Life and Death of an Information Package: Implementing the Lifecycle in a Multi-Purpose Preservation System

Bertrand Caron
Department of Metadata
Bibliothèque nationale de France
Paris, France
bertrand.caron@bnf.fr

Jordan de La Houssaye
Department of Information Technology
Bibliothèque nationale de France
Paris, France
jordan.de-la-houssaye@bnf.fr

Thomas Ledoux
Department of Information Technology
Bibliothèque nationale de France
Paris, France
thomas.ledoux@bnf.fr

Stéphane Reecht
Department of Conservation and Preservation
Bibliothèque nationale de France
Paris, France
stephane.reecht@bnf.fr

ABSTRACT

This paper aims to explain how the National Library of France (BnF) faced the question of the information lifecycle, during the implementation of its digital preservation system, in the light of the experience it acquired managing and using this system, from a theoretical approach following the OAIS reference model to an implementation. It was understood very early that from a preservation point of view the management of subsequent Versions and Editions of an Information Package was a particularly sensitive problem akin to risks management. However it was thought at the time that the system should be able to manage the lifecycle of an AIP (Archival Information Package) in a universal way that would be valid whatever the package considered and the context. These ideas quickly proved to be impractical and didn't resist the test of reality. Some improvements were provided, while taking into account the rest of the digital ecosystem of the BnF.

KEYWORDS

Lessons learned, package lifecycle, information lifecycle, package update, package deletion, preservation policies, data enrichment, package Version, package Edition, OAIS interpretation, PREMIS Events

1 INTRODUCTION

Among the challenges an institution has to face when it puts into practice digital preservation, the management of updates of data isn't the easiest to face. It is necessary to conciliate precautions to avoid the loss of information, requirements for enabling enrichment of data, and financial sustainability of storage. The lifecycle of information has to be considered through the prism of risk management, therefore no turnkey solution is available for those who have to deal with this issue.

The BnF (National Library of France) is assigned with the mission of preserving over the long-term a part of the French cultural heritage, including in its digital forms. It has been building since 2007 its own digital preservation system. The issue of the packages lifecycle had to be taken into consideration.

In this paper we reflect on the history of digital preservation at the BnF in order to account for successive refinements that were made to our original assumptions regarding the lifecycle. In the first part, we detail the way the OAIS reference model was interpreted in order to make implementation choices in the SPAR system. In the second part we detail chronologically how the system originally managed any updates of AIPs, and how new use cases the BnF faced forced us to adapt our preservation system. In the third part we reflect on the way these improvements changed the lifecycle of AIPs in SPAR as well as the broader preservation environment at the BnF.

2 THINKING ABOUT A LIFECYCLE: WHERE, WHY, HOW?

2.1 The Context: SPAR and its Ecosystem

The BnF preservation system is named SPAR (*Scalable Preservation and Archiving Repository*). It is being built in house since 2007 and is in production since 2010. Its scope is to manage all entities that can be automated through modules corresponding to the OAIS entities. The system includes the management of the storage infrastructure.

In SPAR, the sets of documents to be ingested are processed by tracks and channels (sub-tracks), according to their nature (e.g., digitized books, audiovisual files, web archives, administrative records), their legal framework, and the way the BnF plans to apply preservation strategies. At the present time, SPAR can ingest objects through six tracks: digitized documents and associated files, audiovisual objects, web legal deposit (ARC or WARC files), negotiated legal deposit (ebooks and other born-digital documents), administrative records, acquisitions and donations (born-digital documents out of the scope of the legal deposit), and third-party archiving (various kinds of files, from partners outside the institution).

Each track has a leader (named a 'track manager'), someone in the library, generally a librarian, who is more than a representative. This person is responsible for preserving the collections, and is in charge of negotiations with the IT department. The result of the negotiations is formalized, for each channel of the track, in Service Level Agreements (SLAs) that formally rule the interaction between the Producer, the Archive and the Consumer (see Figure 1). More precisely, they define the terms of ingest, preservation and dissemination (e.g., formats accepted, maximum size of packages, availability of service, number of copies etc.). Each SLA is transcribed in XML files that configure the system.

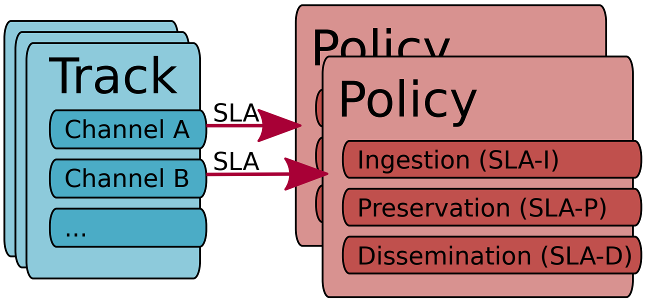


Figure 1: Relationships between tracks, channels, policies and Service Level Agreements (SLAs)

As of March, 2017, more than 7 million Archival Information Packages (AIPs) are preserved in the system, containing about 300 million files, for a volume of 3 Petabytes.

2.2 Original Principles of the Implementation of OAIS Lifecycle Concepts

The SPAR system is strongly inspired, from the beginning, by OAIS [1], and its implementation follows closely the principles of this standard[1]. Therefore, not only the responsibilities, the data model, the entities and their functions were supposed to follow the reference model (part 3 and part 4), but also the practices described in the standard (part 5, *Preservation perspectives*).

