La diffusion des archives à l’ère des humanités numériques. Enquête et perspectives sur la transformation numérique d’un patrimoine mémoriel romand.

Mémoire de Master réalisé par :   
Raphaël ROLLINET

E-mail [raphael.rollinet@unine.ch](mailto:raphael.rollinet@unine.ch)

Matricule : 21290572

(Image : Raphael Rollinet/Wikimédia Commons)

*Sous la direction de*

Lionel Bartolini

*Mémoire soumis le .. janvier 2026*

*Dans le cadre du Master*

MA - pilier patrimoine régional et humanités numériques

Université de Neuchâtel, semestres d’automne 2024-2025

Faculté  
Lettres et sciences humaines



Résumé

Ce Mémoire de Master dans le cadre du pilier « Patrimoine régional et humanités numériques », se concentre sur la fonction archivistique de la diffusion appliquée à notre ère dite numérique. Ce sujet sera abordé par l’axe de l'accessibilité et de la sélection des archives et vise à apporter une meilleure vision des évolutions professionnelles actuelles, mais aussi futures pour les archivistes, et l’impact du numérique sur cet environnement.

Cette étude introduira les thématiques par des éléments quantitatifs avec les indicateurs de la diffusion des archives en Suisse issue directement des institutions. Mais aussi des éléments qualitatifs notamment avec une enquête effectuée auprès des directions d’institutions cantonales d’archives en Suisse romande. Cette étude analysera l’impact de stratégies numériques appliquées à la diffusion d’archives, notamment celui des Archives fédérales suisses et la stratégie numérique suisse 2025.

Les résultats montrent que les stratégies numériques comprenant l’amélioration des systèmes d’information, la numérisation des archives renforce et dynamise la diffusion des archives. Néanmoins le numérique soulève de nombreuses questions pour la profession d’archiviste notamment pour s’adapter à ses enjeux, mais également en termes de moyens très inégaux entre les institutions romandes et des risques liés à la recherche et l’accessibilité des archives vis-à-vis des usagers.

**Mots-clés** : *Archivistiques, diffusion des archives, transformation numérique, humanités numériques,*

Remerciements

Mes remerciements particuliers pour monsieur Lionel Bartolini pour avoir accepté d’être mon directeur de mémoire. Monsieur Beretta également pour le soutien sur les questions numériques, notamment pour la base de données.

Mes remerciements pour leurs contributions à ce travail dans le cadre de l’enquête vont aux membres des directions romandes qui ont répondu à mon enquête et dont certains on était d’accord d’être cité nominativement, soit Delphine Friedmann directrice des Archives cantonales vaudoises, Fabienne Lutz-Studer directrice des Archives de l'État du Valais, le répondant des Archives de l'État de Fribourg, le répondant des Archives cantonales jurassiennes, Sébastien Bischoff archivistes responsables des questions numériques aux Archives de l'état de Neuchâtel, et le répondant des Archives d'État de Genève.

Des remerciements également aux membres de la Conférence des directrices et directeurs d'Archives suisses pour la mise à disposition en open data des données statistiques de leurs institutions.

Table des matières

[Résumé i](#_Toc211890764)

[Remerciements ii](#_Toc211890765)

[Liste des tableaux v](#_Toc211890766)

[Liste des figures vi](#_Toc211890767)

[Liste des abréviations vii](#_Toc211890768)

[1. Introduction 1](#_Toc211890769)

[1.1 Problématique et questions de recherche 3](#_Toc211890770)

[1.1.1 Problématique 3](#_Toc211890771)

[1.1.2 Délimitation du sujet 3](#_Toc211890772)

[1.1.3 Questions de recherche 4](#_Toc211890773)

[2. Contexte 5](#_Toc211890774)

[2.1.1 Institutions 5](#_Toc211890775)

[2.1.2 Fonctions de l’archiviste 5](#_Toc211890776)

[2.1.3 Fonction de diffusions 6](#_Toc211890777)

[2.1.4 Enjeux 7](#_Toc211890778)

[2.1.5 Cadre légal 8](#_Toc211890779)

[2.1.5.1 Loi sur les archives (Larch) 8](#_Toc211890780)

[2.1.5.2 La Loi sur la transparence des activités étatiques (LTAE) 9](#_Toc211890781)

[2.1.6 Cadre stratégique 9](#_Toc211890782)

[2.1.6.1 Code de déontologie professionnelle 9](#_Toc211890783)

[2.1.6.2 Livre blanc sur L’apprentissage automatique dans les archives 10](#_Toc211890784)

[2.1.6.3 Stratégie numérique suisse 12](#_Toc211890785)

[2.1.7 Transformation numérique en archives 15](#_Toc211890786)

[3. Méthodologie 18](#_Toc211890787)

[3.1 Revue de la littérature technique et professionnelle 18](#_Toc211890788)

[3.2 Utilisation de données en open data 18](#_Toc211890789)

[3.3 Enquête 19](#_Toc211890790)

[3.4 Création d’une base de données 19](#_Toc211890791)

[3.5 Visualisation de données 19](#_Toc211890792)

[3.6 Cas pratique 20](#_Toc211890793)

[4. Traitement du sujet 21](#_Toc211890794)

[4.1 Étude sur l’impact des évolutions numériques en archives 21](#_Toc211890795)

[4.1.1 Indicateurs liés aux compétences numériques des archivistes 22](#_Toc211890796)

[4.1.2 Indicateurs d’infrastructures techniques en archives 25](#_Toc211890797)

[4.1.3 Indicateurs liés aux moyens en archives 27](#_Toc211890798)

[4.1.4 Indicateurs liés aux collaborations institutionnelles 31](#_Toc211890799)

[4.1.5 Indicateurs liés à la gestion documentaire 32](#_Toc211890800)

[4.1.6 Indicateurs liés à l’accès aux archives 32](#_Toc211890801)

[4.2 Entretiens 41](#_Toc211890802)

[4.2.1 Entretien de Julien Raemy, data-archiviste aux AFS 41](#_Toc211890803)

[4.2.2 Entretiens de Baptiste de Coulon, data-archiviste à la Fondation SAPA 41](#_Toc211890804)

[4.3 Les humanités numériques 45](#_Toc211890805)

[4.3.1 Domaine 45](#_Toc211890806)

[4.3.2 Formation 46](#_Toc211890807)

[4.3.3 Projets et compétences applicatives aux Archives 48](#_Toc211890808)

[4.3.3.1 Web sémantique 48](#_Toc211890809)

[4.3.3.2 Indexation automatique 51](#_Toc211890810)

[4.3.3.3 Intelligence artificielle 55](#_Toc211890811)

[5. Analyse et mise en perspective 57](#_Toc211890812)

[6. Conclusion 58](#_Toc211890813)

[Bibliographie 60](#_Toc211890814)

[Annexe 1 : Déclaration sur l’honneur 63](#_Toc211890815)

[Annexe 2 : Entretien semi-conduit de Baptiste de Coulon 64](#_Toc211890816)

[Annexe 3 : Entretien semi-conduit de Julien Rémy 69](#_Toc211890817)

Liste des tableaux

[**Tableau 1 : Les fonctions archivistiques** 6](#_Toc211890818)

[**Tableau 2: Concept des 5 étoiles de l'open data** 49](#_Toc211890819)

Liste des figures

[**Figure 1: Répartition des titres d'étude chez les directrices et directeurs d’archives** 21](#_Toc211890820)

[**Figure 2 : Estimation moyenne de l’importance de compétences clés pour l’archiviste (score sur 100)** 22](#_Toc211890821)

[**Figure 3: Responsabilité de la gestion des projets numériques dans les services d'archives romands** 23](#_Toc211890822)

[**Figure 4: Compétences numériques les plus utiles en archives** 24](#_Toc211890823)

[**Figure 5: Infrastructure de numérisation dans les Archives cantonales/d'États romands** 25](#_Toc211890824)

[**Figure 6: Archivage électronique dans les institutions d'archives cantonales** 27](#_Toc211890825)

[**Figure 7: Evolution des dépenses des archives fédérales suisses par années (1951-2024)** 28](#_Toc211890826)

[**Figure 8: Variation des employés à durée indéterminée et temps plein par année en archives** 29](#_Toc211890827)

[**Figure 9: Ratio d’employés en archives pour 1000 employés de l’administration** 30](#_Toc211890828)

[**Figure 10: Types de collaborations institutionnelles déclarées par les 6 Archives cantonales** 31](#_Toc211890829)

[**Figure 11: Volume total des fonds d'archives traités (papier), en mètres linéaires, par année** 32](#_Toc211890830)

[**Figure 12: Évolution du nombre de places dans les salles de lecture en archives** 33](#_Toc211890831)

[**Figure 13: Nombre d'heures d'ouverture par semaine de la salle de lecture** 34](#_Toc211890832)

[**Figure 14: Nombre d'usagers uniques en salle de lecture par année** 34](#_Toc211890833)

[**Figure 15: Nombre de visiteurs uniques utilisant les inventaires en ligne par année** 35](#_Toc211890834)

[**Figure 16: Nombre d'articles recherchables en ligne par année** 36](#_Toc211890835)

[**Figure 17: Évolution du nombre d'unités physiques commandées en salle de lecture** 38](#_Toc211890836)

[**Figure 18: Analyse bivariée du nombre d'unités physiques commandées en salle de lecture par année en résidus pondérés du test d'indépendance** 39](#_Toc211890837)

[**Figure 19: Illustration du concept des 5 étoiles de Tim Berners-Lee** 50](#_Toc211890838)

[**Figure 20: Images numérisées de la commune de Milvignes** 53](#_Toc211890839)

Liste des abréviations

|  |  |
| --- | --- |
| **Abréviation** | **Intitulé** |
| AAS | Association des archivistes suisse |
| ACV | Archives cantonales vaudoises |
| AEF | Archives de l'État de Fribourg |
| AEG | Archives d'État de Genève |
| AEN | Archives de l'état de Neuchatel |
| AEV | Archives de l'État du Valais |
| AFS | Archives fédérales suisses |
| AIS | Archiv information system ou système d’information archivistique |
| ArCJ | Archives cantonales jurassiennes |
| CDA | Conférence des directrices et directeurs d'Archives suisses |
| GLAM | Galleries, Libraries, Archives and Museums |
| RiC | Record in Context |
| SHS | Sciences humaines et sociales |

# Introduction

Quel est le rôle du patrimoine, si ce n’est de conserver et transmettre l’information et les connaissances amassées à travers le temps, non seulement à un public actuel, mais également futur. À cette fin, les archives tiennent un rôle central tant dans son soutien à la recherche comme passeur de sources auprès des chercheurs qu’auprès du citoyen comme garant de la transparence de l’état.

La profession d’archiviste tient sur des fonctions dont la finalité est celle de la diffusion, encadrée par une pratique professionnelle bien établie, des normes, une législation et une déontologie professionnelle. La diffusion auprès d’un public est donc au cœur de la réflexion de l’archiviste tant pour les besoins actuels que futur.

Nos sociétés modernes actuelles sont en évolution constante au sein de ce que nous pouvons décrire comme la 4e révolution de l’information, soit la communication informatique. Dans ce contexte les archives tiennent un rôle de médiateur dans cette diffusion et s’adaptent à de nouveaux médiums documentaires numériques.

En Suisse les archives sont comprises dans la stratégie numérique 2025 de la Confédération (Chancellerie fédérale 2025) visant directement à permettre l’accès numérique des archives indépendamment du temps et du lieu. Cela deviendra la norme d'ici 2025, tandis que l'accès analogique deviendra l'exception. Cette stratégie numérique met également l’accent sur l’utilisation de l’IA, l’utilisation de logiciel ouvert et la sécurité de l’information.

Dans ce contexte, cette recherche tend à montrer les évolutions de cette fonction archivistique de diffusion à l’ère numérique, par un regard d’un archiviste humaniste. Cette recherche vise également à comprendre les évolutions actuelles qui impactent les archives, les tendances des institutions locales, les innovations en la matière, d’en donner un panorama à l’aune du domaine des humanités numériques.

En effet, les humanités numériques sont une discipline relativement jeune, aux croisements des sciences humaines et de l’informatique. Elle vise comme transdiscipline à ouvrir de nouvelles perspectives de recherches grâce aux développements de méthode numérique. Il s’agit d’une discipline particulièrement pertinente dans le cadre de cette recherche, car les archives sont souvent un passage inévitable pour les chercheurs de SHS et un grand nombre d’archivistes sorte de formation initiale en science humaine ou sociale, particulièrement l’histoire dont l’archivistique en est une science connexe.

Les humanités numériques sont donc pertinentes pour fonder un discours sur les évolutions numérique touchant l’ensemble des GLAM notamment les archives tant du point de vue du chercheur en SHS que du professionnel des archives. De plus les archives ont toujours constitué un laboratoire pour les humanités numériques pour appliquer des méthodes tant du point de vue du traitement automatique du langage, de l’indexation automatique qu’aujourd’hui avec le web sémantique auquel les archives s’adaptent notamment avec la nouvelle norme Record in Context.

Mais il s’agit au sens large d’un ensemble de méthode computationnelle tel que l’intelligence artificielle générative, le machine learning, le deep learning pratiqué depuis longtemps en humanités numériques que l’on peut aujourd’hui retrouver cité comme objectif prioritaire de développement pour les archives. L’on peut évoquer récemment le « Livre blanc de l’apprentissage automatique dans les archives » rédigé par l’Association suisse des archivistes, qui tend à montrer des synergies importantes entre les humanités numériques et les objectifs stratégiques des archives en Suisse.

Ces méthodes numériques ont un point commun, elles permettent l’exploitation et l’enrichissement documentaires pour offrir une meilleure diffusion à l’ère numérique que nous connaissons actuellement. Dans ce cadre cette recherche menée par l’axe d’évolutions numériques touchant les archives, s’intéressera à la fonction de diffusion dans un contexte institutionnel régional.

Cette recherche intègrera une réflexion sur la place des humanités numériques tant sur le plan disciplinaire que de l’importance de la formation, aussi bien pour la pratique de la recherche que de celle des professions du patrimoine. Mais également l’importance du point de vue de l’archiviste d’avoir une convergence entre la technique ou pratique professionnelle et celle des enjeux liés aux technologies dans le cadre de son rôle de gardien d’une mémoire commune.

Finalement, ces réflexions mêlant patrimoine régional et humanités numériques, provenant d’un archiviste de tradition humaniste, tendent à valoriser à la fois la place du patrimoine dans la société, son rôle pour la recherche que celui d’inscrire les sciences humaines dans un débat légitime et contemporain sur le numérique. Mais également sur notre consommation intelligente des technologies de l’information en tant qu’outils aux services tant des humanités que de l’Humanité.

## Problématique et questions de recherche

### Problématique

La réflexion initiale de ce travail autour des évolutions de la diffusion archivistique en lien avec le numérique est lié en grande partie à mon expérience académique dans le domaine des sciences de l’information et des humanités numériques, mais également professionnellement en tant qu’archiviste.

En effet cette fonction de diffusion passe par la numérisation, l’archivage électronique et la diffusion des inventaires par l’intermédiaire de systèmes d’information archivistiques (AIS) et de portail sur internet. Ces moyens numériques appliqués au domaine des archives permettent de questionner l’impact et les évolutions de la diffusion en archives par le biais du numérique.

Le rôle de l’Archiviste lui-même dans cet environnement s’en trouve transformé n’étant plus seulement le gardien du patrimoine qui sur le plan de la conservation numérique est plutôt du ressort de l’informaticien. Mais l’archiviste intervient plutôt sur la gestion du cycle de vie dès le début de ce cycle afin de limiter les pertes d’informations, les risques judiciaires et d’en améliorer la gestion et la diffusion[[1]](#footnote-1). C’est également ces raisons qui mettent l’archiviste au cœur de la transformation numérique des administrations.

Le numérique est devenu un enjeu central pour les archives, mais également sensible pour les archivistes, car extrêmement complexe à gérer pour des professionnelles devant aujourd’hui s’occuper de fonds d’archives dites « hybride », c’est-à-dire à la fois physique et numérique. Il est donc intéressant dans ce contexte de se poser la question sur les choix stratégiques, les évolutions du métier pour s’adapter à ces évolutions et l’impact de celle-ci sur le métier d’archiviste. Mais également afin de garantir l’accès et la qualité des services d’archives aux usagers et chercheurs.

### Délimitation du sujet

Le sujet se limitera à l’étude des évolutions produites par le numérique sur la diffusion des archives en Suisse romande. La période observée par les statistiques en possession concerne les années 2013-2024. L’objectif principal de ce travail est de faire un bon « arrêt sur image » de la diffusion des archives pour mieux analyser les enjeux pour le domaine.

Des entretiens avec des experts viendront viseront à apporter des solutions, tendances ou applications pratiques afin d’améliorer cette fonction archivistique en Suisse romande, cela tant au profit des archivistes que des usagers notamment les chercheurs issus des SHS. Le présent travail ne vise pas à apporter des solutions techniques clé en main, mais plutôt des réflexions stratégiques.

Bien que ce travail se concentre sur la fonction de diffusions des archives dans un contexte de plus en plus numérique, ce mémoire abordera les autres fonctions archivistiques dans un objectif de contextualiser l’impact du numérique qui doit être considéré depuis la création du document d’archives jusqu’à sa diffusion qui si elle est faite numériquement demandera aussi des moyens numériques pour conserver les images issues d’une numérisation par exemple.

Ce travail ne prétend pas offrir de réponse à l’ensemble des institutions d’archives, il se concentre sur les archives d’états/cantonales à un niveau local ayant des missions similaires. Les compares à d’autres institutions notamment les archives fédérales, cela afin d’esquisser les enjeux et d’éventuelle proposition d’amélioration applicable à la situation locale.

### Questions de recherche

**Question de recherche principale**

* Quel est l’impact du numérique dans la diffusion archivistique en Suisse romande ?

**Question de recherche secondaire**

* Les stratégies institutionnelles du domaine des archives sont-elles en adéquation avec les évolutions numériques actuelles ?
* Quels sont les apports du domaine et des formations en humanités numériques pour ces évolutions technologiques aux domaines du patrimoine ?

