

Le Programme interministériel d'archivage VITAM

Par **Jean-Séverin LAIR**

Chargé du programme Tech.Gouv

Direction interministérielle du numérique (DINUM)

Le lancement d'un programme sur l'archivage numérique public

L'archivage, un besoin essentiel... même dans le numérique

Pour la grande majorité des lecteurs, et encore plus sans doute pour ceux férus de numérique, à la seule évocation du mot *archivage*, arrivent instantanément à l'esprit des images de parchemins, de rayonnages poussiéreux et de gestionnaires grisonnants. Un peu de modernité, que diable !, avec le numérique on n'a plus besoin de cela, la mémoire est à l'échelle de la planète, diront certains. Et pourtant...

Je ne jette la pierre à personne car je n'étais pas loin d'entériner cette vision caricaturale des archives (et des archivistes !) avant d'avoir été « éduqué »⁽¹⁾. Revenons aux fondements des archives publiques, retranscrites par le béotien que je suis⁽²⁾ :

La raison d'être des archives publiques est de garder la trace des travaux et décisions administratives ayant un impact sur les droits et devoirs des personnes physiques et morales, pour les faire valoir durant leur durée d'effet, pouvoir les invoquer avec force probante en justice, assurer leur transparence et la capacité à rendre des comptes des décideurs et finalement, pour certaines archives, en garder la mémoire historique.

Je tiens à souligner que :

- n'en déplaise aux plus intéressés par l'histoire et contrairement à la croyance populaire, la première utilité de l'archivage n'est pas l'étude scientifique, mais bien l'activité administrative, au profit des organisations et des personnes, et ce, depuis l'aube même des archives (la conservation des titres nobiliaires et de propriétés au niveau royal...). L'archivage est un élément-clé de la continuité de l'État et de la validité des décisions prises en son nom.
- nulle part dans cette définition je n'ai évoqué de papier et il est évident que de plus en plus de travaux et décisions sont maintenant sous forme numérique. Alors qu'il y a encore dix ans on dématérialisait mais avec des étapes qui restaient encore sous forme papier, on atteint enfin maintenant le tout numérique dans les procédures administratives. Ceci veut dire qu'en l'absence d'archivage numérique, nous perdons des traces essentielles ou du moins qu'elles perdent de leur force. *Quid* du cadastre avec ses tracés et ses servitudes, des décisions de justice, des naturalisations... ? Sans archivage numérique nous risquons un « Alzheimer d'État » – sans parler de la perte historique que l'on pourra constater dans quelques décennies.

(1) « Sorti de l'état sauvage » pour les étymologistes avertis.

(2) Je prie les archivistes qui pourront lire ces lignes d'excuser cette vision simpliste, mais il s'agit de vulgariser...

L'archivage est important, mettons ! Mais dans le numérique on sait conserver les traces sur nos disques durs, et même plus, puisque dans le *cloud* on a des capacités énormes. Alors pourquoi ce besoin d'archivage ?

L'archivage consiste à garder de manière sûre et suffisamment bien ordonnée une information pour pouvoir la retrouver plusieurs dizaines d'années plus tard, le tout garanti par des processus éprouvés et des personnes mandatées pour garantir toute la valeur probante et historique nécessaire. Empiler les informations sur des disques, comme on a l'a fait depuis vingt ans, ou les entasser dans des gisements de fichiers avec un moteur de recherche ne répond que très partiellement au besoin. Le numérique standard est parfait pour le temps court et la performance, mais les archives ont besoin de temps long et de garantie. Faire de l'archivage numérique nécessite de réconcilier la nanoseconde avec le siècle, le flux et la preuve. Il s'agit de projeter une pratique professionnelle ancestrale sur d'autres médias avec d'autres outils, en maintenant l'exigence du temps long et de la preuve qui manque tant à l'informatique.

Le constat au niveau de l'Etat français en 2011 : nécessité d'agir

Le ministère de la Culture, qui est garant des aspects interministériels de l'archivage, a été le seul dès 1982 à se doter d'un procédé d'archivage du numérique, dit « Constance ». Reposant essentiellement sur des procédures manuelles et des moyens jugés maintenant artisanaux (des bandes, des bases documentaires...), il a permis d'archiver des contenus importants, comme certaines grandes enquêtes INSEE (recensement de 1962, par exemple) ou des données statistiques de différents ministères. Mais les volumes d'informations qui pouvaient être versés ont grandi de manière exponentielle. En 2011, les Archives Nationales ont constaté que les contenus prêts à être versés par les ministères étaient plus volumineux que l'ensemble des archives collectées depuis 1982. Il fallait agir !

