



HAL
open science

Le numérique est-il un territoire hostile pour l'archiviste ?

Céline Guyon

► **To cite this version:**

Céline Guyon. Le numérique est-il un territoire hostile pour l'archiviste?. Les fondamentaux de l'archivistique : un génome en évolution ?, Association des archivistes français - AAF, Feb 2017, Montpellier, France. hal-04371414

HAL Id: hal-04371414

<https://hal.science/hal-04371414v1>

Submitted on 3 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



RASAD 2017

Les fondamentaux de l'archivistique : un génome en évolution ?

Montpellier, 2-3 février 2017

Le numérique est-il un territoire hostile pour l'archiviste ?

Céline Guyon

1^{er} Constat : Le paradigme des « 4 C » (Collecte, classement, conservation et communication) traduit une forme de linéarité dans l'enchaînement de nos missions depuis la collecte jusqu'à la communication des archives au public en salle de lecture. Le numérique et notamment Internet viennent bouleverser cet ordonnancement linéaire et redistribuer les cartes. Aujourd'hui, les archives circulent, sont commentées, partagées, réappropriées, réutilisées. Sur le terrain de l'accès aux archives, les Archives se retrouvent en concurrence avec de nombreux acteurs.

2^{ème} constat, l'archiviste n'est plus le seul à manipuler des données : les CDO¹ (Chief data officer), PRADA (personnes responsables de l'accès aux documents administratifs), CIL (correspondant informatique et liberté), chargé de mission Open data sont-ils les meilleurs amis ou ennemis de l'archiviste... ? Alors même que le discours sur les données, entre Open data et Big data, est devenu un enjeu de pouvoir ?

3^{ème} constat : la polysémie du terme archives est une source de malentendu ; généralement les archives sont associées aux vieux papiers ; le « sens commun » des archives me semble aujourd'hui avoir dévié vers un nouvel objet, les archives pour désigner les informations disponibles et accessibles sur le Web.². Il y a comme une mise en concurrence des archives (en tant que support d'information) avec les données, documents, traces et information, entretenue par une forme de confusion terminologique.

¹ Sur l'articulation « archiviste-CDO », voir l'intervention de Jean-Daniel Zeller lors de la seconde édition du Forum des archivistes (30, 31 mars et 1^{er} avril 2016) : <http://www.slideshare.net/AssociationAF/le-cdo-est-il-le-meilleure-ami-ennemi-de-l-archiviste>

² Voir à ce sujet l'émission de radio *Du grain à moudre* « Faut-il mettre les archives à l'abri du changement politique ? » sur France Culture en date du 1^{er} février au cours de laquelle les archives ont été désignées comme les informations disponibles sur le Web : <https://www.franceculture.fr/emissions/du-grain-moudre/faut-il-mettre-les-archives-labri-du-changement-politique>

Communication, diffusion, réutilisation

Les Archives, en tant qu'institution participent au mouvement de circulation et de diffusion des archives comme en témoignent les millions d'images d'archives accessibles depuis les sites Internet des services d'archives ou les nouvelles formes de médiation et d'appropriation avec notamment les plateformes d'annotation collaborative.

Dans le même temps, la question de la réutilisation des données publiques (et donc des archives) est une question irritante pour la profession dont le point de crispation porte notamment sur l'état civil³. En l'occurrence, des fonds largement numérisés et accessibles sur les sites Internet des services publics d'archives. Cette friction des archivistes avec la réutilisation des données publiques ne serait-elle pas le signe d'une crainte, inconsciente, de perte d'une forme de monopole dans l'accès aux archives, après avoir vu la fréquentation des salles de lecture baisser ?

Sur le terrain de l'accès aux archives, les Archives se retrouvent en concurrence avec de nombreux acteurs, y compris les services producteurs qui mettent à disposition certaines de leurs données en Open data. Si on se place du point de vue de l'utilisateur, cette concurrence renvoie à la visibilité des acteurs et notamment à leur visibilité sur le Web. Le projet de portail France Archives s'inscrit bien, me semble-t-il, dans cette recherche de visibilité.

