

Preservation Events Controlled Vocabulary

Release date: 22 June 2017

Background

The preservation event controlled vocabulary was last modified in August 2012 and is part of the preservation vocabularies maintained at the Library of Congress Linked Data Service [<http://id.loc.gov/vocabulary/preservation/eventType.html>]. It contains sixteen event types.

A review of the controlled vocabulary for preservation events has been undertaken, with requests for new events and suggestions for edits and changes to existing events. We have undertaken public consultation and have generated this new version of the vocabulary.

Scope Statement

The scope for the controlled vocabulary has been adapted from the scope of the PREMIS Data Dictionary¹:

Preservation events record activities that affect the long term preservation of digital objects.

Principles

During the review, the team kept to a few principles:

1. Existing events should not be changed in either terminology or semantics.
2. If there was no clear use case for a suggested event, it was not included.
3. The controlled vocabulary will not cover all functions of a repository. Institutions are encouraged to enrich the vocabulary locally with their own specific events.

Note on Usage

Every event should be assigned a timestamp which records the end of all event actions. A timestamp may be recorded for the beginning of the event.

¹ PREMIS v3, p.7.

Term	Alternative label	Existing link	Definition/description	Detail
accession			The process of adding an object to the inventory of a repository. This provides a clear delineation point for the assumption of responsibility for the digital content's preservation.	See also the antonym "deaccession".
appraisal			The process of evaluating whether an object should be included or retained in the repository.	Appraisal can include the process of determining the length of time digital material should be retained. This can be based on legal requirements, appraisal policy and current and potential usefulness.
capture		http://id.loc.gov/vocabulary/preservation/eventType/cap	The process whereby a repository actively obtains an object through means other than a transfer from the creator/donor.	One use case for this event is the capture of websites during crawling functions of a repository.
compression		http://id.loc.gov/vocabulary/preservation/eventType/com	The process of encoding data to save storage space or transmission time.	Within a PREMIS context, information about the lossiness of the compression should be in the eventDetailInformation or eventOutcomeInformation. See also the antonym "decompression".
creation		http://id.loc.gov/vocabulary/preservation/eventType/cre	The process of creating a new object.	"Normalization" and "migration" are narrower terms. However, this is not limited to an object created through a derivation event. It can include the results of digitisation.

Term	Alternative label	Existing link	Definition/description	Detail
deaccession		http://id.loc.gov/vocabulary/preservation/eventType/dea	The process of formally removing an object from the inventory of a repository. This may be by transfer to another repository, return to the depositor or by permanent deletion.	Note the relationship of deletion to this event. Deletion is a technical undertaking that may be a response to deaccession. An outcome of deaccessioning may be the "deletion" of objects. See also the antonym "accession".
decompression		http://id.loc.gov/vocabulary/preservation/eventType/dec	The process of reversing the effects of compression.	See also the antonym "compression".
decryption		http://id.loc.gov/vocabulary/preservation/eventType/der	The process of converting encrypted data to data.	See also the antonym "encryption".
deletion		http://id.loc.gov/vocabulary/preservation/eventType/del	The process of permanently and wilfully removing an object in a repository.	Some objects may be deaccessioned as a precursor to "deletion".
digital signature generation			The process of generating a digital signature from a message using a private key.	Digital signatures are verified through the "digital signature validation" event.
digital signature validation		http://id.loc.gov/vocabulary/preservation/eventType/dig	The process of determining that a decrypted digital signature matches an expected value.	
dissemination			The process of transmitting or providing access to a copy of the object.	It is assumed that a SIP is created in this process. Note the relationship with "information package creation".
encryption			The process of converting data into encrypted data.	See also the antonym "decryption".

Term	Alternative label	Existing link	Definition/description	Detail
filename change			The process of modifying a filename.	Either a removal of prohibited character or a partial or entire replacement of the original filename. This can be used to note changes such as removing characters, or where a system removes the filename entirely and replaces it with a system generated name.
fixity check		http://id.loc.gov/vocabulary/preservation/eventType/fix	The process of verifying that an object has not been changed in a given period.	This will most likely utilise the results of the "message digest calculation" event.
forensic feature analysis			The process of forensically analysing raw bitstreams.	Forensic analysis is the process of determining structure, characteristics, or format of an unknown digital object
format identification			The process of determining the object's file format and version.	This event is different from "validation" which compares the object to known format specifications. Format identification could of course have the result of <i>unidentified</i> which could in turn trigger further investigation/processes.
imaging	disk image creation		The process of extracting a disk image from a physical information carrier.	

