19
Transfer
submission of submission information packages or SIPs by a *Transferring Agency* (3.20) to an *Archive* (3.2) in order to hand over the responsibility for preservation

Note 1 to entry: This transaction is equivalent to the OAIS ingest functional entity.

3.20
Transferring Agency
individual or organization that submits submission information packages (SIPs) to an *Archive* (3.2) but does not own the Data Objects submitted

Note 1 to entry: In the OAIS model, the term “producers” has a narrower meaning (people, or more likely, the organizations, which provide the objects to be archived). From an OAIS point of view, a Transferring Agency is a kind of producer, who is acting on behalf of third parties and is not responsible for the content per se, but may have responsibility for creating the Submission Information Packages (SIPs).

Note 2 to entry: A Transferring Agency could sometimes also be the *Originating Agency* (3.11). In that scenario, it is the owner of the Data Objects it is transferring, and it also has the responsibility for creating the submission information packages (SIPs).

4 Context

4.1 Roles played by individuals or organizations involved in transactions

The following are the principal roles in the DEPIP context (defined in Clause 3): Archive, Transferring Agency, Originating Agency, Control Authority and Consumer.

NOTE Three roles identified in DEPIP are not directly part of the basic OAIS model (Control Authority, Transferring Agency and Originating Agency) but may be seen as aspects of OAIS Producer and Management.

4.2 Types of transactions

The transactions described in detail in Clause 5 are:

— Transfer: Transfer of information by a Transferring Agency to an Archive in order to hand over the responsibility for preservation. The Transfer may be preceded by a Transfer Request for agreement;

— Deliver: Delivery of information by an Archive to a Consumer, with the authorization, if required, of the Originating Agency and/or of a Control Authority;

— Modify: Notification by an Archive to an Originating Agency to inform it that the transferred information has been modified. These modifications may be necessary in order to ensure proper storage of information (e.g. changing the file format). Note that the entire activity of modification as would be undertaken by an Archive would comprise many more steps than merely a notification being sent, but the scope of "Modify" as focused here is merely the notification;

— Dispose: Notification by an Archive to an Originating Agency to inform it that the requested information has been disposed of. The disposal may be preceded, if applicable, by a Disposal request to the Control Authority and by an Authorization request to the Originating Agency. Note that the entire activity of disposition as would be undertaken by an Archive would comprise many more steps than merely a notification being sent, but the scope of "Dispose" as focused here is merely the notification;

— Restitute: Transfer of information from an Archive to the Originating Agency for the purpose of returning the responsibility for preservation to the Originating Agency. Restitution should not be confused with data recovery, that is to say, the full Restitution of the relevant information in a reusable way and as contracted.

4.3 Exchanged objects

4.3.1 General

The objects exchanged during DEPIP transactions are Data Objects (including technical metadata) accompanied by descriptive and administrative metadata. Since DEPIP is based on OAIS, technical metadata relates to structural metadata (representation information) and is considered as part of the Data Object, while descriptive and administrative metadata is left open and is only accompanying information. The types of these objects (indicated in brackets) and the cardinality of their components are presented in class diagrams.

4.3.2 Exchanged Object packages (DataObjectPackageType)

4.3.2.1 General

A package of Data Objects (DataObjectPackageType) is composed of a set of Data Objects accompanied by descriptive and administrative metadata.

4.3.2.2 Data Objects (BinaryDataObjectType and PhysicalDataObjectType)

A Data Object is either a digital (sequence of bits) or a physical object which is to be preserved, containing technical metadata, i.e. representation information (for instance, format), integrity information (for instance, hash code) and identification information (for instance, identifier).

This document makes a distinction between:

— Binary Data Objects (BinaryDataObjectType): for instance, an electronic file, i.e. a named and ordered sequence of bytes that the file system of an operating system may handle as a unit;

— Physical Data Objects (PhysicalDataObjectType): for instance, a file, a box, a CD-ROM, etc.

A Binary Data Object may be characterized by its format (e.g. "PDF 1.4"), its encoding (e.g. "UTF-8" for a text file) and its size (in bytes). The digital content may be physically included (encapsulated) within a message, or it may be bound by a reference (e.g. a file name).

The decision to either encapsulate digital content in the information package or to leave it outside of the package is implementation-specific. The decision may be based on criteria such as the size of the Data Object.

A Data Object on a physical medium (e.g. paper document or analogue recording) is characterized by specific technical metadata. The main related metadata elements are its size (number of folders, boxes, linear meters, etc.) and its medium or container using its technical identifier and/or its storage location.

4.3.3 Administrative metadata of exchanged Data Objects (AdministrativeMetadataType)

Administrative metadata applies to all the Data Objects in a SIP package and includes the following information:

— Submission Agreement;

— Preservation profile;

— Service Level;

— Rights metadata;

— Disposition rule.

NOTE Rights metadata is bound to the Data Object. It is controlled by the Control Authority, but it is not held by it.

4.3.4 Descriptive metadata of the exchanged Data Objects (DescriptiveMetadataType)

Descriptive metadata includes information related to the Data Objects (data origin, description, date, keywords, etc.). Descriptive metadata should apply to all Data Objects in a SIP. Descriptive metadata may be based on different metadata formats, depending on the domain (e.g. MARC 21 in libraries, EAD in archives, ONIX for book trade, and so on).

5 Modelling

5.1 General

The model used for description of transactions is Unified Modeling Language (UML). Three types of diagrams are used.

— The use case diagrams provide an overview of the system by representing the individuals or organizations involved and their actions on the system.

— The sequence diagrams include each use case and provide a temporal representation of the progress of each transaction. These diagrams represent the scenarios involving the Archive and its partners.

— The class diagrams are used to define the set of elements and their properties used in different transactions.

5.2 Use case diagrams

5.2.1 General

Five transactions may occur between the Archive and its partners: Transfer, Deliver, Modify, Dispose and Restitute. These transactions are shown in Figure 1.

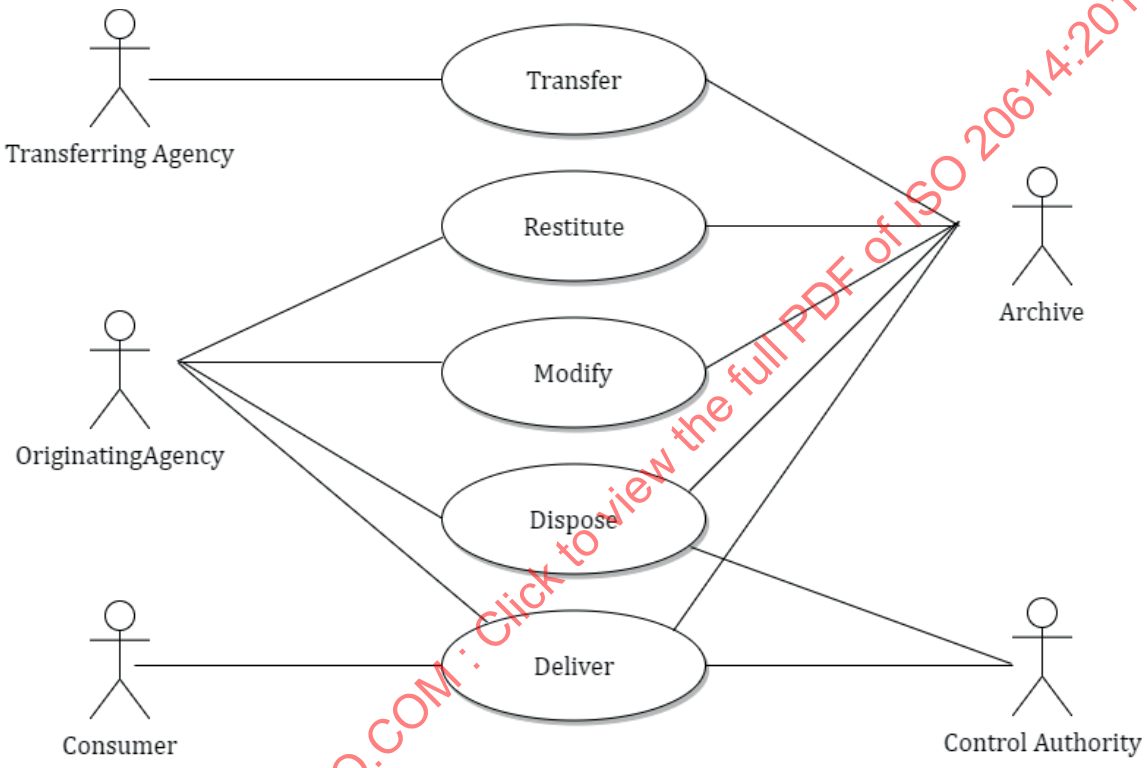

Figure 1 — General use case diagram

5.2.2 Transfer

During Transfer, the Transferring Agency should transmit to the Archive the following information:

— information concerning the transfer itself (identification of the Transferring Agency and of the Archive, date of sending the message);

— management information (identification of the Submission Agreement between these two parties);

— information on the Data Objects to be preserved (administrative and descriptive metadata).

If they are digital, the Data Objects themselves may be joined to the message. After the transfer, in case of an acceptance by the Archive, it is its responsibility to retain the transferred information. The Submission Agreement shall specify whether there is also a transfer of responsibility.

The Transfer may be preceded by a Transfer request for agreement that allows the Transferring Agency to check with the Archive that the planned transfer is acceptable by sending, for instance, only the metadata for authorization. Figure 2 shows the Transfer preceded by a Transfer request.

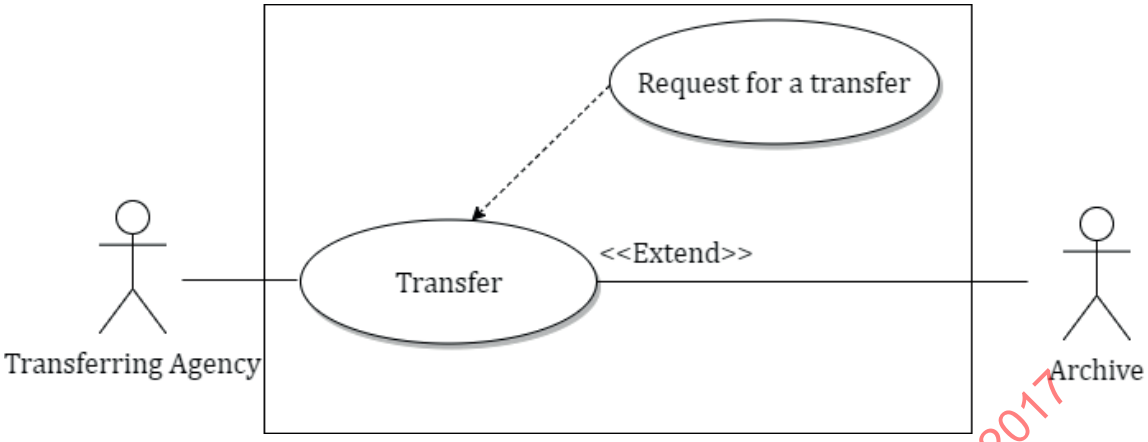

Figure 2 — Use case diagram: Transfer

5.2.3 Deliver

A request to deliver a preserved Data Object may be sent by the Originating Agency or, more indirectly, from any person with an interest in consulting the Data Object (i.e. a Consumer) for administrative, legal, litigious or historical reasons. The Originating Agency may always access the Data Objects it has submitted and which have been archived unless legal, regulatory or contractual exceptions requiring authorization of the Control Authority exist.

NOTE 1 Authorization from the Control Authority shall be requested for the delivery of personal data. Once the retention periods defined for the initial purpose of the processing have expired, they are no longer of use for their initial purpose. The Authorization request also applies to the Originating Agency.

NOTE 2 If the Consumer is the Originating Agency, the authorization from the Originating Agency is considered tacit.

Consumers may need an authorization from the Originating Agency, the Archive and/or from the Control Authority because of legal, regulatory or contractual conditions.

Figure 3 shows the Delivery preceded by authorization requests sent to the Originating Agency or to the Control Authority.