# Contexte

### Institutions

Dans le cadre de ce travail, mon intérêt c’est limiter un à un niveau local, la Suisse romande. Les réflexions abordées durant ce travail seront centrées sur la situation romande et les institutions concernées sont essentiellement les archives d’états/cantonales des cantons romands, soit :

* Archives cantonales vaudoises
* Archives de l'État de Fribourg
* Archives d'État de Genève
* Archives de l'état de Neuchâtel
* Archives de l'État du Valais
* Archives cantonales jurassiennes
* Archives fédérales suisses

D’autres institutions seront utilisées également comme point de comparaison sur le plan du numérique, notamment les Archives fédérales suisses dont rayonnement national concerne la suisse romande, mais également d’autres institutions telles que les archives de la fondation SAPA présente notamment à Lausanne, pour l’application de moyen novateur sur le plan de la diffusion des archives qui pourrait tout à fait être appliquées plus largement en Suisse romande.

### Fonctions de l’archiviste

En archives, les fonctions archivistiques reposent sur la conception des missions de l’archiviste, il existe de nombreuses conceptions de ces fonctions archivistiques variant selon l’origine ou la formation de l’archiviste, provenant souvent des sciences historiques, sciences de l’organisation ou sciences de l’information. Des traditions se sont également développées selon les pays notamment la tradition française ou la tradition québécoise qui est la plus citée dans les sources consultées pour ce travail, reposant essentiellement sur les travaux d’archivistes québécois tels que l’éminent Carol Couture[[2]](#footnote-2).

**Tableau 1 : Les fonctions archivistiques**

|  |  |  |
| --- | --- | --- |
| Tradition française | Tradition québécoise | Modèle OAIS (numérique) |
|  | Création |  |
|  | Evaluation | Pré-ingest |
| Collecte | Accroissement | Entrée (ingest) |
| Classement | Classement | Gestion des données |
| Conservation | Conservation | Stockage |
|  | Description | Gestion du système |
| Communication | **Diffusion** | **Accès** |

(Tableau dessiné à partir des données de Wikipédia[[3]](#footnote-3))

L’on constate sur le tableau ci-dessus que peu importe les différentes traditions ou que les archives soient physiques ou numériques, la finalité reste la même soit la communication, la diffusion ou l’accès aux archives. Il s’agit finalement de l’utilisation des archives qui est la raison d’être de la conservation des archives et du travail perpétuel de l’archiviste sur une durée historique, soit indéterminée.

### Fonction de diffusions

La diffusion est une fonction archivistique qui inclut les activités de promotion des archives ou de l’institution, la mise en valeur ou valorisation des archives et l’exploitation des archives. Dans le cadre de ce rapport, la définition de cette fonction est issue de l’article « La diffusion des archives ou les 12 travaux des archivistes à l'ère du numérique »[[4]](#footnote-4) :

« La diffusion est l’action de faire connaître, de mettre en valeur, de transmettre ou de rendre accessibles une ou des informations contenues dans des documents d’archives à des utilisateurs (personnes ou organismes) connus ou potentiels pour répondre à leurs besoins spécifiques. »

**Normand Charbonneau**

*(Archiviste québécois et chef de l'exploitation chez Bibliothèque et Archives Canada)*

Il s’agit donc d’une finalité pour l’archiviste, mais c’est également la fonction directement en lien avec l’usager et notamment les chercheurs dont les archives restent un préalable à de nombreuses recherches en SHS.

En termes de moyen numérique, comme le précise Carol couture et Marcel Lajeunesse dans leur ouvrage de référence[[5]](#footnote-5), le début du 21e siècle a vu naitre et se développer à toute allure de nouveau moyens de diffusion. A la fin des années 1990, de nombreux établissements d’archives avaient ajouté un site web à la panoplie d’outils d’information et de promotion qu’ils possédaient déjà, mais c’est véritablement dans les années 2000 que l’on constate l’adoption massive du site web institutionnel.

D’abord simple reproduction numérique des instruments de recherche traditionnels, le site évolue tant dans les contenus qu’il véhicule que dans ses fonctionnalités. Il devient plus complexe et interactif grâce à de nouvelles applications favorisant le partage d’information et le réseautage social. La contribution des utilisateurs est de plus en plus sollicitée ; ils sont invités à commenter, décrire diffuser les documents d’archives.

Toujours plus proactifs et exigeant ils s’attendent à pouvoir tout trouver et consulter, instantanément, gratuitement, et à partir d’un seul point d’accès. Ce nouveau contexte et les exigences élevées du public ont amené les établissements d’archives à transformer radicalement leurs services aux usagers et à réaménager leurs salles de consultation afin de garder leur clientèle traditionnelle, mais aussi d’attirer de nouveaux publics. Et cela va sans tenir compte du fait que maintenant grâce à la numérisation d’un nombre de plus en plus imposant de documents, les chercheurs ont souvent accès à l’information de leur domicile.

Cette évolution de la diffusion des archives depuis les années 1990-2000 se poursuis et s’accélère avec de nombreux enjeux mais également un contexte stratégique et légal qui se complexifie à mesure que la transformation numérique s’accroit tant aux niveaux des archives que de l’ensemble des administrations et au sens large de la société.

### Enjeux

Parmi les enjeux liés à la diffusion, évoqués déjà en 2012 dans l’article « *La diffusion des archives ou les 12 travaux des archivistes à l’ère du numérique »[[6]](#footnote-6)*. L’on peut citer la nécessité d’encourager une culture de collaboration pour trouver des synergies, notamment en établissant des partenariats entre les archives, les bibliothèques et les musées afin de s’adapter aux nouveaux concurrents, nouveaux lieux et nouvelles pratiques.

En termes de système d’information, développer des outils de gestion, c’est notamment le cas de nos jours avec l’arrivée de nouveau système basé sur le web sémantique. Le numérique permettant d’accroitre la visibilité du domaine des archives sur le web. Au niveau de l’information, l’enjeu est de faciliter le partage et la réutilisation notamment avec ce que l’on nomme aujourd’hui l’open data. Mais également de défendre l’espace public et garantir le respect de la vie privée, raison pour laquelle les inventaires d’archives sont aujourd’hui anonymisés.

L’on retrouve également le fait de connaitre les usagers et les usages des archives et encourager le contact direct avec le public. C’est un élément motivé par la médiation culturelle afin de considérer la dimension émotive des archives en élargissant le cadre de référence des archives. De nouvelles méthodes de médiation ont amené à des projets d’archives collaboratives grâce aux sciences participatives. Cela permet de créer un nouveau lien avec le public tout en exploitant la force de l’intelligence collective. C’est typiquement la raison de projets d’enrichissement de Wikipédia ou Wikidata en archives.

Il est intéressant que ces enjeux pourtant liés spécifiquement à l’ère numérique ne soient pas strictement numériques, mais peuvent en partie être pris en compte dans le développement de système d’information ou dans le cadre de la médiation culturelle autour des archives. L’ensemble de ces enjeux montre que le numérique est profondément transversal et touche l’ensemble des fonctions de l’archiviste.

### Cadre légal

Les missions et fonctions de l’archiviste étant d’intérêt public, elles sont encadrées par un cadre légal. En Suisse, avec un système fédéral basé sur le principe de subsidiarité garantie une souveraineté des cantons sur leur législation y compris sur le plan des archives. L’on peut citer quelques exemples ayant un impact sur la diffusion particulièrement à notre ère numérique.

#### Loi sur les archives (Larch)

En Suisse la plupart des cantons disposent de leurs lois sur l’archivage. Pour donner quelques exemples ayant un impact avec le numérique. L’on peut citer la Loi sur les archives (Larch)[[7]](#footnote-7) neuchâteloise qui évoque notamment à l’Article 2 de la loi sur l’archivage que « *l’archivage des documents, organisé dans l’intérêt public, contribue à assurer la gestion et la traçabilité des activités des autorités cantonales et communales mentionnées à l’article 4, la justification des droits des personnes physiques ou morales.* *Mais également la sauvegarde et l’étude du patrimoine historique, économique, social et culturel du canton de Neuchâtel. »* Il s’agit là de la définition de la valeur primaire et secondaire des archives.

#### La Loi sur la transparence des activités étatiques (LTAE)

Très important également celle de l'accès du public aux archives, notion essentielle pour la diffusion. La Loi sur la transparence des activités étatiques (LTAE)[[8]](#footnote-8) précise à l’article premier que *« la présente loi a pour but de garantir la libre formation de l'opinion publique et de favoriser la participation à la vie publique en veillant à la transparence des activités des autorités »,* élément essentiel pour le rôle démocratique des archives.

L’article 11 ajoute *« selon les moyens dont elles disposent, les autorités peuvent mettre à disposition du public, par le biais des technologies modernes d’information et de communication, les informations qu'elles ont transmises aux médias et d’autres documents jugés importants. »* Il s’agit d’un élément notable pour la diffusion par des moyens numériques.

Ces extraits montrent que le cadre légal ou déontologique de la fonction de diffusion des archivistes est déjà bien encadré, fournissant aux archivistes un cadre menant à la rédaction d’objectifs stratégiques découlant de la législation.

### Cadre stratégique

Le cadre stratégique pour le numérique en archives en suisse est très intéressant, car le modèle fédéral suisse permet l’émergence de multiple stratégie. Certaines stratégies sont basées sur des recommandations professionnelles visant à définir des tendances aux domaines avec la déontologie professionnelle ou des standards.

D’autres stratégies sont plus contraignantes et basées sur une législation, c’est notamment le cas de la stratégie numérique suisse aux AFS. Il est intéressant de faire ressortir quelques-uns des aspects stratégiques les plus importants en Suisse pour contextualiser introduire ce travail.

#### Code de déontologie professionnelle

Sur le plan stratégique en premier lieu l’on peut citer l’importance de la déontologie professionnelle de l’archiviste en faveur d’une diffusion. Ce code de déontologie permet d’encadrer le rôle de l’archiviste dans la durée et de garantir les fondamentaux du métier. L’alinéa 6 précise *« Les archivistes facilitent l’accès aux archives du plus grand nombre possible d’utilisateurs et offrent leurs services avec impartialité à tous les usagers »[[9]](#footnote-9).*

#### Livre blanc sur L’apprentissage automatique dans les archives

L’association des archivistes suisse (AAS), c’est récemment doté d’un livre blanc sur « L’apprentissage automatique dans les archives : l'indexation en profondeur au service de l'accès aux archives »[[10]](#footnote-10). Cet outil stratégique est important aux vues des changements apportés par le numérique en archives en proposant des méthodes et outils afin d’encadrer les archivistes dans ces évolutions numériques notamment avec l’arrivée des intelligences artificielles.

Parmi les applications potentiellement intéressantes dans ce document stratégique pour le domaine des archives, l’on peut citer :

* Le résumé du contenu d'un ensemble de documents : il devient possible que les méthodes d'indexation en profondeur ne s'appliquent plus uniquement aux documents individuels, mais à un ensemble ou un réseau de documents ou de dossiers apparentés, comme par exemple un ensemble de résultats (Record set) pour une requête spécifique de l'utilisateur[[11]](#footnote-11).
* La création de formulations créatives de questions de recherche à partir des métadonnées ou des textes complets : cela implique, d'une part, que les documents d'archives soient disponibles sous forme numérique et lisible par machine et, d'autre part, que les métadonnées et les données primaires soient utilisées comme matériel d'entraînement pour les modèles linguistiques ou que de tels modèles linguistiques soient intégrés dans les systèmes d'information des archives.[[12]](#footnote-12)
* Chatbots et systèmes de recherche qui génèrent directement des réponses aux questions des utilisateurs : de tels systèmes existent déjà aujourd'hui, ils devraient être intégrés dans les systèmes d'archives et être entraînés à l'aide de données provenant des fonds d'archives. Une possibilité attrayante est la composition automatique d'images ou de films appropriés (comme des expositions virtuelles). Ces idées peuvent même être poussées jusqu'à la création automatique d'un arbre généalogique, qui remplacera ainsi la simple mise à disposition de données sur des ancêtres individuels.[[13]](#footnote-13)

Ce livre blanc met également en lumière les principales opérations mais également contrainte, attentes pour l’utilisation de méthode de Machine Learning en archives. Comme il s’agit d’un élément central dans la mise en œuvre d’une transformation numérique en archives visant à l’enrichissement des métadonnées pour leurs diffusion. Il me parait important de les détailler dans ce mémoire en reprenant succinctement ces informations.

L’on retrouve des opérations visant à analyser et permettre l’exploitation du document d’archives. Notamment la reconnaissance optique de caractères (OCR), la reconnaissance de texte manuscrit (HTR), ou la reconnaissance de la parole au texte (ASR). Des opérations de segmentation de documents avec la reconnaissance de la mise en page et des modèles, par exemple de tableaux visant à un remplissage automatique de champs à partir d’un modèle type.

Le Machine Learning en Archives concerne également des opérations de classification ou indexation automatique des documents. Il permet de générer des étiquettes ou mots-clés sur la base des caractéristiques extraites. Il permet notamment la reconnaissance des entités (NER) à l'aide des mots clés extraits, de l'analyse d'images et du traitement du langage naturel (NLP).

Des méthodes permettent également d’aller plus loin en permettant la mise en correspondance des métadonnées entre des données liées ou Linked Data et les notices d’autorité. Cela permet de relier les entités reconnues par la reconnaissance d'entités nommées (NER) aux entités identifiées dans les sources de connaissances basées sur les technologies de données ouvertes liées (Linked Open Data) par le biais de la liaison d'entités nommées (NEL). Les perspectives sont très intéressantes pour sortir les inventaires d’une logique de silo, en effectuant des reconnaissances avec les entités nommées de Wikidata par exemple.

Il est possible d’aller plus loin qu’une simple indexation avec des modèles de Machine Learning. Notamment en condensant les informations présentes sur le document numérique en métadonnées (Abstracting), cela peut aller jusqu’à la description de l'affaire et résumé du contenu du dossier, soit création de résumés ou d’incipit. Cela peut facilement être automatisé par des outils grand public comme ChatGPT ou pdf.ai.

L’on retrouve également des opérations utilisant des méthodes de Machine Learning directement en liens avec l’usager et visant à faciliter ou « augmenter » l’accès au document d’archives. Cela passe par la traduction automatique des métadonnées des inventaires d’archives permettant un accès multilingue. Cela peut être mis en œuvre par des outils connues tel que DeepL. Grâce à l’implémentation de *« chatbots »* directement sur les portails d’accès d’archives, il est également possible de fournir un soutien à la recherche. Ces *« chatbots »* peuvent également garder en mémoire les préférences et recherche de l’utilisateurs pour affiner les résultats de recherche afin de mieux les adaptés aux besoins individuels des usagers.

Le livre blanc précise également qu’une application réussie du Machine Learning dans la chaîne de processus archivistique présuppose dans la plupart des cas une précision élevée à très élevée tant des données primaires que des dérivés qui en résultent. Il peut s'agir de fichiers numérisés avec une haute résolution pour l'analyse réussie d'images, de sons et d'images animées en vue de générer des dérivés de qualité contrôlée, ou de fichiers Office convertis avec précision sur une longue période avec des textes et des polices intégrés, des données complètes et des métadonnées et, pour les images, des données EXIF. Il est également précisé que à la fin de la chaîne de processus archivistique, il peut en résulter différentes formes, en partie nouvelles, de métadonnées ou de dérivés crée grâce aux Machine Learning.

Si l’on résume l’utilisation du Machine Learning en archives, l’on peut dire qu’il permet plus que l’automatisation des tâches de l’archiviste mais augmente sont impact, au moment ou le domaine est au cœur d’une transformation numérique. En partant de la question de recherche « Les archives peuvent-elles se permettre de recourir à l'intelligence artificielle – ou à l'apprentissage automatique ? ». Ce livre blanc conclu avec la question de savoir s’il faudrait plutôt se demander, à l'inverse, si une archive peut se permettre de renoncer à l'utilisation du ML.[[14]](#footnote-14)

C’est donc en partant de ce livre blanc pour encourager et tester la mise en œuvre de technique issue de l’Intelligence artificielle au niveau opérationnel dans les archives, qu’il intéressant de s’interroger sur l’intégration de ces méthodes dans une stratégie institutionnelle plus large.

#### Stratégie numérique suisse

Au niveau de l’administration publique en Suisse, une transition numérique est à l’œuvre à tous les niveaux politique. En effet depuis 2022, le conseil fédéral et les gouvernements cantonaux ont créé l’Administration numérique suisse assurant en une stratégie et un pilotage commun la coordination de la transformation numérique menée aux trois échelons de l’état, soit fédérale, cantonale et communale.[[15]](#footnote-15)

La stratégie fixée par ce nouveau cadre est la stratégie numérique suisse qui fixe quant à elle des lignes directrices pour la transformation numérique en Suisse. Les principes guident la mise en œuvre dans le contexte de la transformation numérique[[16]](#footnote-16) :

* Orientation vers les utilisateurs et inclusion : les interactions avec l’administration sont simplifiées pour tous les groupes concernés.
* « Priorité au numérique » et « tout numérique » : les processus administratifs destinés à la population sont principalement conçus pour la fourniture de prestations numériques. L’économie et l’administration sont à pied d’œuvre pour exécuter de manière entièrement numérique leurs prestations et processus.
* Interopérabilité et réutilisation : les solutions sont conçues de manière à être interopérables et les données de l’administration peuvent être utilisées à plusieurs reprises (principe de saisie unique).
* Sécurité, confiance et transparence : des normes et des directives établies en matière de sécurité des données sont appliquées.
* Globalité et durabilité : les prestations sont développées de manière durable sur les plans économique, écologique et social.

Cette stratégie est fortement corrélée à l’agenda 2030 de l’Organisation des Nations Unies. Elle donne aux acteurs de la numérisation qu’ils soient publics ou privés, un cadre sur lequel s’appuyer. Il s’agit d’un cadre stratégique non contraignant, mais qui peut être repris à tous les niveaux et adapter aux objectifs sectoriels des organisations comme la Stratégie administration fédérale numérique pour l’administration fédérale.

Cette stratégie numérique suisse donne lieu à des thèmes prioritaires, qui sont actuellement l’Intelligence artificielle, la sécurité de l’information et la cybersécurité et les Logiciels ouverts. Il s’agit d’une stratégie qui impacte directement le domaine des archives à plusieurs niveaux, et c’est un élément central dans la compréhension globale de ce mémoire.