Among the practices that the system had to take into account, was the answer to the question: what to do when an update of a package is proposed for ingestion? Previously to SPAR, updates were managed on a case-by-case basis, and generally the new digitized or digital document replaced the older one, that was erased on the storage capacity or became inaccessible. Updates didn't happen very often, but we foresaw that over time it would turn into one of the usual cases: quality improvement on the Producer's initiative (especially in digitization contexts), format migration for preservation purpose... The preservation system had to implement a lifecycle in order to prevent the loss of data and to enable the management of digital document at a large scale.

For this purpose, the OAIS standard offers some valuable concepts. Section 5.1 (and more precisely 5.1.3) distinguishes four types of Digital Migration that can affect the Archival Information Package (AIP) for preservation purposes. Refreshment and Replication are operations that do not affect the Content Information, the Packaging Information or the Preservation Description Information (PDI) of the package; they are directly handled by the Storage layer [6] and are beyond the scope of this paper. Repackaging implies changes to the Packaging Information, and Transformation implies changes to the Content Information and/or to the PDI; both cases enter in the field of the lifecycle of the Information Package[2]. The standard then specifies what the Archive is supposed to do (5.1.3.5): only in case of a Transformation, the Archive has to create a new AIP Version, i.e. an AIP that "is a candidate to replace the source AIP" (1.7.2). Concerning what to do with the replaced AIP, there is no constraining rule: "The first version of the AIP is referred to as the original AIP and may be retained for verification of information preservation." (5.1.3.4).

In case of an upgrade or an improvement of the AIP, at the Producer's initiative, the OAIS standard specifies that "This is not a Digital Migration in that the intent is not to preserve information, but to increase or improve it." Thus, a new AIP Edition has to be created, but once more the standard leaves the choice to the Archive to retain or not the previous AIP (5.1.3.5). At the BnF, it has been understood that these choices were a matter of risk management regarding data loss.

Because implementation choices also had to be done, the interpretation for an AIP Edition was: increasing or improving information means adding Data Objects (for example: supplement image mode digitization with OCR files) or adding/modifying metadata (such as Descriptive Information, PDI…). In every case, there is no necessity to preserve the previous AIP, because there are no risks of losing data, and because the SPAR system includes a lot of functionalities to avoid the ingestion of packages with metadata too poor for preservation. The understanding for an AIP Version was: modifying Data Objects can be thought of as a Transformation, and therefore leads to a new Version of the AIP. In this case, because Data Objects were changed, it has been considered by default that deleting the previous AIP Version was a risk: as the OAIS says that "the new AIP is viewed as a replacement for the source AIP" (5.1.3.5), it would be too big a responsibility for the Archive to remove the source AIP. Therefore a general preservation rule was defined: the first, the latest and the penultimate Version of a package has to be preserved ("0, N-1 & N Rule", see Figure 2), while a new Edition always removes the latest one.

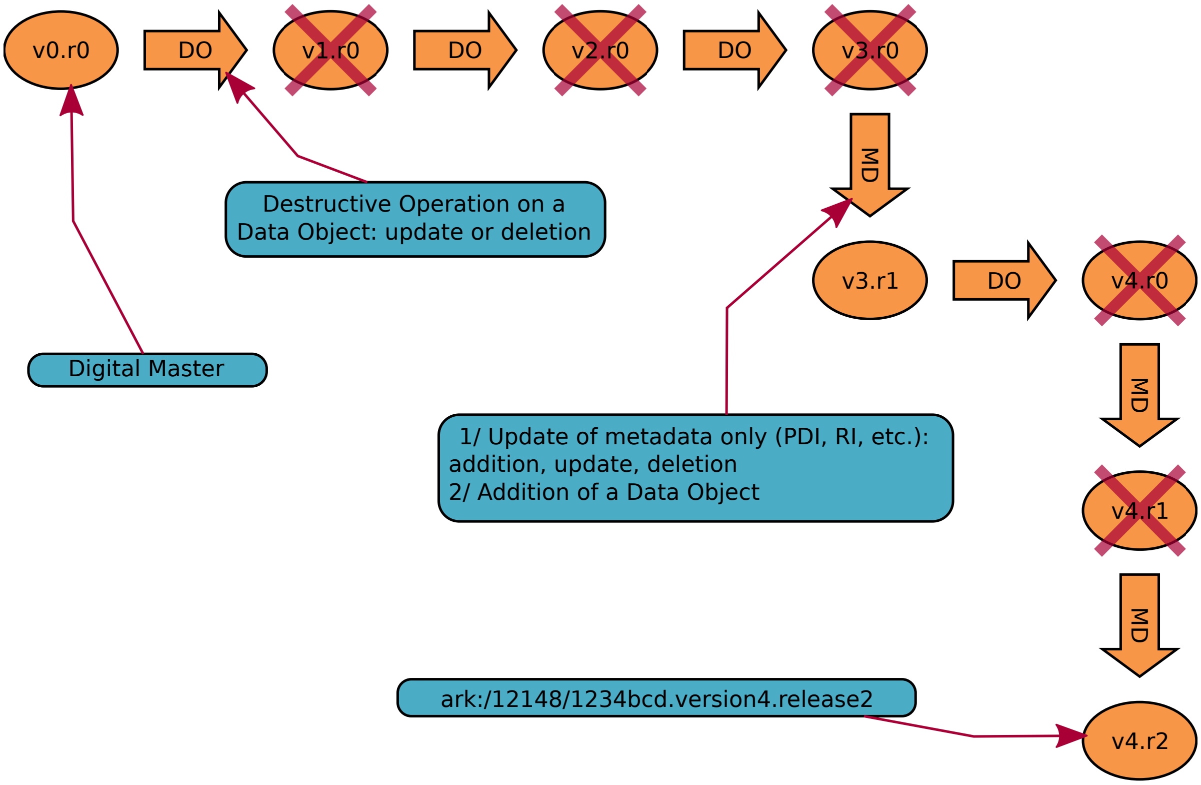
As OAIS requires (sub-section 4.2.1.4.2), this part of the history of the package is recorded as Provenance Information. Moreover, every instance of the AIP is addressable at the level of the Reference Information, through its identifier: the ARK [5] identifier is systematically suffixed with two qualifiers, one to mark the Version and the other to mark the Edition. Initially, it appears in the form of `version0.release0`[3]. When an update occurs, the number of version or release, following the case, is increased for the new package; for example, if the first update produces a new version, the qualifiers for the new package are `version1.release0`. From a preservation point of view, this enables to apply automatically the "0, N-1 & N rule". From an Access point of view, this enables to establish a simple and global policy: if the Consumer asks for an AIP in absolute (i.e. only the ARK identifier without qualifiers), the system retrieves the last Version and Edition; if he asks for a particular Version or Edition (i.e. specifying qualifiers), the system retrieves the requested Version or Edition if it still exists.