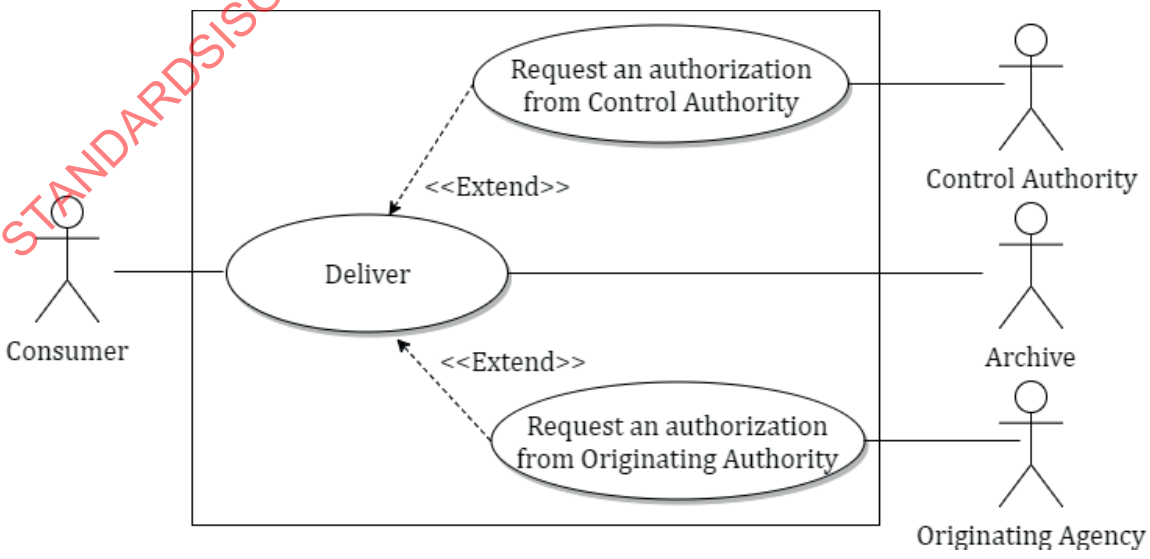

Figure 3 — Use case diagram: Deliver

5.2.4 Modify

The Archive shall maintain, in accordance with the Submission Agreement, a list of authorized modification operations, such as:

— modification of metadata;

— migration of Data Objects (format conversion in case of obsolescence of the format in which the data has been submitted);

— notification [this transaction allows the Archive to send a Dissemination Information Package (DIP) containing the migrated Data Objects and modified metadata to the Originating Agency].

When the Archive has migrated the preserved Data Objects, or modified the associated metadata, the Archive shall inform the Originating Agency about the changes. Figure 4 shows the notification of the modification operations.

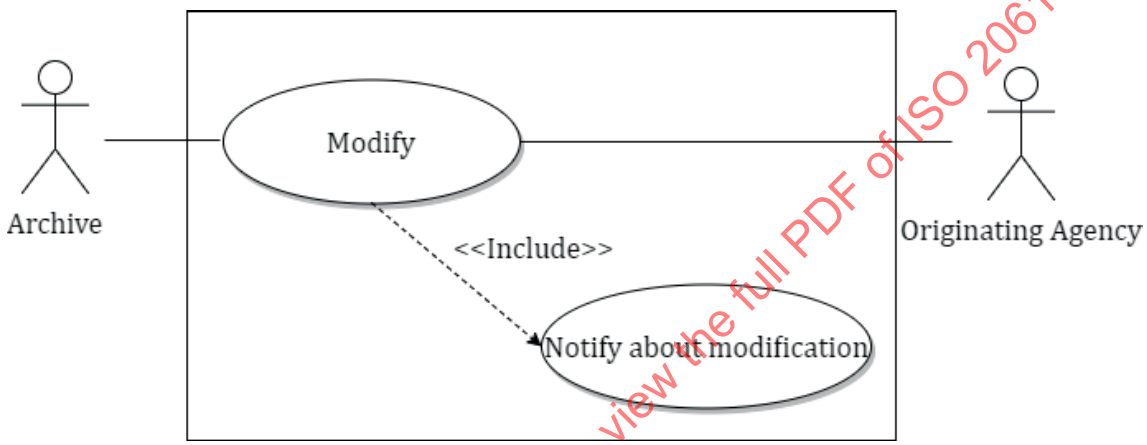

Figure 4 — Use case diagram: Modify

5.2.5 Dispose

The Disposal procedure applies when an Archive disposes of Data Objects whose retention period has expired. Disposal is usually based either on the Submission Agreement or a separate transaction process agreement. The Archive may still check if the Originating Agency wants the data to be deleted before proceeding with the disposal[1)]. An authorization from the Control Authority may also, according to the current legal, regulatory or contractual framework, be required.

Figure 5 shows the Disposal notification, preceded by the Authorization request sent to the Originating Agency and eventually to the Control Authority.

1) See ISO 14641-1 for instance.

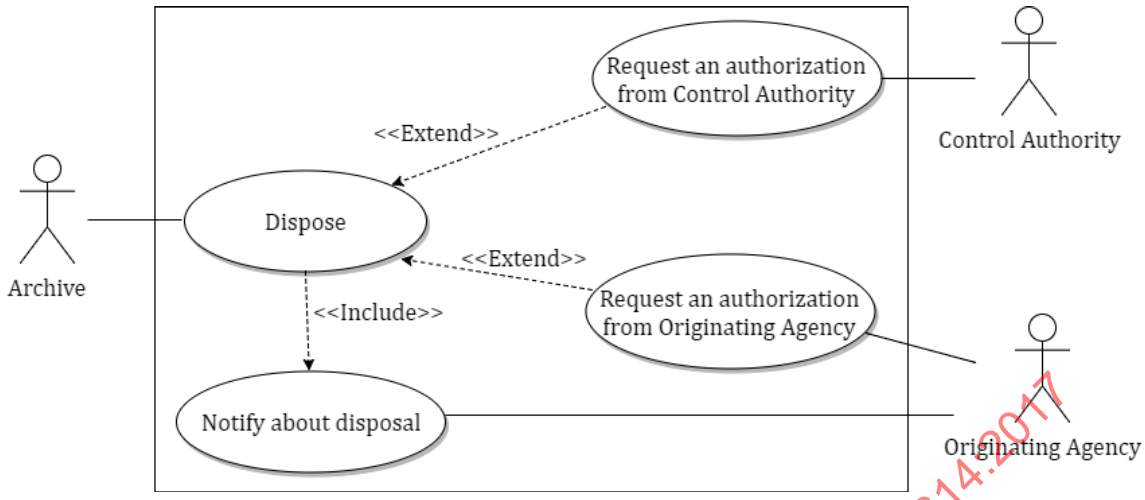

Figure 5 — Use case diagram: Dispose

If the Originating Agency has a copy of the expired Data Object in its own information system, it is the responsibility of the Originating Agency to dispose of it. The Archive is not able to request such an operation since it may not know that a copy exists. However, the Originating Agency may need to obtain the authorization of the Control Authority to dispose of the Data Object. Like the Disposal notification sent from the Archive to the Originating Agency, the Authorization request to the Originating Agency shall be considered tacit.

5.2.6 Restitute

Restitution is a request to return the archived Data Object to the Originating Agency or a third party appointed by it. Restitution may also concern another case, namely, the reactivation, at the request of an Originating Agency, of a preserved file. In this case, the Restitution may be partial and may not cover all the information contained in the original transfer.

This Restitution may be done at the request of the Originating Agency or at the request of the Archive, for example, at the end of the contract binding an Originating Agency and a storage service supplier. The transaction is carried out in two steps: a request, followed (in case of agreement) by a transfer from the Archive to an Originating Agency. These steps are shown in Figure 6.

If Restitution is supported, the Submission Agreement shall specify the characteristics of the Dissemination Information Package to be created (Data Objects in their original format and/or in different file formats after successive migrations, metadata including descriptions of migrations as events, etc.). Once the Restitution is made, the Archive, according to the applicable legal, regulatory or contractual provisions, shall eventually delete the Data Objects and other information concerned.

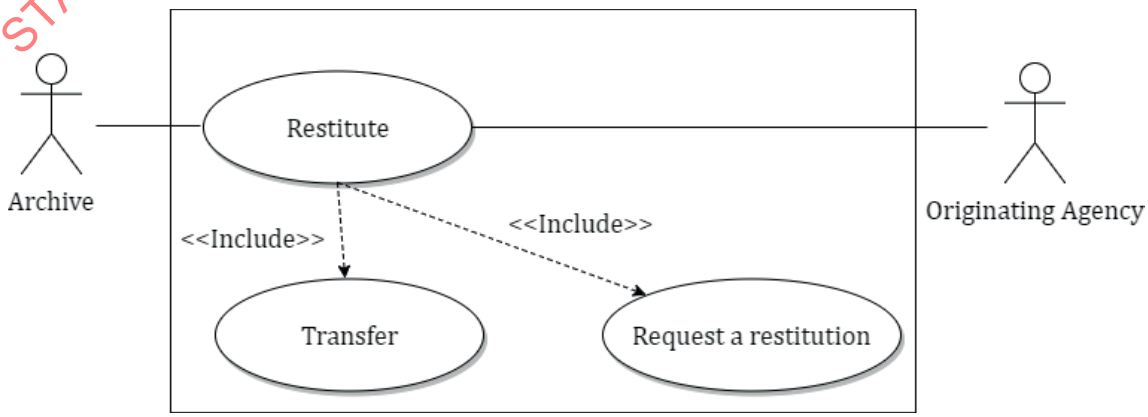

Figure 6 — Use case diagram: Restitute

5.3 Sequence diagrams

5.3.1 General

The sequence diagrams presented below describe the dialogue between the individuals and organizations involved in the context of a transaction. These diagrams identify messages that are sent by the Archive and its partners and describe the sequence of these messages. To facilitate interoperability between information systems, compliance with the order in which exchanges shall be done within each use case is particularly important.

The sequence diagrams link together four messages: a request, an acknowledgment of the request, a reply to the request and an acknowledgement of the reply. It should be noted that the formalism of acknowledgments is only given as a proposal and should be used only when needed, that is, when acknowledgement is not supported by another protocol used. Identifying whether acknowledgments must be included in exchanges, as well as their form or style, should be specified in a Submission Agreement or other formal agreement between the parties.

5.3.2 Transfer

Figure 7 shows the sequence of the different messages that are sent by the Archive and the Transferring Agency during Transfer.

The Transfer request is optional. It allows the Transferring Agency to check with the Archive that the intended transfer meets the requirements of the Submission Agreement (regarding the content of the Data Objects, their volume, the frequency or scheduling of transfers). The content of the request may be, for example, just some information about the agreement.

The Transfer request should be followed by an acknowledgement (optional) and a reply (acceptance or error message) sent by the Archive to the Transferring Agency which in turn acknowledges that it has received the reply.

If the Transfer request is accepted, the Business Identifier of the reply to the Transfer request for agreement should be used in the following Transfer message.

The Transfer message comprises the Data Objects to be transferred and their metadata. The Transfer message is sent by the Transferring Agency to the Archive.

The Archive sends an acknowledgement to the Transferring Agency immediately after it receives the Transfer message.

The Archive checks that the transferred information meets all the conditions specified in the Submission Agreement or some other service contract previously accepted by both parties. Then, either an acceptance notification or an error message is sent by the Archive to the Transferring Agency that acknowledges its receipt.

The reply for the acceptance may include the metadata of the transferred information object, so that the Transferring Agency may keep track of what it has sent.

At the end of the transaction, in the case of acceptance, the information object has been transferred from the Transferring Agency to the Archive and the responsibility for information retention shall lie with the Archive, if the ingest process is completed successfully.