Pour illustrer l’aspect sensible des enjeux lié au numérique en archives, le cas des archives fédérales suisses (AFS) est tout trouver avec la récente mise en œuvre de la stratégie numérique suisse. Cette stratégie sera centrale dans ce mémoire puisqu’elle recouvre l’ensemble du champ numérique applicable aux archives pour l’accès à celle-ci. Cela en fait un très bon exemple et point de comparaison du point de vue de stratégies et mis en œuvre de méthode numérique dans une institution d’archives.

En effet, celle-ci fixe spécifiquement aux AFS un pivot de la diffusion des archives vers le principe de la *« priorité au numérique »* soit *« L'accès numérique aux informations des archives, indépendamment du temps et du lieu, deviendra la norme d'ici 2025, tandis que l'accès analogique deviendra l'exception »*. [[17]](#footnote-17)

La stratégie numérique suisse mise en œuvre entre 2021 et 2025 est très intéressante pour illustrer ce sujet, car elle est l’héritière d’une vision de longue date et d’effort visant à améliorer les prestations des archives au service d’une société de plus en plus numérique. Si les AFS avaient déjà de longue date un système d’information archivistique avec ScopeArchive depuis 2003 et un référentiel de préservation avec Preservica depuis 2009, un renouvellement de cet environnement était de plus en plus nécessaire pour s’adapter aux exigences actuelles.

La mise en œuvre de cette stratégie est donc passée par les étapes suivantes, le renforcement de la structure organisationnelle et des compétences internes, l’amélioration de la qualité des données notamment par l’anonymisation au niveau de la description des notices. La mise en œuvre du nouveau Viewer soit le portail en ligne dès 2019.

Mais également l’automatisation des processus avec notamment la fin de la mise en place de l’infrastructure de numérisation en 2021, cela afin que toutes les unités d’archives commandées soient progressivement mises à disposition gratuitement numériquement auprès de l’usager à la demande, d’un simple clic de commande.

Éléments intéressant du point de vue de la stratégie, il est prévu dans les objectifs de cette stratégie que les AFS *« jouent de leur influence dans l’élaboration des stratégies et de la législation relatives à la transformation numérique de la Confédération, afin de garantir l’archivage des informations de l’administration fédérale et de prévenir d’éventuelles lacunes »[[18]](#footnote-18).*

*Il est noté également « évaluent la mise en œuvre des recommandations résultant de l’évaluation de la loi fédérale sur l’archivage (LAr) de concert avec les acteurs concernés, et veillent à ce que la nécessité et les bénéfices escomptés d’une révision de loi soient clarifiés. »[[19]](#footnote-19)* Ces éléments confirment la place centrale des spécialistes en information et archiviste dans la transformation numérique du secteur public, en l’occurrence ici de la confédération.

En termes de transformation numérique, elle donne la part belle aux nouvelles technologies et à l’automatisation notamment en *« combinent l’expertise humaine et l’intelligence artificielle en matière d’analyse de données et de recherche historique pour fournir à la clientèle la meilleure assistance possible »[[20]](#footnote-20).* L’on peut également citerles objectifs *« automatisent progressivement leurs processus d’archivage standard ; s’inscrivant dans une perspective globale, l’automatisation vise à l’amélioration constante de l’efficacité, au bénéfice de la clientèle » ou encore « traitent les données sensibles de manière sécurisée tout au long du processus de leur reprise, de leur prise en charge et de leur utilisation numérique ; au besoin, elles renforcent la protection des documents numériques classifiés »[[21]](#footnote-21).*

Dans le cadre de la mise en œuvre de cette stratégie, cela a notamment été le cas pour l’automatisation de l’anonymisation des notices d’inventaires permettant leurs mises à disposition en respectant le cadre légal actuel et facilitant la réutilisation de ces données dans une perspective d’open data.

Open data qui prendra la forme dès 2011 par un pilote d’un portail de donnée ouverte « opendata.admin.ch » qui sera mis en service en 2016 et auquel les AFS ont participé. Mais l’open data appliquer directement aux données d’inventaire est aussi d’actualité avec le projet LINDAS. D’abord née d’un projet pilote aLod (Archival Linked Open Data)[[22]](#footnote-22) les AFS poursuivent leurs efforts de mise en œuvre d’une infrastructure en Linked data afin de mettre à disposions sur le WEB les données en Linked data et conservé sous forme de triplet RDF.

### Transformation numérique en archives

D’une manière générale, le cadre légal et les stratégies mises en place pour le domaine des archives visent à tirer la gestion des archives vers le haut en la mettant au cœur de la transformation numérique des institutions et administration publique. Transformation ou l’on retrouve souvent au cœur du numérique, l’information.

La mission principale des archives reste la gestion et conservation de la production documentaire de leur administration. Une administration aujourd’hui fonctionnant autour du numérique avec des documents nativement numériques amenant donc les archives depuis maintenant plusieurs années à s’adapter à ces nouveaux médiums documentaires.

Cette adaptation passe par la gestion électronique des documents, l’on peut notamment citer GEVER aux AFS rendus obligatoires par une ordonnance fédérale en 2019[[23]](#footnote-23). Dès lors, les avec une production documentaire nativement numérique, vient la question de la conservation du document numérique aux archives, à ne pas confondre avec un simple stockage. L’archivage électronique est un sujet d’actualité en archives.

Hormis la conservation vient la question de l’accès ou de la diffusion des archives. En effet du moment où les archives ne sont plus seulement constituées de vieux manuscrit comme l’image du métier peux encore le véhiculer, mais d’une vaste gamme de médium documentaire, il devient nécessaire de traiter leurs diffusions de la même façon afin que l’accès à tous les types de documents soit garanti.

La numérisation joue un rôle crucial dans la diffusion des archives physiques souvent plus anciennes. Attention en archives, on ne parle pas vraiment de dématérialisation, en effet le document original est toujours conservé. L’on peut dire que les motivations d’une numérisation d’archives sont multiples, d’une part créer des copies de consultation numérique afin de les diffuser de la même manière qu’un document nativement numérique.

La numérisation sert également pour des raisons de soutiens à la conservation en limitant son utilisation physique et en garantissant une trace de l’information présente sur le document. Pour que l’information numérique soit pérennisée, l’archiviste a besoin d’une infrastructure technique garantissant sa préservation, mais également en garantissant son authenticité, son intégrité, et son utilisabilité.

Qu’ils soient nés numériques ou en format physique, une fois numérisé un document numérique à besoin d’une infrastructure numérique pour le stocker et l’archiver. L’archivage ou le stockage électronique sont deux concepts distincts. Si le stockage d’un document numérique est assez facile. L’archiviste doit quant à lui s’assurer sa pérennisation. Dans le cadre de ce mémoire focalisé sur la diffusion, l’archivage électronique étant un sujet en soi il ne sera que contextualisé comme un prélude nécessaire à la diffusion d’un document numérique pour les institutions d’archives.

L’utilisation de système d’archivage électronique (SAE) est actuellement un enjeu majeur de la transformation numérique en archives. Utilisant souvent un modèle OAIS, assujetti à des normes ISO telles qu’ISO 14721:2025 pour le cadre de référence pour l’archivage à long terme. ISO 16363:2025 pour effectuer des audit et certification des dépôts numériques de confiance (TDR). ISO 23081-1:2017 et 23081-2:2021 pour les éléments de métadonnées pour les records. ISO 13008:2022 pour la conversion et migration des documents numériques (préserver authenticité, intégrité, utilisabilité).

Mais encore ISO 14641:2018 pour l’Archivage électronique : conception et exploitation d’un système garantissant lisibilité, intégrité, traçabilité. (Souvent cité comme « équivalent international » de NF Z42-013. Ces normes se retrouveront également dans les normes nationales eCH.

Les archives se sont par ailleurs dotées d’un pôle de compétence spécialisée pour l’archivage électronique, le Kost-ceco. Ce pôle de compétence propose des conseils dans la mise en œuvre d’un système d’archivage électronique, mais également un ensemble d’outils libres dans cet objectif.[[24]](#footnote-24)

L’archivage électronique pose également de nombreuses questions tant sur la forme de celui-ci que des moyens de stockage. On peut notamment citer le projet d’archivage sur ADN effectué notamment aux ACV.[[25]](#footnote-25)

Au niveau des systèmes d’information archivistique (SIA), une part importante d’entre eux sont vieillissants, en voie d’obsolescence technologique comme le montre un rapport d’expert du consortium DISTAM directement lié aux humanités numériques. Ce rapport fait état de l’obsolescent d’un système de gestion d’archives très connue et faisant également office de portail Web pour la diffusion des archives, AtoM. Le rapport fait état d’une obsolescence technique à court terme, mais plus largement des limites des modèles EAD utilisés par l’ensemble des archives sur le plan de l’interopérabilité des données.[[26]](#footnote-26)

Pour de nombreuses institutions d’archives mettant au cœur du changement numérique l’archivage électronique, la question du renouvellement des systèmes d’information se pose dans sa globalité. C’est notamment le cas aux AEN ou encore aux ACV. Face aux limites des modèles EAD, le domaine des archives s’est doté d’une nouvelle norme de description compatible avec le web sémantique, la norme « Record in Context (RiC) ».[[27]](#footnote-27)

# Méthodologie

## Revue de la littérature technique et professionnelle

En termes de source principale, dans le cadre de ce travail, elles sont essentiellement l’œuvre d’archiviste ayant un impact probant professionnellement ou académiquement tant sur la littérature scientifique que les documents institutionnels ou statistiques provenant directement des institutions et donc de professionnel de terrain.

Pour citer quelques exemples dans l’ouvrage « Les fonctions de l’archivistique contemporaine » (Couture, Carol 1999), l’éminent archiviste Carol couture aborde les questions de fonctions de la profession d’archiviste, mais également déjà une composante déontologique que l’on retrouve dans le code déontologique des archivistes notamment de l’ICA (ICA 1996) ou encore celui de l’AAS (Association des archivistes suisses 2022).

Les enjeux liés à la fonction de diffusion ont été défini à partir l’ouvrage « La valorisation des archives: Une mission, des motivations, des modalités, des collaborations Enjeux et pratiques actuels - Enjeux et pratiques actuels » (Hiraux, Mirguet 2012) et l’article « La diffusion des archives ou les 12 travaux des archivistes à l’ère du numérique » (Lemay, Klein 2012) et sur le plan des sciences participatives avec un article de Siham Alaoui (Alaoui 2021).

Pour les questions de prospectives et d’évolution de la profession d’archiviste, les ouvrages suivants ont été utilisés, soit pour les aspects numériques et de numérisation l’ouvrage « Les chantiers du numérique : Dématérialisation des archives et métiers de l’archiviste » (Hiraux, Mirguet, Delpierre 2012) et plus globalement les ouvrages « *Archivistes de 2030 »* (Servais, Mirguet 2015) *et « L’archiviste dans quinze ans: Nouvelles attentes, nouvelles responsabilités, nouveaux défis »*(Servais, Mirguet 2016)*.*

## Utilisation de données en open data

Les données utilisées dans le cadre de mon étude proviennent du site Conférence des directrices et directeurs d’Archives suisses (CDA-ADK)[[28]](#footnote-28). Il s’agit des statistiques d’archives suisses que j’ai ensuite restreintes aux institutions romandes. Sur la base d’indicateurs statistiques fournis directement par les institutions en open data, il a été possible de faire un état de la question de la diffusion des archives entre 2013 et 2024 en Suisse romande. Mais également de mettre en exergue les tendances et anomalies.

## Enquête

Afin de faire le lien entre ce travail avec le terrain, un questionnaire Qualtrics a été envoyé aux directions des archives d’états/cantonales romandes. Ce questionnaire constitué d’une vingtaine de questions permet un retour global de professionnelle à des positions stratégiques ou de directions de l’ensemble des institutions visé en Suisse romande.

## Création d’une base de données

Afin de conserver de manière structurée les données issues de mes recherches, une base de données est utilisée à cette fin. Il s’agit d’une base SQL créée à partir du système de gestion de base de données SQLite et de DBeaver.

En prélude de sa création, une modélisation des besoins liés à ce mémoire a été effectuée sur le cœur de ses données produites. Ces besoins ont ensuite fait l’objet d’une modélisation UML, ou entité-relation.

[Modélisation UML]

L’objet de cette recherche est essentiellement tourné sur l’étude stratégique des institutions d’archives sur le plan des pratiques numériques. Au cœur de cette modélisation, nous retrouvons donc les institutions qui ont répondu ou mis à disposition des données sous la forme d’entretiens ou de questionnaires. Ces données apparaissent donc dans ma modélisation et sont mises en relation entre elle et les institutions romandes.

## Visualisation de données

Les données de ce mémoire comportent un nombre important d’indicateurs statistiques permettant de décrire factuellement la situation des institutions d’archives romandes dans le cas des transformations numériques passées ou en cours de mise en œuvre. Afin d’en faciliter l’analyse, l’utilisation de visualisation de données sera une part importante de ce mémoire.

Une phase de curation de données a été effectuée sur les données sources, parfois automatisée en partie avec du code Python. Il s’agit d’une tâche fondamentale afin de rendre exploitable les données recueillies en open data durant ce mémoire.

La majorité des visualisations à était générer en langage python, principalement utilisé avec des librairies classiques *« pandas »*, *« matplotlib.pyplot »* et *« seaborn ».* Ces librairies python sont toutes les trois utilisées pour la création de visualisation à partir de données statistiques.

## Cas pratique

Afin d’ajouter une part d’expérience personnelle dans ce mémoire, des liens seront effectués entre les besoins du domaine des archives sur le plan de la transformation numérique et des compétences enseignées par les humanités numériques dans les universités romandes. Ces compétences ayant été validées dans le cadre de mes cours, une synthèse des résultats de mini-projet effectué dans ce cadre viendra renforcer par un aspect de mise en œuvre des enjeux actuels en archives.

Parmi ces exemples et cas pratiques de mini-projet l’on retrouvera la création d’une Ontologie, son alignement avec d’autres ontologies grâce à l’outil « Protégé ». Un exemple de traitement de fonds d’image issu d’un fonds d’archives sera également effectué avec des méthodes de Deep Learning et leur indexation automatique ainsi qu’un exemple de mise en œuvre d’implémentation de LLM.

Ces méthodes découlant des humanités numériques, elles sont un bon moyen de faire le lien avec les stratégies numériques du domaine, que cela soit la *« Stratégie numérique suisse »* ou le *« Livre blanc – L’apprentissage automatique dans les archives : l’indexation en profondeur au service de l’accès aux archives ».* Cela permet également de montrer les compétences et également les futures attentes des chercheurs en SHS ayant aujourd’hui accès à des cours en humanités numériques. Ces méthodes sont tant un enjeu pour le domaine des archives que pour le monde de la recherche qui côtoie également les instituions d’archives.

# Traitement du sujet

## Étude sur l’impact des évolutions numériques en archives

En prélude à cette enquête, les représentants des institutions cantonales romandes qui ont répondu à mes questions ont tous affirmé connaitre les humanités numériques[[29]](#footnote-29), ce qui tend à conforter la légitimité d’étudier les évolutions numériques du domaine des archives dans ce cadre, au même titre qu’un data-scientifique le fait pour les sciences économiques par exemple.

Les archives et les archivistes jouent un rôle pour la recherche en sciences humaines en étant les gardiens du patrimoine et de sources historiques fiables, c’est l’une des raisons pour laquelle l’archiviste est au cœur de cette étude. Même si nous n’avons pas d’indicateur pour l’ensemble de la profession, le site de la conférence des directrices et directeurs d’Archives suisse (CDA) souligne la proximité de ceux-ci avec les disciplines des Lettres et sciences humaines.

**Figure 1: Répartition des titres d'étude chez les directrices et directeurs d’archives**

Une image contenant texte, capture d’écran, cercle, Police

Le contenu généré par l’IA peut être incorrect.

Les titres d’études des directeurs et directrices d’institutions d’archives en Suisse sont majoritairement issus des sciences humaines et sociales (SHS) dont beaucoup proviennent des facultés d’histoire. À noter qu’un tiers ont un doctorat, soit un lien privilégié avec la recherche scientifique.[[30]](#footnote-30)

À ce titre, les humanités numériques sont une discipline particulièrement intéressante dans le cadre de cette recherche, car les archives sont souvent un passage inévitable pour les chercheurs des sciences humaines et sociales (SHS). Les humanités numériques sont donc adéquates pour fonder un discours sur les évolutions numériques touchant l’ensemble des GLAM. Cela est particulièrement pertinent pour les archives tant du point de vue du chercheur en SHS que du professionnel des archives, soit à la croisée des préoccupations scientifiques et professionnelles.

Cette enquête vise à définir des tendances se détachant au sein des archives cantonales romandes et des Archives fédérales suisses (AFS), intéressantes pour sa mise en œuvre récente de la stratégie numérique suisse entre 2019 et 2025. Nous disposant à ce titre de statistiques nous permettant d’en étudier l’impact sur la diffusion entre 2013 et 2024 et de comparer les différentes institutions cantonales et fédérales en Suisse à l’aide d’indicateurs statistiques définis et comparables. À noter que la mission de ces institutions est similaire, mais pas leurs moyens ou leur taille. Il s’agira donc de définir des tendances entre les AFS ayant mis en œuvre une stratégie numérique globale et d’autres institutions avec une maturité et des moyens variables sur le plan numérique. L’interprétation de ces tendances sera de nature qualitative.

### Indicateurs liés aux compétences numériques des archivistes

En premier lieu, il est intéressant d’estimer la place des compétences dans le profil d’archiviste actuellement en poste.

Une image contenant diagramme, ligne, cercle, origami

Le contenu généré par l’IA peut être incorrect.Il s’agit d’une approximation, mais la visualisation ci-dessous nous permet de placer le « numérique » qui devra être défini plus en détail, dans un environnement de compétences. Ces catégories généralistes sont inspirées de la littérature professionnelle[[31]](#footnote-31).

**Figure 2 : Estimation moyenne de l’importance de compétences clés pour l’archiviste (score sur 100)**

Les répondants estiment en moyenne que ces compétences numériques sont aujourd’hui « importantes »[[32]](#footnote-32), les plaçant juste en dessous des compétences rédactionnelles, analytiques, d’éthiques et relationnelles qui ont toujours fait des personnes issues des Lettres des profils tout indiqués pour des reconversions dans le domaine des archives.

Une image contenant texte, capture d’écran, cercle, diagramme

Le contenu généré par l’IA peut être incorrect.Dans un contexte où le numérique prend de plus en plus de place dans le domaine des archives, il est intéressant de questionner la place de l’archiviste dans cet environnement. Un indicateur adéquat est celui de la responsabilité de l’archiviste au sein de projets numériques en archives.