En matière de communication, nous avons transposé, dans l'environnement numérique, nos pratiques et convictions en faisant des sites Internet des services d'archives, des salles de lecture virtuelles. Ce modèle est-il encore pertinent aujourd'hui ? Les sites Internet des services d'archives sont-ils le meilleur média pour diffuser les fonds numérisés, alors même que les formes d'accès aux archives sont complètement renouvelées ? Les Archives n'auraient-elles pas intérêt à s'appuyer sur d'autres médias, sur d'autres relais qui ont plus de visibilité sur le Web, pour diffuser leurs ressources ?⁴

Les enjeux liés à la protection de la vie privée

Les enjeux liés à la communication ne s'expriment plus tant aujourd'hui en termes d'accès à l'information qu'en termes de protection de la vie privée, dans un équilibre en constante renégociation entre transparence, droit à l'oubli et la volonté du législateur de redonner au citoyen (et au consommateur) la maîtrise de ses données. Par ailleurs, les lois relatives à la transparence administrative et à la protection de la vie privée ont de plus en plus d'impacts sur le domaine des archives.

Les trois textes emblématiques que sont la loi du 3 janvier 1978, la loi du 17 juillet 1978 (codifiée dans le Code des relations entre le public et l'administration) et la loi sur les archives du 3 janvier 1979 (codifiée dans le Code du patrimoine) se font échos par des références croisées.

L'accès aux archives s'exerce dans les conditions définies pour les documents administratifs (article L231-1 du Code du patrimoine) ; et l'article 36 de la loi Informatique et libertés précise que « Les données à caractère personnel ne peuvent être conservées au-delà de la durée prévue au 5° de l'article 6 qu'en vue d'être traitées à des fins historiques, statistiques ou scientifiques ; le choix des données ainsi conservées est opéré dans les conditions prévues à l'article L. 212-3 du code du patrimoine ».

Au-delà de l'articulation de ces trois lois par références croisées, comment s'articulent, en termes de périmètres, les fonctions de PRADA, CIL et archiviste ? Car tous les trois manipulent des données et des documents et donc des archives au sens du Code du patrimoine.

Les PRADA sont les personnes, au sein des administrations, responsable de l'accès aux documents administratifs ; la fonction existe depuis 2005 ; elle a été introduite par le décret n° 2005-1755 du 30

³ Renvoyer à Filae et autres

⁴ Voir à ce sujet les exemples de partenariats avec Wikimedia

décembre 2005 relatif à la liberté d'accès aux documents administratifs et à la réutilisation des informations publiques.

Le décret précise que le PRADA a pour mission « de réceptionner les demandes de communication et les éventuelles réclamations, de veiller à leur instruction, d'assurer la liaison entre leur administration et la CADA, et qu'elle peut être également chargée d'établir un bilan annuel des demandes d'accès aux documents administratifs et de licence de réutilisation des informations publiques ».

La fonction de CIL a également été créée en 2005 à la suite de la révision de la loi Informatique et libertés en août 2004 ; le CIL, en tant référent sur les questions de protection des données personnelles veille à la sécurité juridique et informatique de son organisation⁵. Le CIL tient notamment à jour le registre des traitements mis en œuvre dans son organisme. Avec le récent Règlement européen sur la protection des données personnelles, la fonction de CIL va être renforcée et le CIL devient le DPO (délégué à la protection des données)⁶.

Le PRADA a pour mission de faciliter l'accès aux données publiques ; le CIL d'encadrer le cycle de vie des données à caractère personnel (depuis leur collecte jusqu'à leur destruction, en passant par leur conservation et leur communication).

Les champs d'intervention du PRADA et du CIL sont clairement délimités, tant du point de vue des finalités associées (la protection de la vie privée ou la transparence administrative) que de la nature des données traitées (documents administratifs et données publiques ou données à caractère personnel). Et l'archiviste ? Il s'occupe de tout cela à la fois.....Comment articuler les missions de l'archiviste avec les missions qui sont celles des PRADA, CIL ?

Les CILs et les archivistes interviennent dans un champ commun, celui de la définition des durées de conservation et avec deux approches fondamentalement différentes dans leurs finalités : le droit à l'oubli et la mémoire collective...ce qui, par nature, est susceptible de créer des zones de frictions et d'incompréhension de part et d'autre.