In order to document and preserve PDI, Representation Information and Packaging Information, the BnF initially chose the METS format. A single METS file, called "manifest", is used as a wrapper for descriptive, technical and provenance metadata. For provenance

[1] Unless otherwise stated, all the references in this section are from the OAIS.

[2] Repackaging has not been on the agenda for the BnF yet, given the youth of the system.

[3] The string "release" has been preferred to "edition", because the word "edition" has another meaning in the context of a library and was deemed too ambiguous.



v0.r0
DO
v1.r0
DO
v2.r0
DO
v3.r0
MD
v3.r1
DO
v4.r0
MD
v4.r1
MD
v4.r2
Destructive Operation on a Data Object: update or deletion
Digital Master
1/ Update of metadata only (PDI, RI, etc.): addition, update, deletion
2/ Addition of a Data Object
ark:/12148/1234bcd.version4.release2

Figure 2: Original "0, N-1 & N" rule

metadata, PREMIS was chosen to document operations that affected the Content Information before and after ingestion. Among these operations are particular events occurring in the package lifecycle: SIP creation, ingestion and package update. When a package update is processed, the whole set of events which affects the Information Content from its creation is preserved as the package audit trail.

This general policy, already defined in 2008 [3], can be viewed as a simplification regarding what the reference model offers, but it presents the advantages of enabling an automated implementation.

3 THE LIFECYCLE ACTUAL IMPLEMENTATION AND ITS EVOLUTION

3.1 Initial Implementation

During the initial development of the SPAR system, a decision tree was implemented during the ingestion phase, in order to assess automatically if a package proposed for an update had to lead to a new Version, a new Edition or a failure. A quite complex sequence diagram was specified to describe SPAR's decision making, which occurs in an internal process of the ingestion phase named "ACT_10" (see Figure 3). This sequence combines the SIP manifest analysis and comparisons between the AIP and the SIP manifest. Some steps were a transcription of a simple rule (for example, if there is no `fileSec` in the SIP METS manifest, then it is only a metadata modification, and the system has to create a new Edition); some implied complex comparisons, should the SIP contain Data Objects.

Regarding the "0, N-1 & N" rule, it has not been implemented, so that every Version is preserved in the system. As for the Edition, the system is able to really remove the previous one.

As the system is in production since 2010, with the first track "Digitized documents for preservation", the BnF faced several use cases. That led to reconsidering the whole package lifecycle, not only in the SPAR system but also in the broader environment of digital management in the library.

3.2 A First Exception to the Automated Lifecycle: Requesting Explicitly for a New Version

In the context of mass-digitization, it often happens that a defect is detected after ingestion, or that digitized documents are supplemented with OCR, which causes a reprocessing of the package. It appeared that the automated mechanism was not sufficient to address with absolute confidence the different use cases, because of their complexity and variety. One use case is when, at the same time, the update consists in modifying the existing Data Objects and adding new ones, for example when a book digitized in black and white is newly digitized in color and OCRed. The system was then supposed to create at the same time a new Edition and a new Version, which means that the automated implementation of the lifecycle was not complete. At this point, when this case occurred, a risk existed that the system would create a new Edition and delete