**Figure 3: Responsabilité de la gestion des projets numériques dans les services d'archives romands**

Comme nous pouvons le constater dans la visualisation ci-dessus, la responsabilité dans la gestion de projets numériques dans les institutions cantonales romandes en archives se fait de plus en plus souvent en collaboration avec le service d’informatique cantonale entre un archiviste et un informaticien (50%). Dans le cas des ACV, le chef de projet numérique est directement un informaticien sous la responsabilité du service d’archives.[[33]](#footnote-33)

L’archiviste est donc en contact avec des questions de gestion de l’information dans un environnement numérique et se doit de collaborer avec les métiers connexes. Dans un métier où la conservation et la diffusion du patrimoine s’inscrivent dans une temporalité longue, l’émergence récente des supports nativement numériques et des évolutions technologiques rapides fait du numérique un enjeu stratégique majeur pour les archives.

Dans le cadre de cette enquête le répondant des AEN Monsieur Sébastien Bischoff à très justement précisé le cadre de ces compétences :

*« Le terme de "compétences numériques" me semble très large, tout comme "métier d'archiviste". D’après moi il faudrait d'ailleurs parler "des métiers d'archiviste", l'archiviste informaticien ayant d'autres compétences que le/la gestionnaire d'information, records manager, salle de lecture, ou resp. de la communication (expositions) des archives. Ceci étant dit, je pense qu'il est important d'avoir des notions de formats de fichiers; d'indexation et de classement; de métadonnées au sens large; mais aussi une compréhension des enjeux de sécurité et traçabilité. Au quotidien, une maîtrise d'Excel (ou autre logiciel de tableur) voire idéalement des notions de bases de données et leur export/import me semblent importantes. »*

Il est donc pertinent d’interroger les compétences numériques jugées les plus utiles dans ce domaine, afin de préciser la notion même de « numérique » auprès de l’ensemble des répondants à l’enquête.[[34]](#footnote-34).

**Figure 4: Compétences numériques les plus utiles en archives**

Une image contenant texte, capture d’écran, nombre, Police

Le contenu généré par l’IA peut être incorrect.

Au-delà de pouvoir faire un parallèle réjouissant avec les compétences enseignées par les Humanités numériques. Il y a une pluralité de compétences perçues comme importantes pour l’archiviste, devant s’occuper d’évolution des systèmes d’information, de numérisation, de reconnaissance automatique ou encore d’archivages électroniques. Le numérique dans sa globalité est clairement devenu un enjeu critique pour garantir des sources à l’histoire de demain.

### Indicateurs d’infrastructures techniques en archives

La transformation numérique en archives est un enjeu majeur en parallèle de celle des administrations et plus largement celle de notre société afin de pouvoir conserver et exploité une production documentaire comprenant des médiums documentaires numérique. Comme nous l’avons vu dans la partie « contexte » de ce mémoire, la numérisation est souvent un prélude à une diffusion numérique pour de nombreuses archives physiques.

**Figure 5: Infrastructure de numérisation dans les Archives cantonales/d'États romands**

Une image contenant texte, carte

Le contenu généré par l’IA peut être incorrect.

Comme nous pouvons le constater sur la cartographie, les archives cantonales/d’états se sont dans leurs majorités équipées de matériels de numérisation au moins pour les documents les plus souvent représentés en archives. Les moyens alloués à la numérisation d’archives fluctuent selon les institutions. Par exemple si les Archives cantonales vaudoises disposes de photographes professionnelles et d’un pôle de numérisation tout équipé, les Archives du canton du Jura (ArCJ) confient les tâches de numérisation à un prestataire privé.

Le répondant des ArCJ à confirmer dans le cadre de cette enquête que même si cette la numérisation a été délégué à des prestataires, une politique de numérisation et des projets de numérisation ont bien lieu est sont encadré par une stratégie.

L'Office de la culture jurassien a notamment lancé un programme de numérisation en 2012, appelé Sigma, et s'est doté d'une politique de numérisation approuvée par le Gouvernement en 2014. La priorité des numérisations est établie sur la base de critères définis dans le cadre de cette politique[[35]](#footnote-35).

Actuellement les institutions d’Archives en Suisse romande ne dispose pas de moyens de numérisation similaire à ceux déployé dans le cadre de la stratégie numérique suisse aux AFS. Si les institutions romandes ne vont pas jusqu’à une numérisation offerte à la demande pour facilité la consultation par défaut en numérique, elles disposent pour la plupart de moyens pour numériser les documents les plus rependu en archives.

Qu’ils soient numérisés ou nés numériques, les documents numériques ont besoin d’un environnement numérique pour exister et être conservé de manière à garantir sa pérennisation, son authenticité, son intégrité et finalement dans le cadre de leur diffusion, leur utilisabilité. En archives cela passe par des systèmes d’archivage électronique (SAE). Il est donc intéressant dans le cadre de cette enquête de pouvoir donner l’état actuelle des projets d’archivage électronique dans les institutions d’archives romandes.

**Figure 6: Archivage électronique dans les institutions d'archives cantonales**

Une image contenant carte, texte, atlas

Le contenu généré par l’IA peut être incorrect.

Comme observer toutes les institutions cantonales d’archives sont concernées par l’archivage électronique soit déjà déployé ou en projet. Il s’agit donc d’un des gros « chantier » du moment avec beaucoup d’enjeux sur le long terme. Certains cantons le système d’archivage électronique n’est qu’une partie d’un renouvellement plus large des systèmes d’information des archives ou plus largement de l’administration cantonales.

### Indicateurs liés aux moyens en archives

Nous pouvons également questionner l’évolution des moyens alloués aux archivistes. En effet dans le cadre de vaste projet de transformation numérique et de changement important pour la profession d’archiviste, un projet se base sur des objectifs, une temporalité définie, mais également des moyens.

À ce titre, les ressources financières alloué ont de tel projet été capitale pour remplir ses objectifs. Dans le cadre de la stratégie numérique des AFS, je me suis donc intéressé aux dépenses de celle-ci à travers le temps à partir d’un jeu de donnée provenant de « Lindas » concernant les années 1951-2014, que j’ai compléter par une recherche sur les comptes de la confédération pour les années 2015-2024.

**Figure 7: Evolution des dépenses des archives fédérales suisses par années (1951-2024)**

Une image contenant texte, Tracé, ligne, diagramme

Le contenu généré par l’IA peut être incorrect.

Le constat est assez marquant, il y a une très forte augmentation des dépenses de plus en plus marqué au rythme de la numérisation de l’administration. Cette augmentation est directement lié à l’augmentation globale de la taille de l’administration et de ses besoins spécifiques sur la gestion de l’information dont il est clairement fait référence dans les rapports de la confédération[[36]](#footnote-36)

Pour l’année 2024, il y est notamment notifié les priorités stratégiques relative au compte en cours. Soit compléter les archives numériques en développant l’accès en ligne aux Archives fédérales. « Développer l’archivage numérique (nouveaux types d’information, évolutions technologiques, accroissement de la rentabilité). Assister et conseiller l’administration fédérale dans la gestion de l’information (y c. GEVER) afin de contribuer à une gestion efficace et garante de la sécurité du droit. Mais également préparer la fin de la reprise des documents papier par les Archives fédérales suisses (AFS) »[[37]](#footnote-37)

Dans ce cadre de transformation numérique, il est intéressant de se poser la question de l’attribution de ces moyens, est-il question d’accroitre l’infrastructure des AFS ou cette transformation nécessite-t-elle un accroissement des moyens humains. A ce titre le personnel des archives est également un excellent indicateur pour définir les moyens puisqu’ils sont au cœur de la mise en œuvre de la transformation numérique et des stratégies définies. L’indicateur du « Nombre d’employés à durée indéterminée et à temps plein »[[38]](#footnote-38) inclut l’ensemble des collaborateurs du service d’archives, même si ceux-ci n’accomplissent pas des tâches de nature strictement archivistique. La variation des employés en archives sur les dernières années est la suivante.

**Figure 8: Variation des employés à durée indéterminée et temps plein par année en archives**

Une image contenant texte, ligne, Tracé, diagramme

Le contenu généré par l’IA peut être incorrect.

Cette visualisation montre que les postes fixes dans les archives publiques fluctuent très peu depuis plus de dix ans. Nous pouvons le confirmer avec des tests statistiques afin de déterminer si ces données sont significatives ou issues d’une distribution stable en créant un tableau de contingence de ces variables et calculant ensuite des tests d’indépendance, dont le détail est disponible dans le repository Github. Le Chi2 (9.68) et la p-value (1) ainsi qu’un test de Cramer (0.030) ne montrent aucune relation significative, confirmant une stabilité globale du personnel à travers le temps dans nos institutions d’archives romandes[[39]](#footnote-39).

Nous pouvons également mettre cela en pondération avec les postes d’archiviste par rapport à la taille de l’administration concernée. En effet les archives publiques ont pour principale mission de gérer les documents produits par leurs propres administrations comme préciser dans la loi sur l’archivage (LArch) (voir : 2.2 Cadre légal.). Il y a donc un lien et une pondération potentiellement intéressante entre les deux.

**Figure 9: Ratio d’employés en archives pour 1000 employés de l’administration**

Une image contenant capture d’écran, Rectangle, texte, carré

Le contenu généré par l’IA peut être incorrect.

Stratégie numérique suisse

Un premier constat concerne les données extrêmes telles que les AEV et qui sont facilement explicable par la taille de l’administration. Les ratios les plus importants entre le personnel en archives pour 1000 employés de l’administration concernent les administrations les plus petites.

Il s’agit donc d’un biais, mais montre qu'il y a potentiellement un nombre minimal d'archivistes nécessaire pour accomplir leur mission indépendamment de la taille de l'administration concernée. À l’inverse, les ratios les moins importants pour des administrations de taille élevée peuvent induire des gains d’échelle pour gérer la production documentaire d’une administration. À tâche égale, l’on peut également constater là aussi des différences importantes en termes de moyen humain entre les différents services d’archives en Suisse romande.

Du point de vue de nos questions de recherches pour déterminer l’impact du numérique sur le métier, notamment pendant la durée de mise en œuvre de la Stratégie numérique suisse aux AFS entre 2021 et 2024. Nous pouvons constater une stabilité du personnel durant cette période.

Autrement dit, même face à une transformation numérique importante, les archivistes accomplissent davantage avec les mêmes moyens internes, ou alors les institutions s’appuient potentiellement sur des ressources externes. Il pourrait en effet être envisageable qu’elle s’appuie sur des collaborations avec d’autres services de l’État comme évoqué dans les enjeux du domaine[[40]](#footnote-40) ou encore possiblement en partie par une délégation des compétences à des prestataires privés.

### Indicateurs liés aux collaborations institutionnelles

Face à une stagnation des moyens des institutions d’archives, la question est comment les archives s’adaptent, crée-t-elle des collaboration ou pôle de compétence ? Lors de mon enquête auprès des institutions romandes, il m’a étais possible de poser la question.

**Figure 10: Types de collaborations institutionnelles déclarées par les 6 Archives cantonales**

Une image contenant texte, capture d’écran, diagramme, ligne

Le contenu généré par l’IA peut être incorrect.

Il ne s’agit probablement pas d’une liste exhaustive, mais ces résultats nous donnent des tendances. Comme nous pouvons le constater les projets numériques sont souvent effectués en collaboration avec d’autres institutions. En premier lieu, toutes les Archives cantonales romandes déclare collaborer directement avec leurs services d’informatiques cantonales pour les aspects numériques.

Pratiquement toutes les Archives cantonales font appel à des prestataires dans le cadre de délégation de compétences spécialisé ou simplement des prestations de service. Des collaborations institutionnelles notamment avec les GLAM semble être également une pratique courante pour 4 institutions sur les 6 en Suisse romande. Cela fait écho aux enjeux défini en début de ce mémoire.

L’on retrouve également dans une moindre mesure des collaborations avec le milieu académique, des groupes de travail ou associatif. Le Kost-Ceco comme pôle de compétence pour l’archivage électronique est également cité.

### Indicateurs liés à la gestion documentaire

L’outil au cœur du métier d’archiviste est l’inventaire. Il est donc intéressant dans un contexte ou l’évolution du métier d’archiviste d’interroger la progression documentaire aux seins de nos institutions romandes à travers le temps.

**Figure 11: Volume total des fonds d'archives traités (papier), en mètres linéaires, par année**

Une image contenant texte, ligne, Tracé, capture d’écran

Le contenu généré par l’IA peut être incorrect.

Comme nous pouvons le voir globalement la situation est assez stable, si certain élément devrait être investigué plus longuement, l’on peut déjà observer une augmentation significative de 10,1% des fonds papier traités en 2023-2024 aux AFS. Il s’agit d’un élément rassurant qui met en exergue que le traitement des arriérés ou versement d’archives papier n’est pas impacté par la mise en œuvre de la stratégie suisse et cela avec des moyens humains qui sont resté stable au sein de l’institution.

L’on peut effectivement constater que le traitement d’archives hybride ou l’archives papier et l’archives numériques coexiste dans les services de gestion d’archives et semble dans le cas des AFS être même en augmentation durant la mise en œuvre de la stratégie numériques suisse.

### Indicateurs liés à l’accès aux archives

Un autre indicateur intéressant concerne les salles de lecture. Après tout, une interrogation que nous pourrions avoir avec une stratégie du *« Tout numérique »* et *« une consultation analogique qui deviendrait l’exception »* comme c’est le cas avec la stratégie numérique suisse[[41]](#footnote-41), c’est de voir disparaitre les salles de consultation, soit un contact direct entre l’usager et l’archiviste.

**Figure 12: Évolution du nombre de places dans les salles de lecture en archives**

Une image contenant texte, ligne, diagramme, Tracé

Le contenu généré par l’IA peut être incorrect.

La tendance générale est actuellement stable ou à la baisse en termes de place dans les salles de lectures. Il faut noter une baisse notable aux archives de l’état de Fribourg et aux Archives fédérales suisses. Pour les AFS, une mise à disposition par défaut en numérique tend à logiquement induire une baisse raisonnable (24,4%) des prestations en salle de consultations[[42]](#footnote-42) comme envisagé dans la mise en œuvre de la stratégie numérique suisse[[43]](#footnote-43).

Un autre indicateur intéressant est celui des heures d’ouverture, qui nous permet de juger de l’accessibilité réelle des archives en salle de lecture aux chercheurs. En effet comme le mentionne Sébastien Bischoff des Archives de l’état de Neuchâtel (AEN). *« […] la matérialité dans l'analyse d'une source/d'un document peut selon la recherche avoir toute son importance : les encres, reliures, traces d'usages, mais aussi la composition des cahiers sont plus difficiles à analyser dans une version numérisée, sans même parler de la valeur probatoire ou l'authenticité. »[[44]](#footnote-44)*. La consultation en salle de lecture devrait donc toujours être garantie aux chercheurs. Leurs heures d’ouverture restent en définitive un indicateur significatif dans la qualité de la diffusion d’un service d’archives et sa relation aux usagers.

**Figure 13: Nombre d'heures d'ouverture par semaine de la salle de lecture**

Une image contenant texte, ligne, Tracé, diagramme

Le contenu généré par l’IA peut être incorrect.

Nous pouvons constater mis à part pour des périodes exceptionnelles tels que le covid19, une faible variation des heures d’ouverture aux usagers. Cela garantit donc la qualité du service offert aux chercheurs dans la mesure où ces heures d’ouverture garantissent une aide et un accompagnement dans les recherches et l’accès physique aux sources, nécessaire pour de nombreuses recherches.

Cette prestation est à l’évidence toujours parfaitement garantie par les AFS, y compris depuis la mise en œuvre de leurs stratégies numérique suisse, ce qui est rassurant pour pérenniser un contact direct avec l’archiviste tout en proposant plus de moyens d’accès avec le numérique. Il est à ce titre intéressant de regarder l’évolution du nombre d’usagers en salle de lecture.

**Figure 14: Nombre d'usagers uniques en salle de lecture par année**

Une image contenant texte, diagramme, Tracé, ligne

Le contenu généré par l’IA peut être incorrect.

Nous pouvons observer que mis à part des fluctuations pendant la période du covid19, la fréquentation des salles de lecture des institutions d’archives est relativement stable pour l’ensemble des institutions.

Un autre indicateur similaire, mais relatif cette fois aux visiteurs des inventaires en lignes. Cette valeur décrit le nombre d’adresses IP à partir desquelles au moins une requête est effectuée. Il n’est pas possible de savoir si un ordinateur est utilisé par plusieurs personnes. Si les ordinateurs dans un réseau d’entreprise utilisent un proxy pour l’accès à Internet, toutes les requêtes paraissent provenir du même ordinateur.

**Figure 15: Nombre de visiteurs uniques utilisant les inventaires en ligne par année**

Une image contenant texte, diagramme, ligne, Tracé

Le contenu généré par l’IA peut être incorrect.

Au niveau des AFS, il est raisonnable de dire que la stratégie numérique employée aux AFS n’a pas d’impact particulièrement significatif sur le plan du nombre d’usagers uniques en salle de lecture ou des visiteurs uniques des inventaires en lignes. En sommes, la stratégie numérique suisse semble apporter plutôt un saut qualitatif à la diffusion des archives plutôt que quantitatif.

Une autre institution se détache cependant fortement celle des ACV. Après quelques recherches l’explication la plus probable trouvée dans le rapport d’activités 2022 est la création d’une plateforme temporaire à partir du logiciel libre AtoM au mois de novembre 2022, dans le but de mettre en ligne une petite partie des sources les plus consultées. Les besoins du public s’étant manifestés de manière plus fréquente, notamment durant la pandémie de Covid 19, elle permet la consultation en ligne des registres paroissiaux et des registres de l'état civil[[45]](#footnote-45). L’augmentation de l’offre en ligne permet aux ACV une augmentation notable des visiteurs de leurs inventaires en ligne de 59% pendant l’année 2022-2023 et 134,4% entre 2023-2024[[46]](#footnote-46). Cela montre clairement l’intérêt des usagers pour un accès numérique aux archives.

Cette observation nous amène au constat d’une augmentation des visiteurs sur le portail en lien avec l’augmentation de l’offre de ressource disponible en ligne. À ce titre un autre indicateur sur le nombre d’articles recherchables en lignes est intéressant afin de déterminer l’augmentation de cette offre.