Et la zone de recouvrement est importante car le concept d'archives englobe la notion de données à caractère personnel ; quant au périmètre du CIL, il ne se limite pas, contrairement à une idée reçue, au support électronique. L'article 5 de la loi Informatique et libertés issu de la refonte de 2004, précise : « La présente loi s'applique aux traitements automatisés de données à caractère personnel, ainsi qu'aux traitements non automatisés de données à caractère personnel contenues ou appelées à figurer dans des fichiers, à l'exception des traitements mis en œuvre pour l'exercice d'activités exclusivement personnelles, lorsque leur responsable remplit les conditions prévues à l'article 5 ».

Le CIL a compétence pour apprécier ce qu'on qualifie de DUA (durée d'utilité administrative) : il définit la durée de conservation des données avant l'application de leur sort final, la destruction ou le transfert vers un service d'archives.

Un point essentiel de la loi informatique et liberté est la notion de finalité : les traitements de données à caractère personnel doivent répondre à une finalité déterminée et légitime. Et c'est la qualification de cette finalité qui va permettre d'une part de définir le type de données collectées (les données collectées doivent être pertinentes par rapport à la finalité du traitement) et d'autre part leur durée de conservation. La durée de conservation ne doit pas être excessive au regard à la finalité du traitement. La notion de finalité renvoie directement à la notion de valeur primaire. Et c'est bien la valeur primaire du document sur laquelle on s'appuie pour déterminer la DUA.

Dans son article 6, la loi Informatique et libertés précise que « les données à caractère personnel sont conservées (...) pendant une durée qui n'excède pas la durée nécessaire aux finalités pour lesquelles elles sont collectées et traitées ». Cette rédaction fait directement écho, en première lecture, à l'âge courant, ce que la CNIL qualifie de conservation en base active. Pour autant, l'article L212-3 du Code du patrimoine, précise qu'à l'issue de la durée de conservation prévue à l'article 6 de la loi Informatique et libertés, les données font l'objet d'une sélection. On est clairement ici à la fin de la DUA et donc de l'âge intermédiaire. Par ailleurs, l'article 36 de la loi Informatique et libertés précise que les données ne peuvent être conservées au-delà de la durée de conservation prévue à l'article 6

⁵ Dans la plaquette de présentation de la fonction de CIL, il est précisé que le CIL « permet d'optimiser la politique d'archivage »

⁶ <http://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L:2016:119:TOC>

qu'à des fins historiques, statistiques ou scientifiques : on est encore ici clairement à la fin de la DUA. On peut donc en déduire que la durée de conservation prévue à l'article 6 de la loi Informatique et libertés couvre l'âge courant et l'âge intermédiaire.

Même si les principes qui sous-tendent les deux lois sont opposés (mémoire et le droit à l'oubli), les enjeux liés à la protection de la vie privée peuvent au contraire être un levier pour nos projets d'archivage électronique. L'archivage est d'ailleurs présenté par la CNIL, dans son *Guide de la sécurité des données personnelles*⁷ comme participant à la sécurisation des données.

Le concept de *Privacy by design* est aujourd'hui porté par la CNIL. Il s'agit d'intégrer dans les outils informatiques, dès la phase de leur conception des mesures liées à la protection de la vie privée

Ce concept veut que chaque nouvelle solution informatique ou technologie traitant des données personnelles ou permettant d'en traiter doive garantir dès sa conception, et lors de chaque utilisation, le plus haut niveau possible de protection des données et de préservation dans le temps en fonction de leur finalité d'usage. C'est une belle opportunité de gérer le cycle de vie des données dès leur création et sans rupture, tout en identifiant très en amont les modalités d'extraction des données.

Pratique archivistique et Open data

Les archives « permettent aux citoyens d'exercer leurs droits dans le cadre d'une transparence qui est considérée aujourd'hui comme une condition de la démocratie » ainsi Guy Braibant introduisait-il son rapport sur « Les archives en France », en 1996. Aujourd'hui, « les données en Open data succèdent dans ce rôle aux archives »⁸.