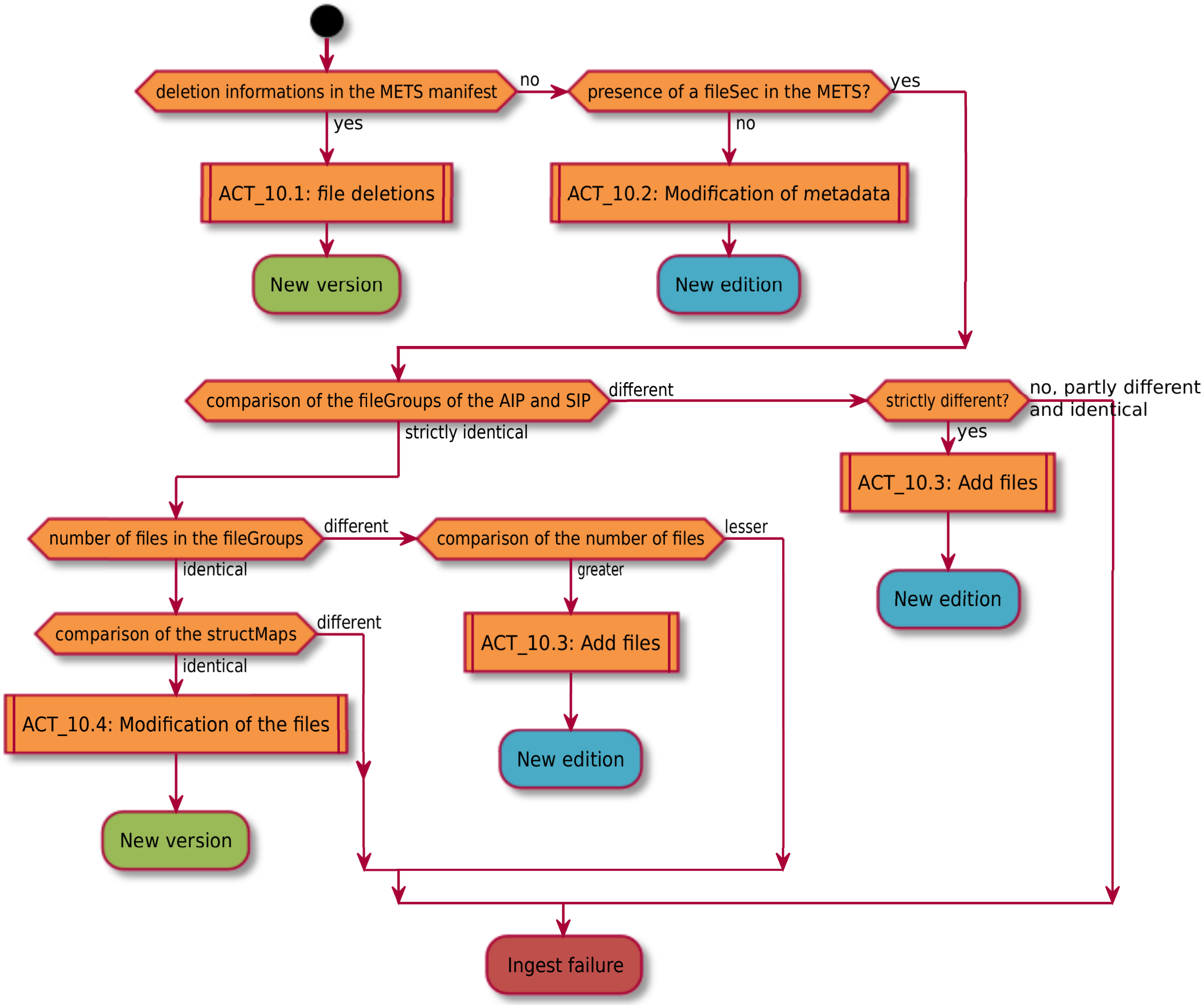

deletion informations in the METS manifest
no
yes
presence of a fileSec in the METS?
yes
no
ACT_10.1: file deletions
ACT_10.2: Modification of metadata
New version
New edition
comparison of the fileGroups of the AIP and SIP
different
strictly identical
strictly different?
no, partly different and identical
yes
ACT_10.3: Add files
New edition
number of files in the fileGroups
different
identical
comparison of the number of files
lesser
greater
comparison of the structMaps
different
identical
ACT_10.3: Add files
ACT_10.4: Modification of the files
New edition
New version
Ingest failure

Figure 3: Simplified view of the decision tree to process a SIP for update (2008 version)

the previous Data Objects, while it was not certain that this was what the Producer wanted.

It was therefore decided that the decision to create or not a new version had to be taken outside of the preservation system, earlier in the workflow. A specific request functionality for a new Version was implemented, coming explicitly from the outer "delivery chain for digitization", that knows the context of the delivery or the update better than the preservation system. This request is submitted to SPAR by means of a PREMIS `updateRequest` Event, which keeps track of the request in the manifest after the update is processed.

It is interesting to note that in the context of mass digitization at the BnF, every subsequent digitization leads to a new Version of the target AIP. From an OAIS perspective, this suggests that a digitization is a kind of migration, and indeed, it can be argued that digitizing a book for example is a media migration from physical to digital. Consequently, each subsequent digitization can be thought as a new migration, albeit from the same original source, and therefore has to lead to a new Version of the corresponding AIP.