**Figure 16: Nombre d'articles recherchables en ligne par année**

Une image contenant texte, ligne, Tracé, diagramme

Le contenu généré par l’IA peut être incorrect.

Nous pouvons noter à travers cette visualisation que la mise en œuvre de la stratégie numérique suisse aux AFS s’est dès ses prémices concentrée sur l’augmentation de l’offre d’accès aux ressources en ligne. Cette augmentation de 1,4 million de dossiers supplémentaires a pu être diffusée grâce à une phase d’anonymisation des métadonnées sensibles figurant dans les notices de description. Cette phase d’automatisation était prévue par la stratégie numérique suisse et a été automatisée par l’utilisation d’outil d’intelligence artificielle. [[47]](#footnote-47)

Ce n’est pas sans rappeler les conseils du Livre blanc sur l’apprentissage profond en archives publiées par l’AAS[[48]](#footnote-48), encourageant l’utilisation de telles méthodes et qui pourrait être appliquée à peu de frais à l’ensemble des institutions afin de faciliter l’accès à des sources d’ordinaire restreintes pour des raisons de délais légaux. Sur le plan documentaire, l’usager des AFS est donc clairement gagnant montrant également une volonté de l’archiviste de faciliter l’accès en trouvant des moyens de contourner les délais légaux tout en respectant le cadre légal de consultations.

En termes de données restreintes, *d’infobésité* et de recherche documentaire, la directrice des ACV, Delphine Friedmann fait très justement état d’un risque de bais que *«[…] le numérique peut renforcer dans la mesure où il donne des résultats basés sur une masse d'informations rarement identifiables et limitées à ce qui existe en ligne, excluant de ce fait une part du savoir réel »*[[49]](#footnote-49).

En l’occurrence jusqu’à récemment aux AFS, les 1,4 million de notices[[50]](#footnote-50) restreintes jusque-là n’auraient pas été connues sans une demande de recherche auprès d’un archiviste. L’accès numérique est un facilitateur pratique, mais toutes les informations que cela soit pour des raisons légales ou encore pour les fonds d’archives pas encore inventoriés ou très sommairement traités s’en trouve invisibilisé du public, or ils sont physiquement bien présents dans les archives et connues des archivistes.

L’indicateur le plus significatif dans le cadre des statistiques des institutions d’archives en suisse[[51]](#footnote-51) est celui des unités d’archives commandées. Il y a avec cet indicateur un lien très direct avec la fonction archivistique de diffusion dans la mesure où la commande passe par un archiviste et qu’une commande répond forcément à un besoin spécifique de l’usager, il ne s’agit donc pas d’une simple visite de curiosité sur l’inventaire en ligne.

Cet indicateur donne le nombre d’unités archivistiques physiques commandées en salle de lecture au cours de l’année sous rapport. Toutes les commandes sont comptabilisées, qu’elles aient été transmises par oral, par écrit ou en ligne. Les fonds d’archives consultés en ligne ne sont pas comptabilisés, même si cette consultation a lieu dans la salle de lecture. Il s’agit d’un très bon indicateur sur le dynamisme tant de la salle de lecture qui doit recevoir et préparer les archives commandées que globalement de la fonction archivistique de diffusion.

**Figure 17: Évolution du nombre d'unités physiques commandées en salle de lecture**

Une image contenant texte, ligne, diagramme, Tracé

Le contenu généré par l’IA peut être incorrect.

Comme nous pouvons le constater, les AFS se distinguent très fortement du reste des institutions. Même si son rôle d’archives nationales favorise d’être bien placé dans ce classement, son évolution est très intéressante par rapport à ses propres résultats depuis 2021, date de la mise en œuvre de la stratégie numérique suisse. Il est donc intéressant de développer particulièrement cet aspect. À cette fin des tests statistiques sont intéressants afin de vérifier que cette évolution n’est pas due au hasard.

D’abord même si c’est un constat logique, le test d’indépendance chi2 (124554.2) montre une relation statistiquement significative entre les variables d’unités d’archives commandées par années et les institutions d’archives. Une valeur aussi élevée montre également que certains services ont connu des variations très importantes certaines années. Le test d’inertie V de Cramer (0.131) permettant une interprétation normalisée indépendamment de la taille du tableau, montre un lien modéré entre ces variables.

Cela signifie que le comportement de consultation en archives varie modérément selon le temps, mais pas de façon aléatoire. Autrement dit, il y a des évolutions temporelles structurées qui ont des causes avec de multiples facteurs sous-jacents qui sont à développer. Ce facteur peut-être notamment un nouveau système d’information, une infrastructure de numérisation à la demande, mais également un élément externe tel que la covid19.

L’indépendance étant écarté, à partir de ces tests, le calcule de résidus du test d’indépendance semble intéressant afin de mettre en exergue ce qui sort de l’ordinaire en déterminant l’écart avec le résultat attendu statistiquement. C’est pour résumé une manière de visualiser l’écart entre un modèle prédictif si ces deux variables étaient indépendantes et la réalité de ce qui sait réellement produit avec des changements tels que des choix stratégiques, un renouvellement des systèmes d’information ou une politique de numérisation par exemple.

**Figure 18: Analyse bivariée du nombre d'unités physiques commandées en salle de lecture par année en résidus pondérés du test d'indépendance**

Une image contenant capture d’écran, Rectangle, carré, Caractère coloré

Le contenu généré par l’IA peut être incorrect.

Stratégie numérique suisse

Si cette méthode statistique comporte des biais dans le cadre d’une analyse comportant un facteur temporel et des instituions aux moyens et à la taille variable. Elle nous permet ici de montrer les années et les institutions qui s’écartent fortement de la tendance attendue. C’est notamment le cas pour les ACV en 2015 avec la création d’un nouveau pôle numérique comprenant la numérisation des archives[[52]](#footnote-52).

Les AEN sont également concernés par un écart significatif en 2020. Après avoir pu enquêter sur cela auprès du directeur des AEN monsieur Bartolini, sa réponse du 19 août[[53]](#footnote-53) confirme que la hausse est imputable à la fermeture prolongée de la salle de lecture pendant le covid19 et son remplacement par une prestation de numérisation sur demande à distance.

Un autre écart significatif est celui des AFS entre 2021 et 2024, soit en corrélation avec la mise en œuvre de la stratégie numérique suisse. Sans prendre en compte les moyens mis en œuvre, cela nous permet néanmoins en regardant ces écarts statistiques de l’attendu, de confirmer l’impact positif d’une stratégie numérique globale et cohérente dans un service d’archives pour dynamiser la diffusion des archives.

Si le numérique est partie intégrante aujourd’hui de la diffusion des archives, les répondants de l’enquête mettent en évidence un autre enjeu central. En effet tous les répondants font état de la question de moyens hétérogène entre les institutions rendant difficile l’application de la stratégie numérique suisse telle que mise en œuvre aux AFS.[[54]](#footnote-54) L’un des plus importants services d’archives concernés par cette enquête après les AFS et celui des ACV qui créa dès 2015 un pôle numérique. Sur la question d’appliquer une stratégie équivalente, sa directrice Delphine Friedmann précise : *« Au vu des coûts (de réalisation et de stockage numérique) et de la complexité que représente la rétro-numérisation, celle-ci n'est pas envisageable actuellement aux ACV, qui conservent actuellement environ 40kml d'archives. On pourrait bien sûr envisager, comme les AFS, une politique de rétro-numérisation liée à la demande, mais cela reste complexe et très coûteux. Actuellement : pas envisagé aux ACV. »[[55]](#footnote-55)*

L’application de la stratégie numérique suisse à l’ensemble des institutions cantonales ne semble pas raisonnable dans l’immédiat, mais des avancées en ce sens sont certaines pour la profession d’archiviste. Néanmoins, il est peu probable que les moyens humains évoluent pour le domaine, ce qui fait du numérique d’autant plus un impératif.

C’est en effet à mon sens en envisageant des gains de productivité ou du moins en amplifiant son impact pour l’archiviste grâce à de l’automatisation avec des IA notamment sur les tâches traditionnelles telles que la description ou l’indexation pour les fonds déjà numérisés que l’archiviste pourra déployer le plein potentiel de ces données.

Ce que j’entends par automatisation, n’est certainement pas le remplacement de l’archiviste par de l’IA, mais comme pour de nombreux outil de gestion électronique de documents au niveau du record management, l’IA est déjà implantée dans la majorité d’entre eux comme GEVER ou OneGov.

C’est en mettant en place de nouveau processus tournés vers la donnée que les archives pourront à terme se permettre un réemploi des moyens comme ce fut le cas aux AFS pour l’anonymisation des notices notamment. Il s’agit en effet d’un changement de paradigme dans les pratiques institutionnelles, mais qui restent dans la continuité des évolutions numériques de ces dernières années. C’est d’ailleurs un intérêt majeur défini par l’association des archivistes suisse (AAS) qui rédigea un Livre blanc sur « L’apprentissage automatique dans les archives : l’indexation en profondeur au service de l’accès aux archives ».[[56]](#footnote-56)

Un autre enjeu est celui de la mutualisation, une tendance depuis de nombreuses années face à la stagnation des moyens afin de trouver des synergies institutionnelles[[57]](#footnote-57). L’on peut notamment citer les archives de l’état de Bâle-Ville et du Musée cantonal des sciences naturelles, le centre de stockage interinstitutionnel cantonal de Fribourg, le rassemblement des bibliothèques, de la médiathèque et des archives aux anciens arsenaux de la Ville de Sion. Mais encore le centre « Archives et Patrimoine » à La Chaux-de-Fonds pour l’État de Neuchâtel. Ces mutualisations rendant l’accès documentaire en un seul site pourront peut-être un jour se développer sur de nombreux aspects y compris les aspects numériques.

## Entretiens

Kost-ceco ?

### Entretien de Julien Raemy, data-archiviste aux AFS

### Entretiens de Baptiste de Coulon, data-archiviste à la Fondation SAPA

Dans le cadre de ce mémoire tourné vers la transformation numériques des archives, particulièrement en lien avec la diffusion des archives, une institution présente en Suisse romande est particulièrement intéressant. Il s’agit des Archives suisses des arts de la scène ou « SAPA ». A ce titre Baptiste de Coulon[[58]](#footnote-58) est archiviste au sein de la Fondation SAPA[[59]](#footnote-59) depuis 2021.

Le parcours de Baptiste de Coulon est aussi très intéressant dans le cas des humanités numériques car il commença par des études en géographie à l’Université de Genève (UNIGE) puis un DEA en recherche à l’EHESS (Paris) suivi par un certificat en nouvelles technologies du WEB à l’UNIGE et enfin un MAS ALIS (Archival Library and Information Science) aux universités de Berne et de Lausanne. Il débuta dans le domaine des sciences en information par un stage dans les Archives de l’Institut d’architecture de l’Université de Genève en 2004.

Baptiste de Coulon travailla en tant qu’archiviste indépendant pour les communes neuchâteloises puis devint fondateur du service intercommunal d’archivage (SIAr) [[60]](#footnote-60). Il travailla ensuite en tant que responsable de l’antenne romande de l’entreprise docuteam, puis en tant que membre du conseil de direction de cette dernière. Il poursuivi par un poste chez Memoriav. Il est actuellement employé à la Fondation SAPA en tant que responsable de la plateforme numérique. Ses tâches concernent particulièrement la gestion des AIS, mais également la gestion des données numériques. Dans le cadre de son métier, il porte un intérêt important pour l’open data et l’open source.

Son temoignage est d’un interet majeur sur le plan de la transformation numérique en cours dans le domaine des archives. SAPA est en effet une fondation privée avec des besoins spécifiques pour gérer et diffuser son patrimoine, faisant d’elle un précurseur sur la mise en œuvre d’une stratégie numérique institutionnelle. Plusieurs éléments sont très intéressants à analyser à ce titre.

Dans ce cadre Baptiste de Coulon est responsable du portail performing-arts.ch[[61]](#footnote-61) qui publie à la fois des données générales sur les arts de la scène en Suisse (qui, quoi, où), mais aussi les descriptions des collections conservées par la fondation. Les descriptions des collections dans ce portail sont compatibles avec le web sémantique plus précisément les ontologies, notamment la nouvelle norme de description Records in Contexts (RiC) [[62]](#footnote-62). Les inventaires y sont exprimés sous forme de données liées (RDF).

En effet la Fondation SAPA est la successeure de plusieurs institutions qui lui ont transféré leurs collections en 2017 au moment de sa création. Il en découle une pluralité d’outil de description tel que Scope Archiv, mais également des outils moins normés comme FileMaker.

Le passage au RDF est le fruit de réflexions initiées dès 2014 au sein de la Collection suisse du théâtre (CST) qui souhaitait fusionner ses différentes bases de données et y donner accès aux publics via Internet. Ces réflexions ont été nourries notamment par des consultants externes tel que Beat Estermann spécialiste de l’open data en Suisse et membre du groupe sur l’ouverture des GLAM[[63]](#footnote-63).

Entre 2014 et 2017, Beat Estermann travailla sur des solutions et fournit au CST un rapport dans lequel il recommandait qu’avec une telle variété de bases la seule solution pour une mise en commun serait l’utilisation du RDF. À noter que cela fut énoncé avant l’arrivée du modèle RiC pour les archives. Un grand effort fut d’ailleurs fait pour modéliser à partir de là les activités des arts de la scène. Cela a donné lieu à des publications scientifiques[[64]](#footnote-64). On appelle à présent cet ensemble de données le « Panorama» chez SAPA.

Les travaux ont commencé par la migration en RDF des données contenues au sein de l’AIS Scope Archiv. Ces données étaient déjà normées selon ISAD(G) ce qui facilitait la tâche.

Scope Archiv a été créé à l’origine sur la demande des institutions publiques suisses, en particulier des Archives fédérales. Ces institutions auraient soutenu la création de la société privée éditrice de ce logiciel Scope Solutions AG. L’outil était centré autour de de la description. Toutefois, profitant d’une situation quasi-monopolistique cette entreprise a mené jusqu’à aujourd’hui des barèmes tarifaires très élevés avec des développements informatiques toujours sous licence Les institutions se sont trouvées piégées et beaucoup d’argent public a été englouti par cette société sans qu’aucune contrepartie, par exemple en termes d’ouverture du code informatique ou de soutien aux petites institutions ne soit réellement exigé. Au niveau financier, pour la Fondation SAPA il était très intéressant de remplacer rapidement cet outil si possible par des outils open source, ce chantier est à présent terminé. C’est également un élément faisant le lien avec la Stratégie numérique des AFS sur le plan de favoriser des outils open source.

Le deuxième gros travail de migration, qui lui est toujours en cours, ce sont les bases FileMaker dont les données sont moins normées. FileMaker est un système de base de données où l’on peut rajouter librement un champ. De fait, cela peut donc partir dans tous les sens. C’est à la fois un avantage et un inconvénient. Les travaux de migration depuis FileMaker ne font que commencer au sein de la Fondation SAPA.

L’ensemble de ces migrations nécessite un mapping spécifique vers le RDF de SAPA. La Fondation utilise en partie RiC pour décrire ses collections, mais également beaucoup d’autres ontologies, dont une ontologie maison[[65]](#footnote-65). RiC est donc seulement une partie de la réponse lorsque l’on passe aux données liées (RDF).

Sortir d’une logique de silos signifie créer une certaine transversalité et interopérabilité. A ce titre, SAPA effectue des *« matching »* entre nos entités RDF et celles des référentiels externes, comme Wikidata. On peut le faire au niveau des autorités, mais aussi sur les vocabulaires contrôlés[[66]](#footnote-66). Cela permet réellement une circulation entre les bases et c’est tout à fait nouveau pour le domaine des archives. On sort de la logique des silos vers la création d’ontologie et l’alignement entre plusieurs ontologie. Wikidata joue un rôle de noyau informationnel (hub) dans ce système[[67]](#footnote-67). Wikidata permet évidemment d’expérimenter aussi le crowd sourcing. Attention toutefois, le temps de travail pour réaliser ces matching n’est pas à sous-estimer.

Le passage au RDF permet d’envisager aussi des recherches fédérées à travers plusieurs bases de données L’utilisation du langage SPARQL est obligatoire pour interroger des graphes RDF. C’est un langage qui permet vraiment de redécouvrir ses données et de poser des questions complexes.

L’un des problèmes qui se pose actuellement à la Fondation SAPA est que sa base AIS et que son portail d’accès public ne sont pas séparés. Avec RDF, sans séparation strictement des deux bases, il est difficile de cacher de l’information lors sa diffusion, en particulier si on veut exposer les encodages techniques. La solution pourrait venir de l’utilisation de 2 bases dont une serait dédiée à la consultation.

Par ailleurs, ces bases sont encore des outils complexes pour les publics et même pour les archivistes. Ces outils demandent pour en tirer pleinement parti de développer un intérêt pour leurs caractéristiques techniques.

Dans le portail performing-arts.ch, il n’existe pas pour l’instant de véritables outils de médiation (type guide de recherche) pour faciliter la découverte, mais la Fondation SAPA souhaite changer cela prochainement.

Une problématique très actuelle dans la gestion des données sur les objets patrimoniaux, c’est la circulation et duplication des métadonnées. On peut compter au moins 4 lieux de stockage différents d’enregistrement de ces métadonnées : les AIP dans le dépôt d’archivage numérique, les descriptions dans l’AIS, les métadonnées embarquées dans les fichiers offertes en consultation et même les Manifest IIIF si on a choisi de mettre en place cette brique technologique. Les données liées pourraient simplifier cela, mais on attend encore les retours d’expérience pour pouvoir vraiment l’affirmer.

Par ailleurs, ce qui est primordial, c’est d’utiliser des systèmes d’identifiants pérennes (par exemple ARK) pour les notices et assurer que l’URI soit redirigé sur la bonne page. Cela est une des grandes réussites du portail performing-arts.ch. Son système d’URI lui permet de réellement de participer au monde des données liées. Et même d’y obtenir 5 étoiles[[68]](#footnote-68) !

Pour le domaine des archives une part de cette transofmation numérique est rendu possible grâce à la création d’une nouvelle norme (RiC) et de son Ontologie (RiC-O). La Fondation SAPA est précurseur face au constat de la faible visibilité des données de ses archives dans les recherches internet. SAPA tourna à présent sa stratégie et ses moyens plus sur la qualité de ses données (et leur matching avec des données externes) que sur l’ergonomie et le graphisme de son portail.