Le cadrage du programme VITAM : un long chemin

En France, trois ministères couvrent globalement l'essentiel de l'archivage historique de l'État, à grosse maille les ministères des Armées et des Affaires étrangères pour leurs besoins propres, le ministère de la Culture pour les autres (sauf Bercy...). En 2011, ils réfléchissaient tous à l'archivage numérique et faute de moyens suffisants, la seule voie était la coopération. Fait important, ce sont non seulement les grandes entités gestionnaires d'archives des trois ministères qui ont collaboré et élaboré ce qui n'était au début qu'un projet, mais aussi leurs DSI.

Il a fallu attendre près de quatre années pour réunir les moyens nécessaires. Un acteur essentiel a permis finalement le lancement, le Directeur Interministériel des Systèmes d'Information et de Communication de l'époque, Jacques Marzin. Fort sponsor, il a su mobiliser le Programme d'Investissement d'Avenir du CGI⁽³⁾ pour compléter l'apport des ministères.

Début 2015, le programme Valeurs Immatérielles Transmises aux Archives pour Mémoire (VITAM) était né.

Le programme VITAM

L'essentiel du programme VITAM en quelques zooms

Les objectifs du programme :

La réalisation de la solution logicielle libre VITAM d'archivage numérique permettant la prise en charge, la conservation et la consultation sécurisée de très grandes volumétries d'archives numériques, pouvant même être classifiées ;

(3) Commissariat Général à l'Investissement

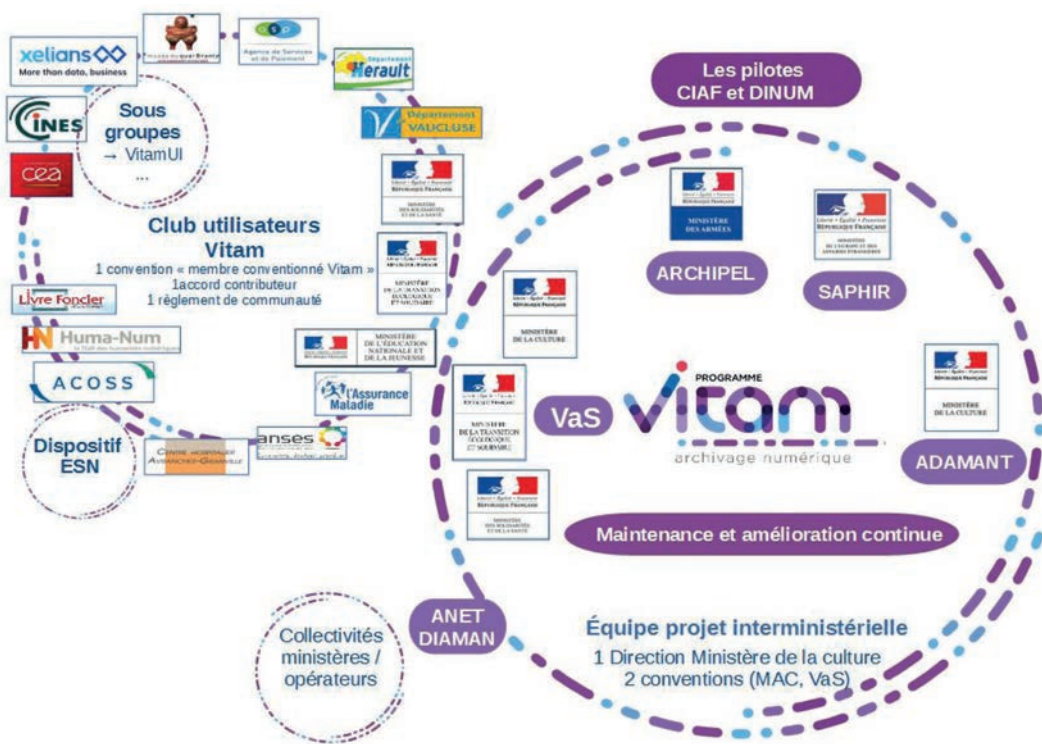
La mise en place de plateformes d'archivage utilisant la solution logicielle VITAM, dans chacun des trois ministères, *via* les projets ministériels : Saphir (MEAE), Adamant (MC) et Archipel (MinArm) ;

La diffusion et la réutilisation la plus large possible de la solution logicielle VITAM pour des usages d'archivage historique et intermédiaire.