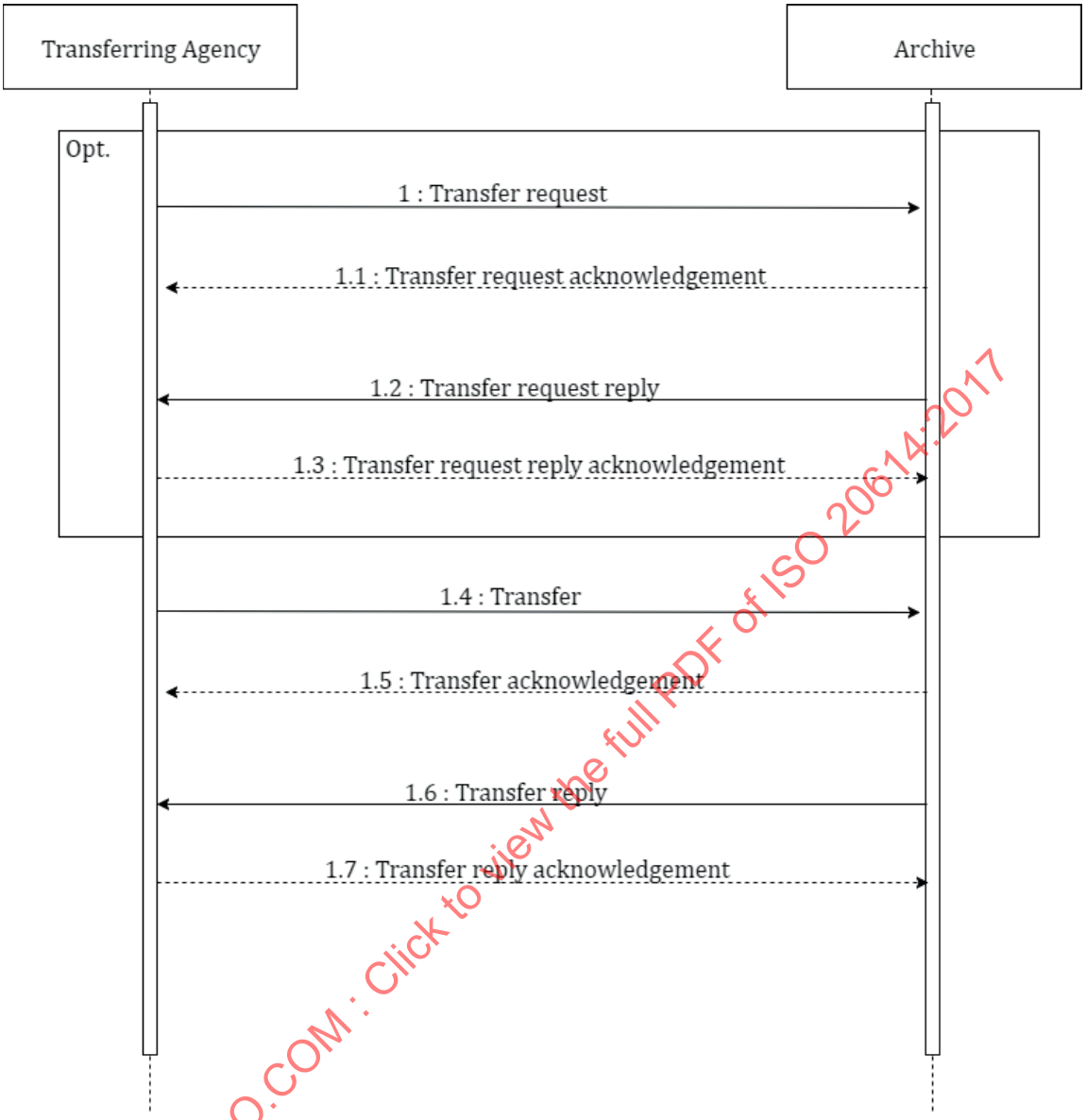

Figure 7 — Sequence diagram: Transfer

5.3.3 Deliver

Figure 8 illustrates the sequence of the messages that are exchanged by the Archive and its partners involved in the Delivery transaction.

The Delivery request is made by a Consumer (either the producer or a third party) that wishes to consult the preserved information. A Dissemination Information Package (DIP) may contain both data and related metadata or only the former or the latter, or the entire archived Data Object or just part of it.

The Archive should send an acknowledgement to the Consumer immediately after the Delivery request has been received.

If delivery requires authorization, one or more Authorization requests shall be made following the posting of acknowledgement message but prior to sending the corresponding Delivery request reply message.

After the examination of the request, and if necessary after authorization has been received, a Delivery request reply shall be sent by the Archive to the Consumer. This reply may be negative (for instance,

where the requested information does not exist, or where the Control Authority objects to the delivery) or affirmative (in which case, the reply includes the requested information).

After reception of the reply, the Consumer sends back an acknowledgement message.

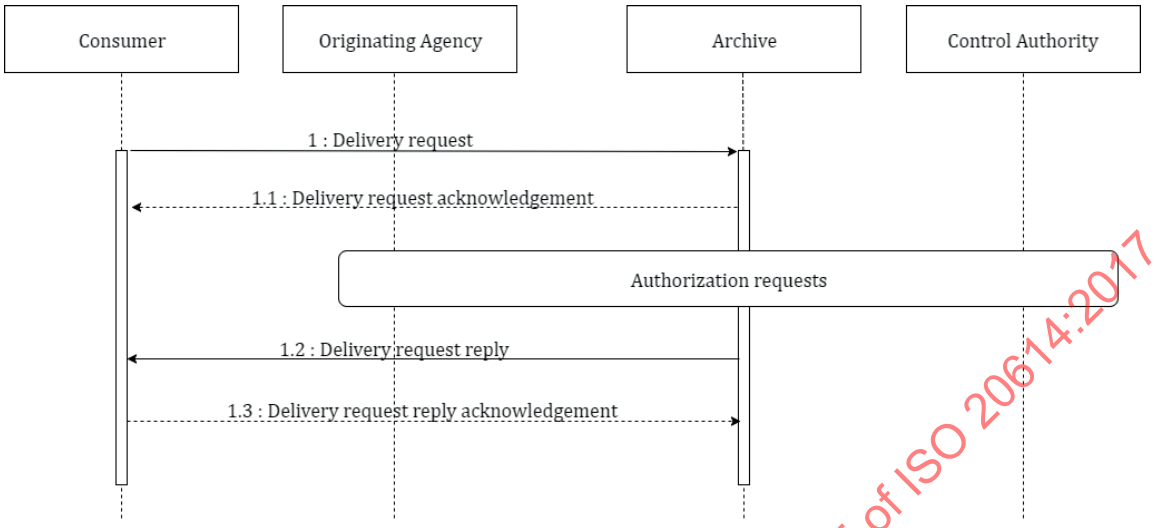

Figure 8 — Sequence diagram: Deliver

5.3.4 Modify

Figure 9 shows the exchange of messages between the Archive and the Originating Agency during the Modification transaction.

The Submission Agreement should allow the Archive to maintain a list of authorized modifications that may be made on the archived or disseminated data (especially file format migrations when the original format of the archived data becomes obsolete).

The Modification notification message comprises the technical identifiers of the modified Data Objects, and specifies the nature of the modification applied (e.g. format migration, metadata updates). Modified Data Objects may also be included in the notification.

Upon receiving a Modification notification, the Originating Agency may respond with an acknowledgement message.

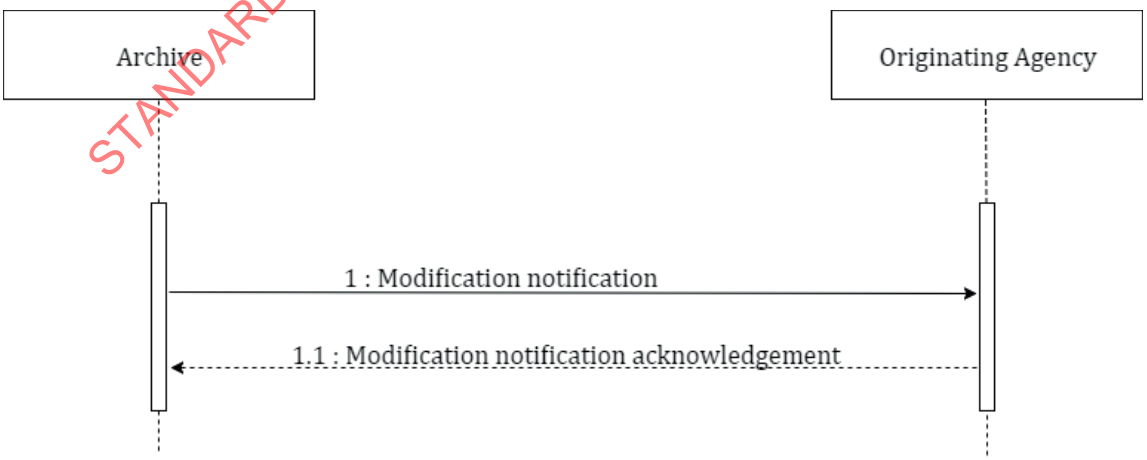

Figure 9 — Sequence diagram: Modify

5.3.5 Dispose

Figure 10 shows the sequence of the messages that are exchanged during the Disposal transaction.

This exchange of messages may occur only after the Archive has obtained an authorization from the Originating Authority to proceed with the disposal of a digital object. Authorizations are obtained either by making Authorization requests to the Originating Agency or by specifying the disposal terms in the Submission Agreement.

Once the authorization has been obtained, the Archive may proceed with the disposal in compliance with the procedure laid down by the Submission Agreement. Once the process is complete, the Archive notifies the Originating Agency of the disposal. The Originating Agency responds with an acknowledgement. The notification may, if necessary, include a reference to the authorization of the Control Authority or to the corresponding Submission Agreement.

When an Originating Agency requests disposal of an information object that it holds in its own information system, an Authorization request to the Control Authority may be sent, if required by legal, regulatory or contractual provisions.

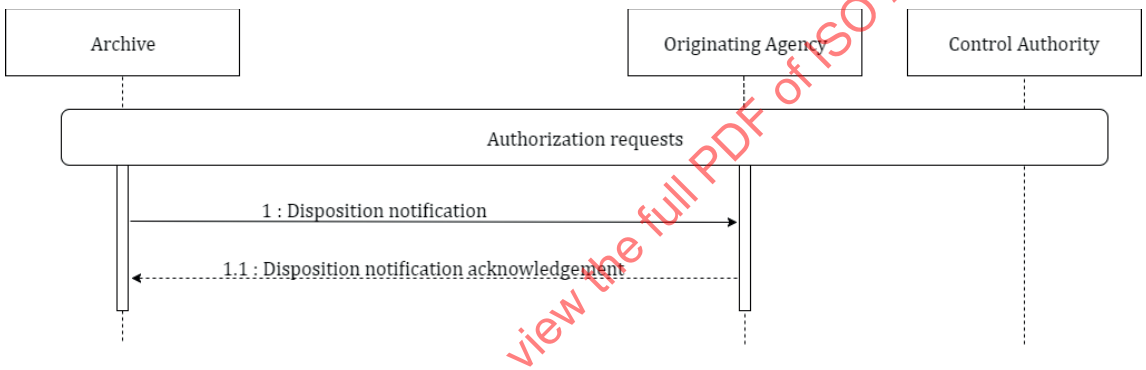

Figure 10 — Sequence diagram: Dispose

5.3.6 Restitute

The Restitution transaction is divided into two sequences: a sequence of Restitution request followed by a sequence of Transfer. These two sequences are shown in Figure 11.

The Restitution request may be initiated

— by the Archive that holds information to be returned. The Archive may routinely return information objects to the Originating Agency when their preservation period has expired, or the Archive may return all the information objects belonging to an Originating Agency at the end of the agreement they had together, or

— by the Originating Agency, for instance, when an Originating Agency needs to reactivate a business file which is no longer present in its production systems, or when an Originating Agency has used a third-party service to preserve its digital information but has decided to take the responsibility back to itself (or to pass it to another third party).