Sur ce point le passage aux données liées et à Records in Contexts peuvent apporter beaucoup d’après l’expérience de Bapiste de Coulon. Via leur exposition dans Wikidata, les données liées relatives à nos collections peuvent être ensuite réutilisées par d’autres services ou par de nouveaux publics qui n’auraient pas nécessairement pensé à utiliser des archives.

La Fondation SAPA avec sa multitude de médium notamment beaucoup d’image met également en œuvre les recommandations IIIF, en particulier l’élaboration de Manifest IIIF. Elle utilise la visionneuse IIIF Clover. Cela a permis un gain de qualité dans la diffusion, non seulement des images fixes[[69]](#footnote-69), mais également du son[[70]](#footnote-70) et bientôt des vidéos. Nous avons bénéficié en cela des conseils précieux de Julien Rémy. L’élaboration de Manifest IIIF permet d’utiliser le potentiel des visionneuses IIIF et de les partager plus facilement via leur identifiant unique. Il s’agit d’un outil de valorisation qui sort véritablement les archives d’une logique de silos.

## Les humanités numériques

### Domaine

Les Humanités numériques sont aujourd’hui une discipline ou transdiscipline à part entière basée sur l’application des technologies de l’information comme auxiliaire des Humanités ou sciences humaines. Elles sont à mon sens un moyen d’accès tant à des moyens d’implémenter des méthodes numériques et technologies de l’information au sein des archives et GLAM en général. Mais également elle permet de mieux cerner les besoins des chercheurs en SHS souvent usagers des archives dans leurs utilisations des données d’archives.

Dans le cadre de ce travail, une question de recherche porte sur l’apport du domaine et des formations en humanités numériques pour ces évolutions technologiques aux domaines du patrimoine et sera donc un axe pour la transformation numérique en archives. Il ne s’agit ici pas d’une thématique nouvelle, il est donc intéressant d’interroger la littérature scientifique à ce sujet pour nous permettre ensuite de mieux contextualiser les applications des humanités numériques à un niveau local.

Un article est particulièrement intéressant pour définir la complémentarité des sciences de l’information comprenant également les archives et des humanités numériques, celui du professeur Maklouf intitulé *« Humanités numériques & sciences de l’information : ressemblance, complémentarité et développements mutualisés »[[71]](#footnote-71).*

L’article fait état de plusieurs éléments notables entre ces deux disciplines. Leurs apparitions formelles sont d’abord relativement jeunes, aux alentours du 20e siècle. Ce sont des disciplines dotées d’une nature accueillante, diversifiée et inclusive. Tous deux en lien avec l’objet informationnel et culturel en ayant des liens privilégiés avec le patrimoine historique. Elles sont intrinsèquement liées au numérique, plaçant la donnée au centre de ces disciplines et de son environnement numérique. Ces disciplines mettent au centre de leur sujet d’étude le comportement humain en prenant compte de l’utilisation, la création, la mise à disposition des   
contenus et information ainsi qu’une pratique orientée utilisateur.[[72]](#footnote-72)

L’archivistique est une science à part entière connexe des sciences historiques, mais plutôt affiliées aux sciences de l’information. Néanmoins beaucoup d’Archivistes sont issues des Lettres et sciences humaines, le domaine des humanités numériques est un moyen pour les archives de voir arriver des archivistes ayant un bagage numérique particulièrement intéressant pour les fonctions archivistiques dans notre ère de l’information.

Mais plus encore les usagers des archives sont aussi des chercheurs des SHS qui ont de plus en plus de besoins numériques pour leurs recherches. Les humanités numériques peuvent donc être un axe important pour offrir un bagage et des compétences numériques aux archivistes et s’adapter aux nouvelles manières de consommer l’information des usagers.

### Formation

La suisse romande compte maintenant plusieurs formations en humanités numériques, l’on peut citer le Master en Humanités numériques de l’Unil, Le Master en Digital humanities de l’EPFL, le Master Patrimoine régional et humanités numériques à l’Unine, mais également le certificat de spécialisation en humanités numériques de l’Unige.

Ayant effectué le parcours du Master en Patrimoine régional et humanités numériques, mais également de nombreux cours en mobilité, en prenant en compte des plans d’étude de ces différentes formations, il est intéressant de comparer les cours proposés avec les compétences numériques évoquées notamment durant l’enquête ou les entretiens [source].

Tiré de mon expérience voici un parcours de cours particulièrement utile en archives.

|  |  |  |
| --- | --- | --- |
| Cours | Universités | Compétences utiles en archives |
| Programmation pour le Web sémantique | UNIL | Création d’une Ontologie et alignement avec d’autres ontologies, utile pour Record in Context (RiC) |
| Humanités numériques I : production et gestion des données | UNINE | Gestion des métadonnées, Base de données |
| Humanités numériques II : analyse de données | UNINE | Curation des données, utilisation du Web sémantique |
| Sciences participatives et visualisation | UNINE | Valorisation de l’information |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

L’ensemble des compétences évoqué durant l’enquête, seul l’archivage électronique, compétence très professionnelle n’est pas évoqué, mais plusieurs cours préparent à comprendre cette compétence tant sur les bases de données que sur le plan des « ingest » avec l’utilisation régulière de Github.

Comme annoncé dans les questions de recherches de ce mémoire, il est question d’interroger quels sont les apports du domaine et des formations en humanités numériques pour les évolutions technologiques liés aux domaines du patrimoine, particulièrement des archives.

L’enquête préalable à montrer que tous les répondants connaissent le domaine des humanités numérique et nous avons pu observer la proximité des archivistes avec le domaine des Lettres et sciences humaines.

Les enjeux du domaine et le besoin de certaines compétences ont aussi été abordés montrant l’intérêt direct des humanités numériques pour les professionnelles des archives. Afin d’en déterminer la plus-value, ayant moi-même travaillé en tant qu’archiviste pendant plusieurs années et issues tant d’une formation en Science de l’information que d’un présent Master en Humanités numériques. Il est adéquat pour moi de mettre en évidence le domaine des humanités numériques, ses formations en Suisse romande et les compétences et projets pouvant être appliqués aux domaines des archives.

### Projets et compétences applicatives aux Archives

Maintenant que nous avons identifié les besoins numériques métiers des archivistes provenant notamment dans le livre blanc de l’association des archivistes ou de l’enquête effectuée, il est intéressant de montrer des exemples d’application des humanités à celle-ci. Étant archiviste de profession et étudiant en humanités numériques, il m’a paru propice d’illustrer quelques applications apporter par des compétences en humanités numériques issues des cours dans les universités en Suisse romande sous la forme de mini-projet représentatif effectué durant mon cursus. Cela vise à démontrer l’apport des humanités numériques et son enseignement dans les Universités et hautes écoles en Suisse romande aux services d’archives romands, en étant un pôle de compétence pour les archives. Cela permet également de mettre en évidence ces compétence à des besoins archivistiques.

#### Web sémantique

Le Web sémantique est une compétence très intéressante enseignée en humanités numériques tant pour la recherche d’information et la constitution de corpus de données que de la création d’ontologie. L’on retrouve d’ailleurs une application directe aujourd’hui en archives avec la nouvelle norme de description archivistique « Record in Context ».

Conceptualisé par Tim Berners-Lee, le web sémantique est une extension du World Wide Web qui vise à rendre l'information lisible et compréhensible par les machines, en ajoutant du sens aux données grâce à des technologies comme RDF et OWL. Les données sous format RDF en Données liées (Linked Data) du Web sémantique se construit sur la capacité à relier des données entre elles, en utilisant des standards et des protocoles communs. Les informations sont enrichies de métadonnées qui décrivent leur signification et leurs relations avec d'autres concepts. Des vocabulaires formels, tels que OWL, sont utilisés pour représenter des connaissances, des concepts, des relations et des catégories d'entités.[[73]](#footnote-73)

En termes de technologies de l’information l’on retrouve le RDF (Resource Description Framework), un cadre standard pour structurer les informations et représenter les données sous forme de triplets (sujet-prédicat-objet).

Des Ontologies tel OWL (Web Ontology Language) sont un langage qui étend le RDF pour permettre des descriptions plus riches, incluant des relations complexes entre les classes et les propriétés. La puissance de cette technologie pour la recherche repose sur SPARQL, un langage de requête pour rechercher, ajouter, modifier ou supprimer des données dans un graphe RDF, eux même conserver dans des bases de données graphes ou « triplestore ».[[74]](#footnote-74)

Ces technologies permettent de mettre en relation des données issues de silos différents, facilitant la connexion et l'intégration de données provenant de sources diverses et hétérogènes. Elles permettent également un accès plus rapide et plus efficace en questionnant directement l’information avec des requêtes SPARQL.

Le Web sémantique est avant tout orienté machine, permettant l’accès à des informations autodescriptive sémantiquement par la nature même des triplets, désambigüisant l’information. Il s’agit du but de l’utilisation d’ontologies comme socle de description intellectuelle de l’information. D’autres application sera également possible grâce à l’accès à des inventaires RiC notamment la lecture et consultation par des intelligences artificielles permettant ainsi une aide à la recherche tout en limitant les hallucinations dû au manque de contexte de l’information donnée à l’IA.

Il ne s’agit ici d’un simple éventail préliminaire du potentielle du web sémantique pour les archives. Cette transformation de la structure de l’information est motivée par le concept des 5 étoiles de l’open data.

**Tableau 2: Concept des 5 étoiles de l'open data**

|  |  |
| --- | --- |
| ★ | Publiez vos données sur le web (peu importe leur format) avec une licence ouverte. |
| ★★ | Publiez-les en tant que données structurées (par exemple, un document Excel au lieu d’une image scannée d’un tableau). |
| ★★★ | Publiez-les dans un format ouvert et non propriétaire (par exemple, en CSV plutôt qu’en Excel). |
| ★★★★ | Utilisez des identifiants uniques « URI » pour désigner des choses dans vos données, afin que les utilisateurs puissent faire des références à celles-ci. |
| ★★★★★ | Liez vos données à d’autres données pour y ajouter du contexte. |

(5stardata 2015)

**Figure 19: Illustration du concept des 5 étoiles de Tim Berners-Lee**

Une image contenant texte, capture d’écran, Icône d’ordinateur, logiciel

Description générée automatiquement

(5stardata 2015)

Les humanités numériques sont au cœur de cette transformation d’abord dans l’intérêt de la recherche scientifique qui comme le domaine des archives a les mêmes objectives de sortir d’une logique de silos de l’information. A notre niveau local des compétences dans le Web sémantiques sont mise en pratique dans le cadre des enseignements « Humanités numériques I : Production et gestion des données » et « Humanités numériques II : Introduction à l’analyse des données » de l’Université de Neuchâtel.

L’objectif pédagogique se concentre sur la constitution de corpus en utilisant des bases de connaissances participatives telles que DBpédia et Wikidata. Il s’agit de cours très orienté sur la recherche puisqu’elle vise à former les chercheurs en SHS de demain, mais également pour beaucoup de futurs usagers des archives dans le cadre de recherche de sources. En plus de la bonne gestion de données numériques en triplet RDF, cela montre également les bonnes pratiques que les étudiants attendent dans la consultation des archives souvent indispensable à leur recherche.

Les compétences mobilisées sont notamment l’utilisation de langage de requête SPARQL, l’import des données, leurs curations par des notebook python et leurs migrations vers des bases de données relationnelles ou SQL, mais également des bases de données graphes. Le langage de requête SPARQL est très intéressant puisqu’il exploite les technologies issues du web sémantique, les données liées. SPARQL peut interroger des « endpoints » distants tels que Wikidata ou DBpedia, mais également agréger des résultats issus de plusieurs sources, et profiter d’ontologies (OWL, SKOS, RiC-O, CIDOC-CRM…).

Il s’agit d’un nouveau paradigme pour la recherche qui essaye de sortir l’information d’une logique de silos. C’est également l’objectif de l’ontologie RiC-O pour le domaine des archives notamment. Il s’agit donc d’une compétence fondamentale enseignée en humanités numériques, à maitriser pour un archiviste souhaitant sortir ses inventaires d’une logique de silos et profiter de données en open data pour enrichir des bases de connaissances telles que Wikidata par exemple.

En Suisse romande un service d’archives s’est particulièrement distingué à cet égard celui de la fondation SAPA.

#### Indexation automatique

Si en archives l’indexation de document est moins présente que chez les bibliothèques par exemple, c’est qu’une archives est principalement constitué de document unique. L’indexation automatique a donc toujours trouvé des échos importants dans la profession avec des tentatives d’innovation connu au niveau romand tel que pour le projet Panorama aux ACV.[[75]](#footnote-75)

Les récentes innovations issues de nouveaux modèles d’IA génératives ont permis un retour d’intérêt pour l’indexation automatique des documents d’archives. Le récent Livre blanc de l’association des archivistes suisse à fait la part belle à cet enjeu.

Si les documents papier ont passablement était traité dans le livre blanc, un autre support documentaire en forte croissance actuellement et qui devrait être encore plus important à l’avenir avec les infrastructures d’archivage électronique, celle des images.

En effet, comme la professeure Béatrice Joyeux-Prunel, le souligne dans son travail de recherche « Visual contagions », *« depuis une vingtaine d’années, un déluge d’images nous submerge, images fixes et animées circulant sur internet, les réseaux sociaux, nos écrans personnels comme ceux des espaces publics. Ce déferlement s’avère sans comparaison avec celui, déjà puissant, de la circulation imprimée et matérielle des images »*.

En termes de valeur secondaire des archives, l’image est un medium documentaire qui est très parlant sur l’histoire d’une époque est dont les archives ont pour objectif de conserver, mais également de diffuser dans les meilleures conditions possibles. L’arrivé de visualisateur IIIF sur les portails d’archives en sont un bon exemple.

Les images sont très présentes en archives qu’elles soient physiques ou numériques. En effet, le métier d'archiviste évolue vers une numérisation et automatisation croissante, passant obligatoirement par une rétronumérisation des archives anciennes et donc par la gestion d'images numérisées. En plus de pouvoir mettre numériquement à disposition du public les archives, le traitement des archives par des méthodes de reconnaissance automatique de caractères (OCR) ou des écritures manuscrites avec le HTR demande que les documents d’archives soient préalablement numérisés.

L'arrivée actuelle d'outil performant de traitement automatique de l'image fait de l'automatisation une priorité pour la profession symbolisée par de nombreux articles comme "Du Machine Learning aux archives" [Source] (Reyrodriguez 2021). Mais encore au niveau stratégique l'écriture d'un livre blanc sur l'apprentissage automatique dans les archives par l'Association des Archivistes suisses (Arnold et al. 2024) dont il décrit la numérisation comme un préalable.

L’utilisation de méthode permettant l’indexation ou la description automatique des images est parfaitement représentative des perspectives des humanités numériques aux archives. Le code utilisé dans le cadre de ce mini-dossier est adapté en partie du code vu dans le cours « Du Pixel aux images : une Introduction au traitement numérique en 2D » à l’Université de Genève.

Mon code vise à effectuer une détection d'objet sur des images, tout en répondant à des contraintes métiers qui sont par exemple un besoin d'automatiser le traitement à partir d'un workflow. Mais également en termes de rendu, un archiviste aura besoin d'extraire les détections sous format CSV afin qu'elles soient intégrées aux notices d'un système d'information archivistique par exemple sous la forme d'une indexation. Le code permet à n'importe quel archiviste avec des notions informatiques minimales de l'exécuter.

Les méthodes utilisées dans le cadre de ce projet sont un mélange de méthodes de deep learning ou apprentissage profond pour le traitement d'image. Pour être exactes, les méthodes fonctionnent en mode inférence uniquement, et exploitent des poids figés issus de l’entraînement supervisé de grands corpus d’images. Leurs spécifications sont les suivantes :

[Tableau]

Les modèles de détections et classifications des images seront appliqués à un fonds d'images datant des années 1900 environ et provenant de l'inventaire des archives de la commune de Milvignes (NE), membre du service intercommunal d'archives de Neuchâtel. La plateforme de diffusion des inventaires est gérée par docuteam SA, entreprise de gestion d'archives qui m'a employé de 2016 à 2021. Les images ont été numérisées par mes soins en format de conservation (TIFF) et archivées électroniquement dans un repository de préservation numérique. Les images sont également disponibles librement en téléchargement dans un format dégradé de diffusion (JPEG).

**Figure 20: Images numérisées de la commune de Milvignes**

Une image contenant arbre, Papier photographique, collage, maison

Le contenu généré par l’IA peut être incorrect.

Les résultats appliqués à un fonds d'images anciennes ont d'abord été mitigés. Avec des résultats à seulement 33% de détection correcte les modèles SSD MobileNet v2 (TensorFlow Hub) et ViT (Vision Transformer) ne sont pas adapté a mon fonds d'images. Le code a parfaitement fonctionné et le traitement de détection et classification a parfaitement été appliqué sur le fonds d'images. Cependant les images utilisées mettent en évidence certaines limitations de cette méthode.

La première provient probablement des images elles-mêmes, il s'agit de JPEG utilisé pour la diffusion sur le web avec une faible résolution. Les modèles se basant sur les pixels de l'image pour détecter les objets, la qualité de l'image influence le résultat final.

Deuxièmement les modèles utilisés ont été entrainés sur des images récentes, cela peut être observé notamment avec les classes "COCO". Le modèle fonctionnant par inférence peu importe l'image, il y a un taux d'erreur plus important sur des images anciennes. Une première solution a été de pallier ce biais en enlevant des classes sans amélioration probante. Néanmoins cela reste un biais inhérent à ces modèles. En effet, une étude datant de 2021 utilisant le modèle SSD MobileNet v2 effectué sur une base de données de 50000 images confirme un taux de précisions variant entre 32 et 47% environ (Cohen et al. 2021) soit dans la marge d'erreur de mon résultat de 33% de précision.

À ce stade avec les résultats de ces 2 premiers modèles très modestes, un test de modèle plus récent avec un LLM multimodal "llama-4-maverick" m’a paru intéressant. Ce modèle d'un tout autre type que les deux précédents est adapté tant au texte qu'aux images. Son point fort est ça polyvalence tant sur le plan des supports qu'elles soient textuelles ou sous forme d'images, mais également pour sa compréhension multilingue comprenant le français.