Le « manifeste »⁽⁴⁾ de la solution logicielle :

*Pour les services de l'Administration,
Afin de satisfaire les enjeux d'accès dans le temps à leurs informations numériques,
Vitam propose une solution logicielle libre d'archivage, évolutive, simple et facilement interfaçable,
Qui permet la gestion unitaire et sécurisée de milliards d'objets
Et vise son adoption par le plus grand nombre d'acteurs publics.*

La carte des acteurs début 2020 :



Des premières implémentations mises en production :

- 2018 : Culture (Adamant)
- 2019 : Affaires Etrangères (Saphir) et Xelians-Locarchives

Le développement logiciel :

- Rythme Agile
 - Sprint 3 semaines, Release tous les 3 mois, Versions majeures tous les ans

(4) Phrase résumant l'essentiel d'un projet dans les standards de développement Agile.

- V1 début 2018, V2 début 2019, V3 début 2020
- Jusqu'à 4 équipes Scrum en parallèle, plus de 10000 commits
- Architecture :
 - Microservices full REST
 - Souches « BigData et schema-less » : MongoDB, ElasticSearch
 - Stockage « agnostique et extensible » : File system, Objet (SWIFT, S3), Bande (LTO)
 - Preuves : Chaînage de journaux, arbre de Merkle, horodatage sûr de lots d'opérations (proche des concepts de Blockchain)
 - Multi-site, multi-tenant, de 3 à 75 VM, voire plus.
 - Automatisé par Ansible
 - Langages et OS supportés : Java 8 et 11, Debian et CentOS
- Performances
 - Conçue pour des milliards d'archives, manipulables unitairement, constituant des pétaoctets.
 - Tenue en charge testée régulièrement sur plus de 400 millions d'objets
 - Exemple de test d'injection : 160 000 objets bureautiques pris en compte pour archivage (donc avec toutes les garanties de conservation légale) en 18 min.
- Communauté
 - Une souche principale et des satellites (bibliothèques, griffons, IHM) libres (CECILL) sur GitHub
 - Plus de 15 partenaires publics utilisateurs (implémentation propre ou mutualisée), dont 2 contributeurs actifs (CEA, CINES)
 - Plus de 10 ESN, et des plus grands, formés sur VITAM
 - Le leader de l'archivage privé en France, Xelians, a construit son offre numérique sur VITAM et contribue au code

Un programme volontairement mené de manière inhabituelle...

Tirant les leçons de l'échec de l'Office National de Paye et utilisant VITAM comme démonstrateur d'une « autre façon » de faire et de réussir les grands projets, la DISIC a exigé que le projet :

1. soit mené en Agile,
2. associe plusieurs prestataires pour éviter d'être sous le contrôle d'un seul dominant,
3. implique des agents publics experts fonctionnels et techniques maîtrisant les choix,
4. et que les projets d'implémentation soient intégrés au programme autour du projet de développement logiciel, chacun ayant du coup des comptes à rendre aux autres.

Outre les exigences de la DISIC, vinrent s'ajouter les exigences du CGI. Prodiguant des fonds qui ont vocation à soutenir le développement économique, le CGI a demandé une approche au moins de développement de l'usage, et si possible, d'activité économique autour de la solution.

Fort de ces deux injonctions atypiques dans les projets de l'époque, le programme VITAM est donc devenu un programme Agile, dont les choix étaient maîtrisés par l'administration et dont la priorité était les besoins des utilisateurs, garants d'une large réutilisation. Une approche fortement collaborative a été mise en place avec des archivistes de tous horizons (ministères, collectivités...) pour capter au mieux leurs réels besoins (et non envies), mais aussi avec des informaticiens pour garantir le réalisme technique des solutions proposées dans leur environnement technique.

Ainsi le programme VITAM n'était plus un projet informatique, mais une opération de construction d'un produit à l'usage d'une communauté professionnelle investie. Gagnant sa crédibilité, il a aussi su attirer les acteurs du monde économique.

Un travail collaboratif de transformation...

Si le programme a pu exister, c'est parce qu'il répondait à un besoin fort de toute une communauté professionnelle obligée d'évoluer face à la révolution numérique. L'information changeant de média, il fallait de nouveaux outils pour la manipuler et assurer sa pérennité. Mais qu'en était-il des pratiques archivistiques ?

L'essentiel des notions séculaires utilisées dans la gestion des archives papier avaient une raison d'être, bien au-delà de la forme de l'information, et avaient toujours un sens dans le numérique. Quel que soit le média, l'objectif de l'archivage reste de récupérer l'information utile à conserver, dans des conditions assurant sa validité dans le temps et la capacité à la retrouver au moment utile même des années, des dizaines d'années ou des siècles plus tard.