L'Open data s'inscrit dans une tendance qui considère l'information publique comme un bien commun dont la diffusion est d'intérêt public et général. Les archives appartiennent au domaine public mobilier et leur conservation « est organisée dans l'intérêt public tant pour les besoins de la gestion et de la justification des droits des personnes physiques ou morales, publiques ou privées, que pour la documentation historique de la recherche ».⁹ Les valeurs partagées par les partisans de l'Open data et les archivistes sont similaires, les finalités de l'ouverture des données publiques et de l'archivage sont proches : transparence administrative, libre accès aux sources, respect de la vie privée, propriété collective des données. Pour autant, la profession apparaît très frileuse vis-à-vis du mouvement Open data qui renverse le dispositif de communication des données publiques puisqu'il encourage la diffusion et non plus simplement l'accès aux données. Faut-il y voir une contradiction de plus ?

Avant d'être ouvertes, les données passent par une série d'opérations qui font écho avec notre propre pratique professionnelle : l'identification des données, l'extraction des données et la mise en œuvre de l'intelligibilité des données¹⁰.

Sur le terrain de l'identification, de l'extraction et de l'intelligibilité, nous aurions tant à partager : l'archiviste a une bonne connaissance des données et de la production documentaire ; Quant à l'extraction des données, elle soulève exactement les mêmes difficultés, que les données soient destinées à être versées dans une plateforme d'Open data ou dans une plateforme d'archivage électronique. Par ailleurs, les acteurs de l'Open data s'accordent aujourd'hui sur la nécessité de pouvoir disposer de données fiables et de qualité.

Open data et archivage entretiennent un compagnonnage, dans les principes, les intentions et les modes opératoires. Mais aussi le vocabulaire.

Dans un article intitulé « Open data? : Data, information, document or record? », Erik Borglund et Tove¹¹, partagent les résultats d'une enquête sur le vocabulaire utilisé dans les documents de

⁷ http://www.cnil.fr/fileadmin/documents/Guides_pratiques/Guide_securite-VD.pdf

⁸ Intervention d'Aude Roelly et Marie Ranquet à l'occasion de la seconde édition du Forum des archivistes, 30, 31 mars et 1^{er} avril 2016.

⁹ Art L. 211-2 du Code du patrimoine.

¹⁰ Denis, J. & Goëta S. (2016) "Brutification" et instauration des données. La fabrique attentionnée de l'open data. i3 Working Papers Series, 16-CSI-01. http://www.i-3.fr/wp-content/uploads/2016/07/WPI3_16-CSI-01-Denis-Goeta.pdf

référence des discours de l'Open data et de l'archivage électronique. Les auteurs concluent que si le vocabulaire utilisé est différent dans les termes qui sont choisis, pour autant, un sens commun se dégage et c'est celui de *Records*, c'est-à-dire que le discours de l'Open data quand il utilise les termes de données et informations, c'est au sens de *Records*

Il apparaît indispensable et urgent de penser l'articulation entre l'Open data et la pratique archivistique au travers notamment de la question de la qualité et de la pérennisation des données.

Archives, données, documents

Le terme de données a été explicitement introduit par la loi LCAP en 2016¹². Cette insertion permet de réaffirmer que le cadre réglementaire de la gestion des archives demeure et doit demeurer le même, quel que soit le support des documents. Qu'il soit papier ou électronique, un document est porteur d'un ensemble d'informations déterminées par un contexte et un environnement de production. En aucun cas le changement de support n'interfère sur le contenu informationnel : la dématérialisation d'un acte par exemple n'entraîne pas modification des informations que contient cet acte. Par ailleurs, la représentation formelle des informations (sous la forme d'un document, de données dans une base de données ou un flux XML par exemple) ne peut pas être et ne doit pas être un critère pour introduire une distinction, notamment dans le régime juridique, entre les documents (sous-entendu papier) et les données (sous-entendu électroniques). C'est bien tout l'enjeu de l'introduction formelle du terme « données » dans la définition des archives.

Pour autant, l'environnement numérique et la pratique professionnelle quotidienne des archivistes interrogent l'actualité du concept d'archives tel que défini dans le Code du patrimoine, notamment dans la sphère publique. Malgré un cadre juridique clair et explicite, il y a de plus en plus de frictions entre le cadre théorique et la pratique quotidienne de la gestion documentaire et archivistique, au sein des administrations et services publics d'archives.