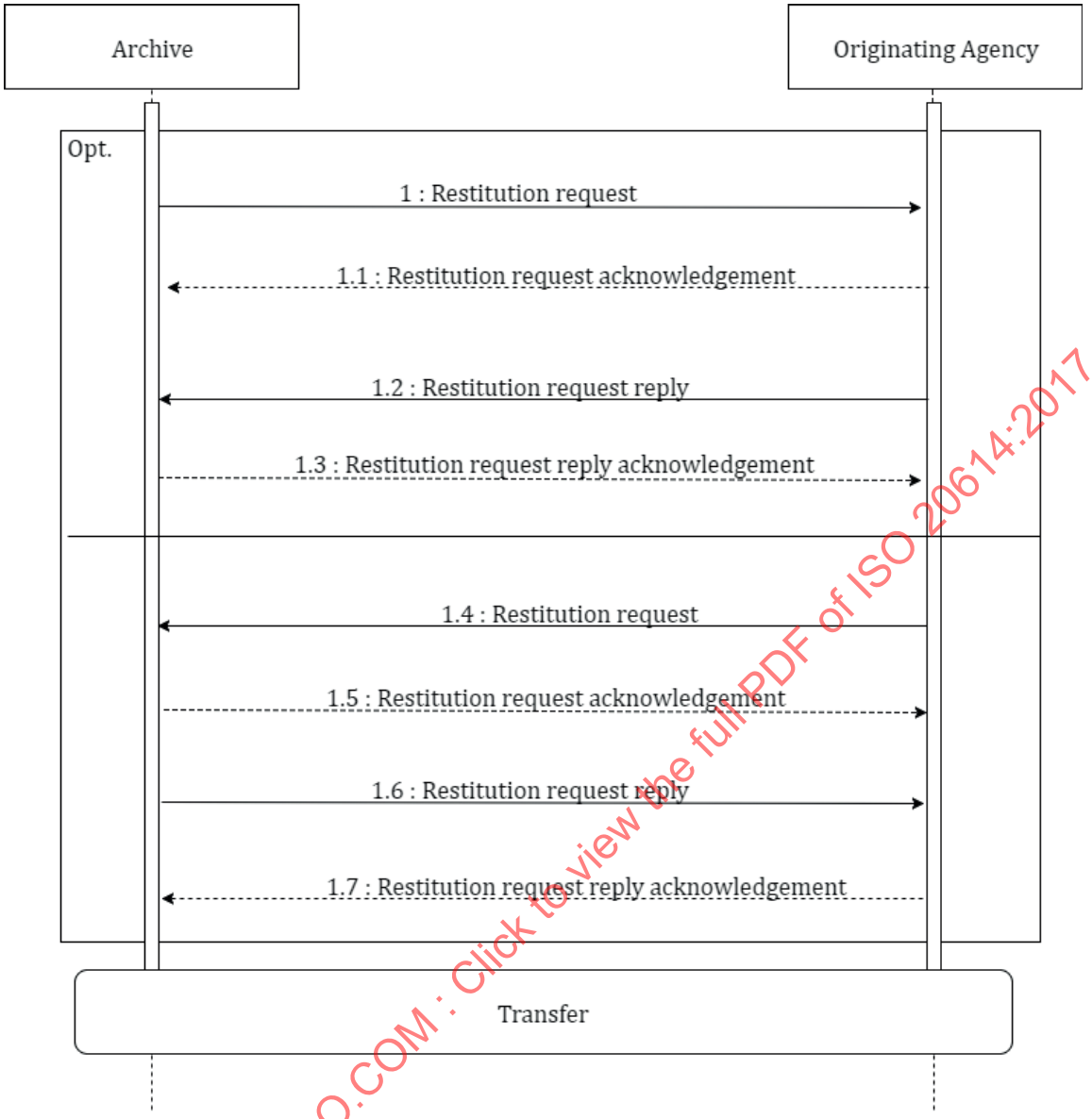

Figure 11 — Sequence diagram: Restitute

The request made includes a list of the technical identifiers of the Data Objects requested (potentially accompanied by metadata about the objects). The agency receiving the request shall respond with an acknowledgement, followed by a reply (either acceptance or refusal of the Restitution request) and the initiator of the request sends back an acknowledgement.

If the Restitution request is accepted, the actual transfer of information objects between the Archive and the Originating Agency follows the normal Transfer procedure. Note that in this case, the Archive that holds information to be returned acts as a Transferring Agency, while the Originating Agency of this information acts as an Archive. At the end of the transfer, information and the responsibility for its retention have been transferred from an individual or organization to another.

5.3.7 Authorization requests

5.3.7.1 General

Figures 12 and 13 (Authorization request to the Originating Agency and Authorization request to the Control Authority) have been isolated from the others, as they are used as additional sequences in Dispose and Deliver.

5.3.7.2 Authorization request to the Originating Agency

An Archive may obtain an authorization from the Originating Agency by sending to it an Authorization request message. The Originating Agency responds immediately with an acknowledgement. After the request has been considered, the Originating Agency notifies the Archive of its decision by sending an Authorization request reply message, which may contain either an acceptance or a rejection. The Archive shall respond to the message with an acknowledgement. Figure 12 shows the Authorization request to the Originating Agency.

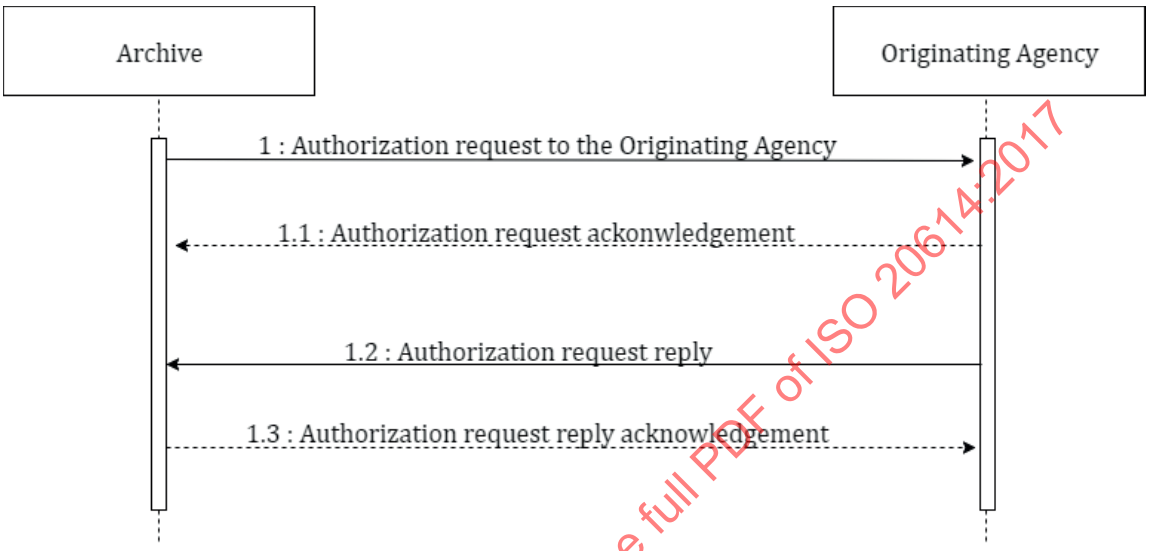

Figure 12 — Authorization request to the Originating Agency

5.3.7.3 Authorization request to the Control Authority

An Archive may obtain an authorization from the Control Authority by sending to it an Authorization request message. The Control Authority responds with an acknowledgement. After the request has been considered, the Control Authority notifies the Archive of its decision by sending an Authorization request reply, which may contain either an acceptance or a rejection. The Archive shall respond to the message with an acknowledgement.

An Authorization request to the Control Authority may include reply messages to earlier authorization requests concerning the same information object (for instance, the AuthorizationOriginatingAgencyRequestReply message) that have been obtained in advance.

Figure 13 shows the Authorization request to the Control Authority.

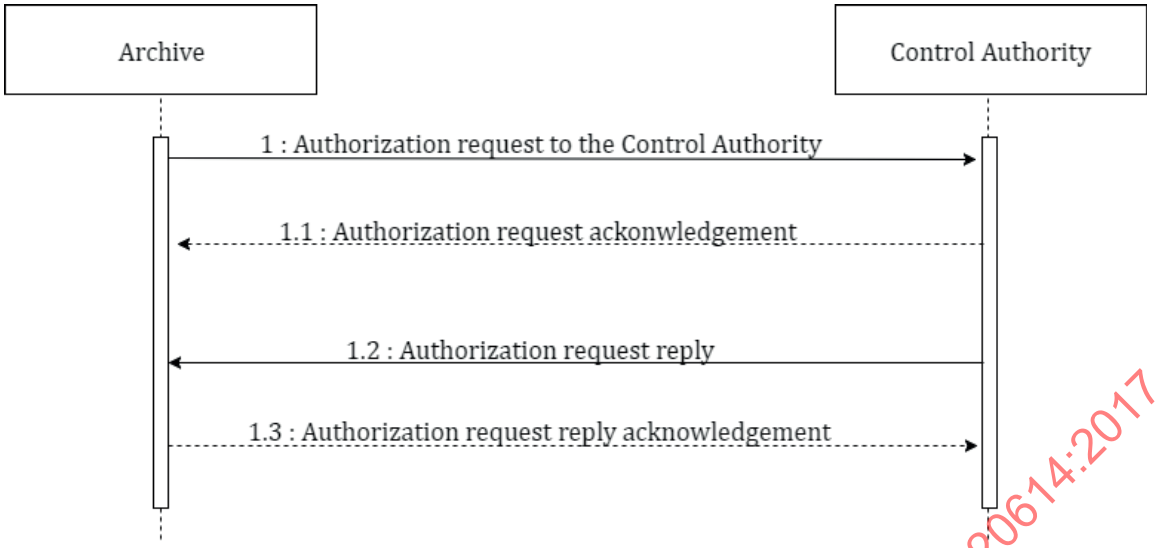

Figure 13 — Authorization request to the Control Authority

5.3.8 List of messages

Table 1 shows for each sequence the messages to use in the order in which they occur, giving each time the name of the corresponding class.

For the details of each message, refer to the sequence diagrams in the previous subclauses.

Table 1 — Protocol messages

Sequence	Message to use
Transfer	
Transfer request	PackageTransferRequest
Transfer request acknowledgement	Acknowledgement
Transfer request reply	PackageTransferRequestReply
Transfer request reply acknowledgement	Acknowledgement
Transfer	PackageTransfer
Transfer acknowledgement	Acknowledgement
Transfer reply	PackageTransferReply
Transfer reply acknowledgement	Acknowledgement
Deliver	
Delivery request	PackageDeliveryRequest
Delivery request acknowledgement	Acknowledgement
Delivery request reply	PackageDeliveryRequestReply
Delivery request reply acknowledgement	Acknowledgement
Modify	
Modification notification	PackageModificationNotification
Modification notification acknowledgement	Acknowledgement
Dispose	
Disposal notification	PackageDestructionNotification
Disposal notification acknowledgement	Acknowledgement
Restitute	

Table 1 *(continued)*

Sequence	Message to use
Restitution request	PackageRestitutionRequest
Restitution request acknowledgement	Acknowledgement
Restitution request reply	PackageRestitutionRequestReply
Restitution request reply acknowledgement	Acknowledgement
Authorization to the Originating Agency	
Authorization request to the Originating Agency	AuthorizationOriginatingAgencyRequest
Authorization request acknowledgement	Acknowledgement
Authorization request reply	AuthorizationOriginatingAgencyRequestReply
Authorization request reply acknowledgement	Acknowledgement
Authorization request to the Control Authority	
Authorization request to the Control Authority	AuthorizationControlAuthorityRequest
Authorization request acknowledgement	Acknowledgement
Authorization request reply	AuthorizationControlAuthorityRequestReply
Authorization request reply acknowledgement	Acknowledgement

5.4 Class diagrams

5.4.1 General

The following class diagrams describe the structure of messages exchanged by the Archive and its partners in the context of transactions, as well as the structure of the information objects handled within these messages.

5.4.2 Organizations

The class Organization (see Figure 14) makes it possible to describe the relevant actors (such as Transferring Agency, Archive and Originating Agency). Organizations shall be identified (Identifier) and it is also possible to describe them in the class OrganizationDescriptiveMetadata. The model (metadata format) to be used for this is undefined.