3.3 Requesting a Deletion of the AIP

The second evolution came directly from the integration of Administrative records. Indeed, handling archival processes forces taking into account the elimination of the content of a package. Regulation rules or laws make such destruction of information mandatory. However, since SPAR gives a permanent identifier to each package, it couldn't just erase the record: the elimination is thought as an update. In SPAR, updating a package requires a re-ingestion, in order to apply the policy defined in the SLAs (as the OAIS specifies that "AIPs should never be deleted unless allowed as part of an approved policy; there should be no ad-hoc deletions"[4]). Here, this update will replace the actual content with a certificate of elimination, called "tombstone". This file contains the reason for the elimination as well as the requester and the authorizer. Doing so enables an appropriate response in case someone wants to retrieve this package.

Similarly to an update request, this request is submitted to SPAR by means of a PREMIS `deletionRequest` Event. In fact, the act of elimination is viewed as an update of the package with a new content (the 'tombstone') that will replace the latest one while erasing all the previous Versions of the package. This is the first time the system accepts that the original Version (Version 0) can be deleted. Of course, in order to keep such an operation safe, the permission of doing it must be given in the ingest SLA as well as who can initiate such a task and in which conditions.

At the end of the operations, a unique version is kept, the latest, containing the tombstone but also, in its manifest, all the audit trail of the operations applied to the package. No one can retrieve any previous content but anyone can be informed as to why it has been discarded.

It is worth noting that the result of a `deletionRequest` is a new Version of the AIP, whose content is the so-called "tombstone". It could be argued that it should be a new Edition, and not a new Version, because its Content Information has been upgraded (albeit not to improve it but to destroy it). However it has been understood as a Transformation (and therefore a Migration) because the Content Information has been altered and the resulting AIP is meant to replace the source AIP.

3.4 Requesting a Deletion with Redirection of the AIP

The previous evolution was fine from an archival perspective but, seen from a preservation one, losing the information content is not acceptable. In the context of the digitization programs, the same analog item might have been digitized by different means (old black and white digitization vs. new high resolution color one) and with different identifiers, generating sorts of duplicates. When keeping both is seen as unneeded, choosing the "best" one cannot be an automated process but is directly related to the management of the collection: the weeding procedure. In order to make it possible, the system was enhanced by adding to the `deletionRequest` the obligation of specifying another package. During the process of selection, the collection manager generates a tombstone not only stating the reasons for his decision but also providing the alternate package holding the same informational content. During the ingestion of this tombstone, the existence of the substitution package is verified. Then, in case someone wants to access the discarded package, the system will automatically redirect its request to the associated one. Once again the linear lifecycle is here clearly eluded but following a clear and stated decision.

This functionality came from the needs of digitization, but turned out to be useful for other tracks, such as Web legal deposit, to also address the case of a Producer's mistake that causes a duplicate and that is detected after the ingestion of both packages. If the duplication is intentional, then the Producer can use the same identifier for both packages in order to create a new Edition of the package. In fact, this case matches exactly what OAIS calls "improvement" of information.

3.5 Requesting a New Edition Explicitly

Following the `updateRequest`, a new functionality was then implemented to address the case of data enrichment, for example when OCR files are added to still images. Here, new data and metadata are delivered simultaneously, and the automatic detection of a new Edition turns out to be very difficult. The explicit decision to create a new Edition (in PREMIS terms: `replacementRequest`) is now possible on the track manager's initiative. For consistency reasons, this functionality was implemented on the model of the `updateRequest` (see 3.2).

3.6 Requesting a Channel Switch

Given that every channel is ruled by SLAs defining ingestion, preservation and dissemination policies (see 2.1), the need for changing those policies for a package implies changing its channel. A `channelSwitchRequest` has been developed, resulting in a new Edition of the package in the target channel. This functionality can be viewed as a mix of a `replacementRequest` (as it is an explicit request for a new Edition) and a `deletionRequest` (as it also requires authorization, explanation and documentation). The Channel Switch is far from being a simple operation performed by an administrator; on the contrary, it means a new ingestion into the system, so it has to be allowed at the same time by the SLAs of the source channel and of the target channel. Switching a package from one channel to another only means moving it and not modifying it, consequently this operation must not imply a doubling of the storage size. That's why it results in a new Edition of the package.

While working on this matter, it was deemed beneficial to expand the possibility to document such an operation. When the switch is defined, it is now mandatory to indicate either the reference of a BnF internal document or the identifier of a preservation plan also preserved in the system [4], that explains the decision of switching the package or a set of packages from one channel to another. In addition, in an iterative approach, this constraint has been extended to the `deletionRequest` functionality.

Almost ten years after its first version, the sequence diagram of the "ACT_10" process has been substantially enriched (see Figure 4). Now, the first steps involve detecting one of the four PREMIS Events that explicitly lead to a new Version or a new Edition of the package. Only after these steps begins the analysis from 2007, that has been simplified because some of the steps were no longer useful after implementation of the new functionalities. Thus, the risk that the system creates an unwanted Edition, is significantly reduced.