Comme il s'agit d'un LLM fonctionnant via une API importée directement dans le code, cela permet une plus grande liberté d'interaction avec l'utilisation d'un prompt. Le prompt est capital pour la qualité du résultat, je me suis ici limité à une indexation des images soit une liste de mots-clés représentant les détections d'objets.

Le prompt utilisé est le suivant :

"Vous êtes expert en indexation d'archives photographiques. Votre tâche consiste à effectuer une indexation descriptive de 3 à 5 mots-clés en français pour chaque image. Basez votre analyse uniquement sur le contenu visuel de l'image ; ne formulez aucune hypothèse et n'incluez aucune information supplémentaire. Formulez votre réponse par des mots-clés d'indexation. Exemple : mot-clé1, mot-clé2, mot-clé3. Analysez maintenant l'image fournie et générez les mots-clés d'indexation."

Après analyse des résultats, Llama-4 obtient d'excellents résultats avec 88% de détection parfaitement correcte. Il faut également faire ressortir la nature des erreurs par Llama-4, soit la détection d'un "banc" pour décrire des enfants assis sur un escalier et une "église", le bâtiment en question est en fait un château avec un clocher de chapelle. Il est probable qu'un opérateur humain aurait également pu commettre ces erreurs.

Une image contenant texte, capture d’écran, diagramme, Rectangle

Le contenu généré par l’IA peut être incorrect.

Ces résultats sont significatifs dans le sens où ils sont en mesure d'égaler un opérateur humain sans vérification systématique des résultats des détections d'objets. Il s’agit d’imprécisions qui ne sont pas nécessairement rédhibitoires selon l’objectif de l’utilisation. En effet pour les images sont très présentes dans les fonds d’archives, mais ne sont pas forcément décris en détail à la « pièce », souvent la description s’arrête à un titre relativement descriptif, parfois plus détailler au niveau du « dossier ».

L’utilisation d’une indexation ou d’une description automatique par un LLM d’un grand nombre d’images permettrait donc d’offrir une description minimale au niveau de la pièce ou d’enrichir celle de l’archiviste afin d’augmenter les points d’entrées de l’information aux notices à des fins de recherche.

Il s’agit dans ce cas d’un mini projet appliqué à des images d’archives, mais qui est intéressant dans l’optique de considérer les applications des humanités numériques au domaine du patrimoine et particulièrement aux archives. De plus il permet d’exercer un avis critique sur les différents modèles et de constater l’évolution rapide et le gain de performance entre des modèles entrainés il y a quelques années et un LLM « grand public » récent.

#### Intelligence artificielle

Aujourd’hui, il serait difficile de parler de la transformation numérique dans le cadre de la diffusion des archives sans aborder la question de l’intelligence artificielle. Le Machine Learning sous toutes ses déclinaisons tel que les Transformers, LLM, Deep Learning ont toujours été un outil très important dans l’établissement de la discipline des humanités numériques, attaché à automatiser le traitement automatiquement de données massives.

Comme nous l’avons ultérieurement, les tâches en archives nécessitent un haut niveau de précision et fiabilité dans la plupart des tâches. Tant le Livre blanc de l’association des archivistes suisses que l’exemple de la stratégie numérique suisse montrent l’importance de ces technologies dans la transformation numérique du domaines.

# Analyse et mise en perspective

# Conclusion

En conclusion même s’il y aurait encore beaucoup à dire et de nombreux indicateurs à étudier, durant cette enquête nous avons pu répondre aux questions de recherches en mettant en exergue des tendances. D’abord en nous demandant quel est l’impact du numérique sur le métier d’Archiviste. C’est-à-dire des compétences numériques devenues très importantes voir essentielles aujourd’hui pour des tâches très diverses, allant de la gestion des données, à la numérisation et l’archivage électronique. Le numérique est dans le fond devenu transversal à l’ensemble du métier d’archiviste.

Sans pour autant faire de l’archiviste un expert numérique, il doit s’adapter à l’image des archives elles-mêmes, devenant de plus en plus hybride entre médium physique et numérique. Les humanités numériques sont donc aujourd’hui une discipline de choix pour les humanistes souhaitant se reconvertir professionnellement dans le domaine du patrimoine et des archives.

Ensuite cette enquête nous a permis d’entrevoir comment le numérique influence la fonction archivistique de diffusion. Le numérique permet sans le moindre doute de dynamiser la diffusion des archives. Le web et les technologies de l’information sont aujourd’hui le vecteur principal de diffusion des archives. L’arrivée aujourd’hui de nouvelles normes telles que « RiC » et des technologies de l’information telles que le web sémantique, des portails avec visionneuse « iiif » et la généralisation de la numérisation permettent un saut qualitatif d’accès aux sources pour les chercheurs.

Enfin l’enquête nous a permis d’observer, notamment grâce aux Archives fédérales suisses (AFS), quelles stratégies le métier développe pour s’adapter à ces évolutions numériques. La stratégie numérique suisse montre un impact initial positif pour les AFS sans pour autant fermer la porte à l’accès physique.

Néanmoins la question sur le long terme dépendra des moyens alloués pour garantir cet accès. Avec une stagnation des postes d’archivistes depuis plus de 10 ans en Suisse romande, mais en parallèle d’une multiplication des missions qui leur sont allouées, souvent en lien avec des enjeux numériques. Le risque est de diluer ou d’essouffler le métier d’archiviste face à la multiplication d’enjeu au lieu de le faire évoluer.

Il parait également difficile une généralisation des moyens des AFS à l’ensemble des archives cantonales en suisse, ce qui fait des moyens alloués comme souvent le déterminant d’une politique fédéraliste à plusieurs vitesses. Les moyens seront donc centraux à la qualité de la conservation du patrimoine, mais aussi à sa diffusion et à son accès auprès du public.

Si une augmentation de ces moyens n’est pas envisageable dans le contexte actuel, des enjeux déjà évoqués en 2012[[76]](#footnote-76) tels que la collaboration institutionnelle ou encore la collaboration directe avec les chercheurs par l’intermédiaire des sciences participatives seront déterminant au développement de stratégie numérique en archives. Il s’agira de trouver des synergies afin de s’appuyer sur les pôles de compétence de chacun.

Mutualiser les moyens et les discussions sur les questions numériques communes pourrait être crucial pour l’ensemble du domaine des GLAM pour faire exister et former des professionnelles ayant un socle spécialisé propre, mais interdisciplinaire sur le plan numérique auxquelles les humanités numériques ont à mon sens un rôle à jouer.

À l’avenir, sans renier ses tâches traditionnelles, il s’agit donc pour l’archiviste de devenir un *« médiateur culturel et de révélateur de sens »* selon les mots de l’archiviste Diane Baillargeon[[77]](#footnote-77) qui précise également *« […] sans tomber dans le piège du tout à la technologie. L’approche humaniste qui permet aux archivistes d’avoir une vision holistique de leur profession et du monde qui les entoure est le seul moyen de ne pas être assimilé à de super-techniciens qui savent tout du comment, mais rien du pourquoi. »*

C’est-à-dire, en utilisant la numérisation, mais aussi des ontologies, le web sémantique (Linked data) et l’intelligence artificielle afin d’organiser l’information pour qu’elle soit la plus diffusable et réutilisable par l’usager dans un environnement numérique. Mais de le faire dans le cadre de stratégie cohérente en prenant en compte les moyens et le contexte de chacun.

À ce titre le domaine des archives semble très bien s’adapter aux enjeux actuels et la faitière de la profession soit l’Association des archivistes suisse (AAS) a très bien su cerner dans son dernier "Livre blanc"[[78]](#footnote-78) les enjeux des nouvelles technologies pour une utilisation responsable dans le domaine. Aujourd’hui chaque archiviste, à l’opportunité d’appliquer ces stratégies à son contexte professionnel. Mais plus encore comme ce travail le montre d’en mesurer l’impact tant pour des besoins de management d’un service d’archives que d’en faire des indicateurs *d’advocacy* pour défendre les moyens attribuer aux archives et au patrimoine.

Ce n’est que de cette façon que l’apport du numérique en archives pourra favoriser sa diffusion, assurer son rôle démocratique, et être à même de toucher l’ensemble des usagers. Cela passera par une numérisation et plus d’automatisation, qui font aujourd’hui et pour longtemps encore, du numérique un enjeu contemporain majeur pour le domaine des archives.

Bibliographie

AFS, Archives fédérales suisses. Stratégie 2021-2025. [en ligne]. Disponible à l’adresse : https://www.bar.admin.ch/bar/fr/home/ueber-uns/das-bundesarchiv/strategie.html [consulté le 16 juillet 2025].

CHANCELLERIE FÉDÉRALE, 2025. Suisse numérique. *Suisse numérique* [en ligne]. 2025. Disponible à l’adresse : https://digital.swiss/fr/plan-d-action/mesures/acces-en-ligne-aux-archives [consulté le 14 mars 2025].

**Archives cantonales vaudoises (ACV), Rapport d’activité 2015**, [en ligne], Archives cantonales vaudoises, disponible depuis la page *Publications* du site officiel de l’État de Vaud : <https://www.vd.ch/acv/publications> (consulté le 10 août 2025)

**Archives cantonales vaudoises (ACV), Rapport d’activité 2022**, [en ligne], ACV, disponible à l’adresse : <https://www.vd.ch/acv/publications> (section « Publications » > Rapports d’activité 2022) (consulté le 10 août 2025)

**Arnold, Jonas ; Lüpold, Martin ; Theilkäs, Lorenz ; Kansy, Lambert**, *L’apprentissage automatique dans les archives : l’indexation en profondeur au service de l’accès aux archives*, livre blanc, [en ligne], Verein Schweizerischer Archivistinnen und Archivare (VSA-AAS), juin 2024, 41 pages, disponible à l’adresse : <https://vsa-aas.ch/wp-content/uploads/2024/08/MachineLearning_im_Archiv_Whitepaper_2024-08-08_fr.pdf> (consulté le 13 août 2025).

**Clavaud, Florence ; Wildi, Tobias**, « ICA Records in Contexts-Ontology (RiC-O) : a Semantic Framework for Describing Archival Resources », in *Linked Archives 2021 : Proceedings of Linked Archives International Workshop 2021*, co-located with 25th International Conference on Theory and Practice of Digital Libraries (TPDL 2021), Aachen, CEUR Workshop Proceedings (nᵒ 3019), 2021, pp. 79-92 [en ligne], disponible à l’adresse : <https://enc.hal.science/hal-03965776> (consulté le 28 mars 2025).

**Conférence des directrices et directeurs d’Archives suisses (CDA-ADK), Statistiques d’archives suisses**, [en ligne], CDA-ADK, 2025 ; disponible à l’adresse : <https://www.adk-cda.ch/fr/statistiques-darchives> (consulté le 13 août 2025).

**« Contexte des fonctions archivistiques »**, [en ligne], Wikipédia, consulté via la page *Archivistique* (section « Fonctions archivistiques ») sur le site Wikipédia, disponible à l’adresse :<https://fr.wikipedia.org/wiki/Archivistique#:~:text=La%20tradition%20québécoise%20distingue%20les,Diffusion> (consulté le 13 août 2025).

**« Contexte de l’ère de l’information »**, [en ligne], *Wikipédia*, article *Ère de l’information* (consulté via la section correspondante), disponible à l’adresse : <https://fr.wikipedia.org/wiki/%C3%88re_de_l%27information> (consulté le 13 août 2025).

**Couture, Carol (dir.)**, *Les fonctions de l’archivistique contemporaine*, Sainte-Foy : Presses de l’Université du Québec (collection « Gestion de l’information »), 1999, xxii + 559 p., ISBN 2-7605-0941-9, [en ligne], disponible à l’adresse : <https://www.jstor.org/stable/j.ctv18phbgp> (consulté le 13 août 2025)

**Françoise Hiraux ; Françoise Mirguet**, *La valorisation des archives : Une mission, des motivations, des modalités, des collaborations. Enjeux et pratiques actuels*, Louvain-la-Neuve : Academia – Publications des archives de l’UCL, 2012, ISBN 978-2-8061-0059-7, [en ligne], <https://doi.org/10.4000/questionsdecommunication.8564> (consulté le 13 août 2025).

**International Council on Archives (ICA), Code de déontologie de l’ICA**, adopté à Pékin lors de l’Assemblée générale du 6 septembre 1996, [en ligne], International Council on Archives, 1996, disponible à l’adresse : <https://www.ica.org/fr/resource/code-de-deontologie-de-lica/> (consulté le 2 août 2025).

**Rollinet, Raphaël**, *Etude-sur-la-diffusion-des-archives-a-l-ere-numerique* [en ligne], 12 août 2025, consulté le 13 août 2025, disponible à l’adresse : <https://github.com/Raphix93/Etude-sur-la-diffusion-des-archives-a-l-ere-numerique> (consulté le 13 août 2025).

**Rollinet, Raphaël**, *Etude sur la diffusion des archives à l’ère numérique. Version 1 (v1.0.0)*, dépôt Zenodo, 2025, [en ligne], <https://doi.org/10.5281/zenodo.16890209> (consulté le 17 août 2025).

**Servais, Paul ; Mirguet, Françoise (dir.)**, *Archivistes de 2030 : Réflexions prospectives*, Louvain-la-Neuve / Paris, Academia-L’Harmattan – Publications des archives de l’Université catholique de Louvain, février 2015, 427 p., ISBN 978-2-8061-0214-0, [en ligne], disponible à l’adresse : <https://www.harmatheque.com/ebook/archivistes-de-2030-reflexions-prospectives-45801> (consulté le 16 août 2025).

**Suisse, Canton de Neuchâtel — Loi sur l’archivage (LArch)**, Recueil systématique de la législation neuchâteloise (RSN 442.20), du 22 février 2011 (entrée en vigueur le 1er janvier 2012), [en ligne], mis à jour au 1ᵉʳ janvier 2021 ; disponible à l’adresse : <https://rsn.ne.ch/DATA/program/books/RSN2021/20211/pdf/44220.pdf> (consulté le 13 août 2025).

**Suisse, Canton de Neuchâtel — Loi sur la transparence des activités étatiques (LTAE)**, adoptée le 28 juin 2006 (RSN 150.50), entrée en vigueur le 1ᵉʳ octobre 2007, Recueil systématique de la législation neuchâteloise, [en ligne], disponible à l’adresse : <https://rsn.ne.ch/DATA/program/books/RSN2010/20075/htm/15050.htm> (consulté le 13 août 2025).

**Crédits iconographiques**

**Rollinet, Raphaël**, *Boites d’archives (photographie personnelle)*, prise en 2020. Illustration de couverture (droite).  
**Slovinský, NOIRLab/NSF/AURA/T**, *English: Server racks at the NOIRLab Headquarters Computer Server Room in Tucson, Arizona*, 2022, [en ligne], , disponible à l’adresse : <https://commons.wikimedia.org/wiki/File:NOIRLab_HQ_Server_Racks_(6V6A0404-CC).jpg>, consulté le 13 août 2025. Illustration de couverture (gauche).

1. Déclaration sur l’honneur

Une image contenant écriture manuscrite, calligraphie, Police, écrit à la main

Le contenu généré par l’IA peut être incorrect.Une image contenant texte, capture d’écran, Police, document

Le contenu généré par l’IA peut être incorrect.

Rollinet

Raphaël

MA - Patrimoine régional et humanités numériques

FLSH

Yverdon, le 17 août 2025

1. Entretien semi-conduit de Baptiste de Coulon

**Partie 1**

*Compte rendu d’un premier entretien établi le 23 mai 2024 sur Teams dans le cadre de mon travail de bachelor*

**Question 1 : Quel est votre parcours dans le domaine ?**

Baptiste de Coulon[[79]](#footnote-79) est archiviste au sein de la Fondation SAPA[[80]](#footnote-80), Archives suisses des arts de la scène, depuis 2021. Il est responsable du portail performing-arts.ch[[81]](#footnote-81) qui publie à la fois des données générales sur les arts de la scène en Suisse (qui, quoi, où), mais aussi les descriptions des collections conservées par la fondation. Les descriptions des collections dans ce portail sont compatibles avec la nouvelle norme de description Records in Contexts (RiC) [[82]](#footnote-82). Elles y sont exprimées sous forme de données liées (RDF). Il commença par des études en géographie à l’Université de Genève (UNIGE) puis un DEA en recherche à l’EHESS (Paris) suivi par un certificat en nouvelles technologies du WEB à l’UNIGE et enfin un MAS ALIS (Archival Library and Information Science) aux universités de Berne et de Lausanne. Il débuta dans le domaine des sciences en information par un stage dans les Archives de l’Institut d’architecture de l’Université de Genève en 2004.

Il travailla en tant qu’archiviste indépendant pour les communes neuchâteloises. Fondateur du service intercommunal d’archivage (SIAr) [[83]](#footnote-83). Il travailla ensuite en tant que responsable de l’antenne romande de l’entreprise docuteam, puis en tant que membre du conseil de direction de cette dernière. Il poursuivi par un poste chez Memoriav. Il est actuellement employé à la Fondation SAPA en tant que responsable de la plateforme numérique. Ses tâches concernent particulièrement la gestion des AIS, mais également la gestion des données numériques. Dans le cadre de son métier, il porte un intérêt important pour l’open data et l’open source.

**Question 2 : Quelles étaient les réflexions autour de la migration à SAPA de la base Scope sur un outil compatible RiC ?**

La Fondation SAPA est la successeure de plusieurs institutions qui lui ont transféré leurs collections en 2017 au moment de sa création. Il en découle une pluralité d’outil de description tel que Scope Archiv, mais également des outils moins normés comme FileMaker.

Le passage au RDF est le fruit de réflexions initiées dès 2014 au sein de la Collection suisse du théâtre (CST) qui souhaitait fusionner ses différentes bases de données et y donner accès aux publics via Internet. Ces réflexions ont été nourries notamment par des consultants externes tel que Beat Estermann spécialiste de l’open data en Suisse et membre du groupe sur l’ouverture des GLAM[[84]](#footnote-84).

Entre 2014 et 2017, Beat Estermann travailla sur des solutions et fournit au CST un rapport dans lequel il recommandait qu’avec une telle variété de bases la seule solution pour une mise en commun serait l’utilisation du RDF. À noter que cela fut énoncé avant l’arrivée du modèle RiC pour les archives. Un grand effort fut d’ailleurs fait pour modéliser à partir de là les activités des arts de la scène. Cela a donné lieu à des publications scientifiques[[85]](#footnote-85). On appelle à présent cet ensemble de données le « Panorama» chez SAPA.