Toutefois, au fur et à mesure des travaux, des notions essentielles furent interrogées par les moyens numériques.

Par exemple, l'archivage du nativement numérique a ouvert la voie à un traitement à l'échelle du document, voire de l'information élémentaire. Cela était inenvisageable en papier en s'appuyant sur une gestion manuelle. Tout l'intérêt du travail conjoint archiviste/informaticien fut de trouver la bonne façon d'utiliser cette puissance du numérique. Fallait-il s'en remettre complètement à de la recherche sur du vrac comme savent le faire les grands moteurs de recherche ? En fait, une vision plus riche, mariant vision hiérarchique et métadonnées sélectionnées à tous les points de la hiérarchie jusqu'à l'information finale, fut retenue. Les moteurs de recherche donnaient ainsi toute leur puissance sur des données de qualité.

D'autres notions, comme la prise en charge, la destruction, le changement de dépositaire, la gestion du cycle de vie aussi, furent revues et, à chaque fois, il s'avérait que la pratique pouvait être enrichie par le numérique, plutôt que remise en cause ou abolie.

Certains concepts nouveaux ont été ajoutés et nécessitent encore de s'enrichir avec le temps. Le plus complexe est celui de la pérennisation de l'information. Un vieux papier conservé proprement, quitte à prendre des cours de paléographie, on peut toujours le lire. Mais un vieux fichier ? Qui n'a jamais fait l'expérience d'un document, d'une photo, voire d'un ancien mémoire, retrouvé sur un disque qu'on n'arrive plus à ouvrir ou dont l'ouverture avec les applications actuelles donne une forme presque illisible ?

Le travail effectué avec la communauté des archivistes dans le cadre du programme VITAM a permis de faire monter en maturité un grand nombre d'archivistes sur l'adaptation au numérique des processus d'archivage. Il a aussi permis aux informaticiens qui construisaient la solution de comprendre les fondements et les vrais besoins, de partager les contraintes et d'apporter les pistes d'évolution et de transformation ouvertes par le numérique.

Une conclusion s'impose : la compétence des archivistes sur la gouvernance de l'information de manière probante et sur le temps long a bien toujours sa raison d'être, et peut-être encore plus dans le monde volatile du numérique.

Un programme qui continue...

La phase initiale du programme VITAM était prévue jusqu'à la V3. Cette version riche de toutes les fonctionnalités essentielles pour l'archivage numérique est sortie début 2020. Mais le programme ne s'arrête pas là :

- D'une part, il a été décidé d'assurer la maintenance et l'amélioration continue de la solution logicielle, avec le maintien d'une équipe et de moyens apportés par les trois ministères d'origine

(Culture, Défense, Affaires étrangères), mais aussi au travers de l’animation d’une communauté d’usagers et de contributeurs ;

- D’autre part, a été lancé un projet interministériel de mutualisation d’une plateforme d’archivage en mode service, « VITAM Accessible en Service » (VAS), embarquant à ce jour les ministères de la Culture (pour ses besoins non patrimoniaux), de l’Écologie et des Affaires sociales. Ce projet vient compléter le développement du *back-office* VITAM par un enrichissement fort des IHM.

Une réelle communauté d’usage est constituée autour de VITAM, elle continue à porter et à faire vivre le produit. Celui-ci pourra ainsi continuer à évoluer, en lien avec la transformation numérique du métier et des pratiques des archivistes.

J’ai eu la très grande chance d’être à la barre de cette aventure de mi-2015 à fin 2019. J’y ai découvert le monde des archives avec des professionnels extraordinaires, porteurs d’une mission de service public sous-estimée et pourtant essentielle.

Je ne peux malheureusement pas citer toute l’équipe VITAM mais je tiens à souligner l’action d’Edouard Vasseur⁽⁵⁾ et Frédéric Brégier, pour la réalisation collaborative de la solution, et de Mélanie Rebours, pour sa diffusion⁽⁶⁾. Je tiens aussi à saluer l’équipe qui a repris le flambeau sur VITAM, notamment Emmanuel Laborde, qui en a repris la tête et qui était déjà directeur technique depuis 2018, et Alice Grippon, en charge de la diffusion. Toute une équipe d’agents publics et de prestataires a relevé ce défi technologique et humain, et peut être sincèrement fière du résultat. Merci à tous.

Pour tout savoir sur VITAM et son actualité : www.programmevitam.fr ou Twitter : @progvitam

(5) Maintenant professeur à l’Ecole Nationale des Chartes.

(6) Et la cohésion de l’équipe.