4 ENRICHMENT AND ENHANCEMENT OF DIGITAL CONTENT: LIFE BEYOND INGESTION

After having produced digital contents for more than twenty years, the BnF had to take into account the fact that its digitization policy is no longer limited to the creation of new digital copies of analog

[4] see [1, section 3.2.5, p. 3-5]

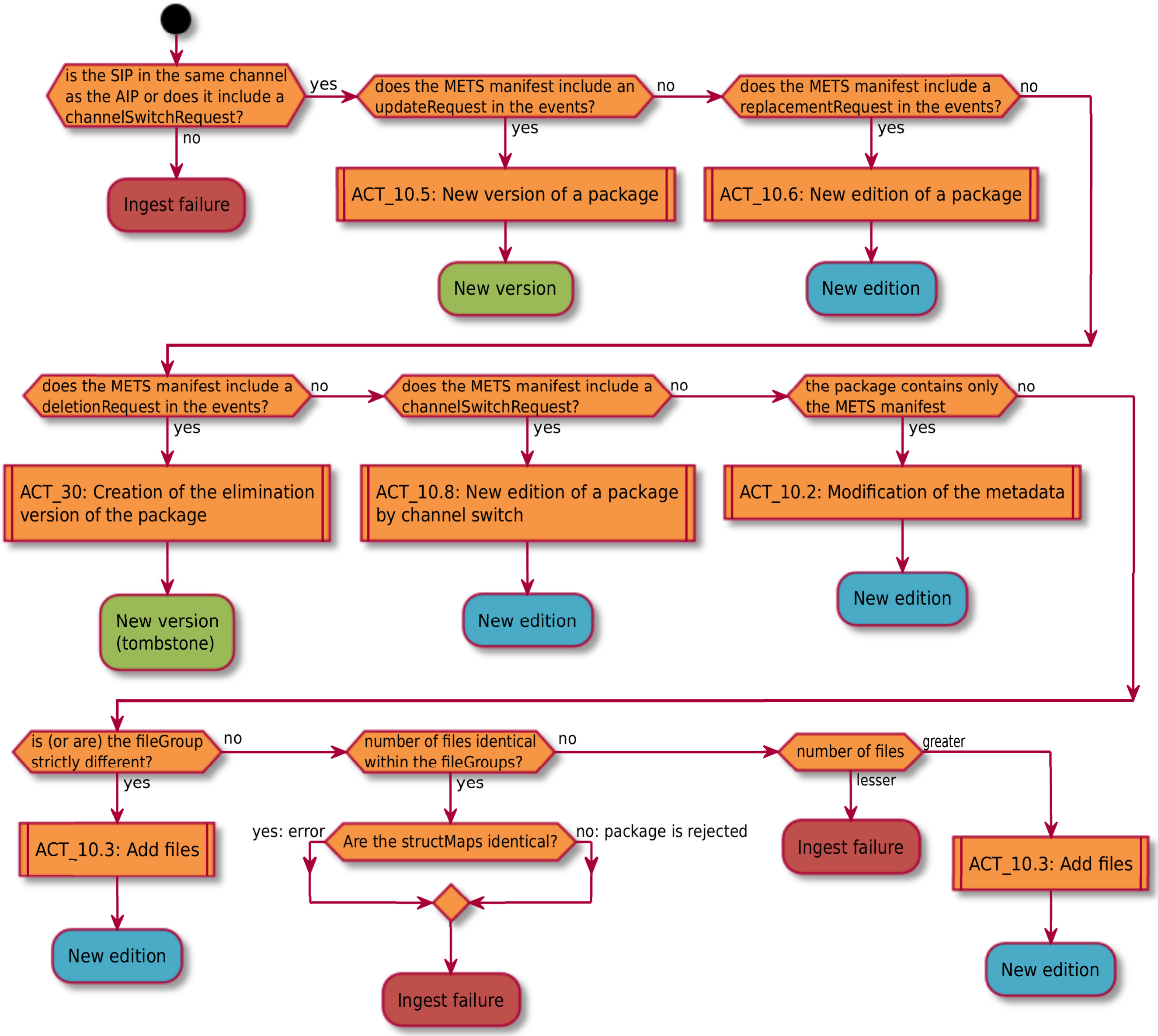

is the SIP in the same channel as the AIP or does it include a channelSwitchRequest?
yes
no
Ingest failure
does the METS manifest include an updateRequest in the events?
no
yes
ACT_10.5: New version of a package
New version
does the METS manifest include a replacementRequest in the events?
no
yes
ACT_10.6: New edition of a package
New edition
does the METS manifest include a deletionRequest in the events?
no
yes
ACT_30: Creation of the elimination version of the package
New version (tombstone)
does the METS manifest include a channelSwitchRequest?
no
yes
ACT_10.8: New edition of a package by channel switch
New edition
the package contains only the METS manifest
no
yes
ACT_10.2: Modification of the metadata
New edition
is (or are) the fileGroup strictly different?
no
yes
ACT_10.3: Add files
New edition
number of files identical within the fileGroups?
no
yes
Are the structMaps identical?
yes: error
no: package is rejected
Ingest failure
number of files
greater
lesser
Ingest failure
ACT_10.3: Add files
New edition

Figure 4: Simplified view of the decision tree to process a SIP for update (2017 version)

material, but also aims at enriching or replacing older digitization and adding new derived products. As a consequence, the package lifecycle, at first designed to be linear, is becoming circular (see Figure 5): Information Packages subjected to enrichment or enhancement have to be disseminated, reprocessed or enriched, and then ingested again. Services and systems which until then had no conscience of their role in digital information preservation (e.g., QA delivery chains, dissemination services, etc.) had to endorse responsibility for actions that would affect the Content Information and its quality. These services were affected by this new principle and had to evolve accordingly in order to ensure that no loss could affect the packages fixity and quality.