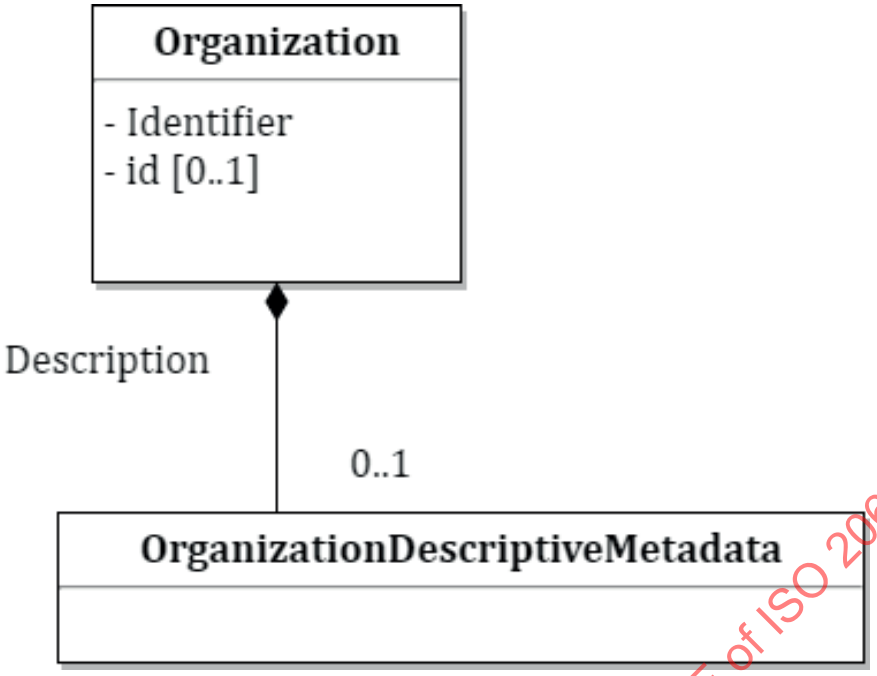

Figure 14 — Organization class

5.4.3 Data Object packages

5.4.3.1 General

The class DataObjectPackage (see Figure 15) represents an unordered set of Data Objects (DataObject) with its administrative (AdministrativeMetadata) and descriptive metadata (DescriptiveMetadata).

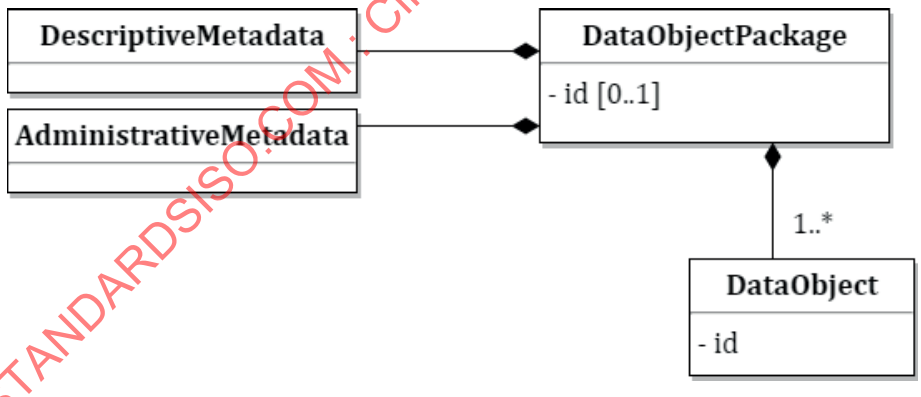

Figure 15 — DataObjectPackage class

5.4.3.2 Data Objects

Figure 16 shows the class DataObject.

Data Object (DataObject) is either binary (BinaryDataObject) or physical (PhysicalDataObject), such as a printed document.

A Data Object may contain links (Relationship) to other Data Objects in the same information package. A link shall specify the technical identifier of the target (target) which is a Data Object in the same package, and the nature of the relationship (type) between the Data Objects.

A Data Object shall have a technical identifier for unambiguous referencing from the metadata record to the Data Object or between Data Objects.

A Binary Data Object contains (directly embedded in the body of the message in base64) or references (in the form of the file name or a URI) a specific Data Object (Attachment) with some technical information about it:

— its file format (Format);

— its digest (MessageDigest);

— its size in bytes (Size);

— the status indicating digital signature characteristics (presence or absence of an electronic signature, verified electronic signature, etc.) (SignatureStatus).

A Physical Data Object does not have any other technical information than size (Size) of the medium or the container, expressed in units such as number of folders, number of boxes, linear meters, etc.

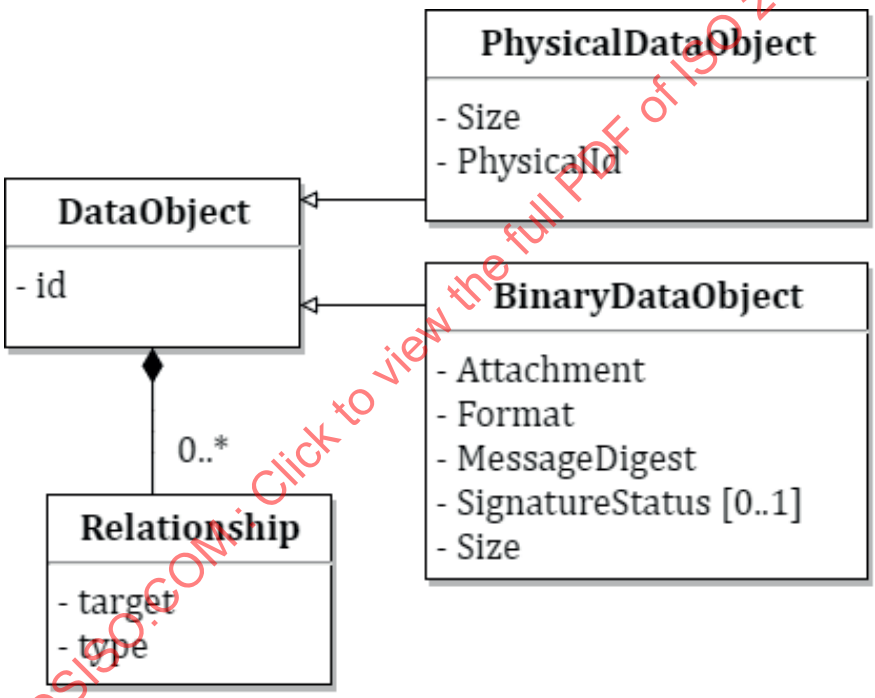

Figure 16 — DataObject class

5.4.3.3 Administrative metadata

The AdministrativeMetadata class (see Figure 17) includes all information required for the preservation of the Data Objects in the Submission Information Package.

NOTE Submission Information Packages can be incomplete; they can contain, for instance, descriptive metadata about the preserved Data Object or the Data Object in a format which is not suitable for long term preservation.

This metadata concerns all the Data Objects in the package, and it consists of the Preservation profile (PreservationProfile), the Service Level (ServiceLevel), Rights metadata (AccessRule) and the rule for calculating the disposal (AppraisalRule). All these information details are optional.

The AppraisalRule class allows someone to indicate the preservation period (Duration), the start date for calculating the period (StartDate) and the disposition (AppraisalCode) which shall be applied at the end of the preservation period (preservation or disposal). The three items of information as constituent

parts of this rule are optional. An Archive may have also other sources for acquiring preservation related information. For instance, Submission Agreements may specify default preservation periods and disposition policies.

The AccessRule class allows the authorized parties to specify regulations concerning access to the Data Objects and to express in particular if access restrictions shall be implemented by giving all the required information. Access information may be expressed in different Rights metadata formats depending on the practices of the partners of the exchanges.

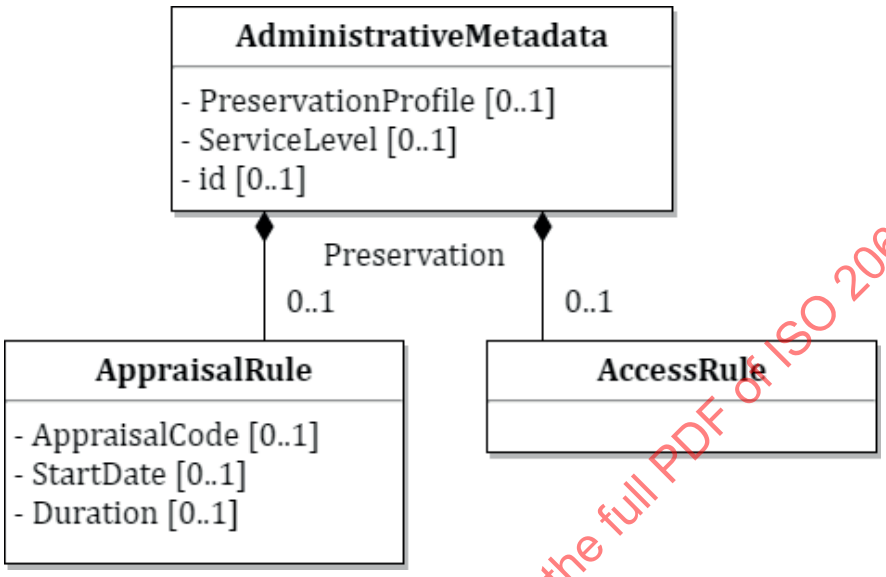

Figure 17 — AdministrativeMetadata class

5.4.3.4 Descriptive metadata

The DescriptiveMetadata class allows someone to describe all the Data Objects contained in the package with indexing, context, provenance, arrangement information, etc. If some metadata in this class is similar to metadata already available in other classes in the same information package, the Submission Agreement should specify how to interpret the metadata and how to manage conflicts, in any.

Metadata may be based on different formats based on the preferred practices of the participating organizations.

5.4.4 Specification of the version of the lists of codes used

The CodeListVersions class (see Figure 18) allows the users to specify versions of the code lists used in a message. These lists are

— appraisal code list (AppraisalCodeListVersion),

— authorization reason code list (AuthorizationReasonCodeListVersion),

— file encoding code list (FileEncodingCodeListVersion),

— file format code list (FileFormatCodeListVersion),

— digest algorithms code list (MessageDigestAlgorithmCodeListVersion),

— code list for relationships between files (RelationshipCodeListVersion),

— list of reply codes used in the reply messages (ReplyCodeListVersion), and

— signature status code list for a Data Object (presence or absence of electronic signature, verified electronic signature, etc.) (SignatureStatusCodeListVersion).

The specified code lists shall be used in all transactions. Only the codes contained in these lists may be used in messages; using invalid codes or other code lists increases the risk of having the requests rejected. These lists may be modified to meet the changing requirements of the market and the new technical developments (such as updated and new file formats and metadata formats, new algorithms, etc.). Therefore, each list shall be identified by a unique identifier and its current version number. If the name of the code list is well-known, the Archive and its partners may agree to include the name in the messages. This agreement may be recorded in the Submission Agreement.

This document does not mandate the use of specific code lists. However, some lists are industry standards and, therefore, their use is recommended. These code lists are currently published and maintained in various ways. The organizations using DEPIP should, if necessary, define their own processes for retrieving, updating and formatting the code lists relevant to them.

To identify file formats, the PRONOM registry[2)], maintained by The National Archives of the United Kingdom (TNA), is recommended. It is commonly used by, for example, libraries, archives and museums, and many software tools use it to identify and validate file formats. PRONOM identifies a large number of file formats and is able to distinguish different versions of these formats even when they are based on differing requirements (for instance, the specifications of PDF by Adobe[3)] and the ISO standards derived from them). PRONOM registry is frequently updated (approximately every month).

Regarding the identification of file encodings, there is no unique code list covering all needs of all families of file formats. For text files, however, the recommended solution is the registry containing the names for character codes compiled by IANA[4)]. UN/CEFACT publishes a version of the IANA registry in the form of an XML schema every six months[5)].

There are many algorithms which may be used to calculate message digests. The vocabulary "Cryptographic Hash Functions" maintained by the Library of Congress[6)] is the recommended code list of these algorithms. It is also possible to use the identification system of the OID archive[7)], in particular, for certificates for electronic signatures.

There are many kinds of possible relationships between Data Objects. Dublin Core, maintained by DCMI[8)], contains several metadata terms which are relevant in this context, such as Source, Relation, IsPartOf, hasPart, hasFormat, hasVersion, isReplacedBy. Generic vocabularies such as DCMI may be used as a basis for more specific relationship definitions if and when necessary. Another option is the preservation-oriented list of relationships maintained by the Library of Congress[9)].

Regarding the identification of the return codes of the reply messages, it is recommended to use and to adapt according to the local requirements a part of the list of the status codes of the http protocol[10)].

2) See: http://www.nationalarchives.gov.uk/PRONOM/.