La définition très (trop ?) englobante des archives telle que précisée dans la loi et telle que mise en œuvre aujourd'hui par les professionnels, n'est-elle pas de nature à desservir les professionnels des archives eux-mêmes et les archives patrimoniales. ? Comment trouver un équilibre entre une définition très englobante des archives et une volumétrie sans cesse grandissante des documents et données créés ?¹³

Dans la sphère publique en France, l'archiviste ne s'autorise, pas me semble-t-il, à se désintéresser (consciemment, volontairement et de façon assumée) d'une partie des documents pour concentrer, en retour, son attention sur certains d'entre eux...La loi impose en effet une même attention à tous les documents. Cette situation est-elle réellement tenable aujourd'hui avec le numérique?¹⁴

Une piste de réflexion pourrait être justement une approche par le risque. Une telle approche présuppose de hiérarchiser les informations selon leur(s) fonction(s) (c'est-à-dire leur rôle et leur valeur pour les utilisateurs) et de graduer les exigences en termes de conservation et de sécurisation des documents et des données en conséquence. On pourrait, en écho avec les traitements de données à des fins archivistiques¹⁵, introduire l'expression de « documents à valeur archivistique ». Les documents à valeur archivistique seraient soumis à un régime juridique plus contraignant que les documents sans valeur archivistique, en termes de conservation et de contrôle par l'administration des conditions de conservation. Et donc qualifier, dans la loi, les archives du point de vue de leur valeur et non plus du point de vue de leurs usages/utilités (avec les archives courantes, intermédiaires et définitives).

¹¹ Erik Borglund Tove Engvall , (2014), "Open data?", *Records Management Journal*, Vol. 24 Iss 2 pp. 163 -180Permanent link to this document : <http://dx.doi.org/10.1108/RMJ-01-2014-0012>

¹² Loi n° 2016-925 du 7 juillet 2016 relative à la liberté de la création, à l'architecture et au patrimoine :

¹³ Celine Guyon. Le concept d'archives : d'une définition à l'autre. 2016. <https://hal.archives-ouvertes.fr/hal-01401138>

¹⁴ L'était-elle dans l'environnement papier ?

¹⁵ Expression introduite par le règlement européen 2016/679 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation des données.

Conclusion

Nicolas Martin, journaliste à France Culture et producteur de l'émission La méthode scientifique introduisait par ces mots l'émission du 1^{er} février 2017 intitulée « Comment repenser Internet et février comme un lieu d'archivage de l'histoire » : « On aborde une nouvelle ère, celle de la seconde révolution d'Internet qui a pour but d'ajouter le temps comme principe d'indexation »¹⁶

« Avec leur savoir-faire historique et leurs outils conceptuels renouvelés, les archivistes sont à même d'identifier les données de référence, archives de demain, et de promouvoir leur archivage pérenne. Il importe donc qu'ils participent à la réflexion et fassent entendre leur voix singulière dans le débat actuel sur les sources de la connaissance en devenir » (Banat-Berger, Nougaret, 2014).¹⁷

Nous n'avons plus le monopole ni de la collecte, ni de la conservation, ni de la communication des archives. Alors que nous reste-t-il ? De nous affirmer comme « créateur de valeur dans la chaîne documentaire » (Banat-Berger, 2010)¹⁸ et de participer à la culture de la donnée en partageant nos concepts et savoir-faire.

A nous d'inventer une nouvelle équation entre celui qui produit, celui qui utilise et celui qui conserve les archives. La recherche de cet équilibre passe, me semble-t-il, par un renforcement de la gestion archivistique des données : c'est-à-dire par leur contextualisation, description, évaluation et par la traçabilité de leur conservation.

¹⁶<https://www.franceculture.fr/emissions/la-methode-scientifique/comment-transformer-internet-en-machine-remonter-le-temps>

¹⁷ Banat-Berger Françoise, Nougaret Christine, Faut-il garder le terme archives ? Des « archives » aux « données », *La Gazette des archives* n°233 (2014-1)

¹⁸ BANAT-BERGER (Françoise), Les archives et la révolution numérique, *Le Débat*, n° 158, janvier-février 2010, Gallimard