4.1 Dissemination

The act of disseminating an IP with the intention of enhancing some of its components or enriching it with new representations consequently appeared to be a critical phase in the package lifecycle. Particularly, the context (date, reason, agents involved) had to be taken into account to determine the systems behavior when the SIP expected to update the corresponding AIP will be submitted. A new PREMIS Event, named `disseminationCompletion`, has been

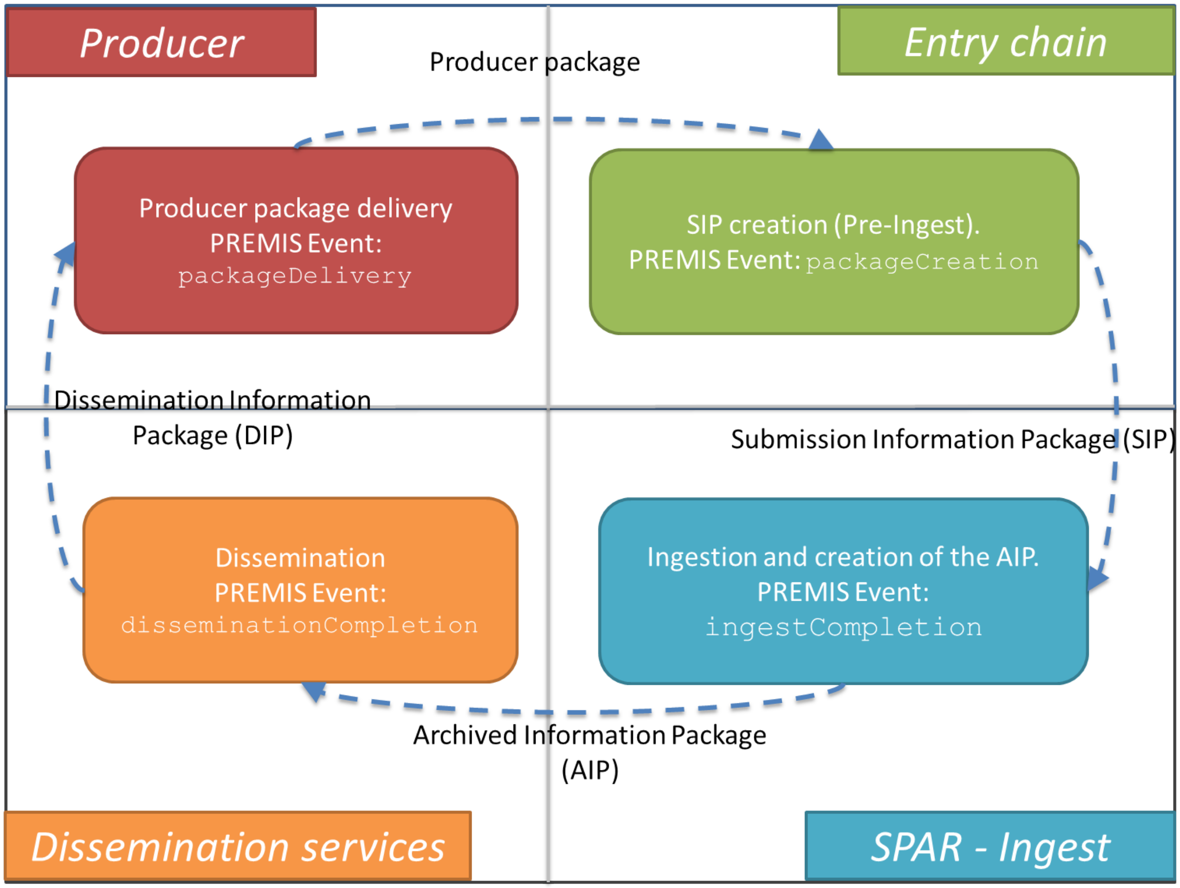

Producer
Entry chain
Producer package
Producer package delivery
PREMIS Event:
packageDelivery
SIP creation (Pre-Ingest).
PREMIS Event: packageCreation
Dissemination Information
Package (DIP)
Submission Information Package (SIP)
Dissemination
PREMIS Event:
disseminationCompletion
Ingestion and creation of the AIP.
PREMIS Event:
ingestCompletion
Archived Information Package
(AIP)
Dissemination services
SPAR - Ingest

Figure 5: Circular lifecycle management

chosen to record the dissemination operation, which created a DIP delivered to a Consumer[5]. The responsibility for disseminating the package and for recording this PREMIS Event had to be endorsed by services aware of the operation context and goals, that is, dissemination services directly in contact with Consumers.

4.2 Package Delivery

As the preservation system is unaware of the production context, specified by digitization contracts for example, the creation of a SIP for update had to be made before ingestion by services that have the understanding of the reason why an AIP had to be updated. The `packageDelivery` PREMIS Event that records the transfer of a producer package from a Producer to the BnF mentions the intended use of this producer package (replacement or addition of Data Objects). Only QA delivery chains, combining the intended use and the policy defined by track managers, can determine how to create a SIP and which kind of request (for a new Version, cf. 3.2 or a new Edition, cf. 3.5) should be submitted to SPAR. Hence, merging the Producer package and the corresponding AIP to produce a new SIP for update is performed by the delivery chains, which, unlike the preservation system, have an understanding of the update expected results. At the ingestion phase, SPAR just creates a new Version or Edition by taking all Data Objects from the SIP, though preserving the whole audit trail coming from the AIP.