3) Adobe is the trade name or trademark of a product supplied by Adobe System Incorporated. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.

4) Internet Assigned Numbers Authority.

5) United Nations Centre for Trade Facilitation and Electronic Business (http://www.unece.org/cefact/xml_schemas).

6) See: http://id.loc.gov/vocabulary/cryptographicHashFunctions.

7) See: http://www.oid-info.com/.

8) Dublin Core Metadata Initiative.

9) See: http://id.loc.gov/vocabulary/preservation/relationshipSubType.

10) RFC2616 section 10.

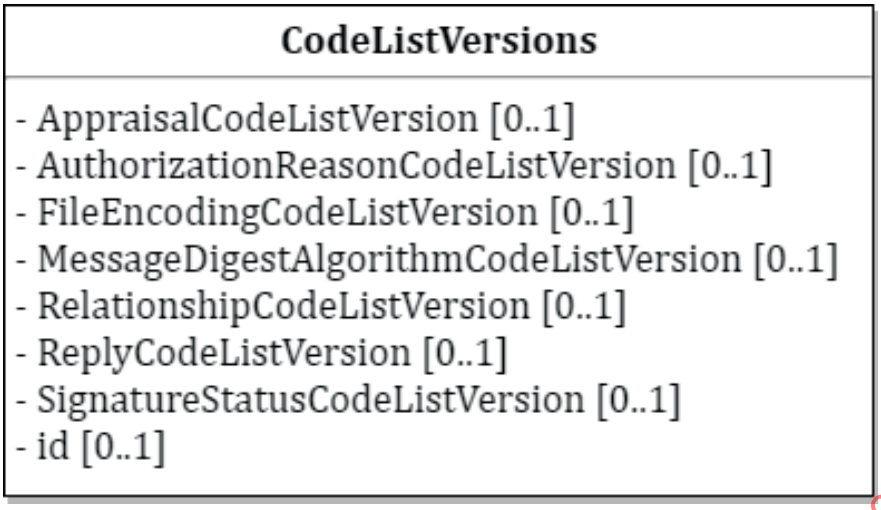

Figure 18 — CodeListVersions class

5.4.5 Signature

An optional electronic signature may be added in all messages. This enables the authentication of senders and non-repudiation of the transmitted messages. These signatures, depending on the practices of the partners, may be based on different standards such as Xades, XMLDsig and PKCS7.

5.4.6 Objects of non-specified types

There are four metadata objects identified by this document which share the same broad definition (see Figure 19). These objects are descriptive metadata of the information package (DescriptiveMetadata), descriptive metadata of organizations (OrganizationDescriptiveMetadata), signature of messages (Signature) and Rights metadata for the Data Objects in the package (AccessRule). The formats used shall be defined by the actors of the exchange. The format specification and any additional information required of these objects should ideally be included in the Submission Agreement.

The encoding of these objects may be done either in a separate and referenced file, or directly, in the body of the message.

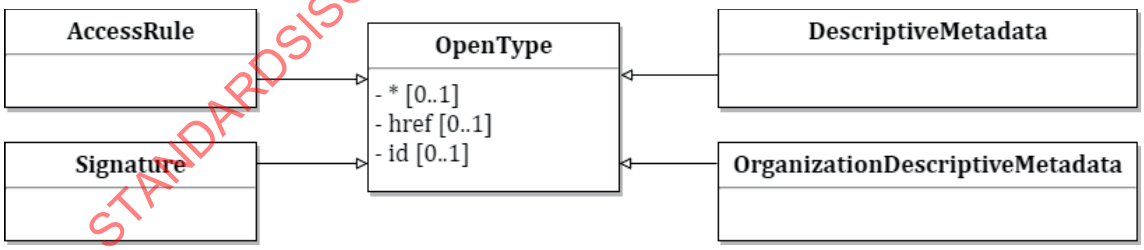

Figure 19 — Objects of non-specified types

5.4.7 Description of messages

5.4.7.1 General

All messages share common properties defined in the class Message (see Figure 20). These properties are

— the date (Date) the message was sent,

— message identifier (MessageIdentifier),

— comments (Comment), which may be used to explain the reason for the message (for instance, for a rejection, comments may specify the reason for the rejection and what needs to be done to fix the request), and

— signature of the message file (Signature).

The acknowledgement message inherits properties from the class Message, plus the identifier of the acknowledged message.

Other messages used during the transactions are grouped within the same class (BusinessMessage), which inherits all the properties from the class Message, plus

— the list of the versions of the code lists used,

— the Submission Agreement (Agreement), and

— an information package containing Data Objects with their associated metadata.

The business messages are classified in three categories:

— request messages (BusinessRequestMessage);

— notification messages (BusinessNotificationMessage);

— reply messages (BusinessReplyMessage). In addition to the common properties of messages, these also include the identifier of the message they reply to and a return code in case of rejection.

Lastly, all these messages always include information about the actors involved in the process (the sender and the recipient whose labels vary depending on the messages). They may also include specific information (identifiers of the requested Data Objects for the Delivery request, identifiers of the modified Data Objects for the Modification notification, etc.).

All messages are presented below in alphabetical order based on their names.

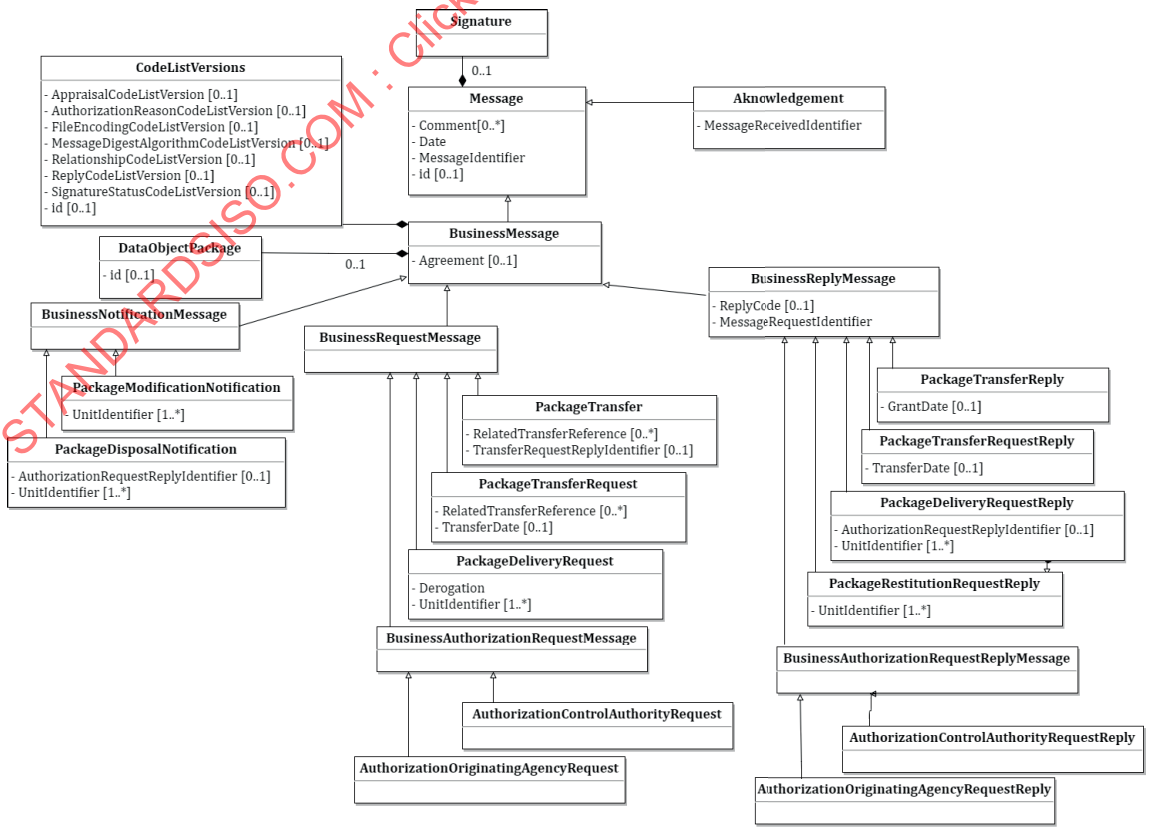

Figure 20 — Message class

5.4.7.2 Message Acknowledgement (Acknowledgement message)

Figure 21 shows the properties of an acknowledgement message.

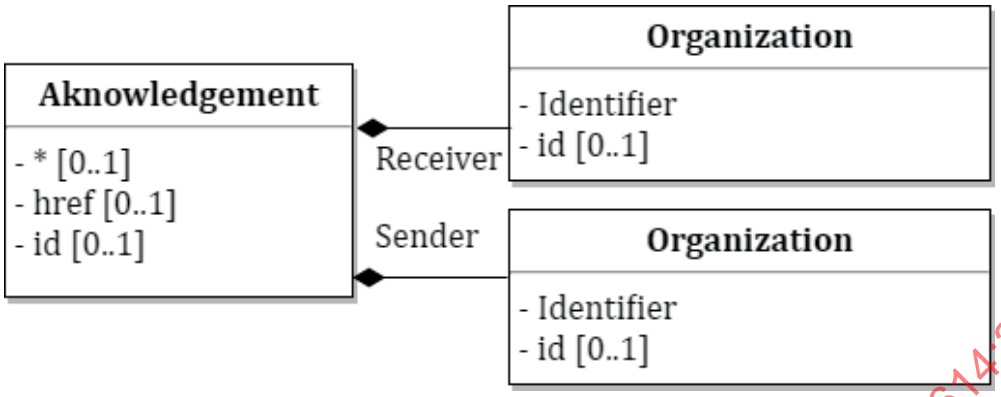

Figure 21 — Message Acknowledgement

5.4.7.3 Message PackageDeliveryRequest (Delivery request)

Figure 22 shows the properties of a Delivery request message.

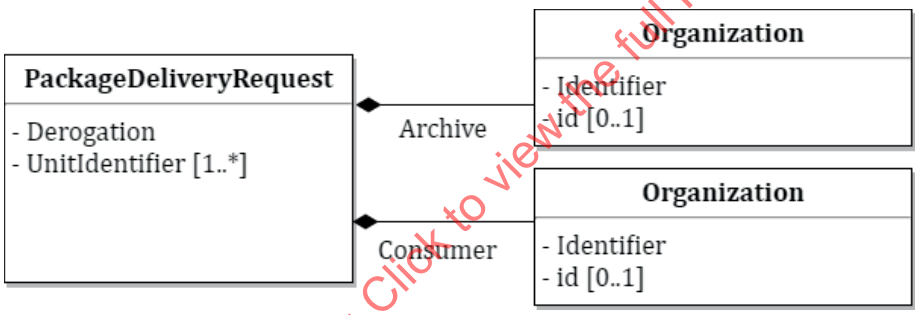

Figure 22 — Message PackageDeliveryRequest

5.4.7.4 Message PackageDeliveryRequestReply (Delivery request reply)

Figure 23 shows the properties of a Delivery request reply message.

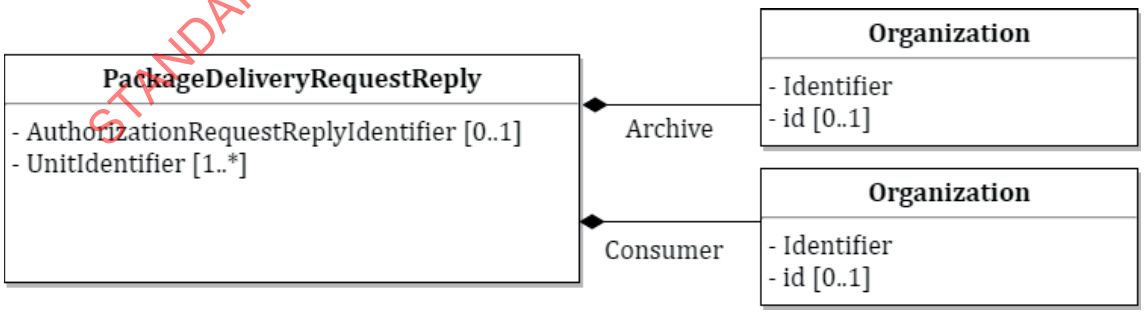

Figure 23 — Message PackageDeliveryRequestReply

5.4.7.5 Message PackageDisposalNotification (Disposal notification)

Figure 24 shows the properties of a Disposal notification message.