Term	Alternative label	Existing link	Definition/description	Detail
information package creation			The process of generating an Information Package (SIP, AIP or DIP).	This event can describe the creation of an SIP, AIP or DIP. The eventOutcomeDetailNote can be used to determine what package type has been created. Note that other event types can result in information package creation (such as “ingestion” and “dissemination”). It is left to the discretion of the user to decide what event(s) are recorded in line with their own requirements and standards.
information package merging			The process of merging two or more Information Packages (SIP, AIP or DIP) into one Information Package of the same type.	This event covers functions that merge the contents of multiple information packages of one type into one information package: for example, multiple SIPs into one SIP. It does not cover moving multiple information packages across types (i.e. multiple SIPs into one AIP). Such merging may for example be undertaken as a result of a repository wishing to reshape the structure of SIPs to better align with its object model.
information package splitting			The process of dividing an Information Package (SIP, AIP or DIP) into two or more Information Packages of the same type.	This event covers functions that divide the contents of one information package into more than one information package of the same type: for example, one AIP into multiple AIPs. It does not cover dividing an information package across types (e.g. one SIP into multiple AIPs). Such splitting may as an example be undertaken as a result of a repository wishing to divide a single large AIP into smaller AIPs to accommodate storage systems with size limitations.

Term	Alternative label	Existing link	Definition/description	Detail
ingestion	ingest	http://id.loc.gov/vocabulary/preservation/eventType/ing	The process of adding objects to a preservation repository. More detail can be gained by utilising "Ingestion Start" and "Ingestion End" rather than this one event.	It is assumed that a DIP is created in this process. Note the relationship with "information package creation".
ingestion end	ingest end		The process of completion of the total ingest process.	A subset of "ingestion". Should be used if "ingestion start" is recorded.
ingestion start	ingest start		The process of starting the ingest process.	A subset of "ingestion".
message digest calculation		http://id.loc.gov/vocabulary/preservation/eventType/mes	The process by which a message digest ("hash") is created.	Message digest is also commonly referred to as a "checksum". The event "fixity check" checks the message digest.
metadata extraction			The process of extracting metadata from an object. This includes technical, administrative and descriptive metadata.	
metadata modification			The process of making changes to the metadata of an object.	
migration	transformation	http://id.loc.gov/vocabulary/preservation/eventType/mig	The act of transforming an object from one file format(s) into another file format(s).	Migration events should always create a new object. This term is narrower than the OAIS definition.
modification			The act of changing a file or bitstream after receipt of the object.	This event does not create a new object. The modification takes place on the object.

Term	Alternative label	Existing link	Definition/description	Detail
normalization		http://id.loc.gov/vocabulary/preservation/eventType/bnor	An act of transforming of an object into an institutionally supported preservation format.	
packing			The process of putting objects into container files (e.g. .zip, .tar).	Compression may be used in this packing. If so, an additional event of "compression" should be recorded.
policy assignment			The process of assigning a policy to an object.	Policies can be related to rights, preservation, access, etc.
quarantine			The process of segregating objects for designated period of time.	Usually undertaken as part of virus checking to allow time for virus definition updates, or as a response to finding malware.
recovery			The act of regaining one or more files after a disaster. Usually occurs as part of a disaster recovery process.	
redaction			The process of modifying the content of a digital object to remove or mask information considered to be sensitive in nature (that is, the information cannot be viewed by non-authorized users of the repository).	Redaction will usually take place on a copy of the object.
refreshment	storage migration		The process of moving objects to a new storage medium.	
replication		http://id.loc.gov/vocabulary/preservation/eventType/rep	The process of creating a copy of an object that is, bit-wise, identical to the original.	
transfer			The process of transmitting metadata and/or digital object(s) to a repository.	

Term	Alternative label	Existing link	Definition/description	Detail
unpacking			The process of extracting objects from packages (e.g. .zip, .tar).	The archive file may be compressed. If so, an additional event of "decompression" should be recorded.
unquarantine			The process of releasing a file from quarantine.	
validation		http://id.loc.gov/vocabulary/preservation/eventType/validation	The process of comparing an object with a standard and noting compliance or exceptions.	The object being validated may be a file or an information package.
virus check	malware check	http://id.loc.gov/vocabulary/preservation/eventType/virus-check	The process of scanning a file for malicious programs.	