4.3 Policy about Versions Retainment

As previously mentioned, no policy was globally defined concerning the choice whether to retain or not previous package Versions. Preservation experts considered that this choice had to be made by people particularly aware of digital collections management. In the case of heritage digitization for example, the following policy was adopted:

(1) whenever a reprocessing due to quality failure in the course of an ongoing contract generated a new Producer package, delivery chains will request the preservation system for a new Version;
(2) whereas, in the case of a new Producer package intended for enhancement of the existing AIP, delivery chains will request a new Edition.

At the end of the digitization contract, a general decision has to be made to define which policy should be adopted for previous

[5] The Consumer of the DIP happens, in this case, to be the Producer of the SIP for update.

Versions, taking into account the causes that led to requesting a new Producer package.

4.4 Managing the Risk of Quality Loss

Though creating AIPs with several SIPs delivered progressively is a common problem that Archives have to face for several years[6], the iterative process of Data Objects production, based on previously produced Data Objects (OCR files produced from image files, or accessible Daisy files produced from OCR files) became extremely tricky when both production processes had to be managed simultaneously.

Indeed, the risk of quality loss arose because Producer packages intended for updates could be delivered concurrently from different Producers, either to enrich or to enhance the original AIP. The decision was taken to reject a SIP for update intended for enhancement if the corresponding AIP has been disseminated for enrichment.

Moreover, in order to warn contract managers if some case of quality loss should arise, the delivery chains have to compare the last AIP Version or Edition and the Version or Edition of the DIP used to generate the Producer package. Whenever the last AIP Version/Edition is more recent than that of the DIP, QA would raise a warning flag and alert contract managers and track managers.

5 CONCLUSION

In the course of our daily operations, two trends emerged:

First, we tend to explicit more and more our intentions in the SIP: if a new Edition or a new Version is required, we state as much explicitly. It is the business of the track manager to decide if such or such an action to the data should lead to a new Version, a new Edition or an entirely new Package. This trend reached the point where we started to invoke directly the business intent into SPAR.

Second, we now acknowledge that the question of the management of subsequent Versions should be treated track by track, as part of the broader question of preservation policy. This question is once again in the hands of the track manager.

The enrichment of the AIP lifecycle leads to putting our track managers more and more at the center of our daily operations. We foresee the time when this question of the AIP lifecycle will be treated explicitly in the SLAs of our tracks, alongside the treatments associated with business related actions on the packages.

From an OAIS point of view, this trend is perfectly logical. We first needed to build a working system, and therefore we took some shortcuts in our OAIS implementation. 10 years in the making, both the system and the business matured. SPAR is now clearly understood as the technical part of our OAIS, completed by the BnF as an organization dedicated to preserving part of the French cultural heritage.

ACKNOWLEDGMENTS

The authors would like to thank all the members of the SPAR team for their commitment and their invaluable contributions, John Thompson for his proofreading, and Claire Caillet from the CNES as well as Claire Sibille from the SIAF for the fruitful talks about OAIS.

[6] In particular, the CCSDS standard PAIS (Producer-Archive Interface Specification) [2] addresses this need.

REFERENCES

[1] 2012. *Reference Model for an Open Archival Information System (OAIS).* Recommended Practice. CCSDS. 135 pages. https://public.ccsds.org/Pubs/650x0m2.pdf CCSDS 650-M-2.

[2] 2014. *Producer-Archive Interface Specification (PAIS).* Recommended Standard. CCSDS. 104 pages. https://public.ccsds.org/Pubs/651x1b1.pdf CCSDS 651-B-1.

[3] Emmanuelle Bermès, Isabelle Dussert-Carbone, Thomas Ledoux, and Christian Lupovici. 2008. Digital preservation at the National Library of France: A technical and organizational overview. In *Proceedings of the 74th IFLA General Conference and Council, meeting 84, Digital preservation.* http://www.ifla.org/IV/ifla74/papers/084-Bermes_Carbone_Ledoux_Lupovici-en.pdf

[4] Bertrand Caron, Thomas Ledoux, Stéphane Reecht, and Jean-Philippe Tramoni. 2015. Experiment, Document & Decide: a Collaborative Approach to Preservation Planning at the BnF. In *iPRES 2015 12th International Conference on Digital Preservation.* Chapel Hill, NC, United States. https://hal-bnf.archives-ouvertes.fr/hal-01288699

[5] John A. Kunze and R. P. C. Rodgers. 2013. *The ARK Identifier Scheme.* Internet-Draft draft-kunze-ark-18. Internet Engineering Task Force. https://datatracker.ietf.org/doc/html/draft-kunze-ark-18 Work in Progress.

[6] Thomas Ledoux. 2012. SPAR: From Design to Operations. (January 2012). Retrieved 2017-06-15 from http://c.bnf.fr/fyL Presentation at the Preservation and Archiving Special Interest Group (PASIG) (Austin, USA).