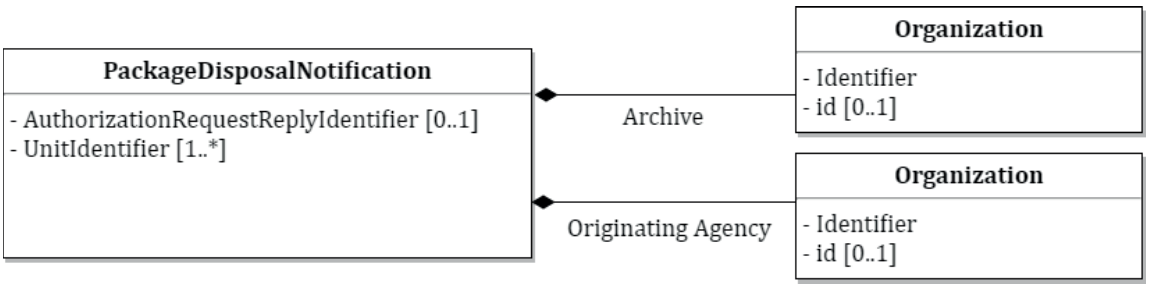

Figure 24 — Message PackageDisposalNotification

5.4.7.6 Message PackageModificationNotification (Modification notification)

Figure 25 shows the properties of a Modification notification message.

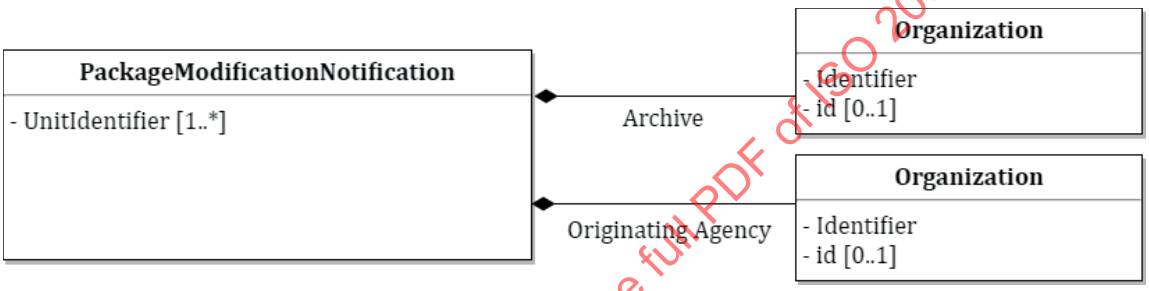

Figure 25 — Message PackageModificationNotification

5.4.7.7 Message PackageRestitutionRequest (Restitution request)

Figure 26 shows the properties of a Restitution request message.

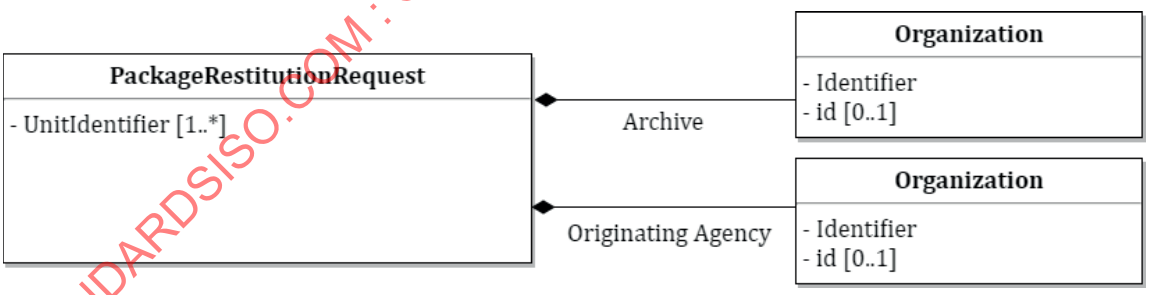

Figure 26 — Message PackageRestitutionRequest

5.4.7.8 Message PackageRestitutionRequestReply (Restitution request reply)

Figure 27 shows the properties of a Restitution request reply message.

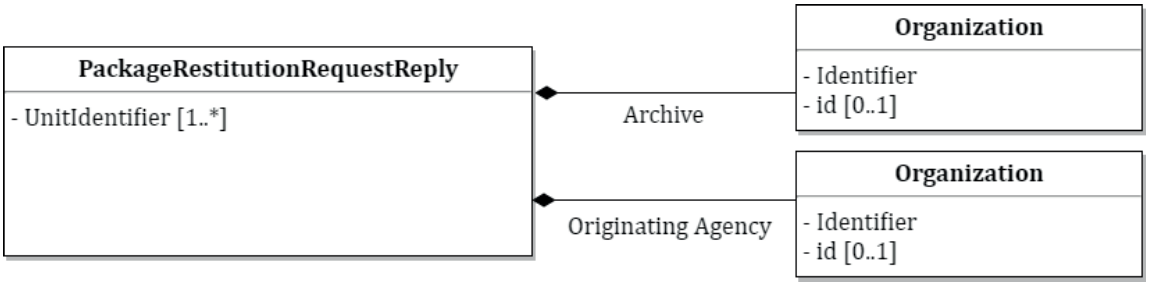

Figure 27 — Message PackageRestitutionRequestReply

5.4.7.9 Message PackageTransfer (Transfer)

Figure 28 shows the properties of a Transfer message.

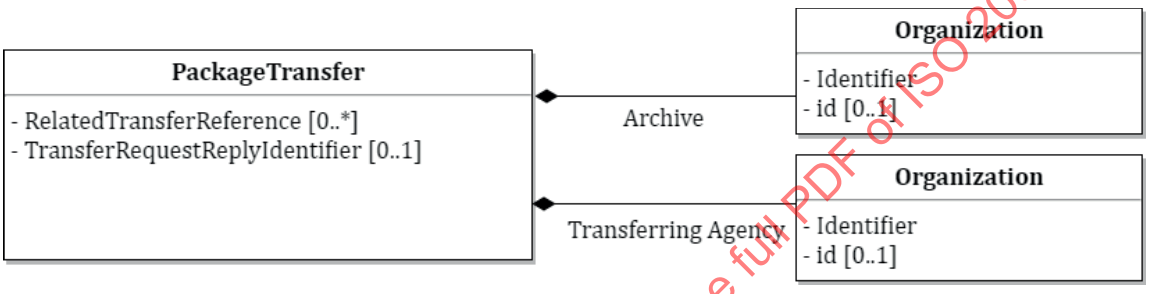

Figure 28 — Message PackageTransfer

5.4.7.10 Message PackageTransferReply (Transfer reply)

Figure 29 shows the properties of a Transfer reply message.

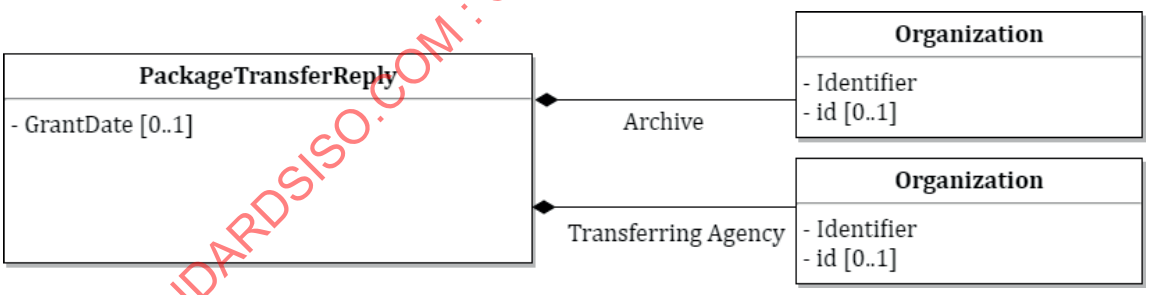

Figure 29 — Message PackageTransferReply

5.4.7.11 Message PackageTransferRequest (Transfer request)

Figure 30 shows the properties of a Transfer request message.

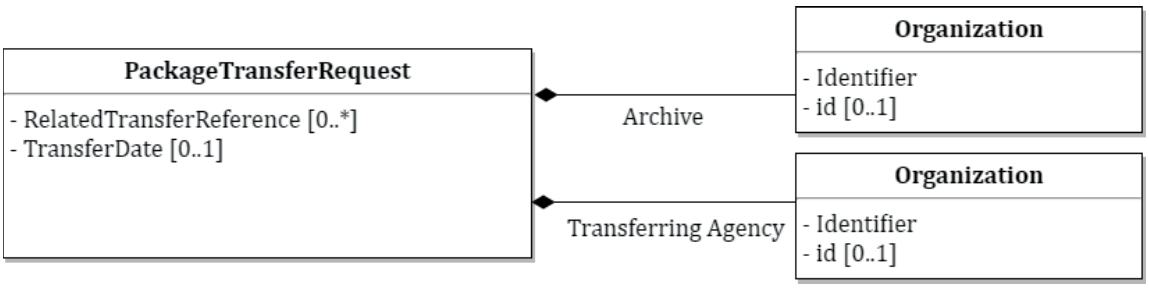

Figure 30 — Message PackageTransferRequest

5.4.7.12 Message PackageTransferRequestReply (Transfer request reply)

Figure 31 shows the properties of a Transfer request reply message.

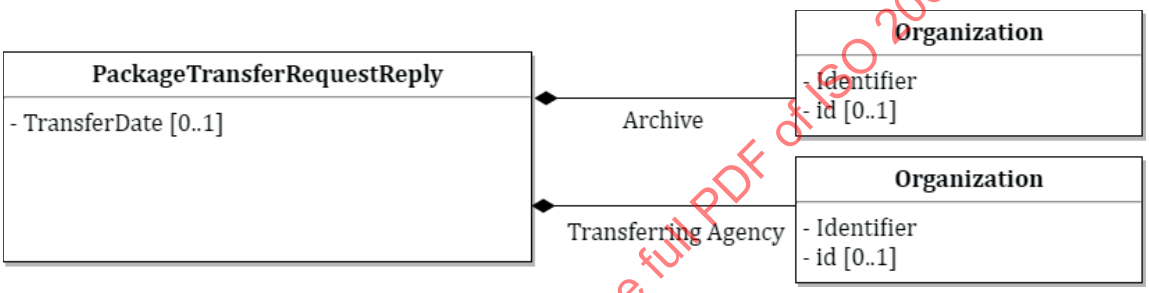

Figure 31 — Message PackageTransferRequestReply

5.4.7.13 Message AuthorizationControlAuthorityRequest (Authorization request to the Control Authority)

Both Authorization requests (AuthorizationControlAuthorityRequest and AuthorizationOriginatingAgencyRequest) include a class AuthorizationRequestContent, which allows the requester to specify the content of an authorization request. This class includes

— code corresponding to the reason of the authorization (AuthorizationReason),

— comments (Comment) that allow the requester to provide additional information such as the reason for the request,

— date on which the request has been made (RequestDate),

— technical identifier(s) of the Data Object(s) the request relates to (UnitIdentifier),

— Business Identifier and description of the requester (Requester), and

— reply messages already obtained from other services (BusinessReplyMessage).

Figure 32 shows the properties of the class AuthorizationControlAuthorityRequest.

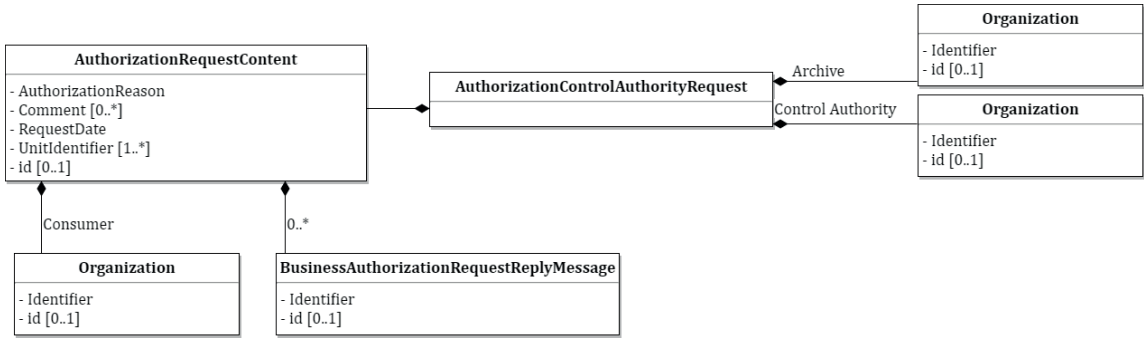

Figure 32 — Message AuthorizationControlAuthorityRequest

5.4.7.14 Message AuthorizationControlAuthorityRequestReply (authorization request reply)

Figure 33 shows the properties of the class AuthorizationControlAuthorityRequestReply.