**Question 3 : Comment c’est passé la migration de vos bases de données ?**

Les travaux ont commencé par la migration en RDF des données contenues au sein de l’AIS Scope Archiv. Ces données étaient déjà normées selon ISAD(G) ce qui devait faciliter la tâche.

Scope Archiv a été créé à l’origine sur la demande des institutions publiques suisses, en particulier des Archives fédérales. Ces institutions auraient soutenu la création de la société privée éditrice de ce logiciel Scope Solutions AG. L’outil était centré autour de de la description. Toutefois, profitant d’une situation quasi-monopolistique cette entreprise a mené jusqu’à aujourd’hui des barèmes tarifaires très élevés avec des développements informatiques toujours sous licence Les institutions se sont trouvées piégées et beaucoup d’argent public a été englouti par cette société sans qu’aucune contrepartie, par exemple en termes d’ouverture du code informatique ou de soutien aux petites institutions ne soit réellement exigé. Au niveau financier, pour la Fondation SAPA il était très intéressant de remplacer rapidement cet outil. Ce chantier est à présent terminé.

Le deuxième gros travail de migration, qui lui est toujours en cours, ce sont les bases FileMaker dont les données sont moins normées. FileMaker est un système de base de données où l’on peut rajouter librement un champ. De fait, cela peut donc partir dans tous les sens. C’est à la fois un avantage et un inconvénient. Les travaux de migration depuis FileMaker ne font que commencer au sein de la Fondation SAPA. C’est un gros morceau.

L’ensemble de ces migrations nécessite un mapping spécifique vers le RDF de SAPA. La Fondation utilise en partie RiC pour décrire ses collections, mais également beaucoup d’autres ontologies, dont une ontologie maison[[86]](#footnote-86). RiC est donc seulement une partie de la réponse lorsque l’on passe aux données liées (RDF).

**Question 4 : Comment créer une transversalité et une interopérabilité avec RiC ?**

Il est possible d’effectuer un matching entre nos entités RDF et celles des référentiels externes, comme Wikidata. On peut le faire au niveau des autorités, mais aussi sur les vocabulaires contrôlés[[87]](#footnote-87). Cela permet réellement une circulation entre les bases et c’est tout à fait nouveau. On sort de la logique des silos. Wikidata joue un rôle de noyau informationnel (hub) dans ce système[[88]](#footnote-88). Wikidata permet évidemment d’expérimenter aussi le crowd sourcing. Attention toutefois, le temps de travail pour réaliser ces matching n’est pas à sous-estimer.

Le passage au RDF permet d’envisager aussi des recherches fédérées à travers plusieurs bases de données

L’utilisation du langage SPARQL est obligatoire pour interroger des graphes RDF. C’est un langage qui permet vraiment de redécouvrir ses données et de poser des questions complexes.

L’un des problèmes qui se pose actuellement à la Fondation SAPA est que sa base AIS et que son portail d’accès public ne sont pas séparés. C’est un problème que l’on retrouve par exemple lors de l’utilisation du logiciel AtoM. Avec RDF, sans séparation strictement des deux bases, il est difficile de cacher de l’information lors sa diffusion, en particulier si on veut exposer les encodages techniques. La solution pourrait venir de l’utilisation de 2 bases dont une serait dédiée à la consultation.

Par ailleurs, ces bases sont encore des outils complexes pour les publics et même pour les archivistes. Ces outils demandent pour en tirer pleinement parti de développer un intérêt pour leurs caractéristiques techniques.

Dans le portail performing-arts.ch, il n’existe pas pour l’instant de véritables outils de médiation (type guide de recherche) pour faciliter la découverte, mais la Fondation SAPA souhaite changer cela prochainement.

Une problématique très actuelle dans la gestion des données sur les objets patrimoniaux, c’est la circulation et duplication des métadonnées. On peut compter au moins 4 lieux de stockage différents d’enregistrement de ces métadonnées : les AIP dans le dépôt d’archivage numérique, les descriptions dans l’AIS, les métadonnées embarquées dans les fichiers offertes en consultation et même les Manifest IIIF si on a choisi de mettre en place cette brique technologique. Les données liées pourraient simplifier cela, mais on attend encore les retours d’expérience pour pouvoir vraiment l’affirmer.

Par ailleurs, ce qui est primordial, c’est d’utiliser des systèmes d’identifiants pérennes (par exemple ARK) pour les notices et assurer que l’URI soit redirigé sur la bonne page. Cela est une des grandes réussites du portail performing-arts.ch. Son système d’URI lui permet de réellement de participer au monde des données liées. Et même d’y obtenir 5 étoiles[[89]](#footnote-89) !

**Question 5 : Qu’apportera la norme RiC dans la diffusion et la valorisation des archives ?**

La Fondation SAPA face au constat de la faible visibilité des données sur les archives dans les recherches internet, axe à présent sa stratégie et ses moyens plus sur la qualité de ses données (et leur matching avec des données externes) que sur l’ergonomie et le graphisme de son portail. Sur ce point le passage aux données liées et à Records in Contexts peuvent apporter beaucoup. Via leur exposition dans Wikidata, les données liées relatives à nos collections peuvent être ensuite réutilisées par d’autres services ou par de nouveaux publics qui n’auraient pas nécessairement pensé à utiliser des archives.

**Question 6 : Au niveau de la diffusion des images de ta base, est-ce que vous utilisez IIIF et que vous apporte-t-elle ?**

La Fondation SAPA met en œuvre les recommandations IIIF (en particulier l’élaboration de Manifest IIIF) et utilise une visionneuse IIIF (Clover). Cela a permis un gain de qualité dans la diffusion, non seulement des images fixes[[90]](#footnote-90), mais également du son[[91]](#footnote-91) et bientôt des vidéos. Nous avons bénéficié en cela des conseils précieux de Julien Rémy. L’élaboration de Manifest IIIF permet d’utiliser le potentiel des visionneuses IIIF et de les partager plus facilement via leur identifiant unique. Il s’agit d’un outil de valorisation qui sort véritablement les archives d’une logique de silos.

**Partie 2**

*Compte rendu d’un second entretien établi le 24 octobre 2025 à 11h sur Teams dans le cadre de mon mémoire de Master.*

**Remarques**

Il s’agit d’une retranscription d’un entretien oral. Certaines imprécisions peuvent encore subsister.

1. Entretien semi-conduit de Julien Rémy

1. Paul Servais ; Françoise Mirguet (dir.), *Archivistes de 2030 : Réflexions prospectives*, Louvain-la-Neuve / Paris : Academia L’Harmattan – Publications des archives de l’Université catholique de Louvain, février 2015, pp. 40-41 (consulté le 10 août 2025). [↑](#footnote-ref-1)
2. Carol Couture (dir.), *Les fonctions de l’archivistique contemporaine***,** Sainte‑Foy : Presses de l’Université du Québec, coll. « Gestion de l’information », 1999, xxii + 559 p. (ISBN 2‑7605‑0941‑9), [en ligne], consulté le 13 août 2025. [↑](#footnote-ref-2)
3. *« Contexte des fonctions archivistiques », section « Fonctions archivistiques »,* *Archivistique*, Wikipédia [en ligne], consultée le 13 août 2025. [↑](#footnote-ref-3)
4. Yvon Lemay ; Anne Klein, *« La diffusion des archives ou les 12 travaux des archivistes à l’ère du numérique »*, in *Les Cahiers du numérique*, vol. 8, n° 3, 2012, pp. 15-48, https://doi.org/10.3166/LCN.8.3.15-48. [↑](#footnote-ref-4)
5. [↑](#footnote-ref-5)
6. Yvon Lemay ; Anne Klein*, « La diffusion des archives ou les 12 travaux des archivistes à l’ère du numérique »*, in *Les Cahiers du numérique*, vol. 8, n° 3, 2012, p. 16. [↑](#footnote-ref-6)
7. Canton de Neuchâtel, *Loi sur l’archivage (LArch)*, RSN 442.20, du 22 février 2011 (entrée en vigueur le 1ᵉʳ janvier 2012), Recueil systématique de la législation neuchâteloise [en ligne], mise à jour au 1ᵉʳ janvier 2021, consultée le 13 août 2025 [↑](#footnote-ref-7)
8. Canton de Neuchâtel, *Loi sur la transparence des activités étatiques (LTAE)*, adoptée le 28 juin 2006, entrée en vigueur le 1er octobre 2007, RSN 150.50 [en ligne], consultée le 13 août 2025. [↑](#footnote-ref-8)
9. International Council on Archives (ICA), Code de déontologie de l’ICA, adopté à Pékin (Assemblée générale, 6 septembre 1996) [en ligne], International Council on Archives, 1996, consulté le 13 août 2025. [↑](#footnote-ref-9)
10. Jonas Arnold ; Martin Lüpold ; Lorenz Theilkäs ; Lambert Kansy, L’apprentissage automatique dans les archives : l’indexation en profondeur au service de l’accès aux archives, livre blanc, VSA‑AAS, juin 2024 [en ligne], pp. 1-41 (consulté le 13 août 2025) [↑](#footnote-ref-10)
11. [↑](#footnote-ref-11)
12. [↑](#footnote-ref-12)
13. [↑](#footnote-ref-13)
14. [↑](#footnote-ref-14)
15. [↑](#footnote-ref-15)
16. [↑](#footnote-ref-16)
17. *Archives fédérales suisses (AFS), Stratégie 2021–2025, Département fédéral de l’intérieur [en ligne], consultée le 13 août 2025.* [↑](#footnote-ref-17)
18. [↑](#footnote-ref-18)
19. [↑](#footnote-ref-19)
20. [↑](#footnote-ref-20)
21. [↑](#footnote-ref-21)
22. [↑](#footnote-ref-22)
23. [↑](#footnote-ref-23)
24. [↑](#footnote-ref-24)
25. [↑](#footnote-ref-25)
26. DISTAM, 2023. [étude] Quel avenir pour les instruments de recherche au format EAD produits avec le logiciel Access to Memory (AtoM)? *DISTAM* [en ligne]. 12 février 2023. Disponible à l’adresse : <https://distam.hypotheses.org/4941> [consulté le 28 mars 2025]. [↑](#footnote-ref-26)
27. [↑](#footnote-ref-27)
28. [↑](#footnote-ref-28)
29. *Rollinet, Etude-sur-la-diffusion-des-archives-a-l-ere-numerique…, op. cit., Q4 : « Connaissez-vous le champ disciplinaire des “Humanités numériques” ou “digital humanities” ? »* [↑](#footnote-ref-29)
30. Conférence des directrices et directeurs d’Archives suisses (CDA-ADK), *Adresses des archives représentées dans la CDA*, CDA-ADK [en ligne], consultée le 13 août 2025. [↑](#footnote-ref-30)
31. Servais ; Mirguet, Archivistes de 2030…, op. cit., pp. 162-165. [↑](#footnote-ref-31)
32. *Rollinet, Etude-sur-la-diffusion-des-archives-a-l-ere-numerique…, op. cit., Q20 : « Sélectionnez l’importance que vous accordez à chacune des compétences générales suivantes dans l’exercice du métier d’archiviste, en plaçant le curseur entre 0 (pas du tout importante) et 100 (absolument essentielle).»* [↑](#footnote-ref-32)
33. Rollinet, *Etude-sur-la-diffusion-des-archives-a-l-ere-numerique*…, op. cit.,, Q17 : « Dans votre service d’archives, la gestion des projets numériques est de la responsabilité de quel profil professionnel ?» [↑](#footnote-ref-33)
34. Rollinet, *Etude-sur-la-diffusion-des-archives-a-l-ere-numerique*…, op. cit.,, Q6 : « Quelles sont d'après vous les compétences numériques les plus utiles en archives aujourd'hui ?» [↑](#footnote-ref-34)
35. [↑](#footnote-ref-35)
36. [↑](#footnote-ref-36)
37. [↑](#footnote-ref-37)
38. CDA-ADK, Statistiques d’archives..., op. cit. [↑](#footnote-ref-38)
39. *Rollinet, Etude-sur-la-diffusion-des-archives-a-l-ere-numerique…, Notebook\_visualisation\_statistiques\_archives.ipynb, op. cit* [↑](#footnote-ref-39)
40. Lemay ; Klein, *op. cit.* p. 16. [↑](#footnote-ref-40)
41. AFS, Stratégie 2021–2025, op. cit. [↑](#footnote-ref-41)
42. CDA-ADK, Statistiques d’archives suisses, op. cit. [↑](#footnote-ref-42)
43. AFS, Stratégie 2021–2025, op. cit. [↑](#footnote-ref-43)
44. *Rollinet, Etude-sur-la-diffusion-des-archives-a-l-ere-numerique…, op. cit., Q14 : « Pensez-vous que le numérique peut amener des biais dans l’accès aux archives, si oui donnez un exemple ? »* [↑](#footnote-ref-44)
45. Archives cantonales vaudoises (ACV), *Rapport d’activité 2022*, ACV [en ligne], p. 44, consulté le 13 août 2025. [↑](#footnote-ref-45)
46. CDA-ADK, Statistiques d’archives suisses, op. cit. [↑](#footnote-ref-46)
47. AFS, Stratégie 2021–2025, op. cit. [↑](#footnote-ref-47)
48. Arnold et al., L’apprentissage automatique dans les archives…, op. cit.. [↑](#footnote-ref-48)
49. Rollinet, *Etude-sur-la-diffusion-des-archives-a-l-ere-numerique*…, op. cit., Q14 « Pensez-vous que le numérique peut amener des biais dans l’accès aux archives, si oui donnez un exemple ? » [↑](#footnote-ref-49)
50. AFS, Stratégie 2021–2025, op. cit. [↑](#footnote-ref-50)
51. CDA-ADK, Statistiques d’archives..., op. cit. [↑](#footnote-ref-51)
52. Archives cantonales vaudoises (ACV), *Rapport d’activité 2015*, ACV [en ligne], consulté le 10 août 2025 [↑](#footnote-ref-52)
53. Réponse par mail : *« […] cette hausse importante des documents consultés en 2020 s’explique par la fermeture prolongée de la salle de lecture pendant le COVID et la transmission de nombreux documents numérisées à distance. Le compatage des pièces s’est fait à la page numérisée. »* [↑](#footnote-ref-53)
54. Archives fédérales suisses (AFS), Stratégie 2021–2025, AFS [en ligne], consultée le 16 juillet 2025. [↑](#footnote-ref-54)
55. Rollinet, Etude-sur-la-diffusion-des-archives-a-l-ere-numerique…, op. cit., Q13 « Dans le cadre de la stratégie numérique suisse (voir:https://digital.swiss/fr/), la confédération prévoit l’accès aux archives par défaut en numérique, que pensez-vous de cette politique d’accès ? » [↑](#footnote-ref-55)
56. Arnold et al., L’apprentissage automatique dans les archives…, op. cit. [↑](#footnote-ref-56)
57. Lemay ; Klein, La diffusion des archives…, op. cit., p. 16. [↑](#footnote-ref-57)
58. [baptiste.decoulon@sapa.swiss](mailto:baptiste.decoulon@sapa.swiss) [↑](#footnote-ref-58)
59. https://sapa.swiss/ [↑](#footnote-ref-59)
60. https://www.siar.ch/ [↑](#footnote-ref-60)
61. https://www.performing-arts.ch/ [↑](#footnote-ref-61)
62. http://shapes.performing-arts.ch/collections/ [↑](#footnote-ref-62)
63. https://opendata.ch/fr/projets/openglam-working-group/ [↑](#footnote-ref-63)
64. Voir https://hcommons.org/deposits/item/hc:30581 [↑](#footnote-ref-64)
65. Voir article de bilan à ce sujet <https://doi.org/10.55790/journals/ressi.2024.e1511> [↑](#footnote-ref-65)
66. http://vocab.performing-arts.ch/ [↑](#footnote-ref-66)
67. Voir article de Baptiste de Coulon à paraître prochainement dans la Gazette des archivistes sur le sujet [↑](#footnote-ref-67)
68. https://5stardata.info/en/ [↑](#footnote-ref-68)
69. Voir http://data.performing-arts.ch/r/712a8e0a-4f24-4bb4-8015-392d228d60e7 [↑](#footnote-ref-69)
70. Voir http://data.performing-arts.ch/r/9559762d-cf89-4cf2-beed-9b383cdd8fb6 [↑](#footnote-ref-70)
71. [↑](#footnote-ref-71)
72. [↑](#footnote-ref-72)
73. [↑](#footnote-ref-73)
74. [↑](#footnote-ref-74)
75. [↑](#footnote-ref-75)
76. Lemay ; Klein, La diffusion des archives…, op. cit., p. 16. [↑](#footnote-ref-76)
77. Servais ; Mirguet, *Archivistes de 2030…*, op. cit., p. 32. [↑](#footnote-ref-77)
78. Arnold et al., L’apprentissage automatique dans les archives…, op. cit. [↑](#footnote-ref-78)
79. [baptiste.decoulon@sapa.swiss](mailto:baptiste.decoulon@sapa.swiss) [↑](#footnote-ref-79)
80. https://sapa.swiss/ [↑](#footnote-ref-80)
81. https://www.performing-arts.ch/ [↑](#footnote-ref-81)
82. http://shapes.performing-arts.ch/collections/ [↑](#footnote-ref-82)
83. https://www.siar.ch/ [↑](#footnote-ref-83)
84. https://opendata.ch/fr/projets/openglam-working-group/ [↑](#footnote-ref-84)
85. Voir https://hcommons.org/deposits/item/hc:30581 [↑](#footnote-ref-85)
86. Voir article de bilan à ce sujet <https://doi.org/10.55790/journals/ressi.2024.e1511> [↑](#footnote-ref-86)
87. http://vocab.performing-arts.ch/ [↑](#footnote-ref-87)
88. Voir article de Baptiste de Coulon à paraître prochainement dans la Gazette des archivistes sur le sujet [↑](#footnote-ref-88)
89. https://5stardata.info/en/ [↑](#footnote-ref-89)
90. Voir http://data.performing-arts.ch/r/712a8e0a-4f24-4bb4-8015-392d228d60e7 [↑](#footnote-ref-90)
91. Voir http://data.performing-arts.ch/r/9559762d-cf89-4cf2-beed-9b383cdd8fb6 [↑](#footnote-ref-91)