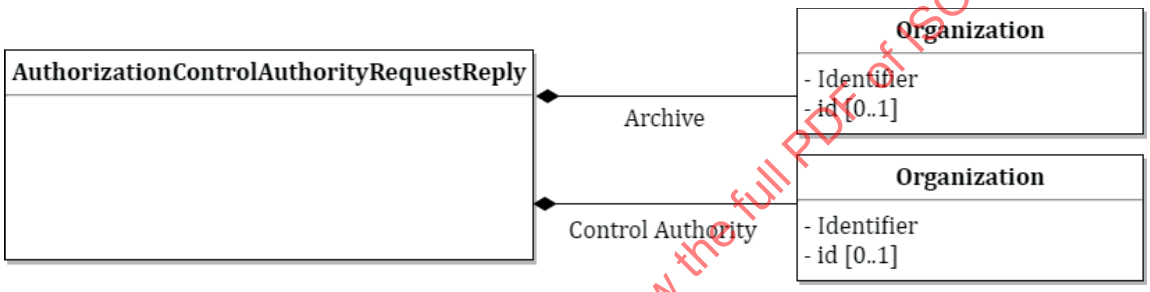

Figure 33 — Message AuthorizationControlAuthorityRequestReply

5.4.7.15 Message AuthorizationOriginatingAgencyRequest (authorization request)

Figure 34 shows the properties of the class AuthorizationOriginatingAgencyRequest.

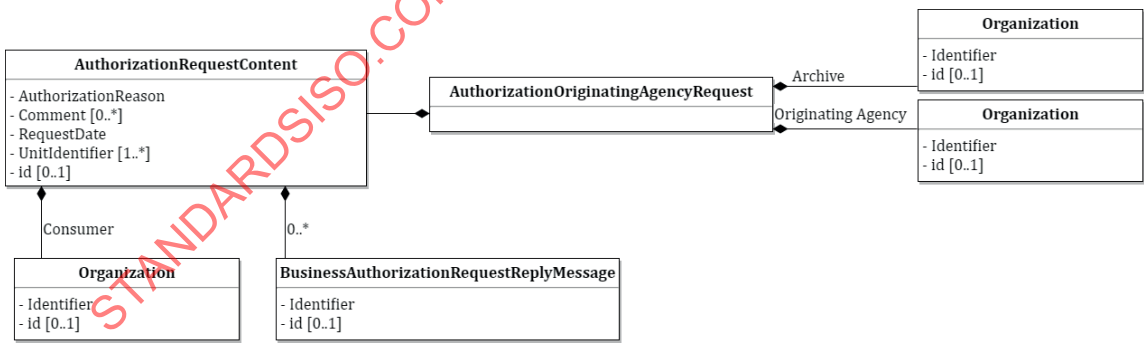

Figure 34 — Message AuthorizationOriginatingAgencyRequest

5.4.7.16 Message AuthorizationOriginatingAgencyRequestReply (authorization request reply)

Figure 35 shows the properties of the class AuthorizationOriginatingAgencyRequestReply.

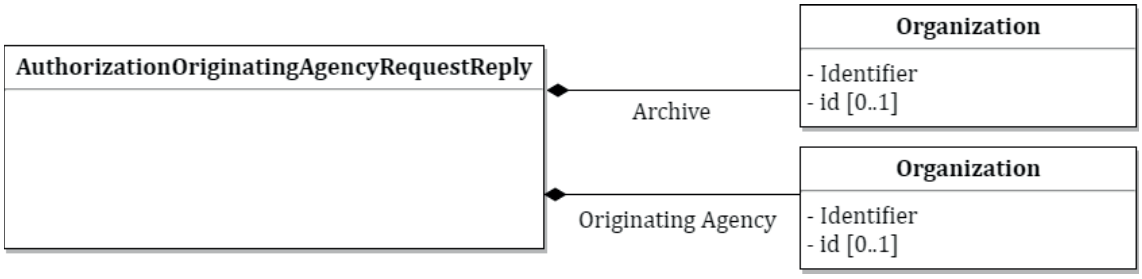

Figure 35 — Message AuthorizationOriginatingAgencyRequestReply

6 Implementation model

6.1 General

The messages may be encoded using either XML or JSON. Other encodings may be added in the future.

A website with XML or JSON schemas and additional information is available. Information about this website is available in Annex A.

6.2 Definition of types

Two types of identifiers are used in this document: Business Identifiers (IdentifierType) and technical identifiers (ID Type).

Business Identifiers may be used to identify actors, messages, Submission Agreements, etc. The proposed model allows the user to specify for each Business Identifier, in addition to the identifier string, the relevant identification schema (containing the identifier of the identifier system, its name, version, path and the Business Identifier and name of the organization which maintains it).

Technical identifiers shall be used for Data Objects; they may be used for other objects as well.

Codes (CodeType) allow the users to represent values listed in controlled vocabularies. They are used in this document to indicate versions of the code lists used. Such a model makes it possible to specify, in addition to a value itself, its textual equivalent, the language in which it is expressed, the schema used (its identifier, its name, its version, its location, and the identifier and the name of the organization which maintains it and the location of the code list). See Annex B, which gives rules for the use of code lists.

The type BinaryObjectType may be used to encode a unit of binary information. The proposed model allows the users to express the value of binary information in a "base64" encoding or to indicate a reference in the form of a URI or a file name.

The type MessageDigestBinaryObjectType may be used to encode a message digest. It is a binary information object (BinaryObjectType) to which an identification code shall be associated to specify the message digest algorithm used.

MeasureType may be used to express the size of the Physical Data Objects.

The type SizeInBytesType is a quantitative measure (MeasureType) expressed in bytes. It may be used to describe the size of Binary Data Objects.

6.3 Elements metadata

Metadata elements listed in Table 2 are categorized according to the types defined in this document.

Table 2 — Elements metadata

Type	Element	Cardinalities	Type	Definitions and comments
AccessRuleType			OpenType	Rule to be applied regarding the access to Data Objects (access)
Acknowledgement-Type			MessageType	Acknowledgement message
	MessageReceive-dIdentifier	[1..1]	IdentifierType	Identifier of the acknowledged message
	Sender	[1..1]	OrganizationType	Sender: Organization that acknowledges the message
	Receiver	[1..1]	OrganizationType	Receiver: Agency to which the acknowledgement message is sent
Administra-tiveMetadataType				Administrative metadata
	PreservationProfile	[0..1]	IdentifierType	Preservation profile: rules governing the creation of De-scriptive metadata depend-ing on the type of documents or the application concerned
	ServiceLevel	[0..1]	IdentifierType	Required service level (avail-ability, security), referring to the different levels planned by the Submission agreement
	AccessRule	[0..1]	AccessRuleType	Rule to be applied about the access to the Data Objects (access)
	AppraisalRule	[0..1]	AppraisalRuleType	Rule about the lifecycle of the information content
AppraisalRuleType	id			Rule about the lifecycle of the information content
	AppraisalCode	[0..1]	TextType	Code corresponding to the disposition authority to be applied
	Duration	[0..1]	DurationType	Retention period
	StartDate	[0..1]	DateType	Start date for calculating the period
AuthorizationCon-trolAuthorityRe-questReplyType			BusinessAuthori-zationRequestRe-plyMessageType	Authorization request reply message
	Archive	[1..1]	OrganizationType	Archive
	ControlAuthority	[1..1]	OrganizationType	Control Authority
AuthorizationCon-trolAuthorityRe-questType			BusinessAuthor-izationRequest-MessageType	Authorization request message
	Archive	[1..1]	OrganizationType	Archive
	ControlAuthority	[1..1]	OrganizationType	Control Authority
AuthorizationOrig-inatingAgencyRe-questReplyType			BusinessAuthori-zationRequestRe-plyMessageType	Authorization request reply message
	Archive	[1..1]	OrganizationType	Archive
	OriginatingAgency	[1..1]	OrganizationType	Originating Agency

Table 2 *(continued)*

Type	Element	Cardinalities	Type	Definitions and comments
AuthorizationOriginatingAgencyRequestType			BusinessAuthorizationRequestMessageType	Authorization request message
	Archive	[1..1]	OrganizationType	Archive
	OriginatingAgency	[1..1]	OrganizationType	Originating Agency
AuthorizationRequestContentType	id			Authorization request content
	AuthorizationReason	[1..1]	TextType	Authorization request reason (for disposal, for delivery)
	Comment	[0..n]	TextType	Comment
	RequestDate	[1..1]	DateType	Authorization request date
	UnitIdentifier	[1..n]	IdentifierType	Any identifier allowing someone to identify the Data Objects for which the authorization request is being made
	Requester	[1..1]	OrganizationType	The organization that asks for authorization
	AuthorizationRequestReply	[0..n]	BusinessAuthorizationRequestReplyMessageType	Authorization request reply message
BinaryDataObjectType			DataObjectType	Digital Data Object: for instance an electronic file, i.e. a named and ordered sequence of bytes, handled by the file system of an operating system as a unit
	Attachment	[0..1]	BinaryObjectType	Digital data content (a sequence of bytes)
	Format	[1..1]	TextType	Information representation format
	MessageDigest	[1..1]	MessageDigestBinaryObjectType	Digest of the Data Object
	SignatureStatus	[0..1]	TextType	Code indicating the status of the signature (presence or absence of electronic signature, verified electronic signature, etc.)
	Size	[1..1]	SizeInBytesType	Size in bytes
BinaryObjectType	filename, uri		Base64Type	Digital data content (a sequence of bytes)
BusinessAuthorizationRequestMessageType			BusinessRequestMessageType	Authorization request message
	AuthorizationRequestContent	[1..1]	AuthorizationRequestContentType	Authorization request content
BusinessAuthorizationRequestReplyMessageType			BusinessReplyMessageType	Authorization request reply message

Table 2 *(continued)*

Type	Element	Cardinalities	Type	Definitions and comments
BusinessMessageType			MessageType	Business message
	Agreement	[0..1]	IdentifierType	Submission: agreement, or some other contract or regulation providing a framework for the relationships between the actors of the exchange
	CodeListVersions	[1..1]	CodeListVersionsType	Versions of the code lists used
	DataObjectPackage	[0..1]	DataObjectPackageType	Data Object package
BusinessNotificationMessageType			BusinessMessageType	Notification message
BusinessReplyMessageType			BusinessMessageType	Reply message
	ReplyCode	[0..1]	TextType	Reply code Error warning shall be recorded in the Comment element, if the request is rejected.
	MessageRequestIdentifier	[1..1]	IdentifierType	Request message identifier
BusinessRequestMessageType			BusinessMessageType	Request message
CodeListVersionsType	id			Versions of the code lists used
	AuthorizationReasonCodeListVersion	[0..1]	CodeType	Authorization reason code list version
	FileEncodingCodeListVersion	[0..1]	CodeType	File encoding code list version
	FileFormatCodeListVersion	[0..1]	CodeType	File format code list version
	MessageDigestAlgorithmCodeListVersion	[0..1]	CodeType	Digest algorithm code list version
	RelationshipCodeListVersion	[0..1]	CodeType	Version of the list of codes of relationships between Data Objects
	ReplyCodeListVersion	[0..1]	CodeType	Reply code list version
	SignatureStatusCodeListVersion	[0..1]	CodeType	Version of the code lists indicating the signature status for a Data Object (presence or absence of electronic signature, verified electronic signature, etc.)
CodeType	listID, listAgencyID, listAgencyName, listName, listVersionID, name, languageID, listURI, listSchemeURI		TextType	Code